

Review

Accurate thermochemistry from quantum chemical calculations?

Walter M. F. Fabian

Institut für Chemie, Karl-Franzens Universität Graz, Graz, Austria

Received 6 August 2007; Accepted 4 September 2007; Published online 14 March 2008
© Springer-Verlag 2008

Abstract The answer to the title question is definitely “yes” – at least for fairly small molecules. Computational procedures, namely the *Weizmann* (Wn) and Gaussian-3 (G3) family of methods, the complete basis set extrapolation scheme (CBS- x), the “high accuracy extrapolated *ab initio* thermochemistry” (HEAT) as well as the “correlation consistent composite approach” (ccCA), aimed at energies with chemical accuracy or even better (sub kJ mol^{-1}) are described and several applications illustrating the level of accuracy that can be achieved are presented.

Keywords Thermodynamics; *Ab initio* calculations; Computer chemistry; Enthalpy of formation.

Introduction

Thermochemical quantities such as enthalpy of formation or entropy are among the most important chemical data. Reliable values for stable molecules are relatively well established and have been collected in several compilations [1–4]. The situation is less favourable for reactive intermediates such as free radicals. Also, experimental determination of entropy to high accuracy is demanding, especially at temperatures other than the boiling point. Here, accurate computational chemistry methods to allow a reliable prediction of thermochemical data of such species would be highly desirable. Actually, in recent compilations of thermochemical properties of selected radicals, results from quantum chemical

calculations have been included or used to resolve discrepancies between various experimental data [5, 6]. To be useful, chemical accuracy for energies, $\leq 4\text{--}8 \text{ kJ mol}^{-1}$, must be achievable by computational methods. Progress in the development of such methods also depends on the accuracy of the available experimental data. With the introduction of the so-called “Active Thermochemical Tables” (ATcT) benchmarking of quantum chemistry methods has been greatly facilitated [7].

Several reviews of computational methods specifically designed to achieve energies and/or thermochemical quantities with chemical accuracy or even better, have been published [8]. Consequently, first only a short outline of the various available procedures and methods will be given. The main emphasis is on the calculation of standard enthalpies of formation, *i.e.*, $\Delta_f H^\circ$ (298.15). Secondly, some recent applications to illustrate the achievable accuracy as well as pitfalls and directions of further developments will be described.

Description of methods and procedures

The microscopic–macroscopic connection

Quantum chemistry deals with single molecules, thermodynamics with molar quantities. The interpretation of macroscopic thermochemical data on the molecular level, *i.e.*, the microscopic–macroscopic connection, is provided by statistical thermodynam-

Correspondence: Walter M. F. Fabian, Institut für Chemie, Karl-Franzens Universität Graz, A-8010 Graz, Austria.
E-mail: walter.fabian@uni-graz.at

ics. The necessary equations for computing thermochemical data can be found in textbooks on statistical thermodynamics or white papers of software vendors [9] and need not be repeated here. In the following we only briefly mention the underlying assumptions and possible drawbacks resulting therefrom.

Zero point vibrational energies (ZPVE) as well as thermal corrections to energy usually are calculated by the harmonic oscillator-rigid rotor approximation. Since the least harmonic behavior is expected for low-frequency vibrations which make only small contributions to ZPVE, generally the harmonic oscillator approximation is considered to be a good approximation, provided a sufficiently reliable computational procedure is used and some scaling to account for anharmonic effects is done. Different scaling factors for different methods/basis sets [10] or even frequency ranges [11] are in use. However, accumulating evidence suggests that going beyond the harmonic approximation for the ZPVE contribution might be necessary not only in very accurate calculations of enthalpies of formation in the sub-kJ mol⁻¹ range (see the applications to ethane and *n*-alkanes below). Likely, the grid density used in the now popular DFT methods might have a larger impact on the accuracy of the ZPVE contribution than the harmonic oscillator approximation. The thermal contribution of the vibrational partition function to both internal energy as well as entropy can significantly depend on the harmonic vs. anharmonic treatment [12]. Here especially the terms arising from low-frequency modes are problematic with increasing errors at higher temperatures. Frequently, these low-frequency modes are associated with hindered internal rotations; for these automatic detection and correction procedures, based on the *Pitzer–Gwinn* tables [13], have been developed [12, 14]. Besides hindered internal rotations, other low-frequency (large amplitude) motions include ring inversion, pseudorotation, or umbrella-type inversion, *e.g.*, in ammonia. Thermochemical data beyond the harmonic approximation for molecular systems exhibiting this type of large amplitude motions have been calculated by *Katzer and Sax* [12]. As expected, the importance of anharmonic contributions to internal energy *U*, entropy *S*, and heat capacity *C_v* increases with temperature. Furthermore, for a series of silicon hydrides these authors have found that in some cases these effects will cancel whereas in others reinforce each other. For instance, the temperature

dependence of the enthalpy of isomerization, 1-pentasilene ↔ cyclopentasilane, is entirely due to anharmonic effects. Inclusion of anharmonic contributions leads to a change from Δ_R*H* (298) = −182.1 kJ mol⁻¹ to Δ_R*H* (3000) = −169.0 kJ mol⁻¹. Considering only harmonic vibrational frequencies leads to Δ_R*H* (298) = −180.0 kJ mol⁻¹ and Δ_R*H* (3000) = −182.6 kJ mol⁻¹. In contrast to this example, the anharmonic correction owing to hindered internal rotation of the two NO₂ moieties in N₂O₅ is nearly negligible, ΔΔ*H*[°] = 0.05 (100 K), 0.21 (298 K), and −0.07 kJ mol⁻¹ (500 K) [15].

Standard enthalpies of formation

Experimental standard enthalpies of formation refer to the reaction energy for formation of a molecule from its constituents atoms in the respective elemental standard state, *e.g.*, for methane Δ_f*H*[°] (CH₄) is obtained from the equation



Almost all direct computations of enthalpies of formation rely on the calculation of atomization energies. To reconcile the two different energy zeros, infinitely separated particles at rest and elements in their most stable, pure form under standard conditions for the computational and experimental energy zero, respectively, the *experimental* standard enthalpies of formation of the constituent atoms from the respective elemental standard states have to be added to the calculated enthalpy of atomization (Eq. 1):

$$\begin{aligned} \Delta_f H^\circ(T)(\mathbf{M}) = & E(\mathbf{M}) + \text{ZPVE}(\mathbf{M}) + [H(T) \\ & - H(0)](\mathbf{M}) - \sum \{E(\mathbf{X}_i) + [H(T) \\ & - H(0)](\mathbf{X}_i) + \Delta E_{\text{SO}}(\mathbf{X}_i)\} \\ & + \sum \Delta_f H(T)(\mathbf{X}_i) \end{aligned} \quad (1)$$

Thus, the so-called direct calculation of enthalpies of formation requires quantum chemical computations of the molecule *M* and the constituent atoms *X_i*. Usually the spin orbit correction to atoms, Δ*E*_{SO}, and if necessary also for molecular species, *e.g.*, the OH• radical, as well as the enthalpies of formation of the constituent atoms from their elemental standard states, Δ_f*H*(*T*)(*X_i*), are taken from experiment [3b, 16]. For molecules containing only first-row elements Δ*E*_{SO} is quite small; however, for larger molecules and especially those containing second- or higher row elements, this correction can become

quite substantial. For instance, in polychlorinated benzodioxines ΔE_{SO} amounts to $\geq 20 \text{ kJ mol}^{-1}$ [17]. The calculation of zero point vibrational energy and thermal corrections, $ZPVE(\mathbf{M})$, $[H(T) - H(0)](\mathbf{M})$, and $[H(T) - H(0)](\mathbf{X}_i)$ has been outlined in the previous section. Thus, we are left with the calculation of atomization energies, $E(\mathbf{M}) - \Sigma E(\mathbf{X}_i)$.

Standard methods for *ab initio* thermochemistry

In order to overcome the problems associated with electron correlation, all procedures devised to yield accurate thermochemical data use a composite energy approach (multilevel methods). The basic idea is the (approximate) additivity not only of basis set effects but also of the electron correlation energy. In the following, “standardized” variants of such multilevel approaches, *viz.* the Gaussian-3 (G3) family [8a, 8b, 18], the “complete basis set model chemistries” (CBS-*x*) [8c, 19], the *Weizmann* family W_n , $n = 1-4$, of methods [8d, 8e, 20], the “high accuracy extrapolated *ab initio* thermochemistry” (HEAT) [21], and the “correlation consistent composite approach” (ccCA) [22], will be described.

We commence by a short outline of the G3 procedures. First, it should be stressed that G3 is limited to systems containing only first- and second-row elements. Geometries are optimized with $MP2(\text{full})/6-31\text{G(d)}$, followed by a series of $MP4$ single point calculations to estimate basis set effects (diffuse and polarization functions); additional basis set and core-valence correlation effects are approximated by $MP2(\text{full})/\text{G3large}$ (G3large is a basis set specifically designed for the G3 procedure); correlation effects beyond fourth order are estimated by $\text{QCISD(T)}/6-31\text{G(d)}$ single point calculations and residual deficiencies in the correlation treatment are accounted for by an empirical correction (higher-level correction, HLC) of the form $-An_\beta - B(n_\alpha - n_\beta)$, where n_α and n_β are the numbers of α and β valence electrons with $n_\alpha \geq n_\beta$. The constants A and B are fitted to experimental data. Their actual values are different for molecular and atomic species and also depend on the respective G3 variant. This use of empirical correction factors allows for a reduction in remaining deficiencies of the G3 procedure. However, it also critically depends on the reliability of the available experimental data. In this context it is worthwhile to note that the enthalpy of formation of the carbon atom is still under intense development and any error

in this value might lead to a systematic degradation of calculated standard enthalpies of formation for molecules containing an increasing number of carbon atoms [23]. A further possible drawback is the use of different values of A and B for atomic and molecular species which makes the application of G3 procedures to potential energy surface calculations for *e.g.*, atom abstraction reactions, problematic. Finally, $ZPVE$ and thermal corrections are obtained from scaled $HF/6-31\text{G(d)}$ calculations. As discussed below, this might be quite inaccurate. To improve on this possible error source, an extended version of G3, designated as G3X, has been proposed, where $B3LYP/6-31\text{G(2df,p)}$ calculations are used to obtain both geometries as well as scaled $ZPVE$ s [8b]. These new geometries showed substantially better agreement with experiment. Furthermore, the scale factor was determined by fitting to published zero point energies [10] rather than experimental frequencies. Other subsequent modifications of the G3 procedure were aimed at increasing their range of applicability without too much loss of reliability. Probably the most versatile G3 variant is G3MP2B3 . Here, both geometries as well as $ZPVE$ s and thermal corrections are obtained from scaled $B3LYP/6-31\text{G(d)}$ calculations. No $MP4$ single point computations are done; instead only $MP2/\text{G3MP2large}$ and $\text{QCISD(T)}/6-31\text{G(d)}$ calculations are used to estimate correlation contributions and an analogous empirical HLC scheme is applied. The electronic energy component from G3MP2 and G3MP2B3 , thus, is: $E(\text{G3MP2}) = E[\text{QCISD(T)}/6-31\text{G(d)}] + \{E(\text{MP2}/\text{G3MP2large}) - E[\text{MP2}/6-31\text{G(d)}]\} + \text{HLC}$

Standard enthalpies of formations of a series of nearly 300 organic molecules containing up to 10 carbon atoms calculated by G3MP2B3 were nearly as accurate as those obtained by the original G3 procedure (mean absolute deviations between experimental and calculated $\Delta_f H^\circ$ (298) values are 3.0 kJ mol^{-1} and 3.5 kJ mol^{-1} for G3 and G3MP2B3). However, much larger systems could be treated with G3MP2B3 [24].

The “complete basis set model chemistries” (CBS-*x*) of *Petersson* and co-workers [8c, 19] also apply an empirical correction scheme as well as a pair correlation extrapolation. Like G3, the CBS methods are restricted to first- and second-row elements. The most sophisticated variant, CBS-QCI/APNO, allows for an average accuracy of $\leq 2.1 \text{ kJ mol}^{-1}$ but is applicable only to first-row elements and limited to small molecules. For the same set of organic molecules

mentioned above, CBS-Q [19a] was found to give unsatisfactory results, especially for aromatic compounds and, thus, no data were reported [24]. In contrast, with CBS-QB3 [19b] which uses both *B3LYP*/6-311G(2d,d,p) geometries and frequencies instead of *MP2*/6-31G[†] geometries and *UHF*/6-31G[†] frequencies as well as CCSD(T) instead of QCISD(T), nearly chemical accuracy could be achieved for these test molecules, MAD = 8.1 kJ mol⁻¹ [24]. Among the members of CBS family, thus, in cases where CBS/QCI-APNO is not applicable, CBS-QB3 should be chosen. However, this procedure is more limited in the size of molecules that can be treated than with *G3MP2B3*.

No empirical correction scheme is applied in the *Wn* family; more precisely, only the *W1* variant contains one single molecule- and experiment-independent parameter in the *l*-extrapolation scheme, $A + B/l^\beta$, with $\beta = 3.22$ determined by optimization of the total atomization energies against calculations for the training molecules done without using such an empirical parameter [20]. Thus, this parameter does not explicitly depend on experimental data. However, even with this simplification *W1* theory is limited to rather small molecules. Nevertheless, it has found widespread use because of its “black-box” implementation in some popular software packages. Because of this popularity, a brief outline of this procedure is given below. Throughout, (modified) correlation consistent basis sets [25], *VnZ* and *AVnZ* with $n = 2-6$, are used. Geometries are optimized by *B3LYP*/VTZ + 1 (the notation +1 indicates the presence of high-exponent *d*-type “inner polarization functions on second-row elements”); zero point energies are obtained from scaled harmonic *B3LYP*/VTZ + 1 frequencies. It had been found that this approach can quite safely replace the very costly CCSD(T)/VQZ + 1 geometries and CCSD(T)/VTZ + 1 anharmonic frequencies used in the *W2* protocol. Using this geometry, a series of coupled cluster single point calculations, CCSD(T)/AVDZ + 2*d*, CCSD(T)/AVTZ + 2*d1f*, and CCSD/AVQZ + 2*d1f* is done. The SCF component of the total atomization energy, TAE (*not* the individual SCF energies of the molecule and the constituent atoms), is obtained by extrapolating the SCF/AVDZ + 2*d*, SCF/AVTZ + 2*d1f* and SCF/AVQZ + 2*d1f* components of TAE by $A + B/C^l$ ($l = 2, 3$, and 4). The CCSD valence correlation component is extrapolated by $A + B/l^\beta$ ($\beta = 3.22$) from the CCSD/AVTZ +

2*d1f* and CCSD/AVQZ + 2*d1f* valence correlation energies ($l = 3$ and 4). Similarly, $A + B/l^\beta$ ($\beta = 3.22$) is used for estimating the (T) valence component from CCSD(T)/AVDZ + 2*d* and CCSD(T)/AVTZ + 2*d1f* values for the (T) contribution ($l = 2$ and 3). In contrast to the SCF component, here extrapolation might be either for the individual energies or the total atomization energy. Core correlation energies are calculated by CCSD(T)/*MT* small (the *Martin–Taylor* basis sets are designed specifically to treat core correlation effects [26]; despite the notation “small” the *MT* small basis set is still *larger* than the G3large basis). Inclusion of inner-shell correlation contributions was found to be absolutely essential; it is not recommended to use a lower-level procedure for their treatment than CCSD(T). Inclusion of scalar relativistic effects, calculated by the averaged coupled-pair functional (ACPF/*MT* small) turned out to be essential to obtain good results for second-row compounds, especially highly polar ones. If necessary, spin-orbit coupling effects are also calculated by ACPF/*MT* small (or have to be added manually, if the software package used does not allow this type of calculation). Overall, even *W1*, *i.e.*, the simplest member of the *Wn* family, uses quite high level methods and rather large basis sets. Consequently, it is restricted to fairly small molecules. Mean absolute deviations for total atomization energies of the training set were 1.38 kJ mol⁻¹, compared with 3.60 and 3.79 kJ mol⁻¹ for G3 and CBS-QB3, respectively. *W1* errors for second-row compounds are larger (1.86 kJ mol⁻¹) than for molecules containing only first-row elements (1.07 kJ mol⁻¹) [20a]. The importance of the core correlation and scalar relativistic contributions even in a simple molecule consisting only of light elements, *e.g.*, CH₄ can be seen from their fairly large amount of ≈ 4.5 kJ mol⁻¹. Subsequent modifications aimed at increasing the accuracy rather than broadening the applicability of *Wn* theory. The main weaknesses of *W1* and *W2* are a) the use of CCSD(T), which makes these procedures prone to failure in systems with significant non-dynamical electron correlation; and b) the simplified scalar relativistic treatment which apparently works quite well for first- and second-row elements but is expected to be insufficient for heavier elements. A molecule well outside the range of CCSD(T) is ozone because of its intrinsic multireference character: *W2* total atomization energies were in error by 12.6 kJ mol⁻¹. To make *Wn* theory more

robust also for such systems, better estimates for the contribution of connected quadruple excitations, inclusion of connected quintuples, and improved basis set extrapolation schemes were introduced. In addition, a more accurate scalar relativistic treatment (*Douglas-Kroll*) and the diagonal *Born–Oppenheimer* correction (*DBOC*) are used. Post-CCSD(T) contributions to the core-valence correlation energy were found to be only significant in systems with severe non-dynamical electron correlation [20c]. For highly accurate total atomization energies, accurate geometries, *e.g.*, CCSD(T)/cc-pVQZ, and *ZPVE* contributions beyond the harmonic approximation are essential (see the discussion for ethane in the next section). Finally, for validation revised experimental data obtained from the Active Thermochemical Tables [7] were used.

In contrast to other schemes for the calculation of thermochemical quantities, the “high accuracy extrapolated *ab initio* thermochemistry (HEAT) procedure [21] does not rely on atomization reactions. Rather, isodesmic, or nearly isodesmic schemes with an at least approximate conservation of bond types, are used. No empirical scaling factors or adjustments are used and all terms except the spin-orbit coupling contribution are size extensive, guaranteeing no degradation of the calculated results with increasing size of the molecules treated. The method has been designed with the aim to do the best what can be done (at the time of publication), even at the expense of being restricted to all but the smallest first-row systems (<5 atoms). Geometries as well as harmonic vibrational frequencies are obtained by CCSD(T)(full)/cc-pVQZ optimizations; an anharmonic correction is applied to the *ZPVE* contribution. Benchmark studies had shown that using the cc-pVQZ basis instead of the cc-pCVQZ negligibly affected geometries of molecules containing only first- or second-row elements. The SCF and CCSD(T) correlation energies are extrapolated separately using a three- and two-point extrapolation, respectively, with the aug-cc-pCV n Z basis sets [$n = 3, 4$, and 5 for SCF; $n = 4$ and 5 for CCSD(T)]. It should be pointed out that in contrast to other approaches, in HEAT the CCSD(T) correlation treatment includes the core electrons; none are omitted from the individual CCSD(T) calculations. The contribution arising from a complete treatment of triple excitations is estimated by extrapolating CCSDT(fc)/cc-pV n Z, $n = 3$ and 4, energies,

$\Delta E_{\text{CCSDT}} = E^{\text{34}}_{\text{CCSDT}}(\text{fc}) - E^{\text{34}}_{\text{CCSD(T)}}(\text{fc})$. Since the contribution of quadruple excitations converges quite rapidly with basis set size [27], higher-level correlation effects are obtained in the HEAT protocol as the difference of CCSDTQ and CCSDT correlation energies with the cc-pVDZ basis set (frozen core approximation). In the HEAT test molecules, ΔE_{CCSDTQ} contributes 1–2 kJ mol^{−1} to the enthalpies of formation with opposite sign to the ΔE_{CCSDT} term [21a]. The diagonal *Born–Oppenheimer* correction is obtained at the *Hartree-Fock* level. *HF*/aug-cc-pVTZ deemed to be sufficiently accurate (for problems associated with this approach, see below). The spin-orbit contribution is obtained by a spin-orbit configuration interaction calculation with the cc-pVDZ basis set and relativistic effective core potentials which include spin-orbit effects. Only first-order spin-orbit interactions are taken into account; thus, the only molecules with $\Delta E_{\text{SO}} \neq 0$ are radicals with degenerate ground states, *e.g.*, the $\cdot\text{OH}$ radical. Scalar relativistic effects (*Darwin* and mass-velocity terms [28]) are treated at the CCSD(T)/aug-cc-pCVTZ level. Since the application of the HEAT method is restricted at present to first-row systems, this approximation should be accurate enough. The MAD for the enthalpies of formation at 0 K of a series of simple molecules based on ATcT experimental data is 0.34 kJ mol^{−1}. Some further improvements (treatment of electron correlation up to quintuple excitations) have recently been described [21b].

The correlation consistent composite approach (ccCA) has been designed along similar lines as Gaussian-3. The main differences to this latter scheme are the use of the correlation consistent basis sets [25], which allow for a more straightforward extrapolation to the basis set limit than Pople-style basis sets, and the exclusion of *any* semi-empirical or optimized parameters [22]. Depending on the reference energy, *MP4*/cc-pVDZ, *MP4*/cc-pVTZ, *MP4*/aug-cc-pVTZ, or *MP2*/CBS aug-cc-pV n Z ($n = 2, 3$, and 4 using two different extrapolations schemes), several variants have been implemented. All of them use *B3LYP*/6-31G(d) geometries and scaled harmonic frequencies for the *ZPVE* contribution. In the ccCA CBS variant, the reference energy is obtained by extrapolating *MP2*/aug-cc-pV n Z energies to the basis set limit with two different extrapolation formulae, denoted as CBS-1 and CBS-2. Higher correlation effects are estimated with the aid of QCISD(T)/cc-pVTZ, and core correlation contri-

butions are obtained from *MP2(full)/aug-cc-pCVTZ* single point calculations. Mean absolute deviations for computed enthalpies of formation, $\Delta_f H^\circ$ (298), for the G2-1 test set are 6.07 (CBS-1) and 5.56 kJ mol⁻¹ (CBS-2). This latter value can be reduced to 3.39 kJ mol⁻¹ when atomic spin-orbit and scalar relativistic corrections are added. A further assessment of the performance of this theoretical thermochemistry scheme is provided by the enthalpies of formation for the series of *n*-alkanes, C_{*n*}H_{2*n*+2}, *n* = 1–8 (see next section). In contrast to the previous procedures which are restricted to first- and second-row systems (or even only first-row molecules), recently the ccCA scheme has been extended to simple transition metal compounds [22b]. Depending on the extrapolation formula used, for a test set of ten such species, *e.g.*, TiO₂, TiF₂, VO₂, FeCl₃, and Zn(CH₂CH₃)₂, *etc.*, a mean absolute deviation of 36.0–40.2 kJ mol⁻¹ was obtained. A composite DK-CCSD(T) approach comparable to that used by Dixon, Feller, and Peterson [29], resulted in a MAD of 34.3 kJ mol⁻¹. The largest deviations, up to >130 kJ mol⁻¹, were found for TiF₂ and TiF₃. Omitting these two species reduces the MAD to 15.5–19.2 kJ mol⁻¹ [13.0 for CCSD(T)].

Besides the procedures described above, other composite methods have been described in the literature, notably the *focal point* extrapolation scheme of Allen and co-workers [30]. Finally, one of the first successful approximate *ab initio* procedures for first-principles thermochemistry, the bond-additivity correction (BAC) methodology, *e.g.*, BAC-*MPx* (*x* = 2, 4), should be mentioned [31]. An empirical variant of multilevel approaches is the *multicoefficient* method of Truhlar *et al.* [32], where the energy components of, *e.g.*, the G3 procedure, are weighted by optimized coefficients (MCG3). The omission of the empirical HLC used in G3 makes MCG3 size-extensive, allowing the calculation of potential energy surfaces.

Examples and applications

The following examples are intended to illustrate (i) the level of accuracy of enthalpies of formation that can be reached with *ab initio* quantum chemical methods; (ii) steps that might be necessary beyond “standard” implementations; (iii) directions for possible further improvements; and (iv) problems or successes with larger molecules.

Total atomization energy of ethane

The accuracy that can be achieved with the most sophisticated of the methods described above can be illustrated by a calculation of the atomization energy of ethane by the W4 procedure [33]. The calculated W4 total atomization energy at 0 K, $TAE_0 = 2788.89$ kJ mol⁻¹ to be compared with an experimental value of 2787.30 ± 0.25 kJ mol⁻¹ [1, 7]. This 1.6 kJ mol⁻¹ discrepancy was deemed to need explanation. Internal rotation in ethane could be ruled out as a major contributor. However, the (in)accuracy of the zero point vibrational energy was found to be significant. Using CCSD(T)/cc-pVQZ harmonic frequencies with anharmonic corrections (quartic force field) from *B3LYP*/pc-2 (pc-2 is a basis set comparable to cc-pVTZ but optimized for DFT methods) calculations and employing very large grids, the difference between experiment and calculation could be reduced to 0.84 kJ mol⁻¹. Especially for molecules containing many hydrogen atoms, post-SCF contributions to the diagonal *Born–Oppenheimer* correction had already been found to be significant. Therefore, CISD/cc-pVDZ calculations were performed for these DBOC corrections, leading to perfect agreement with experiment, TAE_0 (calc) = 2787.34 kJ mol⁻¹. The main conclusion emerging from this study is that for small molecules for which this type of high-level electronic structure calculations are possible, nuclear motion related issues, like anharmonic *ZPVE* corrections and/or diagonal *Born–Oppenheimer* corrections, may become increasingly important for highly accurate thermochemical data.

The $\cdot\text{NO}_3$ radical

Experimental enthalpies of formation of the nitrate radical $\cdot\text{NO}_3$ are $\Delta_f H^\circ$ (298) = 71.13 ± 20.9 [3b], 73.2 ± 4.2 [34a], and 73.72 ± 1.38 kJ mol⁻¹ [34b]. *G3MP2B3* as well as CCSD(T) calculations with basis set extrapolation (aug-cc-pVIZ, *l* = 2–5) led to a significant deviation from these experimental values, $\Delta_f H^\circ$ (298) = 83.30 (*G3MP2B3*) and 83.34 [CCSD(T)] kJ mol⁻¹ [15]. In view of the close agreement obtained by these two procedures as well as previous experience with *G3MP2B3* for a set of 32 free radicals with a mean absolute deviation between experimental and calculated standard enthalpies of formation <4 kJ mol⁻¹ [35], a purely adi-

adiabatic description of the nitrate radical was supposed to be inadequate. This assertion is further corroborated by the fact that – although calculations resulted in a D_{3h} molecular geometry of $^{\bullet}\text{NO}_3$ – a tenacious symmetry-breaking of the *HF* wave function had been observed which could not be overcome by inclusion of just dynamic electron correlation [36]. A MR-ACPF/cc-pVTZ (multireference averaged coupled-pair functional) treatment of the diabatic coupling indicated a strong interaction (*pseudo-Jahn-Teller-effect*) between the $^2\text{A}'_2$ ground state and the second excited state, $^2\text{E}'$. Although this excited state lies significantly higher (calculated vertical excitation energy $T_e^v = 15755\text{ cm}^{-1}$) than the first excited state $^2\text{E}''$ ($T_e^v = 5834\text{ cm}^{-1}$), diabatic coupling with this latter turned out to be negligible. The $^2\text{A}'_2 - ^2\text{E}'$ interaction leads to an energy lowering of the $^2\text{A}'_2$ ground state by $\Delta E = -7.19\text{ kJ mol}^{-1}$, resulting in $\Delta_f H^\circ(298) = 76.15\text{ kJ mol}^{-1}$.

Enthalpies of formation of *n*-alkanes

The enthalpies of formation of a series of *n*-alkanes $\text{C}_n\text{H}_{2n+2}$, $n = 1-16$ have been used to assess the reliability of G3 procedures [37]. An increase in the deviation between experimental and calculated enthalpies of formation with increasing chain length has been observed, *ca.* 0.2 kJ mol^{-1} for G3 and 0.4 kJ mol^{-1} for G3MP2B3 per bond. *MADs* for the series $n = 1-8$ are quite small and well below the desired chemical accuracy, 2.05 (G3) and 0.75 kJ mol^{-1} (G3MP2B3). Similarly, CCSD(T) calculations with basis set extrapolation and inclusion of core correlation and scalar relativistic effects [38] gave a *MAD* = 0.96 kJ mol^{-1} for *n*-alkanes with $n = 1-8$ excluding *n*-heptane. The largest correction to the total atomization energy was found to be the *ZPVE* contribution ($120-620\text{ kJ mol}^{-1}$) followed by the core-valence correction ($4-35\text{ kJ mol}^{-1}$). This clearly demonstrates that for achieving chemical accuracy, the *ZPVE* correction must be better than $\pm 4\text{ kJ mol}^{-1}$. Similar to the G3 procedures, the correlation consistent composite approach also showed an increased deviation with increasing chain length, *e.g.*, from -7.5 (methane) to -33.1 kJ mol^{-1} (*n*-hexane, ccCA-aTZ variant), and from -4.2 (methane) to -16.7 kJ mol^{-1} (octane, ccCA-CBS-1). In contrast, ccCA-CBS-2 enthalpies of formation of *n*-alkanes are too small, leading to positive deviations, $+0.4$ (methane) to $+9.2\text{ kJ mol}^{-1}$ (octane); inclusion of

ΔE_{SO} and relativistic corrections significantly improves the ccCA-CBS-2 results, *MAD* = 0.9 kJ mol^{-1} in excellent agreement with the CBS-CCSD(T) results [38].

The hydrotrioxy radical $^{\bullet}\text{O}_3\text{H}$

Computational thermochemistry is especially important for short-lived reactive intermediates, *e.g.*, free radicals, where accurate experimental data are difficult if not impossible to obtain. Consequently, a number of calculations on a series of small [5, 35] or medium-sized radicals [39], including reactive species involved in peroxyxynitrite chemistry [40], combustion, soot formation, and atmospheric processes [29d, 30, 41, 42], have been published. As an example for the interplay between theory [42] and experiment [43] to derive accurate enthalpies of formation for highly unstable species, results for the hydrotrioxy radical $^{\bullet}\text{O}_3\text{H}$ will be described in some detail. This radical has long been postulated as a possible intermediate in the reaction of hydrogen atoms with ozone,



or oxygen atoms with hydroperoxy radicals,



playing, thus, an important role in atmospheric ozone chemistry as well as in hydrocarbon combustion and biological oxidations. Calculated [35b, 42] values for $\Delta_f H^\circ(298)$ ($^{\bullet}\text{O}_3\text{H}$), 26.4 (CBS-QCI/APNO), 31.1 (G3MP2B3), 35.1 (G3), and 32.9 kJ mol^{-1} (W1), substantially deviate from the experimental value of *Speranza*, $\Delta_f H^\circ(298)$ ($^{\bullet}\text{O}_3\text{H}$) = $-4.2 \pm 21\text{ kJ mol}^{-1}$ [43a]. However, *Ventura et al.* argued in view of the convergence of the calculated $\Delta_f H^\circ(298)$ ($^{\bullet}\text{O}_3\text{H}$), both with respect to the computational procedure used as well as whether atomization or isodesmic/isogyric reactions were used, that it is the experimental value which needs reconsideration. They recommend a value of $\Delta_f H^\circ(298)$ ($^{\bullet}\text{O}_3\text{H}$) = 26.4 kJ mol^{-1} [42c]. Since the hydrotrioxy radical might exhibit significant non-dynamical electron correlation, a multireference averaged coupled-pair functional treatment (MR-ACPF-CBS) including not only basis set but also a CAS extrapolation has been performed yielding $\Delta_f H^\circ(298)$ ($^{\bullet}\text{O}_3\text{H}$) = 21.2 kJ mol^{-1} [42d]. Recently, infrared action spectroscopy has provided an upper limit for the HO-

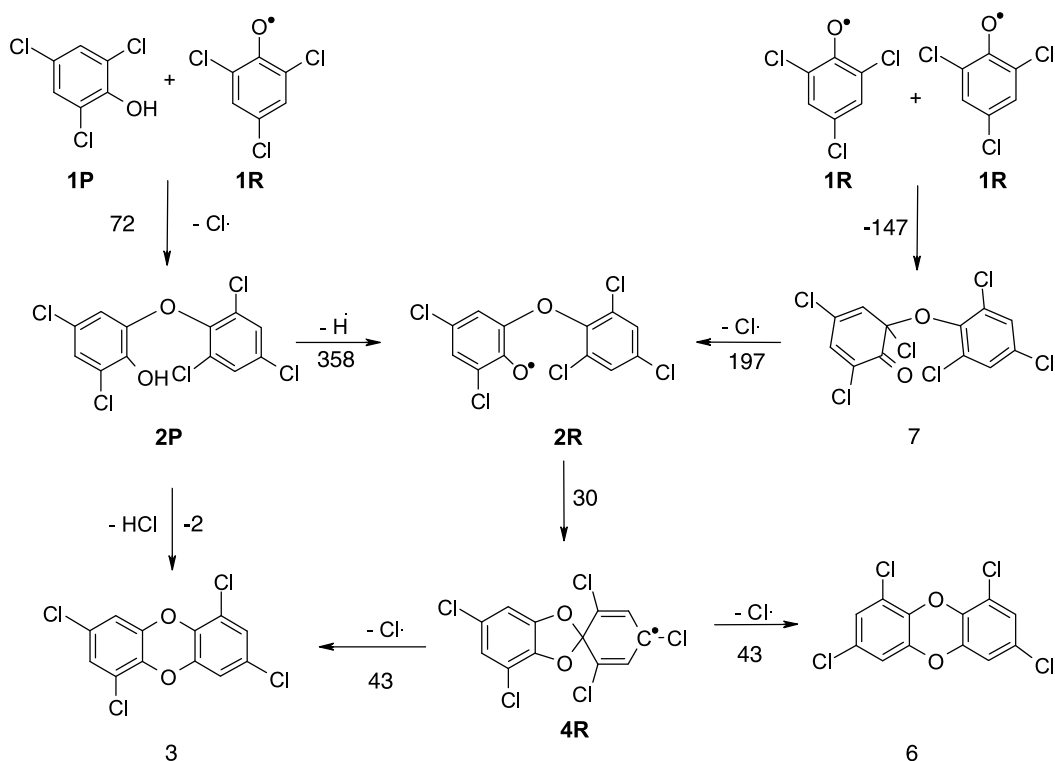
O_2 binding energy, $D_0 \leq 25.6 \text{ kJ mol}^{-1}$, resulting in a lower limit for the enthalpy of formation for the $\cdot O_3H$ radical of $\Delta_f H^\circ$ (298) ($\cdot O_3H$) $\geq 7.1 \text{ kJ mol}^{-1}$ [43c] – much closer to the theoretical results.

Chlorinated dibenzo-*p*-dioxins

The examples and applications discussed so far had been restricted to rather small molecules, allowing treatment with high-level procedures. Larger systems require the use of less sophisticated procedures, hopefully without loss of too much accuracy. In this respect, the G3MP2B3 variant of the G3 family seems to be quite promising, and the following example describes its use for the treatment of polychlorinated dibenzo-*p*-dioxins. These compounds are produced from a variety of sources and in order to better understand their formation and/or removal, several theoretical studies concerning their electronic, physicochemical, and thermodynamic properties have appeared [17, 44]. A detailed elementary kinetic model, consisting of >70 individual reaction steps, for the gas phase formation of isomeric tetrachlorinated dibenzo-*p*-dioxins from

2,4,6-trichlorophenol, has been published [45]. Both radical-molecule as well as radical-radical pathways have been considered. The key steps for these two pathways are shown in Scheme 1. The G3MP2B3 enthalpies of reaction are also indicated there.

From previous experience these data are expected to be sufficiently reliable to be useful in kinetic simulations. Calculated G3MP2B3 enthalpies of formation for 1,3,6,8-tetrachlorodibenzo-*p*-dioxine (**3**) ($\Delta_f H^\circ$ (298) = $-174.1 \text{ kJ mol}^{-1}$) and 1,3,7,9-tetrachlorodibenzo-*p*-dioxine (**6**) ($\Delta_f H^\circ$ (298) = $-174.1 \text{ kJ mol}^{-1}$) show good agreement with the corresponding experimental estimates (group additivity method), $\Delta_f H^\circ$ (298) = $-173.0 \text{ kJ mol}^{-1}$ for **3** and **6** [46]. Both isomers, **3** and **6**, can be formed by radical-molecule and radical-radical reactions. All steps but radical combinations are endothermic; however, subsequent reactions of the expelled radicals, H^\cdot and/or Cl^\cdot , should provide sufficient “driving force” for the reactions to proceed. The estimated activation energy for the formation of **2P** via the radical-molecule path, $1P + 1R \rightarrow 2P$ (Scheme 1), $E_A \approx 100 \text{ kJ mol}^{-1}$, is on the high end side, where-



Scheme 1 (Reprinted from Combustion and Flame 145, Fabian WMF, Janoschek R, Thermochemical properties of stable intermediates and derived radicals for the gas-phase formation of 1,3,6,8-G3MP2B3 study, 289–289, Copyright 2007 with permission from Elsevier (<http://www.sciencedirect.com/science/journal/00102180>))

as that for the reverse reaction $7 \rightarrow 1R + 1R$, $E_A \approx 150 \text{ kJ mol}^{-1}$, is considerably smaller than the value used in the kinetic modelling [45]. Repeating the kinetic modelling with these updated values would be worthwhile.

Conclusions and outlook

Significant progress has been achieved in *ab initio* thermochemistry and the question mark in the title safely can be omitted now – at least for small molecules. For these enthalpies of formation can be obtained in the sub-kJ mol⁻¹ range. In favourable cases revision of inaccurate experimental data or resolving discrepancies between them, has become possible. However, to push the accuracy that far is not only computationally quite demanding but requires also the availability of several software packages, and of course, considerable skill and experience. Besides these rather sophisticated procedures also several largely black-box implementations exist which in favourable cases also allow the calculation of enthalpies of formation with chemical accuracy ($\leq 4\text{--}8 \text{ kJ mol}^{-1}$). In future developments increased attention should be paid to nuclear motion related corrections, *e.g.*, more reliable vibrational frequencies and *ZPVE* corrections beyond the rigid rotor – harmonic oscillator approximation. Replacing *HF*/6-31G(d) or *B3LYP*/6-31G(d) *ZPVE* corrections used in several of the standardized approaches by more accurate contributions might also be useful, *e.g.*, the improvement brought about in the G3X variant. For larger molecules where even the G3 method barely can be used, the *G3MP2B3* offers an attractive alternative.

Finally, in this review we have concentrated exclusively on enthalpies, no mention has been made of procedures for the calculation of entropy. Generally, entropy is obtained from the rigid rotor-harmonic oscillator approximation used to derive the thermal contribution to enthalpy. Shortcomings of these approximations, especially the problem of low-frequency vibrations, are even more severe for entropy than for enthalpy. Anharmonic corrections appear to be essential, especially for molecules with internal rotations, inversions, and the like, or at higher temperatures. Recently, a consistent method for the calculation of entropy and heat capacity of organic molecules without adjustable parameters has been proposed [11].

References

1. Gurvich LV, Veyts IV, Alcock CB (1989, 1991) *Thermodynamic Properties of Individual Substances*, vols. 1, 2. Hemisphere, New York
2. Cox JD, Wagman DD, Medvedev VA (1989) *CODATA Key Values for Thermodynamics*, Hemisphere, New York
3. a) Afeefy HY, Liebmann JF, Stein SE (2005) *Neutral Thermochemical Data*. In: Linstrom PJ, Mallard WG (eds) *NIST Chemistry Webbook*, NIST Standard Reference Database 69, NIST, Gaithersburg; b) NIST-JANAF Thermochemical Tables, 4th edn. (1998). In: Chase MW Jr (ed) *J Phys Chem Ref Data*, Monograph 9, NIST, Gaithersburg; c) Johnson RD III (ed) *NIST Computational Chemistry Comparison and Benchmark Database*, NIST Standard Reference Database Number 101 Release 12, Aug 2005, NIST, Gaithersburg, <http://srdata.nist.gov/cccbdb>
4. Burcat A (2001) *Third Millenium Ideal Gas and Condensed Phase Thermochemical Database for Combustion*. TAE Report No 867, Technion, Haifa
5. Ruscic B, Boggs JE, Burcat A, Császár AG, Demaison J, Janoschek R, Martin JML, Morton ML, Rossi MJ, Stanton JF, Szalay PG, Westmoreland PR, Zabel F, Bérces T (2005) *J Phys Chem Ref Data* 34:573
6. Luo Y-R (2003) In: Lide DR (ed) *CRC Handbook Handbook of Chemistry and Physics*, 84th edn. CRC Press, Boca Raton
7. Ruscic B, Pinzon RE, Morton ML, von Laszewski G, Bittner SJ, Nijssure SG, Amin KA, Minkoff M, Wagner AF (2004) *J Phys Chem A* 108:9979
8. a) Curtiss LA, Redfern PC, Frurip DJ (2000) In: Boyd DB, Lipkowitz KB (eds) *Rev Comp Chem* 15:147; b) Curtiss LA, Raghavachari K (2002) *Theor Chem Acc* 108:61; c) Petersson GA (2002) In: Mezey PG (ser ed) *Understanding Chemical Reactivity*, vol. 22, Kluwer Academic Publishers (Springer), Dordrecht, p 99; d) Martin JML, Parthiban S (2002) In: Mezey PG (ser ed) *Understanding Chemical Reactivity*, vol. 22, Kluwer Academic Publishers (Springer), Dordrecht, p 31; e) Martin JML (2005) In: Spellmeyer D (ed) *Annual Reports in Computational Chemistry*, vol. 1, Elsevier, Amsterdam, p 31
9. a) Gasser RPH, Richards WG (1995) *An Introduction to Statistical Thermodynamics*. World Scientific, Singapore; b) McQuarry DA, Simon JD (1999) *Molecular Thermodynamics*, University Science Books, Sausalito; c) Cramer CJ (2004) *Essentials of Computational Chemistry*, J Wiley & Sons, Chichester, p 355; d) Ochterski JW (2000) *Thermochemistry in Gaussian* http://www.gaussian.com/g_whitepap/thermo.htm
10. Scott AP, Radom L (1996) *J Phys Chem* 100:16502
11. DeTar DF (2007) *J Phys Chem A* 111:4464
12. a) Katzer G, Sax AF (2002) *J Phys Chem A* 106:7204; b) Katzer G, Sax AF (2002) *J Chem Phys* 117:8219; c) Katzer G, Sax AF (2005) *J Comput Chem* 26:1438
13. Pitzer KS, Gwinn WD (1942) *J Chem Phys* 10:428
14. Ayala PY, Schlegel HB (1998) *J Chem Phys* 108:2314
15. Janoschek R, Kalcher J (2002) *Z Anorg Allg Chem* 628:2724

16. Moore CE (1952) Natl Bur Stand Circ (US) 467
17. Fabian WMF, Janoschek R (2006) Combust Flame 145:282
18. a) Curtiss LA, Raghavachari K, Redfern PC, Rassolov V, Pople JA (1998) J Chem Phys 109:7764; b) Curtiss LA, Redfern PC, Raghavachari K, Rassolov V, Pople JA (1999) J Chem Phys 110:4703; c) Baboul AG, Curtiss LA, Redfern PC, Raghavachari K (1999) J Chem Phys 110:7650
19. a) Ochterski JW, Petersson GA, Montgomery JA Jr (1996) J Chem Phys 104:2598; b) Montgomery JA Jr, Frisch MJ, Ochterski JW, Petersson GA (1999) J Chem Phys 110:2822
20. a) Martin JML, de Oliveira G (1999) J Chem Phys 111:1843; b) Boese AD, Oren M, Atasoylu O, Martin JML, Kállay M, Gauss J (2004) J Chem Phys 120:4129; c) Karton A, Rabinovich E, Martin JML, Ruscic B (2006) J Chem Phys 125:144108
21. a) Tajti A, Szalay PG, Császár AG, Kállay M, Gauss J, Valeev EF, Flowers BA, Vázquez J, Stanton JF (2004) J Chem Phys 121:11599; b) Bomble YJ, Vázquez J, Kállay M, Michauk C, Szalay PG, Császár AG, Gauss J, Stanton JF (2006) J Chem Phys 125:064108
22. a) DeYonker NJ, Cundari TR, Wilson AK (2006) J Chem Phys 124:114104; b) DeYonker NJ, Peterson KA, Steyl G, Wilson AK, Cundari TR (2007) J Phys Chem A ASAP article DOI: 10.1021/jp0715023
23. Tasi G, Izsák R, Matisz G, Császár AG, Kállay M, Ruscic B, Stanton JF (2006) Chem Phys Chem 7:1664
24. Bond D (2007) J Org Chem 72:5555
25. a) Dunning TH Jr (1989) J Chem Phys 90:1007; b) Woon DE, Dunning TH Jr (1993) J Chem Phys 98:1358; c) Woon DE, Dunning TH Jr (1994) J Chem Phys 100:2975; d) Woon DE, Dunning TH Jr (1995) J Chem Phys 103:4572; e) Balabanov NB, Peterson KA (2005) J Chem Phys 123:064107
26. a) Martin JML, Taylor PR (1994) Chem Phys Lett 225:473; b) Martin JML (1995) Chem Phys Lett 242:343
27. Ruden TA, Helgaker T, Jørgensen P, Olsen J (2003) Chem Phys Lett 371:62
28. a) Cowan RD, Griffin DC (1976) J Opt Soc Am 66:1010; b) Martin RL (1983) J Phys Chem 87:750
29. a) Feller D, Dixon DA (2000) J Phys Chem A 104:3048; b) Dixon DA, Feller D, Sandrone G (1999) J Phys Chem A 103:4744; c) Feller D, Peterson KA, de Jong WA, Dixon DA (2003) J Chem Phys 118:3510; d) Feller D, Dixon DA, Francisco JS (2003) J Phys Chem A 107:1604
30. a) East ALL, Allen WD (1993) J Chem Phys 99:4638; b) Wheeler SE, Allen WD, Schaefer HF III (2004) J Chem Phys 121:8800
31. a) Allendorf MD, Melius CF, Ho P, Zachariah MR (1995) J Phys Chem 99:15285; b) Zachariah MR, Westmoreland PR, Burgess DR Jr, Tsang W, Melius CF (1996) J Phys Chem 100:8737
32. Fast PL, Sánchez ML, Truhlar DG (1999) Chem Phys Lett 306:407
33. Karton A, Ruscic B, Martin JML (2007) J Mol Struct (Theochem) 811:345
34. a) Burrows JP, Tyndall GS, Moortgat GK (1985) Chem Phys Lett 119:193; b) Davis HF, Kim B, Johnston HS, Lee YT (1993) J Phys Chem 97:2172
35. a) Janoschek R, Rossi MJ (2002) Int J Chem Kinet 34:550; b) Janoschek R, Rossi MJ (2004) Int J Chem Kinet 36:661
36. Eisfeld W, Morokuma K (2000) J Chem Phys 113:5587
37. Redfern PC, Zapol P, Curtiss LA, Raghavachari K (2000) J Phys Chem A 104:5850
38. Pollack L, Windus TL, de Jong WA, Dixon DA (2005) J Phys Chem A 109:6934
39. Janoschek R, Fabian WMF (2003) J Mol Struct 661–662:635
40. Pfeiffer S, Mayer B, Janoschek R (2003) J Mol Struct (Theochem) 623:95
41. a) Senosiain JP, Musgrave CB, Golden DM (2003) Int J Chem Kinet 35:464; b) Golden DM, Smith GP, McEwen AB, Yu C-L, Eiteneer B, Frenklach M, Vaghjiani GL, Ravishankara AR, Tully FP (1998) J Phys Chem A 102:8598; c) Duncan TV, Miller CE (2000) J Chem Phys 113:5138; d) Kim EH, Bradforth SE, Arnold DW, Metz RB, Neumark DM (1995) J Chem Phys 103:7801; e) Rauk A, Yu D, Borowski P, Roos B (1995) Chem Phys 197:73; f) Yu H-G, Muckerman JT, Sears TJ (2001) Chem Phys Lett 349:547; g) Zhu RS, Diau EGW, Lin MC, Mebel AM (2001) J Phys Chem A 105:11249; h) Fabian WMF, Janoschek R (2005) J Mol Struct (Theochem) 713:227
42. a) Jungkamp TPW, Seinfeld JH (1996) Chem Phys Lett 257:15; b) Setokuchi O, Sato M, Matuzawa S (2000) J Phys Chem A 104:3204; c) Yu HG, Varandas AJC (2001) Chem Phys Lett 334:173; d) Denis PA, Kieninger M, Ventura ON, Cachau RE, Dierksen GHF (2002) Chem Phys Lett 365:440 (Erratum: *ibid* (2003) 377:483); e) Fabian WMF, Kalcher J, Janoschek R (2005) Theor Chem Acc 114:182
43. a) Speranza M (1998) J Phys Chem A 102:7535; b) Nelander B, Engdahl A, Svensson T (2000) Chem Phys Lett 332:403; c) Murray C, Derro EL, Sechler TD, Lester MI (2007) J Phys Chem A 111:4727
44. a) Tuppurainen K, Halonen I, Ruuskanen J (1996) Chemosphere 32:1349; b) Han J, Deming RL, Tao F-M (2004) J Phys Chem A 108:7736; c) Lee JE, Choi W, Mhin BJ, Balasubramanian K (2004) J Phys Chem A 108:607; d) Wang Z-Y, Zhai Z-C, Wang L-S, Chen J-L, Kikuchi O, Watanabe T (2004) J Mol Struct (Theochem) 672:97; e) Okamoto Y, Tomonari M (1999) J Phys Chem A 103:7686
45. Khachatryan L, Burcat A, Dellinger B (2003) Combust Flame 132:406
46. a) Dorofeeva OV, Iorish VS, Moiseeva NF (1999) J Chem Eng Data 44:516; b) Iorish VS, Dorofeeva OV, Moiseeva NF (2001) J Chem Eng Data 46:286