

## Semiempirical NDDO Calculations with STO-3G and 4-31G Basis Sets

Walter Thiel

Fachbereich Physikalische Chemie, Universität Marburg, Hans-Meerwein-Strasse, D-3550 Marburg, Federal Republic of Germany

Possible refinements of semiempirical methods include the use of larger basis sets and of correlated wave functions. These possibilities are investigated in semiempirical NDDO SCF calculations with the STO-3G and 4-31G basis sets, and in correlated calculations at the STO-3G level. The present approach is characterized by the analytical evaluation of all one-center terms and two-electron integrals, and the semiempirical adjustment of the remaining one-electron integrals and the nuclear repulsions. The NDDO SCF results tend to reproduce the corresponding *ab initio* results more closely than experimental data, even if they are parametrized with respect to experiment. The explicit inclusion of electron correlation at the STO-3G level improves the calculated results only slightly.

**Key words:** NDDO calculations – Basis sets – Electron correlation

### 1. Introduction

From a formal point of view, current semiempirical methods designed for the study of potential surfaces [1–3] involve a number of drastic approximations, among them the use of a minimal basis set of valence orbitals and the neglect of electron correlation. These approximations are believed to be compensated by the parametrization procedure. On the other hand, *ab initio* calculations have demonstrated [4–6] that extension of the basis set usually improves the results in a well-defined manner and that inclusion of electron correlation is often essential for reliable predictions. This experience would suggest that the performance of semiempirical methods may also be improved by the use of larger basis sets and correlated wave functions.

In the present paper, we examine this possibility in semiempirical SCF calculations with two standard *ab initio* basis sets, STO-3G [7] and 4-31G [8], and in correlated calculations at the STO-3G level. The NDDO integral approximation [9] is used throughout because it is more refined and receives more theoretical justification [10–16] than its alternatives CNDO or INDO [9]. Since unparametrized NDDO versions [17–19] do not seem to be very accurate, the present approach follows the spirit of the MNDO method [3] including a parametrization primarily with respect to experimental data. It differs from MNDO [3] and other parametrized NDDO versions [20, 21] in the use of larger basis sets and the analytical evaluation of most terms in the Fock matrix.

The present paper does not intend to introduce a new semiempirical treatment for general use. It is mainly concerned with methodical development and aims at exploring the effects of possible refinements within a semiempirical framework.

## 2. Theoretical Approach

We consider an all-electron treatment for closed-shell molecules. The explicit inclusion of the core electrons and the possible use of a split-valence basis set make the NDDO formalism slightly more complicated than in the MNDO case [3].

For a given set of atomic orbitals (AOs)  $\phi_{\nu}$ , the LCAO coefficients  $c_{\nu i}$  and the eigenvalues  $\varepsilon_i$  are found from the Roothaan [22]–Hall [23] equations:

$$\sum_{\nu} (F_{\mu\nu} - \varepsilon_i S_{\mu\nu}) c_{\nu i} = 0. \quad (1)$$

In NDDO approximation, all nonzero one-center elements  $S_{\mu\nu}$  of the overlap matrix are retained whereas all two-center terms  $S_{\mu\nu}$  vanish. The elements  $F_{\mu\nu}$  of the Fock matrix are the sum of a one-electron part  $H_{\mu\nu}$  and a two-electron part  $G_{\mu\nu}$ , and the electronic energy  $E_{el}$  is given by:

$$E_{el} = \frac{1}{2} \sum_{\mu} \sum_{\nu} P_{\mu\nu} (H_{\mu\nu} + F_{\mu\nu}) \quad (2)$$

where  $P_{\mu\nu}$  is an element of the bond order matrix.

From now on, we shall assume that the AOs  $\phi_{\mu}$ ,  $\phi_{\mu'}$ ,  $\phi_{\nu}$  and  $\phi_{\nu'}$  are centered at atom  $A$  and the AOs  $\phi_{\lambda}$  and  $\phi_{\sigma}$  at atom  $B$  ( $A \neq B$ ). Superscripts  $A$  and  $B$  will assign a particular symbol to atom  $A$  or  $B$ , respectively. In this notation, the NDDO Fock matrix elements are:

$$F_{\mu\nu} = U_{\mu\nu} + \sum_B V_{\mu\nu,B} + \sum_{\mu',\nu'}^A P_{\mu'\nu'} [(\mu\nu, \mu'\nu') - \frac{1}{2}(\mu\mu', \nu\nu')] \quad (3)$$

$$+ \sum_B \sum_{\lambda,\sigma} P_{\lambda\sigma} (\mu\nu, \lambda\sigma) \quad (4)$$

$$F_{\mu\lambda} = \beta_{\mu\lambda} - \frac{1}{2} \sum_{\nu} \sum_{\sigma}^A P_{\nu\sigma} (\mu\nu, \lambda\sigma).$$

The following terms appear in the Fock matrix:

- (a) One-center energies  $U_{\mu\nu}$  which represent the sum of the kinetic energy of an electron in the distribution  $\mu\nu$  at atom  $A$  and its potential energy due to the attraction by nucleus  $A$ .
- (b) One-center repulsion integrals  $(\mu\nu, \mu'\nu')$  and  $(\mu\mu', \nu\nu')$ .
- (c) Two-center repulsion integrals  $(\mu\nu, \lambda\sigma)$ .
- (d) Two-center resonance integrals  $\beta_{\mu\lambda}$ .
- (e) Two-center core-electron attractions  $V_{\mu\nu,B}$ .

In current semiempirical methods with a minimal basis set of valence AOs [1–3] the one-center terms are either derived from atomic spectroscopic data or treated as adjustable parameters. This is no longer feasible for larger basis sets, e.g. 4-31G. We have therefore decided to evaluate all one-center terms analytically which makes the present procedure equivalent to an *ab initio* calculation in the case of atoms.

To be consistent with the treatment of the one-center repulsion integrals, the two-center repulsion integrals must also be evaluated exactly. The actual computations are carried out using modified versions of subroutines in the GAUSSIAN 70 program system [24].

The two-center one-electron integrals  $\beta_{\mu\lambda}$  and  $V_{\mu\nu,B}$  are the only remaining terms in the Fock matrix which may be expressed by parametric functions. The choices for these functions were partly guided by experiences from previous parametrizations [3].

The resonance integrals  $\beta_{\mu\lambda}$  were initially assumed to be proportional to the overlap integrals  $S_{\mu\lambda}$  [1–3] which mostly led to acceptable results. In general, however, the following expression proved to be superior:

$$\beta_{\mu\lambda} = Z_{\text{eff}}^A (\mu^A \lambda^B, A) + Z_{\text{eff}}^B (\mu^A \lambda^B, B). \quad (5)$$

The matrix elements  $(\mu^A \lambda^B, A)$  and  $(\mu^A \lambda^B, B)$  represent the attraction between the electron distribution  $\mu^A \lambda^B$  and a unit charge at the positions of the nuclei  $A$  and  $B$ , respectively. These matrix elements are evaluated analytically [24] while the effective charges  $Z_{\text{eff}}^A$  and  $Z_{\text{eff}}^B$  are treated as adjustable parameters. As a possible refinement of Eq. (5), inclusion of an additional term  $\beta_{\text{kin}} T_{\mu\lambda}$  was considered where  $\beta_{\text{kin}}$  is a scaling factor between 0 and 1 and  $T_{\mu\lambda}$  the exact matrix element [24] of the kinetic energy operator. No substantial improvement, however, was obtained for various choices of  $\beta_{\text{kin}}$ . Therefore the kinetic energy term was omitted in the final expression, Eq. (5), for the resonance integral.

An adequate representation of the core-electron attractions  $V_{\mu\nu,B}$  is crucial in semiempirical methods because the diagonal terms  $V_{\mu\mu,B}$  are essential for the prediction of the charge distribution and for the balance between attractive and repulsive forces in the molecule.

For electron distributions  $\mu^A \nu^A$  which involve at least one core AO, the analytical evaluation [24] of the interaction with the nuclear charge  $Z_{Be}$  was found to be

satisfactory, and is therefore adopted:

$$V_{\mu\nu,B}^{\text{th}} = Z_B(\mu^A\nu^A, B). \quad (6)$$

For valence electron distributions  $\mu^A\nu^A$ , however, this approach is not feasible because it leads to an overestimate of the attractions and hence to molecular geometries with very short bond lengths. These problems have been encountered previously [25, 26], and they are usually met by the neglect of penetration integrals. In analogy to MNDO [3], the core-electron attractions  $V_{\mu\nu,B}^{\text{pen}}$  may thus be expressed in terms of the two-center two-electron integrals. Using the indices  $c$  and  $v$  for core and valence AOs, respectively, one may write in the case of the STO-3G basis:

$$V_{\mu\nu,B}^{\text{pen}} = -Z_B(\mu^A\nu^A, s_v^B s_v^B). \quad (7)$$

This expression, however, underestimates the core-electron attractions relative to the electronic repulsions because  $(\mu^A\nu^A, s_c^B s_c^B)$  is normally greater than  $(\mu^A\nu^A, s_v^B s_v^B)$ . We have therefore adopted the following interpolation scheme in the case of the valence electron distributions:

$$V_{\mu\nu,B} = V_{\mu\nu,B}^{\text{pen}} + \gamma(V_{\mu\nu,B}^{\text{th}} - V_{\mu\nu,B}^{\text{pen}}). \quad (8)$$

The calculation of the scaling factor  $\gamma$  is based on the criterion that the interactions between all valence electrons at atom  $A$  and all electrons at atom  $B$  should balance the corresponding interactions with nucleus  $B$  (and vice versa). Denoting the number of core electrons by  $n_c$  and of valence electrons by  $n_v$ , the following expressions are used:

$$\gamma = (V_{\text{sum}}^{\text{rep}} - V_{\text{sum}}^{\text{pen}})/(V_{\text{sum}}^{\text{th}} - V_{\text{sum}}^{\text{pen}}) \quad (9)$$

$$V_{\text{sum}}^{\text{rep}} = -n_v^A n_c^B (s_v^A s_v^A, s_c^B s_c^B) - n_c^A n_v^B (s_c^A s_c^A, s_v^B s_v^B) - 2n_v^A n_v^B (s_v^A s_v^A, s_v^B s_v^B) \quad (10)$$

$$V_{\text{sum}}^{\text{pen}} = -(n_v^A Z_B + n_v^B Z_A)(s_v^A s_v^A, s_v^B s_v^B) \quad (11)$$

$$V_{\text{sum}}^{\text{th}} = n_v^A Z_B (s_v^A s_v^B, B) + n_v^B Z_B (s_v^B s_v^B, A). \quad (12)$$

Eq. (7)–(12) refer to the STO-3G basis. For the split-valence 4-31G basis, completely analogous expressions are employed; Eq. (7) and (10)–(12) are modified by introducing the appropriate average over the two valence shells assigning  $n_v/2$  electrons to each of the two shells.

Several other parametric functions for the core-electron attractions have been tested. The scheme described above turned out to be superior to the alternatives investigated in that it provided a better balance between attractions and repulsions in the molecule and more realistic charge distributions. It has the additional advantage that it does not involve a single adjustable parameter.

Having defined all terms in the Fock matrix, the total energy  $E_{\text{tot}}$  of the molecule can be computed as the sum of the electronic energy  $E_{\text{el}}$ , see Eq. (2), and the nuclear repulsions  $E_{AB}$ . As a consequence of the approximations for the core-electron integrals, the nuclear repulsions cannot be represented by their point-charge values, but must be treated as parametric functions [10]. In analogy to

MNDO [3], they are calculated from the expression

$$E_{AB} = Z_A Z_B (s_v^A s_v^A, s_v^B s_v^B) (1 + e^{-\alpha_A R_{AB}} + e^{-\alpha_B R_{AB}}) \quad (13)$$

where  $\alpha_A$  and  $\alpha_B$  are adjustable parameters and  $R_{AB}$  the internuclear distance. Eq. (13) is valid for the STO-3G basis; for the 4-31G basis, the term  $(s_v^A s_v^A, s_v^B s_v^B)$  is replaced by the appropriate average of valence-shell repulsion integrals. Several alternative functions for  $E_{AB}$  have been studied, but none of them proved to be superior to the MNDO-type expression of Eq. (13).

The heat of formation  $\Delta H_f^{mol}$  of the molecule is obtained from its total energy  $E_{tot}$  by subtracting the electronic energies  $E_{el}^A$  and adding the experimental heats of formation  $\Delta H_f^A$  of the atoms in the molecule [3, 27]:

$$\Delta H_f^{mol} = E_{tot} - \sum_A E_{el}^A + \sum_A \Delta H_f^A. \quad (14)$$

The electronic energies  $E_{el}^A$  of the atoms are determined from half-electron [28] SCF calculations. It has been checked, however, that the results are almost unchanged when optimized atomic energies from UHF calculations [7, 8] are used throughout.

This completes the description of the NDDO SCF calculations with the STO-3G and 4-31G basis sets. As discussed above, all two-electron repulsion integrals are presently evaluated analytically. Therefore, the effects of electron correlation may be expected to be more important than in other methods [2, 3] which feature a semiempirical reduction of the two-electron integrals to account for correlation effects in an average manner. To examine this possibility further, electron correlation is included at the STO-3G level by a second-order Rayleigh-Schrödinger perturbation treatment [29-31] with Møller-Plesset [32] energy denominators (RSMP2). This yields a correlation correction  $E_{corr}^{(2)}$  to the total energy:

$$E_{corr}^{(2)} = - \sum_{ij \rightarrow uv} \frac{|\langle \Psi_0 | \mathcal{H} | \Psi_{ij}^{uv} \rangle|^2}{\epsilon_u + \epsilon_v - \epsilon_i - \epsilon_j} \quad (15)$$

where the summation extends over all double excitations  $ij \rightarrow uv$  from occupied valence MOs  $i, j$  to virtual MOs  $u, v$ .  $\Psi_0$  denotes the SCF configuration,  $\Psi_{ij}^{uv}$  a doubly-excited spin-adapted configuration, and  $\mathcal{H}$  the Hamiltonian operator. The matrix elements appearing in Eq. (15) are available in the literature [30].

Concluding this section we briefly discuss some computational aspects. For the NDDO SCF calculations, both with the STO-3G and 4-31G basis, the rate-determining step is the diagonalization of the Fock matrix which accounts for more than half of the computation time used. The RSMP2 correlation treatment requires less time than an SCF calculation for all molecules studied; for larger molecules, however, it is expected to become the most time-consuming step. Taking propane as an example, the relative computation times are: MNDO 1.0, STO-3G NDDO SCF 1.9, STO-3G NDDO RSMP2 2.4, and 4-31G NDDO SCF 11.9; estimated *ab initio* SCF times are: STO-3G 30, 4-31G 200. The present NDDO SCF calculations are thus intermediate in complexity between current

semiempirical and *ab initio* methods. While the NDDO calculations with the STO-3G basis are still fairly efficient, those with the larger 4-31G basis appear to be rather slow and would thus require a considerable gain in accuracy for their justification.

### 3. Parametrization

The parametrizations were carried out in analogy to the MNDO one, using a nonlinear least-squares optimization procedure [33] to minimize the deviations between calculated results and reference data. Since the MNDO parametrization has been described in detail [3], we shall only remark on some essential aspects of the present parametrizations, and refer to our previous discussion [3] for further information. Parameters were determined for hydrocarbons.

For the STO-3G basis, the standard molecules used in the parametrization were hydrogen, methane, ethane, ethylene, acetylene, cyclopropane, cyclobutane, and benzene. The reference properties included experimental heats of formation (all molecules) and geometrical variables (all molecules except cyclobutane). The calculations were carried out at the experimental geometries. Since the gradients of the energy with respect to the geometrical variables were minimized, the parameters were adjusted to reproduce equilibrium structures without optimizing molecular geometries.

For the 4-31G basis, the first five molecules in the above list were used in the parametrization. Two sets of reference data were employed: The first set comprised experimental heats of formation and geometries, and the second one total energies and geometries from 4-31G *ab initio* SCF calculations [34]. The final parameters were obtained from the second set (see Chapter 5).

Table 1 contains the optimized parameters for the three approaches studied, i.e. the orbital exponents  $\zeta$ , the effective charges  $Z_{\text{eff}}$ , and the repulsion parameters  $\alpha$ .

For the STO-3G basis, the optimized exponents for the valence orbitals are rather low and differ strongly from the corresponding *ab initio* scaling factors (H 1.24, C

**Table 1.** Optimized parameters

	STO-3G NDDO SCF	STO-3G NDDO RSMP2	4-31G NDDO SCF
$\zeta_{\text{H}}$ , a.u.	1.045601	1.150000	—
$\zeta_{\text{C}}$ , a.u.	1.536932	1.564960	—
$Z_{\text{eff}}^{\text{H}}$ , a.u.	0.140239	0.167821	0.137555, 0.096350
$Z_{\text{eff}}^{\text{C}}$ , a.u.	0.291210	0.296160	0.300907, 0.178359
$Z_{\text{eff}}^{\text{C}\pi}$ , a.u.	0.380830	0.362974	—
$\alpha_{\text{H}}^{\text{H}}$ , $\text{\AA}^{-1}$	3.620510	3.010159	3.967564
$\alpha_{\text{H}}^{\text{HC}}$ , $\text{\AA}^{-1}$	2.936650	2.799700	3.080712
$\alpha_{\text{C}}$ , $\text{\AA}^{-1}$	3.103572	3.137020	3.299399

1.72 [7]). The exponent for the carbon 1s AO is assumed to be the same as in the *ab initio* case [7]. As indicated in Table 1, the carbon 2p $\pi$  AO is assigned an effective charge  $Z_{\text{eff}}^{C\pi}$  which is different from that for the  $\sigma$  AOs,  $Z_{\text{eff}}^C$ . This was found necessary to reproduce the relative stabilities of saturated and unsaturated systems. Similar distinctions between  $\sigma$  and  $\pi$  AOs in a local coordinate system have previously been introduced in the CNDO/S [35] and LNDO/S [36] methods. With regard to the  $\alpha$  parameters, two different values are used for hydrogen in H—H and H—C atom pairs ( $\alpha_{\text{H}}^{\text{HH}}$ ,  $\alpha_{\text{H}}^{\text{HC}}$ ). This refinement serves to adjust the H—H bond length in the hydrogen molecule, but it is not essential for other properties. Comparing the two parameter sets for STO-3G NDDO SCF and STO-3G NDDO RSMP2 as a whole, the overall similarity is remarkable. The higher orbital exponents for RSMP2 are due to the fact that electron correlation tends to increase bond lengths [37, 38] which is compensated by higher exponents.

**Table 2.** Heats of formation (in kcal/mol)

Molecule	Exp. <sup>a</sup>	STO-3G <sup>b</sup>	STO-3G <sup>b</sup>
		NDDO SCF	NDDO RSMP2
hydrogen	0	-1.5	-0.6
methane	-17.9	-12.7	-13.7
ethane	-20.2	-21.8	-19.8
ethylene	12.5	22.8	23.0
acetylene	54.3	40.2	43.4
propane	-24.8	-30.6	-26.7
propene	4.9	9.4	10.7
propyne	44.4	23.7	26.9
allene	45.6	54.2	56.3
<i>n</i> -butane	-30.4	-39.3	-33.6
isobutane	-32.4	-38.9	-33.9
1-butene	-0.2	1.0	3.9
2-butene, trans	-3.0	-4.0	-1.5
2-butene, cis	-1.9	-3.6	-0.5
isobutene	-4.3	-3.1	-1.5
1,3-butadiene	26.0	39.8	39.8
1-butyne	39.5	14.2	18.6
2-butyne	34.7	7.1	10.4
but-1-yn-3-ene	72.8	56.1	58.2
butadiyne	113.0	78.5	83.8
cyclopropane	12.7	20.7	22.1
cyclopropene	66.2	82.1	81.0
methylene cyclopropane	47.9	54.1	55.0
cyclobutane	6.8	-9.3	-8.1
cyclobutene	37.5	40.7	37.0
bicyclobutane	51.9	79.3	77.2
benzene	19.8	24.8	22.1

<sup>a</sup> For references to experimental data, see Table II of Ref. [51].

<sup>b</sup> Calculated at optimized geometries.

For the 4-31G basis, the standard scaling factors from the *ab initio* calculations are employed [8] since the optimization of these factors led to no substantial improvement. With respect to the effective charges, the first value for each atom (Table 1) refers to the core and inner valence AOs, and the second one to the other valence AOs; there is no special treatment of  $2p\pi$  AOs. The use of two different values seems reasonable for a split-valence basis set since the diffuse outer valence AOs are expected to see a different effective charge than the inner valence AOs.

#### 4. Results and Discussion: STO-3G Basis

Table 2 compares calculated and observed heats of formation for a set of 27 hydrocarbon molecules. The mean absolute error for these compounds is 10.8 kcal/mol in the NDDO SCF calculations, and 9.4 kcal/mol in the NDDO RSMP2 calculations which is about twice as high as the average MNDO error (4.6 kcal/mol for the molecules in Table 2). Some systematic deviations from the experimental values are obvious: In particular, molecules with C=C bonds are usually predicted too unstable, and those with C≡C bonds too stable<sup>†)</sup>.

A comparison with *ab initio* STO-3G results is not feasible for heats of formation which cannot be expected to be reproduced by *ab initio* MO theory due to the neglect of electron correlation. With regard to relative energies, the *ab initio* STO-3G method has mainly been applied to the calculation of the energies of isodesmic reactions [39, 40], e.g. bond separation reactions



with stoichiometric factors  $a, b, c, d \geq 0$ . These reactions preserve the number of formal bonds and thus provide a measure for the interaction between bonds.

Table 3 lists the energies of bond separation reactions for 24 hydrocarbon molecules, the *ab initio* values being calculated from literature data [34, 41, 42]. Relative to experiment the mean absolute error for the calculated energies is 10.1 kcal/mol for *ab initio* SCF, 10.3 kcal/mol for NDDO SCF, and 6.8 kcal/mol for NDDO RSMP2. A closer comparison between the *ab initio* and NDDO SCF results reveals striking analogies: With one exception, the bond separation energies always show the same sign, and the values for almost each molecule are very similar, particularly also for ring compounds (Table 3). In fact, the average absolute deviation between the two sets of SCF results is only 3.1 kcal/mol, i.e. far below the error compared to experiment (see above). This seems quite remarkable because the NDDO SCF method has been parametrized to reproduce experimental data, and not *ab initio* SCF results (see Chapter 3).

Table 4 contains a statistical analysis of the results for bond lengths and bond angles. Table 5 shows the optimized geometries of 25 hydrocarbon molecules,

<sup>†)</sup> This is related to the effective atomic charges chosen for carbon (Table 1). For  $Z_{\text{eff}}^C = Z_{\text{eff}}^{C\pi}$ , unsaturated systems are generally too high in energy relative to saturated ones. The necessary correction  $Z_{\text{eff}}^C < Z_{\text{eff}}^{C\pi}$  removes most of this error but it stabilizes triple bonds somewhat more than double bonds.



**Table 3.** Energies of bond separation reactions (in kcal/mol)

Molecule	Exp. <sup>a</sup>	STO-3G <sup>b,c</sup>	STO-3G <sup>b</sup>	STO-3G <sup>b</sup>
		<i>ab initio</i> SCF	NDDO SCF	NDDO RSMP2
propane	2.2	0.6	-0.3	0.8
propene	5.2	4.4	4.3	6.2
propyne	7.6	8.4	7.4	10.4
allene	-2.8	0.4	4.2	3.4
<i>n</i> -butane	5.5	0.7 <sup>d</sup>	0.6	1.6
isobutane	7.5	0.6 <sup>d</sup>	0.2	1.9
1-butene	8.0	3.2 <sup>d</sup>	3.7	6.9
2-butene, trans	10.8	7.0 <sup>d</sup>	8.7	12.3
2-butene, cis	9.7	5.3 <sup>d</sup>	8.3	11.3
isobutene	12.1	6.8 <sup>d</sup>	7.8	12.3
1,3-butadiene	14.4	12.5	9.5	13.8
1-butyne	10.1	6.6 <sup>d</sup>	7.8	12.6
2-butyne	14.9	13.7 <sup>d</sup>	14.9	20.8
butatriene	—	3.7	6.7	7.9
but-1-yn-3-ene	9.5	14.8	10.6	15.8
butadiyne	11.2	19.3	5.5	10.6
cyclopropane	-19.8	-45.1	-47.9	-40.4
cyclopropene	-40.6	-65.6	-64.7	-56.5
methylene cyclopropane	-24.6	-46.4	-45.8	-36.6
cyclobutane	-16.2	-27.2	-27.0	-16.3
cyclobutene	-14.2	-28.1	-32.4	-18.6
cyclobutadiene	—	-65.7	-65.2	-44.2
bicyclobutane	-45.8	-105.3	-111.9	-94.0
tetrahydrane	—	-197.3	-215.9	-189.6

<sup>a</sup> Calculated from the experimental data in Table 2. The values refer to 298 K.

<sup>b</sup> Calculated at optimized geometries, unless noted otherwise.

<sup>c</sup> Computed from published total energies [34, 41, 42].

<sup>d</sup> Partial geometry optimization, flexible rotor approximation [43]. Full optimization would tend to increase the *ab initio* value.

**Table 4.** Mean absolute errors for optimized geometries<sup>a</sup>

Geometrical variable	No.	STO-3G	STO-3G	STO-3G	MNDO
		<i>ab initio</i> SCF	NDDO SCF	NDDO RSMP2	
C—H, Å	19	0.005	0.004	0.004	0.011
C—C, Å	22	0.016	0.017	0.017	0.012
C=C, Å	13	0.024	0.013	0.013	0.010
C≡C, Å	4	0.038	0.040	0.006	0.010
All CC, Å	39	0.021	0.018	0.014	0.011
All bond lengths, Å	59	0.016	0.014	0.011	0.012
All angles at C, deg	25	0.8	1.2	1.0	1.9

<sup>a</sup> Based on the results in Table 5.

**Table 5.** Molecular geometries<sup>a</sup>: STO-3G basis

Molecule	Point group	Variable	Exp. <sup>b</sup>	STO-3G <i>ab initio</i> SCF	STO-3G NDDO SCF	STO-3G NDDO RSMP2	Ref. <sup>c</sup>
H <sub>2</sub>	<i>D</i> <sub>∞h</sub>	H—H	0.742	0.712	0.726	0.737	34
CH <sub>4</sub>	<i>T</i> <sub>d</sub>	C—H	1.085 <sup>d</sup>	1.083	1.078	1.085	34
C <sub>2</sub> H <sub>6</sub>	<i>D</i> <sub>3d</sub>	C—C	1.536	1.538	1.514	1.526	34
		C—H	1.091	1.086	1.088	1.093	
		HCC	110.9	110.7	110.9	110.5	
C <sub>2</sub> H <sub>4</sub>	<i>D</i> <sub>2h</sub>	C=C	1.339	1.306	1.311	1.338	34
		C—H	1.086	1.082	1.082	1.085	
		HCC	121.2	122.1	122.8	122.1	
C <sub>2</sub> H <sub>2</sub>	<i>D</i> <sub>∞h</sub>	C≡C	1.203	1.168	1.161	1.194	34
		C—H	1.060	1.065	1.058	1.064	
propane <sup>e</sup>	<i>C</i> <sub>2v</sub>	C <sup>1</sup> —C <sup>2</sup>	1.526	1.541	1.523	1.531	41
		C <sup>1</sup> —H <sup>4</sup>	1.089	1.086	1.087	1.094	
		C <sup>1</sup> —H <sup>5</sup>	1.094	1.086	1.087	1.093	
		C <sup>2</sup> —H <sup>7</sup>	1.096	1.089	1.096	1.100	
		C <sup>1</sup> C <sup>2</sup> C <sup>3</sup>	112.4	112.4	113.8	113.1	
		H <sup>7</sup> C <sup>2</sup> H <sup>8</sup>	106.1	107.2	107.9	107.4	
		H <sup>4</sup> C <sup>1</sup> C <sup>2</sup>	111.8	110.7	110.2	110.0	
		H <sup>5</sup> C <sup>1</sup> H <sup>6</sup>	107.3	108.2	107.9	108.3	
propene <sup>e</sup>	<i>C</i> <sub>s</sub>	C <sup>1</sup> =C <sup>2</sup>	1.336	1.308	1.319	1.343	41
		C <sup>2</sup> —C <sup>3</sup>	1.501	1.520	1.507	1.510	
		C <sup>1</sup> —H <sup>4</sup>	1.081	1.081	1.081	1.086	
		C <sup>1</sup> —H <sup>5</sup>	1.091	1.081	1.081	1.085	
		C <sup>2</sup> —H <sup>6</sup>	1.090	1.085	1.091	1.093	
		C <sup>3</sup> —H <sup>7</sup>	1.085	1.085	1.087	1.092	
		C <sup>3</sup> —H <sup>8</sup>	1.098	1.088	1.088	1.093	
		C <sup>1</sup> C <sup>2</sup> C <sup>3</sup>	124.3	125.1	125.4	124.5	
		H <sup>4</sup> C <sup>1</sup> C <sup>2</sup>	121.5	121.9	122.3	121.8	
		H <sup>5</sup> C <sup>1</sup> C <sup>2</sup>	120.5	122.2	123.2	122.7	
		H <sup>6</sup> C <sup>2</sup> C <sup>1</sup>	119.0	119.8	120.3	119.9	
		H <sup>7</sup> C <sup>3</sup> C <sup>2</sup>	111.2	111.1	112.0	111.6	
H <sup>8</sup> C <sup>3</sup> H <sup>9</sup>	106.2	107.6	108.0	108.6			
propyne	<i>C</i> <sub>3v</sub>	C <sup>1</sup> ≡C <sup>2</sup>	1.206	1.170	1.168	1.202	41
		C <sup>2</sup> —C <sup>3</sup>	1.459	1.484	1.484	1.484	
		C <sup>1</sup> —H	1.056	1.064	1.057	1.064	
		C <sup>3</sup> —H	1.105	1.088	1.086	1.092	
		HC <sup>3</sup> C <sup>2</sup>	110.2	110.5	109.1	108.7	
allene	<i>D</i> <sub>2d</sub>	C=C	1.308	1.288	1.300	1.320	41
		C—H	1.087	1.083	1.081	1.086	
		HCH	118.2	116.2	116.4	117.8	
<i>n</i> -butane	<i>C</i> <sub>2</sub>	C <sup>1</sup> —C <sup>2</sup>	1.533	1.54 <sup>f</sup>	1.523	1.531	43
		C <sup>2</sup> —C <sup>3</sup>	1.533	1.54 <sup>f</sup>	1.532	1.537	
		C <sup>1</sup> C <sup>2</sup> C <sup>3</sup>	112.4	112.2	113.2	112.6	

Table 5. Continued

Molecule	Point group	Variable	Exp. <sup>b</sup>	STO-3G <i>ab initio</i> SCF	STO-3G NDDO SCF	STO-3G NDDO RSMP2	Ref. <sup>c</sup>
isobutane	$C_{3v}$	C—C	1.525	1.54 <sup>f</sup>	1.532	1.536	43
		CCC	111.2	110.8	111.5	112.0	
2-butene	$C_{2h}$	C <sup>1</sup> —C <sup>2</sup>	1.508	1.52 <sup>f</sup>	1.506	1.510	43
		C <sup>2</sup> =C <sup>3</sup>	1.347	1.34 <sup>f</sup>	1.327	1.350	
		C <sup>1</sup> C <sup>2</sup> C <sup>3</sup>	123.8	124.5	124.9	124.2	
isobutene	$C_{2v}$	C <sup>1</sup> =C <sup>2</sup>	1.330	1.34 <sup>f</sup>	1.327	1.350	43
		C <sup>2</sup> —C <sup>3</sup>	1.507	1.52 <sup>f</sup>	1.517	1.518	
		C <sup>1</sup> C <sup>2</sup> C <sup>3</sup>	122.4	122.2	122.0	121.6	
1,3-butadiene	$C_{2h}$	C <sup>1</sup> =C <sup>2</sup>	1.341	1.313	1.321	1.347	42
		C <sup>2</sup> —C <sup>3</sup>	1.463	1.488	1.496	1.487	
		C <sup>1</sup> C <sup>2</sup> C <sup>3</sup>	123.3	124.2	123.7	122.9	
2-butyne	$D_{3d}$	C <sup>1</sup> —C <sup>2</sup>	1.467	1.484	1.482	1.483	44
		C <sup>2</sup> ≡C <sup>3</sup>	1.213	1.171	1.174	1.209	
butatriene	$D_{2h}$	C <sup>1</sup> =C <sup>2</sup>	1.318	1.296	1.305	1.331	42
		C <sup>2</sup> =C <sup>3</sup>	1.283	1.257	1.279	1.285	
but-1-yn-3-ene	$C_s$	C <sup>1</sup> ≡C <sup>2</sup>	1.208	1.171	1.168	1.203	42
		C <sup>2</sup> —C <sup>3</sup>	1.431	1.459	1.474	1.464	
		C <sup>3</sup> =C <sup>4</sup>	1.341	1.320	1.319	1.344	
		C <sup>2</sup> C <sup>3</sup> C <sup>4</sup>	123.1	124.0	122.5	121.7	
butadiyne	$D_{\infty h}$	C <sup>1</sup> ≡C <sup>2</sup>	1.205 <sup>g</sup>	1.175	1.167	1.202	42
		C <sup>2</sup> —C <sup>3</sup>	1.376	1.408	1.439	1.424	
cyclopropane	$D_{3h}$	C—C	1.510	1.502	1.516	1.523	41
		C—H	1.089	1.081	1.084	1.089	
		HCH	115.1	113.8	112.2	112.8	
cyclopropene	$C_{2v}$	C <sup>1</sup> =C <sup>2</sup>	1.296	1.277	1.302	1.368	41
		C <sup>1</sup> —C <sup>3</sup>	1.509	1.493	1.511	1.521	
		C <sup>1</sup> —H	1.072	1.075	1.070	1.076	
		C <sup>3</sup> —H	1.088	1.087	1.090	1.093	
		HC <sup>1</sup> C <sup>2</sup>	149.9	150.3	151.4	150.7	
		HC <sup>3</sup> H	114.6	112.5	112.9	114.2	
methylene cyclopropane	$C_{2v}$	C <sup>1</sup> =C <sup>4</sup>	1.332 <sup>h</sup>	1.298	1.309	1.330	42
		C <sup>1</sup> —C <sup>2</sup>	1.457	1.474	1.499	1.501	
		C <sup>2</sup> —C <sup>3</sup>	1.542	1.522	1.529	1.540	
cyclobutane	$D_{2d}$	C—C	1.548	1.554	1.545	1.547	42
		C <sup>1</sup> C <sup>2</sup> C <sup>4</sup> C <sup>3</sup>	153.0	173.1	180.0	180.0	
cyclobutene	$C_{2v}$	C <sup>1</sup> =C <sup>2</sup>	1.342	1.314	1.337	1.361	42
		C <sup>1</sup> —C <sup>4</sup>	1.517	1.526	1.533	1.532	
		C <sup>3</sup> —C <sup>4</sup>	1.566	1.565	1.560	1.562	
bicyclobutane	$C_{2v}$	C <sup>1</sup> —C <sup>2</sup>	1.498	1.501	1.519	1.523	42
		C <sup>1</sup> —C <sup>3</sup>	1.497	1.469	1.530	1.560	
		C <sup>2</sup> C <sup>3</sup> C <sup>1</sup> C <sup>4</sup>	121.7	117.4	122.5	124.6	

Table 5. Continued

Molecule	Point group	Variable	Exp. <sup>b</sup>	STO-3G <i>ab initio</i> SCF	STO-3G NDDO SCF	STO-3G NDDO RSMP2	Ref. <sup>c</sup>
benzene	$D_{6h}$	C=C	1.397	1.387	1.403	1.412	45
		C—H	1.084	1.083	1.086	1.091	

<sup>a</sup> Bond lengths  $A^iB^j$  (in Å), bond angles  $A^iB^jC^k$  (in deg), dihedral angles  $A^iB^jC^kD^l$  (in deg) of A–B with respect to C–D measured counterclockwise along B–C, and angles  $A^i-B^jC^kD^l$  (in deg) of A–B with plane BCD. The superscripts  $i, j, k, l$  number the atoms in the molecule according to IUPAC rules; they are left out when the atoms can be specified unambiguously without them.

<sup>b</sup> For references to experimental data, see Table VII of Ref. [51], unless noted otherwise.

<sup>c</sup> References to *ab initio* results.

<sup>d</sup> Gray, D. L., Robiette, A. G.: Mol. Phys. **37**, 1901 (1979).

<sup>e</sup> For the numbering of the H atoms, see Table VII of Ref. [51].

<sup>f</sup> Assumed standard bond length, not optimized.

<sup>g</sup> Assumed value.

<sup>h</sup> Laurie, V. W., Stigliani, W. M.: J. Am. Chem. Soc. **92**, 1485 (1970).

along with experimental data and *ab initio* STO-3G results from the literature [34, 41–45]. Full structures are given for small molecules, but usually only the molecular skeletons are given for the larger ones.

The NDDO SCF and NDDO RSMP2 results are in excellent agreement with experiment. Bond lengths are reproduced with a similar accuracy as in MNDO whereas the average deviations for bond angles are reduced by a factor of about 2 (Table 4). The present NDDO calculations thus provide the most accurate semiempirical structural predictions.

A comparison between the *ab initio* and NDDO SCF geometries again shows remarkable agreement. The calculated structures for particular molecules are usually quite similar (Table 5), even with regard to the small errors encountered, e.g. C=C and C≡C bonds being somewhat too short and C—C bonds between multiple bond somewhat too long. The average absolute deviations between the two sets of SCF results are 0.009 Å for bond lengths and 0.6° for bond angles which is again below the errors compared to experiment (Table 4).

For the molecules studied, the first ionization potential obtained from NDDO SCF via Koopmans' theorem [46] is usually too low, typically by 0.5–1.0 eV and particularly for unsaturated systems. The calculated values often lie between the experimental and the *ab initio* STO-3G values, the deviations showing similar trends as in the *ab initio* case. The ordering of the higher ionizations is normally satisfactory.

Judging from the available evidence, the present NDDO SCF calculations with the STO-3G basis tend to reproduce the *ab initio* STO-3G results more closely than the experimental data for which they have been parametrized. This suggests that the underlying mathematical framework is rather similar in both methods, probably due to the analytical evaluation of all one-center terms and all two-

electron integrals in the present NDDO calculations. This similarity seems to influence the semiempirical results more strongly than the parametrization procedure.

We now turn to a comparison between the NDDO SCF and NDDO RSMP2 results to assess the importance of electron correlation. As already mentioned, the average absolute errors for heats of formation, bond separation energies, bond lengths, and bond angles (Table 4) are all slightly lower in NDDO RSMP2 than in NDDO SCF. In particular, the bond separation energies for ring compounds and the C $\equiv$ C bond lengths are improved by the explicit inclusion of electron correlation. In general, however, the differences between the NDDO SCF and NDDO RSMP2 results are fairly small, and most trends are quite similar, especially for heats of formation. Hence, we have to conclude that a semiempirical parametrization can adequately account for correlation effects in closed-shell ground-state molecules, even if all two-electron integrals are evaluated analytically.

This view is supported by a more detailed analysis of the correlation energies. Average contributions from C—H, C—C, C=C, and C $\equiv$ C bonds to the NDDO RSMP2 correlation energy can be determined from the results for methane, ethane, ethylene, and acetylene, the C—H value from methane being used for the other molecules, too. Assuming additivity of these bond contributions, the NDDO RSMP2 correlation energies for the molecules in Table 3 have been estimated. The average relative deviations between estimated and calculated NDDO RSMP2 correlation energies are about 5% which amounts to 2.2 kcal/mol on an absolute scale. To a good approximation, the NDDO RSMP2 correlation energies can thus be partitioned into transferable bond contributions. This is exactly the situation where one would expect that the correlation corrections can, on the average, be well incorporated by a semiempirical parametrization at the SCF level. Therefore the similarity between NDDO SCF and NDDO RSMP2 results is not surprising.

Since the preceding argument holds for a theoretical approach with analytical evaluation of the two-electron integrals, it should be even stronger for methods using a semiempirical reduction of the two-electron integrals, e.g. MNDO [3]. The MNDO RSMP2 correlation energies are indeed smaller than the NDDO RSMP2 ones, typically by a factor of 3. Application of the above analysis leads to average deviations between estimated and calculated MNDO RSMP2 correlation energies of about 5% or 0.8 kcal/mol. Hence, the SCF approximation is justified even more for MNDO than for the present NDDO calculations.

## 5. Results and Discussion: 4-31G Basis

Considering the similarities between the *ab initio* and NDDO SCF results for the STO-3G basis, one may expect that the introduction of the 4-31G basis in NDDO will improve the results significantly, as in the *ab initio* case [5]. We have therefore attempted to parametrize the 4-31G NDDO SCF approach with respect to

experimental reference data. In about 20 parametrization runs with various parametric functions, however, we have only been able to find parameter sets which yield results of similar accuracy as in the STO-3G NDDO SCF calculations (Chapter 4), in spite of the higher computational effort involved. The most serious systematic shortcoming of all 4-31G NDDO SCF calculations was the overestimate of the stability of saturated systems relative to unsaturated systems.

In this situation we decided to parametrize this approach directly with respect to *ab initio* 4-31G reference data which would allow for more meaningful methodical comparisons. The optimized parameters in Table 1 and all following results refer to such a parametrization.

Table 6 lists the total energies of 15 small molecules and cations, the *ab initio* data being taken from the literature [34, 41, 47–49]. The average deviation between the *ab initio* and NDDO total energies is 0.022 a.u. (13.9 kcal/mol). This agreement may be considered satisfactory since it is far better than in unparametrized minimal-basis-set NDDO calculations [17–19] designed to mimic *ab initio* methods. On the other hand, the present 4-31G NDDO SCF calculations do not always reliably reproduce the relative energies from the *ab initio* calculations. While there is good agreement for the classical and nonclassical vinyl and ethyl cations, the predictions differ for the relative stabilities of the C<sub>3</sub>H<sub>4</sub> and C<sub>3</sub>H<sub>6</sub> isomers, the three-membered rings being too stable in NDDO (Table 6).

**Table 6.** Total energies (in atomic units)<sup>a</sup>

Molecule <sup>b</sup>	4-31G <i>ab initio</i> SCF	4-31G NDDO SCF	Ref.
H <sub>2</sub>	-1.12683	-1.13081	34
CH <sub>4</sub>	-40.13977	-40.11539	34
C <sub>2</sub> H <sub>6</sub>	-79.11593	-79.12319	34
C <sub>2</sub> H <sub>4</sub>	-77.92216	-77.89037	34
C <sub>2</sub> H <sub>2</sub>	-76.71141	-76.68031	34
propane	-118.09360 <sup>c</sup>	-118.13718	41
propene	-116.90459 <sup>c</sup>	-116.90641	41
propyne	-115.70110	-115.69363	47
allene	-115.69880	-115.66147	47
cyclopropane	-116.88350 <sup>c</sup>	-116.92342	41
cyclopropene	-115.64260	-115.66199	48
C <sub>2</sub> H <sub>3</sub> <sup>+</sup> IV	-76.97973	-76.95042	49
C <sub>2</sub> H <sub>3</sub> <sup>+</sup> V	-76.94913	-76.91745	49
C <sub>2</sub> H <sub>3</sub> <sup>+</sup> VI	-78.19852	-76.18746	49
C <sub>2</sub> H <sub>3</sub> <sup>+</sup> VIII	-78.18680	-78.17331	49

<sup>a</sup> Calculated at optimized geometries, unless noted otherwise.

<sup>b</sup> The cations are labelled as in Ref. [49] IV and VI are the classical forms, and V and VIII the nonclassical bridged isomers.

<sup>c</sup> Calculated at the *ab initio* STO-3G geometry.

Table 7 compares optimized geometries from the 4-31G NDDO SCF calculations to *ab initio* results [34, 49]. The trends of the results are generally rather similar, but occasional discrepancies may reach several hundredths of an angstrom for bond lengths and several degrees for bond angles. In particular, C≡C bond lengths seem to be too short and HCH angles too large in NDDO.

**Table 7.** Molecular geometries<sup>a</sup>: 4-31G basis

Molecule <sup>b</sup>	Point group	Variable	Exp.	4-31G <sup>c</sup> <i>ab initio</i> SCF	4-31G NDDO SCF
H <sub>2</sub>	<i>D</i> <sub>∞h</sub>	H—H	0.742	0.730	0.729
CH <sub>4</sub>	<i>T</i> <sub>d</sub>	C—H	1.085	1.081	1.085
C <sub>2</sub> H <sub>6</sub>	<i>D</i> <sub>3d</sub>	C—C	1.536	1.529	1.525
		C—H	1.091	1.083	1.087
		HCC	110.9	111.2	108.9
C <sub>2</sub> H <sub>4</sub>	<i>D</i> <sub>2h</sub>	C=C	1.339	1.316	1.307
		C—H	1.086	1.073	1.077
		HCC	121.2	122.0	120.8
C <sub>2</sub> H <sub>2</sub>	<i>D</i> <sub>∞h</sub>	C≡C	1.203	1.190	1.163
		C—H	1.060	1.051	1.041
C <sub>2</sub> H <sub>3</sub> <sup>+</sup> IV	<i>C</i> <sub>2v</sub>	C <sup>1</sup> —C <sup>2</sup>		1.263	1.225
		C <sup>1</sup> —H <sup>1</sup>		1.068	1.051
		C <sup>2</sup> —H <sup>2</sup>		1.083	1.080
		C <sup>1</sup> C <sup>2</sup> H <sup>2</sup>		121.1	121.8
C <sub>2</sub> H <sub>3</sub> <sup>+</sup> V	<i>C</i> <sub>2v</sub>	C <sup>1</sup> —C <sup>2</sup>		1.214	1.192
		C <sup>1</sup> —H <sup>1</sup>		1.065	1.053
		C <sup>1</sup> —H <sup>3</sup>		1.301	1.332
		C <sup>1</sup> C <sup>2</sup> H <sup>2</sup>		177.0	174.0
C <sub>2</sub> H <sub>5</sub> <sup>+</sup> VI	<i>C</i> <sub>s</sub>	C <sup>1</sup> —C <sup>2</sup>		1.440	1.444
		C <sup>1</sup> —H <sup>1</sup>		1.077	1.063
		C <sup>2</sup> —H <sup>3</sup>		1.110	1.100
		C <sup>2</sup> —H <sup>4</sup>		1.091	1.077
		H <sup>1</sup> C <sup>1</sup> H <sup>2</sup>		116.4	118.9
		C <sup>1</sup> C <sup>2</sup> H <sup>3</sup>		105.3	102.3
		H <sup>4</sup> C <sup>2</sup> H <sup>5</sup>		113.6	115.1
		C <sup>2</sup> —C <sup>1</sup> H <sup>1</sup> H <sup>2</sup>		177.5	178.4
		C <sup>1</sup> —C <sup>2</sup> H <sup>4</sup> H <sup>5</sup>		136.3	130.0
C <sub>2</sub> H <sub>5</sub> <sup>+</sup> VIII	<i>C</i> <sub>2v</sub>	C <sup>1</sup> —C <sup>2</sup>		1.375	1.396
		C <sup>1</sup> —H <sup>1</sup>		1.074	1.065
		C <sup>1</sup> —H <sup>3</sup>		1.329	1.337
		H <sup>1</sup> C <sup>1</sup> H <sup>2</sup>		117.8	121.3
		C <sup>2</sup> —C <sup>1</sup> H <sup>1</sup> H <sup>2</sup>		176.8	175.8

<sup>a</sup> Footnotes a, b, d in Table 5 also apply to Table 7.

<sup>b</sup> The labelling of the cations and the numbering of atoms in them follow Ref. [49] (cf. footnote b of Table 6).

<sup>c</sup> *Ab initio* results for neutral molecules from Ref. [34], and for cations from Ref. [49].

The calculated ionization potentials at Koopmans' level [46] for the molecules  $\text{H}_2$ ,  $\text{CH}_4$ ,  $\text{C}_2\text{H}_6$ ,  $\text{C}_2\text{H}_4$ ,  $\text{C}_2\text{H}_2$  show similar behaviour. For the valence ionizations, the average absolute deviation between the NDDO and *ab initio* 4-31G results is 0.019 a.u. (0.53 eV). With regard to experiment, both methods predict ionization potentials for  $\sigma$ -MOs which are systematically too high, while those for the  $\pi$ -MOs in  $\text{C}_2\text{H}_4$  and  $\text{C}_2\text{H}_2$  are slightly too low.

The calculated net atomic charges at the carbon atom show reasonable agreement for the molecules  $\text{CH}_4$ ,  $\text{C}_2\text{H}_6$ ,  $\text{C}_2\text{H}_4$ ,  $\text{C}_2\text{H}_2$  (*ab initio*: 6.61, 6.45, 6.33, 6.30, NDDO: 6.61, 6.44, 6.41, 6.38). Closer inspection reveals, however, that the populations of the inner and outer valence shells are different for the two methods, the outer valence shell being more favored in NDDO. Going along the series  $\text{C}_2\text{H}_6$ ,  $\text{C}_2\text{H}_4$ ,  $\text{C}_2\text{H}_2$ , the *ab initio* and NDDO calculations predict changes in the populations of the inner and outer valence shells which are of the same direction, but much more pronounced in the *ab initio* case. This smaller flexibility of the 4-31G NDDO SCF wave function compared to the *ab initio* one may limit the accuracy of the 4-31G NDDO SCF calculations.

Summarizing our results with the 4-31G basis, the present NDDO calculations seem to reproduce *ab initio* 4-31G results better than experimental data. However, the similarities between the *ab initio* and NDDO SCF results for the 4-31G basis are not as close as for the STO-3G basis (see Tables 3–7). To obtain an improved agreement in the 4-31G case it may be necessary to reproduce the detailed structure of the *ab initio* Fock matrix rather than attempting an overall parametrization. Such an approach has e.g. been followed in the PRDDO method [50] at the minimal-basis-set level.

## 6. Conclusions

Semiempirical NDDO calculations with STO-3G and 4-31G sets can be carried out in a consistent manner. The calculations involve the analytical evaluation of all one-center terms and two-electron integrals, and the semiempirical adjustment of the remaining one-electron integrals and the nuclear repulsions. Within this framework, the STO-3G NDDO SCF approach predicts excellent molecular geometries while molecular energies are of similar accuracy as in *ab initio* STO-3G calculations. Explicit inclusion of electron correlation in STO-3G NDDO RSMP2 leads to a slight overall improvement of the results, but without dramatic changes. Contrary to the expectations, the NDDO SCF calculations with the 4-31G basis are not more accurate than those with the STO-3G basis, for the approximations investigated. In general, the present NDDO SCF results tend to reproduce the corresponding *ab initio* results more closely than experimental data.

*Acknowledgement.* This work was supported by the Deutsche Forschungsgemeinschaft and the Verband der Chemischen Industrie (through a Liebigstipendium). The calculations were carried out on the TR 440 computer of the Universität Marburg and the IBM 370-168 computer of the Technische Hochschule Darmstadt.



## References

1. Segal, G. A., ed.: Modern theoretical chemistry, Vol. 7-8. New York: Plenum Press 1977
2. Bingham, R. C., Dewar, M. J. S., Lo, D. H.: *J. Am. Chem. Soc.* **97**, 1285 (1975)
3. Dewar, M. J. S., Thiel, W.: *J. Am. Chem. Soc.* **99**, 4899 (1977)
4. Schaefer, H. F., III: The electronic structure of atoms and molecules. Reading: Addison-Wesley 1972
5. Hehre, W. J.: *Acc. Chem. Res.* **9**, 399 (1976)
6. Schaefer, H. F., III, ed.: Modern theoretical chemistry, Vol. 3-4. New York: Plenum Press 1977
7. Hehre, W. J., Stewart, R. F., Pople, J. A.: *J. Chem. Phys.* **51**, 2657 (1969)
8. Ditchfield, R., Hehre, W. J., Pople, J. A.: *J. Chem. Phys.* **54**, 724 (1971)
9. Pople, J. A., Segal, G. A.: *J. Chem. Phys.* **43**, S129 (1965)
10. Dewar, M. J. S.: The molecular orbital theory of organic chemistry. New York: McGraw-Hill 1969
11. Cook, D. B., Hollis, P. G., McWeeny, R.: *Mol. Phys.* **13**, 553 (1967)
12. Brown, R. D., Roby, K. R.: *Theoret. Chim. Acta (Berl.)* **16**, 175 (1970)
13. Roby, K. R.: *Chem. Phys. Letters* **11**, 6 (1971); **12**, 579 (1972)
14. Chandler, G. S., Grader, F. E.: *Theoret. Chim. Acta (Berl.)* **54**, 131 (1980)
15. Nelson, P. G.: *J. Chem. Res. (S)* 106, (M) 1701 (1980)
16. Thiel, W.: *J.C.S. Faraday II* **76**, 302 (1980)
17. Brown, R. D., Burden, F. R., Williams, G. R.: *Theoret. Chim. Acta (Berl.)* **18**, 98 (1970)
18. Brown, R. D., Burden, F. R., Williams, G. R., Phillips, L. F.: *Theoret. Chim. Acta (Berl.)* **21**, 205 (1971)
19. Chandrasekhar, J., Mehrotra, P. K., Subramanian, S., Manoharan, P. T.: *Theoret. Chim. Acta (Berl.)* **41**, 243 (1976); **52**, 303 (1979)
20. Sustmann, R., Williams, J. E., Dewar, M. J. S., Allen, L. C., Schleyer, P. v. R.: *J. Am. Chem. Soc.* **91**, 5350 (1969)
21. Birner, P., Köhler, H.-J., Weiss, C.: *Chem. Phys. Letters* **27**, 347 (1974)
22. Roothaan, C. C. J.: *Rev. Mod. Phys.* **23**, 69 (1951)
23. Hall, G. G.: *Proc. R. Soc. London, Ser. A*, **205**, 541 (1951)
24. Hehre, W. J., Lathan, W. A., Ditchfield, R., Newton, M. D., Pople, J. A.: GAUSSIAN 70, Program No. 236, Quantum Chemistry Program Exchange, University of Indiana, Bloomington, Indiana
25. Pople, J. A., Segal, G. A.: *J. Chem. Phys.* **44**, 3289 (1966)
26. Pople, J. A., Beveridge, D. L.: Approximate molecular orbital theory. McGraw-Hill: New York 1970
27. Flanigan, M. C., Komornicki, A., McIver, J. W., Jr., in: Ref. 1, Vol. 8, Chapter 1
28. Dewar, M. J. S., Hashmall, J. A., Venier, C. G.: *J. Am. Chem. Soc.* **90**, 1953 (1968)
29. Claverie, P., Diner, S., Malrieu, J. P.: *Int. J. Quant. Chem.* **1**, 751 (1967)
30. Ostlund, N. S., Bowen, M. F.: *Theoret. Chim. Acta (Berl.)* **40**, 175 (1976)
31. Pople, J. A., Binkley, J. S., Seeger, R.: *Int. J. Quant. Chem. Symp.* **10**, 1 (1976)
32. Møller, C., Plesset, M. S.: *Phys. Rev.* **46**, 618 (1934)
33. Weiner, P. K.: Ph.D. Thesis, University of Texas, Austin, 1975
34. Lathan, W. A., Hehre, W. J., Pople, J. A.: *J. Am. Chem. Soc.* **93**, 808 (1971)
35. Del Bene, J., Jaffé, H. H.: *J. Chem. Phys.* **48**, 1807, 4050 (1968)
36. Lauer, G., Schulte, K.-W., Schweig, A.: *J. Am. Chem. Soc.* **100**, 4925 (1978)
37. Lauer, G., Schulte, K.-W., Schweig, A., Thiel, W.: *Theoret. Chim. Acta (Berl.)* **52**, 319 (1979)
38. DeFrees, D. J., Levi, B. A., Pollack, S. K., Hehre, W. J., Binkley, J. S., Pople, J. A.: *J. Am. Chem. Soc.* **101**, 4085 (1979)
39. Hehre, W. J., Ditchfield, R. D., Pople, J. A.: *J. Am. Chem. Soc.* **92**, 4796 (1970)
40. Radom, L., Hehre, W. J., Pople, J. A.: *J. Am. Chem. Soc.* **93**, 289 (1971)
41. Radom, L., Lathan, W. A., Hehre, W. J., Pople, J. A.: *J. Am. Chem. Soc.* **93**, 5339 (1971)
42. Hehre, W. J., Pople, J. A.: *J. Am. Chem. Soc.* **97**, 6941 (1975)
43. Radom, L., Pople, J. A.: *J. Am. Chem. Soc.* **92**, 4786 (1970)
44. Binkley, J. S., Pople, J. A., Hehre, W. J.: *Chem. Phys. Letters* **36**, 1 (1975)

45. Radom, L., Schaefer, H. F., III: *J. Am. Chem. Soc.* **99**, 7522 (1977)
46. Koopmans, T.: *Physica (Utrecht)* **1**, 104 (1934)
47. Hopkinson, A. C., Lien, M. H., Yates, K., Mezey, P. G., Csizmadia, I. G.: *J. Chem. Phys.* **67**, 517 (1977)
48. Hopkinson, A. C., Lien, M. H., Yates, K., Csizmadia, I. G.: *Int. J. Quant. Chem.* **12**, 355 (1977)
49. Hariharan, P. C., Lathan, W. A., Pople, J. A.: *Chem. Phys. Letters* **14**, 385 (1972)
50. Halgren, T. A., Lipscomb, W. N.: *J. Chem. Phys.* **58**, 1569 (1973)
51. Dewar, M. J. S., Thiel, W.: *J. Am. Chem. Soc.* **99**, 4907 (1977)

Received October 6, 1980