

# 用 PM3 量化参数预测多氯代二苯并呋喃的 $\log K_{ow}$ \*

饶火瑜\*\* , 丁健桦 , 乐长高 , 朱霞萍 , 罗明标

( 东华理工学院应用化学系, 抚州 344000 )

**摘要:** 从现有的 PCDFs 分子的正辛醇/水分配系数 ( $\log K_{ow}$ ) 实验数据出发, 建立定量结构-性质关系方程 (QSPR). 采用 G98W 程序包中的 PM3 方法对 135 个多氯代二苯并呋喃 (PCDFs) 分子和二苯并呋喃进行了优化计算, 作业命令为 #p PM3 opt freq scf( conver = 9 ), 以计算所得的分子轨道能量、碳原子电荷作为 PCDFs 分子结构描述符, 运用多元线性回归技术建立了 PCDFs 的  $\log K_{ow}$  与分子结构描述符的四元方程, 最优相关系数为 0.9507, 标准偏差为 0.1737. 经检验该模型的稳健性好, 并对未有实验数据的 85 个 PCDFs 的  $\log K_{ow}$  进行预测.

**关键词:** 多氯代二苯并呋喃; 正辛醇/水分配系数; 量化参数; QSPR

中图分类号: O625.2, X820.6 文献标识码: A

## Predicting the $\log K_{ow}$ of PCDFs with PM3 Quantum Chemical Descriptors \*

Rao Huoyu\*\* , Ding Jianhua , Le Zhanggao , Zhu Xiaping , Luo Mingbiao

( Department of Applied Chemistry , East China Institute of Technology , Fuzhou 344000 )

**Abstract** Restricted by experimental determination means, the establishment of quantitative structure-properties relationship (QSPR) from the existing n-Octanol/water partition coefficient ( $\log K_{ow}$ ) experimental values of PCDFs. The geometrical optimization on 135 polychlorinated dibenzofurans (PCDFs) and dibenzofuran have been performed with the PM3 method included in G98W package. Job control is #p PM3 opt freq scf( conver = 9 ). The calculated molecular orbital energies and mulliken atomic charges of carbon atoms are used as structural descriptors for PCDFs. A quantitative equation between the  $\log K_{ow}$  and the structural descriptors has been established with the multiple linear regression method. The best multiple regression coefficient is 0.9507, and the standard deviation is 0.1737. The robustness of the model is verified to be excellent. The  $\log K_{ow}$  of 85 PCDFs without experimental values is predicted with the model.

**Keywords** PCDFs, n-Octanol/water partition coefficient, Quantum chemical descriptors, QSPR

### 1 引言

多氯代二苯并呋喃 (PCDFs) 属二噁英类化合物, 是持久性有机污染物的一类, 由于它对环境、生物和人体极其有害, 近年来已引起广大环境科学家

和公众的特别关注. 按照《关于持久性有机污染物 (POPs) 的斯德哥尔摩公约》, 正辛醇/水分配系数 ( $\log K_{ow}$ ) 是用于判断某种化合物是否属于持久性有机污染物的重要判据. 就 PCDFs 而言, 因氯原子取代个数及位置的不同, PCDFs 共有 135 个异构

\* Project supported by the Natural Science Foundation of Jiangxi province (0120019) and the East China Institute of Technology (DHB0406).

\*\* Corresponding author, E-mail: rhy10@163.com Received 20 May 2003; in final form 18 August 2003.

体<sup>[1]</sup>. 而单一的 PCDF 纯物质通常难以获得, 而且有些共存的 PCDFs 即使是用 GC/MS 方法仍不容易分开, 因此, 还有相当多的 PCDFs 无实验数据. 从 PCDFs 现有的  $\log K_{ow}$  实验数据出发, 找出其与 PCDFs 分子结构之间的定量关系, 有着十分重要的意义. 近年来, 黄俊等分别应用分子距边矢量结合神经网络法<sup>[2]</sup>和单苯环氯取代指数法<sup>[3]</sup>对未有实验值的 85 个 PCDFs 的  $\log K_{ow}$  进行了预测, 两法对同一 PCDFs 分子的  $\log K_{ow}$  的预测值普遍相差较大, 个别相差高达 1.1. 最近, Wang 等报道了 PCDFs 的气相色谱保留指数与  $\log K_{ow}$  的定量关系<sup>[4]</sup>, 该法对 135 个 PCDFs 的  $\log K_{ow}$  的预测值与 SOFA 法<sup>[5]</sup>的预测值绝对平均误差达到 0.34, 两法对一氯代二苯并呋喃的  $\log K_{ow}$  的预测值平均相差 0.80. 上述报道说明, 分子连接性指数与气相色谱保留指数等方法仍有一定的局限性, 对 PCDFs 的三环结构区分程度不很理想, 仍有必要采用其它参数表征 PCDFs 分子结构并建立新的 QSPR 模型.

在近期的 QSAR/QSPR 研究中, 已经越来越多地独自应用 PM3 法计算的量子化学参数或将其与其它常规的参数联合使用<sup>[6-8]</sup>. 将量化参数应用到 QSPR 研究中的优点为化合物分子及其子结构、取代物都能直接地在其分子结构的基础之上被表征. 本工作以 PM3 方法计算得到的量化参数作为 PCDFs 的分子结构描述符, 利用多元线性回归技术建立了估算 PCDFs 的  $\log K_{ow}$  的新模型.

## 2 理 论

由 Kamlet 等人发展起来的线性溶解能相关 (LSER) 理论<sup>[9]</sup>在有机污染物的 QSPR 研究中得到了广泛的应用. Wilson 和 Famini 等应用理论计算的参数<sup>[10]</sup>, 提出了理论线性溶解能相关 (TLSE) 模型. 笔者在前人基础上, 结合 PCDFs 分子的结构 (见图 1), 提出其理化性质与最低空轨道能量、环上碳

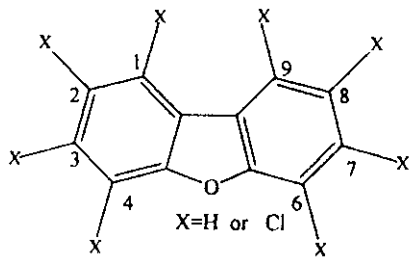


图 1 PCDFs 分子及碳原子编号

Fig. 1 The PCDFs molecules and the number of carbon atoms

原子电荷线性相关模型, 即:

$$XYZ = XYZ_0 + a_0 E_{LUMO} + a_1 Q_1 + a_2 Q_2 + a_3 Q_3 + a_4 Q_4 + a_6 Q_6 + a_7 Q_7 + a_8 Q_8 + a_9 Q_9 \quad (1)$$

式中:  $XYZ$  为分配性质;  $XYZ_0$  为常数项;  $E_{LUMO}$  为 PCDFs 分子的最低空轨道能量 (hartree);  $Q_1$ 、 $Q_2$ 、 $Q_3$ 、 $Q_4$ 、 $Q_6$ 、 $Q_7$ 、 $Q_8$ 、 $Q_9$  为 PCDFs 分子中氯连或氢连碳原子的形式电荷;  $a_0$ 、 $a_1$ 、 $a_2$ 、 $a_3$ 、 $a_4$ 、 $a_6$ 、 $a_7$ 、 $a_8$ 、 $a_9$  为常数.

## 3 量化参数的计算

在 P4/2.4GHz 微机上, 运用 Chem3D 软件包, 采用图形输入方式, 选用该软件包中的 MM2 方法对输入的 PCDFs 分子进行构象优化. 对优化好的 PCDFs 分子, 利用 G98W 程序包中的 PM3 方法进行优化, 收敛控制采用 scf (conver = 9), 并加入关键词 freq 进行频率分析, 以确证所得最优结构无虚频, 原子电荷的计算采用程序默认的 Mulliken 算法, 计算结果见表 1. 从表 1 可以看出, 碳原子电荷的变化反映了取代方式的变化, 是 PCDFs 分子结构的微观表征.

## 4 量化参数与 PCDFs 的 $\log K_{ow}$ 的相关性

以 Ruelle 等收集整理的含有 50 种 PCDFs 的数据集<sup>[11]</sup>, 再加上母体化合物二苯并呋喃共 51 种化合物构成建模样本, 如表 1 所示. 选择不同个数的量化参数进行多元线性回归, 在充分满足相关系数高、标准偏差小、自变量个数少以及自变量之间无自相关关系的条件下, 得到如下四元 QSPR 方程:

$$\begin{aligned} \log K_{ow} = & 1.4559 - 103.47 E_{LUMO} - \\ & 4.9781 Q_1 - 3.9217 Q_4 + 2.0609 Q_8 \quad (2) \\ n = & 51, s = 0.1901, r = 0.9266, \\ F = & 69.77, p < 0.0001 \end{aligned}$$

其中,  $n$  为建模样本数;  $s$  为标准偏差;  $r$  为复相关系数;  $F$  为方差比;  $p$  为显著水平. 式中系数表示各种作用的贡献程度, 从式 (2) 中可以看出,  $E_{LUMO}$  的贡献非常大. 事实上, 单独以  $E_{LUMO}$  与 PCDFs 分子的  $\log K_{ow}$  进行一元相关, 其相关系数已达 0.8921, 相当于文献 [3] 的九元相关.

为了识别“异常值”及考察模型是否存在机会相关, 用修正的 Jackknife 法对回归方程的稳健性进行检验. 鉴于所研究的样本是一个中等大小样本, 检验方式采用保留二苯并呋喃依次抽取 5 个化合物, 以余下的 46 个样本进行回归, 共得 10 个复相关系

表 1 PCDFs 的量化参数及  $\log K_{ow}$  的实验值、预测值Table 1 The quantum chemical descriptors and the experimental and predicted values of  $\log K_{ow}$  of PCDFs

No	Substitution pattern	$E_{LUMO}$ /hartree	$Q_1$	$Q_4$	$Q_8$	$\log K_{ow}$		
						Exp.	Equ.(2)	Equ.(3)
0	Dibenzofuran	-0.01750	-0.12592	-0.20178	-0.21683	4.31	4.24	4.18
1	2 8 -	-0.02887	-0.12602	-0.18659	-0.15455	5.44	5.48	5.47
2	1 2 7 8 -	-0.03830	-0.08105	-0.18833	-0.16310	6.23	6.22	6.26
3	1 3 4 7 -	-0.04017	-0.06594	-0.16362	-0.21590	6.23	6.14	6.18
4	1 3 4 8 -	-0.03931	-0.06461	-0.16481	-0.15171	6.13	6.18	6.25
5	1 3 4 9 -	-0.03869	-0.06509	-0.16809	-0.22407	5.89	5.98	6.02
6	1 3 6 8 -	-0.03941	-0.05779	-0.20674	-0.13960	6.37	6.34	6.40
7	1 4 6 9 -	-0.03770	-0.07531	-0.15637	-0.21143	5.60	5.91	5.95
8	2 3 4 6 -	-0.03760	-0.11903	-0.14739	-0.20033	6.11	6.10	6.12
9	2 3 6 7 -	-0.03909	-0.11390	-0.18924	-0.20529	6.31	6.39	6.39
10	2 3 6 8 -	-0.03874	-0.11128	-0.19033	-0.14218	6.73	6.47	6.49
11	2 3 7 8 -	-0.03949	-0.11270	-0.19107	-0.16497	6.53	6.51	6.53
12	2 4 6 7 -	-0.03822	-0.13129	-0.13512	-0.20576	6.25	6.17	6.18
13	1 2 4 8 -	-0.03794	-0.08717	-0.14173	-0.15272	6.31	6.06	6.12
14	1 3 4 6 -	-0.03860	-0.06494	-0.16364	-0.19797	6.31	6.01	6.06
15	1 3 7 8 -	-0.04019	-0.05899	-0.20735	-0.16223	6.34	6.39	6.43
16	1 3 7 9 -	-0.04052	-0.05942	-0.21093	-0.23050	6.34	6.30	6.32
17	2 3 4 9 -	-0.03801	-0.12151	-0.14774	-0.21836	6.17	6.12	6.13
18	1 2 3 4 -	-0.03794	-0.07693	-0.15290	-0.21192	6.17	5.93	5.97
19	2 4 6 8 -	-0.03781	-0.12875	-0.13628	-0.14267	6.17	6.25	6.29
20	1 2 3 8 -	-0.03851	-0.06985	-0.19602	-0.15213	6.15	6.24	6.29
21	1 4 6 7 -	-0.03830	-0.07637	-0.15170	-0.20345	6.15	5.97	6.02
22	1 2 3 6 -	-0.03797	-0.07007	-0.19418	-0.19842	6.15	6.09	6.11
23	1 2 7 9 -	-0.03876	-0.08119	-0.19203	-0.23136	6.25	6.15	6.16
24	1 2 6 7 -	-0.03792	-0.08138	-0.18639	-0.20371	6.25	6.10	6.12
25	2 3 4 7 -	-0.03918	-0.11997	-0.14735	-0.21844	6.06	6.23	6.25
26	1 2 3 9 -	-0.03801	-0.07015	-0.20009	-0.22449	6.06	6.06	6.08
27	1 2 3 6 9 -	-0.04223	-0.06888	-0.19808	-0.21063	6.65	6.51	6.54
28	1 2 3 6 7 -	-0.04278	-0.07024	-0.19300	-0.20270	6.26	6.57	6.61
29	1 2 4 6 7 -	-0.04210	-0.08792	-0.14003	-0.20320	6.27	6.38	6.43
30	1 2 4 6 8 -	-0.04178	-0.08640	-0.14128	-0.14054	6.34	6.47	6.55
31	1 2 4 7 8 -	-0.04266	-0.08747	-0.14122	-0.16345	6.26	6.52	6.59
32	1 3 4 6 8 -	-0.04303	-0.06356	-0.16459	-0.13949	6.24	6.58	6.66
33	1 3 4 6 9 -	-0.04266	-0.06393	-0.16773	-0.21017	6.34	6.41	6.46
34	1 3 4 8 9 -	-0.04244	-0.06362	-0.16848	-0.16978	6.51	6.47	6.54
35	2 3 4 6 7 -	-0.04242	-0.11864	-0.14673	-0.20477	6.47	6.59	6.62
36	2 3 4 7 8 -	-0.04296	-0.11733	-0.14797	-0.16513	6.92	6.73	6.77

续表 1.

No	Substitution pattern	$E_{LUMO}$ /hartree	$Q_1$	$Q_4$	$Q_8$	$\log K_{ow}$		
						Exp.	Equ.( 2 )	Equ.( 3 )
37	2 3 4 8 9 -	-0.04181	-0.11946	-0.14798	-0.16528	6.42	6.62	6.65
38	1 3 4 7 9 -	-0.04403	-0.06501	-0.16758	-0.23013	6.33	6.52	6.56
39	1 2 3 6 8 -	-0.04248	-0.06868	-0.19427	-0.14001	6.33	6.67	6.73
40	1 2 4 7 9 -	-0.04297	-0.08755	-0.14448	-0.23106	6.19	6.43	6.47
41	1 3 4 6 7 -	-0.04326	-0.06512	-0.16319	-0.20228	6.19	6.48	6.54
42	2 3 4 6 9 -	-0.04214	-0.11998	-0.14730	-0.20432	6.53	6.57	6.59
43	1 2 3 4 7 -	-0.04323	-0.07696	-0.15238	-0.21583	6.53	6.46	6.51
44	1 2 3 4 8 -	-0.04242	-0.07556	-0.15357	-0.15213	6.79	6.51	6.58
45	1 2 3 7 8 -	-0.04320	-0.06996	-0.19495	-0.16278	6.79	6.70	6.75
46	2 3 4 6 8 -	-0.04216	-0.11599	-0.14795	-0.14209	6.59	6.68	6.73
47	1 2 4 6 9 -	-0.04153	-0.08649	-0.14470	-0.21107	6.59	6.32	6.36
48	1 2 3 4 7 8 -	-0.04673	-0.07580	-0.15318	-0.16287	7	6.93	7.01
49	1 2 3 4 6 7 8 -	-0.04959	-0.07508	-0.15262	-0.15036	7.4	7.25	7.34
50	1 2 3 4 6 7 8 9 -	-0.05251	-0.07392	-0.15630	-0.16970	7.97	7.52	7.61
51	1 -	-0.02341	-0.07143	-0.20382	-0.21428		4.59	4.58
52	2 -	-0.02338	-0.12919	-0.18742	-0.21623		4.81	4.76
53	3 -	-0.02474	-0.11250	-0.20544	-0.21542		4.94	4.90
54	4 -	-0.02288	-0.12904	-0.14814	-0.21598		4.60	4.58
55	1 2 -	-0.02801	-0.08212	-0.19020	-0.21367		5.07	5.06
56	1 3 -	-0.03030	-0.06070	-0.20859	-0.21284		5.27	5.28
57	1 4 -	-0.02883	-0.07718	-0.15287	-0.21344		4.98	5.00
58	1 6 -	-0.02828	-0.07029	-0.20153	-0.20018		5.11	5.11
59	1 7 -	-0.03003	-0.07121	-0.20213	-0.21762		5.26	5.26
60	1 8 -	-0.02881	-0.07006	-0.20334	-0.15212		5.27	5.29
61	1 9 -	-0.02845	-0.07060	-0.20745	-0.22601		5.10	5.09
62	2 3 -	-0.02937	-0.11667	-0.19281	-0.21514		5.39	5.36
63	2 4 -	-0.02844	-0.13332	-0.13662	-0.21550		5.15	5.14
64	2 6 -	-0.02838	-0.12723	-0.18497	-0.20218		5.33	5.31
65	2 7 -	-0.03012	-0.12812	-0.18554	-0.21985		5.48	5.45
66	3 4 -	-0.02905	-0.11617	-0.15866	-0.21471		5.22	5.21
67	3 6 -	-0.02956	-0.11013	-0.20334	-0.20129		5.45	5.42
68	3 7 -	-0.03125	-0.11101	-0.20391	-0.21897		5.59	5.56
69	4 6 -	-0.02767	-0.12712	-0.14759	-0.20186		5.11	5.11
70	1 2 3 -	-0.03375	-0.07133	-0.19627	-0.21257		5.63	5.65
71	1 2 4 -	-0.03317	-0.08832	-0.14128	-0.21295		5.44	5.47
72	1 2 6 -	-0.03259	-0.08114	-0.18792	-0.19958		5.56	5.57
73	1 2 7 -	-0.03427	-0.08212	-0.18841	-0.21724		5.70	5.71
74	1 2 8 -	-0.03313	-0.08090	-0.18965	-0.15253		5.72	5.75
75	1 2 9 -	-0.03276	-0.08114	-0.19379	-0.22548		5.54	5.55

续表 1.

No	Substitution pattern	$E_{\text{LUMO}}$ /hartree	$Q_1$	$Q_4$	$Q_8$	$\log K_{\text{ow}}$		
						Exp.	Equ.( 2 )	Equ.( 3 )
76	1 3 4 -	-0.03465	-0.06610	-0.16398	-0.21215		5.58	5.61
77	1 3 6 -	-0.03470	-0.05932	-0.20667	-0.19868		5.74	5.76
78	1 3 7 -	-0.03629	-0.06034	-0.20714	-0.21636		5.88	5.90
79	1 3 8 -	-0.03527	-0.05910	-0.20845	-0.15170		5.90	5.94
80	1 3 9 -	-0.03477	-0.05969	-0.21241	-0.22466		5.72	5.73
81	1 4 6 -	-0.03314	-0.07622	-0.15239	-0.19928		5.45	5.49
82	1 4 7 -	-0.03488	-0.07713	-0.15235	-0.21708		5.60	5.63
83	1 4 8 -	-0.03385	-0.07587	-0.15345	-0.15230		5.62	5.68
84	1 4 9 -	-0.03334	-0.07632	-0.15690	-0.22530		5.44	5.46
85	2 3 4 -	-0.03350	-0.12090	-0.14792	-0.21448		5.66	5.66
86	2 3 6 -	-0.03389	-0.11450	-0.19062	-0.20104		5.87	5.86
87	2 3 7 -	-0.03549	-0.11548	-0.19114	-0.21889		6.00	5.99
88	2 3 8 -	-0.03446	-0.11328	-0.19228	-0.15424		6.02	6.03
89	2 3 9 -	-0.03428	-0.11709	-0.19218	-0.21900		5.89	5.87
90	2 4 6 -	-0.03290	-0.13163	-0.13594	-0.20138		5.63	5.64
91	2 4 7 -	-0.03467	-0.13251	-0.13591	-0.21941		5.78	5.79
92	2 4 8 -	-0.03359	-0.13030	-0.13687	-0.15481		5.80	5.82
93	2 4 9 -	-0.03339	-0.13387	-0.13627	-0.21941		5.66	5.66
94	3 4 6 -	-0.03339	-0.11413	-0.15826	-0.20055		5.69	5.69
95	3 4 7 -	-0.03509	-0.11496	-0.15827	-0.21849		5.83	5.83
96	3 4 8 -	-0.03412	-0.11276	-0.15932	-0.15356		5.86	5.88
97	3 4 9 -	-0.03390	-0.11640	-0.15884	-0.21851		5.72	5.71
98	1 2 3 7 -	-0.03947	-0.07117	-0.19470	-0.21626		6.21	6.24
99	1 2 4 6 -	-0.03724	-0.08754	-0.14076	-0.19882		5.89	5.93
100	1 2 4 7 -	-0.03890	-0.08850	-0.14062	-0.21679		6.03	6.07
101	1 2 4 9 -	-0.03743	-0.08733	-0.14523	-0.22490		5.87	5.90
102	1 2 6 8 -	-0.03746	-0.07995	-0.18754	-0.14052		6.18	6.22
103	1 2 6 9 -	-0.03738	-0.08013	-0.19158	-0.21172		6.04	6.05
104	1 2 8 9 -	-0.03684	-0.07989	-0.19314	-0.17059		6.07	6.10
105	1 3 6 7 -	-0.03977	-0.05928	-0.20536	-0.20277		6.25	6.28
106	1 3 6 9 -	-0.03919	-0.05829	-0.21048	-0.21080		6.19	6.22
107	1 4 6 8 -	-0.03792	-0.07493	-0.15301	-0.14013		6.06	6.13
108	1 4 7 8 -	-0.03886	-0.07598	-0.15301	-0.16292		6.12	6.18
109	2 3 4 8 -	-0.03830	-0.11769	-0.14839	-0.15429		6.27	6.30
110	3 4 6 7 -	-0.03853	-0.11348	-0.15773	-0.20481		6.20	6.22
111	1 2 3 4 6 -	-0.04174	-0.07590	-0.15243	-0.19775		6.34	6.40
112	1 2 3 4 9 -	-0.04179	-0.07574	-0.15691	-0.22395		6.31	6.35
113	1 2 3 7 9 -	-0.04351	-0.07007	-0.19854	-0.23050		6.61	6.64
114	1 2 3 8 9 -	-0.04182	-0.06866	-0.19953	-0.17020		6.56	6.60

续表 1.

No	Substitution pattern	$E_{LUMO}$ /hartree	$Q_1$	$Q_4$	$Q_8$	$\log K_{ow}$	
						Exp.	Equ.( 2 ) Equ.( 3 )
115	1 2 4 8 9 -	-0.04125	-0.08613	-0.14539	-0.17081	6.37	6.43
116	1 3 4 7 8 -	-0.04383	-0.06469	-0.16452	-0.16238	6.62	6.70
117	2 3 4 7 9 -	-0.04354	-0.12110	-0.14710	-0.22437	6.68	6.70
118	1 2 3 4 6 7 -	-0.04618	-0.07620	-0.15179	-0.20214	6.79	6.85
119	1 2 3 4 6 8 -	-0.04600	-0.07459	-0.15320	-0.13988	6.90	6.98
120	1 2 3 4 6 9 -	-0.04559	-0.07476	-0.15652	-0.21007	6.73	6.78
121	1 2 3 4 7 9 -	-0.04690	-0.07584	-0.15627	-0.23015	6.82	6.88
122	1 2 3 4 8 9 -	-0.04539	-0.07444	-0.15725	-0.17025	6.79	6.86
123	1 2 3 6 7 8 -	-0.04640	-0.06904	-0.19318	-0.15038	7.05	7.11
124	1 2 3 6 7 9 -	-0.04685	-0.06910	-0.19662	-0.21700	6.97	7.01
125	1 2 3 6 8 9 -	-0.04585	-0.06745	-0.19771	-0.15805	6.99	7.04
126	1 2 3 7 8 9 -	-0.04637	-0.06876	-0.19847	-0.18203	7.00	7.05
127	1 2 4 6 7 8 -	-0.04576	-0.08694	-0.14065	-0.15100	6.86	6.94
128	1 2 4 6 7 9 -	-0.04621	-0.08692	-0.14389	-0.21759	6.79	6.84
129	1 2 4 6 8 9 -	-0.04518	-0.08529	-0.14483	-0.15866	6.80	6.87
130	1 3 4 6 7 8 -	-0.04682	-0.06391	-0.16408	-0.14990	6.95	7.04
131	1 3 4 6 7 9 -	-0.04716	-0.06413	-0.16707	-0.21657	6.86	6.92
132	2 3 4 6 7 8 -	-0.04608	-0.11611	-0.14736	-0.15263	7.07	7.12
133	1 2 3 4 6 7 9 -	-0.04989	-0.07516	-0.15587	-0.21667	7.16	7.22
134	1 2 3 4 6 8 9 -	-0.04901	-0.07339	-0.15681	-0.15797	7.18	7.26
135	1 2 3 4 7 8 9 -	-0.04961	-0.07465	-0.15678	-0.18220	7.20	7.27

数是 0.9196、0.9309、0.9353、0.9298、0.9306、0.9324、0.9370、0.9382、0.9315 和 0.9063 ,只有两次比式( 2 )的相关系数略低 ;进一步保留二苯并呋喃依次抽取 10 个样本以余下的 41 个样本进行建模 ,所得相关系数依次为 0.9238、0.9396、0.9363、0.9507 和 0.9116. 经过多次随机删除  $n$  个样本进行  $m$  次回归检验表明该模型稳健可靠 ,不是随机相关的 ,可用于对尚无实验数据的 85 个 PCDFs 的  $\log K_{ow}$  进行预测.

为进一步说明该模型的可靠性 ,我们列出对 41 个样本建模所得的相关系数最好的方程 ,以同式( 2 )及现有文献方法<sup>[3 4 5 11]</sup>对 51 个样本的  $\log K_{ow}$  的计算结果进行比较 ,比较结果见表 2.

$$\log K_{ow} = 1.6330 - 105.46E_{LUMO} - 4.2801Q_1 - 3.3839Q_4 + 2.4204Q_8 \quad (3)$$

$n = 41$  ,  $s = 0.1737$  ,  $r = 0.9507$   
 $F = 84.58$  ,  $p < 0.0001$

表 2 5 种方法预测结果比较

Table 2 Comparisons of results obtained by the five methods

	Equ.( 2 )	Equ.( 3 )	Single-phenyl-ring Cl-substitution index <sup>[3 1]</sup>	MOD <sup>[11]</sup>	SOFA <sup>[5]</sup>	GC-RF <sup>[4]</sup>
Mean absolute deviation	0.1448	0.1483	0.2143	0.1790	0.3461	0.1770
RMS deviation	0.1805	0.1867	0.2618	0.2331	0.4033	0.2240
Max absolute deviation	0.45	0.36	0.70	0.61	0.74	0.68

与现有文献方法<sup>[3,4,5,11]</sup>相比,本工作方法具有预测平均绝对误差小、均方根误差小、最大绝对误差小的优点,而且本法无需做复杂的测试. 综合比较,本法在对 PCDFs 的  $\log K_{ow}$  的预测中优于现有方法. 在对 PCDFs 的  $\log K_{ow}$  的预测中,式(3)比式(2)具有较小的最大绝对误差,比较接近的平均偏差和均方根误差,说明式(3)比式(2)的预测效果更胜一筹. 运用式(2)和式(3)对全部 PCDFs 的  $\log K_{ow}$  的预测结果列于表 1.

## 5 结 论

PM3 方法计算的最低空轨道能量和 Mulliken 原子电荷反映了 PCDFs 分子的原子间的相互作用,揭示了不同取代方式的 PCDFs 分子结构的细微差别,比分子连接性指数更具优越性,该方法所用到的描述符计算简单、所需机时少,仅需运用量子化学中的半经验程序就可得到满意的结果,可用于对 PCDFs 分子其它理化性质的建模与预测.

## 参 考 文 献

[1] Zheng Minghui (郑明辉), Liu Pengyan (刘鹏岩), Bao Zhicheng (包志成), *et al. Chin. Sci. Bul.* (科学通报),

1999, **44**:455

- [2] Huang Jun (黄俊), Yu Gang (余刚), Zhang Pengyi (张彭义). *Comput. Appl. Chem.* (计算机与应用化学), 2002, **19**:103
- [3] Huang Jun (黄俊), Yu Gang (余刚), Zhang Pengyi (张彭义). *Res. of Environ. Sci.* (环境科学研究), 2002, **15**:1
- [4] Wang Y H, Wong P K. *Chemosphere*, 2003, **50**:499
- [5] Govers H A J, Krop H B. *Chemosphere*, 1998, **37**:2139
- [6] Chen Jingwen (陈景文), Wang Shijie (王师杰), Quan Xie (全燮), *et al. Journal of Dalian University of Technology* (大连理工大学学报), 2000, **40**:305
- [7] Liu Yihui (刘毅慧), Quan Xie (全燮), Chen Jingwen (陈景文), *et al. Journal of Dalian University of Technology* (大连理工大学学报), 2000, **40**:301
- [8] Yan Yun (阎云), Chen Jingwen (陈景文), Quan Xie (全燮), *et al. Acta Scientiae Circumstantiae* (环境科学学报), 2001, **21**:286
- [9] Kamlet M J, Doherty R M, Abboud J L M. *Chemtech*, 1986, **16**:566
- [10] Wilson L Y, Famini G R. *J. Med. Chem.*, 1991, **34**:1668
- [11] Ruelle P. *Chemosphere*, 2000, **40**:457