

# 拓扑指数法在QSPR研究中的应用

张晶, 赫敏, 杨林\*

(青海师范大学 化学系, 青海 西宁 810008)

**摘要:** 目前拓扑指数方法在定量构效关系(QSPR)研究中属活跃领域。简述了Wiener指数、Randic-Kier指数、Hosoya指数等拓扑指数, 介绍了拓扑指数在定量构效关系研究、物化性质研究、药物设计研究、环境污染预测等方面的应用。表明拓扑指数法在QSPR研究中取得显著效果, 具有良好的应用前景和实用价值。

**关键词:** 拓扑指数法; QSPR; 应用

**中图分类号:** O65 **文献标志码:** A **文章编号:** 1673-0143(2012)06-0038-04

拓扑指数法是定量构效关系(QSPR)研究领域一种常用的方法, 在分析化学研究中具有广阔的应用前景。结构决定性质, 性质反映结构, 结构和性质的关系在化学研究中一直占据着重要地位。QSPR是依据化合物的性能与分子结构密切相关的原理, 寻求分子结构和物质性质之间的内在定量关系。拓扑指数是分子结构的数学描述符, 通过描述分子的拓扑指数揭示分子的基本拓扑性。目前已有200余种指数, 其中著名的有Wiener指数、Hosoya指数、Randic-Kier指数、Balaban指数等, 这些指数在物质的QSPR研究领域发挥了重要的作用。本文综述了拓扑指数法在QSPR研究中的应用, 主要介绍拓扑指数法在QSPR研究、物化性质研究、药物设计研究、环境污染预测等方面的应用研究<sup>[1-5]</sup>。

## 1 拓扑指数法

### 1.1 Wiener指数

Wiener拓扑指数也称Wiener指数, 1947年由美国化学家Herald Wiener在研究烷烃(即石蜡)分子沸点的过程中, 对烷烃分子中键空间相互作用与物理性质的关系进行研究的基础上提出的, 是一种重要的分子拓扑指标, 也是最古老的一个分子拓扑指数。Wiener指数是用来描述有机化合物的物理化学性质与它们的分子图的拓扑结构之间关系的一种不变量。Wiener指数

自提出后就在化学和数学等方面得到了广泛的应用和研究<sup>[6]</sup>。

### 1.2 Randic-Kier指数

分子的连接性指数目前被广泛应用, 它能定量表示一个分子结构特征, 用希腊字母 $\chi$ 表示。分子连接性指数首先由Randic所建议, 他最初提出的分子连接性指数是一阶连接性指数, 后来Kier等发展了分子连接性指数概念, 使其计算边长由 $l$ 扩展到 $h$ , 因此, 除了一阶连接性指数, 还会有零阶和二阶连接性指数, 可以反映分子的整体属性, 而高阶路径项指数对于描述分子的局部特征十分有效<sup>[7]</sup>。

### 1.3 Hosoya指数

Hosoya指数概念是由Hosoya于1971年在日本杂志《Bull Chem Soc》中提出的, 这个指数与分子图的特征多项式紧密相关, 是研究物质分子结构与物理和化学性质之间关系的拓扑参数。Hosoya指数是用来研究分子图的独立集总数, 它们与分子的总 $\pi$ -电子能、沸点等物理化学性质有密切的关系。近年来, 许多数学家和化学家对Hosoya指数进行了研究和刻划, 从而达到对某类分子的一些化学性质的研究<sup>[8]</sup>。

### 1.4 广义 $a_N$ 指数

在量子力学的基础上, 杨家安和江元生1983年提出了 $a_N$ 指数, 是可用于区分分子支化度的一个指数并且与饱和链烃的多种物理化学性质相

收稿日期: 2012-03-22

基金项目: 青海省科技厅项目(2009-Z-719)

作者简介: 张晶(1989—), 女, 硕士生, 研究方向: 分析化学。

\*通信作者: 杨林(1962—), 男, 教授, 研究方向: 环境分析化学。E-mail: yanglin@qhnu.edu.cn

关。在此基础上,许禄、王华云等进行了扩充,称之为广义 $a_N$ 指数。在广义 $a_N$ 指数的计算案中考虑了两种d型轨道作用方式,即通过d型轨道作用的权重 $d_{ij}$ 和 $d'_{ij}$ 的不同提供了区分顺、反异构体的依据,该指数不仅可以区分空间顺反异构体,而且可以用于含杂原子含多重键的复杂化合物复杂体系<sup>[9]</sup>。

### 1.5 平均距离和连通性指数 $J$

分子图可用距离矩阵 $D$ 表示,把距离矩阵 $D$ 的一行或一列的元素相加所得结果叫做相应于该顶点的距离度。1983年,Balaban按照Randic连通性指数 $X$ 的类似形式定义了新的拓扑指数 $J$ ,它是建立在分子图加权后的距离矩阵基础上构造的,对芳香键与单键给予区分,比较全面地反映了分子的大小和形状等结构信息<sup>[10-11]</sup>。

## 2 应用方面

### 2.1 QSPR研究

Raouf Ghavami等<sup>[12]</sup>建立新的拓扑指数: $Sh$ 指数,并应用主成分回归(PCR)、主成分—人工神经网络(PC-ANN)等研究了有机化合物的正常沸点和摩尔折光率与拓扑指数间的相关性,建立了QSPR模型,预测了一系列的有机化合物如烃类、脂类、醇类、苯类等的正常沸点和摩尔折光率;Porto L C等<sup>[13]</sup>建立半经验的拓扑指数:1-ET指数,用线性回归方法获得122个苯系物的色谱保留值与拓扑指数间良好的QSPR模型,具有较好的预测能力;Hu等<sup>[14]</sup>建立 $Lu$ 指数、DAI指数和分子电负性指数 $ep$ ,研究几种结构类型不同的有机农药在4种不同色谱柱上的色谱保留指数和 $Lu$ 指数,DAI指数,分子电负性指数 $ep$ 的模型,很好地预测了不同类型农药的气相色谱保留指数;刘凤萍等<sup>[15]</sup>提出修正的分子极化效应指数( $MPEI_m$ )、烷基的极化效应指数( $PEI$ )、立体效应指数( $SV_{ij}$ )等拓扑结构参数,运用回归方法获得了饱和酯类化合物在7种不同极性色谱柱上的QSPR模型,得出的模型方程用于预测各饱和酯类化合物的气相色谱保留指数,其稳定性和准确性俱佳,且较好地揭示了饱和酯类化合物在不同极性色谱柱上气相色谱保留指数的变化规律;林英武等<sup>[16]</sup>利用分子力学MM2和MM+的方法将脂肪族醛酮分子作构象优化,获得其分子体积,同时将分子内的甲基数作为分子结构参数,建立分子体积和分子内的甲基数与脂肪族醛酮沸点间的

QSPR模型,表明脂肪族醛酮的沸点与其分子体积及甲基数具有很好的相关性,在预测脂肪族醛酮的沸点方面将有可观的应用前景。

### 2.2 物化性质研究

Liu F P等<sup>[17]</sup>研究新的拓扑指数——极化效应指数( $PEI$ ),奇偶指数(OEI)和空间效应指数( $SV_{ij}$ )。研究饱和酯在7个固定相上的气相色谱保留指数,通过多元线性回归方法获得拓扑指数与气相色谱保留指数间的关系模型,可知饱和酯的色谱保留指数与极化效应、分子大小和分支有重要关系;陈桂珍等<sup>[18-19]</sup>根据分子图的邻接矩阵和距离矩阵提出新的拓扑指数 $X$ ,该指数与饱和烷烃的沸点、摩尔折光率、 $\Delta H_a$ 等均有良好的相关性,可知 $X$ 指数不仅具有良好的结构选择性,而且与饱和烷烃的多种物理化学性质均有良好的相关性;杨林等<sup>[20-22]</sup>利用分子中各个原子的三维坐标数据,定义出新的3D拓扑指数( $Y_j$ 、 $Y_h$ 和 $Y_g$ )进行多元线性回归分析,得到39种饱和烷烃的部分物理化学性质与自定义的3D拓扑指数的QSPR模型,具有良好的结构与性质的相关性;刘凤萍等<sup>[23]</sup>应用键连接矩阵特征根( $SX_{ICH}$ 、 $SX_{ICC}$ )、立体效应指数( $SV_{ij}$ )等拓扑结构参数,用多元线性回归方法获得醛酮化合物的沸点及摩尔折光率与拓扑指数间QSPR模型,并定量解释了各参数对其性质的影响;刘先军等<sup>[24]</sup>运用定义的分子拓扑指数 $WN$ 和位置拓扑指数 $SG$ ,采用回归分析方法建立了硫醇的气态标准生成焓、气态标准生成自由能、沸点、色谱保留值等与拓扑指数 $WN$ 和 $SG$ 的QSPR模型,所建立的模型方程具有良好的稳定性和预测能力,较好地揭示了这些物理化学性质的变化规律。

### 2.3 药物设计研究

Norinder等<sup>[25]</sup>使用表示理化性质的描述符如脂性、极性、氢键、计算的分子表述符及拓扑指数描述符,并用偏最小二乘法对Caco-2细胞通透性进行预测,所得模型有较好的预测功能;李婷婷等<sup>[26]</sup>收集了967个结构不同的化合物,并用4种软件计算得到4种描述符集,包括表征分子大小、形状、表面积等的2D描述符和表征构象能量、带电性状等的3D描述符,用于计算化合物的分子性状,且运用主成分分析方法筛选描述符,建立分子描述符与对CYP3A4酶半数抑制浓度的大小( $IC_{50}$ )之间的模型,能较准确地预测药物分子的 $IC_{50}$ ;肖爱靖<sup>[27]</sup>利用量子化学参数结合分子

拓扑指数以及疏水性参数等研究42个3-芳基喹啉酮类选择性雌激素受体调节剂的QSPR,并建立其激活性的线性和非线性模型,找到了对活性有重要影响的结构因素,这对于结构修饰和发现具有活性的药物具有重要的指导意义;堵锡华等<sup>[28]</sup>利用量子化学AM1算法获得了拓扑系列指数,建立了能预测血管紧张素转换酶抑制剂摩尔折光率的QSPR模型,研究了血管紧张素转换酶抑制剂(ACEI)的分子结构与摩尔折光率的关系。

## 2.4 环境污染预测

包锦渊<sup>[29]</sup>借助元素的有效主量子数改造了原子点价,定义了新的表征卤代苯结构的分子连接性指数,研究了17种卤代苯化合物对生物的毒性,通过最小二乘法建立了相应方程,预测结果令人满意,为该类化学品的环境危险性评价提供了依据;李鸣健等<sup>[30]</sup>基于表征生物降解性的原子点价( $g_i$ )构建新的自相关拓扑指数( ${}^mG$ ),其中的 ${}^1G$ 不仅对取代芳烃呈现良好的结构选择性,而且与10种取代苯甲酸一级生物降解速率常数显著相关,可据此研究取代苯甲酸类物质对于环境的影响;冯长君等<sup>[31]</sup>依据价连接性指数 $X_j$ 、电性拓扑态指数( $e_j$ )及电性距离矢量( $m_k$ ),构建239种有机污染物生物富集因子( $F_{BC}$ )的6参数模型,不仅相关程度高,而且所用自变量数少,由模型可知有机污染物BCF的主要影响因素是 $-C-$ 、 $>C-$ 、 $-O-$ 、 $-S-$ 、 $-X$ 等分子结构碎片以及分子的柔韧性、折叠程度等空间因素,新建的连接性指数 ${}^1F$ 以及电性距离矢量与有机物的生物富集因子具有良好的相关性,可在QSPR研究中获得广泛的应用;刘先军等<sup>[32]</sup>根据直链烷基芳基磺酸盐分子结构的特点,用距离矩阵表征分子中原子的连接性,据此建立相应的定量结构性能关系式,计算了直链烷基芳基磺酸盐的等效烷烃碳数和生物降解性,计算结果表明具有较好的相关性;冯长君<sup>[33]</sup>采用拓扑指数法计算多酚类分子的结构参数,运用最佳量子集回归法初步建立它们4种抗氧化活性的QSAR模型,所建模型具有良好的稳健性,分子内 $=O$ 、 $=C<$ 、 $-CH_3$ 、 $-COO-$ 等基团对多酚类化合物的抗氧化活性具有很大影响。

## 3 展望

随着各种不同的拓扑指数的相继发现,对于分子结构的描述也更加准确,根据不同的拓扑指数,应用多元线性回归,人工神经网络等建立

QSPR模型,对于物化性质研究、药物设计研究、环境污染预测等方面将起到积极作用。综上所述,拓扑指数法在QSPR研究中具有广泛的应用前景。

## 参考文献:

- [1] Liu F P, Liang Y Z, Cao C Z. Theoretical prediction of the Kovat's retention index for oxygen-containing organic compounds using novel topological indices[J]. Analytica Chimica Acta, 2007, 594(2):279-289.
- [2] Panneerselvam K, Antony M P, Srinivasan T G, et al. Estimation of normal boiling points of trialkyl phosphates using retention indices by gas chromatography[J]. Thermochimica Acta, 2010, 511(1/2):107-111.
- [3] Gonzalez D H, Romaris F, Duardo S A, et al. Predicting drugs and proteins in parasite infections with topological indices of complex networks: theoretical backgrounds, applications, and legal issues[J]. Current Pharmaceutical Design, 2010, 16(24):2737-2764.
- [4] Jyoti S, Basheerulla S, Vijay K A, et al. Modeling of mutagenicity of aromatic and heteroaromatic amines in Salmonella typhimurium TA98: Role of hydrophobicity and topological indices[J]. Journal of the Indian Chemical Society, 2008, 85(5):517-535.
- [5] Vijay K A, Vinay K D, Basheerulla S, et al. Modeling of lipophilicity of some organic compounds using structural and topological indices[J]. Journal of the Indian Chemical Society, 2009, 86(4):337-345.
- [6] Wiener H. Structural determination of paraffin boiling points[J]. J Am Chem Soc, 1947, 69(1):17. 20.
- [7] Randic M. On characterization of molecular branching[J]. J Am Chem Soc, 1975, 97(23):6609-6615.
- [8] Hosoya H. Topological index a proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons[J]. Bull Chem Soc Jpn, 1971, 44(9):2332-2339.
- [9] 许禄,邵学广. 化学计量学方法[M]. 北京:科学出版社, 2004:389.
- [10] Balaban A T. Applications of graph theory in chemistry[J]. J Chem Inf Comput Sci, 1985, 25(3):334-343.
- [11] Balaban A T. Chemical graphs: looking back and glimpsing ahead[J]. J Chem Inf Comput Sci, 1995, 35(3):339-350.
- [12] Raouf G, Amir N, Bahram H. QSPR studies on normal boiling points and molar refractivities of organic compounds by correlation - ranking - based PCR and PC - ANN analyses of new topological indices[J]. Canadian Journal of Chemistry, 2009, 87(11):1593-1604.
- [13] Porto L C, Souza E S, Junkes B D, et al. Semi - empirical topological index: Development of QSPR/QSRR

- and optimization for alkylbenzenes[J]. *Talanta*, 2008, 76(2):407-412.
- [14] Hu R Z, Yin C S, Wang Y, et al. QSPR study on GC relative retention time of organic pesticides on different chromatographic columns[J]. *Journal of Separation Science*, 2008, 31(13):2434-2443.
- [15] 刘凤萍,曹晨忠,吴湘江,等.用分子拓扑指数预测饱和酯类化合物的气相色谱保留指数[J].*计算机及应用化学*, 2008, 25(6): 711-716.
- [16] 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.
- [17] Liu F P, Liang Y Z, Cao C Z, et al. QSPR study of GC retention indices for saturated esters on seven stationary phases based on novel topological indices[J]. *Talanta*, 2007, 72(4):1307-1315.
- [18] 陈桂珍,张树棠,黄正国.一种新拓扑指数X用于烯烃的QSAR研究[J].*唐山师范学院学报*, 2009, 31(2): 11-15.
- [19] 朱显杰,张树棠,黄正国.拓扑指数X用于脂肪醇的QSPR研究[J].*计算机及应用化学*, 2009, 26(9): 1197-1202.
- [20] 杨娜,杨林.多氯联苯分子空间坐标与气相色谱相对保留时间的QSPR研究[J].*计算机及应用化学*, 2012, 29(2):219-222.
- [21] 侯恩卿,谢占川,李秀庆,等.拓扑指数在烃类化合物定量构效关系中的研究进展[J].*甘肃联合大学学报:自然科学版*, 2011, 25(6): 54-56.
- [22] 程学峰,张巍巍,杨林,等.饱和烷烃的3D拓扑指数与其性质相关性研究[J].*青海师范大学学报:自然科学版*, 2010, 33(1): 35-38.
- [23] 刘凤萍,曹晨忠,程彬.脂肪族醛酮的定量构效关系研究[J].*湖南科技大学学报:自然科学版*, 2010, 25(2): 102-105.
- [24] 刘先军,王宝辉,崔宝臣,等.直链硫醇拓扑结构与性质之间的关系[J].*石油炼制与化工*, 2012, 43(1): 32-35.
- [25] Norinder U, Osterberg T. Theoretical calculation and prediction of drug transport processes using simple parameters and partial least squares projections to latent structures(PLS) statistics: The use of electrotopological state indices[J]. *J Pharm Sci*, 2001, 90(8): 1076-1085.
- [26] 李婷婷,季晖,陈西敬,等.药物设计和研发中的定量构效关系研究进展[J].*药学进展*, 2009, 33(3): 98-103.
- [27] 肖爱婧.3-芳基喹啉酮类和四氢异喹啉类选择性ER调节剂的QSAR及对接研究[D].北京:首都师范大学, 2008: 1-10.
- [28] 堵锡华,冯长君.ACEI的分子结构与摩尔折光率的定量关系[J].*湖南师范大学自然科学学报*, 2007, 30(3): 80-83.
- [29] 包锦渊.一种新的连接性拓扑指数与卤代苯毒性的定量关系[J].*邢台学院学报*, 2006, 21(4): 112-113.
- [30] 李鸣建,冯长君.取代苯甲酸的生物降解性与拓扑相关性研究[J].*化工科技*, 2006, 14(1): 24-26.
- [31] 冯长君,沐来龙,杨伟华,等.有机污染物的生物富集因子与拓扑指数的数学模型[J].*物理化学学报*, 2008, 24(6): 1053-1057.
- [32] 刘先军,崔宝臣.拓扑指数法研究直链烷基芳基磺酸盐的构效关系[J].*日用化学工业*, 2009, 39(4): 238-240.
- [33] 冯长君.多酚类化合物抗氧化活性的电拓扑模型[J].*营养学报*, 2010, 32(6): 524-527.

## Application of Topological Index Method in QSPR

ZHANG Jing, HE Min, YANG Lin

(Department of Chemistry, Qinghai Normal University, Xining 810008, Qinghai, China)

**Abstract:** Topological index method in the study of QSPR is a very active field at present. Briefly describes the Wiener index, Randic-Kier index, Hosoya index such topological index. The applications of topological index in QSPR is researched, such as quantitative structure-activity relationship, physical and chemical properties, drug design research, environmental pollution prediction are summarized. The topological index method has made significant effect in QSPR research. It has a good application prospect and practical value.

**Key words:** topological index method; QSPR; application

(责任编辑:叶冰)