Chinese Science Bulletin

Quantitative structure-retention relationship for polychlorinated dibenzofurans based on molecular interaction field analysis

ZHAO JinSong† & YU ShuXia

Key Laboratory of Subtropical Agriculture and Environment, Ministry of Agriculture, College of Resources and Environment, Huazhong Agricultural University, Wuhan 430070, China

A new quantitative structure-retention relationship (QSRR) model is developed for polychlorinated dibenzofurans (PCDFs) based on molecular interaction field (MIF) analysis. The MIF of all 135 PCDFs is calculated using DRY, C1= and C3 probe, characterizing the hydrophobic and steric interaction between PCDFs and different groups of stationary phase. Then QSRR model is constructed by multiblock partial least squares (MBPLS), and the significance of each block is evaluated by the block importance in the prediction (BIP) method. The model used for prediction is statistically significant, with calibration and cross-validation correlation coefficients 0.9990 and 0.9980 respectively, and relative error less than 1.0%. The results of MBPLS and BIP show that the steric properties have dominant influence on the retention behavior of PCDFs, and then the hydrophobic effects.

polychlorinated dibenzofurans, molecular interaction field, quantitative structure-retention relationship, retention index

Polychlorinated dibenzofurans (PCDFs) are a class of persistent organic pollutants (POPs), and present globally in many of environmental matrixes[1,2]. PCDFs are released into environments as by-products of various thermal and chemical processes such as production of steel and copper, gasoline, coal and municipal waste combustion, and as intermediates in the fabrication of chlorophenol-based products[2,3]. PCDFs can induce various toxicological responses including immunotoxicity, carcinogenicity, and adverse effects on reproduction, development and endocrine functions[4,5]. Therefore, their environmental fates have received scientific concerns.

There are 135 PCDFs in total, and the nomenclature for the numbering of the chlorine positions is given in Figure 1. Although they share similar structural and chemical characteristics, these compounds vary greatly in their biological properties and toxicities[5]. The accurate identification of each isomer plays important roles in quantitative analysis and environmental risk assessment[6].

Though various techniques have been used for PCDFs analysis, gas chromatography remains the primary separation technique[7]. Gas chromatographic retention parameters, which are typically used to qualitatively identify PCDFs congeners[7], are determined by electronic, chemical and structural properties of compounds[8]. Therefore, descriptors encoding significant

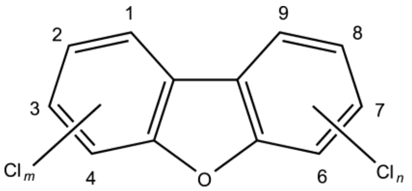

Figure 1 General formula and atom numbering of PCDFs.

Received January 12, 2009; accepted March 8, 2009
doi: 10.1007/s11434-009-0552-1

† Corresponding author (email: jszhao@mail.hzau.edu.cn)

Supported by the National Natural Science Foundation of China (Grant No. 40601085), and National Key Technology R&D Program (Grant No.2008BADA7B01)

Citation: Zhao J S, Yu S X. Quantitative structure-retention relationship for polychlorinated dibenzofurans based on molecular interaction field analysis. Chinese Sci Bull, 2009, 54: 4644—4650, doi: 10.1007/s11434-009-0552-1

structural information such as topology, geometry, and electronic environment can be used to model the retention behavior of compounds[8,9].

The quantitative structure-retention relationship (QSRR) method used to interpret the separation from a molecular structure viewpoint, is of great analytical value[7]. Many QSRR models for PCDFs have been reported previously using various descriptors[7,8,10,11]. However, there are no studies attempting to develop QSRR models using the molecular interaction field (MIF) as descriptors.

MIF is the interaction between a molecule and a probe, which is moved through a regular grid of points in a specific region around the target molecule. At each point, the interaction energy is calculated as the sum of Lennard-Jones, hydrogen bond, and electrostatic interactions. For specific probes, entropic contribution is also included [12–14]. According to the presented studies, the differences of the chemical properties are often related to the differences of MIF.

The purpose of this study is to develop a QSRR model based on MIF, then, to explore the retention mechanisms of the PCDFs separation.

1 Materials and methods

1.1 Data sources

The gas chromatographic Kováts retention indexes (RI) of 115 out of 135 PCDFs on column DB-5 are obtained from ref. [7] (Table 1). The column parameter for DB-5 is a 30 m×0.25 mm×0.25 μm fused silica capillary column, and the stationary phase is methyl polysiloxane with 5% phenyl substitution. The column is temperature programmed from 175℃ with an initial hold of 1 min to 300℃ at a rate of 6℃ per min. Helium is used as a carrier gas at 48263.299 Pa head pressure.

1.2 Computation of MIF

The geometry of 135 PCDFs is optimized using MOPAC 6.0[15] with keywords: PM3 EF LET GNORM=0.001. All molecules are then superimposed on the dibenzofurans rings with the most toxic 2,3,7,8-TCDF[16] as a reference.

MIF is obtained with GRID version 22a[17]. For PCDFs interacting with stationary phase of column DB-5, it is reasonable to use the DRY probe to represent hydrophobic interactions, and the C1= (sp^2, phenyl) and

Table 1 The 135 PCDFs congeners, experimental and calculated retention index

No.	Congener a)	CAS No.	RI$_{\text{exp}}$ b)	RI$_{\text{cal}}$ c)	RE% d)
1	1	84761-86-4	1739	1740	–0.06
2	2	51230-49-0	1749	1764	–0.86
3	3	25074-67-3	1749	1764	–0.86
4	4*	74992-96-4	1760	1758	0.11
5	1,2	64126-85-8	1934	1933	0.05
6	1,3	94538-00-8	1884	1904	–1.06
7	1,4*	94538-01-9	1913	1916	–0.16
8	1,6	74992-97-5	–	1917	–
9	1,7	94538-02-0	1910	1910	0.00
10	1,8	81638-37-1	1925	1916	0.47
11	1,9	70648-14-5	1975	1951	1.22
12	2,3	64126-86-9	1939	1967	–1.44
13	2,4	24478-74-8	1912	1927	–0.78
14	2,6	60390-27-4	1946	1936	0.51
15	2,7	74992-98-6	1930	1940	–0.52
16	2,8	5409-83-6	1935	1934	0.05
17	3,4	94570-83-9	1959	1978	–0.97
18	3,6*	74918-40-4	1944	1935	0.46
19	3,7	58802-21-4	1930	1921	0.47
20	4,6	64560-13-0	1953	1937	0.82
21	1,2,3	83636-47-9	2113	2135	–1.04
22	1,2,4	24478-73-7	2085	2107	–1.06
69	1,3,6,8	71998-72-6	2227	2246	–0.85
70	1,3,6,9	83690-98-6	2296	2280	0.70
71	1,3,7,8*	57117-35-8	2263	2277	–0.62
72	1,3,7,9	64560-17-4	2273	2276	–0.13
73	1,4,6,7	66794-59-0	2288	2285	0.13
74	1,4,6,8	82911-58-8	2242	2255	–0.58
75	1,4,6,9	70648-19-0	2314	2295	0.82
76	1,4,7,8	83704-29-4	2290	2289	0.04
77	1,6,7,8	83704-33-0	2308	2300	0.35
78	2,3,4,6*	83704-30-7	2339	2332	0.30
79	2,3,4,7	83704-31-8	2337	2329	0.34
80	2,3,4,8	83704-32-9	2340	2322	0.77
81	2,3,6,7	57117-39-2	2354	2352	0.08
82	2,3,6,8*	57117-37-0	2297	2303	–0.26
83	2,3,7,8	51207-31-9	2338	2327	0.47
84	2,4,6,7	57117-38-1	2305	2311	–0.26
85	2,4,6,8	58802-19-0	2254	2263	–0.40
86	3,4,6,7*	57117-40-5	2362	2311	2.16
87	1,2,3,4,6*	83704-47-6	2496	2498	–0.08
88	1,2,3,4,7*	83704-48-7	2495	2492	0.12
89	1,2,3,4,8*	67517-48-0	2508	2492	0.64
90	1,2,3,4,9	83704-49-8	–	2555	–

(To be continued on the next page)

Continued

No.	Congener[a)]	CAS No.	RI_{exp}[b)]	RI_{cal}[c)]	RE%[d)]	No.	Congener	CAS No.	RI_{exp}	RI_{cal}	RE%
23	1,2,6	64560-15-2	2125	2112	0.61	91	1,2,3,6,7	57117-42-7	2540	2514	1.02
24	1,2,7	83704-37-4	2109	2104	0.24	92	1,2,3,6,8	83704-51-2	–	2479	–
25	1,2,8	83704-34-1	2129	2108	0.99	93	1,2,3,6,9	83704-52-3	2546	2549	–0.12
26	1,2,9	83704-38-5	–	2164	–	94	1,2,3,7,8	57117-41-6	2507	2495	0.48
27	1,3,4	82911-61-3	2088	2115	–1.29	95	1,2,3,7,9	83704-53-4	–	2533	–
28	1,3,6*	83704-39-6	2072	2072	0.00	96	1,2,3,8,9	83704-54-5	2593	2581	0.46
29	1,3,7	64560-16-3	2057	2061	–0.19	97	1,2,4,6,7*	83704-50-1	2465	2480	–0.61
30	1,3,8	76621-12-0	2070	2079	–0.43	98	1,2,4,6,8	69698-57-3	–	2448	–
31	1,3,9	83704-40-9	2124	2113	0.52	99	1,2,4,6,9	70648-24-7	2497	2512	–0.60
32	1,4,6	82911-60-2	2094	2081	0.62	100	1,2,4,7,8	58802-15-6	–	2481	–
33	1,4,7	83704-41-0	2086	2076	0.48	101	1,2,4,7,9	71998-74-8	2479	2496	–0.69
34	1,4,8	64560-14-1	2100	2090	0.48	102	1,2,4,8,9*	70648-23-6	2559	2557	0.08
35	1,4,9	70648-13-4	2151	2130	0.98	103	1,2,6,7,8*	69433-00-7	2521	2494	1.07
36	1,6,7	83704-46-5	2125	2120	0.24	104	1,2,6,7,9	70872-82-1	–	2544	–
37	1,6,8	82911-59-9	2082	2080	0.10	105	1,3,4,6,7	83704-36-3	2469	2462	0.28
38	1,7,8	58802-18-9	2111	2113	–0.09	106	1,3,4,6,8	83704-55-6	–	2456	–
39	2,3,4	57117-34-7	2148	2153	–0.23	107	1,3,4,6,9	70648-15-6	–	2495	–
40	2,3,6	57117-33-6	2141	2145	–0.19	108	1,3,4,7,8	58802-16-7	2469	2489	–0.81
41	2,3,7	58802-17-8	2134	2140	–0.28	109	1,3,4,7,9	70648-20-3	2473	2493	–0.81
42	2,3,8*	57117-32-5	2132	2133	–0.05	110	1,3,6,7,8	70648-21-4	2467	2466	0.04
43	2,4,6	58802-14-5	2101	2096	0.24	111	1,4,6,7,8	83704-35-2	2476	2476	0.00
44	2,4,7	83704-42-1	2099	2102	–0.14	112	2,3,4,6,7	57117-43-8	2555	2539	0.63
45	2,4,8	54589-71-8	2097	2101	–0.19	113	2,3,4,6,8	67481-22-5	2495	2490	0.20
46	2,6,7	83704-45-4	2151	2150	0.05	114	2,3,4,7,8	57117-31-4	2551	2538	0.51
47	3,4,6	83704-43-2	2152	2144	0.37	115	1,2,3,4,6,7	79060-60-9	2706	2703	0.11
48	3,4,7	83704-44-3	2150	2132	0.84	116	1,2,3,4,6,8	69698-60-8	2650	2667	–0.64
49	1,2,3,4	24478-72-6	2310	2323	–0.56	117	1,2,3,4,6,9	91538-83-9	–	2736	–
50	1,2,3,6	83704-21-6	2307	2311	–0.17	118	1,2,3,4,7,8	70648-26-9	2708	2708	0.00
51	1,2,3,7*	83704-22-7	2294	2302	–0.35	119	1,2,3,4,7,9	91538-84-0	2720	2721	–0.04
52	1,2,3,8*	62615-08-1	2307	2304	0.13	120	1,2,3,4,8,9	92341-07-6	–	2768	–
53	1,2,3,9	83704-23-8	2369	2369	0.00	121	1,2,3,6,7,8	57117-44-9	–	2684	–
54	1,2,4,6	71998-73-7	2264	2273	–0.40	122	1,2,3,6,7,9	92341-06-5	–	2749	–
55	1,2,4,7	83719-40-8	2264	2269	–0.22	123	1,2,3,6,8,9	75198-68-8	–	2762	–
56	1,2,4,8*	64126-87-0	2274	2283	–0.40	124	1,2,3,7,8,9	72918-21-9	–	2766	–
57	1,2,4,9	83704-24-9	2335	2342	–0.30	125	1,2,4,6,7,8	67562-40-7	–	2669	–
58	1,2,6,7	83704-25-0	2329	2315	0.60	126	1,2,4,6,7,9	75627-02-0	–	2711	–
59	1,2,6,8*	83710-07-0	2281	2274	0.31	127	1,2,4,6,8,9*	69698-59-5	2686	2724	–1.41
60	1,2,6,9	70648-18-9	2364	2344	0.85	128	1,3,4,6,7,8	71998-75-9	–	2677	–
61	1,2,7,8	58802-20-3	2322	2307	0.65	129	1,3,4,6,7,9	92341-05-4	–	2680	–
62	1,2,7,9	83704-26-1	2341	2329	0.51	130	2,3,4,6,7,8	60581-34-5	2748	2726	0.80
63	1,2,8,9*	70648-22-5	2406	2376	1.25	131	1,2,3,4,6,7,8	67562-39-4	2898	2895	0.10
64	1,3,4,6*	83704-27-2	2262	2282	–0.88	132	1,2,3,4,6,7,9*	70648-25-8	2913	2935	–0.76
65	1,3,4,7	70648-16-7	2257	2270	–0.58	133	1,2,3,4,6,8,9	69698-58-4	2922	2948	–0.89
66	1,3,4,8	92341-04-3	2276	2290	–0.62	134	1,2,3,4,7,8,9	55673-89-7	2986	2981	0.17
67	1,3,4,9	83704-28-3	2325	2329	–0.17	135	1,2,3,4,6,7,8,9	39001-02-0	3147	3167	–0.64
68	1,3,6,7	57117-36-9	2272	2274	–0.09						

a) Dataset for testing indicated by *; b) RI_{exp}, experimental retention index; and “–”, not available; c) RI_{calc}, calculated retention index; d) RE%, relative error.

C3 (sp^3, methyl) probes to represent the steric interaction[13]. A grid spacing of 0.1 nm is used with an extended 0.5 nm beyond all molecules. This results in a box with dimension of 2.1 nm×1.7 nm×1.1 nm, thus giving 3927 probe-target interactions for each compound. To produce a more symmetrical distribution of energy values, a cutoff of 20.6 kJ mol^{-1} is applied [14].

1.3 Statistical analysis

The MIF data are pretreated before statistical modeling. Firstly, those data with absolute values smaller than 1.648 kJ mol^{-1} are zeroed; secondly, any variables with a standard deviation below 0.1 are removed [14]. The threshold values are obtained by experiments based on a $L_{25}(5^6)$ orthogonal arrays. After pretreatment, the data sets for final modeling contain 1101, 1007, 319 variables, for C1=, C3 and DRY probes, respectively.

The multiblock partial least squares (MBPLS)[18] method is used to establish the QSRR model, for the MIF matrix consists of three blocks from the C1=, C3 and DRY probes. In addition to conventional PLS[19] with the same capability in building a QSRR model and in predicting, MBPLS can also be used to determine the contribution of each block to the QSRR model, and provide the chance for further interpretation of mechanisms associated with the QSRR model. The algorithm of super score deflation of **Y** is used to implement the MBPLS analysis, because it overcomes the disadvantages of block score deflation of **X** method in prediction[20,21], and of super score deflation of **X** in interception[21,22].

All 115 PCDFs are divided randomly into two groups, training set and testing set. The QSRR model is built using training set of 92 PCDFs. And its predictive ability is evaluated by leave-more-out cross-validation, using 10 groups of approximately the same size in which the objects are assigned randomly. The whole procedure is repeated 46 times. This cross-validation procedure provides a safer alternative to the more widely preferred leave-one-out (LOO) method and gives more conservative results: a smaller cross-validated squared correlation coefficient (q^2) and a higher standard deviation of error of predictions (SDEP)[14]. In addition to the internal validation, 23 PCDFs in the testing set are used as an external validation set.

The MBPLS program is coded in R language with version 2.8.0[23].

2 Results and discussion

2.1 QSRR model

The QSRR model is obtained from the training set using MBPLS. The predictive ability of the models is evaluated by internal validation and by external validation using the testing set. As shown in Figure 2, the amount of variance explained by the model increases, while the corresponding predicted residual sum of squares (PRESS) decreases rapidly. The model reaches stabilization in its predictive ability after 7 latent variables (*K*) and is not improved significantly by the addition of further components. Thus the final model is built with a *K* of 7.

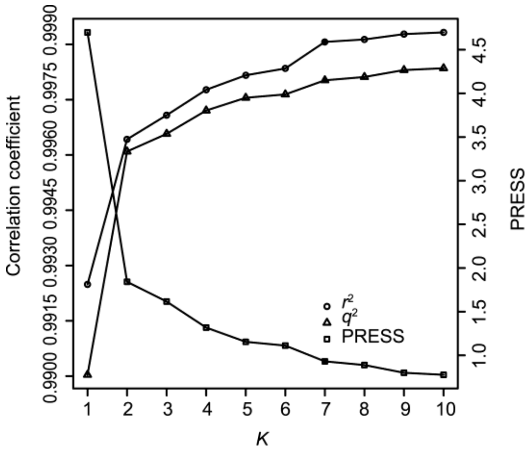

Figure 2 r^2, q^2 and PRESS versus different *K* for the QSRR model.

The quality of the final model in fitting is expressed as the squared correlation coefficient (r^2, 0.9990), and as the standard deviation of error of calibration (SDEC, 8.288 or 0.364% of the mean). The result of the internal validation is expressed as the cross-validated squared correlation coefficient (q^2, 0.9980), and as the standard deviation of error of predictions (SDEP, 12.938 or 0.568%). The external SDEP (11.388 or 0.500%) expresses the ability of the model to predict the RI of the compounds in the external testing set. The results of MBPLS analysis show that the QSRR model has a significant ability to fit and predict.

Predicted values, as well as the relative error for all 135 PCDFs are listed in Table 1. Since the experimental error of the DB-5 RI column is known to be approximately 0.3%, thus the mean error 0.310% of this MBPLS model is considered excellent. It agrees with previous studies on the same dataset[7,8,11]. The scatter plot of RI (Figure 3) shows excellent agreement between the experimental data and the calculated values.

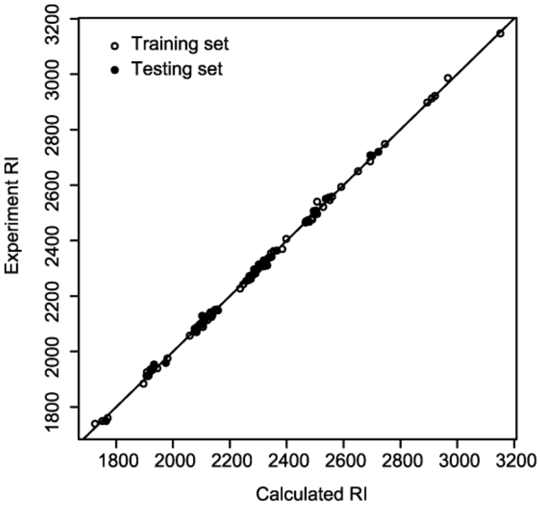

Figure 3 Scatter plot showing the calculated versus experimental RI of the QSRR model.

When compared with traditional QSRR model[7,8,11], the QSRR model developed in this study can give more information than prediction, e.g., the retention mechanisms. The model is built on not only the chemical structure descriptors, but also the interaction between chemicals and the matrix that those chemicals presented, e.g., the stationary in this study. With proper probes, new MIF can be calculated, based on that novel QSRR model would be developed, when the stationary changed. It is very important to QSRR models used for prediction and mechanism research.

Polychlorinated dibenzo-*p*-dioxins (PCDDs) have similar physiochemical properties and environmental effects with PCDFs[1,4,5]. It seems that RI of PCDDs would be predicted by the QSRR models based on the MIF of PCDFs. Three scenes should be taken into consideration. Firstly, the RI and MIF of PCDD are not included into the training set, and just use the presented QSRR model to predict. The results are very poor. The variation between the RI of PCDD and PCDF is not considered during construction of QSRR model is the main reason. Secondly, the IR and MIF calculated alone are included in the model construction. And, thirdly, the MIF of PCDDs is calculated along with that of PCDFs. In fact, the latter two situations are only different in the position of aligned chemical structures. The third way is preferred, for the structural alignment is critical to the calculation of MIF. Note that there is significant difference between the structures of PCDDs and PCDFs, thus, the QSRR based on data sets of PCDDs and PCDFs are inferior to the one presented in this study. The detailed comparison is given in the supplemental material.

2.2 Mechanisms of chromatographic retention

The above QSRR model has presented a significant predictability of MIF descriptors to RI. It implies that the interactions between PCDFs and the methyl, phenyl on stationary phase and the hydrophobic characteristics, are significant during the separation of PCDFs in DB-5 column. However, it does not give a clear perspective on which is more important. In fact, it is a problem of variable selection.

The block importance in the prediction (BIP)[18], which generalizes the variable importance for projection (VIP) coefficients wildly used in PLS regression[19], can be used here to interpret the importance of different probes in the interaction of separation. It measures the contribution of each block in the model to the variance explanation of response.

In this study, the BIP at K = 7 is 1.120, 1.084, and 0.755 for C1=, C3 and DRY probes, respectively. It indicates that the MIF descriptors of C1= probe has the highest contribution to RI, but just slightly larger than that of C3 probe. Considering the absolute content of methyl on the stationary phase, the interaction between PCDFs and C3 probes may be more important. And the MIF of DRY is much smaller in contribution when compared with the other probes.

In order to give a more clear perspective on the interactions between of substructure of each molecule and stationary phase, the coefficients of each grid from every probe are projected to molecular structures. In the coefficients contour, the light color indicates the regions where increasing MIF would give positive contribution to IR, while, the deep color indicates the regions where increasing MIF would give negative contribution to IR.

C1= and C3 probes are intended to characterize the interaction between solute and stationary phase in DB-5 column. Figure 4 shows the coefficients contour map of C1= probe (that of C3 probe not shown here, because of similar pattern), and the positive coefficients are on the both sides of benzene ring plane. It reveals that as the polarity of the molecule increases, i.e., the number of chlorine atom, RI increases. The results demonstrate that

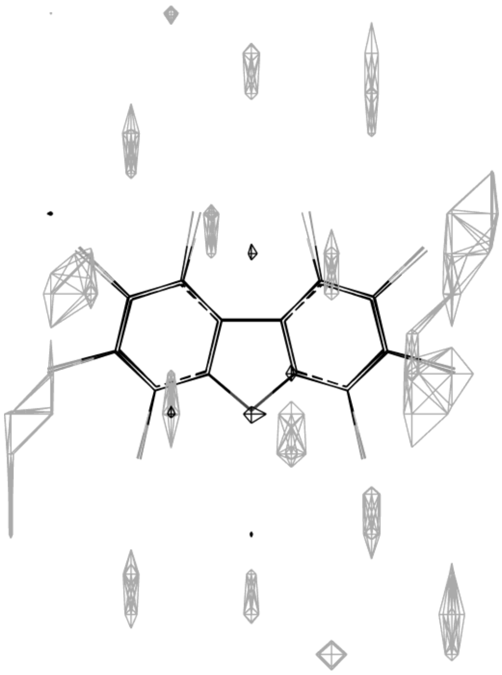

Figure 4 Coefficient contour plot of C1= probe.

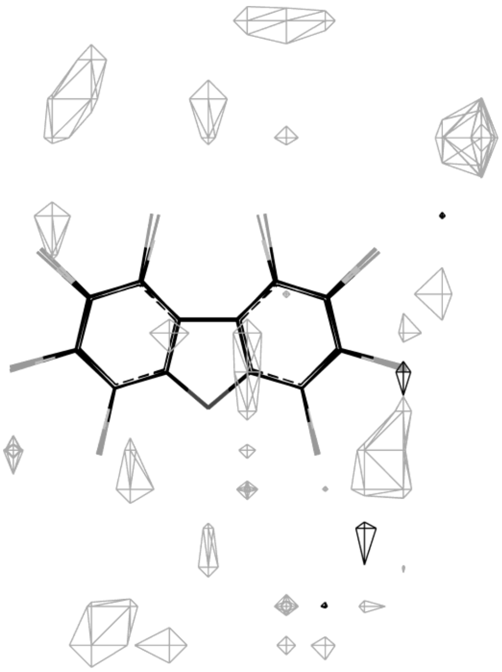

Figure 5 Coefficient contour plot of DRY probe.

interactions between PCDFs and stationary phase are induction and dispersive, with the energy of the interactions dependent on the electronic polarizability, contributed by chlorine substituent.

Figure 5 shows the coefficients contour of DRY probe. The absolute values of coefficients are much smaller than those of C1= and C3, thus, hydrophobic effects are not as significant as steric effects. However, all the coefficients are positive, which reveals that as the chlorine increases, the hydrophobic and IR increase. The DB-5 stationary has small polarizability, and PCDFs have strong hydrophobicity, therefore the variance of MIF of DRY probe is not significant for every PCDFs.

3 Conclusions

The retention behavior for PCDFs is successfully modeled based on MIF using MBPLS method. The model is statistically valid and correlated highly with observed data. The results of MBPLS and BIP show that the steric properties have a dominant influence on the retention behavior of PCDFs, and then the hydrophobic effects. This study also implies that the application of MIF descriptors to QSRR modeling may provide a potential way to interpret and predict the chromatographic retention behavior from intermolecular interaction.

The authors are indebted to Dr. M. Vivien and Dr. J. Westerhuis for their kind help in programming MBPLS.

1 Van den Heuvel J P, Lucier G. Environmental toxicology of polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzofurans. Environ Health Perspect, 1993, 100: 189—200[doi]

2 Whyte J J, Jung R E, Schmitt C J et al. Ethoxyresorufin-O-deethylase (EROD) activity in fish as a biomarker of chemical exposure. Crit Rev Toxicol, 2000, 30: 347—570[doi]

3 Rappe C, Andersson R, Bergqvist P A, et al. Overview on environmental fate of chlorinated dioxins and dibenzofurans. Sources, levels and isomeric pattern in various matrices. Chemosphere, 1987, 16: 1603—1618[doi]

4 Safe S. Polychlorinated dibenzo-*p*-dioxins and related compounds: Sources, environmental distribution and risk assessment. Environ Carcinog Ecotoxicol Rev, 1991, C9: 261—302

5 Grimwood M J, Dobbs T J. A review of the aquatic ecotoxicology of polychlorinated dibenzo-*p*-dioxins and dibenzofurans. Environ Toxicol Water Qual, 1995, 10: 57—75[doi]

6 Makino M. Dependence of GC-RRTs on the electrical charge at substituents of chlorinated dioxins and related compounds. Chemosphere, 2001, 44: 1299—1305[doi]

7 Hale M D, Hileman F D, Mazer T, et al. Mathematical modeling of temperature programmed capillary gas chromatographic retention indexes for polychlorinated dibenzofurans. Anal Chem, 1985, 57: 640—648[doi]

8 Needham M D, Adams K C, Jurs P C. Quantitative structure-retention relationship studies of polychlorinated dibenzofurans on gas chromatographic stationary phases of varying polarity. Anal Chim Acta, 1992, 258: 199—218[doi]

9 Hodjmohammadi M R, Ebrahimi P, Pourmorad F. Quantitative structure-retention relationships (QSRR) of some CNS agents studied on DB-5 and DB-17 phases in gas chromatography. QSAR Comb Sci, 2004, 23: 295—302[doi]

10 Rayne S, Ikonomou M G. Development of a multiple-class high-resolution gas chromatographic relative retention time model for halogenated environmental contaminants. Anal Chem, 2003, 75: 1049—1057[doi]

11 Liang X, Wang W, Schramm K W, et al. A new method of predicting of gas chromatographic retention indices for polychlorinated dibenzofurans (PCDFs). Chemosphere, 2001, 41: 1889—1895[doi]

12 Cruciani G, Benedetti P, Caltabiano G, et al. Structure-based rationalization of antitumor drugs mechanism of action by a MIF approach. Euro J Med Chem, 2004, 39: 281—289[doi]

13 Goodford P J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J Med Chem, 1985, 28: 849—857[doi]

14 Pastor M, Cruciani G, Watson K A. A strategy for the incorporation of water molecules present in a ligand binding site into a three-dimensional quantitative structure-activity relationship analysis. J Med Chem, 1997, 40: 4089—4102[doi]

15 Stewart J J P. Mopac: A semi empirical molecular orbital program. J Comput-Aided Mol Des, 1990, 4: 1—105[doi]

16 Hites R A. Environmental behavior of chlorinated dioxins and furans. Acc Chem Res, 1990, 23: 194—201[doi]

17 GRID 22a. Molecular Discovery Ltd., http://www.moldiscovery.com, 2004

18 Vivien M, Sabatier R. Generalized orthogonal multiple co-inertia analysis (-PLS): New multiblock component and regression methods. J Chemometrics, 2003, 17: 287—301[doi]

19 Wold S. PLS for multivariate linear modeling. In: van de Waterbeemd H ed. Chemometric Methods in Molecular Design. Weinheim: Verlag Chemie, 1995. 195—218

20 Wangen L E, Kowalski B R. A multiblock partial least squares algorithm for investigating complex chemical system. J Chemometrics, 1989, 3: 3—20[doi]

21 Westerhuis J A, Kourti T, MacGregor J F. Analysis of multiblock and hierarchical PCA and PLS models. J Chemometrics, 1998, 12: 301—321[doi]

22 Westerhuis J A, Smilde A K. Deflation in multiblock PLS. J Chemometrics, 2001, 15: 485—493

23 R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org