

# A Method for Calculating the Heats of Formation of Medium-Sized and Large-Sized Molecules

Bing He<sup>1,2</sup>, Hongwei Zhou<sup>1,3\*</sup>, Fan Yang<sup>1</sup>, Wai-Kee Li<sup>4</sup>

<sup>1</sup>Molecular Design Institute, Chengdu Normal University, Chengdu, China

<sup>2</sup>State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University, and Collaborative Innovation Center for Biotherapy, Chengdu, China

<sup>3</sup>IPNL, UMR IN2P3-CNRS-UCBL 5822, Villeurbanne, France

<sup>4</sup>Department of Chemistry, The Chinese University of Hong Kong, Hong Kong, China

Email: \*[jcbzhou@sina.com](mailto:jcbzhou@sina.com)

Received 3 July 2015; accepted 25 July 2015; published 28 July 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

A calculation method for heats of formation (HOF, referred to as  $\Delta H_f$ ) based on the density functional theory (DFT) is presented in this work. Similar to Gaussian-3 theory, the atomic scheme is applied to calculate the heats of formation of the molecules. In this method, we have modified the formula for calculation of Gaussian-3 theory in several ways, including the correction for diffuse functions and the correction for higher polarization functions. These corrections are found to be significant. The average absolute deviation from experiment for the 164 calculated heats of formation is about 1.9 kcal·mol<sup>-1</sup>, while average absolute deviation from G3MP2 for the 149 (among the 164 molecules, 15 large-sized molecules can not be calculated at the G3MP2 level) calculated heats of formation is only about 1.9 kcal·mol<sup>-1</sup>. It indicates that the present method can be applied to predict the heats of formation of medium-sized and large-sized molecules, while the heats of formation of these molecules using Gaussian-3 theory are much difficult, even impossible, to calculate. That is, this method provides a choice in the calculation of  $\Delta H_f$  for medium-sized and large-sized molecules.

## Keywords

Heats of Formation, Gaussian-3 Theory, Energy, Absolute Deviation

---

\*Corresponding author.

## 1. Introduction

Quantum chemical methods for the calculation of thermochemical data have been developed beyond the level of just reproducing experimental data and can now make accurate predictions where the experimental data are unknown or uncertain. The more accurate one in these methods is the Gaussian- $n$  theory [1]-[8], which has been widely used to estimate the heats of formation [7] [8] of small-sized molecules. For example, in an assessment [9] of Gaussian-3 (G3) theory on the 148 calculated heats of formation of neutral molecules, the average absolute deviation from experiment is less than 1 kcal·mol<sup>-1</sup>. This means that G3 theory can be used to predict heats of formation of molecules accurately. However, there are some deficiencies in G3 theory and its variation (commonly referred to as G3MP2 theory and G3B3 theory), such as, i) they can only be used to calculate the heats of formation of small-sized molecules, but become computationally intensive with the increasing number of atoms in molecules, and ii) there are large deviations for some molecules, especially for polynitrogen compounds, which are the potential candidates of high energy density materials. Especially, Gaussian-4 (G4) theory [8] and various modifications that recently come out show good accuracy for the calculation of heats of formations, the aforementioned deficiencies still exist.

The correlation method for calculation of heats of formation has drawn tremendous interest to find better ways to match the computational requirements of medium-sized and large-sized molecules, including isodesmic reaction schemes [9]-[14], group additive method, molecular mechanics and semiempirical methods [13] [15]-[17], and linear regression correction approach [18], etc. For the isodesmic reactions method, it is important to construct an appropriate bond separation reaction in which  $\Delta H_f$  for all components, except the target component, are known. A bond separation reaction is a reaction which breaks down any molecule composed of three or more heavy atoms, and which can be represented in classical valence structure, into its simplest set of two heavy atom molecules containing the same type of bond, *i.e.* the number and types of all bonds are retained. Sometimes this approach is very difficult. Of cause, it does not incorporate the energy stabilization effect caused by conjugate bonds in polyene or aromatic compounds. For group additive method, molecular mechanics, semiempirical methods [13] [15]-[17] and linear regression correction approach [18], the results are strongly dependent on the parameters used and thus are less reliable because they are all parameterized methods. For example, the thermochemical parameters can be obtained easily by the semiempirical methods, but the heats of formation based on these parameters are either underestimated or overestimated. The deviations are so large that a set of terms are introduced to correct the heats of formation in agreement with experimental values. So, semiempirical methods cannot be used to predict heats of formation of compounds if the experimental data are unknown.

*Ab initio* MO method and density functional theory (DFT), on the other hand, are independent on the experimental results and parameters, and have emerged as a very reliable method to calculate geometries, energies, and frequencies of molecules. Hence, they have been used to evaluate the  $\Delta H_f$  of interested molecules [15] [16] [19] [20]. Dunning's correlation consistent basis sets [21]-[23] (cc-pV\*Z, where \* denotes double, triple, quadruple, quintuple-zeta and sextuple-zeta, respectively) have the redundant functions removed and have been rotated [24] in order to increase computational efficiency. By combining the DFT with cc-pVDZ, the calculation results will be reliable. However, DFT/cc-pVDZ calculations do not produce  $\Delta H_f$  directly, so special model reactions have to be designed to derive the  $\Delta H_f$  (referred to as DFT  $\Delta H_f$ ) from the calculated total energy and vibrational analysis results [25]-[27]. This is also the goal we will pursue.

Our objective is to develop a procedure applicable to any molecular system in an unambiguous manner, which can reproduce experimental data to an accuracy of about of 2 kcal·mol<sup>-1</sup> even to species having larger experimental uncertainty. Recently, we have investigated the relative stabilities of N<sub>2n</sub> (N<sub>6</sub> (D<sub>3h</sub>), N<sub>8</sub> (O<sub>h</sub>), N<sub>10</sub> (D<sub>5h</sub>), N<sub>12</sub> (D<sub>6h</sub>), N<sub>12</sub> (D<sub>3d</sub>), N<sub>16</sub> (D<sub>4d</sub>), N<sub>18</sub> (D<sub>3h</sub>), N<sub>20</sub> (I<sub>h</sub>), N<sub>24</sub> (D<sub>3d</sub>), N<sub>24</sub> (D<sub>4h</sub>), N<sub>24</sub> (D<sub>6d</sub>), N<sub>30</sub> (D<sub>3h</sub>), N<sub>30</sub> (D<sub>5h</sub>), N<sub>32</sub> (D<sub>4d</sub>), N<sub>36</sub> (D<sub>3d</sub>), N<sub>40</sub> (D<sub>4h</sub>), N<sub>42</sub> (D<sub>3h</sub>), N<sub>48</sub> (D<sub>4d</sub>), N<sub>48</sub> (D<sub>3d</sub>), N<sub>54</sub> (D<sub>3h</sub>), N<sub>56</sub> (D<sub>4h</sub>), N<sub>60</sub> (D<sub>3d</sub>) and N<sub>72</sub>(D<sub>3d</sub>)) [28] [29] molecules at B3LYP/cc-pVDZ. As the potential candidates of high energy density materials, one important issue is to calculate the  $\Delta H_f$  of the molecules. However, the calculations of  $\Delta H_f$  of the molecules from N<sub>16</sub> to N<sub>72</sub> are very difficult, even impossible using Gaussian- $n$  theory because these molecules are medium-sized or large-sized and the experimental energies have not been well established. Furthermore, we found that Gaussian- $n$  theory performed poorly on the polynitrogen compounds (about 2 kcal·mol<sup>-1</sup> for each nitrogen atom in the molecules). In such case, the computational method for heats of formation based on DFT (referred to as DFT method) was conceived as the first in a series of well defined methods that could be routinely applied to the calculation of molecular energies of these medium-sized and large-sized molecules in a systematic manner and indeed, the results agreed with experimental values and so were reliable.

## 2. Theoretical and Computational Method

For the reaction Reactants  $\rightarrow$  Product:

The heats of formation at 298 K ( $\Delta H_f$ ) can be calculated by Equation (1).

$$\Delta H_f = H_{\text{rxn}} + \Delta H_{\text{exp},0} + \Delta H_m - \Delta H_{\text{atom}} \quad (1)$$

where  $H_{\text{rxn}} = E_{\text{product},0} - \Sigma E_{\text{atom},0}$ ;

$H_{\text{exp},0} = \Sigma H_{\text{atom},0}$ , which can be obtained from Ref. [30];

$\Delta H_m = H_{\text{product},0} - \Sigma E_{\text{product},0}$ ;

$\Delta H_{\text{atom}} = \Sigma H_{\text{atom}}$ , which can be obtained from Ref. [30].

Thereof, terms  $H_{\text{exp},0}$  and  $\Delta H_{\text{atom}}$  in Equation (1) are constants for the specified product whatever calculation methods are used to obtain the thermodynamic data. While terms  $H_{\text{rxn}}$  and  $\Delta H_m$  vary with different computational levels.

Equation (1) is applied to calculate the  $\Delta H_f$  of a compound in G3 theory and G3MP2 theory (referred to as G3MP2  $\Delta H_f$ ), where total energy of the product ( $E_{\text{product},0}$ ) and total energy of each atom of the reactants ( $\Sigma E_{\text{atom},0}$ ) are referred to as ‘‘G3 (0 K)’’ or and ‘‘G3MP2 (0 K)’’. G3 (0 K) or G3MP2 (0 K) are modified by a series of corrections (referred to as  $E_c$ ) from additional calculations, including a correction for diffuse functions [9] [10]

$$\Delta E(+)=E[\text{MP4/6-31+G(d)}]-E[\text{MP4/6-31G(d)}] \quad (2)$$

and a correction for higher polarization functions on nonhydrogen atoms and  $p$ -functions on hydrogens, [9] [10] etc.

$$\Delta E(2\text{df,p})=E[\text{MP4/6-31G(2df,p)}]-E[\text{MP4/6-31G(d)}]. \quad (3)$$

It can be found that the key issue is to obtain  $E_{\text{product},0}$  and  $E_{\text{atom},0}$ . In our work, only the total energy at the level B3LYP/cc-pVDZ can be obtained. Similar to the G3 theory and G3MP2 theory, the total energy at the level B3LYP/cc-pVDZ is modified by a correction for diffuse functions

$$\Delta E(+)=E[\text{B3LYP/6-31+G(d)}]-E[\text{B3LYP/6-31G(d)}] \quad (4)$$

and a correction for higher polarization functions on nonhydrogen atoms and  $p$ -functions on hydrogens.

$$\Delta E(2\text{df,p})=E[\text{B3LYP/6-31G(2df,p)}]-E[\text{B3LYP/6-31G(d)}]. \quad (5)$$

Comparing to the 6-31G (d) basis set, the cc-pVDZ basis set has the redundant functions removed. So, the corrected total energy is described as

$$E_0(\text{DFT})=E[\text{B3LYP/cc-pVDZ}]-\Delta E(+)-\Delta E(2\text{df,p}) \quad (6)$$

where  $E_0$  (DFT) is the energy of each atom of the reactants that Equation (1) requires. The correction energy is defined as  $E_c$ , which can be written as

$$E_c=\Delta E(+)+\Delta E(2\text{df,p}). \quad (7)$$

Note that for H (Hydrogen) to O (Oxygen) atoms,  $\Delta E(+)$  will be removed from  $E_0$  (DFT), for fluorine atom, and  $\Delta E(2\text{df,p})$  will be removed from  $E_0$  (DFT). According to the above corrections,  $E_c$  for the first and second row atoms are listed in **Table 1**.

A number of deficiencies in the method should be noted and future developments to alleviate them are proposed. In particular, this method works poorly on dissociation energies of ionic molecules such as LiF, on inorganic molecules such as  $\text{CO}_2$  (5.6 kcal·mol<sup>-1</sup> too low),  $\text{NH}_3$  (3.7 kcal·mol<sup>-1</sup> too large). Also, it works poorly on the hypervalent species, such as -SO<sub>2</sub> group and -NO<sub>2</sub> group, where their energies are high by 19 - 21 kcal·mol<sup>-1</sup> for the -SO<sub>2</sub> group and low by 9 - 10 kcal·mol<sup>-1</sup> for the -NO<sub>2</sub> group. It was found that additional group corrections might reduce discrepancy so that experimental values could be fitted perfectly.

Now, the total energy and the enthalpy of the product can be obtained from quantum chemistry calculation

**Table 1.** The atomic energies of the first row and the second row.

Atom	$E_0$ (au)	$H_0$ (kcal·mol <sup>-1</sup> )	$H_m$ (kcal·mol <sup>-1</sup> )	$E_c$ (au)
H	-0.501258	51.63	1.01	-0.001393
He	-2.907054	0.00	0.00	-0.002036
Li	0.000000	37.69	1.10	-0.000211
Be	0.000000	76.48	0.46	-0.002312
B	-24.660873	136.20	0.29	-0.005346
C	-37.851975	169.98	0.25	-0.005383
N	-54.589136	112.53	1.04	-0.003555
O	-75.068499	58.99	1.04	-0.007213
F	-99.726602	18.47	1.05	-0.001885
Ne	-128.909463	0.00	0.00	-0.021749
Na	0.000000	25.69	1.54	-0.000032
Mg	0.000000	34.87	1.19	-0.001603
Al	-242.382859	78.23	1.08	-0.003215
Si	-289.388651	106.60	0.76	-0.013175
P	-341.276438	75.42	1.28	-0.003958
S	-398.125081	65.66	1.05	-0.006948
Cl	-460.158464	28.59	1.10	-0.006929
Ar	-527.542275	0.00	0.00	-0.006574

$E_0$ : energy of each atom of the reactants (au);  $H_0$ : the experimental heats of each atom of the reactants (kcal·mol<sup>-1</sup>);  $H_m$ : the correction value of the experimental heat of each atom of the reactants (kcal·mol<sup>-1</sup>);  $E_c$ : the correction energy (au).

directly. The  $\Delta H_{\text{exp},0}$  and  $\Delta H_{\text{atom}}$  can be obtained from correlative books [30]. The  $\Delta H_f$  of a molecule at the level B3LYP/cc-pVDZ can be calculated by Equation (1) via Equation (6).

### 3. Results and Discussion

In this work, 164 compounds are selected for testing. They are divided into four test sets: i) **G2/97** test set, ii) **CH** test set, iii) **NOS** test set, and iv) **LARGE** test set.

#### 3.1. G2/97 Test Set

There are 70 neutral molecules in this test set. The structures are taken from Ref. [9]. All calculations are carried out using the GAUSSIAN 98 program package [31]. Density Function Theory has been applied to optimize the structures at basis set cc-pVDZ. The basis sets are the correlation-consistent basis sets of Dunning, specifically the polarized valence double- $\zeta$  (cc-pVDZ). The convergence criterion is  $10^{-8}$ . The optimized structures of the 70 species at the level B3LYP/cc-pVDZ and G3MP2 are shown in **Table 2**. The harmonic vibrational frequencies have been predicted in these optimized structures. All the vibrational frequencies of the molecules both at the levels B3LYP/cc-pVDZ and G3MP2 are positive (not listed). This indicates that the molecules are at a local minimum at the levels B3LYP/cc-pVDZ and G3MP2.

In **Table 2**, the experimental  $\Delta H_f$  (Exp. column) are taken from Ref. [9]. Some values have been updated by values from Ref. [30], such as the experimental value for **02** is changed from  $-118.4$  kcal·mol<sup>-1</sup> to  $-119.4$  kcal·mol<sup>-1</sup>, for **48** from  $8.9$  kcal·mol<sup>-1</sup> to  $5.0$  kcal·mol<sup>-1</sup>, etc.

It can be found that the DFT  $\Delta H_f$  deviations from experiment in some molecules are comparatively high (It is noted that the absolute  $\Delta H_f$  deviations which are greater than  $2.5$  kcal·mol<sup>-1</sup> are in bold and italic in **Table 2** and the subsequent tables): **01** ( $-7.1$  kcal·mol<sup>-1</sup>), **12** ( $-2.5$  kcal·mol<sup>-1</sup>), **13** ( $-3.1$  kcal·mol<sup>-1</sup>), **31** ( $-6.0$  kcal·mol<sup>-1</sup>), **37** ( $-2.5$  kcal·mol<sup>-1</sup>), **41** ( $-2.6$  kcal·mol<sup>-1</sup>), **50** ( $-2.7$  kcal·mol<sup>-1</sup>), **51** ( $-3.1$  kcal·mol<sup>-1</sup>), **53** ( $-4.9$  kcal·mol<sup>-1</sup>), **56**

**Table 2.** The  $\Delta H_f$  and the deviations from experiment of the 70 selected molecules of the **G2/97** test set. All are in kcal·mol<sup>-1</sup>.

No.	Mol.	Exp.	G3MP2	DFT	G3Dev	DFTDev
01	C <sub>2</sub> F <sub>4</sub>	-158.0	-161.1	-165.1	-3.1	-7.1
02	C <sub>2</sub> Cl <sub>4</sub>	-3.0	-7.4	-2.5	-4.4	0.5
03	CF <sub>3</sub> CN	-119.4	-118.8	-121.1	0.6	-1.7
04	CH <sub>3</sub> CCH (propyne)	44.2	44.5	46.3	0.3	2.1
05	CH <sub>2</sub> =C=CH <sub>2</sub> (allene)	45.5	44.6	43.2	-0.9	-2.3
06	C <sub>3</sub> H <sub>4</sub> (cyclopropene)	66.2	68.4	68.1	2.2	1.9
07	CH <sub>3</sub> CH=CH <sub>2</sub> (propylene)	4.8	4.9	4.3	0.1	-0.5
08	C <sub>3</sub> H <sub>6</sub> (cyclopropane)	12.7	14.1	13.5	1.4	0.8
09	C <sub>3</sub> H <sub>8</sub> (propane)	-25.0	-24.5	-25.4	0.5	-0.4
10	CH <sub>2</sub> CHCHCH <sub>2</sub> (butadiene)	26.3	29.1	29.3	2.8	3.0
11	C <sub>4</sub> H <sub>6</sub> (2-butyne)	34.8	35.5	35.9	0.7	1.1
12	C <sub>4</sub> H <sub>6</sub> (methylene cyclopropane)	47.9	46.5	45.4	-1.4	-2.5
13	C <sub>4</sub> H <sub>6</sub> (bicyclobutane)	51.9	54.8	55.0	2.9	3.1
14	C <sub>4</sub> H <sub>8</sub> (cyclobutane)	6.8	7.4	7.0	0.6	0.2
15	C <sub>4</sub> H <sub>10</sub> (trans butane)	-30.0	-29.6	-29.8	0.4	0.2
16	C <sub>5</sub> H <sub>8</sub> (spiropentane)	44.3	45.4	45.2	1.1	0.9
17	C <sub>6</sub> H <sub>6</sub> (benzene)	19.7	19.2	17.8	-0.5	-1.9
18	CH <sub>2</sub> F <sub>2</sub>	-107.7	-107.3	-106.6	0.4	1.1
19	C <sub>4</sub> H <sub>6</sub> (cyclobutene)	37.4	39.4	38.9	2.0	1.5
20	C <sub>4</sub> H <sub>8</sub> (isobutene)	-4.0	-3.8	-3.5	0.2	0.5
21	CHF <sub>3</sub>	-166.6	-165.9	-166.2	0.7	0.4
22	CH <sub>2</sub> Cl <sub>2</sub>	-22.8	-22.3	-23.4	0.5	-0.6
23	C <sub>4</sub> H <sub>10</sub> (isobutane)	-32.1	-31.5	-30.6	0.6	1.5
24	CH <sub>3</sub> Cl	-19.6	-19.0	-20.8	0.6	-1.2
25	CHCl <sub>3</sub>	-24.7	-24.7	-24.0	0.0	0.7
26	CH <sub>3</sub> NH <sub>2</sub> (methylamine)	-5.5	-3.5	-4.5	2.0	1.0
27	CH <sub>3</sub> CN (methyl cyanide)	18.0	18.5	17.0	0.5	-1.0
28	CH <sub>3</sub> NO <sub>2</sub> (nitromethane)	-17.8	-16.1	-17.5	1.7	0.3
29	CH <sub>3</sub> SiH <sub>3</sub> (methyl silane)	-7.0	-6.1	-8.9	0.9	-1.9
30	HCOOH (formic acid)	-90.5	-85.4	-89.4	5.1	1.1
31	HCOOCH <sub>3</sub> (methyl formate)	-85.0	-85.1	-91.0	-0.1	-6.0
32	CH <sub>3</sub> CONH <sub>2</sub> (acetamide)	-57.0	-54.3	-59.0	2.7	-2.0
33	C <sub>2</sub> H <sub>4</sub> NH (aziridine)	30.2	32.2	30.3	2.0	0.1
34	NCCN (cyanogen)	73.3	74.1	71.1	0.8	-2.2
35	(CH <sub>3</sub> ) <sub>2</sub> NH (dimethylamine)	-4.4	-2.4	-5.3	2.0	-0.9
36	CH <sub>3</sub> CH <sub>2</sub> NH <sub>2</sub> (trans ethylamine)	-11.3	-10.4	-11.0	0.9	0.3

## Continued

37	CH <sub>2</sub> O	-26.2	-25.9	-28.7	0.3	-2.5
38	CH <sub>3</sub> CHO (acetaldehyde)	-39.8	-38.9	-42.9	0.9	<b>-3.1</b>
39	HCOCOH (glyoxal)	-50.7	-46.0	-52.9	<b>4.7</b>	-2.2
40	CH <sub>3</sub> CH <sub>2</sub> OH (ethanol)	-56.2	-55.2	-55.1	1.0	1.1
41	CH <sub>3</sub> OCH <sub>3</sub> (dimethylether)	-44.0	-43.1	-46.6	0.9	<b>-2.6</b>
42	C <sub>2</sub> H <sub>4</sub> S (thiirane)	19.6	18.7	20.5	-0.9	0.9
43	(CH <sub>3</sub> ) <sub>2</sub> SO (dimethyl sulfoxide)	-36.2	-34.2	-35.7	2.0	0.5
44	C <sub>2</sub> H <sub>5</sub> SH (ethanethiol)	-11.1	-10.8	-9.9	0.3	1.2
45	CH <sub>3</sub> SCH <sub>3</sub> (dimethyl sulfide)	-8.9	-8.9	-7.4	0.0	1.5
46	CH <sub>2</sub> =CHF (vinyl fluoride)	-33.2	-34.0	-34.8	-0.8	-1.6
47	C <sub>2</sub> H <sub>5</sub> Cl (ethyl chloride)	-26.8	-26.3	-27.6	0.5	-0.8
48	CH <sub>2</sub> =CHCl (vinyl chloride)	5.0	5.1	4.3	0.1	-0.7
49	CH <sub>2</sub> =CHCN (acrylonitrile)	43.2	45.0	43.7	1.8	0.5
50	CH <sub>3</sub> COCH <sub>3</sub> (acetone)	-51.9	-50.9	-54.6	1.0	<b>-2.7</b>
51	CH <sub>3</sub> COOH (acetic acid)	-103.4	-101.9	-106.5	1.5	<b>-3.1</b>
52	CH <sub>3</sub> COF (acetyl fluoride)	-106.7	-104.6	-108.5	2.1	-1.8
53	CH <sub>3</sub> COCl (acetyl chloride)	-58.0	-57.4	-62.9	0.6	<b>-4.9</b>
54	CH <sub>3</sub> CH <sub>2</sub> CH <sub>2</sub> Cl (propyl chloride)	-31.5	-31.4	-31.8	0.1	-0.3
55	(CH <sub>3</sub> ) <sub>2</sub> CHOH (isopropanol)	-65.2	-64.5	-63.5	0.7	1.7
56	C <sub>2</sub> H <sub>5</sub> OCH <sub>3</sub> (methyl ethyl ether)	-51.7	-51.4	-54.4	0.3	<b>-2.7</b>
57	(CH <sub>3</sub> ) <sub>3</sub> N (trimethylamine)	-5.7	-4.6	-7.4	1.1	-1.7
58	C <sub>4</sub> H <sub>4</sub> O (furan)	-8.3	-7.7	-11.6	0.6	<b>-3.3</b>
59	C <sub>4</sub> H <sub>4</sub> S (thiophene)	27.5	26.1	31.6	-1.4	<b>4.1</b>
60	C <sub>4</sub> H <sub>5</sub> N (pyrrole)	25.9	27.0	24.0	1.1	-1.9
61	H <sub>2</sub> S	-4.9	-4.9	-5.6	0.0	-0.7
62	CH <sub>4</sub>	-17.9	-17.3	-18.2	0.6	-0.3
63	HCN	31.5	31.8	31.7	0.3	0.2
64	CO	-26.4	-26.8	-26.6	-0.4	-0.2
65	HCO	-26.0	-25.9	-28.7	0.1	<b>-2.7</b>
66	ClH	-22.1	-21.8	-22.8	0.3	-0.7
67	H <sub>3</sub> COH	-48.0	-47.1	-47.0	0.9	1.0
68	C <sub>2</sub> H <sub>4</sub>	12.5	12.5	12.4	0.0	-0.1
69	C <sub>2</sub> H <sub>6</sub>	-20.1	-19.5	-21.0	0.6	-0.9
70	H <sub>2</sub> NNH <sub>2</sub>	22.8	25.9	24.2	<b>3.1</b>	1.4

Exp.: experimental  $\Delta H_f$  taken from Ref. [30]; G3MP2:  $\Delta H_f$  obtained at the level G3MP2; DFT:  $\Delta H_f$  obtained at the level B3LYP/cc-pVDZ; G3Dev: G3MP2  $\Delta H_f$  deviation from experiment; DFTDev: DFT  $\Delta H_f$  deviation from experiment. Deviations which exceed 2.5 are in bold and italic.

( $-2.7 \text{ kcal}\cdot\text{mol}^{-1}$ ), **58** ( $-3.3 \text{ kcal}\cdot\text{mol}^{-1}$ ) and **65** ( $-2.7 \text{ kcal}\cdot\text{mol}^{-1}$ ). The **01** ( $\text{C}_2\text{F}_4$ ) is a halide. As known, Gaussian- $n$  theory and other method work poorly on this species. For example, the calculated enthalpy of formation of  $\text{C}_2\text{F}_4$  at G3 [9] is too negative by  $4.9 \text{ kcal}\cdot\text{mol}^{-1}$ , whereas at G3MP2 [10] is too negative by  $3.1 \text{ kcal}\cdot\text{mol}^{-1}$ . Our method works poorly on the molecules **05**, **10**, which contain cumulated double-bond ( $-\text{X}=\text{C}=\text{Y}-$ ) because the cumulated double-bond  $-\text{X}=\text{C}=\text{Y}-$  can also be written as  $>\text{X}-\text{C}\equiv\text{Y}$ . There should be different  $\Delta H_f$  between  $-\text{X}=\text{C}=\text{Y}-$  and  $>\text{X}-\text{C}\equiv\text{Y}$ . It can be found that the present method works poorly on the species which contain functional group  $>\text{C}=\text{O}$ . The calculated enthalpies of formation are underestimated too negative by 2.5 to  $5.6 \text{ kcal}\cdot\text{mol}^{-1}$ . The molecules **31**, **37**, **41**, **50**, **51**, **53**, **56**, **58** and **65** belong to this category. It can also be found that the present method works poorly on the inorganic species. The molecules **01**, **03** and **65** belong to this category. The sum of absolute deviation from experiment for the 70 calculated heats of formation is  $110.1 \text{ kcal}\cdot\text{mol}^{-1}$ . The average absolute deviation from experiment is about  $1.6 \text{ kcal}\cdot\text{mol}^{-1}$ .

The G3MP2  $\Delta H_f$  deviations of some molecules from experiment value are also comparatively high: **01** ( $-3.7 \text{ kcal}\cdot\text{mol}^{-1}$ ), **02** ( $-4.4 \text{ kcal}\cdot\text{mol}^{-1}$ ), **10** ( $2.8 \text{ kcal}\cdot\text{mol}^{-1}$ ), **13** ( $2.9 \text{ kcal}\cdot\text{mol}^{-1}$ ), **30** ( $5.1 \text{ kcal}\cdot\text{mol}^{-1}$ ), **32** ( $2.7 \text{ kcal}\cdot\text{mol}^{-1}$ ), **39** ( $4.7 \text{ kcal}\cdot\text{mol}^{-1}$ ) and **70** ( $3.1 \text{ kcal}\cdot\text{mol}^{-1}$ ). It can be found that G3MP2 does poorly on the halides, too. **01** and **02** belong to this category. G3MP2 also works poorly on the molecules which contain cumulated double-bond ( $-\text{X}=\text{C}=\text{Y}-$ ), **01** ( $-0.9 \text{ kcal}\cdot\text{mol}^{-1}$ ) and **10** belong to this category. Both DFT and G3MP2 work poorly on the bicyclobutane (**13** in Table 2). The sum of absolute deviation from experiment for the 70 calculated heats of formation is only  $78.6 \text{ kcal}\cdot\text{mol}^{-1}$ . The average absolute deviation from experiment is about  $1.1 \text{ kcal}\cdot\text{mol}^{-1}$ .

The G3MP2  $\Delta H_f$  deviations and the DFT  $\Delta H_f$  deviations from experiment value are shown in Figure 1. It can be found that the trends of the two lines are identical for the same molecule if the deviation is neglected. Most of the G3MP2  $\Delta H_f$  deviations from experiment are positive, while most of the DFT  $\Delta H_f$  deviations from experiment are negative.

It is noted that the molecule structures are taken from the original test set of G3 theory [9] (G2/97 test set), where a ‘‘higher level correction’’ (HLC) [9] is added to take into account some deficiencies in the energy calculations.

$$Ee(\text{G3}) = E(\text{combined}) + E(\text{HLC})$$

The HLC is  $-An_\beta - B(n_\alpha - n_\beta)$  for molecules and  $-Cn_\beta - D(n_\alpha - n_\beta)$  for atoms (including atomic ions). The  $n_\beta$  and  $n_\alpha$  are the number of  $\beta$  and  $\alpha$  valence electrons, respectively, with  $n_\alpha \geq n_\beta$ . The number of valence electron pairs corresponds to  $n_\beta$ . Thus,  $A$  is the correction for pairs of valence electrons in molecules,  $B$  is the correction for unpaired electrons in molecules,  $C$  is the correction for pairs of valence electrons in atoms, and  $D$  is the correction for unpaired electrons in atoms. The use of different corrections for atoms and molecules can be justified,

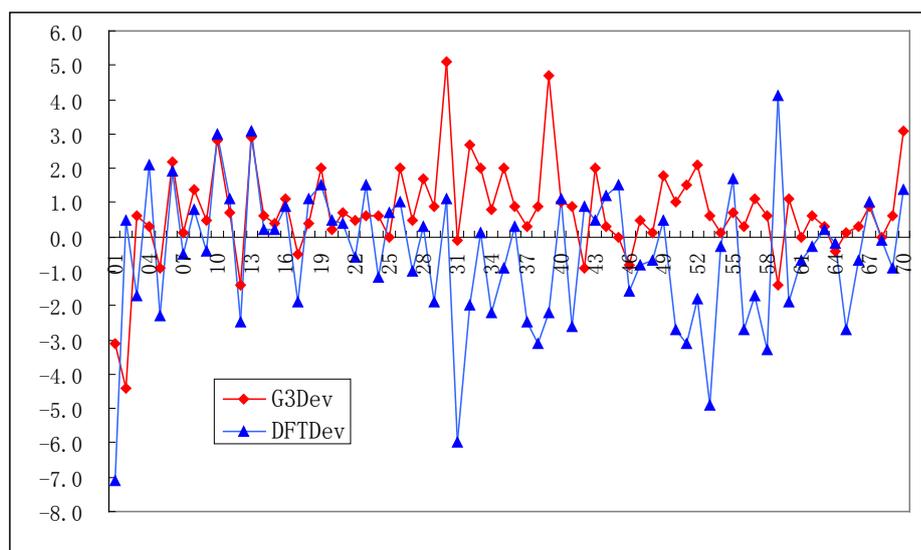


Figure 1. DFT  $\Delta H_f$  and G3MP2  $\Delta H_f$  deviations from experiment of the G2/97 test set.

in part, by noting that these extrapolations take some account of effects of basis functions with higher angular momentum, which are likely to be of more importance in molecules than in atoms. For G3 theory,  $A = 6.386$  mhartrees,  $B = 2.977$  mhartrees,  $C = 6.219$  mhartrees,  $D = 1.185$  mhartrees. The  $A$ ,  $B$ ,  $C$ ,  $D$  values are chosen to give the smallest average absolute deviation from experiment for the **G2/97** test set. Obviously,  $A$ ,  $B$ ,  $C$  and  $D$  are the fit parameters which are taken into account the electron structures of molecules in **G2/97** test set, and in turn, they are used to calculate the energies of molecules in the same test set. That is, the precisions for calculation energies, especially for the molecules in the test set, are improved by introducing the fit parameters  $A$ ,  $B$ ,  $C$ ,  $D$ . In this circumstances, it is not strange that the average absolute deviation of G3MP2  $\Delta H_f$  from experiment is less than that of DFT  $\Delta H_f$ .

### 3.2. CH Test Set

There are 20 neutral molecules which are all typical hydrocarbons in this test set. All calculations are carried out using the GAUSSIAN 98 program package. Density Function Theory (DFT) has been applied to optimize the structures at basis set cc-pVDZ. The optimized structures of the 20 species at the levels B3LYP/cc-pVDZ and G3MP2 are shown in **Table 3**. The harmonic vibrational frequencies have been predicted in these optimized structures. All the vibrational frequencies of the molecules both at the levels B3LYP/cc-pVDZ and G3MP2 are positive (not listed). This indicates that the molecules are at local minimum at the levels B3LYP/cc-pVDZ and G3MP2.

In **Table 3**, the experimental  $\Delta H_f$  (**Exp.** column) are taken from Ref [30]. It can be found that the DFT  $\Delta H_f$  deviations of some molecules from experiment value are comparatively large: **01** (3.7 kcal·mol<sup>-1</sup>), **15** (-4.5 kcal·mol<sup>-1</sup>) and **18** (4.0 kcal·mol<sup>-1</sup>). For **01** and **18**, both contain a functional group -C≡C-. It indicates that this method works poorly on the species. As is known, the isodesmic method for calculation does not incorporate the energy stabilization effect caused by conjugated bonds in polyene or aromatic compounds. It can also be found that the present method works poorly on polyene or aromatic species. The molecules **02**, **03**, **05**, **15**, **18** belong to conjugated category. The sum of absolute deviation from experiment for the 20 calculated heats of formation is 36.3 kcal·mol<sup>-1</sup>. The average absolute deviation from experiment is about 1.8 kcal·mol<sup>-1</sup>.

In this test set, the G3MP2  $\Delta H_f$  deviations of some molecules from experiment are also comparatively high: **02** (-4.2 kcal·mol<sup>-1</sup>), **04** (-3.3 kcal·mol<sup>-1</sup>), **13** (-3.3 kcal·mol<sup>-1</sup>), **15** (-5.7 kcal·mol<sup>-1</sup>), **16** (3.0 kcal·mol<sup>-1</sup>), **19** (-5.6 kcal·mol<sup>-1</sup>) and **20** (-3.8 kcal·mol<sup>-1</sup>). These results show that G3MP2 theory, which is known as the isodesmic method, for calculation does not incorporate the energy stabilization effect caused by conjugated bonds in polyene or aromatic compounds. Whereas **02**, **04**, **013**, **15**, **16**, **19**, **20** belong to conjugated category. It can be found that the number of G3MP2  $\Delta H_f$  deviations is more than that of the DFT  $\Delta H_f$  deviations. And comparing the G3MP2  $\Delta H_f$  deviations and the DFT  $\Delta H_f$  deviations, one can find that the former is higher than that of the later. The sum of absolute deviation from experiment for the 20 calculated heats of formation is 45.4 kcal·mol<sup>-1</sup>. The average absolute deviation is about 2.3 kcal·mol<sup>-1</sup>.

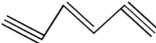
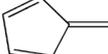
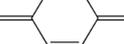
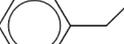
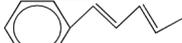
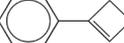
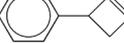
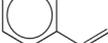
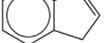
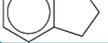
The G3MP2  $\Delta H_f$  deviations and the DFT  $\Delta H_f$  deviations from experiment are shown in **Figure 2**. It can be found that the trends of the two lines are identical for the same molecule if the deviation sign is neglected. Most of the G3MP2  $\Delta H_f$  deviations from experiment are negative, while most of the DFT  $\Delta H_f$  deviations from experiment are possibly negative or positive. From the view of point of average absolute deviation from experiment, DFT  $\Delta H_f$  method is more preferable than the G3MP2  $\Delta H_f$  method in this test set.

### 3.3. NOS Test Set

There are 60 neutral molecules in this test set. All calculations are carried out using the GAUSSIAN 98 program package. Density Function Theory (DFT) has been applied to optimize the structures at basis set cc-pVDZ. The optimized structures of the 60 species at the levels B3LYP/cc-pVDZ and G3MP2 are shown in **Table 4**. The harmonic vibrational frequencies have been predicted in these optimized structures. All the vibrational frequencies of the molecules both at the level B3LYP/cc-pVDZ and G3MP2 are positive (not listed). This indicates that the molecules are at local minimum at B3LYP/cc-pVDZ and G3MP2.

In **Table 4**, the experimental  $\Delta H_f$  (**Exp.** column) are taken from Ref [30]. In this test set, we selected some typical molecules which contain special functional groups, such as -NO<sub>2</sub>, -SO<sub>2</sub>, -X=C=Y-, etc. The calculation results show that each -NO<sub>2</sub> group may be low by 9.6 kcal·mol<sup>-1</sup>, and each -SO<sub>2</sub> group may be high by 20.0 kcal·mol<sup>-1</sup>. In order to fit for the experimental values, 9.6 kcal·mol<sup>-1</sup> is added for the DFT  $\Delta H_f$  for each -NO<sub>2</sub>

**Table 3.** The  $\Delta H_f$  and the deviations from experiment of the 20 molecules of the CH test set. All are in kcal·mol<sup>-1</sup>.

No.	Mol.	Exp.	G3MP2	DFT	G3Dev	DFTDev
01		54.5	54.0	58.2	-0.5	3.7
02		73.0	68.8	70.6	-4.2	-2.4
03		128.0	129.5	130.2	1.5	2.2
04		85.0	81.7	82.8	-3.3	-2.2
05		90.0	88.4	87.9	-1.6	-2.1
06		19.8	19.2	17.8	-0.6	-2.0
07		53.5	50.8	52.2	-2.7	-1.3
08		80.4	80.7	80.9	0.3	0.5
09		135.7	134.6	136.8	-1.1	1.1
10		53.0	51.4	53.5	-1.6	0.5
11		35.3	34.3	34.5	-1.0	-0.8
12		48.0	46.9	47.1	-1.1	-0.9
13		150.0	146.7	150.9	-3.3	0.9
14		49.0	47.0	47.0	-2.0	-2.0
15		66.0	60.3	61.5	-5.7	-4.5
16		60.0	57.0	58.6	-3.0	-1.4
17		64.0	62.0	65.7	-2.0	1.7
18		28.0	28.5	32.0	0.5	4.0
19		33.0	27.4	31.8	-5.6	-1.2
20		36.0	32.2	35.1	-3.8	-0.9

Exp.: experimental  $\Delta H_f$  taken from Ref. [30]; G3MP2:  $\Delta H_f$  obtained at the level G3MP2; DFT:  $\Delta H_f$  obtained at the level B3LYP/cc-pVDZ; G3Dev: G3MP2  $\Delta H_f$  deviation from experiment; DFTDev: DFT  $\Delta H_f$  deviation from experiment.

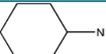
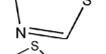
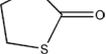
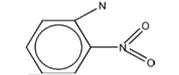
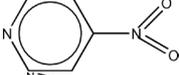
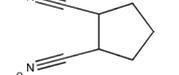
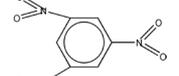
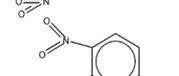
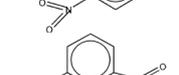
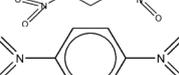
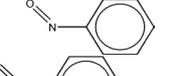
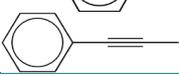
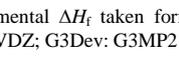
group a molecule contains, and 20.0 kcal·mol<sup>-1</sup> is subtracted from the DFT  $\Delta H_f$  for each -SO<sub>2</sub> group a molecule contains. The listed DFT  $\Delta H_f$  values in **Table 4** are corrected by the two values, 9.6 kcal·mol<sup>-1</sup> and 20.0 kcal·mol<sup>-1</sup>.

In **Table 4**, it can be found that the DFT  $\Delta H_f$  deviations of some molecules from experiment are comparatively large: **03** (3.7 kcal·mol<sup>-1</sup>), **07** (-4.0 kcal·mol<sup>-1</sup>), **08** (-6.3 kcal·mol<sup>-1</sup>), **11** (-5.7 kcal·mol<sup>-1</sup>), **12** (-3.1 kcal·mol<sup>-1</sup>), **18** (-3.2 kcal·mol<sup>-1</sup>), **23** (3.3 kcal·mol<sup>-1</sup>), **25** (-4.0 kcal·mol<sup>-1</sup>), **27** (-4.1 kcal·mol<sup>-1</sup>), **28** (-3.2 kcal·mol<sup>-1</sup>), **29** (-4.6 kcal·mol<sup>-1</sup>), **33** (3.5 kcal·mol<sup>-1</sup>), **34** (-2.7 kcal·mol<sup>-1</sup>), **41** (-2.7 kcal·mol<sup>-1</sup>), **42** (-2.7 kcal·mol<sup>-1</sup>), **48** (-3.4 kcal·mol<sup>-1</sup>), **50** (3.4 kcal·mol<sup>-1</sup>), **51** (-6.0 kcal·mol<sup>-1</sup>), **53** (-5.0 kcal·mol<sup>-1</sup>), **54** (-4.9

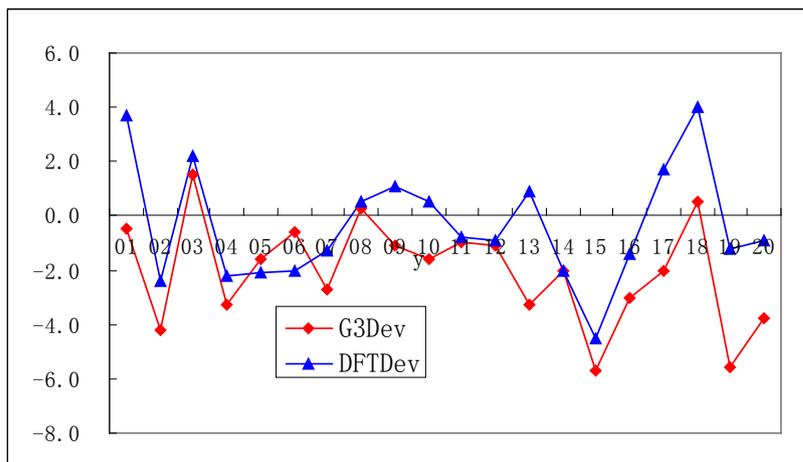
**Table 4.** The  $\Delta H_f$  and the deviations from experiment of the 60 molecules of the NOS test set. All are in kcal·mol<sup>-1</sup>.

No.	Mol.	Exp.	G3MP2	DFT	G3Dev	DFTDev
01	CH <sub>3</sub> NO <sub>2</sub>	-17.9	-15.6	-17.5	2.3	0.4
02	CH <sub>3</sub> ONO	-15.9	-13.3	-15.7	2.6	0.2
03	CH <sub>3</sub> ONO <sub>2</sub>	-29.5	-26.8	-25.8	2.7	3.7
04	(NH <sub>2</sub> ) <sub>2</sub> CS	4.5	6.0	2.9	1.5	-1.6
05	CH <sub>3</sub> SiH <sub>3</sub>	-8.0	-6.1	-8.9	1.9	-0.9
06		111.0	117.5	111.5	6.5	0.5
07	HOCCOOH	-175.7	-169.5	-179.7	6.2	-4.0
08	CH <sub>2</sub> =C=S	39.0	45.3	45.3	6.3	6.3
09	CH <sub>3</sub> NHNH <sub>2</sub>	22.7	27.8	24.0	5.1	1.3
10	CH <sub>3</sub> NCO	-31.0	-24.7	-32.6	6.3	-1.6
11	CH <sub>3</sub> NCS	31.0	29.7	25.3	-1.3	-5.7
12	CH <sub>3</sub> SCN	38.0	32.8	34.9	-5.2	-3.1
13		59.0	65.7	59.6	6.7	0.6
14		45.5	48.7	43.6	3.2	-1.9
15	NH <sub>2</sub> COCONH <sub>2</sub>	-95.0	-90.4	-96.5	4.6	-1.5
16		19.0	20.2	21.0	1.2	2.0
17		-7.0	-3.3	-7.2	3.7	-0.2
18	CH <sub>3</sub> COOH	-103.3	-101.9	-106.5	1.4	-3.2
19	S=C=S	19.9	18.7	20.4	-1.2	0.5
20	C <sub>2</sub> H <sub>5</sub> NO <sub>2</sub>	-24.4	-22.8	-25.1	1.6	-0.7
21	NH <sub>2</sub> CH <sub>2</sub> COOH	-92.0	-89.9	-94.2	2.1	-2.2
22	C <sub>2</sub> H <sub>5</sub> ONO	-25.0	-21.5	-23.7	3.5	1.3
23	C <sub>2</sub> H <sub>5</sub> ONO <sub>2</sub>	-36.8	-35.1	-33.5	1.7	3.3
24	(CH <sub>3</sub> ) <sub>2</sub> NNO	15.0	16.4	13.1	1.4	-1.9
25	(CH <sub>3</sub> S) <sub>2</sub>	-5.9	-5.8	-1.9	0.1	4.0
26	(CH <sub>3</sub> ) <sub>2</sub> SO	-36.2	-34.2	-35.7	2.0	0.5
27	(CH <sub>3</sub> ) <sub>2</sub> SO <sub>2</sub>	-89.0	-84.5	-84.9	4.5	4.1
28	(CH <sub>3</sub> O) <sub>2</sub> SO	-115.5	-106.3	-118.7	9.2	-3.2
29	(CH <sub>3</sub> O) <sub>2</sub> BH	-138.8	-134.3	-143.4	4.5	-4.6
30	(CH <sub>3</sub> ) <sub>2</sub> SiH <sub>2</sub>	-22.0	-20.7	-20.1	1.3	1.9
31	NCCN	73.3	74.5	71.1	1.2	-2.2
32	N≡C-S-S-C≡N	85.0	89.2	84.1	4.2	-0.9
33	CF <sub>2</sub> =C=CF <sub>2</sub>	-142.0	-131.8	-138.5	10.2	3.5
34	CH <sub>2</sub> (CN) <sub>2</sub>	63.5	62.1	60.8	-1.4	-2.7
35		-3.6	-8.3	-4.5	-4.7	-0.9

Continued

36		20.8	21.5	18.9	0.7	-1.9
37		40.0	37.5	41.3	-2.5	1.3
38		35.0	35.6	37.1	0.6	2.1
39		-29.0	-30.6	-27.1	-1.6	1.9
40		22.4	24.8	24.8	2.4	2.4
41	C <sub>2</sub> H <sub>5</sub> COOH	-107.5	-106.1	-110.2	1.4	-2.7
42	HCOOC <sub>2</sub> H <sub>5</sub>	-92.0	-88.3	-94.7	3.7	-2.7
43	n-C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub>	-29.7	-27.9	-29.1	1.8	0.6
44	i-C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub>	-33.2	-31.7	-31.9	1.5	1.3
45	n-C <sub>3</sub> H <sub>7</sub> ONO	-28.0	-26.6	-28.0	1.4	0.0
46	i-C <sub>3</sub> H <sub>7</sub> ONO	-32.0	-30.8	-31.4	1.2	0.6
47	B(CH <sub>3</sub> ) <sub>3</sub>	-28.0	-23.1	-27.9	4.9	0.1
48		12.5	-	9.1	0.0	-3.4
49		33.0	31.6	34.7	-1.4	1.7
50		96.0	93.8	99.4	-2.2	3.4
51		15.0	15.5	9.0	0.5	-6.0
52		21.0	21.6	19.6	0.6	-1.4
53		14.0	14.0	9.0	0.0	-5.0
54		14.0	14.0	9.1	0.0	-4.9
55		16.1	15.0	11.8	-1.1	-4.3
56		48.0	47.6	47.9	-0.4	-0.1
57		48.0	49.4	46.3	1.4	-1.7
58		52.0	50.9	49.6	-1.1	-2.4
59		85.0	84.9	83.7	-0.1	-1.3
60		64.0	65.0	66.0	1.0	2.0

Exp.: experimental  $\Delta H_f$  taken from Ref. [30]; G3MP2:  $\Delta H_f$  obtained at the level G3MP2; DFT:  $\Delta H_f$  obtained at the level B3LYP/cc-pVDZ; G3Dev: G3MP2  $\Delta H_f$  deviation from experiment; DFTDev: DFT  $\Delta H_f$  deviation from experiment.



**Figure 2.** DFT  $\Delta H_f$  and G3MP2  $\Delta H_f$  deviations from experiment of the CH test set.

kcal·mol<sup>-1</sup>) and **55** (-4.3 kcal·mol<sup>-1</sup>). Among these molecules, **03**, **23**, **48**, **51**, **53**, **54**, **55** contain the -NO<sub>2</sub> group, **27** and **28** contain the -SO<sub>2</sub> group, and **07**, **18**, **41**, **42** contain the -CO<sub>2</sub> group, while **08**, **11**, **12**, **33**, **34** contain the -X=C=Y- group. As mentioned above, the DFT  $\Delta H_f$  method works poorly on these species. The sum of absolute deviation from experiment for the 60 calculated heats of formation is 132.3 kcal·mol<sup>-1</sup>. The average absolute deviation is about 2.2 kcal·mol<sup>-1</sup>.

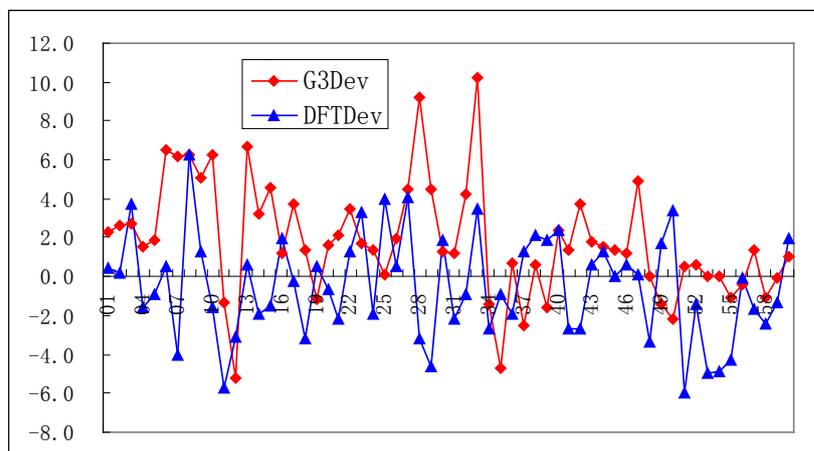
In this test set, the G3MP2  $\Delta H_f$  deviations of some molecules from experiment value are also comparative high: For the molecules contain the -NO<sub>2</sub> group, **02** (2.6 kcal·mol<sup>-1</sup>), **03** (2.7 kcal·mol<sup>-1</sup>) and **22** (3.5 kcal·mol<sup>-1</sup>); for the molecules contain the -SO<sub>2</sub> group, **17** (3.7 kcal·mol<sup>-1</sup>), **27** (4.5 kcal·mol<sup>-1</sup>), **28** (9.2 kcal·mol<sup>-1</sup>) and **32** (4.2 kcal·mol<sup>-1</sup>); for the molecules contain the -X=C=Y- group, **07** (6.2 kcal·mol<sup>-1</sup>), **08** (6.3 kcal·mol<sup>-1</sup>), **10** (6.3 kcal·mol<sup>-1</sup>), **12** (-5.2 kcal·mol<sup>-1</sup>), **33** (10.2 kcal·mol<sup>-1</sup>) and **35** (-4.7 kcal·mol<sup>-1</sup>); For the molecules contain the -CO<sub>2</sub> group, **07** (6.2 kcal·mol<sup>-1</sup>), **15** (4.6 kcal·mol<sup>-1</sup>) and **42** (3.7 kcal·mol<sup>-1</sup>). Furthermore, the G3MP2  $\Delta H_f$  deviations of polynitrogen compounds, **06** (6.5 kcal·mol<sup>-1</sup>), **09** (5.1 kcal·mol<sup>-1</sup>), **10** (6.3 kcal·mol<sup>-1</sup>), **13** (6.7 kcal·mol<sup>-1</sup>), **14** (3.2 kcal·mol<sup>-1</sup>), **15** (4.6 kcal·mol<sup>-1</sup>) and **32** (4.2 kcal·mol<sup>-1</sup>), and of boron compounds, **29** (4.5 kcal·mol<sup>-1</sup>) and **47** (4.9 kcal·mol<sup>-1</sup>), are high. These results show that G3MP2 theory works poorly on these species. The sum of absolute deviation from experiment for the 59 calculated heats of formation, wherein the molecule **48** cannot be calculated at G3MP2, is 157.0 kcal·mol<sup>-1</sup>. The average absolute deviation from experiment for the 59 calculated G3MP2  $\Delta H_f$  is 2.7 kcal·mol<sup>-1</sup>.

The G3MP2  $\Delta H_f$  deviations and the DFT  $\Delta H_f$  deviations from experiment value are shown in Figure 3. It can be found that most of the G3MP2  $\Delta H_f$  deviations from experiment value are positive, while most of the DFT  $\Delta H_f$  deviations from experiment value are possibly negative or positive. From the judgment of average absolute deviation from experiment value, the DFT  $\Delta H_f$  method is more preferable than that of G3MP2  $\Delta H_f$  method in the test set because the average absolute deviation from experiment of the DFT  $\Delta H_f$  is lower than that of the G3MP2  $\Delta H_f$ .

The sum of the absolute deviations from experiment is 278.7 for the above 150 calculated DFT  $\Delta H_f$ . While the sum of the absolute deviations from experiment is 281.0 for the above 149 calculated G3MP2  $\Delta H_f$ . Both of the average absolute deviations are about 1.9 kcal·mol<sup>-1</sup> (1.89 kcal·mol<sup>-1</sup> for G3MP2 theory, 1.86 kcal·mol<sup>-1</sup> for DFT method). The average absolute deviation of G3MP2 theory for the 70 molecules in **G2/97** test set is only 1.1 kcal·mol<sup>-1</sup>, while the average absolute deviations of the remaining two test sets are very high (2.3 kcal·mol<sup>-1</sup> for **CH** test set, and 2.7 kcal·mol<sup>-1</sup> for **NOS** test set) because the former is the original test set while the later are not. Whereas, the average absolute deviations of DFT method the results are from 1.6 kcal·mol<sup>-1</sup> to 2.2 kcal·mol<sup>-1</sup> for all the three test sets. By taking this into account, we can conclude that the DFT method is the same effective as the G3MP2 theory in predication of  $\Delta H_f$  of compounds.

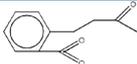
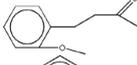
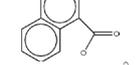
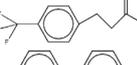
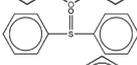
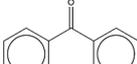
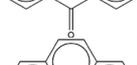
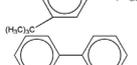
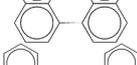
### 3.4. LARGE Test Set

There are 14 neutral molecules in this test set. All calculations are carried out using the GAUSSIAN 98 program package. DFT has been applied to optimize the structures at basis set cc-pVDZ. The optimized structures of the



**Figure 3.** DFT  $\Delta H_f$  and G3MP2  $\Delta H_f$  deviations from experiment of the NOS test set.

**Table 5.** The  $\Delta H_f$  and the deviations from experiment of the 14 molecules of the LARGE test set. All are in kcal·mol<sup>-1</sup>.

No.	Mol.	Mol.	Exp.	DFT	DFTDev
01	C <sub>10</sub> H <sub>9</sub> NO <sub>3</sub>		-10.0	-9.3	0.7
02	C <sub>10</sub> H <sub>10</sub> O <sub>3</sub>		-94.0	-98.9	-4.9
03	C <sub>11</sub> H <sub>8</sub> O <sub>2</sub>		-55.9	-57.6	-1.7
04	C <sub>11</sub> H <sub>9</sub> F <sub>3</sub> O		-168.0	-171.1	-3.1
05	C <sub>12</sub> H <sub>8</sub> F <sub>2</sub>		-48.0	-53.8	-5.8
06	C <sub>12</sub> H <sub>8</sub> O <sub>2</sub>		-15.0	-20.6	-5.6
07	C <sub>12</sub> H <sub>10</sub> OS		26.7	30.1	3.4
08	C <sub>13</sub> H <sub>9</sub> N		59.0	61.7	2.7
09	C <sub>14</sub> H <sub>8</sub> O <sub>2</sub>		-24.4	-26.1	-1.7
10	C <sub>14</sub> H <sub>10</sub>		49.2	49.2	0.0
11	C <sub>18</sub> H <sub>30</sub>		-32.6	-31.5	1.1
12	C <sub>20</sub> H <sub>12</sub>		75.0	78.3	3.3
13	C <sub>22</sub> H <sub>14</sub>		81.0	82.0	1.0
14	C <sub>24</sub> H <sub>12</sub>		77.0	75.5	-1.5

Exp.: experimental  $\Delta H_f$  taken from Ref. [30]; DFT:  $\Delta H_f$  obtained at the level B3LYP/cc-pVDZ; DFTDev: DFT  $\Delta H_f$  deviation from experiment.

14 species at the level B3LYP/cc-pVDZ are shown in **Table 5**. The harmonic vibrational frequencies have been predicted in these optimized structures. All the vibrational frequencies of the molecules at the level B3LYP/cc-pVDZ are positive (not listed). This indicates that the molecules are at local minimum at the level B3LYP/cc-pVDZ.

In **Table 5**, the experimental  $\Delta H_f$  (**Exp.** column) are taken from Ref. [30]. In this test set, we selected some medium-sized and large-sized molecules, of which the calculation of heats of formation of these molecules using G3 or G3MP2 theory is much difficult, even impossible.

From **Table 5**, it can be found that the DFT  $\Delta H_f$  deviations of some molecules from experiment value are comparative large: **02** ( $-4.9 \text{ kcal}\cdot\text{mol}^{-1}$ ), **04** ( $-3.1 \text{ kcal}\cdot\text{mol}^{-1}$ ), **05** ( $-5.8 \text{ kcal}\cdot\text{mol}^{-1}$ ), **06** ( $-5.6 \text{ kcal}\cdot\text{mol}^{-1}$ ), **07** ( $3.4 \text{ kcal}\cdot\text{mol}^{-1}$ ), **08** ( $2.7 \text{ kcal}\cdot\text{mol}^{-1}$ ) and **12** ( $3.3 \text{ kcal}\cdot\text{mol}^{-1}$ ). Among them, the deviations of the molecules **04** and **05** are mainly caused by the halogen atoms in the molecules. While the deviations of the molecules **02** and **06** are mainly caused by the  $-\text{CO}_2$  group. The sum of absolute deviation from experiment for the 14 calculated heats of formation is  $36.5 \text{ kcal}\cdot\text{mol}^{-1}$ . The average absolute deviation from experiment for the 14 calculated heats of formation is about  $2.6 \text{ kcal}\cdot\text{mol}^{-1}$ . It seems that the average absolute deviation is comparatively high in this test set. However, the high absolute deviation  $5.8 \text{ kcal}\cdot\text{mol}^{-1}$ , for example in 05, is acceptable because the molecules are the medium-sized and large-sized.

## 4. Conclusion

In this work, we have developed a method for calculating the heats of formation of medium-sized and large-sized molecules. This method has the following characteristics: i) The calculation formula for the heats of formation is derived from the famous G3 and G3MP2 theory. The atomic energies are obtained from the calculated results. There are no empirical parameters or fit parameters to be introduced to eliminate the deficiencies in the calculation of the heats of formation except the corrections of the chemical functional groups  $-\text{NO}_2$  and  $-\text{SO}_2$ . ii) The average absolute deviation from experiment for the 150 calculated DFT  $\Delta H_f$  is  $1.5 \text{ kcal}\cdot\text{mol}^{-1}$ . While the average absolute deviation from experiment for the 149 calculated G3MP2  $\Delta H_f$  is  $1.7 \text{ kcal}\cdot\text{mol}^{-1}$ . The average absolute deviation from experiment for the whole 164 calculated DFT  $\Delta H_f$  is also  $1.9 \text{ kcal}\cdot\text{mol}^{-1}$ . The G3MP2  $\Delta H_f$  and DFT  $\Delta H_f$  can be used to predict the heats of formation when the experimental data are unknown or uncertain. iii) The present method can be applied to predict the heats of formation of medium-sized and large-sized molecules. The heats of formation of a molecule containing 100 up to 200 heavy atoms can be calculated by this method. Under economical consideration, this method is expected to impact the applications in the calculations of heats of formation of large-sized molecules.

## References

- [1] Curtiss, L.A., Raghavachari, K., Trucks, G.W. and Pople, J.A. (1991) Gaussian-2 Theory for Molecular Energies of First- and Second-Row Compounds. *Journal of Chemical Physics*, **94**, 7221-7230. <http://dx.doi.org/10.1063/1.460205>
- [2] Curtiss, L.A. and Raghavachari, K. (1995) In: Langhoff, S.R., Ed., *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, Kluwer Academic, Netherlands, 139.
- [3] Pople, J.A., Head-Gordon, M., Fox, D.J., Raghavachari, K. and Curtiss, L.A. (1989) Gaussian-1 Theory: A General Procedure for Prediction of Molecular Energies. *Journal of Chemical Physics*, **90**, 5622-5629. <http://dx.doi.org/10.1063/1.456415>
- [4] Curtiss, L.A., Jones, C., Trucks, G.W., Raghavachari, K. and Pople, J.A. (1990) Gaussian-1 Theory of Molecular Energies for Second-Row Compounds. *Journal of Chemical Physics*, **93**, 2537-2545. <http://dx.doi.org/10.1063/1.458892>
- [5] Curtiss, L.A., Raghavachari, K., Redfern, P.C. and Pople, J.A. (1997) Assessment of Gaussian-2 and Density Functional Theories for the Computation of Enthalpies of Formation. *Journal of Chemical Physics*, **105**, 1063-1079. <http://dx.doi.org/10.1063/1.473182>
- [6] Curtiss, L.A., Redfern, P.C., Raghavachari, K. and Pople, J.A. (1998) Assessment of Gaussian-2 and Density Functional Theories for the Computation of Ionization Potentials and Electron Affinities. *Journal of Chemical Physics*, **109**, 42-55. <http://dx.doi.org/10.1063/1.476538>
- [7] Lau, C.-K., Li, W.-K., Wang, X., Tian, A.M. and Wong, N.B. (2002) A Gaussian-3 Study of  $\text{N}_7^+$  and  $\text{N}_7^-$  Isomers. *Journal of Molecular Structure (THEOCHEM)*, **617**, 121-131. [http://dx.doi.org/10.1016/S0166-1280\(02\)00411-6](http://dx.doi.org/10.1016/S0166-1280(02)00411-6)
- [8] Curtiss, L.A., Redfern, P.C. and Raghavachari, K. (2011) Gn Theory. *Wireless Communications & Mobile Computing*,

1, 810-825.

- [9] Curtiss, L.A., Raghavachari, K., Redfern, P.C., Rassolov, V. and Pople, J.A. (1998) Gaussian-3 (G3) Theory for Molecules Containing First and Second-Row Atoms. *Journal of Physical Chemistry*, **109**, 7764-7775. <http://dx.doi.org/10.1063/1.477422>
- [10] Curtiss, L.A., Redfern, P.C., Raghavachari, K., Rassolov, V. and Pople, J.A. (1999) Gaussian-3 Theory Using Reduced Moller-Plesset Order. *Journal of Chemical Physics*, **110**, 4703-4709. <http://dx.doi.org/10.1063/1.478385>
- [11] Haworth, N.L. and Bacskay, G.B. (2002) Heats of Formation of Phosphorus Compounds Determined by Current Methods of Computational Quantum chemistry. *The Journal of Chemical Physics*, **117**, 11175-11187. <http://dx.doi.org/10.1063/1.1521760>
- [12] Gong, X.D., Zhang, J. and Xiao, H.M. (1999) Studies on the Synthesis of (2S,3R)-3-Hydroxy-3-Methylproline via C-2-N Bond Formation. *Proceedings of the 26th International Pyrotechnics Seminar*, 136.
- [13] Chen, Z.X., Xiao, J.M., Xiao, H.M. and Chiu, Y.N. (1999) Studies on Heats of Formation for Tetrazole Derivatives with Density Functional Theory B3LYP Method. *The Journal of Chemical Physics*, **103**, 8062-8066. <http://dx.doi.org/10.1021/jp9903209>
- [14] Hehre, W.J. (1995) Practical Strategies for Electronic Structure Calculation. Wavefunction, Inc., Irvine, 102-134.
- [15] Xu, X.J., Xiao, H.M., Ma, X.F. and Ju, X.H. (2006) Looking for High-Energy Density Compounds among Hexaazadamantane Derivatives with Bond CN, Bond NC, and Bond ONO<sub>2</sub> Groups. *International Journal of Quantum Chemistry*, **106**, 1561-1568. <http://dx.doi.org/10.1002/qua.20909>
- [16] Wang, G.X., Gong, X.D. and Xiao, H.M. (2009) Theoretical Investigation on Density, Detonation Properties, and Pyrolysis Mechanism of Nitro Derivatives of Benzene and Aminobenzenes. *International Journal of Quantum Chemistry*, **109**, 1522-1530. <http://dx.doi.org/10.1002/qua.21967>
- [17] Ruzsinszky, A., van Alsenoy, C. and Csonka, G.I. (2002) Optimal Selection of Partial Charge Calculation Method for Rapid Estimation of Enthalpies of Formation from Hartree-Fock Total Energy. *The Journal of Physical Chemistry*, **106**, 12139-12150. <http://dx.doi.org/10.1021/jp026913s>
- [18] Duan, X.M., Song, G.L., Li, Z.H., Wang, X.J., Chen, G.H. and Fan, K.N. (2004) Accurate Prediction of Heat of Formation by Combining Hartree-Fock/Density Functional Theory Calculation with Linear Regression Correction Approach. *The Journal of Chemical Physics*, **121**, 7086-7095. <http://dx.doi.org/10.1063/1.1786582>
- [19] Jursic, B.S. (2003) Density Functional Calculation of the Heats of Formation for Various Aromatic Nitro Compounds. *Journal of Molecular Structure (THEOCHEM)*, **634**, 215-224. [http://dx.doi.org/10.1016/S0166-1280\(03\)00345-2](http://dx.doi.org/10.1016/S0166-1280(03)00345-2)
- [20] Chen, P.C., Chieh, Y.C. and Tzeng, S.C. (2000) Computing Heats of Formation for Cubane and Tetrahane with Density Functional Theory and Complete Basis Set *ab Initio* Methods. *Journal of Molecular Structure (THEOCHEM)*, **499**, 137-140. [http://dx.doi.org/10.1016/S0166-1280\(99\)00293-6](http://dx.doi.org/10.1016/S0166-1280(99)00293-6)
- [21] Dunning, T.H. (1989) Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *The Journal of Chemical Physics*, **90**, 1007-1023. <http://dx.doi.org/10.1063/1.456153>
- [22] Peterson, K.A., Woon, D.E. and Dunning Jr., T.H. (1994) Benchmark Calculations with Correlated Molecular Wave Functions. IV. The Classical Barrier Height of the H + H<sub>2</sub> → H<sub>2</sub> + H Reaction. *The Journal of Chemical Physics*, **100**, 7410-7415. <http://dx.doi.org/10.1063/1.466884>
- [23] Wilson, A., van Mourik, T. and Dunning Jr., T.H. (1997) Gaussian Basis Sets for Use in Correlated Molecular Calculations. VI Sextuple Zeta Correlation Consistent Basis Sets for Boron through Neon. *Journal of Molecular Structure (THEOCHEM)*, **388**, 339-349. [http://dx.doi.org/10.1016/S0166-1280\(96\)80048-0](http://dx.doi.org/10.1016/S0166-1280(96)80048-0)
- [24] Davidson, E.R. (1996) Comment on "Comment on Dunning's Correlation-Consistent Basis Sets". *Chemical Physics Letters*, **220**, 514-518. [http://dx.doi.org/10.1016/0009-2614\(96\)00917-7](http://dx.doi.org/10.1016/0009-2614(96)00917-7)
- [25] Berry, R.J., Burgess Jr., D.R.F., Nyden, M.R., Zacharian, M.R., Melius, C.F. and Schwarz, M. (1996) Halon Thermochemistry: Calculated Enthalpies of Formation of Chlorofluoromethanes. *The Journal of Physical Chemistry*, **100**, 7405-7410.
- [26] Raghavachari, K., Stefanov, B.B. and Curtiss, L.A. (1997) Accurate Thermochemistry for Larger Molecules: Gaussian-2 Theory with Bond Separation Energies. *The Journal of Chemical Physics*, **106**, 6764-6767. <http://dx.doi.org/10.1063/1.473659>
- [27] Baboul, A.G., Curtiss, L.A., Redfern, P.C. and Raghavachari, K. (1999) Gaussian-3 Theory Using Density Functional Geometries and Zero-Point Energies. *The Journal of Chemical Physics*, **110**, 7650-7657. <http://dx.doi.org/10.1063/1.478676>
- [28] Zhou, H.W., Wong, N.B., Zhou, G. and Tian, A.M. (2006) Theoretical Study on "Multilayer" Nitrogen Cages. *The Journal of Physical Chemistry A*, **110**, 3845-3852. <http://dx.doi.org/10.1021/jp056435w>
- [29] Zhou, H.W., Wong, N.B., Zhou, G. and Tian, A.M. (2006) What Makes the Cylinder-Shaped N<sub>72</sub> Cage Stable? *The*

*Journal of Physical Chemistry A*, **110**, 7441-7446. <http://dx.doi.org/10.1021/jp062214u>

- [30] Lias, S.G., Bartmess, J.E., Liebman, J.F., Holmes, J.L., Levin, R.D. and Mallard, W.G. (1988) Gas-Phase Ion and Neutral Thermochemistry. *Journal of Physical and Chemical Reference Data*, **17**.
- [31] Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G.A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H.P., Izmaylov, A.F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery Jr., J.A., Peralta, J.E., Ogliaro, F., Bearpark, M., Heyd, J.J., Brothers, E., Kudin, K.N., Staroverov, V.N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Rega, N., Millam, J.M., Klene, M., Knox, J.E., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Martin, R.L., Morokuma, K., Zakrzewski, V.G., Voth, G.A., Salvador, P., Dannenberg, J.J., Dapprich, S., Daniels, A.D., Farkas, Ö., Foresman, J.B., Ortiz, J.V., Cioslowski, J. and Fox, D.J. (2009) Gaussian 09. Revision C.01. Gaussian, Inc., Wallingford.