

## 取代芳烃对五种生物急性毒性的 QSAR 研究

冯长君\* 李鸣建 陈 艳 唐自强

(徐州教育学院化学系 徐州 221006)

**摘要** 提出成键原子的生物活性点价( $A_i$ ),并建构蕴涵有机物的生物活性的自相关拓扑指数( ${}^1Q$ ).使用 ${}^1Q$ ,距离指数( $L$ )分别与 16 ~ 30 种取代芳烃(含 F, Cl, Br, I,  $\text{NO}_2$ , OH,  $\text{NH}_2$ , 甲氧基, 甲醛基, 甲酰胺基, 乙酰基, 烷基等)对绿藻、斜生栅列藻、日本长腿蛙蝌蚪、发光菌、呆鲦鱼等的毒性数据关联,相应二元相关系数依次为 0.9361, 0.9559, 0.9733, 0.9354, 0.9599. 以上 QSAR 模型的估算结果均优于相应的文献成果.

**关键词** 生物活性点价, 自相关拓扑指数, 取代芳烃, 绿藻, 斜生栅列藻, 日本长腿蛙蝌蚪, 发光菌, 呆鲦鱼, 急性毒性

## The Acute Toxicities of Substituted Aromatics to Five Kinds of Organisms and QSAR Studies

FENG Chang-Jun\* LI Ming-Jian CHEN Yan TANG Zi-Qiang

(Department of Chemistry, Xuzhou Education College, Xuzhou, 221006)

**Abstract** A novel autocorrelation topological index( ${}^1Q$ ) based on the bioactivity valence ( $A_i$ ) of atom  $i$  and substituent distance index ( $L$ ) are derived in this paper. The quantitative structure - activity relationships (QSAR) studies are performed with the acute toxicities of 16 ~ 30 substituted aromatics (containing substituents such as: F, Cl, Br, I,  $\text{NO}_2$ , OH,  $\text{NH}_2$ ,  $\text{OCH}_3$ , CHO,  $\text{CONH}_2$ ,  $\text{COCH}_3$  and alkyl) and  ${}^1Q$ ,  $L$ . Their bivariate correlation coefficients for green algae, *Scenedsmus obliquus*, *Rana Japonica tadpoles*, *Photobacterium phosphoreum*, *Pimephales* are 0.9361, 0.9559, 0.9733, 0.9354, 0.9599 respectively. The estimated results with above QSAR models are all better than the relevant achievements of literatures.

**Keywords** bioactivity valence, autocorrelation topological index, substituted aromatics, green algae, *Scenedsmus obliquus*, *Rana Japonica tadpoles*, *Photobacterium phosphoreum*, *Pimephales*, acute toxicity

近年来,环境中有机污染物的定量结构-生物活性相关(QSAR)已引起人们的普遍关注<sup>[1,2]</sup>. QSAR 为人们提供一种简便,经济,可以信赖的方法,被用于有机物的生态风险评价<sup>[3]</sup>. 并在预测有机物的生物活性,探讨毒性作用机理中发挥重要作用<sup>[4,5]</sup>. 目前具有代表性的 QSAR 方法有:辛醇/水分配系数

法,线性溶剂化能法,拓扑指数法等<sup>[5]</sup>. 笔者根据有机物的生物活性与其组成原子的关系,建立揭示有机物生物活性的自相关拓扑指数( ${}^1Q$ );并以苯环为母体,提出取代基间的距离指数( $L$ ). 将 ${}^1Q$ ,  $L$ 用于多种类型取代芳烃的生物活性相关,结果为:与 16 种氯代芳烃(含氨基,羟基,硝基,氰基等)对绿藻毒

\* E-mail: wxyl68@pub.xz.jsoinfo.net

收稿日期:2000-10-08,修回日期:2000-12-18,定稿日期:2001-03-18

(Received October 8, 2000. Revised December 18, 2000. Accepted March 18, 2001)

性的相关系数( $R$ )为0.9361,与18种苯酚、苯胺类化合物对斜生栅列藻急性毒性的 $R$ 为0.9559,与21种苯酚类化合物对日本长腿蛙蝌蚪急性毒性的 $R$ 为0.9733,与30种取代芳烃(氯代苯,烷基苯,苯胺类,苯酚类)对发光菌急性毒性的 $R$ 为0.9354,与20种取代芳烃(同上)对呆鲮鱼急性毒性的 $R$ 为0.9599.所得的QSAR模型均优于相应的文献报道结果<sup>[1,6-8]</sup>,为定量评估和预测同类其他化合物的生物毒性提供了参考依据.

## 1 ${}^mQ$ 及 $L$ 的建立

拓扑指数是以图论为基础,根据分子的某些结构特征构造的数学量,籍以定量表征不同分子.自Wiener<sup>[9]</sup>以来,迄今已报道了100多种拓扑指数.其中以由Randic<sup>[10]</sup>提出,Kier等<sup>[11]</sup>进一步发展的分子连接性指数( ${}^m\chi$ )应用最为广泛,但对自相关拓扑指数( ${}^mF$ )<sup>[12]</sup>研究较少.

在数学上,函数 $Y=f(x)$ 的自相关函数 $F(t)$ 定义为:

$$F(t) = \int f(x)f(x+t)dx \quad (1)$$

若式中 $F(x)$ 是描述随 $x$ 变化的某种理化性质,则可认为 $F(x)$ 反映了该种理化性质的分布情况.如把分子以隐氢图表示,其自相关函数 $F(t)$ 可表示为 ${}^mF$ ,并称为自相关拓扑指数:

$${}^mF = F(t) = \sum f(i)f(j) \quad (2)$$

式中 $i, j$ 为隐氢图中相距路径长度为 $t$ 的两个(非氢)原子, $f(i), f(j)$ 为 $i, j$ 原子的某种理化性质.众所周知,每种成键原子都具有一定的理化性质,正是这些原子理化性质的分布情况影响了整个分子的理化性质.因此, ${}^mF$ 描述了这种分布情况,从而揭示了分子的有关结构信息.

作者发现有机化合物的生物活性与其组成原子的结构特征密切相关,由此定义原子的生物活性点价( $A_i$ )为:

$$A_i = m_i \left[ n_i - w_i (1 + k_i)^{0.125} \times (n_i - 1)^{0.3\beta_i} \right] - h_i^{1.5} \quad (3)$$

式中:  $w_i = x_{pi} p_i e_i / x_{pc} p_c e_c$

$$\beta_i = (n_i - 2)^{g_i}$$

$$g_i = 0.5(n_i - 3), n_i \geq 3 \quad (4)$$

式(3),(4)中: $m_i, n_i, p_i, e_i$ 依次为原子 $i$ 的价电子数,电子层数,价电子层中 $p$ 电子数,自旋平行的单电子数, $x_{pi}$ 为原子 $i$ 的电负性(Pauling标度), $k_i$ 为

成键原子 $i$ 的孤电子对数. $x_{pc}, p_c, e_c$ 为碳原子的相应值.氯原子的 $m, n, p, e, x_p, k$ 依次为7, 3, 5, 1, 3, 16, 3, 碳原子的 $x_{pc}, p_c, e_c$ 依次为2.55, 2, 2, 将这些数据代入式(4)及式(3)得Cl原子的 $A_{Cl} = 5.12472$ .同法求得F, Br, I的 $A_i$ 依次为-2.24084, 8.74948, -2.79732, 硝基中N, O原子的 $A_i$ 为-4.62563, -6.57119, 羟基中O原子的 $A_O$ 为-7.57119, 氨基中N原子的 $A_N$ 为-7.45406.对于碳原子,式(3)可简化为:

$$A_{ci} = m_c - h_{ci}^{1.5} = 4 - h_{ci}^{1.5} \quad (5)$$

相应伯,仲,叔,季碳原子的 $A_{ci}$ 依次为-1.19615, 1.17157, 3, 4.

在分子图的邻接矩阵基础上,将 $A_i$ 代 $f(i)$ ,  $A_j$ 代 $f(j)$ ,建构蕴含有机物生物活性因素的自相关拓扑指数( ${}^mQ$ ):

$${}^mQ = \sum A_i \cdot A_j \quad j - i = 0, 1, 2, \dots, m \quad (6)$$

式中“ $j - i$ ”为 $i$ 与 $j$ 两原子间的路径长度,即自相关拓扑指数的阶数. ${}^mQ$ 是一个系列,由 $m + 1$ 个指数构成.其中0阶项( ${}^0Q$ ),一阶项( ${}^1Q$ )的计算公式为:

$${}^0Q = \sum A_i^2 \quad j = i \quad (7)$$

$${}^1Q = \sum A_i \cdot A_j \quad j = i + 1 \quad (8)$$

式(7)中“ $\sum$ ”是对分子图中每个节点(即非氢原子)求和,式(8)中“ $\sum$ ”是对分子图中每条边(即化学键)求和.例如邻氯苯酚的 ${}^0Q, {}^1Q$ 为:

$${}^0Q = (-7.57119)^2 + 5.12472^2 + 2 \times 4^2 + 4^3 = 151.586$$

$${}^1Q = -7.57119 \times 4 + 5.12472 \times 4 + 4 \times 4 + 2 \times (4 \times 3) + 3 \times (3 \times 3) = 57.214$$

由于 ${}^0Q$ 对有机物的结构选择性及生物活性相关性较低,本文仅研究 ${}^1Q$ .以苯环为母体的取代基距离指数( $L$ )的计算规则:(1)遵照Ingold等<sup>[13]</sup>提出的“取代基次序规则”,对同一苯环上不同取代基排序,以最优取代基为始点.如苯环上连有:OH, Cl, NH<sub>2</sub>,这3个基团的顺序为:Cl > OH > NH<sub>2</sub>,应以Cl为始点.若有2个或以上个最优基团,应以至其它取代基的距离之和最小的那个最优基团为始点.(2)最优基团与另一基团中离苯环最远的非氢原子(K)之间的最短距离以“ $L_k$ ”表示,则:

$$L = \sum L_k \quad (9)$$

例如对氯邻硝基苯酚:氯原子与硝基中氧原子之间的最近距离为5(即其间的边数),Cl与OH中氧原子(不考虑氢原子)之间的边数为5,则该分子的 $L = 2 \times 5 + 5 = 15$ .

## 2 氯代芳烃衍生物对绿藻的毒性与<sup>1</sup>Q, L 的相关性

将文献[6]中 16 种氯代芳烃衍生物对绿藻的毒性 ( $-\lg EC_{50}$ ,  $EC_{50}$  的单位为  $\text{mmol} \cdot \text{L}^{-1}$ ) 实验值及其<sup>1</sup>Q, L 列于表 1. 用最小二乘法拟合它们之间的线性回归方程为:

$$-\lg EC_{50} = 0.0136 + 0.0143 {}^1Q \quad (10)$$

$$n' = 16, r = 0.8642, F = 41.32, S = 0.3372$$

式中  $n'$ ,  $r$ ,  $F$ ,  $S$  依次为化合物数, 相关系数, Fischer 检验值, 估计标准误差.

$$-\lg EC_{50} = 0.2819 + 0.1110 L \quad (11)$$

$$n' = 16, r = 0.7598, F = 19.21, S = 0.4358$$

而文献[7]对所研究的化合物分别使用 Kier 的<sup>2</sup>X<sup>+</sup>及 $\Delta {}^2X^+$ (非色散力因子)进行回归的最佳(一元)相关系数仅为 0.736. 建立其二元曲线模型, 相关程度有很大改善:

$$-\lg EC_{50} = 0.9048 + 3.1567 \times 10^{-8} {}^1Q^{3.5} - 2.0777 \times 10^{-6} L^4 \quad (12)$$

$$n' = 16, R = 0.9361, F = 46.05, S = 0.2358$$

与文献[7]逐步回归所得的方程(2)比较,  $F$  值提高了 63.8%,  $S$  值下降了 26.8%. 按式(12)给出的计算值(表 1 中的 Cal. 1)与其实验值(Exp.)较为吻合. 为便于比较, 表 1 中还列出文献[6]的拟合值(Cal. 2).

表 1 氯代苯衍生物对绿藻的毒性与<sup>1</sup>Q, L 的相关性

Table 1 A correlativity between <sup>1</sup>Q, L and toxicity data of chlorinated aromatics to Green Algae

No.	氯代苯衍生物	<sup>1</sup> Q	L	$\lg K_{OW}$	$-\lg EC_{50}$		
					Exp.	Cal. 1	Cal. 2
1	2,4-二氯苯胺	84.182	7	2.910	0.818	1.073	1.132
2	3,4-二氯硝基苯	156.287	12	3.090	2.283	2.368	2.042
3	2,4-二氯苯酚	83.713	7	3.060	1.183	1.069	1.005
4	4-氯苯甲醛	78.785	6	2.210	0.659	1.039	0.728
5	3-氯硝基苯	128.788	10	2.470	1.870	1.649	1.446
6	3,4-二氯苯胺	84.182	7	2.690	1.292	1.073	1.132
7	4-氯-1,3-二硝基苯	178.077	20	2.141	2.804	2.951	2.539
8	五氯苯酚	168.210	19	5.120	2.770	2.583	2.794
9	4-氯-2-硝基苯胺	105.972	15	2.720	1.005	1.186	1.629
10	5-氯-2-羟基苯胺	33.398	9	1.810	0.845	0.898	0.593
11	4-氯-2-硝基苯酚	105.503	15	2.470	1.328	1.180	1.503
12	2-氯苯甲酰胺	47.398	8	0.640	0.960	0.920	0.875
13	4-氯苯甲腈	83.996	6	2.288	0.583	1.074	1.103
14	3,4-二氯苯甲腈	111.495	8	2.881	1.464	1.358	1.700
15	4-氯-3-氯苯胺	54.719	7	2.060	1.358	0.938	1.132
16	4-氯碘苯	75.310	5	3.978	1.158	1.021	0.948

## 3 苯酚、苯胺及其衍生物对斜生栅列藻的急性毒性与<sup>1</sup>Q, L 的数学模型

将 18 种苯酚、苯胺及其衍生物对斜生栅列藻的急性毒性 ( $-\lg EC_{50}$ ) [7] 及其<sup>1</sup>Q, L 列于表 2. 拟合它们之间的线性回归方程为:

$$-\lg EC_{50} = 2.0473 + 0.0170 {}^1Q \quad (13)$$

$$n' = 18, R = 0.9284, F = 99.87, S = 0.2636$$

$$-\lg EC_{50} = 2.7768 + 0.0763 L \quad (14)$$

$$n' = 18, R = 0.8160, F = 31.87, S = 0.4100$$

$$-\lg EC_{50} = 2.1257 + 0.0134 {}^1Q + 0.0247 L \quad (15)$$

$$n' = 18, R = 0.9443, F = 61.80, S = 0.2333$$

式(13)和(15)的相关性都明显优于文献[8]中式(5)的 0.903. 建立二元曲线模型为:

$$-\lg EC_{50} = 0.0721 + 0.6955 {}^1Q^{0.35} - 0.0225 L \quad (16)$$

$$n' = 18, R = 0.9559, F = 79.48, S = 0.2083$$

其  $S$  仅有文献[7]中式(5)的 66.3%. 表 2 中的“Cal. 1”为式(16)的计算值, “Cal. 2”为文献[7]的估

算值.

表 2 18 种酚类, 胺类化合物对斜生栅列藻的毒性值及其<sup>1</sup>*Q*, *L*Table 2 The parameters <sup>1</sup>*Q*, *L* of 18 aniline and phenol derivatives and the toxicities to *Scenedmus obliquus*

No.	化合物	<sup>1</sup> <i>Q</i>	<i>L</i>	lg <i>K</i> <sub>OW</sub>	-lgEC <sub>50</sub>		
					exp.	Cal. 1	Cal. 2
1	苯酚	29.715	1	1.46	2.46	2.37	2.72
2	2,4-二氯苯酚	83.713	8	2.92	3.62	3.53	3.69
3	2,4,6-三氯苯酚	111.212	11	3.69	3.81	3.94	4.18
4	五氯苯酚	168.210	19	5.12	4.63	4.68	5.00
5	对-硝基苯酚	78.005	12	1.92	3.57	3.54	3.81
6	2,4-二硝基苯酚	127.294	20	1.54	4.16	4.31	4.18
7	苯胺	30.184	1	1.03	2.56	2.39	2.36
8	邻-甲苯胺	32.399	3	1.40	2.34	2.49	2.57
9	邻-氯苯胺	57.683	3	1.76	2.89	3.01	2.94
10	间-氯苯胺	56.683	4	1.88	2.79	3.02	3.00
11	2,3-二氯苯胺	85.182	6	2.44	3.98	3.50	3.40
12	2,4-二氯苯胺	84.182	7	2.80	3.74	3.51	3.50
13	2,5-二氯苯胺	84.182	8	2.80	3.82	3.53	3.55
14	间-溴苯胺	71.182	4	2.10	2.80	3.26	3.05
15	2,4,6-三溴苯胺	155.177	11	4.03	4.37	4.38	3.97
16	间-硝基苯胺	78.473	10	1.37	3.48	3.50	3.56
17	对-硝基苯胺	78.473	12	1.39	3.40	3.54	3.41
18	2,4-二硝基苯胺	127.762	32	2.38	4.68	4.59	4.2

表 3 21 种苯酚类化合物对蝌蚪的急性毒性与<sup>1</sup>*Q*, *L* 的相关性Table 3 The relationships between <sup>1</sup>*Q*, *L* and the toxicity data of phenol derivatives to *Rana Japonica tadpoles*

No.	化合物	<sup>1</sup> <i>Q</i>	<i>L</i>	lg <i>K</i> <sub>OW</sub>	24 h - LC <sub>50</sub>		
					exp.	Cal. 1	Cal. 2
1	邻硝基酚	79.005	8	1.91	3.686	3.788	3.940
2	间硝基酚	78.005	10	1.91	3.482	3.816	3.731
3	对硝基酚	78.008	12	1.91	4.052	3.915	3.935
4	4-氯-2-硝基酚	105.503	15	2.55	4.880	4.706	4.648
5	邻氯酚	57.214	3	2.15	3.024	3.370	3.385
6	对氯酚	56.214	5	2.16	3.308	3.352	3.183
7	2,4-二氯苯酚	83.713	8	2.8	3.996	3.869	3.979
8	对溴酚	70.713	5	2.4	3.619	3.623	3.478
9	2-溴-4-甲基苯酚	72.929	8	2.95	3.753	3.682	3.941
10	对氯苯酚	26.752	5	1.71	2.670	2.709	2.391
11	对甲氧基酚	17.291	6	1.59	2.371	2.456	2.511
12	邻甲氧基酚	18.291	4	1.34	2.500	2.480	2.415
13	对甲基酚	30.931	5	2.06	2.928	2.811	2.946
14	邻甲基酚	31.931	3	2.06	2.819	2.833	2.951
15	2,6-二甲基酚	34.146	6	2.61	3.331	2.890	3.325
16	对羟基苯甲醛	28.002	6	1.23	2.804	2.743	2.783
17	间苯二酚	5.431	4	1.03	2.077	2.038	2.062
18	间氨基苯酚	5.899	4	0.24	2.064	2.058	2.028
19	对羟基苯乙酮	20.646	12	1.19	2.509	2.712	2.799
20	对特丁基苯酚	37.361	18	3.31	4.170	4.217	4.134
21	苯酚	29.715	1	1.51	2.804	2.780	2.611

表 4 取代芳烃的毒性数据及 ${}^1Q, L$ Table 4 The parameters  ${}^1Q, L$  and toxicity data of substituted aromatics

No.	取代芳烃	${}^1Q$	$L$	$\lg K_{OW}$	$-\lg EC_{50}$			$-\lg LC_{50}$		
					Exp.	Cal. 1	Cal. 2	Exp.	Cal. 1	Cal. 2
1	1,2,4,5-四氯苯	161.995	12	5.05	5.51	5.35	5.54	5.85	5.62	5.14
2	1,2,4-三氯苯	134.497	8	4.29	4.50	4.80	4.72	5.00	5.01	4.71
3	1,2,3-三氯苯	135.497	7	4.27	4.53	4.80	4.68	4.89	4.95	4.67
4	1,4-二氯苯	106.998	5	3.59	4.39	4.32	4.34	4.62	4.46	4.32
5	1,3-二氯苯	106.998	4	3.38	4.24	4.30	4.21	4.30	4.39	4.19
6	1,2-二氯苯	107.998	3	3.35	4.38	4.29	4.31	4.40	4.33	4.28
7	氯苯	80.499	1	2.81	3.86	3.88	3.90	3.77	3.86	3.86
8	1,3-二溴苯	135.996	4	3.75	4.99	4.73	4.41			
9	对溴氯苯	121.497	5	3.83	4.50	4.53	4.47			
10	2,4,5-三氯甲苯	136.712	12	4.93	4.86	4.93	5.10			
11	2,5-二氯甲苯	109.213	8	4.04	4.38	4.42	4.61			
12	对氯甲苯	81.714	5	3.31	3.88	3.98	4.22	4.33	4.16	4.21
13	对二甲苯	56.431	5	3.09	3.68	3.70	4.14	4.21	3.86	4.12
14	间二甲苯	56.431	4	3.09	3.65	3.67	4.10			
15	溴苯	94.998	1	2.99	3.78	4.06	3.99	3.89	4.04	3.96
16	苯	54.000	0	2.13	3.34	3.56	3.50	3.40	3.48	3.45
17	2,4,5-三氯苯胺	111.680	12	3.04	4.51	4.55	4.35			
18	2,6-二氯苯胺	85.182	6	2.75	4.16	4.06	4.21			
19	2,4-二氯苯胺	84.182	7	2.75	4.09	4.06	4.22			
20	3,4-二氯苯胺	84.182	8	2.55	4.20	4.08	4.10	4.33	4.41	4.12
21	3-氯-4-氟苯胺	54.719	8	2.04	3.28	3.75	3.85			
22	对氯苯胺	56.683	4	1.90	3.57	3.68	3.76			
23	对溴苯胺	71.182	5	2.05	3.92	3.86	3.84	3.56	4.04	3.85
24	苯胺	30.184	1	0.98	3.28	3.37	3.26	2.84	3.27	3.26
25	五氯苯酚	168.210	19	5.04	5.69	5.62	5.25	6.06	6.19	5.27
26	2,4-二氯苯酚	83.713	8	2.90	4.45	4.08	4.11	4.30	4.40	4.11
27	邻氯苯酚	57.214	3	2.18	4.14	3.66	3.71	4.02	3.73	3.69
28	邻甲基苯酚	31.931	3	1.37	3.75	3.43	3.30	3.77	3.43	3.28
29	间二苯酚	5.431	4	0.80	3.00	3.30	3.02	3.04	3.19	2.99
30	苯酚	29.715	1	1.46	3.64	3.37	3.31	3.51	3.26	3.29

#### 4 苯酚类化合物对日本长腿蛙蝌蚪的急性毒性与 ${}^1Q, L$ 的构效关系

将文献[8]中 21 种苯酚类化合物对日本长腿蛙蝌蚪急性毒性(24 h- $LC_{50}$ , 24 h 的每个化合物的半数致死浓度的负对数值)实验值与 ${}^1Q, L$ 列于表 3. 用最小二乘法拟合的线性方程为:

$$24\text{ h} - LC_{50} = 2.1330 + 0.0228 {}^1Q \quad (17)$$

$$n' = 21, r = 0.8897, F = 72.17, S = 0.3303$$

文献[8]选取了 5 种理化参数: 正辛醇/水分配系数( $\lg K_{OW}$ ), 分子最低未占轨道能( $E_{LUMO}$ ), 电子能( $E_E$ ), 离解常数( $pK_a$ )和分子生成热(HOF), 它们与 24 h- $LC_{50}$ 的  $r$  依次为: 0.7987, 0.5878, 0.6202, 0.5728, 0.5784. 均比 ${}^1Q$ 的相关性差.

$$24\text{ h} - LC_{50} = 2.3487 + 0.1184 L \quad (18)$$

$$n' = 21, r = 0.6784, F = 16.20, S = 0.5316$$

$L$  的相关性仅比  $\lg K_{OW}$  差, 而优于其它 4 种参数.

$$24\text{ h} - LC_{50} = 1.8866 + 0.0188 {}^1Q + 0.0612 L \quad (19)$$

$$n' = 21, R = 0.9434, F = 72.86, S = 0.2399$$

文献[9]用  $\lg K_{OW}$  分别和余下 4 种参数同 24 h- $LC_{50}$  回归的最佳二元相关系数为 0.9411, 比本文结果稍差. 建立 ${}^1Q, L$  的曲线模型, 相关程度有所改善:

$$24\text{ h} - LC_{50} = 1.6694 + 0.1225 {}^1Q^{0.65} + 6.6625 \times 10^{-7} L^5 \quad (20)$$

$$n' = 21, R = 0.9733, F = 161.59, S = 0.1662$$

优于文献[8]中式(13)的相关性(该式含 3 个自

变量.笔者按其数据拟合的  $R$  仅为 0.9448.),其  $S$  只有文献[8]中式(13)的 75.2%.表 3 中的“Cal.1”为式(20)的计算值,其中的“Cal.2”为文献[8]的拟合值.

## 5 取代芳烃对水生生物的急性毒性与 $^1Q, L$ 的定量关系

将文献[1]中 30 种取代芳烃化合物对发光菌的  $-\lg EC_{50}$  和呆鲦鱼的  $-\lg LC_{50}$  及其 $^1Q, L$ 列于表 4.它们之间的回归方程如下:

$$-\lg EC_{50} = 2.8910 + 0.0142 ^1Q \quad (21)$$

$$n' = 30, r = 0.9225, F = 160.04, S = 0.2366$$

$$-\lg EC_{50} = 3.5022 + 0.1103 L \quad (22)$$

$$n' = 30, r = 0.7236, F = 30.77, S = 0.4233$$

$$-\lg EC_{50} = 2.8931 + 0.0121 ^1Q + 0.0313 L \quad (23)$$

$$n' = 30, R = 0.9354(0.88^{[1]}),$$

$$F = 94.57, S = 0.2167$$

$$-\lg LC_{50} = 2.7722 + 0.0170 ^1Q \quad (24)$$

$$n' = 20, r = 0.9205, F = 99.88, S = 0.3124$$

$$-\lg LC_{50} = 3.4007 + 0.1561 L \quad (25)$$

$$n' = 20, r = 0.8493, F = 46.57, S = 0.4222$$

$$-\lg LC_{50} = 2.8378 + 0.0119 ^1Q + 0.0717 L \quad (26)$$

$$n' = 20, R = 0.9599(0.95^{[1]}),$$

$$F = 99.52, S = 0.2243$$

文献[1]使用疏水性参数( $\lg K_{OW}$ ),最高占据分子轨道能( $E_{HOMO}$ )分别与  $-\lg EC_{50}$ 、 $-\lg LC_{50}$  关联,相关系数依次为 0.88, 0.95, 均差于本文的 0.9354, 0.9599.按式(23)、(26)给出的计算值为表 4 中的“Cal.1”,表 4 中的“Cal.2”为文献[1]的计算值.对于未参与回归的六氯苯,间氯苯甲醛,对氯苯甲醛,蔡,  $\alpha$ -氯蔡,式(23)给出的预测值<sup>[14]</sup>依次为: 6.11, 3.95, 3.96, 4.32, 4.67, 与其实验值: 6.32, 4.00, 4.15, 4.86, 4.99 较好地吻合.

## 6 结果与讨论

### 6.1 $A_i$ 是表征成键原子生物活性的特征值

考察表 1~表 4 中 85 种取代芳烃(含少数重复,实为 68 种)对绿藻,斜生栅列藻,日本长腿蛙蝌蚪,发光菌,呆鲦鱼等生物的毒性发现:

(1)苯环上连有氯,溴等官能团的取代芳烃,其生物毒性较大,通常比其母体的毒性还大;并按 Cl, Br 的顺序增强.笔者称此类基团为增毒基团.例如表

3 中的苯酚,对氯酚,对溴酚的 24h- $LC_{50}$  依次为 2.804, 3.308, 3.619.在其它表中也可看到类似现象.苯环上的增毒基团越多,其生物毒性(一般)也越大.例如表 4 中氯代苯,随着苯环上氯原子数增多,其  $-\lg EC_{50}$ 、 $-\lg LC_{50}$  依次增大.

(2)苯环上连有羟基,氨基,氟,碘,伯碳等官能团的取代芳烃,其生物毒性较低,通常比其母体芳烃还低.这类基团可谓降毒基团.苯环上的解毒基团越多,其毒性越小.例如表 3 中的苯酚,对氟苯酚,间氨基苯酚,间苯二酚的 24h- $LC_{50}$  依次为 2.804, 2.670, 2.064, 2.077.又如表 2 中的苯酚,苯胺,邻甲苯胺的  $-\lg EC_{50}$  为 2.46, 2.56, 2.34.

为了表征成键原子的生物毒性行为,笔者按式(2)定义生物活性点价( $A_i$ ).  $A_i > 0$  的非氢原子具有毒性,且  $A_i$  值越大,毒性越强.如 Cl, Br 的  $A_i$  值依次增大,相应毒性也增强.  $A_i < 0$  的原子没有毒性,且  $A_i$  越小,降低毒性(即似解毒)作用越大.如  $CH_3$ , F, I,  $NH_2$ , OH 中 C, F, I, N, O 的  $A_i$  值依次减小,相应解毒作用也越强.

根据式(2)将影响  $A_i$  的各种因素列于表 5.由表 5 可见,在电子层数( $n_i$ )相同情况下,  $A_i$  与  $m_i$ ,  $x_{pi}$ ,  $e_i$ ,  $h_i$ ,  $k_i$ ,  $p_i$  等因素负相关,如序号 4~10.在  $m_i$ ,  $k_i$ ,  $p_i$ ,  $e_i$ ,  $h_i$  等因素相同情况下,  $A_i$  与  $x_{pi}$  负相关,与  $n_i$  呈抛物线(性)相关.如序号 1~4 的  $A_i$  与  $n_i$ ,  $x_{pi}$  即为抛物线关系.总体上看,影响  $A_i$  的主要因素为  $n_i$ ,  $m_i$ ,  $x_{pi}$ .

Hall 等<sup>[5]</sup>指出,苯环上所连取代基的吸电子能力及其离去能力影响取代芳烃的毒性.连在苯环上的取代基的吸电子能力应为其诱导效应( $I$ )与共轭效应( $C$ )的综合效果.通常,  $I$  与  $\chi_p$  正相关,  $C$  与  $n$  负相关.对于 F, Cl, Br, I 的  $m_i$ ,  $k_i$ ,  $p_i$ ,  $e_i$ ,  $h_i$  均相同,仅有  $n_i$ ,  $\chi_{pi}$  不同.其  $n_i$  依次增大,提供电子给苯环的能力依次降低,相应  $C$  依次减小.然而其  $\chi_{pi}$  依次减小,拉苯环上电子的能力依次降低,相应  $I$  减小.卤原子对苯环的吸电子能力应为其  $I$  与  $C$  的综合效果.即 F, I 较小, Cl, Br 较大.结合他们以  $X^-$  离子形式的离去能力按 F, Cl, Br, I 顺序增大,因此, Cl, Br 的毒性较大,而 F, I 没有毒性.

对于  $CH_3$ ,  $NH_2$ , OH 中 C, N, O 的  $n_i$  相同,而  $\chi_{pi}$ ,  $p_i$  依次增大,对苯环上的  $I$ ,  $C$  递增,综合效果是对苯环的供电子能力递增.结合其离去的产物为  $CH_4$ ,  $NH_3$ ,  $H_2O$ , 故其不仅没有毒性,而且有解毒的作用.综上所述,按式(3)建立的  $A_i$  不仅符合 Hall 等的观点,也与实验事实基本一致.

表 5 影响  $A_i$  的各种因素Table 5 The manifold factors affecting  $A_i$ 

序号	原子	$n_i$	$m_i$	$k_i$	$\tau_i$	$p_i$	$e_i$	$h_i$	$A_i$
1	Cl	3	7	3	3.16	5	1	0	5.12472
2	Br	4	7	3	2.96	5	1	0	8.74948
3	I	5	7	3	2.66	5	1	0	-2.79732
4	F	2	7	3	3.98	5	1	0	-2.24084
5	伯 C	2	4	0	2.55	2	2	3	-1.19615
6	仲 C	2	4	0	2.55	2	2	2	1.17157
7	叔 C	2	4	0	2.55	2	2	1	3.0
8	季 C	2	4	0	2.55	2	2	0	4.0
9	NH <sub>2</sub> 中 N	2	5	1	3.04	3	3	2	-7.45406
10	OH 中 O	2	6	2	3.44	4	2	1	-7.57119

## 6.2 $^1Q'$ 是表征取代基生物活性的特征值

按式(8)求得本文涉及的各种取代基的 $^1Q'$ (包括取代基所连芳环上的一个碳原子)为:

C—NH<sub>2</sub>: -29.81624, C—OH: -30.28476, C—CHO: -10.71357, C—CONH<sub>2</sub>: -40.101, C—CN: 2.50252, C—OCH<sub>3</sub>: -18.42463, C—CH<sub>3</sub>: -4.7846, C—C(CH<sub>3</sub>)<sub>3</sub>: 1.6462, C—F: -8.96336, C—I: -11.18428, C—COCH<sub>3</sub>: -15.06936, C—Cl: 20.04988, C—Br: 34.99792, C—NO<sub>2</sub>: 42.28927.

根据 $^1Q'$ 可将上述 14 种取代基分为两类:

(1)  $^1Q' > 0$  的取代基, 芳环上连有此类取代基的芳烃衍生物, 它的生物毒性通常比其母体芳烃(或取代芳烃)还大. 如表 3 中的苯酚, 对氯苯酚, 对溴苯酚, 对特丁基苯酚的 24h-LC<sub>50</sub> 依次为 2.804, 3.308, 3.619, 4.170, 4.052, 与其 $^1Q'$ 的相关系数为 0.9828(不含对特丁基苯酚), 显著正相关. 可称此类为生毒基团.

(2)  $^1Q' < 0$  的取代基: 连有此类取代基的取代芳烃的生物毒性通常比其母体还小. 如表 3 中的苯酚, 对甲苯酚, 对氟苯酚, 对甲氧基酚, 对羟基苯甲醛, 对羟基苯乙酮的 24h-LC<sub>50</sub> 依次为: 2.804, 2.928, 2.670, 2.371, 2.804, 2.509, 与其 $^1Q'$ 的相关系数为 0.845, 亦呈显著相关. 可称此类为解毒基团. 又如文献[6]认为氰基可被羊角月牙藻代谢, 促进其生长而具降毒作用. 这与氰基的 $^1Q' < 0$ (不含芳环上的碳原子)是一致的. 取代基的 $^1Q'$ 与其生物毒性密切相关, 因此,  $^1Q'$ 谓之取代基的生物活性指数. 显然, Kier 等的 $^1\chi'$ 是不能如此反映基团的生物功能, 说明 $^1Q$ 在揭示原子间的连接性方面优于 $^1\chi'$ .

廖宜勇等<sup>[1]</sup>利用 Zhao 报道的取代苯类对发光菌的毒性数据, 用 Free-Wilson 方法<sup>[15]</sup>计算出各种取代基对发光菌的毒性贡献值为  $F_j$ :

Br(0.73) > NO<sub>2</sub>(0.61) > COOH(0.59) > Cl(0.48) >

CN(0.39) > CH<sub>3</sub>(0.19)  $\approx$  OH(0.19)  $\approx$  CHO(0.19)  $\approx$  NH<sub>2</sub>(-0.19) > F(-0.34).

这与 $^1Q'$ 的顺序基本一致, 证明 $^1Q'$ 的建构具有合理性, 可用其表征取代基的生物活性. 两者比较尚存不一致现象, 究其原因有二: 其一, 同种取代芳烃在不同生物体内的中毒机理及敏感程度不同,  $F_j$ 只对发光菌, 而 $^1Q'$ 涉及 5 种生物, 两者必然不会完全吻合. 其二,  $^1Q'$ 的建立具有较大的非经验性.

## 6.3 $^1Q$ 蕴含影响化合物生物活性的本质因素

一般认为, 有机物对生物的毒性作用包含二个过程. 其一, 有机物通过扩散作用到达生物体内的靶分子, 并在水相与生物相之间达成分配平衡; 其二, 进入生物相的有机物分子与靶分子发生可逆性相互作用而使生物呈现毒性效应<sup>[16]</sup>. 如果不考虑有机物在生物体内的代谢, 其生物活性(以  $\lg C$  表示)可表示为该物质的疏水效应, 电性效应和立体效应的函数:

$$-\lg C = a \lg K_{OW} + b (\lg K_{OW})^2 + dH + eE_s + f \quad (27)$$

以辛醇/水分配系数( $K_{OW}$ )表征第一过程. 在  $K_{OW}$  的一定范围内, 有机物的  $K_{OW}$  越大, 其亲脂性越强, 在生物相中溶解越多, 相应毒性越大; 然而, 对于  $K_{OW}$  很大的有机物, 进入生物相的阻力增大, 会随  $K_{OW}$  增大而溶解量减小, 相应毒性减弱. 故第一过程描述为  $K_{OW}$  的二次曲线关系. 第二过程以电性效应( $H$ ), 立体效应( $E_s$ )表征, 它反映有机物对生物作用的真实毒性, 与有机物的分子结构密切相关. 主要表现为有机物与受体间的非色散力作用如静电作用, 氢键作用等.

$A_i$  包含影响成键原子  $i$  尺寸大小的  $n_i, m_i, k_i, e_i, h_i$  等主要因素, 即  $A_i$  与成键原子的体积相关. 前已述及,  $A_i$  与原子  $i$  的生物活性密切相关, 由此可称  $A_i$  为原子  $i$  的生物体积(这与空间体积不同). 则由  $A_i$  建构的 $^1Q$ 应蕴含分子的生物体积大小.  $^1Q$  越大, 其生物体积越大, 相应  $K_{OW}$  越大. 例如表 1~表 4 中的 $^1Q$ 与  $\lg K_{OW}$  的线性相关系数依次为: 0.5094, 0.7804, 0.6276, 0.9443. 说明 $^1Q$ 具有反映有机物的疏水性功能.

根据线性自由能相关理论, 有机物分子与靶细胞内分子的相互作用可以用最高占有轨道能( $E_{HOMO}$ )来描述. 据量子化学理论可知,  $E_{HOMO}$  与最低未占轨道能( $E_{LUMO}$ )正相关. 将文献[7,8]中的  $E_{LUMO}$  分别与相应的 $^1Q$ 回归, 得线性相关系数为 0.7764, 0.7437. 说明在 $^1Q$ 中蕴含  $E_{LUMO}$ , 这是因为在  $A_i$  中包

含影响原子的  $E_{\text{LUMO}}$  的各种因素, 如  $n_i, m_i, x_{pi}, e_i$  等.

文献[1]选用 Kier 的 1 阶价连接性指数 ( $^1\chi^v$ ) 表征化合物在立体结构上的差异. 根据  $^1Q$  的计算公式, 可以认为  $^1Q$  也是以分子图的邻接矩阵为基础的类似于  $^1\chi^v$  的连接性指数 (只是方次上的差异:  $^1\chi^v$  为  $-0.5$ ,  $^1Q$  为 1. 但对于  $m=0$  及  $m \geq 2$ , 两者差别很大), 因此,  $^1Q$  也反映了有机物的立体效应. 综合以上分析, 式(27)可写为:

$$-\lg C = B^1Q + D \quad (28)$$

按此公式, 笔者建立了方程(10), (13), (17), (21), (24), 其  $r$  依次为: 0.8642, 0.9284, 0.8897, 0.9225, 0.9205. 相应  $\lg K_{\text{OW}}$  与之的  $r$  依次为: 0.4375, 0.6790, 0.7987, 0.8612, 0.9408. 对于方程(13), (17),  $E_{\text{LUMO}}$  与之的  $r$  为 0.8164, 0.5878; 对于方程(21), (24),  $E_{\text{HOMO}}$  与之的  $r$  为 0.207, 0.226. 它们的相关性 (几乎) 均低于  $^1Q$ , 并且方程(21)的  $r$  大于文献[1]使用  $\lg K_{\text{OW}}$ ,  $E_{\text{HOMO}}$  的二元相关系数 0.88. 以上事实说明,  $^1Q$  不仅蕴含  $\lg K_{\text{OW}}$ ,  $E_{\text{LUMO}}$  (或  $E_{\text{HOMO}}$ ), 而且证明含有立体效应. 即  $^1Q$  揭示了影响有机物生物活性的本质因素.

由表 1~表 4 可见, 在含有多种基团的分子与靶标作用时, 存在一个整体分子和各基团间的协同作用. 可以认为, 这种“协同作用”已隐含在  $^1Q = \sum A_i, A_j$  的“ $\sum$ ”之中. 如表 4 中苯, 苯胺, 3,4-二氯苯胺, 3-氯-4-氟苯胺的  $^1Q$  依次为: 54.000, 30.184, 84.182, 54.719, 这与其对发光菌的  $-\lg EC_{50}$  为 3.34, 3.28, 4.20, 3.28 基本一致. 其中不完全一致的原因, 除与  $A_i$  有关外, 在  $^1Q$  中把它们看作是简单的线性相加也是重要原因.

$^1Q$  与  $^1\chi^v$  相比还具有其  $A_i$  值可正可负优点, 而  $^1\chi^v$  的  $\delta_i^v$  只能大于零. 另外,  $^1Q$  对有机物分子的结构选择性接近于  $^1\chi^v$  (因计算形式相似).

#### 6.4 距离指数 ( $L$ ) 的功能

距离指数 ( $L$ ) 与苯环上取代基数目及取代基中非氢原子数正相关, 因此  $L$  与有机物的分子体积正相关. 相应  $L$  与其  $\lg K_{\text{OW}}$  亦为正相关. 表 1 至表 4 中有机物的  $\lg K_{\text{OW}}$  与  $L$  的相关系数依次为: 0.2622, 0.3409, 0.4584, 0.6405, 均大于零. 因此,  $L$  与有机物的生物毒性亦为同向变化. 式(11), (14), (18), (22), (25) 的  $r$  依次为 0.7598, 0.8160, 0.6784, 0.7236, 0.8493. 其相关程度与  $\lg K_{\text{OW}}$  相近, 证实上述讨论的正确性.

另外, 由于  $L$  不考虑非氢原子的种类, 从另一角度反映有机物分子的空间结构差异即立体效应. 因

此, 在  $^1Q$  方程中引入  $L$  都使相关程度有较大幅度改善.  $^1Q$  与  $L$  同时并用, 也基本实现对有机物分子结构的唯一性表征.

根据 72 种取代芳烃的  $^1Q$ ,  $L$  与 5 种生物的毒性关系, 可得出以下结论:

(1) 成键原子的生物活性点价 ( $A_i$ ), 其正、负值, 与非氢原子的生物活性基本一致.  $A_i > 0$  的为增毒基团,  $A_i < 0$  的为降毒基团.

(2) 自 Wiener 以来, 迄今建立的 100 余种拓扑指数所使用的点价中尚无负值, 包括自相关拓扑指数. 本文  $A_i$  的取值可以是一切实数.

(3) 物质的理化性质与生物活性通常表现为加和性和构成性. 由  $A_i$  建构的  $^1Q$ , 其公式中的“ $\sum$ ”体现加和性, 由“ $\sum$ ”及“ $A_i \times A_j$ ”反映构成性.

(4)  $^1Q$  蕴涵有机物的疏水性, 电性效应与立体效应;  $L$  蕴涵有机物的疏水效应, 立体效应.  $^1Q$  与  $L$  同时使用, 不仅实现有机物分子结构差异的定量化表征, 而且使相关程度有较大改善.

(5)  $^1Q^*$  为取代基的生物活性指数,  $^1Q^* > 0$  的为生毒基团,  $^1Q^* < 0$  的为解毒基团.

(6) 由  $^1Q, L$  建立的有机物生物活性的数学模型, 涉及苯环上连有: F, Cl, Br, I, OH, NH<sub>2</sub>, CHO, COCH<sub>3</sub>, CONH<sub>2</sub>, CN, OCH<sub>3</sub>, 烷基等众多官能团的取代芳烃; 涉及的生物有: 绿藻 (羊角月牙藻), 斜生栅列藻, 日本长腿蛙蝌蚪, 发光菌, 呆鲱鱼等. 因此, 本文的 QSAR 模型具有化合物类型及生物种类的广泛性, 生物活性的多样性特点. 并且估算的准确程度都优于相应文献的研究成果, 即还具有显著相关性.

#### References

- Wang, L. - S.; Han, S. - Q. *Progress in Organic Pollution Chemistry*, Chemical Industry Press, Beijing, 1998, pp. 1 - 199, 274 - 348 (in Chinese).
- Feng, C. - J.; Cai, K. - Y. *J. Natural Sci. Nanjing Normal University*, 2000, 2(1), 39.
- Bradbury, S. P. *Qsar Environ. Res.* 1994, 2, 89.
- Wang, L. - S. *Progress in Environmental Chemistry*, Chemical Industry Press, Beijing, 1995, pp. 101 - 153, 285 - 344 (in Chinese).
- Wang, L. - S.; Han, S. - Q. *Molecular Structure, Property and Activity*, Chemical Industry Press, Beijing, 1997, pp. 1 - 139, 177 - 271 (in Chinese).
- Zhang, A. - Q.; Wei, D. - B.; Wang, L. - S. *Environ. Chem.* 2000, 19(3), 220 (in Chinese).
- Lu, G. - H.; Yuan, X.; Zhao, Y. - H. *Environ. Chem.* 2000, 19(3), 225 (in Chinese).
- Wang, X. - D.; Dong, J. - S.; Wang, L. - S. *Environ. Chem.*,



- 2000, 19(3), 230 (in Chinese).
- 9 Wiener, H. *J. Am. Chem. Soc.*, **1947**, 69(1), 17.
- 10 Randic, M. *J. Am. Chem. Soc.*, **1975**, 97(25), 6609.
- 11 Kier, L. B.; Hall, H. *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, **1976**, p. 82.
- 12 Wang, P.; Su, J. - C.; Shan, J. - J. *Chemistry*, **1998**, 10, 40 (in Chinese).
- 13 Reusch, W. H. *An Introduction to Organic Chemistry*, H. D. Inc. London, **1977**, pp. 193 - 196.
- 14 Fang, C. - J.; Cai, K. - Y.; Li, M. - J. *Environ. Chem.* **2000**, 19(5), 408 (in Chinese).
- 15 Free, S. M.; Wilson, J. W. *J. Med. Chem.*, **1964**, 7, 395.
- 16 McFarland, J. W. *J. Med. Chem.*, **1970**, 13, 1092.

(Ed XIE Hui - Qin)

(ZHENG Guo - Cheng)