

# 苯酚类化合物光解量子产率与结构性质的定量关系

牛军峰,余刚,施玮 (清华大学环境科学与工程系,持久性有机污染物研究中心,北京 100084)

**摘要:**采用量子化学 PM3 算法计算得到的苯酚类化合物的量子化学参数,应用偏最小二乘(PLS)算法,建立了能预测苯酚类化合物光解量子产率( $Y$ )的定量结构-性质关系(QSPR)模型.影响苯酚类化合物光解量子产率的主要因素是分子最高占据轨道能( $E_{\text{HOMO}}$ )和分子生成热( $HOF$ ), $\log Y$  值随着  $E_{\text{HOMO}}$  的增大而增大,随着  $HOF$  的增大而减小.分子的电负性和 $(E_{\text{LUMO}}-E_{\text{HOMO}})^2$  对  $\log Y$  值也有一定的影响,电负性和 $(E_{\text{LUMO}}-E_{\text{HOMO}})^2$  较小的分子,其  $\log Y$  值也较大.

**关键词:** 定量结构-性质关系(QSPR); 光解量子产率; 苯酚类化合物; 偏最小二乘(PLS)

中图分类号: X131 文献标识码: A 文章编号: 1000-6923(2004)01-0081-04

**Quantitative structure-property relationship for photodegradation quantum yields of phenols.** NIU Jun-feng, YU Gang, SHI Wei (Department of Environmental Science and Engineering, Research Centre for Persistent Organic Pollutants, Tsinghua University, Beijing 100084, China). *China Environmental Science*, 2004,24(1): 81~84

**Abstract:** Adopting quantum chemical parameters of phenols computed with quantum chemical PM3 algorithm, quantitative structure-property relationship (QSPR) model which could predict photodegradation quantum yields ( $Y$ ) of phenols was established using partial least squares (PLS) algorithm. It was considered that main factors affecting  $\log Y$  of phenols were the energy of the highest occupied molecular orbital ( $E_{\text{HOMO}}$ ) and the heat of formation ( $HOF$ ). The  $\log Y$  values increased with the increase of  $E_{\text{HOMO}}$  and decreased with the increase of  $HOF$ . Besides, element electronegativity and  $(E_{\text{LUMO}} - E_{\text{HOMO}})^2$  also affected definitely the  $\log Y$  value; the smaller the former, the greater the latter.

**Key words:** quantitative structure-property relationship (QSPR); photodegradation quantum yield; phenols; partial least squares (PLS)

苯酚类化合物是环境中分布广泛的重要有机污染物<sup>[1]</sup>.由于一些酚类化合物具有较强的毒性和持久性,因此被美国环境保护局列入 129 种环境优先控制污染物黑名单<sup>[2]</sup>.

光解是苯酚类化合物在水体环境中的重要转化过程<sup>[3-5]</sup>.其中,光解量子产率是研究这些物质光化学行为的一个重要参数,它对于有机污染物的环境归宿和风险性评价有重要的作用.迄今为止,对于水体中苯酚类化合物光解量子产率( $Y$ )的研究报道甚少.同时,由于时间和经费的限制,难以逐个测定苯酚类化合物的光解量子产率.有机污染物定量结构-性质关系(QSPR)研究是解决这种问题的必要手段.作者以容易计算得到且具有明确物理-化学意义的分子的量子化学参数为自变量<sup>[6,7]</sup>,采用近年来在化学计量学方面得到了广泛应用的偏最小二乘(PLS)算法<sup>[8]</sup>,对苯酚类化

合物的光解量子产率进行 QSPR 研究,考察苯酚类化合物光解量子产率与它们分子描述符之间的定量关系.

## 1 材料与方法

### 1.1 研究对象

研究对象为 18 种苯酚类化合物,其中 15 种苯酚类化合物的光解量子产率( $Y$ )取自 Serpone 等人<sup>[5]</sup>测定的暴露于大气中的水-TiO<sub>2</sub>(Degussa P-25; 2g/L; pH 3)平衡分散体系中不同苯酚类化合物的光解量子产率.这 15 种苯酚类化合物分别是苯酚、氯代苯酚(4 种)和甲基取代苯酚(10 种),如表 1 所示.

收稿日期: 2003-05-06

基金项目: 国家自然科学基金资助项目(50238020)

\* 责任作者,教授, yg-den@tsinghua.edu.cn

表 1 苯酚类化合物的光解量子产率

Table 1 The phenol compounds under study and their quantum yields

编号	化合物	$Y$ (实测值)*	$\log Y$ (实测值)	$\log Y$ (拟合值)	$SE-pred.^{\#}$	残差**	$Y$ (拟合值)
1	苯酚	0.11	-0.959	-0.956	$\pm 0.029$	-0.003	0.11
2	2-甲基苯酚	0.13	-0.886	-0.793	$\pm 0.018$	-0.093	0.16
3	3-甲基苯酚	0.14	-0.854	-0.806	$\pm 0.019$	-0.048	0.16
4	4-甲基苯酚	0.18	-0.745	-0.707	$\pm 0.018$	-0.038	0.20
5	2,3-二甲基苯酚	0.22	-0.658	-0.649	$\pm 0.021$	-0.009	0.22
6	2,4-二甲基苯酚	0.30	-0.523	-0.555	$\pm 0.027$	0.032	0.28
7	2,5-二甲基苯酚	0.25	-0.602	-0.634	$\pm 0.021$	0.032	0.23
8	2,6-二甲基苯酚	0.33	-0.481	-0.627	$\pm 0.022$	0.146	0.24
9	3,4-二甲基苯酚	0.28	-0.553	-0.568	$\pm 0.026$	0.015	0.27
10	3,5-二甲基苯酚	0.18	-0.745	-0.673	$\pm 0.019$	-0.072	0.21
11	2,3,5-三甲基苯酚	0.31	-0.509	-0.491	$\pm 0.033$	-0.018	0.32
12	2-氯苯酚	0.13	-0.886	-0.992	$\pm 0.032$	0.106	0.10
13	3-氯苯酚	0.11	-0.959	-1.005	$\pm 0.033$	0.046	0.10
14	4-氯苯酚	0.13	-0.886	-0.845	$\pm 0.020$	-0.041	0.14
15	2,4-二氯苯酚	0.11	-0.959	-0.904	$\pm 0.024$	-0.055	0.12
16	3,5-二氯苯酚			-1.127	$\pm 0.045$		0.07
17	2,4,5-三氯苯酚			-0.907	$\pm 0.025$		0.12
18	五氯苯酚			-0.852	$\pm 0.021$		0.14

注:\* 数据取自文献[5]; #  $SE-Pred$ : 拟合值标准误差; \*\* 残差= $\log Y$ (实测值) -  $\log Y$ (拟合值)

## 1.2 分子性质参数和结构参数

采用 MOPAC(6.0 版)软件中的 PM3 算法计算这些化合物的量子化学参数.控制 MOPAC 计算的关键词为:PM3、ESP、POLAR、DIPOLE、PRECISE、NOINTER.共选取了 11 个量子化学参数.这些量子化学参数是反映分子整体特征的分子结构参数,如相对分子质量( $M_w$ )、分子生成热( $HOF$ )、平均分子极化率( $a$ )和偶极矩( $\mu$ )、分子最高占据轨道能( $E_{\text{homo}}$ )、分子最低未占据轨道能( $E_{\text{lumo}}$ )、分子总的芯排斥能( $CCR$ )、电子能( $EE$ )、分子总能量( $TE$ )、苯酚类分子中 C 原子所带最大负电荷( $q_C^-$ )、苯酚类分子中 H 原子所带最大正电荷( $q_H^+$ )等.对有机物分子光解 QSPR<sup>[6,7]</sup>的研究表明,分子的前线轨道能  $E_{\text{lumo}} - E_{\text{homo}}$ 、 $(E_{\text{lumo}} - E_{\text{homo}})^2$  和  $E_{\text{lumo}} + E_{\text{homo}}$  对于模型的建立也有着重要的影响.因此将这 3 个参数也作为自变量进行考察.其中, $E_{\text{lumo}} - E_{\text{homo}}$  和  $E_{\text{lumo}} + E_{\text{homo}}$  分别和分子的绝对硬度和电负性有关.

## 1.3 统计分析

采用 SIMCA(6.0)软件进行 PLS 分析,计算条

件按程序的缺省设置值.当某个 PLS 主成分交叉验证的  $Q^2$ (当前 PLS 主成分所能解释的因变量的总方差的比例)值大于程序缺省设置的显著性水平限(0.097)时,即认为该主成分有意义.当  $Q^2_{\text{cum}}$ (所提取的所有 PLS 主成分所能解释的因变量总方差的比例)大于 0.5 时,认为所建立的模型有较好的预测可靠性.采用 PLS 主成分数( $A$ )、 $Q^2_{\text{cum}}$ 、拟合值和实测值的相关系数( $R$ )和模型的显著性水平( $p$ )等参数来表征模型.除了 PLS 分析提供的拟合值的标准误差( $SE-Pred.$ )外,同时采用另外一个标准误差( $SE$ )以评价模型的预测精度. $SE$ 可由下式计算得到:

$$SE = \sqrt{\frac{\sum_{i=1}^n [\log Y(\text{实测值})_i - \log Y(\text{拟合值})_i]^2}{n - A - 1}} \quad (1)$$

式中: $n$  为所建立的模型中训练集的个数.

## 2 结果与讨论

在 PLS 分析中, $VIP$  是一个表征模型中自变

量相对重要性大小的参数,如果某自变量的  $VIP$  值大于 1,表示这个自变量对因变量有较大的影响.以表 1 中所列的 15 种苯酚类化合物的  $\log Y$  为因变量,以 1.2 节所述的量子化学参数为自变量,采用 SIMCA(6.0 版)软件进行 PLS 分析,并在分析中逐步累计剔除  $VIP$  值最小的变量,最后得到最佳的 PLS 模型(1).最佳 PLS 模型的判定标准是,在交叉验证的条件下,所选取的主成分所解释的因变量的方差的比例( $Q^2_{cum}$ )及拟合值和实测值的相关系数( $R$ )达到最大,模型的显著性水平( $P$ )达到最小.由 PLS 分析得到 QSPR 模型(1).对于模型(1),化合物个数  $n=15, A=1, Q^2_{cum}=0.839, R=0.931, P<4.775\times 10^{-7}$ . 模型(1)的拟合结果为  $R^2_{X(adj)(cum)}=0.679, R^2_{Y(adj)(cum)}=0.856, Eig=3.809, Q^2_{cum}=0.839, SE=0.065$ ,其中  $R^2_{Y(adj)(cum)}$  和  $R^2_{X(adj)(cum)}$  分别表示所提取的所有主成分所能解释的矩阵  $X$  和矩阵  $Y$  的方差的比例, $Eig$  是表征 PLS 主成分重要性大小的本征值.可见,所提取的 1 个 PLS 主成分解释了自变量矩阵方差的 67.9%,解释了因变量矩阵方差的 85.6%.

对模型(1)中包括的 15 种化合物,该模型给出的拟合值与实测值之间的相关系数  $R$  为 0.931,相应的显著性水平  $P<0.000001$ ,说明模型(1)代表的 QSPR 是统计学上相当显著的.从表 1 可以看出,该模型给出的拟合值与实测值之间的差别较小.该模型经过交叉验证的  $Q^2_{cum}$  值远远大于 0.50,因此具有很好的预测可靠性.应用该模型,对任意选取的其他 3 个苯酚类化合物的  $\log Y$  值进行了预测(表 1).

模型(1)中各个自变量的  $VIP$  值列于表 2 中,表 2 同时列出了经过对 PLS 分析结果进行转化所得到的各个自变量的准回归系数和常数项.根据这些未经标准化的准回归系数,得出类似于多元回归分析结果的多元方程,根据准回归系数前的正负号,可以判定各个自变量对因变量影响的性质.根据 PLS 分析得到的可用于预测苯酚类物质光解量子产率( $Y$ )的多变量方程如方程(2)所示:

$$\log Y = 3.906 - 5.653 \times 10^{-3} HOF + 4.569 \times 10^{-5} CCR + 0.3479 E_{\text{homo}} + 0.1462 (E_{\text{lumo}} + E_{\text{homo}}) - 8.134 \times 10^{-3} (E_{\text{lumo}} - E_{\text{homo}})^2 \quad (2)$$

$$n = 15, Q^2_{cum} = 0.839, R = 0.931, P < 0.000001$$

表 2 模型(1)中各参数的  $VIP$  值和准回归系数  
Table 2 The  $VIP$ s and pseudo-regression coefficients in Model (1)

自变量	$VIP$	系数 $a$	系数 $b$
$E_{\text{homo}}$	1.051	0.2243	$3.479 \times 10^{-1}$
$HOF$	1.011	-0.2157	$-5.653 \times 10^{-3}$
$E_{\text{lumo}} + E_{\text{homo}}$	0.9806	0.2093	$1.462 \times 10^{-1}$
$CCR$	0.9802	0.2092	$4.569 \times 10^{-5}$
$(E_{\text{lumo}} - E_{\text{homo}})^2$	0.9753	-0.2081	$-8.134 \times 10^{-3}$
常数项		-4.174	3.906

注:  $a$  标准化系数;  $b$  未经标准化的系数

$E_{\text{homo}}$  和  $HOF$  的  $VIP$  值较大,均大于 1,说明这些分子结构参数是影响苯酚类化合物  $\log Y$  大小的主要因素.苯酚类化合物的  $\log Y$  随着分子  $E_{\text{homo}}$  的增大而增大,随着分子  $HOF$  的增大而减小.因此,所研究的苯酚类化合物的  $\log Y$  主要取决于  $E_{\text{homo}}$  和  $HOF$  的大小,  $E_{\text{homo}}$  较大的苯酚类分子的光解量子产量较大.这样的分子更容易光解.在太阳光照射下,对表层水中 PAH 光解的 QSPR 研究表明<sup>[7]</sup>, PAH 的光解半衰期随着  $E_{\text{homo}}$  的增大而减小,  $E_{\text{homo}}$  较大的分子容易光解,这和本实验的结论相一致.然而,气溶胶中 PAH 的直接光解量子产率随着  $E_{\text{homo}}$  的增大而减小<sup>[9]</sup>,这表明在不同的环境介质中和不同类型光源的照射下,有机污染物的光解机理有所不同<sup>[7]</sup>.  $E_{\text{lumo}} + E_{\text{homo}}$ 、 $CCR$  和  $(E_{\text{lumo}} - E_{\text{homo}})^2$  的  $VIP$  值接近于 1,说明这些参数对苯酚类化合物  $\log Y$  的大小也有一定的影响.由表 3 可见,  $CCR$  和  $HOF$  有显著的负相关性.因此,  $CCR$  较大和  $HOF$  较小的分子,其  $\log Y$  值也较大.苯酚类化合物的  $\log Y$  随  $E_{\text{lumo}} + E_{\text{homo}}$  的增大而增大,随  $(E_{\text{lumo}} - E_{\text{homo}})^2$  的增大而减小.由于  $E_{\text{lumo}} + E_{\text{homo}}$  和分子的电负性有关<sup>[10]</sup>.因此,电负性越小的分子,其光解量子产率越大,这样的分子越

容易光解,对 PAH 直接光解半衰期的 QSPR 研究也表明,电负性越小的 PAH 分子,其光解速率也越快<sup>[6]</sup>,此外,苯酚类化合物分子随 $(E_{\text{lumo}}-E_{\text{homo}})^2$  的增大而更加稳定,对其他物质的研究结果也得出了类似的结论<sup>[7,11]</sup>.作者研究结果表明,对苯酚类化合物光解量子产率的 QSPR 研究不仅可以对这类化合物的光解量子产率进行初步的预测,同时也有助于对这些物质在水-TiO<sub>2</sub> 平衡分散体系中的光解机理进行进一步的解释.

表 3 部分分子结构参数的相关系数

Table 3 Correlation coefficients between some quantum chemical descriptors

参数	<i>M<sub>w</sub></i>	<i>HOF</i>	<i>TE</i>	<i>EE</i>	<i>CCR</i>	<b>a</b>
<i>M<sub>w</sub></i>	1.000					
<i>HOF</i>	-0.406	1.000				
<i>TE</i>	-0.980	0.578	1.000			
<i>EE</i>	-0.598	0.968	0.745	1.000		
<i>CCR</i>	0.505	-0.986	-0.666	-0.994	1.000	
<b>a</b>	0.681	-0.942	-0.812	-0.990	0.971	1.000

注:  $P < 0.05$

### 3 结语

应用 PLS 算法得到部分苯酚类化合物的光解量子产率的 QSPR 模型,得到的多变量方程为:

$$\log Y = 3.906 - 5.653 \times 10^{-3} HOF + 4.569 \times 10^{-5} CCR + 0.3479 E_{\text{homo}} + 0.1462 (E_{\text{lumo}} + E_{\text{homo}}) - 8.134 \times 10^{-3} (E_{\text{lumo}} - E_{\text{homo}})^2$$

应用该模型,可以对其他苯酚类分子的  $\log Y$  进行初步的预测.模型的  $Q^2_{\text{cum}}$  值为 0.839,表明所得到的 PLS 模型具有较好的稳定性和预测能力.影响苯酚类化合物光解量子产率的主要因素是分子最高占据轨道能( $E_{\text{homo}}$ )和分子生成热( $HOF$ ).  $\log Y$  值随着  $E_{\text{homo}}$  的增大而增大,随着  $HOF$  的增大而减小.此外,分子的电负性和 $(E_{\text{lumo}}-E_{\text{homo}})^2$  对  $\log Y$  值也有一定的影响,电负性和 $(E_{\text{lumo}}-E_{\text{homo}})^2$  较小的分子,其  $\log Y$  值也较大.

- [1] Meharg A A, Wright J, Osborn D. Chlorobenzenes in rivers draining industrial catchments [J]. *Sci. Total Environ.*, 2000, 251/252:243-253.
- [2] U.S. Environmental Protection Agency. National pollutant discharge elimination system [R]. Washington, DC: U.S. Government Printing Office, 1988.
- [3] Hwang H-M, Hodson R E, Lee R F. Degradation of phenol and chlorophenols by sunlight and microbes in estuarine water [J]. *Environ. Sci. Technol.*, 1986,20:1002-1007.
- [4] Wong A S, Crosby D G. Photodecomposition of pentachloro-phenol in water [J]. *J. Agric. Food Chem.*, 1979,29:125-130.
- [5] Serpone N. Relative photonic efficiencies and quantum yields in heterogeneous photocatalysis [J]. *J. Photochem. Photobiol. A: Chem.*, 1997,104(1):1-12.
- [6] Chen J W, Quan X, Peijnenburg W J G M, *et al.* Quantitative structure-property relationships (QSPRs) on direct photolysis quantum yields of PCDDs [J]. *Chemosphere*, 2001,43:235-241.
- [7] Chen J W, Peijnenburg W J G M, Quan X, *et al.* Is it possible to develop a QSPR model for direct photolysis half-lives of PAHs under irradiation of sunlight [J]. *Environ. Pollut.*, 2001,114:137-143.
- [8] Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics [J]. *Chemom. Intell. Lab. Syst.*, 2001,58:109-130.
- [9] Chen J W, Quan X, Yan Y, *et al.* Quantitative structure-property relationship studies on direct photolysis of selected polycyclic aromatic hydrocarbons in atmospheric aerosol [J]. *Chemosphere*, 2001,42:263-270.
- [10] Pearson P G. Absolute electronegativity and hardness correlated with molecular orbital theory [J]. *Proc. Natl. Acad. Sci.*, 1986,83: 8440-8441.
- [11] Faucon J C, Bureau R, Faisant J, *et al.* Predicting of the fish acute toxicity from heterogeneous data coming from notification files [J]. *Chemosphere*, 1999,38:3261-3276.

作者简介:牛军峰(1974-),男,河南开封人,清华大学环境科学与工程系在站博士后,主要从事有机污染物的环境行为、生态毒理及定量结构-性质-活性关系、难降解有机物污染控制、污染治理等方面的研究工作.参加编写专著 1 部,发表论文 19 篇.

参考文献: