用分子形状指数和苯环因子预测有机物在 鱼体中的生物富集因子

冯长君1,杨伟华2,沐来龙2

(1. 徐州工程学院 化学化工学院, 江苏 徐州 221008; 2. 徐州师范大学 化学化工学院, 江苏 徐州 221116)

摘 要:通过分子形状指数(*TK*)及苯环因子(G)建立有机物生物富集因子(B_{GP})的良好 QSAR 模型,计算了 165 种非离子性有机物的分子形状指数(¹K、²K).基于分子母体及取代基的结构特征——苯环因子(G).将 ¹K、G 与 122 种非离子性有机物在鱼体中的生物富集因子(lgB_{GP})拟合,建立数学模型,相关系数(R)取 0.955. 该模型通过 Jackknife 法检验具有总体稳健性,对其他有机物的生物富集因子具有良好的估算与预测能力.结果 表明,分子形状指数与苯环因子不仅对于非离子性有机物在鱼体中的生物富集因子的表征具有合理性与有效 性,而且与其疏水性参数具有良好的相关性.

关键词:非离子性有机物;生物富集因子;分子形状指数;苯环因子;定量构效关系 中**把分类号:**O625 **文献标识码:**A **文章编号:**0254-0037(2008)06-0607-10

有机污染物对水生生物尤其是鱼类的危害表现为对生物的直接毒害及在生物体内的累积,即生物富 集.生物富集因子(B)是衡量水体中有机物在生物体(如鱼体)内富集程度的物理量,定义为污染物在生 物体中与水体中体积浓度的比值.显然,B_{CF}大于1为正富集,也叫生物浓缩.通过生物放大作用,导致处 于食物链塔顶的人类成为最终受害者.由于实测 B_{CF}成本高、周期长及人力的限制,对有机污染物的环境 行为及归宿进行评价的估算方法主要有疏水性(正辛醇/水分配系数 K_{ow})、溶解度(S_w)、片段常数、拓扑 指数等方法.其中以拓扑指数法最为简便,就是根据分子隐氢图建立各种数学矩阵(最常用的是邻接矩 阵、距离矩阵等),结合图中结点(即非氢原子)结构特征衍生出表征分子某种结构特征的描述符(即拓扑指 数)^[1-3].自 Wiener^[4]提出第1个拓扑指数(w)以来,迄今已有 400 多种图论参数问世^[5],并在物质构效 学领域中发挥主要作用^[6].国内外已有较多学者运用连接性指数("X_i")对有机物的 B_{CF}进行了研究.

Sacan 等人^[7]对 122 种非离子性有机物(涉及取代苯、取代联苯及烷烃、烯烃衍生物;取代基团有烷 基、CL-、Br-、一NH₂、一OH 及一NO₂等)的 B_{CF} ,运用 CRI 拓扑指数、分子的最高占有轨道能(E_{home})、分子 的最低空轨道能(E_{humo})以及分子的偶极距(μ)与之拟合,获得良好的相关性,其最佳二元回归方程的相关 系数(R)为0.921,估计标准误差(S)为0.599(个对数单位).本文根据 Kier 等人^[8]的分子形状指数("K) 及作者建立的苯环因子(G_B)关联这 122 种非离子性有机物的 B_{CF} ,建立最佳二元回归方程, R = 0.955、 S = 0.454.运用该模型对未进入回归的 43 种有机物的 B_{CF} 值进行预测.

1 分子形状指数(^mK)

1.1 饱和烃的 Kappa 形状指数

对于烷烃及其环烷烃分子中碳原子均为 sp³ 杂化,其分子隐氢图中的每条边均是 C---C 单键, 据此结

收稿日期: 2007-03-14.

基金项目:环境模拟与污染控制国家重点联合实验室开放基金资助项目(KJ2007001); 江苏省高校自然科学基金资助项目 (05KJD150221);徐州师范大学培育课题资助项目(05PLY04).

作者简介:冯长君(1954-),男,江苏徐州人,教授.

构特征,在分子邻接矩阵基础上,Kier 定义 Kappa 形状指数

$$K = [n_i - (m-1)][n_i - m]^2 m p_i^{-2}$$
(1)

式中, m 为 Kappa 形状指数的阶数, 本文只讨论 $m = 1, 2; n_j$ 为分子j 中非氢原子数; $^{m}P_j$ 为分子j 中具有 路径数为m 的个数. 由式(1)得

$${}^{1}K = n_{j}(n_{j}-1)^{2}/({}^{1}P_{j})^{2}$$
⁽²⁾

$${}^{2}K = (n_{i} - 1)(n_{i} - 2)^{2} / ({}^{2}P_{i})^{2}$$
(3)

例如甲基环已烷(A),其碳原子数 n'=7, P=7,²P=8,代入式(2)、(3)得

$$^{1}K = 5.143, \quad ^{2}K = 2.344$$

对于直链烷烃(不含支链),式(2)、(3)可还原为¹ $K = n_j$,² $K = n_j - 1$.

1.2 不饱和烃及含杂原子分子的 Kappa 形状指数

为了扩展饱和烃的 Kappa 形状指数的应用范围, Kier 于式(1)中引入修正项 *a*_j, 分子 *j* 中原子 *i* 的修 正值

$$\alpha_i = r_i r^{-1} - 1 \tag{4}$$

式中, r 为 sp³ 杂化碳原子的原子半径(r = 77 nm); r_i 为非 sp³ 杂化碳原子或其它杂原子的原子半径.例 如羰基中的氧原子, 其 sp² 杂化的原子半径为 62 nm, 则 $\alpha_0 = -0.20$. 分子 j 的修正值 a_j 则为所有的非 sp³ 杂化碳原子及杂原子的 α_i 之代数和 $\alpha_j = \Sigma \alpha_{ji}$.例如硝基苯的 $\alpha = -1.38$.

将 aj 引入式(1)得

$${}^{m}K = (n_{j} + a_{j} - (m - 1))(n_{j} + a_{j} - 1)^{2}({}^{1}P_{j} + a_{j})^{-2}$$
(5)

同理

$${}^{1}K = (n_{i} + a_{i})(n_{i} + a_{i} - 1)^{2}({}^{1}P_{i} + a_{i})^{-2}$$
(6)

$${}^{2}K = (n_{j} + a_{j} - 1)(n_{j} + a_{j} - 2)^{2}({}^{2}P_{j} + a_{j})^{-2}$$
(7)

例如硝基苯的 $^{1}K = 2.096 8, ^{2}K = 2.259 3$

文献[7]中 122 种非离子性有机物的 B_{CF} 及其¹K、²K 列于表 1.

表 1 122 种非离子性有机物的生物富集因子与分子结构描述符

Table 1 Bioconcentation factors (lgB_{CF}) and molecular structure descriptors for 122 nonionic organic chemicals

/ 本 日	(), A 4L	1	2		1 17	lgB _{CF}		Fr
编兮	化合物	¹ K	²K	G	lgK _{ow}	Exp.	Cal.	Er.
1	1, 1, 1-trichloroethane	5.8700	1.545 4	0.000 0	2.49	0.95	1.32	0.37
2	1, 1, 2, 2-tetrachloroethane	7.1600	3.1993	0.0000	2.39	0.90	1.38	0.48
3	1, 1, 2, 3, 4, 4-hexachloro-1, 3-butadiene	7.8479	4.9711	1.024 5	4.78	3.76	3.46	-0.30
4	trichloroethylene	3.7883	2.8269	0.724 4	2.42	1.59	1.36	-0.23
5	1, 2-dichloroethane	4.5800	3.5800	0.0000	1.48	0.30	-0.10	-0.40
6	tetrachloromethane	6.1600	1.7418	0.0000	2.83	1.48	1.40	-0.08
7	trichloromethane	4.8700	2.1284	0.0000	1.97	0.78	0.57	-0.21
8	hexachloroethane	9.7400	2.7734	0.0000	4.14	2.92	2.65	-0.27
9	pentachloroethane	8.4500	2.838 2	0.0000	3.22	1.83	2.09	0.26
10	tetrachloroethylene	5.0449	2.975 4	0.724 4	3.40	1.74	2.06	0.32
11	benzene	1.3758	1.6058	1.0000	2.19	0.64	0.48	-0.16
12	toluene	1.9937	1.7833	1.0000	2.73	1.12	1.02	-0.10
13	ethylbenzene	2.6743	2.5083	1.0000	3.15	1.19	1.32	0.13

,

٠

		Tabl	e coutinued					
(A) [n. A 44	1	2		1 72		lgB _C	F
编号	化合物	¹ K	2Κ	G	lgK _{ow}	Exp.	Cal.	Er.
14	o-xylene	2.674 3	1.9937	1.000 0	3.12	1.24	1.51	0.27
15	<i>m</i> -xylene	2.674 3	1.9937	1.000 0	3.20	1.27	1.51	0.24
16	p-xylene	2.674 3	1.9937	1.000 0	3.15	1.27	1.51	0.24
17	isopropylbenzene	3.4038	2.6743	1.0000	3.72	1.55	1.80	0.25
18	biphenyl	3.0588	3.2249	1.4142	4.09	2.64	2.32	-0.32
19	1, 2, 3, 4-tetrachlorobenzene	5.101.4	3.1855	1.2547	4.64	3.72	3.25	-0.47
20	1, 2, 3, 5-tetrachlorobenzene	5.1014	3.1855	1.2547	4.92	3.36	3.25	-0.11
21	1, 2, 3-trichlorobenzene	4.0701	2.7822	1.2547	4.05	2.90	2.78	-0.12
22	1, 2, 4, 5-tetrachlorobenzene	5.1014	3.1855	1.2547	4.82	3.61	3.25	-0.36
23	1, 2, 4-trichlorobenzene	4.0701	2.7822	1.2547	4.02	2.95	2.78	-0.17
24	1, 2-dichlorobenzene	3.0922	2.3818	1.2547	3.43	2.43	2.27	-0.16
25	1, 3, 5-trichlorobenzene	4.0701	2.7822	1.2547	4.19	3.26	2.78	-0.48
26	1, 3-dichlorobenzene	3.0922	2.3818	1.2547	3.60	2.65	2.27	-0.38
27	1, 4-dichlorobenzene	3.0922	2.3818	1.2547	3.52	2.47	2.27	-0.20
28	hexachlorobenzene	7.2778	3.996 5	1.254 7	5.31	4.16	4.07	-0.09
29	2, 4, 5-trichlorotoluene	4.8656	2.9967	1.2547	4.56	3.87	3.18	-0.69
30	chlorobenzene	2.1854	1.9871	1.2547	2.84	1.85	1.70	-0.15
31	pentachlorobenzene	6.1736	3.5904	1.2547	5.18	3.45	3.68	0.23
32	1, 2, 4, 5-tetrabromobenzene	5.7288	3.6955	1.298 5	5.13	3.81	3.51	-0.30
33	1, 2, 4-tribromobenzene	4.5201	3.1704	1.298 5	4.66	3.66	3.02	-0.64
34	1, 3, 5-tribromobenzene	4.5201	3.1704	1.298 5	4.51	3.70	3.02	-0.68
35	1, 3-dibromobenzene	3.3738	2.6461	1.298 5	3.75	2.80	2.48	-0.32
36	1, 4-dibromobenzene	3.3738	2.6461	1.298 5	3.79	2.83	2.48	-0.35
37	bromobenzene	2.3136	2.1237	1.298 5	2.99	1.70	1.87	0.17
38	1, 2-dibromobenzene	3.3738	2.6461	1.298 5	3.64	2.70	2.48	-0.22
39	2, 2', 4, 5-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	5.69	5.00	4.62	-0.38
40	2, 2', 5, 5'-tetrachlorobiphenyl	6.5105	4.8485	1.7744	5.79	4.63	4.62	-0.01
41	2, 2', 5-trichlorobiphenyl	5.5825	4.4415	1.7744	5.55	4.01	4.28	0.27
42	2, 2', 4, 4'-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	6.29	4.02	4.62	0.60
43	2, 2'-dichlorobiphenyl	4.6934	4.0350	1.7744	4.90	3.26	3.93	0.67
44	2, 3′, 4′, 5-tetrachlorobipenyl	6.5105	4.848 5	1.7744	6.07	4.62	4.62	0.00
45	2, 3-dichlorobiphenyl	4.6934	4.0350	1.7744	5.02	4.25	3.93	-0.32
46	2, 4, 4'-trichlorobiphenyl	5.5825	4.4415	1.7744	5.58	4.63	4.28	-0.35
47	2, 4', 5-trichlorobiphenyl	5.5825	4.4415	1.7744	5.68	3.75	4.28	0.53
48	2, 4, 5-trichlorobiphenyl	5.5825	4.4415	1.7744	5.90	4.02	4.28	0.26
49	2,4'-dichlorobiphenyl	4.6934	4.0350	1.7744	5.10	3.55	3.93	0.38
50	2, 5-dichlorobiphenyl	4.6934	4.0350	1.7744	5.16	4.00	3.93	-0.07
51	3, 3', 4, 4'-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	6.63	3.90	4.62	0.72
52	3, 5-dichlorobiphenyl	4.6934	4.0350	1.7744	5.41	3.78	3.93	0.15

续表

续表

		Tabl	e coutinued					
	N. A 4L	1	3		1 17		lgB _{CF}	
编号	化合物	^{1}K	^{2}K	G	lgK _{ow}	Exp.	Cal.	Er.
53	4-chlorobiphenyl	3.849 7	3.629 4	1.774 4	4.63	2.69	3.56	0.87
54	4, 4'-dichlorobiphenyl	4.6934	4.0350	1.7744	5.58	3.28	3.93	0.65
55	2, 2', 3, 3'-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	5.67	4.23	4.62	0.39
56	2, 2', 4, 5, 5'-pentachlorobiphenyl	7.4721	5.2559	1.7744	6.65	5.40	4.93	-0.47
57	2, 2', 4, 4', 5, 5'-hexachlorobiphenyl	8.4631	5.6637	1.7744	7.75	4.83	5.23	0.40
58	2, 2', 4, 4', 6, 6'-hexachlorobiphenyl	8.4631	5.6637	1.7744	7.55	4.93	5.23	0.30
59	2, 2', 3, 3', 4, 4', 5, 5'-octachlorobiphenyl	10.5190	6.4798	1.7744	8.68	5.08	5.79	0.71
60	2, 2', 3, 5'-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	5.73	4.84	4.62	-0.22
61	2, 2', 4, 5'-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	5.87	4.84	4.62	-0.22
62	2, 2', 6, 6'-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	5.94	3.85	4.62	0.77
63	2, 2', 3, 4, 5'-pentachlorobiphenyl	7.4721	5.2559	1.7744	6.23	5.38	4.93	-0.45
64	2,2', 3', 4, 5-pentachlorobiphenyl	7.4721	5.2559	1.7744	6.67	5.43	4.93	-0.50
65	3, 3', 4, 4', 5-pentachlorobiphenyl	7.4721	5.2559	1.7744	6.98	5.81	4.93	-0.88
66	2, 2', 3, 3', 4, 4'-hexachlorobiphenyl	8.4631	5.6637	1.7744	6.96	5.77	5.23	-0.54
67	2, 2', 3, 3', 6, 6'-hexachlorobiphenyl	8.4631	5.6637	1.7744	7.03	5.43	5.23	-0.20
68	2, 2', 3, 4, 4', 5-hexachlorobiphenyl	8.4631	5.6637	1.7744	6.82	5.88	5.23	-0.65
69	2, 2', 3, 4, 5, 5'-hexachlorobiphenyl	8.4631	5.6637	1.7744	6.75	5.81	5.23	-0.58
70	2, 2', 3, 4, 4', 5', 6-heptachlorobiphenyl	9.4796	6.0717	1.7744	7.04	5.84	5.52	-0.32
71	2, 2', 3, 3', 4, 4', 5, 6-octachlorobiphenyl	10.5190	6.4798	1.7744	7.35	5.92	5.79	-0.13
72	2, 2', 3, 3', 4, 5, 5', 6-octachlorobiphenyl	10.5190	6.4798	1.7744	8.91	5.88	5.79	-0.09
73	2, 2', 3, 3', 5, 5', 6, 6'-octachlorobiphenyl	10.5190	6.4798	1.7744	7.73	5.82	5.79	-0.03
74	2, 2', 3, 3', 4, 4', 5, 5', 6-nonachlorobiphenyl	11.5780	6.8882	1.7744	9.14	5.71	6.05	0.34
75	2, 2', 5, 5'-tetrabromobiphenyl	7.0733	5.3376	1.8364	7.31	4.80	4.86	0.06
76	2, 4, 6-tribromobiphenyl	5.9881	4.8098	1.8364	6.42	3.93	4.51	0.58
77	4,4'-dibromobiphenyl	4.9509	4.2818	1.8364	5.72	4.19	4.14	-0.05
78	2, 4-dichlorophenol	3.814 3	2.5641	1.0777	5.53	2.00	2.29	0.29
79	pentachlorophenol	6.9927	3.786 3	1.0777	5.12	2.99	3.60	0.61
80	2, 4, 6-trichlorophenol	4.8333	2.9709	1.0777	3.69	2.43	2.76	0.33
81	2-chlorophenol	2.8526	2.1586	1.0777	2.15	2.33	1.76	-0.57
82	3-chlorophenol	2.8526	2.1586	1.0777	2.50	1.30	1.76	0.46
83	2-methylphenol	2.6461	1.9677	1.0777	1.95	1.03	1.67	0.64
84	phenol	1.9677	1.7556	1.0777	1.46	1.24	1.18	-0.06
85	4-t-buthylphenol	4.9386	2.6657	1.0777	3.31	2.07	2.93	0.86
86	2, 4-dimethylphenol	3.3738	2.1939	1.0777	2.30	2.18	2.12	-0.06
87	4-bromophenol	2.9899	2.2864	1.0777	2.59	1.56	1.82	0.26
88	p-sec-buthylphenol	4.9386	3.5361	1.0777	3.08	1.57	2.63	1.06
89	2-chloroaniline	2.8526	2.1586	0.8975	1.90	1.18	1.34	0.16
90	3-chloroaniline	2.8526	2.1586	0.8975	1.88	1.06	1.34	0.28
91	diphenylamine	3.6420	3.8748	0.8975	3.50	1.48	1.30	-0.18

•

		Tab	终衣 le coutinued					
	<u> </u>						lg Born	
编号	化合物	^{1}K	^{2}K	G	$\lg K_{\mathrm{ow}}$	Exp.	Cal.	Er
92	pentachloroaniline	6.9927	3.786 3	0.897 5	4.82	3.78	3.18	-0.60
93	2, 3, 4, 5-tetrachloroaniline	5.8960	3.3784	0.897 5	4.27	3.28	2.78	- 0.50
94	2, 3, 5, 6-tetrachloroaniline	5.8960	3.378 4	0.897 5	4.10	3.03	2.78	-0.25
95	2, 3, 4-trichloroaniline	4.833 3	2.9709	0.897 5	3.68	2.31	2.34	0.03
96	2, 4, 5-trichloroaniline	4.833 3	2.970 9	0.8975	3.45	2.61	2.34	-0.27
97	2, 4, 6-trichloroaniline	4.833 3	2.970 9	0.8975	3.52	2.73	2.34	-0.39
98	3, 4, 5-trichloroaniline	4.833 3	2.970 9	0.8975	3.32	2.70	2.34	-0.36
99	4-chloroaniline	2.8526	2.1586	0.8975	1.88	0.91	1.34	0.43
100	2, 4-dichloroaniline	3.814 3	2.5641	0.8975	2.78	1.98	1.87	-0.11
101	3, 4-dichloroaniline	3.814 3	2.5641	0.8975	2.78	1.48	1.87	0.39
102	aniline	1.9677	1.7556	0.8975	0.90	0.41	0.76	0.35
103	2-nitrophenol	2.6732	2.4474	0.9062	1.79	1.60	1.13	-0.47
104	2-chloronitrobenzene	2.8812	2.6626	1.0551	2.24	2.10	1.56	-0.54
105	3-chloronitrobenzene	2.8812	2.6626	1.0551	2.46	1.89	1.56	-0.33
106	4-chloronitrobenzene	2.8812	2.6626	1.0551	2.39	2.00	1.56	-0.44
107	2, 3-dichloronitrobenzene	3.7367	3.0679	1.0551	3.05	2.16	2.01	-0.15
108	2, 4-dichloronitrobenzene	3.7367	3.0679	1.0551	3.07	2.07	2.01	-0.06
109	2, 5-dichloronitrobenzene	3.7367	3.0679	1.0551	3.09	2.05	2.01	-0.04
110	3, 4-dichloronitrobenzene	3.7367	3.0679	1.0551	3.12	2.07	2.01	-0.06
111	3, 5-dichloronitrobenzene	3.7367	3.0679	1.0551	3.09	2.23	2.01	-0.22
112	2-methyl-4, 6-dinitrophenol	4.0093	3.382 9	0.5308	2.12	0.16	0.85	0.69
113	3-nitrophenol	2.6732	2.4474	0.9062	2.00	1.40	1.13	-0.27
114	pentachloronitrobenzene	6.6091	4.2890	1.0551	4.77	2.40	3.19	0.79
115	2, 3, 4, 5-tetrachloronitrobenzene	5.6099	3.8814	1.0551	4.57	1.89	2.82	0.93
116	2, 3, 5, 6-tetrachloronitrobenzene	5.6099	3.8814	1.0551	3.89	3.20	2.82	-0.38
117	2, 3, 4-trichloronitrobenzene	4.6498	3.474 3	1.0551	3.68	2.20	2.43	0.23
118	2, 4, 5-trichloronitrobenzene	4.6498	3.474 3	1.0551	3.48	1.84	2.43	0.59
119	4-nitroaniline	2.6732	2.4474	0.7547	1.39	0.64	0.77	0.13
120	3-nitroaniline	2.6732	2.4474	0.7547	1.37	0.92	0.77	-0.15
121	2-nitroaniline	2.6732	2.4474	0.7547	1.85	0.91	0.77	-0.14
122	2, 4, 6-tribromophenol	5.2981	3.344 2	1.0777	4.13	2.71	2.89	0.18

2 苯环因子

由表 2 可见有机物的 B_{CF} 与分子中所含 C — C 键数(f_j)以及苯环上所连取代基的类型密切相关.因此定义一个分子 j 的结构参数——苯环因子(G_j)

$$G_j = (f_j X_j)^{0.5}$$
 (8)

式中, f_j 为分子 j 中所含的 C — C 键数, X_j 为其苯环(或 C — C 体系)上所连取代基的校正系数. X_j 定义 为

$$X_{j} = \left[a^{b} \frac{x_{PE}}{x_{PC}} \left(\frac{N-1}{1+N_{X}^{4}} \right)^{0.5} \right]_{j}^{0.5d}$$

式中, $a = t_C/t_E$; $b = a^{-1}$; t_E 为取代基中最主要元素 E 的原子的自旋平行的成单电子数; x_{PF} 为 E 的 Pauling 电负性; N 为其最外层主量子数, N_x 为分子中所含硝基($-NO_2$)的数目; t_C, x_{PC} 为碳原子的相应 值. d 为苯环上取代基个数修正值,若苯环上无取代基, d=0;若有取代基, 则 d=1. 尚需指出,若苯环上 有多种取代基,则以有机物中取代基的先后顺序规则^[9]确定母体化合物,计算 X_i 是以母体化合物中的取 代基为准. 例如 3-甲氧基-2-氯苯胺,其母体为苯胺,则氨基($-NH_2$)的 X 求法为 $t_N = 3_x x_{PN} = 3.04$ 、 $t_{\rm C}=2_{X_{PC}}=2.55, N=2, d=1, N_{\rm X}=0; 其 a=2/3, b=3/2; 其 X_{\rm NH}=0.80556.$ 相应其分子 G=0.8975. 同法可得当 f=1 时取代基为—Cl、—Br、—OH 的 G; 值依次为 1.254 7、1.298 5、1.077 7, 当 f=2(即联 苯)时取代基为--Cl、--Br的 G; 分别为 1.774 5、1.836 4, 对于 f=1/3(单烯烃)、f=2/3(双烯烃)--Cl 的 G, 值依次分别为 0.724 4、1.024 5.

表 2 $\lg B_{CF} = E_* \Delta G$ 的最佳变量子集回归结果 Table 2 The results of F C and laR with Leans-and-Bounds regression

	Table 2 The I	$(Surfs of E_n)$ of unit	Prof. and period and	Doundo regression	
 序号	变量	n'	R	F	S
10	² K	122	0.881	417.97	0.724
11	^{1}K G	122	0.955	623.47	0.454
12	$G_{2}^{2}K_{1}^{1}K$	122	0.959	334.83	0.440

3 非离子性有机物的 Bcr与电拓扑指数的相关性

3.1 非离子性有机物的 Bcr 与电拓扑指数的数学模型

将文献[7]中 122 种非离子性有机物的 Bcr实验值及其 G、2K、1K 引入 SPSS 统计软件,通过逐步回 归和最佳变量子集回归,其结果见表 2. 表中 n'、R、F、S 依次为样本容量、相关系数, Fisher 检验值及估 计标准误差. 由表 3 可见, $\lg B_{CF} = 1^{K} \cdot G$ 二元模型的相关性系数为 0.955, 因此, 本文选用二元回归模型

$$\lg B_{\rm CF} = -1.433 + 0.401^{1}K + 1.897G \tag{10}$$

n' = 122, R = 0.955, F = 623.47, S = 0.454

式(10)为非离子性有机物的 Bcr-与电拓扑指数,苯环因子的数学模型按式(10)计算的 lgBcr值列于表 1 (Cal.),其与相应实验值基本吻合.本文模型(10)以及文献[7]中方程(4)对 122 种有机物 lgBcr的估算误 差(RES)均列于表 3. 本文的估算误差呈现出良好的正态分布.

	Table 3	Comparison of calculated	errors	
计算误差	0~0.3	0.3~0.6	0.6~0.9	0.9~
模型(10)	61	42	17	2
文献[7]	43	36	21	22

事? 计管理单比约

3.2 模型(10)的稳健性检验

为了检验模型(10)中是否存在"异常值",采用 Jackknifed 法^[10]对模型(10)进行稳健性检验.考虑所 研究的是大样本(样本容量大于 30),采用逐组剔除法.即每次分别从样本中删除化合物序号的个位上是 1、2、3、……、0的化合物,用余下的化合物 lg Bcr的建模,其结果见表 4. 这 10 个相关系数的简单平均值为 0.9554, 与式(10)的0.955非常接近, 并且R的波动范围很小. 说明模型(10)具有总体可接受的稳健性,

2008年

(9)

对这 122 个有机物,不存在明显的异常值.

表 4 模型(10)的稳健性检验

Table 4 The robustness tests of model (10) with Jackknife method

——————————————————————————————————————	1	2	3	4	5	6	7	8	9	0
R	0.952	0.957	0.957	0.957	0.957	0.955	0.954	0.957	0.957	0.951
<u> </u>	0.468	0.441	0.448	0.446	0.449	0.458	0.468	0.453	0.440	0.474

4 结果讨论

4.1 分子形状指数及G蕴含分子结构信息

根据表1可知,¹K 及²K 均随分子中非氢原子数增多而增大;对于结构相同的分子,¹K 及²K 是按--OH、---NH₂、--Cl、--Br 顺序递增的. 这与其化合物的疏水性参数(正辛醇/水分配系数)的递变趋势一致.

从 G 的计算公式可知,苯环因子蕴含分子结构信息,G 与分子中(折合)苯环数正相关;反映苯环上有 无取代基以及取代基的疏水性.如单个苯环上连有一Cl、一Br、一OH、一NH₂的 G 值依次为 1.254 7、 1.298 5、1.077 7、0.897 5,这些基团的正辛醇/水分配系数是依次减小的.如苯环上无取代基或只是由碳 原子组成的取代基,其 G 值为 1.

4.2 分子形状指数及共轭参数

目前非离子性有机物在鱼体内生物富集实验的数据处理多采有"二室模型"^[11](见图 1).即将实验测 定的模拟池分为水室和鱼室,假定有机物在池内的迁移转化为一级动力学过程.



图1 生物富集的"二室模型"

Fig.1 "Two compartments model" of bioconcentration

$$\frac{dC_{W}}{dt} = -(k_{2} + k_{0})C_{W} + k_{2}C_{f}$$
(11)
$$\frac{dC_{f}}{dt} = k_{1}C_{W} - k_{2}C_{f}$$
(12)

式中, C_{W}^{0} 为有机物在水体中的起始体积浓度, C_{W} 、 C_{t} 为 t 时刻有机物在水体及鱼体内体积浓度; k_{1} 、 k_{2} 为鱼体对有机物的吸收, 解吸速率常数, k_{V} 为有机物在水体中的挥发速率常数, 当此过程达到吸收与释放 平衡即生物富集平衡: $dC_{t}/dt = 0$, 即 $B_{CF} = C_{t}/C_{w} = k_{1}/k_{2}$. 生物富集模型与有机物在正辛醇相(相当于 鱼室)及水相(相当于水室)的平衡分配过程极为相似. 因此, 表征这 2 个过程的平衡常数(生物富集因子 与正辛醇/水分配系数)之间应存在密切相关系.

1) 分子的大小. 化合物分子被生物受体所富集的量的大小,显然与其间的分子力相关. 对于极性不太大的分子是以色散力为主,而色散力则与分子相对质量同向变化. 有机物分子的 *Mr* 越大,其与受体分子间的色散力越大,越易被受体吸附,其 *B*_{CF}越大. 新建拓扑指数¹K 及²K 均呈现与*M*,递增的趋势.

2) 不饱和键上所连取代基的性质,如苯环上连有一Br、一Cl、一OH、烷基、一NH₂ 等基团,相应化合物的 B_{CF}值是递减的.原因是这些基团与水分子间的作用力按此序递增;其中一OH 排在一R 之前,可能是一OH 更易与受体分子形成氢键所至.对于联苯也呈现类似规律.

上述 2 个因素恰是影响有机物疏水性(K_{OW})的主要因素,因此,有机物的 B_{CF} 应与 K_{OW} 密切相关.将 文献[7]中的 122 种有机物的正辛醇-水分配系数($\lg K_{ow}$,见表 1.)与其 $\lg B_{CF}$ 用最小二乘法拟合得其线性 方程为

$$lgB_{CF} = -0.407 + 0.781 lgK_{ow}$$
(13)

$$n' = 122, R = 0.930, F = 765.89, s = 0.564$$

其 R > 0.85,2 者呈显著相关,因此, $\lg K_{ow}$ 确是影响本文化合物在鱼体中生物富集的主要因素.笔者将 $\lg K_{ow} = h^1 K_{\Lambda} G$ 回归,建立的线性方程

 $\lg K_{\rm ow} = -0.959 + 0.525^{1}K + 2.088G \tag{14}$

n' = 122, R = 0.968, F = 891.38, s = 0.458

说明这 2 种结构参数高度蕴含 lgK_{ow},也必然会与这 122 种有机物的 lgB_{CF}高度相关,方程(10)证实此推断的正确性.另外"2.2"也证明方程(10)具有总体可接受的稳健性.可以认为这二种结构参数与常数项 共同揭示了影响有机物生物富集因子(B_{CF})的本质因素.

4.3 模型(10)的预测能力

一个好的数学模型,不仅要具有明确的物理意义、显著相关性及总体稳建性,而且应具有良好的预测 能力.笔者用模型(10)预测了文献[7]中43种未参与回归有机物的 lgB_{CF}值见表 5 中的"Pre.2",此值比 文献[7]给出的预测值(Pre.1)更为可信.理由之一:方程(10)的相关性明显好于文献[7]的式(4),即方程 (10)蕴含了影响 B_{CF}更为丰富的结构信息;理由之二:这 43 个预测值都与 122 个有机物中相应同分异物 体的 lgB_{CF}实验值更为接近.例如 2,2′,3,3′,4,5′,6,6′-八氯联苯的预测值为 6.15,与同分异物体 2,2′,3, 3′,4,5,5′,6′-八氯联苯的实验值 5.88 较为接近(这 2 个分子的形状应是非常相似的),而文献[7]给出的 预测值为 6.27,与 5.88 相差较远.

	化合物		F	0	lgB_{CF}		
骗兮		E ₃₇	Eg	G	Pre. 1	Pre.2	
1	4-chlorophenol	2.8526	2.1586	1.077 7	1.61	1.76	
2	2, 3-dichlorophenol	3.814 3	2.5641	1.0777	1.97	2.14	
3	2, 6-dichlorophenol	3.814 3	2.5641	1.0777	2.04	2.14	
4	3, 4-chlorophenol	3.814 3	2.5641	1.0777	2.12	2.14	
5	3, 5-dichlorophenol	3.814 3	2.5641	1.0777	1.97	2.14	
6	2, 3, 4-chlorophenol	4.833 3	2.9709	1.0777	2.57	2.55	
7	2, 3, 5-trichlorophenol	4.833 3	2.9709	1.0777	2.60	2.55	
8	2, 3, 6-trichlorophenol	4.833 3	2.9709	1.0777	2.49	2.55	
9	2, 4, 5-trichlorophenol	4.833 3	2.9709	1.0777	2.66	2.55	
10	3, 4, 5-trichlorophenol	4.833 3	2.9709	1.0777	2.59	2.55	
11	2, 3, 4, 5-tetrachlorophenol	5.8960	3.3784	1.0777	3.08	2.98	
12	2, 3, 4, 6-tetrachlorophenol	5.8960	3.3784	1.0777	3.04	2.98	
13	2, 3, 5, 6-tetrachlorophenol	5.8960	3.3784	1.0777	3.02	2.98	
14	2, 6-dichlorobiphenyl	4.6934	4.0350	1.7744	3.11	3.82	
15	3, 3'-biphenyl	4.6934	4.0350	1.7744	3.15	3.82	
16	3, 4-biphenyl	4.6934	4.0350	1.7744	3.24	3.82	
17	2, 2', 3-trichlorobiphenyl	5.582 5	4.4415	1.7744	3.68	4.17	

表 5 模型(10)对取代酚、多氯联苯 lgB_{CP}的预测值 Table 5 Predicted lg B_{CP} values of chlorophenols, PCBs, and alkylphenols with model(10)

编号

18 19

20

21

22

23

24

25

26

27

28

29 30

31

32

33

34 35

36

37

38

39

40

41

42

43

2, 2', 4, 4', 5-pentachlorobiphenyl

2, 2', 4, 6, 6'-pentachlorobiphenyl

2, 3, 3', 4, 4'-pentachlorobiphenyl

2, 3, 3', 4', 6-pentachlorobiphenyl

2, 3', 4, 4', 5-pentachlorobiphenyl

2, 2', 4, 5, 5', 6-hexachlorobiphenyl

2,2',3,3',4,5',6,6'-octachlorobiphenyl

2, 3, 4, 5, 6-pentachlorobiphenyl

2, 3-dimethylphenol 2, 5-dimethylphenol

2, 6-dimethylphenol

3, 4-dimethylphenol

	Table coutinued					
<i>μ</i> ~ μ	5	F	~	lgB _{CF}		
化合物	£ 37	Eg	G	Pre. 1	Pre.2	
2, 2', 4-trichlorobiphenyl	5.582 5	4.4415	1.7744	3.64	4.17	
2, 3, 3'-trichlorobiphenyl	5.5825	4.4415	1.7744	3.63	4.17	
2, 3, 4'-trichlorobiphenyl	5.582 5	4.4415	1.7744	3.74	4.17	
2, 4, 6-trichlorobiphenyl	5.5825	4,441 5	1.7744	3.58	4.17	
3, 4, 4'-trichlorobiphenyl	5.5825	4.4415	1.7744	3.72	4.17	
2, 2′, 3, 4-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	4.15	4.54	
2, 2′, 3, 5-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	4.26	4.54	
2, 2', 5, 6-tetrachlorobiphenyl	6.5105	4.8485	1.7744	4.28	4.54	
2, 3, 4, 4'-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	4.01	4.54	
2, 3, 4, 5-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	4.13	4.54	
2, 4, 4′, 5-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	4.26	4.54	
2, 3, 5, 6-tetrachlorobiphenyl	6.5105	4.848 5	1.7744	4.01	4.54	
2, 2', 3, 4, 4'-pentachlorobiphenyl	7.4721	5.2559	1.7744	4.63	4.93	
2, 2', 3, 5', 6-pentachlorobiphenyl	7.4721	5.2559	1.7744	4.76	4.93	

5.2559

5.2559

5.2559

5.2559

5.2559

5.2559

5.6637

6.4798

2.1939

2.1939

2.1939

2.1939

1.7744

1.7744

1.7744

1.7744

1.7744

1.7744

1.7744

1.7744

1.0777

1.0777

1.0777

1.0777

4.73

4.64

4.73

4.68

4.77

4.64

5.26

6.27

1.89

1.93

1.92

1.95

4.93

4.93

4.93

4.93

4.93

4.93

5.33

6.15

1.96

1.96

1.96

1.96

续表

综上所述,本文建立的苯环因子及分子形状指数不仅对于非离子性有机物在鱼体中的生物富集因子 的表征具有合理性与有效性,而且与其正辛醇-水分配系数也有良好的相关性.因此,苯环因子及分子形 状指数有望在物质构效关系研究中获得广泛的应用.

7.4721

7.4721

7.4721

7.4721

7.4721

7.4721

8.4631

10.5190

3.3738

3.3738

3.3738

3.3738

参考文献:

- [1] 徐士友, 储先萍, 汪海燕. 轨道平均能拓扑指数"K 及其应用[J]. 北京工业大学学报, 2006, 32(12): 1102-1105. XU Shi-you, CHU Xian-ping, WANG Hai-yan. Topological index "K of orbital average energy and its application [J]. Journal of Beijing University of Technology, 2006, 32(12): 1 102-1 105. (in Chinese)
- [2] 杨伟华, 冯长君. 脂肪醇的溶解度及生物活性的 QSPR/QSAR 研究[J]. 哈尔滨工业大学学报, 2005, 37(11): 1 552-1 554.

YANG Wei-hua, FENG Chang-jun. QSPR/QSAR study on the wa ter solubility and biological toxicity for fatty alcohol compounds[J]. Journal of Haerbing University of Technology, 2005, 37(11): 1 552-1 554. (in Chinese)

- [3] RANDIC M. On characterization of molecular branching[J]. J. Amer. Chem. Soc., 1975, 97(23); 6 609-6 615.
- [4] WIENER H. Structural determination of paraffln boihing points [J]. J. Am. Chem. Soc., 1947, 69(1); 17-20.

- [5] 向铮,梁逸曾,胡黔南. 甲基烷烃结构与色谱保留指数相关性的拓扑指数法研究[J]. 色谱, 2005, 23(2): 117-122. XIANG Zheng, LIANG Yi-zeng, HU Qian-nan. Relativity study on topoligical index of methyl alkane structures and chromatographic retention index[J]. Chin. of Chromatography, 2005, 23(2): 117-122. (in Chinese)
- [6] 许禄, 邵学广. 化学计量方法(第2版)[M]. 北京: 科学出版社, 2004: 389-405.
- [7] 沐来龙,冯长君,何红梅. 无机氢化物酸性的拓扑研究[J]. 武汉理工大学学报, 2006, 28(1): 119-122.
 MU Lai-long, FENG Chang-jun. HE Hong-mei. Topological research on the acidity of inorganic hydrides[J]. Journal of Wuhan University of Technology, 2006, 28(1): 119-122. (in Chinese)
- [8] 王青松,刘娟,徐美荣,等. 不饱和链烃沸点的 QSPR 研究[J]. 武汉理工大学学报, 2006, 28(3): 117-120.
 WANG Qing-song, LIU Juan, XU Mei-rong, et al. QSPR research on the boiling points of unsaturated Chain hydrocarbons
 [J]. Journal of Wuhan University of Technology, 2006, 28(3): 117-120. (in Chinese)
- [9] SACAN M T, ERDEM S S, OZPINAR G A. QSPR study on the bioconcentration factors of nonionic organic compounds in fish by characteristic root index and semi-empirical molecular descriptors[J]. J. Chem. Inf. Comput. Sci., 2004, 44(3): 985-992.
- [10] 许禄, 胡昌玉. 应用化学图论[M]. 北京: 科学出版社, 2000: 247-258.
- [11] 徐寿昌. 有机化学(第2版)[M]. 北京: 高等教育出版社, 1993: 248.
- [12] 沐来龙, 冯长君. 环价连接性指数与饱和烃沸点的 QSPR 研究[J]. 化工学报, 2004, 55(10): 1 702-1 705.
 MU Lai-long, FENG Chang-jun. Connectivity index of environment valence and QSPR research for boiling points of saturated hydrocarbon[J]. Journal of Chemical Industry and Engineering, 2004, 55(10): 1 702-1 705. (in Chinese)
- [13] 侯玲, 胡长敏, 赵晓明, 等. 江河中有机物在鲤鱼体内富集与释放的模拟研究[J]. 环境科学, 1997, 18(2): 34-37.
 HOU Ling, HU Chang-min, ZHAO Xiao-ming, et al. Simulating studies on bioaccumulation and elimination of the organic compounds of river[J]. Environmetal Science, 1997, 18(2): 34-37. (in Chinese)

Estimation and Prediction of Bio-concentration Factors of Nonionic Organic Chemicals in Fish by Molecular Shape Index and Benzene Ring Factor

FENG Chang-jun¹, YANG Wei-hua², MU Lai-long²

School of Chemistry & Chemical Engineering, Xuzhou Institute of Technology, Jiangsu Xuzhou 221008, China;
 School of Chemistry & Chemical Engineering, Xuzhou Normal University, Jiangsu Xuzhou 221116, China)

Abstract: Kier's molecular shape indices $({}^{1}K, {}^{2}K)$ are calculated for 165 nonionic organic compounds. On the basis of the characteristics of substituents and conjugated matrix, a novel molecular structure parameterbenzene ring factor (G) is defined and calculated for 165 molecules in this paper. A satisfactory relationship between the fish bio-concentration factor (B_{CF}) of 122 nonionic organic compounds and ${}^{1}K$, G is expressed as: $lgB_{CF} = -1.433 + 0.401 {}^{1}K + 1.897G$, R = 0.955, F = 623.47, S = 0.454, which could provide estimation and prediction for the B_{CF} of nonionic organic chemicals. Furthermore, the model is examined to validate overall robustness with Jackknife tests. All these regression results showed that the new parameter G and molecular shape indices have good rationality and efficiency for the fish bio-concentration factor. It is concluded that the ${}^{1}K, {}^{2}K$ and G will be used widely in quantitative structure-property/activity relationship (QSPR/QSAR) research.

Key words: nonionic organic compound; bio-concentration factor; Kier's molecular shape indices; benzene ring factor; QSAR

(责任编辑 张士瑛)