Bull. Chem. Soc. Jpn. Vol. 80, No. 4, 721–723 (2007)

Fragmentation Position Dependency of the Total Energy and Atomic Charge Difference between the Fragment MO Method and Conventional Ab Initio SCF-MO Method. A Case of (−)-Epicatechin Gallate with STO-3G Basis Set

Katsuhiro Tamura,[1,2] **Yuichi Inadomi,**[3,4] **and Umpei Nagashima***[1,3,4]

[1]Department of Chemistry, Tsukuba University, 1-1-1 Tennodai, Tsukuba 305-8571

[2]Shizuoka Industrial Research Institute of Shizuoka Prefecture, 2078 Makigaya, Aoi-ku, Shizuoka 421-1298

[3]Research Institute of Computational Science, National Institute of Advanced Industrial Science and Technology, 1-1-1 Umezono, Tsukuba 305-8561

[4]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi 332-0012

Received August 11, 2006; E-mail: u.nagashima@aist.go.jp

The fragmentation pattern dependence of the total energy and atomic charge difference between the fragment MO method and conventional ab initio Hartree–Fock SCF-MO method (HFMO) was investigated using 44 patterns of (−)-Epicatechin gallate. Both differences were sensitive to the fragmentation pattern. A "best" pattern for which results comparable to HFMO can be achieved, exists.

The fragment molecular orbital (FMO) method[1–4] has been proposed as a method to calculate the electronic structure for large molecular systems such as proteins and molecular clusters. In the FMO method, a target molecule is divided into small fragments, and the electronic structure for each fragment and fragment pair is calculated to obtain the total energy and other one-electron properties of the entire molecule. Because the FMO method does not involve MO calculations for the whole molecule, computational cost of the FMO method is drastically reduced in comparison to that of the conventional MO method, especially for large molecules. Because MO calculations for each fragment and fragment pair in the FMO method can be performed independently, the FMO method is ideally suited for parallel processing.[2,4] Based on calculations for several polypeptides using the ab initio Hartree–Fock SCF MO method (HFMO) as a reference,[2] the loss in accuracy when using the FMO method is less than 1 kcal mol^{-1} for the total energy. Highly accurate FMO calculations for large molecules can be carried out on computer systems as small as PC cluster, and are computationally inexpensive for evaluations the reactivities and chemical properties of large molecules.

In the FMO method, selection of the fragmentation position of a molecule and the distribution pattern of two electrons at a breaking bond is crucial for achieving highly accurate electronic structure calculations. The accuracy of the FMO method has only been confirmed for fragmentation at a sp^3 C–C bond, especially the Cα–C* bond of a protein and enzyme[4] and the C4′–C5′ bond of deoxyribose in DNA, because applications of the FMO method have focused mainly on proteins and DNA. The accuracy of the FMO method must still be confirmed for fragmentation of large molecular systems in general, such as autacoids and polymers.

In this study, to develop a guide for fragmentation of general molecules in the FMO method, the fragmentation pattern dependency of the total energy difference (ΔE) and atomic charge difference (ΔQ) between the FMO and HFMO methods were evaluated. (−)-Epicatechin gallate ((−)-ECg) was used as the representative molecule in the calculation because it is a type of tea catechin, and is known as an autacoid for physiologically active substances to undergo anti-oxidative reactions in biosystems. First, 44 fragmentation patterns of (−)-ECg were selected for the calculations. Then, for each pattern, ΔE and ΔQ between the FMO and HFMO methods were evaluated. Finally, errors for C–O and C–N bond fragmentation were investigated by applying the C–C bond fragmentation parameter set. For this best fragmentation pattern, the error in total energy was less than 1.0 kcal mol^{-1} and that in atomic charge was less than 0.0002. When the best fragmentation pattern was used in the FMO calculation, the parameter set for the C–C bond breaking could also be applied to C–N bond and C–O bond fragmentations.

Calculation

ABINIT-MP Ver.20021029[2,4,5] was used for the FMO calculation, and Gaussian98W Rev.-A.7[6] was used for the HFMO calculation. The threshold for the SCF convergence (1.0×10^{-8} and 1.0×10^{-6} in energy and the maximum difference of diagonal elements in density matrix, respectively) and accuracy of molecular integrals (1.0×10^{-12}) were the same for both calculations. All calculations were performed using the STO-3G basis set.

Figure 1 shows the molecular structure of (−)-ECg, which has 52 atoms, 230 electrons, and 180 basis functions. The geometrical parameters were fully optimized using Gaussian98W. For each of the 44 fragmentation patterns, the total energy and

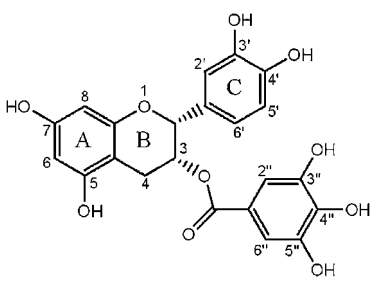

Fig. 1. Molecular structure of (−)-ECg.

Published on the web April 13, 2007; doi:10.1246/bcsj.80.721

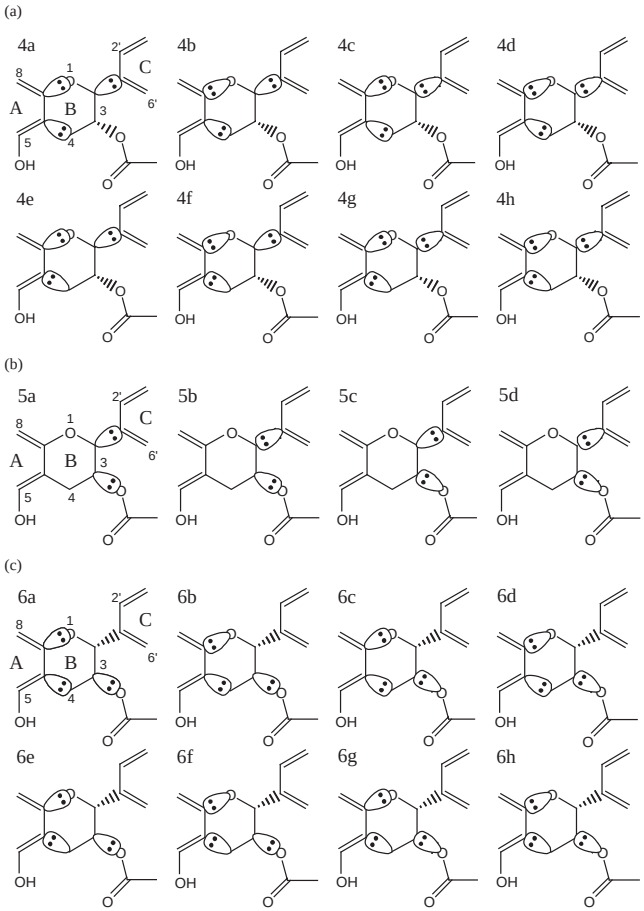

Fig. 2. Molecular fragmentation series 4 (a), 5 (b), and 6 (c), and patterns of bonding-electron distribution.

the Mulliken charge at the optimized geometry were calculated. Each pattern had three fragments and had three bonds broken. There were six fragmentation series consisting of a bond breaking in ring B and seven single bonds around ring B. For each fragmentation series, there were either four or eight electron-redistribution patterns. The fragmentation series are described as follows for a total of 44 fragmentation patterns.

Series 1: Fragmentation of two bonds in ring B and a bond connecting ring B and gallate ester part with eight electron-redistribution patterns.

Series 2: Fragmentation of three bonds in ring B with eight electron-redistribution patterns.

Series 3: Fragmentation of two bonds in ring B and a bond connecting ring B and ring C with eight electron-redistribution patterns.

Series 4: Fragmentation of two bonds in ring B and a bond connecting ring B and ring C with eight electron-redistribution patterns. The positions of these two bonds in ring B differed from the positions in Series 3 (Fig. 2a).

Series 5: Fragmentation of a bond connecting ring B and the gallate ester part of the molecule and of a bond connecting ring B and ring C with four electron-redistribution patterns (Fig. 2b).

Series 6: Fragmentation of three bonds in ring B with eight electron-redistribution patterns. The positions of the three bonds in ring B differed from the positions in Series 2 (Fig. 2c).

Note that series 5 is an exception in that only two bonds are broken, whereas three bonds are broken in all the other series.

Table 1. Average Values of ΔE (kcal mol^{-1}) and rmsΔQ

Series	No. of Patterns	ΔE	rmsΔQ
1	8	25.1	0.0031
2	8	41.9	0.0053
3	8	205.3	0.0452
4	8	13.7	0.0030
5	4	14.1	0.0020
6	8	13.9	0.0019

Table 2. ΔE (kcal mol^{-1}) and rmsΔQ for Different Electron-Distribution Patterns of Series 4, 5, and 6

	Series 4		Series 5		Series 6	
Pattern	ΔE	rmsΔQ	ΔE	rmsΔQ	ΔE	rmsΔQ
a	6.99	0.0018	1.76	0.0003	5.27	0.0006
b	−0.25	0.0021	5.61	0.0007	5.23	0.0004
c	33.27	0.0047	6.67	0.0011	40.96	0.0053
d	17.09	0.0032	42.39	0.0058	39.95	0.0053
e	6.88	0.0020	—	—	0.94	0.0004
f	−0.38	0.0025	—	—	0.87	0.0004
g	20.87	0.0043	—	—	8.66	0.0012
h	16.06	0.0030	—	—	9.08	0.0015

The dependence of the fragmentation parameter in bond fragmentation was evaluated by executing the same FMO calculations for model molecules, in which an oxygen atom at the 1-position of ring B of (−)-ECg was replaced by NH or by CH_2. The fragmentation parameter set for C–C bond fragmentation was applied in both the NH or CH_2 cases.

All calculations using the FMO and HFMO were performed on a notebook type PC, IBM ThinkPad X22, which had Microsoft Windows 2000, a Pentium III-M 800 MHz processor, 256 MB of memory, and a 40 GB hard drive.

Result and Discussion

Table 1 summarizes the averaged values of ΔE and root mean square of ΔQ (rmsΔQ) for the fragmentation series with the number of patterns. ΔE and rmsΔQ depended on the fragmentation pattern, suggesting the existence of a "best/better" fragmentation pattern for general molecular systems. For (−)-ECg, fragmentation series 3 had the largest ΔE and rmsΔQ. Although series 4, 5, and 6 had smaller ΔE and rmsΔQ, the values were similar among the three series (Table 2). We, therefore, focus our discussion on series 4, 5, and 6.

Figure 2 shows the patterns of fragmentation series 4, 5, and 6. In this figure, a drop shape indicates the location of the bond fragmentation position and the way to assign two electrons in the bond in fragmentation. Two electrons in the broken bond are put on atom at root of the drop shape.

Compared with other patterns, patterns 4b and 4f had small ΔE, but had similar rmsΔQ. A possible explanation for this result is that in patterns 4b and 4f, the electrons in a broken bond are redistributed to the oxygen atom at the 1-position and to the carbon atom adjoining this oxygen atom. The carbon atom adjoining the oxygen atom had sp^3 configuration and usually had a δ^+ atomic charge.

In series 5, although the number of bonds broken in the fragmentation was 2, which is smaller than that for the other series,

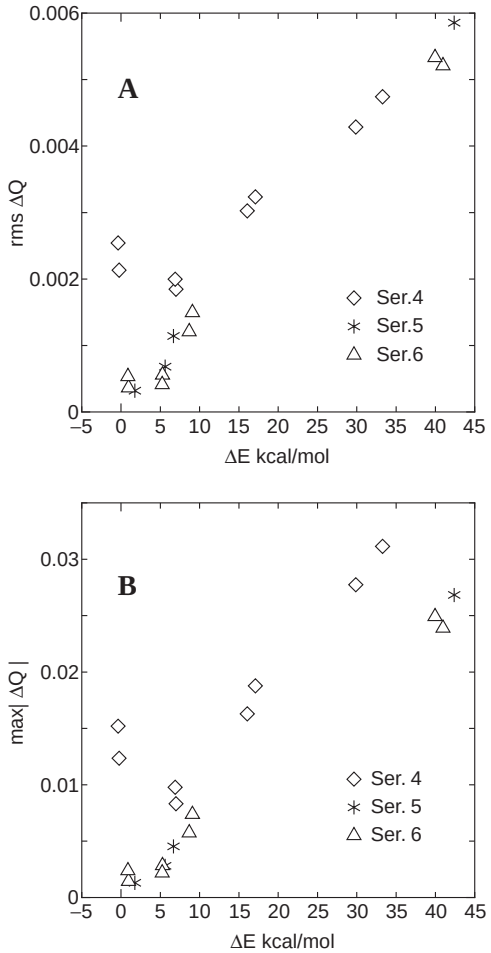

Fig. 3. Relationship between ΔE (kcal mol^{-1}) and root mean square (A) and maximum (B) of ΔQ for Series 4, 5, and 6.

ΔE and rmsΔQ were similar to those for the other series. This result suggests that the number of bonds broken in fragmentation does not significantly affect ΔE and rmsΔQ. ΔE for pattern 5a was reasonably small and rmsΔQ was almost 7 times smaller than that for patterns 4b and 4f. A possible explanation for this result is that electrons in 5a are also redistributed to the carbon atom that adjoins the oxygen atom and that has sp^3 configuration.

In series 6, ΔE and rmsΔQ for patterns 6e and 6f were reasonably small. The patterns, in which the electrons are redistributed into a fragment, such as in patterns 6c and 6d, had large ΔE and rmsΔQs.

Figure 3 shows the relation between ΔE and ΔQ for series 4, 5, and 6. Both rmsΔQ (Fig. 3A) and the largest ΔQ (max $|\Delta Q|$) (Fig. 3B) were proportional to ΔEs and showed almost the same trends. For patterns 4b and 4f, however, ΔE was small, whereas rmsΔQ and max $|\Delta Q|$ were large. Both the ΔE and ΔQ strongly depended on the fragmentation pattern and the electron redistribution. The atoms at a broken bond usually had the largest value of max $|\Delta Q|$ in each fragmentation pattern (data not shown in detail). For patterns 6a, 6b, and 6e, however, max $|\Delta Q|$ occurred on carbonic oxygen in the ester base. These carbonic oxygen atoms are far from the bond-breaking positions.

Based on the ΔE and ΔQ results, the best fragmentation pattern for the FMO method was pattern 6e or 6f, in which ΔE was less than 1 kcal mol^{-1}, and rmsΔQs and max $|\Delta Q|$ were less than 0.0005 and 0.005, respectively. Therefore, to achieve highly accurate FMO calculations of the electronic structure of large molecules, (1) the number of electrons for each fragment should be close to that for neutral one, (2) the electrons in the fragmentation bond should be redistributed to a carbon that has a sp^3 configuration, and (3) the electrons in the fragmentation bond should be redistributed to an atom that has a large electronegativity.

The effect of substitution was analyzed by replacing the oxygen at the 1-position of ring B (1O in Fig. 1) with NH and with CH_2. Patterns c and d showed large ΔE and rmsΔQ. Patterns g and h showed relatively large ΔE and rmsΔQ, almost 2 kcal mol^{-1} and 0.0003, respectively. In contrast, patterns e and f showed small ΔE and rmsΔQ, less than 1 kcal mol^{-1} and 0.0003, respectively. This change in accuracy caused by replacing the oxygen atom was small. Thus the accuracy of the FMO calculation was insensitive to the fragmentation parameters, although it was extremely sensitive to the fragmentation pattern itself.

References

1 K. Kitaura, E. Ikeo, T. Asada, T. Nakano, M. Uebayasi, *Chem. Phys. Lett.* **1999**, *313*, 701.

2 T. Nakano, T. Kaminuma, T. Sato, Y. Akiyama, M. Uebayasi, K. Kitaura, *Chem. Phys. Lett.* **2000**, *318*, 614.

3 K. Kitaura, S. Sugiki, T. Nakano, Y. Komeiji, M. Uebayasi, *Chem. Phys. Lett.* **2001**, *336*, 163.

4 T. Nakano, T. Kaminuma, T. Sato, K. Fukuzawa, Y. Akiyama, M. Uebayasi, K. Kitaura, *Chem. Phys. Lett.* **2002**, *351*, 475.

5 T. Nakano et al., http://moldb.nihs.go.jp/abinitmp/.

6 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, J. A. Pople, *Gaussian98W, Rev. A.7*, Gaussian, Inc., Pittsburgh PA, **1998**.