Original Research

Multiple-Site Molecular Modification of Dioxin-Like PCBs to Eliminate Bioconcentration

Minghao Li, Xiaolei Wang, Zhenhua Chu, Yu Li*

The Moe Key Laboratory of Resources and Environmental Systems Optimization,
North China Electric Power University, Beijing 102206, China

Abstract

The environmentally friendly modification of chemical pollutants lacks the research of multiple-site simultaneous modification. In addition, there is a lack of efficient screening methods for multiple-sites of molecules, such as polychlorinated biphenyls and their derivatives. In this study, three-dimensional quantitative structure-activity relationship (3D-QSAR) models were established using experimental bioconcentration factor (BCF) data on 58 polychlorinated biphenyls (PCBs). Based on the experimental and predicted values by 3D-QSAR, a 2^{10-3} fractional factorial design with a resolution of V was used to evaluate the effects of molecular modification (single, double, or triple) at different positions on PCB bioconcentration. Finally, molecular docking technique was used to explore the mechanism of bioconcentration in the receptor of PCBs before and after multiple-site modification. The results showed that ten substitution solutions of main effect, second-order interaction effect and third-order interaction effect that significantly influenced the bioconcentration of dioxin-like PCBs were obtained. Two derivatives prepared by modification of the substituents at two sites were decoupled from the bioconcentration effect. In addition, Molecular docking results showed that the main factors affecting PCB bioconcentration were the degree of matching between hydrophilic/hydrophilic amino acid residues near the binding site and non-covalent forces formed after binding of the ligand to the receptor. The above results explore the rationality of modification schemes for low bioconcentration molecules, and provide theoretical design reference for non-bioconcentration PCBs substitutes.

Keywords: dioxin-like polychlorinated biphenyl, bioconcentration, three-dimensional quantitative structure-activity relationship, factorial experiment, molecular modification

*e-mail: liyuxx8@hotmail.com

Introduction

The polychlorinated biphenyls (PCBs) are 209 congeners with distinct chlorine-substituted biphenyl structures, and were among the first of 12 compounds to be designated as persistent organic pollutants under the Stockholm Convention. Because of their extremely stable physical and chemical properties, good electrical insulation, high heat resistance, and low flammability, PCBs have been used extensively as dielectric fluids for capacitors and transformers, flame retardants, and in a variety of the other commercial products [1]. From the 1930s through the mid-1970s, the production of PCBs was used in a wide range of industrial applications [2]. Since the 1980s, various countries have gradually stopped production of PCBs [3], but large quantities of commercial PCB mixtures are still present in electrical systems, landfills, and stored on land [4]. Consequently, humans will be exposed to these persistent pollutants in the environment for centuries [5].

The Cl isomer ratios in a commercial mixture of PCBs [4] indicated that most were trichlorobiphenyl to heptachlorobiphenyl isomers, followed by chlorobiphenyl and dichlorobiphenyl isomers, and finally octachlorobiphenyl and decachlorobiphenyl isomers. It is estimated that about 65% of the PCBs produced in the world are present in electrical systems, landfills, or in stores on land, 31% are lost to the environment, and the rest are degraded or incinerated [4]. Studies on PCBs in lichens in parts of Ontario, Canada have found that the predominant PCBs that accumulate in lichens are trichlorobiphenyl to octachlorobiphenyl isomers [6]. In Arctic seabirds, the main PCBs that accumulated were trichlorobiphenyl to decachlorobiphenyl isomers [7]. The PCBs present in sharks in Greenland were mainly trichlorobiphenyl to nonachlorobiphenyl isomers [8]. These studies indicate that although more chlorobiphenyl and dichlorobiphenyl isomers are produced and enter the environment than octachlorobiphenyl to decachlorobiphenyl isomers, the amounts of octachlorobiphenyl to decachlorobiphenyl isomers detected in vivo are much less. This shows that a high degree of chlorination makes it easier for a PCB to concentrate in an organism than a low degree of chlorination.

The concentration of PCBs detected in soil from the center of Pristina, Kosovo and Metohija was 168.57 µg/kg [9]. PCBs have been detected in the air in Brasilia, Brazil, and attributed to persistent use of old equipment containing PCBs [10]. In US and Canadian rivers, PCBs have been detected at 0.63 ng/L and 0.09 ng/L, respectively [11]. The concentration range for PCBs detected in vegetables on British organic farms was 0.83-2.68 µg/kg [12]. PCBs have been also detected in egg yolks [13], fish [14], birds [15], and human blood [16], adipose tissue [17], and hair [18]. PCB concentrations in breastmilk from women in urban and rural areas in Zhejiang Province, China [19] were 2.66-3.90 pg/g and 1.83-2.27 pg/g, respectively, and may

threaten the health of the next generation. In Belgium, PCBs were detected in human adipose tissue and livers at average concentrations of 490 ng/g and 380 ng/g, respectively, and the concentrations were observed to increase with age, which suggested that food and air may be the main sources for human exposure to these chemicals [20]. Accumulation of PCBs in the human body can indirectly cause lesions in internal organs [21], the endocrine system [22], and the reproductive system [23].

The bioconcentration potential of a chemical is normally expressed as a bioconcentration factor (BCF), which is defined as the concentration ratio of a chemical inside an organism to its concentration in the surrounding environment. The BCF of a chemical can be used to assess its environmental risks and for classification of dangerous substances [24]. Ivanciuc et al. [25] developed a quantitative superstructure-activity relationship (QSSAR) method to predict $\log BCF$ values for 151 PCBs. However, conducting many molecular structure alignment analyses using the QSSAR method is time-consuming. In addition, this method does not consider the spatial orientation of the ligand and receptor interaction. The QSSAR method cannot quantitatively correlate the force field distribution around the molecule with active sites that affect the BCF values of PCBs for molecular design. Melo [26] used a two-dimensional quantitative structure-property relationship method to predict the $\log BCF$ values of 209 PCBs. Because of the large number of quantum chemical descriptors of PCBs, the established two-dimensional quantitative structure-property (2D-QSPR) relationship model could not accurately represent the influence of structural parameters on the BCFs. A three-dimensional quantitative structure–activity relationship (3D-QSAR) model was established by Liu to analyze PCB bioconcentration [27], but it did not use both the CoMFA and CoMSIA methods. Compared with the CoMSIA method, the CoMFA method lacks contribution rate analysis of the hydrophobic, hydrogen bond donor, and hydrogen bond acceptor fields. The use of only the CoMFA method to describe the structures and properties of PCBs has limitations. Xu et al. [28] used a fractional factorial design method to perform molecular modification studies of substitution at one or two sites (e.g., the 2, 5, or 5' sites) that affect the persistence of PCBs. This proves that the double substitution has certain effect in the molecular modification. However, they did not consider the effect of substitution at the 3, 3', 6, and 6' sites or the effect of substitution at three sites on the environment persistence of PCBs. In view of the limitations of these methods, this study attempts to make a breakthrough and research the potential advantages of multiple-sites molecular modification.

The aim of this study is to explore the influence of third-order interactions on the bioconcentration of PCBs using 3D-QSAR and fractional factorial design methods, and determine an optimum design for
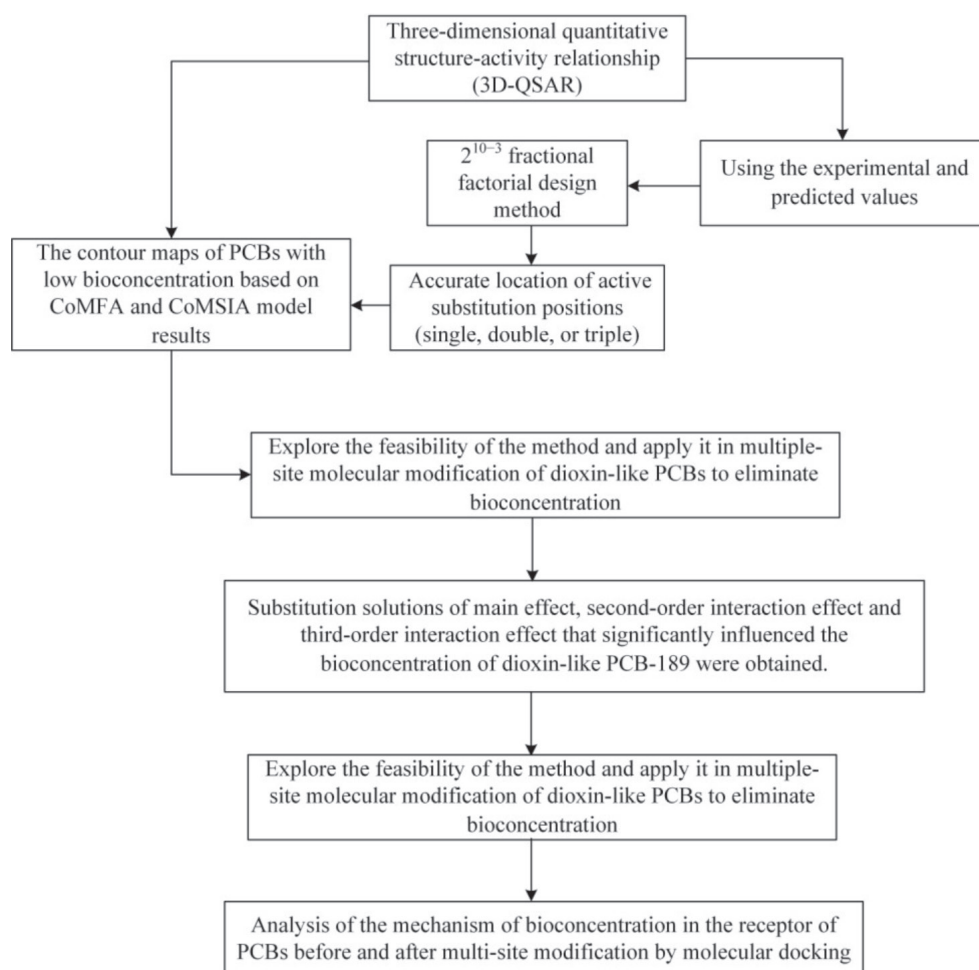
Fig. 1. The design route of novel PCBs multiple-site modifications.

PCBs that do not bioconcentrate, aiming to provide a different theoretical method for future research of PCBs derivatives. These methods were combined with molecular docking to analyze the reasons for the decrease in bioconcentration after binding of PCBs with receptors. The design route of novel PCBs multiple-site modifications has shown in Fig. 1.

Materials and Methods

Construction and Analysis the 3D-QSAR Model of PCB Bioconcentration

Experimental logBCF data for 58 PCBs [25] and their structural parameters were used as dependent and independent variables, respectively, to establish 3D-QSAR models. The data set was divided into a training set (48 compounds) for 3D-QSAR model generation and a test set (10 compounds) for model validation. Construction of the 3D-QSAR model was based on CoMFA and CoMSIA methods.

In this study, drawing the PCBs molecular structures using SYBYL-X 2.0 molecular modeling software package (Fig. 2). The resulting structures

were not the most stable conformations, so the energy minimized structures were used as the initial conformations for molecular alignment process. The geometry of PCBs molecular was then optimized using the Tripos force field with Gasteiger-Hückel charges. Repeated minimizations were performed using the Powell method with a maximum iteration of 10,000 to achieve an energy convergence gradient value of 0.005 kJ/mol. Ligand-based alignment was used to derive the best possible 3D-QSAR statistical model. The 3,3',4,4',5,5'-Hexachlorobiphenyl (PCB-169) molecule which has the highest logBCF value was used as the
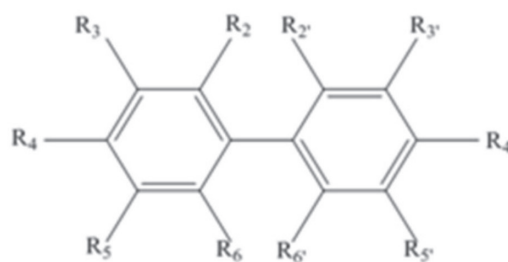

Fig. 2. Molecular structures of PCBs (R and R' represent H or Cl atoms).

template to align the other compounds using the Align Database command in the SYBYL software.

The CoMSIA model was calculated using a carbon probe atom that added the hydrophobic, hydrogen bond donor, and hydrogen bond acceptor fields compared with those in the CoMFA method. Although the CoMSIA method can overcome inherent defects of the CoMFA method, it does not necessarily provide better results [29]. In this study, the CoMFA and CoMSIA methods were used to supplement each other and obtain reliable prediction models.

In the PLS analysis, the leave-one-out procedure was performed to determine the optimum number of components (n) and the cross-validation correlation coefficient (q^2) for the correlation models. Simultaneously, non-validation analysis has performed [30], and the conventional multiple correlation coefficient (r^2), standard error of the estimate, Fisher test value (F), and the contribution rate of each molecular field have calculated. The cross-validation method was used to test whether the model had high external predictive ability [31], and the correlation coefficient for test set predictions (r^2pred) and the standard error of prediction were calculated.

Multi-Site Substitution Solutions for Reducing PCB Bioconcentration Based on a 2^{10-3} Fractional Factorial Design with a Resolution of V

A fractional factorial design was conducted using the factorial design module of Minitab. The results were used to identify substituent positions that affected the log*BCF* values of the PCBs. In an experimental table, the factor represents the number of active substitution sites and each factor contains two different levels. To ensure the accuracy of generated results, we used a design with 10 factors, 2 levels, and the highest possible resolution (V) to generate a table of active PCB substitution sites. The log*BCF* data were input as controllable fixed factors in the experimental table to generate normal distribution maps of PCB bioconcentration. The main (one substitution site), second-order interaction (two substitution sites), and third-order interaction (three substitution sites) effects on PCB bioconcentration were analyzed.

The replacement characteristics of the PCBs were analyzed with multi-order effect factor maps, and only the model of this paper could supply more substituent positions (single, double and triple binding sites) than previous literatures (no triple binding sites) (Table 1). The model was considered reliable and acceptable if q^2 (the cross-validated value) was higher than 0.50 and r^2 (the coefficient of determination) was higher than 0.90 [28]. A value of 1 implies that a linear correlation coefficient (R) describes the relationship between experimental data and predicting ones perfectly [32]. A residual graph was established for the model to verify the reliability of the PCB data. We analyzed the graph of residual normal distribution, the residual histograms, the distribution graph of residual and fitting values, the distribution graph of residuals, and the data sequence.

Dioxin-like PCBs are the most toxic of 209 congeners, including alteration of immune system function, developmental and neurodevelopmental effects, and increased risk of cancer. Four nonortho substituted congeners and eight mono-ortho congeners (3,3',4,4'-Tetrachlorobiphenyl (PCB-77), 3,4,4',5-Tetrachlorobiphenyl (PCB-81), 3,3',4,4',5-Pentachlorobiphenyl (PCB-126), 3,3',4,4',5,5'-Hexachlorobiphenyl (PCB-169), 2,3,3',4,4'-Pentachlorobiphenyl (PCB-105), 2,3,4,4',5-Pentachlorobiphenyl (PCB-114), 2,3',4,4',5-Pentachlorobiphenyl (PCB-118),

Table 1. Comparison of model parameters in this manuscript with those in previous predictions.

Molecular	q^2/ (q^2>0.5)	r^2/ (r^2>0.9)	R	Model	Molecular modification	Reference
PCBs	0.875	0.946	0.9779	CoMFA	Single, double, and triple	In this manuscript
	0.888	0.943	0.9799	CoMSIA		
PCNs	0.996	0.968	0.986	CoMSIA	Single, double	[33]
PCNs	0.966	0.992	0.939	CoMFA	Single	[34]
	0.976	0.991	0.945	CoMSIA		
PCNs	0.586	1	0.998	CoMFA	Single, double	[35]
	0.833	1	1	CoMSIA		
PCPs	0.718	0.985	0.992	CoMSIA	Single	[36]
PCBs	-	-	0.908	Cluster expansion	-	[25]
	-	-	0.959	Splinoid poset		
PCBs	-	-	0.901	QSPR	-	[26]
PCBs	0.821	0.926	0.962	CoMFA	-	[27]

2,3',4,4',5'-Pentachlorobiphenyl (PCB-123), 2,3,3',4,4',5-Hexachlorobiphenyl (PCB-156), 2,3,3',4,4',5'-Hexachlorobiphenyl (PCB-157), 2,3',4,4',5,5'-Hexachlorobiphenyl (PCB-167) and 2,3,3',4,4',5,5'-Heptachlorobiphenyl (PCB-189)) constitute 12 dioxin-like PCBs. PCB-189 is the molecule with the largest number of substitution sites. As the most representative molecule of dioxin-like PCBs, this study selected PCB-189 as an example to conduct the study of multiple-site molecular modification of dioxin-like PCBs to eliminate bioconcentration.

Calculation of Quantum Chemical Descriptors for the PCB-189 Derivatives

Before and after modification of PCB-189, six quantum chemical descriptors (total energy, energy gap, C-Cl bond dissociation enthalpy, frequency, Gibbs free energy, and energy barrier) were calculated at the B3LYP/6-31G* level [37] using Gaussian software. The results were used to detect PCB-189 molecules with low bioconcentration, and determine their functional changes.

Evaluation of the Persistent Organic Pollutant Characteristics of the PCB-189 Derivatives

We used 3D-QSAR models of PCB environmental persistence ($\log t_{1/2}$), long-distance migration ($\log K_{OA}$), and biological toxicity (pEC_{50}) established by Xu et al. [28], Chen et al. [4], and Li et al. [38] to evaluate persistent organic pollutant (POPs) characteristics before and after modification of PCB-189. We analyzed the results from the viewpoint of the number of sites of modification (one, two, or three).

Analysis of Changes that Affect Bioconcentration of the PCB Derivatives

To study the correlation between the number of substituted Cl atoms and the $\log BCF$ value, the PCBs were divided into chlorobiphenyl and decachlorobiphenyl substitution groups, and the relationship with the average $\log BCF$ was plotted. Non-covalent interactions play an important role in determining the stability of a protein structure, and most chemical reactions are dominated by weak interactions [39].

The molecular docking module in Discovery Studio 4.0 software was used for molecular docking between the molecules, both before and after modification, and six receptors present in the human body. The Libdock module was selected to perform the molecular docking of ligands and receptors. The option of docking preferences selected the high quality. The docking tolerance and the energy threshold selected default values that were 0.25 and 20, respectively. The binding sites of the ligands and receptors were extracted to study changes in the bioconcentration after docking of the 20 novel PCB-189 derivatives with the receptors by the software. The mechanism of interaction between each derivative and amino acid residue in the receptor was analyzed using the hydrophilicity/hydrophobicity of the amino acid residues and details for any non-covalent interactions near the docking molecule (type of interaction and average bond length). To determine the main factors affecting PCB bioconcentration, we investigated the bioconcentration after modification of PCB-189. These results could be used to provide important theoretical knowledge for molecular design of PCB derivatives.

Results and Analysis

The Feasibility Exploration of the Method and Application in Multiple-Site Molecular Modification of Dioxin-Like PCBs to Eliminate Bioconcentration

Evaluation of the logBCF Values of PCBs Based on the CoMFA and CoMSIA Models

In the CoMFA and CoMSIA models, the optimum numbers of components (n) were 6 and 9, respectively. The cross-validated correlation coefficients (q^2) were greater than 0.5 at 0.875 (CoMFA) and 0.888 (CoMSIA). The non-cross-validated (R^2) were greater than 0.9 at 0.946 (CoMFA) and 0.943 (CoMSIA. The standard errors of the estimate were 0.225 and 0.240 for the CoMFA and CoMSIA models, respectively. The F values were 119.126 (CoMFA) and 64.443 (CoMSIA). These results showed that the two models had good fit and predictive abilities. The standard error of prediction and r²pred values for the CoMFA model were 0.943 and 0.946, respectively. The corresponding values for the CoMSIA model were 0.225 and 0.240. These values indicate that the two models have good external predictive abilities (Table 2).

The CoMFA analysis revealed that the percentage contributions of steric and electrostatic fields were 29.8% and 70.2%, respectively. Therefore, the steric and electrostatic effects of the groups may influence the

Table 2. Evaluation of the parameters of the CoMFA and CoMSIA models based on PCB log*BCF* values.

Model	q^2	n	SEE	R^2	F	r^2_{pred}	SEP
CoMFA	0.875	6	0.225	0.946	119.126	0.943	0.212
CoMSIA	0.888	9	0.240	0.943	64.443	0.946	0.206

Table 3. Contributions of the molecular field to PCB bioconcentration estimated using the CoMFA and CoMSIA models.

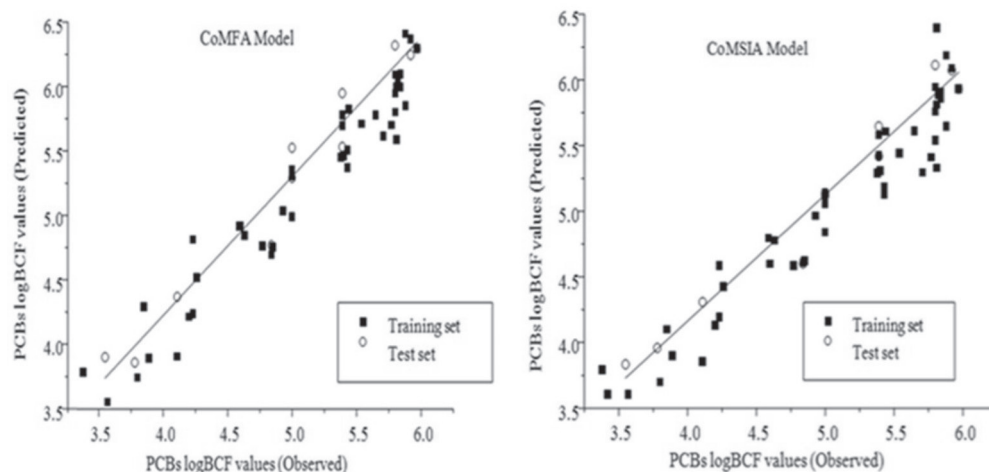| Model	S	E	H	D	A
CoMFA	29.8%	70.2%	-	-	-
CoMSIA	2.2%	72.7%	25.1%	0.00%	0.00%

Fig. 3. The plot of observed versus predicted logBCF values from the CoMFA and CoMSIA models.

logBCF values of the PCB homologs. The electrostatic interaction was a major contributor to the logBCF values of PCBs. The CoMSIA analysis revealed that the percentage contributions of the steric, electrostatic, hydrogen, hydrogen bond donor, and hydrogen bond acceptor fields were 2.2%, 72.7%, 25.1%, 0.00%, and 0.00%, respectively. These results indicate that the hydrogen bond donor and acceptor fields have negligible contributions, and the electrostatic field has the largest influence on the logBCF values of the PCBs (Table 3).

Analysis the correlations between the experimentally determined and predicted logBCF values for the training and test sets. The results showed that the CoMFA and CoMSIA models had a fine linear dependence and the correlation coefficients R were 0.9779 and 0.9799, which can predict the PCBs logBCF values well (Fig. 3).

Prediction of PCB logBCF Values and Verification of the CoMFA and CoMSIA Models

The CoMFA and CoMSIA models were predicting logBCF values for 209 PCBs (Table 4). The predicted logBCF values of 58 PCBs from the CoMFA and CoMSIA models had relatively small errors (<10%) compared with the experimental values, and were acceptable [30].

Ivanciuc [25], Melo [26], and Liu [27] reported logBCF values for PCBs that were predicted using QSSAR, 2D-QSPR, and CoMFA methods, respectively. Their results were compared with logBCF values predicted using our CoMFA and CoMSIA methods.

The correlation coefficients (R) for the CoMFA and CoMSIA methods, respectively, were: 0.9470 and 0.9174, 0.9207 and 0.9442, and 0.9226 and 0.8632. These results showed good linear dependence and validated the predicted values for the PCBs.

Analysis of the Contour Maps of PCBs with Low Bioconcentration Based on CoMFA and CoMSIA Model Results

Because PCB-169 had the highest logBCF value in the training set, it was used as an example to analyze the contour maps of the CoMFA and CoMSIA models (Fig. 4).

In the CoMFA steric field contour map (Fig. 4a), the green regions around R_2, R_3, R_5, $R_{2'}$, $R_{3'}$, $R_{4'}$ indicated that introducing small groups at these positions would decrease the logBCF values of the PCBs. In the CoMFA electrostatic field contour map (Fig. 4b), blue contours close to R_4 and indicated that introducing electropositive substituents at the R_4 position would decrease the logBCF values of the PCBs. Red contours around the R_2, R_5, R_6, $R_{2'}$, $R_{3'}$, $R_{5'}$, $R_{6'}$ positions indicated that introducing electropositive substituents would decrease the logBCF values of the PCBs.

In the CoMSIA steric field contour map (Fig. 4c), green regions located at R_2, R_3, R_5, $R_{3'}$, $R_{5'}$, and $R_{6'}$ and yellow regions close to one of the benzene rings indicated that introducing small groups at these positions would decrease the logBCF values of the PCBs. In the CoMSIA electrostatic field contour map (Fig. 4d), the blue contours close to R_4 and $R_{4'}$ and

Table 4. Predicted logBCF values for the 58 PCBs from the CoMFA and CoMSIA models.

NO.	Compounds	Obs.	CoMFA		CoMSIA	
			Pred.	Relative error (%)	Pred.	Relative error (%)
0	Biphenyl	2.64	2.476	-6.21	2.64	0
3	4-Chlorobiphenyl	2.77	2.834	2.31	2.929	5.74
4	2,2'-Dichlorobiphenyl	3.38	3.421	1.21	3.397	0.5
5	2,3-Dichlorobiphenyl	4.11	3.864	-5.99	3.852	-6.28
6	2,3'-Dichlorobiphenyl	3.80	3.719	-2.13	3.698	-2.68
7[a]	2,4-Dichlorobiphenyl	3.55	3.862	8.79	3.834	8
8	2,4'-Dichlorobiphenyl	3.57	3.546	-0.67	3.604	0.95
9	2,5-Dichlorobiphenyl	3.89	3.854	-0.93	3.898	0.21
14[a]	3,5-Dichlorobiphenyl	3.78	3.825	1.19	3.959	4.79
15	4,4'-Dichlorobiphenyl	3.28	3.589	9.42	3.453	5.27
18[a]	2,2',5-Trichlorobiphenyl	4.11	4.283	4.21	4.308	4.82
28	2,4,4'-Trichlorobiphenyl	4.20	4.142	-1.38	4.13	-1.67
29	2,4,5-Trichlorobiphenyl	4.26	4.416	3.66	4.424	3.85
31	2,4',5-Trichlorobiphenyl	4.23	4.163	-1.58	4.192	-0.9
40	2,2',3,3'-Tetrachlorobiphenyl	4.23	4.562	7.85	4.584	8.37
44[a]	2,2',3,5'-Tetrachlorobiphenyl	4.84	4.644	-4.05	4.601	-4.94
47	2,2',4,4'-Tetrachlorobiphenyl	4.85	4.631	-4.52	4.622	-4.7
48	2,2',4,5-Tetrachlorobiphenyl	5.00	4.843	-3.14	4.838	-3.24
49	2,2',4,5'-Tetrachlorobiphenyl	4.84	4.575	-5.48	4.605	-4.86
52	2,2',5,5'-Tetrachlorobiphenyl	4.63	4.712	1.77	4.775	3.13
54	2,2',6,6'-Tetrachlorobiphenyl	3.85	4.211	9.38	4.099	6.47
64	2,3,4',6-Tetrachlorobiphenyl	4.60	4.778	3.87	4.597	-0.07
70	2,3',4',5-Tetrachlorobiphenyl	4.77	4.64	-2.73	4.585	-3.88
77	3,3',4,4'-Tetrachlorobiphenyl	4.59	4.775	4.03	4.795	4.47
87	2,2',3,4,5'-Pentachlorobiphenyl	5.38	5.257	-2.29	5.288	-1.71
90[a]	2,2',3,4',5-Pentachlorobiphenyl	5.00	5.115	2.3	5.118	2.36
97	2,2',3,4',5'-Pentachlorobiphenyl	5.43	5.187	-4.48	5.129	-5.54
99	2,2',4,4',5-Pentachlorobiphenyl	5.00	5.125	2.5	5.139	2.78
101	2,2',4,5,5'-Pentachlorobiphenyl	5.40	5.264	-2.52	5.309	-1.69
105	2,3,3',4,4'-Pentachlorobiphenyl	5.00	5.13	2.6	5.058	1.16
109[a]	2,3,3',4,6-Pentachlorobiphenyl	5.00	5.324	6.48	5.139	2.78
118	2,3',4,4',5-Pentachlorobiphenyl	5.00	5.173	3.46	5.111	2.22
126	3,3',4,4',5-Pentachlorobiphenyl	5.81	5.384	-7.33	5.331	-8.24
128	2,2',3,3',4,4'-Hexachlorobiphenyl	5.77	5.483	-4.97	5.409	-6.26
136	2,2',3,3',6,6'-Hexachlorobiphenyl	5.43	5.307	-2.27	5.183	-4.55
137	2,2',3,4,4',5-Hexachlorobiphenyl	5.88	5.618	-4.46	5.646	-3.98
138	2,2',3,4,4',5'-Hexachlorobiphenyl	5.39	5.477	1.61	5.423	0.61
141	2,2',3,4,5,5'-Hexachlorobiphenyl	5.81	5.753	-0.98	5.811	0.02

Table 4. Continued.

148[a]	2,2',3,4',5,6'-Hexachlorobiphenyl	5.39	5.331[a]	-1.09	5.418	0.52
151	2,2',3,5,5',6-Hexachlorobiphenyl	5.54	5.492	-0.87	5.442	-1.77
153	2,2',4,4',5,5'-Hexachlorobiphenyl	5.65	5.551	-1.75	5.61	-0.71
155	2,2',4,4',6,6'-Hexachlorobiphenyl	4.93	4.883	-0.95	4.963	0.67
156[a]	2,3,3',4,4',5-Hexachlorobiphenyl	5.39	5.706	5.86	5.645	4.73
157	2,3,3',4,4',5'-Hexachlorobiphenyl	5.39	5.555	3.06	5.578	3.49
169	3,3',4,4',5,5'-Hexachlorobiphenyl	5.97	6.018	0.8	5.931	-0.65
174	2,2',3,3',4,5,6'-Heptachlorobiphenyl	5.80	5.572	-3.93	5.541	-4.47
180[a]	2,2',3,4,4',5,5'-Heptachlorobiphenyl	5.80	6.04	4.14	6.113	5.4
182	2,2',3,4,4',5,6'-Heptachlorobiphenyl	5.80	5.835	0.6	5.946	2.52
183	2,2',3,4,4',5',6-Heptachlorobiphenyl	5.84	5.75	-1.54	5.861	0.36
187	2,2',3,4',5,5',6-Heptachlorobiphenyl	5.80	5.704	-1.66	5.756	-0.76
191	2,3,3',4,4',5',6-Heptachlorobiphenyl	5.84	5.838	-0.03	5.907	1.15
194	2,2',3,3',4,4',5,5'-Octachlorobiphenyl	5.81	6.357	9.41	6.381	9.83
195	2,2',3,3',4,4',5,6-Octachlorobiphenyl	5.92	6.082	2.74	6.087	2.82
196[a]	2,2',3,3',4,4',5,6'-Octachlorobiphenyl	5.92	5.973	0.9	6.072	2.57
198	2,2',3,3',4,5,5',6-Octachlorobiphenyl	5.88	6.12	4.08	6.184	5.17
202	2,2',3,3',5,5',6,6'-Octachlorobiphenyl	5.82	5.807	-0.22	5.884	1.1
208	2,2',3,3',4,5,5',6,6'-Nonachlorobiphenyl	5.71	5.405	-5.34	5.291	-7.34
209	Decachlorobiphenyl	5.44	5.595	2.85	5.606	3.05

([a]Test set)

red contours close to R_3, R_5, $R_{2'}$, and $R_{3'}$ indicated that introducing electropositive substituents at the R_4 and $R_{4'}$ positions would decrease the logBCF values of the PCBs, whereas electropositive substituents at the R_3, R_5, $R_{2'}$, and $R_{3'}$ positions would have the same effect. In the CoMSIA hydrophobic field contour map (Fig. 4e), yellow regions close to one of benzene rings and white regions around R_2, R_3, R_4, R_5, R_6, $R_{2'}$, $R_{3'}$, $R_{4'}$, and $R_{5'}$ indicated that introducing hydrophobic substituents at these positions would decrease the logBCF values of the PCBs.

The optimum design for active substitution could not be determined based on the contour maps. Therefore, another method was required to accurately determine
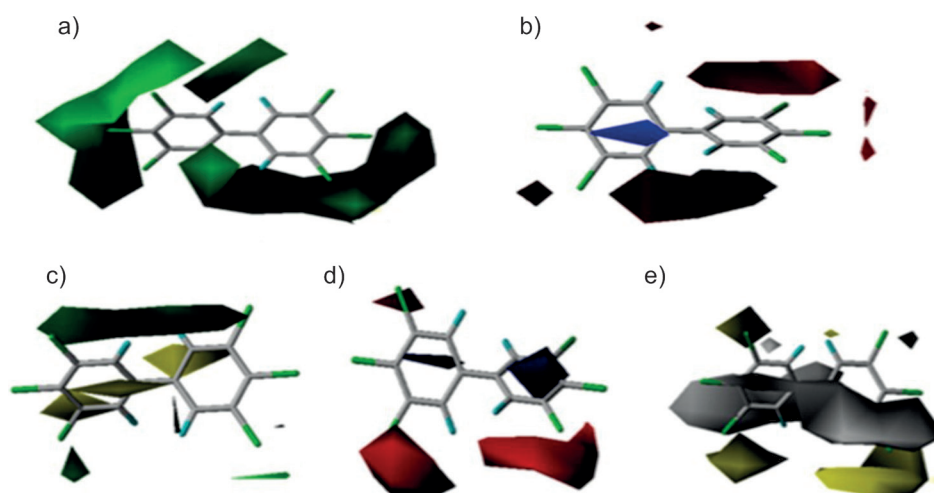

Fig. 4. Contour maps of the CoMFA and CoMSIA models.

Table 5. Main and significant second-/third-order interaction effects of the substitution positions.

Effect	Factor	Effect value		P(sig.)
		+	-	
Main	Ortho-R_2,$R_{2'}$ (A, F)	0.552		(0.000, 0.000)
	Meta-R_3,$R_{3'}$ (B, G)	0.882		(0.000, 0.000)
	Para-R_4,$R_{4'}$ (C, H)	0.780		(0.000, 0.000)
	Meta-R_5,$R_{5'}$ (D, J)	0.830		(0.000, 0.000)
	Ortho-R_6,$R_{6'}$ (E, K)	0.568		(0.000, 0.000)
The second-order interaction	Ortho-R_2, Ortho-R_6(AE)		-0.114	0.012
	Ortho-R_2, Meta-R_3(AG)		-0.095	0.026
	Meta-R_5, Ortho-R_6(DE)		-0.100	0.022
	Ortho-$R_{2'}$, Meta-$R_{3'}$(FG)		-0.106	0.017
	Ortho-$R_{2'}$, Meta-$R_{6'}$(FK)		-0.061	0.123
The third-order interaction	Ortho-R_2, Meta-R_5, Ortho-R_6(ADE)		-0.093	0.029

the active substitution sites, which could be used to optimize substitution and decrease bioconcentration of the PCBs.

Application of a 2^{10-3} Fractional Factorial Design with Resolution of V for Accurate Location of Active Substitution Positions to Reduce PCB Bioconcentration

We used 10 factors and 2 levels for the PCB substitution, which gave a totally of 1024 different substitution patterns. However, because of spatial symmetry and rotation, some different substitution patterns corresponded to the same PCBs homologs. The number of substitutions for each PCBs homolog could be determined from the symmetry of substitution between PCBs and the symmetry of a single benzene ring. In total, there were 209 unique PCBs. We input log*BCF* data as controllable fixed factors. Because of the limited experimental data and the large number of supplementary values required, both the CoMFA and CoMSIA methods were evaluated. The best model was selected to obtain supplementary data for the factorial design.

Although both CoMFA and CoMSIA models have good stability and predictive ability, the CoMFA model research has some limitations in the field of different forces on the impact of the molecular and molecular modification compared to the CoMSIA model. The CoMFA model includes only have the steric field and the electrostatic field, and the CoMSIA model includes five kinds of force fields (steric field, electrostatic field, hydrophobic field, hydrogen bond-donor field
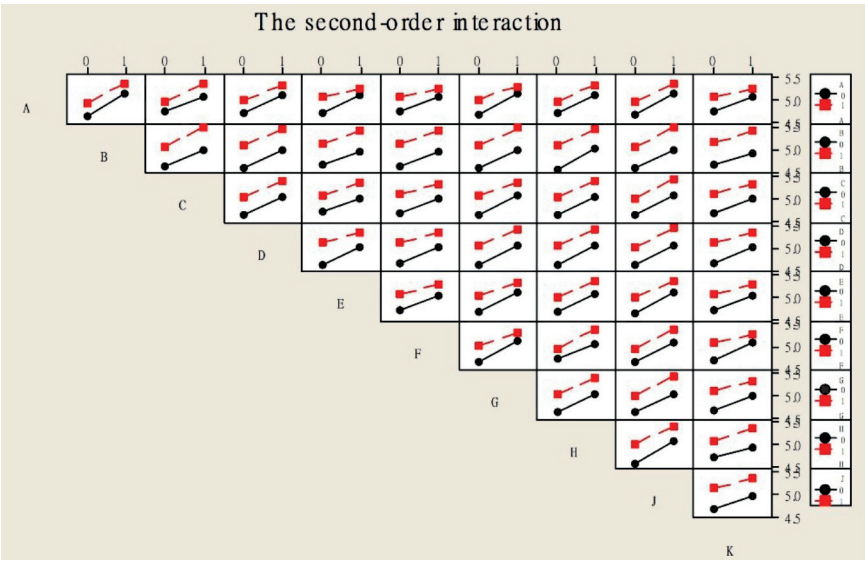

Fig. 5. Second-order interaction of substituent positions for PCB bioconcentration.

and hydrogen bond-acceptor field). The CoMSIA model can more fully understand the impact of the force fields around the molecule, the generated active substituent positions design table can be combined with the CoMSIA contour map to choose more suitable substituent groups, so we finally selected CoMSIA model predictive value as the input data source for low bioconcentration of PCBs active substituent sites analysis.

The important sites for modification with one substituent were R_2, $R_{2'}$, R_3, $R_{3'}$, R_4, $R_{4'}$, R_5, $R_{5'}$, R_6, and $R_{6'}$. The important sites for modification with two substituents were (R_2, R_6), ($R_{2'}$, R_3), ($R_{2'}$, $R_{6'}$), and (R_5, R_6), and those for modification with three substituents were (R_2, R_5, and R_6).

Analysis of the Replacement Characteristics of Multi-Order Interaction Sites for Reducing PCB Bioconcentration

The effects of PCB substitution based on the fractional factorial results were estimated. The estimated values for ortho-, meta-, and para-substitution were all significantly positively correlated with the log*BCF* (Table 5). The estimates for meta-substitution at $R_3/R_{3'}$ and $R_5/R_{5'}$ were similar and the largest among all the sites. The effects of substitution at different sites were in the order meta->para->ortho-substitution.

The results for the five second-order interactions were basically the same, and showed a negative correlation with the log*BCF* value (Fig. 5). The simultaneous introduction of substituent groups at both sites will produce an antagonistic effect on the molecular bioconcentration of PCBs.

Bioconcentration increases when there is no chlorine atom in the $R_2/R_{2'}$ positions but a chlorine atom is present in the $R_6/R_{6'}$ positions. Substitutions at two positions (A/E and F/K positions) (Table 5) was significantly negatively correlated with the log*BCF*, and negatively related to bioconcentration. For example, replacing the Cl atom at position 5 and 6 with other substituents resulted in a decrease in the log*BCF* value of PCB-10 compared with that of PCB-9. For second-order interaction effects at the D/E positions and F/G positions, the second-order interaction lines for (R_5, R_6) and ($R_{2'}$, $R_{3'}$) showed the opposite effect to the A/E and F/K results. Simultaneous replacement of the Cl atom substituents at both positions was negatively correlated with the log*BCF* value, and the bioconcentration reduced. For example, the log*BCF* value of PCB-93 was smaller than that of PCB-92, and the log*BCF* value of PCB-23 was smaller than the log*BCF* of PCB-29. Second-order interaction effects were observed for bioconcentration of the PCBs (e.g., ortho-R_2 and para-$R_{3'}$ substitution). This is because two benzene rings were substituted at the same time. Substitution at positions A and G was significantly negatively correlated with log*BCF*, and this could greatly reduce bioconcentration of the PCBs. For example, PCB-6 had a smaller log*BCF* value than PCB-7.

Suitability of the Factorial Design for Decreasing PCB Bioconcentration

The residual graph can be used as a goodness-of-fit graph for multiple linear regression analysis and analysis of variance to help determine if the ordinary least-squares hypothesis is satisfied. If these assumptions are satisfied, the ordinary least-squares regression will generate an unbiased estimate of the smallest variance.

In the normal distribution graph of the factorial experimental design model, the residuals of the PCBs log*BCF* data and fitting values were in an almost straight line with a normal distribution. This indicated that the experimental data used satisfied the model's normal distribution assumption. In the residual histograms graph of log BCF data and fitting values of the PCBs, the distribution of the patterns was uniform and no graph skew or abnormal values appeared. This indicated that the selected experimental data were stable. In the relationship graph between the residuals and the fitting values, the relative fitting values of the residuals were unevenly distributed in a fan shape, and the points did not show any regularity, which indicated that the error between the log BCF values and the fitting values of the PCBs satisfied the assumption of constant variance. In the residuals and data graph, most of the data fluctuated randomly around the centerline, confirming the assumptions about the log*BCF* values of PCBs and the error of fitting data. The residual data of [-0.1, +0.1] indicated that the PCB log*BCF* data satisfied data independence requirements (Fig. 6). In summary, the selected data meets the reliability requirements of the fractional factorial design and can be used to analyze multi-order interaction effects of the replacement of substituents on the log*BCF* values of the PCBs.

Design of New Dioxin-Like PCB-189 Molecules by Substitution at Multiple Sites to Reduce Bioconcentration

According to the CoMSIA contour map, R_2 and R_3 of the PCB-189 molecule were in green regions, R_2, R_3, $R_{3'}$, R_5, and $R_{5'}$ were in red regions, R_4 and $R_{4'}$ were in a blue region, and R_2, R_3, R_4 R_5, $R_{3'}$, $R_{4'}$, and $R_{5'}$ were in white regions. The parent PCB-189 was modified by substitution at one, two, or three sites. Position R_6 in PCB-189 is occupied by a stable H atom, and substitution at three sites only used the R_2 and R_5 positions.

The log*BCF* value of the designed molecule was lower than that of the parent PCB-189 (Table 6). After addition of one, two, or three substituents, the log*BCF* values were 10.69%-38.57%, 43.35%-58.90% and 44.18%-49.46% lower than that of PCB-189, respectively. According to the Toxic Substances Control Act issued by the US Environmental Protection Agency in 1998,
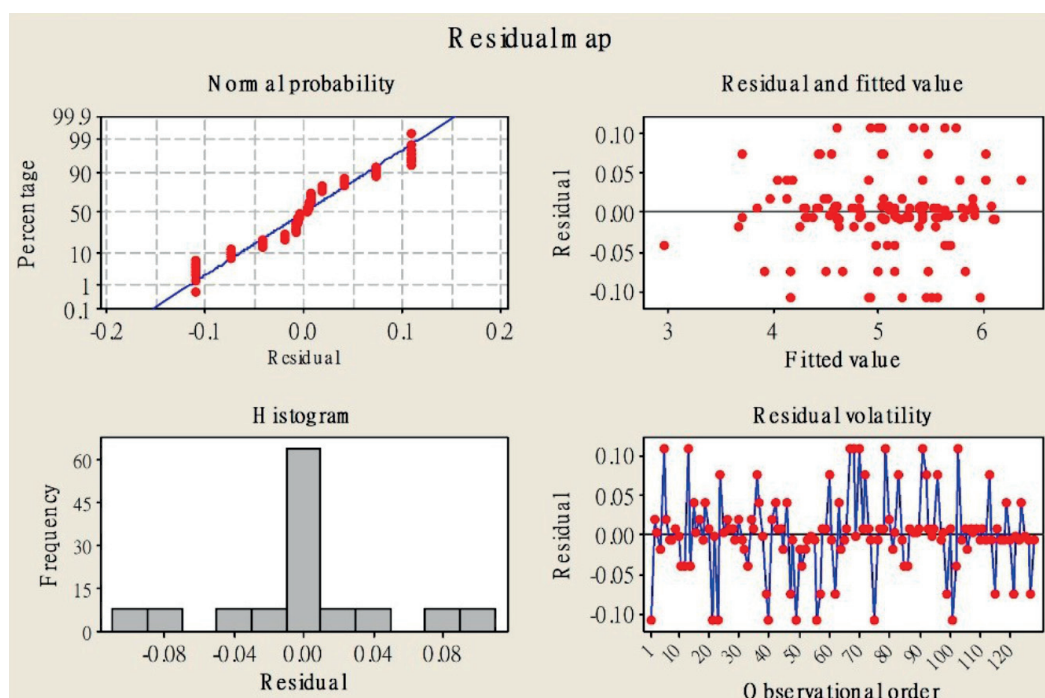
Fig. 6. The reliability test residual graph of the fractional factorial design model.

BCF values below 1000 indicate that a molecule will not bioconcentrate and BCF values between 1000 and 5000 indicate that a molecule could bioconcentrate [40]. The BCF values of the molecules with substitution at two or three sites all dropped below 5000, and those for two of the molecules with substitution at two sites were below 1000. These results show that the molecules have been decoupled from the bioconcentration effect.

Evaluation of functional property changes and the difficulty of substitution reactions to form the PCB-189 derivatives

Quantum chemistry descriptors for the 20 new PCB-189 molecules were calculated. The energy gap value [41] was used as the energy difference between the highest occupied molecular orbital and lowest unoccupied molecular orbital. The energy gap is the lowest energy required for the electron excitation process and is an important parameter reflecting the conductivity and luminescent properties of a material. A larger energy gap is correlated with weaker electrical conductivity. The total energy of the PCB-189 derivative was much lower than that of PCB-189, but the energy gap did not change. This shows that derivatization of PCB-189 will not affect the insulation performance or stability. Because PCBs contain multiple Cl atoms, we used the minimum C-Cl bond dissociation enthalpy as a parameter to evaluate the flame retardancy [35]. The results showed that the C-Cl bond dissociation enthalpy did not changed, indicating that the flame retardancies of the PCB-189 derivatives were not affected (Table 6).

To evaluate the difficulty of the PCB-189 substitution reaction, thermodynamics calculations were performed using PCB-189 and the substituent as reactants and the PCB-189 derivatives as products [42]. The Gibbs free energies (ΔG) of substitution reaction pathways for the PCB-189 derivatives were negative, and the reactions for formation of the 20 derivatives would occur spontaneously. To compare the difficulty of replacing chlorine atoms with new substituents, the energy barrier of the reaction was obtained by calculating the difference between the product transition state and the reactant energy [33]. There were no major differences in the ease of substitution among the various substituents (Table 6).

Long-Distance Migration, Environmental Persistence, and Biological Toxicity Assessment of the PCB-189 Derivatives

Compared with the parent PCB-189, the long-range migration, environmental persistence, and biotoxicity of the PCB-189 derivatives were reduced to varying degrees (Table 7). The following reductions were observed for PCB-189 derivatives produced by modification with substituents at one, two, and three sites, respectively: long-range migration, 0.19%-11.36%, 11.24%-17.15%, and 20.60%-24.35%; environmental persistence, 0.00%-30.76%, 3.19%-16.21%, and 0.06%-16.34%; and biological toxicity 0.95%-27.49%, 1.98%-18.54%, and 4.17%-20.01%. The effect of modification at three sites on long-distance migration was more effective than that of modification at one or two sites. The effect of modification at one site on environmental persistence was relatively good compared with the effects of modification at two or three sites. The effects of modification at two or three sites on environmental persistence were similar. The best reduction in biotoxicity was achieved with

Table 6. The logBCF values and quantum chemistry descriptors for 20 new PCB-189 derivatives.

	Compounds	COMSIA		Frequency (cm⁻¹)	ΔG (a.u.)	ΔE (a.u.)	Total energy (a.u.)	Energy gap (a.u.)	C-Cl bond dissociation enthalpy (kcal/mal)	Change rate (%)
		Pred.	Fall rate (%)							
0	Target molecule PCB-189	6.173					-3680.4391	0.19016	87.66	
1	2-ethyl-PCB-189	4.145	32.85	15.15	-0.1281	0.1559	-3299.4812	0.20014	87.6	-0.07
2	2-methyl-PCB-189	5.344	13.43	19.34	-0.1383	0.1631	-3260.1694	0.193741	87.66	0.00
3	3-aminomethyl-PCB-189	5.402	12.49	19.68	-0.1193	0.1399	-3315.507	0.189157	87.73	0.08
4	3-methyl-PCB-189	5.489	11.08	21.28	-0.1406	0.1681	-3260.17	0.190992	87.66	0.00
5	4-ethyl-PCB-189	5.545	10.17	19.23	-0.1327	0.1326	-3299.48	0.192881	87.66	0.00
6	4-hydroxymethyl- PCB-189	5.49	11.06	19.53	-0.1262	0.1531	-3335.368	0.191321	87.66	0.00
7	5-butyl-PCB-189	4.876	21.01	19.07	-285.8644	0.1592	-3378.1052	0.191342	88.17	0.58
8	5-allyl-PCB-189	4.884	20.88	20.45	-285.8394	1.3909	-3337.5576	0.190937	87.29	-0.42
9	3'-allyl-PCB-189	4.828	21.79	19.04	-0.107	0.1355	-3337.5564	0.189219	87.79	0.15
10	3'-butylene-PCB-189	5.419	12.21	17.75	-0.1065	0.0567	-3376.8769	0.188594	89.86	2.51
11	4'-ethyl-PCB-189	5.506	10.81	18.34	-0.1343	0.1601	-3299.4841	0.192528	88.29	0.72
12	4'-hydroxymethyl- PCB-189	5.513	10.69	19.27	-0.1278	1.07	-3335.3681	0.193294	86.85	-0.92
13	5' - ethyl-PCB-189	5.412	12.33	19.89	-0.1319	0.1579	-3299.4816	0.189969	88.23	0.65
14	5'-silyl-PCB-189	3.792	38.57	19.61	-0.1265	0.1473	-3091.2697	0.192109	132.03	50.62
15	2- methyl-3-ethyl-PCB-189	2.537	58.9	19.56	-0.27	0.3162	-3166.4743	0.189324	90.98	3.79
16	2-hydroxymethyl-3-silyl-PCB-189	2.834	54.09	25.93	-0.2632	0.1167	-2934.5463	0.196223	92.22	5.20
17	2- silyl-3'-pentyl- PCB-189	3.161	48.79	14.05	-0.2783	0.2032	-3168.6645	0.188727	89.16	1.71
18	2-methyl-3'-butyl- PCB-189	3.497	43.35	16.8	-0.269	0.2984	-2997.1455	0.203377	90.98	3.79
19	2-methyl fluoride-5-propyl-PCB-189	3.446	44.18	13.3	-0.1259	0.3277	-3017.7423	0.194303	76.71	-12.49
20	2-hydroxyl-5-pentyl-PCB-189	3.12	49.46	24.8	-0.1364	0.4562	-3032.9956	0.187501	101.47	15.75

Table 7. Prediction of the environment persistence, long-distance migration, and biotoxicity of the PCB-189 derivatives.

		$\log K_{OA}$		$\log t_{1/2}$		pEC50	
		CoMFA	Change rate (%)	CoMFA	Change rate (%)	CoMFA	Change rate (%)
0	Target molecule PCB-189	11.517		1.567		5.998	
1	2-ethyl- PCB-189	11.103	-3.59%	1.567	0.00%	5.941	-0.95%
2	2-methyl- PCB-189	11.36	-1.36%	1.388	-11.42%	5.478	-8.67%
3	3- aminomethyl- PCB-189	10.789	-6.32%	1.451	-7.40%	5.435	-9.39%
4	3-methyl-PCB-189	10.937	-5.04%	1.254	-19.97%	5.78	-3.63%
5	4-ethyl-PCB-189	10.815	-6.10%	1.294	-17.42%	5.267	-12.19%
6	4-hydroxymethyl- PCB-189	10.914	-5.24%	1.085	-30.76%	5.159	-13.99%
7	5- butyl- PCB-189	11.27	-2.14%	1.275	-18.63%	5.087	-15.19%
8	5- allyl- PCB-189	11.495	-0.19%	1.362	-13.08%	5.802	-3.27%
9	3'- allyl- PCB-189	10.209	-11.36%	1.318	-15.89%	5.681	-5.29%
10	3'- butylene - PCB -189	10.566	-8.26%	1.249	-20.29%	5.4	-9.97%
11	4'-ethyl- PCB-189	10.746	-6.69%	1.257	-19.78%	4.349	-27.49%
12	4'-hydroxymethyl- PCB-189	11.132	-3.34%	1.261	-19.53%	5.064	-15.57%
13	5'- ethyl- PCB-189	11.221	-2.57%	1.486	-5.17%	5.494	-8.40%
14	5'- silyl - PCB-189	10.608	-7.89%	1.508	-3.77%	5.753	-4.08%
15	2-methyl-3-ethyl-PCB-189	10.222	-11.24%	1.313	-16.21%	4.886	-18.54%
16	2-hydroxymethyl-3-silyl-PCB-189	9.626	-16.42%	1.393	-11.10%	5.879	-1.98%
17	2-silyl-3'-entyl-PCB-189	10.038	-12.84%	1.517	-3.19%	5.006	-16.54%
18	2-methyl-3'-butyl-PCB-189	9.542	-17.15%	1.412	-9.89%	5.611	-6.45%
19	2-methyl fluoride -5-propyl -PCB-189	4.642	-24.35%	1.311	-16.34%	5.758	-4.17%
20	2-hydroxyl-5pentyl-PCB-189	4.872	-20.60%	1.473	-0.06%	4.998	-20.01%

modification at one site, followed by three sites, and then two sites. In summary, for the PCB-189 derivatives all investigated properties showed better improvements with modification at three sites than at one or two sites. Therefore, the effect of modification of PCBs at three sites cannot be ignored.
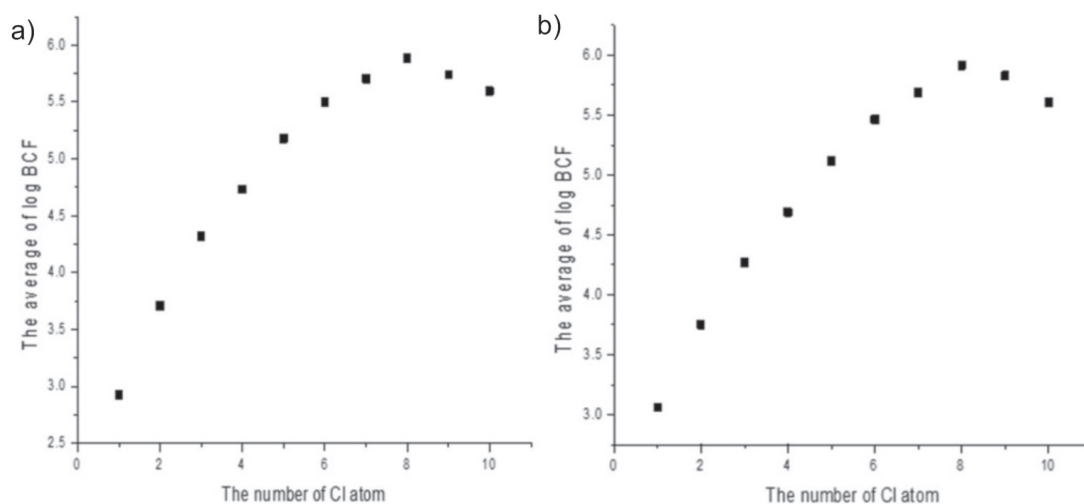

Fig. 7. A plot of Cl atoms versus log BCF values for the CoMFA a) and CoMSIA b) models.

Table 8. Variation in the score function before and after PCB-189 modification.

NO.	Compounds	AHR (5NJ8)		CYPIA (2HI4)		SOD (1UER)		EROD (1JPH)		CAT (3LGD)		GST (5LC2)	
		Dock score	Change rate (%)	Dock score	Change rate (%)	Dock score	Change rate (%)	Dock score	Change rate (%)	Dock score	Change rate (%)	Dock score	Change rate (%)
0	Target molecule PCB-189	77.27		67.65		55.55		57.80		57.07		53.80	
1	2-ethyl-PCB-189	70.11	-9.27	64.36	-4.86	54.59	-1.73	57.19	-1.06	56.52	-0.96	56.92	5.80
2	2-methyl-PCB-189	68.99	-10.72	62.21	-8.04	56.43	1.58	48.25	-16.52	56.84	-0.40	55.47	3.10
3	3-aminomethyl-PCB-189	86.60	12.07	70.40	4.07	61.58	10.86	60.78	5.16	69.16	21.18	60.79	12.99
4	3-methyl-PCB-189	69.68	-9.82	67.10	-0.81	55.80	0.45	57.00	-1.38	63.75	11.70	53.75	-0.09
5	4-ethyl-PCB-189	73.32	-5.11	67.77	0.18	57.17	2.92	62.13	7.49	65.34	14.49	43.75	-18.68
6	4-hydroxymethyl-PCB-189	73.56	-4.80	75.15	11.09	63.25	13.86	64.89	12.27	68.00	19.15	57.00	5.95
7	5-butyl-PCB-189	93.22	20.64	89.52	32.33	63.08	13.56	66.94	15.81	71.97	26.11	62.58	16.32
8	5-allyl-PCB-189	92.19	19.31	81.28	20.15	61.49	10.69	63.62	10.07	70.91	24.25	54.54	1.38
9	3'-allyl-PCB-189	86.34	11.74	81.59	20.61	62.00	11.61	65.07	12.58	68.80	20.55	60.75	12.92
10	3'-butylene-PCB -189	88.22	14.17	88.42	30.70	65.64	18.16	68.45	18.43	72.23	26.56	68.74	27.77
11	4'-ethyl-PCB-189	78.65	1.79	73.37	8.46	60.28	8.51	61.07	5.66	67.03	17.45	52.29	-2.81
12	4'-hydroxymethyl-PCB-189	74.53	-3.55	74.63	10.32	66.00	18.81	66.43	14.93	70.81	24.08	56.30	4.65
13	5'-ethyl-PCB-189	73.51	-4.87	73.70	8.94	59.87	7.78	62.17	7.56	65.01	13.91	50.25	-6.60
14	5'-silyl-PCB-189	75.75	-1.97	66.97	-1.01	54.76	-1.42	56.28	-2.63	61.11	7.08	53.30	-0.93
15	2-methyl-3-ethyl-PCB-189	83.76	8.40	67.45	-0.30	59.39	6.91	61.82	6.96	67.24	17.82	58.63	8.98
16	2-hydroxymethyl-3-silyl-PCB-189	75.86	-1.82	71.67	5.94	58.38	5.09	61.68	6.71	65.30	14.42	60.62	12.68
17	2-silyl-3'-entyl-PCB-189	100.92	30.61	89.41	32.17	59.11	6.41	71.98	24.53	78.56	37.66	75.71	40.72
18	2-methyl-3'-butyl-PCB-189	97.76	26.52	83.76	23.81	65.53	17.97	65.44	13.22	75.50	32.29	70.79	31.58
19	2-methyl fluoride-5-ethylene- PCB-189	79.62	3.04	72.33	6.92	58.31	4.97	62.80	8.65	63.16	10.67	51.37	-4.52
20	2-hydroxyl-5-pentyl-PCB-189	79.08	2.34	75.76	11.99	65.39	17.71	61.85	7.01	70.00	22.66	47.71	-11.32

Molecular Docking Analysis of the Mechanisms of PCB Bioconcentration

Molecular Docking Correlation between Chlorine Atom Substitution and PCB Bioconcentration

The PCBs can be classified into groups from chlorobiphenyl to decachlorobiphenyl isomers according to their number of chlorine atoms. In this study, the average log*BCF* for each class of PCBs was used to study the relationship between the log*BCF* and the number of Cl atoms (Fig. 7). As the number of Cl atoms in the PCB increased, the log*BCF* value also increased. Larger BCF values were indicative of stronger PCB bioconcentration. The bioaccumulation capacities of the PCBs gradually increased as the number of Cl atoms increased.

Analysis of the Causes of Decreased Bioaccumulation after Molecular Modification of PCB-189 Using the Molecular Docking Technique

Molecular docking technique was used to evaluate the affinity between receptor and ligands [43]. The binding effects between the modified molecules and six receptors were analyzed by molecular docking technique (Fig. 8). Although the 20 PCB-189 derivatives were theoretically predicted to have lower bioconcentration than their parent, the scoring functions of some of the PCB-189 derivatives increased after binding (Fig. 9). Therefore, binding of some of the PCB-189 derivatives to the receptors was enhanced (Table 8). The receptors CYPIA, EROD, and GST are involved in exogenous pollutant metabolism in organisms. When some of the derivatives bound to these three receptors, the scoring function increased, which may be because the receptor enhanced metabolic detoxification of these derivatives

and this could reduce their bioconcentration. SOD and CAT remove superoxide anions and hydrogen peroxide in organisms. In this study, some of the derivatives bound to them and their scoring functions increased. This means that the binding abilities of the derivatives are enhanced and the cells of the organism may be susceptible to oxygen damage.

The AHR [44] is more susceptible to activation by aromatic compounds than the other five receptors that we studied, and has associated toxicity. Therefore, AHR (5NJ8) was selected as the main target for the analysis of receptors. We selected three derivatives as examples: the first with modification at one substitution site, and that bound with the receptor and showed the largest decrease in the scoring function; the second with modification at two substitution sites, and could be used to predict the decrease in bioconcentration; and the third with modification at three substitution sites, and could be used to predict potential bioconcentration effects. We analyzed the average bond lengths for non-covalent interactions (hydrogen bonds, electrostatic interactions, hydrophobic interactions, and halogen bonds) between the three derivatives and the receptors. The hydrophilicities and hydrophobicities of amino acid residues at the docking sites were also analyzed to explore changes in the scoring functions of the derivatives after binding to protein receptors.

The parent PCB-189 forms a halogen bond (average bond length of 3.26Å) with the AHR receptor (Fig. 6). By contrast, 2-ethyl-PCB-189, which is formed by modification of PCB-189 with one substituent, does not form a halogen bond with the AHR receptor. Halogen bonds are weak interaction forces similar to weak hydrogen bonds, and can greatly increase the affinity of a receptor to a ligand [39]. Compared with PCB-189, the 2-ethyl-PCB-189 molecule had 28.57% more hydrophilic amino acid residues near the binding
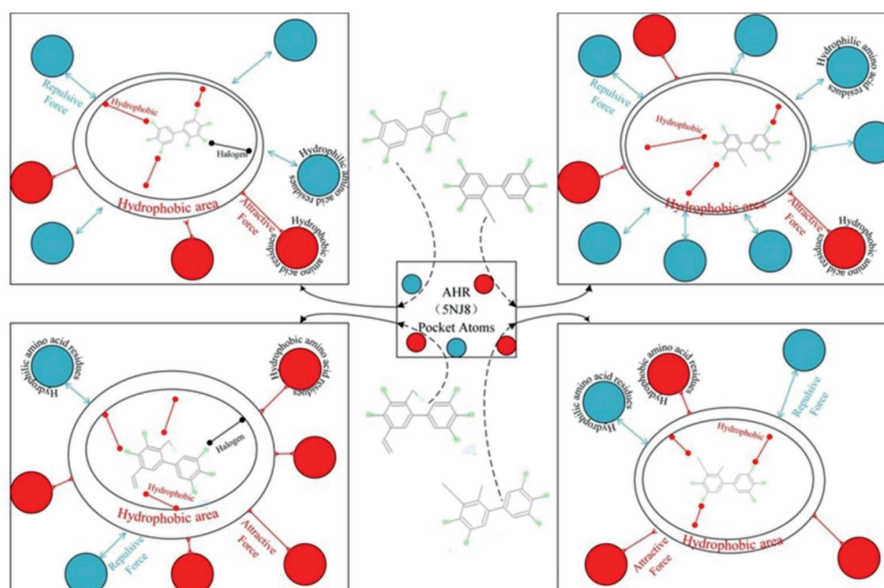

Fig. 8. Ligand interaction diagram of PCB-189 (one, two, or three substituents) docked with AHR (5NJ8) protein.
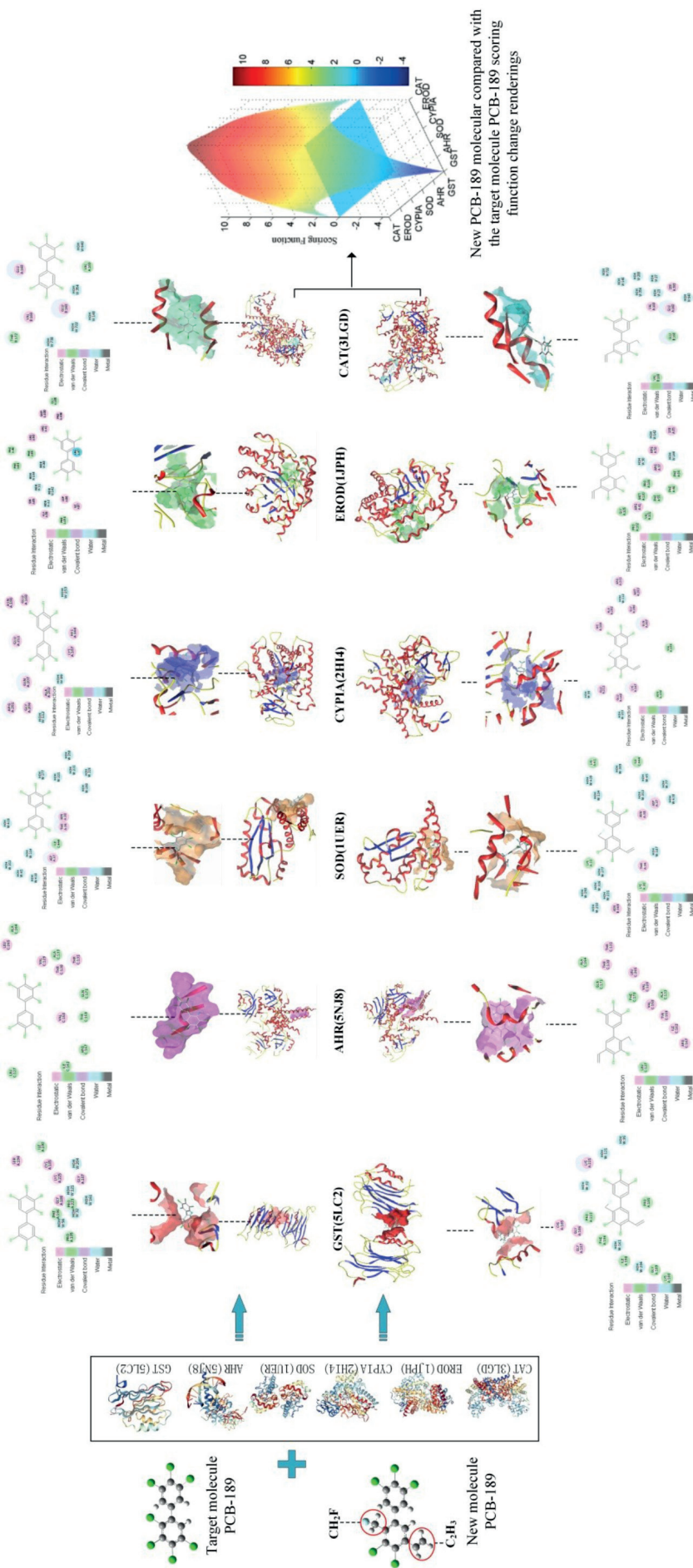
Fig. 9. Binding effect diagram of PCB-189 with one substituents and six protein receptors.

Table 9. Average bond lengths of weak interaction forces near the binding sites of PCB-189 for one, two, and three substituents.

Molecular	Category	AHR (5NJ8) Average bond length (Å)	AHR (5NJ8) Change rate (%)	CYP1A (2HI4) Average bond length (Å)	CYP1A (2HI4) Change rate (%)	SOD (1UER) Average bond length (Å)	SOD (1UER) Change rate (%)	EROD (1JPH) Average bond length (Å)	EROD (1JPH) Change rate (%)	CAT (3LGD) Average bond length (Å)	CAT (3LGD) Change rate (%)	GST (5LC2) Average bond length (Å)	GST (5LC2) Change rate (%)
Target molecular PCB-189	Hydrogen Bonds							3.43					
	Electrostatic					4				4.21			
	Hydrophobic	4.21		4.3		3.88				4.93		4.56	
	Halogen	3.26		2.95								3.11	
2-ethyl-PCB-189	Hydrogen Bonds												
	Electrostatic							3.64	-	3.95	-6.18		
	Hydrophobic	4.49	6.65	4.66	8.37	4.15	6.96	4.44	-1.55	3.61	-26.77	4.58	0.44
	Halogen					3.05	-	3.05	-			3.02	-2.89
2-methyl-3-ethyl-PCB-189	Hydrogen Bonds					2.16	-						
	Electrostatic							4.5	-	3.92	-6.88		
	Hydrophobic	4.29	1.9	4.45	3.49	3.61	-6.96	4.75	5.32	4.92	-0.2	4.39	-3.72
	Halogen							3.42	-			3.12	0.32
2-methyl fluoride-5-ethylene-PCB-189	Hydrogen Bonds					2.26	-						
	Electrostatic			4.04	-			4.51	-	4.24	0.71		
	Hydrophobic	4.3	2.14	4.27	-0.7	3.99	2.84	4.42	-2	4.19	-15.01	4.45	2.41
	Halogen	3.3	1.23			3.19	-					2.72	-12.54

Table 10. Changes in the number of hydrophilic/hydrophobic amino acid residues near the binding sites of target PCB-189 molecular single-, double-, and triple-substituted modifications.

Molecular	Amino acid residues type	AHR (5NJ8)		CYP1A (2HI4)		SOD (1UER)		EROD (1JPH)		CAT (3LGD)		GST (5LC2)	
		Count	Change rate (%)	Count	Change rate (%)	Count	Change rate (%)	Count	Change rate (%)	Count	Change rate (%)	Count	Change rate (%)
Target molecular PCB-189	Hydrophilic	7		6		3		6		3		4	
	Hydrophobic	5				1		6		2		3	
	Neutral	2		3				2				2	
2-ethyl-PCB-189	Hydrophilic	9	28.57	4	-33.33	5	66.67	5	-16.67	3	0.00	4	0.00
	Hydrophobic	5	0.00			2	100	5	-16.67	2	0.00	3	0.00
	Neutral	2	0.00	4	33.33	1	-	1	-50.00	1	-	2	0.00
2-methyl-3-ethyl-PCB-189	Hydrophilic	4	-42.86	5	-16.70	7	133.33	6	0.00	3	0.00	3	-25
	Hydrophobic	5	0.00			0	-100	7	16.67	2	0.00	4	33.33
	Neutral	2	0.00	3	0.00			1	-50.00	1	-	2	0.00
2-methyl fluoride-5- ethylene-PCB-189	Hydrophilic	4	-42.86	6	0.00	7	133.33	6	0.00	3	0.00	4	0.00
	Hydrophobic	6	20.00	2	-	1		5	-16.67	2	0.00	3	0.00
	Neutral	3	50.00	3	0.00			1	-50.00			3	50.00

site (Table 9). The hydrophobic interaction between 2-ethyl-PCB-189 and the AHR receptor had an average bond length that was 6.65% times longer than that of PCB-189 (Table 10), and the strength of the hydrophobic interaction decreased. However, the number of hydrophilic amino acid residues in the vicinity of the binding site increased, which in turn increased the repulsive force of the hydrophobic effect formed by the molecule. Therefore, because of weak interaction forces with the ligand and changes in the number of amino acid residues near the docking site, 2-ethyl-PCB-189 had decreased affinity for the AHR compared with the parent PCB-189.

The 2-methyl-3-ethyl-PCB-189 molecule, which was modified at two substitution sites, did not form a halogen bond on binding to the receptor (Fig. 6). However, the strength of the hydrophobic interaction between the ligand and the receptor for 2-methyl-3-ethyl-PCB-189 was lower than that for the parent PCB-189. In addition, the number of hydrophilic amino acid residues near the 2-methyl-3-ethyl-PCB-189 binding site decreased by 42.86% compared with the parent PCB-189 (Table 9). The presence of hydrophilic amino acid residues in the receptor greatly reduces repulsion of the hydrophobic region formed by 2-methyl-3-ethyl-PCB-189, which increases the affinity of the receptor to the ligand and increases the scoring function (Table 9).

Compared with the parent PCB-189, the average bond length of the halogen bond formed on binding of 2-methyl fluoride-5-ethylene-PCB-189 (substitution at three sites) with the AHR (5NJ8) was only extended by 1.23% (Fig. 6). The change in energy of the halogen bond was small (Table 9). The decrease in affinity of 2-methyl fluoride-5-ethylene-PCB-189 after binding to the receptor was limited. The number of hydrophilic amino acid residues in the vicinity of the 2-methyl fluoride-5-ethylene-PCB-189 binding site decreased by 42.86%, and the number of hydrophobic amino acid residues increased by 20.00% (Table 9). Since the average bond length of the hydrophobic interaction force formed between the ligand and amino acid residue was only extended by 2.14%, the hydrophobic force compared with the parent PCB-189 was decreased less. Therefore, the hydrophobic amino acid residues of the receptor protein greatly enhanced the hydrophobic interaction region formed by the ligand, and the binding affinity of the molecule to the AHR receptor protein and the scoring function increased (Table 9).

The changes in the scoring functions of the three receptors CYPIA (2HI4), CAT (3LGD), and GST (5LC2) were the same as that of the AHR (5NJ8). On binding of the three types of derivatives (one, two, or three substitution sites) to SOD (1UER) and EROD (1JPH), two or more complex forces were generated. At the same time, the number of the hydrophilic/hydrophobic amino acid residues changed, and the molecules docking situation became more complex. If only the average bond length and hydrophobic amino acid residues are used, only general conclusions can be reached from the

molecular docking results. Complex docking situations needs to be analyzed with a combination of molecular docking and other methods.

In summary, although various weak interaction forces greatly influence the affinity of the receptor for the ligand, the number of hydrophilic/hydrophobic amino acid residues near the binding site cannot be ignored. These residues affect the stability of the binding of the ligand with the receptor. Molecular docking can be used to screen for molecules with low bioconcentration after molecular design is performed for specific receptors. In this study, molecular docking screening identified 11 derivatives with low bioconcentration. Their scoring functions after docking with more than half of the receptors changed by less than 10%, and the affinities of the receptors for the ligands did not change significantly. These data provide a reference for theoretical studies of other PCBs and their derivatives in the future.

Conclusions

In this paper, multiple-site molecular modification of the dioxin PCB-189 was performed using 3D-QSAR and fractional factorial experiments to reduce bioconcentration. Among these derivatives, six that were modified at two or three substitution sites were decoupled from the bioconcentration effect. Therefore, the effects of third-order interactions on the POPs properties of PCBs cannot be ignored.

Molecular docking results indicated that the main factors affecting the affinity of the receptor to the ligand were the relationship between the number of hydrophobic amino acids near the binding site and the non-covalent interaction forces generated when the ligands docked. In addition, molecular docking technology can be used for further screening of PCBs with low bioconcentration, and to provide reference data for the design of PCBs with low bioconcentration.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. AFE S. Toxicology, structure-function relationship, and human and environmental health impacts of

polychlorinated biphenyls: progress and problems. Environmental Health Perspectives, **100** (4), 259, **1993**.

2. BORJA J., TALEON D.M., AURESENIA J., GALLARDO S. Polychlorinated biphenyls and their biodegradation. Process Biochemistry, **40** (6), 1999, **2005**.

3. FU J., MAI B., SHENG G., ZHANG G., WANG X., PENG P., XIAO X., RAN R., CHENG F., PENG X., WANG Z., TANG U.W. Persistent organic pollutants in environment of the Pearl River Delta, China: an overview. Chemosphere, **52** (9), 1411, **2003**.

4. CHEN Y., CAI X., JIANG L., LI Y. Prediction of octanol-air partition coefficients for polychlorinated biphenyls (PCBs) using 3D-QSAR models. Ecotoxicology and Environmental Safety, **124**, 202, **2016**.

5. JÖNSSON A., GUSTAFSSON O., AXELMAN J., SUNDBERG H. Global accounting of PCBs in the continental shelf sediments. Environmental Science and Technology, **37** (2), 245, **2003**.

6. MUIR D.C.G., SEGSTRO M.D., WELBOURN P.M., TOOM D., EISENREICH S.J., MACDONALD C.R., WHELPDALE D.M. Patterns of accumulation of airborne organochlorine contaminants in lichens from the Upper Great Lakes Region of Ontario. Environmental Science and Technology, **27** (6), 1201, **1993**.

7. HEGSETH M.N., CAMUS L., HELGASON L.B., BOCCHETTI R., GABRIELSEN G.W., REGOLI F. Hepatic antioxidant responses related to levels of PCBs and metals in chicks of three Arctic seabird species. Comparative Biochemistry and Physiology Part C, **154** (1), 28, **2011**.

8. MARCO V., STEFANO Z., ELENA A., CARLO B., ANDREA G., ROSSANO P. Polychlorinated biphenyls (PCBs) and polybrominated diphenyl ethers (PBDEs) in Antarctic ice-free areas: Influence of local sources on lakes and soils. Microchemical Journal, **120**, 26, **2015**.

9. GULAN L., MILENKOVIC B., ZEREMSKI T., MILIC G., VUCKOVIC B. Persistent organic pollutants, heavy metals and radioactivity in the urban soil of Priština City, Kosovo and Metohija. Chemosphere, **171**, 415, **2016**.

10. 10. GASIC B., Macleod M., KLANOVA J., SCHERINGER M., ILIC P., LAMMEL G., PAJOVIC A., BREIVIK K., HOLOUBEK I., HUNGERBÜHLER K. Quantification of sources of PCBs to the atmosphere in urban areas: a comparison of cities in North America, Western Europe and former Yugoslavia. Environmental Pollution, **158** (10), 3230, **2010**.

11. DROUILLARD K.G., JEZDIC I., O'ROURKE S.M., GEWURTZ S.B., RAESIDE A.A., LEADLEY T.A., DRCA P., HAFFNER G.D. Spatial and temporal variability of PCBs in Detroit River water assessed using a long-term biomonitoring program. Chemosphere, **90** (1), 95, **2013**.

12. ZOHAIR A., SALIM A.B., SOYIBO A.A., BECK A.J. Residues of polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs) and organochlorine pesticides in organically-farmed vegetables. Chemosphere, **63** (4), 541, **2006**.

13. LUZARDO O.P., RODRÍGUEZ-HERNÁNDEZ A., QUESADA-TACORONTE Y., RUIZ-SUÁREZ N., ALMEIDA-GONZÁLEZ M., HENRÍQUEZ-HERNÁNDEZ L.A., ZUMBADO M., BOADA L.D. Influence of the method of production of eggs on the daily intake of polycyclic aromatic hydrocarbons and organochlorine contaminants: an independent study in the Canary Islands (Spain). Food and Chemical Toxicology, **60** (10), 455, **2013**.

14. BERNTSSEN M.H.G., LUNDEBYE A., TORSTENSEN B.E. Reducing the levels of dioxins and dioxin-like PCBs in farmed Atlantic salmon by substitution of fish oil with vegetable oil in the feed. Aquaculture Nutrition, **11** (3), 219, **2015**.

15. LUZARDO O.P., RUIZ-SUÁREZ N., HENRÍQUEZ-HERNÁNDEZ L.A., VALERÓN P.F., CAMACHO M., ZUMBADO M., BOADA L.D. Assessment of the exposure to organochlorine pesticides, PCBs and PAHs in six species of predatory birds of the Canary Islands, Spain. Science of the Total Environment, **472**, 146, **2014**.

16. JOTAKI T., FUKATA H., MORI C. Confirmation of polychlorinated biphenyl (PCB) distribution in the blood and verification of simple quantitative method for PCBs based on specific congeners. Chemosphere, **82** (1), 107, **2011**.

17. ARREBOLA J.P., FERNANDEZ M.F., PORTA M., ROSELL J., MARTINEZ D.L.O.R., OLEA N. Multivariate models to predict human adipose tissue PCB concentrations in Southern Spain. Environment International, **36** (7), 705, **2010**.

18. NAKAO T., AOZASA O., OHTA S., MIYATA H. Assessment of human exposure to PCDDs, PCDFs and Co-PCBs using hair as a human pollution indicator sample I: Development of analytical method for human hair and evaluation for exposure assessment. Chemosphere, **48** (8), 885, **2002**.

19. SHEN H., DING G., WU Y., PAN G., ZHOU X., HAN J., LI J., WEN S. Polychlorinated dibenzo-p-dioxins/furans (PCDD/Fs), polychlorinated biphenyls (PCBs), and polybrominated diphenyl ethers (PBDEs) in breast milk from Zhejiang, China. Environment International, **42**, 84, **2012**.

20. COVACI A., VOORSPOELS S., ROOSENS L., JACOBS W., BLUST R., NEELS H. Polybrominated diphenyl ethers (PBDEs) and polychlorinated biphenyls (PCBs) in human liver and adipose tissue samples from Belgium. Chemosphere, **73** (2), 170, **2008**.

21. CHEN X., CHEN J.S., ZHANG L., LI J.G., YAO L., SELF S.G., SUN X., TANG N.J. Levels of PCDDs, PCDFs and dl-PCBs in the blood of childbearing-aged women living in the vicinity of a chemical plant in Tianjin: a primary study. Chemosphere, **118** (1), 1, **2015**.

22. GAUGER K.J., KATO Y., HARAGUCHI K., LEHMLER H.J., ROBERTSON L.W., BANSAL R., PAJOVIC A., BREIVIK K., HOLOUBEK I., HUNGERBÜHLER K. Polychlorinated biphenyls (PCBs) exert thyroid hormone-like effects in the fetal rat brain but do not bind to thyroid hormone receptors. Environmental Health Perspectives, **112** (5), 516, **2004**.

23. MEEKER J.D., HAUSER R. Exposure to Polychlorinated Biphenyls (PCBs) and Male Reproduction. Systems Biology in Reproductive Medicine, **56** (2), 122, **2010**.

24. ARNOT J.A., GOBAS F.A. A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms. Environmental Reviews, **14** (4), 257, **2006**.

25. IVANCIUC T., IVANCIUC O., KLEIN D.J. Modeling the bioconcentration factors and bioaccumulation factors of polychlorinated biphenyls with posetic quantitative super-structure/activity relationships (QSSAR). Molecular Diversity, **10** (2), 133, **2006**.

26. MELO E.B. A new quantitative structure-property relationship model to predict bioconcentration factors of polychlorinated biphenyls (PCBs) in fishes using E-state

index and topological descriptors. Ecotoxicology and Environmental Safety, **75** (1), 213, **2012**.

27. LIU H., LIU H., SUN P., WANG Z. QSAR studies of bioconcentration factors of polychlorinated biphenyls (PCBs) using DFT, PCS and CoMFA. Chemosphere, **114**, 101, **2014**.

28. XU Z., CHEN Y., QIU Y., GU W., LI Y. Prediction of Stability for Polychlorinated Biphenyls in Transformer Insulation Oil Through Three-dimensional Quantitative Structure-activity Relationship Pharmacophore Model and Full Factor Experimental Design. Chemical Research in Chinese Universities, **32** (3), 348, **2016**.

29. LI M., WEI D., ZHAO H., DU Y. Genotoxicity of quinolones: substituents contribution and transformation products QSAR evaluation using 2D and 3D models. Chemosphere, **95** (1), 220, **2014**.

30. GU W., CHEN Y., ZHANG L., LI Y. Prediction of octanol-water partition coefficient for polychlorinated naphthalenes through three-dimensional QSAR models. Human and Ecological Risk Assessment, **23** (1), 40, **2016**.

31. SALAHINEJAD M., GHASEMI J.B. 3D-QSAR studies on the toxicity of substituted benzenes to Tetrahymena pyriformis: CoMFA, CoMSIA and VolSurf approaches. Ecotoxicology and Environmental Safety, **105** (1), 128, **2014**.

32. RAMACHANDRAN M., LAAKSO A., HARBAUGH R.E., RAGHAVAN M.L. On the role of modeling choices in estimation of cerebral aneurysm wall tension. Journal of Biomechanics, **45** (16), 2914, **2012**.

33. GU W., CHEN Y., LI Y. Attenuation of the Atmospheric Migration Ability of Polychlorinated Naphthalenes (PCN-2) Based on Three-dimensional QSAR Models with Full Factor Experimental Design. Bulletin of Environmental Contamination and Toxicology, **99** (2), 276, **2017**.

34. WANG X., CHU Z., YANG J., LI Y. Pentachlorophenol molecule design with lower bioconcentration through 3D-QSAR associated with molecule docking. Environmental Science and Pollution Research, **24** (18), 25114, **2017**.

35. ZHAO Y., GU W., LI Y. Molecular design of 1,3,5,7-TetraCN derivatives with reduced bioconcentration using 3D-QSAR modeling, full factorial design, and molecular docking. Journal of Molecular Graphic and Modelling, **84**, 197, **2018**.

36. WANG X., GU W., GUO E., CUI C., LI Y. Assessment of long-range transport potential of polychlorinated Naphthalenes based on three-dimensional QSAR models. Environ Science and Pollution Research, **24**, 14802, **2017**.

37. ZENG X.L., QU R.J., FENG M.B., CHEN J.S., WANG L., WANG Z.Y. Photodegradation of Polyfluorinated Dibenzo-p-Dioxins (PFDDs) in Organic Solvents: Experimental and Theoretical Studies. Environmental Science and Technology, 50 (15), 8128, 2016.

38. LI F., LI X., LIU X., ZHANG L., YOU L., ZHAO J., WU H. Docking and 3D-QSAR studies on the receptor binding affinities of polychlorinated biphenyls (PCBs), dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs). Environmental Toxicology and Pharmacology, **32** (3), 478, **2011**.

39. JOHNSON E.R., KEINAN S., MORISÁNCHEZ P., CONTRERAS-GARCÍA J., COHEN A.J., YANG W. Revealing Noncovalent Interactions. Journal of the American Chemical Society, **132** (18), 6498, **2010**.

40. KELLY B.C., IKONOMOU M.G., BLAIR J.D., MORIN A.E., GOBAS F.A. Food Web-Specific Biomagnification of Persistent Organic Pollutants. Science, **317** (5835), 236, **2007**.

41. GU W., ZHAO Y., LI Q., LI Y. Environmentally friendly polychlorinated naphthalenes (PCNs) derivatives designed using 3D-QSAR and screened using molecular docking, density functional theory and health-based risk assessment. Journal of Hazardous Materials, **363**, 316, **2018**.

42. DOLFING J., Harrison B.K. Gibbs free energy of formation of halogenated aromatic compounds and their potential role as electron acceptors in anaerobic environments. Environmental Science and Technology, **26** (11), 2213, **1992**.

43. ZHAO X.H., QIU Y.L., JIANG L., LI Y. Analysis of Affinity Energy Between Biphenyl Dioxygenase and Polychlorinated Biphenyls Using Molecular Docking. Chemical Research in Chinese Universities, **35** (2), 325, **2019.**

44. MANDAVIA C. TCDD-induced activation of aryl hydrocarbon receptor regulates the skin stem cell population. Medical Hypotheses, **84** (3), 204, **2015**.