

文章编号: 1000-4556(2008)02-0228-06

吡啶酮生物碱的核磁共振碳谱模拟研究

何 留¹, 梅 虎^{1,2,3}, 李志良^{1,2,3*}(1. 重庆大学 化学化工学院, 重庆 400044; 2. 重庆大学 生物工程学院, 重庆 400044;
3. 湖南大学 化学生物传感与计量学国家重点实验室, 湖南 长沙 410082)

摘 要: 采用表征分子内部化学微环境及原子所处杂化状态的结构描述子: 原子电性作用矢量(AEIV)和原子杂化状态指数(AHSI), 对 42 个吡啶酮生物碱分子 792 个共振碳原子进行结构表征, 以多元线性回归技术建立¹³C 核磁共振化学位移定量结构波谱关系模型, 所得回归模型的复相关系数为 $R=0.957$, 标准偏差为 $SD=12.247$. 采用留一法交互检验的结果为 $R_{cv}=0.956$, 标准偏差 $SD_{cv}=12.331$. 对外部样本集预测结果表明, AEIV 和 AHSI 具有表征能力强、物化意义明确等优点, 所建模型具有良好的稳定性和估计能力.

关键词: ¹³C 核磁共振波谱模拟; 原子电性作用矢量; 原子杂化状态指数; 吡啶酮

中图分类号: O482.53; O626.32 **文献标识码:** A

吡啶酮生物碱是一类重要的生物活性物质, 在抗癌、抗肿瘤、抗病毒等方面有重要作用^[1-3]. 近年来, 随着吡啶酮生物碱生物活性和功能评价研究的不断深入, 对吡啶酮分子结构的研究显得非常必要. 核磁共振(NMR)在有机物结构鉴定中起着重要作用. 可在不破坏分子生理功能条件下高精度地解析有机物分子结构^[4], 因此广泛应用于有机化学、生物分析和化学化工等领域. 核磁共振波谱模拟始见于 Grant^[5] 和 Lindeman^[6] 对烷烃的研究; 近年来取得众多成果^[7-10]. 本研究组提出原子电性作用矢量(AEIV), 预测 20 种天然氨基酸¹³C 核磁共振化学位移^[11], 效果良好. 在此基础上, 本研究应用原子电性作用矢量, 结合原子杂化状态指数(AHSI)表征吡啶酮生物碱分子 792 个共振碳原子化学微环境和自身杂化状态, 应用多元线性回归(MLR)建立定量构谱关系(QSSR)模

收稿日期: 2007-05-21; 收修改稿日期: 2007-06-26

基金项目: 化学生物传感与计量学国家重点实验室基金(05-12-1)资助项目、重庆大学创新基金(04-10-10)资助项目.

作者简介: 何留(1982-), 女, 重庆人, 重庆大学分析化学专业在读硕士生, 主要从事波谱模拟、化学信息和计量学等方面的研究. 电话: 023-65110584, E-mail: heliu8212@163.com. * 通讯联系人: 李志良, 电话: 023-65106677, E-mail: zlli-cqu@163.com.

型, 同时采用留一法和外部预测对模型进行检验, 所得结果颇为满意.

1 原理及方法

1.1 基本概念

碳原子化学位移与其所处微化学环境及自身状态密切相关. 分子中不同化学性质原子对目标碳原子作用效果与两原子电负性及相隔距离的 6 次方成反比^[8]. 本文将原子按元素周期表主族分为 5 类, 得到不同主族原子对目标原子电性作用计算方法. 对于原子电性, 取该原子相对电负性即该原子与碳原子 Pauling 电负性的比值, 例如氧原子相对电负性为: $3.44/2.55=1.3490$. 对于原子相对键长, 这里仿照原子相对电性的处理方法, 取其相对于 C-C 键长大小, 例如 C-O 相对键长为: $0.143/0.154=0.9286$. 本研究中原子实际电负性和化学键长取自文献^[12,13]. 由此得到不同类型原子对目标原子的原子电性作用矢量(AEIV)计算公式为:

$$\nu_{ik} = \sum_{j \in k, j \neq i}^{all(j)} \frac{x_j}{d_{i,j}^6} \quad (1 \leq k \leq 5) \quad (1)$$

其中 ν_{ik} 为第 i 原子 AEIV 描述子, i 为目标原子, j 为第 k 类所有原子 ($j \neq i$); x 为相对电负性; $d_{i,j}$ 为 i, j 两原子相对键距, 即最短连接路径相对键长之和.

对于原子杂化状态, 本研究在原子固有状态^[14](I)的基础上进行一定修正, 得到原子杂化状态指数(AHSI)用于表征原子自身杂化类型, 其计算式为:

$$AHSI = \sqrt{\nu/4}((2n)^2\delta_{\sigma+\pi} + 1)/\delta_{\sigma} \quad (2)$$

式中 ν 为该原子价层电子数, n 为原子价层主量子数, $\delta_{\sigma+\pi}$ 是参与成 σ 和 π 键总电子数; δ_{σ} 为 σ 键电子数. 各类原子不同杂化类型对应 AHSI 值列于附表 3. 其中 C 原子 3 种不同杂化类型(SP 、 SP^2 、 SP^3 杂化)对应 AHSI 值分别为 2.5000、1.6667、1.2500.

1.2 计算实例

以化合物 I (示例结构见图 1) 为例, 对于 1 号碳原子, (VA) 族(氮原子)对其作用项为: $\nu_{14} = 1 \times 1.1922 / (0.9026 \times 2 + 0.8701)^6 = 0.0033$, 同理可分别计算出 H、C 原子对该 C 原子作用项为: $\nu_{11} = 6.5475$, $\nu_{12} = 3.7909$. 该化合物分子中无卤素原子, 故 ν_{15} 为零. 由于 1 号碳原子为 SP^2 杂化类型, 对应 AHSI 值为 1.6667. 其他碳原子 AEIV 描述子均按此方法计算得到.

1.3 QSSR 模型建立

从文献^[15]中取得 42 种吡啶酮生物碱分子共计 792 个碳原子 ^{13}C 核磁共振化学位移. 对每个分子中各非氢原子进行编号, 并输入非氢原子数、原子属性及连接方式, 利用 C 语言程序 AEIVMP.exe 计算出每一个碳原子对应 AEIV 描述子. 以多元线性回归 (MLR) 方法建立该 792 个等价共振碳原子化学位移与其 AEIV 及 AHSI 描述子的回归模型. 采用留一法 (LOO) 对模型进行交互检验 (CV), 评价模型内部预测能力.

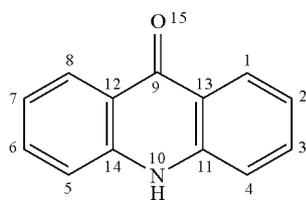


图 1 吡啶酮类生物碱分子结构
Fig. 1 Molecular structure of acridone alkaloids

2 结果与讨论

2.1 吡啶酮生物碱¹³C NMR 波谱模拟

利用 MLR 技术将计算得到的 792 个等价共振碳原子对应的 AEIV 描述子与¹³C NMR 化学位移初步建立起 4 参数回归模型 M1, 所得回归方程如下: (“±”后数字为回归系数取值范围为 95%置信度下的置信区间)

$$CS = (120.285 \pm 5.430) - (3.863 \pm 0.252) \cdot \nu_H + (1.594 \pm 1.180) \cdot \nu_C + (4.595 \pm 1.075) \cdot \nu_N + (7.892 \pm 0.723) \cdot \nu_O$$

采用 SPSS13.0 统计学软件检验各变量的 *t* 分布值、偏相关系数等相关统计结果见表 1. 结果表明, 仅用 AEIV 描述子表征碳原子所处化学环境就可较好地计算出相应化学位移值. 由于碳原子自身杂化状态对化学位移的影响不可忽略, 故引入原子杂化状态指数 AHSI, 结合 AEIV 描述子与共振碳原子化学位移建立起 5 参数回归模型 M2, 同时采用留一法交互检验 (LOOCV) 对模型进行验证, 所得回归方程及相关统计结果如下: (“±”后数字为回归系数取值范围为 95%置信度下的置信区间)

$$CS = -(200.381 \pm 7.157) - (0.897 \pm 0.141) \cdot \nu_H - (3.627 \pm 0.601) \cdot \nu_C + (6.952 \pm 0.541) \cdot \nu_N + (11.475 \pm 0.369) \cdot \nu_O + (201.413 \pm 4.158) \cdot AHSI$$

表 1 模型 M1、M2 中有关变量的部分统计量

Table 1 Partial statistics about variables of model 1 and model 2

模型	参数	a_0	ν_H	ν_C	ν_N	ν_O	AHSI	其他
M1	标准误差	5.430	0.252	1.180	1.075	0.723	—	$R=0.814$, $SD=24.437$
	偏相关系数	—	-0.479	0.048	0.151	0.366	—	$F=387.508$; $R_{cv}=0.812$
	<i>t</i> 值	22.150	-15.302	1.351	4.274	11.042	—	$SD_{cv}=24.541$, $F_{cv}=382.544$
M2	标准误差	7.157	0.141	0.601	0.541	0.369	4.158	$R=0.957$, $SD=12.247$
	偏相关系数	—	-0.222	-0.210	0.417	0.742	0.866	$F=1703.448$; $R_{cv}=0.956$
	<i>t</i> 值	-27.998	-6.381	-6.033	12.848	31.060	48.445	$SD_{cv}=12.331$, $F_{cv}=1678.573$

比较相关统计结果可知, M2 拟合结果优于 M1: R 和 R_{cv} 分别由 0.814 和 0.812 提高到 0.957 和 0.956, 整体变量显著性 F 亦有明显提高. 结果表明: (1) 采用 AEIV 和 AHSI 能较好地表征碳原子所处化学环境和杂化状态, 对碳原子化学位移具有较高拟合能力. (2) 由 M2 中各参数回归系数可知, AHSI 对碳原子化学位移值影响最为明显, 由实测结果可以看出: SP^3 杂化碳原子化学位移在 20~80 之间, 而 SP^2 杂化碳原子化学位移大多处于 120 以上. 这里将 M2 对 792 个共振碳原子化学位移留一法交互检验预测值和实验观测值相关绘于图 2 中.

从该图可以看到绝大多数样本密集分布在过原点与 x 轴成 45° 角的直线上, 表明 M2 对碳原子化学位移拟合较好, 对内部样本预测能力较高. 只有少数样本预测值误差较大, 进一步分析发现, 它们是处于环外支链上的少数 C 原子, 由于其化学位移值与其他多数碳原子化学位移值有较大差异, 可能超出线性范围从而存在一定系统误差, 故估计结果误差较大.

2.2 模型验证

研究表明^[16,17], 模型的预测能力只能通过外部预测样本来评价. 故从 42 个吡啶酮生物碱分子中随机选取 5 个样本分子共计 86 个共振碳原子作为测试集, 余下 37 个吡啶酮分子共 706 个碳原子作为训练集建立 5 参数回归模型. 所建模型的相关统计结果为: $R=0.959$ 、 $RMS=10.021$ 、 $F=1602.424$ 、 $R_{cv}=0.958$ 、 $RMS_{cv}=12.507$ 、 $F_{cv}=1583.574$, 对 86 个共振碳原子化学位移预测结果为: $R_{ext}=0.936$ 、 $RMS_{ext}=13.624$. 结果表明, 该模型较准确地模拟和预测了 86 个共振碳原子化学位移值, 模型对外部样本预测能力较高.

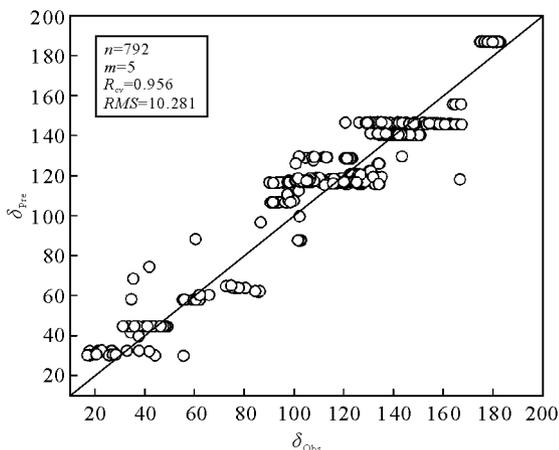


图 2 模型 M2 留一法预测值与观测值相关

Fig. 2 Plot of observed versus the predicted values of model 2 with leave-one-out cross-validations

3 结语

采用 AEIV 描述子表征原子所处化学微环境, 结合原子杂化状态指数 AHSI, 较准确地模拟和预测了 42 个吡啶酮生物碱分子¹³C 核磁共振化学位移值. 研究表明, AEIV 和 AHSI 可较好地反映有机物分子中目标原子所处化学微环境和原子存在状态, 具有物化意义明确、计算简便等优点. 在今后的工作中, 拟将 AEIV 和 AHSI 推广到包含大量杂原子的复杂有机物分子的核磁共振波谱模拟研究中, 以期探索该方法在 QSSR 研究中更为广阔的应用前景.

参考文献:

- [1] Chun M M, Olmstead K K, Choi Y S, *et al.* Synthesis and biological activities of truncated acridones: Structure-activity relationship studies of cytotoxic 5-hydroxy-4 quinolone [J]. *Bioorg Med Chem Lett*, 1997, 7(7): 789-792.
- [2] Itoigawa M, Ito C, Wu T S, *et al.* Cancer Chemopreventive activity of acridone alkaloids on Epstein-Barr virus activation and two-stage mouse skin carcinogenesis [J]. *Cancer Lett*, 2003, 193(2): 133-138.
- [3] Ahua K M, Ioset J R, Ransijn A, *et al.* Antileishmanial and antifungal acridone derivatives from the roots of *Thamnosma rhodesica* [J]. *Phytochemistry*, 2004, 65(7): 963-968.
- [4] Beger R D, Freeman J P, Lay Jr J O, *et al.* Use of ¹³C NMR spectrometric data to produce a predictive model of estrogen receptor binding activity[J]. *J Chem Inf Comput Sci*, 2001, 41(1): 219-224.
- [5] Grant D M, Paul E G. Carbon-13 magnetic resonance II. Chemical shift data for the alkanes [J]. *J A Chem Soc*, 1964, 86: 2 984-2 990.
- [6] Lindeman L P, Adams J Q. Carbon-13 nuclear magnetic shifts for the resonance spectrometry. Chemical shifts for the paraffins through C9 [J]. *Anal Chem*, 1971, 43: 1 245-1 252.
- [7] Meiler J, Meusinger R, Will M. Fast determination of ¹³C NMR chemical shifts using artificial neural networks

- [J]. *J Chem Inf Comput Sci*, 2000, 40(1): 1 692–1 176.
- [8] Liu S S, Liu H, Yu B M, *et al.* Investigation on quantitative relationship between chemical shift of carbon-13 nuclear magnetic resonance spectra and molecular topological structure based on a novel Atomic Distance-Edge Vector[J]. *J Chemom*, 2001, 15(5): 427–438.
- [9] Zhang Meng-jun(张梦军), Liao Chun-yang(廖春阳), Zhou Li-ping(周丽平), *et al.* On preliminary approach to estimation and prediction of chemical shift of ^{13}C NMR spectroscopy in flavonoids on vector of atomic electronegative distance (VAED) (黄酮类化合物的原子电距矢量表达及核磁共振碳谱模拟) [J]. *Chinese J Magn Reson (波谱学杂志)*, 2002, 19(3): 293–300.
- [10] Liang Gui-zhao(梁桂兆), Mei Hu(梅虎), Zhou Yuan(周原), *et al.* Simulation of ^{13}C nuclear magnetic resonance spectra for derivatives of bases and nucleotides(碱基和核苷及其衍生物的 ^{13}C 核磁共振波谱模拟)[J]. *Chinese J of Anal Chem(分析化学)*, 2006, 34(3): 329–332.
- [11] Mei Hu(梅虎), Zhou Peng(周鹏), Qin Ren-hui(覃仁辉), *et al.* Calculation of ^{13}C chemical shifts of amino acids using atomic electronegativity interaction vector (原子电性作用矢量用于氨基酸化学位移计算) [J]. *Chinese J Magn Reson (波谱学杂志)*, 2005, 22(2): 163–172.
- [12] Aylward G H, Findlay T J. *SI Chemical Data (2nd ed) (SI 化学数据手册)* [M]. Beijing (北京): High Education Press (高等教育出版社), 1985, 94.
- [13] Zhang Han-qi(张寒琦), Chen Zheng(陈铮), Lin Ying-jie(林英杰), *et al.* *Handbook of Chemical Data (实用化学手册)* [M]. Beijing(北京): Science Press (科学出版社), 2001. 682–723.
- [14] Hall L H, Kier L B. Electrotological state indices for atom types; a novel combination of electronic, topological, and valence state information [J]. *J Chem Inf Comput Sci*, 1995, 35: 1 039–1 045.
- [15] Yu D Q(于德泉), Yang J S(杨峻山), Xie J X(谢晶曦). *Handbook of Analytical Chemistry: Nuclear Magnetic Resonance(分析化学手册: 核磁共振波谱分析)* [M]. Beijing(北京): Chemical Industry Press(化学工业出版社), 1999. 554–555.
- [16] Golbraikh A, Tropsha A. Beware of $q(2)$! [J]. *J Mol Graphics Mod*, 2002, 20(4): 269–276.
- [17] Tropsha A, Gramatica P, Gombar V K. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models [J]. *QSAR Comb Sci*, 2003, 22(1): 69–77.

Estimation and Prediction of ^{13}C Chemical Shifts of Acridone Alkaloids: A Quantitative Structure-Spectrum Relationship Study

HE Liu¹, MEI Hu^{1, 2, 3}, LI Zhi-liang^{1, 2, 3*}

(1. College of Chemistry and Chemical Engineering, Chongqing University, Chongqing 400044, China;

2. College of Bioengineering, Chongqing University, Chongqing 400044, China;

3. State Key Laboratory for Chemobiosensors and Chemobiometrics, Hunan University, Changsha 410082, China)

Abstract: A multiple linear regression model for calculating ^{13}C chemical shifts in acridone alkaloids from the structural indices namely the atomic electronegative interaction (AEIV) vector and the atomic hybridization state index (AHSI) was established. ^{13}C chemical shifts of 792 equivalent carbon atoms in 42 acridone alkaloids were calculated based on their AEIV and AHSI indices. The correlation coefficients and standard deviation of modeling estimation and leave-one-out cross-validation are 0.957, 0.956 and 12.247, 12.331, respectively. An external prediction set was applied to validate the prediction capabilities of the model. The results show that both AEIV and AHSI are excellent topological indices with satisfactory estimation stability and favorable predictive ability.

Key words: ^{13}C nuclear magnetic resonance, atomic electronegativity interaction vector, atomic hybridization state index, acridone alkaloid

* Corresponding author; Li Zhi-liang, Tel: 023-65106677, E-mail: zlli-cqu@163.com.