# An Efficient Method to Compute Partial Atomic Charges of Large Molecules Using Reassociation of Fragments

**Jung-Goo Lee,[*] Ho Young Jeong,[†] and Hosull Lee[‡]**

*Department of Physics, North Carolina State University, Raleigh, NC 27695, USA*
*[†]Gatsby Computational Neuroscience Unit, University College London, London WC1N 3AR, UK*
*[‡]Kumho Chemical Laboratories, P.O. Box 64, Yuseong, Daejeon 305-600, Korea*

*Received May 2, 2002*

Coulson (ZINDO), Mulliken (MP2/6-31G*) and Natural (MP2/6-31G*) population analyses of several large molecules were performed by the Fragment Reassociation (FR) method. The agreement between the conventional ZINDO (or conventional MP2) and FR-ZINDO (or FR-MP2) charges of these molecules was excellent. The standard deviations of the FR-ZINDO net atomic charges from the conventional ZINDO net atomic charges were 0.0008 for $C_{10}H_{22}$ (32 atoms), 0.0012 for $NH_2$-$C_{16}O_2H_{28}$-COOH (53 atoms), 0.0014 for $NH_3^+$-$C_{16}O_2H_{28}$-COOH (54 atoms), 0.0017 for $NH_2$-$C_{16}O_2H_{28}$-$COO^-$ (52 atoms), 0.0019 for $NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$ (53 atoms), 0.0024 for a conjugated model (O=CH-$(CH=CH)_{15}$-C=O-$(CH=CH)_{12}$-CH=$CH_2$), 118 atoms), 0.0038 for aglycoristocetin ($C_{60}N_7O_{19}H_{52}^+$, 138 atoms), 0.0023 for a polypropylene model complexed with a zirconocene catalyst ($C_{68}H_{121}Zr^+$, 190 atoms) and 0.0013 for magainin ($C_{112}N_{29}O_{28}SH_{177}$, 347 atoms), respectively. The standard deviations of the FR-MP2 Mulliken (or Natural) partial atomic charges from the conventional ones were 0.0016 (or 0.0016) for $C_{10}H_{22}$, 0.0019 (or 0.0018) for $NH_2$-$C_{16}O_2H_{28}$-COOH and 0.0033 (or 0.0023) for $NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$, respectively. These errors were attributed to the shape of molecules, the choice of fragments and the degree of ionic characters of molecules as well as the choice of methods. The CPU time of aglycoristocetin, conjugated model, polypropylene model complexed with zirconocene and magainin computed by the FR-ZINDO method was respectively 2, 4, 6 and 21 times faster than that by the normal ZINDO method. The CPU time of $NH_2$-$C_{16}O_2H_{28}$-COOH and $NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$ computed by the FR-MP2 method was, respectively, 6 and 20 times faster than that by the normal MP2 method. The largest molecule calculated by the FR-ZINDO method was B-DNA (766 atoms). These results will enable us to compute atomic charges of huge molecules near future.

**Key Words :** Population analysis, Fragment reassociation

## Introduction

Net atomic charges or partial atomic charges play a significant role to analyze the polarization effects of molecules and the electrostatic interactions in chemical reactions. Also, atomic charges are widely used as one of the QSAR parameters for drug design. The electrostatic potential is very important as one of the force fields for molecular modeling.[1] Classically, atomic charges can be obtained, for example, by the charge equilibration method,[2] where empirical parameters such as atomic ionization potentials, electron affinities and atomic radii are employed. Quantum mechanically, many attempts[3-21] have been made to calculate atomic charges or to perform population analyses. Typical population analyses are based on the basis set (the Coulson, Mulliken[3] and Natural[4] population analysis, or briefly CPA, MPA and NPA), the ElectroStatic Potential (ESP),[4-13] and the wave function (the topological theory of Atoms In Molecules (AIM)).[14-16] The Coulson and Mulliken charges are simple to calculate but vary a lot depending on basis functions. On the other hand, ESP and AIM charges are rather independent of basis set but take more CPU time. Natural charges[17-21] are based on the natural atomic orbital whose derivation involves diagonalizing the localized block of the electron density matrix with atomic basis set. Natural charges are basis set independent and take slightly more CPU time than Mulliken charges do. Therefore, for analysis purposes the NPA procedure is an attractive method, but for modeling purposes (*i.e.* force field charges) ESP charges are clearly the logical choice.[1] Even in NPA, however, computations of a large molecular system is still limited to less than 100 atoms at the MP2 level with the 6-31G* basis set, and to a few hundred atoms at the Hartree-Fock (HF) level. As a practical sense, atomic charges of large molecules obtained by NPA or even CPA and MPA are used for the analyses described earlier.

Some time ago, we proposed so called the Fragment Density Matrix (FDM) method[22] where the HF electrostatic potentials of medium (29 to 32 atomic) molecules were constructed from the density matrix elements of two fragment molecules. When the FDM method is applied to the calculations of CPA, MPA and NPA, atomic charges of large molecules are obtained by the reassociation of atomic charges of fragments rather than by that of density matrix elements of fragments, as derived in the next section. Therefore, the CPU time of the Coulson, Mulliken or

---

[*]To whom correspondence should be addressed. Phone: +919-513-0317; Fax: +919-513-4804; E-mail: jglee@nemo.physics.ncsu.edu

Natural charges expressed by the method of Fragment Reassociation, or the FR method, will become even shorter than that obtained by the FDM method.

In this work, the Coulson charges of molecules with up to 766 atoms or the Mulliken and Natural charges of molecules with up to 53 atoms were obtained by the semiempirical (ZINDO) or ab initio (MP2/6-31G*) calculations using the FR method. We will abbreviate these methods as FR-ZINDO and FR-MP2, respectively. Here, the FR-MP2 method means either FR-MP2 MPA or FR-MP2 NPA computation, or collectively both calculations. The difference between charges obtained by the FR-ZINDO (or FR-MP2) method and those obtained by the conventional ZINDO (or MP2 method) is also discussed. We have chosen neutral, cationic, anionic, and zwitterionic models as well as a conjugated model as target molecules.

## Computational Details

**A. Net Atomic Charges**. Net atomic charges mainly consist of the following four types. The first one is by the Coulson type,

$$q_A = Z_A - \sum_{\mu \in A} P_{\mu\mu} \tag{1}$$

where $q_A$ is the net atomic charge on $A^{th}$ nucleus, $Z_A$ the charge of $A^{th}$ nucleus, and $P_{\mu\mu}$ the density matrix given by the summation of multiplications of M.O. coefficients ($c$) over occupied orbitals;

$$P_{\mu\mu} = \sum_{i=1}^{occ} c_{\mu i} c_{\upsilon i}\,. \tag{2}$$

The second one is by the Mulliken type,

$$q_A = Z_A - \sum_{\mu \in A} \sum_{\upsilon \in A} P_{\mu\upsilon} S_{\mu\upsilon} \tag{3}$$

where $S_{\mu\upsilon}$ is the overlap matrix. Therefore, the Mulliken net atomic charges give similar values to the Coulson net atomic charges.

The third one is by the Natural charge,

$$q_A = Z_A - N_A \tag{4}$$

Here, $N_A$ is the total number of electrons on center A and obtained from the sum of the diagonal elements ($\Sigma_{\mu \in A} q_\mu$) of the first-order reduced density matrix in the orthonormal Natural Atomic Orbital (NAO) basis. The NAOs on each center A typically separate into two distinct sets: (1) the high-occupancy natural minimal basis set, equal in number and type to the nominal minimal basis of occupied HF AOs in the ground state configuration, and (2) the low-occupancy natural Rydberg basis set, corresponding to all residual orbitals lying outside the formal valence shell.

The fourth one, which represents the net atomic charges more accurately than the previous three methods, is by the ESP fit,

$$(q_1, q_2, \ldots, q_N) = R^{-1} V_i\,. \tag{5}$$

In eq 5, $R$ is the matrix of inverse distance from grid point to nuclei, and $V_i$ is the electrostatic potential surrounding a molecule at a grid point, $r_i$, given by

$$V_i = \sum_{A=1}^{N} \frac{Z_A}{|r_i - R_A|} - \sum_{\mu} \sum_{\upsilon} P_{\mu\upsilon} \int \frac{\chi_\mu \chi_\upsilon}{|r_i - r'|} dr'\,. \tag{6}$$

Here, $N$ is the number of atoms in a molecule, $R_A$ the position of $A^{th}$ nucleus, $r'$ the position of electron, and $\chi_\mu$ and $\chi_\upsilon$ are the basis functions used.

In all these four types, density matrix terms of a molecule have to be computed after SCF iteration. Therefore, it takes a large amount of cpu time to calculate atomic charges of large molecules by ordinary SCF methods. In order to avoid the computation of whole density matrix, the Fragment Density Matrix (FDM) method was proposed and explained in the following section.

**B. FDM Approach**. Instead of calculating accurate density matrix elements for a large molecular system, approximate density matrix terms $P_{\mu\upsilon}{}'$, are constructed from those of two or more fragment molecules in the FDM method. In the evaluation of $P_{\mu\upsilon}{}'$, it is important for all atoms to include neighboring effects (usually 2 to 4 nearest neighbors for a neutral molecule). The approximate density matrix elements $P_{\mu\upsilon}{}'$ are then scaled in order to satisfy the total atomic charge requirement. Therefore, the final density matrix $P_{\mu\upsilon}{}'$ can be expressed by

$$P_{\mu\upsilon} = k_{\mu\upsilon} P_{\mu\upsilon}{}' \tag{7}$$

where $k_{\mu\upsilon}$ are the scale factors. Or more simply,

$$P_{\mu\upsilon} = k P_{\mu\upsilon}{}' \tag{8}$$

In eq. 8, $k$ is a constant scale factor and the optimum $k$ ($= k_o$) is derived as

$$k_o = \frac{N_S - \alpha_S}{N_S - \Delta q|_{k=1}}, \quad \Delta q|_{k=1} = \sum_{A=1}^{N} q_A|_{k=1}\,. \tag{9}$$

Here, $N_s$ (and $\alpha_s$) are the summation of atomic numbers (and net atomic charges) over all atoms. The constant, $\alpha_s$, is equal to 0 for a neutral molecule, +1 for a cationic compound, and -1 for an anionic compound, respectively. $\Delta q|_{k=1}$ represents the summation of net atomic charges over all atoms when $k = 1$.

Since the FDM method constructs the density matrix,[23] it can be applied to evaluate atomic charges of the Coulson-Type, Mulliken-Type, Natural-Type, potential derived electrostatic-Type and other types in which atomic charges are derived from density matrix elements. When the FDM method is applied to compute the Coulson, Mulliken and Natural atomic charges, these charges can be expressed not by the density matrix elements of fragments but by partial atomic charges of fragments as derived next. This method, or the Fragment Reassociation (FR) method, will further reduce the cpu time for calculations of the Coulson, Mulliken and Natural net atomic charges. Hereafter, “net atomic charges” are abbreviated as “charges” for simplicity unless specified otherwise.

**C. FR Method to Compute Coulson, Mulliken and Natural Charges**. First, the Coulson charge is considered. The Coulson charge on atom $A$ obtained by the FDM method, $q_A'$, is

$$q_A' = Z_A - \sum_{\mu \in A} P_{\mu\upsilon}' \tag{10}$$

From eq. 1 and eq. 8,

$$kq_A' = kZ_A - \sum_{\mu \in A} P_{\mu\mu} \tag{11}$$

$$= kZ_A - (Z_A - q_A) \tag{12}$$

Then,

$$q_A = Z_A - k(Z_A - q_A') \tag{13}$$

Therefore, $q_A$ is expressed by the charges of the fragments $q_A'$ and the scale factor $k$ without including the density matrix terms.

The Mulliken charge is obtained in a similar manner if one assumes that

$$k = \frac{\sum_{\mu \in A} \sum_{\upsilon \in A} P_{\mu\upsilon} S_{\mu\upsilon}}{\sum_{\mu \in A} \sum_{\upsilon \in A} P_{\mu\upsilon}' S_{\mu\upsilon}'} \tag{14}$$

Then,

$$q_A = Z_A - k \sum_{\mu \in A} \sum_{\upsilon \in A} P_{\mu\upsilon}' S_{\mu\upsilon}' \tag{15}$$

$$= Z_A - k(Z_A - q_A') \tag{16}$$

Therefore, the Mulliken charge (eq. 16) is expressed by the same equation as the Coulson charge (eq. 13).

The Natural charge is also obtained in a similar way if one assumes that

$$k = \frac{N_A}{N_A'} \tag{17}$$

Then,

$$q_A = Z_A - kN_A' \tag{18}$$

$$= Z_A - k(Z_A - q_A') \tag{19}$$

Therefore, the Natural charge (eq. 19) is also expressed by the same equation as the Coulson charge (eq. 13). Although our FR method employs eq. 9 and eq. 13, other methods of reassociation of fragments were proposed elsewhere.[24-26]

**D. Target Molecules and Fragments**. A non-polar molecule ($C_{10}H_{22}$) was taken to investigate the error between the normal- and FR-ZINDO charges. Then, a hypothetical molecule ($NH_2$-$C_{16}O_2H_{28}$-COOH) was taken to study the error between the ZINDO and FR–ZINDO charges by the polar effect. Furthermore, the charges of cationic ($NH_3^+$-$C_{16}O_2H_{28}$-COOH), anionic ($NH_2$-$C_{16}O_2H_{28}$-$COO^-$) and zwitterionic ($NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$) forms were calculated to estimate the error by the ionic effect. The geometries of above compounds were arbitrarily chosen. The charges of several larger compounds, a conjugated model (O=CH-$(CH=CH)_{15}$-C=O-$(CH=CH)_{12}$-($CH=CH_2$), or $C_{58}O_2H_{58}$), aglycoristocetin ($C_{60}N_7O_{19}H_{52}^+$), a polypropylene model complexed with zirconocene ($Zr^+(C_5H_2C_4H_8CH_2)_2(CH_2$-$CH$-$CH_3)_{14}H(CH_2$=CH-$CH_3$), or $C_{68}H_{121}Zr^+$) and magainin ($C_{112}N_{29}O_{28}SH_{177}$), were computed by both FR-ZINDO and normal- ZINDO methods. All models had no symmetry although $C_{10}H_{22}$ and the conjugated model had nearly planar symmetry. ZINDO calculations of the target and fragment molecules except for B-DNA were carried out by the ZINDO program[27] in Cerius2 on Silicon Graphics (R10000) workstations. Because the maximum number of atoms in the ZINDO program is limited to 400, calculations of only B-DNA fragments were carried out. MP2 computations of $C_{10}H_{22}$, $NH_2$-$C_{16}O_2H_{28}$-COOH, $NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$ and their fragments were done by the Gaussian98 package[28] on a Cray-C94 at the KUMHO Supercomputing Center.

In FR calculations, the charges of first target molecule, $C_{10}H_{22}$, were constructed from those of two $C_7H_{16}$ fragment molecules as shown in Figure 1. In other words, the charges of the atoms left to the junction (solid line) of $C_{10}H_{22}$ (Figure 1a) were taken from those left to the junction of the left fragment molecule (Figure 1b) while charges of the atoms right to the junction of $C_{10}H_{22}$ (Figure 1a) were taken from those right to junction of the right fragment molecule (Figure 1c). To consider the neighboring effect, both right and left $C_7H_{16}$ fragments included 2 nearest neighbors (2 nearest carbons and 5 nearest hydrogens) from the junctions as shown in Figure 1b and 1c. For example, in FR-ZINDO calculations, Coulson charges of $C_{10}H_{22}$, ($q_1$, $q_2$, …, $q_{16}$, $q_{17}$, …, $q_{32}$; here the atomic numbering being shown in Figure 1a), were constructed from those of two fragments; (i) Coulson charges of the left fragment $C_7H_{16}$, ($q_1'$,$q_2'$,… $q_{16}'$; Figure 1b) and (ii) Coulson charges of the right fragment, $C_7H_{16}$, ($q_{17}'$, … $q_{32}'$; Figure 1c). Each charge, $q_A$ (A=1,2, …, 16, 17, …, or 32), was calculated from $q_A'$ according to eq.

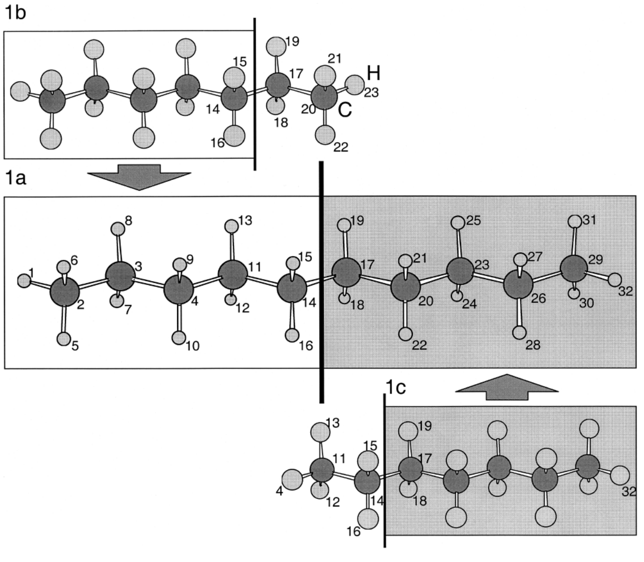

**Figure 1**. $C_{10}H_{22}$(a) and its fragments (b, c).

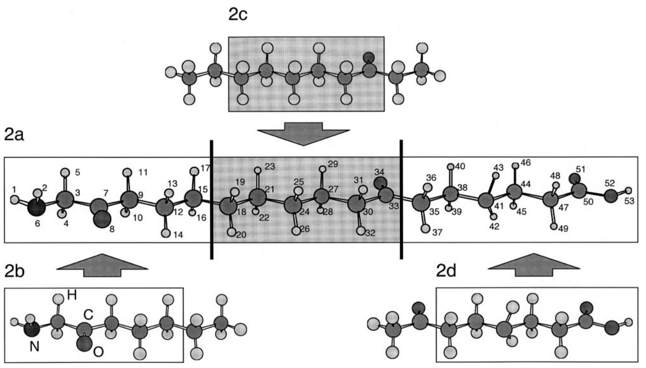

**Figure 2**. $NH_2$-$C_{16}O_2H_{28}$-COOH (a) and its fragment (b, c, d).

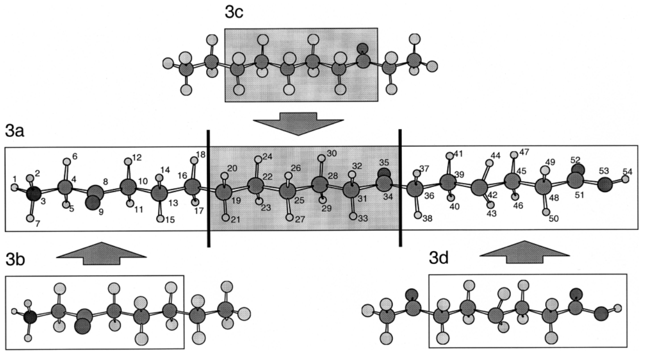

**Figure 3**. $NH_3^+$-$C_{16}O_2H_{28}$-COOH (a) and its fragment (b, c, d).

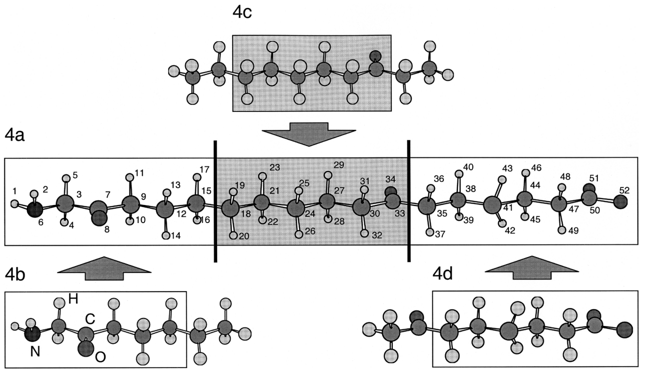

**Figure 4**. $NH_2$-$C_{16}O_2H_{28}$-$COO^-$ (a) and its fragment (b, c, d).

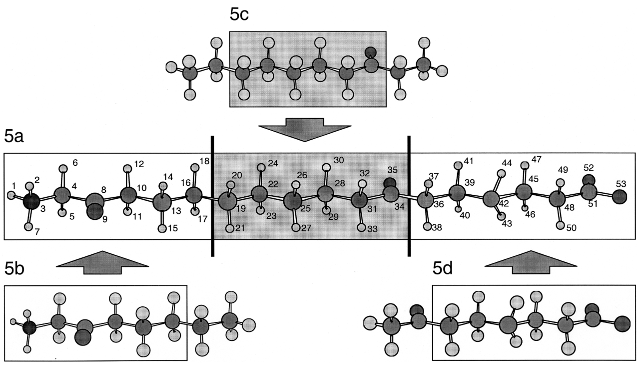

**Figure 5**. $NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$ (a) and its fragment (b, c, d).

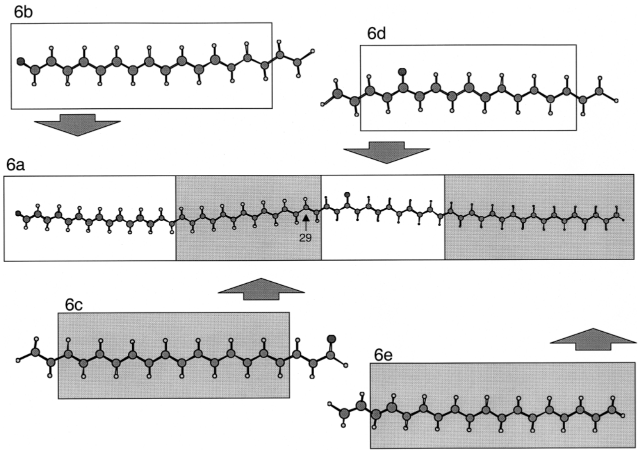

**Figure 6**. Conjugated model (a) and its fragment (b, c, d, e).

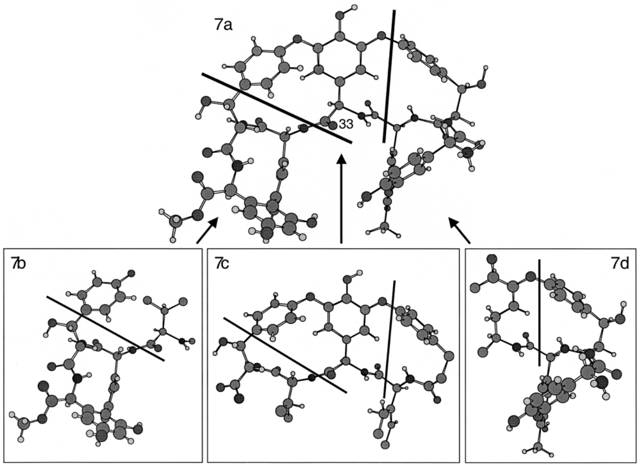

**Figure 7**. Aglycoristocetin (a) and its fragment (b, c, d).

13. The scale factor in eq. 13, *k*, was computed by eq. 9 where $\Delta q|_{k=1}$ was obtained from the summation of $q_A$', ($q_1$'+$q_2$'+ $\cdots$ $q_{32}$'). Note that Coulson charges of the neighboring atoms of these 2 fragments (from $q_{17}$' to $q_{23}$' in Figure 1b, and $q_4$' and $q_{11}$' to $q_{16}$' in Figure 1c) were not involved in the construction of those of $C_{10}H_{22}$.

Next four target molecules, $NH_2$-$C_{16}O_2H_{28}$-COOH (Figure 2) and its cationic (Figure 3), anionic (Figure 4) and zwitterionic (Figure 5) forms, were constructed from 3 fragment molecules. The left (Figure 2b, 3b, 4b and 5b), middle (Figure 2c, 3c, 4c and 5c) and right (Figure 2d, 3d, 4d and 5d) fragment molecules of these compounds also included 2 nearest neighbors. The FR charges of the conjugated model (Figure 6), aglycoristocetin (Figure 7), the polypropylene model complexed with zirconocene (Figure 8), magainin (Figure 9) and B-DNA (Figure 10) were composed of 4 (Figure 6b, 6c, 6d and 6e), 3 (Figure 7b, 7c and 7d), 4 (Figure 8b, 8c, 8d and 8e), 6 (Figure 9b, 9c, 9d, 9e, 9f and 9g) and 4 (Figire 10b, 10c, 10d and 10e) fragments, respectively. For the conjugated model, aglycoristocetin, the polypropylene model complexed with zirconocene and magainin, each fragment molecule included 2 to 4 nearest neighbors. For B-DNA, each fragment molecule included the nearest amino acid pair(s).

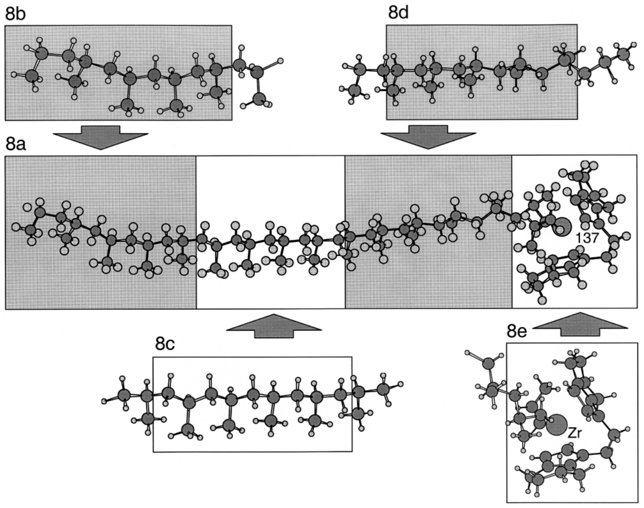

**Figure 8**. Polypropylene model complexed with a zirconocene (a) and its fragment (b, c, d, e).

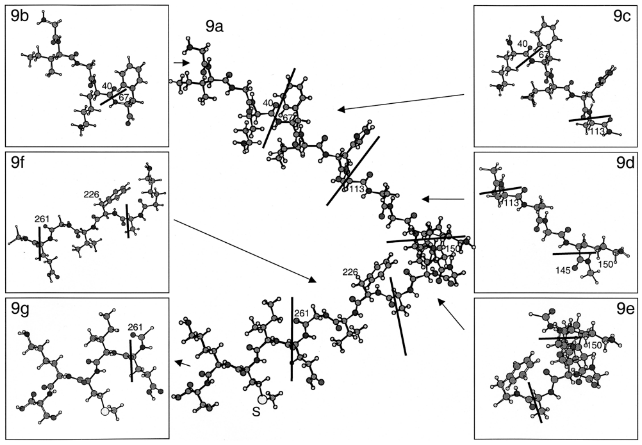

**Figure 9**. Magainin (a) and its fragment (b, c, d, e, f, g).

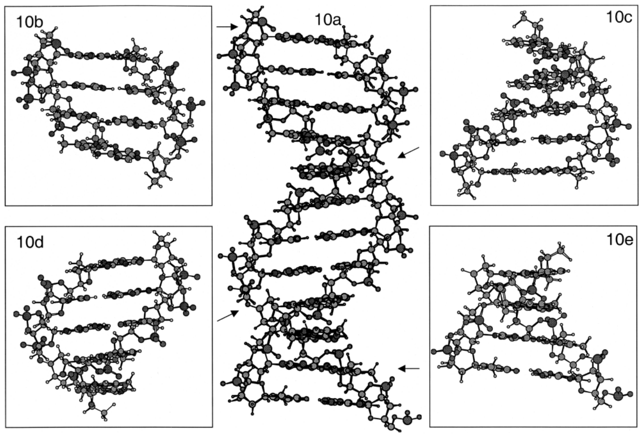

**Figure 10**. B-DNA (a) and its fragment (b, c, d. e).

## Results and Discussion

The normal-ZINDO and FR-ZINDO charges of $C_{10}H_{22}$ are listed in Table 1. Maximum differences between these net atomic charges occurred on two carbon atoms (C14 and C17) near the junction (solid line) in Figure 1a. The standard deviation (SD) between the ZINDO and FR-ZINDO charges turned out to be very small (0.0008). In comparison, the SD between the normal and FDM ab initio charges of $C_{10}H_{22}$ derived from the ESP fit was much larger (0.0173).[22]

**Table 1**. Partial Atomic Charges of $C_{10}H_{10}$ (ZINDO)

| Atom | Element | ZINDO | FR-ZINDO | Difference[a] |
|---|---|---|---|---|
| 1 | H | 0.0320 | 0.0317 | -0.0002 |
| 2 | C | -0.1130 | -0.1142 | -0.0012 |
| 3 | C | -0.0650 | -0.0662 | -0.0012 |
| 4 | C | -0.0680 | -0.0692 | -0.0012 |
| 5 | H | 0.0360 | 0.0357 | -0.0002 |
| 6 | H | 0.0360 | 0.0357 | -0.0002 |
| 7 | H | 0.0370 | 0.0367 | -0.0002 |
| 8 | H | 0.0370 | 0.0367 | -0.0002 |
| 9 | H | 0.0360 | 0.0357 | -0.0002 |
| 10 | H | 0.0360 | 0.0357 | -0.0002 |
| 11 | C | -0.0690 | -0.0692 | -0.0002 |
| 12 | H | 0.0350 | 0.0357 | 0.0007 |
| 13 | H | 0.0350 | 0.0357 | 0.0007 |
| 14 | C | -0.0700 | -0.0682 | 0.0017 |
| 15 | H | 0.0350 | 0.0357 | 0.0007 |
| 16 | H | 0.0350 | 0.0357 | 0.0007 |
| 17 | C | -0.0700 | -0.0682 | 0.0017 |
| 18 | H | 0.0350 | 0.0357 | 0.0007 |
| 19 | H | 0.0350 | 0.0357 | 0.0007 |
| 20 | C | -0.0690 | -0.0692 | -0.0002 |
| 21 | H | 0.0350 | 0.0357 | 0.0007 |
| 22 | H | 0.0350 | 0.0357 | 0.0007 |
| 23 | C | -0.0670 | -0.0682 | -0.0012 |
| 24 | H | 0.0360 | 0.0357 | -0.0002 |
| 25 | H | 0.0360 | 0.0357 | -0.0002 |
| 26 | C | -0.0660 | -0.0672 | -0.0012 |
| 27 | H | 0.0350 | 0.0347 | -0.0002 |
| 28 | H | 0.0350 | 0.0347 | -0.0002 |
| 29 | C | -0.1140 | -0.1152 | -0.0012 |
| 30 | H | 0.0350 | 0.0347 | -0.0002 |
| 31 | H | 0.0350 | 0.0347 | -0.0002 |
| 32 | H | 0.0320 | 0.0317 | -0.0002 |
| Standard deviation | 0.0008 | | | |
| Scale factor | 1.0003065 | | | |

[a] FR-ZINDO-ZINDO

In Table 2, both ZINDO and FR-ZINDO net atomic charges of $NH_2$-$C_{16}O_2H_{28}$-COOH are listed. The SD of the FR-ZINDO charges of this polar molecule was again smaller (0.0012) but larger than that of the non-polar molecule ($C_{10}H_{22}$). In order to see the fragment effect, we also computed the SD of the FR-ZINDO charges of $NH_2$-$C_{16}O_2H_{28}$-COOH with 2 fragments. This SD turned out to be 0.0007. Therefore, the error of the polar molecule by the FR-ZINDO method was affected by the choice of fragments rather than the polarity of molecules. The SDs of the FR-ZINDO charges of the cationic form $NH_3^+C_{16}O_2H_{28}$-COOH and the anionic form $NH_2$-$C_{16}O_2H_{28}$-$COO^-$ were 0.0014 and 0.0017, respectively. The SD of the zwitterionic form, $NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$, became even larger (0.0019) than that of the singly anionic or cationic form. Compared to the SD of the neutral form ($NH_2$-$C_{16}O_2H_{28}$-COOH), the SD's of the

**Table 2**. Partial Atomic Charges of $NH_2$-$C_{16}O_2H_{28}$-COOH (ZINDO)

| Atom | Element | ZINDO | FR-ZINDO | Difference[a] |
|---|---|---|---|---|
| 1 | H | 0.1660 | 0.1657 | -0.0002 |
| 2 | N | -0.4430 | -0.4447 | -0.0017 |
| 3 | C | 0.0520 | 0.0507 | -0.0012 |
| 4 | H | 0.0490 | 0.0486 | -0.0003 |
| 5 | H | 0.0530 | 0.0526 | -0.0003 |
| 6 | H | 0.1800 | 0.1797 | -0.0002 |
| 7 | C | 0.3900 | 0.3898 | -0.0001 |
| 8 | O | -0.5080 | -0.5090 | -0.0010 |
| 9 | C | -0.0560 | -0.0562 | -0.0002 |
| 10 | H | 0.0440 | 0.0436 | -0.0003 |
| 11 | H | 0.0450 | 0.0446 | -0.0003 |
| 12 | C | -0.0570 | -0.0573 | -0.0003 |
| 13 | H | 0.0460 | 0.0466 | 0.0006 |
| 14 | H | 0.0460 | 0.0466 | 0.0006 |
| 15 | C | -0.0660 | -0.0643 | 0.0016 |
| 16 | H | 0.0340 | 0.0346 | 0.0006 |
| 17 | H | 0.0350 | 0.0346 | -0.0003 |
| 18 | C | -0.0660 | -0.0653 | 0.0006 |
| 19 | H | 0.0350 | 0.0336 | -0.0013 |
| 20 | H | 0.0350 | 0.0336 | -0.0013 |
| 21 | C | -0.0660 | -0.0663 | -0.0003 |
| 22 | H | 0.0350 | 0.0366 | 0.0016 |
| 23 | H | 0.0350 | 0.0366 | 0.0016 |
| 24 | C | -0.0680 | -0.0693 | -0.0013 |
| 25 | H | 0.0320 | 0.0306 | -0.0013 |
| 26 | H | 0.0320 | 0.0306 | -0.0013 |
| 27 | C | -0.0630 | -0.0633 | -0.0003 |
| 28 | H | 0.0450 | 0.0456 | 0.0006 |
| 29 | H | 0.0450 | 0.0456 | 0.0006 |
| 30 | C | -0.0570 | -0.0562 | 0.0007 |
| 31 | H | 0.0450 | 0.0456 | 0.0006 |
| 32 | H | 0.0450 | 0.0456 | 0.0006 |
| 33 | C | 0.3840 | 0.3878 | 0.0038 |
| 34 | O | -0.5160 | -0.5130 | 0.0029 |
| 35 | C | -0.0520 | -0.0502 | 0.0017 |
| 36 | H | 0.0460 | 0.0466 | 0.0006 |
| 37 | H | 0.0450 | 0.0456 | 0.0006 |
| 38 | C | -0.0850 | -0.0853 | -0.0003 |
| 39 | H | 0.0190 | 0.0196 | 0.0006 |
| 40 | H | 0.0190 | 0.0196 | 0.0006 |
| 41 | C | -0.4390 | -0.4404 | -0.0014 |
| 42 | H | 0.2110 | 0.2117 | 0.0007 |
| 43 | H | 0.2110 | 0.2107 | -0.0002 |
| 44 | C | 0.0160 | 0.0147 | -0.0012 |
| 45 | H | 0.0410 | 0.0406 | -0.0003 |
| 46 | H | 0.0410 | 0.0406 | -0.0003 |
| 47 | C | -0.0310 | -0.0322 | -0.0012 |
| 48 | H | 0.0640 | 0.0637 | -0.0003 |
| 49 | H | 0.0640 | 0.0637 | -0.0003 |
| 50 | C | 0.5490 | 0.5478 | -0.0011 |
| 51 | O | -0.5990 | -0.6011 | -0.0021 |
| 52 | O | -0.3230 | -0.3250 | -0.0020 |
| 53 | H | 0.2630 | 0.2627 | -0.0002 |
| Standard deviation | 0.0012 | | | |
| Scale factor | 1.0003204 | | | |

[a] FR-ZINDO-ZINDO

cationic, anionic and zwitterionic forms were increased by 17, 42 and 58 %, respectively. For these compounds, large differences between the ZINDO and FR-ZINDO charges were found on the carbon and oxygen atoms of the >C=O group near the right junction in Figure 2 to Figure 5.

The SD (=0.0024) of the conjugated model turned out to be rather large for a neutral molecule. The SD (=0.0038) of aglycoristocetin (cationic form) was largest among all models. Here, we wanted to check how much error was contributed from the ionic effect only. So, the ZINDO and FR-ZINDO charges of the neutral form ($C_{60}N_7O_{19}H_{51}$, 137 atoms) were calculated. The SD of the FR-ZINDO charges of the neutral form was greatly reduced (SD=0.0013) compared with that of the cationic form. The error caused by the ionic effect of aglycoristocetin was much larger than that of $NH_3^+C_{16}O_2H_{28}$-COOH. For the polypropylene model complexed with zirconocene (cationic form), the SD of the FR-ZINDO charges was 0.0023, which was larger than that (SD=0.0014) of $NH_3^+C_{16}O_2H_{28}$-COOH, but smaller than that of aglycoristocetin. In other words, as the degree of linearity of a molecule (such as $NH_3^+C_{16}O_2H_{28}$-COOH or $C_{68}H_{121}Zr^+$) increased, the error of the charges obtained by the FR-ZINDO method decreased. For magainin, the SD of the FR-ZINDO charges was also small (SD=0.0013). The reason for the small SD is that magainin has a V-shape being considered as a combination of two linear molecules. The CPU time of the conjugated model, aglycoristocetin, the polypropylene model complexed with zirconocene or magainin by the FR-ZINDO turned out to be, respectively, 4, 2, 6 or 21 times faster than that by the normal ZINDO method. In general, the more the molecular size is increased, the more the CPU saving by the FR method is. Also, the more the number of fragments of a molecule is increased, the more the CPU saving by the FR method is.

As shown in Table 3, the SD (=0.0016) of $C_{10}H_{22}$ between the MP2 and FR-MP2 Mulliken charges was larger than that (=0.0008) between the ZINDO and FR-ZINDO charges. The SD of $C_{10}H_{22}$ between the MP2 and FR-MP2 Natural charges was also 0.0016. However, these SD's of of $C_{10}H_{22}$ were still much smaller than in the cases of the FDM-ESP fit as described earlier. As shown in FR-ZINDO calculations of $C_{10}H_{22}$, large differences between the MP2 and FR-MP2 charges in both MPA and NPA also took place on the carbon atoms (C14 and C17) near the junction. The SD's (=0.0019 and 0.0033) of $NH_2C_{16}O_2H_{28}$-COOH and $NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$ between the MP2 and FR-MP2 Mulliken charges were again larger than those (=0.0012 and 0.0019) between the ZINDO and FR-ZINDO charges. The SD's (=0.0014 and 0.0023) of $NH_2C_{16}O_2H_{28}$-COOH and $NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$ between the MP2 and FR-MP2 Natural charges were smaller than those between MP2 and FR-MP2 Mulliken charges. The CPU time of $NH_2C_{16}O_2H_{28}$-COOH and $NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$ by the FR-MP2 calculations was, respectively, 6 and 20 times faster than that by the conventional MP2 calculations. Large errors of the FR-MP2 charges of these two compounds were found on the carbon and oxygen atoms near the right junctions in Figure 2 and Figure 5.

**Table 3**. Partial Atomic Charges of $C_{10}H_{10}$ (MP2)

| Atom | Element | Mulliken Charge | | | Natural Charge | | |
|---|---|---|---|---|---|---|---|
| | | MP2 | FR-MP2 | Difference[a] | MP2 | FR-MP2 | Difference[a] |
| 1 | H | 0.1561 | 0.1567 | 0.0006 | 0.2041 | 0.2045 | 0.0004 |
| 2 | C | -0.4756 | -0.4769 | -0.0013 | -0.5978 | -0.5982 | -0.0004 |
| 3 | C | -0.2822 | -0.2837 | -0.0015 | -0.3839 | -0.3844 | -0.0005 |
| 4 | C | -0.2854 | -0.2870 | -0.0016 | -0.3926 | -0.3931 | -0.0005 |
| 5 | H | 0.1532 | 0.1534 | 0.0002 | 0.1973 | 0.1974 | 0.0001 |
| 6 | H | 0.1532 | 0.1534 | 0.0002 | 0.1974 | 0.1976 | 0.0002 |
| 7 | H | 0.1480 | 0.1485 | 0.0005 | 0.1936 | 0.1939 | 0.0003 |
| 8 | H | 0.1481 | 0.1485 | 0.0004 | 0.1948 | 0.1951 | 0.0003 |
| 9 | H | 0.1446 | 0.1452 | 0.0006 | 0.1942 | 0.1945 | 0.0003 |
| 10 | H | 0.1446 | 0.1452 | 0.0006 | 0.1921 | 0.1924 | 0.0003 |
| 11 | C | -0.2898 | -0.2903 | -0.0005 | -0.3893 | -0.3883 | 0.0010 |
| 12 | H | 0.1448 | 0.1468 | 0.0020 | 0.1936 | 0.1949 | 0.0013 |
| 13 | H | 0.1448 | 0.1468 | 0.0020 | 0.1950 | 0.1963 | 0.0013 |
| 14 | C | -0.2898 | -0.2937 | -0.0039 | -0.3888 | -0.3949 | -0.0061 |
| 15 | H | 0.1450 | 0.1456 | 0.0006 | 0.1934 | 0.1933 | -0.0001 |
| 16 | H | 0.1449 | 0.1456 | 0.0007 | 0.1967 | 0.1979 | 0.0012 |
| 17 | C | -0.2900 | -0.2946 | -0.0046 | -0.3892 | -0.3946 | -0.0054 |
| 18 | H | 0.1449 | 0.1458 | 0.0009 | 0.1954 | 0.1964 | 0.0010 |
| 19 | H | 0.1449 | 0.1458 | 0.0009 | 0.1957 | 0.1960 | 0.0003 |
| 20 | C | -0.2897 | -0.2901 | -0.0004 | -0.3884 | -0.3875 | 0.0009 |
| 21 | H | 0.1451 | 0.1471 | 0.0020 | 0.1961 | 0.1973 | 0.0012 |
| 22 | H | 0.1451 | 0.1471 | 0.0020 | 0.1924 | 0.1937 | 0.0013 |
| 23 | C | -0.2855 | -0.2873 | -0.0018 | -0.3941 | -0.3947 | -0.0006 |
| 24 | H | 0.1445 | 0.1452 | 0.0007 | 0.1903 | 0.1906 | 0.0003 |
| 25 | H | 0.1445 | 0.1452 | 0.0007 | 0.1956 | 0.1961 | 0.0005 |
| 26 | C | -0.2852 | -0.2867 | -0.0015 | -0.3844 | -0.3848 | -0.0004 |
| 27 | H | 0.1467 | 0.1472 | 0.0005 | 0.1934 | 0.1937 | 0.0003 |
| 28 | H | 0.1467 | 0.1472 | 0.0005 | 0.1952 | 0.1955 | 0.0003 |
| 29 | C | -0.4759 | -0.4772 | -0.0013 | -0.5961 | -0.5965 | -0.0004 |
| 30 | H | 0.1518 | 0.1520 | 0.0002 | 0.1966 | 0.1967 | 0.0001 |
| 31 | H | 0.1518 | 0.1520 | 0.0002 | 0.1980 | 0.1981 | 0.0001 |
| 32 | H | 0.1557 | 0.1563 | 0.0006 | 0.2035 | 0.2040 | 0.0005 |
| | | Standard deviation | 0.0016 | | Standard deviation | 0.0016 | |
| | | Scale factor | 1.0001647 | | Scale factor | 1.0000744 | |

[a] FR-MP2-MP2.

## Conclusion

Net atomic charges of several model compounds were calculated by the FR-ZINDO and FR-MP2 methods. These charges were compared with the conventional ZINDO and MP2 methods, respectively. In the FR-ZINDO calculations, the standard deviations of neutral molecules (SD = 0.0008-0.0013) except for the conjugated model turned out to be smaller than those of ionic, cationic and zwitterionic compounds (SD = 0.0014-0.0038). In both FR-MP2 Mulliken and FR-MP2 Natural charge computations, the SD's (=0.0016-0.0033) of $C_{10}H_{22}$, $NH_2C_{16}O_2H_{28}$-COOH and $NH_3^+$-$C_{16}O_2H_{28}$-$COO^-$ were larger than those (SD = 0.008-0.0019) in the corresponding FR-ZINDO Coulson charge computations. These errors depended on the shape of molecules, the degree of ionic characters of molecules, the choice of fragment molecules as well as the choice of methods. The SDs (= 0.0008-0.0038) of all models by the FR-ZINDO and FR-MP2 methods are, however, much smaller than those (SD = 0.014-0.021)[22] of medium molecules by the FDM ab initio charges derived from the ESP fit. The CPU time by the FR-ZINDO (or FR-MP2) method turned out to be much faster than that by the normal ZINDO (or normal MP2) method. Because of the smaller SD and more CPU saving, the FR method, especially the FR-MP2 NPA method, would become a choice for systematic charge analyses of huge molecules.

**Supplementary Materials Available**. The FR-ZINDO and FR-MP2 atomic charges of the large molecules calculated in this work are available upon request.

**Acknowledgment**. The GAUSSIAN98 calculations were carried out at the KUMHO Supercomputing Center in Kwangju, Korea.

## References

1. Jensen, F. *Introduction to Computational Chemistry*; Wiley & Sons: Chichester, 1999; Chapter 9.
2. Rappe, A. K.; Goddard III, W. A. *J. Phys. Chem.* **1991**, *95*, 3358.

3. Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.
4. Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
5. Stone, A. J.; Alderton, M. *Mol. Phys.* **1985**, *56*, 1047.
6. Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894.
7. Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.
8. Williams, D. E. *Rev. Comput. Chem.* **1991**, *2*, 219.
9. Reynolds, C. A.; Essex, J. W.; Richards, W. G. *J. Am. Chem. Soc.* **1992**, *114*, 9075.
10. Aleman, C.; Orozro, M.; Luque, F. J. *Chem. Phys.* **1994**, *189*, 573.
11. Koch, U.; Egert, E. *J. Comput. Chem.* **1995**, *16*, 937.
12. Marynick, D. S. *J. Comput. Chem.* **1997**, *18*, 955.
13. Swart, M.; Duijnen, P. T. v.; Snijders, J. G. *J. Comput. Chem.* **2001**, *22*, 79.
14. Bader, R. F. W. *Atoms in Molecules - A Quantum Theory*; Oxford: London, 1990.
15. Cioslowski, J. *J. Am. Chem. Soc.* **1989**, *111*, 8333.
16. Cioslowski, J. *J. Chem. Phys. Lett.* **1992**, *189*, 524.
17. Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211.
18. Reed, A. E.; Weinhold, F. *J. Chem. Phys.* **1983**, *78*, 4066.
19. Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.
20. Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899.
21. Weinhold, F.; Carpenter, J. E. *The Structure of Small Molecules and Ions*; Plenum: 1988; p 227.
22. Lee, J.-G.; Friesner, R. A. *J. Phys. Chem.* **1993**, *97*, 3515.
23. Recently, the semiempirical devide and conquer method similar to our FDM approach was applied to a biological system. See Van der Vaart, A.; Gogonea, V.; Dixon, S. L.; Merz, K. M., Jr. *J. Comput. Chem.* **2000**, *21*, 1494.
24. Bellido, M. N.; Rullmann, J. A. C. *J. Comput. Chem.* **1989**, *10*, 479.
25. Young, L.; Topol, I. A.; Rashin, A. A.; Burt, S. K. *J. Comput. Chem.* **1997**, *18*, 522.
26. Dardenne, L. E.; Werneck, A. S.; Oliveira Neto, M.; Bisch, P. M. *J. Comput. Chem.* **2000**, *22*, 689.
27. ZINDO, User Guide, version 95.0/3.0.0, BIOSYM/Molecular Simulations: 1995.
28. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, Jr., R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98 (Revision A.7)*, Gaussian, Inc.: Pittsburgh, PA, 1998.