

Density functional Gaussian-type-orbital approach to molecular geometries, vibrations, and reaction energies

J. Andzelm and E. Wimmer

Cray Research, Inc., Eagan, Minnesota 55121

(Received 3 July 1991; accepted 3 October 1991)

We present the theory, computational implementation, and applications of a density functional Gaussian-type-orbital approach called DGauss. For a range of typical organic and small inorganic molecules, it is found that this approach results in equilibrium geometries, vibrational frequencies, bond dissociation energies, and reaction energies that are in many cases significantly closer to experiment than those obtained with Hartree–Fock theory. On the local spin density functional level, DGauss predicts equilibrium bond lengths within about 0.02 Å or better compared with experiment, bond angles, and dihedral angles to within 1–2°, and vibrational frequencies within about 3%–5%. While Hartree–Fock optimized basis sets such as the 6-31 G** set can be used in DGauss calculations to give good geometries, the accurate prediction of reaction energies requires the use of density functional optimized Gaussian-type basis sets. Nonlocal corrections as proposed by Becke and Perdew for the exchange and correlation energies are found to be essential in order to predict bond dissociation energies and reaction energies within a few kcal/mol. The computational efficiency of the present method together with its accuracy, which is comparable to correlated Hartree–Fock based methods, promises a great usefulness of the DGauss approach for the study of large and complex molecular structures.

I. INTRODUCTION

One of the central tasks in computational chemistry is the accurate and efficient prediction of molecular geometries, vibrational frequencies, and energies of chemical reactions. Such predictions are of great value in areas such as synthetic chemistry, molecular biology, catalysis, and materials science. From a great number of systematic calculations, mostly on organic molecules, it has been established that useful energetic, structural, vibrational, and electronic properties can be obtained from Hartree–Fock theory, even on the single-configuration self-consistent-field (SCF) level.¹ Today, molecules with about 10–20 carbonlike atoms are being investigated routinely on that level of theory using polarized double-zeta basis sets. For smaller molecules, it is possible to carry out correlated calculations using single-reference as well as multireference wave functions.

Recent advances in the implementation of direct SCF schemes,^{2,3} combined with the increasing performance of computer hardware, are moving the frontier to molecules of increasing size and complexity. Although the computational requirements for Hartree–Fock theory increase formally with a fourth power in the number of basis functions, this scaling is significantly less for large molecules, especially when their geometry is extended. Novel algorithms for integral evaluations have been pursued and implemented successfully.^{4–6} Nevertheless, the calculation of four-index, two-electron integrals remains a major computational bottleneck. Recent developments of pseudospectral methods make it possible to carry out Hartree–Fock calculations with an algorithm that scales with a third power in the number of basis functions.⁷

However, these advances in the implementation of Hartree–Fock calculations do not address the fundamental issue

of electron correlation and the need for multireference representations, which is known to be important for many systems such as transition metal complexes and organometallics, but also for relatively simple molecules such as nitromethane. Inclusion of correlation typically increases the computational requirements dramatically thus limiting the range of correlated methods to rather small molecules. For these reasons, there is an urgent need to explore and develop new theoretical and computational methods that include electron correlation and, at the same time, are practical for large systems.

Today, there is increasing evidence^{8–11} that density functional theory (DFT)^{12–15} offers a promising alternative to the Hartree–Fock approach. DFT includes electron correlation in a form that does not lead to the scaling problem of Hartree–Fock-based methods. Thus, DFT has the potential to be applicable to fairly large systems. While there is sufficient evidence that DFT provides an accurate description of electronic and structural properties of solids, surfaces, and interfaces,¹⁶ relatively little is known about the systematic performance of DFT applied to typical organic molecules.

One of the reasons for this lack of knowledge is the fact that analytic gradient calculations, which are critical for molecular geometry optimizations, have only been developed recently for molecular DFT calculations, while this feature has been available and used widely in Hartree–Fock calculations for well over a decade.¹ Furthermore, a number of well-tested Hartree–Fock-based programs are available to the scientific community, which is not the case for molecular DFT programs. There are also other obstacles and limiting factors: (i) There is no systematic theoretical way to improve the accuracy of the so-called local spin density (LSD) approximation; (ii) besides the expansion of molecular orbi-

tals in finite basis sets, there are issues of adequate representations of the electron density and the exchange-correlation potential; (iii) more than one explicit form for the exchange-correlation potential have been proposed; and (iv) in the past, simplifications in the shape of the potential such as the muffin tin potential have introduced a somewhat uncontrolled approximation and an element of arbitrariness.

Many of the obstacles mentioned above have now been overcome and there is the opportunity for molecular DFT methods to become a useful and accurate tool in the investigation of molecular structures and chemical reactions. The practical implementation of DFT leads to effective one-electron Schrödinger equations or "Kohn-Sham equations" which are very similar to the Hartree-Fock equations, except that in the Kohn-Sham equations, the orbital-dependent exchange operator of the Hartree-Fock equations are formally replaced by an exchange-correlation operator that depends only on the total electron density (and spin density in spin-polarized calculations). This makes the form of the matrix elements simpler and the one-particle wave functions can be represented not only by Gaussian-type orbitals, but also by a variety of other functions such as Slater-type orbitals,^{17,18} numerical functions,¹⁹⁻²¹ plane waves,^{22,23} or augmented plane waves.²⁴⁻²⁶ Completely numerical ("basis set free") solutions of the density functional equations have been proposed as well.²⁷ It turns out, however, that Gaussian-type basis functions are appealing in molecular DFT calculations for the same reasons as in Hartree-Fock theory. The calculation of multiple-center integrals is fast and contracted basis sets can be found which are accurate, but still relatively small and efficient. The simple analytic form of the wave functions enables the analytic calculation of energy gradients. Furthermore, there is a wealth of experience from Hartree-Fock calculations on the choice of basis sets. This experience can be transferred directly to molecular DFT calculations. As pointed out earlier, there is a great need for systematic comparisons between Hartree-Fock-based results and DFT results. Using the same type of basis set facilitates this comparison.

For these reasons, we have developed a highly efficient molecular density functional implementation using a linear combination of Gaussian-type orbitals (LCGTO) and an overall approach that is amenable to systematic comparisons with existing Hartree-Fock and correlated *ab initio* methods. In this way, we hope a bridge can be built between the wealth of experience gained from molecular Hartree-Fock calculations and the newer molecular DFT approach. This similarity should help gain a clearer picture of the strengths and weaknesses of both the Hartree-Fock and the density functional approaches so that each method can be applied in the best possible way to address the many challenging problems of theoretical and computational molecular science. The present work builds conceptually on previous efforts, especially by Sambe and Felton,²⁸ Dunlap,²⁹ and Salahub.³⁰ Currently, the use of Gaussian-type orbitals in density functional calculations is also pursued by a number of groups.²⁹⁻³⁹ For the present implementation, many algorithmic and computational aspects have been reconsidered, further developed, and coded in a new computer program,

which is called DGauss. In essence, the present LCGTO density functional approach relies on a variational, analytic representation of the electron density and the fact that the exchange-correlation terms are a smooth function of the density, which can be accurately fitted by using a numerical integration. Together with efficient algorithms for integral and gradient evaluations, all basic computational components are thus in place for accurate geometry optimizations, evaluation of vibrational frequencies, and computation of reaction energies.

In the following sections of this paper, we will first present the theoretical background and the computational implementation of the present density functional Gaussian-type-orbital approach. Then, various computational levels, including basis sets and grid resolutions, are defined and the sensitivity of the computed molecular properties to these levels is tested for methylamine. The accuracy and reliability of this approach is then illustrated for a number of typical molecules containing C, N, O, H, and F atoms. The results include optimized molecular structures, vibrational frequencies and IR intensities, bond dissociation energies, hydrogenation reactions, and energies of isodesmic reactions. In addition, results are presented for two molecular systems CH_3NO_2 and C_2F_2 , where correlation effects are known to be important. While the present approach is targeted for the study of large molecular structures requiring 1000 basis functions and more, the aim of the present work is the assessment of the accuracy of DGauss. To this end, various classes of small molecules are investigated where systematic comparisons with experiment, Hartree-Fock, and correlated methods are possible.

II. THEORETICAL ASPECTS

A. SCF equations and single point energies

In density functional theory,¹²⁻¹⁵ the total energy, including electron correlation effects, is written in the form

$$E[\rho] = T[\rho] + U[\rho] + E_{xc}[\rho]. \quad (1)$$

Here, T is a kinetic energy term, U is the electrostatic interaction energy between all electrons and nuclei, and E_{xc} is the exchange-correlation energy of the system.

The total electron density ρ in Eq. (1) can be related to single-particle wave functions by

$$\rho(\mathbf{r}) = \sum_{\text{occ}} |\psi_i(\mathbf{r})|^2, \quad (2)$$

where the summation extends over all occupied electronic levels.

A variational principle applied to Eq. (1), together with the definition of the one-particle wave functions (2), leads to effective one-particle Schrödinger equations, usually referred to as Kohn-Sham equations, of the form

$$\mathbf{H}\psi_i = \epsilon_i\psi_i, \quad (3)$$

where \mathbf{H} represents a one-particle Hamiltonian operator, ψ_i are one-electron wave functions [molecular orbitals (MOs)], and ϵ_i can be interpreted as one-electron energies (MO energies). DFT can be generalized to spin-unrestricted systems. In that case, the Hamiltonian operator also depends on the spin and separate MOs and MO energies are obtained

for the spin-up and spin-down levels. For simplicity, we discuss here only the spin-restricted case, although both the spin-restricted and the spin-unrestricted cases are implemented in the present program.

As a consequence of the form for the total energy expression in Eq. (1), the effective one-particle Hamiltonian operator \mathbf{H} can be written as

$$\mathbf{H} \equiv \left[-1/2\nabla^2 + V_C(\mathbf{r}) + \mu_{xc}(\mathbf{r}) \right]. \quad (4)$$

Hartree atomic units are used here with $\hbar^2/(4\pi^2m) = 1$ and $e^2 = 1$. V_C is the electrostatic (or Coulomb) potential

$$V_C = V_N + V_e \quad (5)$$

consisting of the electron-nuclear attraction

$$V_N(\mathbf{r}) = - \sum_{\alpha} Z_{\alpha}/|\mathbf{R}_{\alpha} - \mathbf{r}| \quad (6)$$

and the electron-electron repulsion

$$V_e(\mathbf{r}) = \int \rho(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'|d\mathbf{r}'. \quad (7)$$

Here, \mathbf{R}_{α} denotes the position of atom α with the atomic number Z_{α} . The sum in Eq. (6) extends over all atoms of the molecular system.

The exchange-correlation potential is expressed by the term μ_{xc} which is related to the exchange-correlation energy by

$$\mu_{xc} = \partial E_{xc}/\partial \rho. \quad (8)$$

In the so-called local density approximation (LDA),¹³ the total exchange-correlation energy is approximated by

$$E_{xc} \approx \int \rho(\mathbf{r})\epsilon_{xc}[\rho(\mathbf{r})]d\mathbf{r}, \quad (9)$$

where $\epsilon_{xc}[\rho(\mathbf{r})]d\mathbf{r}$ is the exchange-correlation energy in a volume element $d\mathbf{r}$ in which the local density is $\rho(\mathbf{r})$; $\epsilon_{xc}[\rho]$ is the exchange-correlation energy per electron in a correlated (i.e., interacting) electron system of constant density ρ . In the present implementation, we use the form for $\epsilon_{xc}[\rho]$ as given by Vosko, Wilk, and Nusair.⁴⁰ It is generally believed that this represents one of the best analytic functional forms available for LDA potentials. It should be noted that no empirically adjustable parameters are incorporated in this potential.

The molecular orbitals ψ_i are represented by Gaussians in the same way as in the Hartree-Fock method

$$\psi_i = \sum_p c_{ip}g_p \quad (10)$$

with $\{g_p; p = 1, \dots, N\}$ being a set of contracted Gaussian basis functions.

Following Sambe and Felton,²⁸ the electron density is also expanded in a set of Gaussian-type functions. Because of the finite number of Gaussians in this auxiliary set, this representation amounts to the approximation

$$\rho(\mathbf{r}) \approx \rho'(\mathbf{r}) = \sum_r \rho_r g_r \quad (11)$$

with $\{g_r; r = 1, \dots, N_r\}$ being a set of auxiliary basis functions. Similarly, the exchange-correlation potential $\mu_{xc}(\mathbf{r})$ is expanded in another auxiliary set of Gaussian-type functions

$\{g_s; s = 1, \dots, N_s\}$ in the form

$$\mu_{xc}(\mathbf{r}) = \sum_s \mu_s g_s. \quad (12)$$

Substituting expressions (4)–(12) in Eq. (3) and applying a variational principle analogous to that used in Hartree-Fock theory leads to a system of equations that determines the coefficients in expansion (10):

$$(H_{pq} - \epsilon_i S_{pq})c_{iq} = 0. \quad (13)$$

These equations have to be solved in a self-consistent procedure. The matrix elements in Eq. (13) are given as follows:

$$H_{pq} = h_{pq} + \sum_r \rho_r [pq||r] + \sum_s \mu_s [pqs] \quad (14)$$

and

$$S_{pq} = [pq]. \quad (15)$$

Equations (13) and (14) contain one-electron and two-electron integrals defined by

$$h_{pq} \equiv \int g_p(\mathbf{r}) \left(-1/2\nabla^2 - \sum_{\alpha} Z_{\alpha}/|\mathbf{R}_{\alpha} - \mathbf{r}| \right) g_q(\mathbf{r}) d\mathbf{r}, \quad (16)$$

$$[pq||r] \equiv \int \int g_p(\mathbf{r}) g_q(\mathbf{r}) (1/|\mathbf{r} - \mathbf{r}'|) g_r(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \quad (17)$$

$$[pqs] \equiv \int g_p(\mathbf{r}) g_q(\mathbf{r}) g_s(\mathbf{r}) d\mathbf{r}, \quad (18)$$

$$[pq] \equiv \int g_p(\mathbf{r}) g_q(\mathbf{r}) d\mathbf{r}. \quad (19)$$

It should be noted that the one-electron integrals h_{pq} , as well as the overlap integrals S_{pq} , are identical to those found in the Roothaan-Hartree-Fock equations.¹ Without the density expansion of Eq. (11), the term $\sum_r \rho_r [pq||r]$ of Eq. (14) would contain

$$\sum_r c_{ir} c_{is} \int \int g_p(\mathbf{r}) g_q(\mathbf{r}) (1/|\mathbf{r} - \mathbf{r}'|) g_r(\mathbf{r}') g_s(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \quad (20)$$

which represent the familiar four-index, two-electron Coulomb integrals of the Hartree-Fock theory. The last term in Eq. (14) can be compared formally with the exchange integrals of the Hartree-Fock theory. However, as discussed above, the term includes correlation effects.

Following Dunlap,²⁹ the density fitting coefficients ρ_r in Eq. (11) are defined such that the Coulomb energy

$$\Delta \equiv \iint \delta\rho(\mathbf{r}) (1/|\mathbf{r} - \mathbf{r}'|) \delta\rho(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \quad (21)$$

arising from the difference between the fitted and original density

$$\delta\rho(\mathbf{r}) \equiv \rho(\mathbf{r}) - \rho'(\mathbf{r}) \quad (22)$$

is minimized while maintaining charge conservation. Using the definition of the density matrix \mathbf{P} ,

$$P_{pq} = \sum_i c_{ip} c_{iq} \quad (23)$$

with the summation extending over all occupied molecular orbitals, we find the coefficients ρ_r to be determined by

$$\rho_r = \sum_{r'} C_{rr'}^{-1} \left\{ \sum_{pq} P_{pq} [pq||r'] - \Lambda \int g_r(\mathbf{r}) d\mathbf{r} \right\}. \quad (24)$$

The Lagrange multipliers Λ guarantee charge conservation.²⁹ The matrix \mathbf{C} is defined by its elements

$$C_{rr'} = [r||r'] \equiv \iint g_r(\mathbf{r})(1/|\mathbf{r}-\mathbf{r}'|)g_{r'}(\mathbf{r}')d\mathbf{r}d\mathbf{r}'. \quad (25)$$

All the integrals necessary to calculate the density fitting coefficients ρ_r can be obtained from analytic, Coulomb-type integrals.

The fitting coefficients for the exchange-correlation potential in Eq. (11) are given by the following relations:

$$\mu_s = \sum_{s'} S_{ss'}^{-1} \int g_{s'}(\mathbf{r})\mu_{xc}(\mathbf{r})d\mathbf{r} \quad (26)$$

with $S_{ss'}$ being overlap matrix elements defined in Eq. (12). The evaluation of the integrals in Eq. (26) is done numerically as will be discussed below. Recently, a method for the variational fitting of the exchange-correlation potential has been formulated,⁴¹ but not yet implemented in the present approach.

Based on the self-consistent charge density, Dunlap²⁹ derived an expression for the LDA total energy which is exact to second order in the error of the density fit of Eq. (11). The explicit form, which resembles the corresponding expression of the Hartree-Fock theory, is

$$E_{\text{LDA}} = \sum_{pq} P_{pq} \left(h_{pq} + \sum_r \rho_r [pq||r] + \sum_s \epsilon_s [pqs] \right) - \frac{1}{2} \sum_{rr'} \rho_r \rho_{r'} [r||r'] + U_N. \quad (27)$$

Similar to the expansion of the exchange-correlation potential given in Eq. (12), the exchange-correlation energy needed for the total energy expression (27) is expanded in a set of Gaussian-type functions in the form

$$\epsilon_{xc}(\mathbf{r}) = \sum_s \epsilon_s g_s. \quad (28)$$

In fact, the same set of functions $\{g_s\}$ is used as in the expansion of μ_{xc} . The integrals $[r||r']$ are defined in Eq. (25)

$$U_N = 1/2 \sum_{\alpha\alpha'} Z_\alpha Z_{\alpha'} / |\mathbf{R}_\alpha - \mathbf{R}_{\alpha'}| \quad (29)$$

denotes the Coulomb repulsion energy between all nuclei.

B. Analytic energy gradients

One of the advantages in using Gaussian-type basis functions is the possibility of evaluating the first derivatives of the total energy with respect to nuclear displacements analytically.⁴¹⁻⁴⁷ The energy gradient can be written as

$$\partial E_{\text{LSD}}/\partial x = F_{\text{HFB}} + F_D, \quad (30)$$

where F_{HFB} is the Hellman-Feynman force with a correction arising from the incompleteness of the orbital basis

$$F_{\text{HFB}} = \sum_{pq} P_{pq} \left\{ \partial h_{pq}/\partial x + \sum_r \rho_r [\partial(pq)/\partial x||r] + [\partial(pq)/\partial x \mu_{xc}(\mathbf{r})] \right\} + \partial U_N/\partial x - \sum_{pq} W_{pq} \partial [pq]/\partial x. \quad (31)$$

The term F_D originates from the incompleteness of the density fit

$$F_D = \sum_r \rho_r [\partial(r)/\partial x||(\rho - \rho')] = \sum_r \rho_r \left[\partial(r)/\partial x \left(\sum_{pq} P_{pq} pq - \sum_t \rho_t t \right) \right]. \quad (32)$$

Expressions (31) and (32) contain in essence derivatives of three-index integrals, which can be calculated analytically in the same way as the integrals in the SCF calculations. The exchange-correlation term in Eq. (31) is evaluated by a numerical integration (33) which turns out to be more accurate than using the fitted form (34) of μ_{xc} provided a sufficiently accurate grid is used for the numerical integration

$$[\partial(pq)/\partial x \mu_{xc}(\mathbf{r})] = \int \partial/\partial x [g_p(\mathbf{r})g_q(\mathbf{r})] \mu_{xc}(\mathbf{r})d\mathbf{r} \quad (33) = \sum_s \epsilon_s \int \partial/\partial x [g_p(\mathbf{r})g_q(\mathbf{r})] g_s(\mathbf{r})d\mathbf{r}. \quad (34)$$

Spurious one-center contributions to the exchange-correlation forces are eliminated in a similar way as has been done by Versluis and Ziegler.⁴⁸ The matrix \mathbf{W} is an energy weighted density matrix similar to that used in the gradients in the Hartree-Fock theory.

C. Nonlocal exchange-correlation corrections

It is known that the local density approximation tends to overestimate binding energies.⁹⁻¹¹ One way to correct for this error is the inclusion of nonlocal correction terms using the density gradient. To this end, the exchange-correlation energy is written in the form

$$E_{\text{NLS}} = E_{\text{LSD}} + E_x^G + E_c^G, \quad (35)$$

where E_x^G and E_c^G are gradient corrections to the exchange and the correlation energies, respectively. Two different forms for these gradient corrections are considered here.

One form uses a correction to the exchange energy proposed by Becke^{49,50} and an expression for the correlation corrections given by Langreth and Mehl, which was later refined by Perdew.⁵¹ This approach is referred to as the Becke-Perdew (BP) gradient correction. In the BP correction, the exchange part is written in the form

$$E_x^G = b \sum_\sigma \int \rho_\sigma x_\sigma^2 / (1 + 6bx_\sigma \sinh^{-1} x_\sigma) d\mathbf{r}, \quad (36)$$

where

$$x_\sigma \equiv |\nabla\rho|/\rho_\sigma^{4/3}. \quad (37)$$

The label σ denotes spin up or spin down. Expression (36)

has the correct asymptotic behavior and contains a constant b , which is determined by fitting the exact exchange energies of inert gases. The correlation part in the BP correction has the form

$$E_c^G = \int f(\rho_\alpha, \rho_\beta) \exp\{-g(\rho)\} |\nabla\rho| |\nabla\rho|^2 dr. \quad (38)$$

The definition of the functions f and g is given by Perdew.⁵¹

The other form of the nonlocal corrections to the total energy is based on the work of Becke⁵² and Stoll, Pavlidou, and Preuss⁵³ and thus is referred to as the BSPP approach. Here the exchange energy is due to Becke.⁵² It is asymptotically correct and contains two parameters b and c in the form

$$E_x^G = b \sum_\sigma \int \rho_\sigma^{4/3} x_\sigma^2 / (1 + cx_\sigma^2) dr. \quad (39)$$

The parameters b and c are determined by fitting the exact atomic exchange energies for all atoms in the Periodic Table and using an appropriate average to make the expression independent of any particular atom. The correlation energy in this correction has the form⁵³

$$E_c^G = - \int \rho_\alpha \epsilon_c(\rho_\alpha, 0) dr - \int \rho_\beta \epsilon_c(0, \rho_\beta) dr, \quad (40)$$

where ϵ_c is the LSD correlation energy.

III. COMPUTATIONAL IMPLEMENTATION

A. Basis sets

While Gaussian basis sets are convenient for the calculation of multicenter integrals, great care has to be used in the construction of these basis sets in order to find the right balance between accuracy and efficiency. While the familiar Hartree–Fock optimized basis sets such as Pople's 6-31G** basis set^{1,54} give reasonable molecular geometries when used in molecular DFT calculations, reaction energies are more sensitive and Hartree–Fock optimized basis sets are inadequate. To this end, new all-electron LSD optimized basis sets have been developed⁵⁵ for all elements from H to Xe and are used in the present investigation. The optimization procedure is based on the work of Tatewaki and Huzinaga⁵⁶ as implemented for LSD optimization by Andzelm *et al.*⁵⁷ For the elements B to Ne, this double-zeta basis set with polarization functions (DZVP) is built from nine s -type and five p -type primitive Cartesian Gaussians, augmented by one d -type polarization function. These primitives, written as (621/41/1), are contracted to three s -type and two p -type functions. Together with the polarization functions, this results in a [3/2/1] contracted set. For hydrogen, a (41) set is used in this DZVP set. These LSD-optimized Gaussian basis sets cause only small basis set superposition errors (BSSE) and exhibit a high quality for the valence orbitals, as judged by comparison with energies and orbitals obtained from a numerical procedure.^{47,57}

By adding p -polarization functions to the H basis and using the DZVP basis sets for the other atoms, a so-called DZVPP basis is defined. A more accurate basis set, called DZVP2, is obtained by adding one s - and one p -type primitive to the set for the atoms B to Ne. This results in a (721/51/1) basis set. In the DZVP2 basis, polarization functions are also added to H atoms resulting in a (41/1) basis for hydrogen.

Corresponding to these orbital basis sets are auxiliary basis sets to represent the electron density [Eq. (11)], the exchange-correlation potential [Eq. (12)] and the exchange-correlation energy [Eq. (28)]. It was found that even-tempered expansions of uncontracted s , p , and d functions provide a flexible enough representation. Following a suggestion of Dunlap,²⁹ the first term in the even-tempered expansion corresponds to the most diffuse orbital function. For carbonlike atoms, seven s -type, three p -type, and three d -type functions are found to be adequate. The same number of functions is also used to represent the exchange-correlation terms. Since the most diffuse auxiliary s , p , and d functions are used to fit the same region in the valence space, it is possible to construct auxiliary basis sets with shared exponents.⁵⁸ This facilitates the calculation of molecular integrals. The auxiliary basis sets corresponding to the DZVP and DZVP2 orbital basis sets are called $A1$ and $A2$, respectively.

B. Analytical integrals

In contrast to the usual implementations of the Hartree–Fock method, the present DFT method requires the evaluation of only two- and three-index integrals. A recursive scheme, developed originally by Obara and Saika (OS)⁴ for the computation of four-index integrals over Cartesian Gaussian functions, has been reformulated for three-index integrals to meet the needs of DFT.^{42,43} The resulting scheme turns out to be computationally highly efficient on vector and parallel computers.

Two kinds of three-index integrals are needed—Coulomb integrals I_C and overlaplike integrals I_{xc} for the evaluation of exchange-correlation terms. Following the notation of OS, these two types of integrals have the form

$$I_C = [a(1)b(1)||c(2)] \quad (41)$$

and

$$I_{xc} = [a(1)b(1)c(1)] \quad (42)$$

with a and b denoting orbital basis functions and c representing an auxiliary basis function used in the expansion of the electron density, the exchange-correlation potential, and the exchange-correlation energy. Again, the symbol $||$ stands for the Coulomb operator $1/|r - r'|$.

Equation (39) of OS can be rewritten for three-index Coulomb integrals to give the following recursive expression:

$$\begin{aligned} [ab||c + 1_i]^{(m)} &= (W_i - Q_i)[ab||c]^{(m+1)} + 1/2\eta N_i(c)\{[ab||c - 1_i]^{(m)} - (\rho/\eta)[ab||c - 1_i]^{(m+1)}\} \\ &+ 1/2(\zeta + \eta)\{N_i(a)[a - 1_i]b||c]^{(m)} + N_i(b)[a(b - 1_i)||c]^{(m+1)}\}. \end{aligned} \quad (43)$$

Here, l_i denotes a p_x , p_y , or p_z function and the function $c + 1_i$ has an angular momentum number one order higher than c . The superscript (m) refers to the order of the incomplete Γ function which is needed in the evaluation of electron repulsion integrals. The prefactors W and Q are determined by the distances between the centers of the Gaussian functions and their exponents. The values for η , ζ , and ρ depend only on the Gaussian exponents. N_i is a generalized Kronecker delta as defined by OS. Following a procedure implemented in the Asterix program,⁵⁹ the incomplete Γ function is evaluated from an efficient four-term interpolation formula yielding an accuracy of 10^{-12} in the integrals.

In essence, Eq. (43), together with "transfer formulas" of the type

$$\begin{aligned} [(a + 1_i)b \|c]^{(m)} &= [a(b + 1_i) \|c]^{(m)} \\ &- (A_i - B_i)[ab \|c]^{(m)} \end{aligned} \quad (44)$$

and explicit formulas for integrals with low angular momenta, allow the evaluation of all Coulomb integrals involving higher angular momentum quantum numbers. For example, integrals containing d functions can be built from integrals with only s and p functions. Similar relationships can be obtained for the overlaplike integrals needed in the exchange-correlation terms. For the integrals in the gradient calculations, a modified OS method due to Head-Gordon and Pople⁵ has been used.

The evaluation of the Coulomb and exchange-correlation integrals as needed in the construction of the Hamiltonian matrix elements (14) lends itself to parallelism in the index r of Eq. (14). To this end, shared exponents of s , p , and d fitting functions are highly advantageous. A similar strategy for parallelism is used in the evaluation of the density fitting coefficients given in Eqs. (24) and (25).

The computing times for calculating the Coulomb integrals for the Hamiltonian matrix elements (14) and for calculating the integrals for the density fits (24) and (25) is approximately the same, whereas the evaluation of the overlaplike integrals for the exchange-correlation fit (18) requires about half that time. In the current vectorized and parallel implementation of a direct SCF scheme, approximately 15 million three-index integrals per second can be generated on an eight-processor CRAY Y-MP system at a sustained computational rate of over 1 billion floating point operations per second (GFLOPS).⁶⁰ Compared with a one-processor system, a speedup of approximately 7.3 has been observed. Thus, the evaluation of about 350 million integrals for a molecular system with about 1000 basis functions can be accomplished in about 160 s of real time by using all eight processors of a CRAY Y-MP system simultaneously.

C. Grid design and numerical integrations

Numerical integrations are used to determine the exchange-correlation terms in Eqs. (12), (26), (28), (31), and (34). Furthermore, in the present program, the multipole moments are obtained by using a numerical integration. These integrations use a finite adaptive grid which is con-

structed in the following way^{21,61}: around each atom, a polar coordinate system is defined. Within these local coordinate systems, a radial mesh is constructed leading to a sequence of radial shells. On each shell, an angular grid is defined such that its resolution is adapted to the variation of the exchange-correlation potential. For example, close to the nuclei, the radial shells are densely spaced, but each shell contains relatively few angular points. Further away from the nuclei, the situation is reversed and a higher number of angular points is required, while the radial shells might be more widely spaced.

Using this kind of grid, any three-dimensional molecular multicenter integral I with the integrand $F(\mathbf{r})$ can be decomposed into a sum of one-center integrals of the following form:

$$I = \int F(\mathbf{r}) d\mathbf{r} = \sum_A I_A = \sum_A \int \omega_A(\mathbf{r}) F(\mathbf{r}) d\mathbf{r}. \quad (45)$$

The weighting function ω_A is close to 1 near the center A and diminishes in the vicinity of other centers. At every point of real space, the weighting functions have to be normalized to 1,

$$\sum_A \omega_A(\mathbf{r}) = 1. \quad (46)$$

The one-center integrals are then solved within a polar coordinate system using the finite grid described above

$$I_A = \int \int \omega_A(r, \Omega) F(r, \Omega) r^2 dr d\Omega. \quad (47)$$

Here, the radial integration is accomplished using a method by Becke⁶¹ applying a Gauss-Chebyshev quadrature, while a Stroud-Lebedev quadrature is employed for the angular integration.^{62,63}

The angular grid is constructed in the following adaptive way: at the beginning of a calculation, the atomic electron densities are superposed resulting in an electron density which has in essence the shape of the molecular density. Then, polar coordinate systems are defined around each atom center and radial grids are defined to meet the criteria for the Gauss-Chebyshev quadrature. On each radial shell, a set of angular points is created. These sets correspond to the Stroud-Lebedev quadrature and contain the following number of points: 12, 32, 50, 72, 110, 194, and 302. Once a set is selected, the values of the exchange-correlation energy $\epsilon_{xc}(\mathbf{r})$ corresponding to the superposed electron densities are calculated and an assessment is made how accurately the current angular grid can integrate the function $\epsilon_{xc}(\mathbf{r})$. Angular grids of increasing resolution are chosen until the integration is guaranteed to have an error of less than 10^{-8} . In order to minimize any sensitivity of the numerical integration on the specific orientation of the molecule, angular grids on subsequent radial shells are rotated with respect to each other.⁴⁷ Typically, about 1000 grid points per atom are obtained from this procedure.

The weighting function of Eqs. (45)–(47) is chosen to be

$$\omega_A(\mathbf{r}) = \rho_A^2(\mathbf{r}) / \sum_A \rho_A^2(\mathbf{r}) \quad (48)$$

with $\rho_A(\mathbf{r})$ being the atomic density of center A .

D. Direct SCF procedure

The SCF procedure involves the following sequence:

$$\rho^{(0)} \rightarrow \mu_{xc}^{(0)} \rightarrow \mathbf{H}^{(0)} \rightarrow \{\psi_i^{(0)}\} \rightarrow \rho^{(1)} \rightarrow \mu_{xc}^{(1)} \rightarrow \mathbf{H}^{(1)} \rightarrow \{\psi_i^{(1)}\} \dots, \quad (49)$$

or in terms of the expansion coefficients given by Eqs. (10)–(12),

$$\{\rho_r^{(0)}\} \rightarrow \{\mu_s^{(0)}\} \rightarrow \{c_{ip}^{(0)}\} \rightarrow \{\rho_r^{(1)}\} \rightarrow \{\mu_s^{(1)}\} \rightarrow \{c_{ip}^{(1)}\} \rightarrow \dots \quad (50)$$

The procedure (see Ref. 47 for a preliminary description) is started with a superposition of atomic densities which defines the coefficients $\{\rho_r^{(0)}\}$. In the case of a geometry optimization, the SCF results from the previous geometry are taken as a SCF starting point for the new geometry.

The Hamiltonian matrix elements of two subsequent iterations i and $i+1$ can be written in the form

$$H_{pq}^{(i+1)} = H_{pq}^{(i)} + \Delta H_{pq}^{(i+1)}. \quad (51)$$

Using Eq. (14),

$$\begin{aligned} \Delta H_{pq}^{(i+1)} = & \sum_r (\rho_r^{(i+1)} - \rho_r^{(i)}) [pq||r] \\ & + \sum_s (\mu_s^{(i+1)} - \mu_s^{(i)}) [pqs]. \end{aligned} \quad (52)$$

An update of a Hamiltonian matrix element H_{pq} is not necessary if ΔH_{pq} is smaller than a certain numerical threshold. In the present calculations, a threshold of 10^{-10} is used, although it was found that a value of 10^{-8} still leads to acceptable molecular properties of chemical accuracy.

In order to determine the threshold, only an upper bound or estimate and not the exact values for the integrals are required. This fact leads to significant savings in the computational effort. The following conditions for the thresholds are obtained from Eq. (52):

$$(\rho_r^{(i+1)} - \rho_r^{(i)}) \text{est}([pq||r]) < T_\rho, \quad (53)$$

$$(\mu_s^{(i+1)} - \mu_s^{(i)}) \text{est}([pqs]) < T_\mu, \quad (54)$$

where the symbol $\text{est}[\]$ denotes an estimate of an integral. Hence, a Hamiltonian matrix element of a new iteration is updated only if it can be expected to be larger than T_ρ and T_μ defined by Eqs. (53) and (54).

The estimation of the integral $(pq||r)$ is based on the Schwarz inequality³

$$[pq||r] \leq [pq||pq] [r||r]. \quad (55)$$

By definition,²⁹ $[r||r] = 1$ and Eq. (55) becomes

$$[pq||r] \leq [pq||pq] \equiv \text{est}([pq||r]), \quad (56)$$

where $[pq||pq]$ represents a standard four-index, two-electron integral. Since there are only N^2 integrals of this type, their calculation is inexpensive while leading to an accurate upper bound for the three-index integrals $[pq||r]$. In the present implementation, integrals of the type $[dp||dp]$ and

$[dd||dd]$ are not calculated explicitly, but estimated by the simple formula^{3,4}

$$\text{est}([ab||ab]) \equiv N_a N_b \exp[-ab/(a+b)] |\mathbf{A} - \mathbf{B}|^{-2}, \quad (57)$$

where N_a and N_b are normalization factors, a and b exponents, and \mathbf{A} and \mathbf{B} the positions of Gaussian functions.

An estimate of the values for the three-index overlaplike integrals $[pqs]$ in Eq. (54) is obtained from the prefactors occurring in the calculation of these integrals.⁴

A similar strategy as used for the update of the Hamiltonian matrix elements can also be employed for the fitting coefficients ρ_r as given by Eq. (24). An update of a density fitting coefficient is not necessary if

$$(P_{pq}^{(i+1)} - P_{pq}^{(i)}) \text{est}([pq||r]) < T_d. \quad (58)$$

Expression (58) contains the same types of three-index Coulomb integrals as in Eq. (54) and hence the same procedure for estimating their values is used. In the present calculation, the value for T_d is set to 10^{-10} . This concept for updating as used in the direct SCF can also be applied to the evaluation of the total energy.

E. Geometry optimizations and vibrational analysis

The geometry optimizations and the calculation of vibrational frequencies follow well-established techniques of quantum chemistry.⁶⁴ Specifically, the present approach uses a procedure for geometry optimizations as implemented in the GRADSCF program.⁶⁵ Three differences compared with Hartree–Fock calculations are important to mention: (i) Because of the nonvariational character of the present formulation of energy gradients, the energy minimum does not coincide strictly with vanishing gradients. The resulting inconsistencies are tolerable. For example, for equilibrium bond lengths, this discrepancy is only several thousands of one Ångström, which is one order of magnitude smaller than the accuracy of the method compared with experiment. (ii) Gradient calculations are faster than SCF calculations and thus gradients are evaluated for each new geometry in the optimization procedure. (iii) Due to the numerical integrations, there is residual numerical noise in the total energy. Close to equilibrium, this noise is relatively smaller for the gradients than for the total energy. Therefore, the final steps in the geometry optimization are based on the gradient alone.

The vibrational frequencies are calculated by a finite difference technique using a displacement of 0.02 a.u. The infrared intensities are evaluated also by finite differences in the dipole moments. In the present program, dipole moments are calculated by straight numerical integrations using the same grid as for the gradients, which is finer than that used for the SCF energies.

IV. COMPUTATIONAL LEVELS AND NUMERICAL SENSITIVITY

In the present implementation, the following major computational parameters need to be defined: the orbital basis set; the auxiliary basis sets; the grid resolution for the numerical integrations; the accuracy in the analytical inte-

grals; the convergence in the SCF procedure; and the convergence in the gradient optimization.

In order to enable systematic comparisons, we define several computational levels, as given in Table I. The DZVP level is characterized by a double-zeta orbital basis set with polarization functions except for hydrogen atoms. On this level, a set of auxiliary functions, called *A* 1, is used as defined in Table I. The grid resolution, the integral accuracy, the SCF convergence, and the gradient convergence are chosen to be on a medium level as explained in Table I. The DZVPP level differs from the DZVP level only in the orbital basis sets for hydrogen, where *p*-polarization functions are included on the DZVPP level. Most of the calculations reported in this paper were carried out on this level.

A somewhat cruder level is denoted DZVP-1c. Here, the same orbital and auxiliary basis sets are used as on the DZVP level, but a coarser grid is used for the numerical integration resulting in about half the number of grid points as on the DZVP level. The integral accuracy is lower and the convergence criteria for the SCF procedure and the gradient optimizations are less stringent than those of the DZVP level. An even more economic level is obtained by reducing the number of auxiliary functions used in the fitting of the electron density and the exchange-correlation terms. This level is denoted DZVP-3c since a so-called *A* 3 auxiliary basis set is used (cf. Table I.)

DZVP-1f denotes a level which has the same orbital and auxiliary sets as the DZVP level, but applies more stringent criteria for the grid resolution, the integral accuracy, and the SCF and gradient convergence criteria. Note that about four times more grid points are used for the numerical integration as on the DZVP-1c level. The DZVP2-2f level can be considered to be close to the limits of a double-zeta orbital basis set with one polarization function. Here, a larger auxiliary basis

set, denoted *A* 2, is used than on the DZVP level. The grid resolution, integral accuracy, and the convergence criteria are the same as on the DZVP-1f level.

Table II shows the sensitivity of calculated geometric, vibrational, and energetic properties on the computational levels for methylamine. The calculated bond lengths differ by 0.004 Å or less comparing the various levels listed in Table I. The largest variation in the bond angles is 1.4° for the hypernetted chain (HNC) angle. Overall, there is relatively little sensitivity of geometric variables on the choice of computational levels. This is gratifying since, e.g., the numerical grid resolution varies by a factor of four between DZVP-1c and DVPP-1f. It should be pointed out, though, that the gradient convergence criterion may have to be tightened from the default value of 8×10^{-4} to about 5×10^{-5} a.u. in order to obtain accurate dihedral angles and vibrational frequencies for low frequency or strongly anharmonic torsional modes.

The frequencies reported in Table II were obtained by a finite difference technique using nuclear displacements of 0.02 bohr and analytic first derivatives. For all but the lowest frequency modes, the different computational levels give frequencies that are consistent to within better than 4% which is approximately the same size as the difference between theory and experiment (note that the experimental values are not the harmonic frequencies and a direct comparison can be misleading). For the lowest frequency, the DZVP-3c and DZVP-1c levels deviate substantially from the other results. The vibrational zero point energy provides a convenient measure for the quality of the vibrational frequencies. While the results from DZVP and higher are very consistent, some deviations are found for the calculations using fairly rough numerical grids and convergence criteria. However, even for these economic choices of computational parameters, the

TABLE I. The definition of computational levels including orbital basis sets, auxiliary basis sets, grid resolutions, integral accuracies, SCF convergence thresholds, and gradient convergence criteria. The most common levels are DZVP and DZVPP with *A* 1 auxiliary basis sets and medium selections for the numerical grid, integral accuracy, and convergence criteria.

Level	Basis ^a	Aux. ^b	Grid ^c	Int. ^d	CVSCF ^e	CVGRAD ^f
DZVP-3c	DZVP	<i>A</i> 3	c	l	l	l
DZVP-1c	DZVP	<i>A</i> 1	c	l	l	l
DZVP	DZVP	<i>A</i> 1	m	m	m	m
DZVPP	DZVPP	<i>A</i> 1	m	m	m	m
DZVP-1f	DZVP	<i>A</i> 1	f	h	t	t
DZVP2-2f	DZVP2	<i>A</i> 2	f	h	t	t

^a Following the notation of Huzinaga, the basis sets for carbonlike atoms are: DZVP—(621/41/1)/[3/2/1]; DZVP2—(721/51/1)/[3/2/1]. The basis set denoted DZVPP is the same as DZVP except that *p*-polarization functions are added for H atoms. The corresponding Pople basis sets are 6-31G* (DZVP) and 6-31G** (DZVPP).

^b The auxiliary basis sets are uncontracted and defined as follows: *A* 1 (7/3/3); *A* 2 (8/4/4); *A* 3 (7/3/2).

^c The grid selection can be coarse (c) with about 500 points/atom, medium (m) with about 1100 points per atom, or fine (f) with about 2500 points/atom.

^d The accuracy in the analytical integral evaluation is low (l) at 10^{-8} , medium (m) at 10^{-10} , or high (h) at 10^{-12} ; the corresponding accuracy in the numerical integration is 10^{-10} , 10^{-12} , or 10^{-14} , respectively.

^e The SCF convergence threshold for the density (total energy) is loose (l) at 10^{-4} (10^{-6}), medium (m) at 5×10^{-5} (5×10^{-7}), or tight (t) at 10^{-5} (10^{-7}).

^f The convergence criterion for the largest gradient component is loose (l) at 10^{-3} , medium (m) at 8×10^{-4} , and tight (t) at 5×10^{-4} .

TABLE II. Ground-state geometries, vibrational frequencies, dipole moments, and C–N bond dissociation energies of methylamine obtained with various basis sets and selections of computational parameters. An explanation of the computational levels is given in Table I.

	DZVP-3c	DZVP-1c	DZVP	DZVPP	DZVP-1f	DZVP2-2f	Expt. ^a
Bond lengths (Å)							
CN	1.456	1.453	1.452	1.452	1.453	1.454	1.471
CH ₁	1.117	1.114	1.116	1.115	1.116	1.114	1.099
CH ₂	1.106	1.105	1.106	1.105	1.107	1.105	1.099
NH	1.026	1.026	1.027	1.025	1.027	1.025	1.010
Bond angles (degrees)							
HNC	109.1	109.3	109.4	109.4	109.4	109.5	110.3
HNH	107.0	107.1	107.1	107.2	107.0	107.4	107.1
Angle between methyl top axis and CN bond (degrees)							
θ	7.3	7.1	7.1	6.9	7.0	6.4	3.0
Frequencies (cm ⁻¹)							
NH ₂ α stretch	3524	3543	3516	3533	3514	3551	3427
NH ₂ s stretch	3432	3428	3418	3445	3417	3461	3361
CH ₃ d stretch	3058	3067	3048	3049	3043	3058	2985
	3009	3009	2998	3004	2995	3013	2961
CH ₃ s stretch	2878	2895	2880	2893	2878	2906	2820
NH ₂ scissor	1624	1602	1621	1588	1620	1591	1623
CH ₃ d deform	1458	1493	1453	1450	1450	1438	1483
	1442	1439	1430	1423	1430	1420	1473
CH ₃ s deform	1377	1382	1396	1385	1395	1382	1430
NH ₂ twist	1311	1296	1300	1281	1297	1274	1419
	1132	1133	1131	1127	1130	1119	1195
CN stretch	1074	1077	1085	1084	1084	1085	1130
CH ₃ rock	963	877	948	939	941	934	1044
NH ₂ wag	789	809	810	785	811	777	780
Torsion	342	331	306	309	308	302	268
ZPE	39.19	39.14	39.08	39.00	39.04	39.04	
Dipole moment (D)							
	1.52	1.49	1.50	1.41	1.51	1.46	
Total energies (hartree)							
CH ₃ NH ₂							
LSD	-95.023 12	-95.025 50	-95.025 50	-95.028 98	-95.025 59	-95.036 10	
NLSD	-95.879 50	-95.877 67	-95.876 75	-95.881 55	-95.876 56	-95.890 37	
CH ₃ (doublet)							
LSD	-39.436 07	-39.435 73	-39.435 71	-39.437 47	-39.435 70	-39.439 02	
NLSD	-39.845 96	-39.843 16	-39.843 03	-39.845 00	-39.842 89	-39.848 11	
NH ₂ (doublet)							
LSD	-55.408 80	-55.405 71	-55.405 70	-55.409 36	-55.405 68	-55.414 88	
NLSD	-55.891 63	-55.888 71	-55.888 13	-55.892 31	-55.888 07	-55.898 41	
C–N bond dissociation energy (kcal/mol)							
LSD	111.9	115.5	115.5	114.3	115.6	114.3	
NLSD	89.1	91.5	91.4	90.5	91.4	90.3	
Expt.	93						

^a From Ref. 1; geometric data from M. D. Harmony, V. W. Laurie, R. L. Kuczkowski, R. H. Schwendeman, D. A. Ramsay, F. J. Lovas, W. J. Lafferty, and A. G. Maki, *J. Phys. Chem. Ref. Data* 8, 619 (1979).

results are still reasonable.

The calculated dipole moment ranges between 1.41 and 1.52 D. A significant change arises from the inclusion of *p*-polarization functions on the hydrogen atoms. In fact, the sensitivity on the orbital basis set is larger than on the resolution of the grid or the choice for auxiliary basis sets.

The C–N bond dissociation energy provides a gauge for the sensitivity of calculated reaction energies as a function of the computational level. This energy is calculated by separately optimizing the fragments $\cdot\text{CH}_3$ and $\cdot\text{NH}_2$ as well as the complete molecule and then subtracting the total energy of the complete molecule from the sum of the total energies of the fragments. In these calculations, the methylamine molecule is in a singlet state, whereas the fragments are in a doublet state. The step from the compact *A* 3 auxiliary basis set, used in the DZVP-3c calculation, to the more flexible *A* 1 set, employed in the DZVP-1c results, causes a significant change in the calculated bond dissociation energy. The other modifications of basis sets and computational parameters lead to changes of less than 1.2 kcal/mol.

In summary, both the DZVP and DZVPP levels of computational parameters provide accurate and computationally efficient choices of orbital basis sets, auxiliary basis sets, grid resolutions, integral accuracies, and SCF and gradient convergence criteria. While bond distances, bond angles, dipole moments, and total energy differences are fairly insensitive within the range of computational levels defined in Table I, low frequency modes deserve special attention if a high accuracy is needed.

V. RESULTS FOR MOLECULES CONTAINING C, N, O, H, AND F

A. Ground state geometries

Table III shows calculated and experimental ground state geometries of small molecules and fragments containing the elements C, N, O, H, and F. All calculations were done within local spin density functional theory using a DZVPP computational level as described in the previous section. For the cases in which the spin state is other than a singlet, the multiplicity is given explicitly in Table III.

The calculated C–C single bond lengths are typically 0.01–0.02 Å too short compared with experiment. The formal C–C single bond in cyclopropane is an exception. In this case, theory and experiment agree to within 0.001 Å. The C = C double bond and the aromatic C–C bond are very well described by this level of theory and agree typically within a few thousandths of an Å with experiment. For example, the computed C = C bond lengths in propene and butadiene agree to within 0.003 and 0.002 Å with experiment. LSD theory on the DZVPP level overestimates the C≡C triple bond in acetylene by 0.014 Å. Also, the theoretical bond length in the carbon dimer is too long by about 0.02 Å.

In many cases, the calculated C–H bond length is too long^{10,11} by about 0.01 to 0.02 Å, which is also found, possibly even more pronounced, for O–H, N–H, and H–F distances. The overestimation of the C–H bond length is particularly noticeable in the small fragments CH, CH₂, and CH₃. These systems contain only a few electrons and the

electron gas approximation underlying the LSD theory can be expected to be less appropriate than for bonds with more electrons (such as double bonds and aromatic systems).

For single and double bonds between carbon and oxygen atoms, similar trends are found as between carbon atoms. The computed lengths for the C–O single bonds in methanol and dimethylether are too short by about 0.01 Å. In contrast, the C = O bond in formaldehyde is close to experiment (0.004 Å too long) and about 0.01 Å too long in formic acid and carbon dioxide. The C = O bond length in ketene is overestimated by 0.016 Å and by the surprising amount of 0.030 Å in formamide. The calculated C≡O bond in the CO molecule is 0.016 Å too long.

The bonds between C and N seem to repeat this pattern. The single C–N bond in methylamine is too short by about 0.02 Å, by 0.012 Å in trimethylamine, and by 0.018 Å in formamide. The C≡N bond in hydrogen cyanide is too long by 0.012 Å and by 0.022 Å in methyl isocyanide. The polar C–F bonds in CH₃F, CHF₃, and CF₄ fluoromethane are within 0.007 Å of experiment. The bonds between N and O in NO, NO₂, and HNO₃ are within 0.016 Å of experiment. On the other hand, the N–O distance in nitrosyl fluoride is too long by 0.025 Å and the N–F distance in NF₃ is too long by the unusual amount of 0.064 Å.

The equilibrium distances in the molecules O₂, N₂, and F₂ are fairly well described by the LSD/DZVPP approach and deviate from experiment by 0.007, 0.021, and 0.020 Å, respectively. The ambiguity in the experimental values for the O–O distance in hydrogen peroxide make an assessment of the theoretical value difficult.

Calculated bond angles deviate in many cases by less than 1° from the experimental values. A notable exception is the O–N–O angle in nitrogen dioxide, where the present theoretical approach yields a value which is 2.5° smaller than experiment (133.5° vs 136.0°). Possibly, LDF theory does not include enough repulsion between the O atoms in NO₂.

B. Vibrational frequencies

Table IV provides a comparison of computed and measured vibrational frequencies for typical small molecules. As is well known,¹ Hartree–Fock theory on the SCF level fairly systematically overestimates vibrational frequencies by about 10%. The MP2 level of theory provides vibrational frequencies which are closer to experiment, but still mostly too high. On the other hand, LSD theory on the DZVPP level yields frequencies which are overall too low, but at least as close to experiment as those obtained from second-order Møller–Plesset (MP2) theory.

The present calculations give C≡C and C = C bond stretching frequencies which are remarkably close to the harmonic frequencies deduced from experiment. The frequency of the single C–C bond stretching mode is about 1% too high, which might be related to the fact that the present calculations result in a C–C bond length which is slightly too short compared with experiment (cf. Table III). The LSD C–H stretching frequencies are typically too low by about 1%–2%, which might be related to the overestimation of the C–H equilibrium bond length. In many cases, but not all, the low frequency modes involving bending and torsion are cal-

TABLE III. A comparison of calculated and experimental ground state geometries of small molecules containing C, N, O, H, and F. The LSD calculations were carried out on the DZVPP level of accuracy as defined in Table I.

Molecule		Variable	LSD	Expt. ^a	Expt. ^b
CH	<i>Doublet</i>	CH	1.148	1.120	
CH ₂	<i>Methylene, triplet</i>	CH	1.094	1.029	1.078
		HCH	137.0	144.7	136.0
CH ₃	<i>Methyl, doublet</i>	CH	1.093	1.079	
		HCH	120.0	120.0	
CH ₄	<i>Methane</i>	CH	1.101	1.094	
C ₂	<i>Carbon dimer</i>	CC	1.263	1.242	
C ₂ H ₂	<i>Acetylene</i>	CC	1.217	1.203	
		CH	1.081	1.060	
C ₂ H ₄	<i>Ethylene</i>	CC	1.336	1.339	
		CH	1.098	1.086	
		HCC	121.4	121.2	
C ₂ H ₆	<i>Ethane</i>	CC	1.519	1.536	1.531
		CH	1.104	1.091	1.096
		HCC	111.6	110.9	
C ₃ H ₄	<i>Cyclopropene</i>	C1C2	1.305	1.296	
		C2C3	1.510	1.509	
		C1H	1.091	1.072	
		HC1C2	149.5	149.9	
C ₃ H ₆	<i>Cyclopropane</i>	CC	1.504	1.510	
		CH	1.095	1.089	
C ₃ H ₆	<i>Propene</i>	C = C	1.339	1.336	
		C-C	1.487	1.501	
		C1H	1.098	1.091	
		C2H	1.101	1.090	
		C3H	1.106	1.085	
		CCC	125.8	124.3	
		C2C3H	110.7	111.2	
		C1C2H	118.3	119.0	
		C2C1H	121.0	121.5	
		CC	1.517	1.526	
		C1H	1.103	1.096	
C ₄ H ₆	<i>1,3-butadiene</i>	C2H	1.106	1.115	
		CCC	112.9	112.4	
		C1C2H	109.5	109.5	
		C2C1H	111.9	111.8	
		C1C2	1.346	1.344	
C ₄ H ₈	<i>Cyclobutane</i>	C2C3	1.446	1.467	
		CCC	122.5	122.9	
		CC	1.541	1.548	
C ₄ H ₁₀	<i>n-butane</i>	CH	1.104	1.105	
		C1C2C3C4	152.6	153.0	
		C1C2	1.517	1.533	
		C2C3	1.532	1.533	
CH(CH ₃) ₃	<i>Isobutane</i>	CCC	113.1	112.8	
		CC	1.518	1.525	
		CC	1.523	1.539	
C(CH ₃) ₄	<i>Neopentane</i>	CH	1.107	1.120	
		HCC	110.2	110.0	
		CC	1.396	1.399	
C ₆ H ₆	<i>Benzene</i>	CH	1.098	1.084	
		CO	1.144	1.128	
CO	<i>Carbon monoxide</i>	CO	1.212	1.208	
CH ₂ O	<i>Formaldehyde</i>	CH	1.124	1.116	
		HCO	121.9	121.8	
		CO	1.411	1.425	1.421
CH ₃ OH	<i>Methanol</i>	CH	1.104	1.094	1.094
		OH	0.974	0.945	0.963
		HCO	106.8	108.5	107.2
		COH	108.8	108.0	108.0
		CO	1.177	1.161	
CH ₂ CO	<i>Ketene</i>	CC	1.316	1.314	
		CH	1.094	1.083	
		CCH	119.6	118.7	
		CC	1.316	1.314	

TABLE III. (Continued.)

Molecule	Variable	LSD	Expt. ^a	Expt. ^b			
CH ₃ OCH ₃	Dimethyl ether	CO	1.401	1.410			
		COC	110.7	111.3			
CO ₂	Carbon dioxide	CO	1.175	1.162			
HCOOH	Formic acid	C=O	1.212	1.202			
		C-O	1.342	1.343			
		CH	1.115	1.097			
		OH	0.988	0.972			
		OCO	124.7	124.9			
		HOC	106.5	106.3			
		HC-O	125.6	124.1			
CN	Cyanide	CN	1.183	1.175			
HCN	Hydrogen cyanide	CN	1.166	1.154			
		CH	1.085	1.063			
CH ₃ CN	Methyl isocyanide	CN	1.446	1.424			
		CH	1.103	1.101			
		CN	1.170	1.166			
		HCN	110.2	109.1			
(CH ₃) ₃ N	Trimethylamine	CN	1.439	1.451			
		CNC	111.8	110.9			
		CN	1.358	1.376			
HCONH ₂	Formamide	NH	1.022	1.002			
		CH	1.122	1.102			
		CO	1.223	1.193			
		OCN	124.5	123.8			
		C ₄ H ₄ N ₂	Pyrimidine	C4C5	1.393	1.393	
		N3C4		1.336	1.350		
C2N3	1.336	1.328					
C2H	1.101	1.082					
C5H	1.102	1.087					
C6H	1.098	1.079					
CH ₃ F	Fluoromethane	N3C4C5	122.2	121.2			
		CH	1.106	1.098			
		CF	1.378	1.382			
		FCH	109.2	108.5			
CHF ₃	Trifluoromethane	CH	1.108	1.098			
		CF	1.340	1.333			
		FCH	110.5	110.3			
CF ₄	Carbon tetrafluoride	CF	1.326	1.321			
H ₂ O	Water	OH	0.974	0.957			
		HOH	105.2	104.5			
		OO	1.441	1.475	1.452		
H ₂ O ₂	Hydrogen peroxide	OH	0.981	0.950	0.965		
		HOO	100.4	94.8	100.0		
		HOOH	117.8	119.8	119.1		
		OO	1.223	1.216			
O ₂	Oxygen, triplet	NH	1.025	1.012			
NH ₃	Ammonia	HNH	107.3	106.7			
		NO	1.167	1.151			
NO	Nitrogen oxide	NO	1.208	1.197			
NO ₂	Nitrogen dioxide	ONO	133.5	136.0			
		N=O	1.222	1.206			
HNO ₃	Nitric acid	N-O	1.402	1.405			
		OH	0.988	0.960			
		O=N=O	130.4	130.0			
		NOH	101.5	102.0			
		NN	1.114	1.094			
		NF	1.499	1.520			
N ₂	Nitrogen	NO	1.155	1.130			
		FNO	110.4	110.2			
HF	Hydrogen fluoride	FF	1.392	1.412			
NOF	Nitrosyl fluoride	NF	1.381	1.317			
		FNF	101.7	102.2			
F ₂	Fluorine						
NF ₃	Nitrogen trifluorine						

^aFrom J. J. P. Stewart, *J. Computer-Aided Mol. Design* **4**, 1 (1990), Table 10.

^bFrom Ref. 1, Table 6.6.

TABLE IV. A comparison of calculated and experimental vibrational frequencies in cm^{-1} for small molecules containing C, N, O, H, and F. The LSD calculations were carried out on the DZVPP level as defined in Table I. The last column gives the harmonic frequencies derived from experiment. The following abbreviations are used to characterize the vibrational modes: antisymmetric stretch (as), bend (b), *d* deform (*dd*), *d* stretch (*ds*), rock (r), stretch (s), scissor (sc), symmetric stretch (ss), twist (t), and wag (w). ZPE denotes the vibrational contribution to the zero point energy in kcal/mol.

Molecule	Mode	HF ^a	MP2 ^a	LSD	Expt. ^a	
						Harm.
C ₂ H ₂	CH _s	3719	3593	3441	3374	3497
		3607	3516	3343	3289	3415
	CC _s	2247	2006	2011	1974	2011
	CH _b	883	783	705	730	747
		794	444	560	612	624
ZPE	18.5	16.5	16.2	16.2	16.7	
C ₂ H ₄	CH _{ss}	3420	3323	3181	3106	3234
		3394	3297	3156	3103	3233
	CH _{ss}	3344	3231	3081	3026	3153
		3321	3222	3067	2989	3147
	CC _s	1856	1724	1654	1623	1655
	CH _{2sc}	1610	1523	1388	1444	1473
		1499	1425	1320	1342	1370
	CH _{2r}	1353	1265	1177	1236	1245
		897	873	799	826	843
	CH _{2t}	1155	1083	1019	1023	1044
	CH _{2w}	1095	980	910	949	969
	ZPE	1099	931	909	943	959
	C ₂ H ₆	ZPE	34.4	32.7	31.0	30.9
CH _{ds}		3271	3215	3048	2985	3140
		3242	3228	3070	2969	3175
CH _{ss}		3194	3104	2978	2986	3061
		3201	3086	2975	2954	3043
CH _{3dd}		1646	1520	1432	1469	1526
		1652	1604	1428	1468	1552
CH _{3sd}		1584	1493	1358	1388	1449
		1546	1494	1332	1375	1438
CH _{3r}		1338	1264	1155	1190	1246
		894	783	790	822	822
CC _s		1063	1040	1024	995	1016
Torsion		331	452	316	289	303
ZPE	50.0	48.5	45.5	45.5	47.5	
HCN	CH _s	3679	3517	3376	3311	3442
	CN _s	2438	2038	2139	2097	2129
	Bend	889	702	692	712	727
	ZPE	11.3	9.9	9.9	9.8	10.0
HNC	NH _s	4092	3844	3718	3620	3842
	NC _s	2311	2038	2041	2029	2067
	Bend	519	389	494	477	490
	ZPE	10.6	9.5	9.6	9.4	9.8
CH ₂ NH	NH _s	3719	3463	3355	3297	
	CH _{2ss}	3347	3254	3040	3036	
	CH _{2as}	3254	3116	2937	2924	
	CN _s	1901	1724	1690	1640	
	Bend	1628	1542	1426	1453	
		1496	1412	1289	1347	
		1164	1100	1116	1059	
	Torsion	1270	1159	1040	1123	
	Bend	1223	1107	1025	1063	
	ZPE	27.2	25.6	24.2	24.2	
CH ₃ NH ₂	NH _{2as}	3813	3641	3533	3427	
	NH _{2ss}	3730	3508	3445	3361	
	CH _{3ds}	3281	3228	3049	2985	
		3245	3155	3004	2961	
	CH _{3ss}	3156	3063	2893	2820	
	NH _{2sc}	1841	1745	1588	1623	
	CH _{3dd}	1665	1596	1450	1483	
		1648	1539	1423	1473	
	CH _{3ds}	1607	1469	1385	1430	

TABLE IV. (Continued.)

Molecule	Mode	HF ^a	MP2 ^a	LSD	Expt. ^a		
						Harm.	
H ₂ CO	NH _{2t}	1479	1405	1281	1419		
	CH _{3r}	1052	915	939	1195		
		1289	1237	1127	1130		
	CN _s	1149	1113	1084	1044		
	NH _{2w}	946	941	785	780		
	Torsion	341	351	309	268		
	ZPE	43.2	41.3	39.0	39.2		
	CH _{2as}	3231	3064	2877	2843	3009	
	CH _{2ss}	3159	3019	2815	2783	2944	
	CO _s	2028	1786	1794	1746	1764	
	CH _{2sc}	1680	1567	1449	1500	1563	
	CH _{2r}	1384	1249	1196	1249	1287	
CH ₃ OH	CH _{2w}	1336	1194	1120	1167	1191	
	ZPE	18.3	17.0	16.1	16.1	16.8	
	OH _s	4117	3785	3759	3681		
	CH _{3ds}	3305	3201	3048	3000		
		3231	3140	2964	2960		
	CH _{3ss}	3185	3065	2901	2844		
	CH _{3dd}	1664	1552	1440	1477		
		1652	1562	1425	1477		
	CH _{3ds}	1638	1542	1413	1455		
	OH _b	1508	1424	1308	1345		
	CH _{3r}	1289	1160	1121	1165		
		1187	1120	1026	1060		
CH ₃ F	CO _s	1164	1082	1100	1033		
	Torsion	348	250	317	295		
	ZPE	34.7	32.7	31.2	31.1		
	CH _{3ds}	3312	3205	3048	3006	3132	
	CH _{3ss}	3232	3110	2951	2930	3031	
	CH _{3dd}	1653	1556	1419	1467	1498	
	CH _{3ds}	1652	1549	1414	1464	1490	
	CH _{3r}	1312	1213	1140	1182	1206	
	CF _s	1186	1102	1095	1049	1059	
	ZPE	26.6	25.3	23.8	23.9	24.7	
	N ₂ H ₂	NH _{ss}	3613	3334	3144	3131	
		NH _{ss}	3578	3386	3125	3128	
NH _b		1911	1749	1635	1583		
		1472	1308	1276	1286		
NN _s		1763	1523	1517	1529		
Torsion		1472	1328	1291	1359		
ZPE		19.7	18.1	17.1	17.2		
HNO		NH _s	1551	2999	2709	3039	3039
	Bend	1979	1586	1657	1593	1505	
	NO _s	1733	1479	1479	1505	1564	
	ZPE	10.4	8.7	8.4	8.8	8.7	
H ₂ O ₂	OH _s	4095	3731	3661	3608		
		4091	3710	3658	3599		
	OH _b	1630	1390	1387	1402		
		1494	1294	1269	1266		
	OO _s	1161	926	968	877		
	Torsion	397	329	397	371		
	ZPE	18.4	16.3	16.2	15.9		

^a From Ref. 1, Table 6.41.

culated to be too low on the LSD/DZVPP level of theory.

A comparison of the vibrational zero point energies obtained from experiment and various levels of theory provide a global measure for the agreement between theory and ex-

periment. The values obtained from the LSD calculations are consistently smaller than those obtained from Hartree-Fock and MP2 computations. The differences between the LSD results and experiment are about the same as between

the MP2 calculations and experiment. For C_2H_4 , C_2H_6 , H_2CO , and CH_3F , the LSD and MP2 values for the zero point energies bracket the (harmonic) experimental values.

It should be noted that the values reported in Table IV were obtained by a finite difference technique using analytic first derivatives and atomic displacements of 0.02 a.u. Given the numerical sensitivity in this method, there is a residual uncertainty in the values. Given the evidence from Table II, this uncertainty should be less than about 30 cm^{-1} .

C. C–H bond energies

In order to compute the bond energies for the successive removal of H atoms from methane, calculations have been performed on the H(2S) and C(3P) atoms and the molecules CH($^3\Pi$), CH_2 (3B), CH_3 (2A), and CH_4 (1A). Full geometry optimizations were carried out leading to the structures reported in Table V. As in the previous cases, all calculations were performed on the LSD/DZVPP level. For the final geometries, nonlocal corrections to the total energy (NLSD) were evaluated using the functional forms proposed by Becke and Perdew (BP) as discussed in Sec. II C.

In Table V, the LSD and NLSD bond energies are compared with results from Hartree–Fock and correlated *ab initio* methods.^{1,66} This comparison is meaningful since the same type of basis sets, namely a valence double-zeta basis with polarization functions, has been used in all calculations given in Table V.

In the case of the CH–H bond, the LSD level of theory overestimates the experimental value by as much as 15 kcal/mol. In contrast, the NLSD approach agrees with the experimental values within 6 kcal/mol for the CH fragment and 2 kcal/mol for CH_3 –H and CH_2 –H. All correlated *ab initio* methods experience difficulties in calculating the bond energy of the CH radical. The best approach closed-coupled configuration interaction (CCCI), underestimates the energy by 6 kcal/mol.

In order to compare the various methods, the error (defined here as the average absolute deviation from experiment) is presented in the last column of Table V. While the LSD bond energies are consistently overestimated by about 10 kcal, the results show that, nevertheless, the energies obtained from LSD theory are closer to experiment than those computed at the Hartree–Fock level. A dramatic improvement in the calculated bond energies is found when nonlocal corrections are included. In fact, the NLSD energies are substantially better than the Hartree–Fock and ground valence bond (GVB) results, and are comparable with the sophisticated and computationally demanding MP4 and CCCI treatments. The vibrational frequencies obtained with LSD theory give vibrational zero point energies (ZPE) within 0.2 kcal/mol of the corresponding experimental values.

D. C–C bond energies and dissociation potential curves for ethylene

Table VI provides bond energies for the single, double, and triple bonds between carbon atoms in ethane, ethylene, and acetylene obtained with Hartree–Fock based methods^{1,66} and density functional theory. Again, all density

TABLE V. Calculated and experimental bond energies D_e for CH_3 –H, CH_2 –H, CH–H, and C–H defined with respect to the ground states of the corresponding fragments. The DGauss calculations were carried out on the DZVPP level. ZPE stands for the vibrational zero point energy. All energies are given in kcal/mol.

Theory	CH_3 –H	CH_2 –H	CH–H	C–H	Error ^a
HF ^b	87	88	101	55	22.0
MP2 ^b	109	110	109	73	5.5
MP4 ^b	110	112	107	76	3.5
CISD ^c	106	109	107	74	6.5
GVB-PP ^c	98	98	91	66	16.5
CCCI ^c	111	113	102	78	3.8
LSD	124	127	122	93	11.8
NLSD (BP)	114	118	113	90	3.8
Expt. ^c	112	116	107	84	
ZPE					
LSD	27	18	11	4	
Expt.	27	18	11	4	

^aThe average over absolute deviations between calculated and experimental results.

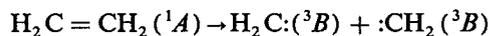
^b6-31G**//6-31G* results from Hehre *et al.* (Ref. 1, p. 274).

^cFrom Ref. 66.

functional calculations were carried out on the LSD/DZVPP level of theory with nonlocal corrections evaluated for the optimized LSD geometries. The CH_3 , CH_2 , and CH fragments were calculated in their electronic ground state as discussed above.

Whereas the LSD level of theory gives bond energies which are too large by 20–30 kcal/mol, the NLSD results lead to a surprising agreement with experiment. In fact, none of the correlated Hartree–Fock based methods shown in Table VI comes as close to experiment as the present NLSD approach.

An important aspect in the study of chemical bonding is a theory's ability to describe the entire energy hypersurface throughout a bond dissociation or bond formation process. To this end, the potential curve for the dissociation of the reaction



has been calculated using LSD and NLSD theory (cf. Fig. 1). The C–C bond distance was varied in steps between 1.0

TABLE VI. Calculated and experimental C–C bond energies D_e in kcal/mol for H_3C – CH_3 , $H_2C = CH_2$, and $HC \equiv CH$. The DGauss calculations were carried out on the DZVPP level.

Theory	H_3C – CH_3	$H_2C = CH_2$	$HC \equiv CH$
HF ^a	69	122	132
CISD ^a		154	194
GVB-PP ^a		146	165
CCCI ^a		174	214
LSD	115	204	268
NLSD	95	178	235
Expt. ^a	97	179	236

^aFrom Refs. 1 and 66.

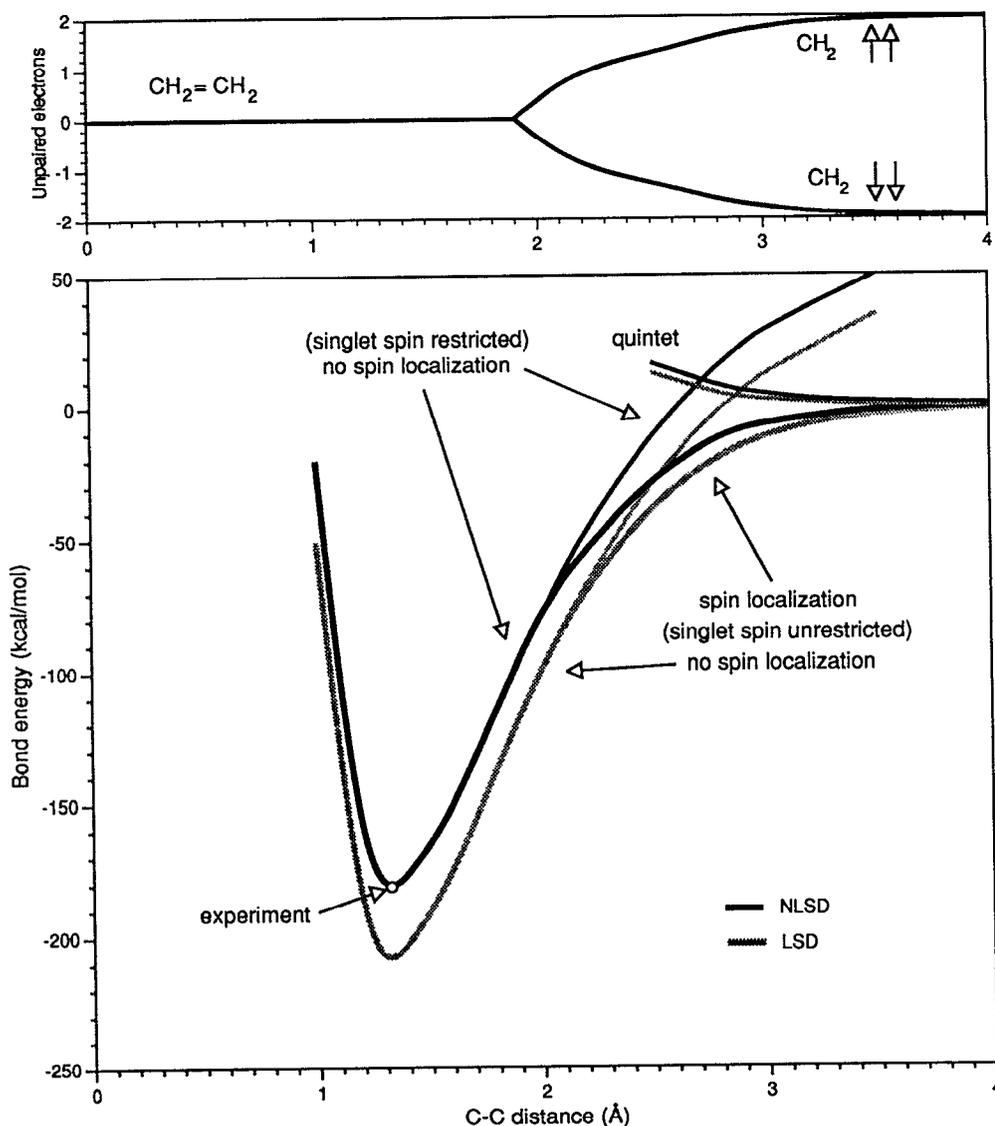


FIG. 1. The main panel shows the bond dissociation curves for the reaction $\text{CH}_2 = \text{CH}_2 \rightarrow 2\text{CH}_2$ obtained from LSD and NLSD calculations on the DZVPP level. "Singlet spin restricted" does not allow any spin polarization. "Singlet spin unrestricted" provides spin polarization (antiparallel spin) for C-C distances larger than about 2.0 Å. The "quintet" calculations assume spin polarization with parallel spin. The upper panel gives the number of unpaired electrons (negative sign for β spin) on the two CH_2 fragments as obtained from Mulliken populations.

and 4.0 Å and for each C-C distance, all other geometric parameters were optimized. As can be seen from Fig. 1, the overall shape of the potential curve is the same for both the LSD and NLSD approaches with the LSD energies being consistently lower than the NLSD values.

The assumption of a singlet state within a spin-restricted calculation and maintaining D_{2h} symmetry of the system throughout the dissociation process leads to a wrong dissociation limit. In fact, using this artificial restriction, SCF convergence problems appear at a C-C separation of about 4 Å, while the total energy is still rising with increasing C-C separation (cf. Fig. 1). A similar problem occurs for GVP-PP calculations as reported by Carter and Goddard.⁶⁶

In local spin density functional theory, the spin state of a molecule is defined by the occupation numbers of the spin α and β one-particle eigenstates.^{9,11,15} For example, a state is formally denoted quintet ($S = 2$) when there is an excess of four electrons in one of the spin manifolds. Calculations on the dissociated ethylene molecule in the quintet spin state (in the above mentioned sense) lead to the correct dissociation

limit since the σ and π electrons of each CH_2 fragment are now, correctly, triplet coupled. However, as the separation between the two C atoms diminishes, the energy of the quintet state of ethylene is higher than that of the singlet state (cf. Fig. 1). An antiparallel coupling of the triplet states of the CH_2 fragments is needed to describe correctly the dissociation process of ethylene. Spin unrestricted LSD calculations assuming a formal singlet state achieves this goal. In these calculations, the α and β spin densities around the two carbon atoms turn out to be different as the distance is increased. As a consequence, the symmetry of the system is reduced to C_{2v} . Now, at large distances, each fragment goes locally into a triplet state, but the spin densities on the two fragments have opposite signs and hence the spin state of the total system is formally still a singlet. As the two fragments approach each other, the magnitude of the spin density on each fragment diminishes until it vanishes for C-C distances smaller than 2.0 Å.

The upper panel in Fig. 1 shows the number of unpaired electrons on each CH_2 fragment as a function of C-C separa-

TABLE VII. Mulliken charges for the $\text{H}_2\text{C}=\text{CH}_2$ dissociation reaction as a function of the C=C bond distance $R(\text{CC})$ in Ångstroms.

$R(\text{CC})$	1.8	2.0	2.2	2.5	2.8	3.0	3.2	3.5	4.0
C1	3.18	3.37	3.67	3.86	4.03	4.10	4.15	4.19	4.21
H1	0.41	0.40	0.39	0.39	0.39	0.39	0.39	0.39	0.39
C2	3.18	3.00	2.71	2.52	2.35	2.27	2.22	2.18	2.16
H2	0.41	0.41	0.42	0.42	0.42	0.42	0.42	0.43	0.43
ΔE^b	0.0	0.1	2.3	9.2	17.8	24.0		38.5	

^a Mulliken charges for spin α are given; charges for spin β can be obtained from the ones for spin α interchanging the atom numbers.

^b Energy gain due to the localization of the spins $E_{\text{LSD}}(\text{localized}) - E_{\text{LSD}}(\text{nonlocalized})$ in kcal/mol.

tion. Starting from the equilibrium distance of 1.336 Å, the spin densities are zero up to a separation of about 2 Å. For increasing distances, the number of unpaired electrons on each fragment rises quickly to a value close to 2 near 4 Å. At this point, there are essentially two CH_2 fragments, each in a triplet state, but with antiparallel spin. This spin localization effect can be monitored conveniently by examining the Mulliken charges for spin α and β obtained from spin polarized calculations (cf. Table VII). Near the equilibrium distance, the Mulliken charge from the α electrons is the same on both carbon atoms. At larger distances, this charge increases on one carbon atom while it diminishes on the other. The opposite is true of the Mulliken charge from the β electrons. This spin separation is practically completed for a distance of about 4.0 Å. There is a considerable gain in energy due to this spin localization as can be seen from Table VII amounting to about 40 kcal/mol at a separation of 3.5 Å.

It is computationally advantageous to start this series of calculations with large C-C distances where the localization of the spins occurs spontaneously and then to diminish the separation step by step, using the densities of α and β electrons from the previous step as input for the smaller distance. It is gratifying to see that LSD theory describes this "antiparallel" spin system correctly. In fact, a somewhat similar case has been found for the Cr_2 molecule, where the spins on

the two Cr atoms have opposite signs (antiferromagnetic coupling). Also in the case of Cr_2 , LSD theory correctly describes the binding curve.⁶⁷

E. Bond dissociation energies involving the atoms C, N, O, and F

The direct calculation of dissociation energies for processes of the form $A - B \rightarrow A + B$ is known to be a difficult task for *ab initio* approaches. In fact, Hartree-Fock theory yields bond energies which are in a very poor agreement with experiment. In some cases, such as the dissociation of the F_2 molecule, Hartree-Fock theory predicts a negative dissociation energy and the use of correlated methods is mandatory. Often, the MP2 level of theory is sufficient to bring the agreement between calculated and experimental data within a few kcal/mol.

Table VIII provides a comparison of calculated and experimental bond dissociation energies. The LSD and NLSD results using the DZVPP computational level are presented together with HF, MP2, and MP4 results. The DZVPP basis set used in the density functional calculations is comparable to the 6-31G** basis underlying the HF and MP calculations.

The LSD results are too large by about the same amount as the HF results are too small. Both the LSD and HF levels

TABLE VIII. Bond dissociation energies D_e involving the atoms C, N, O, and F. The LSD and NLSD calculations were carried on the DZVPP level. All energies are given in kcal/mol.

Reaction	HF ^a	MP2 ^a	MP4 ^a	LSD	NLSD ^b	Expt. ^c
$\text{CH}_3-\text{CH}_3 \rightarrow \cdot\text{CH}_3 + \cdot\text{CH}_3$	69	99	97	115	95	97
$\text{CH}_3-\text{NH}_2 \rightarrow \cdot\text{CH}_3 + \cdot\text{NH}_2$	58	93	88	114	91	93
$\text{CH}_3-\text{OH} \rightarrow \cdot\text{CH}_3 + \cdot\text{OH}$	58	98	92	122	98	98
$\text{CH}_3-\text{F} \rightarrow \cdot\text{CH}_3 + \cdot\text{F}$	69	113	108	141	123 (117)	114
$\text{NH}_2-\text{NH}_2 \rightarrow \cdot\text{NH}_2 + \cdot\text{NH}_2$	34	73	67	101	74	73
$\text{HO}-\text{OH} \rightarrow \cdot\text{OH} + \cdot\text{OH}$	0	53	47	88	61	55
$\text{HO}-\text{F} \rightarrow \cdot\text{OH} + \cdot\text{F}$	-11	48	43	86	67 (61)	54
$\text{F}-\text{F} \rightarrow \cdot\text{F} + \cdot\text{F}$	-33	35	30	75	63 (51)	38
$\Delta E(\text{Expt.})^c$	47	2	6	28	7 (5)	0

^a 6-31G**//6-31G* and experimental results from Hehre *et al.* (Ref. 1, Table 6.61).

^b The values in parentheses refer to nonspherically symmetric F atoms as explained in the text.

^c The average error in bond energies comparing calculated and experimental results.

of theory are inadequate for quantitative statements about bond dissociation energies. The NLSD results, with the exception of molecules containing F, are comparable to the results from MP2 and MP4 calculations. On average, the NLSD values are comparable with the MP4 results. The F_2 molecule is a noticeable exception. It is possible that the limitations in the present basis set for F causes the relatively large error.^{48,50} In addition, the assumption of spherical symmetry in the F atom used in atomic calculations leads to a higher energy than in the case when this spherical symmetry is removed. If this effect is taken into account, the energy per fluorine atom is reduced by about 6 kcal, which improves the values for the bond dissociation energies significantly.

F. Energies of hydrogenation reactions

It is known that Hartree–Fock theory is quite successful in predicting the hydrogenation energies of compounds containing first row elements.¹ For many systems, the 6-31G* level of HF theory yields hydrogenation energies that are within 3–5 kcal/mol of the experimental values. The MP2 calculations bring improvements to the energetics, although there exist cases where even the MP4 level of theory gives errors of over 10 kcal/mol, e.g., in the reaction of hydrogen and oxygen. In Table IX, LSD and NLSD (BP) energies of selected hydrogenation reactions are presented as calculated using the DZVPP computational level. The resulting energies are compared with the HF, MP2, MP4, and experimen-

tal values as reported by Hehre *et al.*¹ Becke–Perdew (BP) corrected NLSD energies are listed in Table IX. The results obtained with Becke–Perdew–Pavlidou–Stoll corrections (cf. Sec. II C) show a somewhat larger error compared with experiment than the BP corrections.⁴⁷

The results given in Table IX are grouped into reactions involving single, double, and triple bonds, respectively. With the exception of a few reactions, notably the hydrogenation of O_2 , *ab initio* methods reproduce experimental values within ~4 kcal/mol. Surprisingly, the HF energies are on the average better than correlated results. In fact, the hydrogenation energies for molecules with single bonds seem to be particularly well captured by Hartree–Fock theory, although it is rather disturbing that systematic improvements over the single determinant description provided by MP2 and MP4 calculations increase the discrepancy between theory and experiment.

The LSD approach shows serious deficiencies for reactions involving triple bonds, while the hydrogenation of single bonds is fairly well described. On the other hand, the NLSD (BP) method provides a more uniform account for exchange and correlation for a variety of different bonds. On average, the NLSD approach performs as well as the correlated MP2 and MP4 methods. The energetics obtained with the NLSD approach are much closer to MP4 results than those obtained using the LSD method.

The comparison with experimental data cannot be done

TABLE IX. Energies of hydrogenation reactions involving C, N, O, and F atoms in units of kcal/mol. All LSD and NLSD calculations were carried out on the DZVPP level as defined in Table I.

Reaction	HF ^a	MP2 ^a	MP4 ^a	LSD	NLSD ^b	Expt. ^{a,c}
$CH_3-CH_3 + H_2 \rightarrow 2CH_4$	21	18	18	18	19 (16)	19 (16)
$CH_3-NH_2 + H_2 \rightarrow CH_4 + NH_3$	28	25	25	24	26 (23)	26 (23)
$CH_3-OH + H_2 \rightarrow CH_4 + H_2O$	30	28	29	28	28 (25)	30 (27)
$CH_3-F + H_2 \rightarrow CH_4 + HF$	27	26	25	27	26 (23)	29 (27)
$NH_2-NH_2 + H_2 \rightarrow 2NH_3$	48	46	45	43	44 (40)	48 (45)
$HO-OH + H_2 \rightarrow 2H_2O$	87	83	82	80	77 (73)	86 (83)
$F-F + H_2 \rightarrow 2HF$	134	126	121	129	122(118)	133(130)
ΔE^d	1.1	2.7	3.7	3.1	4.1	
$CH_2 = CH_2 + 2H_2 \rightarrow 2CH_4$	64	61	60	67	60 (48)	57 (48)
$CH_2 = NH + 2H_2 \rightarrow CH_4 + NH_3$	62	58	57	67	59 (47)	64 (55)
$CH_2 = O + 2H_2 \rightarrow CH_4 + H_2O$	58	55	52	67	57 (45)	59 (50)
$HN = NH + 2H_2 \rightarrow 2NH_3$	80	76	72	89	77 (64)	68 (58)
$HN = O + 2H_2 \rightarrow NH_3 + H_2O$	105	100	95	114	100 (86)	103 (93)
$O_2 + 2H_2 \rightarrow 2H_2O$	105	114	109	127	110 (98)	125(116)
$H_2C = CH_2 + 3H_2 \rightarrow 2CH_4$	118	111	111	131	114 (93)	105 (90)
ΔE^d	7.1	5.4	7.7	10.0	6.1	
$HC \equiv N + 3H_2 \rightarrow CH_4 + NH_3$	79	71	70	102	81 (61)	76 (61)
$C \equiv O + 3H_2 \rightarrow CH_4 + H_2O$	59	58	54	93	70 (51)	63 (49)
$N \equiv N + 3H_2 \rightarrow 2NH_3$	33	28	26	71	46 (25)	37 (22)
ΔE^d	3.7	6.3	8.7	30.0	7.0	

^a6-31G**//6-31G* results from Ref. 1, Table 6.65.

^bNLSD (BP) energies corrected for vibrational zero point energies are given in parentheses.

^cHeats of reaction corrected for zero point energies and extrapolated to 0 K are given in parentheses (from Ref. 1).

^dThe average error with respect to experimental data.

TABLE X. Energies, in kcal/mol, of reactions relating multiple and single bonds. All LSD and NLSD calculations were carried out on the DZVPP level.

Reaction	HF ^a	MP2 ^a	MP4 ^a	LSD	NLSD ^b	Expt. ^{a,c}
$\text{CH}_2 = \text{CH}_2 + 2\text{CH}_4 \rightarrow 2\text{CH}_3 - \text{CH}_3$	-22	-26	-24	-32	-22 (-17)	-20 (-17)
$\text{CH}_2 = \text{NH} + \text{CH}_4 + \text{NH}_3 \rightarrow 2\text{CH}_3 - \text{NH}_2$	-7	-8	-7	-20	-8 (-3)	-12 (-9)
$\text{CH}_2 = \text{O} + \text{CH}_4 + \text{H}_2\text{O} \rightarrow 2\text{CH}_3 - \text{OH}$	1	2	6	-11	0 (6)	1 (5)
$\text{HN} = \text{NH} + 2\text{NH}_3 \rightarrow 2\text{NH}_2 - \text{NH}_2$	16	16	18	4	17 (19)	28 (32)
$\text{O} = \text{O} + 2\text{H}_2\text{O} \rightarrow 2\text{HO} - \text{OH}$	70	52	49	33	44 (48)	47 (51)
$\text{HC} \equiv \text{CH} + 4\text{CH}_4 \rightarrow 3\text{CH}_3 - \text{CH}_3$	-54	-58	-57	-77	-58 (-46)	-49 (-44)
$\text{HC} \equiv \text{N} + 2\text{CH}_4 + 2\text{NH}_3 \rightarrow 3\text{CH}_3 - \text{NH}_3$	4	4	5	-29	-5 (5)	4 (9)
$\text{C} \equiv \text{O} + 2\text{CH}_4 + 2\text{H}_2\text{O} \rightarrow 3\text{CH}_3 - \text{OH}$	30	27	33	-9	14 (24)	27 (33)
$\text{N} \equiv \text{N} + 4\text{NH}_3 \rightarrow 3\text{NH}_2 - \text{NH}_2$	110	109	109	69	93 (97)	107 (112)
ΔE^d	6	4	5	23	7 ---	0

^a6-31G**//6-31G* results from Ref. 1, Table 6.68.

^bNLSD (BP) energies corrected for ZPE are given in parentheses.

^cHeats of reaction corrected for ZPE and extrapolated to 0 K are given in parentheses (from Ref. 1).

^dThe average error with respect to experiment.

exactly since the temperature corrections to vibrations were not calculated. We assume that the translational and rotational temperature contributions for hydrogenation reactions can effect the energetics by about 1 kcal/mol. The zero point energy contributions have been calculated and are shown in Table IX. In the earlier discussion of vibrational properties, it has been shown that vibrational zero point energies obtained with the LSD approach agree very well with experimental data for organic molecules (within 1 kcal/mol). Thus, a comparison of the theoretical and experimental values (given in parentheses in Table IX), which are corrected for zero point energies, can be assumed to be significant within about 2 kcal/mol. While the agreement between NLSD and experiment is quite reasonable for a substantial number of reactions, significant errors of up to about 10% do exist including the hydrogenation of H_2O_2 , F_2 , and O_2 .

G. Energies of isodesmic and other reactions

A set of typical organic reactions involving H, C, N, O atoms in various types of bonds are presented in Tables X and XI. The reaction energies were calculated by comparing reactants and products which were fully optimized at the LSD/DZVPP level.

The results reported in Table X confirm that the LSD method is not reliable in studying the energetics of reactions which are typical for organic chemistry. On the other hand, the NLSD approach overcomes most of the deficiencies of the local density approximation. In fact, the average error of 7 kcal/mol (see Table X) is close to the error of Hartree-Fock-based methods. In general, the largest error appears for reactions involving triple bonds.

Table XI contains several examples for isodesmic reactions as calculated by the HF, LSD, and NLSD methods. The results are compared with experimental heats of reactions corrected for zero-point vibrational energy and extrapolated to 0 K. Uncorrected heats are given in parentheses. On average (see Table XI), the NLSD method performs much better than the LSD approach. The NLSD results are comparable in accuracy to HF calculations. The average error of the NLSD approach in this series is 1.7 kcal/mol. The

largest discrepancy is 8.5 kcal/mol for the reaction with neopentane.

There is a need for a more systematic study of the influence of different nonlocal gradient corrections and the basis sets on the calculated heats of reactions. For example, the reaction involving HCN (cf. Table X), when calculated⁶⁸ with a triple-zeta basis for C and N, yields energies of -24 and 0 kcal/mol for LSD and NLSD (BP) calculations, respectively, compared with values of -29 and -5 kcal/mol which are obtained from the DZVPP level.

H. Nitro compounds— CH_3NO_2

Molecules containing nitro groups are difficult to treat using single determinant wave functions. For example, the SCF calculation incorrectly predicts the ground state of nitromethane to be a triplet. Multiconfiguration SCF (MCSCF) calculations correctly give the ground state as a singlet. In recent studies,⁶⁹ it was found that the optimized geometry obtained from LSD/6-31G** calculations agrees better with experiment than the results of MCSCF calculations using the same basis set. Although Pople's 6-31G* and 6-31G** basis sets seem to be adequate for LSD geometry optimizations, they are inadequate to describe reaction energies.³⁸

To this end, the geometry and energetics of nitromethane are studied on the DZVP2 level. The results, given in Table XII, are compared with MCSCF calculations and experiment as reported by Redington and Andzelm.⁶⁹ For comparison, LSD/6-31G** calculations are also presented. The small differences in the optimized geometries between the present results and those reported in Ref. 69 are due to more accurate gradient calculations performed in the present study.

The geometries obtained using both 6-31G** and DZVP2 basis sets are quite similar and compare well with MCSCF and experimental data. The differences in the dipole moments are more pronounced. The LSD-optimized basis set is more diffuse than the 6-31G** set and this leads to larger delocalization of the charge distributions. Contrary to the energetics of isodesmic reactions, which are im-

TABLE XI. Calculated and experimental bond separation energies in kcal/mol. All LSD and NLSD calculations were carried out on the DZVPP level.

Reaction	HF ^a	LSD		NLSD ^b	Expt. ^{a,c}
CH ₃ CH ₂ CH ₃ + CH ₄ → 2CH ₃ CH ₃	0.8	3.0	1.3	(2.0)	(2.6)
CH(CH ₃) ₃ + 2CH ₄ → 3CH ₃ CH ₃	1.9	7.5	3.0	(4.9)	(7.5)
C(CH ₃) ₄ + 3CH ₄ → 4CH ₃ CH ₃	7.3	14.8	4.6	(7.9)	(13.1)
CH ₃ CH ₂ NH ₂ + CH ₄ → CH ₃ CH ₃ + CH ₃ NH ₂	2.5	4.9	3.2	(3.9)	2.9 (3.6)
CH ₃ NHCH ₃ + NH ₃ → 2CH ₃ NH ₂	2.1	4.2	2.8	(3.5)	3.8 (4.4)
(CH ₃) ₃ N + 2NH ₃ → 3CH ₃ NH ₂	5.2	10.4	6.8	(8.6)	9.5 (11.1)
CH ₃ CH ₂ OH + CH ₄ → CH ₃ CH ₃ + CH ₃ OH	4.1	6.5	4.7	(5.4)	5.0 (5.7)
CH ₃ OCH ₃ + H ₂ O → 2CH ₃ OH	2.8	5.1	3.1	(4.1)	4.4 (5.4)
CH ₂ F ₂ + CH ₄ → 2CH ₃ F	12.5	15.3	12.6	(13.0)	13.3 (13.9)
CHF ₃ + 2CH ₄ → 3CH ₃ F	32.3	37.8	30.6	(32.0)	32.9 (34.6)
CF ₄ + 3CH ₄ → 4CH ₃ F	49.6	57.0	45.1	(47.9)	49.3 (52.8)
CH ₃ CHCH ₂ + CH ₄ → CH ₃ CH ₃ + CH ₂ CH ₂	3.9	7.0	5.2	(6.1)	4.7 (5.4)
CH ₃ CHO + CH ₄ → CH ₃ CH ₃ + H ₂ CO	9.9	13.5	11.8	(12.4)	10.7 (11.4)
CH ₃ CN + CH ₄ → CH ₃ CH ₃ + HCN	11.7	14.0	12.5	(13.1)	14.4 (14.7)
CH ₂ CO + CH ₄ → CH ₂ CH ₂ + H ₂ CO	13.3	21.3	18.6	(19.2)	15.0 (15.7)
CH ₂ CHCHCH ₂ + 2CH ₄ → CH ₃ CH ₃ + 2CH ₂ CH ₂	11.2	17.2	14.7	(15.9)	11.3 (14.2)
C ₃ H ₆ (cyclopropane) + 3CH ₄ → 3CH ₃ CH ₃	-26.2	-26.2	-22.6	(-17.5)	-22.7 (-19.6)
C ₂ O ₂ (oxacyclopropane) + 2CH ₄ + H ₂ O → CH ₃ CH ₃ + 2CH ₃ OH	-19.1	-19.2	-13.8	(-8.5)	-13.7 (-10.5)
C ₃ H ₄ (cyclopropene) + 3CH ₄ → 2CH ₃ CH ₃ + CH ₂ CH ₂	-50.4	-50.1	-44.8	(-38.7)	-43.9 (-40.5)
ΔE ^d	2.0	3.4	1.7	...	0

^a6-31G**//STO-3G results from Ref. 1, Table 6.71.

^bNLSD (BP) energies corrected for ZPE are given in parentheses.

^cExperimental heats of reactions corrected for ZPE and extrapolated to 0 K are given in parentheses (from Ref. 1).

^dThe average error as compared with experimental values.

TABLE XII. Geometries, dipole moments, and relative energies of singlet and triplet states of CH₃NO₂, comparing LSD, NLSD (using both 6-31G** and DZVP2 basis sets), MCSCF, and experimental results.

	Triplet		Singlet			Expt. ^a
	LSD		LSD	MCSCF ^a	Expt. ^a	
	6-31G**	DZVP2	6-31G**	DZVP2		
Geometry						
C-N	1.441	1.441	1.479	1.476	1.477	1.489
N-O	1.304	1.309	1.225	1.230	1.222	1.224
C-H1	1.109	1.108	1.102	1.102	1.080	1.088
C-H2	1.105	1.102	1.102	1.097	1.076	1.088
N-C-H1	110.0	109.9	108.4	108.4	108.5	107.2
N-C-H2	108.2	108.2	107.2	107.0	107.1	107.2
O-N-O	104.4	104.5	126.1	125.7	125.5	125.3
Dipole moment	3.25	3.54	3.36	3.59	...	3.46
Singlet-triplet energy difference (eV)						
LSD/6-31G**		2.75 ^b				
LSD/DZVP2		2.74				
NLSD/DZVP2		2.48				
MCSCF ^a		2.52				

^aFrom Ref. 68.

^bThe value of 3.63 eV reported in Ref. 69 is in error. The correct value within the computational level used in Ref. 69 is 2.75 eV.

TABLE XIII. Ground state geometries and relative energy differences of $F_2C=C$ and $FC=CF$ obtained from LSD, Hartree-Fock SCF, CISD, and CCSD calculations using basis sets of double-zeta + polarization quality. Bond lengths are given in Ångstroms, energy differences in kcal/mol. NLSLSD refers to the Becke-Perdew nonlocal corrections.

Level ^a	$FC\equiv CF$		$F_2C=C$		ΔE	
	CC	CF	CC	CF	LSD	NLSLSD
DZVP-3c	1.203	1.286	1.351	1.317	37.6	25.7
DZVP-1c	1.204	1.284	1.355	1.315	39.5	30.8
DZVP	1.204	1.284	1.355	1.314	39.2	30.6
DZVP-1f	1.204	1.284	1.355	1.315	39.5	30.8
DZVP2-2f	1.205	1.288	1.358	1.320	39.5	31.3
DZP SCF ^b			1.331	1.294	30.0	
DZP CISD ^b			1.348	1.311		
DZP CCSD ^b	1.196	1.296	1.362	1.322	25.4	

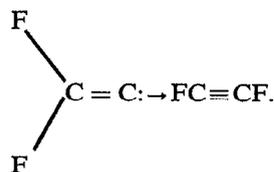
^aSee Table I for an explanation of the computational levels.

^bFrom Ref. 71.

proved significantly when LSD-optimized basis sets are used, the difference between singlet and triplet states of nitromethane does not depend as much on the level of basis set. The singlet state is found to be lower in energy than the triplet state by 2.74 eV on the LSD level and 2.48 eV on the NLSLSD (BP) level. This result corresponds very well with MCSCF results of 2.52 eV by Marynick *et al.*⁷⁰

I. Isomerization of unsaturated fluorocarbons— $F_2C=C$ and $FC\equiv CF$

Recent accurate DZP CCSD calculations on difluorovinylidene $F_2C=C:$ and difluoroacetylene $FC\equiv CF$ by Gallo and Schaefer,⁷¹ together with recent measurements of a matrix isolated IR spectrum for $FC\equiv CF$ ⁷² provide an excellent reference to assess the accuracy of the present DGauss approach. From correlated *ab initio* calculations,⁷¹ an isomerization energy of about 25 kcal/mol is predicted for the following ground state rearrangement:



In Table XIII, geometric and energetic properties of $FC\equiv CF$ and $F_2C=C$ are reported, while the vibrational frequencies and infrared spectrum are shown in Table XIII. In order to establish the dependence of the results on the choice of basis sets and accuracy of computational parameters, a series of DGauss calculations have been carried out using basis sets and computational parameters as given in Table I.

The examination of the results in Tables XIII and XIV lead to the following conclusions: The equilibrium bond distances vary by less than 0.01 Å within the range of the chosen computational parameters. The most economical level

TABLE XIV. Harmonic vibrational frequencies (in cm^{-1}) and infrared intensities (in km/mol) for $F_2C=C$ and $FC\equiv CF$. For the latter molecule, only IR active frequencies are given. The vibrational modes are denoted as follows: asymmetric stretch (as); stretch (s); scissor (sc); rock (r); symmetric stretch (ss); and wag (w).

Level ^a	$FC\equiv CF$		$F_2C=C$				
	CF_{as}	CC_s	CF_{as}	CF_{ss}	CF_{2w}	CF_{2sc}	CF_{2r}
DZVP-3c	1391	1708	1291	941	530	483	265
	355	46	298	60	4	3	37
DZVP-c1	1403	1692	1302	940	535	486	257
	362	49	291	62	3	4	32
DZVP	1406	1693	1298	941	539	491	257
	362	49	288	65	3	4	31
DZVP-1f	1402	1692	1298	942	539	490	258
	365	48	292	63	3	4	33
DZVP2-2f	1405	1684	1286	929	535	483	252
	364	44	301	65	3	5	32
CCSD ^b	1377	1724	1326	939	540	512	339
	309	71	302	74	5	5	33

^aSee Table I for an explanation of the computational levels.

^bFrom Ref. 71.

DZVP-3c performs almost as well as the standard DZVP level. Apparently, the choice of orbital basis set influences the geometry more than other parameters, including fitting, auxiliary functions, and size of the grid.

The relative energies of difluoroacetylene and difluoro-vinylidene defined as

$$\Delta E = E(\text{F}_2\text{C}=\text{C}\cdot) - E(\text{FC}\equiv\text{CF})$$

show a greater dependence on the choice of computational parameters. The computationally economic DZVP-3c level is inadequate, whereas the standard DZVP level seems to be converged sufficiently within the class of polarized double-zeta basis sets.

The harmonic vibrational frequencies and infrared intensities (see Table XIV) show satisfactory consistency of the results. The largest relative deviation of 6% is found for the frequency of the lowest mode, which is a rocking mode of the CF_2 group. Again, the DZVP-3c level may be inadequate especially for lower frequencies. Apparently the auxiliary basis A_3 , which is more constrained than the A_1 set, is not sufficient to describe the low motion of the CF_2 group. As expected, infrared intensities are more dependent on the choice of orbital basis set than the corresponding frequencies.

A direct comparison of the DFT calculations with the CCSD results of Gallo and Schaefer⁷¹ is meaningful since double-zeta plus polarization functions were used in both theoretical approaches. Excluding the crude DZVP-3c level, the LSD and CCSD C–C and C–F bond distances agree within 0.009 and 0.012 Å, respectively. Using the NLSD (BP) corrections, the energetics is considerably closer to the CCSD results compared with that resulting from the LSD level. The difference of 5 kcal/mol between NLSD and CCSD theories could be attributed to the incomplete treatment of correlation by the CCSD method and to the approximations inherent in the BP corrections.

The vibrational frequencies, obtained on the LSD level, agree quite well with the CCSD results except for the low CF_2 rocking mode. For this vibration, the LSD results are markedly lower than CCSD values. The same main features of the IR spectrum are predicted by both CCSD and LSD theories. The infrared intensities of the normal modes also agree quite well between the two different theoretical approaches with the exception of the CC symmetric stretch of $\text{F}_2\text{C}=\text{C}$, for which LSD theory predicts a significantly weaker intensity.

VI. SUMMARY AND FUTURE PERSPECTIVES

In this work, we have presented the theory, implementation, and applications of a density functional Gaussian-type orbital approach for the calculation of molecular geometries, vibrational properties, and reaction energies. The current implementation, called DGauss, uses contracted Gaussian-type orbitals including d functions as orbital basis. The electron density is expanded variationally in auxiliary basis sets, which are constructed from even-tempered uncontracted Gaussians. A similar expansion is used for the exchange-correlation terms, where the expansion coefficients are obtained from numerical integration using a density-adaptive

grid. A direct SCF method is employed to solve the Kohn–Sham equations. Energy gradients are evaluated analytically thus allowing the efficient optimization of molecular geometries. Second derivatives are calculated by a finite difference method. An important capability of the present implementation is its ability to calculate nonlocal corrections to the exchange-correlation energy.

The calculations reported here have been performed with LSD optimized orbital basis sets of double-zeta quality for the valence electrons plus polarization functions. These basis sets are known to be necessary for accurate predictions of reaction energies.^{38,47,57} Together with auxiliary basis sets of the type (7/3/3) for C-like atoms, a grid resolution of about 1000 points per atom, SCF convergence thresholds of 5×10^{-5} a.u. for the density fitting coefficients and 5×10^{-7} hartree for the total energy, and a gradient convergence threshold of 8×10^{-4} a.u., an accurate, yet computationally efficient computational level is established which is called DZVPP. This level corresponds to 6-31G** Hartree–Fock calculations. Dropping the polarization functions on the H atoms, a DZVP level is defined which is comparable with the 6-31G* level.

For small molecules containing C, N, O, H, and F atoms, equilibrium bond distances are predicted within about 0.01–0.02 Å. For example, the DZVPP level typically underestimates the single C–C bond and overestimates the $\text{C}\equiv\text{C}$ bond length by about 0.01–0.02 Å, while the calculated C=C bond distances agree within about 0.001 Å. H–X bond lengths ($X = \text{C}, \text{N}, \text{O}, \text{F}$) are typically overestimated by about 0.01–0.02 Å. Bond angles and dihedral angles are predicted with 1° – 2° . Equilibrium geometries predicted by LSD/DZVPP (and also LSD/DZVP) calculations are comparable to those obtained from the Hartree–Fock theory for typical organic and small inorganic molecules. However, LSD theory maintains this level of accuracy also for molecules such as FOOF,⁷³ C_2F_2 , and nitromethane, where the Hartree–Fock theory leads to much larger errors.

LSD/DZVPP calculations give vibrational frequencies which are consistently closer to experiment than those obtained with Hartree–Fock theory. In fact, the LSD results are comparable to MP2 results. In general, LSD/DZVPP frequencies are somewhat lower than experiment, while Hartree–Fock and MP2 calculations tend to be too high.

Nonlocal corrections to the total energy are found to be essential for the quantitative prediction of dissociation energies. For example, the local spin density (LSD) approximation gives a dissociation energy of 204 kcal/mol for ethylene, whereas the nonlocal corrections (NLSD) proposed by Becke and Perdew^{51,52} give a value of 178 kcal/mol which is almost suspiciously close to the experimental result of 179 kcal/mol. It is gratifying to see that the complete curve for the dissociation of ethylene into two methylene fragments is described properly by NLSD calculations, provided that a spin unrestricted approach is taken that allows an antiparallel orientation of the spins in the two ethylene fragments.

Density functional theory on the NLSD/DZVPP level gives absolute bond dissociation energies which are comparable to results from correlated methods such as MP2, MP4, and even MCSCF and coupled cluster methods. This accu-

racy, combined with the computational efficiency of the present density functional implementation, i.e., a scaling with less than a third power in the number of basis functions, promises to make this a powerful approach for the quantitative study of complex chemical reactions.

For example, in the current implementation, the direct SCF procedure for a system with 119 basis functions takes 55 s per processor on a CRAY Y-MP computer. In contrast to Hartree–Fock methods, the analytic evaluation of energy gradients is faster than the SCF. In fact, for the example mentioned above, the gradients take only 28 s of processor time. A molecule with 551 basis functions requires about 20 min for one SCF and about 8 min for the gradients.⁴⁷ Both the SCF and the gradient calculations lend themselves to computational parallelization.⁶⁰ Thus, it is conceivable to use the present implementation of density functional theory for the investigation of molecular dynamics, yet without the shortcomings of quasiclassical force–field models. Using today's computational technology, an energy and force calculation for a system consisting of 100 atoms can be done within about 1 h of computing time. If we assume a gain of three orders of magnitude in computing power over the next five to ten years by using novel high-speed computer architectures, a single step (energy and gradient) could then be accomplished within a few seconds. Using time steps of a few femtoseconds in a molecular dynamics calculations, the evolution of a 100 atom system over a few hundred picoseconds could be computed within about 100 h of computing time.

However, a number of major challenges remain to be met. These include improvements over the local spin density approximation that would allow the systematic approach of the correct density functional. Although the accuracy of the current approach in terms of molecular properties is quite impressive, it would be highly desirable to gain another order of magnitude. For example, equilibrium bond distances should be predicted within a few thousandths of one Ångstrom rather than a few hundredths and reaction energies should be predicted within fractions of 1 kcal/mol rather than within several kcal/mol. Another challenge is the inclusion of the molecular environment such as a solvent or a protein. To this end, accurate and efficient embedding methods need to be developed and tested. In the present approach, second derivatives are calculated by a finite difference technique using analytic first derivatives. This is computationally cumbersome for large systems, yet analytic second derivatives are needed for the optimization of complex molecular structures including the search for transition states.

Several aspects should be addressed in the near future. Those include the use of fully variational first derivatives⁴¹ and the systematic development and assessment of triple zeta polarized basis sets. Minimal basis sets would be very useful for the study of large molecular structures. The investigation and systematic development of effective core potentials and the extension of basis sets including *f* functions present additional opportunities.

In summary, the present results demonstrate clearly that density functional theory as implemented with Gaussian-type orbitals has become an accurate and efficient ap-

proach for the study of molecular geometries, vibrational properties, and reaction energies. The accuracy of this approach is comparable to elaborate correlated Hartree–Fock-based method while its computational efficiency surpasses that of single-reference Hartree–Fock calculations, especially for larger molecules. Hopefully, these characteristics will make the present approach a significant enrichment to the repertoire of quantum chemical methods and enable the investigation of complex molecular structures of scientific and technological importance.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the interactions, stimulating discussions, and valuable contributions by Dennis Salahub and his group at the University of Montreal and Brett Dunlap at the Naval Research Laboratory. The work of Nathalie Godbout in the development of the LSD optimized basis sets is highly appreciated. The collaboration with Andrew Komornicki was instrumental to build on the existing experience of geometry optimization techniques and vibrational analysis. Furthermore, the authors would like to thank Alain St. Amant, Axel Becke, Marc Benard, Bernard Delley, David Dixon, Stephane Doridot, René Fournier, Jan Gryko, Miroslav Kohout, Imre Papai, and Andreas Savin for their contributions in various stages of this effort. The chemistry team at the Industry, Science & Technology Department at Cray Research was helpful in this work through many stimulating and fruitful discussions. Last, but not least, the authors would like to thank the management at Cray Research, Inc. for enabling and supporting this work and to Corporate Computing and Networking at Cray Research for providing a highly professional computational environment.

¹ W. J. Hehre, L. Radom, P. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1986).

² J. Almlöf, K. Faegri, Jr., and K. Korsell, *J. Comp. Chem.* **3**, 385 (1982).

³ M. Haser and R. Ahlrichs, *J. Comp. Chem.* **10**, 104 (1989).

⁴ S. Obara and A. Saika, *J. Chem. Phys.* **84**, 3963 (1986).

⁵ M. Head-Gordon and J. A. Pople, *J. Chem. Phys.* **89**, 5777 (1988).

⁶ P. M. W. Gill, M. Head-Gordon, and J. A. Pople, *J. Phys. Chem.* **94**, 5564 (1990).

⁷ R. Friesner, *Chem. Phys. Lett.* **116**, 39 (1985).

⁸ E. Wimmer, A. J. Freeman, C.-L. Fu, P.-L. Cao, S.-H. Chou, and B. Delley, *ACS Symp. Ser.* **353**, (1987).

⁹ R. O. Jones and O. Gunnarsson, *Rev. Mod. Phys.* **61**, 689 (1990).

¹⁰ *Density Functional Methods in Chemistry*, edited by J. Labanowski and J. Andzelm (Springer, New York, 1991).

¹¹ T. Ziegler, *Chem. Rev.* **91**, 651 (1991).

¹² P. Hohenberg and W. Kohn, *Phys. Rev. B* **136**, 864 (1964).

¹³ W. Kohn and L. J. Sham, *Phys. Rev. A* **140**, 1133 (1965).

¹⁴ M. Levy, *Proc. Natl. Acad. Sci. (USA)* **76**, 6062 (1979).

¹⁵ R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University, New York, 1989).

¹⁶ E. Wimmer, H. Krakauer, and A. J. Freeman, in *Advances in Electronics and Electron Physics*, edited by P. Hawkes (Academic, New York, 1985), Vol. 65, p. 35.

¹⁷ E. J. Baerends, D. E. Ellis, and P. Ros, *Chem. Phys.* **2**, 41 (1973).

¹⁸ A. Rosen, and D. E. Ellis, *J. Chem. Phys.* **65**, 3629 (1976).

¹⁹ F. W. Averill and D. E. Ellis, *J. Chem. Phys.* **59**, 6412 (1973).

²⁰ B. Delley and D. E. Ellis, *J. Chem. Phys.* **76**, 1949 (1982).

²¹ B. Delley, *J. Chem. Phys.* **92**, 508 (1990).

²² R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).

²³ M. P. Teter, M. C. Payne, and D. C. Allan, *Phys. Rev. B* **40**, 12255 (1989).

- ²⁴ J. C. Slater, *Phys. Rev.* **51**, 846 (1937).
- ²⁵ O. K. Andersen, *Phys. Rev. B* **12**, 3060 (1975).
- ²⁶ E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, *Phys. Rev. B* **24**, 864 (1981).
- ²⁷ A. D. Becke, *Int. J. Quantum Chem. Symp.* **23**, 599 (1989).
- ²⁸ H. Sambe and R. H. Felton, *J. Chem. Phys.* **62**, 1122 (1975).
- ²⁹ B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, *J. Chem. Phys.* **71**, 3396 (1979); **71**, 4993 (1979).
- ³⁰ D. R. Salahub, in *Ab Initio Methods in Quantum Chemistry-II*, edited by K. P. Lawley (Wiley, New York, 1987), p. 447.
- ³¹ G. S. Painter and F. W. Averill, *Phys. Rev. B* **28**, 5536 (1983).
- ³² J. W. Mintmire and C. T. White, *Phys. Rev. Lett.* **50**, 101 (1983).
- ³³ K. Lee, J. Callaway, K. Kwong, R. Tang, and A. Ziegler, *Phys. Rev. B* **31**, 1796 (1985).
- ³⁴ N. Russo, J. Andzelm, and D. R. Salahub, *Chem. Phys.* **114**, 331 (1987).
- ³⁵ N. Rösch, P. Knappe, P. Sandl, A. Gorling, and B. I. Dunlap, *ACS Symp. Ser.* **394**, 180 (1989).
- ³⁶ B. I. Dunlap, in Ref. 10, p. 49.
- ³⁷ D. R. Salahub, R. Fournier, P. Mlynarski, I. Papai, A. St-Amant, and J. Ushio, in Ref. 10, p. 77.
- ³⁸ D. A. Dixon, J. Andzelm, G. Fitzgerald, E. Wimmer, and P. Jasien, in Ref. 10, p. 33.
- ³⁹ J. W. Mintmire, in Ref. 10, p. 125.
- ⁴⁰ S. H. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).
- ⁴¹ B. I. Dunlap, J. Andzelm, and J. W. Mintmire, *Phys. Rev. A* **42**, 6354 (1990).
- ⁴² J. Andzelm, E. Wimmer, and D. R. Salahub, *ACS Symp. Ser.* **394**, 228 (1989).
- ⁴³ R. Fournier, J. Andzelm, and D. R. Salahub, *J. Chem. Phys.* **90**, 6371 (1989).
- ⁴⁴ B. I. Dunlap and N. Rösch, *J. Chim. Phys. Phys. Chim. Biol.* **86**, 671 (1989).
- ⁴⁵ A. St-Amant and D. R. Salahub, *Chem. Phys. Lett.* **169**, 387 (1990).
- ⁴⁶ F. W. Averill and G. S. Painter, *Phys. Rev. B* **32**, 2141 (1985).
- ⁴⁷ J. Andzelm in Ref. 10, p. 155.
- ⁴⁸ L. Versluis, and T. Ziegler, *J. Chem. Phys.* **88**, 322 (1988).
- ⁴⁹ A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
- ⁵⁰ A. D. Becke, *ACS Symp. Ser.* **394**, 165 (1989).
- ⁵¹ J. P. Perdew, *Phys. Rev. B* **33**, 8822 (1986).
- ⁵² A. D. Becke, *J. Chem. Phys.* **84**, 4524 (1986).
- ⁵³ H. Stoll, C. M. E. Pavlidou, and H. Preuss, *Theor. Chim. Acta* **49**, 143 (1978).
- ⁵⁴ P. C. Hariharan and J. A. Pople, *Chem. Phys. Lett.* **66**, 217 (1972).
- ⁵⁵ N. Godbout, J. Andzelm, E. Wimmer, and D. R. Salahub, *Can. J. Chem.* (in press).
- ⁵⁶ H. Tatewaki and S. Huzinaga, *J. Chem. Phys. B* **71**, 4339 (1979).
- ⁵⁷ J. Andzelm, E. Radzio, and D. R. Salahub, *J. Comput. Chem.* **6**, 520 (1985); **6**, 533 (1985).
- ⁵⁸ J. Andzelm, N. Russo, and D. R. Salahub, *J. Chem. Phys.* **87**, 6562 (1987).
- ⁵⁹ Asterix, *Ab initio* SCF-CI program, Laboratoire de chimie quantique, Strasbourg, 1989.
- ⁶⁰ J. Andzelm, GigaFlop Report, Cray Research Inc., 1989, p. 15.
- ⁶¹ A. D. Becke, *J. Chem. Phys.* **88**, 2547 (1988).
- ⁶² V. I. Lebedev, *Zh. V. Math. Fiz.* **16**, 293 (1976).
- ⁶³ A. H. Stroud, *Approximate Calculations of Multiple Integrals* (Prentice-Hall, Englewood Cliffs, N.J., 1971).
- ⁶⁴ W. McIver and A. Komornicki, *Chem. Phys. Lett.* **10**, 303 (1971).
- ⁶⁵ GRADSCF program, designed and written by A. Komornicki, Polyatomics Research Institute.
- ⁶⁶ E. A. Carter and W. A. Goddard III, *J. Chem. Phys.* **88**, 3132 (1988).
- ⁶⁷ N. A. Baykara, B. N. McMaster, and D. R. Salahub, *Mol. Phys.* **52**, 891 (1984).
- ⁶⁸ G. Fitzgerald and J. Andzelm (to be published).
- ⁶⁹ P. K. Redington and J. Andzelm, in Ref. 10, p. 411.
- ⁷⁰ D. S. Marynick, A. K. Ray, J. L. Fry, and D. A. Kleier, *J. Mol. Struct. (Theochem.)* **108**, 45 (1984).
- ⁷¹ M. M. Gallo and H. F. Schaefer III, *J. Chem. Phys.* **93**, 865 (1990).
- ⁷² J. C. Brahms and W. P. Dailey, *J. Am. Chem. Soc.* **111**, 8940 (1989).
- ⁷³ D. A. Dixon, J. Andzelm, G. Fitzgerald, and E. Wimmer, *J. Phys. Chem.* **95**, 9197 (1991).