A Fast Method for Calculating Geometry-Dependent Net Atomic Charges for Polypeptides

Kwang-Hwi Cho,[†] Young Kee Kang,[‡] Kyoung Tai No,[§,||] and Harold A. Scheraga*[,†]

Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301, Department of Chemistry, Chungbuk National University, Cheongju, Chungbuk 361-763, Korea, and Department of Chemistry and Computer Aided Molecular Design Research Center, Soong Sil University, Seoul 156-743, Korea

Received: June 30, 2000; In Final Form: January 5, 2001

A new method for calculating geometry-dependent atomic charges (GDACs) for polypeptides is presented. It overcomes the limitations of the partial equalization of orbital electronegativity (PEOE) and modified PEOE (MPEOE) methods, which depend only on connectivity, not geometry. Introduction of distance-dependent damping factors helps to include the effect of environment in determining the variation of bond distance (without explicit contribution of the correlated variation of bond angles), and thereby to reduce the number of parameters required to represent different atomic species. Since the correlation between the geometry and the dipole moments of molecules is crucial for this method, ab initio molecular orbital calculations were carried out to obtain the geometries and dipole moments with the 6-31G** basis set at the level of B3LYP theory. When bond distances are fixed prior to a charge calculation, the methodology outlined here leads to a direct calculation of the permanent molecular charge distribution represented as a set of distributed monopoles that depend on the geometry of the molecule. Hence, this method automatically accounts for the transferability of charges of small amino acid residues to build up a large polypeptide molecule, and can therefore provide an approximate description of any redistribution of charge density of large polypeptide molecules. The parameters characterizing the charge transfer in the formation of bonds were optimized by using dipole moment components and total dipole moments of 50 molecules that serve as models for the backbone and side chains of proteins. The calculated total dipole moments of these 50 molecules agree well with the ab initio results within an error of 5%. The new charge scheme has been applied to seven conformers of *N*-acetylalanine-*N*′-methylamide (Ac-Ala-NHMe) with good agreement between ab initio and GDAC dipole moments. This method, however, gives poor results for conjugated systems that are larger than amides.

1. Introduction

Molecular simulation is concerned with interaction energies between or within molecules. Electrostatic interactions between polar and ionic molecules are important in molecular mechanics, molecular dynamics, and Monte Carlo simulations. Thus, it is necessary to have fast, reliable, and simple descriptions of charge distributions for simulations of molecules such as proteins.

Most of the currently available force fields[1−5] for treating proteins involve pairwise potentials in which electrostatic interactions are described with the Coulomb equation, and the net atomic charges are essential to represent the interactions. Because the electron distribution is assigned to individual nuclei, the definition of an atomic charge is somewhat arbitrary. The definition of atomic charges is critical when it is used to correlate or predict physical or chemical properties, such as dipole moments or electrostatic potentials.

Because there is no quantum mechanical operator for net atomic charges, and charges cannot be obtained directly from ab initio molecular orbital calculations, it is essential to develop methods to represent the charge distribution of a molecule.[6−43] Several such methods have received great attention. The most well-known method is the Mulliken population analysis (MPA),[6] which is based on simple partitioning of the electrons among the atomic orbital basis set. The MPA technique has been used in conjunction with both empirical and ab initio molecular orbital calculations. Although the magnitudes of MPA charges represent the polarity of atoms and functional groups reasonably well, they are inappropriate to describe electrostatic properties such as the electric moments of molecules and the electrostatic potentials around molecules. The results depend strongly on the level of the quantum mechanical approach being taken and the basis set chosen.

One of the best methods proposed so far, within the "net atomic charge approximation", is the electrostatic potential-derived (PD) charges method.[11,13,19] In the PD method, the charges are derived to reproduce the electrostatic potential at points around the molecule that are precalculated using ab initio molecular orbital calculations. The charges obtained by this method can reproduce the electrostatic potentials of the molecules well, and the charges depend on the conformations. However, ab initio molecular orbital calculations become difficult and less effective when dealing with large molecules such as polypeptides or proteins. Also, one faces a multiple-minimum problem in finding the best set of charges of such macromolecules, which satisfies the electrostatic potentials from ab initio molecular orbital calculations.

More effective methods that require significantly less computational effort have been proposed. These methods are based

* To whom correspondence should be addressed. Phone: (607) 255-4034. Fax: (607) 254-4700. E-mail: has5@cornell.edu.
[†] Cornell University.
[‡] Chungbuk National University.
[§] Soong Sil University.
[||] Member of the Center for Molecular Science, Korea.

10.1021/jp0023213 CCC: $20.00
Published on Web 04/04/2001

mainly on the concept of electronegativity equalization, viz., the electronegativity equalization method (EEM),[7−9] the partial equalization of orbital electronegativity (PEOE) method,[12,14] the full equalization of orbital electronegativity (FEOE) method,[10,18a] and the modified PEOE (MPEOE) method.[25,37,40] Unlike the PD method, these other methods are all capable of modeling of the redistribution of charges resulting from the assembly of molecular fragments, such as amino acid residues, to form proteins.

In general, the parameters in the electronegativity equalization methods are usually taken from atomic properties, viz., valence-state ionization potentials and electron affinities. In the original PEOE method of Gasteiger et al.,[12,14] these parameters are used without modification for describing atomic orbital electronegativities in molecules. This method has the tendency to produce partial charges whose dipole moments are too low. To overcome this difficulty, Abraham and Smith[20] proposed a parametrization of orbital electronegativity based on experimental dipole moments. This proposal was found to improve the values of calculated dipole moments.

Later, the PEOE was modified (MPEOE) by No et al.[25a] for neutral molecules and extended to ionic and aromatic molecules,[25b] and compounds containing halogens[29] and sulfur and phosphorus.[40] In the MPEOE method, instead of taking the parameters from ionization potentials and electron affinities, both the damping factors and sets of electronegativity parameters are obtained by using experimental dipole and quadrupole moments as constraints. The MPEOE method shows improvements over the PEOE method, and details are described in the Methods.

However, both PEOE and MPEOE charges are not geometry-dependent since the charge transfer depends only on the connectivity. This means that all atoms of a given type in a molecule with the same connectivity have the same parameters and hence the same atomic charges. But, atomic charges in a molecule are geometry-dependent.[33,38,41] In all current major applications,[1−5] however, the electrostatic parameters are considered to be fixed irrespective of the geometry of the molecules. Since geometry-dependent atomic charges are needed for a proper description of molecular interactions, this crude approximation leads to inaccuracies in the force fields as far as electrostatic forces are concerned. Atoms with the same connectivity but in different environments should have different charges for a proper description of the electronic properties of molecules.

Within a given molecule, there is greater charge flux in shorter than in longer bonds since there is more orbital overlap in shorter bonds. It is thus necessary to develop a new charge-calculation method that takes into account the geometric differences among the same species in a molecule. On the basis of the principle of the PEOE method, a formalism has been developed in this work for calculating atomic charges in molecules which are connectivity- and distance-dependent. The method leads to geometry-dependent net atomic charges (GDACs).

2. Methods

On the basis of the PEOE charge scheme, the total net atomic charge Q_A on an atom A is obtained by an iterative procedure[12] that converges when the transfer of charge between all atoms in a molecule is zero. The degree of charge transfer ($dq^{(n)}{}_{AB}$) from atom A to atom B in a given bond during the nth iteration of the iterative procedure is described by

$$dq^{(n)}{}_{AB} = \frac{[\chi^{(n-1)}{}_B - \chi^{(n-1)}{}_A]}{\chi_{A+}}(f_{AB})^n \quad \text{if} \quad \chi^{(n-1)}{}_B > \chi^{(n-1)}{}_A \quad (1)$$

where $\chi^{(n-1)}{}_A$ and $\chi^{(n-1)}{}_B$ are the electronegativities of atoms A and B in the $(n-1)$th iteration and χ_{A^+} is the electronegativity of the positive ion of atom A. f_{AB} represents the damping factor. The original PEOE method[12] used a fixed value of 0.5 for f_{AB}, and the MPEOE method used different values depending on the hybridization states, thereby controlling not only the termination of the iterations but also the charge flux between atoms in a bond. More details about the damping factor will be discussed later in this section.

The net atomic charge on atom A in the nth iteraction, $Q^{(n)}{}_A$, is obtained from

$$Q^{(n)}{}_A = Q^{(0)}{}_A + \sum_n \sum_B dq^{(n)}{}_{AB} \quad (2)$$

where $Q^{(0)}{}_A$ represents the initial net atomic charge, which is set to zero for the atoms in neutral molecules, and the index B sums over the atoms covalently bonded to atom A, while the index n indicates summation over the number of iterative cycles required until convergence is attained. In this calculation, the orbital electronegativity of atom i in the nth iteration is described as a function of the charges associated with the given atomic valence states; thus[12]

$$\chi^n{}_i = a_i + b_i Q^{(n)}{}_i + c_i [Q^{(n)}{}_i]^2 \quad (3)$$

where a_i, b_i, and c_i are determined from ionization potentials and electron affinities in the original PEOE method (see ref 12 for details).

The MPEOE method differs from the original PEOE method in several ways. Instead of using polynomials up to second order in eq 3, a linear equation is used. The values of a_i and b_i in the MPEOE method are parametrized by using a set of suitable gas-phase experimental dipole and quadrupole moments as constraints as part of an optimization procedure, whereas the a_i and b_I values in the PEOE method are obtained from ionization potentials and electron affinities. The MPEOE method uses a set of damping factors f_{AB} between atoms A and B, which differ depending on their hybridization; by contrast, the PEOE method uses a fixed value of 0.5 for all cases. Last, the MPEOE method treats the electronegativity of positively charged hydrogen, χ_{H+}, as an entirely independent parameter in the optimization; by contrast, the PEOE method uses a value of 20.02 for χ_{H+} by extrapolation beyond the first ionization potential of the hydrogen atom because hydrogen does not have the second ionization potential required by the method.

The MPEOE method leads to much improved results for dipole moments and electrostatic potentials compared to those obtained with the PEOE method (see Table 2 in ref 25a). It should be emphasized, however, that, when the PEOE or MPEOE formalisms are used with fixed damping factors, the charge density is assumed to depend only on the connectivity and not on the geometry. While this is computationally advantageous, the method is unable to model changes in the charge density occurring as a result of distortions in protein geometry significantly different from those of the standard geometries of the molecules used in obtaining the MPEOE parameters.

As the iteration proceeds in the PEOE method, f_{AB} decreases to 0.5^n in the nth iteration; i.e., f_{AB} serves to terminate an iteration. After 20 iterations, f_{AB} reaches $\sim 10^{-6}$, and the later iterations do not affect the charge calculation; in the MPEOE method, even though the authors did not stress a precise physical correspondence between this parameter and any molecular properties, f_{AB} serves two roles. One is the termination of an

iteration after a certain iteration as in the PEOE method since all the values lie between 0 and 1, with the fractional values (raised to the *n*th power) decreasing with increasing *n*. Another is the control of charge flux between atoms A and B according to their valence states. Since a large value of f_{AB} vanishes more slowly than a smaller value during the iterations, f_{AB} controls the flow of charge between atoms. The larger the value, the slower the decay; hence, the longer the iteration. A longer iteration means that there is greater charge flux through the bond between atoms according to their electronegativity differences.

The GDAC method suggested in this paper differs from the MPEOE method in three ways. First, instead of using fixed damping factors that were optimized for specific atomic species, the damping factor is a function of the distance between two atoms in a bond to take account of geometric effects. Second, since the charge distribution of a molecule can be represented better by the three components of the dipole moment than only by the total dipole moment, both of these ab initio quantities are used for parametrizations. Third, in the MPEOE method, the parameters for new atomic species are optimized while the parameters for previously obtained species are fixed; in the work reported here, by contrast, all parameters for the atomic species in polypeptides are parametrized simultaneously to avoid propagation of errors.

The bond distance is not the only geometric factor that reflects the conformation of a molecule. Other factors such as bond angles and torsional angles are also important geometric factors. Without using a multitude of parameters, there is no simple way to include all the geometric factors in molecular mechanical charge calculations. Dinur and Hagler[41] included the bond angles as well as the bond distances in their geometry-dependent atomic-charge-calculation method. This procedure led to a 3−5% error in the predicted dipole moments with 332 parameters to represent only alkanes, aldehydes, ketones, and amides, and did not treat aromatic or charged molecules. They also found that the charge flux along the bond is a major factor contributing to the geometry dependence of the atomic charge. We have assumed that all these geometric factors are correlated with each other, and that the bond distance reflects the differences in bond angles and torsional angles. This means that bond distances differ for different bond angles and torsional angles; for example, the bond distance between carbon and hydrogen in a methyl group in the eclipsed form of ethane differs from that in the staggered form. Inclusion of bond distances, which are correlated with other geometric factors such as bond angles and torsional angles, in a charge calculation helps in the development of a geometry-dependent charge distribution.

Two further examples illustrate this effect. One concerns the same atomic species in different environments, and the other concerns different atomic species in different environments. In the first example, the three hydrogens attached to carbon C_1 in Figure 1 have the same charge in the PEOE or MPEOE method (the values are not shown in the figure), since these methods depend on only one geometric factor, the connectivity. However, the bond distances in brackets, especially the one close to the oxygen, differ from each other. The MPA charges (from ab initio calculations) in parentheses show that all hydrogens have different charges. With a charge calculation method which depends only on the connectivity, as in the PEOE and MPEOE methods, there is no way to assign different charges to the hydrogens. In the second example, the *average* distance between the C and O atoms of the carbonyl group for molecules A1−A9 of Table 1 is 1.223 Å in amides, and differs from 1.210 Å in esters for molecules D1−D3 of Table 1 and 1.209 Å in acids

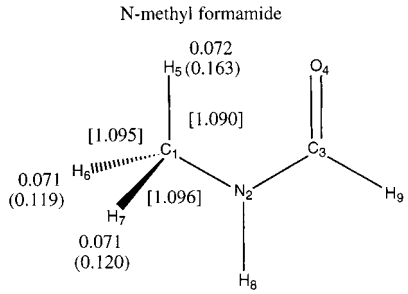

Figure 1. An example of *N*-methylformamide to show that the hydrogens in different environments have different charges (in electron units) in the GDAC method, even though the differences are small in this case. The values without parentheses are the charges from the GDAC method, and the values in parentheses are the charges from the MPA method. The values in brackets are the bond distances (Å) between carbon and hydrogens. Calculations of the geometry and the MPA charges were carried out with the B3LYP/6-31G** algorithm.

for molecules E1, E3, E5, and E7 of Table 1 (a few examples are shown in Figure 2). It is desirable to treat those carbonyl groups differently for each case by assigning different parameters, if one wants to obtain better results in calculating the charges. Instead of using different parameters for each case, the GDAC method uses only one type of parameter for C and O, respectively, in carbonyl groups. The distance between C and O takes into account the difference in the chemical environment.

A new charge-calculation scheme, including the geometric effect, is suggested in this paper on the basis of the PEOE and MPEOE schemes. Instead of using a fixed f_{AB} (0.5 in the PEOE, and varying parameters in the MPEOE method), a new treatment of f_{AB} as a function of the distance R_{AB} between atoms A and B in a bond is introduced. To describe the new f^0_{AB}, which is a function of the bond distance between A and B in the zeroth iteration, two things must be considered. First, the value should lie between 0 and 1; otherwise, convergence cannot be achieved with the PEOE scheme. Second, since atoms in shorter bonds have more orbital overlap than those in longer bonds, this effect is introduced by having a larger value of f^0_{AB} in a shorter bond. Therefore, the new f^0_{AB}, which is a function of R_{AB}, is

$$f^0_{AB} = \left(1 - \frac{R_{AB}}{R^{vdw}_A + R^{vdw}_B}\right) \tag{4}$$

where f^0_{AB} is the damping factor in the zeroth iteration and R^{vdw}_A and R^{vdw}_B are the van der Waals radii of atoms A and B, respectively. Three sets of van der Waals radii[44−46] are available and are listed in Table 2 for comparison. The values of the van der Waals radii are not a crucial factor for this equation as long as they are physically reasonable. The values of Rowland and Taylor[46] have been adopted here because they were obtained by a recent extensive search of 100 000 organic and organometallic crystal structures. The ratio between R_{AB} and $R^{vdw}_A + R^{vdw}_B$ represents the bond strength or orbital overlap between atoms A and B, which controls the charge flux between atoms A and B since a smaller value of this ratio corresponds to a larger value of f^0_{AB} and a larger value decays slower than a smaller value of f^0_{AB} in the iterative procedure. As the iteration proceeds in the calculations, f_{AB} decreases with increasing *n*. The damping factor in the *n*th iteration is given by $(f_{AB})^n$.

Since the new method has been developed for biomolecules such as peptides, 50 molecules representing the backbone and side chains of proteins were chosen in this work (Table 1). The atoms in the 50 molecules are classified into 22 atomic species

TABLE 1: Comparison of Experimental, ab Initio (B3LYP/6-31G), and GDAC Dipole Moments of 50 Molecules That Were Used in the Parametrization**

molecule	exptl[a]	ab initio[b]	GDAC	MPEOE	molecule	exptl[a]	ab initio[b]	GDAC	MPEOE
A1. formamide	3.714	3.8204	3.6253	3.9283	F1. methanethiol	1.53	1.7225	1.5926	1.3269
A2. *N*-methylformamide (cis)	3.82	3.7958	3.5996	4.0645	F2. ethanethiol	1.52	1.8144	1.6014	1.3712
(trans)[c]	–	3.8766	3.5739	4.0314	F3. dimethyl sulfide	1.50	1.7180	1.7797	1.7221
A3. *N*,*N*-dimethylformamide	3.85	3.8323	3.6738	3.8285	F4. ethyl methyl sulfide	1.560	1.7043	1.7994	1.6778
A4. acetamide	3.68	3.7733	3.6999	4.0325	F5. dimethyl disulfide		2.2460	2.2186	2.1595
A5. *N*-methylacetamide	3.75	3.7131	3.6112	4.2069					
A6. *N*,*N*-dimethylacetamide	3.80	3.6611	3.6770	3.9647	mean percent error[d] (%)			5.93	10.61
A7. urea	4.56	4.2485	4.1800	4.2645					
A8. methylurea	4.34	4.0918	4.1669	4.6961					
A9. malonamide	3.5	3.4140	2.9741	3.2512	G1. methylamine	1.238	1.4230	1.3361	1.2880
					G2. methylammonium ion		2.2761	2.1569	2.5101
mean percent error[d] (%)			3.99	6.49	G3. ethylamine	1.304	1.3953	1.3327	1.3291
					G4. ethylammonium ion		3.9561	3.7334	5.3996
B1. propane	0.083	0.0495	0.0022	0.0155	G5. propylamine		1.4365	1.3287	1.3095
B2. butane	0.09	0.0003	0.0009	0.0002	G6. propylammonium ion		6.4485	6.7438	8.4564
B3. isobutane		0.0543	0.0248	0.0187					
					mean percent error[d] (%)			5.59	16.83
mean percent error[d] (%)			119.04	60.20					
					H1. toluene		0.3419	0.2338	0.1377
C1. methanol	1.69	1.6633	1.5925	1.5318	H2. ethylbenzene		0.2967	0.2501	0.1580
C2. ethanol	1.441	1.5295	1.5873	1.4360					
C3. propanol	1.58	1.4617	1.5851	1.4503	mean percent error[d] (%)			23.66	53.20
C4. phenol		1.3354	1.2878	1.5450					
C5. methylphenol		1.2972	1.2635	1.5423	I1. guanidine		3.7858	3.7246	2.7075
					I2. guanidinium ion		0.0002	0.0010	0.0011
mean percent error[d] (%)			4.53	9.88	I3. methylguanidine		3.7262	3.7414	3.3869
					I4. methylguanidinium ion		1.5113	1.5040	5.7908
D1. methyl formate	1.77	1.8726	1.8098	1.0113					
D2. methyl ethanoate		1.7902	1.8679	1.4219	mean percent error[d] (%)			103.22	189.57
D3. methyl propanoate		1.7584	1.8276	1.3658					
					J1. imidazole	3.8	3.7034	3.6621	2.7985
mean percent error[d] (%)			3.88	29.63	J2. imidazolium ion		1.4134	1.3756	0.3190
					J3. methylimidazole		3.9075	3.8612	2.8906
E1. formic acid	1.415	1.4421	1.5355	1.3478	J4. methylimidazolium ion		2.5462	2.5675	0.2946
E2. formate ion		1.5553	1.5525	1.7164	J5. 3-methylindole		2.0399	1.9307	0.6396
E3. acetic acid	1.70	1.6070	1.6800	1.5454					
E4. acetate ion		3.9222	3.6005	3.6863	mean percent error[d] (%)			2.23	56.99
E5. propanoic acid	1.55	1.6998	1.6516	1.5355					
E6. propanoate ion		5.7391	5.7190	5.7914	total mean percent error[e] (%)			4.91	24.29
E7. butanoic acid		1.7613	1.6618	1.5352	root-mean-square error[f]			0.131	0.879
E8. butanoate ion		8.1536	8.4382	8.4998					
mean percent error[d] (%)			3.97	6.80					

[a] Taken from ref 48. [b] B3LYP/6-31G**. [c] Not included in the parameterization. [d] Mean percent error for all molecules in category *i*. Percent error = $(|\mu^{ab} - \mu^{calcd}|/\mu^{ab}) \times 100$. Total mean percent error = $[\sum_{i=1}^{n}(\text{percent error of catagory } i)]/n$, where n is the number of categories. [e] Molecules with total dipole moments smaller than 0.1 D were not included in the calculation. See the text for details. [f] The root-mean-square error in debye units. rms = $[\sum_i^n(\mu^{ab}{}_i - \mu^{calcd}{}_i)^2/n]^{1/2}$, where i ranges from 1 to 50 and n is the number of molecules, 50.

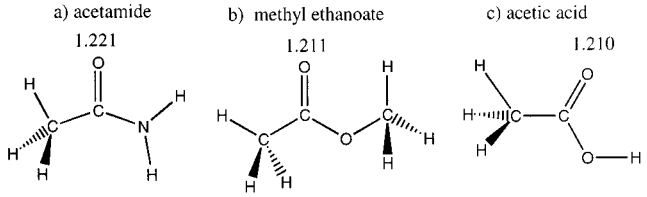

Figure 2. The bond distances (Å) between carbon and oxygen in the carbonyl group in amides, esters, and acids differ from each other. Calculations of the geometry were carried out with the B3LYP/6-31G** algorithm.

TABLE 2: Van der Waals Radii

atom	Pauling[a] (Å)	Bondi[b] (Å)	Rowland and Taylor[c] (Å)
H	1.20	1.20	1.10
C	1.70	1.70	1.77
N	1.50	1.55	1.64
O	1.40	1.52	1.58
S	1.85	1.80	1.81

[a] Taken from ref 44. [b] Taken from ref 45. [c] Taken from ref 46.

according to their atomic number and valence and charged states. The atomic species and the optimized parameters for a_i and b_i are listed in Table 3. Some examples of the classification are shown in Figure 3.

Experimental dipole-moment data for gas-phase charged molecules are usually not available. Since the correlation between the geometry and the dipole moments of the molecules is a crucial point in this work, ab initio molecular orbital calculations were carried out to obtain the geometries and dipole moments with the 6-31G** basis set at the level of B3LYP theory for neutral and charged molecules. The calculations were carried out with the Gaussian 94 program.[47] Those experimental dipole moments which are available,[48] and all ab initio dipole-moment data, are listed in Table 1. The 6-31G** basis set at the level of B3LYP theory was chosen because it leads to good agreement with experimental data, and the comparison is plotted in Figure 4.

In the MPEOE method, once the parameters were obtained for neutral models for polypeptides, the method was extended

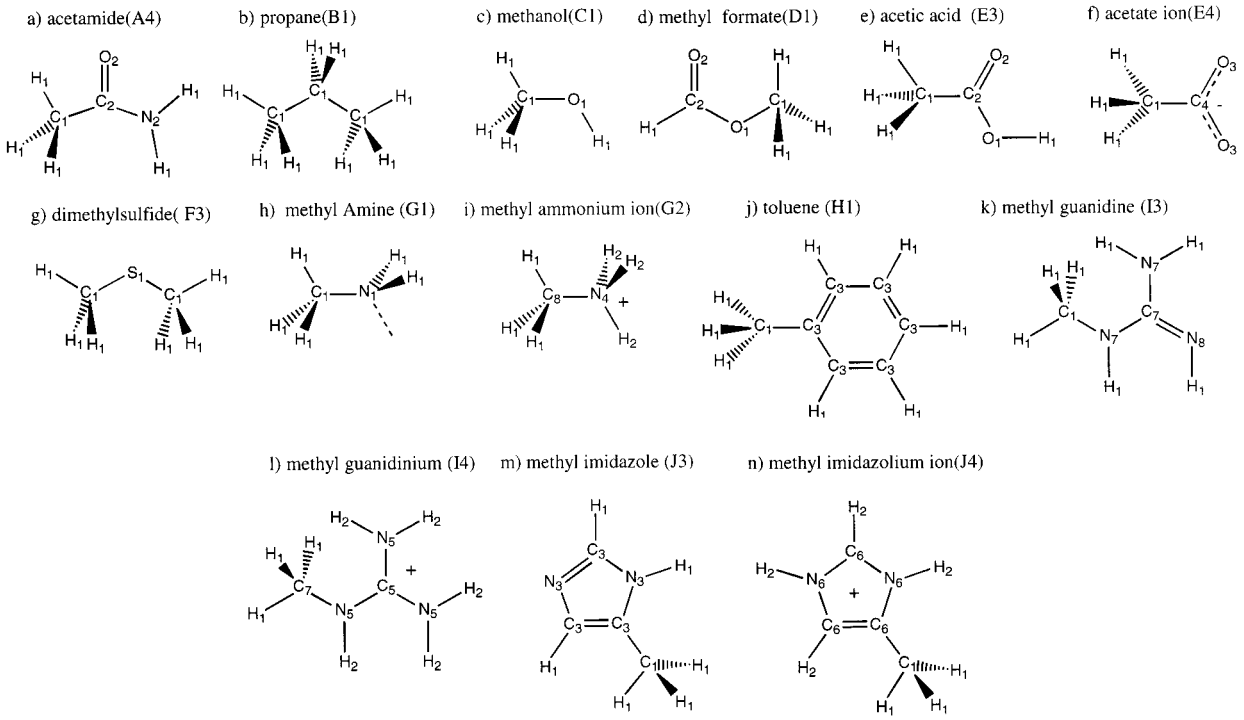

Figure 3. Some examples of the classification of atomic species. The subscripts correspond to the subscripts in Table 3.

TABLE 3: Classified Atomic Species[a] and the Optimized Parameters

atomic species	a	b
H_1 (neutral)	2.648	0.429
H_2 (H atom connected to N^+)	2.435	1.096
C_1 (sp^3)	2.810	0.432
C_2 (sp^2)	2.382	22.999
C_3 (in aromatic ring)	2.648	8.602
C_4 (in $-CO^{2-}$)	3.037	−1.544
C_5 (in Arg^+ side chain)	2.396	18.813
C_6 (in His^+ side chain)	1.121	13.101
C_7 (in Arg side chain)	4.883	52.688
C_8 (in $-C-NH_3^+$)	5.637	11.460
N_1 (sp^3)	2.811	−2.470
N_2 (sp^2)	3.093	−0.810
N_3 (in aromatic ring)	5.547	4.922
N_4 (in NH_4^+)	1.969	2.628
N_5 (in Arg^+)	2.842	0.825
N_6 (in His^+)	3.731	−2.832
N_7 (three-bonded N in Arg side chain)	3.126	−1.097
N_8 (two-bonded N in Arg side chain)	4.071	3.830
O_1 (sp^3 in C−O−H or C−O−C)	3.353	0.953
O_2 (sp^2 in C=O)	5.727	0.646
O_3 (in CO_2^-)	3.957	2.180
S_1 (all neutral cases)	3.354	1.712
number of *a* and *b* parameters	44[b] (40)[c]	
number of damping factors	0[b] (13)[c]	
total number of parameters	44[b] (53)[c]	

[a] See Figure 3 for some examples of the classification of atomic species in the GDAC method. [b] This is the total number for the 22 classified atomic species. The damping factor is not a separate parameter in the GDAC method because it is obtained directly from the bond and van der Waals distances by eq 4. [c] Values in parentheses are for MPEOE. For the classifications of atomic species, see refs 25 and 37.

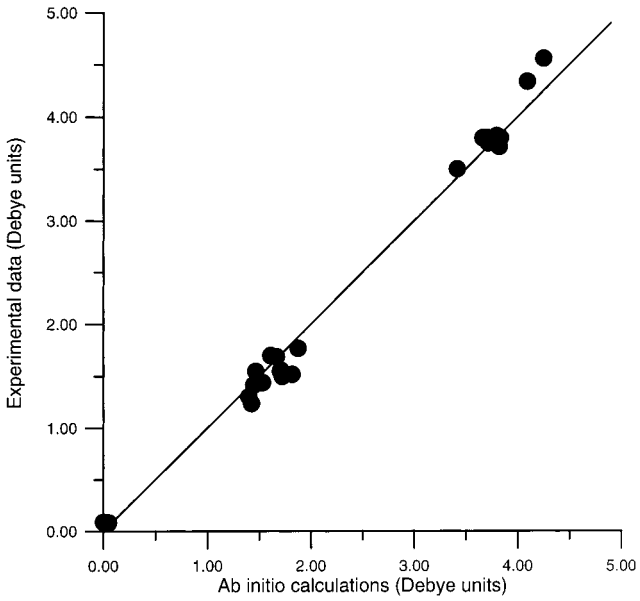

Figure 4. Comparison of dipole moments from ab initio molecular orbital calculations (B3LYP/6-31G**) and from experiment. The correlation coefficient is 0.9986.

to ionic and aromatic molecules, then halogen-containing molecules, and so on. In that procedure, the parameters for *new* atomic species were optimized while the previously obtained parameters were fixed. With such a procedure, the method may represent the total dipole moment well, but it is difficult to represent the three components of each dipole moment since it inherits the errors from the previous procedures. The MPEOE method led to larger errors for ionic and aromatic molecules which were parametrized after the parameters for normal neutral molecules such as amides and acids were fixed. To overcome this shortcoming, all parameters in the GDAC method were optimized simultaneously to avoid any inheritance of errors. This approach, however, suffers from another problem, viz., the multiple-minimum problem. It is difficult to find the best set of parameters if a large number of parameters are involved in the optimization procedure since the procedure is more likely to be trapped in a local minimum. Since the method is based on the electronegativity equalization method, and electronegativity is not a property of atoms in their ground states, but of atoms under the same conditions in which they are found in

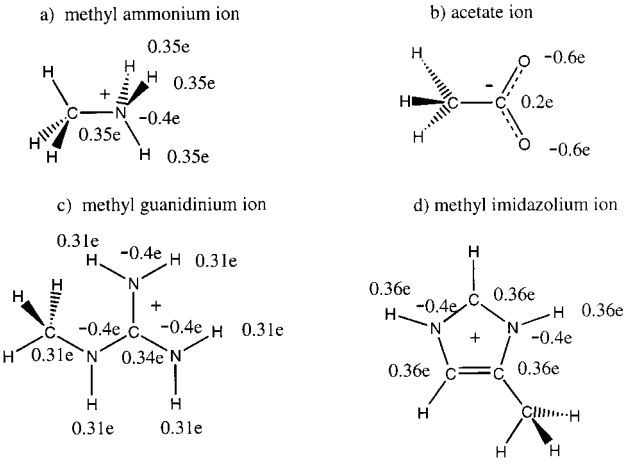

Figure 5. Arbitrary initial charges for charged molecules in the GDAC method. The charges on the atoms without values were set to zero. Atomic species in the same category have the same values, and a few examples are shown in the figure.

molecules, i.e., the valence state,[8a] we introduced as few types of atoms as possible while still taking into account their environment.

For charged species, the initial net atomic charges were assigned arbitrarily to the atoms prior to the calculation because the method required an assignment of initial charges before the iteration procedure was begun. It is more reasonable to distribute the charge around the atoms rather than assign it to a single atom since the excess charges for negatively charged molecules and the reduced number of charges for positively charged molecules are distributed among the atoms in a molecule rather than being localized on one atom. Examples of the assignment of the initial charges are shown in Figure 5. Even though this assignment of the initial charge is arbitrary, the effect of the choice of the initial charge on the final charge is very small because the parameters were determined on the basis of initial charges that were fit to the dipole moments. No et al.[25b] also tried several initial charges for ionic molecules in the MPEOE method and found that the effect of the choice of the initial charge is very small.

The parameters a_i and b_i were obtained during the optimization procedure to minimize the function F, using SUMSL (secant unconstrained minimization solver):[49]

$$F = W_\mu \sum_i \sum_j [\mu^{\mathrm{ab}}{}_{i,j} - \mu^{\mathrm{calcd}}{}_{i,j}(\alpha,r)]^2 + W_{\mathrm{M}} \sum_i \sum_k [q^{\mathrm{ab}}{}_{i,k} - q^{\mathrm{calcd}}{}_{i,k}(\alpha,r)]^2 \quad (5)$$

where the superscript "ab" represents the ab initio molecular orbital calculations and "calcd" represents the calculated values. μ and q represent dipole moments and charges, respectively. q^{ab} represents the charges from the MPA (ab initio) method. The index i represents molecules 1−50 in Table 1, and index j (from 1 to 4) corresponds to the three components of the dipole moments and the total dipole moment for each species. The index k represents the atoms in a molecule. The W values are weight factors. The α represents the parameters a_i and b_i, and r represents the bond length. Equation 5 indicates that dipole moments (and their components) and MPA charges from ab initio molecular orbital calculations are used as constraints in the optimization procedure. The MPA charges were used as constraints because the optimization with dipole moment alone more easily leads to entrapment in local minima. The same value (i.e., 1.0) was used for W_μ and W_{M} at the beginning of the optimization, and W_{M} was decreased gradually to one-tenth of W_μ in the final stage of optimization.

To improve the quality of the minimization and to circumvent the multiple-minimum problem mentioned above, we started with as few as eight atomic species which correspond to the sp^2 and sp^3 hybridization states of carbon, nitrogen, and oxygen and only one hybridization state of hydrogen and sulfur, respectively. Then the number of atomic species was extended, step by step. For example, in the first stage, nitrogen was represented by only two species, viz., $N_1(sp^3)$ and $N_2(sp^2)$. In the next stage, N_2 was split into $N_2(sp^2)$ and N_3(in an aromatic ring), and so on. In that case, the initial values for N_3 were taken from the values of N_2 of the previous optimization. The final stage included 22 atomic species, requiring 44 parameters according to their chemical environments (see Table 3). The procedure was repeated with various weight factors, W, and various orders of classification until it reached a minimum value of F that gave good agreement between GDAC and ab initio dipole moments.

The GDAC parameters were fitted to the dipole moment surface rather than to the ab initio electrostatic potential surface because electrostatic potentials depend on the basis set used, and there are no experimental data available to verify the quality of the electrostatic potential from an ab initio calculation. However, experimental data are available for dipole moments and their components, which can be compared with those from ab initio calculations.

Unlike net atomic charges, the molecular electrostatic potential is a defined quantum mechanical property. Since the density matrix can be calculated accurately by solving the Schrödinger equation, the electrostatic potential $V(\mathbf{r})$ can be obtained easily. The electrostatic potential energy (for a unit positive charge at point $\mathbf{r}$) in the vicinity of the given molecule is calculated from the following equation:[50]

$$V(\mathbf{r}) = \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} - \sum_{mn} P_{mm} \int \frac{\chi_m^* \chi_n}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}\mathbf{r}' \quad (6)$$

Z_A is the charge of the Ath nucleus (e) located at $\mathbf{R}_A$. P_{mn} refers to the (mn)th element of the electron density matrix of the ground-state wave function for the molecule as obtained from an SCF-MO (LCAO) calculation. χ_m^* and χ_n are members of the atomic orbital basis set. The first summation runs over the nuclei of the molecule. The summations over m and n are over the atomic orbitals, and the integration over $\mathbf{r}'$ covers all space.

Alternatively, the electrostatic potential at point $\mathbf{r}$, with discrete charges, q_A, such as GDACs, MPEOE charges, and MPA charges, located at $\mathbf{R}_A$, can be calculated classically with the following equation:

$$V(\mathbf{r}) = \sum_A \frac{q_A}{|\mathbf{r} - \mathbf{R}_A|} \quad (7)$$

Comparison of the electrostatic potential between ab initio (eq 6) and classical (eq 7) calculations with discrete charges such as GDACs, MPEOE charges, and MPA charges is reported in the Results and Discussion. A total of 961 grid points were selected for each calculation on the $x-y$ plane from −6.0 to +6.0 Å with an interval of 0.4 Å.

3. Results and Discussion

Since we optimized a_i and b_i rather than take them from experimental values of ionization potentials and electron af-

TABLE 4: Comparison of ab Initio (B3LYP/6-31G), GDAC, and MPEOE Dipole Moment Components of the 50 Molecules Used To Determine the *a* and *b* Parameters of Eq 3**

	ab initio			GDAC			MPEOE		
	X	Y	Z	X	Y	Z	X	Y	Z
A1	3.7488	0.7360	0.0016	3.3512	0.8210	0.0014	3.7940	1.0182	0.0015
A2	−2.8559	2.5004	0.0030	−2.5412	2.5494	0.0042	−2.1793	3.4308	0.0068
A3	3.8172	−0.3406	0.0000	3.6316	−0.5556	0.0001	3.7084	−0.9513	0.0001
A4	−0.1922	−3.7684	0.0001	−0.1349	−3.6974	0.0002	−0.1223	−4.0306	0.0001
A5	0.6561	−3.6547	0.0002	0.5950	−3.5619	0.0002	−0.0506	−4.2066	0.0002
A6	1.8040	3.1858	0.0005	1.8885	3.1550	0.0020	1.7237	3.5704	0.0052
A7	0.0001	−4.2485	0.0000	0.0001	−4.1800	0.0000	0.0001	−4.2645	0.0000
A8	−0.1105	−4.0903	0.0002	−0.0662	−4.1664	0.0002	−0.7440	−4.6368	0.0003
A9	3.3354	−0.7283	0.0045	2.8530	−0.8169	0.1962	3.1533	−0.7886	0.0719
B1	0.0000	0.0492	0.0005	0.0001	−0.0022	0.0004	0.0000	−0.0155	0.0000
B2	0.0001	−0.0002	0.0002	0.0002	−0.0005	0.0007	0.0000	−0.0001	0.0001
B3	0.0258	−0.0196	−0.0437	−0.0113	−0.0218	0.0027	−0.0035	0.0092	0.0160
C1	0.7338	−1.4927	−0.0005	1.0104	−1.2309	−0.0004	0.1782	−1.5214	−0.0003
C2	−0.2569	1.5078	0.0018	0.0770	1.5854	0.0021	−0.6601	1.2753	0.0019
C3	0.0678	1.4597	0.0359	0.3855	1.5372	0.0321	−0.3900	1.3965	0.0316
C4	−0.0929	1.3321	0.0002	−0.2632	1.2606	0.0002	0.0735	1.5432	0.0002
C5	0.2256	1.2774	0.0004	−0.0174	1.2633	0.0002	0.0786	1.5403	0.0003
D1	−1.0316	1.5628	0.0000	−0.6464	1.6904	−0.0000	−0.1560	0.9992	0.0000
D2	−0.4357	−1.7364	0.0000	−0.3705	−1.8308	0.0000	0.3335	−1.3822	0.0000
D3	0.4280	−1.6985	0.1542	0.2465	−1.8104	0.0444	−0.3336	−1.2894	0.3028
E1	−1.3825	−0.4102	0.0000	−1.5333	−0.0814	0.0000	−1.2832	−0.4109	0.0000
E2	0.0008	1.5553	0.0000	−0.0003	1.5525	0.0000	0.0002	1.7164	0.0000
E3	0.5923	−1.4938	0.0002	0.6627	−1.5437	0.0001	0.4725	−1.4714	0.0001
E4	−3.9193	−0.1529	0.0000	−3.5986	−0.1174	−0.0000	−3.6832	−0.1515	−0.0001
E5	−1.1882	1.1490	0.3969	−1.1687	1.0581	0.4925	−0.9784	1.1183	0.3869
E6	5.4727	−1.7283	−0.0152	5.4562	−1.7136	−0.0092	5.5143	−1.7700	−0.0075
E7	−1.1617	1.2877	0.3072	−1.0534	1.1946	0.4739	−0.8588	1.2134	0.3834
E8	−8.1009	−0.9252	0.0379	−8.3728	−1.0482	0.0193	−8.4281	−1.1017	0.0095
F1	1.4229	0.9709	0.0000	1.1074	1.1447	0.0000	0.9127	0.9631	0.0000
F2	1.6305	0.0683	0.7931	1.2617	−0.1355	0.9768	1.0924	−0.1856	0.8077
F3	0.0000	1.7180	0.0001	0.0001	1.7797	0.0000	−0.0000	1.7221	0.0000
F4	0.0360	1.7040	−0.0009	−0.1339	1.7944	−0.0007	−0.0574	1.6769	−0.0010
F5	−0.0005	2.2460	−0.0010	−0.0007	2.2186	−0.0006	−0.0007	2.1595	−0.0006
G1	−0.1903	−0.0008	1.4120	−0.2349	−0.0012	1.3153	0.8489	−0.0008	0.9686
G2	2.2761	−0.0003	0.0003	2.1569	−0.0005	0.0003	2.5101	−0.0003	0.0002
G3	−1.0206	−0.9514	−0.0001	−0.9717	−0.9120	−0.0002	0.0359	−1.3286	−0.0001
G4	−3.9002	−0.6631	0.0000	−3.6573	−0.7503	0.0000	−5.3559	−0.6856	0.0000
G5	−0.8778	1.1371	0.0020	−0.8058	1.0565	0.0016	0.2429	1.2855	0.0017
G6	−6.4415	−0.3021	−0.0002	−6.7290	−0.4464	0.0001	−8.4558	−0.1037	−0.0003
H1	0.3395	−0.0006	0.0405	0.2337	−0.0000	0.0073	0.1376	−0.0001	0.0045
H2	0.2966	0.0000	−0.0080	0.2460	−0.0001	−0.0451	0.1555	0.0000	0.0282
I1	−0.1339	−3.6120	0.0000	−1.0252	−3.5807	0.0000	−1.6185	−2.1705	0.0000
I2	0.0000	−0.0002	0.0000	−0.0002	−0.0010	0.0000	−0.0002	−0.0011	0.0000
I3	−3.6620	0.6887	0.0000	−3.6821	0.6638	0.0000	−2.7740	1.9432	−0.0000
I4	−0.0616	1.5101	0.0000	0.3463	1.4636	0.0000	0.3825	5.7781	0.0000
J1	1.1140	3.5318	0.0000	0.9504	3.5366	0.0000	1.5278	2.3446	0.0000
J2	−1.3444	0.4362	0.0000	−1.3085	0.4243	0.0000	−0.3035	0.0984	0.0000
J3	2.5510	−2.9599	0.0006	2.3522	−3.0620	0.0008	2.2658	−1.7950	0.0009
J4	2.3802	−0.9043	−0.0003	2.3992	−0.9142	−0.0003	0.1692	−0.2412	−0.0004
J5	−0.9200	1.8207	0.0007	−0.5466	1.8517	0.0003	−0.3267	0.5499	0.0005

finities as in the PEOE method, we cannot attribute physical meaning to these parameters. All parameters are correlated, and according to their values and the damping factors, the charges are transferred through the bonds.

The designated atomic species, optimized parameters, and number of parameters are listed in Table 3. The values in parentheses pertain to the MPEOE method. The GDAC method uses 44 parameters while the MPEOE method uses 53 including 13 damping factors.

Table 1 shows the calculated dipole moments with the GDAC method, compared with experimental data, and data from ab initio molecular orbital calculations and the MPEOE method. The mean percent errors are presented in the table after each category, such as amide, alcohol, acid, and so on. The total mean percent error does not include the molecules whose total dipole moments are smaller than 0.1 D. The results from the GDAC method are in good agreement with the ab initio results, within 5% error (Table 1), while the MPEOE method involves an error of 24.29%. The root-mean-square (rms) errors are given at the end of Table 1.

Since this method is based on the net atomic charge approximation, the charges are placed on the nuclear centers. Nuclear-centered charges do nevertheless suffer from some drawbacks when applied to systems in which the charge density about each atom is not spherically symmetrical. In particular, poorer results are obtained when the method is applied to large conjugated systems such as those in categories H and I in Table 1. (The large mean percent error for the molecules in category B in Table 1 is due to the small dipole moment.) This limitation pertains to all methods that use the net atomic charge approximation.

The components of the dipole moments from the GDAC and MPEOE methods are compared in Table 4 with those from the ab initio molecular orbital calculations that were used to

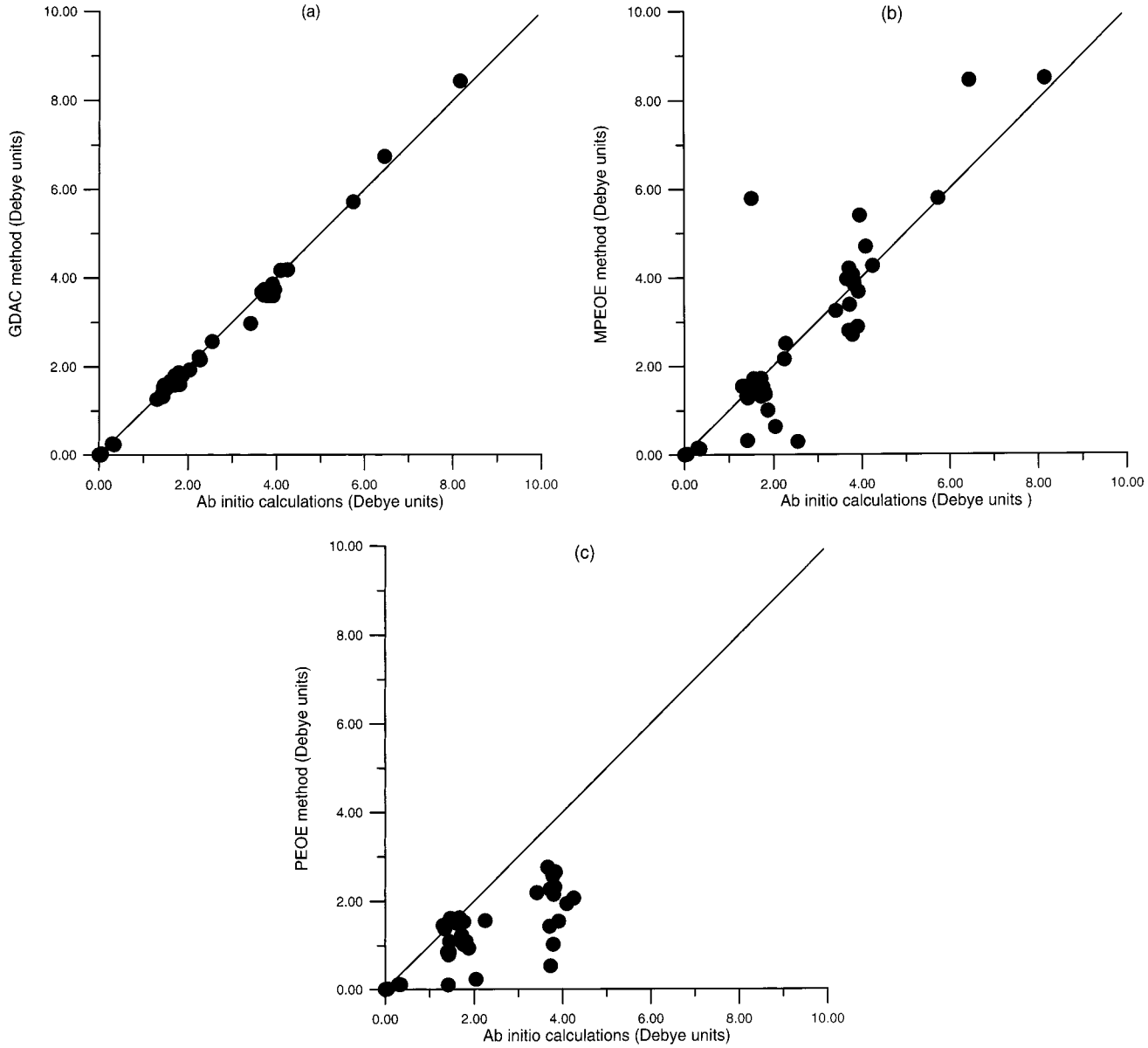

Figure 6. Comparison of total dipole moments from empirical (GDAC, MPEOE, PEOE) and ab initio molecular orbital calculations: (a) GDAC vs ab initio (the correlation coefficient is 0.9990); (b) MPEOE vs ab initio (the correlation coefficient is 0.9599); (c) PEOE vs ab initio (the correlation coefficient is 0.9342).

determine *a* and *b* of eq 3. While these components were used in the parametrization (eq 5), it is difficult to draw conclusions from a comparison of the ab initio values with those from the GDACs or MPEOE charges, because the agreement differs among the various components.

Figure 6a shows a comparison between the dipole moments obtained from the ab initio calculation and those obtained by the GDAC method. Good agreement is seen with ab initio dipole moments, better than MPEOE (Figure 6b) and much better than PEOE (Figure 6c). Since the original PEOE method had been developed only for neutral molecules, only 39 neutral molecules out of the 50 molecules of Table 1 are shown in Figure 6c.

To see how the GDAC method reflects the geometric effect, we show *N*-methylformamide (NMF) in Figure 1 as an example. Hydrogens H_5, H_6, and H_7 attached to C_1 are treated as equivalent in the PEOE and MPEOE methods because those methods consider only connectivity. But the ab initio calculation shows that H_5 differs from H_6 and H_7 in bond distance (in brackets) and MPA ab initio charge (in parentheses). For the values of the charge with the GDAC method, without parentheses, H_5 is different from H_6 and H_7 even though the difference is small. It should be noted that the MPA charges are not a proper set because they reproduce dipole moments and electrostatic potential very poorly. However, they serve to show how the charge depends on environment.

The electrostatic potential of acetamide, calculated with the GDACs and MPEOE, PEOE, and MPA charges, is compared with the ab initio electrostatic potential calculated with the 6-31G** basis set at the level of B3LYP in Figure 7. Acetamide was chosen for comparing the electrostatic potential because it can serve as a model of the backbone of proteins and it is relatively flat so that the contour map can be drawn on a 2D plane. The shapes of the contours rather than the exact values of the electrostatic potential seem to provide a better comparison because the atomic charges derived by fitting the dipole moment surface will not reproduce the electrostatic potential as well as potential-derived charges do.

The GDAC results are apparently better than the MPA or PEOE results, but are comparable with the MPEOE results. For example, the 0.00 kcal/mol contours of the electrostatic potential

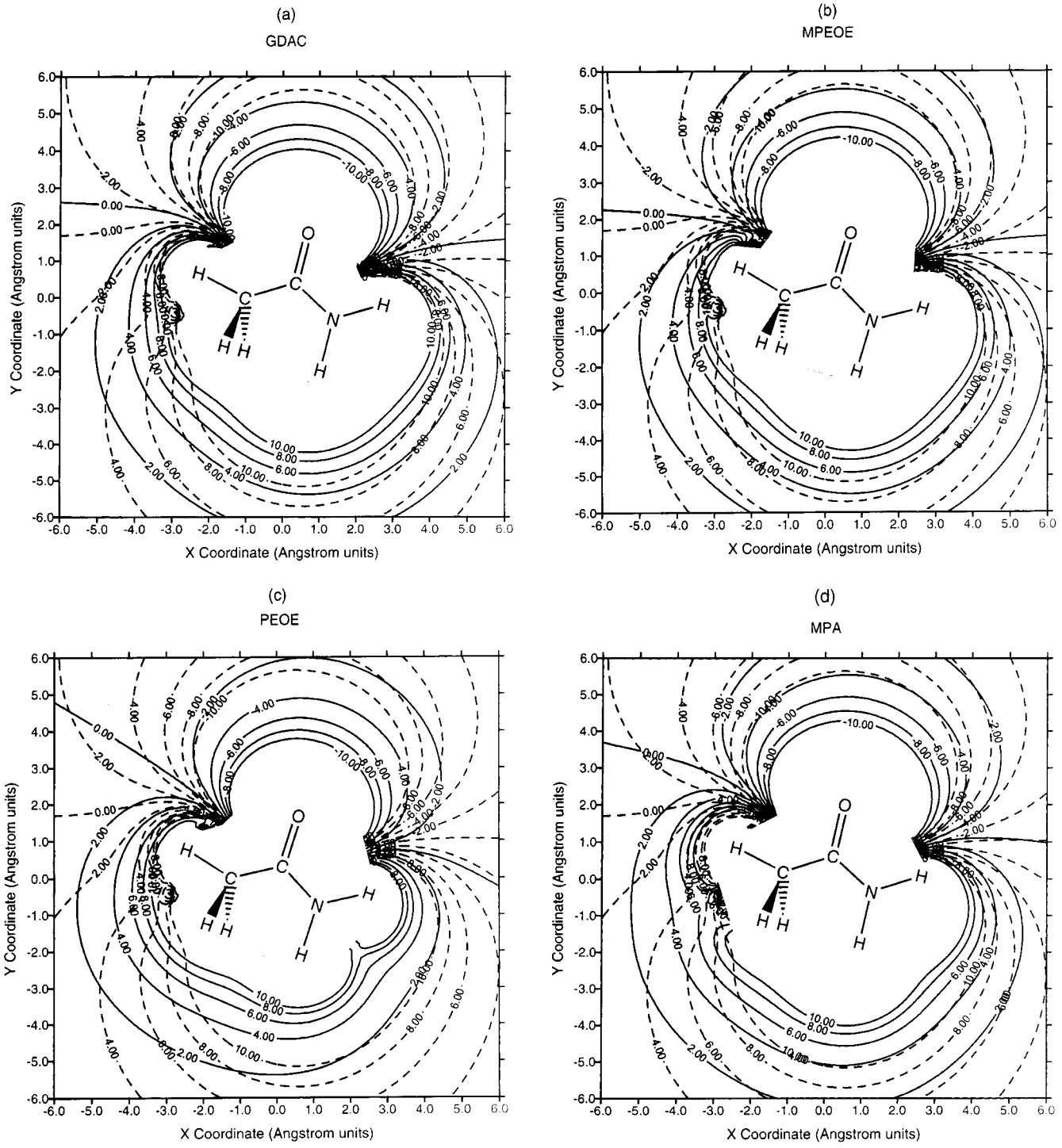

Figure 7. Comparison of the electrostatic potential surface (kcal/mol) of acetamide from the various methods (GDAC, MPEOE, PEOE, and MPA) (solid lines) with the result from the ab initio molecular orbital calculations (B3LYP/6-31G**) (dashed lines). A total of 961 grid points were selected for each calculation on the $x-y$ plane from -6.0 to $+6.0$ Å with an interval of 0.4 Å.

from the GDAC and MPEOE methods agree better than the contours from the PEOE and MPA methods with those from the ab initio calculation. Even though the result from the MPEOE method is quite comparable with that from the GDAC method for acetamide, the components of the dipole moment for the GDAC method (-0.1349, -3.6974, $+0.0002$) agree better with those from the ab initio calculations (-0.1922, -3.7684, $+0.0001$) than do those from the MPEOE method (-0.1223, -4.0306, $+0.0001$), as shown in Table 4.

Since our goal is to develop a charge-calculation method that can describe any redistribution of charge density upon assembly of amino acid residues to form a large polypeptide molecule, the GDAC method has been applied to seven conformations of *N*-acetylalanine-*N*′-methylamide (Ac-Ala-NHMe), which were not included in the parametrization. The dipole moments are compared with the results from ab initio molecular orbital calculations and the MPEOE method in Table 5. The GDAC results are better than the MPEOE results not only in magnitude but also in relative order. Even though the overall GDAC results are better than the MPEOE results, the GDAC dipole moment of C_5 does not agree very well with that from the ab inito calculation. This small discrepancy in charges can be compen-

TABLE 5: Dipole Moments of Ac-Ala-NHMe[a]

conformation	φ, ψ[b]	ab initio[c]	GDAC	MPEOE
C_7^{eq}	−82, 72	2.85	2.62	3.93
C_5	−157, 165	3.08	2.56	1.83
C_7^{ax}	74, −60	3.54	3.13	4.49
β_2	−136, 23	4.63	4.75	5.89
α'	−170, −38	4.89	5.12	5.80
α_L	68, 25	5.65	6.07	7.27
α_R	−60, −40	5.98	6.25	7.59

[a] In Debye units. [b] In degrees. All values of φ and ψ were obtained by optimization with the B3LYP/6-31G** method. [c] B3LYP/6-31G** calculations.

sated by introducing proper nonbonded and hydrogen-bonding potentials in a molecular mechanics force field to predict the geometry and the energy related to the geometry.

The introduction of the GDAC calculation method into molecular force fields requires very small modification. Since this method uses a damping factor which depends on the bond distances, the bond distances must be provided prior to the charge calculation. The bond distances can be obtained in conjunction with a force field calculation or a quantum mechanical calculation. A force field such as ECEPP[1,4] that uses rigid geometry needs only one charge calculation at the beginning, whereas, for a force field using flexible geometry,[2,3,5] the charge calculation can be carried out right after the changes in geometry are calculated, to take full advantage of this method.

A flexible-geometry force field has advantages over a rigid-geometry force field, because the GDAC method takes into full account the variation of bond distances between atoms. However, even in the rigid-geometry force fields, the GDAC method gives reasonable charges as long as it starts with a reliable geometry. New ECEPP parameters have been derived with the GDACs, and the result will be presented elsewhere.

4. Conclusions

The GDAC method described here provides a simple and direct approach for calculating molecular charge distributions in proteins. No assumptions are made concerning the transferability of net atomic charges between molecular fragments: hence, the method accounts in an approximate way at least for the redistribution of charge density upon the assembly of amino acid residues to form a large polypeptide molecule. When the PEOE formalism is used, the charge density is assumed to depend only on the connectivity and not on the geometry. In the GDAC method, distance-dependent damping factors were used to overcome the limitation of the PEOE scheme. Introduction of distance-dependent damping factors helps to include the effect of geometry variation of the same atomic species but in different environments in a molecule and to reduce the number of parameters required to represent the different atomic species. The dipole moments calculated with the GDAC method agree with those from ab initio molecular orbital calculations within a 5% error (Table 1). The GDAC method was also more successful in reproducing the dipole moments of seven conformers of *N*-acetylalanine-*N*′-methylamide (Ac-Ala-NHMe) than the MPEOE method. The GDAC method leads to the direct calculation of atomic charges requiring negligible computational resources, after the parameters have been derived.

Acknowledgment. This work was supported by grants from the U.S. National Science Foundation (MCB95-13167) and from KISTEP/MOST, Korea (1999). The computations were carried out on the IBM SP2 supercomputer of the Cornell Theory Center (CTC), which is funded in part by the National Center for Research Resources of the National Institutes of Health (Grant P41RR-04293), and CTC's Advanced Cluster Computing Consortium.

Supporting Information Available: Tables of atomic species, optimized geometries, and GDACs. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361.

(2) Kollman, P.; Caldwell, J. W.; Ross, W. S.; Pearlman, D. A.; Case, D. A.; DeBolt, S.; Cheatham, T. E., III; Ferguson, D.; Seibel, G. AMBER: A Program for Simulation of Biological and Organic Molecules. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Editor-in-Chief; John Wiley & Sons: New York, 1998; Vol. 1, pp 11−13.

(3) MacKerell, A. D., Jr.; Brooks, B.; Brooks, C. L., III; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. CHARMM: The Energy Function and Its Parameterization. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Editor-in-Chief; John Wiley & Sons: New York, 1998; Vol. 1, pp 271−277.

(4) Ripoll, D. R.; Scheraga, H. A. ECEPP: Empirical Conformational Energy Program for Peptides. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Editor-in-Chief; John Wiley & Sons: New York, 1998; Vol. 2, pp 813−815.

(5) Jorgensen, W. L. OPLS Force Fields. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Editor-in-Chief; John Wiley & Sons: New York, 1998; Vol. 3, pp 1986−1989.

(6) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.

(7) Iczkowski, R. P.; Margrave, J. L. *J. Am. Chem. Soc.* **1961**, *83*, 3547.

(8) Hinze, J.; Jaffé, H. H. *J. Am. Chem. Soc.* **1962**, *84*, 540. (b) Hinze, J.; Whitehead, M. A.; Jaffé, H. H. *J. Am. Chem. Soc.* **1963**, *85*, 148.

(9) Huheey, J. E. *J. Phys. Chem.* **1965**, *69*, 3284. (b) Huheey, J. E. *J. Phys. Chem.* **1966**, *70*, 2086.

(10) Sanderson, R. *Chemical Bonds and Bond Energy*; Academic Press: New York, 1976; p 15.

(11) Momany, F. A. *J. Phys. Chem.* **1978**, *82*, 592.

(12) Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.

(13) Cox, S. R.; Williams, D. E. *J. Comput. Chem.* **1981**, *2*, 304.

(14) Guillen, M. D.; Gasteiger, J. *Tetrahedron* **1983**, *39*, 1331.

(15) Dősen-Mićović, L.; Jeremić, D.; Allinger, N. L. *J. Am. Chem. Soc.* **1983**, *105*, 1716.

(16) Lavery, R.; Zakrzewska, K.; Pullman, A. *J. Comput. Chem.* **1984**, *5*, 363.

(17) Abraham, R. J.; Hudson, B. *J. Comput. Chem.* **1985**, *6*, 173.

(18) (a) Mortier, W. J.; Genechten, K. V.; Gasteiger, J. *J. Am. Chem. Soc.* **1985**, *107*, 829. (b) 24) Mortier, W. J.; Ghosh, S. K.; Shankar, S. *J. Am. Chem. Soc.* **1986**, *108*, 4315.

(19) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894.

(20) Abraham, R. J.; Smith, P. E. *J. Comput. Chem.* **1987**, *9*, 288.

(21) Hammarström, L.-G.; Liljefors, T.; Gasteiger, J. *J. Comput. Chem.* **1988**, *9*, 424.

(22) Mullay, J. *J. Comput. Chem.* **1988**, *9*, 399.

(23) Kim, S.; Jhon, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1988**, *92*, 7216.

(24) (a) Dinur, U.; Hagler, A. T. *J. Chem. Phys.* **1989**, *91*, 2949. (b) Dinur, U.; Hagler, A. T. *J. Chem. Phys.* **1989**, *91*, 2959.

(25) (a) No, K. T.; Grant, J. A.; Scheraga, H. A. *J. Phys. Chem.* **1990**, *94*, 4732. (b) No, K. T.; Grant, J. A.; Jhon, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1990**, *94*, 4740.

(26) Ferenczy, G. G.; Reynolds, C. A.; Richards, W. G. *J. Comput. Chem.* **1990**, *11*, 159.

(27) Woods, R. J.; Khalil, M.; Pell, W.; Moffat, S. H.; Smith, V. H., Jr. *J. Comput. Chem.* **1990**, *11*, 297.

(28) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.

(29) Luque, F. J.; Illas, F.; Orozco, M. *J. Comput. Chem.* **1990**, *11*, 416.

(30) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.

(31) Colonna, F.; Angyan, J. G.; Tapia, O. *Chem. Phys. Lett.* **1990**, *172*, 55.

(32) Ferenczy, G. G. *J. Comput. Chem.* **1991**, *12*, 913.

(33) Stouch, T. R.; Williams, D. E. *J. Comput. Chem.* **1992**, *13*, 622.

(34) Aida, M.; Corongiu, G.; Clementi, E. *Int. J. Quantum Chem.* **1992**, *42*, 1353.

(35) Chipot, C.; Maigret, B.; Rivail, J.-L.; Scheraga, H. A. *J. Phys. Chem.* **1992**, *96*, 10276.

(36) Chipot, C.; Ángyán, J. G.; Ferenczy, G. G.; Scheraga, H. A. *J. Phys. Chem.* **1993**, *97*, 6628.

(37) Park, J. M.; No, K. T.; Jhon, M. S.; Scheraga, H. A. *J. Comput. Chem.* **1993**, *14*, 1482.
(38) Stouch, T. R.; Williams, D. E. *J. Comput. Chem.* **1993**, *14*, 858.
(39) Colonna, F.; Evleth, E. M. *Chem. Phys. Lett.* **1993**, *212*, 665.
(40) Park, J. M.; Kwon, O. Y.; No, K. T.; Jhon, M. S.; Scheraga, H. A. *J. Comput. Chem.* **1995**, *16*, 1011.
(41) Dinur, U.; Hagler, A. T. *J. Comput. Chem.* **1995**, *16*, 154.
(42) Kushwaha, P. S.; Kumar, A.; Mishra, P. C. *Int. J. Quantum Chem.* **1999**, *74*, 271.
(43) Banks, J. L.; Kaminski, G. A.; Zhou, R.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **1999**, *110*, 741.
(44) Pauling, L. *The Nature of the Chemical Bond*; Cornell University Press: New York, 1948; pp 187−193.
(45) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.
(46) Rowland, R. S.; Taylor, R. *J. Phys. Chem.* **1996**, *100*, 7384.
(47) Gaussian 94, Revision D.2: Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A., Gaussian, Inc., Pittsburgh, PA, 1995.
(48) (a) Starck, B. In *Landolt-Börnstein New Series*; Hellwege, K.-H., Hellwege, A. M., Eds.; Springer-Verlag: New York, 1967; Vol. II/4. (b) Demaison, J.; Dubrulle, A.; Hüttner, W.; Tiemann, E. In *Landolt-Börnstein New Series*; Hellwege, K.-H., Hellwege, A. M., Eds.; Springer-Verlag: New York, 1982; Vol. II/14a. (c) Demaison, J.; Hüttner, W.; Tiemann, E.; Vogt, J.; Wlodarczak, G. In *Landolt-Börnstein New Series*; Hüttner, W., Ed.; Springer-Verlag: New York, 1992; Vol. II/19c.
(49) Gay, D. M. *Assoc. Comput. Math. Trans. Math. Software* **1983**, *9*, 503.
(50) Scrocco, E.; Tomasi, J. *Top. Curr. Chem.* **1973**, *42*, 95.