AB INITIO THERMOCHEMISTRY OF ATMOSPHERICALLY IMPORTANT SMALL MOLECULES AND RADICALS

Doctoral (Ph.D.) theses

BALÁZS NAGY

Supervisor:
DR. GYULA TASI
associate professor

Doctoral School of Chemistry

University of Szeged
Faculty of Science and Informatics
Department of Applied and Environmental Chemistry

Szeged
2013

1   Introduction and aims

During the last few decades, the ongoing race in thermochemistry between experiment and theory, with the aim to predict basic quantities to very high accuracy has been widely appreciated. One crucial quantity of central interest in this field is the enthalpy of formation of chemical species. Besides enthalpy, other parameters such as entropy, Gibbs energy, and the various thermodynamic functions for reactions are also of utmost importance in modeling atmospheric and combustion processes as well as designing and analyzing chemical reactions in the laboratory or in an industrial environment.

Among the numerous experimental techniques, it is spectroscopy that provides the most accurate values. Nevertheless, even the most sophisticated methods can have significant uncertainties and therefore the quality of their results is questionable. A further issue is the experimental treatment of radicals and molecules with short lifetime and/or high reactivity. Obviously, increasing the accuracy makes the measurements more and more difficult or, in extreme cases, it is even impossible to carry them out. Because of this only a few laboratories in the world have the opportunity the provide experimental thermodynamic quantities that can be used as reference values.

At the same time, state-of-the-art wavefunction-based quantum chemical methods, when pushed to the available technical limits, can have a similar accuracy to that achievable by spectroscopy. However, for polyatomic and polyelectronic systems the accuracy of the spectroscopic measurements is below 1 cm^{-1} which can hardly be matched by even the most sophisticated and time-consuming quantum chemical procedures. Nevertheless, this latter accuracy is rarely achieved in the experiments for quantities of thermochemical interest. In a computational point of view, the exact treatment of polyatomic and polyelectronic species may also be problematic. The increasing number of atoms in the studied molecule or radical, i.e., the increasing number of the electrons, involves the simplification of the computational approaches in our model. In particular, as the system increases, we have to decrease the level of theory of our calculations. In this manner, obviously, the associated uncertainties, i.e., the error bars of the results are inevitably increasing. Besides the above, it is fair to say that the high-level *ab initio* model chemistries applying the appropriate computational techniques can provide reliable and reference-quality thermochemical parameters.

The continuous and large-scale development of quantum chemistry and computer hardware technology seems to change the time-proven assumption, that it is experiment which provides reference data for quantum chemical approaches. In thermochemistry, as I show in my doctoral study, the situation is just reversed: it is quantum chemistry which provides thermodynamic quantities to higher accuracy.

Various thermochemical parameters for chemical species can be found either in extended spreadsheets built up sequentially or in active thermochemical tables following the conception of active databases. Unfortunately, well-defined uncertainties associated with the tabulated values are scarcely available and for many molecules and even families of molecules the reliability of their listed data is questionable. This fact makes the applicability of these values in high-accuracy calculations limited or even impossible.

The aim of my doctoral study was to calculate highly accurate thermochemical parameters for atmospherically important small radicals and molecules that do not have sufficiently precise

literature data. To this end, three composite *ab initio* model chemistries have been developed and applied. The primary purpose was to define such cost-effective thermochemical protocols which were more or less system specific but generally applicable to each member within a group of chemically similar molecules.

2 Computational methods

In my Ph.D. work three different composite thermochemical protocols have been used in order to calculate thermodynamic functions for the formaldehyde molecule (focal-point analysis, FPA approach), the derivatives of formaldehyde [HEAT-345(Q) model], and several sulfur-containing species [HEAT-345+d(Q) protocol]. The principal difference among these models lies in the levels of theory used to calculate the individual contributions.

In general, composite model chemistries are based on the assumption that the various contributions of the total energy reach their converged values at different levels of theory. For instance, compared to the Hartree-Fock (HF-SCF) part, a smaller basis set is sufficient for the calculation of a higher-order electron correlation contribution.

The FPA approach applied to formaldehyde is conceptually different form the two HEAT type protocols. In FPA all contributions to the total electronic energy are calculated at the highest possible level of theory in order to ensure their correct convergence. The model developed in this study uses CCSD(T)/aug-cc-pCVQZ equilibrium geometries and estimates the total non-relativistic energy as the sum of contributions obtained at the CCSD, CCSD(T), CCSDT, CCSDT(Q), CCSDTQ, and CCSDTQ(P) levels. In order to check the convergence of the CC contributions, explicitly correlated F12-CCSD(T) calculations were also performed. The non-relativistic Born-Oppenheimer total energy was further augmented with several additive corrections. On one hand, scalar relativistic effects were taken into consideration by averaging the results of the Douglas-Kroll-Hess approach with that of the mass-velocity and one-electron-Darwin term (MVD1) calculations. On the other hand, deficiencies of the Born-Oppenheimer approximation were corrected by calculating the diagonal Born-Oppenheimer correction (DBOC). Finally, harmonic and anharmonic contributions of the zero-point vibrational energy (ZPE) were also considered.

For the derivatives of formaldehyde can be obtained from the parent molecule by substituting and/or eliminating its hydrogen atoms the more or less fixed and recipe like HEAT-345(Q) model chemistry has been utilized. This protocol defines the total energy of the chemical species according to the following scheme:

$$E_{\text{TOT}} = E_{\text{HF}}^{\infty} + \Delta E_{\text{CCSD(T)}}^{\infty} + \Delta E_{\text{CCSDT}}^{\infty} + \Delta E_{\text{CCSDT(Q)}} + \Delta E_{\text{ZPE}} + \Delta E_{\text{DBOC}} + \Delta E_{\text{REL}}. \qquad (1)$$

The various contributions in the above equation, in their respective order, are as follows: *i)* HF-SCF energy extrapolated to the basis set limit; *ii-iv)* correlation contributions at the CCSD(T), CCSDT, and CCSDT(Q) levels of theory, respectively; *v)* harmonic and anharmonic ZPE calculated relying on vibrational perturbation theory (VPT2) at the CCSD(T) level; *vi)* DBOC correction determined at the CCSD level; and *vii)* contributions from the scalar relativistic effects calculated by evaluating the expectation value of the mass-velocity and one- and two-electron Darwin operators at the CCSD(T) level. The latter part also involves the spin-orbit correction of the carbon atom which, instead of graphite, was used as the reference state for carbon. The value of this effect was taken from experiment.

The HEAT-345+d(Q) protocol applied to the sulfur-containing species defines the total

molecular energy by analogy with its predecessor HEAT-345(Q) model:

$$E_{\text{TOT}} = E_{\text{HF}}^{\infty} + \Delta E_{\text{CCSD(T)}}^{\infty} + \Delta E_{\text{CCSDT(Q)}} + \Delta E_{\text{core}}^{\infty} + \Delta E_{\text{ZPE}} + \Delta E_{\text{DBOC}} + \Delta E_{\text{REL}}. \qquad (2)$$

The basic difference between the two HEAT type models is the separated treatment of the correlation effects for the core and valence electrons. While HEAT-345(Q) takes all electrons into consideration during the calculation of $\Delta E_{\text{CCSD(T)}}^{\infty}$ in eq (1), HEAT-345+d(Q) uses the frozen-core approximation to compute this term in eq (2). In the latter protocol the error arising from the above separation is compensated by the term $\Delta E_{\text{core}}^{\infty}$ which is the core correlation energy contribution calculated with a smaller basis set. It is important to note that the partitioning of the electrons into core and valence regions causes an error in atomization energies which is not larger than 0.5 kJ mol^{-1}. Because of the presence of the sulfur atom, where it was appropriate, the standard Dunning type basis functions (aug-)cc-pVXZ were replaced by their (aug-)cc-pV(X+d)Z counterparts where the notation +d denotes the augmentation of the corresponding traditional basis set with a set of tight d functions. In this manner the adequate convergence of the various contributions could be ensured. The augmentation is crucial in the HF-SCF calculations primarily, however, it is also important in the correlation calculations as well if the targeted accuracy is high.

For the sulfur-containing species it was necessary to apply two further approximations. For the three largest systems (S$_2$COH, SCSOH, and HOSO$_2$) the CCSDT and the CCSDT(Q) calculations were practically not performable. These contributions were computed by the so-called MP2 FNO method and an additive error to the final results originating from this approximation was also considered. Furthermore, for species with hindered internal rotations where the rigid rotor–harmonic oscillator (RRHO) model is not straightforwardly applicable a one-dimensional hindered rotor approximation was used to correct zero-point vibrational energies, thermal corrections, and standard entropies. In these cases conformational analyses were performed at the CCSD(T)/cc-pVTZ level. By complete analogy with HEAT-345(Q) heat of formation values at 0 K were derived on the basis of the elemental reaction approach. For both sulfur and carbon the gaseous atoms were used as references.

Most of the calculations presented here were performed using the CFOUR program package and its predecessor ACESII. The higher-order coupled-cluster contributions were calculated using the MRCC code. For the HF-SCF/aug-cc-pCV6Z computations the MOLPRO package was invoked. To process the calculated raw data as well as to derive the final results I also used several self-written or self-modified smaller routines and scripts.

3 New scientific results

T.1. Heat of formation of formaldehyde

1.1. According to the philosophy of the focal-point approach the heat of formation and the proton affinity of formaldehyde were calculated with an associated accuracy that surpasses all previous experimental and theoretical values. The total non-relativistic energy of the involved species at 0 K was calculated by means of the FPA scheme where the various contributions (HF-SCF, CCSD, CCSD(T), ...) were determined at the highest possible level of theory using the largest Dunning type basis sets together with extrapolation formulas. The non-relativistic energy was further augmented by several additive corrections. Particularly, scalar relativistic and diagonal Born-Oppenheimer corrections as well as harmonic and anharmonic contributions of the ZPE were calculated also at the highest affordable level of theory. By investigating the convergence properties of the various contributions to the FPA total energy, a well-defined error bar can be easily defined to the calculated thermodynamic functions.

1.2. By means of the FPA procedure defined above, the heat of formation of formaldehyde at 0 K has been calculated on the basis of three different reactions. The corresponding 298.15 K has been obtained by calculating thermal corrections according to the formulas of statistical thermodynamics. The results are summarized in the following table:

Table 1
Heat of formation of formaldehyde at 0 K
(in $kJ\,mol^{-1}$) calculated using three
reaction schemes.

Reaction scheme	$\Delta_f H_0^\circ(H_2CO)$
$H_2 + CO = H_2CO$	-105.42
$H_2O + C = H_2CO$	-105.58
$2\,H + C + O = H_2CO$	-105.63

Out of the three reactions, I accepted the first one as the most accurate, because on one hand, the remaining errors in the contributions of the total energy are the smallest for this reaction, and on the other hand, this scheme involves only closed-shell species. The necessary highly accurate heat of formation value for CO was taken from literature.

In summary, the final FPA results for the heat of formation of formaldehyde are $\Delta_f H_0^\circ(H_2CO) = -105.42 \pm 0.40\,kJ\,mol^{-1}$ and $\Delta_f H_{298}^\circ(H_2CO) = -109.23 \pm 0.40\,kJ\,mol^{-1}$.

T.2. *Ab initio* thermochemistry of the formaldehyde derivatives

2.1. Similar to formaldehyde, the experimentally and theoretically derived enthalpy and entropy values of its derivatives are also inaccurate and in some cases they have unacceptably large uncertainties. Because of this fact, a slightly modified HEAT type model chemistry, the HEAT-345(Q) protocol has been applied to this family of molecules. According to the model, I calculated total relativistic energies for the molecules and their constituent atoms and elements

by a more or less rigid and recipe type algorithm. For all calculations equilibrium geometries obtained at the CCSD(T)/cc-pVQZ level of theory were used.

2.2. In order to check the accuracy of the protocol I studied the extrapolation formulas applied to the HF-SCF energies. I investigated the convergence of the extrapolated HF-SCF part obtained with the two most popular equations: the three-point formula of Feller (eq 3a) and the two-point formula of Karton-Martin (eq 3b).

$$E_{HF}^X = E_{HF}^\infty + b \cdot e^{-cX} \tag{3a}$$

$$E_{HF}^X = E_{HF}^\infty + a\,(X+1)\,e^{-9\sqrt{X}}, \tag{3b}$$

where X denotes the cardinal number of the Dunning type basis set used to calculate the single-point HF-SCF energy E_{HF}^X, E_{HF}^∞ is the basis set limit energy, and a and b are extrapolation parameters. On the basis of my results it is fair to say that extrapolations from the aug-cc-pCV[Q and 5]Z values using the Karton-Martin formula estimates the HF-SCF limit energy to a reasonable accuracy and, therefore, in studies where the targeted precision is in the sub-kJ/mol range the usage of this extrapolation procedure is highly recommended. Considering the HF-SCF results obtained with the triple-ζ basis set aug-cc-pCVTZ it is obvious that these vales are sufficiently far from convergence. Hence, accuracy of limit energies obtained with extrapolations involving these values are not always satisfactory. Nevertheless, if the aimed accuracy is in the 1-2 kJ mol^{-1} range then the triple-ζ results together with the Feller type equation can be used more or less reliably.

2.3. From total energies calculated with the HEAT-345(Q) protocol, heats of formation for formaldehyde derivatives at 0 K were derived using the elemental reaction approach. The corresponding 298.15 K enthalpies along with standard molar entropies at this temperature were computed using thermal corrections obtained with the RRHO approximation. All calculated heats of formation and in almost all cases the new entropy values surpass every previous literature data in accuracy. Therefore, all enthalpies and a total of 4 entropies calculated here were recommended as new reference values. My final results are listed below in table 2.

T.3. *Ab initio* **thermochemistry of several sulfur-containing species. Heat of formation of the sulfur atom.**

3.1. The HEAT-345+d(Q) protocol is a new and more or less system specific *ab initio* model chemistry applicable to sulfur-containing species. It can be derived from its predecessor HEAT-345(Q) by applying several simplifications. The accuracy of this reduced protocol were proven by performing calculations on a test set containing four molecules (H_2S, S_2, SO, and SO_2) for which reliable thermochemical quantities with small, well-defined error bars were available in the literature. During the test calculations it was found that the experimental heat of formation of the sulfur atom taken from literature is quite inaccurate: my calculated results for the test set differed significantly from the corresponding literature data. Therefore, $\Delta_f H_0^\circ(S_{gas})$ value of the NIST-JANAF database was corrected using a recently published experimental result for the dissociation energy of S_2. Using the new atomic heat of formation, $\Delta_f H_0^\circ(S_{gas}) = 277.25 \pm 0.3$ kJ mol^{-1}, the calculated values were in a reasonable agreement with the available literature data.

Based on my findings, the usage of this new quantity in subsequent thermochemical calculations is highly recommended.

3.2. For six members from the family of sulfur-containing molecules systematic conformational analyses were performed at the CCSD(T)/cc-pVTZ level. The potential energy curves for the hindered internal rotations were derived and the barrier heights for the motion were calculated. Based on these results, by solving a one-dimensional Schrödinger equation, the statistical thermodynamical quantities obtained with the RRHO approximation were corrected.

3.3. Enthalpies of formation at 0 K were calculated using the elemental reaction approach. For sulfur the ground state atom in the gas phase was used as reference and its heat of formation was the value obtained in my doctoral work. Enthalpies and entropies at room temperature were derived using the appropriate thermal corrections computed via the RRHO method. In three cases additional corrections due to the hindered internal motions were also considered according to thesis 3.2. Except for the SH radical, enthalpies and entropies of formation calculated here were proven to be the most accurate values in the literature to date, therefore, these vales were recommended as new references. The final results are summarized in table 2.

3.4. For the heat of formation of HOS and its structural isomer HSO there exist inconsistent experimental results in the literature. On the basis of my theoretically calculated values the conclusion was drawn that instead of the assumed HSO it was very likely HOS that was observed in the experiments.

3.5. The experimental results obtained for the reaction $CS_2 + OH$ are also contradictory. Some studies concluded that the product of the reaction is most likely the thermodynamically favored S_2COH radical while others assumed this reaction channel kinetically hindered and proposed the formation of the isomer radical SCSOH. In order to resolve this discrepancy, I calculated the enthalpy, entropy, Gibbs-energy, and the equilibrium constant for both reaction channels and compared these values with those measured in the experiments. My results clearly revealed that the reaction between radicals CS_2 and OH leads to the formation of the thermodynamically less favored isomer SCSOH.

Table 2

Summary of my results: thermal corrections and heats of formation (in $\mathrm{kJ\,mol^{-1}}$) at temperatures of 0 and 298.15 K as well as standard molar entropies (in $\mathrm{J\,K^{-1}\,mol^{-1}}$) at room temperature. Except for the values set as italic I recommended these results as new references in subsequent thermochemical calculations.

| Species | $H^\circ_{298} - H^\circ_0$ | Enthalpy of formation | | S°_{298} |
		$\Delta_f H^\circ_0$	$\Delta_f H^\circ_{298}$	
CF_2O	11.1	-603.4 ± 1.2	-606.5 ± 1.2	*258.6 ± 1.5*
FCO	10.4	-177.2 ± 0.9	-176.7 ± 0.9	248.8 ± 1.5
HFCO	10.4	-378.8 ± 1.2	-382.5 ± 1.2	*246.5 ± 1.5*
HClCO	11.0	-180.8 ± 1.7	-184.2 ± 1.7	*258.7 ± 1.5*
FClCO	11.9	-408.4 ± 1.7	-410.9 ± 1.7	*276.4 ± 1.5*
cisz-HOCO	10.9	-174.6 ± 1.2	-177.8 ± 1.2	251.7 ± 1.5
transz-HOCO	10.9	-181.3 ± 1.2	-184.5 ± 1.2	251.4 ± 1.5
NH_2CO	12.5	-7.2 ± 1.5	-13.1 ± 1.5	256.5 ± 1.5
HSO	10.1	-19.6 ± 2.3	-22.6 ± 2.3	241.4 ± 1.5
HOS	10.1	-2.4 ± 2.3	-5.3 ± 2.3	240.0 ± 1.5
$HOSO_2$	13.6	-367.5 ± 4.4	-375.7 ± 4.4	294.1 ± 1.5
transz-HSNO	12.4	112.4 ± 2.7	107.3 ± 2.7	266.4 ± 1.5
cisz-HSNO	12.4	115.8 ± 2.7	110.7 ± 2.7	266.1 ± 1.5
SH	*9.3*	*142.8 ± 1.9*	*143.4 ± 1.9*	*195.4 ± 1.5*
CH_2SO	11.7	-28.3 ± 3.1	-35.1 ± 3.1	261.4 ± 1.5
CH_2SH	13.4	165.3 ± 3.1	160.2 ± 3.1	270.0 ± 1.5
SCSOH	16.0	111.4 ± 5.0	108.9 ± 5.0	321.8 ± 1.5
S_2COH	13.5	20.1 ± 4.6	15.1 ± 4.6	296.9 ± 1.5

4   Publications related to the dissertation

1. **Proton affinity and enthalpy of formation of formaldehyde**
   G. Czakó, B. Nagy, G. Tasi, Á. Somogyi, J. Šimunek, J. Noga, B. J. Braams, J. M. Bowman and A. G. Császár
   International Journal of Quantum Chemistry, **109**(11), 2393-2409 (2009).
   IF$_{2012}$=1.306

2. **High-Accuracy Theoretical Study on the Thermochemistry of Several Formaldehyde Derivatives**
   B. Nagy, J. Csontos, M. Kállay and G. Tasi
   Journal of Physical Chemistry A, **114**(50), 13213-13221 (2010).
   IF$_{2012}$=2.771

3. **High-Accuracy Theoretical Thermochemistry of Atmospherically Important Sulfur-Containing Molecules**
   B. Nagy, P. Szakács, J. Csontos, Z. Rolik, G. Tasi and M. Kállay
   Journal of Physical Chemistry A, **115**(26), 7823-7833 (2011).
   IF$_{2012}$=2.771

4. **Correction to "High-Accuracy Theoretical Thermochemistry of Atmospherically Important Sulfur-Containing Molecules"**
   B. Nagy, P. Szakács, J. Csontos, Z. Rolik, G. Tasi and M. Kállay
   Journal of Physical Chemistry A, **117**(24), 5220-5220 (2013).
   IF$_{2012}$=2.771

5   Other publications

5. **Comment on "Enthalpy Difference between Conformations of Normal Alkanes: Raman Spectroscopy Study of n-Pentane and n-Butane"**
G. Tasi and B. Nagy
Journal of Physical Chemistry A, **114**(24), 6728-6728 (2010).
IF$_{2012}$=2.771

6. **Similarity analysis of the conformational potential energy surface of *n*-pentane**
G. Tasi, B. Nagy, G. Matisz and T. S. Tasi
Computational and Theoretical Chemistry, **963**(2-3), 378-383 (2011).
IF$_{2012}$=1.139

7. **Benchmarking Experimental and Computational Thermochemical Data: A Case Study of the Butane Conformers**
D. Barna, B. Nagy, J. Csontos, A. G. Császár and G. Tasi
Journal of Chemical Theory and Computation, **8**(2), 479-486 (2012).
IF$_{2012}$=5.389

8. **Dissociation of the fluorine molecule**
B. Csontos, B. Nagy, J. Csontos, and M. Kállay
Journal of Physical Chemistry A, **117**(26), 5518-5528 (2013).
IF$_{2012}$=2.771

6 Conference lectures and posters

1. *Poster:* **Dissociation of the Fluorine Molecule: A Benchmark Study**
   B. Nagy, J. Csontos, B. Csontos, and M. Kállay
   7th Molecular Quantum Mechanics – Electron Correlation: The Many-Body Problem at the Heart of Chemistry. An International Conference in Honour of Rodney J. Bartlett.
   Lugano, Switzerland, **2013**.

2. *Poster:* **Transition State Theory Calculations at Higher Levels of Electronic Structure Theory: The HO_2 + ClO reaction**
   B. Nagy, and M. Kállay
   XIV. International Congress of Quantum Chemistry
   Boulder, CO, USA, **2012**.

3. *Poster:* **Accurate ab initio heats of formation and standard molar entropies for several atmospherically important formyl derivatives**
   B. Nagy, J. Csontos, M. Kállay, and G. Tasi
   9th Central European Symposium on Theoretical Chemistry
   Novy Smokovec, Slovakia, **2010**.

4. *Lecture (in Hungarian): Ab initio* **thermochemistry of formaldehyde**
   B. Nagy
   National Student Association Conference of Hungary (OTDK)
   Debrecen, **2009**.

5. *Lecture (in Hungarian):* **Exploring the conformational potential energy surface of *n*-pentane**
   B. Nagy
   XXXII. Chemistry Presentations for Young Scientists (KEN)
   Szeged, **2009**.

6. *Poster:* **Exploring the full conformational potential energy surface of *n*-pentane**
   G. Matisz, B. Nagy, I. Hannus, I. Kiricsi, and G. Tasi
   Molecular Quantum Mechanics – Analytic Gradients and Beyond. An International Conference in Honour of Professor Peter Pulay.
   Budapest, Hungary, **2007**.

Total peer-reviewed publications:	8	out of this, related to the dissertation:	4
Cumulative impact factor:	21.689	out of this, related to the dissertation:	9.619
Total independent citations:	30	out of this, related to the dissertation:	23