A NOVEL MOLECULAR CONNECTIVITY METHOD

LI Xin-hua(李新华), ZHU Long-guan*(朱龙观), YU Qing-sen(俞庆森)

(*Dept. of Chemistry*, *Zhejiang University*, *Hangzhou*, *310027*, *China*)

Received Mar.28, 1998;revision accepted July.5,1999

Abstract: This paper presents a new molecular connectivity wherein the atomic delta value is weighted by the Mulliken bond order, and thus connects effectively the molecular orbital theory with the topological theory and converts the molecular connectivity method into a novel quantum-topology method. Comparison of the atomic delta value of the molecular connectivity with the value of the Mulliken bond order, showed that they have similar physical meanings. The calculated new molecular connectivity indexes of some typical hydrocarbons showed good correlation with their volumes, hydrophobic constants and thermodynamic properties, respectively.

Key words: Mulliken bond order, hydrophobic constant, molecular connectivity method, quantum-topology method

Document code: A   **CLC number**: O641

INTRODUCTION

The molecule-topology index and quantum chemistry index, two important parameters for study on the structure-activity relationship, can be obtained at different levels of molecular structure expression, respectively. The molecule-topology index is obtained from the molecular ichnographic-topological structure, and is relatively objective and simple (Wang et al., 1992). The molecule-topology index is calculated based only on the framework of a molecule, although some non-bond forces are important in QSAR or QSPR. The quantum chemistry index calculated from the molecular configuration in quantum level, can reveal in principle all the electronic information of the molecule: electron distribution probability, energy, polarity, electronic charge density, etc. The use of semi-empirical quantum chemistry method to calculate molecules has some limitations, so a quantum-topology theory proposed by Bader et al in 1977 (Xin et al., 1991) , may unify the molecular orbital theory and topological theory. Xu Lu et al. presented some quantum-topology indexes (Xu et al., 1985). There are various ways to modify the molecular connectivity index (Ni et al., 1998; Yang et al., 1998; Hall et al., 1990; Pogliani et al., 1992; Galvez et al., 1994).

METHODOLOGY

The molecular connectivity index is a widely-used topological index and has been successfully used in many fields (Kier et al., 1976; Kier et al., 1985), but the limitation of the molecular connectivity index is obvious. Its physical meaning is ambiguous and it cannot distinguish steric-isomers. The quantum chemistry method is inconvenient for extracting indexes from a molecule structure. The vital disadvantage of the quantum chemistry method is that it cannot explain the steric effect, a very important factor in QSAR. Comparison of the above two methods shows they are complementary. We combine the above two methods so as to offset their respective disadvantages.

Mulliken presented a method named Mulliken electron population analysis for determining the charge distribution of atoms. The method is helpful for explaining how the atoms connected by bonds in the molecule. According to the molecular orbital theory (Xu et al., 1985), the total population charge between atoms A and B in a molecule is $\sum_{i}^{B}\sum_{\lambda}^{A}\sum_{\mu} n_i c^*_{A\mu i} c_{B\lambda i} S_{A\mu, B\lambda}$, where $c^*_{A\mu i}$ is the coefficient of the μth atomic orbital of

atom A in the ith molecular orbital, $c_{B\lambda i}$ is the coefficient of the λth atomic orbital of atom B in the ith molecular orbital, $s_{A\mu, B\lambda}$ is the overlap integral between the uth atomic orbital of atom A and the λth atomic orbital of atom B of the ith molecular orbital, and n_i is the number of electrons in the ith molecular orbital, respectively. Mulliken pointed out that the charge population reflects the covalent bond order between atom A and atom B.
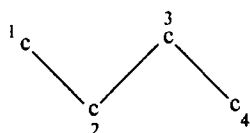

Fig.1   The figure of n-butane(ignoring hydrogen)

Table 1   The molecular connectivity matrix of n-butane

No.	1	2	3	4
1	0	1	0	0
2	1	0	1	0
3	0	1	0	1
4	0	0	1	0

Table 2   The Mulliken bond order matrix of n-butane (Calculated by 6-31G)**

No.	1	2	3	4
1	4.919081	0.344249	-0.054936	-0.054939
2	0.344249	4.957014	0.351815	0.004219
3	-0.054936	0.351815	4.957054	0.34425
4	0.004219	-0.054939	0.344251	4.919122

Comparison of the Mulliken bond order matrix(ignoring hydrogen) with the molecular connectivity matrix of a certain molecule, such as n-butane showed(Fig.1, Table1 and Table 2) that the off-diagonal value of the Mulliken bond order matrix corresponding to disconnected atoms is very small, even negative, which means that there is almost no bonding. For two disconnected carbon atoms the Mulliken bond order between them is approximately zero. In the molecular connectivity matrix the off-diagonal value between two disconnected carbons is absolutely zero. These two matrices look very similar. The value of the diagonal in the Mulliken bond order matrix is the charge density of the atom itself, and other values in the Mulliken matrix refer to the bond orders between two connected atoms. Thus, we find the Mulliken bond orders have good relationship with the delta values of the molecular connectivity matrix. The atomic delta values of molecular connectivity and that weighted by Mulliken bond orders of each carbon in ethane, propane, butane, iso-propane and neo-propane are shown in Table 3 and their correlation is expressed as Equation 1, where δ_i is the ith atomic delta value, δ_i^* is the ith atomic delta value weighted by Mulliken bond order, S is the standard deviation of errors, F is the statistics for correlation and R is the correlation coefficient.

$$\delta_i = 0.084289 + 2.692657\delta_i^* \qquad (1)$$

$$S = 0.05, \quad F = 4366.6, \quad R = 0.999$$

In molecules, atomic connectivity and environment relate closely to their electronic structures. For example, the Mulliken bond order be-

Table 3   The atomic delta values and new delta values weighted by Mulliken bond order of some hydrocarbons

		C1	C2	C3	C4	C5
Ethane	δ_i	1.00000	1.00000	—	—	—
	δ_i^*	0.33106	0.33106	—	—	—
Propane	δ_i	1.00000	2.00000	1.00000	—	—
	δ_i^*	0.34288	0.68576	0.34288	—	—
Butane	δ_i	1.00000	2.00000	2.00000	1.00000	—
	δ_i^*	0.34425	0.69607	0.69607	0.34425	—
Isopropane	δ_i	1.00000	3.00000	1.00000	1.00000	—
	δ_i^*	0.35510	1.0653	0.35510	0.35510	—
Neopropane	δ_i	1.00000	4.00000	1.00000	1.00000	1.00000
	δ_i^*	0.36923	1.476920	0.36923	0.36923	0.36923

tween two carbon atoms in conjugated, cyclic or unsaturated hydrocarbon molecules is different from that of the saturated catenarian hydrocarbon molecule, and their bonds lengths are different too. We think the Mulliken bond order can more reasonably reflect the molecular topological characteristic, than the delta value defined in the molecular connectivity method. So we adopt the Mulliken bond order instead of the delta values of connectivity for a molecule. This enables us to effectively combine the quantum theory and the topology theory, and introduce a kind of new quantum-topology index.

All information obtained from the original molecular connectivity method will be preserved in our new method because the weighted Mulliken values are used. For example, the nonvalence indexes and the Δ term indexes are all the same as the original molecular connectivity indexes, but our new valence indexes are different from the original valence indexes because we only modified the atomic delta value. As a kind of quantum-topology index, it is expected that our new method can make up for the shortcomings of the quantum chemistry and molecular connectivity methods.

CALCULATION METHOD

All the molecules were calculated and optimized by 6-31G** under the Hartree-Fock level (considering sufficiently the effect of the flexible bond on the original construction of a molecule for the calculation). The Mulliken bond order and molecular volume were calculated by the program of Gaussian 92. The molecular connectivity indexes were calculated with the program written by Zhu Longguan (Zhu et al., 1995). All the calculations were performed on a Pentium 586 computer.

RESULTS AND DISCUSSION

In order to validate our new method, we calculated some specific hydrocarbons and correlated them with some physical-chemistry properties. The calculated hydrocarbons were alkyl, alkene, alkyne, aroma and some molecules with conjugate π bond. The correlation among molecular

connectivity indexes and the new indexes and the molecular standard formation heat, free energy, entropy and isobaric thermal capacity (Dean et al., 1985) are described by Equations (2), (3), (4) and (5), respectively.

$$\Delta H_f^0 = 18.515 - 171.770\Delta^0\chi_p + 78.628^0\chi_p \tag{2}$$

$N = 14$, $S = 28.5$, $F = 98.0$, $R = 0.973$

$$\Delta G_f^0 = 33.467 + 64.336^0\chi_p^v - 183.085\Delta^0\chi_p$$
$$+ 6.519\Delta^3\chi_p \tag{3}$$

$N = 14$, $S = 23.2$, $F = 42.2$, $R = 0.963$

$$\Delta S_f^0 = 128.916 + 35.869^0\chi_p^v - 239.494^5\chi_{c_H} \tag{4}$$

$N = 14$, $S = 10.0$, $F = 55.1$, $R = 0.971$

$$\Delta C_f^0 = 8.099 + 14.787^0\chi_p^v - 9.223^4\chi_p \tag{5}$$

$N = 14$, $S = 5.0$, $F = 138.9$, $R = 0.981$

The new index has the characteristics of both the quantum chemistry and the topology method and make up for their shortcomings. We conclude on some points as follows:

1. The vital disadvantage of the quantum chemistry method is that it cannot reflect the steric effect. We can correlate the molecular volume to the Mulliken bond order with our new index (Equation 6), which offsets the disadvantage of the quantum chemistry method.

$$V_m = 2.872 + 6.637^0\chi_p^v + 6.835^0\chi_p \tag{6}$$

$N = 19$, $S = 2.7$, $F = 477.4$, $R = 0.992$

2. The molecular connectivity index is obtained from the molecular ichnographic-topological structure. So it can lose some information about the molecular configuration which has magnitude relationship with its properties. For example, pharmic activities of stereo-isomerous molecules often have distinct difference. The molecular connectivity index cannot distinguish stero-isomerous molecules. Our new index is a kind of topology-quantum index and can deal with this problem efficiently. For example, the volume of trans-2-butylene and cis-2-butylene are 62.789 and 62.263 cm³/mol, respectively, and the values calculated by Equation (6) are 62.135 and 61.914, but they have the same values according to the molecular connectivity index.

3. The molecular connectivity method actu-

ally deals only with the covalent bond, although some non-bond forces are of importance to molecular properties. Furthermore, the molecular volume and electronic properties of molecules, which have the same number of atoms, are different in different of atomic chemical environment, but the molecular connectivity indexes have the same value for the above situation. The above problems can be solved by the quantum chemistry method, so our new index can make up for the shortcoming of the molecular connectivity index. Equation (7) is the relationship between Δ terms and hydrophobic constants of 18 hydrocarbons (Gilles et al., 1994), where the Δ term is the difference of the corresponding molecular connectivity index and our new index. From Equation (7) we can see that they have good relationship, which means that the difference in value caused by the difference in chemical environment of the electronic distribution has important relationship with the molecular properties.

$$\log P = 0.634 + 0.665\Delta^1\chi_p - 0.772\Delta^2\chi_{cH}$$
$$- 0.312\Delta^5\chi_{cH} \qquad (7)$$
$$N = 18, \quad S = 0.26, \quad F = 39.0, \quad R = 0.945$$

4. A main factor that limits the wide application of the molecular connectivity index is that its physical meaning is vague because of the indistinctness of the atomic delta value. It makes clear the physical meaning of the atomic delta value weighted by Mulliken bond order and may be a way to solve the meaning of the topological index. So our new method can not only widen the applied range of the index, but can also make it a kind of scientific and reasonable index for QSAR.

References

Dean, J.A. (ed). 1985. Lange's Hand Book of Chemistry. Thirteenth Edition, Translated by Shang J. F. et al. Scientific Press, Beijing, p.6 – 56.

Galvez, J., Garcia, R., Salabert, M. T., et al. 1994. Charge indexes: New topological descriptors. *J. Chem. Inf. Comput. Sci.*, **34**: 520.

Gilles, K., Li Juyun, Wang Shaomeng, et al, 1994. Computer automated logP calculations based on an extended group contribution approach. *J. Chem. Inf. Comput. Sci.*, **34**: 752.

Hall, L.H., Kier, L.B., 1990. Determination of topological state: structure information at the atomic level for molecular graphs. *Quant. Struct. Act. Relat.*, **9**: 115.

Kier, L.B., Hall, L.H., 1976. Molecular Connectivity in Chemistry and Drug Research. Academic, New York, p.36.

Kier, L.B., Hall, L.H., 1985. Molecular Connectivity in Structure-Activity Analysis. Ltd, Letchworth, Research Studies Press, England, p.75.

Ni Caihua, Feng Zhiyun, Li Liangchao, 1998. Studies on the physico-chemical properties of unsaturated hydrocarbons using topological indexes. *Acta Chimica Sinica*, **56**: 359.

Pogliani, L., 1992. Molecular connectivity model for determination of isoelectric point of amino acids. *J. Pharm. Sci.*, **81**(4): 334.

Wang Liansheng, Zhi Zhengliang, 1992. Molecular Connectivity and Molecular Structure-activity. Environmental Science Press of China, Beijing, p.326.

Xin Houwen, Molecular Topology, 1991. Chinese Science and Technology University Press, Hefei, p.382.

Xu Guangxian, Li Lemin, Wang Demin, 1985. Quantum Chemistry-basic Theory and Ab Nitio Calculation Method. Science Press, Beijing, p.785.

Xu Lu, Jiang Yuangsheng, 1985. Chemometrical Method, Science Press, Beijing, p.522.

Yang Feng, Yan Xiaoci, Ou Yangli, et al, 1998. An improved delta value of connectivity index and its applications. *Chinese Journal of Chemical Physics*, **11**(3), 221.

Zhu Longguan, 1995. Some methods for QSAR and QSAR study on compounds of fluoroquinolone. Ph. D. Thesis, Zhejiang University, Hangzhou, p.98.