

自学之友

# 分子连接性指数 ${}^mX^z$ 与烷基苯标准生成焓的定量关系

张玉林 郭满才

(西北农林科技大学生命科学院 杨陵 712100)

**摘要** 提出了计算有机物分子中成键原子点价  $\delta_i^z$  的新方法, 以  $\delta_i^z$  为核心建构了一个新的分子连接性指数 ${}^mX^z$  ( $m = 0, 1$ ), 并探讨了 ${}^mX^z$  指数与烷基苯标准生成焓 $-\Delta_f H_m^\ominus$  的相关性, 结果表明,  ${}^mX^z$  指数对于烷基苯具有良好的结构选择性和性质相关性。以 ${}^0X^z$ ,  ${}^1X^z$  和碳原子数 $N$  为自变量的三元线性回归方程为:  $-\Delta_f H_m^\ominus = 58.008 {}^0X^z + 27.075 {}^1X^z + 929.836N - 1504.311$ , 该方程对烷基苯标准生成焓的预测结果令人满意。

在有机物的定量结构-性质/活性关系(QSPR/QSAR) 研究中, 分子拓扑指数的应用很广泛。所谓分子拓扑指数, 是采用图论方法从分子图中获得的表征分子结构的一种数学量。目前已知的分子拓扑指数约有 100 余种, 其中 Randic-Kier 连接性指数<sup>[1,2]</sup> 的应用最为广泛, 也最为成功。近年来, 国内很多学者对连接性指数的核心概念——点价( $\delta$ ) 进行改进, 由此建构新的拓扑指数<sup>[2~4]</sup>, 不断取得很有意义的成果。本文介绍了一个计算有机物分子中成键原子点价  $\delta_i^z$  的新方法, 以  $\delta_i^z$  为基础建构了一个新的分子连接性指数 ${}^mX^z$ , 并研究了 ${}^mX^z$  ( $m = 0, 1$ ) 指数与烷基苯标准生成焓 $-\Delta_f H_m^\ominus$  的相关性, 结果表明,  ${}^mX^z$  指数对烷基苯具有良好的结构选择性和性质相关性。

## 1 分子连接性指数 ${}^mX^z$

有机物分子中成键原子的特征, 主要由原子的结构和性质以及原子所处的化学环境等因素决定, 由此我们提出了一个计算成键原子点价  $\delta_i^z$  的新方法, 其公式为:

$$\delta_i^z = \frac{Z_i(Z_i - h_i)^{V_i/E_i}}{n_i^2} \quad (1)$$

式(1)中,  $n_i$ 、 $Z_i$  分别为原子  $i$  的核外电子层数和价电子数,  $h_i$  为与原子  $i$  直接键合的氢原子数,  $V_i$ 、 $E_i$  分别为分子图中原子  $i$  的顶点度和边度。

在有机物的分子图中, 顶点代表原子, 边代表化学键。对分子图的各顶点进行随机编号。与顶点  $i$  直接相连的顶点的个数叫做  $i$  的顶点度, 以  $V_i$  表示; 与顶点  $i$  直接相连的边的数目叫做  $i$  的边度, 以  $E_i$  表示。例如, 叔丁基苯的分子图如图 1 所示, 其中数字 1 到 10 为顶点(C 原子) 的随机编号, 根据以上定义, 顶点 1 到 10 的

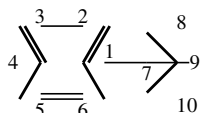


图 1 叔丁基苯的分子图

顶点度分别为 3, 2, 2, 2, 2, 2, 4, 1, 1, 1, 边度分别为 4, 3, 3, 3, 3, 3, 4, 1, 1, 1。按照式(1), 叔丁基苯分子中碳原子 1 到 10 的点价  $\delta_i^z$  依次为  $\sqrt{8}, \sqrt[3]{9}, \sqrt[3]{9}, \sqrt[3]{9}, \sqrt[3]{9}, \sqrt[3]{9}, 4, 1, 1, 1$ 。

在分子邻接矩阵的基础上, 由  $\delta_i^z$  建构分子连接性指数  ${}^mX^z$ 。按照连接性指数的计算形式,  ${}^mX^z$  指数的计算通式为:

$$X^z = \sum (\delta_i^z \cdot \delta_j^z \cdot \delta_k^z \dots)^{-0.5} \quad (2)$$

其 0 阶指数( ${}^0X^z$ ) 和 1 阶指数( ${}^1X^z$ ) 的计算公式为:

$${}^0X^z = \sum (\delta_i^z)^{-0.5} \quad (3)$$

$${}^1X^z = \sum (\delta_i^z \cdot \delta_j^z)^{-0.5} \quad (4)$$

式(3) 中  $\sum$  表示对分子中所有的原子求和; 式(4) 中  $\sum$  表示对分子中全部的化学键求和。

用式(3), 式(4) 计算得到叔丁基苯的  ${}^0X^z, {}^1X^z$  分别为:

$${}^0X^z = (\sqrt{8})^{-0.5} + (\sqrt[3]{9})^{-0.5} \times 5 + 4^{-0.5} + 1^{-0.5} \times 3 = 7.5614$$

$${}^1X^z = (\sqrt{8} \times \sqrt[3]{9})^{-0.5} \times 2 + (\sqrt[3]{9} \times \sqrt[3]{9})^{-0.5} \times 4 + (\sqrt{8} \times 4)^{-0.5} + (4 \times 1)^{-0.5} \times 3 = 4.5449$$

用式(1), 式(3), 式(4) 对文献[5] 中 48 个化合物的  ${}^0X^z, {}^1X^z$  进行计算, 结果列于表 1。表 1 中还列出了文献[5] 中的  $F'$  值, 以便进行比较。由表 1 可见,  ${}^0X^z, {}^1X^z$  蕴含了烷基苯分子的大小以及烷基苯分子中烷基的数目, 烷基的结构和烷基的位置等分子结构信息, 即  ${}^mX^z$  指数反映了烷基苯分子的结构特征。

## 2 ${}^mX^z$ ( $m=0,1$ ) 指数与烷基苯标准生成焓 $-\Delta_f H_m^\ominus$ 的相关性

烷基苯的  $-\Delta_f H_m^\ominus$  主要决定于它们的自身结构, 而  ${}^mX^z$  指数正好反映了烷基苯分子的结构特征, 因此烷基苯的  $-\Delta_f H_m^\ominus$  与  ${}^mX^z$  指数必然存在着相关性。根据散点图, 我们提出如下数学模型:

$$-\Delta_f H_m^\ominus = A {}^mX^z + B \quad (5)$$

式(5) 中  $A, B$  为待定常数。用计算机对文献[5] 中 48 个化合物的  $-\Delta_f H_m^\ominus$  与其  ${}^0X^z$  和  ${}^1X^z$  分别进行拟合, 得到一元线性回归方程为:

$$-\Delta_f H_m^\ominus = 1425.643 {}^0X^z - 2191.547 \quad (6)$$

$$n=48, R=0.998(0.999964), F=10622.96, s=240.10$$

$$-\Delta_f H_m^\ominus = 1937.898 {}^1X^z - 961.575 \quad (7)$$

$$n=48, R=0.999(0.999964), F=36447.85, s=129.82$$

式(6) 和式(7) 中,  $n, R, F, s$  依次为回归方程的样本数、相关系数、 $F$  检验值和标准偏差(括号内的数据为文献[5] 中用  $F'$  指数进行相应的回归所得到的结果, 下同)。用式(7) 计算了上述 48 个化合物的  $-\Delta_f H_m^\ominus$ , 计算值列在表 1 中。为了便于比较, 表 1 中还列出了文献[5] 的计算结果。由表 1 可见, 式(7) 的计算结果比文献[5] 稍差, 但却相当接近。

如果仅对直链烷基苯(即编号为 2, 3, 7, 15, 37~48 的化合物) 进行回归, 得到一元线性回归方程为:

$$-\Delta_f H_m^\ominus = 1392.082 {}^0X^z - 1655.755 \quad (8)$$

$$n=16, R=1.000, F=923594419.90, s=0.60$$

$$-\Delta_f H_m^\ominus = 1965.834 {}^1X^z - 1229.951 \quad (9)$$

$n=16, R=1.000(1.0000), F=1476179.71, s=14.94$   
 式(8)对16个直链烷基苯的计算值与实验值的平均误差为0.43, 平均相对误差为0.0049, 比文献[5]相应的计算结果稍好。式(9)的平均误差为8.38, 平均相对误差为0.11。

为了充分表达分子结构信息, 提高相关性, 将 $^0X^z, ^1X^z$ 作为两个独立变量, 与上述48个化合物的 $-\Delta_fH_m^\ominus$ 关联, 得到二元线性回归方程为:

$$-\Delta_fH_m^\ominus=470.854^0X^z+1301.694^1X^z-1388.470 \tag{10}$$

$n=48, R=1.000, F=77151.09, s=63.12$   
 在式(10)中再引入另一个自变量碳原子数 $N$ , 得到三元线性回归方程为:

$$-\Delta_fH_m^\ominus=58.008^0X^z+27.075^1X^z+929.836N-1504.311 \tag{11}$$

$n=48, R=1.000, F=46893302.93, s=2.09$   
 用式(11)计算了上述48个化合物的 $-\Delta_fH_m^\ominus$ , 计算结果列在表1中。由表1可见, 式(11)的计算值与实验值符合得很好。

表1 苯和烷基苯的 $^mX^z(m=0,1)、F'、N$ 与 $-\Delta_fH_m^\ominus$

序号	化合物	$^0X^z$	$^1X^z$	$F'$	$N$	$-\Delta_fH_{m(\text{exp})}^\ominus/(\text{kJ}\cdot\text{mol}^{-1})^{[5]}$	$-\Delta_fH_m^\ominus(\text{cal})/(\text{kJ}\cdot\text{mol}^{-1})$		
							式(7)	式(11)	文献[5]
1	苯	4.1602	2.8845	5.2008	6	4392.72	4628.29	4394.13	4321.37
2	甲苯	5.0614	3.3422	6.0517	7	5389.27	5515.27	5388.64	5350.22
3	乙苯	5.7685	3.8751	6.8682	8	6373.12	6547.97	6373.92	6337.49
4	邻二甲苯	5.9627	3.8096	6.9026	8	6383.91	6421.04	6383.41	6379.08
5	间二甲苯	5.9627	3.7998	6.9026	8	6385.66	6402.05	6383.15	6379.08
6	对二甲苯	5.9627	3.7998	6.9026	8	6384.95	6402.05	6383.15	6379.08
7	正丙苯	6.4756	4.3751	7.6847	9	7358.72	7516.92	7358.31	7324.75
8	异丙苯	6.6388	4.2455	7.7020	9	7362.61	7265.77	7364.27	7345.66
9	邻甲乙苯	6.6698	4.3425	7.7191	9	7365.33	7453.75	7368.69	7366.34
10	间甲乙苯	6.6698	4.3328	7.7191	9	7368.46	7434.95	7368.43	7366.34
11	对甲乙苯	6.6698	4.3328	7.7191	9	7369.80	7434.95	7368.43	7366.34
12	1,2,3-三甲苯	6.8639	4.2770	7.7535	9	7376.11	7326.81	7378.18	7407.94
13	1,2,4-三甲苯	6.8639	4.2672	7.7535	9	7380.46	7307.82	7377.92	7407.94
14	1,3,5-三甲苯	6.8639	4.2575	7.7535	9	7382.59	7289.03	7377.65	7407.94
15	正丁基苯	7.1827	4.8751	8.5012	10	8343.99	8485.87	8342.70	8312.01
16	异丁基苯	7.3459	4.7309	8.5185	10	8351.72	8206.43	8348.27	8332.93
17	仲丁基苯	7.3459	4.7835	8.5185	10	8347.63	8308.36	8349.69	8332.93
18	叔丁基苯	7.5614	4.5449	8.4579	10	8352.85	7845.98	8355.73	8259.65
19	邻二乙苯	7.3769	4.8755	8.5356	10	8349.13	8486.65	8353.98	8353.60
20	间二乙苯	7.3769	4.8657	8.5356	10	8352.02	8467.66	8353.71	8353.60
21	对二乙苯	7.3769	4.8657	8.5356	10	8352.43	8467.66	8353.71	8353.60
22	邻正丙基甲苯	7.3769	4.8425	8.5356	10	8351.81	8422.70	8353.09	8353.60
23	间正丙基甲苯	7.3769	4.8328	8.5356	10	8354.69	8403.90	8352.82	8353.60
24	对正丙基甲苯	7.3769	4.8328	8.5356	10	8355.15	8403.90	8352.82	8353.60
25	邻异丙基甲苯	7.5400	4.7130	8.5529	10	8355.69	8171.74	8359.04	8374.52
26	间异丙基甲苯	7.5400	4.7032	8.5529	10	8358.58	8152.75	8358.78	8374.52
27	对异丙基甲苯	7.5400	4.7032	8.5529	10	8359.04	8152.75	8358.78	8374.52
28	1,2-二甲基-3-乙基苯	7.5710	4.8099	8.5700	10	8359.92	8359.52	8363.46	8395.20

序号	化合物	${}^0X^z$	${}^1X^z$	$F'$	$N$	$-\Delta_f H_m^\ominus(\text{exp})/(\text{kJ}\cdot\text{mol}^{-1})[5]$	$-\Delta_f H_m^\ominus(\text{cal})/(\text{kJ}\cdot\text{mol}^{-1})$		
							式(7)	式(11)	文献[5]
29	1, 2 二甲基 4 乙基苯	7.5710	4.8002	8.5700	10	8362.01	8340.72	8363.20	8395.20
30	1, 3 二甲基 2 乙基苯	7.5710	4.8099	8.5700	10	8359.92	8359.52	8363.46	8395.20
31	1, 3 二甲基 4 乙基苯	7.5710	4.8002	8.5700	10	8363.26	8340.72	8363.20	8395.20
32	1, 3 二甲基 5 乙基苯	7.5710	4.7904	8.5700	10	8365.73	8321.73	8362.93	8395.20
33	1, 4 二甲基 2 乙基苯	7.5710	4.8002	8.5700	10	8363.26	8340.72	8363.20	8395.20
34	1, 2, 3, 4 四甲苯	7.7651	4.7444	8.6044	10	8372.08	8232.59	8372.95	8436.79
35	1, 2, 3, 5 四甲苯	7.7651	4.7346	8.6044	10	8374.96	8213.60	8372.68	8436.79
36	1, 2, 4, 5 四甲苯	7.7651	4.7346	8.6044	10	8375.42	8213.60	8372.68	8436.79
37	正戊基苯	7.8898	5.3751	9.3177	11	9328.26	9454.82	9327.10	9299.27
38	正己基苯	8.5969	5.8751	10.1342	12	10312.48	10423.77	10311.49	10286.53
39	正庚基苯	9.3041	6.3751	10.9507	13	11296.74	11392.72	11295.88	11273.79
40	正辛基苯	10.0112	6.8751	11.7672	14	12280.97	12361.67	12280.28	12261.05
41	正壬基苯	10.7183	7.3751	12.5837	15	13265.23	13330.62	13264.67	13248.31
42	正癸基苯	11.4254	7.8751	13.4002	16	14249.49	14299.57	14249.06	14235.57
43	正十一烷基苯	12.1325	8.3751	14.2176	17	15233.72	15268.51	15233.45	15222.83
44	正十二烷基苯	12.8396	8.8751	15.0332	18	16217.98	16237.46	16217.84	16210.10
45	正十三烷基苯	13.5467	9.3751	15.8497	19	17202.20	17206.41	17202.23	17197.36
46	正十四烷基苯	14.2538	9.8751	16.6662	20	18186.47	18175.36	18186.63	18184.62
47	正十五烷基苯	14.9609	10.3751	17.4827	21	19170.73	19144.31	19171.02	19171.88
48	正十六烷基苯	15.6680	10.8751	18.2992	22	20154.96	20113.26	20155.41	20159.14
平均误差							93.87	1.54	22.30
平均相对误差							1.19%	0.019%	0.28%

注： $-\Delta_f H_m^\ominus(\text{exp})$  是标准生成焓的实验值； $-\Delta_f H_m^\ominus(\text{cal})$  是标准生成焓的计算值；误差  $\Delta = (-\Delta_f H_m^\ominus(\text{cal})) - (-\Delta_f H_m^\ominus(\text{exp}))$ ；平均误差=  $\sum \Delta / 48$ ，平均相对误差=  $\sum \left[ \frac{|\Delta|}{-\Delta_f H_m^\ominus(\text{exp})} \times 100 \right] / 48$ 。

### 3 结果讨论

(1)  $\mathfrak{F}_i$  是表征成键原子特征的结构参数。

式(1)中,  $n_i$ 、 $Z_i$  可以表征成键原子*i* 的结构和性质;  $V_i$ 、 $E_i$ 、 $Z_i - h_i$  可以表征成键原子*i* 的成键状况; 对于  $\text{sp}^3$  杂化的原子,  $V_i = E_i$ ,  $V_i/E_i = 1$ , 对于  $\text{sp}^2$  和  $\text{sp}$  杂化的原子,  $V_i < E_i$ ,  $V_i/E_i < 1$ , 因此  $V_i/E_i$  可以表征原子*i* 的杂化状态; 所以, 式(1)中  $(Z_i - h_i)^{V_i/E_i}$  可以用来表征原子*i* 在分子中所处的化学环境。由此可见, 式(1)具有明确的物理意义, 由式(1)计算的  $\mathfrak{F}_i$  是一种表征成键原子特征的结构参数。

同一种原子, 由于氧化状态不同或者杂化状态不同, 其成键状况必然不同, 由式(1)计算的  $\mathfrak{F}_i$  值也必然不同, 因此,  $\mathfrak{F}_i$  能够有效地区分同一原子的不同氧化状态或者同一原子的不同杂化状态。叔丁基苯分子中  $\text{C}_1$ 、 $\text{C}_7$  的  $\mathfrak{F}_i$  值分别为 $\sqrt{8}$ 和 4, 就说明了这一点。

(2)  $X^{pz}$  指数是一种比较优良的分子拓扑指数。

“一种新的分子拓扑指数的提出应该同时具备高的选择性(即唯一性)和良好的相关性”<sup>[6]</sup>。文献[5]报道的  $F'$  指数, 虽然与烷基苯 $-\Delta_f H_m^\ominus$  有优良的性质相关性, 但其结构选择性较差, 对同分异构体的分辨率较低。对于碳原子数  $N = 10$  的烷基苯同分异构体,  ${}^0X^z$ ,  ${}^1X^z$

和  $F'$  指数的 Razinger 选择性系数<sup>[6]</sup> 依次为 0.32, 0.68, 0.32, 这就说明  ${}^0X^z$  的选择性与  $F'$  指数相同, 而  ${}^1X^z$  的选择性显著地优于  $F'$  指数。式(6)、(7)、(8)、(9) 的回归结果表明,  ${}^0X^z$ 、 ${}^1X^z$  与 烷基苯- $\Delta_f H_m^\ominus$  的相关性与  $F'$  指数大致相当, 它们对烷基苯- $\Delta_f H_m^\ominus$  都呈现优良的性质相关性。总之,  ${}^mX^z$  指数对于烷基苯是一种结构选择性和性质相关性均较佳的分子拓扑指数。

(3) 本文方法揭示了影响烷基苯- $\Delta_f H_m^\ominus$  的本质因素。

式(11) 的回归结果表明,  ${}^0X^z$ 、 ${}^1X^z$  和  $N$  共同揭示了影响烷基苯- $\Delta_f H_m^\ominus$  的本质因素, 其中碳原子数  $N$  是影响烷基苯- $\Delta_f H_m^\ominus$  变化规律的主要因素。一般来说, 化合物的- $\Delta_f H_m^\ominus$  具有加和性,  ${}^mX^z$  指数也具有加和性, - $\Delta_f H_m^\ominus$  与  ${}^mX^z$  指数之间必然存在着内在联系。 ${}^0X^z$ 、 ${}^1X^z$  都随着碳原子数  $N$  的增加而增大, 因此, 在  ${}^0X^z$ 、 ${}^1X^z$  和  $N$  3 个结构参数中, 碳原子数  $N$  必然成为影响烷基苯- $\Delta_f H_m^\ominus$  变化规律的主要因素。

无论是  $\delta$  值的计算, 还是  ${}^0X^z$ 、 ${}^1X^z$  值的计算, 都不涉及高深复杂的数学知识和数学运算, 也不需要查找任何化学数据, 只需画出分子图便可实现。因此, 本文方法计算简单, 应用方便。应用式(11) 可以方便、准确地预测烷基苯的标准生成焓, 这具有一定的实际意义。

## 参 考 文 献

- 1 Randic M. *J Am Chem Soc*, 1975, 97(23): 6609
- 2 余训民, 余训爽. 有机化学, 2001, 21(9): 658
- 3 冯长君, 陈艳, 李鸣建. 有机化学, 2002, 22(3): 206
- 4 秦正龙. 有机化学, 2002, 22(6): 436
- 5 杨锋, 罗明道, 颜肖慈, 等. 有机化学, 2000, 20(3): 410
- 6 许禄. 化学计量学方法. 北京: 科学出版社, 1995

## 第 2 届全国实验室管理科学研讨会通知

由中国分析测试协会主办、《分析试验室》编辑部承办的第 2 届全国实验室管理科学研讨会定于 2005 年 4 月中旬在安徽黄山市召开。

会议内容包括实验室管理科学、实验室认证认可、实验室比对、实验室信息管理系统(LIMS)、实验室仪器设备运行及考核、实验技术人员培训等有关方面的学术研讨。会议将邀请国内有关专家做大会报告, 欢迎各行业的实验室管理人员及实验室科研工作人员参加会议。

请在 2005 年 2 月 28 日前将录有会议征文文章题目和摘要的软盘邮寄(或用电子邮件发送)到会议筹备组(3 日内收到筹备组的回复邮件方为发送成功), 并注明联系人、详细通信地址、联系电话、传真号码及 E-mail 地址。经审查录用的会议征文可在《分析试验室》期刊上正式发表。

欲参加会议的各界人士请于 2005 年 2 月 28 日前与会议筹备组联系, 以便寄发下一轮通知。

筹备组地址: 北京新外大街 2 号 603 室 实验室管理会议筹备组 邮政编码: 100088

电话/传真: 010 82013328 E-mail: ana-info@263.net 联系人: 孙臣良