文章编号:1673 - 0062(2012)01 - 0082 - 05

含硫苯衍生物对发光细菌生物毒性的 QSAR 研究

段强林,聂长明*,石春雨

(南华大学 化学化工学院,湖南 衡阳 421001)

摘 要: 应用分子力学,半经验量子化学 RM1 方法,对 29 个含硫苯衍生物的几何构型进行全优化,将一系列的量子化学参数用于其结构和对发光细菌生物毒性的 QSAR 研究中. 利用 SPSS13.0 软件,用多元逐步回归法建立了含硫苯衍生物对发光细菌毒性与量化参数的多元回归方程,结果表明,前线轨道能量差 ΔE,辛醇/水分配系数($\lg K_{ow}$),取代基的 Hammett 电荷效应常数 $\sum \sigma$,碳原子 C(6) 上的静电荷 $E(6)$ 及苯环上碳原子的静电之和 $\sum E$ 对其生物毒性有明显的影响,其生物毒性可用下列模型来定量描述:$\lg(1/EC50) = 10.238 - 0.009\triangle E + 3.034E(6) + 2.108\sum E - 0.292\lg K_{ow} + 0.522\sum \sigma, N = 29, R = 0.930.$

关键词:含硫苯衍生物;生物毒性;定量结构活性;量子化学 RM1 方法

中图分类号:O623.54⁺3 **文献标识码**:B

QSAR Studies on Thiobenzene-containing Derivatives's Biological Toxicity to Photobateria

DUAN Qiang-lin, NIE Chang-ming*, SHI Chun-yu

(School of Chemistry and Chemical Engineering, University of South China, Hengyang, Hunan 421001, China)

Abstract: In this work, molecular mechanics and semi-empirical quantum chemistry RM1 methods were used to refine and optimize the geometry of 29 thiobenzene-containg derivartives and then the physicochemical properties parameters, molecular parameters extracted from the geometric model along with indicator parameters were applied to establish the QSAR model for biological toxicity of thiobenzene-containg derivartives to photobacteria. Based on the multiple linear regression analysis, using software SPSS 13.0, equation was established between biological toxicity of thiobenzene-containing de-

收稿日期:2011 - 12 - 20

基金项目:湖南省科技计划基金资助项目(06FJ54104);湖南省优秀博士学位论文基金资助项目(2008YBX01)

作者简介:段强林(1985 -),男,湖南衡阳人,南华大学化学化工学院硕士研究生. 主要研究方向:有机化合物性质与结构关系. *通讯作者.

rivatives to photobacteria and the quantum chemical descriptors. The results indicate that biological toxicity of thiobenzene-containing derivatives to photobacteria was affected obviously by frontier orbital energy gap $\Delta E(E_{LUMO} - E_{HOMO})$, Octanol/water Partition Coefficients($\lg K_{ow}$), Hammett charge effect constant($\sum\sigma$), Electrostatic of C(6) $E(6)$ as well as total electrostatic of benzene ring($\sum E$). Then, we establish a QSAR model($\lg(1/EC50) = 10.238 - 0.009\triangle E + 3.034E(6) + 2.108\sum E - 0.292\lg K_{ow} + 0.522\sum\sigma, R = 0.930$) with satisfactory results.

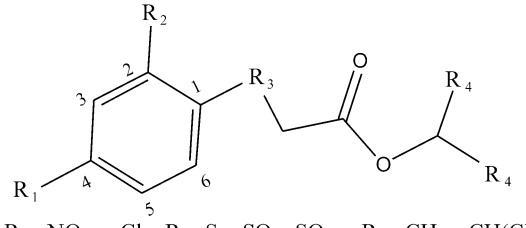
key words: thiobenzene-containing derivatives; biological toxicity; RM1 method for quantum chemistry; QSAR

含硫苯衍生物是一类重要的农药和化工原料,其分子的结构和组成直接影响化合物的活性(毒性)效应常称这两者间的量变规律为定量构效关系(QSAR). 该类化合物,除了少数分子结构比较简单分子量比较小的化合物外,常含有苯基等易产生毒性的基团,使得这类化合物常常具有毒性或有潜在毒性. 由于在生产中的广泛应用,这些化合物常常对环境产生污染,有关含硫苯衍生物毒性的 QSAR 研究受到学术界的广泛关注[2]. 如王连生等人用 29 种含硫苯衍生物对发光细菌的毒性效应及其定量结构活性相关分析发现,此类化合物的毒性主要与其取代基的电荷效应相关,电荷的区域分布对化合物的毒性效应有重要影响,并建立了相关模型,取得较好的预测能力($R = 0.837$)[3];谢前等人对这 29 种含硫苯衍生物对发光细菌的毒性,利用逐步线形回归方法挑选重要自变量,作为人工神经网络方法的输入,进行神经网络计算,建立了比回归方程更为精确的构效关系模型[4]. 以往的 QSAR 分析所研究的系列化合物结构变化通常较小,往往有相同的或类似的活性中心,化合物的辛醇/水分配能力对其毒性具有主导作用,在 QSAR 方程中化合物的毒性参数与分配系数、分子连接指数等参数具有良好的相关性. 本文通过应用分子力学,半经验量子化学 RM1 方法,对 29 个含硫苯衍生物的几何构型进行全优化,将量子化学计算结果用于其结构和生物毒性的研究中. 获得了满意的结果,为含硫苯衍生物对发光细菌的毒性研究供了一种新的思路.

1　理论与方法

1.1　分子结构及其优化

通过 Spartan06 v1.12① 软件构建出各含硫苯衍生物的分子立体结构,然后通过分子力学方法优化到能量最低的几何构象,再使用半经验的 RM1[5](一种更为精确的重新参数化 AM1 的半经验算法)量化方法进行较为精确的几何优化,从而得到各个化合物的优势构象. 化合物分子结构如图 1 所示,其结构优化图如图 2 所示.



R_1, R_2=NO_2, Cl; R_3=S, SO, SO_2; R_4=CH_3, $CH(CH_3)_2$

图 1　含硫苯衍生物的母体结构

Fig. 1　Chemical structure of thiobenzene-containing derivatives

1.2　量化参数计算

一般来说,有毒化合物产生毒性主要涉及两个关键步骤:有毒化合物分配到/穿过生物膜,以及与有毒化合物的作用位点的相互作用[6]. 因此,在QSAR研究中选择合适的分子结构参数来

①Spartan06 for windows. Wavefunction, Inc., Japan Branch Office 3-5-2, Kouji-machi, Suite 608 Chiyoda-ku, Tokyo, Japan 102-0083. 2006.

描述这些过程十分重要. 本文从量化计算的分子模型中,提取易于直接解释作用机理的热力学参数,量子化学参数,分子物化参数建立 QSAR 模型.

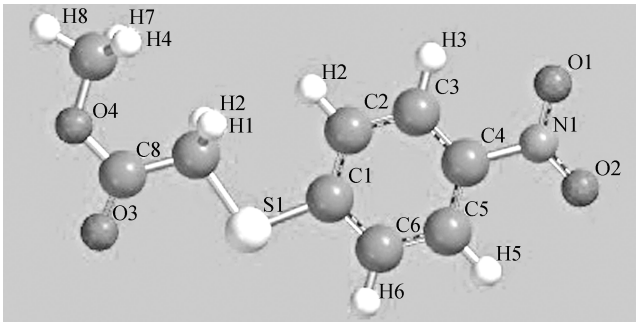


图 2 化合物 1 的结构优化图

Fig. 2 Optimized molecular structure for comopound 1

采用化合物对发光细菌的毒性取值为 lg(1/EC50) 作为QSAR模型的因变量,其中EC50是相对发光率为 50% 时化合物的浓度 mmol/L,其他物理化学参数有:辛醇/水分配系数($\lg K_{ow}$)、取代基的 Hammett 电荷效应常数($\sum \sigma$),以上参数有文献得到[7].

量子化学参数包括:分子最高占据轨道能(E_{HOMO}),分子最低未占据轨道能(E_{LUMO}),前线轨道能量差 ΔE($E_{LUMO} - E_{HOMO}$),绝对硬度(η),电负性(χ),软度(S)和亲电指数(ω),苯环上碳原子的 Mulliken 电荷,苯环上的碳原子静电荷 E,分子表面积 TSA,氢键受体个数 HBA,分子的总能量(E_T),极性曲面面积(PSA),分子体积 V. 热力学参数:内能(E_{th})、恒容热容(CV)、熵(S)、总能(E_T)、焓(H°)、自由能(G°)和零点振动能(ZPVE). E_{th}、CV 和 S 均为系统在标准状态下(T = 298.150 K, P = 101 325 Pa)计算得到,这些参数能量贡献包括平动能、振动能和转动能分量,均由 Spartan06 v1.12 计算得到.

本文中的分子结构优化计算以及量化参数的选取是在 spartan06 v1.12 软件上完成的,对变量参数线性回归分析验证是通过 Spss 13.0① 化学计量学软件完成. 其中 R 表示复相关系数,S 表示标准偏差,F 是 fisher 检验值,R 均在显著性水平 p = 0.06 下进行.

2 结果与讨论

2.1 量化计算结果

含硫苯衍生物部分量子化学参数及对发光细菌生物毒性列于表 1.

2.2 模型的建立

应用 RM1 法计算含硫苯衍生物分子的量子化学参数,作为含硫苯衍生物分子结构描述符,采用 SPSS13.0 软件进行多元逐步线性回归,按照相对标准偏差最小,相关系数最大的原则,建立含硫苯衍生物物对发光细菌生物毒性 lg(1/EC50) 的 QSAR 模型,并对相关关系进行显著性检验,结果表明,含硫苯衍生物物对发光细菌生物毒性 lg(1/*EC*50) 可用下式来定量描述:

$$\lg(1/EC50) = 10.238 - 0.009\Delta E + 3.034E(6) + 2.108\sum E - 0.292\lg K_{OW} + 0.522\sum \sigma$$

$$N = 29, R = 0.930, SD = 0.1843, F = 29.318 \quad (1)$$

其中 N 是化合物的个数,R 是相关系数,F 为 Fisher 检验值,SD 为标准偏差,为了更直观的表现所建模型的显著性,将预测的结果对活性的实验值做散点图,如图 3 所示,预测值与实验结果吻合良好.

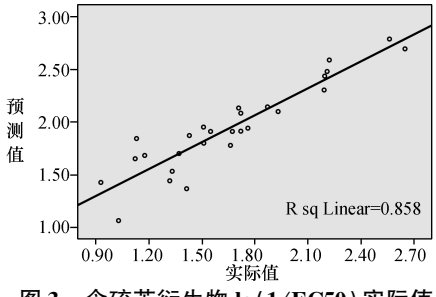


图 3 含硫苯衍生物 lg(1/EC50) 实际值与预测值的回归直线图

Fig. 3 Plot of predicted validation vs reference (log1/*EC*50) for thiobenzene-containing derivatives

①Norman H Nie. SPSS 13.0 for windows[CP]. SPSS Inc, Chicago, 2006.

表 1　含硫苯衍生物的部分量子化学参数及对发光细菌生物毒性

Table. 1　The value of toxicity and some quantum-chemical parmaters for thiobenzene-containing derivatives

编号	结构式	ΔE	E_{LUMO}	$E(6)$	$\sum E$	$\lg K_{ow}$	$\sum \sigma$	lg(1/EC50)
1	4 – NO_2 _ ϕ – S – CH_2 – $COOCH_3$	789.13	–903.44	–0.16	–0.39	2.29	0.81	1.45
2	4 – NO_2 _ ϕ – S – CH_2 – $COOCH(CH_3)_2$	790.40	–908.63	–0.12	–0.36	3.61	0.81	1.02
3	4 – NO_2 _ 2 – Cl – ϕ – S – CH_2 – $COOCH_3$	769.54	–909.77	–0.19	–0.22	3.00	1.04	1.76
4	4 – NO_2 _ 2 – Cl – ϕ – S – CH_2 – $COOCH(CH_3)_2$	799.00	–926.86	–0.07	–0.13	4.32	1.04	1.13
5	2 – NO_2 _ ϕ – S – CH_2 – $COOCH_3$	769.62	–885.82	–0.24	–0.38	2.29	0.81	1.33
6	2 – NO_2 _ ϕ – S – CH_2 – $COOCH(CH_3)_2$	848.31	–920.09	–0.06	–0.35	3.61	0.81	1.03
7	4 – Cl – 2 – NO_2 _ ϕ – S – CH_2 – $COOCH_3$	837.20	–941.95	–0.08	–0.16	3.00	1.04	1.51
8	4 – Cl – 2 – NO_2 _ ϕ – S – CH_2 – $COOCH(CH_3)_2$	837.06	–937.86	–0.09	–0.16	4.32	1.04	1.41
9	2,4 – diNO_2 _ ϕ – S – CH_2 – $COOCH_3$	804.21	–971.16	–0.09	–0.24	2.03	1.59	2.21
10	2,4 – diNO_2 _ ϕ – S – CH_2 – $COOCH(CH_3)_2$	802.59	–966.63	–0.09	–0.24	3.35	1.59	1.72
11	4 – NO_2 _ ϕ – SO – CH_2 – $COOCH_3$	814.24	–948.31	–0.16	–0.53	0.98	1.28	1.37
12	4 – NO_2 _ ϕ – SO – CH_2 – $COOCH(CH_3)_2$	809.26	–944.97	–0.14	–0.52	2.30	1.28	0.93
13	4 – NO_2 _ 2 – Cl – ϕ – SO – CH_2 – $COOCH_3$	807.49	–954.3	–0.05	–0.32	1.69	1.51	2.2
14	4 – NO_2 _ 2 – Cl – ϕ – SO – CH_2 – $COOCH(CH_3)_2$	802.35	–948.67	–0.06	–0.31	3.01	1.51	1.93
15	2 – NO_2 _ ϕ – SO – CH_2 – $COOCH_3$	848.30	–937.8	–0.07	–0.51	0.98	1.28	1.18
16	2 – NO_2 _ ϕ – SO – CH_2 – $COOCH(CH_3)_2$	847.26	–930.89	–0.03	–0.5	2.30	1.28	1.32
17	4 – Cl – 2 – NO_2 _ ϕ – SO – CH_2 – $COOCH_3$	833.82	–953.58	–0.13	–0.35	1.69	1.51	1.43
18	4 – Cl – 2 – NO_2 _ ϕ – SO – CH_2 – $COOCH(CH_3)_2$	831.89	–948.31	–0.08	–0.36	3.01	1.51	1.12
19	2,4 – diNO_2 _ ϕ – SO – CH_2 – $COOCH_3$	801.55	–985.26	–0.1	–0.44	0.58	2.06	2.65
20	2,4 – diNO_2 _ ϕ – SO – CH_2 – $COOCH(CH_3)_2$	799.37	–979.69	–0.09	–0.45	2.04	2.06	2.19
21	4 – NO_2 _ ϕ – SO_2 – CH_2 – $COOCH_3$	862.83	–1048.51	0.3	–0.84	0.92	1.48	1.71
22	4 – NO_2 _ ϕ – SO_2 – CH_2 – $COOCH(CH_3)_2$	862.94	–1039.09	0.32	–0.78	2.27	1.48	1.72
23	4 – NO_2 _ 2 – Cl – ϕ – SO_2 – CH_2 – $COOCH_3$	831.93	–1037.02	0.22	–0.63	0.66	1.71	2.56
24	4 – NO_2 _ 2 – Cl – ϕ – SO_2 – CH_2 – $COOCH(CH_3)_2$	832.38	–1032.48	0.23	–0.63	3.00	1.71	1.87
25	2 – NO_2 _ ϕ – SO_2 – CH_2 – $COOCH_3$	900.76	–1052.99	0.32	–0.78	0.95	1.48	1.51
26	2 – NO_2 _ ϕ – SO_2 – CH_2 – $COOCH(CH_3)_2$	900.63	–1041.3	0.38	–0.77	2.27	1.48	1.66
27	4 – Cl – 2 – NO_2 _ ϕ – SO_2 – CH_2 – $COOCH_3$	887.50	–1088.07	0.16	–0.72	0.66	1.71	1.67
28	4 – Cl – 2 – NO_2 _ ϕ – SO_2 – CH_2 – $COOCH(CH_3)_2$	896.07	–1065.28	0.35	–0.63	3.00	1.71	1.55
29	2,4 – diNO_2 _ ϕ – SO_2 – CH_2 – $COOCH(CH_3)_2$	867.84	–1093.83	0.34	–0.69	2.01	2.26	2.22

2.3　影响毒性的主要因素

前线轨道能量差 $\Delta E(E_{LUMO}-E_{HOMO})$,作为重要的量子化学参数在 lg(1/*EC*50)的值的变化中起着非常重要的作用,由于化合物的反应活性和前线轨道能量差 ΔE 的能量有关,HOMO 与 LUMO 能级差 $\Delta E(E_{LUMO}-E_{HOMO})$能够决定化合物反

应活性的高低. 具有较大 $\Delta E(E_{LUMO}-E_{HOMO})$ 时电子不容易发生跃迁,活性相应的减弱,而具有较小 $\Delta E(E_{LUMO}-E_{HOMO})$ 的化合物,活性较高,对发光菌的生物毒性就越强,这也与模型 1 得到的结论是一致的. 此外化合物的辛醇/水分配系数($\lg K_{ow}$)表征有机化合物的亲脂疏水性,即 $\lg K_{ow}$ 越大,该化合物就越容易穿过细胞膜,就越可能对有机体产生毒害作用,该类化合物对发光细菌的毒性呈现一定的负相关性,可能与化合物引入硝基和氯取代基,结构变化较大且该系列化合物硫基团有异有关. $\sum E$ 越大苯环上的正电荷越正,即被取代基拉走的电荷越多时,化合物的毒性越高,因此越强的吸电子取代基,其相应的化合物的毒性越大. 含硫苯衍生物对发光细菌的毒性大小还取决于苯环上取代基的类型和数目,取代基亲电性越强,对应化合物的毒性越强;取代基数目越多,σ 值越大,电荷密度越高,对应化合物的分子量越大,形成氢键能力越强,毒性越强. 该模型具有参数易得,计算简便,物理意义清楚等优点. 经检验,模型方程描述的发光菌生物毒性$\lg(1/EC50)$与量子化学参数的相关关系是显著的.

3 结 论

应用量子化学 RM1 方法优化了 29 个含硫苯衍生物的分子结构,从中选取有效的参数建立多元线性方程构建了含硫苯衍生物对发光细菌毒性 QSAR 研究模型. 结果表明,前线轨道能量差 $\Delta E(E_{LUMO}-E_{HOMO})$,辛醇/水分配系数($\lg K_{ow}$),取代基的 Hammett 电荷效应常数 $\sum\sigma$,碳原子 $C(6)$ 上的静电荷 $E(6)$ 及苯环上碳原子的静电之和 $\sum E$ 是影响含硫苯衍生物对发光细菌活性的主要因素,所得模型对该类化合物对发光细菌毒性有较好的预测效果. 得到的 QSAR 方程可以较好地评价和预测该类化合物的对发光细菌的生物毒性,为新的具有生物活性的含硫苯衍生物类化合物的设计与合成提供有理论依据.

参考文献:

[1] Yuan Qiong, Yuan Shen-gang, Chai Ge-qing, et al. QSAR study of fluorine-containing pesticides[J]. Acta Chimica Sinica, 1999, 57: 100-107.

[2] 杨江涛. 混沌遗传神经网络算法用于定量构效关系及相关分析体系研究[D]. 长沙:湖南大学, 2003.

[3] 王连生, 黄庆国, 韩朔睽, 等. 含硫苯衍生物对发光细菌的毒性效应及其定量结构活性相关分析[J]. 环境化学, 1994, 13(2): 22-24.

[4] 谢前, 孙红梅, 周家驹. 人工神经网络方法研究含硫芳香衍生物毒性与结构的关系[J]. 计算机与应用化学, 1996, 13(1): 310-319.

[5] Gerd Bruno Rocha, Ricardo Oliveira Freire, Alfredo May-all Simas, et al. RM1: Areparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I[J]. J. Comput. Chem., 2006, 27(10): 1101-1111.

[6] 马卫平. 线性和非线性方法在 QSAR/QSPR 研究中的应用[D]. 兰州:兰州大学, 2007.

[7] 王连生. 有机污染化学进展[M]. 北京:化学工业出版社, 2006: 124-193.