

文章编号: 1673- 0062(2009) 04- 0084- 04

量子拓扑方法对硫醇的定量构效关系研究

陈 炫, 聂长明*, 蒋司同, 张成林, 杨 挺, 杨孝智

(南华大学 化学化工学院, 湖南 衡阳 421001)

摘 要: 应用密度泛函理论 (DFT), 在 B3LYP/6- 31+ G (d) 水平上对 20 硫醇的分子结构进行几何优化和量化计算. 将量子化学计算的部分结果引入分子的拓扑结构中, 构建新型分子量子拓扑指数 (QT), 与计算所得的量化参数 Gibbs 自由能 (G)、恒容摩尔热容 (CV) 一起描述硫醇的量子拓扑规律. QT 指数不仅能有效地区分硫醇的分子结构, 而且能较好的描述硫醇理化性质及色谱保留指数的变化规律. 大部分所建模型的相关系数都大于 0.99. 本文对构建的模型采用留一法 (LOO) 进行验证, 结果表明所构建的模型稳健、合理、有效, 该方法有望在物质的 QSPR/QSAR 研究中得到广泛的应用.

关键词: 量子拓扑指数; 硫醇; 密度泛函理论 (DFT); 量子力学; 定量构效关系

中图分类号: O641- 3 **文献标识码:** A

QSPR/QSAR Study on Mercaptans by Quantum Topological Method

CHEN Xuan, NIE Chang-ming*, JIANG Si-tong
ZHANG Cheng-lin, YANG Ting, YANG Xiao-zhi

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

Abstract 20 molecules of mercaptans were structural optimized and calculated by using density functional theory (DFT) at the B3LYP/6- 31+ G (d) level of theory. In order to depict the quantum topological change rules of mercaptans, a new molecular quantum topological index QT was constructed by the method which is to introduce some calculated results of quantum chemistry into the structure of molecules together with Gibbs free energy (G), Constant volume mole hot melting (CV) calculated by density functional theory (DFT). Index QT can not only efficiently distinguish molecular structures of mercaptans, but also possess good applications of QSPR/QSAR (quantitative structure-property/activity relationships). And most of the correlation coefficients of the models were over 0.99. The LOO CV (leave-one-out cross-validation) method was used to testify the stability and predictive ability of the models. The validation results verified the good stability, reson-

收稿日期: 2009- 10- 22

基金项目: 湖南省优秀博士学位论文基金资助项目 (2008YBX01)

作者简介: 陈 炫 (1984-), 男, 江苏宿迁人, 南华大学化学化工学院硕士研究生. 主要研究方向: 理论及计算化学. * 通讯作者.

able and predictive ability of the models And this method can be widely used in QSPR / QSAR study

Key words quantum topological index; mercaptan; density functional theory (DFT); quantum mechanics; QSPR / QSAR

随着计算化学的发展,近年来定量结构-性质(活性)相关(QSPR/QSAR)的研究越来越受到人们的关注,现已广泛应用于有机化学、药物化学以及环境化学等领域^[1-2].自1947年Wiener指数 $W^{[3]}$ 提出以来,越来越多的拓扑指数已被提出并应用于物质的QSPR/QSAR研究^[4].有学者曾采用拓扑指数对硫醇的色谱保留指数进行定量构效关系研究^[5-6],如周从芝等^[5]采用拓扑指数 N_t 对硫醇等化合物的色谱保留指数进行构效关系研究并取得了较好的效果.此外,应用密度泛函理论(DFT)对物质的构效关系的研究也有不少报道^[2,7],如林英武等^[8]采用MM2及MM+方法对醛酮的沸点性质进行相关性研究.

本文在以前工作^[1,9-10]的基础上,将分子拓扑方法与量子化学方法结合起来,将量子化学计算的部分结果引入分子的拓扑结构中,构建新的分子量子拓扑指数(QT),由此研究硫醇的理化性质及色谱保留指数的构效关系,获得了较好的结果.

1 基本原理与方法

1.1 分子量子拓扑指数的构建

对于由 n 个原子构成的分子隐氢图,距离矩阵可表示为 $D = (d_{ij})_{n \times n}$,其中 n 为分子中的非氢原子个数, d_{ij} 为分子结构中原子 i 和 j 的最短拓扑距离.本文将 d_{ij} 进行改进,既不同于传统矩阵中的定义距离^[3],也不同于作者以前采用的键长的相对距离 $d_{ij} = \sum L_j / L_{C-C}^{[1]}$,而是采用密度泛函理论(DFT),在 6-31+G(d) 水平上对分子结构进行几何优化和量化计算的原子间距离.

电负性是描述原子性质的重要参数之一,主要是表示原子在形成分子过程中得失电子的能力,而平衡电负性可以表征化合物中同种原子在不同化学环境中的电负性不同的特点^[11],为此我们采用分子中顶点原子的平衡电负性对分子图进行着色,构建新矩阵 D :

$$D = \begin{bmatrix} X_1 & d_{12} & \cdots & d_{1(n-1)} & d_{1n} \\ d_{21} & X_2 & \cdots & d_{2(n-1)} & d_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ d_{(n-1)1} & d_{(n-1)2} & \cdots & X_{n-1} & d_{(n-1)n} \\ d_{n1} & d_{n2} & \cdots & d_{n(n-1)} & X_n \end{bmatrix}$$

将新的分子量子拓扑指数定义为 QT :

$$QT = \log V \cdot \log \lambda_{max} \tag{1}$$

式(1)中 N 为分子中非氢原子的个数, λ_{max} 为改进的距离矩阵 D 的最大特征值.

1.2 分子结构的优化和量化计算

运用密度泛函理论(DFT),采用 Gaussian 03 程序,在 6-31+G(d) 水平上对脂肪醛酮的分子结构进行几何优化.并对分子的 Gibbs 自由能(G)、恒容摩尔热容(CV)、熵(S)、配分函数($\lg Q$)、最高占据轨道(HOMO)能量(ϵ_H)、最低空轨道(LUMO)能量(ϵ_L)、羰基碳氧原子的电荷(Q_C, Q_O)、偶极矩(μ)、原子间的距离(d)等进行理论计算.例如 2-丁硫醇的分子优化结构示意图(图1)及其改进的距离矩阵 D .

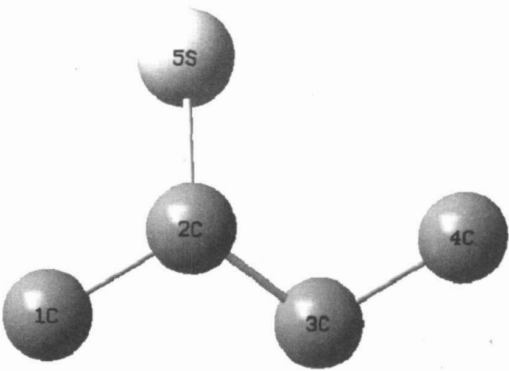


图 1 2-丁硫醇的分子优化结构示意图
Fig 1 Optimized molecular structure for 2-Butyl mercaptan

$$D = \begin{bmatrix} X_{C1} & d_{C1C2} & d_{C1C3} & d_{C1C4} & d_{C1S5} \\ d_{C2C1} & X_{C2} & d_{C2C3} & d_{C2C4} & d_{C2S5} \\ d_{C3C1} & d_{C3C2} & X_{C3} & d_{C3C4} & d_{C3S5} \\ d_{C4C1} & d_{C4C2} & d_{C4C3} & X_{C4} & d_{C4S5} \\ d_{S5C1} & d_{S5C2} & d_{S5C3} & d_{S5C4} & X_{S5} \end{bmatrix}$$
$$= \begin{bmatrix} 2.3025 & 1.5296 & 2.5672 & 3.1559 & 2.8074 \\ 1.5296 & 2.3474 & 1.5387 & 2.5722 & 1.8608 \\ 2.5672 & 1.5387 & 2.3189 & 1.5340 & 2.7329 \\ 3.1559 & 2.5722 & 1.5340 & 2.2953 & 4.1625 \\ 2.8074 & 1.8608 & 2.7329 & 4.1625 & 2.3722 \end{bmatrix}$$

由式(1)得 2-丁硫醇 QT 指数为:

$QT = \log 5 \cdot \log 12\ 3403 = 0\ 7628$

2 结果与讨论

2.1 模型的建立

为得到对硫醇性质影响显著的参数,我们对所计算的量化参数采用逐步回归分析,所得多元线性回归模型如下:

$P = aG + bCV + cQT + d$ (2)

式 (2) 中, a 、 b 、 c 、 d 是回归常数, P 代表硫醇

的性质, G 为硫醇的 Gibbs 自由能、 CV 为硫醇的恒容摩尔热容, QT 为式 (1) 计算的量子拓扑指数, G 和 CV 都是量化计算所得. 构建模型的稳定性采用相关系数 (R), 费舍尔检验值 (F), 标准误差 (S) 进行评价.

2.2 硫醇理化性质的构效关系

根据式 (2) 对硫醇的沸点温度 (T_B), 临界温度 (T_c), 临界压力 (P_c) 和临界体积 (VC)^[12] 进行多元线性回归, 回归结果如表 1

表 1 硫醇理化性质的回归结果

Table 1 Results of regression of physichemical properties for mercaptans

Property(P)	$P = aG + bCV + cQT + d$				R	S	F	n
	a	b	c	d				
TB	- 1. 0914	- 10. 9116	199. 0130	- 125. 7355	0. 9983	3. 6145	2343. 6619	20
TC	- 0. 4499	- 9. 7011	303. 0946	316. 4784	0. 9955	5. 5371	882. 7239	20
PC	- 0. 6696	- 3. 2974	- 82. 2590	- 186. 4664	0. 9822	1. 6308	220. 1324	20
VC	- 1. 6457	- 1. 1753	- 38. 1167	- 550. 4094	0. 9998	2. 6958	18119. 64	20

只要原子在分子中所处的环境有差异, 其在分子中的平衡电负性就不同, 原子间的空间距离也不一样, 由式 (1) 得到的指数 QT 也不会相同, 也即 QT 对物质选择性高, 具有唯一性, 不简并. QT 不仅体现了分子的大小, 同时蕴含分子的成键信息 (如键长)、分子中原子电性信息 (如平衡电负性)、分子间相互作用信息 (如分子体积和分子支化) 等. 从表 1 中我们可看到采用该模型只有一个 R 值低于 0. 99 其他 R 值都超过 0. 995 都达

到了优级标准. 四种物理性质的计算值与实验值的平均绝对误差分别为 0. 21%、0. 09%、3. 85%、0. 17%, 由此可见硫醇四种物理性质的实验值与计算值吻合良好.

2.3 硫醇色谱保留指数构效关系

根据式 (2) 对硫醇在四个不同极性固定相 (Apiezon M, OV - 17, TritonX 305, PEG - 1000)^[13] 的色谱保留指数进行多元线性回归, 回归结果如表 2

表 2 硫醇色谱保留指数回归结果

Table 2 Results of regression of chromatographic retention index for mercaptans

Property(P)	$P = aG + bCV + cQT + d$				R	S	F	n
	a	b	c	d				
Apiezon M	- 10. 0071	- 61. 7175	35. 0017	- 3362. 2254	0. 9995	7. 3635	10451. 2023	20
OV - 17	- 11. 1214	- 70. 6854	32. 0175	- 3680. 7754	0. 9992	9. 1372	6901. 9814	20
TritonX 05	- 12. 3516	- 80. 1812	20. 7424	- 4025. 7885	0. 9990	10. 3870	5496. 3629	20
PEG - 1000	- 13. 6147	- 91. 0157	34. 4582	- 4418. 5877	0. 9988	11. 4553	4589. 2228	20

色谱保留指数是由色谱保留值计算而来, 由于色谱的保留值是由色谱分离过程中的热力学因素控制的, 化合物的色谱保留值主要与其在固定相和流动相中的分配系数有关, 而分配系数主要与系统的热力学性质有关, 如分子的吉布斯函数、配分函数、分子大小等. 在非极性固定相中对同一类有机物而言, 色散力对色谱保留指数的影响较大, 色散力越大色谱保留指数越大, 而色散力与分

子大小具有方向一致性. 从表 2 的结果中可看到模型的相关性非常好, 都大于 0. 99 这表明指数 QT 与量化参数构建的模型非常有效.

2.4 模型验证

为考察模型的稳定性和预测能力, 本文中采用留一检验法 (LOO) 对所建模型进行检验, 采用 LOO 方法计算的相关系数 (R_{cv}), 费舍尔检验值 (F_{cv}), 标准误差 (S_{cv}) 对模型的优良性及稳定性

进行评价. 结果如表 3

表 3 LOO 法检验结果
Table 3 The results of LOO method

Property (P)	Equation of regression	R_{CV}	S_{CV}	F_{CV}	n
T_B	$P_{exp.} = 1.0076 P_{pre} - 2.6879$	0.9950	5.7077	2810.3789	20
T_C	$P_{exp.} = 1.0140 P_{pre} - 7.6749$	0.9859	9.0686	997.7425	20
P_C	$P_{exp.} = 1.0124 P_{pre} - 0.6746$	0.9404	2.7581	221.0652	20
V_C	$P_{exp.} = 0.9966 P_{pre} + 1.2939$	0.9996	3.3678	34822.3549	20
Apiezon M	$P_{exp.} = 1.0013 P_{pre} - 0.7102$	0.9988	10.4708	15495.8259	20
OV-17	$P_{exp.} = 1.0004 P_{pre} + 0.0993$	0.9983	12.9814	10248.2653	20
TritonX-305	$P_{exp.} = 1.0007 P_{pre} - 0.2418$	0.9979	14.4440	8517.3918	20
PEG-1000	$P_{exp.} = 0.9999 P_{pre} + 0.7206$	0.9974	16.1763	6894.2659	20

从表 3 中可以看出留一检验法的多重相关系数 R_{CV} 与回归表 1 和表 2 的 R 值接近. 硫醇模型大多数相关系数波动范围狭窄表示各线性关系良好, 由此可见采用量子拓扑指数 QT 与量化参数 G 、 CV 构建的模型稳健、合理、有效.

3 结论

本文将分子拓扑方法与量子化学方法结合起来, 将量子化学计算结果引入分子的拓扑结构中, 构建新的分子量子拓扑指数 QT . 分子量子拓扑指数 QT 不仅蕴含丰富的分子结构信息、电负性信息、空间信息, 同时将量子化学方法与拓扑学方法结合起来, 使得拓扑方法中某些经验数值进一步上升到理论数值, 从而推动拓扑化学的进步. 本文用分子量子拓扑指数 QT 与量化参数 G 、 CV 一起研究硫醇的性质变化规律, 所建模型的大部分相关系数都大于 0.99 且模型经留一法进行验证均获得了满意的结果, 该方法有望在物质的 QSPR / QSAR 研究中得到广泛的应用.

参考文献:

[1] Cong Yi Zhou, Chang Ming Nie, Shan Li et al. A novel semi-empirical topological descriptor N_t and the application to study on QSPR/QSAR [J]. Journal of Computational Chemistry, 2007, 28(15): 2413–2423.
[2] Pablo R Duchowicz, Martin G Vitek, Eduardo A Castro et al. QSAR modeling of the interaction of avonoids with GABA (A) receptor [J]. European Journal of Medicinal Chemistry, 2008, 43(8): 1593–1602.
[3] Wiener H. Structural determination of paraffin boiling

points [J]. Journal of the American Chemical Society, 1947, 69(1): 17–20.
[4] 许禄, 邵学广. 化学计量学方法 [M]. 北京: 科学出版社, 2004.
[5] 周从艺, 聂长明, 田万福, 等. 碳氢化合物、醛、酮和硫醇的定量结构-色谱保留指数相关性研究 [J]. 分析测试学报, 2007, 26(6): 802–807.
[6] 刘利, 曹晨忠, 谢斌, 等. 硫醚、硫醇色谱保留指数的 QSRR 研究 [J]. 湖南科技大学学报 (自然科学版), 2005, 20(4): 74–80.
[7] Malakhat A Turabekova, Bakhtiyor F Rasulev, Mikhail G Levkovich et al. Aconitum and Delphinium alkaloids as antagonist modulators of voltage-gated Na^+ channels: AM1/DFT electronic structure investigations and QSAR studies [J]. Computational Biology and Chemistry, 2008, 32(2): 88–101.
[8] 林英武, 张秀利, 王中华, 等. 脂肪族醛酮的沸点与其分子体积及甲基数的相关性研究 [J]. 计算机与应用化学, 2007, 24(3): 395–397.
[9] Chang Ming Nie, Ya Xin Wu, Rong Yan Wu et al. Applications of a new topological index ED_m in some aliphatic hydrocarbons [J]. Journal of Theoretical and Computational Chemistry, 2009, 8(1): 19–45.
[10] 彭国文, 聂长明, 肖芳竹, 等. 拓扑指数与路径数预测单硫醚气相色谱保留指数 [J]. 南华大学学报 (自然科学版), 2006, 20(1): 1–4.
[11] 聂长明. 基团电负性 [J]. 武汉大学学报 (自然科学版), 2000, 46(2): 176–180.
[12] Carl L Yaws. Chemical Properties Handbook [M]. Beijing: McGraw-Hill Book Co, 1999.
[13] 李浩春. 分析化学手册: 第五分册 气相色谱分析 [M]. 北京: 化学工业出版社, 2003.