

化合物毒性及三维结构参数的定量构效关系研究<sup>\*</sup>郭亦然 张燕玲 艾路 乔延江<sup>\*\*</sup>

(北京中医药大学中药学院 北京 100102)

**摘要** 目的:建立小分子化合物毒性的三维定量构效关系模型,探索化合物毒性数据和三维结构参数之间关系的方法。方法:利用比较分子力场分析方法(CoMFA),建立了一组对发光菌有急性毒性的小分子的三维定量构效模型。结果:模型的交叉验证相关系数 $q^2=0.731$ ,非交叉验证相关系数 $r^2=0.973$ ,标准偏差 $SE=0.122$ , $F=70.910$ 。结论:该模型具有较好的预测能力,表明在甲基的邻对位减小取代基体积或电负性可以降低化合物毒性。

**关键词**:化合物毒性;三维构效关系;比较力场分析方法

## 3D- QSAR Discovery of Toxicity and Chemical Parameters of Benzene Drugs

GUO Yi-ran, ZHANG Yan-ling, WANG Yun, QIAO Yan-jiang

(Beijing University of Chinese Medicine, Beijing, 100102)

**ABSTRACT Objective:** To predict the relationship between the toxicity and the structural parameters of Benzene drugs, a 3D- QSAR (three dimensional quantitative structure- activity relationship) model was established. **Methods:** By determining their lowest energy conformation, molecules are aligned according to one specific rule and the scope of the molecular field is defined around the aligned molecules. The characters of the molecule field such as the electrostatic and steric interaction energies are calculated. Then Partial Least Square (PLS) method is exploited to analyze the relationship between the toxicity of compounds and characters of the molecular field and finally Comparative Molecular Field Analysis (CoMFA) model was built. **Results:** The cross- validated  $q^2$  of the model is 0.731 and the cross- validated  $r^2$  is 0.973, so the predictive abilities of these models are considerable. **Conclusion:** This model has a good predictive ability. To reduce the bulk of substituent in the para- position and ortho- position of methyl or decrease the negative charge of substituent in the para- position and ortho- position of methyl can increase the toxicity.

**Key Words:** Toxicity of compound; 3D- QSAR(Three Dimensional Quantitative Structure- Activity Relationship); CoMFA

## 1 引言

化合物的“毒性”,泛指化学试剂对生物系统产生的不良影响(此处是否有参考文献,如果有,请标明,下同)。天然有毒化合物广泛存在于动植物中,而(两句是并列关系,不是转折,删去)人工合成有毒物包括各种工业试剂、农药、洗涤剂、添加剂等(删去)等。化合物分子结构和致毒机理的多样性,使得毒理实验研究往往费时久而准确性欠佳。可以利用计算机辅助方法对化合物毒性进行理论预测,从而指导毒理研究。

本文通过比较分子场分析(Comparative Molecular Field Analysis, CoMFA)方法进行小分子芳香类化合物对发光菌的急性毒性与其分子特征之间的三维定量结构-性质相关关系(three dimensional quantitative structure- activity relationship, 3D- QSAR)的研究,探讨了三维定量构效关系方法在毒性化合物性质预测中的应用,建立了较好的3D- QSAR模型。

## 2 仪器与方法(或设备与方法)

## 2.1 化合物的选择

芳香类化合物是最早被发现具有致癌性或潜在致癌性的环境污染物质之一。根据文献报道,部分小分子芳香类化合物

对发光菌具有急性毒性<sup>[1]</sup>。本文以文献实验数据(见表1)为基础进行芳香类化合物三维结构与毒性数据相关性研究。

任选表1中两个化合物为测试集,其余为训练集,用于模型计算。

## 2.2 计算方法的选择

比较分子场分析(Comparative Molecular Field Analysis, CoMFA)方法的基本原理是:一组结构相似的化合物以同一种方式作用于一个受体,它们的生物活性取决于各化合物周围分子场的差别,这种分子场可以反映化合物和受体之间的非键作用特性<sup>[2]</sup>。因此,在不了解受体三维结构的情况下,研究化合物分子周围的力场分布以及其与化合物活性之间的定量关系,可以推测受体性质,建立一个定量模型设计新化合物,并定量预测化合物活性。本文采用CoMFA方法建立模型,计算分为3个步骤:首先,确定活性分子的活性构象,按照一定的规则进行分子的叠合;然后,在叠合好的分子周围定义分子场的空间范围,在定义的空间中按照一定的步长均匀划分格点;在每个格点上用探针离子来评价格点上的分子场特征,在计算分子中采用不同的分子场势能函数;最后,应用偏最小二乘(Partial Least Square, PLS)方法建立化合物的a、c值和分子场特征之间的关系[1, 2]。在CoMFA模型中电荷计算所使用的

\* 基金项目:国家自然科学基金(No. 30472180), 国家中医药管理局科学研究专项课题基金(04- 05zt03)资助项目

作者简介:郭亦然,女,博士研究生

\*\* 通讯作者:乔延江,男,博士生导师, yjqiao@ 263.net

(收稿日期:2006- 04- 08 接受日期:2006- 04- 26)

是Gast-Huck方法<sup>[7]</sup>,这种方法所计算出的电荷数是 $\sigma$ 电荷与 $\pi$ 电荷之和[ iii]。

表1 小分子化合物的毒性数据实测值与预测值<sup>[1]</sup>

Table 1 toxicities of benzene drugs by experimented and predicted

化合物	实测	预测	残差	化合物	实测	预测	残差
苯	3.34	3.3523	-0.0122964	甲苯	3.08	3.20726	-0.127265
氯苯	3.86	3.87526	-0.0152588	1,2-二氯苯	4.38	4.21306	0.166942
1,4-二氯苯	4.39	4.3003	0.0896953	1,2,3-三氯苯	4.53	4.55851	-0.0285049
1,2,4,5-四氯苯	5.51	5.49419	0.0158135	溴苯	3.78	3.67396	0.106039
苯酚	3.64	3.6792	-0.0392036	2-甲基苯酚	3.75	3.84744	-0.0974361
苯胺	3.28	3.23052	0.049781	硝基苯	3.82	3.83737	-0.0173688
对二甲苯	3.68	3.71845	-0.0384501	3,4-二氯甲苯	4.36	4.24866	0.11134
1,3-二氯苯	4.24			4-硝基苯酚	4.05		
1,2,4-三氯苯	4.50	4.72898	-0.228983	2-硝基甲苯	3.91	3.84454	0.0654571

2.2.1 构象分析和分子叠合:应用构象搜索中系统搜索的方法,以10度为间隔搜寻该类化合物的低能构象,最大能量差异为99999kcal/mol。其余参数为默认值,并将最低能量构象默认为优势构象<sup>[3,4]</sup>(当多个构象同时具有最低能量时,选择出现概率最大的构象作为化合物可能的优势构象)。对构象优化好的该类化合物加载了Gasteiger-Huckel电荷<sup>[5]</sup>,用于CoMFA的模型计算。

由于此系列化合物皆为小分子芳香类化合物,因此采用的叠合母核为苯环母核。对训练集分子进行分子叠合。

2.2.2 构效关系的建立:根据分子叠加结果,对每个分子所选取的优势构象进行CoMFA计算,使用探针原子为sp<sup>3</sup>杂化的碳原子,原子半径为0.1nm,带e<sup>+</sup>电荷,网格步长为0.02nm,计算该分子的立体场和静电场势能值。

应用偏最小二乘(Partial Least Square, PLS)方法建立化合物的急性毒性和分子场特征之间的关系<sup>[6,7]</sup>。静电场能,立体场能,以及文献中测得的急性毒性数据保留时间值都用于模型PLS计算。

2.3 仪器软件

本文所有工作均在SGI O2工作站上,基于Sybyl 6.6分子模拟系统下完成,计算过程中若无特别指明,所用参数均为软件中的默认值。(放到前面)

3 结果

3.1 分子叠合

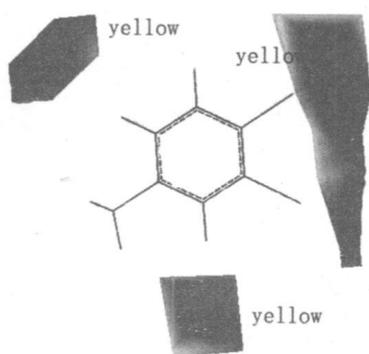


图2 CoMFA 立体场等势图  
Fig. 2 Steric contour map of CoMFA

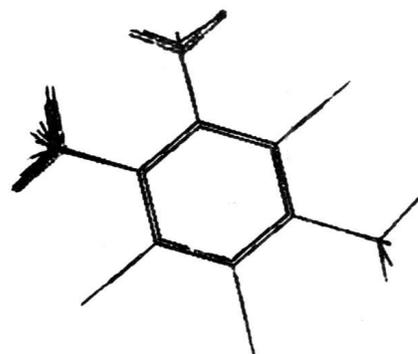


图1 以苯为模板进行叠合图  
Fig. 1 The aligned compounds of benzene

训练集化合物分子结构的叠合图见图1。(此图说明、反映了什么,请简要介绍,写法如3.2)

3.2 构效关系

图2、图3是立体场和静电场等势图,在立体场等势图中绿色区域表示在该区域引入体积大的基团将提高化合物毒性,而黄色区域则表明该区域附近引入体积大的基团将降低化合物毒性。在静电场等势图中红色区域表示引入负电性基团毒性将提高,而蓝色区域则表示引入正电性基团将提高毒性。(在图中似乎只有黑色和浅黑色,请用箭头指示或者用图像处理工具处理)

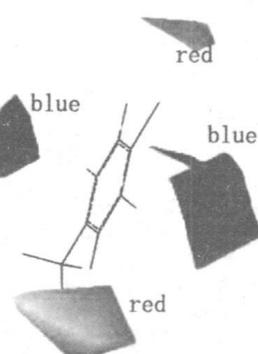


图3 CoMFA 静电场等势图请补充英文对照  
Fig. 3 Electrostatic field contour map of CoMFA

通过训练集所得 CoMFA 模型中: 交叉验证相关系数  $q^2$  为 0.731, 最佳主成分数为 5。由最佳主成分数建立的 CoMFA 模型的传统相关系数  $r^2=0.973$ 、 $F=70.910$ , 标准偏差  $SE(\text{standard error})=0.122$ 。各个参数对所建构效关系方程的贡献分别是: 立体场为 32.6%, 静电场为 67.4%。

该 CoMFA 模型对 16 个训练集化合物的活性预测值如表 1 所示, 预测值与实验测定值的相关图见图 3。

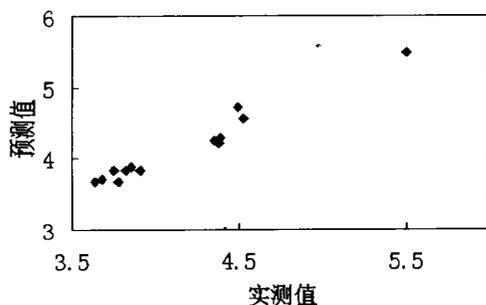


图 4 化合物毒性实验值与 CoMFA 模型预测值相关图

Fig. 4 The predictive and experimental activity of training set

表 3 显示了利用训练集化合物得到的 CoMFA 模型对预测集化合物 1,3- 二氯苯、4- 硝基苯酚进行预测得到的活性预测值与实验测定值的比较, 此预测结果表明此方程是可靠的, 对小分子化合物的毒性数据的预测具有一定的意义。

表 3 CoMFA 模型对测试集的预测结果

Table 3 the experimental and predicted activities and residual values of of test set

化合物	实测值	预测值	残差
1,3- 二氯苯	4.24	4.41961	- 0.17961
4- 硝基苯酚	4.05	4.47538	- 0.42538

## 4 讨论

通常交叉验证相关系数  $q^2 > 0.5$  即表明由相应的变量所建立的模型具有较好的预测能力<sup>[8]</sup>, 本模型  $q^2$  值为 0.731, 表明该模型具有较好的预测能力。

从 CoMFA 等势图中可以看出: 在此 CoMFA 模型中, 体积较大的取代基位于间位时会降低化合物毒性。如 1,3- 二氯苯, 由于取代基位于间位, 所以毒性比对位氯取代的 1,4- 二氯苯的毒性低。

在此 CoMFA 模型中, 在甲基邻位和对位, 取代基负电性越大, 小分子毒性越强, 如 3,4- 二氯甲苯的毒性比对二甲苯的毒性强, 2- 硝基甲苯的毒性比 3,4- 二氯甲苯的毒性弱; 在甲基的间位取代基正电性越大, 则小分子的毒性越强, 如 3,4- 二氯甲苯的毒性比 2- 甲基苯酚的毒性强。

## 5 结论

实验结果表明, 本实验建立的 CoMFA 模型具有较好的预测能力; 而通过在小分子芳香类化合物甲基的邻对位减小取代基体积或电负性都可以降低其生物毒性。

本文通过对芳香类化合物的生物毒性和三维结构的 QSAR 研究, 建立了较好的三维结构- 保留时间的 CoMFA 模型, 探讨了小分子化合物毒性数据预测的方法。通过此实验, 可以指导对同系列未知成分的小分子芳香类化合物毒性数据进行预测, 并提供相关的结构类别信息; 为进一步通过 QSAR 方法指导同类化合物毒性数据预测提供了依据。

### 参考文献

- [1] 王连生, 韩朔睽. 分子结构、性质与活性 [M]. 第一版. 化学工业出版社, 1997: P86- 88
- [2] 刘仲杰, 刘叔倩. 三维定量构效关系最新研究进展 [J]. 河北科技大学学报, 1999, 20(1): 17- 21
- [3] 谷妍, 陈敏伯, 董喜城, 等. 皂甙的三维定量构效关系研究 [J]. 化学学报, 2000, 58(12): 1534- 1539
- [4] 张燕玲, 艾路, 王耘, 等. 黄酮化合物色谱保留时间与其三维结构的关系研究. 现代生物医学进展, 2006, (2): 9- 12
- [5] 邹霞娟, 来鲁华, 金桂玉, 等. 新型含噻嗪酮基胍类化合物的 3D-QSAR 研究 [J]. 物理化学学报, 2002, 18(6): 513- 516
- [6] 刘仲杰, 刘叔倩. 三维定量构效关系最新研究进展 [J]. 河北科技大学学报, 1999, 20(1): 17- 21
- [7] 侯廷军, 徐筱杰. 比较分子场分析方法研究的最新进展 [J]. 化学进展, 2001, 13(6): 436- 440
- [8] 万升标, 易翔, 郭宗儒. 1- 咪唑基苯并二氮杂卓类法呢基蛋白转移酶抑制剂的三维定量构效关系研究 [J]. 药学报, 2001, 36(6): 423- 426