(d, 1 H, J_{AB} = 12 Hz), 4.21 (dd, 1 H, J_1 = 9, J_2 = 4 Hz), 4.03 (d, 1 H, J = 8 Hz), 3.87 (dd, 1 H, J_1 = 12, J_2 = 4 Hz), 3.65 (s, 3 H), 3.59 (m, 1 H), 3.49 (m, 1 H), 3.42 (m, 1 H), 3.23 (s, 3 H), 3.22 (s, 3 H), 3.08 (br s, 1 H), 2.87 (br s, 1 H), 1.87 (s, 3 H).

Supplementary Material Available: An ORTEP drawing of compound **32** as well as the procedures related to the solution of the crystal structure, thermal parameters (Table I), bond distances (Table II), and bond angles (Table III). Experimental protocols for the preparative work, as well as modifications in the synthesis of phenol **7**, are also provided. In addition, experimental procedures for the conversion of purified ether **8** to **15a** through fully characterized intermediates **12a**, **13a**, and **14a** are included, as well as the spectral properties of the crude intermediates **12b**, **13b**, **14b**, **15b**, and selenide **21**. Also included are the preliminary results of transformations of tricyclic amine oxides (17 pages). Ordering information given on any current masthead page.

Comparative Tests of Theoretical Procedures for Studying Chemical Reactions[1]

Michael J. S. Dewar* and Donn M. Storch[2]

Contribution from the Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712. Received October 29, 1984

Abstract: A simple procedure is described for estimating the effective errors in molecular energies calculated by ab initio methods with respect to use of the latter in studies of chemical reactions. The procedure is illustrated by application to the STO-3G, 3-21G, and 6-31G* models. Parallel results from semiempirical models (MINDO/3, MNDO, AM1) are included for comparison.

Introduction

The most basic problem in chemistry is to find out how chemical reactions take place. It is also, unfortunately, a problem that cannot be solved by experiment, because the time a chemical reaction takes is too short ($<10^{-13} \sim$ s) for its course to be observed.[3] Current "experimental" approaches rely on theory to delineate possible mechanisms for a reaction. Experiments are then devised to distinguish between the various possibilities. Such an approach is naturally limited by the efficacy of the theory on which it is based, and until recently only qualitative theories were used in this connection. Studies of reaction mechanisms would clearly be much more effective if they were based on a theoretical approach able to reproduce the properties of molecules, in particular their energies, quantitatively. Reaction mechanisms could indeed be predicted unambiguously if the corresponding potential energy surfaces could be calculated accurately.

To be chemically useful, such a calculation must be carried out properly, i.e., with full geometry optimization, etc., and without making any assumptions, and the method used must also be sufficiently accurate. Unfortunately, no current ab initio procedure comes anywhere near to achieving the needed accuracy, in an a priori sense. The errors in molecular energies, calculated by standard Hartree–Fock-type ab initio models, are indeed comparable with the corresponding heats of atomization.[4] Since this point is still not generally appreciated, some additional examples are shown in Table I. Indeed, since the calculated values refer to molecules at equilibrium geometries, without zero point or thermal energy, the real errors are greater than indicated, by ca. 5%. Some improvement is possible in "beyond HF (Hartree–Fock)" methods, but these are confined to small molecules and the residual errors are still enormous, by chemical standards.

Table I. Errors[a] in Total Energies (E) and Heats of Atomization (HA), Calculated by the RH Method Using the 6-31G* Basis Set[a]

molecule	error in E	error in HA	HA (obsd)
acetylene	324	106	392
propane	578	189	955
cyclopropane	494	176	814
cyclopropene	554	166	656
diacetylene	738	190	676
1,3-butadiene	682	221	972
hydrazine	444	162	412
hydrazoic acid	699	205	320
acetonitrile	523	51	478
cyanogen	799	207	478
methanol	515	130	487
dimethyl ether	704	94	671
ozone	631	68	145
acetone	896	225	938
dimethylamine	554	203	825
benzene	1301	649	1320

[a]All values in kcal/mol. The ab initio energies were taken from ref 14. These refer to calculations carried out with full geometry optimization, using the 6-31G* basis set. Experimental total energies of atoms, relative to nuclei and electrons, were estimated from ionization energies (Weast, R. C. "CRC Handbook of Chemistry and Physics", 65th Ed.; CRC Press: Boca Raton, FL, 1984–5; pp E-63,64. Energies (eV): H, 13.598; C, 1030.080; N, 1486.029; O, 2043.794. For conversion factors, see Table II (footnote). Experimental heats of atomization (standard state, gas phase, 25 °C) were estimated from thermochemical data listed by Cox and Pilcher. [Cox, J. D.; Pilcher, G. "Thermochemistry of Organic and Organometallic Compounds"; Academic Press: New York, 1970]. The experimental total energy of a molecule refers to the sum of its heat of formation and the total energies of the component atoms. It therefore includes kinetic energy terms. Since the calculated values for the molecular energies refer to equilibrium geometries, without corrections for zero point or thermal energy, the errors listed above are correspondingly too small; see text.

The errors are due primarily to neglect of electron correlation. If the correlation energy of a set of atoms did not change significantly when they combine to form a molecule, the errors might then cancel in calculating the differences in energy (heats of

(1) Part 75 of a series of papers reporting the development and use of quantum molecular models. For part 74, see; Dewar, M. J. S.; Grady, G. L.; Merz, K. M., Jr.; Stewart, J. J. P. *J. Am. Chem. Soc.*, in press.

(2) Present address: USAFA/DFC, Colorado Springs, Colo. 80840.

(3) Our inability to observe reactions is due not merely to lack of techniques but to the limits set by the Uncertainty Principle.

(4) Dewar, M. J. S.; Ford, G. P. *J. Am. Chem. Soc.* **1979**, *101*, 5558.
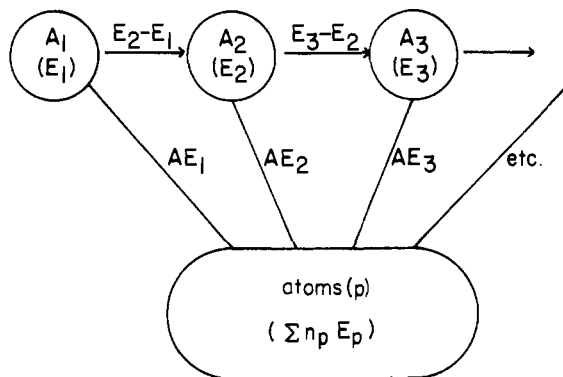
Comparative Tests for Studying Chemical Reactions

J. Am. Chem. Soc., Vol. 107, No. 13, 1985 3899

Figure 1. Heats of reaction for interconversion of species A_i formed from the same set of atoms p; AE_i is the heat of atomization of A_i.

formation, reaction, activation) involved in chemistry. This, however, is not the case.[4,5] Whatever their origin, the errors do *not* cancel. The errors in calculated heats of atomization (Table I), while less than those in total energies, are still very large. Furthermore, the real errors are again greater than those listed in Table I, this time by ca. 10%, because of the neglect of thermal energy. Use of ab initio procedures in chemistry is therefore justifiable only in areas where empirical tests show that the errors *do* cancel.

However, as noted above, the way a chemical reaction takes place cannot be observed.[3] Available experimental data for potential surfaces refer only to sections of them near minima, and to certain properties of saddle points (i.e., transition states). The ability of a given procedure to reproduce potential surfaces cannot therefore be adequately tested. *Any* use of *any* current theoretical method to study a chemical reaction therefore involves extrapolation of an empirical method into an area where it has not been tested. This is true of *all* current procedures, ab initio and semiempirical alike. All we can do is check their efficacy in all contexts where tests are possible. Each ab initio model must moreover be tested separately, because the errors vary unpredictably from one model to another. In the case of models based on the Roothaan[6]–Hall[7] (RH) SCF MO approximation, each basis set must be tested individually, while a "beyond HF" model, where allowance is made for electron correlation, must be tested separately for each basis set to which it is applied.

While these points should be self-evident, they have been generally ignored. While many different basis sets are in current use, none has been tested at all systematically, other than those introduced by Pople's group. Tests of the latter have moreover been confined to comparisons of heats of reaction. Such comparisons are of limited value because results for one reaction provide no indication of the errors to be expected in others. The quantities needed are estimates of the effective errors for individual molecules, effective in the sense that the error in any calculated heat of reaction is equal to the difference between the sums of errors for the reactants and for the products. Our purpose here is to present a simple procedure for obtaining estimates for such errors from calculated total energies. This can be applied to *any* theoretical treatment for which adequate data exist, in the form of geometry-optimized total energies for a sufficient variety of molecules.

Procedure

Consider a set of molecules, or collection of molecules (A_i), formed from the same set (P) of atoms (p). Let the energy calculated by a given procedure for A_i be E_i, and that for p be ϵ_p; see Figure 1. The calculated energy of atomization (AE_i) of A_i is then given by:

$$AE_i = E_i - \sum n_p \epsilon_p \qquad (1)$$

where n_p is the number of atoms p in A_i. If the procedure is to give accurate predictions of chemical behavior, it must predict the relative energies of the A_i correctly.[5] The real difference in energy between A_i and A_j must then be equal to the difference between their calculated energies of atomization. This can be so only if the errors in the latter are all the same, i.e., if the real energy of atomization (AE^0) of A_i is given by:

$$AE_i^0 = AE_i + X_A \qquad (2)$$

where the error, X_A, is the same for all the A_i. If this is to be true for all possible sets of molecules, the error (X_A) must be an additive sum of errors (x_p) in the energies of the individual atoms, i.e.,

$$X_A = \sum_p n_p x_p \qquad (3)$$

It must then be possible to assign an energy (ϵ_p') to each atom (p) which will allow the real energy of atomization (EA_p^0) of A_i to be estimated from its calculated energy, E_i; i.e.,

$$EA_i^0 = E_i - \sum n_p \epsilon_p' \qquad (4)$$

where

$$\epsilon_p' = \epsilon_p + x_p \qquad (5)$$

In most applications of ab initio methods to reactions, it is assumed[5] that the differences in energy between the A_i can be equated to the corresponding differences in enthalpy, the heat of reaction ($\Delta H_{i \to j}$) for $A_i \to A_j$ being set equal to the differences ($E_i - E_j$) between the energies calculated for A_i and A_j. This can be true if, and only if, the kinetic (thermal + zero point) energies of the A_i are additive functions of the atoms in them. If so, the corresponding atomic contributions can also be absorbed into the atomic energies. There must then be a set of values (e_p) for the energies of gaseous atoms that will allow the (real) heat of atomization (HA_i) of *any* molecule (A_i) to be calculated from its calculated total energy (E_i) by the following modification of eq 4:

$$HA_i = E_i - \sum_p n_p e_p \qquad (6)$$

The heat of formation of each gaseous atom, from the corresponding element in its standard state, can moreover be absorbed into the corresponding atomic energy (e_p), use of the resulting values (e_p') in eq 4 then leading directly to standard heats of formation (Δ_i):

$$\Delta H_i = E_i - \sum_p n_p e_p' \qquad (7)$$

In practice, the values of ΔH_i given by eq 7 will, of course, differ from experiment through failure of the assumptions on which it was based. The corresponding errors will then indicate the effective errors in the energies of individual molecules, representing their contributions to the errors in the heats of reaction calculated for reactions involving them. Appropriate values for the e_p' can be found by a least-squares fit of eq 5 to the experimental heats of formation of an appropriately comprehensive selection of molecules.

Estimation of Heats of Formation for Stable Species

It should be emphasized that our object is not to try to obtain the best possible estimates of heats of formation of molecules from their ab initio total energies. This can be done better by applying empirical corrections to the calculated values. Two procedures of this kind have been described, each of which is superior to the one set out above.

(5) Nearly all current studies of reactions are based on the assumption that calculated differences in energy between the reactants and products (transition state) in a reaction can be equated to observed heats of reaction (activation). Since molecular energies calculated by ab initio methods refer to systems at rest, with no vibrational energy, this amounts to the assumption that the thermal energies of molecules are an additive function of the atoms in them, which is certainly not the case. Similar problems also arise in the semiempirical procedures developed here (MINDO/3, MNDO, AM1; see below) because while these contain an implicit allowance for thermal energy, having been parametrized to reproduce observed heats of formation, the allowance is only an average over the set of molecules used in the parametrization.

(6) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.

(7) Hall, G. G. *Proc. R. Soc. London, Ser. A* **1951**, *205*, 541.

Table II. Values (au) of e_p and e_p' in Eq 4 and 5[a]

	basis set		
atom	STO-3G	3-21G	6-31G*
H	−0.49126	−0.48605	−0.48609
	(−0.57429)	(−0.56908)	(−0.56912)
C	−37.13685	−37.40049	−37.61642
	(−37.40983)	(−37.67347)	(−37.88940)
N	−52.99316	−53.39569	−53.71288
	(−53.74645)	(−54.14898)	(−54.46617)
O	−73.37643	−73.96599	−74.38783
	(−73.77352)	(−74.36308)	(−74.78492)

[a] 1 au = 2 rydbergs = 27.2166 eV = 2625.925 kJ/mol = 627.6112 kcal/mol. 1 eV = 96.5 kJ/mol = 23.064 kcal/mol.

Pople et al.[9] related the heats of formation of molecules to the heats of formation of simple hydrides of the type XH_n or H_mXYH_n by appropriate isodesmic reactions. The heats of formation of the hydrides were treated as parameters, being set equal to the experimental values. This approach provides an empirical correction for the effects of the changes in correlation energy that accompany the formation of specific kinds of chemical bonds.

A similar but more elaborate scheme has recently been introduced by Wiberg[10,11] in which corrections are applied for the individual groups present in a molecule, on the analogy of empirical schemes for estimating heats of formation as sums of contributions by individual groups. The results for a number of molecules, using the 6-31G* basis set, are impressive.

While this kind of approach may be of practical value for estimating heats of formation, as have analogous ones for the calculation of molecular geometries[12] and ionization energies,[13] it cannot be used to study reactions because it is restricted to molecules built entirely from standard groups. It cannot be applied to the intermediate sections of potential surfaces, where partial (delocalized) bonds are present, or indeed to other "nonclassical" species, because of the lack of reference compounds to provide the necessary empirical corrections. For studies of reactions we need an analogous procedure in which the corrections depend only on the atoms present in a molecule, not on the bonds they form. The procedure presented here meets these conditions, the reference species being indeed the individual atoms. Considered as a way to estimate heats of formation, it is naturally inferior to the methods devised by Pople and Wiberg because it involves far fewer parameters, one per type of atom rather than one per type of bond or per group of atoms.

Results and Discussion

Our purpose here is not to provide a compendium of tests of ab initio methods because this is not our responsibility. Indeed, we could not do so because of the lack of data for any procedures other than the ones introduced by Pople's group. Those who wish to use other alternatives should first test them, or at least provide the information needed to test them, in the form of geometry-optimized calculations for an adequate range of molecules. The present situation, where papers reporting calculations by essentially untested empirical methods continue to appear in print, is, to say the least, unsatisfactory.

Here we report applications of the approach indicated above to three of the RH procedures introduced by Pople et al., i.e., those

using the STO-3G, 3-21G, and 6-31G* basis sets. These cover a wide range, from a minimum basis set (STO-3G) to a split basis set with polarization functions (6-31G*). These have been extensively used and the results of calculations for quite a wide range of molecules are listed in a compilation[14] by Pople et al. Using them, we determined values of e_p and e_p' for carbon, hydrogen, nitrogen, and oxygen for each of the basis sets indicated above. These are listed in Table II, in atomic units (hartrees), for conformity with published values for the calculated molecular energies. Conversion factors to other units are shown in a footnote to the table.

Table III shows the corresponding errors (in kcal/mol) in the heats of formation calculated for 45 molecules. Analogous values for our semiempirical procedures are included for comparison, i.e., MINDO/3,[15] MNDO,[16] and AM1.[17] Tables IV and V show additional comparisons of 6-31G*, MNDO, and AM1 for some cations and for some larger hydrocarbons, 6-31G* energies for the latter being those reported by Wiberg.[10]

The results in Table III show that the errors for STO-3G are unacceptably large, as many have suspected; cf. ref 4. Since 3-21G calculations require little more computing time and lead to much better results, the STO-3G model should be regarded as obsolete.

The results in Tables III–V suggest that MNDO and AM1 are comparable with the 6-31G model and far superior to STO-3G. Note the surprisingly good performance of MNDO and AM1 for cations, surprising because the parameters in both treatments were derived entirely from properties of neutral molecules. The 6-31G* values are subject to large systematic errors.

Comparisons of this kind cannot be taken too literally, because the errors given by each individual treatment depend individually on the particular set of molecules used in making the comparisons. However, since the errors are fairly uniform over different kinds of molecules, it seems unlikely that use of an alternative "basis set" of molecules would lead to significantly different results. While each procedure has its "good" and "bad" molecules, these do not as a rule follow any visible pattern. The only exception is MINDO/3, which gives energies that are far too negative for molecules where adjacent atoms have pairs of unshared electrons (e.g., H_2O_2 or N_2H_4), an error due to the neglect of one-center overlap in the INDO approximation. If molecules of this kind are omitted, the error for MINDO/3 in Table III becomes similar to those for MNDO or 6-31G*.

The comparisons in Table III indicate that in case of a disagreement between an ab initio procedure and one of ours, it can by no means be assumed that it is the latter that is wrong. This is true even for calculations using the 6-31G* model. In the past, the lack of adequate information concerning the performance of ab initio methods has engendered a false impression of their accuracy and reliability. The results reported here should serve to set the record straight.

The time required for a single SCF calculation, using the 3-21G model, is two orders of magnitude greater than that using MINDO/3, MNDO, or AM1, all of which are similar in this connection, while the time required by 6-31G* is two orders of magnitude greater again. Our procedures are therefore clearly much more cost effective. It should also be noted that our approach could undoubtedly be greatly improved, at relatively little cost in extra computing time, by applying it to a better basic approximation than NDDO. We have refrained from doing so because *any* increase in computing time would be unacceptable, so far as we ourselves are concerned. Even our present procedures need too much computing time for many applications in which we are ourselves interested, using the computer (DEC VAX 11-780) currently available to us. It should be noted that while

(8) For example, few organic chemists are aware that the EH method predicts neither hydrogen nor benzene to exist, the former being predicted to collapse to He^2 and the latter to decompose exothermically and without activation into three molecules of acetylene. For other examples, see ref 4.

(9) (a) Hehre, W. J.; Ditchfield, R.; Radom, L.; Pople, J. A. *J. Am. Chem. Soc.* **1970**, *92*, 4796. (b) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(10) Wiberg, K. B. *J. Comput. Chem.* **1984**, *5*, 197.

(11) Numerous additional examples are given in a forthcoming paper by Schleyer, P. v. R., personal communication.

(12) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550.

(13) Rabalais, J. W. "Principles of Ultraviolet Photoelectron Spectroscopy"; Wiley-Interscience: New York, 1977.

(14) Whiteside, R. A.; Frisch, M. J.; DeFrees, D. J.; Schlegel, H. B.; Raghavachari, K.; Pople, J. A. "Carnegie-Mellon Quantum Chemistry Archive", 2nd ed.; Carnegie-Mellon University: Pittsburgh, 1981.

(15) Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J. Am. Chem. Soc.* **1975**, *97*, 1285, 1294, 1302, 1307.

(16) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899, 4907.

(17) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P.; *J. Am. Chem. Soc.*, following paper in this issue.

Table III. Errors in Heats of Formation[a] Calculated by Various Methods[b]

molecule	STO-3G	3.21G	6-31G*	MINDO/3	MNDO	AM1
CH_4	-5.4	-0.9	0.5	11.6	6.0	9.1
C_2H_6	5.4	0.2	1.9	0.4	0.5	2.8
C_3H_8	14.5	0.4	1.9	-1.7	-0.1	0.5
C_2H_4	-26.9	-1.6	-2.4	6.8	2.8	4.0
$CH_3CH=CH_2$	-4.5	1.9	-2.3	1.6	0.0	1.7
$CH_3CH=CHCH_3$ (cis)	4.3	n/a	3.5	-3.8	-2.1	-0.3
$CH_3CH=CHCH_3$ (trans)	4.9	-3.1	3.1	-3.3	-2.1	-0.3
$H_2C=C(CH_3)_2$	3.6	-3.9	2.9	6.8	2.2	3.1
$H_2C=C=CH_2$	-20.3	-2.6	-6.8	-3.9	-1.8	0.5
$H_2C=CHCH=CH_2$	-14.6	-4.7	12.6	5.8	2.9	3.9
C_2H_2	-16.0	-1.7	-8.0	3.5	3.0	0.5
$CH_3C\equiv CH$	-4.4	0.4	-6.0	-9.3	-3.5	-1.0
$H_3CC\equiv CCH_3$	6.9	0.4	4.9	-22.6	-10.3	-2.7
$HC\equiv C-C\equiv CH$	-7.6	1.4	11.2	21.9	-10.8	-6.9
cyclopropane	7.0	-8.4	-2.4	-4.0	-1.5	5.1
cyclopropene	-12.6	-18.4	-10.1	-6.8	2.0	6.6
cyclobutene	9.4	-11.2	-8.1(d)	-4.3	-6.5	8.3
benzene	11.4	-2.7	10.4	9.0	1.4	2.2
H_2	-19.5	-9.5	-7.2	0.1	0.7	-5.2
N_2	4.9	1.9	7.3	5.2	8.0	11.2
O_2 ($^1\Delta_g$)	24.8	-2.4	-1.1	0.8	-9.8	-21.3
O_3	14.8	-28.6	-24.4	5.0	14.3	3.6
H_2O	-30.3	-4.6	-2.8	4.2	-3.1	-1.4
H_2O_2	11.0	18.6	3.1	0.9	-5.7	-2.8
CO	0.0	9.2	13.5	12.9	20.2	20.7
CO_2	-24.3	7.4	15.7	-1.7	18.7	14.2
CH_3OH	-5.0	5.3	-5.0	-2.0	-9.3	-9.0
C_2H_5OH	2.9	5.4	-4.9	-7.9	-6.8	-6.5
CH_2O	-11.9	3.6	7.8	0.4	-7.0	-5.5
$HCOOH$	-20.0	11.3	-12.9	2.0	-2.1	-0.8
CH_3CHO	-4.8	3.5	7.8	-3.9	-2.7	-1.9
CH_3COCH_3	2.9	4.1	-7.4	-1.6[c]	2.4	2.7
CH_3OCH_3	15.6	11.6	-10.2	2.9	-7.2	-9.2
NH_3	-19.7	-1.0	-4.2	1.5	4.7	3.7
N_2H_4	-3.6	7.6	-1.9	-20.1	-8.5	-9.1
$HN=NH$ (cis)	-2.2	-7.9	-3.7	-44.3[c]	-17.1	-18.5
HN_3	n/a	-17.3	10.6	-84.2[c]	2.7	5.6
CH_3NH_2	-2.3	3.2	0.0	0.6	-2.0	-1.9
CH_3NHCH_3	11.9	5.2	1.9	3.0	-2.2	1.0
HCN	-2.4	8.8	1.2	3.1	2.6	-1.3
CH_3CN	9.9	13.8	3.3	-11.5	-2.0	-1.6
CH_3NC	0.6	8.0	n/a	-14.1[c]	24.4	14.8
$NC-CN$	15.4	22.4	1.5	-35.4[c]	-7.9	-5.9
HNO_2	11.2	-3.6	n/a	5.2	-21.8	-20.6
N_2O	-19.1	-15.2	-3.7	2.2[c]	11.3	8.9
average error:	10.7	6.9	6.1	9.0	6.3	5.8
no. of molecules:	44	44	43	45	45	45

[a] All values in kcal/mol. Heats of formation refer to the formation of molecules in their standard states in the gas phase, at 25 °C, from elements in their standard states. [b] Sources of data: ab initio, ref 14; MINDO/3, ref 15; MNDO, ref 16; AM1, ref 17.

Table IV. Errors in Heats of Formation Calculated for Some Cations by Various Procedures[a]

ion	error (kcal/mol)		
	MNDO	AM1	6-31G*
H_3C^+	-16	-8	-30
$(CH_3)_2CH^+$	9	0	-22
allyl cation	-5	0	-25
$H_2C=NH_2^+$	9	-2	-22
HCO^+	-14	-11	-25
NO^+	-6	-9	-23
H_3O^+	-5	5	-12
NH_4^+	10	-4	-22
mean absolute error	9.2	4.9	22.6

[a] For sources of data and definitions, see Table III.

computers are currently available that are several hundred times faster than the VAX (e.g., CRAY, CDC 205), computing time on them costs comparably more. A calculation taking 1000 h on a VAX is entirely feasible, in terms both of cost and computing time. One taking 1000 h on a CRAY or CDC 205 would be feasible in terms of computing time but not cost.

As we have repeatedly pointed out, the gap to be bridged before current ab initio procedures could be used in place of ours is very

Table V. Errors in Heats of Formation Calculated for Some Hydrocarbons by Various Procedures[a]

molecule	error (kcal/mol)		
	MNDO	AM1	6-31G*
butane	0.7	-0.7	-0.8
2-methylpropane	5.6	3.0	0.9
pentane	0.7	-2.8	-0.5
cyclopentane	-11.9	-10.4	4.0
cyclohexane	-5.3	-9.0	3.1
bicyclobutane	17.2	26.2	11.2
bicyclo[2.1.0]pentane	7.1	8.8	8.1
bicyclo[2.2.1]heptane	2.1	-2.0	8.8
bicyclo[2.2.2]octane	-2.2	-11.9	10.7
cubane	-49.6	2.5	23.6
1,4-pentadiene	0.6	-0.3	11.7
cyclopentene	-8.6	-5.2	6.1
cyclopentadiene	0.1	5.2	12.0
norbornadiene	2.9	8.1	22.0

[a] For definitions, see Table III. Sources of data: MNDO, ref 16; AM1, ref 17; 6-31G*, ref 8.

great. Its magnitude has been generally underestimated because no system has ever been thoroughly examined, using any ab initio method. Calculations have been confirmed to reactions where

the mechanisms were already known or limited to two alternatives, and calculations using the larger basis sets have been further limited to very small molecules. To be useful as a general chemical tool, it must be possible to study rather large systems in detail. This can require an enormous amount of computation.

A striking feature of the results in Table III is the relatively small difference between the errors given by the 3-21G and 6-31G* models and between the ones given by the three semiempirical procedures. The accuracies of all seem indeed to be limited by some common factor. Thermal energy seems the obvious candidate. As noted above, nearly all ab initio studies of reactions have been based on the assumption that the thermal energy of a molecule is an additive function of the atoms in it, so that a heat of reaction or activation can be equated to the corresponding difference in total energy between the reactants and the products or the transition state. The same assumption is made tacitly in our semiempirical methods, where allowance for thermal energy is included via the parametrization, so it applies generally to the results in Table III. Better results could undoubtedly be obtained by making specific allowance for the thermal energy, using partition functions constructed from calculated vibration frequencies, etc..

One last point of interest should be noted. By using eq 5 in reverse, ab initio energies of molecules can be estimated from their experimental heats of formation, with an average error of only ±0.01 au. This could be useful in the case of larger molecules where calculations by the better ab initio methods would be prohibitively expensive. Since these are believed to give energies reasonably close to the HF limit, an indication of the latter could be obtained simply, and at no cost, in this way. Such information would provide a useful indication of the level of accuracy of ab initio procedures relative to Hartree–Fock.

AM1: A New General Purpose Quantum Mechanical Molecular Model[1]

Michael J. S. Dewar,* Eve G. Zoebisch, Eamonn F. Healy, and James J. P. Stewart

Contribution from the Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712. Received October 29, 1984

Abstract: A new parametric quantum mechanical molecular model, AM1 (Austin Model 1), based on the NDDO approximation, is described. In it the major weaknesses of MNDO, in particular failure to reproduce hydrogen bonds, have been overcome without any increase in computing time. Results for 167 molecules are reported. Parameters are currently available for C, H, O, and N.

Introduction

The purpose of the work reported in this series of papers[1] has been the development of a quantitative quantum mechanical molecular model for chemists to use as an aid to experiment in their own research, in particular in studies of chemical reactions and reaction mechanisms. To be useful in this connection, such a procedure must be not only sufficiently accurate but also applicable to the molecules in which chemists are directly interested rather than confined to simple models. These requirements eliminated, and still eliminate, ab initio procedures because such procedures are too inaccurate and/or require far too much computing time.[1] Our approach has accordingly been to use an approximation simple enough for the desired calculations to be feasible, using currently available computers, and to upgrade the accuracy of the results by introducing parameters that can be adjusted to fit the results to experiment. In this way we have been able to develop[2] two effective models, MINDO/3[3] and MNDO,[4] which are being widely used.[5] As the preceding paper[1] shows, the results from MINDO/3 and MNDO are generally comparable with those from ab initio methods that require at least 1000 times more computing time.

It should be emphasized that even MINDO/3 and MNDO are *too slow* for general use in chemistry, using currently available computers. Calculations of reaction mechanisms, using standard computers such as the DEC VAX 11-780, require excessive amounts of computer time for systems containing more than a dozen "heavy" atoms (i.e., other than hydrogen). While much larger systems can be treated using "state-of-the-art" computers, such as the CDC 205 or CRAY, this does not reduce the cost of the calculations, because while these are several hundred times faster than a VAX, the cost of computing time is also greater by an almost equally large factor. A 100-fold increase in the speed of computers, with no increase in the cost of computing time, will be needed to enable our procedures to achieve their full potential, particularly in projected applications to biochemistry and organometallic chemistry.

A major problem in studying reactions by *any* current theoretical model is the lack of experimental data for the intermediate sections of potential surfaces and for the geometries of transition states. Calculations for these consequently involve the extrapolation of an empirical[6] procedure into areas where it has not been, and indeed cannot be, tested. Such an extrapolation is safer, the better the performance of the method in question in all areas where it can be tested. Confidence in a semiempirical procedure is moreover strengthened by demonstrations of its ability to reproduce experimental results unrelated to those used in determining the parameters in it. One of the major assets of MINDO/3 and

(1) Part 76 of a series of papers reporting the development and use of quantum mechanical molecular models. For part 75, see: Dewar, M. J. S.; Storch, D. M. *J. Am. Chem. Soc.*, preceding paper in this issue.

(2) Dewar, M. J. S. *J. Mol. Struct.* **1983**, *100*, 41.

(3) Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J. Am. Chem. Soc.* **1975**, *97*, 1285, 1294, 1302, 1307.

(4) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899, 4907.

(5) A total of 623 papers reporting MNDO calculations have been listed in *Chemical Abstracts* since 1980.

(6) The errors in energies calculated even by "state-of-the-art" ab initio methods are enormous by chemical standards, far too large for any conclusions to be drawn a priori from the results; see ref 1.