

The correlation-consistent composite approach: Application to the G3/99 test set

Nathan J. DeYonker,^{a)} Tom Grimes, Scott Yockel, Adriana Dinescu, Benjamin Mintz, Thomas R. Cundari,^{b)} and Angela K. Wilson^{c)}

Center for Advanced Scientific Computing and Modeling (CASCaM), Department of Chemistry, The University of North Texas, Denton, Texas 76203

(Received 13 April 2006; accepted 28 June 2006; published online 13 September 2006)

The correlation-consistent composite approach (ccCA), an *ab initio* composite technique for computing atomic and molecular energies, recently has been shown to successfully reproduce experimental data for a number of systems. The ccCA is applied to the G3/99 test set, which includes 223 enthalpies of formation, 88 adiabatic ionization potentials, 58 adiabatic electron affinities, and 8 adiabatic proton affinities. Improvements on the original ccCA formalism include replacing the small basis set quadratic configuration interaction computation with a coupled cluster computation, employing a correction for scalar relativistic effects, utilizing the tight-*d* forms of the second-row correlation-consistent basis sets, and revisiting the basis set chosen for geometry optimization. With two types of complete basis set extrapolation of MP2 energies, ccCA results in an almost zero mean deviation for the G3/99 set (with a best value of -0.10 kcal mol⁻¹), and a 0.96 kcal mol⁻¹ mean absolute deviation, which is equivalent to the accuracy of the G3X model chemistry. There are no optimized or empirical parameters included in the computation of ccCA energies. Except for a few systems to be discussed, ccCA performs as well as or better than *Gn* methods for most systems containing first-row atoms, while for systems containing second-row atoms, ccCA is an improvement over *Gn* model chemistries. © 2006 American Institute of Physics. [DOI: 10.1063/1.2236116]

I. INTRODUCTION

Though the successes of sophisticated molecular orbital-based *ab initio* techniques such as Møller-Plesset perturbation theory (MP n) and coupled cluster theory are ubiquitous, the poor computational scaling of such methods (at least N^5 where N is the number of basis functions) with respect to the expansion of the one-particle basis set is still a significant limiting factor on the size of systems that can be examined. The need for high-accuracy computations on larger molecules has brought about the advent of “composite methods.” The philosophy behind composite methods is to reproduce the electronic energy of an expensive and often computationally intractable electronic structure computation using additive approximations of basis set and electron correlation effects obtained from multiple computations of higher efficiency and lower accuracy.

The first widely used composite methods were the Gaussian- n methods, or *Gn* methods.^{1–10} The Gaussian-1 or G1 method^{2,3} was constructed along with a test set of atoms and molecules that had certain experimentally known energetic properties such as enthalpy of formation (ΔH_f°), adiabatic ionization potential (IP), adiabatic electron affinity (EA), and proton affinity (PA). Generally, model chemistries follow the formulation of *Gn* methods by computing equilibrium geometries and harmonic vibrational frequencies at an

efficient but rather low level of theory. Upon this reference geometry, single point energies are computed and additivity assumptions are made regarding basis set, electron correlation, and other molecular effects that influence the electronic energy of an atomic or molecular system.

The overall goal of the G1 method was to compute the energies of the G2-1 test set to within ± 2 kcal mol⁻¹ of the experimental values. Provided the test set is sufficiently diverse, an accuracy of ± 2 kcal mol⁻¹ could be expected for energies of systems outside of the training sets. The *Gn* methods provided an efficient and “black box” manner of computing molecular energies, and as a result, they have also been used in many studies to compute isomeric energy differences and reaction pathway energies (see Refs. 9, 11, and 12, as well as references therein). The original test set used to benchmark the G1 method was called “G2-1” and contained 55 ΔH_f° 's, 38 IPs, 25 EAs, and 8 PAs. Equilibrium geometries for the molecular systems of the various *Gn* test sets and benchmark data in comparison with experimental values for current *Gn* methods are available,¹¹ providing an invaluable way to compare new composite approaches or computational methods directly with the *Gn* model chemistries.

The latest formulation of the *Gn* composite method is called “G3,”^{5,12} where the empirical high-level correction (HLC) was fitted to the G2/97 set of 302 experimentally known quantities. Compared to the G1 method, two important additive corrections were included in the G3 energies, core-valence correlation and atomic spin-orbit splitting. Fur-

^{a)}Electronic mail: ndeyonk@unt.edu

^{b)}Electronic mail: tomc@unt.edu

^{c)}Electronic mail: akwilson@unt.edu

ther modifications to the G3 method have included reducing the basis set size,¹³ employing coupled cluster computations with single, double, and perturbative triple excitations [CCSD(T)] instead of quadratic configuration interaction [QCISD(T)] to create the G3(CC) variant,^{14,15} using B3LYP density functional theory (DFT) equilibrium geometries and vibrational frequencies (G3B and G3X),^{16,17} using second- or third-order MP n theory instead of MP4 to determine the reference electronic energy [G3(MP2) and G3(MP3)],^{18,19} and using scale factors to replace the HLC (G3S and G3SX).^{17,20}

Besides the G_n methods, other composite approaches have been introduced. Many of these model chemistries attempt to approach the complete basis set/full-configuration interaction limit when applied to smaller systems, and report accuracies to within 0.50 kcal mol⁻¹ of experimental data. Some notable composite methods are the complete basis set (CBS- n) theories of Petersson and co-workers,²¹⁻²⁴ the focal point method of Allen *et al.*²⁵⁻²⁹ and Császár *et al.*²⁷⁻³² the W1 and W2 methods of Parthiban and Martin³³ and Martin and de Oliveira,³⁴ and the HEAT method of Stanton and co-workers^{35,36} Lastly, the composite method of Dixon and co-workers³⁷⁻⁴⁵ has been largely successful by computing coupled cluster equilibrium geometries and total energies with large basis sets while using more efficient levels of theory to compute spin-orbit splitting, core-valence correlation, scalar relativistic, and zero-point vibrational anharmonicity effects.

Some popular composite methods, such as the CBS approaches of Petersson and G_n , use parametrized empirical corrections to account for various deficiencies within the model chemistry. The G3B model chemistry performs quite well compared to DFT for determining enthalpies of formation. However, it has been shown that G3 performs far worse than B3LYP if the HLC is not included in the G3 method.⁴⁶ Due to the nature of how the HLC is applied to the total energies, the magnitude of the HLC increases as the number of electron pairs per molecule increases. Thus, with increasing molecular size, the HLC becomes more integral to the success of the G_n approaches.

II. DESCRIPTION OF ORIGINAL ccCA METHOD

Recently, we have successfully formulated a composite method that predicts atomic energies, molecular energies, and thermodynamic properties consistently accurate to within 1 kcal mol⁻¹ of experiment, without the inclusion of large basis set coupled cluster computations, empirical parameters, or energy scaling of additive corrections.⁴⁶ The “correlation-consistent composite approach” (ccCA) was originally based on G3B theory,¹⁶ where the efficiency of DFT and wide availability of both analytic first and second derivatives for DFT were exploited. For larger systems, B3LYP equilibrium geometries and harmonic vibrational frequencies are likely to be more computationally efficient, yet relatively as accurate compared to experimental values as using MP2 or coupled cluster equilibrium geometries and harmonic vibrational frequencies. Replacing all-electron MP2/6-31G(d) geometry optimizations/frequency computations with B3LYP/6-31G(d) was found to save considerable

CPU time without a change in the overall accuracy of the G2 or G3 method.^{16,47,48} A vibrational frequency scaling factor of 0.9854 (equivalent to that used in G3X theory)¹⁷ rather than the scaling factor of 0.96 used in G3B was found to slightly decrease overall errors.

As compared to the G_n methods which employed the Pople-style contracted Gaussian basis sets, such as 6-31G(d) (Refs. 49 and 50) and 6-311G(2 df , p), the ccCA used correlation-consistent basis sets. These basis sets, originally developed by Dunning,⁵¹ were constructed by systematically expanding the angular momentum functions in shells. Correlation-consistent basis sets exist for first-,⁵¹ second-, and third-row elements,⁵²⁻⁵⁴ have been augmented with diffuse functions (aug-cc-pVXZ),^{55,56} and have been modified to account for core-valence electron correlation (cc-pCVXZ).^{57,58}

Once the ccCA equilibrium geometry was obtained, a reference energy was computed upon which further computations provided additive corrections. The first implementation of ccCA, called ccCA-aTZ, used the MP4/aug-cc-pVTZ single point as the reference energy. The ccCA-aTZ method generally performed with an accuracy equivalent to the G3 method for the 125 energies within the G2-1 test set. However, once the ccCA-aTZ model chemistry was applied to a few larger systems within the G3/99 (Ref. 8) test set, the accuracy was found to worsen proportionally with increasing molecular size. Earlier *ab initio* investigations of the stability of Møller-Plesset perturbative expansions have suggested that the inclusion of triple excitations in the MP4 wave function can produce large electron correlation errors.⁵⁹⁻⁶¹ As the MP4 method scales as N^7 , where N is the number of basis functions, large basis set MP4 computations would likely be impractical for the larger molecules of the G3/99 test set.

The first modification of ccCA was to replace the expensive MP4 computations with MP2. Because of the time savings gained from removing the MP4 computations, one-electron basis set error could also be removed by performing CBS extrapolations. The previous ccCA study found that the simple exponential form proposed by Xantheas and Dunning⁶² and Feller,^{63,64}

$$E(x) = A_{\text{CBS}} + B \exp(-Cx), \quad (1)$$

where x = the $zeta$ level of the correlation-consistent basis set [2=DZ, 3=TZ, etc.], was inferior to the mixed exponential/Gaussian functional devised by Peterson *et al.*,⁶⁵ expressed as

$$E(x) = A_{\text{CBS}} + B \exp[-(x-1)] + C \exp[-(x-1)^2]. \quad (2)$$

By implementing the MP2 CBS energy as the reference for additive corrections, the ccCA-CBS-1 method (using the Gaussian functional form), and the ccCA-CBS-2 method (using the Peterson mixed functional form) were constructed. The ccCA-CBS-2 gave an improvement for the mean absolute deviation of the G2-1 test set of 0.1 kcal mol⁻¹ compared to ccCA-CBS-1.

In order to properly account for high-order electron correlation effects, a single point energy was computed at the

triple-zeta level with the QCISD(T) wave function. The $\Delta E(\text{QCI})$ correction for the ccCA-CBS methods can be expressed as

$$\Delta E(\text{QCI}) = E[\text{QCISD(T)/cc-pVTZ}] - E[\text{MP2/cc-pVTZ}]. \quad (3)$$

The final computation was a correction for core-valence correlation effects. This energy, $\Delta E(\text{CV})$, was computed as

$$\Delta E(\text{CV}) = E[\text{MP2(full)/aug-cc-pCVTZ}] - E[\text{MP2/aug-cc-pVTZ}]. \quad (4)$$

Including the zero-point energy [$\Delta E(\text{ZPE})$] determined by the scaled B3LYP/6-31G(2df,p) harmonic vibrational frequencies, the ccCA-CBS energy was defined as

$$E_0(\text{ccCA-CBS}) = E[\text{MP2/aug-cc-pV}\infty\text{Z}] + \Delta E(\text{QCI}) + \Delta E(\text{CV}) + \Delta E(\text{ZPE}). \quad (5)$$

When computing ΔH_f values with ccCA-CBS-2, enthalpies of formation were mostly lower than the experimental values. The sign of the ccCA-CBS-2 error tended to lie in an advantageous direction, as adding atomic spin-orbit splitting and relativistic effects greatly improved the ΔH_f values. The overall ccCA-CBS-2 mean absolute deviation (MAD) for the G2-1 training set was 0.92 kcal mol⁻¹, better than the G3 MAD of 1.02 kcal mol⁻¹ and almost exactly the MAD of the G3B composite method (0.93 kcal mol⁻¹). Both of these G3 methods, of course, include the HLC.

In terms of accuracy, ccCA was found to perform as well if not better than the G3/G3B methods in most respects, and performed satisfactorily when compared to more expensive large basis set coupled cluster computations.⁴⁶ Unlike DFT or G3 without the HLC, the accuracy of ccCA-CBS-2 did not quickly deteriorate with increasing molecular size. While computationally more demanding than the *Gn* methods, ccCA was found to be tractable for significantly large molecules, such as adamantane and C(NO₂)₄, that could not be easily studied with large basis set coupled cluster methods.

III. MODIFICATIONS TO THE ccCA METHOD

In the present contribution, we implement the following modifications to the original ccCA method. They are listed as follows: (1) an atomic spin-orbit coupling correction is added from experimental results⁶⁶ or theoretical computations,⁵ (2) replacement of QCISD(T) energies with CCSD(T) energies, (3) an explicit treatment of relativistic effects using the Douglas-Kroll-Hess Hamiltonian with a MP2 wave function, and (4) use of the tight-*d* augmented basis sets [cc-pV(*x+d*)Z, aug-cc-pV(*x+d*)Z, and aug-cc-pCV(*x+d*)Z] for second-row atoms.^{52,67}

While the overall accuracy of the ccCA-CBS-2 model chemistry was within the target accuracy of $\pm(1-2)$ kcal/mol, there were some systematic improvements that could be made to the model chemistry. In the initial ccCA study, inclusion of atomic spin-orbit corrections was found to significantly improve atomic IPs and EAs, as well as atomization energies. In the present research, the atomic

spin-orbit coupling correction is added to atomic energies and taken from experimental values⁶⁶ or from theoretical computations for a few of the atomic ions.⁵

The second major modification to the original ccCA method is replacement of the QCISD(T) correction with a CCSD(T) correction. Curtiss *et al.* concluded that the G3(CC) variant, which simply replaced the small basis set QCISD(T) computation with a CCSD(T) computation, did little to change the overall accuracy of the G3 method.¹⁴ In fact, using various forms of G3 with QCISD(T) had a slightly lower average absolute deviation. However, coupled cluster corrections should be preferred for black box approaches because of the more widespread availability of parallelized CCSD(T) code (especially for open-shell systems). Also, as QCISD(T) neglects certain product terms containing one-electron excitation operators (T_1),⁶⁸ the wave function can give spurious energies when a system has a large T_1 amplitude. This can be an indication of multireference character. Some examples of this effect have manifested themselves in a transition-metal *ab initio* study,⁶⁹ as well as in the computation of polarizabilities⁷⁰ and enthalpies of formation for alkali earth metal oxides.^{71,72} The coupled cluster correction [$\Delta(\text{CC})$] replaces $\Delta(\text{QCI})$ and is

$$\Delta E(\text{CC}) = E[\text{CCSD(T)/cc-pVTZ}] - E[\text{MP2/cc-pVTZ}]. \quad (6)$$

The third major enhancement of the original ccCA is to add an explicit treatment of scalar relativistic effects. To obtain an estimate of the impact of these effects, the previous ccCA study applied the relativistic shifts in ΔH_f obtained by Keziora *et al.*^{73,74} with the stationary direct perturbation theory⁷⁵ (SDPT) method and the G3Large basis set. While treatment of relativistic effects adds another computation to the ccCA method, it is essential for a complete treatment of molecular energies. In this study, scalar relativistic corrections were obtained from frozen-core MP2 or CCSD(T) wave functions using the cc-pVTZ-DK (Ref. 76) basis sets and the spin-free, one-electron Douglas-Kroll-Hess (DKH) Hamiltonian.⁷⁷⁻⁷⁹ The MP2 relativistic correction to the ccCA energy, $\Delta E(\text{SR-MP2})$, is formulated as

$$\Delta E(\text{SR-MP2}) = E[\text{MP2/cc-pVTZ-DK}] - E[\text{MP2/cc-pVTZ}], \quad (7)$$

whereas the CCSD(T) relativistic correction $\Delta E(\text{SR-CC})$ would be computed as

$$\Delta E(\text{SR-CC}) = E[\text{CCSD(T)/cc-pVTZ-DK}] - E[\text{CCSD(T)/cc-pVTZ}]. \quad (8)$$

The final improvement to ccCA is the use of correlation-consistent tight-*d* basis sets [cc-pV(*x+d*)Z, aug-cc-pV(*x+d*)Z, and aug-cc-pCV(*x+d*)Z] for the second-row atoms.^{52,80-82} With ccCA-CBS-2, molecules containing second-row elements, for example, ClOH and SO₂, had enthalpies of formation with large deviations from experiment (3.1 and 6.9 kcal mol⁻¹, respectively). When ccCA-CBS-2 was applied to a limited set of nonhydrogen molecules from the G3/99 benchmark set, it was found that the MAD of nine

second-row containing molecules was almost four times greater than the MAD of 14 hydrocarbons and substituted hydrocarbons, 2.29 vs 0.68 kcal mol⁻¹.

Dunning *et al.*⁵² have found that reoptimization of the valence-*d* set with an additional *d* function in second-row correlation-consistent basis sets [resulting in the cc-pV(*x* + *d*)Z basis sets] improved the description of molecular core polarization effects. This modification of the basis sets increased the quality of basis set convergence and overall molecular energies without severely impacting computational time. Historically, G3 and its variants perform far worse for nonhydrogen versus organic species, even with the HLC. For example, while the mean absolute deviation for the 60 additional closed-shell organic molecules in the G3/99 test is only 0.69 kcal mol⁻¹, the G3 MAD is 3.24 kcal mol⁻¹ for the 13 added nonhydrogen species.⁸ The smaller Pople basis sets fail to treat the second-row core polarization effects, and this effect is large enough that the *Gn* HLC cannot fully compensate for this problem.

IV. COMMENTS ON SOME OF THE MOLECULES CONTAINED IN THE G3/99 BENCHMARKING SET

The G3/99 (Ref. 8) training set includes 376 atomic and molecular energies. The breakdown of experimental properties to be compared using these energies is 222 ΔH_f 's, 88 IPs, 58 EAs, and 8 PAs. The authors of the *Gn* composite methods further subdivide the 222 enthalpies of formation within the G3/99 set into the following five categories, 38 hydrocarbons, 91 substituted hydrocarbons, 47 nonhydrogens, 15 inorganic hydrides, and 31 radicals. All nonrelativistic computations in this study were performed using the GAUSSIAN03 software package,⁸³ and all relativistic computations were performed using MOLPRO 2002.6.⁸⁴ As MOLPRO does not have completely unrestricted MP2 and CCSD(T) algorithms implemented in the code, the scalar relativistic corrections are computed using the difference between the completely restricted R/ROMP2 and R/ROCCSD(T) relativistic and nonrelativistic wave functions.⁸⁵

The various *Gn* benchmark sets include molecules that have experimental uncertainties of less than 1 kcal mol⁻¹ and have a broad variation in structure, size, and bonding types. This criterion for inclusion within the training sets is based only on firm experimental data and is not influenced by other theoretical studies or by conclusions made from the computation of *Gn* energies. Cioslowski *et al.*⁸⁶ created their own benchmark set of 600 enthalpies of formation, many of which are included in the G3/99 test set.

Periodically, experimental and theoretical work can modify or cast doubt on some of the quantities in the test set. For example, enthalpies of formation for the following radicals, CH, CH₂ (both the ³B₁ and ¹A₁ states), CH₃, NH₂, OH, CH₃CO, CH₃O, and CH₃CH₂O, have been revisited by Ruscic *et al.*,⁸⁷ providing new recommended ΔH_f values. Cioslowski *et al.* recently discussed some of the challenges in gathering such a wide range of experimental data in order to create a benchmarking set of molecules. As is well known, there are many experimental procedures of varying reliability that can be used to determine gas-phase ΔH_f .^{86,88–92} A further

TABLE I. Systems where a different $\Delta H_f^{298.15}$ (in kcal mol⁻¹) value has been used in the G3/99 test set.

Molecule	G3/99 expt. ΔH_f value	ccCA ΔH_f value used	G3/99 expt. reference	ccCA expt. reference
H ₃ COH	-48.0	-48.2 ^a	88	91
SO ₂	-71.0	-70.9 ^b	90	90
CH ₃ CN	18.0	17.7 ^c	92	117
CH ₃ SCH ₃	-8.9	-9.0 ^b	91	91
CF ₃ Cl	-169.5	-169.2 ^d	92	90
1,4-cyclohexadiene	25.0	25.8 ^e	92	92

^aThe ΔH_f value of H₃COH is not directly measured or reported in Ref. 88, therefore we have resorted to the value in the compendium of Pedley *et al.*

^bError in ΔH_f is potentially from round-off error.

^cThe ccCA ΔH_f value is from a more recent experiment than the value cited in the compendium of Lias *et al.*

^dThe JANAF value of ΔH_f is from more recent experiments than those contained in the compendium of Lias *et al.*

^eThe *Gn* value taken from Lias *et al.* has a typographical error.

complication is the vast amount of both experimental and theoretical data compiled in the major thermodynamic compendia, which themselves are subject to occasional typographical errors (which have been known to propagate even within the JANAF tables) as well as variation due to the different statistical schemes used to calculate recommended ΔH_f values. Differences between the training set of Cioslowski *et al.* and the G3/99 set are usually only 0.1–0.2 kcal mol⁻¹ and mostly arise from round-off error or from the use of different thermochemical compendia. We have reexamined the ΔH_f value whenever such a difference occurs. Table I compiles enthalpies of formation where we used an experimental value different from that of the G3/99 set (Refs. 6, 8, and 93), and justification is provided in the footnotes for the preferred value used in our investigation. However, there are cases of larger discrepancies and the question remains as to which thermochemical data are the most reliable.

One quantity in the G3/99 set has been removed because of recent experimental uncertainty, the COF₂ enthalpy of formation. Asher *et al.*⁹⁴ investigated the difference of 7–8 kcal mol⁻¹ between ΔH_f values obtained with computational model chemistries (CBS-QCI/APNO, G2, and G3) (Refs. 5 and 95) and the JANAF value.⁹⁰ Asher *et al.* performed photoionization experiments and were able to redefine the ΔH_f COF₂ value to have an upper bound halfway between the previously tabulated and computed values. However, due to the rather large error bars and the fact that the new value was only an upper bound instead of a recommended ΔH_f , Curtiss *et al.* removed the ΔH_f (COF₂) value from the G3/99 test set.⁸

Table II lists the ccCA and G3 values of a few systems where computational data^{96,97} suggest the tabulated values may need to be revisited. These systems are vinyl chloride, acetyl acetylene, C₂F₄, C₂Cl₄, and B₂F₄⁺. Of these, the cited enthalpy of formation for acetyl acetylene [CH₃C(=O)CCH] especially warrants reconsideration. The ΔH_f value of 15.6±0.2 kcal mol⁻¹ used by Curtiss *et al.*⁸ comes from the thermochemical compilation of Lias *et al.*⁹² While almost all of the thermochemical results in the com-

TABLE II. Molecules in the G3/99 set for which the recommended enthalpy of formation could require additional experimental research.

Molecule	Common name	Expt. ΔH_f°	ccCA ΔH_f° Dev	G3 ΔH_f° Dev	Gn expt. value
COF ₂	Carbonic difluoride	-152.7 ^a -153.1±1.4 ^c -152.95±0.25 ^d	-4.2	-3.4	-149.1 ^b
C ₂ F ₄	Tetrafluoro-ethene	-157.4 ^e -157.9±0.8 ^f -157.9±0.8 ^g -164.0 ^h -162±1 ⁱ	3.3	4.9	-157.4 ^e
C ₂ Cl ₄	Tetrachloro-ethene	-2.971 ^e -3.6 ⁱ	2.8	3.4	-3.0 ^e
CH ₃ COCCH	Acetyl acetylene	7.0 ^k	-3.9	-2.5	15.6 ^j
CH ₂ CHCl	Chloroethene	5.0 ^m 8.43±0.33 ⁿ 9.10±0.20 ^o 5.5 ^p	3.8	3.6	8.9 ^l
Si(CH ₃) ₄	Tetramethyl silane,	-68.499 ^e	-3.0	-2.6	-55.7 ^k
PF ₅	Phosphorus pentafluoride	-38.1073 ^c	-3.0	-7.1	-38.1 ^e
C ₁₀ H ₈	Azulene	73.5 ^q 66.9 ^r	-5.3	-1.6	69.1 ^l
B ₂ F ₄ ⁺	Diboron tetrafluoride cation	Expt. IP 278.3 ^k	7.7	7.0	

^aReference 118.^bReference 94.^cReference 119.^dReference 120.^eReference 90.^fReference 121.^gReference 122.^hReference 123.ⁱReference 124.^jReference 92.^kReference 125.^lReference 91.^mReference 126.ⁿReference 127.^oReference 128.^pReference 129.^qReference 130.^rReference 131.

pilation of Lias *et al.* come from either published literature or other thermochemical compilations (such as the book of Pedley *et al.*),⁹¹ the enthalpy of formation of acetyl acetylene comes from a personal communication. To our knowledge, no detailed thermodynamic work has been published for this system, and the significant deviation from this experimental result using both the ccCA and G3 suggests new experiments would be of great interest. Lastly, while the G3/99 ΔH_f° values of Si(CH₃)₄, PF₅, and azulene have error bars less than ±1 kcal mol⁻¹, as shown in Table II, there exists a wide range of experimental values not considered by the authors of the Gn methods. However, to maintain the benchmarking philosophy of the Gn methods, the aforementioned systems will not be removed from the test set, nor will experimental values be adjusted until convincing experimental evidence changes the recommended values.

The enthalpic data for azulene and many larger hydrocarbons reflect the inherent experimental and computational limitations in terms of obtaining accurate gas-phase enthalpies of formation for materials that exist in a condensed phase under standard conditions. The gas-phase enthalpy is obtained from the solid-state enthalpy by the addition of the enthalpy of sublimation. While solid-state enthalpies are

typically obtained by experimental methods such as combustion calorimetry with high accuracy and precision, experimental enthalpies of sublimation are, however, often fraught with uncertainty being measured by a variety of techniques that require high temperatures to generate sufficient vapor pressure of the material under inquiry;⁹⁸ enthalpies of sublimation are then extrapolated back to a standard temperature such as 298.15 K. To illustrate, in a massive review of enthalpies of sublimation, Chicoks and Acree⁹⁸ provided ΔH_{sub} for azulene ranging from 17 kcal mol⁻¹ (combined correlation gas chromatography–differential scanning calorimetry) to 23 kcal mol⁻¹ (Knudsen effusion). In an application of the ccCA to nitroanilines,⁹⁹ computations have suggested that there may be cases where error in experimental enthalpies of sublimation could propagate to the gas-phase enthalpies of formation.

V. ASSESSMENT OF THE MODIFIED ccCA METHOD FOR THE G3/99 TEST SET

A. Basis set dependence of DFT geometry optimization

The most important modifications to the Gn model chemistries have been attempts to improve the efficiency of

geometry optimizations and the computation of harmonic vibrational frequencies, which are crucial to the efficiency paradigm of most composite methods. Baboul *et al.*¹⁶ tested the G3 method using B3LYP geometry optimizations with a small basis set [6-31G(*d*)] instead of the original implementation of all-electron MP2 geometries with scaled Hartree-Fock frequencies. It was determined that using B3LYP equilibrium geometries and harmonic frequencies changed 27 of the 299 total energies in the G2/97 test set by more than 1 kcal mol⁻¹. Of those 27 molecules, the G3 energies were improved in 18 of the cases. Baboul *et al.* also noted that in many cases, the DFT geometry more closely resembled the QCISD and experimental geometries. The use of DFT for geometry optimizations and harmonic vibrational frequencies is more efficient than all-electron MP2 optimizations, especially for larger systems. Also, obtaining DFT harmonic vibrational frequencies will not be as costly as obtaining MP2 or coupled cluster frequencies and DFT vibrational frequencies are far more accurate than the scaled HF frequencies used in G3. However, any composite method using DFT equilibrium geometries will perform inappropriately when DFT methods cannot provide accurate equilibrium geometries, or optimize to nonphysical stationary points. Examples of this in the G2/97 test set are BF₃⁺, BCl₃⁺, and CH₄⁺, where degenerate electronic states cause a Jahn-Teller distortion.

In the introduction of the G3X method, Curtiss *et al.* explored various equilibrium geometries of a few test cases in the G3/99 set.¹⁷ They studied the 21 molecules from the G3/99 test set. Seven of the molecules (set A) were second-row containing nonhydrides, while the other 14 (set B) were small first-row molecules, including some hydrogen-containing species. Typifying the nonsystematic nature of DFT methods with respect to basis set size, increasing the basis set size from the 6-31G(2*df*,*p*) set to the 6-311+G(2*df*,*p*) gave a higher MAD for set A.¹⁷ Also, a comparison of B3LYP with an all-electron MP2 computation made with the same basis sets showed that B3LYP consistently gave better geometries while at the same time provided a reduction in the computational time requirements. Thus, their conclusion was that B3LYP/6-31G(2*df*,*p*) geometry optimizations provided the best balance between efficiency and accuracy.

In the initial development of the ccCA composite method, B3LYP/6-31G(*d*) equilibrium geometries and scaled harmonic frequencies were implemented, but the geometry dependence of the DFT functional and basis set size was not explicitly investigated. We have examined the DFT basis set dependence of the same 21-molecule G3/99 subset investigated by Curtiss *et al.* in Ref. 17. In Table III, bond distances and bond angles are computed with the B3LYP functional along with the 6-31G(*d*), 6-31+G(*d*), 6-31G(2*df*,*p*), cc-pVDZ, aug-cc-pVDZ, and aug-cc-pVTZ basis sets. For comparison to the original G3 method, MP2(full)/6-31G(*d*) geometries are also presented. For this test set, the accuracy of the Pople-style basis sets increases as the basis set size is increased. The MAD versus experimental geometries is lowest for both set A (0.011 Å) and set B (0.006 Å) when using B3LYP/6-31G(2*df*,*p*) optimizations.

Examining B3LYP optimizations with the medium-sized Pople-style basis sets and the triple-zeta correlation-consistent family of basis sets provides little difference. The MADs of B3LYP/cc-pVDZ and B3LYP/aug-cc-pVDZ methods are far worse (0.042 and 0.043 Å for bond distances and 1.2° and 1.4° for bond angles, respectively) than the other levels of theory for set A and are comparable to B3LYP/6-31G(*d*) for set B, in Table III. Using B3LYP/cc-pVTZ gives a larger MAD for set A than B3LYP/6-31G(2*df*,*p*), but results in significant improvements for set B. To obtain accurate equilibrium geometries with a DFT functional, basis sets of triple-zeta quality appear to be essential, while larger basis sets are likely to substantially lower the efficiency of optimizations and frequency computations with an unsubstantial further gain in accuracy.

In Table IV, experimental properties for a selection of molecules from the G3/99 test set are computed using the ccCA method at various B3LYP geometries. The difference between ccCA energies at B3LYP/6-31G(2*df*,*p*) geometries versus B3LYP/cc-pVTZ is rarely more than a few tenths of a kcal mol⁻¹. In almost all cases, the ccCA energies based on the 6-31G(2*df*,*p*) geometries are even better than those obtained with the cc-pVTZ basis sets. The largest difference is in the computation of ΔH_f for SF₆, where the ccCA energy at the B3LYP/6-31G(2*df*,*p*) geometry is 1.1 kcal mol⁻¹ closer to experiment (but still within the overall MAD of 0.96 kcal mol⁻¹) than the ccCA energy from the B3LYP/cc-pVTZ geometry. Since DFT energies do not systematically improve with respect to increasing basis set size, there are likely cases where the B3LYP/cc-pVTZ geometry provides more accurate values within the G3/99 set. However, the similarities amongst the basis sets as well as the slightly smaller basis set size lead us to recommend B3LYP/6-31G(2*df*,*p*) as the method of choice for ccCA geometry optimizations.

B. MP2 versus CCSD(T) relativistic corrections

While relativistic effects are commonly thought to significantly influence the electronic structure of molecules containing elements starting with the 3*d* transition metals, Kedziora *et al.*^{73,74} found that treatment of scalar relativistic effects could produce a shift in atomization energies of more than 0.5 kcal mol⁻¹ for many main-group systems. Scalar relativistic effects on atomization energies becomes larger with increasing molecular size. As previously mentioned, determining the scalar relativistic contribution to the total ccCA energy requires an additional computation [Eqs. (7) and (8)]. Currently, most high-accuracy *ab initio* studies compute scalar relativistic effects with the CCSD(T) method and either a Douglas-Kroll Hamiltonian (for example, Refs. 40, 46, 100, and 101) or first-order perturbation theory upon the one-electron mass-velocity and Darwin terms¹⁰² with configuration interaction or coupled cluster methods (for example, Refs. 33 and 43).

Unfortunately, comparing the difference between the CCSD(T)/cc-pVTZ-DK and CCSD(T)/cc-pVTZ energies within ccCA will take as much computational time as repeating the $\Delta(\text{CC})$ correction. The $\Delta(\text{CC})$ correction can become

TABLE III. Comparison of basis sets for geometry optimization of selected molecules in the G3/99 test set. (Bond lengths are in angstroms and bond angles are in degrees.)

Parameter	MP2(full)/			B3LYP					Expt. ^b
	6-31G(<i>d</i>) ^a	6-31G(<i>d</i>)	6-31+G(<i>d</i>)	6-31G(2 <i>df</i> , <i>p</i>)	cc-pVDZ	aug-cc-pVDZ	cc-pVTZ		
Set A									
PF ₃	P–F	1.593	1.596	1.610	1.576	1.626	1.630	1.589	1.57
	F–P–F	97.6	97.7	97.5	97.8	97.3	97.0	97.5	97.8
PF ₅	P–F(ax)	1.595	1.597	1.606	1.578	1.622	1.622	1.593	1.577
	P–F(eq)	1.566	1.569	1.574	1.550	1.593	1.590	1.558	1.534
P ₄	P–P	2.195	2.217	2.218	2.211	2.238	2.240	2.214	2.21
PCl ₅	P–Cl(ax)	2.139	2.176	2.181	2.157	2.180	2.183	2.165	2.19
	P–Cl(eq)	2.038	2.071	2.069	2.055	2.085	2.083	2.058	2.04
SO ₂	S–O	1.477	1.464	1.466	1.443	1.481	1.484	1.450	1.432
	O–S–O	119.8	119.1	118.6	119.2	117.7	117.5	118.3	119.5
SO ₃	S–O	1.458	1.453	1.454	1.432	1.468	1.471	1.439	1.43
SF ₆	S–F	1.592	1.600	1.608	1.575	1.615	1.617	1.588	1.564
Deviation ^c	Bonds	0.027	0.025	0.028	0.011	0.042	0.043	0.017	
	Angles	0.3	0.3	0.6	0.2	1.2	1.4	0.8	
Set B									
LiH	Li–H	1.640	1.621	1.620	1.615	1.604	1.602	1.590	1.595
CH	C–H	1.121	1.133	1.128	1.131	1.142	1.136	1.124	1.12
CH ₄	C–H	1.090	1.093	1.094	1.092	1.100	1.097	1.088	1.085
NH ₃	N–H	1.017	1.019	1.018	1.017	1.025	1.019	1.014	1.012
	H–N–H	106.3	105.7	107.7	105.5	104.3	106.8	106.3	106.7
H ₂ O	O–H	0.969	0.969	0.969	0.962	0.969	0.965	0.961	0.957
	H–O–H	104.0	103.6	105.5	103.7	102.7	104.7	104.5	104.5
FH	F–H	0.934	0.934	0.938	0.919	0.927	0.926	0.922	0.917
C ₂ H ₂	C–C	1.216	1.205	1.208	1.199	1.210	1.210	1.196	1.203
	C–H	1.066	1.067	1.068	1.062	1.072	1.071	1.062	1.061
C ₂ H ₄	C–C	1.335	1.331	1.335	1.327	1.334	1.336	1.324	1.339
	C–H	1.085	1.087	1.088	1.086	1.095	1.092	1.083	1.085
	H–C–H	116.6	116.3	116.5	116.1	116.7	116.7	116.6	117.8
HCN	C–N	1.176	1.157	1.158	1.152	1.158	1.157	1.146	1.153
	C–H	1.069	1.070	1.072	1.067	1.077	1.074	1.065	1.065
CO	C–O	1.150	1.138	1.137	1.131	1.135	1.134	1.126	1.128
H ₂ CO	C–O	1.220	1.206	1.209	1.200	1.204	1.208	1.199	1.208
	C–H	1.104	1.110	1.108	1.111	1.120	1.114	1.106	1.116
	H–C–H	115.7	115.2	116.3	115.1	115.2	116.4	115.9	116.5
N ₂	N–N	1.130	1.105	1.105	1.099	1.104	1.104	1.091	1.098
N ₂ H ₄	N–N	1.438	1.437	1.434	1.437	1.436	1.436	1.437	1.447
	N–Ha	1.016	1.017	1.016	1.015	1.022	1.017	1.011	1.008
	N–Hb	1.020	1.022	1.020	1.019	1.027	1.020	1.015	1.008
	Ha–N–N	106.4	106.6	107.7	106.8	106.7	107.6	107.3	109.2
	Hb–N–N	111.5	111.8	112.6	111.6	111.6	112.2	111.9	109.2
	Ha–N–Hb	107.0	106.8	108.4	106.6	105.8	107.6	107.7	113.3
	Ha–N–N–Hb	90.6	90.5	91.4	88.6	88.0	89.9	89.8	88.9
O ₂	O–O	1.246	1.215	1.215	1.206	1.209	1.209	1.206	1.208
Deviation ^d	Bonds	0.013 ^e	0.009	0.009	0.006	0.010	0.008	0.005	
	Angles	2.0 ^e	2.3	2.0	2.1	2.5	1.6	1.7	

^aMP2(full)/6-31G(*d*) optimized geometries are from Ref. 18, Table II.^bExperimental values are from Ref. 90 (set A) and from Ref. 132 (set B).^cMean absolute deviations for bond lengths and bond angles in set A.^dMean absolute deviations for bond lengths and bond angles in set B.^eMean absolute deviations of MP2(full) 6-31G(*d*) set B parameters will slightly differ from Ref. 18 Table II because we have removed N₂H₂ (which is not a molecule in the G3/99 test set) from this subset.

the computational bottleneck of ccCA, and using the coupled cluster relativistic correction can significantly increase the time to compute the ccCA energy. For this reason, the use of the Peterson mixed Gaussian/exponential CBS function has

been examined with the energies of the G3/99 test set using both $\Delta(\text{SR-MP2})$ and $\Delta(\text{SR-CC})$. The mean absolute deviation of the 376 quantities in the G3/99 test set is only 0.004 kcal mol⁻¹ higher with the MP2 DK correction. There-

TABLE IV. Geometry dependence of a few quantities in the G3/99 test set. Properties are computed using the ccCA-P method. Deviations are shown in kcal mol⁻¹.

	Expt.	6-31G(d)	6-31+G(d)	6-31G(2df,p)	cc-pVDZ	aug-cc-pVDZ	cc-pVTZ
NCCN ΔH_f	73.3	-1.8	-1.7	-2.0	-1.8	-1.8	-1.9
H ₂ NNH ₂ ΔH_f	22.8	1.4	1.5	0.7	1.1	1.3	1.5
NH ΔH_f	85.2	-0.6	-0.6	-0.7	-0.6	-0.6	-0.6
Adiabatic EA	8.8	2.5	2.4	2.2	2.4	2.4	2.3
CH ₃ OH ΔH_f	-48.0	1.3	1.4	0.7	1.1	1.3	1.4
Adiabatic IP	250.2	-1.7	-1.5	-1.0	-1.5	-1.7	-1.6
CH ₃ SH ΔH_f	-5.5	1.2	1.5	0.9	1.3	1.3	1.3
Adiabatic IP	217.7	-0.9	-0.6	-0.6	-0.7	-0.7	-0.7
SO ₂ ΔH_f	-70.9	-1.2	-1.2	-0.3	-2.8	-3.2	-0.3
Adiabatic EA	25.5	0.2	0.1	-0.2	0.1	-0.4	0.0
C ₂ Cl ₄ ΔH_f	-3.0	2.5	2.6	2.8	2.3	2.4	2.9
SiCl ₄ ΔH_f	-158.4	-0.7	-0.6	-0.3	-0.4	-0.6	0.0
C ₄ H ₄ S ΔH_f	27.5	0.8	0.9	0.2	0.5	0.5	1.0
SF ₆ ΔH_f	-219.7	-4.8	-6.5	-1.6	-8.5	-9.1	-2.7

fore, the MP2 DK correction can be efficiently computed in lieu of a much more expensive CCSD(T) DK computation.

C. Dependence on CBS fit

In the initial ccCA study, two functions were used to calculate the CBS MP2 energy that was used as the reference energy [Eqs. (1) and (2)]. It was found that the choice of CBS function can have a rather substantial effect on the quality of thermodynamic and experimental properties. Besides the two functions used to fit the CBS energy in the first implementation of the ccCA, there is a variety of other CBS functions that are frequently used. Following investigations of correlation energy on small systems performed by Schwartz^{103,104} and Kutzelnigg and Morgan,¹⁰⁵ functions have been devised based upon inverse powers related to the highest angular momentum represented (l_{\max}) in the basis sets. Two such functions have the forms¹⁰⁶⁻¹⁰⁹

$$E(l_{\max}) = E_{\text{CBS}} + \frac{B}{(l_{\max} + \frac{1}{2})^4} \quad (9)$$

and

$$E(l_{\max}) = E_{\text{CBS}} + \frac{B}{(l_{\max})^3}. \quad (10)$$

A simple analytic formula for Eq. (10) has recently been given by Halkier *et al.*¹⁰⁹ Using triple-zeta and quadruple-zeta energies, Eq. (10) can be reformulated as

$$E_{\text{CBS}} = \frac{E_{\text{QZ}} \times 4^3 - E_{\text{TZ}} \times 3^3}{4^3 - 3^3}. \quad (11)$$

Wilson and Dunning¹¹⁰ also proposed the following function for MP2 CBS extrapolations:

$$E^{(2)}(l_{\max}) = E_{\text{CBS}}^{(2)} + \frac{B}{(l_{\max} + 1)^4} + \frac{C}{(l_{\max} + 1)^5}. \quad (12)$$

The different extrapolation schemes are named ccCA-F [Eq. (1)], ccCA-P [Eq. (2)], ccCA-S4 [Eq. (9)], ccCA-S3 [Eq. (10)], and ccCA-WD [Eq. (12)]. In this study, the total MP2 energies (not just the correlation energies) are extrapolated to the CBS limit. With five different types of CBS extrapolation functions, the full listing of ccCA atomic (Table S1) and molecular properties as well as individual deviations from experiment (Table S2) are given as supplementary material.¹¹¹

TABLE V. Signed mean deviations of ccCA methods using different CBS energy fitting schemes.

	ccCA-P	ccCA-F	ccCA-WD	ccCA-S3	ccCA-S4
Enthalpies of formation (222)	-0.21	-2.04	1.36	1.54	0.04
Nonhydrogens (47)	-0.23	-0.97	1.29	0.91	-0.15
Hydrocarbons (38)	-0.59	-3.78	1.04	1.96	-0.12
Substituted hydrocarbons (91)	-0.11	-2.55	1.87	2.15	0.21
Inorganic hydrides (15)	0.84	0.48	1.57	1.39	0.87
Radicals (31)	-0.51	-1.39	0.30	0.36	-0.39
Ionization energies (88)	-0.54	-0.07	-1.00	-0.69	-0.51
Electron affinities (58)	-0.16	0.04	-0.52	-0.28	-0.09
Proton affinities (8)	1.02	1.20	1.21	1.00	0.98
All quantities (376)	-0.25	-1.19	0.51	0.73	-0.10
Standard deviation	1.30	2.01	1.76	1.72	1.30

TABLE VI. Mean absolute deviations of the G3/99 test set.

	G3	G3X	ccCA-P	ccCA-F	ccCA-WD	ccCA-S3	ccCA-S4
Enthalpies of formation (222)	1.04	0.88	0.86	2.31	1.58	1.72	0.89
Nonhydrogens (47)	2.11	1.49	1.03	1.75	1.60	1.15	0.96
Hydrocarbons (38)	0.57	0.46	0.85	3.55	1.22	2.02	0.86
Substituted hydrocarbons (91)	0.74	0.75	0.79	2.60	1.96	2.22	0.89
Inorganic hydrides (15)	0.87	0.82	0.87	0.60	1.57	1.39	0.90
Radicals (31)	0.80	0.75	0.75	1.45	0.89	0.99	0.77
Ionization energies (88)	1.51	1.07	1.17	1.20	1.44	1.23	1.14
Electron affinities (58)	1.04	0.99	0.98	1.00	1.10	1.01	0.99
Proton affinities (8)	1.33	1.21	1.11	1.28	1.26	1.09	1.08
All quantities (376)	1.16	0.95	0.96	1.83	1.47	1.48	0.97
rms deviation	1.54	1.35	1.32	2.42	1.83	1.87	1.30

The signed deviations of ccCA for the G3/99 test set with the various CBS extrapolation schemes are presented in Table V, along with the breakdown of deviations for the subcategories of atomic and molecular properties. The MADs are given in Table VI, compared to the G3 and G3X values in Table VI, and shown pictorially in Fig. 1. From Table VI, the ccCA-P and ccCA-S4 extrapolation schemes give the lowest mean absolute deviations from experimental values for the 376 energies in the G3/99 test set. Both ccCA-P and ccCA-S4 give overall MADs that are an improvement on G3 and comparable to G3X. Compared to G3X (1.49 kcal mol⁻¹ MAD for the 47 systems), the two best ccCA CBS fits (ccCA-P: MAD=1.03 kcal mol⁻¹; ccCA-S4: MAD=0.96 kcal mol⁻¹) are significantly more reliable at computing ΔH_f values for nonhydrogen systems. On the other hand, the G3X MAD of 0.46 kcal mol⁻¹ for the 38 hydrocarbon systems is quite superior to the MADs of ccCA-P (0.85 kcal mol⁻¹) and ccCA-S4 (0.86 kcal mol⁻¹). All other subcategories and properties included in the G3/99 test set are computed to a nearly equivalent accuracy among ccCA-P, ccCA-S4, and G3X. Shown in Table V, the standard deviations of the ccCA-P and the ccCA-S4 variants (both at 1.30 kcal mol⁻¹) are quite similar to the G3X standard deviation (1.35 kcal mol⁻¹). The root mean squared deviations of

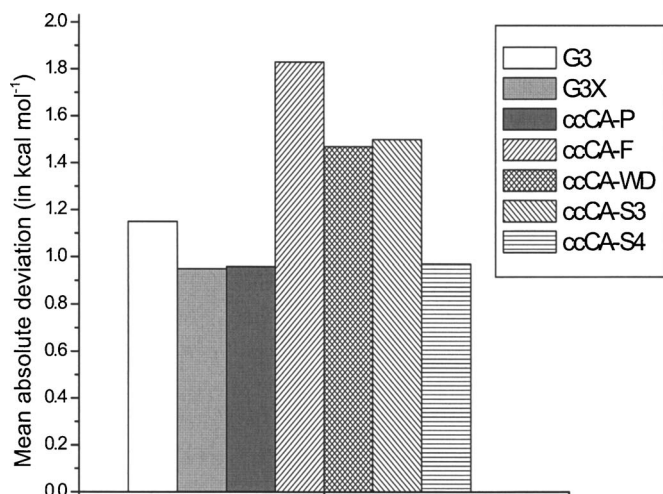


FIG. 1. Mean absolute deviations of G3, G3X, and ccCA energies obtained by various CBS extrapolation procedures for the 376 quantities in the G3/99 test set.

ccCA-P (1.32 kcal mol⁻¹) and ccCA-S4 (1.30 kcal mol⁻¹) are slightly smaller than that of G3X as well. By extrapolating ccCA MP2 energies to the CBS limit and eliminating the need for HLC, ccCA-P and ccCA-S4 generally give energies comparable to G_n methods.

The high-level correction (HLC) of the Gaussian- n composite methods is constructed to minimize the signed mean deviation of the entire G2/97 test set. The success of the HLC is evident when analyzing the mean deviation of the expanded G3/99 set, as it is still quite near zero for both G3 and G3X. Though ccCA-P and ccCA-S4 have almost exactly the same MAD, the signed mean deviation of ccCA-S4 is quite close to zero (-0.10 kcal mol⁻¹), which is equal in magnitude to the mean deviation of the G3X method (+0.10 kcal mol⁻¹). It is important to reiterate that this has been achieved in the ccCA framework *without any optimized parameters*.

The histograms in Fig. 2 illustrate the distribution of G3/99 signed deviations of G3, G3X, ccCA-P, and ccCA-S4. The overall results obtained from G3X, ccCA-P, and ccCA-S4 are remarkably similar. With respect to the total number of systems within the G3/99 test set for which chemical accuracy of ± 1 kcal mol⁻¹ is achieved, ccCA-P performs the best, with 248 of the 376 energies (66.0%) for ccCA-P, versus 244 with G3X (64.9%), 235 with G3 (62.5%), and 229 with ccCA-S4 (60.9%). All three of these composite methods show the same largest positive maximum deviation, the IP of B₂F₄, but this quantity could be an outlier for which more experimental work is necessary. The second largest positive deviation for G3 is the ΔH_f of C₂F₄ (4.9 kcal mol⁻¹). For G3X, this is the ΔH_f of Na₂ (4.5 kcal mol⁻¹), and for ccCA-P and ccCA-S4, the ΔH_f of H₂C=CHCl (3.7 and 3.9 kcal mol⁻¹, respectively). The most negative signed deviation for G3X theory is the IP of CN (-6.5 kcal mol⁻¹), while for both ccCA-P and ccCA-S4 it is the ΔH_f of azulene isomer of C₁₀H₈ (-5.0 and -4.3 kcal mol⁻¹, respectively). Generally, the agreement between ccCA-P/ccCA-S4 and G3X is quite good. On average for the G3/99 test set, the difference of the mean deviations between ccCA-P and G3X is -0.35 kcal mol⁻¹, while the ccCA-S4 and G3X difference in mean deviations is 0.20 kcal mol⁻¹.

Compared to G3X, the performance of ccCA is only noticeably worse for hydrocarbon enthalpies of formation. In

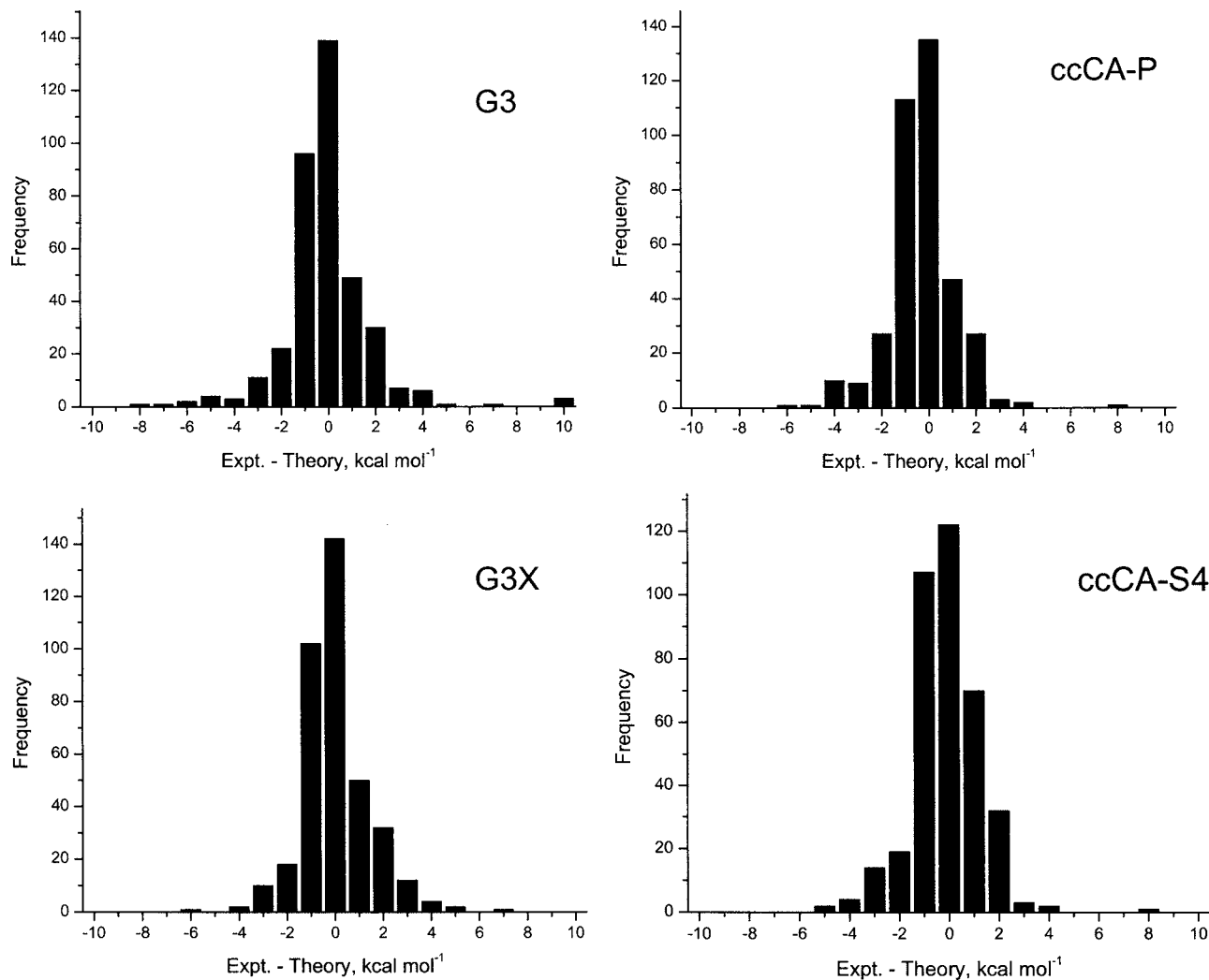


FIG. 2. Histograms showing the distribution of errors in the G3/99 test set.

Fig. 3, the distribution of errors for the 37 hydrocarbons in the G3/99 set is shown for G3, G3X, and ccCA-S4. The two largest absolute hydrocarbon ΔH_f deviations for ccCA are azulene ($-4.3 \text{ kcal mol}^{-1}$ with ccCA-S4) and cyclooctatetraene ($-2.5 \text{ kcal mol}^{-1}$). As noted in Sec. IV, the deviation for azulene may actually be attributed to the experimental data, which is based upon the uncertainty in the reported enthalpies of sublimation. For G3 and G3X, the largest deviation in magnitude, ΔH_f of cyclobutene, is much smaller than with ccCA, $-2.6 \text{ kcal mol}^{-1}$ with G3, and $-2.4 \text{ kcal mol}^{-1}$ with G3X. For the hydrocarbon systems in the G3/99 set, trends are not immediately evident as to why ccCA performs worse than G3 and G3X model chemistries.

In Fig. 4, the distribution of errors for the subcategory of 47 nonhydrogen ΔH_f values is examined. The number of species within chemical accuracy is considerably greater with ccCA, 24 with both ccCA-P and ccCA-S4, in contrast to 19 with G3X and 14 with G3. Even though the G3X treatment of nonhydrogens is significantly improved from G3 due to the B3LYP/6-31G(2df,p) geometry optimizations and the “G3extralarge” Hartree-Fock computation,¹⁷ G3X (Na₂) still has a larger maximum absolute deviation, $4.5 \text{ kcal mol}^{-1}$, than ccCA (C₂F₄), $3.3 \text{ kcal mol}^{-1}$ with

ccCA-P and $3.6 \text{ kcal mol}^{-1}$ with ccCA-S4. G3X also has 11 nonhydrogen enthalpies of formation outside the range of $\pm 3 \text{ kcal mol}^{-1}$ versus only three such deviations with ccCA-P and two with ccCA-S4. Examining the complete set of 134 second-row containing molecules in the G3/99 training set, the signed deviation for ccCA-P is $0.01 \text{ kcal mol}^{-1}$ compared to $0.07 \text{ kcal mol}^{-1}$ with G3X. The MAD for the entire second-row containing set is $0.94 \text{ kcal mol}^{-1}$ with ccCA-P and $1.05 \text{ kcal mol}^{-1}$ with G3X. The ccCA-P standard deviation ($1.25 \text{ kcal mol}^{-1}$) for this subset is an improvement upon G3X ($1.42 \text{ kcal mol}^{-1}$) as well. For group IA- and IIA-containing molecules,^{72,112} which have been considered a severe problem for the *Gn* model chemistries, preliminary results indicate that ccCA proves to be much more reliable than even the G3X composite method.¹¹³

D. Examination of scale factors

The deficiencies of DFT vibrational frequencies and the historical validity of using scale factors have been considered in this work. Previous in-depth studies of scale factors by Sinha *et al.*¹¹⁴ and Scott and Radom¹¹⁵ examined the best combination of scale factors for HF, DFT, and MP2 for the

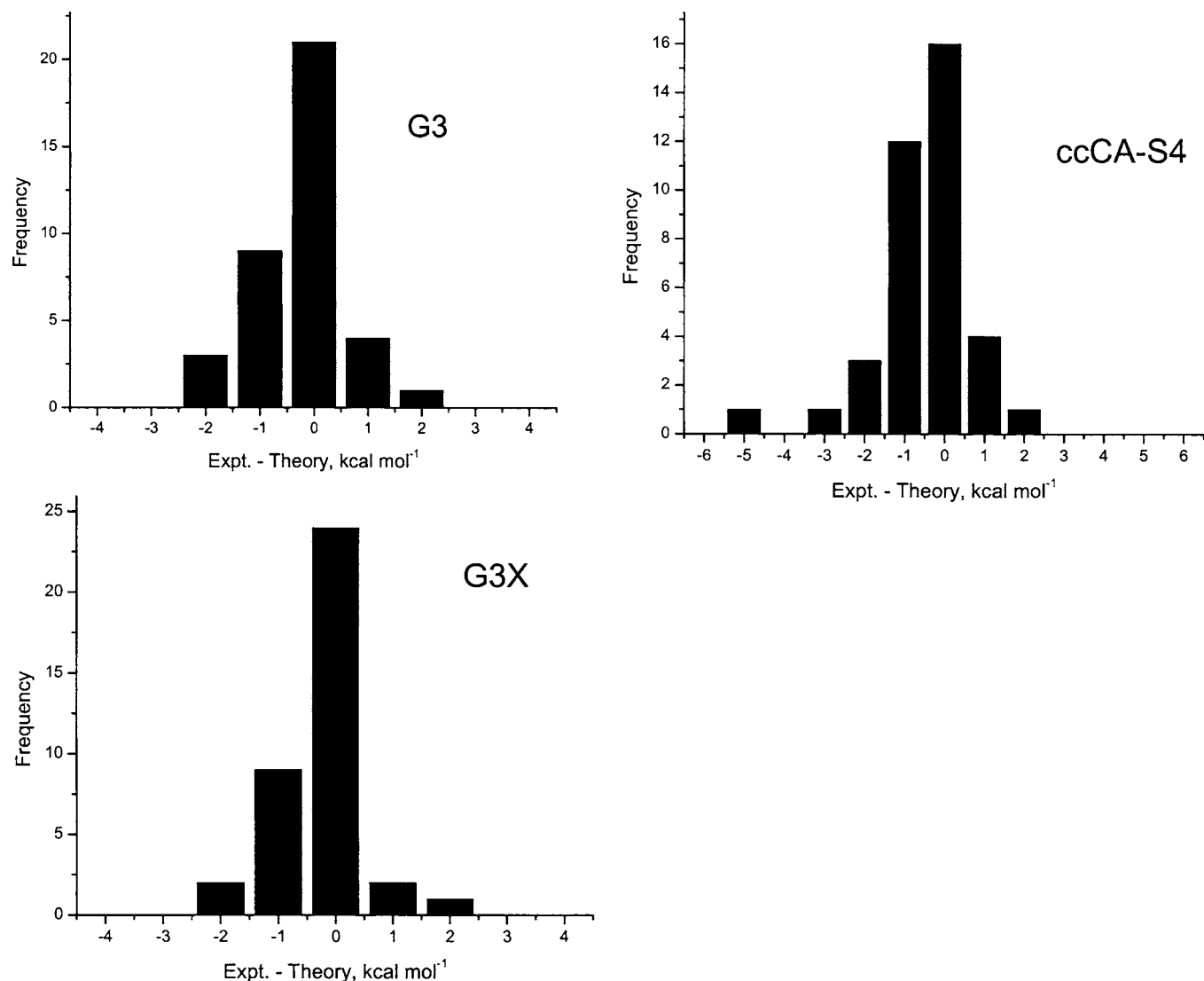


FIG. 3. Distribution of errors for the computation of ΔH_f within the hydrocarbon subcategories of the G3/99 test set.

correlation-consistent and Pople-style basis sets, respectively. Even though scale factors are technically “empirical parameters,” the inclusion of scale factors accounts for the anharmonicity of vibrational frequencies, which is a physical contribution to the electronic energy of a molecule. In order to efficiently compute the quartic force constants necessary to include anharmonicity corrections in the ccCA, analytic DFT third derivatives would be required. Currently, no publicly available software package allows facile computation of analytic third derivatives for any type of wave function for general polyatomic systems. Also, computation of analytical third derivatives would likely be an expensive computational undertaking, even with DFT methods. Scale factors have therefore been used in several high-accuracy theoretical thermochemical investigations such as the *n*-alkane study of Pollack *et al.*⁴⁵

The G3X scale factor of 0.9854 was an optimized parameter fitted by Scott and Radom¹¹⁵ for a training set of 122 molecules. This scale factor was adopted in the first implementation of the ccCA for B3LYP/6-31G(*d*) harmonic frequencies because it provided better ZPEs than the recommended scale factor of 0.965 from the CCCBDB database.¹¹⁶

To determine the impact of different scale factors, the energies of the molecules in the G3/99 set have been recomputed using scale factors of 0.965, 0.9854, and unity.

In Table S3, the ZPE and temperature corrections of the G3/99 set are shown with the three scale factors used in this study.¹¹¹ The signed and mean absolute deviations of the G3/99 set are pictorially shown in Fig. 5. With the scale factor of 0.965, the signed deviation of the G3/99 set with ccCA-P is 0.17 kcal mol⁻¹, while the MAD is 1.13 kcal mol⁻¹. Contrary to the *G_n* necessity of the HLC to achieve chemical accuracy for almost all systems in the G3/99 test set, exclusion of scale factors from the ccCA is not catastrophic. Without vibrational scale factors, the signed deviation with the ccCA-P formalism is -0.57 kcal mol⁻¹, and the MAD is 1.09 kcal mol⁻¹. As an example of how scale factors are a minor influence on the overall accuracy of the ccCA compared to the HLC, for the largest molecule in the G3/99 set (*n*-octane), scaling the frequencies by 0.9854 contributes 2.2 kcal mol⁻¹ to the atomization energy. The G3X HLC on the other hand, contributes 47.3 kcal mol⁻¹ to the atomization energy. In order to attain chemical accuracy, *G_n* methods utilize both the HLC and scale factors. Within

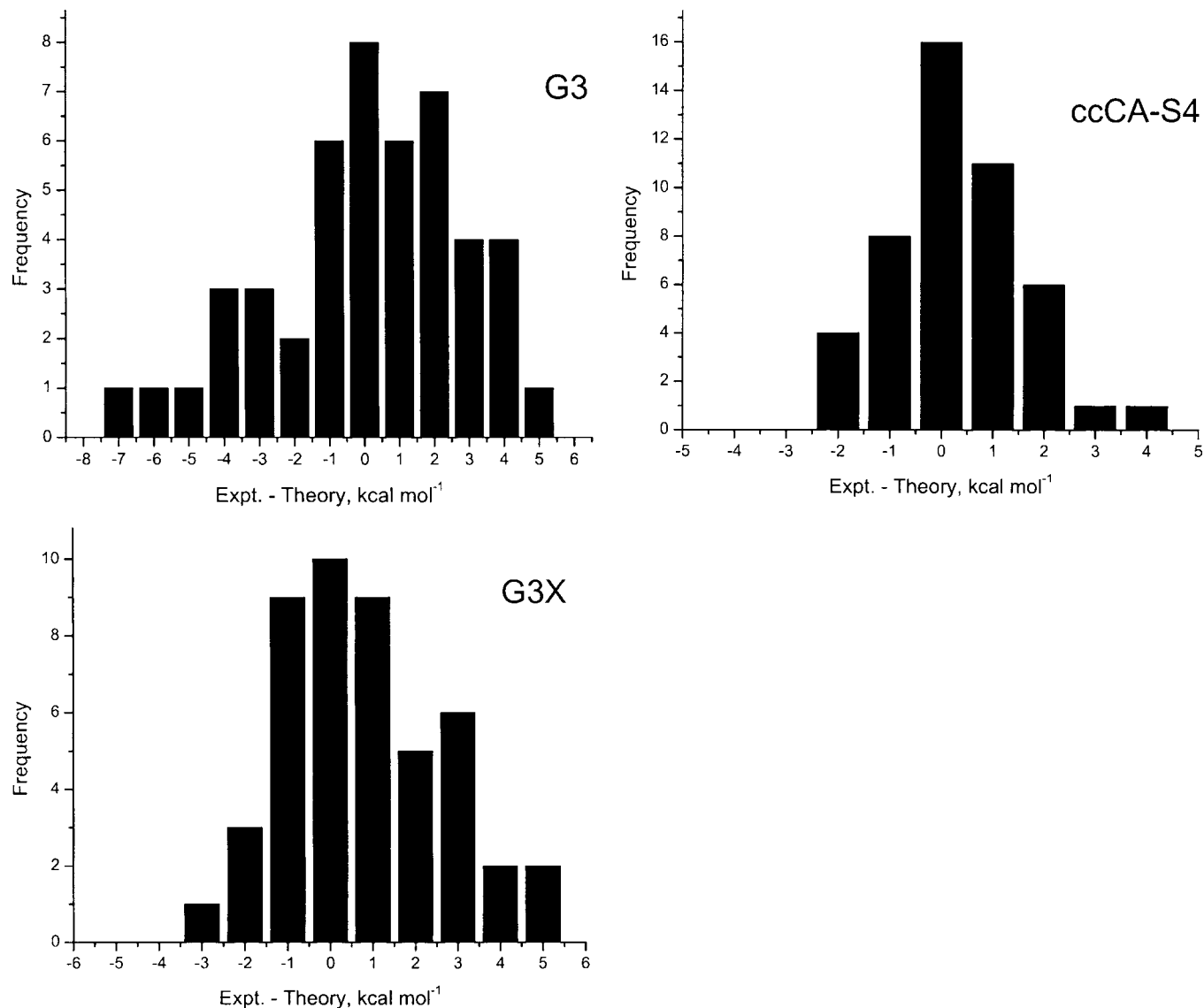


FIG. 4. Distribution of errors for the computation of ΔH_f within the nonhydrogen subcategories of the G3/99 test set.

the ccCA model chemistry, a nonoptimized scale factor improves the accuracy of the method by $0.13 \text{ kcal mol}^{-1}$, taking into account known deficiencies of the vibrational harmonic approximation.

VI. CONCLUSIONS

The modifications to the original implementation of the ccCA lead to an improvement in the description of molecular properties. Four significant modifications have been applied to the ccCA: the effect of spin-orbit coupling is added to the atomic energies, QCISD(T) computations are replaced with CCSD(T), an MP2 Douglas-Kroll computation adds scalar relativistic effects, and the cc-pV($X+d$)Z, aug-cc-pV($X+d$)Z, and aug-cc-pCV($X+d$)Z basis sets have been utilized for second-row atoms. The CBS extrapolation procedure is varied in order to determine which CBS fits perform best. The modified ccCA algorithm is applied to the 376 systems in the G3/99 training set, and the CBS fits in Eq. (2)

(ccCA-P) and Eq. (9) (ccCA-S4) perform best, which give mean absolute deviations of 0.96 and $0.97 \text{ kcal mol}^{-1}$, respectively.

As a composite method, the ccCA generally performs as well or better than the G3X method. Though ccCA has problems with enthalpies of formation for a few hydrocarbon species, it can be used to validate G_n results. When computing enthalpies of formation of molecules containing second-row atoms, ccCA is a substantial improvement over any G3 methods by at least $0.5 \text{ kcal mol}^{-1}$. For hypervalent species such as SF_6 , accuracy of the ccCA benefits from CBS extrapolations and the tight- d family of correlation-consistent basis sets.

Comparing the various MP2 CBS extrapolations, the signed mean deviation of ccCA-S4 is the closest to zero, providing agreement with the G3X model chemistry. Both ccCA-P and ccCA-S4 have a standard deviation smaller than G3X for the 376 quantities in the G3/99 test set. Clearly for the computation of molecular energies, the ccCA model chemistry can be expected to be accurate to at least

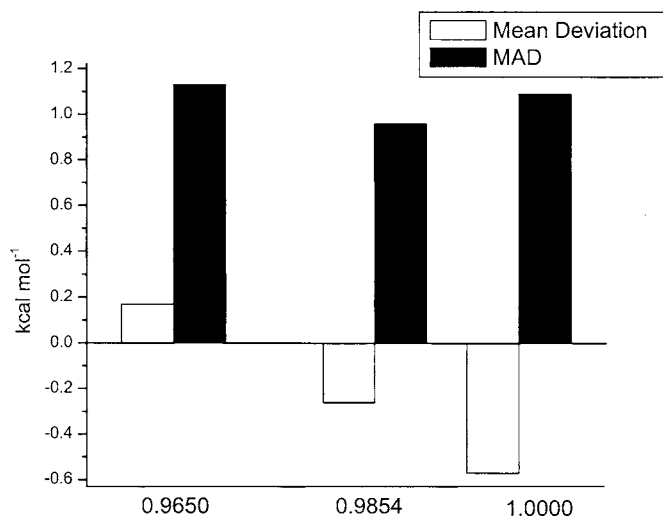


FIG. 5. Deviations of ccCA-P as a function of harmonic vibrational frequency scale factor.

± 1 kcal mol⁻¹. The ccCA could be considered a robust compromise between composite methods that are efficient but parametrized, and those that are highly accurate, but still on the fringe of mainstream utility for molecules with more than 5–15 heavy atoms. In future studies, the ccCA will be applied to molecules significantly larger than azulene and *n*-octane, the largest and most costly members of the G3/99 set, as well as for systems where *Gn* methods give questionable results, such as group IA/IIA oxides and hydroxides.

ACKNOWLEDGMENTS

CASCAM is supported by a grant from the United States Department of Education. This research is partially supported by a grant from the United States Department of Energy, Office of Basic Energy Sciences [to one of the authors (T.R.C.)], Grant No. DE-FG02-03ER15387, and by National Science Foundation CAREER Award No. CHE-0239555 [to another author (A.K.W.)]. Additional support was also provided by the University of North Texas Academic Computing Services for the use of the UNT Research Cluster. Computational resources were provided via the National Science Foundation (CHE-0342824), and by the National Computational Science Alliance under Grant No. CHE010021 and utilized the NCSA IBM p690. Computations employed the UNT computational chemistry resource, whose purchase was supported by a CRIF grant from the U.S. National Science Foundation (CHE-0342824). Professor Kirk A. Peterson (Department of Chemistry, Washington State University) is thanked for providing new group IA and IIA correlation-consistent basis sets prior to their publication. Brian P. Prascher assisted by constructing group IA and IIA Douglas-Kroll relativistic basis sets. Dustin S. Ho is thanked for helping with organization of tables and supplementary material. We would like to thank Dr. Larry A. Curtiss for helping with the subcategorization of ΔH_f° 's.

¹J. A. Pople, B. T. Luke, M. J. Frisch, and J. S. Binkley, *J. Chem. Phys.* **89**, 2198 (1985).

²J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, and L. A. Curtiss, *J. Chem. Phys.* **90**, 5622 (1989).

- ³L. A. Curtiss, C. Jones, G. W. Trucks, K. Raghavachari, and J. A. Pople, *J. Chem. Phys.* **93**, 2537 (1990).
- ⁴L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, *J. Chem. Phys.* **94**, 7221 (1991).
- ⁵L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, and J. A. Pople, *J. Chem. Phys.* **109**, 7764 (1998).
- ⁶L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **106**, 1063 (1997).
- ⁷K. Raghavachari, B. B. Stefanov, and L. A. Curtiss, *J. Chem. Phys.* **106**, 6764 (1997).
- ⁸L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **112**, 7374 (2000).
- ⁹K. Raghavachari and L. A. Curtiss, in *Modern Electronic Structure Theory*, edited by D. R. Yarkony (World Scientific, Singapore, 1995), Vol. 2, p. 991.
- ¹⁰L. A. Curtiss, P. C. Redfern, and D. J. Frurip, in *Reviews of Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (Wiley-VCH, New York, 2000), Vol. 15, p. 147.
- ¹¹L. A. Curtiss, <http://chemistry.anl.gov/compmat/comptherm.htm>
- ¹²L. A. Curtiss and K. Raghavachari, *Theor. Chem. Acc.* **108**, 61 (2002).
- ¹³L. A. Curtiss, P. C. Redfern, B. J. Smith, and L. Radom, *J. Chem. Phys.* **104**, 5148 (1996).
- ¹⁴L. A. Curtiss, K. Raghavachari, P. C. Redfern, A. G. Baboul, and J. A. Pople, *Chem. Phys. Lett.* **314**, 101 (1999).
- ¹⁵J. A. Pople, M. Head-Gordon, and K. Raghavachari, *J. Chem. Phys.* **87**, 5968 (1987).
- ¹⁶A. G. Baboul, L. A. Curtiss, P. C. Redfern, and K. Raghavachari, *J. Chem. Phys.* **110**, 7650 (1999).
- ¹⁷L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, *J. Chem. Phys.* **114**, 108 (2001).
- ¹⁸L. A. Curtiss, K. Raghavachari, and J. A. Pople, *J. Chem. Phys.* **98**, 1293 (1993).
- ¹⁹L. A. Curtiss, P. C. Redfern, K. Raghavachari, V. Rassolov, and J. A. Pople, *J. Chem. Phys.* **110**, 4703 (1999).
- ²⁰L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **112**, 1125 (2000).
- ²¹G. A. Petersson and M. J. Frisch, *J. Phys. Chem. A* **104**, 2183 (2000).
- ²²J. A. Montgomery, M. J. Frisch, J. W. Ochterski, and G. A. Petersson, *J. Chem. Phys.* **110**, 2822 (1999).
- ²³G. A. Petersson, *Abstr. Pap. - Am. Chem. Soc.* **212**, 175 (1996).
- ²⁴J. W. Ochterski, G. A. Petersson, and J. A. Montgomery, *J. Chem. Phys.* **104**, 2598 (1996).
- ²⁵A. L. East and W. D. Allen, *J. Chem. Phys.* **99**, 4638 (1993).
- ²⁶J. P. Kenny, W. D. Allen, and H. F. Schaefer, *J. Chem. Phys.* **118**, 7353 (2003).
- ²⁷J. M. Gonzales, C. Pak, R. S. Cox, W. D. Allen, H. F. Schaefer, A. G. Császár, and G. Tarczay, *Chem.-Eur. J.* **9**, 2173 (2003).
- ²⁸A. G. Császár, W. D. Allen, and H. F. Schaefer, *J. Chem. Phys.* **108**, 9751 (1998).
- ²⁹A. G. Császár, W. D. Allen, Y. Yamaguchi, and H. F. Schaefer, in *Computational Molecular Spectroscopy*, edited by P. Jensen and P. R. Bunker (Wiley, New York, 2000), p. 15.
- ³⁰A. G. Császár, P. G. Szalay, and M. L. Leininger, *Mol. Phys.* **100**, 3879 (2002).
- ³¹A. G. Császár, M. L. Leininger, and A. Burcat, *J. Phys. Chem. A* **107**, 2061 (2003).
- ³²A. G. Császár, M. L. Leininger, and V. Szalay, *J. Chem. Phys.* **118**, 10631 (2003).
- ³³S. Parthiban and J. M. L. Martin, *J. Chem. Phys.* **114**, 6014 (2001).
- ³⁴J. M. L. Martin and G. de Oliveira, *J. Chem. Phys.* **111**, 1843 (1999).
- ³⁵P. G. Szalay, A. Tajti, and J. F. Stanton, *Mol. Phys.* **103**, 2159 (2005).
- ³⁶A. Tajti, P. G. Szalay, A. G. Császár, M. Kallay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vazquez, and J. F. Stanton, *J. Chem. Phys.* **121**, 11599 (2004).
- ³⁷D. Feller and D. A. Dixon, *J. Phys. Chem. A* **107**, 9641 (2003).
- ³⁸D. Feller, D. A. Dixon, and J. S. Francisco, *J. Phys. Chem. A* **107**, 1604 (2003).
- ³⁹D. Feller, K. A. Peterson, W. A. de Jong, and D. A. Dixon, *J. Chem. Phys.* **118**, 3510 (2003).
- ⁴⁰D. A. Dixon, D. Feller, and J. S. Francisco, *J. Phys. Chem. A* **107**, 186 (2003).
- ⁴¹D. A. Dixon, D. Feller, and K. A. Peterson, *J. Chem. Phys.* **115**, 2576 (2001).
- ⁴²D. Feller and D. A. Dixon, *J. Phys. Chem. A* **104**, 3048 (2000).

- ⁴³D. A. Dixon, D. Feller, and G. Sandrone, *J. Phys. Chem. A* **103**, 4744 (1999).
- ⁴⁴D. Feller and D. A. Dixon, *J. Phys. Chem. A* **103**, 6413 (1999).
- ⁴⁵L. Pollack, T. L. Windus, W. A. de Jong, and D. A. Dixon, *J. Phys. Chem. A* **109**, 6934 (2005).
- ⁴⁶N. J. DeYonker, T. R. Cundari, and A. K. Wilson, *J. Chem. Phys.* **124**, 114104 (2006).
- ⁴⁷A. M. Mebel, K. Morokuma, and M. C. Lin, *J. Chem. Phys.* **101**, 3916 (1994).
- ⁴⁸C. W. Bauschlicher, Jr. and H. Partridge, *J. Chem. Phys.* **103**, 1788 (1995).
- ⁴⁹W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972).
- ⁵⁰M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, and J. A. Pople, *J. Chem. Phys.* **77**, 3654 (1982).
- ⁵¹T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- ⁵²T. H. Dunning, Jr., K. A. Peterson, and A. K. Wilson, *J. Chem. Phys.* **114**, 9244 (2001).
- ⁵³A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., *J. Chem. Phys.* **110**, 7667 (1999).
- ⁵⁴A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., *Abstr. Pap. - Am. Chem. Soc.* **213**, 60 (1997).
- ⁵⁵R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- ⁵⁶D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **100**, 2975 (1994).
- ⁵⁷D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **103**, 4572 (1995).
- ⁵⁸K. A. Peterson and T. H. Dunning, Jr., *J. Chem. Phys.* **117**, 10548 (2002).
- ⁵⁹T. H. Dunning, Jr. and K. A. Peterson, *J. Chem. Phys.* **113**, 7799 (2000).
- ⁶⁰M. L. Leininger, W. D. Allen, H. F. Schaefer, and C. D. Sherrill, *J. Chem. Phys.* **112**, 9213 (2000).
- ⁶¹N. C. Handy, P. J. Knowles, and K. Somasundram, *Theor. Chim. Acta* **68**, 87 (1985).
- ⁶²S. S. Xantheas and T. H. Dunning, Jr., *J. Phys. Chem.* **97**, 18 (1993).
- ⁶³D. Feller, *J. Chem. Phys.* **96**, 6104 (1992).
- ⁶⁴D. Feller, *J. Chem. Phys.* **98**, 7059 (1993).
- ⁶⁵K. A. Peterson, D. E. Woon, and T. H. Dunning, Jr., *J. Chem. Phys.* **100**, 7410 (1994).
- ⁶⁶C. E. Moore, *Atomic Energy Levels*, Natl. Bur. Stand. Ref. Data Ser., Natl. Bur. Stand. (U.S.) Circ. No. 35 (U.S. Department of Commerce, Washington D.C., 1971).
- ⁶⁷S. M. Yockel and A. K. Wilson (unpublished).
- ⁶⁸R. J. Bartlett, in *Modern Electronic Structure Theory*, Advanced Series in Physical Chemistry, Vol. 2, edited by D. R. Yarkony (World Scientific, Singapore, 1995), p. 1047.
- ⁶⁹J. Hrusak, S. Tenno, and S. Iwata, *J. Chem. Phys.* **106**, 7185 (1997).
- ⁷⁰J. D. Watts, M. Urban, and R. J. Bartlett, *Theor. Chim. Acta* **90**, 341 (1995).
- ⁷¹J.-P. Blaudeau, M. P. McGrath, L. A. Curtiss, and L. Radom, *J. Chem. Phys.* **107**, 5016 (1997).
- ⁷²M. B. Sullivan, M. A. Iron, P. C. Redfern, J. M. L. Martin, L. A. Curtiss, and L. Radom, *J. Phys. Chem. A* **107**, 5617 (2003).
- ⁷³G. S. Kedziora, J. A. Pople, M. A. Ratner, P. C. Redfern, and L. A. Curtiss, *J. Chem. Phys.* **115**, 718 (2001).
- ⁷⁴G. S. Kedziora, J. A. Pople, V. A. Rassolov, M. A. Ratner, P. C. Redfern, and L. A. Curtiss, *J. Chem. Phys.* **110**, 7123 (1999).
- ⁷⁵W. Kutzelnigg, E. Ottoschowski, and R. Franke, *J. Chem. Phys.* **102**, 1740 (1995).
- ⁷⁶W. A. de Jong, R. J. Harrison, and D. A. Dixon, *J. Chem. Phys.* **114**, 48 (2001).
- ⁷⁷B. A. Hess, *Phys. Rev. A* **33**, 3742 (1986).
- ⁷⁸B. A. Hess, *Phys. Rev. A* **32**, 756 (1985).
- ⁷⁹M. Douglas and N. M. Kroll, *Ann. Phys. (N.Y.)* **82**, 89 (1974).
- ⁸⁰N. X. Wang and A. K. Wilson, *J. Phys. Chem. A* **107**, 6720 (2003).
- ⁸¹R. D. Bell and A. K. Wilson, *Chem. Phys. Lett.* **394**, 105 (2004).
- ⁸²A. K. Wilson and T. H. Dunning, Jr., *J. Phys. Chem. A* **108**, 3129 (2004).
- ⁸³M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.
- ⁸⁴R. D. Amos, A. Bernhardsson, A. Berning *et al.*, MOLPRO, version 2002.6, a package of *ab initio* programs, University of Birmingham, Birmingham, UK, 2002.
- ⁸⁵P. J. Knowles, C. Hampel, and H.-J. Werner, *J. Chem. Phys.* **99**, 5219 (1993).
- ⁸⁶J. Cioslowski, M. Schimeczek, G. Liu, and V. Stoyanov, *J. Chem. Phys.* **113**, 9377 (2000).
- ⁸⁷B. Ruscic, J. E. Boggs, A. Burcat *et al.*, *J. Phys. Chem. Ref. Data* **34**, 573 (2005).
- ⁸⁸J. Berkowitz, G. B. Ellison, and D. Gutman, *J. Phys. Chem.* **98**, 2744 (1994).
- ⁸⁹J. A. M. Simões and J. L. Beauchamp, *Chem. Rev. (Washington, D.C.)* **90**, 629 (1990).
- ⁹⁰M. W. Chase Jr., C. A. Davies, J. R. Downey Jr., D. J. Frurip, R. A. McDonald, and A. N. Syverud, *J. Phys. Chem. Ref. Data Monogr.* **9**, 1 (1998).
- ⁹¹J. B. Pedley, R. D. Naylor, and S. P. Kirby, *Thermochemical Data of Organic Compounds*, 2nd ed. (Chapman and Hall, New York, 1986).
- ⁹²S. G. Lias, J. E. Bartmess, J. F. Liebman, J. L. Holmes, R. D. Levin, and W. G. Mallard, *J. Phys. Chem. Ref. Data Suppl.* **17**, 1 (1988).
- ⁹³L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, *J. Chem. Phys.* **109**, 42 (1998).
- ⁹⁴R. L. Asher, E. H. Appelman, and B. Ruscic, *J. Chem. Phys.* **105**, 9781 (1996).
- ⁹⁵J. J. A. Montgomery, H. H. Michels, and J. S. Francisco, *Chem. Phys. Lett.* **220**, 391 (1994).
- ⁹⁶G. A. Petersson, D. K. Malick, W. G. Wilson, J. W. Ochterski, J. J. A. Montgomery, and M. J. Frisch, *J. Chem. Phys.* **109**, 10570 (1998).
- ⁹⁷B. Ruscic, J. V. Michael, P. C. Redfern, L. A. Curtiss, and K. Raghavachari, *J. Phys. Chem. A* **102**, 10889 (1998).
- ⁹⁸J. S. Chickos and W. E. Acree, *J. Phys. Chem. Ref. Data* **31**, 537 (2002).
- ⁹⁹N. J. DeYonker, T. R. Cundari, A. K. Wilson, C. Sood, and D. H. Magers, THEOCHEM (in press).
- ¹⁰⁰D. Feller and D. A. Dixon, *J. Phys. Chem. A* **107**, 10419 (2003).
- ¹⁰¹D. Feller, K. A. Peterson, and T. D. Crawford, *J. Chem. Phys.* **124**, 054107 (2006).
- ¹⁰²R. D. Cowan and D. C. Griffin, *J. Opt. Soc. Am.* **66**, 1010 (1976).
- ¹⁰³C. Schwartz, *Phys. Rev.* **126**, 1015 (1962).
- ¹⁰⁴C. Schwartz, in *Methods in Computational Physics*, edited by B. J. Alder, S. Fernbach, and M. Rotenberg (Academic, New York, 1963), Vol. 2, p. 262.
- ¹⁰⁵W. Kutzelnigg and J. D. Morgan, *J. Chem. Phys.* **96**, 4484 (1992).
- ¹⁰⁶J. M. L. Martin, *Chem. Phys. Lett.* **259**, 669 (1996).
- ¹⁰⁷T. Helgaker, W. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.* **106**, 9639 (1997).
- ¹⁰⁸J. M. L. Martin and T. J. Lee, *Chem. Phys. Lett.* **258**, 136 (1996).
- ¹⁰⁹A. Halkier, T. Helgaker, P. Jorgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *Chem. Phys. Lett.* **286**, 243 (1998).
- ¹¹⁰A. K. Wilson and T. H. Dunning, Jr., *J. Chem. Phys.* **106**, 8718 (1997).
- ¹¹¹See EPAPS Document No. E-JCPSA6-125-310630 for full atomic and molecular ccCA data. This document can be reached via a direct link in the online article's HTML reference section or via the EPAPS homepage (<http://www.aip.org/pubservs/epaps.html>).
- ¹¹²A. Schulz, B. J. Smith, and L. Radom, *J. Phys. Chem. A* **103**, 7522 (1999).
- ¹¹³D. S. Ho, N. J. DeYonker, T. R. Cundari, and A. K. Wilson, *J. Phys. Chem. A* **110**, 9767 (2006).
- ¹¹⁴P. Sinha, S. E. Boesch, C. M. Gu, R. A. Wheeler, and A. K. Wilson, *J. Phys. Chem. A* **108**, 9213 (2004).
- ¹¹⁵A. T. Scott and L. Radom, *J. Phys. Chem.* **100**, 16502 (1996).
- ¹¹⁶R. D. Johnson III, NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101 Release 12, August 2005, <http://srdata.nist.gov/cccbdb>
- ¹¹⁷X.-W. An and M. Månsson, *J. Chem. Thermodyn.* **15**, 287 (1983).
- ¹¹⁸J. C. Amphlett, J. R. Dacey, and G. O. Pritchard, *J. Phys. Chem.* **75**, 3024 (1971).
- ¹¹⁹H. V. Wartenberg, *Z. Anorg. Chem.* **258**, 354 (1949).
- ¹²⁰D. A. Pittam and G. Pilcher, *J. Chem. Soc., Faraday Trans. 1* **68**, 2224 (1972).
- ¹²¹V. P. Kolesor, I. D. Zenkov, and S. M. Skuratov, *Russ. J. Phys. Chem.* **36**, 45 (1962).
- ¹²²C. A. Nevebaver and J. L. Margrave, *J. Phys. Chem.* **60**, 1318 (1956).
- ¹²³H. V. Wartenberg and J. Schiefer, *Z. Anorg. Allg. Chem.* **278**, 326 (1955).
- ¹²⁴F. W. Kirkbride and F. G. Davidson, *Nature (London)* **174**, 79 (1954).
- ¹²⁵S. V. Levanova, Yu. A. Treger, S. M. Velichko, and A. M. Rozhnov,

- Russ. J. Phys. Chem. **50**, 1148 (1976).
- ¹²⁶ Z. B. Alfassi, D. M. Golden, and S. W. Benson, *J. Chem. Thermodyn.* **5**, 411 (1973).
- ¹²⁷ J. R. Lacher, H. B. Gottlieb, and J. D. Park, *Trans. Faraday Soc.* **58**, 2348 (1962).
- ¹²⁸ J. R. Lacher, E. Emery, E. Bohmfalk, and J. D. Park, *J. Phys. Chem.* **60**, 492 (1956).
- ¹²⁹ V. P. Kolesov and T. S. Papina, *Russ. Chem. Rev.* **52**, 425 (1983).
- ¹³⁰ W. R. Roth, M. Bohm, H.-W. Lenhartz, and F. Vogel, *Angew. Chem.* **95**, 1011 (1983).
- ¹³¹ E. Kovats, H. Gunthard, and A. Plattner, *Helv. Chim. Acta* **38**, 1912 (1955).
- ¹³² J. S. Binkley, J. A. Pople, and W. J. Hehre, *J. Am. Chem. Soc.* **102**, 939 (1980).