

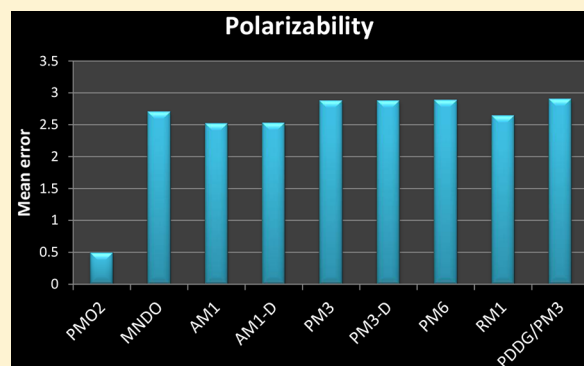
# Polarized Molecular Orbital Model Chemistry 3. The PMO Method Extended to Organic Chemistry

Miho Isegawa, Luke Fiedler, Hannah R. Leverentz, Yingjie Wang, Santhanamoorthi Nachimuthu, Jiali Gao, and Donald G. Truhlar\*

Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455, United States

## S Supporting Information

**ABSTRACT:** The polarized molecular orbital (PMO) method, a neglect-of-diatomic-differential-overlap (NDDO) semiempirical molecular orbital method previously parametrized for systems composed of O and H, is here extended to carbon. We modified the formalism and optimized all the parameters in the PMO Hamiltonian by using a genetic algorithm and a database containing both electrostatic and energetic properties; the new parameter set is called PMO2. The quality of the resulting predictions is compared to results obtained by previous NDDO semiempirical molecular orbital methods, both including and excluding dispersion terms. We also compare the PMO2 properties to SCC-DFTB calculations. Within the class of semiempirical molecular orbital methods, the PMO2 method is found to be especially accurate for polarizabilities, atomization energies, proton transfer energies, noncovalent complexation energies, and chemical reaction barrier heights and to have good across-the-board accuracy for a range of other properties, including dipole moments, partial atomic charges, and molecular geometries.



## 1. INTRODUCTION

Semiempirical electronic structure methods based on the neglect of diatomic differential overlap (NDDO)<sup>1</sup> were originally developed to reduce computational costs to predict the properties of small organic molecules.<sup>1–4</sup> However, with increasing computational power, the demand for these methods has shifted to the predictions of properties for larger molecules such as proteins and to use in direct dynamics calculations where the energy and gradient must be evaluated a large number of times. In the NDDO methods,<sup>1</sup> all three-center and four-center integrals are neglected, and the remaining electronic integrals are approximated. As a consequence, the most time-consuming step becomes the diagonalization of the Fock matrix which has a computational scaling  $O(M^3)$  with  $M$  being the number of basis functions. Therefore, the development of efficient diagonalization methods such as divided-and-conquer strategies<sup>5,6</sup> has become important.

In addition to increasing the computational efficiency, it is also important to raise the accuracy of these semiempirical methods. One possibility for improving the accuracy is increasing the number of basis functions.<sup>7</sup> Conventional semiempirical molecular orbital methods are constructed by using a minimal basis set, and the resulting sparsity of virtual orbitals limits the accuracy of calculated polarizabilities. This is a serious deficiency since one of the main motivations for using a molecular orbital theory method rather than the less expensive molecular mechanics methods is to include polarization effects. It also has been demonstrated that additional

flexibility in modeling core–core repulsions can improve the accuracy of NDDO methods as seen in the progress from MNDO<sup>2</sup> to AM1.<sup>3</sup> However, the core–core repulsions do not affect the character of the molecular orbitals except indirectly through geometry changes, and one must be careful not to introduce spurious effects by making the parametrization of core–core repulsions unphysical.

Another area where semiempirical methods need improvement is the prediction of noncovalent intermolecular interactions. McNamara and Hillier<sup>8</sup> introduced an empirical  $R^{-6}$  dispersion term into the AM1 and PM3 Hamiltonian, where  $R$  is an interatomic distance. By introducing damped dispersion terms and simultaneously partially reoptimizing the electronic integrals in the original AM1 and PM3 Hamiltonian, they obtained significant improvements for biologically important intermolecular interactions such as stacked base pairs and hydrogen-bonded DNA base pairs. These methods are called AM1-D and PM3-D. More recently, Tuttle et al.<sup>9</sup> have included dispersion corrections for a series of orthogonalization models (OMx)<sup>10,11</sup> that have higher accuracy<sup>12</sup> than the standard semiempirical methods such as AM1<sup>3</sup> and PM3,<sup>4</sup> and Řezáč and Hobza<sup>13</sup> developed the D3H4 method that is designed to give an improved treatment of hydrogen bonding as well as dispersion-like interactions.

Received: June 19, 2012

Published: November 13, 2012

The self-consistent-charge density functional tight-binding (SCC-DFTB) method<sup>14</sup> is another development in improving semiempirical molecular orbital theory. This method is derived as a second-order expansion of the Kohn–Sham energy with respect to charge density variation. SCC-DFTB has strong similarities to older semiempirical methods, for example, the eigenvalue problem is solved only for the valence electrons by using a minimal basis, the effect of the core electrons is included in the empirical core repulsion function, all of the electronic integrals are represented by one-center or two-center approximations, and the method employs a number of empirically determined parameters. Sattelmeyer et al<sup>15</sup> reported that the SCC-DFTB method shows similar accuracy to that of AM1 and PM3. Empirical dispersion-like terms may optionally be added to the SCC-DFTB method, and at least two such augmented methods are available, one based on damped dispersion and the Slater–Kirkwood formula<sup>16,17</sup> and one based on modified Lennard-Jones potentials.<sup>18</sup> The need for an explicit treatment of dispersion-like contributions to the interaction energy arises because the single-configuration formalisms such as Hartree–Fock and B3LYP (on which NDDO and SCC-DFTB methods are respectively based) do not include dispersion-like interactions,<sup>19–21</sup> (although density functional theory with state-of-the-art forms for density functionals does include dispersion-like interactions at medium and short-range). Although the parameters of NDDO methods are typically fitted against experimental data, one can achieve transferability of the parameters only if the functional forms used for fitting contain the correct physics, and therefore, if one desires a formalism that is correct at long-range when dispersion makes an important contribution, the Hartree–Fock formalism used as a basis for NDDO must be augmented by functional forms appropriate for long-range dispersion.

In the present paper, we consider another NDDO-based semiempirical molecular orbital method called the polarized molecular orbital (PMO) method. This method was initially developed for molecules composed of H and O atoms,<sup>22</sup> and here it is extended to molecules also containing C atoms. The method is here modified from its original form in two ways: in the resonance term and in the core–core term. All of the parameters except for those of the dispersion term are determined simultaneously using a genetic algorithm (GA) so as to reproduce, in a least-squares sense, a wide range of molecular properties and relative energies. The resulting method, called PMO2 to distinguish it from the original<sup>22</sup> PMO, is compared with several previous NDDO semiempirical methods, in particular, MNDO,<sup>2</sup> AM1,<sup>3</sup> PM3,<sup>4</sup> AM1-D,<sup>8</sup> PM3-D,<sup>8</sup> RM1,<sup>23</sup> PM6,<sup>24</sup> and PDDG/PM3,<sup>25</sup> and with the SCC-DFTB method, with and without dispersion corrections. We also compare to first-principles B3LYP<sup>26</sup> density functional calculations.

## 2. PMO2 METHOD

The main motivation for the construction of the PMO and PMO2 methods is to overcome the small polarizability due to the use of a minimal basis set in conventional NDDO methods. This is important for treating intermolecular interactions, especially those involving hydrogen bonds. To achieve this, we first introduce a subshell of p-orbital basis functions on each hydrogen atom; STO-3G(p), as motivated by a previous study<sup>27</sup> of the efficiency of various small basis sets for including polarization.

The PMO2 method is based on the NDDO formalism, and all interatomic differential overlap is neglected except in the resonance integral. In this section we present the restricted Hartree–Fock formalism<sup>28</sup> used for closed-shell singlets. The formalism is extended to doublets and triplets in the standard way based on the unrestricted Hartree–Fock approximation<sup>29</sup> (which is not reviewed here).

The diagonal elements of the Fock matrix are given by

$$F_{\mu\mu} = H_{\mu\mu}^{\text{core}} + \sum_{\nu \in A} \sum_{\lambda \in A} \left[ P_{\nu\lambda}(\mu\mu|\nu\lambda) - \frac{1}{2} P_{\nu\lambda}(\mu\nu|\mu\lambda) \right] + \sum_{B \neq A} \sum_{\nu \in B} \sum_{\lambda \in B} P_{\nu\lambda}(\mu\mu|\nu\lambda) \quad (1)$$

where

$$H_{\mu\mu}^{\text{core}} = U_{\mu\mu} + \sum_{B \neq A} V_{\mu\mu,B} \quad (2)$$

$A$  and  $B$  label atoms,  $\mu$  and  $\nu$  label basis functions on atom  $A$ ,  $\lambda$  and  $\sigma$  are other basis functions,  $(\mu\nu|\lambda\sigma)$  is a two-electron integral in the standard Mulliken notation, and  $P_{\lambda\sigma}$  is an element of the density matrix. Although eqs 1 and 2 have the same form as in previous NDDO methods, the number of electronic integrals is greater for hydrogen-containing molecules because of the p-subshell on hydrogen. In eq 2,  $U_{\mu\mu}$  is a one-center, one-electron integral,

$$U_{\mu\mu} = \left( \mu \left| -\frac{1}{2} \nabla^2 - \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \right| \mu \right), \quad (3)$$

where  $Z_A$  is the effective nuclear charge (equal to the true nuclear charge minus the number of core electrons on atom  $A$ ),  $\mathbf{R}_A$  is the coordinate of atom  $A$ , and  $\mathbf{r}$  is the coordinate of an electron. In the NDDO formalism,  $U_{\mu\mu}$  is treated as an empirical parameter; therefore, the integral in eq 3 is not explicitly evaluated.  $V_{\mu\mu,B}$  is the two-center, one-electron integral representing attraction between an electron distribution  $\phi_\mu\phi_\mu$  on atom  $A$  and the core of atom  $B$ . This is expressed as the  $\mu = \nu$  case of

$$V_{\mu\nu,B} = -Z_B(\mu\nu|s^B s^B)\tau(R_{AB}), \quad (4)$$

where  $s^B$  is a fictitious atomic orbital used to represent the core electrons on  $B$ ,  $R_{AB}$  is an internuclear distance, and  $\tau$  is a damping function. The charge distribution  $|s^B s^B\rangle$  corresponds to a monopole distribution; then  $(\mu\mu|s^B s^B)$  is represented by monopole-monopole or monopole-multipole interactions in the same manner as in MNDO.<sup>30</sup> The damping factor is non-unity only for the case that  $A$  and  $B$  are both hydrogen atoms, and it is added because the interaction between the electron density on hydrogen and the nucleus of another hydrogen becomes too strong when we add a p-subshell orbital on H atoms; the damping function is given by

$$\tau(R_{AB}) = 1 - \exp(-\gamma_{AB} R_{AB}^2) \quad (5)$$

with  $\gamma_{AB}$  being  $1.4 \text{ \AA}^{-2}$ .

The off-diagonal element of the Fock matrix for the case that atomic orbitals  $\mu$  and  $\nu$  are on the same atom  $A$  is given by

Table 1. Sizes of Databases Used to Optimize and Test Parameters and Units of the Data

database	description	$N_{\text{CH}}^a$	$N_{\text{OH}}^a$	$N_{\text{CO}}^a$	$N_{\text{HCO}}^a$	$N_{\text{total}}^b$	unit
Dip60	dipole moments	26	1	2	31	60	debye
Pol62	polarizabilities	26	1	2	33	62	$\text{\AA}^3$
PAC9	partial atomic charge	4			5	9	electron
VIP30	vertical ionization potentials	7	1		22	30	eV
AEPB60	atomization energies per bond	26	1	2	31	60	kcal/mol
GPD63	gradient norms per DOF <sup>c</sup>	26	1	2	34	63	$\text{kcal mol}^{-1} \text{\AA}^{-1}$
PTE17	proton transfer energies	8		1	8	17	kcal/mol
RE29	chemical reaction energies	12	0	0	17	29	kcal/mol
BH4	chemical reaction barrier heights	0	4	0		4	kcal/mol
CE12	complexation energies	6	4	0	2	12	kcal/mol
ConfE6	conformational energies	2			4	6	kcal/mol
ConfB6	conformational barriers	5			1	6	kcal/mol

<sup>a</sup> $N_{\text{XY}}$ : Number of molecules containing only X and Y. <sup>b</sup> $N_{\text{total}} = N_{\text{CH}} + N_{\text{OH}} + N_{\text{CO}} + N_{\text{HCO}}$ . <sup>c</sup>DOF: Cartesian degree of freedom.

$$F_{\mu\nu} = \sum_{B \neq A} V_{\mu\nu B} + \sum_{\lambda \in A} \sum_{\sigma \in A} \left[ P_{\lambda\sigma}(\mu\nu\lambda\sigma) - \frac{1}{2} P_{\lambda\sigma}(\mu\lambda\nu\sigma) \right] + \sum_{B \neq A} \sum_{\lambda \in B} \sum_{\sigma \in B} P_{\lambda\sigma}(\mu\nu\lambda\sigma). \quad (6)$$

All of the nonzero one-center, two-electron integrals in eqs 1 and 6 are given as parameters,  $G_{ss}^X \equiv (\text{ss}|\text{ss})$ ,  $G_{sp}^X \equiv (\text{ss}|\text{pp})$ ,  $G_{pp}^X \equiv (\text{pp}|\text{pp})$ ,  $G_{pp'}^X \equiv (\text{pp}|\text{p}'\text{p}')$ , and  $H_{sp}^X \equiv (\text{spl}|\text{sp})$ , where X is atom type, e.g., H, C, or O.

The off-diagonal Fock element when  $\mu$  and  $\sigma$  are on different atoms, A and B, is given by

$$F_{\mu\sigma} = S_{\mu\sigma} \beta_{\mu\sigma} \exp(\kappa_{\mu\sigma}^{AB} R_{AB}) - \frac{1}{2} \sum_{\lambda \in B} \sum_{\nu \in A} P_{\lambda\nu}(\mu\nu\sigma\lambda), \quad (7)$$

where  $\beta_{\mu\sigma}$  and  $\kappa_{\mu\sigma}$  are parameters,  $S_{\mu\sigma}$  is an overlap integral, the first term is the resonance term, which is responsible for chemical bonding between two atoms, and we have introduced an exponential function in the resonance term for adjustment of the degree of effective overlap between two orbitals. The resonance term is modeled with an orbital-dependent pairwise scheme, which is an improvement over the MNDO method in which parametrization involves monatomic parameters. Note that the original prescriptions for approximating integrals in NDDO were based on the objective of preserving invariance to both orientation and hybridization, and on that basis it was recommended that  $\beta_{\mu\nu}$  must be the same for all orbitals on a given center. However, violation of this rule only destroys hybridization invariance, which is not essential. It does not violate the essential orientation invariance. Therefore we took advantage of this flexibility to make the theory more accurate.

The two-center two-electron integral in eq 7 is estimated in the same manner as in MNDO,<sup>30</sup> that is, these integrals ( $\mu\nu\lambda\sigma$ ) are represented as interactions of the charge distributions of the first and second atomic orbital products,  $\phi_\mu\phi_\nu$  and  $\phi_\lambda\phi_\sigma$ .

In eqs 1, 6, and 7, the density matrix  $P_{\mu\sigma}$  is defined as

$$P_{\mu\sigma} = 2 \sum_i^{\text{occ}} C_{\mu i} C_{\sigma i}, \quad (8)$$

and the molecular orbital coefficients  $\{C_{\mu i}\}$  are obtained by solving the secular equation,

$$\mathbf{FC} = \mathbf{CE}, \quad (9)$$

where E is a diagonal matrix of molecular orbital energies. Equation 9 is a reduced form of the Roothaan–Hall equation<sup>28</sup>

obtained by neglecting overlap. Finally, the electronic energy of the PMO2 method is given by

$$E_{\text{el}} = \frac{1}{2} \sum_{\mu}^{\text{occ}} \sum_{\sigma}^{\text{occ}} P_{\mu\sigma} (H_{\mu\sigma}^{\text{core}} + F_{\mu\sigma}) \quad (10)$$

where the off-diagonal elements of  $H_{\mu\sigma}^{\text{core}}$  are given by the first term in eq 6 or 7.

The total PMO energy for a system is represented as the sum of the electronic, core–core, and dispersion terms:

$$E^{\text{PMO2}} = E_{\text{el}} + E_{\text{core-core}} + E_{\text{disp}} \quad (11)$$

The core–core repulsion term in the PMO2 method is defined as

$$E_{\text{core-core}} = \frac{1}{2} \sum_A \sum_{B \neq A} E_{AB}^{\text{core-core}} \quad (12)$$

where

$$E_{AB}^{\text{core-core}} = Z_A Z_B (\mu^A \mu^A \nu^B \nu^B) [1 + C_0^{AB} e^{-\alpha'_{AB} R_{AB}} + C_1^{AB} e^{-\alpha_A R_{AB}} + C_2^{AB} e^{-\alpha_B R_{AB}} + C_3^{AB} R_{AB}^{\eta_{AB}} e^{-\alpha''_{AB} R_{AB}}] \quad (13)$$

Here,  $\eta_{AB}$  is equal to 1 for atom pairs involving a hydrogen atom and a noncarbon atom and equal to 0 for all of the remaining atom pairs. Note that in the PMO2 method, both one- and two-body parameters are included in the core–core term to achieve a more accurate description.

Equation 11 also includes an empirical damped dispersion interaction given by<sup>8,31</sup>

$$E_{\text{disp}} = -s_6 \sum_A^N \sum_B^N \frac{C_6^{AB}}{R_{AB}^6} f_{\text{damp}}(R_{AB}) \quad (14)$$

where  $s_6$  is 1.4 as in refs 8 and 31, and  $f_{\text{damp}}(R_{AB})$  is a damping factor<sup>8,31</sup> that is introduced to avoid unphysical attraction at short-range and medium range and is given by,

$$f_{\text{damp}}(R_{AB}) = \frac{1}{1 + e^{[-\alpha(R_{AB}/R_{AB}^0 - 1)]}} \quad (15)$$

Here,  $R_{AB}^0$  is given as a sum of the van der Waals radii:  $R_A^0$  and  $R_B^0$ , respectively,

$$R_{AB}^0 = R_A^0 + R_B^0 \quad (16)$$

and  $C_6^{AB}$  is given by

$$C_6^{AB} = 2 \frac{C_6^A C_6^B}{C_6^A + C_6^B}. \quad (17)$$

The parameters,  $\alpha$  and  $R_A^0$  for hydrogen and carbon are taken to have the values determined by Grimme,<sup>31</sup> whereas  $R_A^0$  for oxygen and  $C_6^A$  for hydrogen, carbon, and oxygen are taken as parameters to be optimized in the present study. The details are described in the following sections.

### 3. BENCHMARK DATABASE

Table 1 lists the databases used for testing the PMO2 parameters and comparing the performance to other methods. It is important to emphasize that not all of the data for a given property were used for training the parameters, so the training and test data are actually mixed in the database. The parametrization involved a large amount of trial and error on various subsets of the data, often with varying weights and by more than one individual with more than one strategy for optimization. Many of the early rounds of parametrization were carried out on very small amounts of data. From this we obtained reasonable parameters, which provided starting points for more complete optimizations. Many calculations were made trying to improve on one or another of the parameter sets obtained in various runs that looked promising. Although the GA in principle finds a global minimum, this is not so when the parametrization process is stopped after a finite amount of time—as it must be, and so the local minima obtained in early rounds of parametrization have an influence on the final results. Furthermore there is a probabilistic element to any round of parametrization; when we increased the amount of training data, it did not necessarily lead to better parameters from the point of view of our overall objective, which was balanced performance across the whole set of databases, not just a particular weighted average of the absolute errors. To pretend that the data could be divided into training data and test data would be to entertain a highly oversimplified view of a very complicated and nonlinear process. Although the polarizability was not included as a target in the original NDDO methods, it is a key property in the present optimization. Since the description of noncovalent interaction energies is a weakness of previous semiempirical methods, this property is also included. All the reference values of properties are obtained from theoretical calculations or experimental data, and most of the theoretical calculations are performed using the M06-2X density functional<sup>32</sup> with the MG3S<sup>33,34</sup> basis set (which is the same as 6-311++G(2df,2p)<sup>35,36</sup> for H, C, and O), although for the smaller molecules, coupled-cluster (CCSD(T)) calculations are performed with an extended basis set such as jun-cc-pVTZ.<sup>37</sup>

The inclusion of partial atomic charges in the database is important to avoid unphysical electron density distributions such as a negative charge on the hydrogen atom in ethylene. Charge model 5 (CM5)<sup>38</sup> is used to generate target values of partial atomic charges, and we also model atomic charge indirectly by including gas-phase dipole moments obtained from experiment or from theoretical calculations.

The ionization potentials that we used are vertical ionization potentials, which are calculated as the electronic energy difference between neutral and charged molecules at the neutral molecule's geometry. For the calculation of ionization potentials, we chose molecules that have an energy gap of 0.4 eV or greater between the highest-occupied molecular orbital (HOMO) and the subjacent occupied molecular orbital

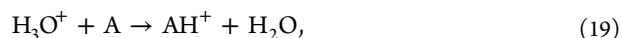
(HOMO-1) to eliminate the need, during parametrization, to monitor SCF convergence to the correct state of the ion.

We normalized the atomization energy of molecules by dividing by the number of bonds in the molecule to prevent large molecules from dominating the statistics. Gradients of the energy at accurate equilibrium geometries were used as fitting targets instead of geometrical parameters such as bond length to reduce the calculation cost and the need to converge geometry optimizations during the parametrization. In particular the quantity we employed as a target for minimization for each molecule is the gradient norm per Cartesian degree of freedom (DOF):

$$G^{\text{PDOF}} = \frac{\left\{ \sum_A^N \left[ \left( \frac{\partial E^{\text{PMO2}}}{\partial x_A} \right)^2 + \left( \frac{\partial E^{\text{PMO2}}}{\partial y_A} \right)^2 + \left( \frac{\partial E^{\text{PMO2}}}{\partial z_A} \right)^2 \right] \right\}^{1/2}}{3N} \quad (18)$$

where  $N$  is the number of atoms in the molecule.

We include four kinds of reaction energies and two kinds of barriers in the parametrization. The reaction energies are for (i) proton transfer from hydronium,



(ii) rearrangement reactions in which bonds are broken and formed, (iii) complexation reactions characterized by hydrogen bonded or van der Waals interactions, and (iv) conformational changes between local minima. The two kinds of barrier we included are (i) chemical reaction barrier heights, also called rearrangement barrier heights, and (ii) barriers to conformational transitions.

The names and sizes of the databases are presented in Table 1. The details of all reference values are given in the Supporting Information.

### 4. PARAMETRIZATION

In the PMO2 method, 82 independent parameters are optimized for compounds composed of H, C, and O. We determine the “best” parameter set using a GA.<sup>39,40</sup> We note that GA algorithms have also been used successfully in parameter optimization in the development of previous semiempirical methods.<sup>41–48</sup> The main advantage of using a GA algorithm is the ability to explore a high-dimensional space without being trapped in a local shallow minimum. In the GA, random numbers are used in creating each new generation of the parameters, and the parameters are evolved from an initial parameter set within user-specified ranges to find their optimum values. The parametrization is executed so as to minimize the penalty function,

$$F = \sum_X \frac{1}{\omega_X} \sqrt{\sum_i (f_{X,i} - f_{X,i}^{\text{ref}})^2 / N_X} \quad (20)$$

where  $X$  labels a database,  $f_{X,i}$  is property  $X$  of molecule  $i$  as calculated by PMO2,  $f_{X,i}^{\text{ref}}$  is the corresponding reference property, and  $N_X$  is the number of data in database  $X$ . The inverse weights,  $\{\omega_X\}$ , are chosen subjectively to obtain reasonable across-the-board accuracy for all properties. The convergence of the parameters for a given set of weights does not necessarily provide the desired properties, because the PMO functional form is not completely flexible, the optimization is stochastic, and the number of generations is finite. Also, it is possible that one or more property has too



Table 2. One-Center Parameters in the Electronic Part of the PMO2, MNDO, AM1-D, and PM3-D Hamiltonians.

atom	PMO2			MNDO		
	C	H	O	C	H	O
$U_{ss}$ [eV]	−48.566000	−10.589000	−91.714000	−52.279745	−11.906276	−99.644309
$U_{pp}$ [eV]	−40.922000	−7.235000	−72.331000	−39.205558	0.000000	−77.797472
$\zeta_s$ [ $\text{\AA}^{-1}$ ]	3.775673	2.246884	5.353594	3.377956	2.517053	5.102081
$\zeta_p$ [ $\text{\AA}^{-1}$ ]	3.093482	1.659180	6.553570	3.377956	0.000000	5.102081
$\alpha$ [ $\text{\AA}^{-1}$ ]	3.014000	3.615000	4.548000	2.546380	2.544134	2.544134
$G_{ss}$ [eV]	11.795000	12.376000	21.437000	12.230000	12.848000	15.420000
$G_{sp}$ [eV]	11.207000	8.160000	12.017000	11.470000	0.000000	14.480000
$G_{pp}$ [eV]	12.541000	7.364000	12.166000	11.080000	0.000000	14.520000
$G_{p2}$ [eV]	9.870000	11.346000	12.686000	9.840000	0.000000	12.980000
$H_{sp}$ [eV]	1.944000	1.882000	4.396000	2.430000	2.430000	3.940000
atom	AM1-D			PM3-D		
	C	H	O	C	H	O
$U_{ss}$ [eV]	−52.183798	−11.223791	−97.610588	−47.275431	−13.054076	−86.960302
$U_{pp}$ [eV]	−39.368413	0.000000	−78.589700	−36.268916	0.000000	−71.926845
$\zeta_s$ [ $\text{\AA}^{-1}$ ]	3.417882	2.245142	5.873330	2.957582	1.828890	7.174429
$\zeta_p$ [ $\text{\AA}^{-1}$ ]	3.184408	0.000000	4.769743	3.481528	0.000000	4.515316
$\alpha$ [ $\text{\AA}^{-1}$ ]	2.625506	3.577756	3.577756	2.721152	3.417532	3.387806
$G_{ss}$ [eV]	12.230000	12.848000	15.420000	11.200708	14.794208	15.755760
$G_{sp}$ [eV]	11.470000	0.000000	14.480000	10.265027	0.000000	10.621160
$G_{pp}$ [eV]	11.080000	0.000000	14.520000	10.796292	0.000000	13.654016
$G_{p2}$ [eV]	9.840000	0.000000	12.980000	9.042566	0.000000	12.406095
$H_{sp}$ [eV]	2.430000	2.430000	3.940000	2.290980	2.290980	0.593883

large a deviation from the reference value because of the improvement of the other properties. Therefore, we try many initial parameter sets, many parameter ranges, and many sets of weights in the GA runs, and monitor the mean unsigned error (MUE) of each property rather than just the value of the penalty function, and we capture several parameter sets that are well balanced for the reproduction of all properties, where each GA run is performed with up to 350–500 generations with a population size of 5. Then, we mix these parameters until we obtain the best parameters. In several semiempirical methods that have been developed in other groups, the parametrization has been executed stepwise starting from the reproduction of atomic properties. On the other hand, in the present parametrization, all of parameters are determined simultaneously so as to reproduce the properties given in Table 1.

## 5. COMPUTATIONAL DETAILS

The parametrization is executed using a program that combines a locally improved version of the GA of Carroll<sup>49</sup> with MOPAC version 5.018mn.<sup>50</sup> To compare PMO2 with earlier semiempirical methods, all of the properties in the present databases are computed for MNDO,<sup>2</sup> AM1,<sup>3</sup> AM1-D,<sup>8</sup> PM3,<sup>4</sup> PM3-D,<sup>8</sup> PM6,<sup>24</sup> RM1,<sup>23</sup> and PDDG/PM3<sup>25</sup> and with DFTB and DFTB+D, where DFTB+D denotes DFTB augmented by dispersion-like terms.

For these previous semiempirical methods as well as for PMO2, the full dipole moment is given by

$$\vec{\mu} = \vec{\mu}^{\text{Mull}} + \vec{\mu}^{\text{hyb}}, \quad (21)$$

where the first term corresponds to dipole moment computed from the Mulliken charges,<sup>51</sup> and the second term originates from atomic orbital hybridization. The hybrid dipole moment is defined as

$$\vec{\mu}^{\text{hyb}} = - \sum_A \sum_{\mu \in A} \sum_{\nu \in A} \vec{\gamma}_{\mu\nu}^A P_{\mu\nu}^A \quad (22)$$

where  $\mu$  and  $\nu$  are atomic orbitals on atom A, and  $\vec{\gamma}_{\mu\nu}^A$  is an integral given by

$$\vec{\gamma}_{\mu\nu}^A = \langle \mu | \vec{r} | \nu \rangle \quad (23)$$

where  $\vec{r}$  is an electronic coordinate. The PMO2 hybrid dipole contributions of eq 23 are calculated using the new orbital exponent parameters and are not hard-wired constants in the code as in some versions of MOPAC. The quantity in the database is the magnitude of the dipole moment vector:

$$\mu = |\vec{\mu}|. \quad (24)$$

SCC-DFTB calculations are performed using DFTB+, version 1.2,<sup>52</sup> the parameters employed are called “mio-0-1” which is a parameter set that was developed for organic molecules. For open-shell systems, spin polarized calculations are performed,<sup>53</sup> where the PBE spin constant (denoted<sup>54</sup>  $W_{\text{All}}$ ) is used. For the dispersion correction in SCC-DFTB-D, the modified Lennard-Jones (LJ) potential<sup>18</sup> is used, where the parameters are taken from the Universal Force Field.<sup>55</sup> Geometry optimization for SCC-DFTB is carried out using the conjugate gradient method.

For all methods, the polarizabilities are obtained as the first derivative of the dipole moment with respect to a weak uniform electric field for the neutral molecules and as the second derivative of the electronic energy for the ionic systems.

## 6. RESULTS AND DISCUSSION

**6.1. PMO2 Parameters.** Table 2 shows optimized PMO2 parameters and compares them to those of MNDO, AM1-D, and PM3-D. The most significant change from MNDO parameters is seen in the one-center two-electron integral  $G_{ss}$  for oxygen with the change being 39%, and the value of  $G_{ss}$

differs greatly from those of AM1-D and PM3-D. The exponent of the p-orbitals of oxygen,  $\zeta_p$ , also changes very much, in particular from 5.10 Å<sup>-1</sup> to 6.55 Å<sup>-1</sup>. Overall, the number of parameters that change more than 10% from MNDO is three for carbon ( $\zeta_s$ ,  $G_{pp}$ , and  $H_{sp}$ ), two for hydrogen ( $U_{ss}$  and  $H_{sp}$ ), and four for oxygen ( $\zeta_p$ ,  $G_{ss}$ ,  $G_{sp}$ , and  $G_{pp}$ ).

Table 3 gives optimized parameters in the resonance term (the first term in eq 7). In the previous PMO parametrization<sup>22</sup>

**Table 3. Orbital Pairwise Parameters in the Resonance Term**

	$\beta_{\mu\sigma}^{AB\alpha}$ [eV]	$\kappa_{\mu\sigma}^{AB\beta}$ [Å <sup>-1</sup> ]
H <sub>s</sub> -H <sub>s</sub> <sup>c</sup>	-6.098	-0.247
H <sub>p</sub> -H <sub>s</sub>	0.000	
H <sub>p</sub> -H <sub>p</sub>	-0.760	0.132
C <sub>s</sub> -H <sub>s</sub>	-13.704	-0.325
C <sub>s</sub> -H <sub>p</sub>	-0.282	0.389
C <sub>s</sub> -C <sub>s</sub>	-13.168	-0.318
C <sub>p</sub> -H <sub>s</sub>	-6.701	0.059
C <sub>p</sub> -H <sub>p</sub>	-0.762	0.021
C <sub>p</sub> -C <sub>s</sub>	-14.640	-0.125
C <sub>p</sub> -C <sub>p</sub>	-6.851	0.100
O <sub>s</sub> -H <sub>s</sub>	-14.923	-0.147
O <sub>s</sub> -H <sub>p</sub>	-0.347	0.252
O <sub>s</sub> -C <sub>s</sub>	-27.678	-0.260
O <sub>s</sub> -C <sub>p</sub>	-17.943	-0.066
O <sub>s</sub> -O <sub>s</sub>	-20.932	-0.913
O <sub>p</sub> -H <sub>s</sub>	-25.614	-0.013
O <sub>p</sub> -H <sub>p</sub>	-1.639	0.308
O <sub>p</sub> -C <sub>s</sub>	-39.434	-0.157
O <sub>p</sub> -C <sub>p</sub>	-26.349	0.036
O <sub>p</sub> -O <sub>s</sub>	-18.304	0.871
O <sub>p</sub> -O <sub>p</sub>	-37.915	-0.907

<sup>a</sup>Resonance integral. <sup>b</sup>Exponent in the resonance term. <sup>c</sup>Pairwise of s-orbital on hydrogen atom and s-orbital on hydrogen atom.

(called version 1) for systems that consist of oxygen and hydrogen,<sup>22</sup> two resonance integrals,  $\beta_{ps}^{HH}$  and  $\beta_{pp}^{HH}$ , are assumed to be zero. However, in PMO2,  $\beta_{pp}^{HH}$  has a value of -0.760 eV. Furthermore, in PMO version 2,  $\beta_{ps}^{HH}$  is constrained to be zero to avoid its taking a positive value. The exponential parameter  $\kappa_{\mu\sigma}$  takes either positive or negative values, and its range is approximately from -1.0 to 0.5 Å<sup>-1</sup>; the overlap between orbitals  $\mu$  and  $\sigma$  increases when  $\kappa_{\mu\sigma}$  takes a positive value. The optimized parameters for the core–core repulsion term are summarized in Table 4, except that one of the core–core parameters,  $\alpha$ , is a one-center parameter, and is therefore included in Table 2.

**Table 4. Atomic Pairwise Parameters in the Core–Core Term<sup>a</sup>**

	$C_0$ [eV]	$C_1$ [eV]	$C_2$ [eV]	$C_3$ [eV]	$\alpha'$ [Å <sup>-1</sup> ]	$\alpha''$ [Å <sup>-1</sup> ]
H-H	0.353	0.388	0.689	0.349	2.494	3.993
H-C	0.538	0.173	0.562	0.530	2.179	2.474
H-O	0.573	0.663	0.477	0.645	2.970	3.071
C-C	0.811	0.297	1.124	0.629	3.038	2.083
C-O	0.530	0.695	0.799	0.350	2.571	2.895
O-O	0.558	0.338	0.631	0.621	3.984	5.438

<sup>a</sup>The one-center parameter in the core–core term is included in Table 2.

We changed four parameters,  $C_6^A$  for hydrogen, oxygen, and carbon and  $R_0^A$  for oxygen, in the dispersion-like term in eq 14. The new values of  $C_6^A$  are 3.32, 9.53, and 26.95 eV Å<sup>6</sup> for hydrogen, carbon, and oxygen as compared to the original values of 1.66, 17.10, and 7.25 eV Å<sup>6</sup>, respectively, and the new value of  $R_0^A$  for oxygen is 1.48 Å as compared to the original value of 1.49 Å.

**6.2. Properties.** We evaluate the accuracy of PMO2 by using the MUE as compared to the reference value for each property. We partition the data into four groups: CH, OH, CO, and HCO, where CH refers to molecules that just contain carbon and hydrogen, OH to molecules with oxygen and hydrogen, CO to molecules with carbon and oxygen, and HCO to molecules with hydrogen, carbon, and oxygen. Table 1 shows the number of data in each database for each of these four classes of molecules. Table 5 summarizes the resultant MUEs for individual properties calculated by PMO2, and it also contains the MUEs computed by ten semiempirical methods: MNDO, AM1, AM1-D, PM3, PM3-D, PM6, RM1, PDDG/PM3, and SCC-DFTB and SCC-DFTB-D. Finally, Table 5 also includes results calculated with the B3LYP density functional with the 6-31G\* basis set.

**6.2.1. Gradients and Geometries.** Since all of the properties listed in Table 1 are computed using the reference geometries, it is important to achieve a small gradient for the reference geometries.

Figure 1 displays the norm of the gradient per Cartesian DOF, for each molecule, and Table 5 gives the corresponding MUEs. (For this property, the error is the deviation from zero.) The MUE of PMO2, 1.99 kcal mol<sup>-1</sup> Å<sup>-1</sup>, is very close to that of PDDG/PM3, 1.98 kcal mol<sup>-1</sup> Å<sup>-1</sup>, and smaller than that of MNDO, AM1, AM1-D, PM3-D, SCC-DFTB, and SCC-DFTB-D. As pointed out in a previous study,<sup>15</sup> CO compounds, in particular carbon dioxide and carbon monoxide, are difficult systems for calculating geometries by semiempirical methods, and PMO2 does not completely solve this although it is considerably better than some of the previous methods.

Table 6 presents the optimized geometry of a water molecule as optimized by each method. PMO2 yields an OH bond length of 0.96 Å, which is consistent with the reference value, and yields a bond angle, 101.8 deg, within 2.7 deg of the reference value, 104.5 deg. For all methods, the deviation from the reference value of the OH bond length is 0.03 Å or less except for AM1-D, and the deviation in bond angle is 4.4 deg or less.

Figure 2 shows binding energy profiles of hydrogen bonded water dimer for PMO2, AM1-D, PM3-D, and CCSD(T)/aug-cc-pVTZ, where the internal coordinates of each water monomer are fixed to the geometry optimized by that method for a single water monomer, and  $\theta$  in Figure 3 is set to 120.0 deg. In the previous PMO study,<sup>22</sup> it was pointed out that PM3, PDDG/PM3, and PM3-D give the qualitatively correct configuration of water dimer; however, MNDO, AM1, RM1, and PM6 are qualitatively incorrect. PMO2 gives the correct configuration, as did version 1 of PMO, although in PMO2 the hydrogen bond length, 2.10 Å, is somewhat longer than the best estimated value, 1.95 Å, and the binding energy at the best estimated bond length is 4.88 kcal/mol as compared to the reference value of 5.02 kcal/mol.

**6.2.2. Dipole Moments and Polarizabilities.** Figure 4 displays the dipole moment for PMO2, AM1-D, PM3-D, PDDG/PM3, and SCC-DFTB-D (the full set of absolute errors is listed in Supporting Information, Table S1) as compared to

Table 5. Summary of MUEs in PMO2 and Several Semiempirical Methods, MNDO, AM1, AM1-D, PM3, PM3-D, PM6, RM1, SCC-DFTB, and SCC-DFTB-D

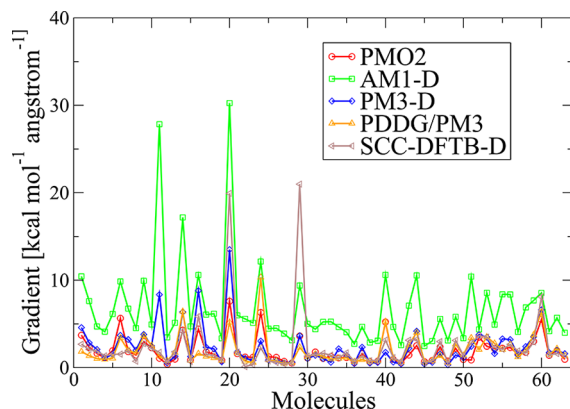
	PMO2					MNDO				
	CH	OH	CO	HCO	total	CH	OH	CO	HCO	total
Dip60	0.14	0.03	0.00	0.15	0.14	0.25	0.27	0.00	0.19	0.21
Pol62	0.47	0.42	0.42	0.53	0.49	2.86	1.12	2.24	2.71	2.70
PAC9	0.03			0.04	0.04	0.08			0.06	0.07
VIP30	0.37	0.14	0.37	0.39	0.37	0.27	0.52	0.62	0.28	0.29
AEPB60	0.96	3.34	8.10	2.35	1.96	5.17	3.97	19.12	4.63	5.34
GPD63	1.96	0.86	2.44	2.02	1.99	4.16	0.61	1.82	3.04	3.44
PTE17	5.00		4.56	10.13	7.39	29.72		12.72	10.69	19.76
RE29	7.44			6.66	6.98	3.90			6.18	5.24
BH4		3.52		2.23	2.88		5.80		4.12	4.96
CE12	1.06	1.20		3.31	1.40	2.70	18.70		24.66	13.76
ConfE6	3.34			2.56	2.95	3.31			2.64	2.97
ConfB6	8.05			0.28	6.76	8.70			0.38	7.31
	AM1					AM1-D				
	CH	OH	CO	HCO	total	CH	OH	CO	HCO	total
Dip60	0.18	0.26	0.00	0.16	0.17	0.13	0.13	0.00	0.11	0.12
Pol62	2.68	1.01	2.11	2.50	2.52	2.70	1.07	2.14	2.52	2.53
PAC9	0.01			0.04	0.03	0.01			0.06	0.04
VIP30	0.23	0.43	0.14	0.28	0.26	0.28	0.24	0.10	0.29	0.28
AEPB60	5.15	4.61	18.54	4.59	5.30	8.85	50.51	12.04	9.91	10.20
GPD63	4.63	1.59	1.77	3.41	3.85	7.80	4.72	6.92	5.92	6.74
PTE17	11.83		11.38	4.45	8.33	20.52		10.95	6.40	13.31
RE29	5.04			6.18	5.71	6.65			14.26	11.11
BH4		2.94		2.81	2.88		16.11		16.95	16.53
CE12	1.46	5.16		7.85	4.17	0.71	1.94		2.21	1.54
ConfE6	1.87			1.94	1.91	2.70			2.69	2.69
ConfB6	8.21			0.05	6.85	8.54			0.18	7.15
	PM3					PM3-D				
	CH	OH	CO	HCO	total	CH	OH	CO	HCO	total
Dip60	0.22	0.15	0.00	0.19	0.19	0.20	0.14	0.00	0.17	0.18
Pol62	3.09	1.02	2.37	2.86	2.88	3.09	1.03	2.37	2.87	2.88
PAC9	0.04			0.07	0.05	0.03			0.07	0.05
VIP30	0.23	0.61	0.38	0.29	0.28	0.23	0.58	0.37	0.29	0.28
AEPB60	4.97	7.62	8.39	4.46	4.86	1.93	6.45	19.56	3.28	3.29
GPD63	2.00	0.86	1.49	1.81	1.87	2.74	1.27	2.26	2.10	2.36
PTE17	9.23		23.14	8.20	9.57	13.34		22.16	7.79	11.25
RE29	6.61			4.86	5.58	5.22			4.84	5.00
BH4		5.85		4.73	5.29		7.48		5.50	6.49
CE12	0.97	2.99		4.86	2.50	0.69	0.49		0.77	0.59
ConfE6	1.91			0.97	1.44	1.90			1.32	1.61
ConfB6	7.94			0.03	6.62	7.84			0.21	6.57
	PM6					RM1				
	CH	OH	CO	HCO	total	CH	OH	CO	HCO	total
Dip60	0.19	0.04	0.00	0.20	0.19	0.18	0.31	0.00	0.14	0.16
Pol62	3.14	0.84	2.49	2.82	2.89	2.83	1.01	2.19	2.63	2.64
PAC9	0.05			0.22	0.15	0.02			0.05	0.04
VIP30	0.45	0.57	0.24	0.58	0.51	0.37	0.60	0.10	0.35	0.36
AEPB60	4.92	7.13	11.30	4.83	5.12	4.93	5.32	12.08	4.63	5.02
GPD63	1.89	0.67	1.30	1.53	1.66	1.95	0.68	1.33	1.92	1.89
PTE17	14.95		8.51	4.24	9.53	8.83		10.56	6.74	7.94
RE29	9.74			3.60	6.14	6.80			2.75	4.43
BH4		8.15		10.82	9.49		1.54		5.71	3.62
CE12	0.95	1.77		3.25	1.66	1.05	5.40		7.59	4.12
ConfE6	2.36			2.10	2.23	1.94			2.03	1.99
ConfB6	8.73			0.30	7.33	8.13			0.12	6.80

Table 5. continued

	PDDG/PM3					SCC-DFTB				
	CH	OH	CO	HCO	total	CH	OH	CO	HCO	total
Dip60	0.15	0.12	0.00	0.15	0.15	0.23	0.16	0.13	0.20	0.20
Pol62	3.12	1.04	2.42	2.87	2.90	2.90	1.47	2.32	2.82	2.78
PAC9	0.01			0.05	0.04	0.02			0.12	0.07
VIP30	0.23	0.48	0.16	0.27	0.26	0.45	0.75	0.39	0.61	0.54
AEPB60	2.14	6.59	11.75	1.82	2.37	2.95	5.09	24.11	4.13	4.30
GPD63	2.10	0.81	1.61	1.94	1.98	2.53	0.85	2.48	1.80	2.12
PTE17	9.11		22.23	9.11	9.88	21.09		9.54	7.88	14.20
RE29	7.52			4.02	5.47	5.36			4.36	4.77
BH4		4.81		3.33	4.07		0.92		3.50	2.21
CE12	0.95	3.39		6.31	2.87	1.09	2.66		2.15	2.05
ConfE6	1.92			1.83	1.88	1.57			1.89	1.73
ConfB6	7.75			0.53	6.55	9.73			0.12	8.13

	SCC-DFTB-D					B3LYP				
	CH	OH	CO	HCO	total	CH	OH	CO	HCO	total
Dip60	0.23	0.16	0.13	0.20	0.20	0.10	0.10	0.06	0.11	0.10
Pol62	2.90	1.47	2.32	2.82	2.78	1.34	0.61	1.19	1.32	1.31
PAC9	0.02			0.12	0.07	0.06			0.12	0.09
VIP30	0.45	0.75	0.39	0.61	0.54	0.39	0.36	0.29	0.50	0.45
AEPB60	0.59	3.98	22.71	1.93	2.08	0.40	7.03	4.51	1.07	0.99
GPD63	2.60	0.68	10.86	1.87	2.45	1.23	0.83	0.54	1.19	1.18
PTE17	21.77		9.84	8.01	14.59	3.42		5.89	3.33	3.52
RE29	6.63			4.49	5.37	3.22			3.97	3.66
BH4		1.02		3.70	2.36		2.90		3.30	3.10
CE12	0.29	3.70		3.91	2.52	1.07	3.37		2.76	2.49
ConfE6	1.61			1.97	1.79	0.43			0.61	0.52
ConfB6	9.11			0.13	7.61	1.21			0.34	1.06



**Figure 1.** Gradient norms per Cartesian degree of freedom (DOF) of 63 molecules calculated by PMO2, AM1-D, PM3-D, PDDG/PM3, and SCC-DFTB-D.

the reference values, and the corresponding MUEs are summarized in Table 5. PMO2, with the MUE being 0.14 debye, performs very well for dipole moments, second only to AM1-D. The largest error in the PMO2 dipole moment, among 60 molecules, is observed for ethylene glycol which has large charge separation within the molecule. It is noted that the dipole moment is reproduced well not only for small molecules, but also for the relatively large molecules as seen in ethyl 3-hydroxybutyrate and adipic acid.

The dipole moment of a water monomer calculated by PMO2 is 1.84 debye as compared to the best estimated value, 1.85 debye. AM1, RM1, and PDDG/PM3 also reproduce well the dipole moment with the value being 1.86, 1.86 and 1.85

**Table 6.** Water Monomer Geometries

	$R_{OH}^a$	$\theta_{HOH}^b$
PMO2	0.96	101.8
MNDO	0.94	106.8
AM1	0.96	103.6
AM1-D	0.83	106.2
PM3	0.95	107.7
PM3-D	0.93	108.9
PM6	0.95	107.0
RM1	0.96	103.7
PDDG/PM3	0.96	105.4
SCC-DFTB	0.97	107.2
SCC-DFTB-D	0.97	107.3
B3LYP	0.97	103.6
ref.	0.96	104.5

<sup>a</sup>In Å. <sup>b</sup>In degree.

debye, respectively. On the other hand, the largest deviation is observed in B3LYP, with the value being 2.08 debye, and AM1-D and PM6 also overestimate the dipole moment at the reference geometry. For the MUEs of the dipole moment in Table 5, it is noted that one of the reasons that CH compounds provide the lowest MUE is that the CH systems include several zero dipole moments such as benzene.

The dipole moments calculated by SCC-DFTB are underestimated for almost all of the molecules in the database, which has also been pointed out in a previous study, where a parametrized charge model was suggested to improve this.<sup>56</sup>

Figure 5 plots the polarizability obtained by PMO2 against those computed by several selected semiempirical methods. This confirms that the polarizability is considerably improved in



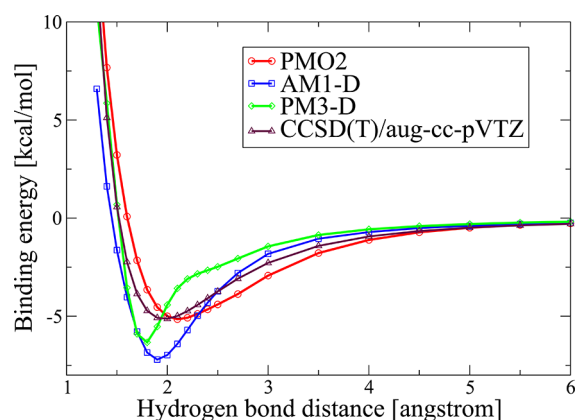


Figure 2. Binding energy profiles of hydrogen bonded water dimer.

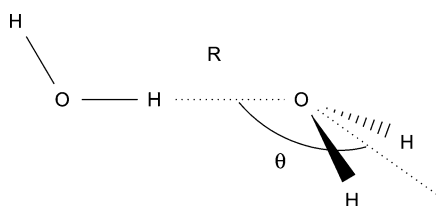


Figure 3. Hydrogen bonded water dimer.

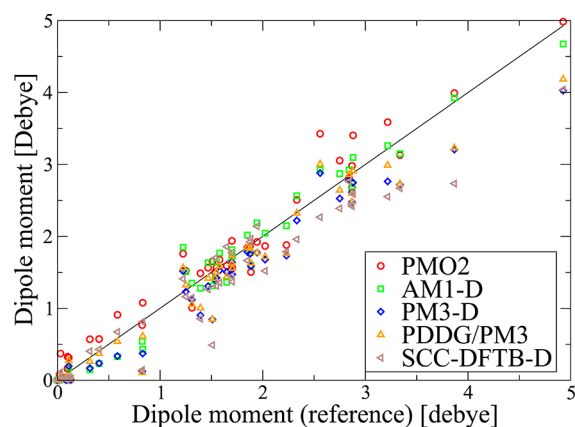


Figure 4. Dipole moments of PMO2, AM1-D, PM3-D, PDDG/PM3, and SCC-DFTB-D against the reference values.

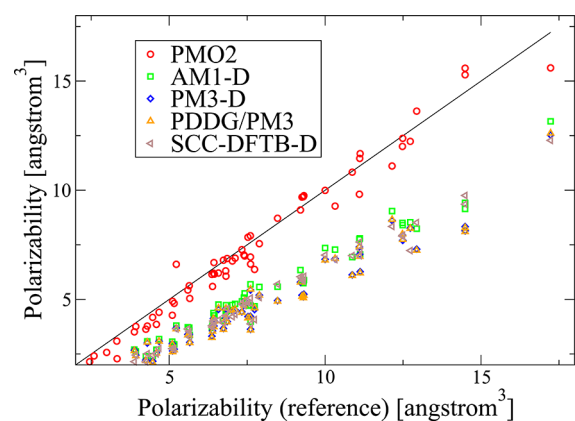


Figure 5. Polarizabilities of PMO2, AM1-D, PM3-D, PDDG/PM3, and SCC-DFTB-D against the reference values.

PMO2, and we attribute this to the presence of p functions on hydrogen. The polarizability of the hydrogen atom is zero for the conventional semiempirical methods because there is only one basis function on H; however, PMO2 gives a nonzero value for that polarizability because of the anisotropic behavior of p-type functions under the electrostatic field. Although the polarizabilities for systems that do not contain hydrogen are not improved when compared with the other semiempirical methods, which use a minimal basis set, most molecules in organic chemistry include hydrogen atoms. Averaged over the 62 molecules in the Pol62 database, PMO2 recovers 92% of the polarizability, whereas AM1 and SCC-DFTB recover only 62% and 58%, respectively. SCC-DFTB is especially poor for the important molecule water, for which it recovers only 26% of the reference polarizability, whereas PMO2 has an error of only 10% for water.

**6.2.3. Partial Atomic Charges.** We include partial atomic charges as target values in the parametrization to avoid an unphysical electron density distribution. A physical distribution of partial atomic charges is particularly important for fragment methods where the effect of one fragment on other subsystems is represented by point charges.

Although it is known that Mulliken population analysis can yield unstable results when one uses extended basis sets or diffuse basis functions, it is reasonable to employ it for semiempirical methods that do not employ diffuse basis functions. Table 7 gives the absolute error of the calculated Mulliken charge of each method for nine atomic charges of five selected molecules, ethane, ethylene, acetone, propane, and benzaldehyde. AM1 provides the closest atomic charges to the CM5 reference values with the MUE being 0.028 electron, and PDDG/PM3 and PMO2 also provide low MUEs. The electron density on the hydrogen atom of ethane is too large, resulting in smaller positive partial charges in PMO2 as compared to CM5, and this trend is also seen in MNDO, PM3, PM3-D, and RM1. On the other hand, the electron density on a hydrogen atom in ethane as estimated by PM6 is too small with respect to that inferred from the CM5 charge. Overall, the signs of atomic charges are consistent with those of CM5 for all the semiempirical methods except for the charge on hydrogen in ethylene estimated by MNDO.

**6.2.4. Vertical Ionization Potentials and Atomization Energies.** MUEs of vertical ionization potential are given in Table 5, and the individual value for each molecule is given in Supporting Information, Table S3. AM1 and PDDG/PM3 give the lowest MUE for the ionization potential; the MUE for PMO2 is almost the same as for RM1 and is much better than PM6, SCC-DFTB, SCC-DFTB-D, and B3LYP. For PMO2, relatively large deviations from the reference are observed for ethylene and isobutane, and the large error for ethylene is also observed for the other methods, suggesting that this is a difficult property for semiempirical methods.

Figure 6 plots the atomization energy per bond of PMO2, AM1-D, PM3-D, PDDG/PM3, and SCC-DFTB-D, against the reference values. Note that the range of plot is set from 90 to 150 kcal/mol to make the plot clear, but as a consequence, two of the 60 sample molecules, CO and CO<sub>2</sub>, are not on the figure. Table 5 shows a summary of MUEs, and each absolute error is listed in Supporting Information, Table S4. The reference values of atomization energies are reproduced quite well by the present calculations, with the resultant MUE being only 1.96 kcal/mol. The MUE of AM1-D is twice as large as that for AM1; on the other hand, PM3 is improved by the dispersion-

Table 7. Mulliken Atomic Charges in Atomic Units

atom	PMO2	MNDO	AM1	AM1-D	PM3	PM3-D	PM6	RM1	PDDG/PM3	SCC-DFTB	SCC-DFTB-D	B3LYP <sup>a</sup>	CM5 <sup>b</sup>
Ethane													
H	0.057	−0.006	0.069	0.067	0.035	0.037	0.141	0.056	0.086	0.064	0.064	0.145	0.082
Ethylene													
H	0.051	0.041	0.109	0.108	0.078	0.080	0.144	0.089	0.124	0.090	0.090	0.143	0.094
Acetone													
C	0.250	0.197	0.228	0.278	0.268	0.275	0.622	0.285	0.278	0.431	0.431	0.450	0.186
O	−0.300	−0.278	−0.280	−0.343	−0.311	−0.318	−0.490	−0.308	−0.308	−0.398	−0.398	−0.424	−0.318
H	0.084	0.024	0.105	0.106	0.069	0.071	0.197	0.099	0.120	0.095	0.095	0.181	0.103
Propane													
H	0.062	−0.006	0.070	0.068	0.037	0.039	0.146	0.057	0.087	0.062	0.062	0.140	0.079
H	0.047	0.009	0.078	0.076	0.047	0.049	0.121	0.059	0.091	0.055	0.055	0.133	0.083
Benzaldehyde													
C	0.175	0.307	0.230	0.284	0.328	0.331	0.392	0.278	0.274	0.336	0.336	0.195	0.127
O	−0.233	−0.296	−0.279	−0.343	−0.315	−0.321	−0.457	−0.303	−0.304	−0.368	−0.368	−0.402	−0.304
MUE	0.038	0.069	0.028	0.040	0.053	0.053	0.148	0.038	0.035	0.075	0.075	0.093	0.000

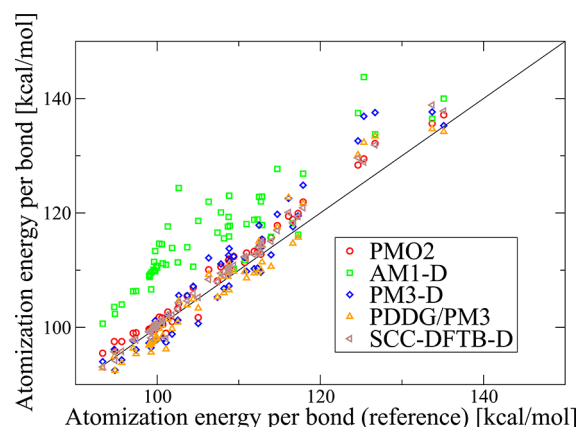
<sup>a</sup>6-31G\* basis set is used. <sup>b</sup>M06-2X/MG3S.

Figure 6. Atomization energies per bond of PMO2, AM1-D, PM3-D, PDDG/PM3, and SCC-DFTB-D against the reference values.

like contribution. This observation indicates that the large error in AM1-D comes from the parameters in the electronic part or from the core–core term in the Hamiltonian, which are partially reoptimized in the AM1-D development.

We suppose that there are two reasons why PMO2 yields the lowest MUE for atomization energies. One of the reasons is that we added an exponential function in the resonance term, and we employ an enhanced functional form in the core–core repulsion term. The former might be more effective than the latter, because the MUE is large in the SCC-DFTB method despite the use of 10 exponential functions in the core–core term. The second reason is that atomization energy is used as a target in the optimization of the PMO2 method, whereas all other methods use heat of formation.

**6.2.5. Proton Transfer Energies.** Table 8 shows the proton transfer energies. PMO2 shows the statistically best performance as compared to the eight other NDDO methods and

Table 8. Absolute Errors of Proton Transfer Energies [kcal/mol]<sup>a</sup>

	PMO2	MNDO	AM1	AM1-D	PM3	PM3-D	PM6	RM1	PDDG/PM3	SCC-DFTB	SCC-DFTB-D	B3LYP	ref. <sup>a</sup>
methanol	17.84	12.77	7.67	4.78	12.39	14.35	5.24	11.61	10.88	7.65	7.61	2.02	−15.55
ethanol	12.53	16.56	5.28	3.02	8.99	11.88	0.83	9.24	6.53	7.57	7.37	1.14	−20.52
dimethyl ether	22.90	18.92	10.19	3.72	17.73	20.95	7.21	18.77	15.56	11.29	11.04	3.65	−24.57
ethane	1.98	45.96	20.60	30.37	14.46	24.28	17.81	11.33	5.95	28.59	30.25	4.25	26.19
ethylene	1.88	24.68	5.63	25.42	1.44	11.49	18.13	7.49	2.91	18.51	20.25	0.55	3.92
acetylene	0.97	3.76	14.70	8.32	17.19	6.61	3.56	9.40	16.53	12.92	13.35	5.34	17.54
acetaldehyde	5.44	8.52	1.25	3.69	2.48	0.46	3.88	4.14	5.42	9.27	9.30	2.64	−19.52
acetone	3.38	13.99	0.60	2.26	1.24	0.83	5.59	5.01	4.78	9.06	9.38	1.88	−29.50
acetic acid	5.29	3.36	2.88	3.53	6.01	5.37	5.25	2.66	9.07	1.75	1.77	2.40	−19.12
cyclohexane (chair)	5.41	42.26	14.56	22.66	6.71	16.47	20.24	8.41	0.57	13.82	11.49	3.53	4.70
cyclopropane	5.11	38.68	20.18	37.08	16.72	26.86	19.26	19.15	17.08	32.02	33.65	0.86	−13.00
propane	1.17	39.96	11.25	16.21	2.88	12.33	14.95	3.40	3.93	14.04	14.12	0.48	18.39
formaldehyde	10.60	0.64	4.87	10.45	7.64	6.83	3.35	0.76	10.97	11.00	10.93	4.45	−5.78
phenol	3.04	10.76	2.88	19.78	9.15	1.66	2.55	1.73	9.68	5.49	6.66	8.46	−29.50
benzene	6.08	7.61	6.06	7.82	11.43	2.47	3.34	5.47	11.36	15.33	16.34	4.97	−12.80
methane	17.39	34.86	1.62	16.25	3.03	6.22	22.29	5.95	14.57	33.53	34.74	7.38	39.70
carbon dioxide	4.56	12.72	11.38	10.95	23.14	22.16	8.51	10.56	22.23	9.54	9.84	5.89	37.50
MUE	7.39	19.76	8.33	13.31	9.57	11.25	9.53	7.94	9.88	14.20	14.59	3.52	0.00

<sup>a</sup>The source of references are listed in the Supporting Information, Table S6.

Table 9. Isomerization Energies [kcal/mol]

	PMO2	MNDO	AM1	AM1-D	PM3	PM3-D	PM6	RM1	PDDG/PM3	ref. <sup>a</sup>
propyne → allene	1.22	2.43	1.67	1.67	6.57	6.57	3.86	5.52	5.49	0.52
propyne → cyclopropane	28.00	28.38	31.29	31.29	28.82	28.82	28.34	31.59	22.78	20.52
<i>trans</i> -2-butene → <i>cis</i> -2-butene	−2.33	1.33	1.20	1.20	0.92	0.92	0.80	1.22	1.30	1.10
<i>trans</i> -2-butene → isobutene	−1.17	2.98	1.59	1.59	−0.01	−0.01	−1.85	−1.46	−1.68	−1.12
acetaldehyde → ethylene oxide	21.23	27.96	32.14	32.14	37.05	37.05	30.56	37.40	35.30	22.88
acetic acid → methylformate	27.48	18.03	13.24	13.24	15.79	15.79	15.98	19.55	19.62	17.03
acetone → oxetane	18.62	11.23	22.39	22.39	26.79	26.79	26.65	27.16	30.77	30.16
MUE	5.04	5.59	5.08	5.08	4.92	4.92	3.49	5.22	3.37	

<sup>a</sup>M06-2X/MG3S//M06-2X/MG3S.

SCC-DFTB and SCC-DFTB-D. However, large deviations are observed for methanol, ethanol, and dimethyl ether, and these three molecules make the MUE for the HCO group large. Thus, the lowest MUEs of PMO2 for the proton transfer energy are found for the CH and CO compounds, for which the MUE is 5.0 and 4.6 kcal/mol, respectively. This is opposite to the trend of AM1, which performs better for the HCO compounds than for CH and CO compounds. For both SCC-DFTB and SCC-DFTB-D, the MUE for the CH group makes the total MUE large.

It is notable that a great improvement of the proton transfer energy is provided by the change of the core–core repulsion term.

**6.2.6. Chemical Reaction Energies and Barrier Heights.** Chemical bond formation and bond breaking are the most important properties in a quantum chemical model, but they cannot be described by conventional classical molecular mechanics force fields, although important progress has been made by using reactive force fields.<sup>48,57,58</sup> The absolute errors in the reaction energies are given in Supporting Information, Table S7, and Supporting Information, Table S8 contains errors in barrier heights. For the reaction energies, PMO2 shows a relatively high MUE compared with the other semiempirical methods. The highest error is seen in AM1-D with the MUE being 11.11 kcal/mol, and RM1 shows the best behavior for the reaction energies with the MUE being 4.43 kcal/mol.

For the four barrier heights, SCC-DFTB and SCC-DFTB-D provide the best performance. The MUEs for these two methods are 2.21 and 2.36 kcal/mol, respectively. Among NDDO methods, PMO2 and AM1 are the best with the MUEs being 2.88 kcal/mol. One of the reasons that most of the other semiempirical methods yield a larger error may be that these barriers were not included as targets in the parametrization procedure.

Table 9 gives the comparison of isomerization energies for several semiempirical methods, where reactions are taken from refs 15 and 25. Since these reactions are not included in our database, and the performance on these reactions was not examined until the parametrization was complete, these data provide a critical test for the optimized parameters. All of these isomerization reaction energies are calculated using geometries optimized by M06-2X/MG3S, and the energies are compared to those obtained by M06-2X/MG3S (rather than to the experimental data that include vibrational effects effects, as in refs 15 and 25). The PMO2 method gives comparable performance to that of AM1 and AM1-D. The best performance is given by PM3/PDDG and PM6 with their MUEs being 3.37 and 3.49 kcal/mol, respectively.

**6.2.7. Complexation Energies.** As McNamara and Hillier have shown,<sup>8</sup> the introduction of the dispersion term greatly improves the calculated complexation energies. This is confirmed for the present database, for which the total error of the complexation energies is much improved as can be seen from the comparison between AM1 and AM1-D and between PM3 and PM3-D.

Table 5 summarizes the MUEs for the 17 complexation energies, and the absolute errors of each complex are given in Supporting Information, Table S9. The best prediction is given by PM3-D, with the MUE being 0.59 kcal/mol, and PM3-D's MUE is lower than that of AM1-D. A large error is seen in the HCO compounds. One of the possible ways that one might try to reduce this error is to introduce pairwise parameters in the dispersion terms instead of using one-center parameters.

MNDO fails to predict the correct sign of the binding energy of water clusters for all types and sizes of water clusters, whereas all of the other semiempirical methods yield bound complexes except for water pentamer by AM1. For the formation of the sandwich form of the benzene dimer, stabilization occurs only for PMO2, AM1-D, PM3-D, and SCC-DFTB-D, which clearly shows the advantage of explicitly including dispersion-like contributions. Although PM6 does not have a dispersion term, the MUE is lower than for all other methods except PMO2, AM1-D, and PM3-D. The low MUE of PM6 comes from the good description of the water clustering energies, in which the electrostatic interactions make larger contributions than the dispersion interactions, but PM6 yields unbound states for two (parallel-displaced and sandwich) of the three types of benzene dimer.

It is noted that adding an empirical dispersion term improves the complexation energies for AM1 and PM3; however, in the SCC-DFTB method, the MUE becomes slightly worse by including the dispersion term. This is because parameters in the electronic part and core–core repulsion of the Hamiltonians are partially reoptimized in AM1-D and PM3-D starting from those of AM1 and PM3; however, the parameters in SCC-DFTB-D are the same as those of SCC-DFTB, which suggests that when one improves the complexation energies, it is required to reoptimize the other parameters.

**6.2.8. Conformational Energies and Conformational Barriers.** Supporting Information, Table S10 shows absolute errors for relative conformational energies. PMO2 gives a relatively large error for conformational energies, and the total error is almost the same as for MNDO and AM1-D. The best prediction is given by PM3 with the MUE being 1.44 kcal/mol. The large error of PMO2 comes from the conformational changes of methyl formate, *s-trans*-butadiene, and cyclohexane.

One can see that the mean unsigned error in the conformational energies becomes worse upon including the dispersion contributions as seen from the comparison between AM1 and AM1-D, PM3 and PM3-D, and SCC-DFTB and SCC-DFTB-D, though the complexation energies are improved for AM1 and PM3. This shows that the dispersion contribution is important for both properties, complexation and conformational change. To resolve the conflict, a special strategy of parametrization should be used as seen in the previous parametrization study for SCC-DFTB.<sup>59</sup>

For the conformational barriers (Supporting Information, Table S11), PMO2 shows comparable performance with the other methods. The largest error is generated in the conformational change of ethylene between planar and twist forms. For this torsion, SCC-DFTB and SCC-DFTB-D give an error greater than the other semiempirical methods by more than 10 kcal/mol. For the ethane barrier height, other studies<sup>60–62</sup> indicate the importance of steric repulsion contributions.

## 7. CONCLUDING REMARKS

We presented a semiempirical electronic structure method, PMO2, whose Hamiltonian is partially changed from the previous PMO method in the resonance and core–core repulsion terms, and we reparametrized the method for compounds containing hydrogen, carbon, and oxygen. As expected, the polarizability shows lower errors than all other semiempirical methods, which are constructed based on minimal basis sets. The MUE for atomization energy per bond is also lower than for any previous semiempirical method, and the chemical reaction barrier heights are more accurate than any previous NDDO method. Overall, the MUEs for nine other sets of properties are comparable to those of other semiempirical methods, showing that good performance for other properties may be retained while improving the performance of an NDDO method for polarizabilities, atomization energies, and barrier heights.

As is well recognized, the quality of semiempirical methods is strongly influenced by the parametrization procedure. Therefore, it is possible that PMO2 could be further improved by using a more efficient parametrization method as seen in the development of PM6, which is improved as compared to AM1 both by improving the functional forms and by better parametrization.

The PMO2 method should be useful for multilayer and fragment models, such as QM/MM,<sup>63,64</sup> ONIOM,<sup>65</sup> and X-Pol,<sup>66</sup> and should be very useful for molecular dynamics simulations. The next step for developing the polarized molecular orbital method is to extend the method to compounds that additionally contain nitrogen and sulfur for application to peptides and proteins.

## ■ ASSOCIATED CONTENT

### Supporting Information

Further details are given in Tables S1–S11. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [truhlar@umn.edu](mailto:truhlar@umn.edu)

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundations under Grant CHE09-56776 and by the National Institutes of Health, Grant RC1-GM091445.

## ■ REFERENCES

- (1) Pople, J. A.; Santry, D. P.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, S129.
- (2) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.
- (3) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
- (4) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.
- (5) Lee, T.-S.; York, D. M.; Yang, W. *J. Chem. Phys.* **1996**, *105*, 2774.
- (6) Xie, W.; Song, L.; Truhlar, D. G. *J. Chem. Phys.* **2008**, *128*, 234108.
- (7) Thiel, W. *Theor. Chim. Acta* **1981**, *59*, 191.
- (8) McNamara, J. P.; Hillier, I. H. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2362.
- (9) Tuttle, T.; Thiel, W. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2159.
- (10) Kolb, M.; Thiel, W. *J. Comput. Chem.* **1993**, *14*, 775.
- (11) Weber, W.; Thiel, W. *Theor. Chem. Acc.* **2000**, *103*, 495.
- (12) Korth, M.; Thiel, W. *J. Chem. Theory Comput.* **2011**, *7*, 2929.
- (13) Rezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 141.
- (14) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.
- (15) Sattelmeyer, K. W.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. A* **2006**, *110*, 13551.
- (16) Kumer, A.; Elstner, M.; Suhai, S. *Int. J. Quantum Chem.* **2003**, *95*, 44.
- (17) Valdés, H.; Řeha, D.; Hobza, P. *J. Phys. Chem. B* **2006**, *110*, 6385.
- (18) Zhechkov, L.; Heine, T.; Patchkovskii, S.; Seifert, G.; Duarte, H. A. *J. Chem. Theory Comput.* **2005**, *1*, 841.
- (19) Elstner, M.; Hobza, P.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149.
- (20) Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515.
- (21) Zimmerli, U.; Parrinello, M.; Koumoutsakos, P. *J. Chem. Phys.* **2004**, *120*, 2693.
- (22) Zhang, P.; Fiedler, L.; Leverentz, H. R.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.* **2011**, *7*, 857.
- (23) Racha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. *J. Comput. Chem.* **2006**, *27*, 1001.
- (24) Stewart, J. J. P. *J. Mol. Model.* **2007**, *13*, 1173.
- (25) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 1601.
- (26) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
- (27) Fiedler, L.; Gao, J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2011**, *7*, 852.
- (28) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.
- (29) Pople, J. A.; Nesbet, R. K. *J. Chem. Phys.* **1954**, *22*, 571.
- (30) Dewar, M. J. S.; Thiel, W. *Theor. Chem. Acta* **1977**, *46*, 89.
- (31) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.
- (32) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.
- (33) Curtiss, L. A.; Raghavachari, K.; Redfern, C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.
- (34) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384.
- (35) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
- (36) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.
- (37) Papajak, E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2011**, *7*, 10.
- (38) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2012**, *8*, 527.
- (39) Goldberg, D. *Genetic Algorithms in Search, Optimization and Machine Learning*; Addison-Wesley: Reading, MA, 1989.



- (40) Goldberg, D. *An Introduction to Genetic Algorithms for Scientists and Engineers*; World Scientific: River Edge, NJ, 1999.
- (41) Rossi, I.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, 233, 231.
- (42) Bash, P. A.; Ho, L. L.; MacKerell, A. D.; Levine, D.; Hallstrom, P. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, 93, 3698.
- (43) Hutter, M. C.; Reimers, J. R.; Hush, N. S. *J. Phys. Chem. B* **1998**, 102, 8080.
- (44) Brothers, E. N.; Merz, K. M., Jr. *J. Phys. Chem. B* **2002**, 106, 2779.
- (45) Lopez, X.; York, D. M. *Theor. Chem. Acc.* **2003**, 109, 149.
- (46) Giese, T. J.; Sherer, E. C.; Cramer, C. J.; York, D. M. *J. Chem. Theory Comput.* **2005**, 1, 1275.
- (47) Iron, M. A.; Heyden, A.; Staszewska, G.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, 4, 804.
- (48) Zhao, M.; Iron, M. A.; Staszewski, P.; Schultz, N. E.; Valero, R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, 5, 594.
- (49) Carroll, D. L. *FORTRAN Genetic Algorithm Driver*; CU Aerospace: Champaign, IL.
- (50) Stewart, J. J. P.; Fiedler, L. J.; Zhang, P.; Zheng, J.; Rossi, I.; Hu, W.-P.; Lynch, G. C.; Liu, Y.-P.; Chuang, Y.-Y.; Pu, J.; Li, J.; Cramer, C. J.; Fast, P. L.; Truhlar, D. G. *MOPAC 5.018mn*; Department of Chemistry, University of Minnesota: Minneapolis, MN, 2012.
- (51) Mulliken, R. S. *J. Chem. Phys.* **1955**, 23, 1833.
- (52) Aradi, B.; Hourahine, B.; Frauenheim, T. *J. Phys. Chem. A* **2007**, 111, 5678.
- (53) Köhler, C.; Seifert, G.; Frauenheim, T. *Chem. Phys.* **2005**, 309, 23.
- (54) Köhler, C.; Seifert, G.; Gerstmann, U.; Elstner, M.; Overhof, H.; Frauenheim, T. *Phys. Chem. Chem. Phys.* **2001**, 3, 5109.
- (55) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, 114, 10024.
- (56) Kalinowski, J. A.; Lesyng, B.; Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, 108, 2545.
- (57) Brenner, D. W. *Phys. Rev. B* **1990**, 42, 9458.
- (58) Van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A., III *J. Phys. Chem. A* **2001**, 105, 9396.
- (59) Gaus, M.; Chou, C.-P.; Witek, H.; Elstner, M. *J. Phys. Chem. A* **2009**, 113, 11866.
- (60) Mo, Y.; Wu, W.; Song, L.; Lin, M.; Zhang, Q.; Gao, J. *Angew. Chem., Int. Ed.* **2004**, 43, 1986.
- (61) Mo, Y.; Gao, J. *Acc. Chem. Res.* **2007**, 40, 113.
- (62) Bickelhaupt, F. M.; Baerend, E. J. *Angew. Chem., Int. Ed.* **2003**, 42, 4183.
- (63) Gao, J. *J. Rev. Comput. Chem.* **1996**, 7, 119.
- (64) Lin, H.; Truhlar, D. G. *Theor. Chim. Acta* **2007**, 117, 1185.
- (65) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, 100, 19357.
- (66) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Phys. Chem.* **2008**, 112, 234108.