

# On the Heats of Formation of Alkanes<sup>1</sup>

Jenn-Huei Lii and Norman L. Allinger\*

Department of Chemistry, Chemistry Annex, University of Georgia, Athens, Georgia 30602

We would like to dedicate this paper to the memory of the good friend and long term colleague of the senior author, Professor Ernest L. Eliel

Received April 21, 2009; accepted August 11, 2009

**Abstract:** A broad diverse test set of alkanes and cycloalkanes previously studied with MM4 calculations has had the heats of formation calculated by several different quantum mechanical methods: Hartree-Fock, MP2, and MP4, and also by B3LYP and B3LYP + dispersion energy. Overall, three computational methods (MM4, MP4, and B3LYP + dispersion) yield results that are generally of experimental accuracy. These results are analyzed and compared in some detail.

**Keywords:** MM4, MP4, B3LYP, Dispersion Energy, Heats of Formation.

**Resumen:** Se describen los cálculos de calores de formación de un conjunto amplio y diverso de alcanos y cicloalcanos que habían sido estudiados previamente por el método MM4. En este trabajo se emplean diversos métodos de mecánica cuántica, entre ellos se encuentran: Hartree-Fock y MP4, así como B3LYP y B3LYP + energía de dispersión. De estos métodos de cálculo, los métodos computacionales MM4, MP4 y B3LYP + energía de dispersión proporcionan resultados que coinciden con los datos experimentales. Estos resultados se analizan y comparan detalladamente.

**Palabras clave:** MM4, MP4, B3LYP, Energía de dispersión, calores de formación.

## Introduction

Heats of formation have a long history, both experimentally, and calculationally. The most simple calculational approach is to use bond energies, and assume additivity, and this gives a qualitatively encouraging result. It was quickly found, however, that "structural features" also had to be included in the calculation if the result was to be more than qualitative. This procedure worked reasonably well for a great many kinds of compounds, and is summarized in the well known book by Sidney Benson [1]. The basic method assumes the additivities of the energies of various structural features in molecules. The problem here is that while there are large numbers of molecules for which these structural features do have additive energies, there are also substantial numbers of molecules where they do not. To some extent one can get around this by introducing new structural feature parameters, but a more general, and more widely useful method was desirable.

When molecular mechanics became available, it was possible to calculate "strain energies" of molecules in a general way [2-5]. If molecules are bent, stretched, or twisted, additional strain energies are introduced. Because of these that the Benson method really comes to a limit of applicability. But if these strains can be calculated and added to the Benson type of calculation, then the method can be extended much further.

Alkanes constitute the most simple class of compounds for such calculations. There are only two different kinds of atoms present (carbon and hydrogen), and only two kinds of bonds (C-C and C-H), where the bonding is between tetrahedrally hybridized carbons. Because the hydrogen has a slightly different electronegativity than the carbon, there are small partial charges on the various atoms. These are sufficiently small, however, that they are usually ignored in heat of formation calculations of alkanes. Thus the molecular mechanics model,

which is classical in behavior, was developed. And it allows the general calculation of the heats of formation for alkanes with experimental accuracy. Much high-quality thermochemistry on these compounds has been carried out since the 1940's, and there are voluminous amounts of accurate experimental data on a very wide variety of compounds [6].

The MM4 force field for alkanes [7] and a number of other classes of organic molecules [8] (functionalized alkanes) was published, beginning in 1996. For the most part MM4 reproduced the experimental information on the heats of formation of alkanes to within chemical accuracy. For a few compounds that force field gave values that disagreed with experiment by relatively large amounts. In those cases we believe that the calculated value is almost always more nearly correct than the experimental value (see later).

As one begins to be concerned with the heats of formation of functionalized molecules, the situation becomes more complicated. MM4 has been extended to a sizable number of such molecules, including classes such as amines, alcohols, carboxylic acids, and others [8]. The results here are less accurate than for hydrocarbons, but the experimental data are also fewer, and they are also generally of lower accuracy. But still, the MM4 accuracy is comparable with that from experiment. Going further to polyfunctional molecules, the calculations become increasingly complex, because the number of parameters required increases with the number of heteroatoms present. And worse, experimental heat of formation data for polyfunctional molecules are scarce, or non-existent for most cases of interest.

Chemistry is still considered by most as an experimental science, although it is becoming increasingly computational in nature. This usefulness of computational chemistry in the mainstream of the science became significant in the 1960's, and has continually increased since that time. In principle, one

can simply calculate the energy of a molecule relative to the energies of the isolated electrons and nuclei by direct solution of the Schrodinger equation, and compare this with the similar calculations for the elements in their standard states. One can thereby obtain the heat of formation directly. In practice, this approach was long stymied by the fact that the energies of atoms and molecules relative to fundamental particles are enormous numbers, and hence to give heats of formation of useful accuracy, the calculations must be accurate to a large number of significant figures. This approach has been, and is being pursued [9]. Also useful is the use of isodesmic and related equations [10]. This method suffers from its own set of problems, but this is certainly a powerful method that will become more useful going forward.

Some things are especially tedious to calculate by quantum mechanics, the vibrational energies of the molecules over the Boltzmann distributions of large numbers of conformations in flexible molecules, for example. But here the calculation is trivial by the methods of molecular mechanics. For practical purposes, a currently useful approach for the utilization of quantum mechanics here is to employ empirical methods where they are sufficiently accurate and reliable, along with quantum mechanical procedures to obtain results. There are various ways that this can be done, and they each have advantages and disadvantages. We will limit our discussion here to what we will refer to as the Wiberg/Schleyer method [11]. This is a method that basically extends the Benson/molecular mechanical method by using the same calculational framework, but obtains the fundamental energy number for each calculation, not by experiment, but by a quantum mechanical calculation. We already know that this method works to the limit of experimental accuracy when molecular mechanics is employed with experimental data with the alkanes [2, 12], and the same formalism can be used for quantum mechanics as was used with molecular mechanics [12]. If we could use this approach employing experimental data with quantum mechanics, we would have a potentially very powerful way to obtain heats of formation. The important advantage of this procedure over the molecular mechanics analog is that it should work equally well for any combination of atoms, whereas the molecular mechanics approach becomes much more complicated and presumably less accurate after one leaves the area of hydrocarbons, or near hydrocarbons.

We have previously described how the procedure developed by Wiberg and Schleyer can be applied to first hydrocarbons, and then extended to functionalized molecules.<sup>12</sup> At the time that work was done the limitations in available computer power were serious, and relatively small basis sets were used at the Hartree-Fock level. The overall results were, however, quite encouraging. Much of the error introduced by using small basis sets and incomplete correlation carries over fairly well from one molecule to another, and therefore one can improve the computational accuracy with the aid of a few well chosen parameters. Originally the 6-31G\* basis set was used, and the results obtained for heats of formation over a broad set of hydrocarbons were comparable in accuracy with

the experimental numbers. It was later realized, however, that part of the good agreement came because the set of data fit to was limited, and did not adequately represent some structural types for which it was more difficult to obtain accurate results.

Later these calculations were extended from Hartree-Fock to the B3LYP level [13], with the same basis set, and the results were somewhat better. But again, there were some hidden errors here. It is now possible, and even easy, to carry out similar calculations to those described above with larger basis sets, with more correlation, and with a much larger (and more complete) data set of compounds.

## Discussion

In the present work we have proceeded in the following way. First, we have the MM4 heat of formation calculations<sup>7</sup> on a test set of 56 compounds (of which four were given zero weight, because we doubted the accuracy of the experimental numbers). (The last two compounds in the table were added subsequently and so are also weighted zero.) The full list is shown in Table 1.

**Table 1.** MM4 Heats of Formation (Gas, 25°), kcal/mol<sup>7,a</sup>

Wt.	Calc.	Exp.	Difference (Calc.-Exp.)	Compound
1	-17.89	-17.89	.00	Methane
6	-19.75	-20.24	.49	Ethane
9	-24.99	-24.82	-.17	Propane
8	-29.97	-30.15	.18	Butane
7	-35.03	-35.00	-.03	Pentane
7	-40.12	-39.96	-.16	Hexane
6	-45.16	-44.89	-.27	Heptane
5	-50.21	-49.82	-.39	Octane
4	-55.24	-54.75	-.49	Nonane
9	-32.36	-32.15	-.21	Isobutane
7	-36.69	-36.92	.23	Isopentane
9	-40.67	-40.27	-.40	Neopentane
7	-42.16	-42.49	.33	2,3-Dimethylbutane
6	-49.01	-48.95	-.06	2,2,3-Trimethylbutane
6	-49.70	-49.20	-.50	2,2-Dimethylpentane
6	-47.86	-48.08	.22	3,3-Dimethylpentane
6	-44.40	-45.25	.85	3-Ethylpentane
6	-48.12	-48.21	.09	2,4-Dimethylpentane
5	-52.85	-53.18	.33	2,5-Dimethylhexane
5	-53.86	-53.92	.06	2,2,3,3-Tetramethylbutane
5	-56.75	-56.64	-.11	2,2,3,3-Tetramethylpentane
5	-57.59	-57.80	.21	Di- <i>tert</i> -butylmethane
7	-55.33	-55.67	.34	Tetraethylmethane
0	-54.06	-56.40	2.34	Tri- <i>t</i> -butylmethane
9	-18.59	-18.74	.15	Cyclopentane

Wt.	Calc.	Exp.	Difference (Calc.-Exp.)	Compound
8	-29.59	-29.43	-.16	Cyclohexane
7	-27.88	-28.22	.34	Cycloheptane
7	-29.72	-29.73	.01	Cyclooctane
6	-31.37	-31.73	.36	Cyclononane
4	-36.74	-36.88	.14	Cyclodecane
0	-53.49	-54.59	1.10	Cyclododecane
6	-33.02	-33.04	.02	1,1-Dimethylcyclopentane
2	-25.70	-25.27	-.43	Methylcyclopentane
5	-30.35	-30.34	-.01	Ethylcyclopentane
9	-36.99	-36.99	.00	Methylcyclohexane
6	-43.43	-43.26	-.17	1,1-Dimethylcyclohexane
6	-41.71	-41.13	-.58	1-ax-2-eq-Dimethylcyclohexane
6	-43.34	-42.99	-.35	1-eq-2-eq-Dimethylcyclohexane
2	-30.04	-30.50	.46	Bicyclo[3.3.1]nonane
3	-22.72	-22.20	-.52	<i>cis</i> -Bicyclo[3.3.0]octane
4	-15.31	-15.92	.61	<i>Trans</i> -Bicyclo[3.3.0]octane
6	-43.60	-43.54	-.06	<i>trans</i> -Decalin
6	-40.91	-40.45	-.46	<i>cis</i> -Decalin
5	-31.73	-31.45	-.28	<i>trans</i> -Hydrindane
5	-31.01	-30.41	-.60	<i>cis</i> -Hydrindane
1	-57.74	-58.12	.38	<i>tsr</i> -perHanthracene
0	-50.27	-52.73	2.46	<i>tat</i> -perHanthracene
5	-13.10	-13.12	.02	Norbornane
5	-30.90	-30.62	-.28	1,4-Dimethylnorbornane
2	-31.85	-31.76	-.09	Adamantane
4	-66.55	-67.15	.60	1,3,5,7-Tetramethyladamantane
2	-20.11	-20.54	.43	Protoadamantane
3	-35.04	-34.61	-.43	Congressane
7	-21.90	-22.58	.68	Bicyclo[2.2.2]octane
3	-24.01	-24.46	.45	Perhydrotriquinacene
0	14.05	18.2 <sup>b</sup>	4.15	Dodecahedrane
<b>0</b>	<b>-58.34</b>	<b>(-59.22)<sup>c,d</sup></b>	<b>0.88</b>	<b>2,2-Di-<i>t</i>-butylpropane</b>
0	-28.14	- <sup>c</sup>	-	Tetra- <i>t</i> -butylmethane

<sup>a</sup> The standard deviation = .355 based on 52 equations.

The weighted standard deviation = .342 based on 52 equations.

<sup>b</sup> This value has been updated from that given earlier [7].

<sup>c</sup> No experimental value is available.

<sup>d</sup> This is the MP4 value.

The set is diverse without being overly large and needlessly redundant, and it omits three- and four-membered rings, which are treated separately. We have used this test set and carried out calculations using several different quantum mechanical methods. These were Hartree-Fock, MP2, and MP4. The calculations were also repeated for the same group of compounds with B3LYP, and then with B3LYP plus dispersion energy. We feel that it has been previously demonstrated that the MM4 results are competitive in accuracy with experiment [7]. There are always problem cases, and these will be individually discussed. But generally speaking, the MM4 results are of chemical accuracy, and the quantum mechanical results

can be compared to those when experimental information is unavailable or of questionable accuracy. Looking ahead, we want to devise a method for obtaining quantum mechanically based heats of formation that can reliably be used in place of the calorimetric values. We might note here that the four compounds given zero weight in the original MM4 treatment also show large deviations from experiment in the quantum mechanical calculations. And these discrepancies are always similar in magnitude, and in the same direction as the MM4 results for those four compounds. We think that there is no doubt but that the experimental accuracy for these compounds is indeed less than desired. We have added two additional compounds to the test set, for reasons that will be discussed later. But they are weighted zero in the heat of formation calculations, so that we can make statistical comparisons between the various calculational methods and the earlier literature.

The general Wiberg/Schleyer method [11] was used here, with some modification [12, 13]. The basic idea is that any quantum mechanical scheme that is at a calculational level lower than the Schroedinger limit will contain errors in the results. If we can identify what these errors are, and why they occur, then to the extent that they are systematic, we can reduce them with the aid of parameterization. One wants to understand insofar as possible just why these parameters are needed. Specifically, what is it that they do?

In Table 2 is given the parameter set used in MM4 to calculate heats of formation. Each of these parameters is required if one is to obtain good results, and we want to understand why.

First, there are the bond energy terms. We have an energy term for each the C-C and the C-H bond. These terms will give us the bonding energy, under a certain set of idealized circumstances. If the bond is deformed in any way, that, of course, changes the energy. In quantum mechanics, it changes the final energy of the system that one calculates. In molecular mechanics it changes, for example, the bond stretching energy, and then the final energy. Either way, we need the standard number, and we need to know how it changes with its environment. In molecular mechanics the calculations of electron correlation

**Table 2.** MM4 Alkane Heat of Formation Parameters<sup>a</sup> (kcal/mol).

C-C	=	-87.1067	Neo	=	-6.9273
C-H	=	-106.7763	R6	=	4.9713
Me	=	2.0108	R5	=	4.4945
Iso	=	-3.3565	TOR	=	.5767

<sup>a</sup>The parameters given in Table II (which are defined later in the text) are in part different from those originally published [7]. This is mainly due to the fact that there was an error in the original work, where methane was accidentally attributed a POP term of -4.20 kcal/mol instead of zero. This led to no difference in the calculated heats of formation, but it did lead to some of the parameters in this table having strange values. The parameter values given in Table 2 have that error removed and are to be regarded as the proper ones for MM4. It is emphasized that the original MM4 heat of formation parameters, and those given in Table 2 yield the same calculated heats of formation for alkanes.

within an atom are avoided by using the energy of the atom as a starting point. The correlation of the bonding electrons with each other and the rest of the atom is taken into account in the bonding parameter. Turning to the quantum mechanical calculation at the Hartree-Fock level, we can use the bonding parameters in exactly the same way, to account for that part of the electron correlation.

When we come to a pair of atoms that are *geminal* (bound to a common atom), the correlation between the electrons in those two atoms is not included at the outset in either the molecular mechanics calculation, nor in the Hartree-Fock calculation. That portion of the correlation can be included in both cases in the same way, by adding specific parameters to account for the bonding pattern at any carbon center. In molecular mechanics we need to include explicitly the 1,3 correlation energy at the reference point. Any change in this correlation with bending is then included in the bending parameter. Thus what is required here is a set of parameters for the five possible substitution arrangements possible in an alkane, where one has four things attached to each carbon. These four things in any particular case include from zero to four hydrogens (a total of five numbers), and a sufficient number of carbons so that the carbon and hydrogen substituents total to four. Whatever is not a hydrogen is a carbon, and so that is not an independent variable. Methane is a special case. But for all other alkane (and cycloalkane) molecules, each carbon in the structure in question is either primary, secondary, tertiary, or quaternary. These are exemplified by the carbon in ethane, and the central carbons in propane, isobutane, and neopentane. We presume that the correlation energy between the geminal electrons will differ from one to another in these systems, and hence we need that many parameters (we need only three\* actually, as one of them can be taken as zero, and methane is treated as a special case).

The long range correlation (1,4 and higher) in molecular mechanics is accounted for in the van der Waals term (dispersion energy). When we switch over to quantum mechanics, if the correlation energy were exactly calculated we would not need these terms. But we know that it will not be calculated at the Hartree-Fock level, and not exactly calculated at the higher approximate levels that we plan to use, so we will use these same terms as we do in molecular mechanics, and for the same reason.

In the present work we have included open-chain systems, and also cyclic (and polycyclic) systems, where there are any number and any size of rings, except the three- and four-membered rings. (The latter have been studied separately [14], and could be included, but they really constitute a separate problem which we will not discuss here. They are mentioned only for completeness.)

It was found in the MM4 molecular mechanics study that there is a set of parameters required to differentiate ring compounds from open-chain compounds. These are called R3, R4,

R5, and R6, and are used, respectively, for three-, four-, and five-membered rings, and rings which contain six or more ring members. These R parameters are simply parameters for rings of different sizes (and in use they are multiplied by integers telling us how many rings containing each of those numbers of members are present. The structural parameter sets, and also these R parameters, required differ markedly from one ring size to the next, except for C<sub>5</sub> and C<sub>6</sub>. Here the parameters are similar. We originally thought that it would be possible to choose our structural parameters so that they would fit both five- and six-membered rings, but it was subsequently found that one can obtain significantly better results if we allow these two rings to be treated as different cases, with different parameter sets. For rings containing six or more members, however, the same R parameter fits all, in a general way.

When quantum mechanical calculations on molecules are discussed, the structure or structures examined are located on the potential energy surface, at the stationary points. But, of course, molecules do not simply exist at points on the potential surface, rather they are in vibrational motion. If we are going to make interconnections between our calculated quantum mechanical structures and their real-world counterparts, we have to decide what to do about this vibrational motion. The original Benson scheme simply ignores it, and hopes that the parameterization will take care of the problem. And that works up to a point. But if one wishes to exploit the inherent accuracy of heat of formation calculations, one really has to do better than that. Specifically, in addition to the energy of the molecule on the potential surface at the global minimum, we need to add the energy from vibrational, rotational, and translational motions. If the molecule contains more than one conformation, then we must also add a Boltzmann distribution of energies of the conformations present. The general expression for the energy used to calculate the heat of formation is given by Equation (1) [12].

$$H_f^\circ = \sum_i n_i BE_i + \sum_i n_i ST_i + E_{(re)} + POP + TOR + T/R + E_{vib} \quad (1)$$

BE and ST are the bond energy and structural terms as in the Benson scheme, and *n* is the number of times each occurs. E<sub>(re)</sub> is the equilibrium energy of the lowest energy conformation. The other terms are from statistical mechanics and are defined below.

It should be clear that one can use any kind of a quantum mechanical calculation with Equation 1 to calculate heats of formation. But one has to decide exactly on the details of the calculation at the outset. The basis set, and method of correlation must be defined. For any such defined method then, one can obtain a parameter set that is specific for that exact method. And with that method then, one obtains the values of E<sub>(re)</sub> for each molecule, and proceeds as described.

For ordinary non-linear molecules, the translational and rotational motion (T/R in Equation 1) can be taken into account classically, with ½ RT for each of the six degrees of freedom. The vibrational contribution to the energy here (E<sub>vib</sub>)

\*The three parameters are given the self-explanatory names Me, Iso, and Neo for convenience.



must be individually calculated for each particular molecule. (In principle, it should be calculated for each conformation of each molecule, and then Boltzmann weighted for these, but it is generally regarded as satisfactory to simply calculate the vibrational energy for the lowest energy species present.) The remaining terms have the following meanings. POP is the *population increment*, or the extra energy that results from the presence of the Boltzmann distribution of the higher energy conformations. TOR is an *anharmonicity correction* for low torsional barriers present in the molecule. If one is interested in chemical accuracy, these corrections must be included. As a simple illustration, consider the two hydrocarbon molecules of similar size, methylcyclohexane and *n*-heptane, at 25°C. For methylcyclohexane, POP and TOR have the values 0.10 and 0.00, or a total of 0.1 kcal/mol. For *n*-heptane, the value for POP is 0.9, and that for TOR is 2.5, a total 3.4 kcal/mol. It just is not good enough to assume that differences like this can be cancelled out in the parameterization.

While one can in principle calculate the Boltzmann distribution by quantum mechanical methods, this would be a needless computational burden. We have (or can readily obtain) the Boltzmann distributions from molecular mechanics, and we simply used those.

The term TOR needs to be discussed briefly. Many torsional vibrations in a molecule (say cyclohexane) occur in such a deep potential well that they can be treated as ordinary vibrations. Some molecules, such as normal heptane, have only low potential barriers (in the range of 3-4 kcal/mol) for each torsion. Since the barriers are so low and anharmonic, the torsional levels are more closely packed than for ordinary vibrational levels. These low barriers lead to energy increases that are large enough to require accounting for. The simplest approximation is simply to count how many such low barriers are present, and assign a constant (parameter) for this correction, multiplied by the number of times that such a barrier occurs. (We exclude methyl groups in the calculation, because the methyl group gets its own parameter in the heat of formation anyway.) We call this parameter TOR. Pitzer and Gwinn [15] carried out an extensive study of these torsional barriers, and how the vibrational frequencies depend on barrier height and the foldness (2-fold, 3-fold, etc.). Looking at their data, it would seem that the value of 0.36 kcal/mol would be adequate to cover any (3-fold) torsion barrier from near 0 to 5 kcal/mol. Those above 5 kcal/mol can be treated as ordinary vibrations, and this correction is neglected. Very tiny barriers (with vibrational frequencies less than 10 wave numbers) should be separately treated as free (or essentially free) rotors.

$E_{\text{vib}}$  is the vibrational energy of the molecule. These vary substantially from one structure to another, and should properly be calculated in full. This is the way we normally proceed in MM4 calculations, and in the work described in this paper. We will refer to this procedure as *with statistical mechanics*. In the present study we have ordinarily included these statistical mechanical calculations throughout, but in a few special instances we have omitted them intentionally (*without statisti-*

*cal mechanics*) because the results thus obtained are of special interest, as will be discussed below.

With these parameters then, we hope to calculate heats of formation of hydrocarbons with the accuracy as previously described in molecular mechanics. We anticipate that we need to use the same group of parameters in quantum mechanics (reevaluated as appropriate, of course), and carry out the calculation in the same manner, except that we use the total quantum mechanically calculated energy ( $E_{\text{(re)}}$ ), in place of what in molecular mechanics is called the *steric energy*.

Next we selected the specific types of quantum mechanical calculations that we wished to carry out. We will discuss these in two parts. First, let us examine the Hartree-Fock calculations, and the results when these energies are further improved by MP2 and then by MP4 correlation calculations. The energies were calculated separately by each of these methods, and a parameter set was derived by least squares to fit the heat of formation data in each case for our test set of compounds in Table 1. These parameters themselves prove to contain a great deal of interesting information, and they are summarized in Table 2.

The first column in Table 3 gives the MM4 parameters, and at the bottom of the column is given the weighted standard deviation (WSD) determined first when the vibrational calculations are fully included (*with stat.*). Below that is given the weighted standard deviation when the vibrational part of the calculation is omitted (*w/o stat.*), although Boltzmann distributions are included. The parameters shown in each case are the set that gives the lowest WSD (with or w/o *stat.*, as the case may be). Note that the proper MM4 calculation (*with stat.*) gives a WSD of 0.34 kcal/mol over our test set of compounds. That is consistent with the overall experimental accuracy. If the vibrational part of the calculation is omitted (*w/o stat.*) the accuracy of the results deteriorates somewhat, and WSD (*w/o stat.*) is 0.45 kcal/mol. We can then compare these results with those in the Hartree-Fock column. Looking first at the accuracy, there are two points of interest. First the Hartree-Fock calculations are much less accurate, but second (and curiously), if one omits the statistics the result WSD (*w/o stat.*) is 0.82 kcal/mol, quite a bit better than if one includes the statistics (0.92 kcal/mol). We will return to that point later. But let's first compare the values for the parameters themselves in that table for the MM4 case, and for the Hartree-Fock case.

The bond energies in the MM4 case are roughly -87 and -106 kcal/mol for the C-C, and C-H bonds, in accord with experimental numbers. In the Hartree-Fock case those numbers are immensely larger, because in that case the energies of the atoms are also included in the so-called *bond energies*.

Then notice the spread of the next three numbers, Me/Neo. In the MM4 calculation, these span a range of 8.9 kcal/mol. The Neo system is about 8.9 kcal/mol more stable than the methyl system, with the iso and secondary are appropriately in between. This is largely a correlation effect between atoms that are 1,3 to one another. In the Hartree-Fock column, the same energy difference is slightly greater (10.5 kcal/mol)

**Table 3.** Alkane Heat of Formation Parameters (kcal/mol) for Various Calculations.

Parameter	MM4 <sup>a</sup>	HF <sup>b</sup>	MP2 <sup>b</sup>	MP4 <sup>b</sup>	B3LYP <sup>b</sup>	B3LYP/D <sup>b</sup>
C-C	-87.1067	11885.826	11928.521	11934.276	11960.060	11958.022
C-H	-106.7763	6301.245	6325.032	6331.240	6351.951	6344.581
Me	2.0108	2.540	-0.173	0.835	1.664	0.961
Iso	-3.3565	-3.537	0.704	-0.969	-2.767	-1.825
Neo	-6.9273	-8.028	2.594	-1.426	-6.366	-3.795
R6	4.9713	8.127	1.409	2.349	4.171	2.693
R5	4.4945	6.935	0.366	1.440	3.374	1.396
TOR	0.5767	1.002	0.121	0.254	0.559	0.265
WSD with stat.	0.34	0.92	0.61	0.42	0.55	0.35
WSD w/o stat.	0.45	0.82	0.65	0.43	0.47	0.44

<sup>a</sup>See footnote *a* in Table 2.

<sup>b</sup>The geometry of each molecule was optimized for each calculation (except as below) using Gaussian 2003 [16] with the “tight” residual force criterion. The basis/correlation for columns two and three are as follows: RHF/6-31G\*; MP2/6-311++G(2d,2p). Column four was a single point calculation, MP4SDQ/aug-cc-PVTZ//MP2/6-311++G(2d,2p); column five was B3LYP/6-31G\*; and column six was the result of adding MM4 dispersion energy in each case multiplied by 0.5 to the energy from column five.

than in the MM4 case. Thus this group of parameters accounts for the 1,3 correlation as previously discussed.

Continuing down the column, there is an R6 parameter 5.0 kcal/mol in MM4, and 8.1 kcal/mol in Hartree-Fock. The value of the R5 parameter is rather similar to the value of the corresponding R6 parameter in each case. We expect that this R5 parameter will probably converge to the R6 parameter in a more accurate quantum mechanical calculation.

And finally is listed the calculated value for TOR, +0.58 kcal/mol from MM4, and +1.00 kcal/mol for Hartree-Fock. This number really should be about +0.36 (from experiment), so the MM4 value is more in line with expectation. Of course, in a least squares fitting with several adjustable parameters, the parameters assume those values which lead to an optimum fit, regardless of the physics of the situation. We would like the Me/Neo values to go to zero. And we would like the TOR value to go to about 0.36. The closer we get to those values, the less adjustment is being made by the parameters to the heat of formation. And the less adjustment that is required by the parameters, the better the result that one expects. But, of course, if the parameters do their job, they will get us near to the place where we want to be anyway. The weighted standard deviation shows how well they are doing that job. The Hartree-Fock values were considered to be pretty good twenty years ago, but now they don't look good at all. Partly that is due to the fact that the test sets of compounds that were used long ago were not sufficiently diverse, and they did not contain sufficient difficult cases (compounds with high dispersion energies).

If we add correlation into a Hartree-Fock calculation, we expect to get better results. If we compare the Hartree-Fock column with the MP2 column that follows it, we note that

there is considerable improvement in the weighted standard deviation in the overall MP2 calculation. Whereas the Hartree-Fock value was 0.92 when statistics was included, or 0.82 without, these numbers went down to 0.61 and 0.65 respectively. Note that there is a crossover, and that the MP2 values are better when the statistics is included, as expected. Other items of interest are these. The Me/Neo difference at the HF level was 10.6 kcal/mol, whereas at the MP2 level this spread has gone down to -2.8 kcal/mol. This is a big step in the right direction. Also note that the sign of the trend has reversed, and that Neo is now a positive number, and Me is negative. This means there has been an over-compensation for correlation in the MP2 calculation (oscillatory convergence). Also note that the values for R6 and R5 have approached more closely to zero, and more closely to each other.

Turning next to the MP4 column in Table 3, we note that the WSD values have again further improved considerably, now being 0.43 *w/o stat.*, and 0.42 *with stat.* These numbers are similar in accuracy to those from MM4. Note that R5, R6, and TOR are showing oscillatory convergence, but perhaps not exactly to the expected values (zero for R5 and R6, and 0.36 for TOR) [17]. (This is probably a result of the nature of the least squares method mentioned earlier.)

The calculations that we previously reported on alkanes that used the Hartree-Fock method also used the B3LYP method [13], which was the limit of our computational abilities at that time. These results are given as the next column in the table, again for the same test set of compounds. Perusing the parameter set in the B3LYP column, and comparing it to the parameter sets in the earlier columns, we note that the bond energies are similar in magnitude to those obtained from the earlier quantum mechanical calculations. The spread between

the Me/Neo numbers is 8.0 kcal/mol, in between HF and MP2, but in the HF order. R6 and R5 are similarly in between the HF and MP2 values. And the TOR value is pretty good, 0.56. The WSD's at the bottom of the column show that one obtains a better result without the statistical corrections (0.47) than one obtains with them (0.55), and these results are better than those obtained from MP2, but not as good as those obtained from MP4. So there do not appear to be any big surprises here.

Density functional theory calculations have certainly become highly useful, but at the current level of development, they have some shortcomings. The B3LYP method is the particular one that we have utilized here. It is said to be inadequate with respect to its lack of inclusion of dispersion energy in molecular calculations [18]. Since we calculate dispersion energies with MM4 routinely (as part of the van der Waals interaction), we have available all of the dispersion energies for the compounds of present interest (Table 4).

**Table 4.** MM4 Heats of Formation, Strain Energies and Dispersion Energies for Selected Alkanes<sup>a</sup>.

Compound	ISE <sup>b</sup>	SE <sup>c</sup>	DE <sup>d</sup>	Compound	ISE <sup>b</sup>	SE <sup>c</sup>	DE <sup>d</sup>	
Methane	-17.89	0.00	0.00	1,1-Dimethylcyclopentane	-33.02	6.59	7.40	-12.66
Ethane	-19.75	-0.03	-0.60	Methylcyclopentane	-25.69	5.60	6.42	-8.79
Propane	-24.99	0.00	0.00	Ethylcyclopentane	-30.35	6.15	7.61	-11.39
Butane	-29.97	0.02	0.86	eq-Methylcyclohexane	-36.99	0.39	0.49	-13.03
Pentane	-35.03	0.02	1.65	1,1-Dimethylcyclohexane	-43.42	2.37	2.37	-17.35
Hexane	-40.12	0.01	2.40	1-ax-2-eq-Dimethylcyclohexane	-41.71	3.14	3.14	-17.29
Heptane	-45.16	0.00	3.20	1-eq-2-eq-Dimethylcyclohexane	-43.34	1.49	1.51	-16.50
Octane	-50.20	-0.01	4.00	Bicyclo[3.3.1]nonane	-30.04	7.65	8.24	-20.39
Nonane	-55.24	-0.03	4.81	cis-Bicyclo[3.3.0]octane	-22.72	9.88	10.67	-13.24
Isobutane	-32.36	0.00	0.00	Trans-Bicyclo[3.3.0]octane	-15.29	18.10	18.10	-12.71
Isopentane	-36.58	0.99	1.62	Trans-Decalin	-43.60	0.52	0.52	-21.22
Neopentane	-40.67	0.00	0.00	cis-Decalin	-40.91	3.21	3.21	-22.57
2,3-Dimethylbutane	-42.16	2.69	3.41	Trans-Hydrindane	-31.73	7.03	7.03	-16.80
2,2,3-Trimethylbutane	-49.01	4.87	4.87	cis-Hydrindane	-31.01	7.70	7.75	-17.82
2,2-Dimethylpentane	-49.69	2.09	2.66	Trans-syn-trans-Perhydroanthracene	-57.73	0.40	0.40	-32.19
3,3-Dimethylpentane	-47.85	4.50	4.50	Trans-anti-trans-Perhydroanthracene	-50.26	7.87	7.87	-32.63
3-Ethylpentane	-44.40	3.67	5.49	Norbornane	-13.10	14.45	14.45	-11.08
2,4-Dimethylpentane	-48.12	1.99	3.29	1,4-Dimethylnorbornane	-30.90	13.26	13.26	-16.18
2,5-Dimethylhexane	-52.85	1.88	4.40	Adamantane	-31.85	2.92	2.92	-22.69
2,2,3,3-Tetramethylbutane	-53.86	8.33	8.33	1,3,5,7-Tetramethyladamantane	-66.55	1.46	1.46	-35.80
2,2,3,3-Tetramethylpentane	-56.74	11.29	11.29	Protoadamantane	-20.11	15.13	15.13	-22.31
Di-t-butylmethane	-57.58	9.29	10.45	Congressane	-35.04	4.37	4.37	-35.19
Tetraethylmethane	-55.33	8.72	8.72	Bicyclo[2.2.2]octane	-21.90	10.54	10.54	-16.37
Tri-t-butylmethane	-54.05	42.86	42.86	Perhydroquinacene	-24.01	12.18	12.18	-18.11
Cyclopentane	-18.59	5.58	6.15	Dodecahedrane	14.05	37.59	37.59	-34.41
Cyclohexane	-29.59	0.53	0.53	Di-t-butylpropane	-58.34	25.37	25.37	-37.27
Cycloheptane	-27.88	7.50	8.07	Tetra-t-butylmethane	-28.14	98.61	98.61	-75.50
Cyclooctane	-29.72	12.08	12.08					
Cyclononane	-31.37	15.51	16.28					
Cyclodecane	-36.74	15.97	16.74					
Cyclododecane	-53.48	10.76	11.68					

<sup>a</sup>In kcal/mol

<sup>b</sup>Inherent Strain Energies

<sup>c</sup>Strain Energies; d. Dispersion Energies.

We wondered if we might not simply add the dispersion energies to the B3LYP energies, and optimize the result to obtain an improved parameter set for heats of formation. It proved not to be quite that simple, but it almost is. We need additionally a scale factor to reduce the dispersion energies by a constant fraction before we utilize them with the B3LYP calculations. This scale factor was chosen empirically so as to give the best fit to three separate pieces of data that we regarded as most important, namely (1); the value of the WSD (*with stat.*) at the bottom of the B3LYP/D column, (2); a fit to the heat of formation of a particular compound (tri-*t*-butylmethane), which has the highest dispersion energy of any compound for which we have an experimental heat of formation, and (3); the value of TOR. This optimization gave us a dispersion energy scale factor of 0.5. So when the dispersion energies (from MM4) multiplied by 0.5 are added to the B3LYP energies of all of the compounds in our test set obtained in the usual way, we get a new set of energies, which were used to optimize a new parameter set. The results are shown in the final column in Table 3 (B3LYP/D). Looking at the WSD (*w/o stat.*) first, note that it has gone from 0.47 down to 0.44, a very good value, but it has gone down much further to 0.35 with statistics. This is even better than the MP4 value, and essentially equivalent to the MM4 value. By the statistical test then, this is a really excellent calculation of the heats of formation of this set of compounds.\* Also, if we look at a few of these parameters, the Me/Neo spread is 4.8 kcal/mol, still sizable, but in the range of the smaller spreads from the other methods shown in the table. R6 and R5 are also quite a bit smaller than they were with the B3LYP method before the dispersion was added. And TOR is quite close to the desired value, +0.26 instead of +0.36. And finally, the WSD is better with the statistical corrections than without.

We have mentioned in passing the differences between the weighted standard deviations with and without the statistical treatment added previously in this discussion, and now with all of the results in hand, we can interpret it. Note that the addition of the statistics improved the results for MM4 by a small amount (measured by the WSD). The addition of the statistical corrections similarly improve the MP2 and MP4 calculations. But, the statistics not only don't improve, but they make the Hartree-Fock and B3LYP results worse. On the other hand, B3LYP/D is better with the statistics added. Why is all of this?

The interpretation here is straightforward. If we consider molecules as belonging to one of two groups, ordinary molecules, or congested molecules, we can see what is happening. Compared to the ordinary molecules, the congested molecules

(like tri-*t*-butylmethane), have too many atoms crammed into too small a space. This means they have unusually high correlation energies. Also, when atoms are congested their vibrational motions are impeded by the presence of their neighbors. They bump into one another. And the result is that their potential wells are narrowed appreciably, which causes the vibrational levels to move upward and spread apart. Thus the zero-point energies of the congested molecules are increased relative to the non-congested, although their thermal energies are also decreased slightly. But these increases in zero-point energies are sizable and lead to significant increases in the heats of formation.

One can then compare the congested molecules with the ordinary molecules. As the correlation energies are increased in congested molecules (becoming more negative), the zero-point energies are also raised, with the former being the larger effect. Thus if we look at the Hartree-Fock calculation to begin with, there are no correlation energies included, and hence the total energies for congested molecules are relatively too high. If we add the correlation energies (in MP2, MP4), this moves these total energies down to nearer where they should be. But if we do not add the correlation energies and stabilize the molecules correctly, then adding the zero-point energies is only going to make things worse, as it makes the energy of the congested molecules still larger when they should be smaller.

Returning to Table 3 again, where the correlation energy is reasonably well accounted for (MM4, MP2, MP4, and B3LYP/D), including the statistical mechanical corrections lowers the WSD in each case, as it should. But in cases where little or no correlation energy is included, making these vibrational corrections simply makes the energies even higher than the too-high values that they already had, and things get worse. One reason that the original Wiberg-Schleyer methods work as well as they do is a result of this fortuitous cancellation of errors.

Finally, we would like to discuss a few specific "problem compounds." There always seem to be compounds that are at the edge or limit of our calculational/experimental abilities, and they tend to be of interest for that reason. We would like to explicitly consider here the four compounds shown in Table 5. They are di-*t*-butylmethane, 2,2-di-*t*-butylpropane, tri-*t*-butylmethane, and tetra-*t*-butylmethane. These are all highly congested compounds, in order of increasing congestion. They therefore show very serious structural deformations, as they arrange themselves into the most comfortable position that they can. And they also show quite unusual heats of formation, as a result of the high strain energies (and attendant high dispersion energies), even after all of their structural relaxations (Table 4).

In Table 5 are given heats of formation for the four compounds mentioned, as determined by various experiments/calculations. From the left, the first column gives experimental values for di-*t*-butylmethane, and tri-*t*-butylmethane. These numbers themselves do not directly convey much information. The strain energies of these compounds are given in Table 4.

\*The difference between the 0.34 – 0.42 WSD values of the different methods is not regarded as significant by the authors, as this is in the range of the average reported experimental errors. The similarly small differences between values of the MM4 calculations with and without statistics probably are significant however, because it seems to result mainly because of large improvements in a few values, rather than the small changes in many values (which also occur).



**Table 5.** Heats of Formation (kcal/mol)<sup>a</sup> (Problem Compounds. Part I).

	EXP	MM4	HF	MP2	MP4	HF	B3LYP	B3LYP/D
Di- <i>t</i> -BuMe	-57.80	-57.58	-56.39	-57.06	-57.28	-56.39	-57.06	-57.45
Di- <i>t</i> -BuPro	–	-58.34	-52.83	-60.17	-59.22 <sup>b</sup>	-52.83	-56.01	-59.33
Tri- <i>t</i> -BuMe	-56.40	-54.06	-42.15	-53.14	-51.69 <sup>c</sup>	-42.15	-48.72	-53.21
Tetra- <i>t</i> -BuMe	–	-28.14	+2.10	-31.79	–	+2.10	-16.23	-27.32

<sup>a</sup>The compound tri-*t*-butylmethane is of particular interest in the present work, because it has a higher dispersion energy than any of the other compounds studied for which an experimental heat of formation is available. We particularly wanted to obtain the MP4 energy of this structure, but were unable to do so after several attempts, because of the lengthy computer running time. We made several attempts, some of which lasted as long as 25 days, but were never able to complete the calculation because of some system or power failure which occurred each time before the calculation was complete. We therefore decided to carry out the calculation in an indirect manner. One might suppose that if a molecule is in a staggered but twisted conformation of low symmetry, twisting the torsion angles so as to put them into eclipsed conformations will sometimes give a structure with a much higher symmetry. Our version of Gaussian utilizes C<sub>2</sub> or C<sub>s</sub> symmetry only. Tri-*t*-butylmethane possesses C<sub>3</sub> symmetry, but Gaussian refuses to utilize it. If we twist the molecule into an all-eclipsed conformation, that conformation has C<sub>3v</sub> symmetry. While the program will not utilize that symmetry either, it will utilize the C<sub>s</sub> symmetry that is also now present, which was expected to shorten the running time for this molecule by tens of days. The question is, how much error is induced in the calculation by calculating the correlation energy for the eclipsed conformation, rather than for the staggered conformation? We carried out a similar calculation for the di-*t*-butylpropane molecule so as to judge the feasibility of this approach. We calculated the energy of the latter molecule in the ground state C<sub>2</sub> conformation, and also in the C<sub>2v</sub> conformation (Table 4). The energy is much higher in the latter, of course, but the correlation energy difference in going from MP2 to MP4, that is DE(MP2 – MP4), was -0.95 kcal/mol for C<sub>2</sub> and -0.92 for C<sub>2v</sub>. The corresponding energy differences were 7.46 and 7.42 for Δ(RHF – MP2). Thus substituting the C<sub>2v</sub> correlation for the C<sub>2</sub> introduced negligible error. Having all of the corresponding numbers for the tri-*t*-butyl compound except the MP4 value for C<sub>3</sub>, the value of the correlation energy of the C<sub>3</sub> form was estimated from that of the C<sub>3v</sub> form. The resulting for tri-*t*-butylmethane thus found was -51.69 kcal/mol at the MP4 level (Table 6). We have thought that this approximation should be accurate to about 0.2 kcal/mol or better, and therefore justifiable in the tri-*t*-butylmethane case.

<sup>b</sup>Est. from C<sub>2v</sub> structure.

<sup>c</sup>Est. from C<sub>3v</sub> structure.

These compounds do not seem to be exceptional in MM4 calculations (column 2, Table 5). The heat of formation of di-*t*-butylmethane is quite well calculated. The tri-*t*-butyl value appears not to be very well calculated, with an apparent error of 2.34 kcal/mol.\* Compounds of this general type (globular hydrocarbons) tend to be extremely difficult to purify, however. It cannot be assumed that the experimental value here is in fact any better than (or even as good as) the MM4 value. The other two compounds in Table 5 do not have experimentally known heats of formation, so the MM4 values will be used for the comparisons in what is to follow.

Taking this information as our basis, we can then look at the calculation of the heats of formation of each of these compounds by these various methods. Consider first the di-*t*-butylmethane, the top line in Table 5. The Hartree-Fock number is a little bit (1.4 kcal/mol) too positive, but the value becomes more negative if we carry out MP2 or MP4 calculations. Continuing along this line, starting from the Hartree-Fock number again (too positive), the B3LYP and B3LYP/D numbers become increasingly negative, and the best calculations (MM4, MP4, and B3LYP/D all agree with the experimental value to within a few tenths of a kcal/mol.

Di-*t*-butylmethane is a congested molecule with a strain energy of 10.45 kcal/mol (Table 4), but if we go to the next compound in the series, the corresponding propane, the extra two methyl groups that have been added in converting the methane to the propane are positioned in a very crowded area, which raises considerably the strain energy (to 25.37 kcal/mol) due to increased van der Waals repulsion. But when the van der Waals repulsion goes up, the dispersion energy goes down (more negative), and the correlation energy thus becomes increasingly important. We do not have an experimental heat of formation here, but taking the MM4 value as our standard, notice that the Hartree-Fock value is some 6 kcal too positive. MP2 brings this down, actually too far, so that it is now 2 kcal too negative. MP4 on the other hand brings it back up again, so that the number now is only 0.9 kcal away from the MM4 number, and we believe that both these values are reasonably accurate. Continuing on with the B3LYP calculations, correlation that is included in B3LYP brings the energy down, but not far enough. The B3LYP/D number (-59.33 kcal/mol) on the other hand, is in reasonable agreement with both the MP4 (-59.22) and MM4 (-58.34) results.

Continuing down Table 5, the tri-*t*-butylmethane has quite a high strain energy (42.86 kcal/mol). Here there is a 2.3 kcal/mol discrepancy between experiment and the MM4 heat of formation. But in any case the Hartree-Fock value is far too positive, by 12 or 14 kcal/mol. MP2 brings this value down into the correct range. The MP4 value raises the heat of formation about 1.4 kcal/mol, and is now 2.37 kcal/mol above (more

\*The MM4 calculations quoted in Table 1 were published in 1996. Subsequently, the heat of formation of tri-*t*-butylmethane was redetermined experimentally.<sup>19</sup> The original value was -56.40 kcal/mol. The redetermined value has moved slightly in the direction of the MM4 value, and is now -56.21 kcal/mol.

**Table 6.** Estimation Data for Di-*t*-Butylpropane and Tri-*t*-Butylmethane

	$\Delta H_f^\circ$ (in kcal/mol)			$\Delta\Delta H_f^\circ$ (in kcal/mol)	
	RHF	MP2	MP4	$\Delta_{(RHF-MP2)}$	$\Delta_{(MP2-MP4)}$
Di- <i>t</i> -butylmethane	-56.10	-57.06	-57.28	0.96	0.22
Di- <i>t</i> -butylpropane (C <sub>2</sub> )	-52.71	-60.17	<b>-59.22<sup>a</sup></b>	7.46	<b>-0.95<sup>a</sup></b>
Di- <i>t</i> -butylpropane (C <sub>2v</sub> )	-47.86	-55.28	-54.36	7.42	-0.92
Tri- <i>t</i> -butylmethane (C <sub>3</sub> )	-41.62	-53.14	<b>-51.69<sup>a</sup></b>	11.52	<b>-1.45<sup>a</sup></b>
Tri- <i>t</i> -butylmethane (C <sub>3v</sub> )	-28.46	-39.77	-38.40	11.31	-1.37

<sup>a</sup>Value estimated as described in footnote (a), Table 5.

positive than) the MM4 value, and 4.71 kcal/mol above the experimental value.

Then looking at the B3LYP value for this molecule, it is some 6 kcal/mol more negative than the Hartree-Fock value, but still not nearly negative enough. But the B3LYP/D value comes down to within 1 kcal of MM4 (or 3 kcal of experiment). So the MM4, MP4, and B3LYP/D values all lie within the range of -51.69 to -54.06 kcal/mol, consistently more positive than the experimental value of -56.40. We believe that the experimental value (-56.40) is too negative, and that the average calculated value (-53.0 ± 1.3) is better.

Then we may look at tetra-*t*-butylmethane. We have no experimental value for the heat of formation of this molecule, and it seems unlikely that we will ever get one. At least, not by heat of combustion measurements in the usual way. The compound has been predicted to be stable enough to isolate at liquid nitrogen temperatures (strain energy 98.6 kcal/mol, more than the bond energy of a C-C bond), but not at room temperature [7]. Carrying out the heat of combustion at liquid nitrogen temperatures is a problem that appears too formidable to warrant the expenditure of effort that would be required to solve it. So let's look at the MM4 value as a reference point (Table 5). The Hartree-Fock value is far from where we might

have expected it to be, actually being a positive number, some 30 kcal/mol too high. The MP2 value drops some 34 kcal/mol, and is some 4 kcal/mol too low. In principle we could obtain an MP4 value here, but in practice we cannot at present because of computational limitations.

When we look at the Hartree-Fock value again (Table 5) and the B3LYP value in comparison, we note the latter is some 18 kcal/mol more negative. So it is well on its way to the MM4 value, but still not nearly good enough. The B3LYP/D method, on the other hand, gives a value that is within 1 kcal of the MM4 value.

Finally, having now established that MM4, MP4, and B3LYP/D calculations each give overall very good agreement between calculated and experimental values for a broad, general set of alkanes, we can look at those experimental values that we were skeptical of when the MM4 work was done, namely tri-*t*-butylmethane, cyclododecane, *tat*-perhydroanthracene, and dodecahedrane (Table 7). A few other compounds are also included here for which the deviation of the MM4 value from experiment was larger than usual.

For the first three of these compounds that are weighted zero (tri-*t*-butylmethane, cyclododecane, and *tat*-perhydroanthracene), we did not believe the experimental value could be

**Table 7.** Heats of Formation (kcal/mol) (Problem Compounds. Part II).

WT	Compound H <sub>f</sub> <sup>o</sup>	Exp	MM4	Δ	MP4	Δ	B3LYP/D	Δ
0	Tri- <i>t</i> -BuMethane	-56.40	-54.06	2.34	-51.69	4.71	-53.21	3.19
6	3-EtPentane	-45.25	-44.40	0.85	-45.09	0.16	-45.28	-0.03
0	Cyclododecane	-54.59	-53.49	1.10	-53.81	0.78	-53.38	1.21
3	<i>c</i> -Bicyclo[3.3.0]octane	-22.20	-22.72	-0.52	-22.83	-0.73	-22.93	-0.73
4	<i>t</i> -Bicyclo[3.3.0]octane	-15.92	-15.31	0.61	-14.88	1.06	-15.16	0.76
5	<i>cis</i> -Hydrindane	-30.41	-31.01	-0.60	-30.58	-0.17	-30.38	0.03
1	<i>Tst</i> -PerHanthracene	-58.12	-57.74	0.38	-57.28	0.84	-57.12	1.00
0	<i>tat</i> -PerHanthracene	-52.73	-50.27	2.46	-50.93	1.80	-50.60	2.13
4	Me <sub>4</sub> Adamantane	-67.15	-66.55	.60	-67.68	-0.53	-67.68	-0.53
0	Dodecahedrane	+18.2	+14.16	-4.04	–	–	+8.87	-9.13

correct, and we considered the MM4 value to be more accurate at the time that MM4 was formulated.<sup>7</sup> The MP4 and B3LYP/D calculations now are also in much better agreement with the MM4 values than with experiment in each case, confirming our suspicions.

Several of the compounds in Table 7 require no special comment. They simply show the range of the more serious disagreements between the various calculational methods and experiment that occur. But a few of compounds here do require comment.

The *cis*- and *trans*-isomers of bicyclo[3.3.0]octane are of interest. The MM4 values for the heats of formation disagree with experiment by  $-0.52$  and  $+0.61$  kcal/mol for the *cis*- and *trans*-isomers respectively. Since the numbers are of opposite sign, this means that the MM4 energy of isomerization is substantially different (by 1.13 kcal/mol) from the experimental value, even though the individual values agree to approximately within the expected experimental errors. But note that this energy difference by MP4 is 1.73 kcal/mol, the B3LYP/D difference is 1.49 kcal/mol, and the discrepancies are all in the same direction. This clearly indicates to us that the bulk of the error in the isomerization energy here is in the experiment, and suggests that the calculated values are to be preferred.

Another reasonably certain error occurs with *tat*-perhydroanthracene. Here the MM4, MP4, and B3LYP/D values are similarly all greater than experiment by 2.46, 1.80, and 2.13 kcal/mol, respectively. This indicates that the experimental value is too negative by about 2.13 kcal/mol, the average discrepancy relative to the three computational values.

The fourth compound weighted zero in Table 7 is dodecahedrane. An experimental value is now available,  $18 \pm 1$  kcal/mol,\* together with a range of (MP2) calculated values (16-21 kcal/mol) [7, 10b]. The previously quoted value<sup>7</sup> (+13.30 kcal/mol) was an estimate from Wiberg/Schleyer calculations which should be disregarded.

The overall agreement between the various calculations for the heat of formation of dodecahedrane is still poor. The calculated and experimental values currently available span the range from  $+8.87$  to  $+21$  kcal/mol. The experimental value thus suffers from some uncertainty of unknown magnitude. The MM4 value is less reliable than one might ordinarily expect from this method, because of the nature of the molecule. The parameters used for this molecule include some that are not very well known, a tertiary center between five-membered rings, and the planar five-membered ring, for example. Compounding the problem, they occur many times in the molecule, so that any errors are multiplied considerably. The other calculated values also contain various uncertainties. Unfortunately the computational demands for an MP4 calculation here cannot be met by our available facilities.

\*Dodecahedrane itself was not used for the combustion. The value was determined in a more convoluted way [20]. The  $\pm 1$  kcal is an experimental standard deviation. The accuracy of the remainder of the determination is unknown.

## References

- Benson, S. W. *Thermochemical Kinetics*, Wiley, New York, 1976.
- Burkert, U.; Allinger, N. L. *Molecular Mechanics*, American Chemical Society, Washington, D.C., 1982.
- Engler, E. M.; Andose, J. D.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1973**, *95*, 8005.
- Allinger, N. L.; Tribble, M. T.; Miller, M. A.; Wertz, D. H. *J. Am. Chem. Soc.* **1971**, *93*, 1637.
- Allinger, N. L.; Hirsch, J. A.; Miller, M. A.; Tyminski, I. J.; and Van-Catledge, F. A. *J. Am. Chem. Soc.* **1968**, *90*, 1199.
- (a) Cox, J. D.; Pilcher, G. *Thermochemistry of Organic and Organometallic Compounds*, Academic Press, London, 1970.  
(b) Pedley, J. B.; Naylor, R. D.; Kirby, S. P. *Thermochemical Data of Organic Compounds*, 2<sup>nd</sup> Edition, Chapman and Hall Ltd, 1986.  
(c) National Institute of Standards and Technology (NIST), [webbook.nist.gov/chemistry](http://webbook.nist.gov/chemistry).
- Allinger, N. L.; Chen, K.; Lii, J.-H. *J. Comput. Chem.* **1996**, *17*, 642.
- (a) Chen, K.-H.; Lii, J.-H.; Fan, Y.; Allinger, N. L. *J. Comput. Chem.* **2007**, *28*, 2391, and references therein.  
(b) Lii, J.-H. *J. Phys. Chem.* **2002**, *106*, 8667.
- Redfern, P. C.; Zapol, P.; Curtiss, L. A.; Raghavachari, K. *J. Phys. Chem. A* **2000**, *104*, 5850, and papers cited therein.
- (a) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*, John Wiley & Sons, Inc., New York, 1986.  
(b) Disch, R. L.; Schulman, J. M. *J. Phys. Chem.* **1996**, *100*, 3504.  
(c) Wheeler, S. E.; Houk, K. N.; Schleyer, P. v. R.; Allen, W. D. *J. Am. Chem. Soc.* **2009**, *131*, 2547.
- (a) Wiberg, K. *J. Comput. Chem.*, **5**, 197 (1984); Wiberg, K. *J. Org. Chem.* **1985**, *50*, 5285.  
(b) Ibrahim, M. R.; Schleyer, P. v. R. *J. Comput. Chem.* **1985**, *6*, 157.
- Allinger, N. L.; Schmitz, L. R.; Motoc, I.; Bender, C.; Labanowski, J. *J. Am. Chem. Soc.* **1992**, *114*, 2880. See especially the Supplementary Material.
- (a) Allinger, N. L.; Sakakibara, K.; Labanowski, J. *J. Phys. Chem.* **1995**, *99*, 9603.  
(b) Schmitz, L. R.; Chen, K.-H.; Labanowski, J.; Allinger, N. L. *J. Phys. Org. Chem.* **2001**, *14*, 90.
- (a) Aped, P.; Allinger, N. L. *J. Am. Chem. Soc.* **1992**, *114*, 1.  
(b) Chen, K.-H.; Allinger, N. L. *J. Mol. Struct. (Theochem)*, **2002**, *581*, 215.
- (a) Pitzer, K. S.; Gwinn, W. D. *J. Chem. Phys.* **1942**, *10*, 428.  
(b) Wertz, D. H.; Allinger, N. L. *Tetrahedron*, **1979**, *35*, 3.
- Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, Jr., T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Wallingford CT, 2004.

17. (a) Labanowski, J.; Schmitz, L. R.; Chen, K.-H.; Allinger, N. L. *J. Comput. Chem.* **1998**, *19*, 1421.  
(b) Schmitz, L. R.; Chen, K.-H.; Labanowski, J.; Allinger, N. L. *J. Phys. Org. Chem.* **2001**, *14*, 90.
18. (a) Kristyan, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175.  
(b) Grimme, S. *J. Chem. Phys.*, **124**, 034108 (2006).  
(c) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397.
19. Verevkin, S. P.; Nolke, M.; Beckhaus, H.-D.; Ruechardt, C. *J. Org. Chem.* **1997**, *62*, 4683.
20. (a) Beckhaus, H.-D.; Ruechardt, C.; Lagerwall, D. R.; Paquette, L. A.; Wahl, F.; Prinzbach, H. *J. Am. Chem. Soc.* **1995**, *117*, 8885.  
(b) Beckhaus, H.-D.; Ruechardt, C.; Lagerwall, D. R.; Paquette, L. A.; Wahl, F. Prinzbach, H. *J. Am. Chem. Soc.* **1994**, *116*, 11775.