J. Phys. Chem. A **2009,** *113,* 11517–11534

Ab Initio Thermochemistry with High-Level Isodesmic Corrections: Validation of the ATOMIC Protocol for a Large Set of Compounds with First-Row Atoms (H, C, N, O, F)[†]

Dirk Bakowies*

Laboratory of Physical Chemistry, ETH Zürich, CH 8093 Zürich, Switzerland

Received: March 27, 2009; Revised Manuscript Received: August 13, 2009

The recently proposed ATOMIC protocol is a fully ab initio thermochemical approach designed to provide accurate atomization energies for molecules with well-defined valence structures. It makes consistent use of the concept of bond-separation reactions to supply high-level precomputed bond increments which correct for errors of lower-level models. The present work extends the approach to the calculation of standard heats of formation and validates it by comparison to experimental and benchmark level ab initio data reported in the literature. Standard heats of formation (298 K) have been compiled for a large sample of 173 neutral molecules containing hydrogen and first-row atoms (C, N, O, F), resorting to several previous compilations and to the original experimental literature. Statistical evaluation shows that the simplest implementation of the ATOMIC protocol (composite model C) achieves an accuracy comparable to the popular Gaussian-3 approach and that composite models A and B perform better. Chemical accuracy (1−2 kcal/mol) is normally achieved even for larger systems with about 10 non-hydrogen atoms and for systems with charge-separated valence structures, bearing testimony to the robustness of the bond-separation reaction model. Effects of conformational averaging have been examined in detail for the series of *n*-alkanes, and our most refined composite model A reproduces experimental heats of formation quantitatively, provided that conformational averaging is properly accounted for. Several cases of larger discrepancy with respect to experiment are discussed, and potential weaknesses of the approach are identified.

1. Introduction

Thermochemistry is a branch of thermodynamics concerned with the energy balance of chemical reactions.[1,2] The elements in their standard states define the reference used in thermochemistry, establishing heats (enthalpies) of formation as the primary quantity relating the heat content of one compound to that of another. Combustion calorimetry is the most common experimental technique to derive heats of formation for compounds in their standard states, and heats of vaporization or heats of sublimation are measured to convert condensed-phase results to those for the ideal gas state. This removal of intermolecular forces permits experimental access to chemical binding enthalpies and evaluation of such important quantities as bond dissociation enthalpies which "measure" the strength of chemical bonds. Gas-phase heats of formation also provide the link to theoretical treatments which start from a description of the chemical bond. Early work in this area was restricted to empirical approaches relating the (gas phase) heat of formation to bond and group increments and applying additivity rules.[3,4]

Significant progress in both computer technology and quantum chemical method development has made it possible, however, to use ab initio electronic structure theory for accurate thermochemical predictions.[5,6] At least for small molecules quantum chemistry can nowadays provide highly accurate estimates of the atomization energy which relates the energy of a molecule to that of the separated atoms. The gas phase heat of formation for a particular temperature T may be obtained from the atomization energy $E_{A,e}$, the zero-point vibrational energy, atomic heats of formation, and thermal enthalpy differences ($0 \rightarrow T$) for both the molecule and the elements. Atomic data are usually taken from experiment, while zero-point energies and molecular thermal enthalpy differences are evaluated quantum mechanically.

One of the major obstacles encountered in ab initio thermochemistry roots in the loss of electron correlation upon bond dissociation and the frustratingly slow convergence of electron correlation energies to the complete basis set limit. The treatment of total atomization is thus very sensitive to errors in both the one- and N-particle description and cannot benefit nearly as much from error cancellation as, e.g., the calculation of conformational energies. Fortunately, theory can still profit from a number of observations: First, convergence of electron correlation energies to the complete basis set limit is slow but fairly regular, making it possible to derive accurate extrapolation formulas. Second, coupled-cluster expansions have proven to be very successful in approaching the N-particle limit quickly, and the de facto standard CCSD(T),[7,8] which includes connected triples excitations only perturbatively, is surprisingly accurate. Third, even higher-order contributions to electron correlation energies tend to be fairly small and less basis-set dependent, allowing us to piece together the entire correlation energy from complete basis set estimates of CCSD(T) correlation energies and higher-order corrections evaluated with smaller basis sets. Further corrections need to be applied to account for relativistic effects and, if very high-accuracy is sought, also for diagonal Born-Oppenheimer terms. Several fully ab initio protocols have been developed, among them focal point analysis,[9−11] Weizmann (W_n) theory,[12−14] the HEAT model,[15−17] and the less standardized approaches presented by Dixon, Feller, and co-workers[18−24] and by Klopper and various collaborators.[25−29] The attainable level of accuracy is impressive, often reaching into the sub-kJ/mol regime. Such benchmark level quality has its price,

---

[†] Part of the "Walter Thiel Festschrift".

* E-mail bakowies@chem-edu.ethz.ch.

10.1021/jp9027782 CCC: $40.75

however, and computational demands limit applications to fairly small molecules.

Less elaborate approaches have thus found considerably more widespread use, and the Gaussian (G*n*)[30−33] composite models are probably the most popular ones among them. Residual errors are often large but so systematic that these approaches attain chemical accuracy (1−2 kcal/mol) after application of "higher-level" corrections. These corrections have initially been formulated in terms of isogyric comparisons (G1[30]), but empirical calibration with large sets of experimental reference data has become standard nowadays. The CBS models of Petersson and co-workers[34−38] are likewise very popular and similar in spirit to the Gaussian approaches, but they include complete basis-set extrapolations based on pair natural orbital expansions of the second-order correlation energy. Truhlar and co-workers introduced a number of parametrized multicoefficient methods, some of them designed as derivative versions of G2 and G3 theories,[39,40] generally mixing scaled energies of several different electron correlation[41−44] and density functional schemes.[45,46] Very recently Ohlinger et al. have reported a dramatically simplified version of the G3(MP2) protocol (T1), which, through the use of a parametrized expression based on bond orders, achieves very respectable accuracy (2 kcal/mol on average) for a large and diverse set of organic molecules.[47] From a theoretical point of view, however, the use of empirical parameters is not entirely satisfactory, no matter whether they are scaling factors or higher-level corrections based on counts of paired and unpaired electrons. The development of correlation-consistent composite approaches (ccCA)[48−50] addresses the goal of devising accurate parameter-free composite schemes which are still efficient enough computationally to be applied to systems with more than ten non-hydrogen atoms.

Recently we have introduced the ATOMIC protocol[51] (Ab initio Thermochemistry using Optimal-balance Models with Isodesmic Corrections) which follows a different strategy to achieve high accuracy without resorting to excessively expensive calculations. It is based on the concept of bond separation reactions[52,53] (BSRs) and treats the process of bond dissociation with high-level methods embedded in a lower-level treatment of the perturbation caused by the chemical environment. Application to a particular molecule only requires that a (Lewis) valence structure can be drawn, and we note that the corresponding BSR is uniquely defined by the valence structure. Avoidance of experimental reference data for the small auxiliary molecules entering BSRs not only ensures that the approach remains truly ab initio but also removes restrictions usually imposed by the limited availability of accurate reference data. We have previously demonstrated that very accurate atomization energies can be obtained with only modest computational effort.[51] The purpose of the present report is to extend the approach to the calculation of heats of formation and to validate it for a large and diverse set of experimental data, including a fair number of larger systems which test the robustness of the BSR model. To this end we have assembled a reference data set consisting of 173 experimental heats of formation (HCNOF), resorting to various standard compilations and to the more recent experimental literature and critically cross-checking data from different sources.

Section 2 summarizes the theoretical concept and section 3 provides computational details. Procedures to calculate zero-point energies and heats of formation are documented in section 4, and comments on the compilation of experimental data may be found in section 5. The validation of the theoretical approach (section 6) includes a statistical evaluation with experimental data (section 6.1), a discussion of cases with larger errors (section 6.2), a comparison with benchmark calculations on small molecules found in the literature (section 6.3), and a more detailed case study on linear alkanes, examining, among other things, the importance of conformational averaging (section 6.4).

2. Outline of the ATOMIC Protocol

The ATOMIC protocol has been discussed in much detail previously,[51] but the essential definitions are given here for the benefit of making this account self-contained. ATOMIC uses geometries optimized at the MP2(fc)/cc-pVTZ level of theory,[54,55] and it is based on a decomposition of the atomization energy $E_{\mathrm{A,e}}$ evaluated at the CCSD(T) level with all electrons correlated and extrapolated to the complete basis set limit. Three valence-shell components are defined in addition to Hartree−Fock (HF, X_1), which cover incremental contributions to the frozen-core CCSD(T) electron correlation treatment, i.e., MP2-HF (X_2), CCSD-MP2 (X_3), and CCSD(T)-CCSD (X_4). Correlation consistent basis sets, cc-pV*X*Z,[55−58] are used for their evaluation, and the four components are defined by the cardinal numbers X_1, X_2, X_3, X_4 of the basis sets employed. An analogous series of energies is defined for core-valence polarized basis sets, cc-pCV*X*Z,[59] and correlation treatments that include all core and valence electrons. Differences between the latter four and the former four components are referred to as inner-shell components and denoted by the respective cardinal numbers X_5, X_6, X_7, X_8. Components are defined such that they sum to the CCSD(T)/cc-pV*X*Z (1−4) and CCSD(T)(full)/cc-pCV*X*Z energies (1−8), respectively, provided that all cardinal numbers are chosen to be equal (X). Hence any composite model with one or several $X_n < X$ provides an approximation to the reference CCSD(T) energy. Arabic numerals are used throughout for all cardinal numbers, and numerals with top bar ($\bar{X}_n$) are used to denote energy components obtained through complete basis set extrapolation based on results for cardinal numbers $X_n - 1$ and X_n. An exponential formula[60] [$E_X = E_\infty + A\ \exp(-\alpha X)$, $\alpha = 1.63$] is used to extrapolate Hartree−Fock contributions and a power-law formula ($E_X = E_\infty + AX^{-\beta}$) with exponents calibrated for atomization energies[61] is used to extrapolate electron correlation contributions ($\beta = 2.1$ for MP2-HF or CCSD-HF using $(X - 1,X) = (2,3)$; $\beta = 2.75$ for MP2-HF and $\beta = 3.15$ for CCSD-HF using $(X - 1,X) \geq (3,4)$; $\beta = 3.15$ for CCSD(T)-CCSD)). Employing the short-hand notation $[X_1X_2X_3X_4 \mid X_5X_6X_7X_8]$, the reference level of theory and all five composite models of the ATOMIC protocol read as follows:

reference	$[\bar{6}\bar{6}\bar{6}\bar{5} \mid \bar{5}\bar{5}\bar{5}\bar{5}]$
A	$[\bar{5}\bar{5}\bar{4}\bar{3} \mid \text{-}3\ \text{-}\ \text{-}]$
B_1	$[\bar{4}\bar{4}\bar{3}\bar{3} \mid \text{-}3\ \text{-}\ \text{-}]$
B_2	$[\bar{4}\bar{4}\bar{3}\bar{3} \mid \text{-}\ \text{-}\ \text{-}\ \text{-}]$
B_3	$[\bar{4}\bar{4}3'2 \mid \text{-}3\ \text{-}\ \text{-}]$
C	$[\bar{3}\bar{3}22 \mid \text{-}\ \text{-}\ \text{-}\ \text{-}]$

The symbol 3′ denotes a mixed basis set (cc-pVDZ on H, cc-pVTZ on first-row atoms) and dashes indicate components neglected entirely. Note that the "inner-shell" HF component only reflects the difference in HF energies evaluated with core-valence polarized and regular correlation-consistent basis sets. By construction it vanishes for $X \rightarrow \infty$, and it is therefore omitted in all approximate composite models.

Bond separation reactions (BSRs) are implemented in an ab initio framework and they are used to reduce basis-set requirements for each of the eight components contributing to the

CCSD(T)(full) atomization energy at the complete basis set limit. As demonstrated previously,[51] BSRs can be mapped uniquely onto a system of corrective bond increments, thus greatly facilitating their computational implementation. Taking acetic acid as an example, the BSR reads

$$\mathrm{CH_3C({=}O)OH} + 2\mathrm{CH_4} \rightarrow \mathrm{H_2C{=}O} + \mathrm{C_2H_6} + \mathrm{CH_3OH} \qquad (1)$$

and an estimate of the high-level atomization energy (H) may be obtained from the reaction energy evaluated at a lower level (L) using precomputed high-level atomization energies of the parent molecules formaldehyde, ethane, methanol, and the simply hydride methane

$$\begin{aligned} E^{\mathrm{H}}_{\mathrm{A,e}}(\mathrm{CH_3C({=}O)OH, est.}) = -\Delta E^{\mathrm{L}}_{\mathrm{reac}} + E^{\mathrm{H}}_{\mathrm{A,e}}(\mathrm{CH_2O}) + \\ E^{\mathrm{H}}_{\mathrm{A,e}}(\mathrm{C_2H_6}) + E^{\mathrm{H}}_{\mathrm{A,e}}(\mathrm{CH_3OH}) - 2E^{\mathrm{H}}_{\mathrm{A,e}}(\mathrm{CH_4}) = \\ E^{\mathrm{L}}_{\mathrm{A,e}}(\mathrm{CH_3C({=}O)OH}) + \Delta E_{\mathrm{A,e}}(\mathrm{CH_2O}) + \Delta E_{\mathrm{A,e}}(\mathrm{C_2H_6}) + \\ \Delta E_{\mathrm{A,e}}(\mathrm{CH_3OH}) - 2\Delta E_{\mathrm{A,e}}(\mathrm{CH_4}) \qquad (2) \end{aligned}$$

Here we have used the short-hand notation $\Delta E_{\mathrm{A,e}}$ for the difference (high − low level) of atomization energies. Decomposing atomization energy differences for simple hydrides and parent molecules into bond contributions finally leads to a formulation in terms of bond contributions

$$\begin{aligned} E^{\mathrm{H}}_{\mathrm{A,e}}(\mathrm{CH_3C({=}O)OH, est.}) = E^{\mathrm{L}}_{\mathrm{A,e}}(\mathrm{CH_3C({=}O)OH}) + \\ \Delta E^{\mathrm{bond}}_{\mathrm{A,e}}(^2\mathrm{C{=}O}) + \Delta E^{\mathrm{bond}}_{\mathrm{A,e}}(^3\mathrm{C{-}C}^3) + \Delta E^{\mathrm{bond}}_{\mathrm{A,e}}(^3\mathrm{C{-}O}^1) + \\ 3\Delta E^{\mathrm{bond}}_{\mathrm{A,e}}(^3\mathrm{C{-}H}) + \Delta E^{\mathrm{bond}}_{\mathrm{A,e}}(^1\mathrm{O{-}H}) \qquad (3) \end{aligned}$$

where integer superscripts denote the number of free valencies of each atom. Note that 8 out of 11 ^{3}C−H bond terms on the reactant side (2CH_4) cancel identical terms on the product side, leaving only those three ^{3}C−H bond terms which originate from the acetic acid molecule.

Basis set requirements have been assessed carefully for each of the eight components contributing to BSR-corrected models A−C (see above) so as to balance computational effort and overall accuracy of the approach. For a set of 73 small to medium-sized neutral molecules containing hydrogen and at least three first-row atoms (C, N, O, F), root-mean-square errors with respect to the reference level of theory amount to 0.1 (A), 0.3 (B), and 1.0 kcal/mol (C), respectively. Corrections beyond CCSD(T), including scalar relativistic (mass-velocity and Darwin, MVD),[62,63] diagonal Born-Oppenheimer (DBOC),[64−66] CCSDT[67,68]-CCSD(T), and CCSDTQ[69]-CCSDT corrections are treated using a greatly simplified BSR model that assumes thermoneutral reactions for each of the contributions and renders their evaluation a trivial task of summing up bond increments. This approach has been found to be surprisingly accurate for relativistic corrections; it yields very reasonable estimates for higher-order electron correlation corrections in most cases, and it is at least qualitatively useful also for the much smaller diagonal Born-Oppenheimer correction. In order to simplify applications of the ATOMIC protocol, we list in Table 1 BSR terms, compiled from our previous study,[51] that include the energies of the atoms, atomic spin-orbit couplings, corrections toward the reference level of theory, $[\bar{6}\bar{6}\bar{6}\bar{5}|\bar{5}\bar{5}\bar{5}\bar{5}]$, and all further corrections just mentioned. Fully corrected “bottom-of-the-well” atomization energies $E_{\mathrm{A,e}}$ can then simply be obtained by subtracting the molecular total energy (example: acetic acid, B_3, −229.07823 au) from the sum of all contributing bond terms (−227.79932 au in this case, $E_{\mathrm{A,e}} = 802.53$ kcal/mol).

3. Computational Details

Geometry optimizations were performed at the RI-MP2/cc-pVTZ level of theory using the resolution of the identity (RI) approximation[70] with optimized auxiliary basis sets[71] as implemented in the TURBOMOLE suite of programs[72] (version 5.5, 2002). All optimized geometries were confirmed to be true minima by standard force-constant analysis (see, however, section 4). Single-point calculations at the MP2[54] and CCSD(T)[7,8,73] levels of theory were generally performed using the MOLPRO program.[74]

We checked the accuracy of the RI approximation for MP2/cc-pVXZ (X = Q, 5) and MP2(full)/cc-pCVXZ (X = T) calculations using the set of all 105 molecules discussed previously.[51] Auxiliary basis sets optimized for cc-pwCVXZ[75] were employed for calculations with all electrons correlated as those optimized for cc-pCVXZ were not available. The accuracy was found to be satisfactory in all cases (mean absolute/maximum errors in molecular total energies, reported in kcal/mol: 0.02/0.04 (MP2/cc-pVQZ), 0.01/0.02 (MP2/cc-pV5Z), 0.01/0.03 (MP2(full)/cc-pCVTZ)); however, auxiliary basis sets cc-pwCV(X+1)Z are needed for cc-pCVXZ calculations in order to eliminate residual errors arising from lack of specific optimization. Following these observations, some of the MP2/cc-pVXZ (X = Q, 5) and MP2(full)/cc-pCVXZ (X = T) calculations of the larger molecules discussed in section 6 were performed using the RI approximation and TURBOMOLE, when normal MP2 calculations appeared too demanding computationally.

4. Evaluation of Heats of Formation

4.1. Statistical Mechanical Procedures. Heats (enthalpies) of formation, $\Delta H_{\mathrm{f}}^0(298.15)$ are evaluated from atomization energies $E_{\mathrm{A,e}}$ using standard procedures of statistical mechanics[76]

$$\Delta H_{\mathrm{f}}^0(T) = -E_{\mathrm{A,e}} + \mathrm{ZPE} + \Delta H^0(T) + \sum_{k}^{\mathrm{atoms}} (\Delta H_{\mathrm{f},k}^{\ 0}(0) - \Delta H_k^{\ 0}(T)) \qquad (4)$$

ZPE is the zero-point energy evaluated from scaled harmonic frequencies ω_i (section 4.2) and $\Delta H^0(T)$ denotes the thermal enthalpy difference between 0 K and T, which is evaluated in the rigid-rotor and harmonic-oscillator (RRHO) approximation[77] without special treatment of low-frequency vibrations

$$\Delta H^0(T) = RT\left(4 + \sum_i \frac{\mu_i}{(\exp(\mu_i) - 1)}\right) \qquad (5)$$

(RT/2 less for linear molecules). Here, μ_i is defined as $\mu_i = hc\omega_i/k_{\mathrm{B}}T$, with h being Planck’s constant, c the speed of light, k_{B} Boltzmann’s constant, and T the temperature. The atomic heats of formation, $\Delta H_{\mathrm{f},k}^{\ 0}(0)$, and elemental thermal enthalpy differences, $\Delta H_k^{\ 0}(298.15)$, are taken from experiment (Table I of ref 76).

In a few cases (2-butyne, cyclopentane, 3-pentanone), optimized geometries are characterized by very small imaginary frequencies ($\leq 30i$ cm^{-1}), indicating numerical problems with extremely flat potential surfaces rather than the occurrence of

TABLE 1: BSR Increments for Atomization Energies (a.u.)[a]

	$[\bar{6}\bar{6}\bar{6}\bar{5}\mid\bar{5}\bar{5}\bar{5}\bar{5}]$	A	B_1	B_2	B_3	C
^{3}C−H	−9.9612101	−9.9591995	−9.9601320	−9.9479889	−9.9603393	−9.9500775
^{1}C−H	−19.4222597	−19.4185450	−19.4201369	−19.3962451	−19.4202071	−19.3985871
^{2}N−H	−18.6963389	−18.6940391	−18.6952027	−18.6778438	−18.6948074	−18.6799377
^{3}N−H (+)	−14.1472221	−14.1453891	−14.1464154	−14.1333886	−14.1463059	−14.1350022
^{1}O−H	−38.0337877	−38.0309980	−38.0324374	−38.0044250	−38.0310840	−38.0061252
^{2}O−H (+)	−25.5224863	−25.5205947	−25.5217803	−25.5030989	−25.5210572	−25.5043379
^{3}C−C^3	−18.9227290	−18.9187502	−18.9198437	−18.8956011	−18.9187444	−18.8956575
^{2}C=C^2	−37.8450694	−37.8370878	−37.8393241	−37.7907574	−37.8373572	−37.7917402
^{1}C≡C^1	−56.7674863	−56.7555379	−56.7590591	−56.6861558	−56.7557450	−56.6878069
^{3}C−N^2	−27.6579253	−27.6535149	−27.6548771	−27.6255177	−27.6533734	−27.6257772
^{2}C=N^1	−55.3153444	−55.3066182	−55.3092965	−55.2505776	−55.3066216	−55.2518046
^{1}C≡N	−82.9726449	−82.9597160	−82.9638852	−82.8755957	−82.9594167	−82.8768009
C≡N^1	−78.7867157	−78.7738379	−78.7781681	−78.6906583	−78.7742455	−78.6926856
^{2}C=N^2 (+)	−46.2171264	−46.2095618	−46.2119117	−46.1617288	−46.2097443	−46.1617189
^{3}C−N^3 (+)	−23.1087271	−23.1049165	−23.1061631	−23.0811740	−23.1048667	−23.0808532
^{1}C≡N^1 (+)	−69.3255111	−69.3139693	−69.3177162	−69.2423113	−69.3142865	−69.2425635
^{3}C−O^1	−46.9953501	−46.9903675	−46.9919347	−46.9520274	−46.9895289	−46.9529249
^{1}C−O^1	−56.4563305	−56.4496680	−56.4516856	−56.3996095	−56.4488120	−56.4006222
^{2}C=O	−93.9902015	−93.9803387	−93.9835055	−93.9035756	−93.9789460	−93.9041178
^{3}C−F	−109.1959610	−109.1894332	−109.1911701	−109.1190559	−109.1867215	−109.1202159
^{1}C−F	−118.6568813	−118.6488111	−118.6508343	−118.5666973	−118.6460305	−118.5684275
^{2}N−N^2	−36.3930245	−36.3882269	−36.3897547	−36.3552632	−36.3876581	−36.3560945
^{1}N=N^1	−72.7854278	−72.7760998	−72.7788654	−72.7099808	−72.7752548	−72.7111988
^{1}N≡N (+)	−95.5304040	−95.5181825	−95.5223466	−95.4317139	−95.5177370	−95.4319712
^{2}N=N	−81.8836993	−81.8734563	−81.8769263	−81.7988710	−81.8725031	−81.7997627
^{3}N−N^2 (+)	−31.8438239	−31.8395543	−31.8413447	−31.8112720	−31.8395342	−31.8113950
^{2}N−O^1	−55.7302489	−55.7248403	−55.7265766	−55.6815928	−55.7234577	−55.6834153
^{1}N=O	−111.4600158	−111.4495441	−111.4530043	−111.3629107	−111.4473037	−111.3635783
^{2}N=O (+)	−102.3618297	−102.3524527	−102.3558775	−102.2744690	−102.3507605	−102.2741764
^{3}N−O	−88.7151759	−88.7074219	−88.7092709	−88.6406044	−88.7046640	−88.6407600
^{3}N−O^1 (+)	−51.1811570	−51.1762960	−51.1783355	−51.1377538	−51.1755673	−51.1386316
^{2}N−F	−117.9307134	−117.9236604	−117.9256469	−117.8484657	−117.9205105	−117.8503002
^{3}N−F (+)	−113.3816469	−113.3752081	−113.3774454	−113.3046892	−113.3726213	−113.3061776
^{1}O−O^1	−75.0673678	−75.0614064	−75.0633376	−75.0077641	−75.0591285	−75.0108301
^{2}O−O	−100.0899392	−100.0819376	−100.0841890	−100.0098244	−100.0781790	−100.0108725
^{1}O=O (+)	−125.1114362	−125.1020691	−125.1054467	−125.0131941	−125.0989448	−125.0119419
^{1}O−F	−137.2677583	−137.2601739	−137.2624211	−137.1745454	−137.2560577	−137.1772016

[a] The atomization energy is obtained by adding to the negative total energy of a molecule the sum of all BSR increments as listed. Increments include atomic energies, all corrections toward the reference level of theory (for A, B_n, and C) as well as MVD, DBOC, CCSDTQ-CCSD(T), and atomic spin-orbit terms. To avoid round-off errors in calculations of atomization energies, increments are reported to 10^{-7} a.u., although actual numerical precision is less.

transition states.[78] Two essentially isoenergetic ($\Delta E \leq 20\mu E_h$) stationary points are obtained for 2-butyne (D_{3h} and D_{3d}) and cyclopentane (C_s and C_2), each characterized by one small imaginary frequency. This problem is irrelevant for the evaluation of molecular energies and zero-point energies but requires some attention in the evaluation of thermal enthalpies. As μ_i approaches 0, the vibrational contribution to thermal enthalpy differences tends to its maximum value of RT (see eq 5). In order to avoid inconsistencies with cases where large-amplitude motions are characterized by small but real frequencies, we add a term of RT also when such modes correspond to small but imaginary frequencies. This procedure only restores consistency with the RRHO treatment, it does not cure for the qualitatively poor treatment of low-frequency vibrations. Obviously, the torsional mode of 2-butyne would be better described as free internal rotor which contributes $RT/2$ (rather than RT) to the thermal enthalpy,[79] but in other circumstances low-frequency modes may require a more specialized treatment,[80−82] depending on the type of normal coordinate involved and the associated barrier height. In most cases, however, errors introduced through the RRHO treatment should be small, and we follow the simpler protocol, also implemented in the Gaussian (G*n*) approaches,[32,76,83] to treat all vibrations as harmonic.

4.2. Harmonic Frequency Scaling. Harmonic frequencies are typically overestimated by about 10% at the Hartree−Fock level of theory and somewhat less if electron correlation is taken into account to second order (MP2).[84] Fortunately, errors tend to be fairly systematic so that scaling factors can be used for accurate estimates of harmonic or fundamental frequencies.[85]

In the harmonic-oscillator approximation, zero-point energies are evaluated as sums over harmonic frequencies, $1/2\sum_i \omega_i$, but the anharmonicity of molecular vibrations affords further terms.[79,86] Accurate work requires consideration of anharmonicity at least in an average sense, and scaling factors specifically derived for the calculation of zero-point energies have been reported for HF/6-31G* and MP2/6-31G*,[87] later for HF, MP2, QCISD, and density functional calculations using various basis sets of Pople and co-workers,[79] and more recently for calculations using correlation-consistent basis sets.[88] Scaling factors are commonly derived by linear least-squares optimization, using experimental, sometimes also high-level ab initio, zero-point energies as reference.

Unfortunately, only very few reference data exist for the calibration of harmonic frequencies to reproduce zero-point energies. Ethane appears to be the largest organic molecule for which an accurate ab initio value has been reported,[89] and most available data refer to di- and triatomic molecules. This poses a significant problem for statistical evaluations, as they are unable to predict the accuracy for larger molecules in which we are interested primarily. The best one can do in this situation

TABLE 2: Zero-Point Energies and Frequency Scaling Factors[a]

		ZPE[b]			ZPE[c]		error in ZPE	
		ref			calcd	f_{scal}^{opt}[d]	$f_{scal} = 1$	$f_{scal} = 0.98$
CF_4	tetrafluoromethane	10.86	a	*	10.95	0.992	0.09	−0.13
CHN	hydrogen cyanide	9.95	b	*	9.92	1.003	−0.03	−0.23
CHN	hydrogen isocyanide	9.71	b		9.82	0.989	0.11	−0.09
CHNO	isocyanic acid, HNCO	13.37	a		13.37	1.000	0.00	−0.27
CH_2	singlet methylene	10.33	b		10.70	0.965	0.37	0.16
CH_2O	formaldehyde	16.52	b	*	16.91	0.977	0.39	0.05
CH_4	methane	27.71	b	*	28.49	0.973	0.78	0.21
CO	carbon monoxide	3.09	b		3.03	1.020	−0.06	−0.12
CO_2	carbon dioxide	7.24	b		7.25	0.999	0.01	−0.14
C_2H_2	acetylene	16.44	b	*	16.65	0.987	0.21	−0.12
C_2H_2O	ketene	19.49	a		19.86	0.981	0.37	−0.03
C_2H_4	ethylene	31.60	b	*	32.29	0.979	0.69	0.04
C_2H_6	ethane	46.39	c	*	47.49	0.977	1.10	0.15
FH	hydrogen fluoride	5.85	b		5.98	0.978	0.13	0.01
FHO	hypofluorous acid	8.58	b		8.76	0.979	0.18	0.00
F_2	fluorine	1.30	b		1.45	0.897	0.15	0.12
HNO	nitroxyl, HNO	8.55	b	*	8.72	0.981	0.17	0.00
H_2	hydrogen	6.20	b		6.47	0.958	0.27	0.14
H_2N_2	diazene (trans)	17.53	a	*	17.87	0.981	0.34	−0.02
H_2O	water	13.25	b	*	13.56	0.977	0.31	0.04
H_2O_2	hydrogen peroxide	16.25	d	*	16.75	0.970	0.50	0.17
H_3N	ammonia	21.27	b	*	21.85	0.973	0.58	0.14
N_2	nitrogen	3.36	b		3.14	1.070	−0.22	−0.28
N_2O	nitrous oxide	6.78	a		6.76	1.003	−0.02	−0.16
O_3	ozone	4.15	b		5.93	0.700	1.78	1.66
MUE/rms (selected)[e]							0.43/0.53	0.11/0.13
MUE/rms (all except ozone)[e]							0.30/0.40	0.12/0.14

[a] All zero-point energies (ZPEs) are given in kcal/mol. [b] Experimental or high-level theoretical value for the total (harmonic+anharmonic) ZPE, taken from compilations. References are as follows: (a) ref 22, (b) ref 90, (c) ref 89, (d) ref 25. [c] Harmonic ZPE calculated at the RI-MP2/cc-pVTZ level. [d] Optimal scaling factor. [e] Mean unsigned (MUE) and root-mean-square (rms) errors for set of 12 selected molecules (as indicated by asterisks) and entire set of 24 molecules (except ozone).

is to analyze optimal scaling factors for individual molecules and to select those molecules for calibration which are thought to be the most representative ones. Table 2 lists results for neutral closed-shell HCNOF molecules with reference data taken from the compilations of Bak et al.,[25] Helgaker et al.,[90] and Feller and Dixon,[22] and the recent study of Karton et al.[89]

Upon inspection of Table 2 one immediately notes the gross overestimation of the zero-point energy for ozone. This error relates to the multireference character of ozone and to the known failure of MP2[87]—and to a lesser extent of several other electron correlation methods as well[91]—to reproduce the antisymmetric-stretch frequency (RI-MP2/cc-pVTZ 2239 cm^{-1} vs experiment,[92] 1089 cm^{-1}).

Apart from this exceptional case, all computed scaling factors fall in the range of 0.9−1.1. If one excludes the elements, i.e., molecular hydrogen, nitrogen, and fluorine, whose bonds are not usually found in other neutral molecules, the observed range narrows down to about 0.96−1.00. Most "typical" organic and inorganic molecules with bonds often found in larger compounds have scaling factors of 0.97−0.99. This suggests a compromise value of $f_{scal} = 0.98$. Averaging scaling factors over the subset of 12 selected molecules judged to be most representative (as marked in Table 2) yields $f_{scal} = 0.981$, least-squares optimization $f_{scal} = 0.978$ for the same subset and also for the entire set (excluding ozone).

The analysis intentionally introduces a bias toward molecules with X−H bonds (X = C, N, O, F), as these are not only characteristic of most larger molecules but also associated with the highest frequencies, and thus potentially the source of large error in estimates of the zero-point energy. It is important to realize that even a fairly small relative error of only 1% introduces errors of 0.5 and 1.0 kcal/mol, respectively, in the heats of formation of ethane and *n*-pentane, two molecules which cannot be considered large by organic chemistry standards. In light of this, it is reassuring to see that the suggested scaling factor $f_{scal} = 0.98$ performs quite well for the most important prototype molecules containing X−H, C−C, C═C, C≡C, and C═O bonds (Table 2).

Sinha et al. have recently determined scaling factors by least-squares fitting to zero-point energies of 24 mostly di- and triatomic species, including several molecules with atoms from the second row (Si, P, S, Cl).[88] We note that their optimized value (0.9832) is reasonably close to the one suggested here, although it appears a little too high for many larger (organic) molecules whose zero-point energies will often be dominated by C−H stretch frequencies.

The same frequency scaling factor (0.98) is also used for vibrational contributions to thermal enthalpy differences. This simplification appears justified as thermal enthalpy contributions to heats of formation at 298.15 K are typically smaller than zero-point energy contributions by at least an order of magnitude. Furthermore, the work of Sinha et al. indicates that explicitly optimized scaling factors for thermal enthalpies and zero-point energies are quite similar in the case of MP2/cc-pVTZ.[88]

A common scaling factor (0.8929) for zero-point energy and thermal contributions is also employed in the Gaussian (G1, G2, G3) approaches,[30−32,76] all of which use vibrational frequencies evaluated at the HF/6-31G* level. This scaling factor is apparently more appropriate for estimates of experimental fundamental frequencies (optimized, 0.8953[87]) than for calculations of zero-point energies (optimized, 0.91,[86] 0.9135[87]).

5. Experimental Reference Data

The assembly of a suitable experimental reference data set should reflect the diversity of bonding situations encountered in the inorganic and organic chemistry of the five elements considered (H, C, N, O, F). Additionally, the data set needs to include a substantial number of larger compounds as only this allows us to see if the proposed computational model is robust and reliable enough to treat species whose bond separation reactions involve many reference (prototype) molecules and thus demand exceptionally high accuracy of the reference level of theory.

Standard test and calibration sets have been compiled for the development of the Gaussian (G*n*) approaches; they are commonly used also for the assessment of alternative thermochemical techniques.[49,93,94] The early version of this calibration set, G2/97,[31,76,95] certainly contains too few larger molecules, but the more recent version, G3/99,[83] provides a useful starting point, while the most recent version, G3/05,[96] mainly adds data for second- and third-row compounds. We have selected the 123 neutral, closed-shell molecules (HCNOF) found in the G3/99 set, and supplemented them with another 50 molecules, mostly larger organic species, but also some smaller systems which extend the range of bond types and were considered previously in the calculation of atomization energies.[51] The largest systems contain 10 or more non-hydrogen atoms and involve up to 34 prototypes (anthracene) in their bond separation reactions.

Experimental heats of formation were taken primarily from the book of Pedley et al.,[97] but two further standard compilations were consulted in all cases.[98,99] Original publications including the more recent literature were looked up in cases of discrepancy, doubt, or lack of data, but cited[100−112] only if the compilations did not provide the required data or were believed to be in error. The evaluations of Gurvich et al.[113] and Chase et al. (JANAF tables)[114] served as additional references mostly for smaller inorganic systems. Not considered were data from purely electronic databases or those that are biased by theoretical input. Discrepancy with ab initio calculations, even at benchmark levels, is at this point not considered reason to revise or reject an experimental value but will be discussed later on (section 6).

In the process of data collection we made the experience that a reportedly small error bar does not necessarily establish an accurately known experimental value. Da Silva et al. recently pointed out that values commonly suggested for formaldehyde and acetaldehyde are disputed although reported error bars suggest they are accurately known.[115] A further example is cubane, for which the well-respected and very careful compilation of Pedley et al. suggests $\Delta H_f^0(298.15) = 148.7 \pm 0.9$ kcal/mol,[97] based on a heat of sublimation (19.2 ± 0.4 kcal/mol)[100] that was recently, after substantial criticism,[116,117] redetermined to be 13.2 ± 0.5 kcal/mol.[111] Unfortunately, the quality of the heat of combustion, quoted as 129.5 ± 0.8 kcal/mol,[100] is also not firmly established,[118] but it has not been redetermined, and the only further available account is indirect and biased by (fairly low-level) quantum-chemical calculations.[119] Carbonyl fluoride is listed with a value of −152.9 ± 0.4 kcal/mol by Pedley et al.[97] and −152.7 ± 0.4 kcal/mol in the JANAF tables,[114] but a more recent investigation[120] shows that it should be ≥−149.1 kcal/mol. Two quite different values for $\Delta H_f^0(298.15)$, both with small error bars, may also be found for hypofluorous acid: −23.5 ± 1.0 kcal/mol,[114] −19.6 ± 0.3 kcal/mol,[99] and the most accurate value still appears to differ by more than 1 kcal/mol from either tabulated value.[109] The generally accepted heat of formation of urea (−58.5 ± 0.5 kcal/mol)[97,98] is based on a fairly old value for the heat of sublimation (21.5 ± 0.5 kcal/mol),[121] while more recent measurements using two different techniques yield values of 23.3 ± 0.2 and 22.6 ± 0.1 kcal/mol, respectively. The authors discuss differences between the techniques and suggest ΔH_f^0 (gas) = −56.3 ± 0.3 kcal/mol.[106] A fairly drastic revision has recently been suggested for the heat of formation of pyridine-*N*-oxide, from 21.0 ± 0.6 kcal/mol as found in a compilation[99] to 29.8 ± 0.5 kcal/mol.[112] And finally we note that three different values of −29.8 ± 0.2 kcal/mol,[97] −31.0 ± 0.3 kcal/mol,[98] and −31.6 ± 0.3 kcal/mol[99,114] are offered by four different compilations for trifluoroamine. These examples are not meant to discredit any of the very valuable compilations of thermochemical data, but they show that considerable care needs to be taken, that different sources should be consulted, and that reported error bars do not always reflect the true accuracy of a value.

In this sense we have made the “best effort” to update previous compilations for the selected set of 173 molecules. In an attempt to achieve a larger diversity of reference data, but also reflecting the experiences detailed above, we have admitted reference data with larger quoted uncertainties (up to 3 kcal/mol) than is standard, noting however, that the vast majority still adheres to the criterion (1 kcal/mol) adopted by the G2/97 and G3/99 test sets.[83] The entire list of molecules, along with experimental heats of formation and ATOMIC results (quoted as deviation “theory-experiment”) for models A, B_1, B_2, B_3, and C may be found in Table 3. For reference, G3 results, obtained with the Gaussian98 series of programs,[122] are given as well. For most of the 123 molecules of the G3/99 subset, previously published heats of formation[32,83] were reproduced to within 0.1 kcal/mol (roundoff errors). Larger discrepancies were found only for phenol (−22.0 vs −21.4), aniline (21.5 vs 22.1), and toluene (12.4 vs 12.9). In each of the three cases, our recomputed heats of formation are lower by 0.5−0.6 kcal/mol and in closer agreement with experiment. Note that the more appropriate statistical-mechanical treatment of the low-frequency torsional mode (around the substituent to benzene axis) as internal rotation would have led to even lower heats of formation.

6. Results: Heats of Formation

In this section, we assess the accuracy of the new thermochemical protocol with comparisons to experimental heats of formation. Statistical analyses are provided first (section 6.1), followed by a fairly detailed discussion of cases with larger error (section 6.2) and comparisons to other theoretical benchmark calculations (section 6.3). Finally, linear *n*-alkanes serve as test cases to study the effects of conformational averaging, the relative importance of various contributions to heats of formation and possible sources of residual error (section 6.4).

6.1. Statistical Analysis. Table 4 provides statistical analyses for ATOMIC models A, B_1, B_2, B_3, C, and the Gaussian protocol G3. Results are reported for our standard approach with all BSR corrections included (columns 1−5) and for raw data obtained from compound methods A to C, corrected only for relativistic, diagonal Born-Oppenheimer, and higher order excitation terms (columns 6−10). In this way, comparisons can be made on an equal footing with the analysis of CCSD(T) atomization energies shown in Table I of ref 51. The first block reports analyses for the complete set of 165 molecules, excluding only 8 molecules for which experimental data are suspect (see below), the second block shows results for the subset of molecules for which level A calculations were performed, and the third and fourth blocks further restrict the data set to molecules with at least three non-hydrogen atoms. The latter data set is the most relevant one, as

TABLE 3: Heats of Formation, ΔH_f^0(298.15)[a,b]

		experiment[c,d]			A	B_1	B_2	B_3	C	G3
CF_2	singlet difluoromethylene	−44.0	±2.0	a	−2.2	−2.4	−2.4	−2.2	−1.7	−2.6
CF_2O	carbonyl fluoride	−149.1	±1.4	a*	4.1	3.7	4.0	4.5	5.2	3.4
CF_4	tetrafluoromethane	−223.1	±0.3	b	−0.5	−0.7	−0.4	0.2	1.3	−0.8
CHF	singlet fluoromethylene	34.2	±3.0	a	1.4	1.4	1.4	1.4	1.4	0.7
CHF_3	trifluoromethane	−166.6	±0.8	a	−0.2	−0.4	−0.2	0.2	0.3	−0.5
CHN	hydrogen cyanide	31.5	±1.0	a	−0.5	−0.5	−0.5	−0.5	−0.5	−0.2
CHN	hydrogen isocyanide, HNC	46.3	±2.2	c	−0.3	−0.3	−0.3	−0.3	−0.3	−0.3
CHNO	isocyanic acid, HNCO	−24.3	±2.0	a*	−3.8	−4.3	−3.8	−3.9	−2.7	−4.5
CH_2	singlet methylene	101.8	±0.5	a	0.9	0.9	0.9	0.9	0.9	0.1
CH_2F_2	difluoromethane	−107.7	±0.4	a	−0.2	−0.3	−0.3	−0.1	−0.3	−0.7
CH_2N_2	3*H*-diazirine	63.3	±2.7	d*	11.9	12.2	12.1	12.3	11.2	12.9
CH_2O	formaldehyde	−26.0	±0.1	b	−0.1	−0.1	−0.1	−0.1	−0.1	−0.6
CH_2O_2	formic acid	−90.5	±0.1	b	0.1	−0.1	0.1	0.3	0.6	−0.1
CH_3NO_2	methylnitrite	−15.9	±0.2	b	0.4	0.5	0.6	0.8	1.6	0.2
CH_3NO_2	nitromethane	−17.8	±0.1	b	1.2	0.4	0.7	0.8	0.8	0.0
CH_4	methane	−17.8	±0.1	b	0.0	0.0	0.0	0.0	0.0	−0.3
CH_4N_2O	urea	−56.3	±0.3	e	0.4	− 0.2	0.4	0.5	1.2	1.3
CH_4O	methanol	−48.2	±0.1	b	0.1	0.1	0.1	0.1	0.1	0.1
CH_5N	methylamine	−5.5	±0.1	b	0.5	0.5	0.5	0.5	0.5	1.0
CO	carbon monoxide	−26.4	±0.0	b	0.1	0.1	0.1	0.1	0.1	−0.3
CO_2	carbon dioxide	−94.1	±0.0	b	0.3	−0.2	0.1	0.2	1.7	−1.2
C_2F_3N	trifluoroacetonitrile	−119.0	±0.7	b	0.4	−0.2	0.1	0.4	0.3	−1.2
C_2F_4	tetrafluoroethylene	−157.5	±1.2	b*	−3.3	−3.6	−3.0	−2.9	−2.3	−4.8
C_2F_6	hexafluoroethane	−321.3	±1.3	b	0.2	−0.8	−0.4	0.4	1.0	−2.8
C_2HF_3	trifluoroethylene	−117.2	±2.0	b	−1.8	−2.1	−1.6	−1.5	−1.2	−2.9
$C_2HF_3O_2$	trifluoroacetic acid	−246.5	±0.4	b	1.5	0.6	1.1	1.7	2.5	−0.5
C_2H_2	acetylene	54.2	±0.2	a	0.3	0.3	0.3	0.3	0.3	0.7
$C_2H_2F_2$	1,1-difluoroethylene	−80.1	±1.1	b	−3.7	−3.9	−3.6	−3.5	−3.3	−4.4
C_2H_2O	ketene	−11.4	±0.4	b	0.0	−0.5	−0.2	−0.1	0.6	−0.7
$C_2H_2O_2$	glyoxal	−50.7	±0.2	b	0.0	−0.3	−0.1	0.0	0.2	−0.9
C_2H_3F	fluoroethylene	−33.2	±0.4	b	−0.9	−1.0	−0.9	−0.8	−0.8	−1.2
C_2H_3FO	acetyl fluoride	−105.7	±0.8	b	0.7	0.3	0.6	0.8	1.3	−0.1
C_2H_3N	acetonitrile	17.7	±0.1	a	0.2	0.1	0.2	0.0	0.1	0.1
C_2H_3N	methyl isocyanide	41.4	±0.2	f	1.3	1.4	1.5	1.2	1.2	0.8
C_2H_4	ethylene	12.5	±0.1	b	−0.1	−0.1	−0.1	−0.1	−0.1	−0.2
C_2H_4O	acetaldehyde	−39.7	±0.1	b	0.2	0.0	0.2	0.1	0.4	−0.1
C_2H_4O	oxirane	−12.6	±0.1	b	0.2	0.6	0.7	0.6	0.0	0.0
$C_2H_4O_2$	acetic acid	−103.4	±0.4	b	0.5	0.1	0.4	0.5	1.0	0.1
$C_2H_4O_2$	methyl formate	−85.0	±0.2	b	−0.7	−0.7	−0.5	−0.4	−0.2	−1.6
C_2H_5N	aziridine	30.2	±0.2	b	0.4	0.7	0.8	0.7	−0.4	1.2
C_2H_5NO	acetamide	−57.0	±0.2	b	0.8	0.5	1.0	0.9	1.5	1.1
$C_2H_5NO_2$	glycine	−93.7	±0.1	b	−0.7	−1.3	−0.9	−0.8	−0.5	−0.4
C_2H_6	ethane	−20.0	±0.1	b	−0.2	−0.2	−0.2	−0.2	−0.2	−0.4
$C_2H_6N_2O$	dimethylnitrosamine	−0.8	±1.6	g*	15.2	15.1	15.5	15.3	15.9	15.3
C_2H_6O	dimethyl ether	−44.0	±0.1	b	−0.1	0.1	0.2	0.0	0.1	−0.4
C_2H_6O	ethanol	−56.2	±0.1	b	0.0	−0.2	−0.1	−0.1	−0.3	−0.1
C_2H_7N	dimethylamine	−4.4	±0.2	b	0.6	0.6	0.6	0.5	0.4	0.9
C_2H_7N	ethylamine	−11.3	±0.2	b	−0.6	−0.8	−0.7	−0.8	−0.8	0.0
C_2N_2	cyanogen	73.3	±0.2	b	1.6	1.4	1.5	1.2	0.2	0.3
C_3H_3N	acrylonitrile	43.2	±0.4	b	2.0	1.8	1.9	1.7	1.3	1.6
$C_3H_3N_3$	1,3,5-triazine	54.0	±0.2	a	0.3	0.2	0.9	0.7	0.5	0.5
C_3H_4	allene	45.5	±0.3	b	−0.1	−0.4	−0.2	−0.4	−0.6	−0.5
C_3H_4	cyclopropene	66.2	±0.6	b	1.6	1.7	1.8	1.7	1.1	2.2
C_3H_4	propyne	44.2	±0.2	b	0.2	0.0	0.2	−0.1	0.1	0.2
C_3H_6	cyclopropane	12.7	±0.1	b	0.2	0.4	0.7	0.5	−0.4	0.7
C_3H_6	propene	4.8	±0.2	b	−0.1	−0.2	−0.1	−0.2	−0.2	−0.1
C_3H_6O	acetone	−51.9	±0.2	b	0.1	−0.2	0.1	0.0	0.5	−0.1
C_3H_6O	oxetane	−19.2	±0.1	b	0.2	0.5	0.5	0.5	−0.3	0.0
$C_3H_6O_2$	methyl acetate	−98.4	±0.4	b	0.5	0.3	0.7	0.6	0.8	−0.7
$C_3H_6O_3$	1,3,5-trioxane	−111.4	±0.1	b	−0.4	−0.1	0.2	0.1	−0.6	−1.7
C_3H_8	propane	−25.0	±0.1	b	−0.3	−0.4	−0.4	−0.4	−0.4	−0.3
C_3H_8O	isopropanol	−65.2	±0.1	b	−0.2	−0.6	−0.4	−0.5	−0.6	−0.5
C_3H_8O	methoxyethane	−51.7	±0.2	b	−0.6	−0.6	−0.4	−0.6	−0.7	−1.1
C_3H_9N	trimethylamine	−5.7	±0.2	b	−0.4	−0.3	−0.1	−0.4	−0.7	−0.2
C_3O_2	carbon suboxide	−22.4	±0.4	a	2.2	1.1	1.6	1.9	2.4	−1.5
C_4H_4	vinylacetylene	70.8	±0.5	a	−1.4	−1.7	−1.5	−1.8	−2.0	−1.5
$C_4H_4N_2$	butanedinitrile	50.1	±0.2	b	0.6	0.1	0.5	0.0	−0.3	0.2
$C_4H_4N_2$	pyrazine	46.9	±0.3	b	2.6	2.5	3.2	2.9	2.1	2.7
$C_4H_4N_2$	pyrimidine	46.8	±0.3	b	−1.7	−1.9	−1.2	−1.5	−2.0	−1.7
C_4H_4O	acetyl acetylene	15.7	±0.2	d	3.2	2.8	3.1	2.8	2.8	2.4
C_4H_4O	furan	−8.3	±0.2	b	−0.1	0.1	0.6	0.4	0.4	0.5

TABLE 3: Continued

		experiment^{c,d}			A	B_1	B_2	B_3	C	G3
C_4H_5N	pyrrole	25.9	±0.1	b	−0.2	−0.1	0.6	0.2	0.3	1.2
C_4H_6	1,2-butadiene	38.8	±0.1	b	0.1	−0.3	0.0	−0.3	−0.5	−0.2
C_4H_6	1,3-butadiene	26.3	±0.3	b	0.3	0.1	0.3	0.1	0.0	0.4
C_4H_6	2-butyne	34.8	±0.3	b	0.8	0.3	0.7	0.1	0.5	0.5
C_4H_6	bicyclo[1.1.0]butane	51.9	±0.2	b	1.9	2.1	2.4	2.0	0.4	2.6
C_4H_6	cyclobutene	37.5	±0.4	b	0.9	0.8	0.9	0.8	0.6	2.0
C_4H_6	methylenecyclopropane	47.9	±0.4	b	−1.6	−1.7	−1.3	−1.7	−2.4	−1.5
C_4H_6O	crotonaldehyde	−24.0	±0.3	b	−0.5	−0.8	−0.5	−0.6	−0.6	−0.9
C_4H_6O	divinyl ether	−3.3	±0.2	b	0.5	0.5	0.9	0.8	0.7	0.3
$C_4H_6O_3$	acetic anhydride	−136.8	±0.4	b	−0.4	−0.9	−0.2	−0.1	0.5	−2.2
C_4H_7N	isobutylnitrile	5.6	±0.3	b	1.5	1.2	1.4	1.1	0.9	1.1
C_4H_8	cyclobutane	6.8	±0.1	b	−0.4	−0.5	−0.5	−0.5	−1.0	0.0
C_4H_8	isobutene	−4.0	±0.2	b	−0.1	−0.3	−0.1	−0.4	−0.4	0.0
C_4H_8O	2-butanone	−57.1	±0.2	b	0.1	−0.4	−0.1	−0.2	0.1	−0.2
C_4H_8O	isobutanal	−51.6	±0.2	b	1.2	0.7	1.0	0.9	0.9	0.6
C_4H_8O	tetrahydrofuran	−44.0	±0.2	b	0.5	0.3	0.5	0.3	−0.1	0.2
$C_4H_8O_2$	1,4-dioxane	−75.5	±0.2	b	−0.1	−0.2	0.1	−0.2	−0.6	−0.9
C_4H_9N	pyrrolidine	−0.8	±0.2	b	0.0	−0.4	−0.2	−0.3	−0.6	0.7
$C_4H_9NO_2$	2-nitrobutane	−39.1	±0.4	b		−0.1	0.3	0.4	−0.8	−1.1
C_4H_{10}	isobutane	−32.1	±0.2	b	−0.2	−0.4	−0.3	−0.5	−0.4	−0.2
C_4H_{10}	*n*-butane	−30.0	±0.2	b	−0.5	−0.7	−0.6	−0.7	−0.6	−0.4
$C_4H_{10}O$	diethyl ether	−60.3	±0.2	b	−0.1	−0.3	−0.1	−0.3	−0.6	−0.8
$C_4H_{10}O$	*tert*-butanol	−74.7	±0.2	b	−0.1	−0.7	−0.4	−0.7	−0.9	−0.8
$C_4H_{10}O_2$	1,1-dimethoxyethane	−93.1	±0.2	b	−0.2	−0.3	0.0	−0.2	−0.8	−1.8
$C_4H_{11}N$	*tert*-butylamine	−28.9	±0.2	b	−0.1	−0.7	−0.4	−0.8	−0.8	0.1
C_4N_2	dicyanoacetylene	126.5	±0.3	b	2.8	2.2	2.6	2.0	0.0	0.2
C_5H_5N	pyridine	33.6	±0.2	b	−0.1	−0.3	0.4	0.0	−0.6	0.1
C_5H_5NO	pyridine-*N*-oxide	29.8	±0.5	h	1.5	1.0	1.7	1.1	−0.9	0.5
C_5H_7N	1-methyl-1*H*-pyrrole	24.6	±0.1	b	0.1	0.1	0.9	0.3	−0.2	0.8
C_5H_8	isoprene	18.0	±0.3	b	0.2	0.0	0.2	−0.1	−0.5	0.3
C_5H_8	spiropentane	44.3	±0.2	b	0.0	0.0	0.6	0.0	−1.1	0.4
C_5H_8O	cyclopentanone	−45.9	±0.4	b	−0.5	−1.1	−0.7	−0.9	−0.6	−0.8
C_5H_{10}	cyclopentane	−18.3	±0.2	b	0.1	−0.3	−0.2	−0.3	−0.4	0.5
$C_5H_{10}O$	3-pentanone	−61.6	±0.2	b	−0.7	−1.3	−0.9	−1.0	−1.1	−1.1
$C_5H_{10}O$	tetrahydro-2*H*-pyran	−53.4	±0.2	b	0.2	−0.1	0.2	0.0	−0.4	−0.3
$C_5H_{10}O_2$	isopropyl acetate	−115.1	±0.2	b		−0.1	0.4	0.3	0.1	−1.4
$C_5H_{11}N$	piperidine	−11.3	±0.1	b	0.5	0.1	0.4	0.1	−0.2	0.9
C_5H_{12}	*n*-pentane	−35.1	±0.2	b	−0.4	−0.8	−0.7	−0.8	−0.7	−0.4
C_5H_{12}	neopentane	−40.2	±0.2	b	−0.1	−0.6	−0.3	−0.7	−0.7	−0.5
$C_5H_{12}O$	*tert*-butyl methyl ether	−67.8	±0.3	b	−0.1	−0.5	−0.2	−0.6	−1.2	−1.4
C_5N_4	tetracyanomethane	160.8	±2.2	b*	7.2	6.7	7.1	6.1	3.3	4.2
C_6F_6	hexafluorobenzene	−228.3	±0.4	b	1.0	−0.3	1.2	0.6	0.9	−4.2
$C_6H_4F_2$	1,3-difluorobenzene	−73.9	±0.2	b	0.7	−0.1	0.9	0.4	−0.1	−0.5
$C_6H_4F_2$	1,4-difluorobenzene	−73.3	±0.2	b	0.7	−0.1	0.9	0.4	−0.2	−0.4
$C_6H_4O_2$	benzoquinone	−29.4	±0.8	b	2.0	1.3	2.0	1.7	1.4	1.2
C_6H_5F	fluorobenzene	−27.7	±0.3	b	0.5	−0.1	0.7	0.2	−0.2	0.3
C_6H_6	benzene	19.7	±0.2	b	−0.2	−0.5	0.2	−0.4	−0.6	0.7
C_6H_6O	phenol	−23.0	±0.2	b	0.7	0.2	1.0	0.5	−0.1	1.0
C_6H_7N	aniline	20.8	±0.2	b	−0.2	−0.7	0.2	−0.5	−0.9	0.7
C_6H_8	1,3-cyclohexadiene	25.4	±0.2	b	0.1	−0.2	0.3	−0.3	−0.4	0.8
C_6H_8	1,4-cyclohexadiene	25.9	±0.5	i	−0.3	−0.8	−0.3	−0.8	−1.0	0.5
C_6H_{10}	cyclohexene	−1.2	±0.2	b		−0.3	0.1	−0.3	−0.6	0.4
C_6H_{12}	cyclohexane	−29.5	±0.2	b	0.0	−0.5	−0.3	−0.5	−0.5	0.2
$C_6H_{12}N_4$	tetraazaadamantane	47.5	±0.7	b	0.1	0.3	1.2	0.2	−2.9	−0.8
C_6H_{14}	3-methylpentane	−41.1	±0.2	b	−0.1	−0.6	−0.4	−0.6	−0.9	−0.2
C_6H_{14}	*n*-hexane	−39.9	±0.2	b	−0.7	−1.2	−1.1	−1.3	−1.2	−0.6
$C_6H_{14}O$	diisopropyl ether	−76.3	±0.4	b	−0.3	−0.8	−0.4	−0.7	−1.4	−1.5
C_7H_5N	benzonitrile	51.6	±0.5	b	0.4	−0.1	0.7	0.0	−1.2	0.1
C_7H_6O	benzaldehyde	−8.8	±0.7	b		−0.5	0.4	−0.2	−1.1	−0.2
$C_7H_6O_2$	benzoic acid	−70.3	±0.5	b		−1.1	0.0	−0.5	−1.2	−0.9
C_7H_8	2,5-norbornadiene	58.8	±0.7	b	−1.5	−1.8	−1.4	−1.9	−2.8	−1.3
C_7H_8	quadricyclane	81.0	±0.6	b	−1.0	−1.2	−0.5	−1.2	−3.3	0.1
C_7H_8	toluene	12.0	±0.1	b	0.0	−0.5	0.3	−0.4	−0.9	0.4
$C_7H_{10}O$	norbornan-7-one	−32.1	±0.7	b	−2.1	−3.0	−2.5	−2.8	−3.3	−3.2
C_7H_{12}	norbornane	−12.8	±0.3	j	−0.7	−1.4	−1.0	−1.4	−2.0	−0.9
C_7H_{16}	*n*-heptane	−44.9	±0.3	b	−0.9	−1.5	−1.3	−1.5	−1.5	−0.7
C_8H_7N	1*H*-indole	37.4	±0.3	b	1.4	1.1	2.4	1.4	0.4	2.0
C_8H_8	1,3,5,7-cyclooctatetraene	70.7	±0.4	b		0.4	1.1	0.2	−1.1	1.4
C_8H_8	cubane	142.7	±0.9	k	2.1	2.2	2.4	2.2	−0.7	4.7
C_8H_8	styrene	35.3	±0.4	b		−0.2	0.6	−0.2	−1.2	0.6
C_8H_{14}	bicyclo[2.2.2]octane	−23.7	±0.3	b		0.0	0.4	0.0	−0.8	0.2

TABLE 3: Continued

		experiment[c,d]			A	B_1	B_2	B_3	C	G3
C_8H_{18}	*n*-octane	−49.9	±0.3	b	−0.9	−1.7	−1.5	−1.7	−1.7	−0.8
$C_{10}H_8$	azulene	69.1	±0.8	b	2.5	2.0	3.0	1.7	−0.2	1.6
$C_{10}H_8$	naphthalene	35.9	±0.4	b	−0.9	−1.4	−0.2	−1.4	−3.0	−0.5
$C_{10}H_{16}$	adamantane	−31.8	±0.3	a	−0.9	−2.1	−1.5	−2.2	−3.0	−2.3
$C_{10}H_{18}$	*trans*-decalin	−43.5	±0.5	b		−1.5	−1.0	−1.6	−2.2	−1.0
$C_{12}H_{10}$	biphenyl	43.4	±0.5	b		−1.6	−0.1	−1.5	−3.3	−0.8
$C_{14}H_{10}$	anthracene	55.2	±0.5	b		−1.6	0.2	−1.7	−5.0	−1.0
FH	hydrogen fluoride	−65.1	±0.2	a	−0.2	−0.2	−0.2	−0.2	−0.2	−0.3
FHO	hypofluorous acid, HOF	−20.7	±0.3	l	−0.1	−0.1	−0.1	−0.1	−0.1	0.3
FNO	nitrosyl fluoride	−15.7	±0.4	a*	−4.4	−4.2	−4.1	−4.1	−3.0	−5.3
F_2	fluorine	0.0	±0.0	D	0.3	0.3	0.3	0.3	0.3	0.7
F_2N_2	difluorodiazene (cis)	16.4	±1.2	a	0.8	0.8	1.1	1.3	2.4	1.6
F_2N_2	difluorodiazene (trans)	19.4	±1.2	a	−0.9	−0.9	−0.6	−0.6	0.5	0.0
F_2O	difluorine monoxide, FOF	5.9	±0.4	a	0.4	0.8	0.8	0.9	2.1	0.6
F_2O_2	perfluoroperoxide, FOOF	4.6	±0.2	a*	4.7	5.3	5.3	5.3	7.8	4.7
F_3N	trifluoroamine	−31.6	±0.3	a	−0.2	0.1	0.2	0.6	2.0	0.0
F_4N_2	tetrafluorohydrazine	−2.0	±2.5	a	−3.4	−3.0	−2.8	−2.5	−0.6	−2.9
HNO	nitroxyl, HNO	25.6	±0.6	m	0.2	0.2	0.2	0.2	0.2	0.8
HNO_2	nitrous acid (cis)	−18.3	±0.3	n	0.4	0.2	0.3	0.5	1.4	1.0
HNO_2	nitrous acid (trans)	−18.8	±0.3	n	0.5	0.5	0.6	0.7	1.4	0.9
HNO_3	nitric acid	−32.1	±0.1	a	0.9	0.1	0.5	1.1	1.8	0.4
H_2	hydrogen	0.0	±0.0	D	0.1	0.1	0.1	0.1	0.1	−0.5
H_2N_2	diazene (trans)	47.1	±0.5	o	1.0	1.0	1.0	1.0	1.0	2.2
H_2O	water	−57.8	±0.0	b	−0.1	−0.1	−0.1	−0.1	−0.1	0.3
H_2O_2	hydrogen peroxide	−32.5	±0.0	a	0.2	0.2	0.2	0.2	0.2	1.2
H_3N	ammonia	−11.0	±0.1	b	0.1	0.1	0.1	0.1	0.1	0.9
H_4N_2	hydrazine	22.8	±0.2	a	0.5	0.5	0.5	0.5	0.5	2.1
N_2	nitrogen	0.0	±0.0	D	0.4	0.4	0.4	0.4	0.4	2.1
N_2O	nitrous oxide	19.6	±0.1	a	1.7	0.6	0.9	1.3	1.3	1.8
N_2O_3	dinitrogen trioxide	19.8	±0.2	a	2.6	2.4	2.6	3.2	3.7	2.7
O_3	ozone	34.1	±0.4	a	4.2	4.1	4.3	3.9	3.0	0.8

[a] All values in kcal/mol. Uncertain experimental data are indicated by asterisks. See text. [b] Results for BSR corrected methods and for G3 are given as difference with respect to experiment. [c] Experimental references are: (a) ref 99, (b) ref 97, (c) ref 99 for HCN and ref 104 for HCN to HNC isomerization, (d) ref 98, (e) ref 106, (f) refs 102 and 105, (g) ref 101, (h) ref 112, (i) ref 103, (j) ref 110, (k) ref 100 (combustion of solid) and ref 111 (sublimation), (l) ref 109 (reevaluation of $\Delta H_f^0(0)$) and ref 114 (conversion to 298 K), (m) ref 108, (n) ref 114, (o) ref 107 ($\Delta H_f^0(0)$) and ref 113 (conversion to 298 K). [d] The entry "D" refers to $\Delta H_f^0(298.15)$ values equal to 0 kcal/mol by thermochemical definition.

it excludes molecules for which models A−C are equivalent to the reference level of theory, $[\bar{6}\bar{6}\bar{6}\bar{5}|\bar{5}\bar{5}\bar{5}\bar{5}]$. Reported statistical measures are mean unsigned (MUE), mean signed (MSE), root-mean-square (rms), and maximum (Max) errors.

BSR corrections reduce both systematic (MSE) and overall (MUE, rms, Max) deviations from experiment and compensate efficiently for weaknesses of individual theoretical approaches (A vs B and C). In this respect, the comparison to experimental heats of formation confirms our earlier conclusions drawn from statistical analyses of CCSD(T) atomization energies,[51] showing that (a) these conclusions are valid also for larger molecules not studied before and that (b) the remaining contributions to heats of formation (ZPE, thermal enthalpies, etc.) have been accounted for quite reasonably.

More detailed inspection reveals, however, that differences in error statistics between the fully corrected models A−C are much smaller than was observed for CCSD(T) atomization energies. rms errors, e.g., vary only from 1.05 to 1.29 kcal/mol, depending on the model and the precise number of reference data considered, while models A and C reproduced reference level CCSD(T) atomization energies to within 0.13 and 0.99 kcal/mol, respectively.[51] The last block shown in Table 4 reports deviations with respect to model A, revealing that deviations between more accurate (A) and less accurate treatments (C) are still quite substantial (0.98 kcal/mol). Uncritical statistical evaluation with experimental reference data obviously obscures real differences between methods, as it conveys information about average errors but no details about error distributions. These considerations only permit the conclusion that contributions other than the atomization energy at the CCSD(T) level, are responsible for most of the discrepancy between high-level theoretical treatments (model A) and experimental data. The major sources of uncertainty appear to be (a) occasional inaccuracies in geometry and zero-point energies and (b) experimental error (for more detailed discussions see sections 6.2 and 6.4.3). Other effects are likely to play a minor role, including residual inaccuracies in relativistic, diagonal Born-Oppenheimer, and higher-order excitation corrections.

Finally, we compare our results to the established and popular G3 protocol and conclude that model C achieves about the same accuracy, while models A and B show small, but consistent improvements. It is quite likely, though, that the statistical evaluation of the more accurate models A (and B) reflects some degree of experimental error, calling for some caution in the interpretation of precise numbers. On the whole, results are very encouraging, however, as they show that a fully ab initio approach is able to achieve the same level of accuracy as the empirically calibrated G3 protocol, at a very reasonable computational cost.

We note at this point, that the use of bond separation reactions with *experimental heats of formation* for the parent molecules (the "BSR prototypes") has been shown to improve the accuracy of G2,[123] CBS-Q, CBS-q, CBS-4,[124] and density functional calculations[125] markedly, and similar improvements should be expected as well for G3. The use of experimental heats of formation infers that not only residual errors in electronic

TABLE 4: Statistical Analysis of Compound Methods A−C for Heats of Formation (ΔH_f^0(298.15))[a]

	fully corrected (ATOMIC)					partially corrected[b]					
	A	B_1	B_2	B_3	C	A	B_1	B_2	B_3	C	G3
165 molecules[c]											
MUE		0.73	0.75	0.75	0.93		1.95	4.06	3.14	3.68	0.90
MSE		−0.11	0.23	0.04	−0.16		1.81	4.05	2.79	−2.11	−0.05
rms		1.05	1.08	1.06	1.28		2.56	4.84	4.18	4.59	1.25
Max (+)		4.1	4.3	3.9	3.7		9.0	13.9	14.7	8.6	4.7
Max (−)		−3.9	−3.6	−3.5	−5.0		−1.3	−0.3	−2.2	−12.7	−4.4
154 molecules[d]											
MUE	0.70	0.73	0.77	0.76	0.89	2.49	1.90	3.79	2.98	3.65	0.90
MSE	0.18	−0.07	0.23	0.08	−0.06	2.41	1.75	3.78	2.61	−2.17	−0.03
rms	1.06	1.06	1.11	1.07	1.21	2.93	2.53	4.50	3.99	4.58	1.27
Max (+)	4.2	4.1	4.3	3.9	3.7	8.5	9.0	12.9	14.7	8.6	4.7
Max (−)	−3.7	−3.9	−3.6	−3.5	−3.3	−2.1	−1.3	−0.3	−2.2	−12.7	−4.4
131 molecules[e]											
MUE	0.77	0.80	0.85	0.84	0.98	2.80	2.08	4.25	3.30	3.88	0.94
MSE	0.18	−0.12	0.24	0.06	−0.11	2.71	1.94	4.25	2.96	−2.21	−0.11
rms	1.13	1.14	1.19	1.14	1.29	3.16	2.71	4.85	4.29	4.82	1.32
Max (+)	4.2	4.1	4.3	3.9	3.7	8.5	9.0	12.9	14.7	8.6	4.7
Max (−)	−3.7	−3.9	−3.6	−3.5	−3.3	−2.1	−1.3	−0.2	−2.2	−12.7	−4.4
131 molecules[e] (deviations with respect to fully corrected model A)											
MUE	0	0.35	0.26	0.31	0.74	2.54	1.79	4.07	3.03	3.50	0.68
MSE	0	−0.30	0.06	−0.12	−0.28	2.54	1.76	4.07	2.79	−2.39	−0.28
rms	0	0.44	0.34	0.38	0.98	2.82	2.24	4.51	3.85	4.46	1.02
Max (+)	0	0.4	1.0	0.9	2.8	6.1	7.9	11.9	13.7	5.7	2.6
Max (−)	0	−1.4	−0.8	−1.4	−3.0		−0.4		−1.4	−11.7	−5.3

[a] Errors are given with respect to experiment (in kcal/mol). [b] Includes BSR terms only for MVD, DBOC, CCSDT-CCSD(T), and CCSDTQ-CCSDT corrections (Table III of ref 51). BSR corrections toward the complete basis set limit CCSD(T) energy (Table II of ref 51) are omitted. [c] Includes all molecules of Table 3 except those for which the experimental value is disputed (marked with an asterisk in Table 3). [d] Excludes molecules for which no "level A" data are available. [e] Excludes additionally BSR prototype molecules (two first-row atoms or less).

energies cancel out but also other imperfections in the theoretical model (ZPEs, thermal enthalpies, geometrical errors), which are expected to affect the molecule under study and its parent compounds in very much the same way. Such correction is not considered in our theoretical model which is truly ab initio, not depending on experimental data for BSR prototypes, and thus much more widely applicable. Note that the investigations of Raghavachari et al.[123,125] and of Curtiss et al.[124] were necessarily limited to molecules with accurate experimental reference data for BSR prototypes and only included six types of non-hydrogen bonds (^{3}C−C^3, ^{2}C=C^2, ^{1}C≡C^1, ^{3}C−N^2, ^{3}C−O^1, ^{2}C=O).

Eight out of 173 molecules were not included in the analysis, as experimental heats of formation appear suspect. Among them are four fluorine compounds for which independent high-level ab initio studies yielded heats of formation close to our calculated values but at variance with experiment and suggested experimental revision: nitrosyl fluoride (−19.1 ± 1.0 kcal/mol, obtained from ab initio data reported by Martin et al.[126] vs model A, −20.1; experiment, −15.7), perfluoroperoxide (9.6 ± 0.9 kcal/mol[23] vs A, 9.3; exp, 4.6), carbonyl fluoride (−145.2 ± 0.8 kcal/mol[127] vs A, −145.0; exp, −149.1), and tetrafluoroethylene (−160.5 ± 1.5 kcal/mol[128,129] vs A, −160.8; exp, −157.5). Likewise, Schuurman et al. have recently reported[130] an ab initio benchmark value, $\Delta H_f^0(0) = -27.6 \pm 0.2$ kcal/mol, for isocyanic acid, HNCO, which corresponds to[114] $\Delta H_f^0(298.15) = -28.2 \pm 0.2$ kcal/mol, fully supporting our calculation (model A, −28.1; exp, −24.3). A revised experimental value,[131] $\Delta H_f^0(0) = -27.8 \pm 0.4$ kcal/mol, was given by Zyrianov et al., which would be in excellent accord with the theoretical data, but Schuurman et al. note that it might require further adjustment.[130]

Three further molecules are not included in the statistical analysis, dimethylnitrosamine, 3*H*-diazirine, and tetracyanomethane. There are no independent high level ab initio calculations available that could support our calculated heats of formation, but discrepancies with respect to experiment are so large that one must assume experimental error. In the case of 3*H*-diazirine (model A, 75.2 kcal/mol), an older experimental estimate is available (79.3 kcal/mol),[132] which, although criticized by Laufer and Okabe,[133] appears to be much more accurate than the revised value (63.3 ± 2.7 kcal/mol).[133]

6.2. Cases of Larger Error.

6.2.1. General Remarks. Inspection of Table 3 reveals several cases where our theoretical results differ from experiment by substantially more than 1 kcal/mol. Among them are molecules for which quoted experimental error bars are large (fluoromethylene, trifluoroethylene, tetrafluorohydrazine) or questionable (acetyl acetylene[49]), and where the theoretical result is likely to be more accurate than experiment. For some larger molecules (benzoquinone, 2,5-norbornadiene, norbornan-7-one, cubane, azulene) experimental values carry error bars of 0.7 kcal/mol or more, and accumulation of errors in the various components of the calculation can easily cause uncertainties of around 1 kcal/mol in the theoretical results as well. Agreement better than about ±2 kcal/mol is thus not expected. There are, however, several smaller molecules, for which disagreement between theory and experiment is larger than expected and which merit some further comment.

The uncertainty quoted by Cioslowski et al.[99] and the JANAF tables[114] for dinitrogen trioxide (19.8 ± 0.2 kcal/mol) seems somewhat optimistic, as a different compilation[113] recommends a value of 20.7 ± 0.2 kcal/mol, which is in substantially better agreement with the calculated heats of formation. The recommended heat of formation for 1,1-difluoroethylene (−80.1 ± 1.1 kcal/mol)[97] results from a weighted average of two independent measurements, reevaluated by Pedley et al.,[97] (−79.8 ± 0.8 kcal/mol[134] and −82.2 ± 2.4 kcal/mol[135]), and we note that the calculations are in reasonable agreement with the later experiments of Kolesov et al.[135] but not with the value recommended by Pedley et al. and listed in Table 3.

Ozone is a special case, where the large discrepancy with respect to experiment is entirely due to a failure of the theoretical method: there are two dominant electronic configurations contributing to the ground state energy,[22] and single-reference coupled cluster methods need to consider at least connected triple and quadruple excitations for semiquantitative agreement with experiment. The CCSDTQ-CCSD(T) corrections to the heat of formation (−2.02 kcal/mol) are not well represented by isodesmic bond contributions (−0.77 kcal/mol);[51] basis-set incompleteness effects for this correction and connected quintuple excitations add another −0.4 kcal/mol each,[14] and the zero-point energy is overestimated by 1.66 kcal/mol (Table 2). The sum of these effects (−3.7 kcal/mol) accounts almost completely for the observed discrepancy of ≈4 kcal/mol between calculated and experimental heats of formation.

While it is beyond the scope of this work to analyze in detail all remaining cases for which disagreement with experiment is non-negligible, it is perhaps worthwhile to study some of the cyano compounds, for which the proposed theoretical approach often yields larger heats of formation than experiment. Among these molecules are cyanogen (method A, +1.6 kcal/mol), acrylonitrile (+2.0), isobutylnitrile (+1.5), and dicyanoacetylene (+2.8). In all these cases, the experimental reference[97] quotes very small error bars. Two of them (cyanogen and dicyanoacetylene), both non-hydrogen compounds with several triple bonds, are analyzed in some detail.

6.2.2. Cyanogen. The most detailed calculations have been performed for cyanogen, although not all data are shown in detail. The CCSD(T) contribution to the total atomization energy $E_{A,e}$ at the reference level, ($[\bar{6}\bar{6}\bar{6}\bar{5}|\bar{5}\bar{5}\bar{5}\bar{5}]$, 500.81 kcal/mol), appears to be well converged and accurate to about 0.1 kcal/mol, as judged from extrapolated CCSD(T)(full)/cc-pCV(TQ)Z (500.57), CCSD(T)(full)/aug-cc-pCV(TQ)Z (500.76), and CCSD(T)(full)/cc-pCV(Q5)Z (500.77) calculations. Using the reference level of theory, one obtains $\Delta H_f^0(298.15) = 74.74$ kcal/mol, which is higher by 1.4 kcal/mol than the experimental value of 73.3 ± 0.2 kcal/mol,[97] but well reproduced by the BSR corrected schemes A and B_n (Table 3). Inaccuracies in the relativistic, diagonal Born-Oppenheimer, and higher order (CCSDTQ-CCSD(T)) corrections are likely to play a minor role. From comparison with explicit calculations for cyanogen,[51] we determined the error inherent in the assumption of thermoneutral isodesmic reactions to be −0.02 kcal/mol for relativistic corrections (at the B3LYP level), −0.01 kcal/mol for diagonal Born-Oppenheimer corrections, and +0.12 kcal/mol for CCSDTQ-CCSD(T) corrections (−0.21 for CCSDT-CCSD(T) and +0.33 for CCSDTQ-CCSDT). Effects of even higher excitations are expected to be in the −0.1 to −0.15 kcal/mol range, based on results of Tajti et al. for hydrogen cyanide.[15] Errors in the evaluated thermal enthalpy term, $\Delta H^0(298.15)$ can be neglected as well (computed, 3.04 kcal/mol; experimental,[136] 3.03 kcal/mol). The only remaining sources of error that can be expected to affect the heat of formation by several tenths of a kcal/mol are inaccuracies in the geometry and in the calculated zero-point energy.

Comparison of experimental (r_0)[137] and RI-MP2/cc-pVTZ (r_e) geometries already indicates that the theoretical level employed may not be accurate enough for CN triple bonds (Table 5). Geometry optimizations at the CCSD(T) level indeed confirm that MP2 overestimates the CN bond length by about 0.02 Å, leading to an error of ≈+0.8 kcal/mol in the heat of formation (Table 5). On the other hand, scaled MP2 zero-point energies are known to be less accurate for molecules with cyano groups (Table 2), and corrections should increase the calculated heat of formation and thus at least partially cancel the overestimation due to inaccuracies in the geometry. Using experimental fundamental frequencies,[136] one obtains a zero-point energy of 9.75 kcal/mol which is expected to be a *lower* bound to the true zero-point energy.[86] The scaled zero-point energy evaluated at the RI-MP2/cc-pVTZ level equals 9.33 kcal/mol, and one may expect a heat of formation higher by at least 0.4 kcal/mol if the (unknown) true zero-point energy was used. In summary we note that the theoretical model proposed here is less accurate for molecules with several cyano groups (probably also for those with other types of triple bonds, see below), but that the most important errors, inaccuracies in geometries and zero-point energies, are likely to cancel each other at least partially. Hence it appears reasonable to assume that the most elaborate calculation ($[\bar{6}\bar{6}\bar{6}\bar{5}|\bar{5}\bar{5}\bar{5}\bar{5}]$, $\Delta H_f^0(298.15) = 74.7$ kcal/mol, see above) carries an error bar of *less* than 1 kcal/mol, although experiment (73.3 ± 0.2 kcal/mol) seems to suggest otherwise.

TABLE 5: Geometries of Cyanogen and Dicyanoacetylene[a]

	C≡N	C−C	C≡C	E_{rel}[b]
cyanogen				
(RI-)MP2/cc-pVTZ	1.176	1.378		0.00
CCSD(T)/cc-pVTZ	1.165	1.389		−0.68
CCSD(T)/cc-pVQZ	1.161	1.388		−0.80
CCSD(T)(full)/cc-pCVTZ	1.162	1.388		−0.78
CCSD(T)(full)/cc-pCVQZ	1.159	1.385		−0.84
exp. (IR, r_0)[c]	1.154	1.389		−0.82
dicyanoacetylene				
(RI-)MP2cc-pVTZ	1.179	1.366	1.225	0.00
CCSD(T)cc-pVTZ	1.168	1.379	1.216	−0.96
CCSD(T)cc-pVQZ	1.164	1.377	1.212	−1.13
CCSD(T)(full)cc-pCVTZ	1.164	1.377	1.213	−1.11
CCSD(T)(full)cc-pCVQZ	1.161	1.375	1.210	−1.20
exp. (electron diffraction, r_α)[d]	1.161(5)	1.367(3)	1.198(11)	−1.03

[a] Both molecules are linear. Bond lengths are reported in Å. [b] Relative energy (kcal/mol) evaluated at the CCSD(T)(full)/cc-pCV(Q5)Z level of theory ($[\bar{5}\bar{5}\bar{5}\bar{5}|\bar{5}\bar{5}\bar{5}\bar{5}]$) for that particular geometry. [c] Reference 137. [d] Reference 140.

The experimental heat of formation for cyanogen (73.3 ± 0.2 kcal/mol) was taken from the book of Pedley et al.[97] and cross-checked with the extensive evaluation of Cioslowski et al.,[99] the monograph of Lias et al.,[98] and the compilation of the G2/97 data set,[76] all of which cite the Pedley value. Closer inspection reveals, however, that Pedley et al. have preferred an older value[138] based on a smaller quoted error bar, discarding the redetermination (73.8 ± 0.4 kcal/mol) reported by Knowlton and Prosen.[139] The more recent value and its associated uncertainty provide better support of the high-level theoretical data discussed above.

6.2.3. Dicyanoacetylene. Discrepancies between calculated and experimental heats of formation are even larger for dicyanoacetylene (Table 3). As for cyanogen, the CCSD(T) atomization energy $E_{A,e}$ appears to be well converged at the reference level of theory ($[\bar{6}\bar{6}\bar{6}\bar{5}|\bar{5}\bar{5}\bar{5}\bar{5}]$, 794.11 kcal/mol), as lower-level results, CCSD(T)(full)/cc-pCV(TQ)Z (794.01 kcal/mol) and CCSD(T)(full)/cc-pCV(Q5)Z (793.98 kcal/mol), agree to within 0.1 kcal/mol. Models A and B_n reproduce the reference value to within 0.5 kcal/mol if isodesmic corrections are taken into account. Disagreement with experiment[97] (126.5 ± 0.3 kcal/mol) amounts to 2.4 kcal/mol, however, even if the theoretical result (128.9 kcal/mol) is based on the reference level of theory (Table 3).

Again, inaccuracies in the relativistic, diagonal Born-Oppenheimer, and higher order (CCSDTQ-CCSD(T)) corrections are expected to play a minor role. From comparison with explicit calculations for dicyanoacetylene,[51] we determined the error inherent in the assumption of thermoneutral isodesmic reactions to be −0.05 kcal/mol for relativistic corrections (at the B3LYP level) and −0.02 kcal/mol for diagonal Born-Oppenheimer corrections. CCSDTQ calculations are not feasible for the size of this system, but the total CCSDTQ-CCSD(T) correction evaluated from the isodesmic scheme (0.49 kcal/mol) should at least be qualitatively reasonable. The calculated thermal enthalpy term $\Delta H^0(298.15) = 4.26$ kcal/mol agrees with the experimental value of 4.24 kcal/mol.[136]

Comparison of MP2, CCSD(T), and experimental[140] geometries shows that MP2 underestimates the level of bond length alternation between single and triple bonds, resulting in C≡C and C≡N bonds too long by almost 0.02 Å (Table 5). From the data reported in Table 5, one may estimate an associated error of +1.2 kcal/mol in the calculated heat of formation.

Like for cyanogen, the highest-frequency modes are triple-bond stretch vibrations, whose experimental frequencies[141] (fundamental, Σ_g^+ 2119, ≈2290 cm^{-1}, Σ_u^+ 2241 cm^{-1}) are underestimated at the RI-MP2/cc-pVTZ level (harmonic, Σ_g^+ 2038, 2246 cm^{-1}, Σ_u^+ 2161 cm^{-1}). The zero-point energy evaluated from experimental fundamental frequencies equals 16.00 kcal/mol and is expected to be a lower bound to the true zero-point energy,[86] while scaled MP2 frequencies lead to a substantially smaller value (15.36 kcal/mol). In summary, one should expect two fairly sizable errors from inaccuracies in the geometry (ca. +1.2 kcal/mol) and inaccuracies in the zero-point energy (at least −0.6 kcal/mol) that cancel each other to a significant extent. While this result is not very satisfying from a theoretical point of view, it also indicates that the experimental heat of formation may be too low.

The compilations of Pedley et al.,[97] Cioslowski et al.,[99] and Lias et al.[98] refer either directly or via other compilations to the heat of combustion (liquid) reported by Armstrong and Marantz[142,143] and the heat of vaporization reported by Saggiomo.[144] Initially, the heat of formation was determined as 127.5 kcal/mol[142] (in closer agreement with our calculations, and cited by Lias et al.), but it was later corrected to 126.5 kcal/mol[143] (cited by Pedley et al. and Cioslowski et al.).

6.3. Predictions and Comparisons to ab Initio Benchmark Studies. For a number of molecules considered earlier,[51] experimental heats of formation are unavailable, known to be inaccurate, or reported as estimates only. Table 6 collects the data found in compilations[98,99,113,114] and original work,[108,145−149] but no attempt has been made to evaluate numbers from different sources. The calculated heats of formation (models A, B_3, and C) are assumed to be much more reliable in these cases and may serve as accurate predictions. We note that some of the molecules are BSR prototypes (two non-hydrogen atoms or less) for which models A, B_3, and C are identical by construction. Table 6 also lists results of previously reported benchmark calculations,[129,130,150−158] with a preference for those with carefully estimated error bars. Results from calibrated models such as G2 and G3 are not considered for this comparison.

In almost all cases, excellent agreement is found between our results—particularly those for models A and B_3—and previously reported calculations. Deviations are generally within or just outside reported error bars and amount to only a few tenths of a kcal/mol. The only exception is formonitrile oxide whose heat of formation is overestimated by at least 1 kcal/mol[130] if the unusually large anharmonic ZPE contribution due to quasi-linearity of the molecule[159,160] is not properly taken into account. Finally, it is noteworthy that the gas phase heat of formation of formamide, the essential building block of protein structure, has never been determined precisely by experimental techniques. The only published study appears to be a calculation of Bauder and Günthard,[145] 50 years back, which was based on thermodynamical data of unknown accuracy. Our calculations show, however, that the estimated value was very accurate indeed.

6.4. Case Study: Linear Alkanes. The thermochemistry of alkanes plays a dominant role in fuel combustion processes and gains further relevance for computational approaches as alkane chains form important scaffolds in organic chemistry. Experimental heats of formation are known to high accuracy, making these molecules an ideal and important test case for ab initio thermochemical procedures. The accuracy of G3 theory (and derivatives thereof) has been tested extensively for heats of formation of linear alkane chains,[161] and more recent work by Pollack et al.[162] reports on large-scale ab initio calculations for alkanes up to *n*-octane, using CCSD(T) calculations with basis sets up to aug-cc-pVQZ. Effects of conformational averaging are discussed in both papers but have been fully taken into account only by Tasi et al.[163] who applied an empirically calibrated effective one-electron Hamiltonian.[164] Here we discuss the importance of various contributions to calculated heats of formation, including effects of conformational averaging and corrections for relativistic, non-Born-Oppenheimer, and higher order electron correlation effects.

6.4.1. Effects of Conformational Averaging. Standard ab initio thermochemistry refers to one particular conformation of a molecule, which is chosen to be the absolute minimum on the potential energy surface. Inclusion of thermal enthalpies in the rigid-rotor and harmonic-oscillator (RRHO) approximation allows for some limited "sampling" of space in the vicinity of the minimum but does not consider any conformational isomers. Experimentally determined heats of formation, on the other hand, reflect a Boltzmann-weighted average over the entire accessible conformational space, and may thus deviate from ab initio estimates for nonrigid molecules such as linear alkane chains. An estimate of the increase in thermodynamical energy $\Delta\langle E\rangle_T$ due to conformational averaging is given by

$$\Delta\langle E\rangle_T = \langle E\rangle_T - E_{\text{all-trans}} = \frac{\sum_i \Delta E_i \cdot \exp(-\Delta E_i/k_B T)}{\sum_i \exp(-\Delta E_i/k_B T)} \tag{6}$$

where $E_{\text{all-trans}}$ refers to the energy of the lowest-energy conformer (the all-trans conformation of alkanes), ΔE_i to the relative energy of conformation *i*, and the sum runs over all conformers *i*, i.e., over all minima of the potential energy surface. This simple treatment rests on the assumption that torsional contributions to thermal enthalpies can be separated into those arising from global motion (Boltzmann-weighted conformational averaging, eq 6) and those arising from local motion (harmonic oscillator treatment based on the set of scaled frequencies for the lowest-energy conformer). Some justification may be found in the fairly high energy barriers separating conformers of *n*-butane (>3 kcal/mol[165,166]), and of longer alkane chains.[166] G3(MP2) estimates of $\Delta\langle E\rangle_{298.15}$ were previously reported by Redfern et al.[161] for *n*-butane (0.26 kcal/mol) and *n*-pentane (0.46 kcal/mol). Following the conformer enumeration

TABLE 6: Heats of Formation $\Delta H_f^0(298.15)$: Comparison to Other Theoretical Benchmark Calculations and Experimental Estimates, and New Predictions[a,b]

		A	B_3	C	experimental estimate[c]		theoretical benchmark[c,d]	
CFN	cyanogen fluoride, FCN	2.5	2.4	2.7	8.6 ± 4.0	a	3.2 ± 1.5	b
CFN	isocyanogen fluoride, FNC	73.1	73.3	73.0			73.0 ± 1.5	c
CHFO	formyl fluoride	−91.4	−91.2	−91.0	−90.0	d		
CHNO	cyanic acid, HOCN	−3.3	−3.4	−3.2			−3.7 ± 0.2	e
CHNO	formonitrile oxide, HCNO	41.9	41.5	40.5			40.3	e
CHNO	isofulminic acid, HONC	56.1	56.3	56.0			55.9 ± 0.2	f
CH_2N_2	cyanamide	32.8	32.7	32.6	32.0	g		
CH_2N_2	diazomethane	65.3	65.4	64.3	55.8 ± 4.5	h	65.3	i
CH_2O	hydroxymethylene	26.2	26.2	26.2				
CH_2O_2	dioxirane	0.9	1.0	0.7				
CH_2O_3	performic acid	−68.7	−68.4	−68.0			−68.2 ± 0.6	j
CH_3F	fluoromethane	−56.4	−56.4	−56.4	−56.0 ± 6.9	k	−56.5	l
CH_3N	methanimine	21.2	21.2	21.2	26.4 ± 3.2	m	21.1 ± 0.5	n
CH_3NO	formamide	−45.1	−44.9	−44.4	−44.5	o		
C_2F_2	difluoroacetylene	1.5	1.3	2.3	5.0 ± 5.0	p	1.4	q
C_2HF	fluoroacetylene	25.3	25.2	25.5	30.0 ± 15.0	r	25.2	s
C_2H_2O	oxirene	64.8	65.5	64.3				
C_2H_5F	fluoroethane	−65.2	−65.3	−65.5	−62.9 ± 0.4	t		
C_4H_4	butatriene	77.4	76.9	76.3				
C_4H_4	cyclobutadiene	102.5	102.6	101.9	96.0 ± 5.0	u		
C_4H_4	tetrahedrane	128.7	129.0	126.1				
FHO_2	fluoroperoxide, HOOF	−10.7	−10.3	−9.1			−10.4 ± 1.0	v
FH_2N	monofluoroamine	−6.2	−6.2	−6.2				
FH_3N_2	fluorohydrazine	8.6	8.9	9.4				
HNO_2	nitrous acid (H-NO_2)	−9.3	−9.6	−9.3				
H_2N_2	diazene (cis)	53.1	53.0	52.8	59.3 ± 7.2	w	52.7 ± 0.4	x
H_2N_2	diazene (iso)	72.1	72.1	72.1			71.6 ± 0.4	y
H_2N_2O	nitrosamine	19.4	19.5	21.0				
H_2O_2	hydrogen peroxide (iso)	14.0	14.0	14.0				
H_3NO	ammonia oxide	14.6	14.6	14.6				
H_3NO	hydroxylamine	−10.4	−10.4	−10.4	−9.6 ± 2.2	z	−10.1 ± 0.3	aa

[a] All values in kcal/mol. [b] Results are given for BSR corrected methods A, B_3, and C. [c] References are as follows: (a) ref 99, (b) ref 150 (0 K) and ref 114 (conversion to 298 K), (c) ref 150 (0 K) and this work (conversion to 298 K +0.4 kcal/mol), (d) ref 114 (est.), (e) ref 130 (0 K) and this work (conversion to 298 K −0.6 kcal/mol), (f) ref 130 (0 K) and this work (conversion to 298 K −0.4 kcal/mol), (g) ref 98 (est.), (h) ref 146, (i) ref 158, (j) ref 156, (k) ref 114, a smaller error bar has apparently been reported by DeMore et al. in a Jet Propulsion Laboratory publication of Caltech, see Table V of ref 155, (l) ref 155, (m) ref 148, (n) ref 153, (o) ref 145 (based on thermochemical data whose accuracy was not firmly established no error bars given), (p) ref 114, (q) ref 129, (r) ref 114 (est.), (s) ref 154 (W2 theory), (t) ref 147 (est.), (u) ref 149 (est.), (v) ref 151, (w) ref 113, (x) ref 152 (0 K) and ref 113 (conversion to 298 K), (y) ref 152 (0 K) and this work (conversion to 298 K, −1.7 kcal/mol), (z) ref 108 (weighted average of two fairly different experimental results, −7.9 ± 1.5 kcal/mol and −12.0 ± 2.4 kcal/mol, considers also theoretical data), (aa) ref 157. [d] In certain cases, authors only reported $\Delta H_f^0(0)$. Conversion to $\Delta H_f^0(298)$ was performed as specified in footnote *c*, but quoted error bars refer to $\Delta H_f^0(0)$.

scheme of Tasi et al.,[163,167] we have performed additional calculations for longer alkane chains (see Figure 1), both at the MP2 (RI-MP2/cc-pVTZ) and density functional (RI-BP86[168−172]/SV(P)[173]) levels of theory. A total of 95 geometry optimizations—most of them in low symmetry—are required for *n*-octane, which seemed prohibitively expensive in the case of MP2 but quite feasible for the simple density functional level of theory. Both levels of theory suggest a heat increment of roughly 0.2 kcal/mol per rotatable C−C single bond, at least for the first few members of the homologous series. Results are inconclusive for long alkane chains, as extrapolation appears to indicate that heat contributions level off slowly in the case of MP2 while an almost perfectly linear increase is observed for BP86. Differences between the two theoretical methods indicate that effects of dispersion on conformational stability may play a role, and more detailed analyses are certainly needed to reach general conclusions. This is less relevant in the current context, however, as the two methods still agree reasonably well for *n*-butane to *n*-heptane, suggesting that results are not overly method-dependent for shorter alkane chains.

6.4.2. Calculated Contributions to Heats of Formation and Comparison with Experiment. Figure 2 illustrates the relative importance of various corrections to computed heats of formation. Calculations at the CCSD(T) (full) level of theory with accurate extrapolation to the complete basis set limit and including atomic spin-orbit corrections as well as zero-point energies and thermal increments (RRHO approximation) yield heats of formation too low by 1.5−3.0 kcal/mol, which is certainly insufficient for quantitative predictions. Consideration of scalar relativistic corrections improves agreement with experiment, but calculated values are still too low by 0.6−1.5 kcal/mol. First-order corrections to the Born-Oppenheimer approximation (DBOC) increase deviations from experiment again, by 0.25−0.4 kcal/mol, but consideration of higher-order excitations (CCSDTQ-CCSD(T)) corrects by another 0.4−1.0 kcal/mol. Finally, corrections for thermal averaging lead to near-quantitative agreement with experiment, and residual errors are well within the tight experimental error bars (0.2−0.3 kcal/mol).

This analysis indicates that effects of conformational averaging cannot be neglected for accurate calculations of standard heats of formation, $\Delta H_f^0(298.15)$, to which they contribute about as much as higher-order excitations. Even relativistic corrections, generally recognized to be essential for accurate thermochemical work, are only about twice as large. All corrections except DBOC increase the computed heats of formation, such that our standard thermochemical protocol—which neglects effects of

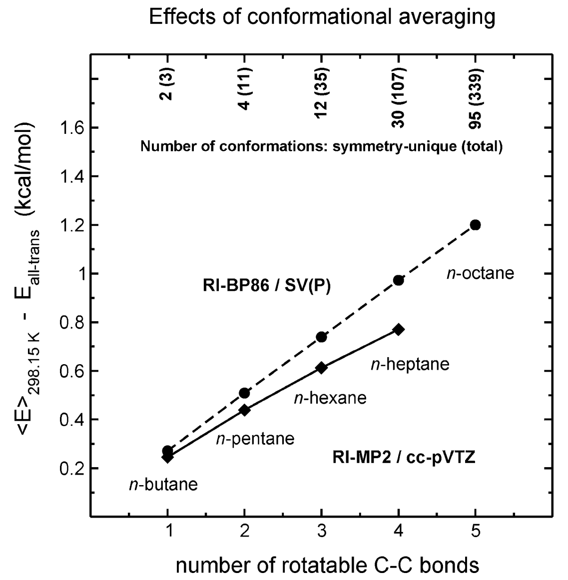

Figure 1. Heats of formation $\Delta H_f^0(298.15)$ of linear *n*-alkanes: contributions from thermal averaging based on RI-BP86/SV(P) (dashed line) and RI-MP2/cc-pVTZ (solid line) calculations. See text for details.

conformational averaging—consistently underestimates heats of formation for *n*-alkanes.

6.4.3. Assessment of Possible Sources of Error. The level of accuracy achieved is quite remarkable in view of the many approximations made, and a more detailed analysis appears mandatory to see if this is due to (partial) error cancellation.

First, the complete basis set limit of the CCSD(T) energy has only been estimated through the use of isodesmic reactions evaluated with model A. Inspection of Table S2 of ref 51 indicates, however, that application of BSR corrections leads to very accurate estimates for propane (error, 0.02 kcal/mol), suggesting that errors will not exceed a few tenths of a kcal/mol even for longer alkane chains. Residual inaccuracies in RI-MP2/cc-pVTZ equilibrium geometries are unlikely to add to this small error, as experimental bond lengths and bond angles are reproduced to within 0.005 Å and 0.5°, respectively: The calculated C−H bond length of methane (1.085 Å) compares well with experimental (1.0870(7) Å,[174] 1.0858(10)[175]) and benchmark-level ab initio r_e reference data (1.08595(30) Å[176]), and known C−C equilibrium bond lengths of ethane (1.528(3) Å,[177] 1.522(2) Å[178]) and propane (1.522(2) Å[179]) are just as well reproduced by the calculations (1.523 Å each). The MP2 equilibrium geometry of propane shows good agreement with experiment (spectroscopic values r_m, r_0, r_s)[179] also for C−C−C (112.4(1)°, 112.1(1)°, 112.2°; calcd 111.9°) and C−C−H bond angles (111.6(3)°/110.6(1)°, 111.3(4)°/110.5(1)°, 111.9°/110.6°; calcd 111.8°/110.7°). And finally, excellent agreement is also observed between RI-MP2/cc-pVTZ and high-level "best estimate" CCSD(T)(full) equilibrium geometries[180] for ethane (MP2/CCSD(T): C−C, 1.523 Å /1.524 Å; C−H, 1.088 Å /1.090 Å; C−C−H, 111.2°/111.2°).

Second, simple BSR estimates are used for relativistic, diagonal Born-Oppenheimer, and CCSDTQ-CCSD(T) corrections to heats of formation. Among these, diagonal Born-Oppenheimer corrections have been found to be the most problematic,[51] and this may be of some concern for molecules with many light (hydrogen) atoms, for which non-Born-Oppenheimer effects may contribute noticeably. Explicit calculations of these terms show, however, that BSR estimates are quite accurate in the case of *n*-alkanes, with errors ranging from 0.01 kcal/mol (propane) to only 0.07 kcal/mol (*n*-octane). Explicit calculations of mass velocity and Darwin terms at the B3LYP/cc-pCVTZ level further indicate that the BSR scheme reproduces relativistic corrections to within 0.02 kcal/mol for all alkanes considered.[51] And finally, good agreement is observed for *n*-propane between the BSR estimate (+0.29 kcal/mol) of CCSDTQ-CCSD(T) corrections and the explicit value derived earlier[51] from CCSD(T), CCSDT, CCSDT(Q), and CCSDT[Q] energies (+0.35 ± 0.10 kcal/mol).

The accuracy of calculated zero-point energies (ZPEs) is less well established. We only note that the scaling procedure affords ZPEs that are slightly too large for methane (0.2 kcal/mol) and ethane (0.15 kcal/mol; see Table 2). It appears reasonable to assume that ZPEs of longer alkane chains are overestimated as well, probably by somewhat larger amounts. Precise estimates would require the accurate knowledge of anharmonic force fields for at least the next two homologues of the series, which are, however, difficult to obtain and have not been reported so far.

The experimental heat of formation of atomic carbon, reported as $\Delta H_f^0(0) = 169.98 \pm 0.1$ kcal/mol in the JANAF tables,[114] is another source of uncertainty. The experimental value carries a fairly large error bar which is of some concern for ab initio thermochemical procedures based on atomization energies. More recent studies using extensive thermochemical networks seem to indicate that the established value should be revised upward by about 0.1 kcal/mol, but several slightly different values (170.12 ± 0.05 kcal/mol,[15] 170.05,[16] empirical and ab initio estimates[181] 170.11 ± 0.11, 170.09 ± 0.08) have been offered, while experimental details or definitively revised values seem to be unavailable so far.[181] We have thus *not* revised the atomic heat of formation of carbon for any of the results reported in this work, but note that revision would increase calculated molecular heats of formation by about 0.1 kcal/mol per carbon atom.

Most of the contributions to heats of formation appear to be estimated quite accurately, lending credence to the decomposition analysis reported in Figure 2. It is quite likely, however, that the excellent agreement with experiment is in part due to some error cancellation between overestimated ZPEs, underestimated atomic heats of formation, and, possibly, some minor inaccuracies in other components. Hence it would seem overly optimistic to expect the same excellent agreement with experiment for other hydrocarbons as well. An error estimate of about ±0.1 kcal/mol per carbon atom (model A) appears more realistic, and inspection of Table 3 shows that this is the level of accuracy observed for most hydrocarbons, with some strained molecules such as bicyclo[1.1.0]butane and methylenecyclopropane being notable exceptions.

7. Discussion and Conclusions

ATOMIC is an ab initio thermochemistry protocol designed to provide accurate estimates of CCSD(T) atomization energies at the complete basis set limit, further corrected for relativistic, diagonal Born-Oppenheimer, and higher-order excitation contributions. The approach is efficient enough to treat molecules with 10−20 non-hydrogen atoms and it avoids dauntingly expensive large-basis set calculations through consistent use of bond separation reactions (BSRs). Implemented in an ab initio framework and mapped onto a system of corrective bond increments, bond separation reactions supply very high-level ab initio data for bond dissociation embedded in a lower-level treatment of the perturbation caused through the chemical environment. Here we have extended the approach to the calculation of standard heats of formation which requires the

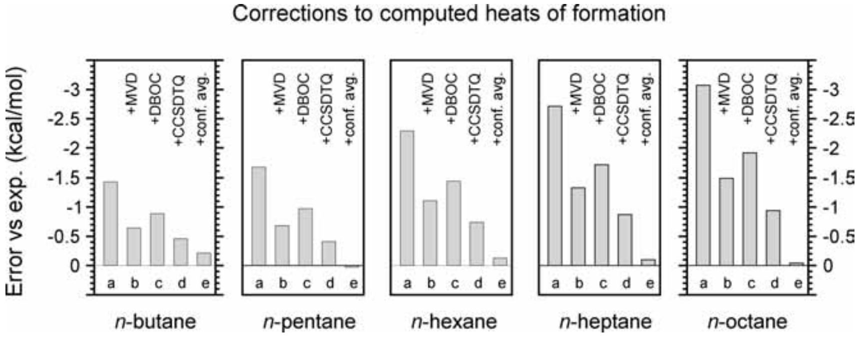

Figure 2. Heats of formation $\Delta H_f^0(298.15)$ of linear *n*-alkanes: Errors (theory − experiment) for BSR-corrected model A. From left to right, each chart refers to calculated values based on (a) CCSD(T)(full) atomization energy, spin-orbit corrections, zero-point energy, and thermal enthalpies (rigid-rotor harmonic-oscillator approximation), (b) plus scalar relativistic (mass-velocity and Darwin, MVD) terms, (c) plus diagonal Born-Oppenheimer (DBOC) terms, (d) plus CCSDTQ-CCSD(T) corrections, (e) plus corrections for thermal averaging. Thermal averaging corrections are taken from RI-MP2/cc-pVTZ data shown in Figure 1: *n*-butane 0.24 kcal/mol, *n*-pentane 0.44 kcal/mol, *n*-hexane 0.61 kcal/mol, *n*-heptane 0.77 kcal/mol, *n*-octane 0.9 kcal/mol (est.). Corrections (b−d) are evaluated from BSR increments, and results denoted as "(d)" including all corrections except thermal averaging are those reported in Table 3. Experimental data are taken from Table 3.

additional evaluation of zero-point energies and thermal enthalpies and the use of experimental thermochemical data for the atoms.

The primary objective of this work was to validate the ATOMIC approach with a large set of experimental heats of formation compiled for neutral molecules containing hydrogen and first row atoms (C, N, O, F). Through the course of this work we realized, however, that uncritical comparison of calculated and experimental data was inappropriate to properly judge the theoretical approach and to document its strengths and weaknesses. The generally high quality of established compendia notwithstanding, we detected several instances where experimental data were suspect. In some cases more recently reported studies established closer agreement with our calculations, in other cases disagreement with either our calculations or results from other theoretical benchmark studies was simply so large that one must assume experimental error. Quite unfortunately, small reported error bars do not always guarantee good accuracy. In this situation we have decided to cross-check various compilations of experimental data, to consult original publications where necessary, to search the more recent experimental literature for revisions, and to update existing compilations accordingly. Experimental data questioned on the grounds of our or other theoretical studies are still listed, but they are indicated as such and excluded from all statistical evaluations.

Overall, the ATOMIC approach performs very satisfactorily. It generally meets the goal of chemical accuracy (1−2 kcal/mol) for all types of bonding situations encountered in the chemistry of hydrogen and first-row elements (C, N, O, F). This is true even for larger systems with about 10 non-hydrogen atoms and for systems with charge-separated valence structure,[182] bearing testimony to the robustness of the BSR model. Comparison with high-level benchmark calculations reported in the literature shows excellent agreement as well. On the whole, the computationally most economical model C is comparable in accuracy to the popular, but empirically calibrated G3 approach, and the more refined models A and B perform better. The generally close agreement between results obtained using the latter models (see Tables 3 and 4) suggests that shortcomings in the BSR model to reproduce CCSD(T) atomization energies may account for only a modest fraction of the observed discrepancy with experiment and that other sources of error dominate. On the basis of these observations we recommend model B_3 for production calculations as it shows the best overall balance between computational cost and accuracy.

While statistical evaluation may not reveal the true accuracy of the thermochemical approach, as it reflects both theoretical and experimental error (see, in particular, sections 6.2.2 and 6.2.3), it has helped us to identify a number of potential problems on the theoretical side: (a) The use of bond increments for higher-order electron correlation contributions is an appropriate means to include effects typically observed for "simple" molecules,[51] but it cannot replace explicit calculation for difficult multireference cases like ozone and will fail in general if excitations are expected to be nonlocal in character. Other nonlocal effects have less bearing on the accuracy of the approach as they are either too small to be significant (diagonal Born-Oppenheimer corrections) or still represented by smaller-basis set calculations (e.g., connected triples corrections). (b) The level of theory employed for geometry optimizations ((RI-)-MP2/cc-pVTZ) is normally sufficient to provide for accurate thermochemistry based on high-level single-point calculations but occasionally falls short of expectations mainly for systems containing several triple bonds. While bond lengths are typically reproduced to within 0.01 Å or better,[183−185] larger errors (0.015−0.02 Å) are observed for C≡C and C≡N triple bonds which thanks to the large associated harmonic force constants (MM3 estimates[186] of 15.25 and 17.33 mdyn/Å) affect the energy by roughly 0.25−0.5 kcal/mol. (c) Zero-point energies are estimated from scaled (RI-)MP2/cc-pVTZ frequencies, using a scaling factor geared toward the standard case that MP2 overestimates harmonic frequencies slightly. Such a procedure seems reasonable in general, but it may produce less accurate zero-point energies for molecules with triple bond stretch modes whose frequencies tend to be underestimated.

On the other hand, results are particularly encouraging for the series of *n*-alkanes. The most refined model A achieves quantitative agreement with experiment, but it does so only *after* proper corrections are applied for conformational averaging. Hence discrepancy between theory and experiment establishes a real physical effect which is never captured in treatments based on single conformations but obviously relevant for nonrigid molecules. Empirically calibrated approaches may absorb these effects in their parameters, as noted previously for G4.[33,51] A similar situation was encountered in the case of ozone where methods not reproducing "nonlocal" excitations beyond CCSD(T) must fail—the ATOMIC protocol included—but empirically calibrated approaches often reproduce experiment very well.[33] This point should be kept in mind when judging the relative merits of the ATOMIC protocol and of empirically calibrated models. A preliminary statistical analysis[51] compiled for a somewhat smaller test set,[187] has indeed indicated that ATOMIC models A, B_1, B_2, and B_3 perform *significantly* better than G3 and comparable to the much improved G4 model, if the known problem case of ozone is excluded and proper thermal corrections are applied for the series of *n*-alkanes.

The ATOMIC protocol is a fully ab initio approach to thermochemistry, first because no empirical parameters other than frequency scaling factors are used, and second because reference to experiment is made only for purposes of validation, but not to guide development. Results reported here and in our previous study[51] are quite encouraging, and they demonstrate that the ATOMIC protocol can provide high-quality estimates of both atomization energies and heats of formation. The new approach is not meant to be a panacea for all problems in thermochemistry. Instead it focuses on molecules with well-defined Lewis valence structures and taps the full potential of bond separation reactions to improve the accuracy of theoretical predictions. It does so quite successfully, even for more difficult molecules which exhibit charge-separated valence structures. Extensions to second-row elements, to ionic systems, and to radicals may require additional basis set considerations and, possibly, rules to generate uniquely defined valence structures, but they appear quite possible otherwise.

Acknowledgment. The author thanks the Swiss National Supercomputing Center (CSCS, Manno TI) for a generous grant of computer time.

Note Added in Proof. After submission of this paper, a Weizmann-theory[12−14] study on linear and branched alkanes has appeared,[188] which provides additional data in support of the analysis documented in section 6.4.

References and Notes

(1) Rossini, F. D. *Chem. Rev.* **1936**, *18*, 233–256.

(2) Cox, J. D.; Pilcher, G. *Thermochemistry of organic and organometallic compounds*; Academic Press: London, 1970.

(3) Benson, S. W.; Buss, J. H. *J. Chem. Phys.* **1958**, *29*, 546–572.

(4) Benson, S. W.; Cruickshank, F. R.; Golden, D. M.; Haugen, G. R.; O'Neal, H. E.; Rodgers, A. S.; Shaw, R.; Walsh, R. *Chem. Rev.* **1969**, *69*, 279–324.

(5) Cioslowski, J. *Quantum-mechanical prediction of thermochemical data*; Kluwer: Dordrecht, 2001.

(6) Helgaker, T.; Klopper, W.; Tew, D. P. *Mol. Phys.* **2008**, *106*, 2107–2143.

(7) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

(8) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(9) East, A. L. L.; Allen, W. D. *J. Chem. Phys.* **1993**, *99*, 4638–4650.

(10) Klippenstein, S. J.; East, A. L. L.; Allen, W. D. *J. Chem. Phys.* **1996**, *105*, 118–140.

(11) Császár, A. G.; Allen, W. D.; Schaefer III, H. F. *J. Chem. Phys.* **1998**, *108*, 9751–9764.

(12) Martin, J. M. L.; de Oliveira, G. *J. Chem. Phys.* **1999**, *111*, 1843–1856.

(13) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kállay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 4129–4141.

(14) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108.

(15) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599–11613.

(16) Bomble, Y. J.; Vázquez, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2006**, *125*, 064108.

(17) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2008**, *128*, 114111.

(18) Feller, D.; Dixon, D. A.; Peterson, K. A. *J. Phys. Chem. A* **1998**, *102*, 7053–7059.

(19) Dixon, D. A.; Feller, D.; Sandrone, G. *J. Phys. Chem. A* **1999**, *103*, 4744–4751.

(20) Feller, D.; Dixon, D. A. *J. Phys. Chem. A* **1999**, *103*, 6413–6419.

(21) Feller, D.; Dixon, D. A. *J. Phys. Chem. A* **2000**, *104*, 3048–3056.

(22) Feller, D.; Dixon, D. A. *J. Chem. Phys.* **2001**, *115*, 3484–3496.

(23) Feller, D.; Dixon, D. A. *J. Phys. Chem. A* **2003**, *107*, 9641–9651.

(24) Feller, D.; Peterson, K. A.; Dixon, D. A. *J. Chem. Phys.* **2008**, *129*, 204105.

(25) Bak, K. L.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Klopper, W. *J. Chem. Phys.* **2000**, *112*, 9229–9242.

(26) Helgaker, T.; Klopper, W.; Halkier, A.; Bak, K. L.; Jorgensen, P.; Olsen, J. In *Quantum-mechanical prediction of thermochemical data*; Cioslowski, J., Ed.; Kluwer: Dordrecht, 2001; pp 1−30.

(27) Helgaker, T.; Ruden, T. A.; Jørgensen, P.; Olsen, J.; Klopper, W. *J. Phys. Org. Chem.* **2004**, *17*, 913–933.

(28) Noga, J.; Valiron, P.; Klopper, W. *J. Chem. Phys.* **2001**, *115*, 2022–2032.

(29) Klopper, W.; Noga, J. *ChemPhysChem* **2003**, *4*, 32–48.

(30) Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622–5629.

(31) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221–7230.

(32) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764–7776.

(33) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *126*, 084108.

(34) Montgomery, J. A., Jr.; Ochterski, J. W.; Petersson, G. A *J. Chem. Phys.* **1994**, *101*, 5900–5909.

(35) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A, Jr. *J. Chem. Phys.* **1996**, *104*, 2598–2619.

(36) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, *110*, 2822–2827.

(37) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **2000**, *112*, 6532–6542.

(38) Petersson, G. A. In *Quantum-mechanical prediction of thermochemical data*; Cioslowski, J., Ed.; Kluwer: Dordrecht, 2001; pp 99−130.

(39) Fast, P. L.; Sánchez, M. L.; Corchado, J. C.; Truhlar, D. G. *J. Chem. Phys.* **1999**, *110*, 11679–11681.

(40) Fast, P. L.; Sánchez, M. L.; Truhlar, D. G. *Chem. Phys. Lett.* **1999**, *306*, 407–410.

(41) Fast, P. L.; Corchado, J. C.; Sánchez, M. L.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 5129–5136.

(42) Fast, P. L.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 6111–6116.

(43) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 3898–3906.

(44) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 1643–1649.

(45) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 4786–4791.

(46) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 43–52.

(47) Ohlinger, W. S.; Klunzinger, P. E.; Deppmeier, B. J.; Hehre, W. J. *J. Phys. Chem. A* **2009**, *113*, 2165–2175.

(48) DeYonker, N. J.; Cundari, T. R.; Wilson, A. K. *J. Chem. Phys.* **2006**, *124*, 114104.

(49) DeYonker, N. J.; Grimes, T.; Yockel, S.; Dinescu, A.; Mintz, B.; Cundari, T. R.; Wilson, A. K. *J. Chem. Phys.* **2006**, *125*, 104111.

(50) DeYonker, N. J.; Wilson, B. R.; Pierpont, A. W.; Cundari, T. R.; Wilson, A. K. *Mol. Phys.* **2009**, *107*, 1107–1121.

(51) Bakowies, D. *J. Chem. Phys.* **2009**, *130*, 144113.

(52) Ditchfield, R.; Hehre, W. J.; Pople, J. A.; Radom, L. *Chem. Phys. Lett.* **1970**, *5*, 13–14.

(53) Hehre, W. J.; Ditchfield, R.; Radom, L.; Pople, J. A. *J. Am. Chem. Soc.* **1970**, *92*, 4796–4801.

(54) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

(55) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(56) Peterson, K. A.; Woon, D. E.; Dunning, T. H, Jr. *J. Chem. Phys.* **1994**, *100*, 7410–7415.

(57) Woon, D. E.; Dunning, T. H, Jr. *J. Chem. Phys.* **1994**, *100*, 2975–2988.

(58) Wilson, A. K.; van Mourik, T.; Dunning, T. H, Jr. *J. Mol. Struct.: THEOCHEM* **1996**, *388*, 339–349.

(59) Woon, D. E.; Dunning, T. H, Jr. *J. Chem. Phys.* **1995**, *103*, 4572–4585.

(60) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437–446.

(61) Bakowies, D. *J. Chem. Phys.* **2007**, *127*, 084105.

(62) Cowan, R. D.; Griffin, D. C. *J. Opt. Soc. Am.* **1976**, *66*, 1010–1014.

(63) Martin, R. L. *J. Phys. Chem.* **1983**, *87*, 750–754.

(64) Handy, N. C.; Yamaguchi, Y.; Schaefer, H. F., III. *J. Chem. Phys.* **1986**, *84*, 4481–4484.

(65) Handy, N. C.; Lee, A. M. *Chem. Phys. Lett.* **1996**, *252*, 425–430.

(66) Valeev, E. F.; Sherrill, C. D. *J. Chem. Phys.* **2003**, *118*, 3921–3927.

(67) Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1987**, *86*, 7041–7050.

(68) Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1988**, *89*, 3401–3401.

(69) Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1992**, *97*, 4282–4288.

(70) Weigend, F.; Häser, M. *Theor. Chem. Acc.* **1997**, *97*, 331–340.

(71) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(72) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.

(73) Hampel, C.; Peterson, K. A.; Werner, H.-J. *Chem. Phys. Lett.* **1992**, *190*, 1–12.

(74) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M. et al.; *MOLPRO, version 2002.6, a package of ab initio programs*; 2002; see http://www.molpro.net.

(75) Hättig, C. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59–66.

(76) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063–1079.

(77) McQuarrie, D. A. *Statistical Mechanics*; Harper Collins: New York, 1976.

(78) Note that force constants have been obtained from numerical differentiation of analytical gradients, as fully analytical procedures were not available for RI-MP2. Closer analysis has shown that the numerical precision of energies and gradients was insufficient to obtain reliable results for extremely low-frequency modes such as those in 2-butyne, cyclopentane, and 3-pentanone. Reoptimization of geometries and *analytical* evaluation of force constants at the conventional (non-RI) MP2 level recovered the expected small but real frequencies for the lowest-frequency modes of the true energy minima of 2-butyne (D_{3h}), cyclopentane (C_2), and 3-pentanone (C_{2v}) and confirmed the calculated (RI-MP2) frequencies of all other modes to within 1 cm^{-1} or better.

(79) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502–16513.

(80) Ayala, P. Y.; Schlegel, H. B. *J. Chem. Phys.* **1998**, *108*, 2314–2325.

(81) Katzer, G.; Sax, A. F. *J. Phys. Chem. A* **2002**, *106*, 7204–7215.

(82) Katzer, G.; Sax, A. F. *J. Comput. Chem.* **2005**, *26*, 1438–1451.

(83) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374–7383.

(84) Hout, R. F., Jr.; Levi, B. A.; Hehre, W. J. *J. Comput. Chem.* **1982**, *3*, 234–250.

(85) Pople, J. A.; Schlegel, H. B.; Krishnan, R.; Defrees, D. J.; Binkley, J. S.; Frisch, M. J.; Whiteside, R. A.; Hout, R. F.; Hehre, W. J. *Int. J. Quantum Chem.* **1981**, *S15*, 269–278.

(86) Grev, R. S.; Janssen, C. L.; Schaefer, H. F., III *J. Chem. Phys.* **1991**, *95*, 5128–5132.

(87) Pople, J. A.; Scott, A. P.; Wong, M. W.; Radom, L. *Isr. J. Chem.* **1993**, *33*, 345–350.

(88) Sinha, P.; Boesch, S. E.; Gu, C.; Wheeler, R. A.; Wilson, A. K. *J. Phys. Chem. A* **2004**, *108*, 9213–9217.

(89) Karton, A.; Ruscic, B.; Martin, J. M. L. *J. Mol. Struct.: THEOCHEM* **2007**, *811*, 345–353.

(90) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley, Chichester, U.K., 2000.

(91) Lee, T. J.; Scuseria, G. E. *J. Chem. Phys.* **1990**, *93*, 489–494.

(92) Barbe, A.; Secroun, C.; Jouve, P. *J. Mol. Spectrosc.* **1974**, *49*, 171–182.

(93) Petersson, G. A.; Malick, D. K.; Wilson, W. G.; Ochterski, J. W.; Montgomery, J. A.; Frisch, M. J. *J. Chem. Phys.* **1998**, *109*, 10570–10579.

(94) Grimme, S. *J. Phys. Chem. A* **2005**, *109*, 3067–3077.

(95) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 42–55.

(96) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2005**, *123*, 124107.

(97) Pedley, J. B.; Naylor, R. D.; Kirby, S. P. *Thermochemical data of organic compounds*, 2nd ed.; Chapman and Hall: London, 1986.

(98) Lias, S. G.; Bartmess, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G. *J. Phys. Chem. Ref. Data* **1988**, *17S1*, 1–861.

(99) Cioslowski, J.; Schimeczek, M.; Liu, G.; Stoyanov, V. *J. Chem. Phys.* **2000**, *113*, 9377–9389.

(100) Kybett, B. D.; Carroll, S.; Natalis, P.; Bonnell, D. W.; Margrave, J. L.; Franklin, J. L. *J. Am. Chem. Soc.* **1966**, *88*, 626–626.

(101) Korsunskii, B. L.; Pepekin, V. I.; Lebedev, Y. A.; Apin, A. Y. *Bull. Acad. Sci. USSR, Div. Chem. Sci.* **1967**, 509–511.

(102) Baghal-Vayjooee, M. H.; Collister, J. L.; Pritchard, H. O. *Can. J. Chem.* **1977**, *55*, 2634–2636.

(103) Shaw, R.; Golden, D. M.; Benson, S. W. *J. Phys. Chem.* **1977**, *81*, 1716–1729.

(104) Pau, C.-F.; Hehre, W. J. *J. Phys. Chem.* **1982**, *86*, 321–322.

(105) An, X.-W.; Månsson, M. *J. Chem. Thermodyn.* **1983**, *15*, 287–293.

(106) Kabo, G. Y.; Miroshnichenko, E. A.; Frenkel, M. L.; Kozyro, A. A.; Simirskii, V. V.; Krasulin, A. P.; Vorob'eva, V. P.; Lebedev, Y. A. *Bull. Acad. Sci. USSR, Div. Chem. Sci.* **1990**, *39*, 662–667.

(107) Biehl, H.; Stuhl, F. *J. Chem. Phys.* **1994**, *100*, 141–145.

(108) Anderson, W. R. *Combust. Flame* **1999**, *117*, 394–403.

(109) Ramachandran, B.; Vegesna, N. S.; Peterson, K. A. *J. Phys. Chem. A* **2003**, *107*, 7938–7944.

(110) Verevkin, S. P.; Emel'yanenko, V. N. *J. Phys. Chem. A* **2004**, *108*, 6575–6580.

(111) Bashir-Hashemi, A.; Chickos, J. S.; Hanshaw, W.; Zhao, H.; Farivar, B. S.; Liebman, J. F. *Thermochim. Acta* **2004**, *424*, 91–97.

(112) Acree, W. E., Jr.; Pilcher, G.; Ribeiro da Silva, M. D. M. C. *J. Phys. Chem. Ref. Data* **2005**, *34*, 553–572.

(113) Gurvich, L. V.; Veyts, I. V.; Alcock, C. B. *Thermodynamic properties of individual substances*; Hemisphere Publishing Corp.: New York, 1989; Vol. 1, part 2.

(114) Chase, M. W., Jr.; Davies, C. A.; Downey, J. R., Jr.; Frurip, D. J.; McDonald, R. A.; Syverud, A. N *J. Phys. Chem. Ref. Data* **1985**, *14S1*, 1–1856.

(115) da Silva, G.; Bozzelli, J. W.; Sebbar, N.; Bockhorn, H. *ChemPhysChem* **2006**, *7*, 1119–1126.

(116) Wang, H.; Law, C. K. *J. Phys. Chem. B* **1997**, *101*, 3400–3403.

(117) Diky, V. V.; Frenkel, M.; Karpushenkava, L. S. *Thermochim. Acta* **2003**, *408*, 115–121.

(118) Kirklin, D. R.; Churney, K. L.; Domalski, E. S. *J. Chem. Thermodyn.* **1989**, *21*, 1105–1113.

(119) Roux, M. V.; Dávalos, J. Z.; Jiménez, P.; Notario, R.; Castaño, O.; Chickos, J. S.; Hanshaw, W.; Zhao, H.; Rath, N.; Liebman, J. F.; Farivar, B. S.; Bashir-Hashemi, A. *J. Org. Chem.* **2005**, *70*, 5461–5470.

(120) Asher, R. L.; Appelman, E. H.; Ruscic, B. *J. Chem. Phys.* **1996**, *105*, 9781–9795.

(121) Suzuki, K.; Ōnishi, S.-i.; Koide, T.; Seki, S. *Bull. Chem. Soc. Jpn.* **1956**, *29*, 127–131.

(122) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A.; *Gaussian 98 (revision A.9)*; Gaussian, Inc.: Pittsburgh, PA, 1998.

(123) Raghavachari, K.; Stefanov, B. B.; Curtiss, L. A. *J. Chem. Phys.* **1997**, *106*, 6764–6767.

(124) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Stefanov, B. B. *J. Chem. Phys.* **1998**, *108*, 692–697.

(125) Raghavachari, K.; Stefanov, B. B.; Curtiss, L. A. *Mol. Phys.* **1997**, *91*, 555–559.

(126) Martin, J. M. L.; François, J.-P.; Gijbels, R. *J. Phys. Chem.* **1994**, *98*, 11394–11400.

(127) Dixon, D. A.; Feller, D. *J. Phys. Chem. A* **1998**, *102*, 8209–8216.

(128) Bauschlicher, C. W., Jr.; Ricca, A. *Chem. Phys. Lett.* **1999**, *315*, 449–453.

(129) Bauschlicher, C. W., Jr.; Ricca, A. *J. Phys. Chem. A* **2000**, *104*, 4581–4585.

(130) Schuurman, M. S.; Muir, S. R.; Allen, W. D.; Schaefer, H. F., III. *J. Chem. Phys.* **2004**, *120*, 11586–11599.

(131) Zyrianov, M.; Droz-Georget, T.; Sanov, A.; Reisler, H. *J. Chem. Phys.* **1996**, *105*, 8111–8116.

(132) Paulett, G. S.; Ettinger, R. *J. Chem. Phys.* **1963**, *39*, 825–827.

(133) Laufer, A. H.; Okabe, H. *J. Phys. Chem.* **1972**, *76*, 3504–3507.

(134) Neugebauer, C. A.; Margrave, J. L. *J. Phys. Chem.* **1956**, *60*, 1318–1321.

(135) Kolesov, V. P.; Martynov, A. M.; Shtekher, S. M.; Skuratov, S. M. *Russ. J. Phys. Chem.* **1962**, *36*, 1118–1120.

(136) Moffat, J. B.; Knowles, A. J. *J. Chem. Eng. Data* **1969**, *14*, 215–217.

(137) Herzberg, G. *Electronic spectra and electronic structure of polyatomic molecules*; Van Nostrand: New York, 1966.

(138) von Wartenberg, H.; Schütza, H. *Z. Phys. Chem., Abt. A* **1933**, *164*, 386–388.

(139) Knowlton, J. W.; Prosen, E. J. *J. Res. Natl. Bur. Stand.(U.S.)* **1951**, *46*, 489–495.

(140) Brown, K. W.; Nibler, J. W.; Hedberg, K.; Hedberg, L. *J. Phys. Chem.* **1989**, *93*, 5679–5684.

(141) Miller, F. A.; Hannan, R. B., Jr.; Cousins, L. R. *J. Chem. Phys.* **1955**, *23*, 2127–2129.

(142) Armstrong, G. T.; Marantz, S. *J. Phys. Chem.* **1960**, *64*, 1776–1778.

(143) Armstrong, G. T.; Marantz, S. *J. Phys. Chem.* **1963**, *67*, 2888–2888.

(144) Saggiomo, A. J. *J. Org. Chem.* **1957**, *22*, 1171–1175.

(145) Bauder, A.; Günthard, H. H. *Helv. Chim. Acta* **1958**, *41*, 670–673.

(146) Laufer, A. H.; Okabe, H. *J. Am. Chem. Soc.* **1971**, *93*, 4137–4140.

(147) Chen, S. S.; Rodgers, A. S.; Chao, J.; Wilhoit, R. C.; Zwolinski, B. J. *J. Phys. Chem. Ref. Data* **1975**, *4*, 441–456.

(148) DeFrees, D. J.; Hehre, W. J. *J. Phys. Chem.* **1978**, *82*, 391–393.

(149) Fattahi, A.; Lis, L.; Kass, S. R. *J. Am. Chem. Soc.* **2005**, *127*, 13065–13069.

(150) Lee, T. J.; Martin, J. M. L.; Dateo, C. E.; Taylor, P. R. *J. Phys. Chem.* **1995**, *99*, 15858–15863.

(151) Lee, T. J.; Rice, J. E.; Dateo, C. E. *Mol. Phys.* **1996**, *89*, 1359–1372.

(152) Martin, J. M. L.; Taylor, P. R. *Mol. Phys.* **1999**, *96*, 681–692.

(153) De Oliveira, G.; Martin, J. M. L.; Silwal, I. K. C.; Liebman, J. F. *J. Comput. Chem.* **2001**, *22*, 1297–1305.

(154) Parthiban, S.; Martin, J. M. L.; Liebman, J. F. *Mol. Phys.* **2002**, *100*, 453–464.

(155) Feller, D.; Peterson, K. A.; de Jong, W. A.; Dixon, D. A. *J. Chem. Phys.* **2003**, *118*, 3510–3522.

(156) Feller, D.; Dixon, D. A.; Francisco, J. S. *J. Phys. Chem. A* **2003**, *107*, 1604–1617.

(157) Feller, D.; Dixon, D. A. *J. Phys. Chem. A* **2003**, *107*, 10419–10427.

(158) Dixon, D. A.; de Jong, W. A.; Peterson, K. A.; McMahon, T. B. *J. Phys. Chem. A* **2005**, *109*, 4073–4080.

(159) Pinnavaia, N.; Bramley, M. J.; Su, M.-D.; Green, W. H.; Handy, N. C. *Mol. Phys.* **1993**, *78*, 319–343.

(160) Koput, J.; Winnewisser, B. P.; Winnewisser, M. *Chem. Phys. Lett.* **1996**, *255*, 357–362.

(161) Redfern, P. C.; Zapol, P.; Curtiss, L. A.; Raghavachari, K. *J. Phys. Chem. A* **2000**, *104*, 5850–5854.

(162) Pollack, L.; Windus, T. L.; de Jong, W. A.; Dixon, D. A. *J. Phys. Chem. A* **2005**, *109*, 6934–6938.

(163) Tasi, G.; Mizukami, F.; Pálinkŏ, I.; Csontos, J.; Győrffy, W.; Nair, P.; Maeda, K.; Toba, M.; Niwa, S.-i.; Kiyozumi, Y.; Kiricsi, I. *J. Phys. Chem. A* **1998**, *102*, 7698–7703.

(164) Tasi, G.; Mizukami, F. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 632–638.

(165) Murcko, M. A.; Castejon, H.; Wiberg, K. B. *J. Phys. Chem.* **1996**, *100*, 16162–16168.

(166) Smith, G. D.; Jaffe, R. L. *J. Phys. Chem.* **1996**, *100*, 18718–18724.

(167) Tasi, G.; Mizukami, F.; Csontos, J.; Győrffy, W.; Pálinkŏ, I. *J. Math. Chem.* **2000**, *27*, 191–199.

(168) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.

(169) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(170) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–289.

(171) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119–124.

(172) von Arnim, M.; Ahlrichs, R. *J. Comput. Chem.* **1998**, *19*, 1746–1757.

(173) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

(174) Hirota, E. *J. Mol. Spectrosc.* **1979**, *77*, 213–221.

(175) Gray, D. L.; Robiette, A. G. *Mol. Phys.* **1979**, *37*, 1901–1920.

(176) Stanton, J. F. *Mol. Phys.* **1999**, *97*, 841–845.

(177) Duncan, J. L.; McKean, D. C.; Bruce, A. J. *J. Mol. Spectrosc.* **1979**, *74*, 361–374.

(178) Harmony, M. D. *J. Chem. Phys.* **1990**, *93*, 7522–7523.

(179) Tam, H. S.; Choe, J.-I.; Harmony, M. D. *J. Phys. Chem.* **1991**, *95*, 9267–9272.

(180) Roberto-Neto, O.; Chakravorty, S.; Machado, F. B. C. *J. Mol. Struct.: THEOCHEM* **2002**, *586*, 29–34.

(181) Tasi, G.; Izsák, R.; Matisz, G.; Császár, A. G.; Kállay, M.; Ruscic, B.; Stanton, J. F. *ChemPhysChem* **2006**, *7*, 1664–1667.

(182) Examples of molecules with charge-separated valence structures include nitro compounds (nitromethane, 2-nitrobutane, nitric acid, dinitrogen trioxide), methyl isocyanide, pyridine-*N*-oxide, nitrous oxide, and ozone from Table 3, and isocyanogen fluoride, formonitrile oxide, isofulminic acid, and diazomethane from Table 6. In all these cases, bond separation reactions involve cationic parent molecules.

(183) Helgaker, T.; Gauss, J.; Jørgensen, P.; Olsen, J. *J. Chem. Phys.* **1997**, *106*, 6430–6440.

(184) Bak, K. L.; Gauss, J.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. *J. Chem. Phys.* **2001**, *114*, 6548–6556.

(185) Riley, K. E.; Op't Holt, B. T.; Merz, K. M *J. Chem. Theory Comput.* **2007**, *3*, 407–433.

(186) Goldstein, E.; Ma, B.; Lii, J.-H.; Allinger, N. L. *J. Phys. Org. Chem.* **1996**, *9*, 191–202.

(187) The test set includes all neutral closed-shell molecules of the G3/99 set (ref 83) composed of hydrogen and at least three first-row atoms (C, N, O, F), except tetrafluoroethylene and carbonyl fluoride for which experimental numbers are suspect. See Table IV of ref 51.

(188) Karton, A.; Gruzman, D.; Martin, J. M. L. *J. Phys. Chem. A* **2009**, *113*, 8434–8447.

JP9027782