

HEAT: High accuracy extrapolated *ab initio* thermochemistry

Attila Tajti

Department of Theoretical Chemistry, Eötvös Loránd University, P.O. Box 32, H-1518 Budapest, Hungary

Péter G. Szalay

*Department of Theoretical Chemistry, Eötvös Loránd University, P.O. Box 32, H-1518 Budapest, Hungary
and Institute for Theoretical Chemistry, Department of Chemistry and Biochemistry,
The University of Texas at Austin, Austin, Texas 78712*

Attila G. Császár

Department of Theoretical Chemistry, Eötvös Loránd University, P.O. Box 32, H-1518 Budapest, Hungary

Mihály Kállay and Jürgen Gauss

Institut für Physikalische Chemie, Universität Mainz, D-55099 Mainz, Germany

Edward F. Valeev

School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400

Bradley A. Flowers, Juana Vázquez, and John F. Stanton

*Institute for Theoretical Chemistry, Department of Chemistry and Biochemistry,
The University of Texas at Austin, Austin, Texas 78712*

(Received 9 August 2004; accepted 9 September 2004)

A theoretical model chemistry designed to achieve high accuracy for enthalpies of formation of atoms and small molecules is described. This approach is entirely independent of experimental data and contains no empirical scaling factors, and includes a treatment of electron correlation up to the full coupled-cluster singles, doubles, triples and quadruples approach. Energies are further augmented by anharmonic zero-point vibrational energies, a scalar relativistic correction, first-order spin-orbit coupling, and the diagonal Born–Oppenheimer correction. The accuracy of the approach is assessed by several means. Enthalpies of formation (at 0 K) calculated for a test suite of 31 atoms and molecules via direct calculation of the corresponding elemental formation reactions are within 1 kJ mol^{-1} to experiment in all cases. Given the quite different bonding environments in the product and reactant sides of these reactions, the results strongly indicate that even greater accuracy may be expected in reactions that preserve (either exactly or approximately) the number and types of chemical bonds. © 2004 American Institute of Physics. [DOI: 10.1063/1.1811608]

I. INTRODUCTION

At a time when the majority of practitioners have essentially abandoned the field in favor of density-functional theory, traditional quantum chemistry (comprising what have come to be known—somewhat pejoratively in some circles—as “wave function methods”) has evolved so that methods are now available that are capable of determining molecular properties at a very high level of accuracy. Instead of providing rough estimates of quantities such as vibrational frequencies and structural parameters, routinely applicable modern techniques are capable of $\approx 10 \text{ cm}^{-1}$ accuracy for *fundamental* vibrational frequencies^{1,2} and $\approx 0.002\text{--}0.003 \text{ Å}$ accuracy in equilibrium bond distances.³ With the seemingly constant improvements made in computer hardware technology, both the level of accuracy and the scope of systems suitable for treatment at a given level of accuracy will continue to grow.

One area where extremely high accuracy is generally useful and has significant impact is in the determination of thermochemical parameters. Enthalpies of formation, heat capacities, and standard entropies of molecular species entirely determine their thermodynamic fate. Experimental

methods based on calorimetry, kinetics, spectroscopy, and various ion cycles have been used for decades to determine these important quantities.^{4,5} As a result of this intense area of research, relatively tight bounds ($<10 \text{ kJ mol}^{-1}$) have been established for enthalpies of formation for many molecules that are stable enough to be studied easily in the laboratory.⁶ However, in two fields that are strongly dependent on accurate thermochemical information—combustion and atmospheric chemistry—many, and perhaps a majority, of the most important compounds are transient species. Accordingly, error bars on the enthalpies of formation for these radicals and other reactive molecules tend to be significantly larger than those for simple closed-shell species. As an example, the enthalpy of formation for the hydroperoxy radical (HO_2) has only recently been determined to better than 1 kJ mol^{-1} .⁷

The extent to which modern high-level *ab initio* (“wave function”) calculations can be competitive with experiment in the precise determination of thermodynamic parameters depends to a large extent on the size of the molecular species in question. For the smallest molecules (a dozen or fewer electrons), there is little question that theory can provide very accurate total electronic energies, irrespective of how

“difficult” the molecule is. For molecules that can be easily studied experimentally, theory can offer competitive precision for thermochemical parameters; for transient species not amenable to experimental characterization, theory is arguably better. A case in point are very recent studies of CH and CH₂,^{8,9} where accuracies estimated to be <0.5 kJ mol⁻¹ were achieved for the enthalpies of formation, as well as others of nearly comparable accuracy.^{7,10–20}

Of course, it is not the total energies of molecules that are relevant for thermochemistry, but rather appropriately defined energy differences. Hence, it is not necessary for a method to obtain a given level of accuracy (say 1 kJ mol⁻¹) in total energies in order to achieve the same level of accuracy for energy differences. Instead, one can benefit from error cancellation; deficiencies of the total energy calculations for the various species have many common sources. Total molecular energies, of course, are those relative to the separated atoms of the molecule in question, all completely ionized with the electrons at rest and the nuclei in their ground states. For example, the total electronic energy of carbon monoxide is $E(\text{C}^{6+}) + E(\text{O}^{8+}) + E(14e^-) - E(\text{CO})$. However, it is a much simpler task to calculate the bond energy of CO, viz. $E[\text{C}(^3\text{P})] + E[\text{O}(^3\text{P})] - E[\text{CO}(^1\Sigma)]$ since the core electrons—which make the largest contribution to the total energy due to strong nuclear attraction forces—of C and O are only slightly perturbed in the molecular environment. The more “similar” the *A* and *B* species involved in the energy difference $E(A) - E(B)$, the less demanding is the calculation needed to achieve a specified level of accuracy. This was realized long ago; so-called “isodesmic” reactions²¹—those in which the number and qualitative “types” of bonds in *A* and *B* are the same—are known to be those in which calculated energy differences tend toward the highest accuracy.²²

The preceding paragraph, while elementary and straightforward, is important to one of the themes of this paper. Most (perhaps all) “theoretical model chemistries”^{23–25} (those in which all species are treated at a consistent and well-defined level of approximation) use atomization energies as the basis for thermochemistry. In this approach, total molecular energies are calculated at some level of theory. Atomization energies, defined for the molecule $M \equiv A_a B_b \cdots Z_z$ as

$$AE(M) = aE(A) + bE(B) + \cdots + zE(Z) - E(M) \quad (1)$$

are then calculated using total atomic energies obtained at the same level of theory. The enthalpy of formation for *M* at 0 K is then given by Hess’s law as

$$\Delta_f H^\circ = a\Delta_f H^\circ(A) + b\Delta_f H^\circ(B) + \cdots + z\Delta_f H^\circ(Z) - AE(M), \quad (2)$$

where the atomic enthalpies of formation are set to literature values. With the notable exceptions of carbon and fluorine,¹² enthalpies of formation are known quite precisely for atoms in the first two rows of the periodic table, so the inherent error in this approach is often localized almost entirely in the atomization energies. However, it is very difficult to calculate atomization energies. By definition, all bonds in the molecule *M* are destroyed in the reaction that is used as the basis for the calculated quantity; it is as far from an isodesmic

reaction as possible. Another approach, rarely used in practice, would be to calculate the enthalpies of formation *directly*, meaning that the molecular energy of *M* and those of its constituent elements in their most stable form are used. For example, if *M* were water, the reaction considered would be



The enthalpy of formation for water is given, by definition, as the difference of ground state energies in the reactants and products in Eq. (3), and its calculation therefore does not make any use of experimental quantities. This method does have some serious limitations, however. For example, any organic molecule would require an (impossible) calculation of graphite; sulfur would necessitate calculations on a molecule with 128 electrons (S₈), chlorine would require *ab initio* calculations of the liquid and the enthalpy of vaporization, etc. Moreover, while chemical bonds can indeed be found on both sides of Eq. (3), it is clearly not an isodesmic reaction, or even nearly so. While it might be easier to calculate this energy difference accurately than that of the atomization reaction $\text{H}_2\text{O} \rightarrow 2\text{H} + \text{O}$, the benefits of doing so are marginal. The atomic enthalpies of formation for H and O are also known so precisely (<0.01 kJ mol⁻¹) that there is really no advantage in eliminating the use of this experimental information. Hence, due to simplicity, the well-defined nature of the process, and the generally straightforward problems posed by atoms to theoretical treatment, the atomization energy approach has been the method of choice for quantum chemical practitioners in estimating enthalpies of formation for molecules with various model chemistries.

One purpose of this work is to present a new theoretical model chemistry. Unlike previous efforts (the only possible exception being the recent definition of W3 theory by Martin and collaborators²⁵), we propose a method that cannot be generally applied to all “small” (loosely defined here as those having five or fewer non-hydrogen atoms) molecules. Rather, it is our intent to define an approach that is essentially the best that can be done—with current computer technology—for very small molecules (four or fewer *total* atoms), and to assess the level of accuracy achieved. It is our belief that this is not an academic exercise with only a narrow practical benefit. First, there are still molecules of this size where accurate (<1 kJ mol⁻¹) enthalpies of formation are not available⁶ [for example, NH, NH₂ and, until recently, OH (Ref. 13) and HO₂ (Ref. 7)]. Second, the accuracy that theory can achieve for small systems is of intrinsic interest, because it is useful to know how large a role is played by some usually neglected effects (coupling of core and valence correlation, relativistic corrections, the diagonal Born–Oppenheimer correction, anharmonic contributions to zero-point vibrational energies, spin–orbit coupling, etc.) in overall accuracy. Third, and most important, since the methods used in our work are “size extensive”^{26,27} (meaning that the quality of the energy calculation is not degraded by the size of the molecule described within a given one-particle basis set), the accuracy achieved for the benchmark systems stud-

TABLE I. Contributions to the HEAT total energies for the 31 species studied in this work. All values are in atomic units. Conversion factor used $2625.4976 \text{ kJ mol}^{-1} = 1 E_h$.

| Species | E_{HF}^{∞} | $\Delta E_{\text{CCSD(T)}}^{\infty}$ | ΔE_{CCSDT} | ΔE_{CCSDTQ} | ΔE_{REL} | ΔE_{ZPE} | ΔE_{DBOC} | ΔE_{SO} | Total |
|-------------------------------|--------------------------|--------------------------------------|---------------------------|----------------------------|-------------------------|-------------------------|--------------------------|------------------------|--------------|
| N ₂ | -108.993 257 | -0.549 274 | 0.000 507 | -0.001 457 | -0.058 658 | 0.005 379 | 0.003 982 | 0.000 000 | -109.592 778 |
| H ₂ | -1.133 661 | -0.040 911 | 0.000 000 | 0.000 000 | -0.000 010 | 0.009 892 | 0.000 460 | 0.000 000 | -1.164 230 |
| F ₂ | -198.774 570 | -0.756 425 | 0.000 100 | -0.001 536 | -0.174 461 | 0.002 095 | 0.005 175 | 0.000 000 | -199.699 622 |
| O ₂ | -149.691 925 | -0.635 217 | 0.000 112 | -0.001 854 | -0.104 607 | 0.003 641 | 0.004 711 | -0.000 012 | -150.425 151 |
| C | -37.693 774 | -0.151 041 | -0.000 466 | -0.000 030 | -0.015 090 | 0.000 000 | 0.001 660 | -0.000 144 | -37.858 885 |
| F | -99.416 800 | -0.318 033 | -0.000 199 | -0.000 116 | -0.087 268 | 0.000 000 | 0.002 591 | -0.000 574 | -99.820 399 |
| H | -0.500 022 | 0.000 000 | 0.000 000 | 0.000 000 | -0.000 007 | 0.000 000 | 0.000 272 | 0.000 000 | -0.499 757 |
| N | -54.404 657 | -0.184 700 | -0.000 344 | -0.000 042 | -0.029 435 | 0.000 000 | 0.002 007 | 0.000 000 | -54.617 171 |
| O | -74.819 232 | -0.248 562 | -0.000 338 | -0.000 078 | -0.052 459 | 0.000 000 | 0.002 366 | -0.000 312 | -75.118 615 |
| CO | -112.790 997 | -0.535 543 | 0.000 097 | -0.000 951 | -0.067 285 | 0.004 945 | 0.004 000 | 0.000 000 | -113.385 734 |
| C ₂ H ₂ | -76.855 684 | -0.480 411 | 0.000 223 | -0.000 911 | -0.029 760 | 0.026 253 | 0.003 674 | 0.000 000 | -77.336 616 |
| CCH | -76.183 645 | -0.428 923 | -0.001 147 | -0.000 928 | -0.029 761 | 0.013 842 | 0.003 503 | 0.000 000 | -76.627 059 |
| CH ₂ | -38.941 051 | -0.207 841 | -0.000 458 | -0.000 082 | -0.014 865 | 0.017 167 | 0.002 161 | 0.000 000 | -39.144 969 |
| CH | -38.284 553 | -0.194 296 | -0.000 628 | -0.000 076 | -0.015 029 | 0.006 463 | 0.002 063 | -0.000 067 | -38.486 123 |
| CH ₃ | -39.581 308 | -0.254 300 | -0.000 408 | -0.000 120 | -0.014 835 | 0.029 573 | 0.002 404 | 0.000 000 | -39.818 994 |
| CO ₂ | -187.725 719 | -0.876 563 | 0.000 567 | -0.001 753 | -0.119 227 | 0.011 580 | 0.006 314 | 0.000 000 | -188.704 801 |
| H ₂ O ₂ | -150.852 930 | -0.711 472 | 0.000 227 | -0.001 280 | -0.104 313 | 0.026 192 | 0.005 078 | 0.000 000 | -151.638 498 |
| H ₂ O | -76.067 761 | -0.371 594 | 0.000 033 | -0.000 453 | -0.052 040 | 0.021 228 | 0.002 710 | 0.000 000 | -76.467 877 |
| HCO | -113.304 223 | -0.553 349 | -0.000 125 | -0.000 943 | -0.067 113 | 0.012 960 | 0.004 336 | 0.000 000 | -113.908 457 |
| HF | -100.071 316 | -0.389 256 | 0.000 056 | -0.000 392 | -0.086 955 | 0.009 391 | 0.002 735 | 0.000 000 | -100.535 737 |
| HO ₂ | -150.253 106 | -0.661 501 | -0.000 447 | -0.001 116 | -0.104 469 | 0.014 110 | 0.004 996 | 0.000 000 | -151.001 533 |
| NO | -129.309 786 | -0.589 993 | -0.000 094 | -0.001 291 | -0.081 577 | 0.004 364 | 0.004 351 | -0.000 275 | -129.974 301 |
| OH | -75.428 343 | -0.310 311 | -0.000 289 | -0.000 259 | -0.052 261 | 0.008 461 | 0.002 619 | -0.000 297 | -75.780 680 |
| HNO | -129.850 244 | -0.634 499 | 0.000 253 | -0.001 469 | -0.081 448 | 0.013 680 | 0.004 732 | 0.000 000 | -130.548 995 |
| CN | -92.242 929 | -0.477 290 | -0.002 005 | -0.001 378 | -0.044 284 | 0.004 858 | 0.003 629 | 0.000 000 | -92.759 399 |
| HCN | -92.915 916 | -0.517 656 | 0.000 413 | -0.001 230 | -0.044 175 | 0.015 898 | 0.003 819 | 0.000 000 | -93.458 847 |
| CF | -137.239 487 | -0.551 585 | -0.000 435 | -0.000 525 | -0.102 084 | 0.003 002 | 0.004 239 | -0.000 178 | -137.887 053 |
| NH ₂ | -55.592 445 | -0.287 829 | -0.000 368 | -0.000 224 | -0.029 194 | 0.018 882 | 0.002 564 | 0.000 000 | -55.888 614 |
| NH ₃ | -56.225 187 | -0.339 348 | -0.000 104 | -0.000 316 | -0.029 045 | 0.034 069 | 0.002 609 | 0.000 000 | -56.557 322 |
| NH | -54.986 522 | -0.235 129 | -0.000 454 | -0.000 127 | -0.029 323 | 0.007 412 | 0.002 353 | -0.000 001 | -55.241 791 |
| OF | -174.211 642 | -0.674 172 | -0.000 938 | -0.000 994 | -0.139 570 | 0.002 426 | 0.004 951 | -0.000 414 | -175.020 353 |

ied here will be the same as that for larger molecules that will be amenable to the treatment in the future.

Finally, the systems studied here become the initial members of a database of compounds that can be used for any number of thermochemical studies. Specifically, we believe that the *best* way to calculate a molecular enthalpy of formation is to use approaches other than atomization enthalpy and direct elemental reaction strategies summarized above. In an ideal world, an isodesmic reaction can be designed in which all participants other than the target molecule *M* have enthalpies of formation that are known precisely from experiment. Then, if total electronic energies from the high-level theoretical model chemistry defined in this paper are available for all species, the reaction energy can presumably be obtained with negligible theoretical error. Adjustment of the reaction energy to the enthalpy of formation of *M* (by appropriate addition and subtraction of experimental enthalpies of formation for the other species), should then give $\Delta_f H^\circ$ for *M* with extraordinary precision. However, it is recognized that this will not usually be possible, either because of inability to design a truly isodesmic reaction (radicals can be difficult in this regard) or the lack of precise thermochemical knowledge about some of the species in an appropriate reaction. Then, alternative strategies can be followed, using reactions that are not isodesmic (but clearly superior to atomization schemes!) but involve at least an

approximate conservation of bond types. Such an approach was recently used by us to determine the enthalpy of formation of HO₂ to an accuracy of $\approx 0.5 \text{ kJ mol}^{-1}$, nearly an order of magnitude better than the definition of “chemical accuracy” (1 kcal mol^{-1} or $4.184 \text{ kJ mol}^{-1}$), and by Schuurman *et al.* in a study of HNCO isomers.¹⁰

The next few sections define the theoretical model chemistry that we have named HEAT. This is an acronym for “high accuracy extrapolated *ab initio* thermochemistry,” which emphasizes two things—apart from energy extrapolation schemes, the approach involves no empirical scaling factors or adjustments, and that the principal area of application that we envision for HEAT will be in the area of thermochemistry. After defining the method, and discussing the various theoretical approaches used to determine the total energies that are the “bottom line” of HEAT for any atom or molecule, the approach will be applied to a test suite of atoms and molecules. While quantities such as atomization energies and enthalpies of formation calculated from them (using the approaches discussed above) as well as from elemental reactions (excepting carbon) will be presented and discussed, we emphasize that it is the total energies obtained by the HEAT protocol that are the most important numbers documented in this work. Hopefully, the total energies for 31 atoms and molecules found later in this paper (Table I)—and straightforwardly calculable (at least in principle) for other

species—will form the initial members of a database that can be used by experimentalists and theorists alike to calculate enthalpies of formation using suitable *reaction-based* approaches.

II. DEFINITION OF HEAT MODEL CHEMISTRY

In order to determine standard enthalpies of formation at 0 K,²⁸ it is always necessary to know the ground state energy of the target species. Within the Born–Oppenheimer approximation, the ground state energy may be partitioned into electronic and vibrational contributions. The former is given by the lowest eigenvalue of the electronic Hamiltonian (electronic energy) at the equilibrium geometry; the latter by the lowest eigenvalue of the nuclear Hamiltonian containing a potential described by electronic energies as a function of position. In some cases, the lowest rotational state is prohibited by nuclear spin statistics, but we ignore this here.

The electronic part, as usually calculated, involves use of the nonrelativistic electronic Hamiltonian. The three simplest improvements upon this are to include “scalar” relativistic effects²⁹ using perturbation theory (which is believed to be entirely adequate for atoms in the first two rows of the periodic table³⁰), splitting of the energy—and lowering of the ground state energy—by spin–orbit interactions, and to compute the so-called diagonal Born–Oppenheimer correction.^{31–34} The latter is given by the expectation value of the nuclear kinetic energy operator over the electronic wave function (that which diagonalizes the electronic Hamiltonian), and is a first-order correction to the simple Born–Oppenheimer approximation that does not spoil the concept of a potential energy surface.³⁵

In practice, all of the contributions above need to be calculated approximately. First, finite basis sets must be used. Even if the treatment of correlation was complete (full configuration interaction, or FCI), the resulting energies would be compromised by limitations of the one-particle basis set. It is nonetheless impossible to do FCI calculations for all but the smallest molecules—even then in necessarily small basis sets—so approximate measures for treating correlation are called for. In HEAT, as in all other model chemistries that we know of, size-extensive many-body methods are used in the treatment of electron correlation.

The total energy defined by the HEAT protocol may be expressed by a formula that contains eight terms

$$E_{\text{HEAT}} = E_{\text{HF}}^{\infty} + \Delta E_{\text{CCSD(T)}}^{\infty} + \Delta E_{\text{CCSDT}} + \Delta E_{\text{CCSDTQ}} + \Delta E_{\text{REL}} + \Delta E_{\text{ZPE}} + \Delta E_{\text{DBOC}} + \Delta E_{\text{SO}}. \quad (4)$$

In Eq. (4), E_{HF}^{∞} and $\Delta E_{\text{CCSD(T)}}^{\infty}$ are the HF-SCF (Hartree–Fock self-consistent field) and correlation energies, the latter given by the coupled-cluster singles and doubles method³⁶ with a perturbative treatment of triple excitations,³⁷ both extrapolated to the basis set limit. The next term is intended to account for deficiencies in the treatment of triple excitations in CCSD(T), the fourth term to account, approximately, for differences between the CCSDT (Ref. 38) and FCI correlation energies, where the latter is approximated by the CCSDTQ method.³⁹ ΔE_{ZPE} is the zero-point vibrational energy. The remaining terms remedy shortcomings of the

simple, nonrelativistic Born–Oppenheimer approximation: ΔE_{DBOC} is the diagonal Born–Oppenheimer correction, ΔE_{SO} is the spin–orbit correction, and ΔE_{REL} is the scalar relativistic contribution to the energy.

We now describe how the individual terms in Eq. (4) are calculated.

A. Molecular geometries

The geometries⁴⁰ of species are taken from optimizations carried out at the CCSD(T) level of theory with the correlation-consistent cc-pVQZ basis sets.⁴¹ A recent benchmark study³ has demonstrated that this level of theory gives equilibrium geometries that are accurate to <0.003 Å (bond lengths) and $<0.5^\circ$ (angles) of experimentally inferred values, where the latter are available. This level of accuracy is comparable to the best that can be achieved solely through analysis of experimental data, since experimental measurements of geometrical parameters never correspond to equilibrium distances and angles, but rather some type of averaged quantities.⁴² Moreover, we correlate *all* electrons in the geometry optimizations, not just the valence electrons. While this is clearly not as good as using a properly defined (and much larger) core correlation basis (cc-pCVXZ),⁴³ it is considerably cheaper and the aforementioned benchmark study³ demonstrated that the approach used here gives geometries that differ only negligibly for molecules containing first- and second-row atoms from those obtained in full-blown CCSD(T)/cc-pCVQZ optimizations. These geometries are then used in all subsequent calculations of quantities contributing to the HEAT energy, and the vibrational problem required for ΔE_{ZPE} is solved with CCSD(T)/cc-pVQZ at this geometry, as well. For closed-shell molecules, the restricted Hartree–Fock (RHF) orbitals are used in all calculations. For open-shell molecules, there are two obvious choices: unrestricted or restricted open-shell Hartree–Fock (UHF and ROHF, respectively). It was our intent to define the HEAT strategy in terms of UHF-based calculations. There are at least two cogent reasons for this choice. At very high levels of theory such as CCSDT, differences between UHF- and ROHF-based total energies are usually very small, so that the choice of reference function should have no impact on the energy calculations. Second, ROHF methods are more prone to symmetry-breaking and related effects,⁴⁴ and its use very often gives rise to nonsensical vibrational frequencies. However, in the course of this work, a rather curious problem was noted for the diatomic NO: UHF-based calculations give absurd parameters for the anharmonic force field; these severely degrade the quality of the calculated vibrational zero-point energy. This problem has been analyzed, and our findings can be found in a separate publication.⁴⁵ Hence, for the moment, the HEAT method will use UHF orbitals as the default for HF-SCF and CCSD(T) calculations on open-shell molecules; systems for which ROHF orbitals turn out to be more appropriate will be so designated.

B. CCSD(T) total energy

Together, the E_{HF}^{∞} and $\Delta E_{\text{CCSD(T)}}^{\infty}$ terms give the estimated exact nonrelativistic electronic energy within the simple Born–Oppenheimer approximation, using the well-

known CCSD(T) method to account for electron correlation effects. Following a relatively common convention, HF-SCF and CCSD(T) correlation energies have been obtained in a hierarchical series of basis sets, and then extrapolated *separately* to obtain estimates of the corresponding basis set limits. For closed-shell molecules, the restricted Hartree–Fock method has been used. Open-shell molecules were treated using the UHF approach.

For the HF-SCF energy, calculations were carried out using the augmented correlation consistent basis sets aug-cc-pCVXZ [$X = T(3)$, $Q(4)$ and 5] (Ref. 46) which are designed to treat core correlation effects properly. These three energies were then extrapolated with the formula advocated by Feller,⁴⁷

$$E_{\text{HF}}^X = E_{\text{HF}}^\infty + a \exp(-bX), \quad (5)$$

where E_{HF}^X is the HF-SCF energy obtained with the aug-cc-pCVXZ basis set. The parameters, a , b , and the extrapolated HF-SCF energy E_{HF}^∞ , are determined uniquely from the three energies.

For the correlation energy, a formula motivated by the atomic partial wave expansion is used,⁴⁸ viz.⁴⁹

$$\Delta E_{\text{CCSD(T)}}^X = \Delta E_{\text{CCSD(T)}}^\infty + \frac{a}{X^3}, \quad (6)$$

where $\Delta E_{\text{CCSD(T)}}^X$ is the CCSD(T) correlation energy [not the total CCSD(T) energy, which includes the HF-SCF contribution] obtained with the aug-cc-pCVXZ basis set. Here, there are two parameters, a and the estimated complete basis set limit CCSD(T) correlation energy $\Delta E_{\text{CCSD(T)}}^\infty$. These are uniquely determined by two correlation energies; the aug-cc-pCVQZ and aug-cc-pCV5Z energies are used in the HEAT protocol.

At this point, it is important to note a distinct difference between the HEAT approach and other model chemistries as well as most isolated efforts to obtain very accurate thermochemical parameters. We *do not* attempt to separate valence correlation effects from those arising from correlation of the core electrons. All of the calculations carried out in determining the extrapolated CCSD(T) energy use basis sets that are designed to treat core correlation as well as valence correlation, and no electrons are dropped from the correlation treatment in the individual CCSD(T) calculations. While it is true that one can obtain similar extrapolated valence-only estimates of the correlation energy and then add a correction for core correlation effects calculated with significantly smaller basis sets, we have chosen not to make the assumption of separation. Again, the point of this work is to do the best calculations possible in a common current computational environment with as few approximations as possible. However, we do recognize that this approach of combining core and valence correlation effects runs counter to the common practice of many of our colleagues, but point out that our strategy is undeniably more rigorous.⁵⁰

C. Higher level correlation effects

Despite the never-ending success story that is the CCSD(T) method, it must not be forgotten that the (T)

correction³⁷ is based on perturbation theory.⁵¹ For cases where triples corrections are large, or alternatively (but not entirely independently) when there are severe problems associated with the reference function, there is cause to investigate the extent to which CCSD(T) differs from the complete treatment of triple excitations defined by the CCSDT approximation.³⁸ However, it is not possible to perform full CCSDT calculations using the large basis sets met with in the extrapolated CCSD(T) energies defined above. Due to both the relatively small differences expected and the fact that it appears that correlation effects beyond CCSD(T) can be estimated with smaller basis sets, we have chosen to estimate this contribution with the cc-pVTZ and cc-pVQZ basis sets and to correlate only the valence electrons. Implicitly assumed here is that effects due to diffuse functions and core correlation are already given sufficiently well by the extrapolated HF-SCF and CCSD(T) energies. Our formula for the CCSDT-CCSD(T) energy difference (ΔE_{CCSDT}) is

$$\Delta E_{\text{CCSDT}} = E_{\text{CCSDT}}^{\text{TQ}}(\text{fc}) - E_{\text{CCSD(T)}}^{\text{TQ}}(\text{fc}), \quad (7)$$

where TQ denotes that the corresponding contribution has been obtained by the correlation energy extrapolation formula [Eq. (6)] using the frozen-core CCSDT and CCSD(T) energies obtained with the cc-pVTZ and cc-pVQZ basis sets. For radicals, both the CCSDT and CCSD(T) calculations were performed using UHF reference functions.

Despite its computational complexity and cost, even CCSDT does not give correlation energies that are sufficiently accurate for the most demanding applications.^{7,25,47,52,53} However, coupled-cluster calculations beyond CCSDT have only been generally possible (for small molecules, of course) with the development of general coupled-cluster codes.^{54–59} Recently, Ruden and collaborators studied the impact of connected quadruples on atomization energies in double- and triple- ζ quality basis sets for six molecules.⁶⁰ The authors found that the contribution of quadruple excitations, as measured by the CCSDTQ-CCSDT energy difference, converges rapidly with basis set size. In absolute terms, changes are largely negligible when going beyond a polarized valence double- ζ basis. In a more recent study focusing on a larger sample of molecules, Boese and co-workers²⁵ reached similar conclusions.

Based on these findings, the effects of higher-level correlation effects (those beyond CCSDT) are estimated in the HEAT protocol by subtracting the CCSDT and CCSDTQ correlation energies obtained with the cc-pVDZ basis set in the frozen-core approximation:

$$\Delta E_{\text{CCSDTQ}} = E_{\text{CCSDTQ}}^{\text{cc-pVDZ}}(\text{fc}) - E_{\text{CCSDT}}^{\text{cc-pVDZ}}(\text{fc}). \quad (8)$$

Due to program limitations that existed while the data were being compiled for this research, the ΔE_{CCSDTQ} correction is based on ROHF reference functions for the radical species. There is an implicit assumption here—that the CCSDT total energies for radicals are independent of the reference function—but this seems to be justified.⁶¹

This approximation, which is intended to account for the difference between CCSDT and an exact treatment of correlation, is clearly important and is further discussed in the appendix.

D. Zero-point vibrational energy

Zero-point vibrational energies for all species (apart from atoms, where it vanishes) were determined from anharmonic force fields calculated at the (all electron) CCSD(T) level of theory with the cc-pVQZ basis sets, using RHF orbitals for closed-shell systems and UHF orbitals for most of the open-shell molecules.⁶² Using standard spectroscopic rovibrational perturbation theory,⁶³ the vibrational energy levels are given by the expression

$$E(v) = G_0 + \sum_i \omega_i \left(v_i + \frac{1}{2} \right) + \sum_{i \geq j} x_{ij} \left(v_i + \frac{1}{2} \right) \left(v_j + \frac{1}{2} \right), \quad (9)$$

where ω_i are the harmonic frequencies. Explicit equations for the anharmonicity constants x_{ij} are given, for example, in Ref. 63. Hence, the ground state energy is given by

$$E_{\text{ZPE}} = G_0 + \sum_i \frac{\omega_i}{2} + \sum_{i \geq j} \frac{x_{ij}}{4} \quad (10)$$

within this model. The second term is the familiar harmonic approximation, while addition of the third and (especially) first terms are rarely included in quantum chemical investigations. Work over the past six years in our laboratories has led to the development of analytic second derivative methods for the CCSD(T) method,^{64,65} and parallel developments have provided the avenue toward accurate and efficient evaluation of quartic force fields via numerical differentiation of analytic second derivatives. In fact, it is now possible to calculate the cubic and quartic force fields, many physical quantities that depend upon them, and the second and third terms of Eq. (10) by simply “pushing a button.”⁶⁶ The first term is an oft-forgotten constant (does not depend on the vibrational state) term^{67,68} which contributes to the vibrational energy. Recent efforts in our laboratory⁶⁹ as well as others^{10,70} have led to the development of explicit formulas for G_0 in terms of quantities calculable from the quartic force field. However, our efforts in this direction have only provided equations that apply to asymmetric tops, although we are working on the required modifications to treat symmetric tops, spherical tops, and polyatomic linear molecules. Because of our inability to calculate G_0 for all molecules, we have chosen to neglect this contribution in the zero-point energies. However, we take some solace from the recent work of Schuurman *et al.*,¹⁰ where this term was included in an exhaustive study of the enthalpies of formation for HNCO and its isomers (which, notably, used a reaction scheme of the sort that we advocate rather than being based strictly on atomization energies). They found that G_0 was typically less than 10 cm^{-1} in magnitude, similar to what we have found in pilot applications,⁶⁹ and below that which is inherent in the calculation of the other two (and numerically more significant) contributions to Eq. (10). Hence, neglect of this constant term is not expected to cause significant errors, although it should be checked in selected cases (cf. H_2O_2 in the following section) and eventually included when general formulas are available. However, for now, the HEAT zero-point vibrational energy contribution will be defined as

$$\Delta E_{\text{ZPE}} = \sum_i \frac{\omega_i}{2} + \sum_{i \geq j} \frac{x_{ij}}{4}. \quad (11)$$

E. Diagonal Born–Oppenheimer correction

It is not always appreciated that the electronic energy, as obtained from traditional electronic structure calculations, is not equivalent to the expectation value of the molecular Hamiltonian over the electronic (clamped-nucleus) wave function. The difference lies in the contribution of the nuclear kinetic energy operator, which can be viewed as a first-order correction to the usual electronic energy. Nonetheless, the simple potential energy surface picture of a molecule is not lost, as each geometry continues to be associated with a specific value of the energy (for a given electronic state) although this surface becomes mass dependent. For atoms, this correction—known as the diagonal Born–Oppenheimer correction (DBOC)—accounts for the finite mass of the nucleus. This diagonal Born–Oppenheimer correction^{31–34,71–73} is calculated by the expectation value

$$\Delta E_{\text{DBOC}} = \langle \Psi_e(\mathbf{r}; \mathbf{R}) | \hat{T}_n | \Psi_e(\mathbf{r}; \mathbf{R}) \rangle, \quad (12)$$

where \hat{T}_n is the nuclear kinetic energy operator and $\Psi_e(\mathbf{r}; \mathbf{R})$ is the normalized electronic wave function obtained at the set of nuclear positions parametrized by \mathbf{R} .

Despite its deceptively simple form, it is not straightforward to calculate the DBOC; efforts in this direction have mostly been (for an exception, see Ref. 72) limited to the HF-SCF or multiconfigurational SCF (MCSCF) level of theory. The landmark paper of Handy, Yamaguchi, and Schaefer³² was the first which reported DBOC energies for a number of polyatomic systems. Recent studies^{71,72} indicate that the use of correlated electronic wave functions has only a modest effect on the DBOC correction. Thus we believe that DBOC corrections calculated with HF-SCF wave functions are sufficiently accurate for the purposes of this study. Therefore, it has been chosen for the HEAT protocol. The aug-cc-pVTZ basis set has been used to calculate ΔE_{DBOC} at the HF-SCF (RHF and ROHF for closed and open-shell systems, respectively) level. This particular choice for the basis set is motivated by a previous study⁷² where the DBOC correction was found to converge relatively rapidly to the one-particle basis limit, provided that diffuse (low-exponent) functions are included in the basis set.

F. Spin–orbit correction

Calculations performed within the framework of a non-relativistic Hamiltonian give a weighted average over the energies of various states involving different coupling of spin and orbital angular momentum. Consideration⁷⁴ of this shortcoming of nonrelativistic theory is necessary for some of the species considered in this work. The relative energies of various states split by the spin–orbit interaction can be calculated from a Hamiltonian that includes the spin–orbit operator. The calculated energy lowering of the lowest spin–orbit state (which is of course the ground state of interest) with respect to the averaged state obtained in a nonrelativistic calculation can be used to adjust the ground state energy.

Since the magnitude of this effect is relatively small for first- and second-row atoms, these calculations need not be performed at levels of theory rivaling those used in the energy extrapolations.

The calculations of ΔE_{SO} (defined here as the energy difference between the ground state and the statistically weighted average of all spin-orbit states—which is not equivalent to the spin-orbit coupling constant!) have been performed with a spin-orbit configuration interaction (CI) procedure. For a detailed description, see Ref. 75. In summary, the core electrons are described by relativistic effective core potentials (RECP) including spin-orbit terms that allow a straightforward calculation of the spin-orbit interaction integrals. The CI wave functions are constructed by considering all single and double excitations out of a valence complete active space reference function. To reduce the computational effort, the double-group symmetry that often facilitates relativistic quantum calculations can be used. The cc-pVDZ basis set developed by Pitzer⁷⁶ together with the corresponding RECPs (Ref. 77) were used in the calculations.

It should be noted that we consider only first-order spin-orbit interactions in this work. Hence, the only molecules for which the ΔE_{SO} contribution is nonzero are radicals in degenerate ground states. Second-order spin-orbit effects, which involve coupling of the ground state with excited states of different spin through the spin-orbit operator, are not included and are expected to be of negligible import for thermochemistry.

One could of course use measured spin-orbit splittings to calculate the correction. It is clearly the contribution appearing in Eq. (4) for which experimental determination is clearly superior to computational estimation. However, since one might want to investigate species where the spin-orbit splitting is not available, we prefer the computational approach. It is also in keeping with the spirit of the HEAT method, in which appeals to experiment and empiricism are kept to a minimum. A quick check reveals that experimentally measured and calculated values of ΔE_{SO} for the molecules in this work differ by less than 10 cm^{-1} in all cases.

G. Scalar relativistic effects

The effect of so-called scalar relativistic contributions²⁹ to the HEAT total energy (ΔE_{REL}) are included by contracting the one-particle density matrix obtained at the CCSD(T)/aug-cc-pCVTZ level with (one-electron) Darwin and mass-velocity terms. As discussed by Davidson, Ishikawa and Malli,⁷⁸ this is a reasonable approximation for relativistic effects when first-row elements are considered. Recently, Boese *et al.*²⁵ compared the sum of Darwin and mass-velocity contributions obtained at high levels of theory to the second-order Douglas-Kroll contributions. While the comparison there is complicated a bit by the lack of a common basis set or method, the results suggest that negligible error is incurred with the simpler first-order treatment for first- and second-row atoms.

H. Computational details

Calculations of the stabilization of the lowest spin-orbit level relative to the weighted average energy were performed with COLUMBUS,⁷⁹ while most of the calculations of ΔE_{DBOC} were carried out with the PSI3.2 electronic structure program suite.⁸⁰ CCSDTQ calculations were done with a string-based many-body code.⁵⁷ All HF-SCF, CCSD, CCSD(T), and CCSDT energy calculations as well as evaluations of ΔE_{REL} and some ΔE_{DBOC} calculations come from a local version of ACES2.⁸¹ Anharmonic force fields used for ΔE_{ZPE} were calculated with ACES2 using the algorithm described in Ref. 82. All calculations, some of which involved more than 500 basis functions, were performed on personal computers running the LINUX operating system.

III. RESULTS

Total HEAT energies, as well as the individual contributions defined in the preceding section, are listed for 31 atoms and molecules in Table I. Species in the test suite contain only hydrogen, carbon, oxygen, nitrogen, or fluorine atoms; there are 12 closed-shell systems and 19 are doublets, triplets, and quartets. The complete documentation of the individual energy contributions should be of use to others interested in expanding the HEAT database, as they can be used to provide a check on calculations. However, the most important numbers in Table I—and indeed in the entire manuscript—are those in the rightmost column: the HEAT total energies for each species. In the following sections, we illustrate use of the HEAT energies in calculating atomization energies and standard enthalpies of formation at 0 K.

A. Atomization energies

Atomization energies calculated from HEAT energies are given in Table II for the 27 molecular species in the test suite, along with the individual contributions. Among other things, the first two columns of the table show the well-known anomalous behavior of fluorine (F_2) and also OF, which are both unbound at the UHF level and owe their stability as molecules to electron correlation. There are also several other species (O_2 , HNO, HO_2 , NO) for which correlation effects account for more than half of the atomization energy; Hartree-Fock and correlation contributions are essentially equal for N_2 , H_2O_2 , and CN. It is not particularly surprising that the magnitude of higher-level correlation contributions—defined here as those beyond CCSD(T)—correlates strongly with the overall correlation contributions identified above. The largest values of ΔE_{CCSDTQ} are found for N_2 , F_2 , O_2 , CO_2 , H_2O_2 , NO, HNO, CN, and HCN, all but HCN mentioned above in the context of having large overall correlation contributions to the atomization energy. For each of these species, the effect of correlation beyond CCSDT—a tract of the quantum chemical landscape that is rarely trod upon in practice—is to increase atomization energies by more than 3 kJ mol^{-1} . Given that the standard definition of chemical accuracy is $4.184 \text{ kJ mol}^{-1}$, one realizes that this time-honored goal of quantum chemistry is still not easily obtained, at least in the context of atomization energies.

TABLE II. Contributions to atomization enthalpies for the molecules in the test set. All values are in kJ mol^{-1} .

| Species | E_{HF}^{∞} | $\Delta E_{\text{CCSD(T)}}^{\infty}$ | ΔE_{CCSDT} | ΔE_{CCSDTQ} | ΔE_{REL} | ΔE_{ZPE} | ΔE_{DBOC} | ΔE_{SO} | Total |
|-------------------------------|--------------------------|--------------------------------------|---------------------------|----------------------------|-------------------------|-------------------------|--------------------------|------------------------|---------|
| N ₂ | 482.94 | 472.26 | -3.14 | 3.61 | -0.56 | -14.12 | 0.08 | 0.00 | 941.07 |
| H ₂ | 350.81 | 107.41 | 0.00 | 0.00 | -0.01 | -25.97 | 0.22 | 0.00 | 432.46 |
| F ₂ | -154.98 | 316.00 | -1.31 | 3.42 | -0.20 | -5.50 | 0.02 | -3.01 | 154.44 |
| O ₂ | 140.36 | 362.56 | -2.07 | 4.46 | -0.82 | -9.56 | 0.06 | -1.61 | 493.39 |
| CO | 729.86 | 356.91 | -2.37 | 2.21 | -0.69 | -12.98 | 0.07 | -1.20 | 1071.82 |
| C ₂ H ₂ | 1228.97 | 468.20 | -3.03 | 2.23 | -1.14 | -68.93 | 0.50 | -0.76 | 1626.06 |
| CCH | 777.34 | 333.02 | 0.56 | 2.28 | -1.12 | -36.34 | 0.23 | -0.76 | 1075.23 |
| CH ₂ | 649.11 | 149.13 | -0.02 | 0.14 | -0.63 | -45.07 | 0.11 | -0.38 | 752.39 |
| CH | 238.28 | 113.57 | 0.43 | 0.12 | -0.18 | -16.97 | -0.34 | -0.20 | 334.70 |
| CH ₃ | 1017.30 | 271.11 | -0.15 | 0.24 | -0.72 | -77.64 | 0.19 | -0.38 | 1209.93 |
| CO ₂ | 1033.08 | 599.66 | -4.49 | 4.11 | -2.05 | -30.40 | 0.20 | -2.02 | 1598.10 |
| H ₂ O ₂ | 562.96 | 562.77 | -2.37 | 2.95 | -1.63 | -68.77 | 0.52 | -1.64 | 1054.81 |
| H ₂ O | 652.40 | 323.02 | -0.97 | 0.99 | -1.14 | -55.73 | 0.53 | -0.82 | 918.26 |
| HCO | 764.53 | 403.66 | -1.78 | 2.19 | -1.16 | -34.03 | -0.10 | -1.20 | 1132.11 |
| HF | 405.62 | 187.00 | -0.67 | 0.72 | -0.84 | -24.66 | 0.34 | -1.51 | 566.01 |
| HO ₂ | 300.93 | 431.57 | -0.60 | 2.52 | -1.20 | -37.05 | 0.02 | -1.64 | 694.56 |
| NO | 225.52 | 411.50 | -1.54 | 3.07 | -0.83 | -11.46 | 0.06 | -0.10 | 626.22 |
| OH | 286.41 | 162.12 | -0.13 | 0.48 | -0.54 | -22.21 | 0.05 | -0.04 | 426.14 |
| HNO | 331.69 | 528.35 | -2.45 | 3.54 | -1.19 | -35.92 | -0.23 | -0.82 | 822.97 |
| CN | 379.38 | 371.64 | 3.14 | 3.43 | -0.63 | -12.75 | 0.10 | -0.38 | 743.92 |
| HCN | 833.50 | 477.62 | -3.21 | 3.04 | -0.94 | -41.74 | 0.32 | -0.38 | 1268.20 |
| CF | 338.46 | 216.63 | -0.60 | 1.00 | -0.72 | -7.88 | 0.03 | -1.42 | 545.50 |
| NH ₂ | 492.92 | 270.76 | 0.06 | 0.48 | -0.67 | -49.57 | -0.03 | 0.00 | 713.95 |
| NH ₃ | 841.38 | 406.03 | -0.63 | 0.72 | -1.08 | -89.45 | 0.56 | 0.00 | 1157.53 |
| NH | 214.88 | 132.40 | 0.29 | 0.22 | -0.31 | -19.46 | -0.19 | 0.00 | 327.83 |
| OF | -64.04 | 282.44 | 1.05 | 2.10 | -0.41 | -6.37 | 0.02 | -1.24 | 213.56 |

The overall difference between the exact correlation contributions (as estimated by HEAT) and those associated with $\Delta E_{\text{CCSD(T)}}^{\infty}$ also includes ΔE_{CCSDT} , which is intended to remedy deficiencies in the CCSD(T) treatment of triple excitation effects. And here, one can only say that CCSD(T) comes through again. While the ΔE_{CCSDTQ} contributions are uniformly positive, those associated with ΔE_{CCSDT} are generally negative, indicating that a HEAT thermochemistry based on CCSD(T) energies instead of estimated exact correlation energies would outperform one based on CCSDT energies.⁸³ When these two contributions are combined [thereby providing an estimate of the difference between CCSD(T) and FCI], excellent cancellation is found in many cases. Exceptions, where the net effect exceeds 1 kJ mol^{-1} , are (kJ mol^{-1} in descending order): CN (6.57), OF (3.15), CCH (2.84), O₂ (2.39), F₂ (2.12), HO₂ (1.92), NO (1.53), and HNO (1.09), most being radicals where the perturbative nature of the (T) correction to the energy is most suspect.⁸⁴ The particularly large value seen for CN also reflects differences in CCSD(T) energies calculated with UHF and ROHF reference functions; it is quite likely that this value would be smaller if the ROHF-based CCSD(T) method^{85,86} was used in determining the $\Delta E_{\text{CCSD(T)}}^{\infty}$ contribution.

Continuing to the right in Table II, we see that scalar relativistic effects systematically reduce the atomization energies by as much as 2 kJ mol^{-1} , corrections to the simple Born–Oppenheimer approximation generally—but not always—increase the atomization energy by no more than 0.5 kJ mol^{-1} . Stabilization of the lowest spin–orbit level can amount to as much as 2 kJ mol^{-1} . Here, it is important to consider the scope of the molecules in the test suite. Scalar relativistic effects will increase with the total number of elec-

trons and the number of bonding electrons (NH, NH₂, and NH₃ form an interesting sequence here, as does OH and H₂O—it seems to indeed be well-approximated by “bond contributions” that can be inferred from the table), and will become more important as one moves down the periodic table. The same is true for spin–orbit effects. The DBOC, on the other hand, is most important for light atoms as can clearly be seen by noting its essentially negligible magnitude for molecules that do not contain hydrogen atoms.

Zero-point vibrational energy contributions to the atomization energies are obviously negative in all cases. Enthalpies of formation calculated from atomization energies and experimental atomic enthalpies of formation will be discussed subsequently.

B. Formation from the elements

Most theoretical model chemistries present atomization energies as the primary thermochemical data (we remind the reader that the total HEAT energies in Table I are the primary data of this paper), but one could also base the calculations on different elemental reference compounds. One choice would be the elements in their standard state, in which case the reaction energies determined would be equivalent to standard enthalpies of formation. However, standard states as defined for many elements are not amenable to computation (carbon, chlorine, boron, to name a few) and this approach has not been followed in the literature of model chemistries. Nevertheless, we will have a go at it here.

Table III lists reaction energies calculated from HEAT data for the formation of the test suite species from molecular reference compounds. For H, N, O, and F, the standard

TABLE III. Contributions to reaction enthalpies for the formation of the test compounds from the elemental reactions, as defined in the text. All values are in kJ mol^{-1} .

| Species | E_{HF}^{∞} | $\Delta E_{\text{CCSD(T)}}^{\infty}$ | ΔE_{CCSDT} | ΔE_{CCSDTQ} | ΔE_{REL} | ΔE_{ZPE} | ΔE_{DBOC} | ΔE_{SO} | Total |
|------------------------|--------------------------|--------------------------------------|---------------------------|----------------------------|-------------------------|-------------------------|--------------------------|------------------------|---------|
| C | 729.86 | 356.91 | -2.37 | 2.21 | -0.69 | -12.98 | 0.07 | -1.20 | 1071.82 |
| F | -77.49 | 158.00 | -0.65 | 1.71 | -0.10 | -2.75 | 0.01 | -1.51 | 77.22 |
| H | 175.41 | 53.71 | 0.00 | 0.00 | -0.01 | -12.99 | 0.11 | 0.00 | 216.23 |
| N | 241.47 | 236.13 | -1.57 | 1.80 | -0.28 | -7.06 | 0.04 | 0.00 | 470.54 |
| O | 70.18 | 181.28 | -1.03 | 2.23 | -0.41 | -4.78 | 0.03 | -0.80 | 246.69 |
| C_2H_2 | 581.57 | 353.03 | -1.70 | 2.19 | -0.26 | 16.99 | -0.14 | -1.64 | 950.04 |
| CCH | 857.79 | 434.51 | -5.30 | 2.15 | -0.27 | -2.61 | 0.01 | -1.64 | 1284.64 |
| CH_2 | 431.57 | 315.19 | -2.34 | 2.08 | -0.08 | 6.12 | 0.18 | -0.82 | 751.89 |
| CH | 666.99 | 297.05 | -2.79 | 2.09 | -0.52 | -9.00 | 0.52 | -1.00 | 953.35 |
| CH_3 | 238.79 | 246.92 | -2.21 | 1.98 | 0.02 | 25.70 | 0.21 | -0.82 | 510.58 |
| CO_2 | -162.86 | 119.81 | 0.05 | 2.56 | 0.54 | 7.86 | -0.08 | -0.79 | -32.90 |
| H_2O_2 | -71.79 | -92.80 | 0.30 | 1.51 | 0.80 | 33.24 | -0.24 | 0.03 | -128.96 |
| H_2O | -231.40 | -34.33 | -0.06 | 1.24 | 0.72 | 24.98 | -0.28 | 0.02 | -239.11 |
| HCO | 210.92 | 188.24 | -1.62 | 2.25 | 0.06 | 3.28 | 0.31 | -0.80 | 402.63 |
| HF | -307.71 | 24.71 | 0.02 | 0.99 | 0.74 | 8.92 | -0.22 | 0.00 | -272.55 |
| HO_2 | 14.83 | -15.30 | -1.47 | 1.94 | 0.38 | 14.50 | 0.14 | 0.03 | 15.05 |
| NO | 86.13 | 5.91 | -1.06 | 0.96 | 0.15 | -0.38 | 0.01 | -0.71 | 91.01 |
| OH | -40.83 | 72.87 | -0.91 | 1.75 | 0.12 | 4.45 | 0.09 | -0.76 | 36.78 |
| HNO | 155.37 | -57.23 | -0.15 | 0.49 | 0.50 | 11.09 | 0.41 | 0.02 | 110.49 |
| CN | 591.96 | 221.40 | -7.07 | 0.59 | -0.34 | -7.29 | 0.01 | -0.82 | 798.44 |
| HCN | 313.24 | 169.13 | -0.72 | 0.98 | -0.04 | 8.71 | -0.09 | -0.82 | 490.38 |
| CF | 313.91 | 298.28 | -2.42 | 2.93 | -0.07 | -7.85 | 0.05 | -1.29 | 603.54 |
| NH_2 | 99.36 | 72.78 | -1.63 | 1.32 | 0.38 | 16.54 | 0.30 | 0.00 | 189.05 |
| NH_3 | -73.69 | -8.78 | -0.94 | 1.08 | 0.79 | 43.43 | -0.19 | 0.00 | -38.30 |
| NH | 202.00 | 157.43 | -1.86 | 1.58 | 0.03 | -0.59 | 0.35 | 0.00 | 358.94 |
| OF | 56.73 | 56.84 | -2.74 | 1.84 | -0.09 | -1.16 | 0.02 | -1.07 | 110.36 |

definition of the corresponding diatomic molecule is used, so that the tabulated reaction energies for molecules containing only these elements are equivalent to the standard enthalpies of formation. For carbon, where the elemental standard state is graphite, we use carbon monoxide in a way most easily communicated by example. For acetylene, the reaction used is



while



is used for the stoichiometrically similar hydrogen peroxide. In the former case, the calculated reaction energy is not equivalent to the enthalpy of formation, differing from it by twice the difference of the enthalpies of formation of CO and O.

All in all, the relative magnitudes of the Hartree-Fock and correlation contributions to the reaction energies is not entirely dissimilar to that found for atomization energies. Correlation corrections dominate in magnitude in a few cases, and others exhibit comparable contributions.

Higher-level correlation contributions are also comparable in magnitude using the atomization and elemental reaction calculations. While the ΔE_{CCSDTQ} contributions are decidedly smaller for the latter (no contributions above 3 kJ mol^{-1} while 8 of the 26 atomization energy contributions exceed this value), the mean absolute post-CCSD(T) correlation contributions (the sum of ΔE_{CCSDT} and ΔE_{CCSDTQ}) are about the same: 0.82 kJ mol^{-1} for the atomization energies and 0.90 kJ mol^{-1} for the elemental reactions.

The remaining contributions are also similar in magnitude in the two cases, except for ΔE_{ZPE} since now there are vibrational modes in all species in the chemical equation.

C. Enthalpies of formation

In Table IV, enthalpies of formation at 0 K ($\Delta_f H_0^\circ$) calculated from HEAT energies are given for all members of the test suite. The calculations were done by two different procedures. In the first (I), HEAT atomization energies were corrected to standard enthalpies of formation according to Eq. (1),⁸⁷ using experimental enthalpies of formation for the atoms. Procedure II is based on the elemental reactions summarized in the preceding section. For all molecules not containing carbon, values of $\Delta_f H^\circ$ obtained by procedure II are equal to the reaction energies given in Table III. This underscores one relative advantage of the elemental reaction approach, specifically that it is based on reference compounds whose enthalpy of formation is precisely zero by definition. This is advantageous only in extremely accurate calculations such as those presented here, since it avoids errors associated with atomic enthalpies of formation for species such as C and F.⁸⁸ Since one cannot do a HEAT calculation on a chunk of graphite, however, an alternative approach was used. The reference for carbon used in approach II was carbon monoxide, as described in the preceding section. CO was chosen since its experimentally determined enthalpy of formation [-113.81 ± 0.17 from Ruscic's Active Thermochemical Tables (ATcT) Refs. 89–91] is not tied to that of the carbon atom⁹² and should therefore be a relatively stable reference. Given the CO value as well as that for O ($\Delta_f H^\circ = 246.84$

TABLE IV. Standard enthalpies of formation (in kJ mol^{-1}), calculated from HEAT atomization energies and experimental atomic enthalpies of formation (I) and the elemental reaction approach (II). Experimental values are from the ATcT (Ref. 89) (superscript a) where available. Remaining values come from the NIST-JANAF compendium. Fixed values are in parenthesis.

| Species | $\Delta_f H^\circ(\text{I})$ | $\Delta_f H^\circ(\text{II})$ | Expt. |
|----------------------------------|------------------------------|-------------------------------|---------------------------|
| N ₂ | 0.11 | 0.00 | 0.00±0.00 |
| H ₂ | -0.39 | 0.00 | 0.00±0.00 |
| F ₂ | -0.02 | 0.00 | 0.00±0.00 |
| O ₂ | 0.3(0) | 0.00 | 0.00±0.00 |
| C | (711.79) | 711.17 | 711.79±0.21 ^a |
| F | (77.21) | 77.22 | 77.21±0.24 ^a |
| H | (216.03) | 216.23 | 216.03±0.00 ^a |
| N | (470.59) | 470.54 | 470.59±0.05 ^a |
| O | (246.84) | 246.69 | 246.84±0.00 ^a |
| CO | (-113.18) | -113.81 | -113.81±0.17 ^a |
| C ₂ H ₂ | 229.59 | 228.74 | 228.20±0.64 ^a |
| CCH | 564.39 | 563.34 | 563.32±0.65 ^a |
| CH ₂ | 391.47 | 391.24 | 390.65±0.54 ^a |
| CH | 593.12 | 592.70 | 593.19±0.36 ^a |
| CH ₃ | 149.96 | 149.93 | 149.94±0.11 ^a |
| CO ₂ | -392.63 | -393.55 | -393.11±0.01 ^a |
| H ₂ O ₂ | -129.05 | -128.96 | -129.82±0.08 ^a |
| H ₂ O | -239.35 | -239.11 | -238.92±0.04 ^a |
| HCO | 42.55 | 41.98 | 42.09±0.38 ^a |
| HF | -272.76 | -272.55 | -272.73±0.24 ^a |
| HO ₂ | 15.16 | 15.05 | 14.96±0.64 ^a |
| NO | 91.22 | 91.01 | 90.53±0.09 ^a |
| OH | 36.74 | 36.78 | 37.09±0.05 ^a |
| HNO | 110.50 | 110.49 | 102.50±0.42 ^b |
| CN | 438.47 | 437.79 | 436.80±10.00 ^b |
| HCN | 130.21 | 129.73 | 135.53±8.40 ^b |
| CF | 243.50 | 242.89 | 251.60±8.00 ^b |
| NH ₂ | 188.71 | 189.05 | 193.25±6.30 ^b |
| NH ₃ | -38.84 | -38.30 | -38.91±0.40 ^b |
| NH | 358.80 | 358.94 | 376.51±16.70 ^b |
| OF | 110.50 | 110.36 | 108.00±10.00 ^b |
| Mean absolute error ^c | 0.37 | 0.24 | ... |
| Mean signed error ^c | -0.28 | -0.06 | ... |
| RMS error ^c | 0.56 | 0.35 | ... |
| Maximum error ^c | 1.39 | 0.86 | ... |

^aFrom active thermochemical tables.

^bFrom NIST-JANAF compilation.

^cBased only on active thermochemical tables data.

±0.002), elemental reaction energies—as defined earlier—need to be adjusted by $360.65 \text{ kJ mol}^{-1}$ per carbon atom to give enthalpies of formation.

The results are startling. For compounds that have well-established and self-consistent enthalpies of formation, as determined by Ruscic's ATcT approach,^{89,91} the HEAT values determined by method II are within 1 kJ mol^{-1} in all cases! Indeed, for only one example—H₂O₂ (-128.96 versus $-129.82 \pm 0.08 \text{ kJ mol}^{-1}$)—does the calculated HEAT value fall more than 0.5 kJ mol^{-1} outside the range estimated by the ATcT approach. Using method I, both C₂H₂ and H₂O₂ fall more than 0.5 kJ mol^{-1} outside the ATcT estimate. For set I, 7 of the 16 HEAT values are within the ATcT error bars, while 8 set II values fall within the estimated bounds. The statistical analysis shown at the bottom of Table IV gives further support that enthalpies of formation determined by the atomization energy approach (I) are not quite as good, although the performance difference is decidedly small.

The astonishing accuracy of HEAT energies in predicting the enthalpies of formation for the well-characterized compounds in Table IV suggests that they are “better” estimates than those found in the NIST-JANAF database for all of the remaining compounds (except perhaps ammonia), in the sense that the error bar associated with the HEAT values (which we assign as 1 kJ mol^{-1} for all but difficult cases) is roughly an order of magnitude smaller. An interesting case that warrants further study is the HNO molecule, where the NIST-JANAF uncertainty is rather small (0.42 kJ mol^{-1}), while the HEAT value is about 8 kJ mol^{-1} higher.

The compounds that appear to be the most difficult cases for HEAT are C₂H₂ and H₂O₂. Calculated enthalpies of formation (0 K) for both are well outside (relatively speaking) the range of values from ATcT, which merits some discussion. For C₂H₂, the error appears to result from a neglect of *connected pentuple excitations* (see appendix), while the discrepancy for H₂O₂ is unclear. It is perhaps significant that H₂O₂ stands alone amongst the test suite molecules in having a low-frequency torsional mode. The way in which ΔE_{ZPE} is calculated in the HEAT protocol is ideally suited for semirigid molecules where both the harmonic approximation and second-order vibrational perturbation theory⁶³ are expected to work well. It may well break down for torsional modes such as that in H₂O₂, at least at the level of tenths of a kJ mol^{-1} . A detailed analysis of H₂O₂ is underway, and will be reported in a separate publication.⁹³

IV. DISCUSSION

As continually stated throughout this paper, the total HEAT energies given in Table I represent the principal (and indeed, the only nonredundant) results of this research. It is hoped that the 31 examples studied here are just the beginning of a fairly significant database of systems for which HEAT results are available. To this end, we have already begun calculations on a few additional systems, and invite other members of the quantum chemical community to join the effort. Although the results presented in the preceding section for the enthalpies of formation are already extremely accurate (more so, in fact, than we had anticipated when this project began), even greater accuracy can be obtained if the theoretical results are used to calculate enthalpy changes for reactions in which the bonding environments of atoms on the left and right side of the chemical equation are similar. There is essentially a continuum between atomization energies (which are undoubtedly the most difficult quantities to calculate accurately) and isodesmic reactions in which the local environment of every atom is preserved on both sides of the equation. In isodesmic and nearly isodesmic reactions, one expects correlation effects to be less severe. Hence, the ΔE_{CCSDT} and ΔE_{CCSDTQ} contributions to the HEAT energy should be considerably smaller in magnitude. This is advantageous, because unlike E_{HF}^∞ and $\Delta E_{\text{CCSD(T)}}^\infty$, these higher-order correlation contributions are not extrapolated as thoroughly nor are they free of the assumption of negligible core correlation effects.

Some evidence that this is the case can be found in Table V, where HEAT thermochemistry values are given for sev-

TABLE V. Individual contributions from HEAT energies to reaction enthalpies at 0 K. All values in kJ mol^{-1} .

| Reaction | E_{HF}^{∞} | $\Delta E_{\text{CCSD(T)}}^{\infty}$ | ΔE_{CCSDT} | ΔE_{CCSDTQ} | ΔE_{REL} | ΔE_{ZPE} | ΔE_{DBOC} | ΔE_{SO} | Total |
|--|--------------------------|--------------------------------------|---------------------------|----------------------------|-------------------------|-------------------------|--------------------------|------------------------|---------|
| $\text{H} + \text{H}_2\text{O}_2 \rightarrow \text{HO}_2 + \text{H}_2\text{O}$ | -103.95 | -29.70 | -0.92 | -0.08 | 0.17 | 1.80 | 0.02 | 0.78 | -131.88 |
| $\text{CH} + \text{C}_2\text{H}_2 \rightarrow \text{CCH} + \text{CH}_2$ | 40.80 | 99.62 | -3.15 | -0.06 | 0.43 | -4.48 | -0.19 | 0.18 | 133.14 |
| $\text{CH}_2 + \text{C}_2\text{H}_2 \rightarrow \text{CCH} + \text{CH}_3$ | 83.44 | 13.20 | -3.47 | -0.14 | 0.08 | -0.01 | 0.19 | 0.00 | 93.29 |
| $\text{HCO} + \text{C} \rightarrow \text{CH} + \text{CO}$ | -203.62 | -66.82 | 0.16 | -0.14 | -0.29 | -4.07 | 0.18 | 0.20 | -274.40 |
| $\text{N} + \text{NH}_3 \rightarrow \text{NH}_2 + \text{NH}$ | 133.58 | 2.86 | -0.98 | 0.02 | -0.10 | -20.41 | 0.79 | 0.00 | 115.75 |
| $2\text{NH} \rightarrow \text{N} + \text{NH}_2$ | -63.16 | -5.96 | 0.51 | -0.03 | 0.04 | 10.65 | -0.36 | 0.01 | -58.29 |
| $\text{HCN} + \text{C} \rightarrow \text{CH} + \text{CN}$ | 215.84 | -7.59 | -6.77 | -0.51 | -0.13 | -12.02 | 0.56 | 0.20 | 189.59 |
| $\text{CH} \rightarrow \text{C} + \text{H}$ | 238.28 | 113.57 | 0.43 | 0.12 | -0.18 | -16.97 | -0.34 | -0.20 | 334.70 |
| $\text{CH}_2 \rightarrow \text{CH} + \text{H}$ | 410.83 | 35.56 | -0.45 | 0.02 | -0.45 | -28.10 | 0.46 | -0.18 | 417.69 |
| $\text{CH}_3 \rightarrow \text{CH}_2 + \text{H}$ | 368.19 | 121.98 | -0.13 | 0.10 | -0.10 | -32.57 | 0.08 | 0.00 | 457.54 |
| $\text{NH} \rightarrow \text{N} + \text{H}$ | 214.88 | 132.40 | 0.29 | 0.22 | -0.31 | -19.46 | -0.19 | 0.00 | 327.83 |
| $\text{NH}_2 \rightarrow \text{NH} + \text{H}$ | 278.04 | 138.36 | -0.23 | 0.25 | -0.36 | -30.11 | 0.16 | 0.00 | 386.12 |
| $\text{NH}_3 \rightarrow \text{NH}_2 + \text{H}$ | 348.46 | 135.26 | -0.69 | 0.24 | -0.41 | -39.87 | 0.60 | 0.00 | 443.58 |
| $\text{OH} \rightarrow \text{O} + \text{H}$ | 286.41 | 162.12 | -0.13 | 0.48 | -0.54 | -22.21 | 0.05 | -0.04 | 426.14 |
| $\text{H}_2\text{O} \rightarrow \text{OH} + \text{H}$ | 365.98 | 160.90 | -0.85 | 0.51 | -0.60 | -33.52 | 0.48 | -0.78 | 492.12 |
| $\text{HF} \rightarrow \text{F} + \text{H}$ | 405.62 | 187.00 | -0.67 | 0.72 | -0.84 | -24.66 | 0.34 | -1.51 | 566.01 |

eral chemical reactions. The first group of reactions are isodesmic or nearly so; the types (partners and bond order) of chemical bonds are preserved in the reactions, although free atoms appear in the last four reactions and the seventh reaction involves a dramatic difference in the location of an unpaired electron. Nonetheless, the ΔE_{CCSDTQ} corrections are less than 1 kJ mol^{-1} in all cases, in contrast to the atomization and elemental reactions documented in Tables II and III, where $\Delta E_{\text{CCSDTQ}} > 1 \text{ kJ mol}^{-1}$ is the norm. Hence, enthalpies calculated from HEAT energies that exclude the ΔE_{CCSDTQ} contribution which is—at least for the larger species—the most expensive calculation met with in the HEAT protocol, should be considerably more accurate than atomization or elemental formation reactions based on the same approximations. We note in passing that some of the ΔE_{CCSDT} contributions are large for these reactions, specifically those involving the isoelectronic CCH and CN radicals for which UHF-CCSD(T) performs poorly due to spin-contamination effects.⁹⁴

Also interesting is the second set of reactions in Table V, all of which involve breaking of an X-H bond, where X is a first-row element. In all of these, there is just one difference in the bonding of products and reactants (there is an extra X-H bond in the former), and these should benefit from cancellation of errors. And indeed, higher-order correlation effects are quite small, with both ΔE_{CCSDT} and ΔE_{CCSDTQ} always below 1 kJ mol^{-1} . Even beyond this, there is a striking regularity in the ΔE_{CCSDTQ} contributions to these reaction energies. The magnitude of this contribution is always positive, and increases systematically with the electronegativity of X. Alternatively, it increases with decreasing X-H distance. Moreover, it is remarkably constant for the sequences $\text{XH} \rightarrow \text{X} + \text{H}$, $\text{XH}_2 \rightarrow \text{XH} + \text{H}$, etc. This can be used advantageously in HEAT studies in which this expensive contribution is excluded. For example, if one was interested in determining the enthalpy of formation of an alkoxy radical from the corresponding alcohol, it would seem pragmatic to skip the CCSDTQ calculation and simply assume that the difference in ΔE_{CCSDTQ} energies of the alcohol and alkoxy radical is 0.5 kJ mol^{-1} . The behavior of ΔE_{CCSDT} is considerably

less systematic, but this is less of a problem since it is comparatively cheap to calculate.

An example of how we feel the HEAT energies are best used is provided by the NH and NH_2 molecules. Uncertainties in NIST-JANAF values for $\Delta_f H^\circ$ for these molecules are rather large (16.7 and 6.3 kJ mol^{-1} , respectively). While the values in Table IV are likely to be within 1 kJ mol^{-1} , there is another approach which is potentially superior. The calculated HEAT reaction energies for $\text{NH}_3 \rightarrow \text{NH}_2 + \text{H}$, $\text{NH}_2 \rightarrow \text{NH} + \text{H}$ and $\text{NH} \rightarrow \text{N} + \text{H}$ in Table V are surely more precise than enthalpies of formation calculated from the relatively difficult atomization energies or elemental formation energies in Table III. These can be combined with the rather precisely known $\Delta_f H^\circ$ value of H and NH_3 (216.03 and $-38.91 \pm 0.4 \text{ kJ mol}^{-1}$) to yield values of $188.64 \text{ kJ mol}^{-1}$ and $358.73 \text{ kJ mol}^{-1}$ for $\Delta_f H^\circ$ of NH_2 and NH , respectively. It should be noted that both of these numbers are also in good agreement with those in Table IV.

Several follow-up studies are in order. First of all, we recognize the computationally demanding nature of the HEAT procedure. The calculations of $\Delta E_{\text{CCSD(T)}}^{\infty}$ and ΔE_{CCSDTQ} , in particular, are the most arduous. Simplification of these steps would certainly result in a theoretical model chemistry that would be more widely applicable. But just how accurate would it be? For $\Delta E_{\text{CCSD(T)}}^{\infty}$, there are some alternatives that come to mind. First of all, the extrapolated energy could be based on aug-cc-pCVTZ and aug-cc-pCVQZ calculations, which would avoid the aug-cc-pCV5Z calculations that involve, for example, 543 basis functions for a molecule as small as CO_2 . Another course of action would be to invoke separation of core and valence correlation. One could even dispense with diffuse functions,⁹⁵ and calculate extrapolated CCSD(T) energies with the cc-pVQZ and cc-pV5Z (or cc-pVTZ and cc-pVQZ) basis sets in dropped core calculations. Effects of core correlation could then be estimated from relatively small basis sets (cc-pCVDZ or cc-pCVTZ) in all electron calculations. These values, when combined, would represent an approximation to $\Delta E_{\text{CCSD(T)}}^{\infty}$. Ultimately, basis set extrapolations are em-

pirical in nature. Explicitly correlated coupled-cluster methods such as the linear R12 methods⁹⁶ could be used to evaluate $\Delta E_{\text{CCSD(T)}}^{\infty}$ more accurately and reliably than by the extrapolation-based schemes used here. For ΔE_{CCSDTQ} , estimation of quadruple effects by a noniterative approximation would be desirable. Although there is no shortage of such noniterative approaches that have been suggested in the literature,⁹⁷ little has been done in the way of testing them for nontrivial systems. Recent work by some of us⁹⁸ has led to the derivation and implementation of a noniterative correction to CCSDT that corresponds—in the sense of perturbation theory—to the (T) correction to CCSD. It is also distinct from any other approach that has been advocated for correcting the CCSDT energy. One of the potential application areas for this “CCSDT(Q)” method will be in the area of thermochemistry. Alternatively, it is also sensible to dispense completely with the quadruples correction *if thermochemical parameters are calculated from (nearly) isodesmic reactions*, especially if accuracies better than 1 kJ mol⁻¹ are not required. Other areas worthy of study are: correlation effects in the diagonal Born–Oppenheimer approximation^{71,72} (where we have also recently focused some theoretical effort⁹⁹), the importance of anharmonicity in the zero-point vibrational correction, and the importance of the thus far ignored constant G_0 contribution to ΔE_{ZPE} . Also of interest would be the use of CC calculations based on ROHF reference functions. Although we believe that differences in the final energies will be negligible, the partitioning of $\Delta E_{\text{CCSD(T)}}^{\infty}$, ΔE_{CCSDT} , and ΔE_{CCSDTQ} will undoubtedly change in interesting ways, which might render more systematic behavior in the individual contributions.

Studies along the lines suggested above will be carried out. The results, however, are predictable. Use of simplified HEAT strategies will likely result in a level of accuracy that is comparable to that obtained in other sophisticated theoretical model chemistries such as G3, CBS-Q, W2, and W3, which have been worked out, benchmarked, and documented in excellent detail in the literature. We are not interested in proliferating the population of roughly equivalent methods, so the community will not be subjected to follow-up papers reporting HEAT2, HEAT3, etc., methods. However, investigation of the approximations suggested above should be interesting, especially with regard to the utility of the noniterative approximation for quadruple excitation effects and for any light that it might shed on the magnitude and systematics of the error introduced by assuming that core and valence correlation effects can be separated.

ACKNOWLEDGMENTS

This work grew from a discussion involving three of the authors (A.G.C., P.G.S., and J.F.S.) in Budapest during a meeting of the IUPAC task force on free radical thermochemistry. We therefore thank IUPAC for providing the stimulus for this excursion into computational thermochemistry. Financial support for this work comes from the U.S. National Science Foundation and the Robert A. Welch Foundation (J.V., B.A.F., and J.F.S.), the Fonds der Chemischen Industrie (M.K. and J.G.), the Hungarian Research Founda-

tion under OTKA Grant No. T047182 (P.G.S. and M.K.), and OTKA Grant No. T047185 (A.G.C.). During the latter stages of this research, P.G.S. was supported by the Fulbright Foundation during a sabbatical at the University of Texas at Austin. The research presented in this paper is part of a current and future work by a Task Group of the International Union of Pure and Applied Chemistry Grant No. (2000-013-2-100) to determine structures, vibrational frequencies, and thermodynamic functions of free radicals of importance in atmospheric chemistry. We also thank Branko Ruscic (Argonne National Laboratory) for his enthusiastic support of our research as well as guidance in the subtle world of high-accuracy thermochemistry.

APPENDIX: COMMENTS ON CALCULATIONS USING CCSDTQ AND BEYOND

We begin with a few comments on the cost of these calculations. The largest CCSDTQ calculation, which was carried out for the HCO radical using the cc-pVTZ basis set, involves the simultaneous solution of roughly 622×10^6 nonlinear equations. The most demanding CCSDTQP (which includes complete treatment of single, double, triple, quadruple, and pentuple excitations) calculation was carried out for H₂O₂, where the dimension of the nonlinear system was slightly greater than 1.3×10^9 . These calculations required several weeks of computer time, and are clearly beyond the scope of anything that could remotely be characterized as a routine application. Despite this, none of the CCSDTQ calculations with the cc-pVDZ basis set (those specified in the HEAT protocol) required more than three days on LINUX-based personal computers.

To test the suitability of the approximation used for ΔE_{CCSDTQ} , additional CCSDTQ calculations were performed for a subset of molecules (the smaller ones, of course) using cc-pVTZ and cc-pVQZ basis sets, and at the CCSDTQP level with the cc-pVDZ basis. The effects of these on values of $\Delta_f H^\circ$ calculated according to method II (elemental reactions) are documented in Table VI. From this, it can be seen that the cc-pVDZ basis set (the only member of the cc-pVXZ hierarchy that can really be used for ΔE_{CCSDTQ} for all of the molecules) gives results that are quite suitable for the present purposes. In fact, from an inspection of Table VII, it appears that the cc-pVDZ basis set is superior to the cc-pVTZ basis set, since it gives CCSDTQ - CCSDT differences that are closer to those based on cc-pVTZ/cc-pVQZ extrapolations for systems small enough to permit CCSDTQ/cc-pVQZ calculations. These calculations also indicate (see Table VI) that the cc-pVDZ contribution to the elemental reaction enthalpies are not very sensitive to further extension of the basis. The fact that the magnitude of the cc-pVDZ CCSDTQ - CCSDT energies appears to be slightly too small with respect to cc-pVTZ, cc-pVQZ, and extrapolated results was in fact the basis for an empirical scaling of quadruple excitation effects by Boese *et al.* in their W3 model.²⁵ However, we have not chosen to scale our calculated quadruple corrections.

Of perhaps greater interest is the effect of connected pentuple excitations. While CCSDTQP calculations are clearly at least a few decades from becoming routinely ap-

TABLE VI. Contribution of quadruple and pentuple excitations to enthalpies of formation in different basis sets. All values are in kJ mol^{-1} .

| Species | CCSDTQ-CCSDT | | CCSDTQP-CCSDTQ |
|-------------------------------|--------------|---------|----------------|
| | cc-pVDZ | cc-pVTZ | |
| C ₂ H ₂ | -2.27 | -2.52 | -0.54 |
| C | -0.01 | 0.08 | -0.11 |
| CCH | -2.31 | -2.62 | -0.54 |
| CF | 0.70 | 0.83 | 0.04 |
| CH ₂ | -0.15 | -0.02 | -0.12 |
| CH | -0.14 | -0.05 | -0.11 |
| CH ₃ | -0.25 | -0.12 | -0.12 |
| CN | -1.64 | -2.03 | -0.40 |
| CO ₂ | 0.33 | | |
| F | 1.71 | 1.66 | 0.08 |
| H ₂ O ₂ | 1.51 | | 0.28 |
| H ₂ O | 1.24 | 1.52 | 0.21 |
| HCN | -1.25 | -1.40 | -0.29 |
| HCO | 0.02 | 0.04 | -0.02 |
| HF | 0.99 | 1.27 | 0.07 |
| HNO | 0.49 | | 0.13 |
| HO ₂ | 1.94 | | 0.27 |
| N | 1.80 | 1.94 | 0.24 |
| NH ₂ | 1.32 | 1.54 | 0.21 |
| NH ₃ | 1.08 | 1.30 | 0.19 |
| NH | 1.58 | 1.77 | 0.23 |
| NO | 0.96 | 0.95 | 0.12 |
| O | 2.23 | 2.22 | 0.24 |
| OF | 1.84 | 1.96 | 0.18 |
| OH | 1.75 | 1.92 | 0.23 |

plicable, their importance in accurate thermochemistry is apparently nonvanishing. With the cc-pVDZ basis set, it is interesting to note that the magnitude of the largest contributions of pentuples to enthalpies of formation calculated by procedure II (the elemental reaction approach) are roughly 0.5 kJ mol^{-1} . These large contributions found for acetylene, CCH, and CN can be applied for “superHEATed” reaction enthalpies of these species. Adding these contribution to those found in Table IV, one finds $\Delta_f H^\circ(\text{C}_2\text{H}_2) = 228.20 \text{ kJ mol}^{-1}$ versus the ATcT value of $228.20 \pm 0.64 \text{ kJ mol}^{-1}$ (!); $\Delta_f H^\circ(\text{CCH}) = 562.78 \text{ kJ mol}^{-1}$ versus the ATcT value of $563.32 \pm 0.65 \text{ kJ mol}^{-1}$; and $\Delta_f H^\circ(\text{CN}) = 437.39 \text{ kJ mol}^{-1}$ versus the NIST-JANAF value of

TABLE VII. Difference of CCSDTQ and CCSDT total energies ($E_{\text{CCSDTQ}} - E_{\text{CCSDT}}$) in different basis sets and with basis set extrapolation (μE_h).

| | | | | DT | TQ |
|--------------------|---------|---------|---------|---------------|---------------|
| | cc-pVDZ | cc-pVTZ | cc-pVQZ | extrapolation | extrapolation |
| C | -30 | -43 | -44 | -49 | -45 |
| F | -115 | -60 | -90 | -37 | -112 |
| N | -42 | -43 | -57 | -43 | -67 |
| O | -78 | -51 | -75 | -40 | -92 |
| CH | -76 | -93 | -105 | -99 | -114 |
| NH | -127 | -107 | -129 | -98 | -145 |
| OH | -259 | -167 | -206 | -128 | -234 |
| HF | -392 | -209 | -258 | -132 | -294 |
| CH ₂ | -82 | -80 | -96 | -80 | -108 |
| NH ₂ | -224 | -196 | -230 | -185 | -254 |
| Error ^a | 29 | 42 | 18 | 58 | ... |

^aMean absolute error with respect to TQ extrapolated values.

$436.8 \pm 10 \text{ kJ mol}^{-1}$. Of some interest here, however, is Ruscic's recommendation⁹² of $\Delta_f H^\circ(\text{CN}) = 438.5 \pm 4.0 \text{ kJ mol}^{-1}$, based on relatively recent experiments.¹⁰⁰ Hence, it appears that pentuple excitation effects are responsible for the relatively poor performance of HEAT for acetylene, which is one of the two prominent outliers seen in Table IV (the other being H_2O_2).¹⁰¹ Fortunately, there is ample evidence that effects of hexuple and higher excitations can be safely ignored. Some recent work^{57,58,60,102} has indicated that contributions to molecular properties from connected excitations decrease by an order of magnitude upon each increase in the excitation level. This, together with the results documented in this appendix, suggests that neglect of higher than pentuple excitations results in errors of $<0.05 \text{ kJ mol}^{-1}$ (which, after all, is less than 5 cm^{-1}), at least for states that are relative free of strong nondynamic correlation effects. Examples such as ozone would be an interesting test, but unfortunately are too large to be calculated beyond CCSDTQ even with the cc-pVDZ basis set.

¹ See, for example, the special issues of *Spectrochim. Acta* **58**, 599–898 (2002).

² For a representative and chemically relevant example, see E. R. Jochowitz, M. R. Nimlos, M. E. Varner, G. B. Ellison, and J. F. Stanton, *J. Am. Chem. Soc.* (to be published).

³ K. L. Bak, J. Gauss, P. Jørgensen, J. Olsen, T. Helgaker, and J. F. Stanton, *J. Chem. Phys.* **114**, 6548 (2001).

⁴ For a review, see S. J. Blanksby and G. B. Ellison, *Acc. Chem. Res.* **36**, 255 (2003).

⁵ B. Ruscic, J. E. Boggs, A. Burcat *et al.*, *J. Phys. Chem. Ref. Data* (in press).

⁶ Well-known tabulations of molecular enthalpies of formation include: M. W. Chase, Jr., in *NIST-JANAF Thermochemical Tables*, 4th ed. Journal of Physical Chemical Reference Data, Monograph Vol. 9 (1998), p. 1; J. D. Cox, D. D. Wagman, and V. A. Medvedev, *CODATA Key Values for Thermodynamics* (Hemisphere Publishing Corp., New York, 1989); L. V. Gurvich, I. V. Veyts, and C. B. Alcock, *Thermodynamic Properties of Individual Substances*, 4th ed. (Hemisphere Publishing Company, New York, 1989).

⁷ B. A. Flowers, P. G. Szalay, J. F. Stanton, M. Kállay, J. Gauss, and A. G. Császár, *J. Phys. Chem. A* **108**, 3195 (2004), and references therein.

⁸ A. G. Császár, P. G. Szalay, and M. L. Leininger, *Mol. Phys.* **100**, 3879 (2002).

⁹ A. G. Császár, M. L. Leininger, and V. Szalay, *J. Chem. Phys.* **118**, 10631 (2003).

¹⁰ M. S. Schuurman, S. R. Muir, W. D. Allen, and H. F. Schaefer, *J. Chem. Phys.* **120**, 11586 (2004).

¹¹ S. Parthiban and J. M. L. Martin, *J. Chem. Phys.* **114**, 6014 (2001).

¹² Before the development of the Active Thermochemical Tables, atomic enthalpies of formation for O and N were also not known precisely. However, uncertainties in these values have been dramatically reduced in the ATcT approach, B. Ruscic (private communication).

¹³ B. Ruscic, A. F. Wagner, L. B. Harding *et al.*, *J. Phys. Chem. A* **106**, 2727 (2002).

¹⁴ A. L. L. East and W. D. Allen, *J. Chem. Phys.* **99**, 4638 (1993).

¹⁵ D. A. Dixon, D. Feller, and G. Sandrone, *J. Phys. Chem. A* **103**, 4744 (1999).

¹⁶ C. W. Bauschlicher, J. M. L. Martin, and P. R. Taylor, *J. Phys. Chem. A* **103**, 7715 (1999).

¹⁷ J. M. L. Martin and G. de Oliveira, *J. Chem. Phys.* **111**, 1843 (1999).

¹⁸ T. Helgaker, W. Klopper, A. Halkier, K. L. Bak, P. Jørgensen, and J. Olsen, *Quantum Mechanical Prediction of Thermochemical Data* (American Chemical Society, Washington DC, 1998).

¹⁹ K. L. Bak, P. Jørgensen, J. Olsen, T. Helgaker, and W. Klopper, *J. Chem. Phys.* **112**, 9229 (2000).

²⁰ *Quantum Mechanical Prediction of Thermochemical Data*, edited J. Cioslowski (Kluwer, Dordrecht, 2001).

²¹ W. J. Hehre, L. Radom, P. V. R. Schleyer, and J. A. Pople, *Molecular Orbital Theory* (Wiley, New York, 1986), p. 298.

²² The reader is reminded that isodesmic reactions are those which would be

- thermoneutral in the limit that additivity of "bond energies" gave a precise picture of chemical thermodynamics.
- ²³ Notable in this regard are the G(aussian) n ($n = 1, 2$, and 3) family of methods developed by Pople and co-workers [K. Raghavachari and L. A. Curtiss in Ref. 20]; the W(eizmann) n ($n = 1, 2$, and 3) methods of Martin and co-workers [Ref. 17 (W1); Ref. 11 (W2); Ref. 25 (W3)]; and the CBS methods of Petersson [G. A. Petersson, in Ref. 20]. A related approach is the focal-point strategy advocated by Allen and collaborators (Ref. 24); see Ref. 10 for a recent example.
- ²⁴ A. G. Császár, W. D. Allen, and H. F. Schaefer III, *J. Chem. Phys.* **1998**, 108 (9751).
- ²⁵ A. D. Boese, M. Oren, O. Atasolyu, J. M. L. Martin, M. Kállay, and J. Gauss, *J. Chem. Phys.* **120**, 4129 (2004).
- ²⁶ R. J. Bartlett, *Annu. Rev. Phys. Chem.* **32**, 359 (1981).
- ²⁷ An exception is our calculation of spin-orbit coupling, which uses a non-size-extensive approach.
- ²⁸ Throughout the text, the term "enthalpy of formation" refers to the difference in ground-state energies, i.e., the enthalpy of formation at 0 K.
- ²⁹ R. D. Cowan and M. Griffin, *J. Opt. Soc. Am.* **66**, 1010 (1976); R. L. Martin, *J. Phys. Chem.* **87**, 750 (1983).
- ³⁰ As is customary in the quantum chemistry world, the second row of the periodic table is taken to be that containing Li, Be, B, C, N, O, F, and Ne.
- ³¹ H. Sellers and P. Pulay, *Chem. Phys. Lett.* **103**, 463 (1984).
- ³² N. C. Handy, Y. Yamaguchi, and H. F. Schaefer, *J. Chem. Phys.* **84**, 4481 (1986).
- ³³ N. C. Handy and A. M. Lee, *Chem. Phys. Lett.* **252**, 425 (1996).
- ³⁴ W. Kutzelnigg, *Mol. Phys.* **90**, 909 (1997).
- ³⁵ It should be noted that inclusion of these effects shifts the position of the equilibrium internuclear distance. Nonetheless, we ignore this complication, and base all of our calculations on equilibrium geometries determined with CCSD(T)/cc-pVQZ. The issue of equilibrium geometry is not of numerical significance. In the vicinity of the true minimum, the energy changes only in second order when the nuclei are displaced. Taking a typical (and large) value for a bond stretching force constant ($10 \text{ aJ } \text{\AA}^{-2}$), an error of 0.005 \AA in the equilibrium bond length translates to a calculated energy error of 0.08 kJ mol^{-1} .
- ³⁶ G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- ³⁷ K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- ³⁸ J. Noga and R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987); G. E. Scuseria and H. F. Schaefer, *Chem. Phys. Lett.* **132**, 382 (1988); J. D. Watts and R. J. Bartlett, *J. Chem. Phys.* **93**, 6104 (1993).
- ³⁹ S. A. Kucharski and R. J. Bartlett, *Theor. Chim. Acta* **80**, 387 (1991); *J. Chem. Phys.* **97**, 4282 (1992); N. Oliphant and L. Adamowicz, *ibid.* **94**, 1229 (1991).
- ⁴⁰ Optimized all-electron CCSD(T)/cc-pVQZ geometries used in all calculations: $\text{C}_2\text{H}_2-r_{\text{CC}}=1.204 \text{ \AA}$, $r_{\text{CH}}=1.061 \text{ \AA}$; $\text{C}_2\text{H}-r_{\text{CC}}=1.205 \text{ \AA}$, $r_{\text{CH}}=1.061 \text{ \AA}$; $\text{CF}-r_{\text{CF}}=1.270 \text{ \AA}$; $\text{CH}_2-r_{\text{CH}}=1.074 \text{ \AA}$, $\theta_{\text{HCH}}=133.73^\circ$; $\text{CH}_3-r_{\text{CH}}=1.075 \text{ \AA}$; $\text{CH}-r_{\text{CH}}=1.116 \text{ \AA}$; $\text{CN}-r_{\text{CN}}=1.167 \text{ \AA}$; $\text{CO}-r_{\text{CO}}=1.128 \text{ \AA}$; $\text{CO}_2-r_{\text{CO}}=1.160 \text{ \AA}$; $\text{F}_2-r_{\text{FF}}=1.411 \text{ \AA}$; $\text{H}_2-r_{\text{HH}}=0.741 \text{ \AA}$; $\text{H}_2\text{O}-r_{\text{OH}}=0.956 \text{ \AA}$, $\theta_{\text{HOH}}=104.25^\circ$; $\text{H}_2\text{O}_2-r_{\text{OH}}=0.960 \text{ \AA}$, $r_{\text{OO}}=1.449 \text{ \AA}$, $\theta_{\text{HOO}}=100.00^\circ$, $\phi_{\text{HOOH}}=112.44^\circ$; $\text{HCN}-r_{\text{CN}}=1.153 \text{ \AA}$, $r_{\text{CH}}=1.064 \text{ \AA}$; $\text{HCO}-r_{\text{CO}}=1.175 \text{ \AA}$, $r_{\text{CH}}=1.116 \text{ \AA}$, $\theta_{\text{HCO}}=124.57^\circ$; $\text{HF}-r_{\text{HF}}=0.915 \text{ \AA}$; $\text{HNO}-r_{\text{NO}}=1.208 \text{ \AA}$, $r_{\text{NH}}=1.050 \text{ \AA}$, $\theta_{\text{HNO}}=108.09^\circ$; $\text{HO}_2-r_{\text{HO}}=0.968 \text{ \AA}$, $r_{\text{OO}}=1.327 \text{ \AA}$, $\theta_{\text{HOO}}=104.30^\circ$; $\text{N}_2-r_{\text{NN}}=1.098 \text{ \AA}$; $\text{NH}-r_{\text{NH}}=1.034 \text{ \AA}$; $\text{NH}_2-r_{\text{NH}}=1.022 \text{ \AA}$, $\theta_{\text{HNN}}=102.89^\circ$; $\text{NH}_3-r_{\text{NH}}=1.010 \text{ \AA}$, $\theta_{\text{HNN}}=106.36^\circ$; $\text{NO}-r_{\text{NO}}=1.147 \text{ \AA}$; $\text{NO}_2-r_{\text{NO}}=1.192 \text{ \AA}$, $\theta_{\text{NOO}}=134.26^\circ$; $\text{O}_2-r_{\text{OO}}=1.205 \text{ \AA}$; $\text{OF}-r_{\text{OF}}=1.350 \text{ \AA}$; $\text{OH}-r_{\text{OH}}=0.967 \text{ \AA}$.
- ⁴¹ T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- ⁴² K. Kuchitsu, in *Accurate Molecular Structures*, edited by A. Domenicano and I. Hargittai (Oxford University Press, Oxford, 1992), pp. 14–46.
- ⁴³ R. A. Kendall and T. H. Dunning, *J. Chem. Phys.* **96**, 6796 (1992).
- ⁴⁴ See, for example, J. F. Stanton and J. Gauss, *Adv. Chem. Phys.* **125**, 101 (2003).
- ⁴⁵ P. G. Szalay, C. S. Simmons, J. Vázquez, and J. F. Stanton, *J. Chem. Phys.* **121**, 7624 (2004).
- ⁴⁶ D. E. Woon and T. H. Dunning, *J. Chem. Phys.* **103**, 4572 (1995).
- ⁴⁷ D. Feller, *J. Chem. Phys.* **96**, 6104 (1992).
- ⁴⁸ It can be shown that the atomic correlation energy converges to the exact result according to Eq. (6), if X is defined such that it corresponds to a basis set in which sets of orbitals associated with all angular momenta up to and including l ($=0$ for s , l , for p , etc.) are saturated. Use of this formula for molecules stands upon a rather flimsy theoretical foundation in the first place, due to the lack of spherical symmetry. It is however an even more extreme (albeit widely used) approximation to treat a basis set such as aug-cc-pCVTZ as though it is saturated with f functions for first-row atoms, since it has but two f functions. Nevertheless, extrapolation schemes based on the formula work well in practice and seem to be as good as any others that have been advocated for this purpose.
- ⁴⁹ T. Helgaker, W. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.* **106**, 9639 (1997).
- ⁵⁰ To test the significance of the combined treatment of core and valence correlation, the total energy of C, O, CO, and CH_2 has been calculated at the frozen core CCSD(T) level extrapolated from aug-cc-pVQZ and aug-cc-pV5Z results. The core correlation effects have been calculated by comparing the extrapolated (Q5) frozen-core and all electron energies at the CCSD(T) level, the latter using aug-cc-pCVXZ basis sets. In all cases, this procedure gives lower total energies compared to the direct inclusion of core correlation effect as used in the HEAT protocol (0.23 , 0.52 , 0.17 , and 0.50 kJ mol^{-1} for C, O, CH_2 , and CO, respectively). From these data the enthalpy of formation of CO changes significantly (by 0.21 kJ mol^{-1}); that of CH_2 shifts by less than 0.1 kJ mol^{-1} .
- ⁵¹ J. F. Stanton, *Chem. Phys. Lett.* **281**, 130 (1997).
- ⁵² D. Feller, *J. Chem. Phys.* **111**, 4373 (1999).
- ⁵³ D. Feller and J. A. Sordo, *J. Chem. Phys.* **113**, 485 (2000).
- ⁵⁴ M. Kállay and J. Gauss, *J. Chem. Phys.* **120**, 6841 (2004).
- ⁵⁵ S. Hirata and R. J. Bartlett, *Chem. Phys. Lett.* **321**, 216 (2000).
- ⁵⁶ J. Olsen, *J. Chem. Phys.* **113**, 7140 (2000).
- ⁵⁷ M. Kállay and P. R. Surján, *J. Chem. Phys.* **115**, 2945 (2001).
- ⁵⁸ M. Kállay and P. R. Surján, *J. Chem. Phys.* **113**, 1359 (2000).
- ⁵⁹ S. Hirata, *J. Phys. Chem. A* **107**, 9887 (2003).
- ⁶⁰ T. A. Ruden, T. Helgaker, P. Jørgensen, and J. Olsen, *Chem. Phys. Lett.* **371**, 62 (2003).
- ⁶¹ Test calculations were carried out for several species, including the CN and CCH radicals that are plagued by significant spin contamination at the UHF level. In no cases did the absolute energies obtained with ROHF and UHF reference functions at the CCSDT level differ by more than 0.1 kJ mol^{-1} .
- ⁶² For NO, CCH, CF, CN, and FO, the ΔE_{ZPE} contribution was obtained at the ROHF/CCSD(T) level with the cc-pVQZ basis set, at the corresponding equilibrium internuclear distances of 1.150 \AA ; 1.208 \AA (CH), and 1.062 \AA (CO); 1.270 \AA , 1.171 \AA , and 1.351 \AA , respectively]. All other contributions to the HEAT energy were obtained with UHF reference functions.
- ⁶³ I. M. Mills, in *Modern Spectroscopy: Modern Research*, edited by K. N. Rao and C. W. Matthews (Academic, New York, 1972), pp. 115–140.
- ⁶⁴ J. Gauss and J. F. Stanton, *Chem. Phys. Lett.* **276**, 70 (1997).
- ⁶⁵ P. G. Szalay, J. Gauss, and J. F. Stanton, *Theor. Chim. Acta* **100**, 5 (1998).
- ⁶⁶ J. F. Stanton and J. Gauss, *Int. Rev. Phys. Chem.* **19**, 61 (2000).
- ⁶⁷ H. H. Nielsen, in *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1959), Vol. 37, Part 1, p. 171. H. H. Nielsen, *Rev. Mod. Phys.* **23**, 90 (1951).
- ⁶⁸ D. G. Truhlar, *J. Comput. Chem.* **12**, 266 (1991).
- ⁶⁹ J. Vázquez and J. F. Stanton (unpublished). For asymmetric tops, G_0 is given in lowest order, by the expression
- $$G_0 = -\frac{1}{4} \sum_{\alpha} B_e^{\alpha} + \frac{1}{64} \sum_k \phi_{kkk} - \frac{7}{576} \sum_k \frac{\phi_{kkk}^2}{\omega_k} + \frac{3}{64} \sum_{k \neq l} \frac{\omega_l \phi_{kkl}^2}{4\omega_k^2 - \omega_l^2} - \frac{1}{4} \sum_{k \neq l \neq m} \frac{\phi_{klm}^2 \omega_k \omega_l \omega_m}{D_{klm}} - \frac{1}{2} \sum_{k \neq l} \sum_{\alpha} B_e^{\alpha} (\xi_{kl}^{\alpha})^2,$$
- where B_e^{α} is the equilibrium rotational constant corresponding to the α th inertial axis, ϕ_{ijk} and ϕ_{ijkl} are cubic and quartic force constants in the dimensionless normal coordinate representation, ω_i is the harmonic frequency of normal mode i , ξ_{kl}^{α} is the Coriolis coupling constant between modes k and l with respect to the α th inertial axis, and D_{klm} is defined as $(\omega_k + \omega_l + \omega_m)(\omega_k - \omega_l - \omega_m)(\omega_k + \omega_l - \omega_m)(\omega_k - \omega_l + \omega_m)$. This equation has been checked by a direct summation procedure, in which the rovibrational Hamiltonian is constructed explicitly and $E_0^{(2)} \equiv \langle 0^{(1)} | H_1 | 0^{(0)} \rangle + \langle 0^{(0)} | H_2 | 0^{(0)} \rangle$ (H_n and $0^{(n)}$ are the n th/order Hamiltonian and ground state wave function, respectively) has been evaluated

- numerically. Our result also agrees with the unpublished result of Allen and co-workers [W. D. Allen (private communication)]. It should be noted that this equation differs from that published in Ref. 70, although the numbers in that paper are consistent with our expression rather than that which was published. We have also derived higher-order terms contributing to G_0 , but they are of no significance for the hydrogen peroxide example addressed later in this paper.
- ⁷⁰ V. Barone, J. Chem. Phys. **120**, 3059 (2004).
- ⁷¹ D. W. Schwenke, J. Phys. Chem. A **105**, 2352 (2001).
- ⁷² E. F. Valeev and C. D. Sherrill, J. Chem. Phys. **118**, 3921 (2003).
- ⁷³ O. L. Polyansky, A. G. Császár, S. V. Shirin, N. F. Zobov, P. Barletta, J. Tennyson, D. W. Schwenke, and P. J. Knowles, Science **299**, 539 (2003).
- ⁷⁴ P. Pykkö, Adv. Quantum Chem. **11**, 353 (1978); P. Pykkö, Chem. Rev. (Washington, D.C.) **88**, 563 (1988); K. Balasubramanian, *Relativistic Effects in Chemistry, Part A: Theory and Techniques and Part B: Applications* (Wiley, New York, 1997).
- ⁷⁵ J. L. Tilson, W. C. Ermler, and R. M. Pitzer, Comput. Phys. Commun. **128**, 128 (2000).
- ⁷⁶ R. M. Pitzer (private communication), see also <http://www.chemistry.ohio-state.edu/pitzer>
- ⁷⁷ W. C. Ermler, R. B. Ross, and P. A. Christiansen, Adv. Quantum Chem. **19**, 139 (1988), and references therein.
- ⁷⁸ E. R. Davidson, Y. Ishikawa, and G. L. Malli, Chem. Phys. Lett. **84**, 226 (1981).
- ⁷⁹ H. Lischka, R. Shepard, I. Shavitt *et al.*, COLUMBUS, an *Ab Initio* Electronic Structure Program, release 5.9, 2003.
- ⁸⁰ T. D. Crawford, C. D. Sherrill, E. F. Valeev *et al.*, Psi 3.2, 2003. Freely available at <http://www.pscod.org/>.
- ⁸¹ J. F. Stanton, J. Gauss, J. D. Watts, W. J. Lauderdale, and R. J. Bartlett, Int. J. Quantum Chem. **S26**, 879 (1992).
- ⁸² J. F. Stanton, C. L. Lopreore, and J. Gauss, J. Chem. Phys. **108**, 7190 (1998).
- ⁸³ For a discussion of CCSDT in the context of atomization energies, see: K. L. Bak, P. Jørgensen, J. Olsen, T. Helgaker, and J. Gauss, Chem. Phys. Lett. **317**, 116 (2000).
- ⁸⁴ Especially when there is strong spin contamination or symmetry breaking (equivalent to unphysical spin localization when appropriate symmetry elements are not present). In this context, it is not surprising to note that CN has by far the largest post-CCSD(T) correction, as it is the most significantly spin-contaminated species in the study. An ROHF-based $E_{\text{HF}}^{\infty} + \Delta E_{\text{CCSD(T)}}^{\infty}$ analysis would be interesting in this context, but is beyond the scope of this initial presentation and study of the HEAT approach.
- ⁸⁵ J. Gauss, W. J. Lauderdale, J. F. Stanton, J. D. Watts, and R. J. Bartlett, Chem. Phys. Lett. **182**, 207 (1991).
- ⁸⁶ J. D. Watts, J. Gauss, and R. J. Bartlett, J. Chem. Phys. **98**, 8718 (1993).
- ⁸⁷ Atomic enthalpies of formation (0 K) were taken from Ruscic's Active Thermochemical Tables (ATcT): H (216.034 ± 0.000 kJ mol⁻¹); C (711.79 ± 0.21 kJ mol⁻¹); N (470.592 ± 0.045 kJ mol⁻¹); O (246.844 ± 0.002 kJ mol⁻¹); F (77.21 ± 0.24 kJ mol⁻¹). These values have considerably smaller error bars than those from the CODATA compilation, and have been chosen in this work for that reason.
- ⁸⁸ Enthalpies of formation for carbon and fluorine atoms given in the NIST-JANAF database (see Ref. 6) are 711.19 ± 0.46 kJ mol⁻¹ and 77.28 ± 0.30 kJ mol⁻¹. Note that these uncertainties are larger than the mean absolute errors in the enthalpies of formation determined with HEAT. Recent work by some of us (B. Ruscic, A. G. Császár, and J. F. Stanton, to be published) has focused on an improved estimate for $\Delta_f H^\circ(C_{\text{gas}})$. Note the relatively large difference between the NIST-JANAF value and that from the active tables (711.79 vs 711.19 kJ mol⁻¹) for carbon.
- ⁸⁹ The idea of the Active Thermochemical Tables is presented in detail in: G. von Laszewski, B. Ruscic, P. Wagstrom *et al.*, in *Lecture Notes in Computer Science*, edited by M. Parashar (Springer, Berlin, 2002), Vol. 2536, pp. 25–38.
- ⁹⁰ The ATcT results presented in this paper were privately communicated by B. Ruscic, and were obtained from Active Thermochemical Tables, ver. 1.20 operating on the Core (Argonne) Thermochemical Network ver. 1.032 (2004). A small subset of these values (H, O, H₂O, CO₂, F, and HF) will appear in: B. Ruscic, R. E. Pinzon, M. L. Morton, G. V. Laszewski, S. J. Bittner, S. G. Nijssure, K. A. Amin, M. Minkoff, and A. F. Wagner, J. Phys. Chem. A. **108**, 9979 (2004).
- ⁹¹ A very basic description of the Active Thermochemical Tables is also scheduled to appear as: B. Ruscic, in *Active Thermochemical Tables*, McGraw-Hill 2005 Yearbook of Science and Technology (McGraw-Hill, New York, 2004).
- ⁹² B. Ruscic (private communication).
- ⁹³ For H₂O₂, the constant contribution to the zero-point energy [G_0 in Eq. (10)] appears to be important. For the elemental formation reaction, H₂ + O₂ → H₂O₂, the calculated contribution (J. Vázquez and J. F. Stanton, unpublished) of G_0 at the CCSD(T)/DZP level of theory is -32.7 cm⁻¹, or about -0.4 kJ mol⁻¹. When this is added to the HEAT value of $\Delta_f H^\circ$ (H₂O₂), agreement with experiment is improved significantly. Note also that an independent variational estimate [M. Mladenovic, Spectrochim. Acta **58A**, 809 (2002)] of the ZPE is 22.4 cm⁻¹ below the value in Table I.
- ⁹⁴ For a discussion concerning CCH, see P. G. Szalay, L. S. Thøgersen, J. Olsen, M. Kállay, and J. Gauss, J. Phys. Chem. A **108**, 3030 (2004).
- ⁹⁵ One could even question why diffuse functions are used in the HEAT approach. If indeed the extrapolation formulas gave the correct result for any sequence of hierarchical basis sets (cc-pCVXZ or aug-cc-pCVXZ), then the use of diffuse functions in the HEAT method would offer no benefit whatsoever. However, the extrapolation formulas are essentially empirical in nature (see Ref. 48). Also, if diffuse functions are of importance, then the sequence of basis sets containing them should converge to the exact result more rapidly than that which omits them.
- ⁹⁶ J. Noga, W. Klopper, and W. Kutzelnigg, in *Recent Advances in Coupled-Cluster Methods* (World Scientific, Singapore, 1997), p. 1.
- ⁹⁷ S. A. Kucharski and R. J. Bartlett, Chem. Phys. Lett. **158**, 550 (1989); J. Chem. Phys. **108**, 9221 (1998).
- ⁹⁸ Y. J. Bomble, M. Kállay, J. Gauss, and J. F. Stanton (unpublished).
- ⁹⁹ A. Tatji, M. Kállay, P. G. Szalay, J. Gauss, and J. F. Stanton (unpublished).
- ¹⁰⁰ J. Berkowitz, G. B. Ellison, and D. Gutman, J. Phys. Chem. **98**, 2744 (1994).
- ¹⁰¹ It should be noted that calculation of $\Delta_f H^\circ$ for C₂H₂ using the second and third isodesmic reaction in Table V together with ATcT values of H_f° for CCH, CH₂, and CH yields 227.70 and 229.32 kJ mol⁻¹, respectively, which clearly differ appreciably. The discrepancy is removed to some degree by using $\Delta_f H^\circ$ values for CH and CH₂ from Császár and co-workers (Refs. 8 and 9) (592.47 and 390.45 kJ mol⁻¹, respectively), with the former notably outside the ATcT range, which gives $\Delta_f H^\circ$ (C₂H₂) = 228.16 kJ mol⁻¹ (second reaction in Table V) and 229.52 kJ mol⁻¹ (third reaction). Error in these calculations arises almost exclusively from uncertainties in the values used in the subtraction procedure; the calculated reaction enthalpies are unlikely to be in error by more than 0.5 kJ mol⁻¹. This example illustrates the pros and cons of using reaction-based approaches, where the theoretically calculated contribution can be excellent but the experimental uncertainty annoying, and the elemental formation reaction where there is absolutely no uncertainty in the “experimental” numbers, but the theoretical calculations are more demanding. Note that pentuple excitations, essential for the calculation of $\Delta_f H^\circ$ (C₂H₂) from elemental reactions, is not at all important with the isodesmic reaction.
- ¹⁰² M. Kállay, P. G. Szalay, and J. Gauss, J. Chem. Phys. **119**, 2991 (2003).