

抗肿瘤喹唑啉衍生物的定量构效关系^{*}

吴文娟^{1,2}, 司徒传艺², 郑康成¹

(1. 中山大学化学与化学工程学院, 广东 广州510275;

2. 广东药学院物理化学教研室, 广东 广州510224)

摘要:用量子化学密度泛函(DFT)及分子力学(MM+)方法, 计算了一系列抗肿瘤喹唑啉衍生物分子的电子结构、几何结构及分子性质(广义结构参数), 并通过回归分析, 筛选出主要因素, 建立定量构效关系(QSAR)方程。发现该类化合物分子的抗肿瘤活性与一些广义结构参数有密切关系, 特别是苯环18位碳上取代基(R)的体积与表面积的比值 r_{18} , 嘧啶环与苯环(C环)间的距离 d , 以及次低未占据分子轨道(NLUMO)的能量 ϵ_{NL} 与抗肿瘤活性显著相关, 可通过取代基的选择与计算来筛选药效最佳的先导物。QSAR方程的拟合相关系数(R^2)及交叉验证相关系数(q^2)分别为0.8516和0.7146, 表明所得模型具有可信的预报能力, 可用以精确预测该系列化合物活性, 并为设计抗肿瘤活性更高的喹唑啉衍生物及作用机理分析提供理论指导。

关键词:喹唑啉; 抗肿瘤活性; 量子化学; 密度泛函法; 定量构效关系

中图分类号: O642.42 **文献标识码:** A **文章编号:** 0529-6579(2005)02-0049-04

近年来, 酶抑制剂的定量构效关系(QSAR)在药物设计^[1]中显示出的重要作用引起人们的广泛关注。尤其是二氢叶酸还原酶(DHFR)抑制剂, 如抗菌剂甲氧苄啶(TMP)、抗疟剂乙胺嘧啶和抗肿瘤的甲氧喋呤(MTX)等的广泛应用大大促进了寻找高度选择性的DHFR抑制剂的研究工作。抗肿瘤喹唑啉类化合物是最近研究得较多的DHFR抑制剂, 由于其亲脂性比MTX强, 可改善肿瘤细胞的耐药性, 因此, 人们对其进行了大量构效关系及结构修饰研究^[2,3], 但相关性欠佳, 有待进一步改进。为此作者应用量子化学密度泛函(DFT)和分子力学(MM+)方法, 将计算得到的电子结构、几何结构与分子性质参数结合起来(称广义结构参数), 探讨2,4-二氨基-6-甲基-7-取代苄氨基喹唑啉类药物分子的广义结构参数与抗肿瘤活性(抑制L1210细胞生长)的关系, 建立定量构效关系方程, 并作预报检验, 得到良好结果, 可为预示该类化合物的药效及分子设计提供理论指导。

1 计算方法与结果

1.1 量子化学与分子力学计算

喹唑啉类衍生物的分子结构示意图如图1所示。每个化合物分子涉及的原子数为46~58个。由于这类化合物分子均无任何对称性, 为节省计算

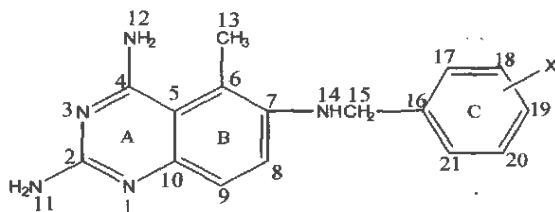


图1 喹唑啉衍生物的分子结构示意图

Fig. 1 Schematic diagram of molecular structures of quinazoline derivatives

机时, 先用ChemDraw软件构建分子, 再用Chem3D软件对各化合物分子结构进行初步几何优化, 最后运用G98量子化学程序包^[4], 用密度泛函(DFT)的B3LYP方法和6-31G基组^[5], 按闭壳层单重态进行全几何优化计算及自然轨道布居分析(NPA)。同时, 进行频率分析, 由所得结果不出现虚频, 确定其稳定构型, 因而可得到化合物分子稳定构型的几何参数(如有关的原子间距离)和电子结构参数(如前沿分子轨道能量、组成及原子的净电荷分布等)。此外, 还应用Hyperchem 7.0软件中的QSAR模块^[6]计算了分子的范德华体积 V , 表面积 A 和油水分配系数 $\log P$ 等分子性质参数。全部计算在Pentium IV微机上进行。

* 收稿日期: 2004-07-02

基金项目: 广东省自然科学基金资助项目(021721)

作者简介: 吴文娟(1966年生), 女, 副教授; 通讯联系人: 郑康成; E-mail: ceszkc@zsu.edu.cn

1.2 相关分析计算

从量子化学的 DFT 计算结果中, 提取了关于喹唑啉类衍生物分子的电子结构的微观参数: 最高占据分子轨道 (HOMO) 能量 (ϵ_H), 次高占据分子轨道 (NHOMO) 能量 (ϵ_{NH}) 最低未占分子轨道 (LUMO) 能量 (ϵ_L), 次低未占分子轨道 (NLUMO) 能量 (ϵ_{NL}), LUMO 和 HOMO 的轨道能量之差 ($\Delta\epsilon_{L-H}$), 分子偶极矩 (DIPO) 以及根据可能的活性部位挑选出的苯环上各碳原子净电荷 (Q_C), 胺基上氮原子的净电荷 (Q_N), 苯环上各取代基的非氢原子净电荷之和 (Q_R), 苯环 (图 1 所示的 B、C 环) 的总电荷 (Q^{Bben} 、 Q^{Cben}), 嘧啶环和苯环相连的 $C_7 \sim C_{16}$ 间的距离 (d) 等。同时, 由分子力学方法计算得到喹唑啉类似物的 V 、 A 和 $\log P$ 及苯环上各取代基的 V 和 A 。此外, 为了充分考虑可能的影响因素, 还选取了查表^[7]得到的苯环上各取代基的疏水参数 (J_I)、摩尔折射率 (R_M) 和诱导参数值等; 共 30 多个广义结构参数。

化合物的抗肿瘤活性以对鼠白血病细胞株 (L_{1210}) 的半数抑制浓度 (IC_{50}) 表示^[8], 选定了同一批实验的 16 个化合物进行了计算, 在定量构效关系中以其负对数 PIC_{50} 表示活性或药效。表 1 列出了 16 种该类化合物的分子式、药效 (PIC_{50}) 及一些主要的广义结构参数 (或指数)。

为选取影响化合物抗肿瘤活性的独立变量, 以上述的广义结构参数为处理变量, 以 PIC_{50} 为因变量进行相关分析以选取回归参数。

1.3 回归分析及交叉验证

在上述相关分析的基础上进行逐步回归分析, 建立该类化合物抗肿瘤活性与广义结构参数的定量构效关系 (QSAR) 方程。用 “leave one out” 法^[9] 检验其预报效果, 有关结果见表 1。同时, 用文献 [10] 提供的公式, 计算交叉检验的相关系数。

2 讨论

通过分析研究, 获得喹唑啉类化合物抗肿瘤活性的定量构效关系方程及其交叉验证结果:

$$PIC_{50} = 1.249 - 5.431r_{18} + 0.017d^5 - 0.568\epsilon_{NL} \quad (1)$$
$$n = 16, R^2 = 0.8516, S = 0.2184,$$
$$F = 25.8, q^2 = 0.7146$$

其中, n 为方程所含化合物的个数, R^2 为拟合的相关系数, S 为标准偏差, F 为 F 检验值, q^2 为交叉验证的相关系数, 它表征模型的预测能力。一般认为, 交叉验证系数值 q^2 大于 0.5 时, 模型具有可信的预报能力。同时, 对式中 4 个广义结构参数还进行了彼此间的相关分析, 发现这些参数间相对独立性较佳。

从回归方程 (1) 可以看出, 经逐步回归所选出的与抑制 L_{1210} 白血病的活性相关的参数包括苯环 (C 环) 上 18 位取代基的体积与表面积的比值 r_{18} , 嘧啶环和苯环 (C 环) 相连的 $C_7 \sim C_{16}$ 间的距离 d 以及次低空未占据的分子轨道 (NLUMO) 的能量 ϵ_{NL} 。

在回归方程中, 抑制活性与苯环 (C 环) 18 位取代基的体积与表面积的比值 r_{18} 呈负相关, 对活

表 1 喹唑啉衍生物的抗肿瘤活性 (PIC_{50}) 与广义结构指数

Tah 1 Antitumor activity (PIC_{50}) and generalized structural indexes of quinazoline derivatives

化合物	分子式	r_{18}	d	ϵ_{NL}	PIC_{50}	回归误差 ¹⁾	预报误差 ²⁾
1	17-Br-20, 21-(OCH ₃) ₂	1.158	3.761	-0.4593	7.84	-0.06	-0.06
2	17-OC ₆ H ₁₃	1.158	3.735	-0.0952	6.89	0.28	0.54
3	17-OC ₃ H ₇	1.158	3.756	0.0106	7.05	0.35	0.46
4	17-OCH ₂ OCH ₃	1.158	3.755	-0.2678	7.94	-0.38	-0.40
5	17-OH	1.158	3.753	-0.2163	7.50	0.02	0.03
6	18, 20-(OCH ₃) ₂	1.124	3.743	-0.0800	7.75	-0.30	-0.43
7	18, 20-Br ₂ , 19-OCH ₂ COOC ₂ H ₅	1.126	3.772	-0.8229	8.25	0.09	0.17
8	18-OC ₆ H ₁₃	1.416	3.745	-0.1059	5.95	-0.03	-0.11
9	18-OCH ₂ OCH ₃	1.269	3.763	-0.1508	6.88	0.14	0.23
10	18 OCH ₃ , 19 OCH ₃	1.124	3.762	-0.1004	7.92	-0.15	-0.15
11	18 OCH ₃ , 19-OCH ₂ COOC ₂ H ₅	1.124	3.761	-0.3878	7.80	0.13	0.15
12	18 OCH ₃ , 19-OCH ₂ OCH ₃	1.124	3.764	-0.2710	8.10	-0.20	-0.21
13	18-OCOCH ₃	1.243	3.762	-0.6866	7.76	-0.28	-0.40
14	19-OC ₆ H ₁₃	1.158	3.762	-0.1205	7.55	0.05	0.09
15	19-OCOCH ₃	1.158	3.763	-0.7214	7.65	0.32	0.45
16	H	1.158	3.763	-0.2256	7.86	-0.16	-0.18

1) 对 16 个化合物的回归误差, 为 $PIC_{50, calc} - PIC_{50, exp}$; 2) 用 15 个化合物回归所得参数对被抽出化合物的预报误差
(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

性影响较大, 即 18 位取代基的表面积 A_{18} 增大和 V_{18} 减少, 化合物所显示的抗肿瘤活性将增大。

这可能与该类化合物的作用机理直接相关。Matthews 等^[11]对 MTX 和 DHFR 分子络合物进行 X 射线衍射研究, 发现 MTX 的嘧啶环和对氨基苯甲酰基分别处于 DHFR 两个疏水性袋形区内, 通过分子间相互作用与酶结合。18 位取代基可能是进入疏水袋中与酶作用的重要部位, 这与文献[3]认为 18 位取代基是重要的活性基团的观点一致。根据匙锁模型可知, 18 位取代基很可能与酶的某一部分镶嵌(或 $\pi-\pi$ 堆积)在一起, A_{18} 越大和 V_{18} 越小, 即平面性越好, 药物与酶的作用越强, 这是因为平面面积越大, 能提供与酶的作用点越多^[12]。相反, 体积 V_{18} 越大和平面面积 A_{18} 越小, 则只能增加空间位阻, 妨碍药物与受体活性点的接触。

此外, 回归方程中还引入了 d^5 , 其系数为正, 即嘧啶环与苯环(C 环)间的距离 d 在一定的范围内增加, 活性明显增大。这可能也与该类化合物的作用机理直接相关, 在一定范围内 d 的增加, 更有利于唑啉类化合物与酶两者活性点的最佳接触, 从而提高其与酶的亲和作用。如活性较大的几个化合物(化合物 7, 10, 12), 其 d 值都较大。说明 d 值的微小变化, 对活性影响的灵敏性, 故在回归方程中用 d^5 而不直接用 d 作参数。

化合物的次低未占据分子轨道(NLUMO)的能量(ϵ_{NL})对抑制活性的影响也较大, 其系数为负, 表明 ϵ_{NL} 值越小, 即能量越低, 活性越大。根据分子轨道理论, 前线轨道(HOMO 和 LUMO)及其附近的分子轨道对活性影响最大。HOMO 及其附近的占据轨道(HOMO- x)等具有优先提供电子的作用, LUMO 及其附近的未占据轨道(LUMO+ x)等具有优先接受电子的作用。但为什么入选于该方程的是 ϵ_{NL} 而不是 ϵ_L 呢? 为了解释这一点需要分析本系列化合物的前线轨道分布, 计算结果表明, 本系列所有化合物的 HOMO 和 LUMO 轨道的分布大体相同, 主要分布在嘧啶环上。本系列唑啉类化合物是由 MTX 修饰而来, 与 MTX 具有相同的嘧啶环和苯环, 所以可认为嘧啶环和 N_{14} 与酶的氢键作用是本系列化合物具有一定抗肿瘤活性的基本因素, 可视为它们的共性(其对活性的影响可归结到 QSAR 方程中的常数项)。由于嘧啶环远离苯环(C 环)及其取代基, 其取代基的改变对其影响不大, 活性的差别主要反映在苯环(C 环)及其取代基上。对于 NLUMO 轨道的分布与 LUMO 轨道不同, 特别是活性较高的化合物(4, 7, 10, 12 号), 其

NLUMO 轨道成分主要分布在苯环的 C_{17} , C_{18} , C_{19} , C_{20} (或 C_{21}) 上, 而活性较低的 8, 9 号化合物, 其 NLUMO 主要成分在 A, B 环上的 C_2 , N_3 , C_5 , C_8 (或 C_9) 原子上。这说明 NLUMO 轨道在 C 环及其取代基上的分布, 更有利于药物分子接受电子而发生亲电反应, 从而抑制二氢叶酸还原酶的活性, 使其不能将二氢叶酸还原为四氢叶酸而阻碍叶酸代谢, 干扰肿瘤细胞的 DNA 和蛋白质的合成, 因而具有抗肿瘤作用。所以, 本系列化合物可在维持基本结构不变情况下进行结构改进, 达到提高抗肿瘤活性目的。

崔智勇^[3]的构效关系方程中所选用的参数如取代基的 J_I 和 R_M 及诱导参数等没有被引入回归方程(1)中, 考虑到参数间可能存在较大的相关性。通过参数间的相关分析发现确实如此, 入选参数与相应的 J_I 、 R_M 和诱导参数间存在着很大的相关性。如 r_{18} 与 J_{I18} 、 R_{M4} , d 与 J_{I17} 、 R_{M4} 以及 ϵ_{NL} 与诱导参数、 R_{M5} 间的相关系数分别为 0.564 5, 0.517 3; 0.477 1, 0.401 4; 0.454 7, 0.385 7。表明它们之间的相关性较大, 所以没有被选进方程中。为作比较性试验, 利用文献[3]中的参数对同一批的 16 个化合物进行多元线性回归分析, 得到 QSAR 方程:

$$PIC_{50} = 7.214 - 0.4249J_{I18} + 0.2067R_{M4} + 0.8301R_{M5} \quad (2)$$

其相关系数 $R=0.768 1$, 标准偏差 S 为 0.362, 比本文方程(1)用相同数量的参数得到方程的相关性差得多[方程(1)的 $R=0.922 8$, S 为 0.218], 说明本文所选的参数比文献[3]所选的参数更能反应本系列化合物的构效关系规律。

由表 1 数据可见, 化合物的抗肿瘤活性并非随某一参数呈规律性变化, 这是因为分子间的作用比较复杂, 不可能用单一参数来全面反映它们之间的相互关系。所以, 建立构效关系方程要综合考虑各方面的影响因素, 选取最佳参数。同时, 回归方程入选的参数全部都是可计算的微观结构或分子性质参数, 不依赖化合物是否已合成出来。即可根据回归方程系数的符号, 通过取代基选择来调控微观参数值, 找出最佳参数值所对应的化合物结构, 从而设计和预见活性更高的化合物。

3 结 论

(1) 本文建立了标题化合物抗肿瘤活性与广义结构参数的定量构效关系(QSAR)方程。该方程表明, 唑啉类化合物苯环(C 环)上 18 位取代基的体积与表面积比值 r_{18} , 嘧啶环和苯环(C

环) 相连的 $C_7 \sim C_{16}$ 间的距离 d , 以及次低未占据分子轨道 (NLUMO) 的能量 ϵ_{NL} 这 3 个结构参数是影响该系列化合物抗肿瘤活性的重要因素, 特别是 18 位碳上取代基 (R) 的 r_{18} 对活性的影响非常重要。可利用 QSAR 方程, 通过取代基选择与计算来筛选药效最佳的先导物。

(2) 作为定量构效关系模型预报效果的检验, 交叉验证法比传统的挑选预报法更加全面合理。本文 QSAR 方程的交叉验证相关系数 q^2 较大, 表明所得模型具有可信的预报能力。

参考文献:

- [1] DONKOR I O, LI H, QUEENER S F. Synthesis and DHFR inhibitory activity of a series of 6-substituted-2, 4-diaminothieno[2, 3-d] pyrimidines[J]. European J Med Chem, 2003, 38(6): 605–611.
- [2] TOBE M, ISOBE Y, TOMIZAWA H, et al. Structure-activity relationships of 6-fluoroquinazolines: Dual-acting compounds with inhibitory activities toward both TNF-production and T cell proliferation[J]. Bioorg Med Chem, 2003, 11: 609–616.
- [3] 崔智勇, 赵维璋, 李仁利. 喹唑啉类化合物抗肿瘤作用的定量构效关系研究[J]. 北京医科大学学报, 1999, 32(1): 84–85.
- [4] FRISCH M J, TRUCKS G W, SCHLEGEL H B, et al. Gaussian 98, Revision A. 11. 4[M]. Pittsburgh: Gaussian Inc, 2002.

- [5] FORESMAN J B, FRISCH A E. Exploring Chemistry with Electronic Structure Methods[M]. 2nd ed. Pittsburgh: Gaussian Inc 1996.
- [6] VISWANADHAN V N, GHOSE A K, REVANKAR G N, et al. Atomic physicochemical parameters for three dimensional structure directed quantitative structure activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition for certain naturally occurring nucleoside antibiotics[J]. J Chem Inf Comput Sci, 1989, 29: 163–172.
- [7] HANSCH C, LEO A J. Substituent constants for correlation analysis in chemistry and biology[M]. New York: Wiley-Interscience, 1979: 48–63.
- [8] 崔智勇, 赵维璋, 李仁利. 喹唑啉类化合物的合成及抗肿瘤活性[J]. 北京医科大学学报, 1999, 31(1): 27–31.
- [9] CENTNER V, MASSART D L, DE NORRDO E, et al. Elimination of uninformative variables for multivariate calibration[J]. Anal Chem, 1996, 68: 3851–3858.
- [10] HAWKINS D M, BASAK S C, MILLS D. Assessing model fit by cross-validation[J]. J Chem Inf Comput Sci, 2003, 43: 579–586.
- [11] MATTHEWS D A. Dihydrofolate reductase: X-ray structure of the binary complex with methotrexate[J]. Science, 1977, 20(197): 452–455.
- [12] 计亮年, 张黔玲, 刘劲刚. 生物医学中 DNA 的结构、构象、作用机制及其生物功能的研究进展[J]. 中国科学 (B 辑), 2001, 31(3): 193–204.

Quantitative Structure activity Relationship of Quinazoline Derivatives with Antitumor Activity

WU Wenjuan^{1,2}, SITU Chuan yi², ZHENG Kang cheng¹

(1.School of Chemistry and Chemical Engineering, Sun Yat Sen University, Guangzhou 510275, China;

2.Department of Physical Chemistry, Guangdong Pharmaceutical College, Guangzhou 510224, China)

Abstract: The computations of the electronic structures and some molecular properties of a series of quinazoline derivatives have been carried out by using the density functional theory (DFT) method of quantum chemistry and molecular mechanism method (MM+). Via a stepwise regression analysis, some main factors affecting the activity of the compounds were selected out, and the quantitative structure activity relationship (QSAR) equation was established. It has been found that the antitumor activity of the quinazoline derivatives is predominantly related to some structural parameters. In particular, the ratio of the volume with respect to the area of the substituent at site 18 on benzene ring, the distance between the C_7 and C_{16} , as well as the energies of the next lowest unoccupied molecular orbitals (NLUMO) of the complexes are closely correlated with the antitumor activity. Since all factors in the QSAR equation are some computed electronic and geometric structural parameters, it will be very advantageous to selection of some precursors with excellent antitumor activity. The fitting relation coefficient R^2 and the cross validation relation coefficient q^2 for the model established by this study are 0.851 6 and 0.714 6, respectively. The results suggest that this model has good predictability. The above results can offer some important theoretical guides in a design for novel quinazoline derivatives having higher antitumor activity and in an analysis of the action mechanism.

Key words: quinazoline derivative; antitumor activity; quantum chemistry; DFT; the quantitative structure activity relationship (QSAR)