

二氯苯类化合物 PCDDs 正辛醇 水分配系数的 QSAR 研究*

QSAR Study on *n*-Octanol / Water Partition Coefficients of Polychlorinated Dibenzo-*p*-Dioxins (PCDDs) Compounds

余训民¹ 唐平²
Yu Xunmin¹ Tang Ping²

(1. 武汉化工学院环境与城市建设学院 湖北武汉 430073;
2. 湖北江汉大学环境保护系 湖北武汉 430010)

(1. Coll. of Environment & Civil Engineering, Wuhan Institute of Chemical Technology, Wuhan, Hubei, 430073, China; 2 Dept. of Environmental Protection, Jiangnan Univ., Wuhan, Hubei, 430010, China)

摘要 根据分子拓扑理论, 提出二氯苯类化合物分子中顶点原子价点价值 W^H 的计算方法, 并由 W^H 构建分子结构信息价连接性拓扑指数 ${}^m i^H (m = 1, 2, 3, \dots, n)$, 提出取代基原子距离熵 S 指数. 用 ${}^m i^H (m = 1, 2)$ 和 S 研究二氯苯类化合物的正辛醇 / 水分配系数 K_{ow} 的 $\log K_{ow}$ 的构效关系 (QSAR), 相关性显著 ($\alpha = 0.05, P = 0.0000$), 其计算值与实验值的吻合程度高于有关文献. 利用建立的模型, 给出无 $\log K_{ow}$ 文献值的其他 33 种 PCDDs 化合物的预测值.

关键词 二氯苯 正辛醇 / 水分配系数 (K_{ow}) QSAR 结构信息指数

中图分类号 0625.22; X132

Abstract Based on the molecular valence connectivity index, a method for calculating new apex atomic valence point values W^H of Polychlorinated Dibenzo-*p*-Dioxins (PCDDs) is proposed. Substituting atomic distance entropy S , we construct molecular structure information connectivity topology index ${}^m i^H (m = 0, 1, 2, \dots, n)$ by W^H . The structural efficiency relationship of the $\log K_{ow}$ of *n*-octanol / water partition coefficients (K_{ow}) of the polychlorinated dibenzo-*p*-dioxins PCDDs compounds is studied by using ${}^m i^H (m = 1, 2)$ and S . Good correlation is obtained ($\alpha = 0.05, P = 0.0000$). The matching of calculated and experimental data is better than that represented in related documents. With the established model, the $\log K_{ow}$ values of the other 33 PCDDs not belonging to the modeling sample set were given. These values has not been reported in literatures.

Key words dioxins, *n*-octanol / water partition coefficients, QSAR, structure information connectivity topological index

化合物构效关系的研究, 对评价化合物的生物效应和环境效应具有重要的意义, 是当前环境化学最前沿的研究领域之一^[1]. 氯代二苯并-*p*-二氯苯 (polychlorinated dibenzo-*p*-dioxins, 以下简称 PCDDs) 有 7 种异构体, 由于对环境、生物和人体极其有害, 近年来引起广大的环境科学家和公众的特别关注^[2~6]. 正辛醇 / 水分配系数 (K_{ow}) 是环境科学中一个重要的参数^[7], 它表征有机污染物在土壤、水体

及生物体内的富集、迁移、分解的能力, 对于评价环境污染及生物污染尤为重要. 75 种 PCDDs 中只有 43 种 PCDDs 有 $\log K_{ow}$ 实验数据, 通过 QSAR 研究, 对未测 $\log K_{ow}$ 值的 PCDDs 进行估算, 可以掌握它们对环境行为方面的许多重要信息. 最近黄俊^[8]等人采用化学图论思想, 通过引入单苯环氯取代指数作为 PCDDs 的结构符, 从而建立相关方程取得了较好的结果. 本文在前人研究^[9]的基础上, 继续对 PCDDs 的 $\log K_{ow}$ 与其结构之间的定量关系进行了研究, 提出了计算分子隐氢图中的顶点原子价点价值 W^H 的新公式, 引入了取代基原子距离熵 S 指数的概念, 并给出

2003-08-06 收稿, 2003-09-28 修回.

* 湖北教育厅自然科学基金重点项目 (2003A1062)

(10)

$$\log K_{ow} = 1.2693^2 i^H - 0.0355S + 0.6572, n = 43, s = 0.28, F = 118.99, R = 0.9252, R^2 = 0.8560,$$

(11)

上述回归方程中 $P = 0.0000$, n , s , R , F 和 P 分别为回归方程样本数、标准偏差、相关系数、方差比 (即 Fisher 检验值) 和显著性水平。分析以上回归方程, 并由 (7) ~ (11) 式得到, $F_{141}^{0.05} = 245.10$ 或 $243.52 \gg 4.08$, $F_{140}^{0.05} = 120.57$ 或 128.46 或 $118.99 \gg 3.23$,

表 1 43 种 PCDDs 化合物 $\lg K_{ow}$ 的实验值、计算值、 $m\chi^H$ ($m = 1, 2$) 和 S 值

Table 1 Experimental and calculated values of $\lg K_{ow}$ from 43 PCDDs compounds and $m\chi^H$ ($m = 1, 2$), S values

化合物 Compounds	1 ^H	2 ^H	S	$\lg K_{ow}^*$								
				Exp	cal ₇	cal ₈	cal ₉	cal ₁₀	cal ₁₁	cal ₁₂	cal ₁₃	
dibenzo-p-dioxin	4.6833	3.2653	0	4.30	4.64	4.69	4.63	4.80	4.80	4.74	4.72	
1-	4.9903	3.6191	2.6975	5.05	5.06	5.10	5.06	5.13	5.16	5.18	5.15	
2-	4.9878	3.5971	2.4859	5.00	5.06	5.08	5.05	5.14	5.13	5.12	5.15	
2, 3-	5.2948	3.9558	4.9414	5.60	5.48	5.49	5.47	5.49	5.50	5.55	5.58	
2, 7-	5.2923	3.9289	4.4643	5.75	5.48	5.46	5.47	5.51	5.49	5.51	5.58	
2, 8-	5.2923	3.9289	4.5154	5.60	5.48	5.46	5.47	5.51	5.48	5.51	5.58	
1, 2, 4-	5.6041	4.3331	7.2136	6.35	5.90	5.92	5.90	5.85	5.90	5.93	6.01	
1, 2, 3, 4-	5.9135	4.7204	9.2100	6.48	6.32	6.37	6.32	6.23	6.32	6.47	6.44	
1, 2, 3, 7-	5.9086	4.6699	8.2381	6.48	6.32	6.31	6.31	6.28	6.29	6.36	6.44	
1, 3, 6, 8-	5.9062	4.6391	7.7315	6.29	6.31	6.28	6.31	6.31	6.27	6.22	6.44	
2, 3, 7, 8-	5.9062	4.6463	6.4351	6.42	6.31	6.28	6.31	6.39	6.33	6.35	6.44	
1, 2, 3, 9-	5.9111	4.6919	9.3207	6.39	6.32	6.11	6.32	6.22	6.03	6.42	6.44	
1, 2, 6, 8-	5.9086	4.6649	8.0496	6.43	6.32	6.31	6.31	6.29	6.29	6.46	6.44	
1, 3, 7, 8-	5.9062	4.6427	7.8701	6.30	6.31	6.28	6.31	6.30	6.27	6.29	6.44	
1, 3, 7, 9-	5.9062	4.6391	7.9336	6.39	6.31	6.28	6.31	6.29	6.26	6.22	6.44	
1, 4, 7, 8-	5.9086	4.6634	8.0494	6.39	6.32	6.30	6.31	6.29	6.29	6.28	6.44	
1, 3, 6, 9-	5.9086	4.6598	8.0231	6.25	6.32	6.30	6.31	6.29	6.29	6.22	6.44	
1, 2, 4, 7-	5.9086	4.6649	8.1957	6.25	6.32	6.31	6.31	6.28	6.29	6.32	6.44	
1, 2, 4, 8-	5.9086	4.6649	8.1875	6.25	6.32	6.31	6.31	6.28	6.29	6.32	6.44	
1, 2, 3, 6-	5.9111	4.6919	8.0496	6.86	6.32	6.34	6.32	6.30	6.33	6.42	6.44	
1, 2, 7, 9-	5.9086	4.6649	8.0654	6.86	6.32	6.31	6.31	6.29	6.29	6.46	6.44	
1, 2, 7, 8-	5.9086	4.6689	7.9951	6.38	6.32	6.31	6.31	6.29	6.30	6.52	6.44	
1, 4, 6, 9-	5.9111	4.6805	8.1217	6.38	6.32	6.32	6.32	6.29	6.31	6.21	6.44	
1, 2, 3, 8-	5.9086	4.6699	8.3665	6.10	6.32	6.31	6.31	6.27	6.29	6.36	6.44	
1, 2, 4, 6-	5.9111	4.6869	8.3911	6.10	6.32	6.33	6.32	6.27	6.31	6.38	6.44	
1, 2, 4, 9-	5.9111	4.6869	8.4491	6.10	6.32	6.33	6.32	6.27	6.31	5.93	6.44	
1, 2, 3, 4, 6-	6.2004	5.0402	9.9371	6.30	6.72	6.74	6.71	6.65	6.70	6.47	6.87	
1, 2, 3, 4, 7-	6.2180	5.0522	9.6751	6.60	6.74	6.75	6.74	6.69	6.73	6.86	6.87	
1, 2, 4, 6, 9-	6.2180	5.0407	9.1994	6.60	6.74	6.74	6.74	6.72	6.73	5.93	6.87	
1, 2, 3, 6, 7-	6.2180	5.0508	9.2297	6.74	6.74	6.75	6.74	6.72	6.74	6.96	6.87	
1, 2, 3, 6, 8-	6.2156	5.0250	9.3610	6.53	6.74	6.72	6.73	6.71	6.70	6.72	6.87	
1, 2, 3, 6, 9-	6.2084	5.0457	9.9871	6.24	6.73	6.74	6.72	6.66	6.71	6.72	6.87	
1, 2, 3, 7, 8-	6.2156	5.0286	9.3906	6.64	6.74	6.72	6.73	6.71	6.71	6.79	6.87	
1, 2, 3, 7, 9-	6.2156	5.0250	9.1700	6.40	6.74	6.72	6.73	6.69	6.71	6.72	6.87	
1, 2, 4, 7, 8-	6.2004	5.0236	9.9677	6.20	6.72	6.72	6.71	6.65	6.68	6.74	6.87	
1, 2, 3, 4, 7, 8-	6.5249	5.4149	9.0103	7.80	7.16	7.17	7.16	7.23	7.21	7.28	7.30	
1, 2, 3, 6, 7, 9-	6.5249	5.4059	9.1101	7.59	7.16	7.16	7.16	7.22	7.20	7.18	7.30	
1, 2, 3, 6, 8, 9-	6.5249	5.4059	9.6950	7.59	7.16	7.16	7.16	7.19	7.17	7.18	7.30	
1, 2, 4, 6, 7, 9-	6.5249	5.4008	9.7101	6.85	7.16	7.15	7.16	7.19	7.17	7.13	7.30	
1, 2, 4, 6, 8, 9-	6.5249	5.4008	9.9101	6.85	7.16	7.15	7.16	7.18	7.16	7.13	7.30	
1, 2, 3, 4, 6, 8-	6.5249	5.4074	10.0112	6.85	7.16	7.16	7.16	7.17	7.17	7.21	7.30	
1, 2, 3, 4, 6, 7, 8-	6.9943	5.9933	10.0454	8.00	7.80	7.84	7.80	7.93	7.91	7.72	7.73	
1, 2, 3, 4, 6, 7, 8, 9-	7.1436	6.1756	9.6954	8.20	8.00	8.05	8.00	8.19	8.15	8.21	8.16	

* cal₇~ cal₁₁分别为 (7) ~ (11) 式的计算值, cal₁₂~ cal₁₃分别为文献 [8] 和 [10] 的计算值. cal₇~ cal₁₁ are the value of prediction form correlation equations (7) ~ (11), cal₁₂~ cal₁₃ calculated value from reference [8] and [10].

3 结果与讨论

3.1 模型值与实验值的对比

利用 (7) ~ (11) 式计算出 43 种 PCDDs 分子的 $\log K_{ow}$ 值, 其结果列在表 1 中. 图 2 是 4 种 PCDDs 和母体的 $\log K_{ow}$ 的模型 (式 (7)) 预测值与实验值的相关图, 从图 2 中可以看出, 两者之间的差异比较小. 从表 1 中的数据可以看出, 尽管我们采用结构信息连接性指数和分子取代基原子距离熵指数, 且没有进行复杂的数学运算, 但所建模型的精度并不比文献 [8] 和 Rulle 等^[10] 建立的基于分子热力学的 MOD 模型逊色, 其计算值与实验值的吻合程度比他们要好. 这说明本文提出的结构信息连接性指数和分子中取代基原子距离熵较好地揭示了 PCDDs 分子结构信息.

表 2 对另外 33 种 PCDDs 的 $\lg K_{ow}$ 值预测

Table 2 Predicted $\lg K_{ow}$ values of other 33 PCDDs compounds

化合物 Compounds	1jH	2jH	S	$\lg K_{ow}^*$					
				cal ₇	cal ₈	cal ₉	cal ₁₀	cal ₁₁	cal ₁₂
1, 2-	5.2972	3.9779	5.1034	5.32	5.51	5.48	5.48	5.53	5.14
1, 3	5.2948	3.9522	4.9784	5.32	5.49	5.47	5.49	5.50	5.29
1, 4-	5.2972	3.9729	5.0435	5.32	5.51	5.48	5.49	5.52	5.14
1, 6-	5.2972	3.9729	4.8973	5.32	5.51	5.48	5.49	5.53	5.10
1, 7-	5.2948	3.9509	4.6774	5.32	5.48	5.47	5.50	5.51	5.20
1, 8-	5.2948	3.9509	4.7354	5.32	5.48	5.47	5.50	5.50	5.20
1, 9-	5.2972	3.9729	4.9643	5.32	5.51	5.48	5.49	5.52	5.10
1, 2, 3-	5.6041	4.3381	7.3205	5.73	5.93	5.90	5.84	5.90	5.52
1, 2, 6-	5.6041	4.3317	6.7665	5.73	5.92	5.90	5.88	5.92	5.31
1, 2, 7-	5.6017	4.3097	6.5423	5.73	5.90	5.89	5.89	5.90	5.41
1, 2, 8-	5.6017	4.3097	6.6515	5.73	5.90	5.89	5.88	5.89	5.41
1, 2, 9-	5.6041	4.3317	6.8914	5.73	5.92	5.90	5.87	5.91	5.83
2, 3, 7-	5.5993	4.2876	6.4273	5.73	5.87	5.89	5.89	5.87	5.54
1, 7, 8-	5.6017	4.3096	6.6515	5.73	5.90	5.89	5.88	5.89	5.44
1, 3, 6-	5.6017	4.3060	6.6995	5.73	5.89	5.89	5.88	5.88	5.47
1, 3, 7-	5.5993	4.2840	6.4685	5.73	5.87	5.89	5.89	5.87	5.57
1, 3, 8-	5.5993	4.2840	6.4754	5.73	5.87	5.89	5.89	5.87	5.57
1, 3, 9-	5.6017	4.3060	6.7085	5.73	5.89	5.89	5.88	5.88	5.47
1, 4, 6-	5.6041	4.3267	6.7845	5.73	5.92	5.90	5.88	5.91	5.31
1, 4, 7-	5.6017	4.3047	6.5446	5.73	5.89	5.89	5.89	5.89	5.41
1, 2, 6, 7-	5.9111	4.6906	8.0655	6.32	6.34	6.32	6.29	6.32	5.53
1, 2, 6, 9-	5.9111	4.6855	8.1481	6.32	6.33	6.32	6.29	6.32	5.53
1, 2, 8, 9-	5.9111	4.6906	8.2995	6.32	6.34	6.32	6.28	6.32	5.53
1, 2, 4, 7, 9-	6.2156	5.0200	9.0921	6.55	6.72	6.74	6.72	6.71	6.21
1, 2, 4, 6, 8-	6.2156	5.0200	9.0853	6.55	6.72	6.74	6.72	6.71	6.21
1, 2, 4, 6, 7-	6.2180	5.0457	9.1591	6.55	6.74	6.74	6.72	6.74	6.05
1, 2, 4, 8, 9-	6.2180	5.0457	9.2682	6.55	6.74	6.74	6.72	6.73	6.05
1, 2, 3, 8, 9-	6.2180	5.0508	9.3529	6.55	6.75	6.74	6.71	6.74	5.91
1, 2, 3, 6, 7, 8-	6.5249	5.4109	9.7683	6.96	7.17	7.16	7.18	7.18	6.29
1, 2, 3, 4, 6, 9-	6.5274	5.4280	10.1360	6.97	7.18	7.16	7.17	7.19	6.21
1, 2, 3, 7, 8, 9-	6.5249	5.4109	9.8352	6.96	7.17	7.16	7.18	7.18	6.29
1, 2, 3, 4, 6, 7-	6.5274	5.4331	10.1361	6.97	7.19	7.16	7.17	7.19	6.21
1, 2, 3, 4, 6, 7, 9-	6.8343	5.7882	10.0067	7.34	7.60	7.58	7.67	7.65	6.73

* cal₇~ cal₁₁分别为 (7) ~ (11) 式的计算值, cal₁₂为文献 [8] 的预测值. cal₇~ cal₁₁ are the value of prediction from correlation equations (7) ~ (11), cal₁₂ calculated value from reference [8].

3.2 其他 33 种 PCDDs 的 $\log K_{ow}$ 的预测值

利用 (7) ~ (11) 式对未见文献报道的另外 33 种 PCDDs 化合物的 $\lg K_{ow}$ 的值进行了预测, 预测结果列在表 2 中. 从预测结果来看, 与文献 [8] 的预测结果是相一致的.

3.3 交叉检验的 R' 与回归模型的 R 呈一致性

本文用留一法 (leave-one-out) 对 4 种 PCDDs 化合物的 $\lg K_{ow}$ 值进行交叉检验, 即每次剔除一个样本, 用此样本作预测集, 其余的 $n-1$ 样本作学习集对它们进行预测. (7) ~ (11) 式交叉检验的统计特征值为 $R' = 0.9150 \sim 0.9280$, $s = 0.27 \sim 0.29$, 与原回归模型的 R 值一致. 交叉检验的结果也表明, 新模型具有较强的外部预测能力.

4 结束语

(1) 结构信息连接性指数可以成功地用于 PCDDs 化合物 $\lg K_{ow}$ 值的预测研究, 是因为 W^m 和 i^H ($m = 1, 2$) S 充分揭示了分子中顶点原子和分子本身的结构信息, 因此计算值和预测值比较接近实验值; (2) 与基于热力学计算 $\lg K_{ow}$ 的 MOD 模型和文献 [8] 的模型相比, 利用 S 和 i^H 的组合建立的新模型具有计算简便、不需要查找任何化学数据, 更适合于 PCDDs 的 $\lg K_{ow}$ 的建模与预测, 其 (复) 相关系数 R 都比相关文献的高; (3) 虽然有个别计算值偏离实验值比较大, 有可能是实验的误差造成的。

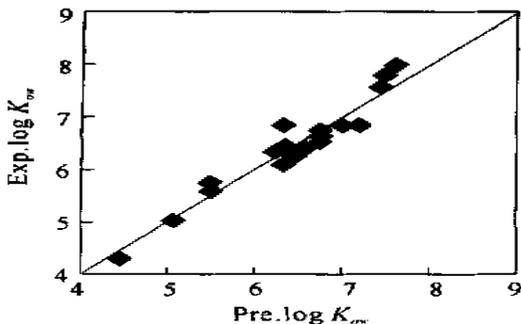


图 2 PCDDs 的 $\lg K_{ow}$ 计算值与实验值的比较

Fig. 2 Predicted value vs. experimental value of $\lg K_{ow}$ of 43 PCDDs in the sample set

参考文献

- 1 蒋展鹏. 新世纪初的环境化学. 科学, 2001, 53(6): 15~ 19.
- 2 岳瑞生. 关于就某些持久性有机污染物采用国际行动的斯

- 德哥尔摩公约及谈判背景. 世界环境, 2000, (1): 24~ 28.
- 3 郑明辉, 刘鹏岩, 包志成, 等. 二噁英的生成及降解研究进展. 科学通报, 1999, 44(5): 455~ 463.
- 4 UN-ECE. Draft protocols to the convention on long rang air pollution on persistent organic pollutants(EB Air/1998/2). Convention on long-range transboundary air pollution, United Nations Economic and Social Council, Economic Commission for Europe, 1998.
- 5 出云谕明 [日]. 威胁人类生存的定时炸弹——环境荷尔蒙. 深圳: 海天出版社, 1999. 192~ 201.
- 6 Sijm D T H M, Wever H, De Vrise P J, Opperhuizen A. Octan-1-ol/water partition coefficients of polychlorinated dibenzo-p-dioxins and dibenzofurans. experiments; experimental values determined with a stirring method. Chemosphere, 1989, 19(1-6): 263~ 266.
- 7 王连生, 韩朔睽. 有机污染化学进展. 北京: 化学工业出版社, 1998. 121, 134.
- 8 黄俊, 余刚, 张彭义, 等. 单苯环氯取代指数法预测二噁英类化合物 PCDDs 的正辛醇/水分配系数. 计算机与应用化学, 2001, 18(4): 289~ 292.
- 9 余训民, 杭义萍. 多氯代二苯并呋喃的结构信息连接性指数与其在不同色谱柱上的气相色谱保留行为的关系. 色谱, 2002, 20(1): 6~ 11.
- 10 Rulle P. The n-octanol and n-hexane/water partition coefficient of environmentally relevant chemicals predicted from the mobile order and disorder (MOD) thermodynamics. Chemosphere, 2000, 40(5): 457~ 512.

(责任编辑: 黎贞崇)

韩国初步研制出禽流感疫苗毒株

韩国科学家最近研制出针对 H5N1 型禽流感病毒的疫苗毒株, 并计划不久后用该疫苗毒株进行动物试验。韩国忠南大学兽医科教授徐相熙在 2003 年 12 月韩国爆发禽流感疫情后开始进行禽流感疫苗毒株的研究。他从 H5N1 型禽流感病毒的 8 个遗传基因中提取生成 H5 型血凝素的遗传基因, 去除其高致病性成分, 并将这种遗传基因取代人类流感病毒的 HA 遗传基因, 然后接种于人体细胞内, 使之重新组合后产生疫苗毒株。

血凝素简称 HA, 是病毒外壳的一种蛋白质, 禽流感病毒的类型就是根据血凝素的类型来划分的。研究者称, 理论上接种这种疫苗毒株后 2 个星期左右动物体内便产生抗体, 具有抗病毒的功能。因此, 接种这种疫苗毒株后, 即使带有 H5 型基因的新型流感病毒侵入, 也不能侵害机体内的正常细胞。

徐相熙计划先在猴子身上进行试验, 然后再进行人体试验。预计在 2004 年 6 月, 这种禽流感病毒疫苗毒株的开发任务将全部完成。

据《科学时报》