J. Serb. Chem. Soc. 70 (4) 669–674 (2005)
JSCS – 3300

UDC 539.193+519.17+54-12
Note

N O T E

Graph of atomic orbitals and the molecular structure-descriptors based on it

ANDREY A. TOROPOV,[a] IVAN GUTMAN[b,* #] and BORIS FURTULA[b]

[a]*Scientific Research Institute "Algortihm-Engineering", 700125, F. Khodjaev Street 25, Tashkent, Uzbekistan and* [b]*Faculty of Science, University of Kragujevac, P. O. Box 60, 34000 Kragujevac, Serbia and Montenegro*

(Received 24 June 2004)

Abstract: The graph of atomic orbitals (GAO) is a novel type of molecular graph, recently proposed by one of the authors. Various molecular structure-descriptors computed for GAO are compared with their analogs computed for ordinary molecular graphs. The quality of these structure-descriptors was tested for correlation with the normal boiling points of alkanes and cycloalkanes. In all the studied cases, the results based on GAO are similar to, and usually slightly better than, those obtained by means of ordinary molecular graps.

Keywords: graph of atomic orbitals (GAO), molecular structure-descriptor, topological index.

INTRODUCTION

In chemical graph theory[1–4] many different types of molecular graphs are encountered. Of these, the graph in which heavy (usually carbon) atoms are represented by vertices, whereas hydrogen atoms are ignored, is employed in the vast majority of cases. Such graphs are simply called "*molecular graphs*". If, however, one wants to distinguish them from other types of molecular graphs, then the former are referred to as "*hydrogen-suppressed* or *hydrogen-depleted molecular graphs*". In "*hydrogen-filled molecular graphs*" hydrogen atoms also are represented by vertices. In what follows hydrogen-suppressed and hydrogen–filled molecular graphs are abbriviated by HSG and HFG, respectively.

The hydrogen-suppressed (HSG) and the hydrogen-filled (HFG) molecular graphs of 2-methylbutane are depicted in Fig. 1.

In a series of recently published articles,[5–13] one of the present authors introduced a novel type of molecular graphs, the "*graph of atomic orbitals*", GAO. In

* Corresponding author.
# Serbian Chemical Society active member.

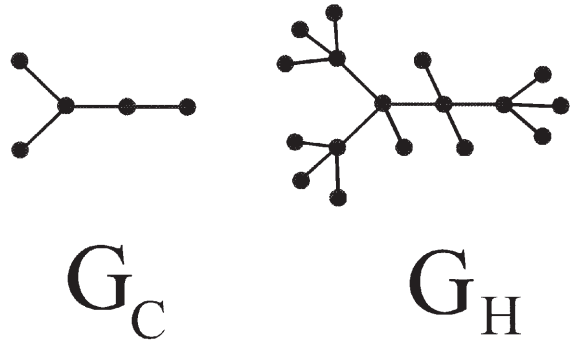

Fig. 1. The usual graph representations of 2-methylbutane; HSG: the hydrogen-suppressed molecular graph, HFG: the bydrogen-filled molecular graph.

all these papers, the vertices of the GAO were weighted by adjustable parameters, reflecting the nature of the group of atomic orbitals that these vertices represented. The properties of GAO, viewed as a simple (unweighted) graph, were not investigated. The aim of the present note is to partially fill this gap.

DEFINITION OF GAO

In both the hydrogen-suppressed and hydrogen-filled molecular graphs, each vertex represents an individual atom. In a GAO, each vertex represents a group of atomic orbitals of the respective atom. These groups of atomic orbitals are the following:

Atom	Groups of atomic orbitals	n_i
H	$1s^1$	1
C	$1s^2\ 2s^2\ 2p^2$	3
N	$1s^2\ 2s^2\ 2p^4$	3
O	$1s^2\ 2s^2\ 2p^4$	3
F	$1s^2\ 2s^2\ 2p^5$	3
S	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^4$	5
Cl	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^5$	5
Br	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 3d^{10}\ 4s^2\ 4p^5$	8
I	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 3d^{10}\ 4s^2\ 4p^6\ 4d^{10}\ 5s^2\ 5p^5$	11

Let M be a molecule and HFG its hydrogen-filled molecular graphs. Then the *graph of atomic orbitals* (GAO) pertaining to M is obtained from the HFG by replacing each of its vertices v_i by n_i mutually non-adjacent vertices, where the value of n_i depends on the type of atom represent by vertex v_i, as specified in the above Table. Two vertices in the GAO are adjacent if and only if they correspond to two different and adjacent atoms of M.

A mathematically more rigorous definition of GAO can be found elsewhere.[14] The GAO of 2-methylbutane is depicted in Fig. 2.

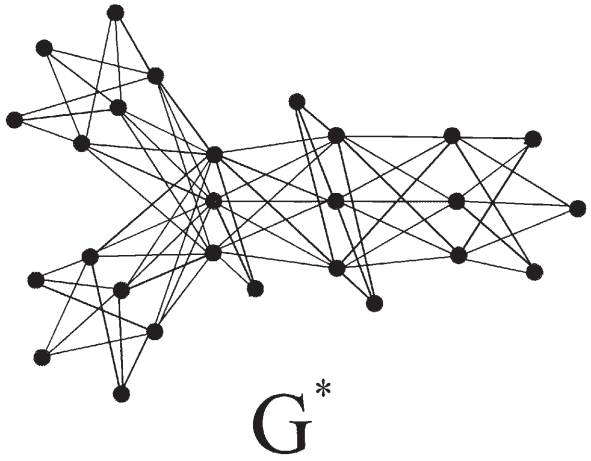

Fig. 2. The graph of atomic orbitals of 2-methylbutane.

COMPARING STRUCTURE-DESCRIPTORS OF GAO AND HSG

Figs. 1 and 2 illustate the fact that the structure of a GAO differs to a great extent from the structure of an ordinary molecular graph. It may, thus, be of some interest to know if molecular structure-descriptors (often referred to as "topological indices"),[3,4] calculated on the basis of a GAO, would be in any way related to the analogous structure-descriptors, calculated on the basis of the corresponding HSG and/or HFG.

In the present study, the following, most frequently employed,[3] graph-based structure-descriptors: the connectivity index[15] $\chi = {}^1\chi$ and its two simplest generalized versions – the zeroth-order ${}^0\chi$ and the second-order ${}^2\chi$ connectivity index,[16] the two Zagreb-group indices[17] M_1 and M_2, and the Wiener index[18] W, are examined.

Alkanes

The numerical values of the structure-descriptors under consideration were computed for a set of 74 alkanes,[19] both for the HSG and GAO. The respective correlation coefficients are as follows:

$R[\chi(\text{HSG}), \chi(\text{GAO})] = 0.9777$

$R[{}^0\chi(\text{HSG}), {}^0\chi(\text{GAO})] = 0.9927$

$R[{}^2\chi(\text{HSG}), {}^2\chi(\text{GAO})] = 0.8897$

$R[M_1(\text{HSG}), M_1(\text{GAO})] = 0.9709$

$R[M_2(\text{HSG}), M_2(\text{GAO})] = 0.9794$

$R[W(\text{HSG}), W(\text{GAO})] = 0.9968$

Thus, all the studied structure-descriptors of HSG and GAO are mutually correlated, especially ${}^0\chi$ and W.

The best three-variable correlation of the normal boiling points of alkanes for the structure-descriptors of HSG was the one based on ${}^0\chi$, χ, and W, having the following statistical characteristics: $R^2 = 0.9917$, $s = 4.2$ °C, $F = 2786$. Here and later the criterion for deciding which correlation was the best, was the correlation coefficient R.

The best three-variable correlation of the normal boiling points of alkanes for the structure-descriptors of GAO was also based on ${}^0\chi$, χ, and W. Its statistical characteristics were $R^2 = 0.9928$, $s = 3.9$ °C, $F = 3214$. It can be seen that GAO gave better results than HSG, but the improvement is not remarkable.

When the structure-descriptors of both HSG and GAO were used simultaneously, then the best three-variable corelation was found to be with ${}^0\chi$(HSG), χ(GAO), and W(GAO), for whith $R^2 = 0.9938$, $s = 3.6$ °C, $F = 3767$. The improvement gained by this mixed approach is not exciting, yet the standard error is reduced by almost 15 %.

Cycloalkanes

An analogous analysis was performed for a set of 45 cycloalkanes.[20] The results obtained are as follows:

$R[\chi(\text{HSG}), \chi(\text{GAO})] = 0.9854$

$R[{}^0\chi(\text{HSG}), {}^0\chi(\text{GAO})] = 0.9909$

$R[{}^2\chi(\text{HSG}), {}^2\chi(\text{GAO})] = 0.9450$

$R[M_1(\text{HSG}), M_1(\text{GAO})] = 0.9659$

$R[M_2(\text{HSG}), M_2(\text{GAO})] = 0.9648$

$R[W(\text{HSG}), W(\text{GAO})] = 0.9984$

Again, a reasonably high correlation exists between the structure-descriptors of HSG and GAO, especially in the case of ${}^0\chi$ and W.

The best three-variable correlation of the normal boiling points of cycloalkanes in the case of HSG is (as was the case for alkanes) for ${}^0\chi$, χ, and W, with $R^2 = 0.9913$, $s = 4.08$ °C, $F = 1548$. In the case of GAO, the best correlation was found for χ, M_1, and W, with $R^2 = 0.9910$, $s = 4.13$ °C, $F = 1508$. Here GAO performs slightly worse than HSG. The best three-variable correlation based on the structure-descriptors of both HSG and GAO is for ${}^0\chi$(HSG), χ(GAO), and W(GAO), with $R^2 = 0.9928$, $s = 3.70$ °C, $F = 1879$. Again, there is only a slight improvement relative to the results obtained using ordinary molecular graphs.

CONCLUDING REMARKS

The above outlined analysis pertains to the graph of molecular orbitals with unweighted vertices and/or edges. It was found that the quantitative structure-property (QSPR) relations obtained by means of such graphs possess essentially the

same precision as the analogous relations based on the ordinary molecular graphs. This is the consequence of the fact that a relatively good linear correlation exists between the structure-descriptors of GAO and of ordinary molecular graphs. In some cases (*e.g.*, for the Wiener index of alkanes)[14] an exact mathematical relation between the respective structure-descriptors could be established.

The finding that relatively complicated molecular graphs (or more generally: relatively complicated mathematical models) do not render any significant improvement in QSPR studies, relative to the simplest graph-representation of molecules, is not new. Several such observations were made in the past, see, for instance, the paper[21] and the references quoted therein.

In any case, our analysis implies that when choosing structure-descriptors for QSPR studies, preference should be given to the (much smaller and simpler) ordinary molecular graph, rather than to the (larger and structurally less transparent) graph of atomic orbitals with unweighted vertices and/or edges.

ИЗВОД

ГРАФ АТОМСКИХ ОРБИТАЛА И НА ЊЕМУ ЗАСНОВАНИ МОЛЕКУЛСКИ СТРУКТУРНИ ДЕСКРИПТОРИ

АНДРЕЈ А. ТОРОПОВ, ИВАН ГУТМАН и БОРИС ФУРТУЛА

Институт за алгоритме и инжењерство, Тошкент, Узбекистан и Природно-математички факултет у Крагујевцу

Граф атомских орбитала (GAO) је нови тип молекулских графова, недавно предложен од стране једног од аутора. Упоређени су разни молекулски структурни дескриптори, израчунати за GAO, са аналогним дескрипторима, израчунатим за обичне молекулске графове. Квалитет ових структурних дескриптора тестиран је на корелацији са нормалним тачкама кључања алкана и циклоалкана. У свим испитаним случајевима, резултати засновани на GAO су слични, и обично мало бољи, од оних добивених помоћу обичних молекулских графова.

(Примљено 24 јуна 2004)

REFERENCES

1. I. Gutman, O. E. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer-Verlag, Berlin, 1986
2. *Chemical Graph Theory – Introduction and Fundamentals*, D. Bonchev, D. H. Rouvray, Eds., Gordon & Breach, New York, 1991
3. R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim, 2000
4. I. Gutman, *Uvod u hemijsku teoriju grafova*, PMF Kragujevac, Kragujevac, 2003
5. A. A. Toropov, A. P. Toropova, *Russ. J. Coord. Chem.* **26** (2000) 398
6. A. A. Toropov, A. P. Toropova, *Russ. J. Coord. Chem.* **26** (2000) 792
7. A. Mercader, E. A. Castro, A. A. Toropov, *Chem. Phys. Lett.* **330** (2000) 612
8. A. A. Toropov, A. P. Toropova, *J. Mol. Struct. (Theochem)* **538** (2001) 287
9. A. A. Toropov, A. P. Toropova, *Russ. J. Coord. Chem.* **27** (2001) 574
10. A. A. Toropov, A. P. Toropova, *Internet Electron. J. Mol. Des.* **1** (2002) 109
11. A. A. Toropov, A. P. Toropova, *Russ. J. Coord. Chem.* **28** (2002) 877

12. A. A. Toropov, A. P. Toropova, *J. Mol Struct. (Theochem)* **637** (2003) 1
13. A. A. Toropov, A. P. Toropova, I. V. Nesterov, O. M. Nabiev, *J. Mol. Struct. (Theochem)* **640** (2003) 175
14. I. Gutman, A. A. Toropov, A. P. Toropova, *MATCH Commun. Math. Comput. Chem.* **53** (2005) 215
15. M. Randić, *J. Am. Chem. Soc.* **97** (1975) 6609
16. L. B. Kier, L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976
17. I. Gutman, B. Ruščić, N. Trinajstić, C. F. Wilcox, *J. Chem. Phys.* **62** (1975) 3399
18. H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 17
19. S. C. Basak, G. J. Niemi, G. D. Veith, *J. Math. Chem.* **7** (1991) 243
20. E. V. Konstantinova, M. V. Diudea, *Croat. Chem. Acta* **73** (2000) 383
21. E. A. Castro, I. Gutman, D. Marino, P. Peruzzo, *J. Serb. Chem. Soc.* **67** (2002) 647.