JOURNAL OF PHYSICAL ORGANIC CHEMISTRY, VOL. 5, 225–229 (1992)

HEATS OF FORMATION OF ORGANIC MOLECULES BY *AB INITIO* CALCULATIONS. ALDEHYDES AND KETONES

LAWRENCE R. SCHMITZ

Department of Chemistry, Marshall University, Huntington, West Virginia 25755, USA

I. MOTOC*

Bristol-Myers Squibb Corporation, Pharmaceutical Research Institute, Core Chemistry Resources, Princeton, New Jersey 08543-4000, USA

CHARLES BENDER AND JAN K. LABANOWSKI

Ohio Supercomputer Center, 1224 Kinnear Road, Columbus, Ohio 43212-1163, USA

AND

NORMAN L. ALLINGER†

School of Chemical Sciences, Department of Chemistry, University of Georgia, Athens, Georgia 30602, USA

A bond and group equivalent scheme that allows the calculation of heats of formation for aldehydes and ketones from *ab initio* 6–31G* energies has been developed. For a group of 24 aldehydes and ketones, the RMS error for the calculated heat of formation was 0·46 kcal mol^{-1}. Heats of formation have been predicted for an additional seven compounds for which the experimental values are believed to be either in error or unknown. There are some problems with the norbornanones.

INTRODUCTION

In previous papers in this series,[1,2] we defined a scheme for the calculation of heats of formation of organic molecules from *ab initio* energies obtained using the Hartree–Fock approximation and the 6–31G* basis set. The pertinent equations were rigorously derived using statistical mechanics. Like similar schemes developed by Wiberg[3] and Ibrahim and Schleyer,[4] the heat of formation is calculated by adding a series of empirically determined equivalents to the Hartree–Fock energy (*HFE*). In our scheme, these equivalents are of two types, bond energy increments (*BE*) and group increments (*GE*). The various bond energy increments correspond to different types of bonds in the molecule and generally have a substantial magnitude and contain the majority of the energy terms needed to convert the HFE to a heat of formation. Group increments correspond to structural units of the molecule that are larger than individual bonds, or to adjustments to the basic bond energy increments. They usually have smaller magnitudes and allow the 'fine tuning' of the calculation based on structure that is needed to give chemical accuracy. They also include corrections for the lack of explicit inclusion of electron correlation. However, this scheme differs from those developed by Wiberg[3] and Ibrahim and Schleyer,[4] since it also explicitly includes terms to account for the statistical mechanical effects of populating higher energy conformers of the molecule (*POP*), low-lying vibrational states (*TOR*), translational and rotational states and a term for the conversion from ΔE to ΔH. Thus the heat of formation is calculated as

$$\Delta H_f^\circ = \Sigma BE + \Sigma GE + HFE + POP + TOR + 4RT$$

where $4RT$ is the classical value for the translational and rotational energy, and the conversion from ΔE to ΔH.

Since the number of experimental determinations of heats of formation has been declining in recent years, we are endeavoring to develop this scheme so that heats of formation may be calculated for molecules containing most of the common functional groups. In

* Deceased.

† Author for correspondence.

0894–3230/92/050225–05$05.00

Received 2 October 1991
Revised 20 December 1991

previous papers, we have shown that it is possible to calculate heats of formation with RMS errors that are competitive with experimental determinations (< 1 kcal mol^{-1}) (1 kcal = 4·184 kJ) for alkanes,[1] amines[1] and simple alcohols[2] and ethers.[2] The results were less good for compounds containing two single-bonded oxygens bound to the same carbon,[2] but the experimental data here also appear to be less reliable. This paper reports our results for aldehydes and ketones.

COMPUTATIONAL METHODS

The *ab initio* calculations described here were carried out using the 6–31G* basis set[5] and either the CADPAC,[6] HONDO8[7] or Gaussian-90[8] program. Starting geometries for the *ab initio* calculations were obtained from MM3[9] using a modified parameter set designed to give r_e rather than r_g geometries (the r_e value of a bond length corresponds to the point in the potential well of minimum energy. The r_g value is the vibrationally averaged distance between the atoms. The latter is usually longer because of anharmonicity, by as much as 0·03 Å in the case of C—H bonds. With this kind of a starting geometry, the *ab initio* geometries were always optimized in less than ten iterations). These geometries were then optimized at the 6–31G* level until the largest component of the gradient was 5×10^{-4} hartree bohr^{-1}.

The *POP* terms were estimated by the usual procedure[2b] using relative energies of the conformations that were calculated by MM2 or MM3 or estimated. The *TOR* terms were estimated using the assumption that this term arises from the population of low-energy internal rotational modes. Therefore, in line with our practice with MM3, for each bond around which there is relatively free rotation (barrier less than 7 kcal mol^{-1} and excluding methyl groups) a contribution of 0·00067 hartree to *TOR* was assigned.[1,2]

The required experimental heats of formation were taken from the compilations of Pedley *et al.*[10a] or Cox and Pilcher.[10b] All heats of formation in this paper are for compounds in the gas phase at 298 K.

RESULTS AND DISCUSSION

We have chosen to fit carbonyl compounds by using three bond energy increments and five group increments. The bond increments consist of the carbon—oxygen double bond (C=O), the bond between the carbonyl and another (saturated) carbon (OC—C) and the bond between the carbonyl and a hydrogen (OC—H). These three variables are not linearly independent. Therefore, we fixed the value of the C=O increment as 93·74056 hartree. This allowed us to determine values for the other two that give the best possible fit to the data.

Table 1. Input data for aldehydes and ketones[a]

Compound	*HFE*(6–31G*)	*POP*	*TOR*	$T/R(=4RT)$
Formaldehyde	−113·86633	0·00000	0·00000	0·00382
Acetaldehyde	−152·91597	0·00000	0·00000	0·00328
Propanal	−191·95227	0·00008	0·00067	0·00382
Butanal	−230·98694	0·00083	0·00134	0·00382
Pentanal	−270·02168	0·00139	0·00201	0·00382
Hexanal	−309·05642	0·00194	0·00268	0·00382
2-Methylpropanal	−230·98695	0·00027	0·00067	0·00382
Acetone	−191·96224	0·00000	0·00000	0·00382
Butan-2-one	−230·99805	0·00000	0·00067	0·00382
Pentan-2-one	−270·03267	0·00043	0·00134	0·00382
Pentan-3-one	−270·03365	0·00000	0·00134	0·00382
Hexan-2-one	−309·06739	0·00099	0·00201	0·00382
Hexan-3-one	−309·06823	0·00043	0·00201	0·00382
3-Methylbutan-2-one	−270·03089	0·00010	0·00067	0·00382
3,3-Dimethylbutan-2-one	−309·06339	0·00000	0·00067	0·00382
3-Methylpentan-2-one	−309·06339	0·00043	0·00134	0·00382
3,3-Dimethylpentan-2-one	−348·09528	0·00043	0·00134	0·00382
4-Methylpentan-2-one	−309·06689	0·00000	0·00134	0·00382
4,4-Dimethylpentan-2-one	−348·09870	0·00000	0·00134	0·00382
2-Methylpentan-3-one	−309·06607	0·00000	0·00134	0·00382
2,2-Dimethylpentan-3-one	−348·09835	0·00000	0·00134	0·00382
2,4-Dimethylpentan-3-one	−348·10038	0·00018	0·00134	0·00382
Cyclopentanone	−268·86641	0·00000	0·00067	0·00382
Cyclohexanone	−307·90591	0·00000	0·00067	0·00382

[a] Energies in hartrees.

Three of the group increments, OC—Me, OC—*sec* and OC—*tert*, can be thought of as adjustments to the OC—C bond increment. The exact value of increment for a carbonyl—carbon bond is actually slightly dependent on the type of carbon (methyl, primary, secondary or tertiary) that the carbonyl is bonded to. Wiberg[11] has shown that the well known 'branched-chain' stabilization is mainly (or completely) due to electron correlation, and so is not found at the Hartree–Fock level; hence the need for these empirical increments here. To model this behavior, we have chosen the OC—C increment to fit attachment of the carbonyl to a primary carbon, and we use these three group increments as small adjustments to this value in the other situations. The OC—Me increment is a special case. Even though there will always be relatively free rotation around the carbonyl—methyl bond, we have chosen not to count a unit of *TOR* for this bond. Rather, we allow the energy associated with this rotation to be included in the increment.

The two remaining increments, OC—*iso* and OC—*neo*, account for branching at the carbons β to the carbonyl. There are only a limited number of compounds containing these structural features in our data set. However, they are fitted much better by including these increments. This is analogous to what we found in parameterizing MM3 for heats of formation of carbonyl compounds.[9]

The values of the increments were determined by a least-squares fitting to the experimental data for 24 carbonyl compounds. These data are shown in Table 1 and the resulting increments are shown in Table 2. In Table 3 are compared the calculated heats of formation and the experimental values for this set of compounds. The standard deviation for this set of compounds is

Table 2. Bond and group increments for aldehydes and ketones[a]

Type	Bond or group	Value
Bond increment	O=C	93·740560
	OC—H	10·040517
	OC—C	18·944324
Group increment	OC—Me	0·000907
	OC—*sec*	−0·005632
	OC—*tert*	−0·007780
	OC—*neo*	−0·003678
	OC—*iso*	−0·002338

[a]Energies in hartrees.

Table 3. Heats of formation for aldehydes and ketones[a]

Compound	ΔH_f° (calc.)	ΔH_f° (exp.)	Difference (calc. − exp.)
Formaldehyde	−25·67	−25·92	0·25
Acetaldehyde	−39·98	−39·73	−0·25
Propanal	−45·06	−45·45	0·39
Butanal	−49·65	−48·94	−0·71
Pentanal	−54·41	−54·45	0·04
Hexanal	−59·16	−59·37	0·21
2-Methylpropanal	−52·44	−52·25	−0·19
Acetone	−52·17	−51·90	−0·26
Butan-2-one	−56·99	−57·02	0·03
Pentan-2-one	−61·75	−61·92	0·18
Pentan-3-one	−61·69	−61·65	−0·03
Hexan-2-one	−66·49	−66·70	0·21
Hexan-3-one	−66·42	−66·51	0·09
3-Methylbutan-2-one	−63·27	−62·76	−0·51
3,3-Dimethylbutan-2-one	−69·72	−69·47	−0·24
3-Methylpentan-2-one	−66·77	−67·90	1·14
3,3-Dimethylpentan-2-one	−73·34	−72·60	−0·73
4-Methylpentan-2-one	−69·60	−69·60	0·00
4,4-Dimethylpentan-2-one	−76·60	−76·60	0·00
2-Methylpentan-3-one	−67·77	−68·38	0·61
2,2-Dimethylpentan-3-one	−74·01	−74·99	0·98
2,4-Dimethylpentan-3-one	−74·92	−74·40	−0·52
Cyclopentanone	−46·11	−46·03	−0·08
Cyclohexanone	−54·63	−54·04	−0·58
Standard deviation			0·46

[a]Energies in kcal mol^{-1}, gas phase.

0·00074 hartree (0·46 kcal mol^{-1}). The largest error in this set of compounds is 0·00181 hartree (1·14 kcal mol^{-1}) for 3-methylpentan-2-one. These errors are of approximately the same magnitude as the experimental errors in the data that they reproduce. Hence we cannot expect to obtain results much better than these.

In addition to the 24 compounds used to evaluate the equivalents, we calculated heats of formation for an additional nine compounds, and the necessary input data are given in Table 4. Three of these compounds, cycloheptanone, norbornan-2-one and norbonan-7-one, have experimentally determined heats of formation. The results are given in Table 5. For these compounds there are substantial deviations between the calculated and experimental heats. MM3 also fails to calculate heats of formation for these compounds that are in agreement with the experimental values.[9] Therefore, we have not included these compounds in the set used to determine the values of the increments.

In the case of cycloheptanone, the present method and MM3 calculate heats that are in good agreement with each other but not with experiment: −55·81, −55·28 and −59·10 kcal mol^{-1}, respectively. This is an ordinary compound for which the MM3 value should be reliable. We therefore believe this conflict with experiment is due to experimental error.

For norbornan-2-one and norbornan-7-one, the *ab initio*-calculated heats are too negative by 1·30 and 3·73 kcal mol^{-1}, respectively. MM3 calculates these values to be more positive than experiment by +3·34 and +3·48 kcal mol^{-1}, respectively. It is known that Hartree–Fock calculations with the 6–31G* basis set give energies of compounds with eclipsed butane interactions that are too positive.[12] However, the present scheme calculates the heats of formation for norbornane well, −12·83 calculated vs −12·40 kcal mol^{-1} experimental. The 7-keto compound has a highly distorted carbonyl group, far beyond anything that was available for the parameterization of MM3. The MM3 value for this molecule is accordingly not well founded. However, it would seem that the 2-keto derivative should be an ordinary compound, and we would expect that MM3 would calculate the value reasonably well. The large discrepancy between the experimental, *ab initio* and MM3 values is therefore very disturbing, and we are unable to account for it. Accordingly, we can only conclude that the heats of formation for these compounds are uncertain.

Table 4. Data for aldehydes and ketone predictions[a]

Compound	*HFE*(6–31G*)	*POP*	*TOR*	*T*/*R*(=4*RT*)
Cycloheptanone	−346·93386	0·00013	0·00067	0·00382
Norbornan-2-one	−345·76152	0·00000	0·00000	0·00382
Norbornan-7-one	−345·75291	0·00000	0·00000	0·00382
2,2-Dimethylpropanal	−270·02184	0·00000	0·00067	0·00382
3,3-Dimethylbutanal	−309·05349	0·00003	0·00134	0·00382
Cyclohexane-1,4-dione	−381·59970	0·00000	0·00067	0·00382
2-Methylcyclohexanone	−346·94201	0·00000	0·00067	0·00382
Heptan-2-one	−348·10213	0·00194	0·00268	0·00382
Heptan-3-one	−348·10298	0·00138	0·00268	0·00382

[a] Energies in hartrees.

Table 5. Predictions for aldehydes and ketones[a]

Compound	ΔH_f° (calc.)	ΔH_f° (exp.)	Difference (calc. − exp.)	MM3	Difference (calc. − MM3)
Cycloheptanone	−55·81	−59·10	3·29	−55·30	−0·51
Norbornan-2-one	−41·50	−40·20	−1·30	−37·43	−4·07
Norbornan-7-one	−35·73	−32·00	−3·73	−28·46	−7·27
2,2-Dimethylpropanal	−60·49	—	—	−59·20	−1·29
3,3-Dimethylbutanal	−65·05	—	—	−64·77	−0·28
Cyclohexahe-1,4-dione	−76·61	—	—	−77·16	0·55
2-Methylcyclohexanone	−63·02	—	—	−62·76	−0·26
Heptan-2-one	−71·00	—	—	−71·70	+0·70
Heptan-3-one	−70·36	—	—	−71·36	+1·00

[a] Energies in kcal mol^{-1}.

Also presented in Table 5 are calculated heats of formation for six compounds for which experimental values could not be found in the literature. We carried out MM3 calculations of these quantities, and in Table 5 are also given the Hartree–Fock vs MM3 comparisons. The agreement is within the expected combined errors of the two methods.

CONCLUSIONS

A combined *ab initio* – empirical scheme for calculating the heats of formation of carbonyl compounds has been developed and shown to give results that in general are competitive with experimental determinations. In addition, heats of formation for six compounds have been predicted. Application of this technique has revealed that the experimental heat of formation for cycloheptanone is seriously in error.

We have now reported how this general method using *ab initio* calculations can be used to obtain heats of formation for a variety of compounds. So far studied have been hydrocarbons, alcohols and ethers, amines, sulfides and, in this work, aldehydes and ketones. In most cases the results of the calculations and of the experiments are in good agreement, in cases where the experiments are believable. The only exceptions to date are with acetals (which shown considerable scatter, and for which it is not clear just how good the experimental data are) and the norbornanones discussed above. For the latter compounds, there is not only a conflict between the experiment and the *ab initio* calculations, but also the MM3 calculations do not agree with either.

Our conclusion concerning the above classes of compounds is that the *ab initio* calculations are probably at least as reliable as the experimental measurements, but there are still some unresolved problems.

ACKNOWLEDGEMENTS

The authors are indebted to the National Institutes of Health, Grant 5 R24RR02165, and the National Science Foundation, Grant CHE 8918839, for financial support of this work, and to the Ohio Supercomputer Center for a grant of Cray-YMP computer time.

REFERENCES

1. N. L. Allinger, L. R. Schmitz, I. Motoc, C. Bender and J. K. Labanowski, *J. Phys. Org. Chem.* **3**, 732 (1990).
2. (a) N. L. Allinger, L. R. Schmitz, I. Motoc, C. Bender and J. K. Labanowski, *J. Am. Chem. Soc.* in press; (b) U. Burkert and N. L. Allinger, *Molecular Mechanics*. American Chemical Society, Washington, DC (1982).
3. (a) K. B. Wiberg, *J. Comput. Chem.* **5**, 197 (1984); (b) K. B. Wiberg, *J. Org. Chem.* **50**, 5285 (1985).
4. M. R. Ibrahim and P. v. R. Schleyer, *J. Comput. Chem.* **6**, 157 (1985).
5. (a) W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.* **52**, 2257 (1972); (b) P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta* **28**, 213 (1972).
6. R. D. Amos and J. E. Rice, *CADPAC4–Analytic Derivative Package*, Issue 4. Cambridge (1987).
7. M. Dupuis, P. Mougenot, J. D. Watts, G. J. B. Hurst and H. O. Villar, in *Modern Techniques in Computational Chemistry*, edited by E. Clementi, Chapt. 7. ESCOM, Leiden (1989).
8. M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. A. Robb, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol and J. A. Pople, Gaussian-90. Gaussian Inc., Pittsburgh, PA (1990).
9. N. L. Allinger, K. Chen, M. Rahman and A. Pathiaseril, *J. Am. Chem. Soc.* **113**, 4505 (1991).
10. (a) J. D. Pedley, R. D. Naylor, and S. P. Kirby, *Thermochemical Data of Organic Compounds*, 2nd ed. Chapman and Hall, London (1986); (b) J. D. Cox and G. Pilcher, *Thermochemistry of Organic and Organometallic Compounds*. Academic Press, London, New York (1970).
11. K. B. Wiberg, *J. Org. Chem.* **56**, 544 (1991).
12. N. L. Allinger, B. F. Yates, R. S. Grev and H. F. Schaefer, III, *J. Am. Chem. Soc.* **112**, 114 (1990).