取代萘系列化合物对发光菌毒性的建模及稳健 性检验

魏东斌^{1,2},张爱茜^2,王连生^2,胡洪营^1(1. 清华大学环境科学与工程系环境模拟与污染控制国家重点联合实验室,北京 100084; 2. 南京大学环境学院,南京 210093)

摘要:遵照标准方法测定了 16 种取代萘系列化合物对明亮发光杆菌的急性毒性,用定量结构-活性相关方法构建了化合物毒性预测模型: -1gEC₅₀ = 5.592 - 7.489qH $^+$ + 0.777 E_{LUMO} + 0.009 (n=16, r_{adj}^2 = 0.9708, s=0.089, p<0.0001),蒙特卡洛模拟检验等方法检验结果表明所得模型稳健. 根据模型中的量子化学参数初步推测出该类化合物对发光菌为反应性毒性,主要包括物理分配和生化反应过程.

关键词:定量结构-活性相关;发光菌;取代萘系列化合物;蒙特卡洛模拟检验

中图分类号:X171.5 文献标识码:A 文章编号:0250-3301(2002)增刊-05-0016

QSAR Modeling and Robustness Tests on Toxicity of Substituted Naphthalenes to Photobacterium phosphoreum

Wei Dongbin¹, Zhang Aiqian², Wang Liansheng², Hu Hongying¹ (1. Environmental Simulation and Pollution Control State Key Joint Laboratory, Department of Environment Science & Engineering, Tsinghua University, Beijing 100084, China; 2. School of Environment, Nanjing University, Nanjing 210093, China)

Abstract: Acute toxicity (15min- EC_{50}) determination of 16 substituted naphthalene compounds to *Photobacterium phosphoreum* was undertaken according to the standard procedures, while the effects of molecular structures of selected compounds on their toxicity to test microorganism were logistically conducted by using quantitative structure-activity relationship (QSAR) technique. The relationship was developed as: - $\log EC_{50} = 5.592 - 7.489 qH^+ + 0.777 E_{LUMO} + 0.009$ (n = 16, $r_{adj}^2 = 0.970$, S = 0.089, p < 0.0001). The prediction model was tested with several different methods and the results indicated that the model was reasonable and robust. Furthermore, the quantum chemical parameters entering into QSAR model were used to discuss the possible toxicity pathways, and the results showed that the selected compounds were of reactivity and their toxicity behaviors were complex processes containing physical partition stages as well as bio-chemical reaction stages.

Keywords: QSARs; Photobacterium phosphoreum; Monte Carlo simulation; substituted naphthalene compounds

环境中有毒污染物的种类和数量日益增多,给生态环境造成巨大威胁,污染物的风险评价和管理非常必要,定量结构-活性相关(QSARs)研究为此开辟了新思路[1].从化合物的结构出发,研究分子结构与其性质或活性之间的定量关系,在此基础上,根据化合物的结构或借助已知特性预测其它性质或活性[2,3].在分子结构的描述方面,量子化学描述符因其显著的优越性而被广泛应用[3].但随方程中参数个数的增加,参数间自相关现象突出,限制了方程的预测能力,使方程变得不够"稳健",失去了本来的意义.

萘系列化合物属多环芳烃类,对生物有较大危害^[4],而我国目前仍采用传统工艺生产该类化合物,排放大量废水,有关这类化合物环境行为方面的研究却相对较少^[5].发光菌系低等微生物,位于食物链低营养级,用它作为毒性指示生物,具有方便、灵敏、经济,与其它毒性数据有较好相关性等优点^[6,7].

Vol. 23 ,Sup. Dec. .2002

本文在测定 16 个取代萘系列化合物对发 光菌急性毒性数据的基础上,筛选出具有明确

基金项目:国家自然科学基金重点项目(29837180)

作者简介:魏东斌(1972~),男,博士,主要从事污染生态化学和生态毒理学方面的研究.

理化意义的结构描述符 .构建稳健的毒性模型 . 推测可能的致毒机理.

1 材料与方法

1.1 材料

表 1 所列 16 个取代萘系列化合物均为分 析纯,用3%氯化钠溶液配成储备液,明亮发光 杆菌冻干粉(Photobacterium phosphoreum, T3 变种) 购自中国科学院南京土壤所.

培养液:酵母浸出汁 0.5g、胰蛋白胨 0.5g、

氯化钠 3.0g、磷酸氢二钠 0.5g、磷酸二氢钾 0.1g、甘油0.3g,加蒸馏水至100mL,pH=7.0 ±0.5;另加1.6g 琼脂粉则可制成固体斜面.

1.2 菌种的培养和毒性测试

菌种的复苏、培养以及毒性测定均参照标 准方法操作[8].求出发光菌相对发光率为 50 % 时所对应的化合物浓度,即 EC_{50} (mol L^{-1})值, 列于表 1.

1.3 结构描述符的计算

表 1 所选择化合物的名称、计算出的量子化学参数和对发光菌的毒性数据

Table 1 The calculated structural parameters and their toxicities of selected chemicals

编号	化合物		$E_{ m LUMO}$	qH ⁺	lg K _{OW}	- 1gEC ₅₀		
						实验值	估算值	残差
1	萘	95.66	- 1.268	0. 1565	3.316	4. 244	4. 280	- 0.036
2	2-溴-萘	101.32	- 0.551	0. 1467	4. 179	4. 968	4. 980	- 0.012
3	1,2-二溴-萘	111.11	- 0.831	0. 1518	4.842	4. 844	4.803	0.041
4	1-硝基-萘	106. 17	- 1.311	0. 1686	3.059	4. 262	4. 248	0.014
5	1-萘胺	99.37	- 0.175	0. 1858	2.089	4.953	4. 973	- 0.020
6	2-萘胺	101.13	- 0.175	0. 1844	2.089	5.034	4. 999	0.035
7	4-氯-1-萘胺	103.89	- 0.642	0. 1581	2.901	4.786	4.828	- 0.042
8	1-萘磺酸	112.01	- 1.269	0. 2703	1.459	3.552	3.572	- 0.020
9	2-萘磺酸	114. 10	- 1.218	0. 2769	1.459	3.607	3.583	0.024
10	2-氨基-1-萘磺酸	121.80	- 0.912	0. 2739	0. 232	3.935	3.919	0.016
11	5-氨基-2-萘磺酸	124. 65	- 1.076	0. 2783	0. 232	3.971	3.780	0. 191
12	7-氨基-4-羟基-2-萘磺酸	132.67	- 1.091	0. 2769	- 0.435	3.648	3.849	- 0.201
13	1-萘酚	94. 67	- 0.368	0. 2191	2. 649	4. 492	4. 527	- 0.035
14	2-萘酚	95.82	- 0.343	0. 2174	2.690	4.470	4. 544	- 0.074
15	5-氨基-1-萘酚	104.88	- 0.279	0. 2198	1.770	4. 634	4. 656	- 0.022
16	2,2-联萘酚	224. 10	- 1.453	0. 2209	4.504	4. 788	4.784	0.004

化合物分子的量子化学描述符用 MOPAC97^[9]软件包的 AM1 算法^[10]计算 .主要 包括: 平均分子极化率(), 偶极矩(µ),单位 为原子单位(A. U.); 分子的最高占据轨道能 (E_{HOMO}) 和最低未占据轨道能 (E_{LUMO}) ,单位为 电子伏特(eV); 分子中最正的氢原子净电荷 (qH⁺)和最负的原子净电荷(q⁻),单位为原子 电荷单位(A.C.U.) 分子生成热(HOF),分 子总能量(TE), 电子能量(EE), 分子的芯-芯 排斥能(CCR),能量单位为电子伏特(eV).部分 描述符列干表 1.

所有统计分析均在"STATGRAPHICS"软 件上实现(STSC, Ver. 4.0.1985),方程中参数 间的共线性用变异膨胀因子(VIF)评价[3].

2 结果与讨论

2.1 模型的建立及评价

一般认为,化学品进入生物体后,如不考虑 代谢作用,其生物效应是疏水性、静电效应和立 体效应的函数^[11]. 量子化学参数容易获得、理化意义明确,有利于揭示化学品的致毒机理. 但随着被选参数个数的增加,参数间可能存在共线性现象. 为此,本文首先运用偏相关分析法对所有参数进行考查,选出共线性低、独立性高的参数,然后通过多元逐步回归分析法构建出lgEC₅₀的预测方程(1):

-
$$\lg EC_{50} = 5.592 - 7.489 qH^{+}$$

+ $0.777 E_{LUMO} + 0.009$ (1)
 $n = 16, r_{adj}^{2} = 0.970, S = 0.089,$
 $F = 161.36, p < 0.0001$

式中,n 代表样本数, r_{adj}^2 是经过自由度校正的复相关系数,S 是方程的标准偏差,F 是 F-检验值,p 为显著性水平.

方程(1) 具有高的相关系数 r_{adj}^2 ,能很好地拟合 16 个取代萘系列化合物的 $lgEC_{50}$ 值. 但方程中只包括描述静电效应和立体效应的描述符,说明该类化合物对发光菌的致毒过程主要与这 2 个效应有关. Hansch 等认为疏水性是影响生物效应的因素之一,但方程(1) 中并未反映出这一点. 故单独考查 $lgEC_{50}$ 和正辛醇/水分配系数 lgK_{OW} 间的相关性:

-
$$1 \text{gEC}_{50} = 3.857 + 0.2307 \text{ lg } K_{\text{OW}}$$
 (2)
 $n = 16$, $r_{\text{adj}}^2 = 0.431$,

S = 0.387, F = 12.34, P = 0.003

结果表明 $I_g EC_{50}$ 和 $I_g K_{OW}$ 间有一定的相关性,即疏水效应对致毒过程也有一定影响,只是没有立体效应和静电效应明显. 另外,从化合物结构考虑,这类化合物母体环相似,大多数含有极性基团如一 NH_2 ,一 SO_3H 和一OH,易与水分子形成分子间氢键而降低了 K_{OW} 参数区分毒性差异的敏感性. 加之,发光菌系单细胞微生物,体内脂肪含量低,化学品对它的毒性效应与化合物本身的疏水性/亲脂性关系不大,而化合物的活性中心与生物体内酶的相互作用显得更为重要.

经检验表明方程(1) 显著,参数间无严重 共线性(表 2),毒性预测结果如图 1 所示. 为进 一步检验方程(1)的稳健性,选择修正的 Jacknife 法^[3] ,表 3 列出了检验时 r_{adj}^2 的变化 ,可以看出 ,方程比较稳健. 从毒性数据看出 :引入取代基 — SO_3H 降低了化合物毒性 ,而引入其它基团都或多或少地增加毒性 ,由此可粗略看出不同取代基对毒性的贡献 : — $NH_2 > -OH > -Br > -NO_2 > -SO_3H$. 另外 ,由于该样本较小 ,用 K-S 方法检验预测模型 (1) 残差的频数分布 (如图 2) ,结果表明 ,残差的分布符合正态函数 (1.401 × 10^{-15} , 0.097 2) ,斜度系数为 - 0.190 , 峭度系数为 3.386.

表 2 方程(1)参数间的自相关矩阵及各参数的 变异膨胀因子(VIF)

Table 2 Correlation matrix and VIF of eq. (1)

		相关矩阵		VIF
		$E_{ m LUMO}$	qH^{+}	eq. (1)
	1.000			1. 202
$E_{ m LUMO}$	0.486	1.000		1. 143
qH +	- 0.149	0. 193	1.000	1.000

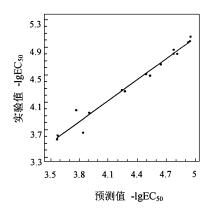


图 1 95 %的置信度时观测值与预测值的关系

Fig. 1 Plot of observed values *vs* predicted values at 95 % confidence level

表 3 方程(1) Jacknife 检验结果

Table 3 Summary of results of Jacknife tests

Tubic 2	Buillinary of	results of suckin	re tests		
删除个数	回归次数	- 1gEC ₅₀			
(<20 % of n)		平均 r_{adi}^2	平均 S		
1	16	0.9695	0.0891		
2	30	0.9679	0.0913		
3	50	0.9666	0.0925		
平均值		0.9680	0.0910		

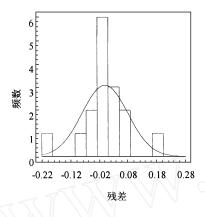


图 2 方程(1) 残差的频数分布

Fig. 2 The frequency distribution of eq. (1)

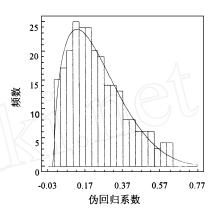


图 3 伪回归系数 r^{2*}的经验分布

Fig. 3 The empirical distribution of r^{2*}

2.2 模型的检验

蒙特卡洛模拟检验法对小样本回归方程的显著性检验具有重要的意义 $^{[12]}$. 本文仅有 16个观测值,采用蒙特卡洛模拟检验法,重复模拟 250 次,得到 250 个伪回归方程和 250 个伪回归系数 r^2 *;所得 r^2 *的范围从 0.008 到 0.631. 对 r^2 *进行经验分布拟合的 Chi-方检验. 结果表明,Chi-方值为 5.080,显著性水平为 0.985. 在 0.05 的显著性水平下, r^2 *拟合的经验分布与 分布(1=1.639,2=5.401)无显著差异(如图 3), 函数的概率分布图上,与概率 0.975(双侧)对应的 r^2 *为 0.580. 由于回归方程(1)的 r^2 域值(0.970)远远大于该 r^2 省值,故可判断回归方程(1)与随机预测的效果差异显著. 据此可得:在一定范围内,可根据量子化学描述符预测该组化合物对发光菌的毒性.

2.3 机理的探讨

发光菌的发光过程可简略描述如下:

 $FMNH_2 + RCHO + O_2$ $\xrightarrow{\text{bacterial fluroenzyne}} FMN + RCOOH + H_2O + light$

FMN 是黄素单核苷酸,FMNH₂ 是其还原形式,FMNH₂是一个重要的辅酶,起到传递氢的作用.在 FMNH₂ 分子中,—NH—易与其它分子发生氢键作用.本文所选化合物中电负性较高的原子如 O, N, F等均可能与 FMNH₂ 形成氢键,阻碍了FMNH₂对氢的传递作用,使发

光菌的发光受到抑制.

此外,运用t-检验法分别对预测模型(1)中 因变量和自变量进行检验,参数 qH⁺, E_{LUMO} 和 的 t 值分别为 - 15.285, 12.559 和 9.985. 说明分子中最正氢原子的净电荷(qH⁺)对毒 性的影响最显著,随着 qH+的增大,化合物毒性 相应降低,化合物中的氢原子和水分子中的氧 原子容易形成氢键,导致 qH+值高的化合物更 倾向于分配在水相而不容易穿过细胞膜,表现 为毒性较低. 例如,含有 - SO₃H 的待测化合物, K_{OW} 值相对较低,而 qH^+ 却较其它化合物高,在 水相中存在较强的溶剂效应. 最低未占据分子 轨道能(E_{LUMO})描述分子接受外来电子对的能 力, ELIMO值越高,化合物与生物体内靶标分子 间电子传递作用越强烈,化合物分子提供电子 对,而靶标分子接受电子对,总体表现为毒性增 大,这一点可作如下解释:化合物分子中的一些 高电负性原子如 CI, O, N 可提供孤对电子,容 易接受来自 FMNH2 的氢质子,形成氢键,抑制 了发光过程. 含有 - SO₃H 的化合物却属例外, 由于该类化合物水溶性较高,分子的平均极化 率()高,毒性大.极化率与化合物分子的体积 呈正相关,而分子的体积大小决定该分子在溶 剂中形成与其匹配的窝穴时所需的能量,因此, 描述符 值越高,分子的体积越大,分子的变形 性越大,更容易穿过细胞壁分配进入极性较小 的细胞内部,与靶标分子作用,毒性变大.

3 结论

综上,该组取代萘系列化合物对发光菌的 毒性不只是简单的分配过程,而是包括生化反 应等在内的复杂过程.另外,该研究证实,量子 化学描述符是一种能很好描述分子结构的参 数,容易计算获得且理化意义明晰,成为建模研 究的重要参数.

参考文献:

- Blum DJ, Speece RE. Determining chemical toxicity to aquatic species. Environ. Sci. Technol., 1990, 24: 284 ~ 293.
- 2 王连生, 韩朔睽. 有机污染化学进展. 北京:化学工业出版 社, 1998. 329~335.
- Wei DB, Zhang AQ, Wu CD, Han SK, Wang LS. Progressive study and robustness test of QSAR model based on quantum chemical parameters for predicting BCF of selected polychlorinated organic compounds. Chemosphere, 2001, 44:1421~1428.
- 4 章思规. 精细有机化学品技术手册(下册). 北京: 科学出

- 版社, 1992. 1353~1375.
- 5 许昭怡. 萘系列化合物的吸附和氧化行为及其对生物处理过程的抑制作用研究. 南京大学博士论文, 1999. 1~5.
- 6 Dezwart D, Slooff W. The Microtox as on alter-native assay in the acute toxicity assessment of water pollutants. Aquat. Toxicol., 1983,4:129~138.
- 7 顾宗濂. 发光细菌法检测水土环境毒性的进展和评价. 环境科学与技术. 1987, **2**:2~6.
- 8 南京大学环境科学系环境生物教研室. 环境生物实验技术与方法. 南京:南京大学出版社,1989. 145~151.
- 9 Stewart JP. CSMOPAC97, Cambridgesoft Co. 1997.
- 10 Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP. AM1. A new general purpose quantum mechanical molecular model. J. Am. Chem. Soc., 1985, 107: 3902 ~ 3909.
- 11 Hansch C, Leo A. Exploring QSAR: Fundamentals and applications in chemistry and biology. American Chemical Society, Washington DC, 1995.
- 12 Lund I A. A Monte Carlo method for testing the statistical significance of a regression equation. J. Appl. Meteor., 1970, $9:330 \sim 332$.