JOURNAL
OF THE AMERICAN CHEMICAL SOCIETY

Registered in U.S. Patent Office.

VOLUME 97, NUMBER 6 MARCH 19, 1975

Ground States of Molecules. XXV.[1] MINDO/3. An Improved Version of the MINDO Semiempirical SCF-MO Method[2]

Richard C. Bingham, Michael J. S. Dewar,* and Donald H. Lo

Contribution from the Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712. *Received July 18, 1974*

Abstract: The problems involved in attempts to develop quantitative treatments of organic chemistry are discussed. An improved version (MINDO/3) of the MINDO semiempirical SCF-MO treatment is described. Results obtained for a large number of molecules are summarized.

The purpose of the work described in this series of papers has been to develop a quantum mechanical treatment of molecular structure and reactivity simple enough, accurate enough, and reliable enough to serve as a practical chemical tool in areas at present inaccessible to experimental study. Typical examples are the properties of transient reaction intermediates, transition states, and other species involved in the course of chemical reactions. We envisage the use of computers as the chemical instruments in a new kind of "experimental" technique to be used by chemists on the same kind of basis as infrared or nmr spectroscopy.

If the Schrödinger equation could be solved accurately for polyatomic molecules, such solutions would at once provide the information we seek. In practice we are forced to use very approximate solutions which by chemical standards are utterly inadequate. The only hope of success must then lie in one or the other of two possible directions. Since chemistry is concerned only with differences in energy between related systems, the errors introduced by our approximations may cancel in estimating such differences. Alternatively, we may be able to upgrade the accuracy of our calculations by introducing adjustable parameters. These represent respectively the so-called *ab initio* and semiempirical approaches to the problem.

It is evident that both approaches will in fact be entirely empirical since the cancellation of errors in *ab initio* calculations, if indeed it exists at all, can be established only on an empirical basis by comparison of calculated molecular properties with experiment. This should be emphasized, for the loaded term "*ab initio*" has certainly conveyed a wholly misleading impression of *a priori* rigor and accuracy to many organic chemists.[3]

Nearly all the *ab initio* studies in this area so far have made use of the treatment introduced simultaneously and independently by Roothaan[4] and Hall,[5] *i.e.*, an orbital approach in which the orbitals are written as linear combinations of a given set (basis set) of functions, usually Slater-Zener or Gaussian AOs. If a sufficiently large and well-chosen basis set is used, the results can approach the theoretical (Hartree-Fock) limit set by the orbital approximation. However, the neglect of electron correlation inherent in this still leads to errors of *ca.* 100 kcal/mol per atom in the calculated energies of organic molecules, *i.e.*, thousands of kilocalories per mole for molecules of quite modest size.

If the correlation energies of a collection of atoms were independent of their relative positions in space, the correlation energies would cancel. Calculations by the Roothaan-Hall[6] (RH) method would then give good estimates of heats of atomization which would be entirely sufficient for our purpose. Unfortunately, however, this is not the case. Very detailed RH calculations for diatomic molecules, certainly approaching the Hartree-Fock limit, have led to calculated heats of atomization that can be in error by as much as +100 (*e.g.*, N_2) or −100% (*e.g.*, F_2). The changes in correlation energy during chemical combination are therefore comparable with the corresponding bonding energies. The situation is made worse by the fact that practical considerations limit the size of the basis set that can be used in calculations for a polyatomic organic molecule. This of course introduces further errors into calculated energies.

If the change in correlation energy on bond formation were constant, RH calculations could still give good estimates of heats of reaction. This possibility has been studied in detail by Pople, *et al.*,[7] who have indeed found that heats of reaction can be reproduced well by RH calculations, even with very small basis sets (*e.g.*, STO-3G), but only if the reactants and products contain not only the same total number of bonds but also the same number of each kind of bond (CH, C—C, C═C, etc.). Otherwise the results are unsatisfactory, particularly if a small basis set is used. Thus the calculated (STO-3G) heat of trimerization of acetylene to benzene ($3HC{\equiv}CH \rightarrow C_6H_6$) is too negative by no less than 67 kcal/mol. Recent work suggests that serious errors also occur in calculations of activation energies[8] and in comparisons of isomeric "classical" and "nonclassical" species.[12]

Allowances for electron correlation can in principle be made by introducing configuration interaction (CI). In practice, however, the convergence is too slow.[13] Other *ab initio* approaches have been suggested that may well prove more successful;[15] however, all seem likely to suffer from an even more serious failing of the RH method, *i.e.*, excessive cost.

The cost of a single RH calculation becomes prohibitive for molecules of quite moderate size, especially if a large enough basis set is used for the results to have any hope of chemical validity. The situation is made very much worse by the need to calculate molecular geometries. The geometry of a molecule must be found by minimizing the energy with respect to the geometrical variables defining its structure, $(3n - 6)$ in number for a molecule containing n atoms. This can be done only by interpolation between energies calculated for a number of specific geometries around the energy minimum. The situation becomes even more troublesome in the case of reactions; here very extensive exploration of the potential surface is often needed to locate the transition state.

Many authors have evaded these difficulties by making assumptions concerning the geometry or symmetry of the reactants and transition states in a reaction, thus reducing the number of independent variables. This procedure is, however, quite unacceptable if the calculations are to be used in the way we envisage, to predict the course of reactions. Apart from the fact[19] that such assumptions can lead to unacceptable errors, their use destroys the predictive value of a calculation, for any such assumption is tacitly based on an assumed mechanism for the reaction and calculations based on it will therefore at least tend to reproduce that mechanism, right or wrong.[20]

For these reasons we have believed for many years that the only hope of success in our project must lie in a semiempirical approach. The basic idea is to take a treatment simple enough for calculations to be carried out at reasonable cost and to try to upgrade its accuracy by introduction of adjustable parameters. This is a familiar procedure; a classical example is the Debye–Hückel theory of strong electrolytes. In the present connection, however, it has been generally believed that such an approach could not possibly succeed and we have been almost alone in pursuing it.

The prejudice against semiempirical treatments has been based on earlier experience which seemed to indicate that different parameters are needed in calculations for different types of molecules or for calculating different molecular properties. If this were true, such calculations would certainly be of very limited significance and would give no real insight into molecular structure and chemical behavior. Our previous work[22] had indicated that the failings of earlier semiempirical treatments were in fact due solely to inefficient or inappropriate parametrization or to the use of procedures too crude to be rescued even by the introduction of parameters. We were able to develop a series of successively better treatments in which a single set of parameters allowed the calculation of varied properties of wide ranges of molecules. These earlier treatments all proved, however, to have Achilles' heels. Here we report a new version (MINDO/3) which has so far survived every test without serious failure, including calculations for several hundred molecules and several dozen chemical reactions. In cases where direct comparisons are possible, MINDO/3 seems to be more accurate than existing RH procedures and the cost of the calculations is less by five orders of magnitude.

Development of MINDO/3

Our initial work[23] was concerned with calculations for conjugated hydrocarbons, using the Hückel σ, π approximation, the Hückel π MO method, and a primitive computer. The results were so encouraging that when better facilities became available, we tried[24] a similar approach based on the Pople[25] π SCF-MO method, parametrized to reproduce heats of atomization and molecular geometries. In its final form[26] this proved astonishingly successful, the accuracy for hydrocarbons being comparable with that of the best thermochemical and structural data and that for heteroconjugated systems almost as good. We therefore felt certain that an analogous procedure, including all the valence electrons, both σ and π, should lead, if properly parametrized, to results of "chemical" accuracy at moderate cost.

The extension of the Pople π approximation to include all the valence electrons has been considered by Klopman[27] and Pople, *et al.*[28,29] The core approximation is retained, the valence electrons being assumed to move in a fixed core composed of the nuclei and inner-shell electrons.[30] The valence-shell electrons are treated using a minimum basis set and certain electron-repulsion integrals involving differential overlap are neglected. Some or all of the remaining integrals are then equated to parametric functions.

Pople, *et al.*, considered three approximations of this type: NDDO,[28] where only diatomic differential overlap is neglected; CNDO,[28] where all differential overlap is neglected; and INDO,[29,31] which differs from CNDO only by inclusion of one-center exchange integrals. Pople, *et al.*,[28,29] and Dixon[31] parametrized their treatments to mimic the results of RH calculations using minimum basis sets and these procedures have been widely used in calculations of structure and reactivity. Since they are approximations to a treatment known to be insufficiently accurate, it is not surprising that they have often given very poor results. Indeed, the prejudice against semiempirical methods has been based largely on this failure. We set out to parametrize them in an entirely different manner, to reproduce the results of experiment rather than those of dubious *ab initio* calculations.

While the neglect of diatomic differential overlap can be shown to be entirely reasonable and probably quite a good approximation,[32] the further simplifications made in CNDO and INDO are suspect. However, it soon became apparent that the parametrization of these treatments presents unexpectedly formidable problems which are much worse in the case of NDDO than CNDO or INDO. Calculations by NDDO are also slower and problems arise with storage of the much larger number of repulsion integrals involved. Therefore, although we believe that NDDO will ultimately form the optimum basis for a parametric treatment and although we have in fact developed a promising preliminary treatment based on it,[33] all our work has been based on INDO,[34] apart from a preliminary investigation[35] based on a treatment (PNDO), intermediate between INDO and NDDO, which we have since abandoned. To distinguish our INDO-based treatments from that[29] of Pople, *et al.*, we term them MINDO (modified INDO).[36,37]

The terms in the expression for the total energy fall into five groups; the coulombic interelectronic repulsions, the electron–core attractions, the core–core repulsions, the one-center exchange (resonance) terms, and the two-center exchange terms. In our approach some or all of these will be set equal to parametric functions and some or all of the latter will contain numerical parameters that can be adjusted to fit experimental data.

In order to keep a treatment of this kind manageable, the number of arbitrary functions and parameters needs to be as small as possible. The same type of function should therefore be used for each energy term, regardless of the atoms involved. Since we wish to fit both energies and geometries of molecules, at least two parameters will be needed for each pair of atoms. We should try to avoid introducing any more. The parametric functions are in principle arbitrary; in practice we believe that there can be no hope of success in a treatment such as this unless the functions chosen conform to physical realities.

In order to make the calculations invariant to rotation, it is necessary[29] in INDO (and so also in MINDO) to set all the repulsion integrals (ii,kk) between an AO i of atom m and an AO k of atom n equal to a common value γ_{mn}. The one-center repulsion integrals g_{ik}, however, can retain their individuality. Following Pople, *et al.*,[29] we use the Goeppert-Meyer-Sklar potential with neglect of penetration integrals for the core-electron attractions, the attraction (V_{in}) between an electron in an AO i of atom m and the core of atom n being given by

$$V_{in} = -C_n\gamma_{mn} \tag{1}$$

where C_n is the core charge of atom n in units of *minus* the electronic charge (e). With these assumptions the elements of the MINDO **F** matrix are given by

$$F_{ii} = U_{ii} + 0.5q_ig_{ii} + \sum_{k \neq i}^{(m)}(q_kg_{ik} - 0.5p_{ik}h_{ik}) - \sum_{n \neq m}\gamma_{mn}\left(C_n - \sum_{k}^{(n)}g_k\right) \tag{2}$$

$$F_{ik}(mm) = -0.5p_{ik}h_{ik} \tag{3}$$

$$F_{ik}(mn) = \beta_{ik}{}^{\mathrm{c}} - 0.5p_{ik}\gamma_{ik} \tag{4}$$

Equation 3 refers to off-diagonal elements between AOs (i,k) of the same atom (m) and eq 4 to those between AOs of different atoms (m,n). $\Sigma^{(p)}$ implies summation over AOs of atom p only, g_{ik} and h_{ik} are respectively the one-center coulomb and exchange integrals between AOs i and k, $\beta_{ik}{}^{\mathrm{c}}$ is the two-center one-electron exchange integral (core resonance integral), U_{ii} is a sum of the kinetic energy of an electron in the AO i of atom m and its potential energy due to attraction by the core of atom m, and q_i and p_{ik} are respectively the electron population of AO i and the bond order between AOs i and k.

The total energy E_{tot} of a molecule is given by a sum of the total electronic energy (E_{el}) and the total intercore repulsion (E_{c})

$$E_{\mathrm{tot}} = E_{\mathrm{el}} + E_{\mathrm{c}} = E_{\mathrm{el}} + \sum_{m>n}\sum CR_{mn} \tag{5}$$

where CR_{mn} is the potential energy of repulsion between the cores of atoms m and n.

The energy of atomization of a molecule is the difference in energy between it and the atoms of which it is composed. The energies of atoms are calculated from single-configuration wave functions, using the same assumptions and the same values for the integrals as in the corresponding molecular calculation. We make the further assumption that the calculated energies of atomization can be equated to experimentally determined heats of formation, the kinetic energy terms in the latter being taken into account by our parameters.

Our parametrization has to take care of three major possible sources of error: *i.e.*, (1) the neglect of coulombic electron correlation inherent in any orbital approximation; (2) additional errors due to the simplifying assumptions made in deriving MINDO from the RH treatment; (3) errors arising from the equation of calculated energies of atomization to measured heats of atomization.

Our approach to the first problem is based on that of Pariser and Parr.[38] It is well known that electron correlation has little effect on the overall electron distribution in a molecule, the first-order density matrices given by good RH calculations apparently corresponding closely to the truth. It should therefore be possible to allow for electron correlation by appropriate reductions in the electron repulsion integrals. The appropriate corrections in the case of an atom can be deduced by fitting the energies of the atom, and of ions derived from it, to corresponding spectroscopic values. If the one-center integrals found in this way are then used in a molecular calculation, allowance will automatically be made for the effects of correlation between a given pair of electrons when they are near the same nucleus (one-center correlation). Long range correlation can then be taken into account by equating the electron repulsion integrals to suitable functions of internuclear distance (R_{mn}) such that as $R_{mn} \rightarrow 0$, the integral approaches an average of the corresponding one-center integrals of atoms m and n. In MINDO the only surviving two-center repulsion integrals are the γ_{mn}. These will then be represented by some function f_1 of R_{mn}

$$\gamma_{mn} = f_1(R_{mn}) \tag{6}$$

such that as $R_{mn} \rightarrow 0, f_1$ tends to an appropriate average of the one-center coulomb integrals g_{ik} of atoms m and n.

The most logical way to use this approach would be first to calculate the MOs of a molecule using theoretical values for the integrals and then to calculate the total electronic energy using the modified (empirical) values. Since, however, the change in electron repulsion due to correlation apparently has little effect on the total electron distribution, it seems reasonable to hope that the same will be true if we use the modified repulsion integrals throughout. We have adopted this course since it is much simpler and since the results obtained using it have proved satisfactory.

In the original versions of MINDO (MINDO/1[36] and MINDO/2[37]) the one-center integrals g_{ik} and h_{ik} were found from the Slater-Condon parameters by the method used by Pople, *et al.*,[29] in INDO. Here simplifying assumptions are made concerning relationships between the integrals. Later an improved procedure was developed,[39] based on Oleari's[40] method, which allows all the g_{ij} and h_{ij} to be evaluated independently. This procedure is used in MINDO/3.

A number of types of function have been suggested for f_1 in eq 6. We have adopted that first proposed and used by Dewar and Sabelli[41] in π calculations by the Pople[25] method, later used in the same connection by Dewar, *et al.*,[24,26] and by Ohno[42] (to whom this approximation is often wrongly attributed; *e.g.* ref 37), extended to all-valence-electron calculations by Klopman,[27] and adapted to MINDO by Baird and Dewar;[36] *viz.*

$$\gamma_{mn} = e^2[R_{mn}{}^2 + 0.25(\rho_m + \rho_n)^2]^{-1/2} \tag{7}$$

where

$$\rho_m = e^2/\bar{g}_m \text{ and } \rho_n = e^2/\bar{g}_n \tag{8}$$

$\bar{g}_m$ and $\bar{g}_n$ being appropriate averages of the one-center coulomb integrals g_{ij} of atoms m and n, respectively. We used this expression because it had given excellent results in our π approximation.[26] It is quite possible that any of the other forms that have been suggested for f_1 might give equally good results but we doubt if they would prove superior.

Next we have to deal with the undoubtedly large errors introduced into the Hartree-Fock method by the simplifying assumptions made in deriving MINDO. This will have to be achieved by appropriate modification of the two remaining quantities in the expression for the total energy (eq 2-5), *i.e.*, the core resonance integrals β_{ik}^c and the core-core repulsions CR_{mn}.

Ruedenberg[43] has shown that the main contribution to the bonding energy of a molecule comes from terms involving the β_{ik}^c. These terms therefore represent the attractive forces that hold the atoms together. The equilibrium bond lengths are determined by a balance between the attractive forces and the overall interelectronic and intercore repulsions. The latter are all determined except for the core-core repulsions CR_{mn}. If then we choose for the β_{ik}^c and CR_{mn} appropriate functions of internuclear distances and bond angles, we should be able to reproduce the energies of molecules as functions of their geometries. Since we should try to use the same functions for β_{ik}^c and CR_{mn}, regardless of the atom pair in question, the functions must contain numerical parameters fitted to each atom pair. To avoid having more than two such parameters for each pair of atoms, we must then include one each in β_{ik}^c and CR_{mn}.

Mulliken[44] has shown that the physical significance of β_{ik}^c in the RH approximation implies that it should be proportional to the corresponding overlap integral S_{ik} and also to the sum of the two corresponding valence state ionization potentials I_i and I_k. It has also been shown[45] that the same should be true in the CNDO/INDO/NDDO group of approximations. These arguments suggest that an appropriate expression for β_{ik}^c will be of the form

$$\beta_{ik}^c = S_{ik}(I_i + I_k)f_2(R_{mn}) \tag{9}$$

where f_2 is a function of the internuclear distance R_{mn}[46] and contains a numerical parameter. Several expressions of this kind have indeed been suggested in the literature.

The repulsion between the cores of atoms m and n (*i.e.*, CR_{mn}) would normally be equated to the repulsion between two point charges (PCR_{mn}); *i.e.*

$$CR_{mn} = PCR_{mn} = C_mC_ne^2/R_{mn} \tag{10}$$

Here, however, we cannot do this for the following reason. Our use of the Goeppert-Meyer-Sklar potential with neglect of penetration integrals (eq 1) leads to an underestimate of the core-electron attractions. If we use the theoretical value (eq 10) for the core-core repulsions, we will then overestimate the net coulomb repulsions between atoms. In particular, we will have spurious repulsions between pairs of neutral atoms. Since coulomb interactions are long range forces, this will lead to large spurious contributions from pairs of neutral atoms in a molecule that are far apart. To avoid this we must ensure that the net coulomb repulsion between two distant neutral atoms vanishes. Since the electron-electron and core-electron interactions are then all equal to $C_mC_n\gamma_{mn}$, so also must be the core-core repulsion. Thus

$$CR_{mn} \longrightarrow C_mC_n\gamma_{mn} \text{ as } R_{mn} \longrightarrow \infty \tag{11}$$

In the π approximation[24,26] this difficulty was met by assuming $CR_{mn} = C_mC_n\gamma_{mn}$ for all R_{mn}. There, however, the geometry of a molecule was determined by a rigid framework of σ bonds. Here we have no such framework. If the net coulomb interactions between pairs of neutral atoms in a molecule vanish at all distances, the attractive exchange forces take over and the molecule consequently collapses. It is therefore necessary to equate CR_{mn} to some function of R_{mn} that becomes greater than $C_mC_n\gamma_{mn}$ for small values of R_{mn} while still obeying eq 11 for large R_{mn}. Given that CR_{mn} must also approximate to PCR_{mn} for small R_{mn}, a reasonable expression for CR_{mn} would seem to be

$$CR_{mn} = C_mC_n[\gamma_{mn} + (e^2R_{mn}^{-1} - \gamma_{mn})f_3(R_{mn})] \tag{12}$$

where f_3 is a function of R_{mn} that vanishes as $R_{mn} \to \infty$ and approaches unity, or some value close to unity, as $R_{mn} \to 0$.

Our problem is to determine the best types of function of R_{mn} to use for f_2 and f_3 in eq 9 and 12 and the best values for the parameters in them.

It might in principle be possible to make this choice on the basis of some detailed theoretical analysis of the MINDO approximation or from a study of the wave functions given by MINDO calculations. We can only say that we have been quite unsuccessful in numerous attempts of this kind and that we are consequently sceptical concerning the value of such an approach.

The parameters in MINDO are so interrelated that it is impossible to predict the effects of a given change in any given one. It also follows that the values of the parameters for different atoms, or for different orbitals in the same atom, need bear no relation to the corresponding "*ab initio*" quantities in an RH calculation. It will for example be noticed that the optimum Slater exponents in MINDO/3 are very different from those used in RH calculations or given by the Slater rules.

All our work has therefore been based on a very laborious purely empirical technique, the parameters for any given choice of f_2 and f_3 being found by the procedure of Dewar and Haselbach[37] in which they are chosen by a least-squares fit of calculated heats of atomization and geometries for a set of standard molecules to those observed. Any given choice of functions is then tested by the closeness of the overall fit to experiment, this being judged on the basis of chemical considerations.

In the original procedure, the fit was to the heat of atomization and the length of one bond in each of the standard molecules. We have modified this in three respects.

First, we also fitted one bond angle in each standard molecule. The necessary additional relations are identical with eq 11-17 of ref 37 with R replaced by the corresponding bond angle θ. A third matrix $[\phi_{ij}^\theta]$ is thus obtained and a corresponding extra term added to the expression for $[\phi_{ij}]$ (eq 18 of ref 37).

Second, as will appear presently, we found it necessary to treat the Slater exponents of our basis set AOs as parameters. This involves additional terms in eq 4, 12, 13, etc., of ref 37.

Third, the derivatives of the energy with respect to the parameters were found by direct differentiation rather than by finite difference.

These changes involved trivial modifications of the original computer program.

In the original parametrization procedure,[37] experimental lengths were assumed for all bonds other than those involved in the parametrization. The energies were thus fitted to a structure in which most of the bond lengths had experimental values. When the geometry was then calculated by minimizing the total energy with respect to all variables, the total energy naturally decreased. Thus the final energies were systematically too negative. To avoid this, we now use an iterative procedure. After the initial parametrization, the geometries of all the standard molecules are calculated by minimizing the energy. The parametrization is then repeated, using the calculated bond lengths and bond angles instead of the experimental ones. The geometries are then recalculated with the new parameters. If they deviate appreciably from those found in the first cycle, the whole process is repeated until self-consistency is reached.

An arbitrary factor in the parametrization is the relative weight to be attached to errors in the energies, in the bond lengths, and in the bond angles. Since this is purely a matter of chemical judgement, the only solution is to run the parametrization with various values for the weighting factors (see eq 18 of ref 37; in the present case there are two weighting factors since bond angles are included in the parametrization). Usually at least ten runs are needed to achieve the desired result.

A rather large amount of computation is therefore required. Thus parametrization for just two atoms, *e.g.*, C and H, using a set of 20 standard molecules, and using our latest program for calculating geometries (see below), requires over 1000 SCF calculations per cycle, *i.e.*, 30,000–50,000 for the whole parametrization. Moreover one such run has to be carried out to test each parametrization scheme.

Our first objective was to remove the two worst failings of MINDO/2 in the case of hydrocarbons, *i.e.*, the systematic overestimation of CH bond lengths and the underestimation of strain energies in small rings. Various expressions of the form of eq 9 have been proposed for the core resonance integrals. We tried all these and many others, together with a number of alternatives for f_3. Altogether over 500 combinations were tested. However, none of these proved significantly better than the simple one first used in the PNDO[35] approximation, *i.e.*

$$f_2 = B_{XY} \text{ and } f_3 = \exp(-\alpha_{XY}R_{mn}) \qquad (13)$$

where B_{XY} and α_{XY} are parameters characteristic of the atoms (X and Y) involved. Thus eq 9 reduces to the standard Mulliken or Wolfsberg–Helmholz form.

One minor failing of MINDO/2 had been a tendency to predict the presence of high-energy σ MOs for which there is no evidence from photoelectron spectroscopic studies. In the case of polyacetylenes these were of CC pσ:pσ type. This seemed to suggest that the core resonance integrals for pσ:pσ interactions were too small in relation to those for pπ:pπ ones and it occurred to us that such a discrepancy might also account for the underestimation of strain energies. However, various attempts to alter the balance, by using different functions (f_2 in eq 9), or different parameters, for pσ and pπ interactions led to no improvement either in strain energies or in the σ-MO energies.

The overlap integrals in eq 9 are calculated using Slater–Zener AOs. Hitherto we had used for the orbital exponents (ζ_i) standard Slater values which do not distinguish between 2s and 2p AOs. We decided to remove the restriction. To avoid additional parameters we first calculated the exponents for the 2s and 2p AOs of carbon from the corresponding core orbital energies U_{ss} and U_{pp}. The results were somewhat better but not much. We therefore finally decided to treat the ζ_i as parameters and to include them in the parametrization. This can be done very simply by adding appropriate terms to the parametrization equations; *e.g.*, by adding an additional sum $(\partial E^m/\partial\zeta_i)\delta\zeta_i$ to eq 4 of ref 37. This change finally solved the strain energy problem, the average error in the calculated heats of formation for microcyclic compounds now being similar to those for compounds of other types. The change in the orbital exponents also corrected two other errors in MINDO/2, the overestimation of CH bond lengths and of the heat of atomization of H_2.

Several additional refinements were also tried. These included attempts to allow for changes in the U_{ii} and ζ_i with changes in the net formal charge at the atom in question (*cf.* the VESCF method[47]) and the use of different orbital exponents for 2p AOs involved in pσ and pπ interactions. Here again these modifications were combined with various alternative forms for the functions f_2 and f_3 in eq 9 and 12. However, none of these changes led to any significant improvement and some were deleterious, leading, *e.g.*, to an incorrect ordering of MOs. In view of these rather extensive studies we doubt if any further improvement in MINDO can be achieved, at any rate along the lines we have followed in our parametrization.

The basic parametrization scheme having thus been established for hydrocarbons, we set out to extend it to heteroatoms. We soon found that the parametrization could not be carried out for a large number of elements at once (*e.g.*, C, H, O, N, F) since it failed to converge. We therefore adopted the values for C and H derived from studies of hydrocarbons and added successive elements one at a time. Thus the parameters for nitrogen (ζ_s^N, ζ_p^N, B_{CN}, B_{HN}, B_{NN}, α_{CN}, α_{HN}, α_{NN}) were found by taking a set of compounds containing C, H, and N, retaining for C and H the parameters found from hydrocarbons. The parameters for O were likewise found from a series of compounds containing C, H, and O. Using these parameters for N and O, the cross parameters (B_{NO}, α_{NO}) were then found from a set of compounds containing C, H, N, and O. Additional atoms were added by a similar stepwise procedure. One modification proved necessary. In the case of HN and HO bonds only, f_3 of eq 12 was replaced by

$$f_3 = \alpha_{HX}e^{-R_{mn}} \text{ (X = N or O)} \qquad (14)$$

Table I shows the values of the one-center integrals for H, B, C, N, O, F, Si, P, S, and Cl and Table II shows the corresponding Slater exponents. Tables III and IV show the corresponding bicentric parameters B_{XY} and α_{XY}.

The values for Si, P, and S, which neglect 3d AOs, were determined by Dewar, Lo, and Ramsden.[48] The parameters for boron are provisional and may need modification.

Calculation of Molecular Geometries

As indicated earlier, the geometry of a molecule has to be found by minimizing its total energy with respect to the corresponding geometrical variables. Procedures for such minimizations are well known[49] and fall into two groups. In the first (*e.g.*, SIMPLEX[50] or Brent[51]) only values of the function being minimized are needed whereas in the second (*e.g.*, Murtagh–Sargent[52] (MS) or Davidon–Fletcher–Powell[53] (DFP)) derivatives of the function with respect to the variables are also required. As might be expected, the latter procedures need far fewer function evaluations, an important factor since here each function evaluation involves an SCF calculation.

In the case of the full RH method, such derivatives can be found only at the expense of very extensive computation. The time involved in each function evaluation then increases so greatly that the advantage of minimization procedures of the second kind is lost. The situation is quite different in the case of approximations of the CNDO/INDO/NDDO type. Here derivatives of the energy with respect to geometrical variables can be found very easily, allowing the use of the more efficient optimization procedures. This of course still further increases the disparity in cost between semiempirical calculations and RH ones.

McIver and Komornicki[54] have developed a procedure for calculating geometries by the MINDO method, using Cartesian coordinates and the MS[52] method. They find the derivatives analytically by direct differentiation of the MINDO expression for the total energy. An alternative procedure has been developed here[55] using internal coordinates (bond lengths, bond angles, dihedral angles) together with the DFP[53] method. The derivatives are found by finite difference.[56] Both procedures are about equally fast and very much faster than the Simplex procedure we had used previously.[57,58] Also while Simplex often failed to converge for $N \geq 20$, no difficulties have been experienced here in complete geometry optimizations, without any assumptions of symmetry, etc., for [18]annulene[59] (N = 102) or protonated lysergic acid diethylamide[60] (LSD) (N = 144). In the latter case only 98 function evaluations were needed.

Table I. One-Center Integrals in MINDO/3

Integral[a]	H	B	C	N	O	F	Si	P	S	Cl
U_s	−12.505	−33.61	−51.79	−66.06	−91.73	−129.86	−39.82	−56.23	−73.39	−98.99
U_p		−25.11	−39.18	−56.40	−78.80	−105.93	−29.15	−42.31	−57.25	−76.43
g_{ss}	12.848	10.59	12.23	13.59	15.42	16.92	9.82	11.56	12.88	15.03
g_{pp}		8.86	11.08	12.98	14.52	16.71	7.31	8.64	9.90	11.30
g_{sp}		9.56	11.47	12.66	14.48	17.25	8.36	10.08	11.26	13.16
$g_{pp'}$		7.86	9.84	11.59	12.98	14.91	6.54	7.68	8.83	9.97
h_{sp}		1.81	2.43	3.14	3.94	4.83	1.32	1.92	2.26	2.42
$h_{pp'}$		0.50	0.62	0.70	0.77	0.90	0.38	0.48	0.54	0.67

[a] All values in eV.

Table II. Slater Exponents for Use in MINDO/3

Element	Slater exponent (Z) 2s	2p
H	1.300000[a]	
B	1.211156	0.972826
C	1.739391	1.709645
N	2.704546	1.870839
O	3.640575	2.168448
F	3.111270	1.419860
Si	1.629173	1.381721
P	1.926108	1.590665
S	1.719480	1.403205
Cl	3.430887	1.627017

[a] For 1s AO.

Configuration Interaction and Open-Shell Systems

Many reactions involve biradicals or biradical-like species as intermediates. For well-known reasons, such species cannot be described satisfactorily by closed-shell single-configuration wave functions. Problems therefore arise in the treatment of such reactions by procedures (such as MINDO) in which single-configuration wave functions are used.

The difficulty can be overcome by including configuration interaction (CI) with the lowest doubly excited configuration and a version of MINDO/2, including such CI throughout, has been described.[61] In spite of its obvious attractions, we have not adopted this expedient in MINDO/3 because of a technical difficulty. In the case of a single-configuration closed shell wave function, the bond order matrix is invariant, to a first approximation, for small changes in the molecular geometry. Derivatives of the energy with respect to the various geometrical parameters can then be found very easily, either by direct differentiation[54] or by finite difference.[55] If CI is included, the overall bond order matrix is now a sum of contributions by different configurations and these contributions are *not* first-order invariant for changes in the geometry. Any rigorous calculation of derivatives must therefore allow for the corresponding first-order changes in the bond order matrices of the contributing configurations and we have found no way to do this without much additional computation. One can of course calculate the derivatives in the usual way and hope for the best. We have indeed carried out such calculations for a large number of molecules. However, although the DFP geometry program does converge under these conditions, the number of function evaluations is greater by almost a power of ten. Including CI throughout in MINDO/3 would therefore increase the cost of calculations tenfold.

These studies showed, however, that inclusion of CI has very little effect on the calculated heats of atomization of "normal" molecules, making them more negative by only 2-3 kcal/mol. Nothing is then gained by including CI in studies of such systems. Larger deviations occur only in cases where a high degree of biradical character would be expected so chemical intuition is a fairly safe guide. If there is any doubt, it can be resolved at trivial cost by comparing heats of atomization calculated with and without CI for the geometry given by the standard MINDO/3 procedure. If the difference is significantly greater than the usual value of 2-3 kcal/mol, then CI should be included. If the species in question is an intermediate or product in a chemical reaction, then of course CI must be included throughout. Since the effect of CI on "normal" molecules is so small, little would be gained by reparametrizing MINDO/3 for CI calculations.

Similar problems arise in the case of open-shell systems, in particular radicals and triplet states. In MINDO/3 these have been treated by the "half-electron" method.[62-64] Here again the bond order matrix is not invariant for small changes in the geometry. Although the DFP optimization still succeeds, convergence is slower than for "normal" molecules. The difference is not great for radicals but becomes serious for triplet states. It is possible that these problems might be overcome by using a MINDO version derived from the unrestricted Hartree-Fock method.

Energies of Atomization *vs.* Heats of Atomization

One problem which we have so far evaded is the validity of equating MINDO/3 energies to heats of atomization. The quantity calculated in MINDO/3 is in principle an equilibrium energy whereas the observed heat of atomization contains in addition vibronic terms corresponding to zero point energy and vibronic energy, plus translational kinetic energy. In our procedure these terms must be absorbed in some way into the parameters.

One might indeed expect this to be possible; for the additivity of bond energies in molecules with localized bonds

Table III. MINDO/3 Parameters (B_{XY} in Equation 13) for Resonance Integrals

Atom pair	Parameter	Atom pair	Parameter	Atom pair	Parameter	Atom pair	Parameter
HH	0.244770	HCl	0.231653	CO	0.464514	NF	0.205347
HB	0.185347	BB	0.151324	CF	0.247494	OO	0.659407
HC	0.315011	BC	0.250031	CSi	0.411377	OF	0.334044
HN	0.360776	BN	0.310959	CP	0.457816	FF	0.197464
HO	0.417759	BO	0.349745	CS	0.284620	SiSi	0.291703
HF	0.195242	BF	0.219591	CCl	0.315480	PP	0.311790
HSi	0.289647	CC	0.419907	NN	0.377342	SS	0.202489
HP	0.320118	CN	0.410886	NO	0.458110	ClCl	0.258969
HS	0.220654						

Table IV. MINDO/3 Parameters (α_{XY} in Equations 12 and 13) for Core Repulsion Functions

Atom pair	Parameter[a]	Atom pair	Parameter[a]	Atom pair	Parameter[a]	Atom pair	Parameter[a]
HH	1.489450	HCl	2.089404	CO	1.820975	NF	2.861667
HB	2.090352	BB	2.280544	CF	2.725913	OO	1.537190
HC	1.475836	BC	2.138291	CSi	1.101382	OF	2.266949
HN	0.589380	BN	1.909763	CP	1.029693	FF	3.864997
HO	0.478901	BO	2.484827	CS	1.761370	SiSi	0.918432
HF	3.771362	BF	2.862183	CCl	1.676222	PP	1.186652
HSi	0.940789	CC	1.371208	NN	2.029618	SS	1.751617
HP	0.923170	CN	1.635259	NO	1.873859	ClCl	1.792125
HS	1.700698						

[a] In eq 12 except for HN and HO (eq 13).

certainly implies that the kinetic energy terms must follow an analogous additivity relationship. Errors due to this assumption may nevertheless be a factor in limiting the accuracy of our approach.

It should be pointed out that this difficulty is not at all insuperable in MINDO/3 because vibrational frequencies can be calculated very easily. This is particularly true in the McIver-Komornicki treatment.[54] It would therefore be quite feasible to parametrize MINDO/3 to reproduce equilibrium energies, these being converted to heat contents by adding calculated (MINDO/3) values of the kinetic energy contributions. This possibility represents a further potential advantage of the semiempirical approach because the calculation of vibration frequencies by the RH procedure would be a formidable undertaking.

Summary of Results

In order to test MINDO/3 as thoroughly as possible, calculations have been carried out for several hundred species, including neutral molecules, ions, radicals, carbenes, and triplet states. Here we will summarize the results which are described in detail in the following papers.

Our computer program automatically calculates the geometry and energy of a molecule by minimizing the energy with respect to all geometrical parameters. The heat of atomization is then found by subtracting the energies of the component atoms. For convenience, the heats of atomization are then converted to heats of formation, using experimental values (Table V) for the heats of formation of gaseous atoms. Thus while we quote energies of molecules in terms of their heats of formation from elements in their standard states at 25°C, the deviations from experiment really represent errors in the calculated heat of atomization, not in the heat of formation. This should be emphasized since an essential criterion[22] that must be satisfied if a given method is to give reliable predictions of mechanisms of reactions is that it should reproduce heats of atomization satisfactorily.

Figure 1 shows a plot of calculated *vs.* observed heats of formation for 193 compounds for which we have carried out MINDO/3 calculations and for which reasonably reliable thermochemical data are available. The line in Figure 1 is the theoretical line of unit slope, passing through the origin. Nearly all the points lie close to it, most of them within ±5 kcal/mol. Note that the linearity would have been even more striking if we had plotted the quantities actually calculated, *i.e.*, heats of atomization; for the spread in heats of atomization is much greater than that in heats of formation.

Table V. Heats of Formation of Gaseous Atoms

Element	ΔH_f[a]	Element	ΔH_f[a]
H	52.102	F	18.86
B	135.7	Si	106.0
C	170.89	P	79.8
N	113.0	S	65.65
O	59.559	Cl	28.95

[a] Heat of formation of gaseous atom in kcal/(g-atom) at 25°.

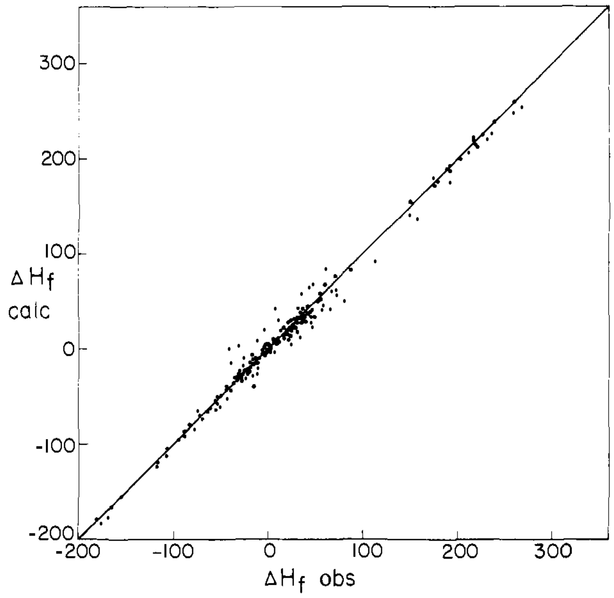

Figure 1. Plot of calculated *vs.* observed heats of formation (ΔH_f; kcal/mol at 25°) for 193 compounds derived from H, C, N, O, F, Si, P, S, and Cl.

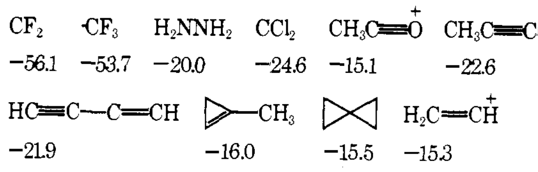

Figure 2. Compounds for which the difference (kcal/mol, shown below each formula) between the calculated and observed heats of formation is <−15.

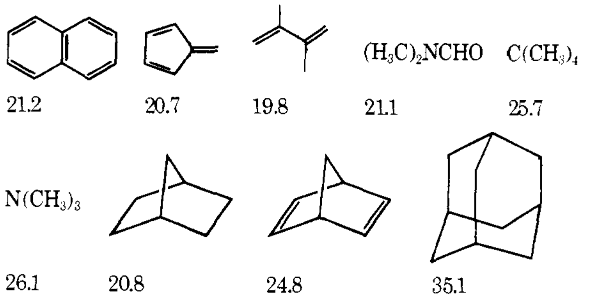

Figure 3. Compounds for which the difference (kcal/mol, shown below each formula) between the calculated and observed heats of formation is >15.

Figures 2 and 3 show the compounds for which the error in the calculated heat of formation is less than −15, or greater than 15, kcal/mol, respectively.

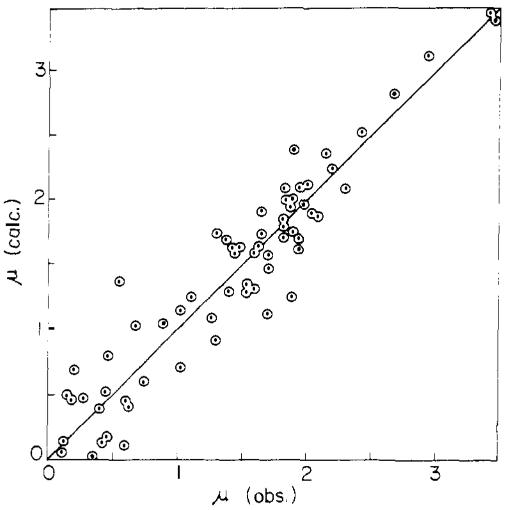

Figure 4. Plot of calculated *vs.* observed dipole moments (μ; Debye) for 65 compounds derived from H, C, N, O, F, Si, P, S, and Cl.

The only gross errors in Figure 2 are for CF_2 and CF_3. Problems arise in perfluorocarbons, as will be discussed in a later paper dealing with the calculations in detail. These may well be due to an unsatisfactory choice of FF parameters. Since the results for other fluorine compounds are uniformly good, we have not felt prepared to attempt the complete reparametrization that would be needed.

The errors for N_2H_4 and CCl_2 are to be expected since no INDO-based treatment can deal satisfactorily with compounds where adjacent atoms have unshared electrons in hybrid AOs.[22] The problem with FF parameters may also be due to this.

The other results suggest that MINDO/3 overestimates the stability of compounds containing triple bonds and that it still tends to underestimate the strain energies of small rings. However, the errors for the two small ring compounds in Figure 2 only just exceed 15 kcal/mol and those for a large number of others are much less.

The results in Figure 3 seem to suggest that MINDO/3 underestimates resonance energies and that it also suffers from a tendency to underestimate the stabilities of compact globular molecules. We think these faults may be further consequences of the neglect of one-center overlap in MINDO because they have persisted in all our parametrizations and because they also seem much less evident in a version of NDDO which has recently been developed.[33]

Clearly MINDO/3 is far from perfect but it does represent a major advance over previous treatments. It should be pointed out that the compounds plotted in Figure 1 include ions, radicals, and triplet states and that its success in treating them is indicated by the fact that few appear in Figures 2 or 3. The geometries given by MINDO/3 are also generally satisfactory, bond lengths usually being correct to 0.02 Å and bond angles to a few degrees. There are some fairly clear trends. Thus CH bonds tend to be too long by 0.01 Å and C–C bonds too short by 0.015 Å. If such corrections are made, the results compare very favorably with those from *ab initio* calculations.

Figure 4 shows a plot of calculated *vs.* observed dipole moments for the compounds for which data are available. Here again the agreement is generally good, most of the points lying within 0.2 D of the theoretical line of unit slope. Most of the larger deviations refer to compounds with adjacent atoms carrying hybrid lone pairs. As has already been pointed out, MINDO/3 cannot be expected to give good results for such compounds. The other deviations are mostly for nitriles. This error is probably related to that in the heats of formation of compounds containing triple bonds. The predicted charge distributions are generally similar to those given by *ab initio* SCF calculations. In short, MINDO/3 seems to give quite realistic electron distributions.

Calculations have also been carried out for over fifty reactions, the calculated activation energies mostly agreeing with experiment to within ±5 kcal/mol. These results will be found in the following papers of this set or in papers and communications in the course of publication. MINDO/3 has also been shown to give good estimates of a variety of other ground state properties of molecules, including first ionization potentials,[65] molecular polarizabilities,[66] ESCA chemical shifts,[67] ^{14}N nuclear quadrupole coupling constants,[68] electronic band structures in polymers,[69] and gas phase proton affinities.[70] It has also given satisfactory results for a number of "nonclassical" carbonium ions.[71] A particularly striking example is provided by the $C_3H_7^+$ ions where the MINDO/3 results[71] are in almost perfect agreement with experiment, whereas the results of a very detailed RH treatment by Pople, *et al.*,[12] using a basis set of double ζ type with inclusion of d AOs (6-31 G*), were much less satisfactory. And finally, some recent studies seem to indicate that MINDO/3 can also give good estimates of the energies of lowest excited singlet and triplet states.

Summary and Conclusions

MINDO/3 has thus proved to be an extraordinarily versatile procedure, giving good results for every ground state property so far studied and apparently offering hope of equally successful extension to excited states and photochemistry. While it has not yet achieved "chemical" accuracy, the average error in the calculated heats of atomization being *ca.* 6 kcal/mol instead of <1 kcal/mol, it has given no unreasonable results except in the one area where the MINDO approximation would be expected to fail.[72] The errors in the heats of atomization are in any case less by orders of magnitude than those given by existing RH calculations or other semiempirical MO procedures. Moreover the results for activation energies of reactions and for the heats of formation of "nonclassical" ions seem definitely superior to those from RH calculations, although admittedly there are few examples for comparison since few meaningful RH calculations have been reported.[73]

Even if cost were not a factor, the available evidence would seem to suggest that MINDO/3 may provide a better overall quantitative guide to chemical behavior than even good RH calculations. Since the cost of MINDO/3 calculations is in fact less by a factor of at least 100,000, there is really no contest in areas where MINDO/3 can be applied.

Admittedly organic chemistry as such probably no longer has many exciting aspects still to be discovered; the main value of an approach of the kind we seek would lie in the areas of organometallic chemistry, catalysis by transition metals, photochemistry, and biochemistry. It seems that photochemistry may already be within reach of MINDO/3 while a moderate increase in the power of computers would make studies of enzyme reactions feasible at reasonable cost. We can already handle molecules with up to 80 atoms (*cf.* recent calculations for LSD[60]), using a computer (CDC 6600) which is an order of magnitude slower than others now available, and even this would be sufficient for an adequate model of an active site. Organometallic chemistry, however, presents serious problems due to the deplorable lack of the thermochemical data that are needed to determine MINDO/3 parameters. Expenditure on research directed to the provision of such data would probably be of much more chemical value than a much greater expenditure on *ab initio* calculations made necessary by the unavailability of MINDO/3.

As regards the future; we think it likely that MINDO/3 represents the limit attainable by an INDO-based semiempirical treatment and we doubt if further efforts in this area will prove rewarding, apart from extensions to additional elements. However, the scope of semiempirical MO treatments remains almost unlimited since one can parametrize successively more sophisticated derivatives of the RH method. As noted above,[33] we have already started on NDDO.

Computer Programs

A complete MINDO/3 program, including options for open-shell calculations (radicals and triplets) by the "half-electron" method, and for inclusion of CI with the lowest doubly excited configuration (for biradical-like species), together with the DFP geometry program has been deposited with QCPE.

References and Notes

(1) Part XXIV: N. Bodor, M. J. S. Dewar, and Z. B. Maksić, *J. Amer. Chem. Soc.*, **95**, 5245 (1973).
(2) This work was supported by the Air Force Office of Scientific Research through Contract F44620-71-C-0119 and by the Robert A. Welch Foundation through Grant F-126. The calculations were carried out using the CDC 6400/6600 computer at the University of Texas Computation Center.
(3) Judging, *e.g.*, by comments at the 23rd National Organic Symposium, Tallahassee, Fla., June 1973, where this work was presented.
(4) C. C. J. Roothaan, *Rev. Mod. Phys.*, **23**, 69 (1951).
(5) G. G. Hall, *Proc. Roy. Soc., Ser. A*, **205**, 541 (1951).
(6) This treatment has been commonly described in the literature as the *ab initio* SCF method. For reasons indicated above, the term *ab initio* carries unfortunately misleading connotations. It seems to us much better to follow the usual practice of naming prodecures for their inventors and we have accordingly followed this course.
(7) See J. A. Pople, *Accounts Chem. Res.*, **3**, 217 (1970); W. J. Hehre, R. Ditchfield, L. Radom, and J. A. Pople, *J. Amer. Chem. Soc.*, **92**, 4796 (1970).
(8) Thus the activation energy calculated[9] for the rearrangement of methyl isocyanide to acetonitrile, using a large (double ζ) basis set, was too large by 20 kcal/mol. Another example is the conrotatory electrocyclic opening of cyclobutene to 1,3-butadiene; here a single-configuration treatment, using a minimum basis set, gave[10] an activation energy (65 kcal/mol) more than double that observed (32 kcal/mol). Introduction of CI improved the situation but the calculated activation energy (48 kcal/mol[11]) was still much too large.
(9) D. H. Liskow, C. F. Bender, and H. F. Schaffer III, *J. Amer. Chem. Soc.*, **94**, 5178 (1972).
(10) Reported by Dr. S. D. Peyerimhoff at the CECAM Colloquium on "Calculation of Reaction Paths and Reaction Mechanisms," Paris, Sept 1972.
(11) K. Hsu, R. J. Buenker, and S. D. Peyerimhoff, *J. Amer. Chem. Soc.*, **94**, 5639 (1972).
(12) See, *e.g.*, a recent very detailed calculation for the isomeric $C_3H_7^+$ ions: P. C. Hariharan, L. Radom, J. A. Pople, and P. v. R. Schleyer, *J. Amer. Chem. Soc.*, **96**, 599 (1974).
(13) It is interesting to note that Boys, who was the first to write computer programs for polyatomic RH calculations, the first to use Gaussian orbitals[14] in this connection, and the first to introduce CI into such calculations, abandoned this approach in 1958 as unpromising for chemical purposes. See, S. F. Boys, G. B. Cook, C. M. Reeves, and I. Shavitt, *Nature (London)*, **178**, 1207 (1956); S. F. Boys and G. B. Cook, *Rev. Mod. Phys.*, **32**, 285 (1960).
(14) The advantages of using Gaussian functions were first recognized by McWeeny; see R. McWeeny, *Nature (London)*, **166**, 21 (1960).
(15) For example, Boys'[16] transcorrelated wave function method, Goddard's[17] generalized valence bond method, and the ingenious procedure developed by Conroy.[18]
(16) S. F. Boys and N. C. Handy, *Proc. Roy. Soc., Ser. A*, **310**, 43, 63 (1969).
(17) W. A. Goddard III and R. C. Ladner, *J. Amer. Chem. Soc.*, **93**, 6750 (1971); P. J. Hay, W. J. Hunt, and W. A. Goddard III, *ibid.*, **94**, 638, 8293 (1972).
(18) H. Conroy, *J. Chem. Phys.*, **41**, 1327, 1331, 1336, 1341 (1964).
(19) Our own calculations have shown that even apparently innocuous assumptions concerning the geometry of parts of a molecule apparently uninvolved in a given reaction can easily lead to errors in the calculated activation energy of 20 kcal/mol.
(20) A good example is provided by calculations by Buenker, *et al.*,[21] for forbidden electrocyclic reactions of cyclopropyl ions and cyclobutene; see M. J. S. Dewar and S. Kirschner, *J. Amer. Chem. Soc.*, **96**, 5244 (1974).
(21) K. Hsu, R. J. Buenker, and S. D. Peyerimhoff, *J. Amer. Chem. Soc.*, **93**, 2117 (1971); P. Merlet, S. D. Peyerimhoff, R. J. Buenker, and S. Shih, *ibid.*, **96**, 959 (1974).
(22) See M. J. S. Dewar, *Fortschr. Chem. Forsch.*, **23**, 1 (1971); *Pure Appl. Chem.*, **23**, 1 (1971).
(23) M. J. S. Dewar and H. N. Schmeising, *Tetrahedron*, **11**, 96 (1960).
(24) A. L.-H. Chung and M. J. S. Dewar, *J. Chem. Phys.*, **42**, 756 (1965); M. J. S. Dewar and G. J. Gleicher, *J. Amer. Chem. Soc.*, **87**, 685 (1965).
(25) J. A. Pople, *Trans. Faraday Soc.*, **49**, 1375 (1953); A. Brickstock and J. A. Pople, *ibid.*, **50**, 901 (1954).
(26) M. J. S. Dewar and C. de Llano, *J. Amer. Chem. Soc.*, **91**, 789 (1969); M. J. S. Dewar and T. Morita, *ibid.*, **91**, 796 (1969); M. J. S. Dewar and A. J. Harget, *Proc. Roy. Soc., Ser. A*, **315**, 443, 457 (1970).
(27) G. Klopman, *J. Amer. Chem. Soc.*, **86**, 4550 (1964); **87**, 3300 (1965).
(28) J. A. Pople, D. P. Santry, and G. A. Segal, *J. Chem. Phys.*, **43**, S129 (1965); J. A. Pople and G. A. Segal, *ibid.*, **43**, S136 (1965); **44**, 3289 (1966).
(29) J. A. Pople, D. L. Beveridge, and P. A. Dobosh, *J. Chem. Phys.*, **47**, 2026 (1967).
(30) This is probably a very good approximation in view of the success of our π approximation where the σ electrons were also included in the core.
(31) This approximation was also suggested (EMZDO) by R. N. Dixon, *Mol. Phys.*, **12**, 83 (1967).
(32) See M. J. S. Dewar, "The Molecular Orbital Theory of Organic Chemistry," McGraw-Hill, New York, N.Y., 1969, p 84.
(33) M. J. S. Dewar, D. H. Lo, and C. A. Ramsden, to be published.
(34) The difference in complexity between CNDO and INDO is negligible and INDO has the advantage of distinguishing between states of different multiplicity in atoms. For example, CNDO involves the unrealistic assumption that the singlet and triplet configurations $(1s)^2(2s)^2(2p)^2$ of carbon have identical energies.
(35) M. J. S. Dewar and G. Klopman, *J. Amer. Chem. Soc.*, **89**, 3089 (1967).
(36) N. C. Baird and M. J. S. Dewar, *J. Chem. Phys.*, **50**, 1262 (1969).
(37) M. J. S. Dewar and E. Haselbach, *J. Amer. Chem. Soc.*, **92**, 590 (1970).
(38) R. Pariser and R. G. Parr, *J. Chem. Phys.*, **21**, 466, 767 (1953).
(39) M. J. S. Dewar and D. H. Lo, *J. Amer. Chem. Soc.*, **94**, 5296 (1972).
(40) L. Oleari, L. DiSipio, and G. De Michelis, *Mol. Phys.*, **10**, 97 (1966).
(41) M. J. S. Dewar and N. L. Hojvat (Sabelli), *J. Chem. Phys.*, **34**, 1232 (1961); *Proc. Roy. Soc., Ser. A*, **264**, 431 (1961); M. J. S. Dewar and N. L. Sabelli, *J. Phys. Chem.*, **66**, 2310 (1962).
(42) K. Ohno, *Theor. Chim. Acta*, **2**, 219 (1964).
(43) See K. Ruedenberg, *Rev. Mod. Phys.*, **34**, 326 (1962).
(44) See R. S. Mulliken, *J. Chem. Phys.*, **46**, 675 (1949); *J. Phys. Chem.*, **56**, 295 (1952).
(45) Reference 32, p 87.
(46) Angular variables must be excluded from f_2 since otherwise the resulting treatment would not be invariant to rotation; see ref 28.
(47) R. D. Brown and M. L. Heffernan, *Trans. Faraday Soc.*, **54**, 757 (1958).
(48) M. J. S. Dewar, D. H. Lo, and C. A. Ramsden, part XXIX of this series, following.
(49) See R. Fletcher, "Optimization," Academic Press, New York, N.Y., 1969.
(50) See J. A. Nelder and R. Mead, *Comput. J.*, **7**, 308 (1965).
(51) R. P. Brent, Ph.D. Thesis, Stanford University, 1971.
(52) B. A. Murtagh and R. W. H. Sargent, *Comput. J.*, **13**, 185 (1970).
(53) W. C. Davidon, *Comput. J.*, **10**, 406 (1968); R. Fletcher, *ibid.*, **8**, 33 (1965); R. Fletcher and M. J. D. Powell, *ibid.*, **6**, 163 (1963).
(54) J. W. McIver, Jr., and A. Komornicki, *Chem. Phys. Lett.*, **10**, 303 (1971).
(55) M. J. S. Dewar, H. W. Kollmar, D. H. Lo, H. Metiu, P. J. Student, and P. K. Weiner, to be published.
(56) These calculations are very fast since no additional SCF calculations are required, the bond order matrix for a closed shell molecule being invariant in first approximation to changes in the molecular geometry. In estimating a derivative by finite difference, the energies can thus be found using the same bond order matrix for both points.
(57) See A. Brown, M. J. S. Dewar, and W. W. Schoeller, *J. Amer. Chem. Soc.*, **92**, 5516 (1970).
(58) Simplex requires *ca.* N^2 function evaluations to optimize N parameters whereas the two gradient methods need only 30–40 for molecules with $N < 60$. The number of function evaluations also increases only slowly with increasing molecular size.
(59) M. J. S. Dewar and R. C. Haddon, *Chem. Commun.*, in press.
(60) P. K. Weiner, unpublished work.
(61) R. C. Bingham and M. J. S. Dewar, *J. Amer. Chem. Soc.*, **94**, 9107 (1972).
(62) This procedure was first developed[63] for treating radicals in the π approximation and later extended[64] to triplet states in the same connection. The further extensions to all-valence-electron calculations are self-evident.
(63) M. J. S. Dewar, J. A. Hashmall, and C. G. Venier, *J. Amer. Chem. Soc.*, **90**, 1953 (1968).
(64) M. J. S. Dewar and N. Trinajstić, *Chem. Commun.*, 646 (1970); *J. Chem. Soc. A*, 1220 (1971).
(65) See the following papers of this series.
(66) M. J. S. Dewar, S. H. Suck, and P. K. Weiner, *Chem. Commun.*, in press.
(67) M. J. S. Dewar and D. H. Lo, to be published.
(68) M. J. S. Dewar, S. H. Suck, and P. K. Weiner, to be published.
(69) M. J. S. Dewar, S. H. Suck, and P. K. Weiner, to be published.
(70) P. Bischoff, M. J. S. Dewar, and P. J. Student, to be published.
(71) P. Bischoff and M. J. S. Dewar, to be published.
(72) *I.e.*, molecules in which pairs of adjacent atoms have hybridized lone pairs; see ref 32.
(73) "Meaningful" in the sense that geometries were optimized without any simplifying assumptions and an adequate basis set used.