Heats of Formation of Curved PAHs and C_{60}: Graph Theoretical vs MM3(92) Predictions

Gordon G. Cash

Risk Assessment Division (7403), U.S. Environmental Protection Agency, 401 M Street, S. W., Washington, D.C. 20460

Received: April 15, 1997; In Final Form: July 8, 1997[⊗]

Correlations of graph theoretical parameters with $\Delta H_f°$ of buckminsterfullerene and 28 proposed precursor molecules were investigated. The 29 molecules examined divided cleanly into groups of 19 and 10 structures, with $\Delta H_f°$ for the smaller group depending only on the number of carbon atoms ($r = 0.9730$). $\Delta H_f°$ for the larger group was successfully predicted with a two-parameter equation ($r = 0.9987$). One of the parameters had been used previously to predict $\Delta H_f°$ for all-hexagon polycyclic aromatic hydrocarbons (PAHs), but not for PAHs containing pentagons. The other parameter had not previously been used with PAHs.

Introduction

Pope et al.[1] proposed a 28-step mechanism for the formation of buckminsterfullerene, C_{60}, from fluoranthene, $C_{16}H_{10}$, in flames. The first 19 steps were additions of C_2 to give a bowl-shaped $C_{54}H_{10}$ isomer. The remaining 9 steps were rearrangements, losses of H_2, and additions of C_2, finally resulting in the familiar isomer of C_{60}. In a subsequent paper,[2] Pope and Howard presented thermochemical properties for all 29 structures calculated from MM3(92), MOPAC (several methods), and a group additivity method of their own devising.

The present study investigated a method based on graph theoretical calculations for predicting standard heats of formation, $\Delta H_f°$, for these 29 structures. The goal was to overcome a drawback of all group additivity methods, namely, that they are restricted to structures composed of groups for which group contributions to the subject parameter are known. Most graph theoretical parameters, on the other hand, are calculable from a (usually hydrogen-depleted) molecular graph of the structure with arbitrarily numbered vertices (atoms). Correlation of graph theoretical parameters with $\Delta H_f°$ of 152 polyhex PAHs[3] and with arrangements of contiguous pentagons in fullerenes has previously been demonstrated.[4] Contiguous pentagons are calculated to make major contributions to steric strain in fullerenes.[5] For comparison with graph theoretical predictions, the MM3(92) results were chosen in preference to those from MOPAC, primarily because MOPAC hugely overestimated $\Delta H_f°$ for C_{60}, while MM3(92) underestimated it but came much closer to experimental values.[6]

Methods

Graph Theoretical Parameters. Perhaps the most widely used mathematical object in chemical graph theory is the adjacency matrix of the hydrogen-depleted graph.[7] For a PAH containing n carbon atoms, this is an $n \times n$ matrix **A** with $\mathbf{A}_{i,j} = 1$ if atoms i and j are directly bonded, zero otherwise.

Numbering of the atoms is arbitrary. The distance matrix **D** is another $n \times n$ matrix in which $\mathbf{D}_{i,j}$ is the number of bonds in the shortest through-bond path from atom i to atom j. **D** is computable from **A** without further reference to the structure.[8−10] Specifically, $\mathbf{D}_{i,j}$ is equal to the smallest value of k for which $(\mathbf{A}^k)_{i,j} \neq 0$. The author has described a method for dealing with the overflow problem in computing $\mathbf{A}^k$ for large k.[9] Schultz et al. introduced two graph theoretical indices based on the sum $\mathbf{A} + \mathbf{D}$, namely, the square root of the logarithm of the determinant, $(\ln|\mathbf{A} + \mathbf{D}|)^{1/2}$, and the logarithm of the principal eigenvalue $\ln[\lambda_1(\mathbf{A} + \mathbf{D})]$.[10] The Kekulé structure count K has long been known to correlate with energy, usually in the form of $E \propto \ln K$. For PAHs, Klein and Liu,[11] following the lead of Kasteleyn,[12−15] showed that $K^2 = |\mathbf{S}|$, the determinant of the signed adjacency matrix. See ref 11 for details on the construction of this matrix. The author used the Kekulé structure count as part of the parameter $\ln[\mathrm{per}(\mathbf{A})]/\ln K$, the ratio of the logarithms of the permanent of the adjacency matrix and the Kekulé structure count, which correlates with arrangements of contiguous pentagons.[4]

Parameter Selection. The present study attempted to correlate $\Delta H_f°$ as calculated by MM3(92) in ref 2 with six graph theoretical indices, namely, $\ln N$, where N was the number of carbon atoms in the structure, $\ln K$, $\ln[\mathrm{per}(\mathbf{A})]$, $\ln[\mathrm{per}(\mathbf{A})]/\ln K$, and the two Schultz indices $(\ln|\mathbf{A} + \mathbf{D}|)^{1/2}$ and $\ln[\lambda_1(\mathbf{A} + \mathbf{D})]$. These parameters were selected on the basis of prior success as predictors of $\Delta H_f°$, from both the references cited above and the author's unpublished results.

Calculations. **A** and **S** matrices were prepared by hand. Hand-calculation of large **S** matrices is prone to error, but in the present circumstance it was possible to build stepwise from the simple fluoranthene structure to buckminsterfullerene. It was already known[13] that $K = 12\,500$ for buckminsterfullerene, so the correct result $|\mathbf{S}| = 156\,250\,000$ indicated that the **S** matrix for this structure, and therefore for all the intermediate structures, was correct. Determinants and Schultz indices were computed in Mathematica. Per(**A**) values for structures $C_{\leq 52}H_{10}$ were computed on a 90 MHz Pentium-class PC using code written by the author.[14] Per(**A**) values for larger structures were computed on a Cray C94 at the National Environmental Supercomputing Center (NESC), Bay City, MI, using a different code developed by Kallman.[15−19] Kallman's algorithm was recommended[16] as the best available for computation of per(**A**) in the present application after a literature search comparative evaluation against other alternatives.[17,17−22] For the PAH matrices studied here, CPU time to compute per(**A**) on either machine approximately doubles with each addition of C_2. On the 90 MHz machine, per(**A**) for $C_{52}H_{10}$ took exactly three weeks. On the Cray, C_{60} took 15.6 h using a single processing element and executing in predominantly scalar mode. Further performance enhancement of Kallman's algorithm should be possible by exploiting parallelism, but it would be difficult to adapt the current code because of its complex branching structure.

[⊗] Abstract published in *Advance ACS Abstracts,* September 15, 1997.

S1089-5639(97)01302-9

TABLE 1: Relevant Parameters for the 29 Compounds in this Study. Schlegel Diagrams for These Compounds Are Located in Ref 1

					$\Delta H_f°$ (kcal/mol)		
compound	K	per(**A**)	ln[per(**A**)]/ ln K	ln[λ_1 (**A**+**D**)]	ref 2	eq 1	Δ
$C_{16}H_{10}$	6	36	2.0000	3.8790	71.14	61.33	9.81
$C_{18}H_{10}$	8	64	2.0000	4.0260	94.42	93.64	0.78
$C_{20}H_{10}$	11	125	2.0136	4.1663	119.93	125.63	−5.70
$C_{22}H_{10}$	12	156	2.0322	4.3092	155.84	158.63	−2.79
$C_{24}H_{10}$	13	221	2.1046	4.4302	188.09	191.43	−3.34
$C_{26}H_{10}$	15	397	2.2096	4.5456	219.70	225.79	−6.07
$C_{28}H_{10}$	17	589	2.2513	4.6543	250.55	253.24	−2.69
$C_{30}H_{10}$	22	1532	2.3728	4.7488	277.33	284.42	−7.09
$C_{32}H_{10}$	29	2445	2.3169	4.8462	301.21	301.04	0.17
$C_{34}H_{10}$	42	4620	2.2576	4.9313	323.37	314.66	8.71
$C_{36}H_{10}$	64	9384	2.1993	5.0134	341.63	327.71	13.92
$C_{38}H_{10}$	110	21 452	2.1218	5.0920	343.90	338.24	5.62
$C_{40}H_{10}$	213	64 653	2.0661	5.1621	345.06	348.97	−3.91
$C_{42}H_{10}$	275	104 597	2.0577	5.2346	361.05	364.19	−3.14
$C_{44}H_{10}$	373	192 217	2.0546	5.2997	375.23	378.23	−3.00
$C_{46}H_{10}$	520	374 660	2.0521	5.3621	389.29	391.73	−2.44
$C_{48}H_{10}$	735	752 793	2.0503	5.4220	403.52	404.74	−1.22
$C_{50}H_{10}$	1125	1 863 365	2.0551	5.4767	416.57	417.17	−0.60
$C_{52}H_{10}$	1335	2 731 689	2.0593	5.5337	433.03	430.06	2.97
$C_{54}H_{10}$	1410	3 523 276	2.0789	5.5911	463.73	444.37	19.36
$C_{54}H_{10}(R)$[a]	1935	5 784 773	2.0575	5.5858	463.33	441.37	21.96
$C_{54}H_8$	2175	8 071 753	2.0695	5.5617	482.86	437.10	45.76
$C_{56}H_8$	2985	15 063 725	2.0656	5.6114	503.66	447.69	55.97
$C_{56}H_6$	3225	18 422 357	2.0708	5.5874	519.90	442.86	77.04
$C_{58}H_6$	4840	44 063 184	2.0745	5.6331	539.73	453.22	86.51
$C_{58}H_4$	5500	61 587 000	2.0825	5.6166	559.06	450.28	108.78
$C_{60}H_4$	7885	127 598 633	2.0801	5.6612	587.40	459.87	127.53
$C_{60}H_2$	9000	182 936 160	2.0895	5.6465	611.58	457.44	154.14
C_{60}	12 500	395 974 320	2.0986	5.6384	573.72	456.45	117.27

[a] Rearranged $C_{54}H_{10}$; see Figure 2.

Results and Discussion

All values of $\Delta H_f°$ are in kcal/mol. Stepwise multiple regression on the first 19 structures, $C_{16}H_{10}$ to $C_{52}H_{10}$, with $F = 3$ to enter or remove parameters, gave

$$\Delta H_f° = 220 \pm 3\{\ln[\lambda_1(\mathbf{A} + \mathbf{D})]\} + 85.7 \pm 12.7\{\ln[\text{per}(\mathbf{A})]/\ln K\} - 962 \quad (1)$$

$$n = 19, \quad r = 0.9987, \quad s = 6.10, \quad F = 3099$$

The success of ln[per(**A**)]/ln K here was somewhat surprising because the only previously published use of this parameter showed its correlation with counts of various arrangements of contiguous pentagons in fullerenes.[4] None of the structures in the present study have any contiguous pentagons. Equation 1 therefore presents a new use of this parameter. Inputs and outputs for eq 1 are listed in Table 1.

Figure 1 illustrates the result of applying this equation to the $C_{\geq 54}H_x$ structures. Clearly, something happens energetically when carbons 53 and 54 are added that had not happened in previous steps. Also from Figure 1, it is apparent that the $\Delta H_f°$ value calculated by MM3(92) increases more from C_{52} to C_{54} that it did for prior C_2 additions. The numerical tabulations in Table 1 also clearly disclose these features.

Figure 2 shows why this addition should be so much different from the others. Steric crowding between the two hydrogens is much more severe in the $C_{54}H_{10}$ structure (upper right, Figure 2) than in the $C_{52}H_{10}$ structure, or in any of the predecessor C_nH_{10} structures. As ref 1 states, the buckminsterfullerene structure builds up naturally by addition of C_2 units to fluoranthene up through $C_{50}H_{10}$. The addition of another C_2 unit must form a hexagon that is not present in buckminsterfullerene, but must convert later to a pentagon by rearrangement. The $C_{52}H_{10}$ structure thus formed (upper left, Figure 2) does not, however, display any obvious steric strain problems, and the MM3(92) calculations reported in ref 2 indicate about the same increase in $\Delta H_f°$ as for the preceding few additions of C_2 (ca. 14 kcal/mol). The increase in going from $C_{52}H_{10}$ to $C_{54}H_{10}$, on the other hand, is about twice that value.

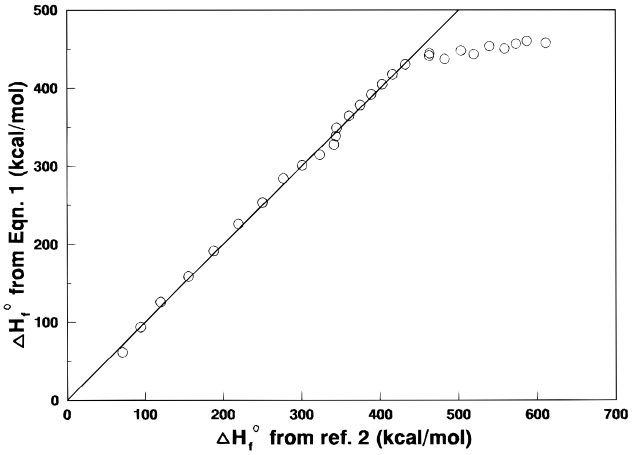

Figure 1. $\Delta H_f°$, MM3(92) calculations from ref 2 (X-axis) and as predicted by eq 1 (Y-axis). The straight line is the regression line through the first 19 points.

Indeed, there is no obvious reason why the C_2 unit should undergo the proposed 3+2 addition to $C_{52}H_{10}$ instead of one of the two equivalent 4+2 additions that are available, analogous to the formation of $C_{52}H_{10}$. Following the 4+2 addition, of course, there would be no low-energy pathway available for the structure to continue on its way toward the familiar I_h form of C_{60}. If the scheme proposed in ref 1 is correct, perhaps this detour at the $C_{54}H_{10}$ stage is the fate of some of the carbon that does not become C_{60}. It is also worth noting that the addition of C_2 to $C_{52}H_{10}$ is the only 3+2 addition proposed in the entire sequence. The others are all 4+2.

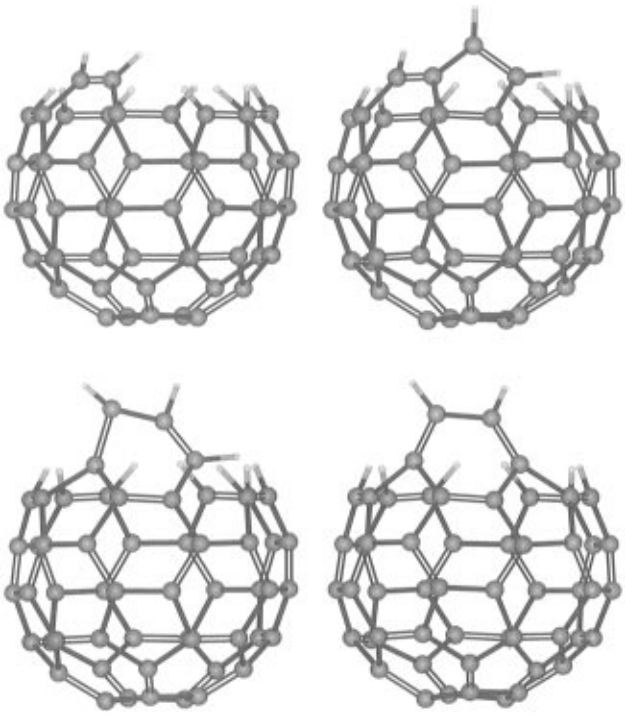

Figure 2. $C_{52}H_{10}$, upper left; $C_{54}H_{10}$, upper right; rearranged $C_{54}H_{10}$, lower left; $C_{54}H_8$, lower right.

Rearranged $C_{54}H_{10}$ (lower left, Figure 2) has not relieved the steric crowding found in its predecessor. The distortion of the protruding six-membered ring in the rearranged $C_{54}H_{10}$ structure is obvious. In fact, MM3(92) predicts nearly identical energies for these two structures. Given the energies, the question arises of what is driving the rearrangement. If the structures in Figure 2 are actual precursors of buckminsterfullerene, then they may exist in equilibrium. The rearranged $C_{54}H_{10}$ can dehydrogenate to $C_{54}H_8$ (lower right, Figure 2), while similar dehydrogenation of the unrearranged $C_{54}H_{10}$ would give a structure with two contiguous pentagons, an energetically unfavorable situation.

It should be noted that eq 1 predicts $\Delta H_f^\circ = 456$ kcal/mol for buckminsterfullerene, a value about 25% below experiment.[6] As stated above, MM3(92) comes much closer. Thus, even without reference to the scheme proposed by Pope et al.,[1] it is clear that the structures above C_{52} differ from the smaller ones in some way that causes a large deviation from eq 1. Inspection of the structures in Figure 2 is important in determining what sorts of interactions are not modeled by eq 1. At first sight, Figure 2 suggests that, for the $C_{54}H_{10}$ through C_{60} structures, the predictor variables from eq 1 are still valid, but the coefficients and intercept are different. Use of these two variables with the 10 structures in question, however, gives eq 2, for which the correlation coefficient is less than desirable.

$$\Delta H_f^\circ = 1087 \pm 348\{\ln[\lambda_1(\mathbf{A}+\mathbf{D})]\} + 1445 \pm 925\{\ln[\mathrm{per}(\mathbf{A})]/\ln K\} - 8574 \quad (2)$$

$$n = 10, \quad r = 0.9017, \quad s = 25.6, \quad F = 15.2$$

In eq 2, $C_{54}H_{10}$ is an outlier, with a residual nearly twice that of any other point. If this point is excluded as a transitional structure between the $C_{16}H_{10}$−$C_{52}H_{10}$ set and the rearranged $C_{54}H_{10}$−C_{60} set, the situation improves somewhat:

$$\Delta H_f^\circ = 1087 \pm 348\{\ln[\lambda_1(\mathbf{A}+\mathbf{D})]\} + 1445 \pm 925\{\ln[\mathrm{per}(\mathbf{A})]/\ln K\} - 8574 \quad (3)$$

$$n = 9, \quad r = 0.9503, \quad s = 17.8, \quad F = 28.0$$

Stepwise multiple regression with the entire set of five predictor variables listed above on the same 10 structures in eq 2, however, provided an equation containing only ln N:

$$\Delta H_f^\circ = 1143 \pm 96(\ln N) - 4094 \quad (4)$$

$$n = 10, \quad r = 0.9730, \quad s = 12.8, \quad F = 142$$

It is important to note here that extrapolating eq 4 back to smaller structures gives $\Delta H_f^\circ = 0$ at $N \approx 36$. Thus, the entire data set of 29 structures can be cleanly divided into two types: those that have ΔH_f° as described by eq 1 and those that are described by eq 4. It is also noteworthy that eq 4 predicts $\Delta H_f^\circ = 590$ kcal/mol for buckminsterfullerene, a value quite close to experiment.[6]

Perhaps the most significant feature of the plot in Figure 1 in terms of molecular structure is that the slope does not go back to that of the C_{16}−C_{52} subset after the break between C_{52} and C_{54}. This must mean that, by the C_{54} stage, steric crowding around the open part of the structure has become so severe that all subsequent additions of C_2 are qualitatively different in terms of energetics than the previous additions. Also, this crowding is not completely relieved by dehydrogenation and formation of new carbon−carbon bonds. Note, however, that C_{60} is neither the rightmost nor second rightmost point on the plot, but the point before that. The order of ΔH_f° from MM3(92) is $C_{60}H_2 > C_{60}H_4 > C_{60}$. Equation 4 predicts the same energy for all three structures because they all have the same value of N. The high value of r for eq 4 indicates that the energetics of C_2 addition is at least approximately the same for $C_{52}H_{10} + C_2$ and beyond and that rearrangement and dehydrogenation are less significant energetically than C_2 addition. The larger energies associated with adding the last tier of C_2 units to close C_{60} may explain why macroscopic yields of smaller fullerenes are not produced experimentally. More work remains to be done, however, on alternate pathways leading to fullerenes larger than C_{60}. In concert with other computational techniques, chemical graph theory may play a role in such studies.

Conclusions

ΔH_f° of 19 postulated buckminsterfullerene precursors C_NH_{10} is well predicted by an equation involving two graph theoretical parameters up through $N = 52$. ΔH_f° for 9 remaining precursors $C_{\geq 54}H_{\leq 10}$, as well as C_{60} itself, are predicted by an equation involving only ln N. The abrupt change in the trend of ΔH_f° above $N = 52$ appears to be due to steric crowding associated with closing the fullerene. The crowding is to be expected, but the abruptness of its onset is surprising. This study further defines some of the capabilities and limitations of chemical graph theory in predicting physical properties of PAHs. A usual advantage of graph theoretical parameters is that they are much easier to calculate than quantum mechanical parameters. The permanent of the adjacency matrix, per(**A**), has been an exception, but the present work shows that per(**A**) values for matrices up to $N = 52$ are now accessible with a fast desktop computer.

This study indicates that, in the formation of buckminsterfullerene from $C_{16}H_{10}$ and C_2 units, addition of the last tier to close the fullerene is qualitatively different from prior steps. This fact may explain the experimental observation that the forming C_NH_{10} does not close prematurely to give fullerenes smaller than C_{60} in significant yields.

Disclaimer. This document has been reviewed by the Office of Pollution Prevention and Toxics, USEPA, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency, nor

does the mention of trade names or commercial products constitute endorsement or recommendation for use.

Acknowledgment. The author wishes to thank the NESC and its Scientific Customer Support Group for providing resources that enabled the computational solution to the $N \geq 54$ problems. Special appreciation is due to Dr. George Delic for identifying the Kallman algorithm, invaluable assistance in porting the computer code, and generating performance evaluation data on the NESC Cray C94.

References and Notes

(1) Pope, C. J.; Marr, J. A.; Howard, J. B. *J. Phys. Chem.* **1993**, *97*, 11001.

(2) Pope, C. J.; Howard, J. B. *J. Phys. Chem.* **1995**, *99*, 4306.

(3) Cash G. G., *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 815.

(4) Cash, G. G. *Polycyclic Arom. Compd.*, in press.

(5) Fowler; P. W.; Manolopoulos, D. E. *An Atlas of Fullerenes*; Clarendon Press: Oxford, 1995; pp 73−74.

(6) Cioslowski, J. *Electronic Structure Calculations on Fullerenes and Their Derivatives*; Oxford University Press: New York, 1995; p 54.

(7) Trinajstić, N. *Chemical Graph Theory*, 2nd ed.; CRC: Boca Raton, FL, 1993; Chapter 4.

(8) Bersohn, M. A. *J. Comput. Chem.* **1982**, *4*, 110.

(9) Müller, W. R.; Szymanski, K.; Knop, J. V.; Trinajstić, N. *J. Comput. Chem.* **1987**, *8*, 170.

(10) Mohar, B.; Pisanski, T. *J. Math. Chem.* **1988**, *2*, 267.

(11) Cash, G. G. *Chemosphere* **1996**, *33*, 2081.

(12) Schultz, H. P.; Schultz, E. B.; Schultz, T. P. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 27.

(13) Klein, D. J.; Liu, X. *J. Comput. Chem.* **1991**, *12*, 1260.

(14) Kasteleyn, P. W. *J. Math. Phys.* **1963**, *4*, 287.

(15) Kasteleyn, P. W. In: *Graph Theory and Theoretical Physics*; Harary, F., Ed.; Academic Press: New York, 1967; Chapter 2.

(16) Klein, D. J.; Schmalz, T. G.; Hite, G. E.; Seitz, W. A. *J. Am. Chem. Soc.* **1986**, *108*, 1301.

(17) Cash, G. G. *J. Math. Chem.* **1995**, *18*, 115.

(18) Kallman, R. *Math. Comput.* **1982**, 167.

(19) Kallman, R. Computer Programs for Evaluating Permanents of 0,1 Matrices. Department of Mathematical Sciences, Technical Report #48, Ball State University, Muncie, IN, July 20, 1980.

(20) Delic, G. Private communication.

(21) Ryser, H. J. *Combinatorial Mathematics*; Carus Mathematical Monographs, Mathematical Association of America, Wiley: New York, 1963.

(22) Nijenhuis, A.; Wilf, H. S. *Combinatorial Algorithms for Computers and Calculators*, 2nd ed.; Academic Press: New York, 1978.