Do composite methods achieve their target accuracy?

Rebecca Weber, Angela K. Wilson*

*Department of Chemistry and Center for Advanced Scientific Computing and Modeling
(CASCaM), University of North Texas, Denton, Texas 76203-5017*

ABSTRACT

Enthalpies of formation have been calculated for the 147 molecules of the G2/97 data set using the correlation consistent Composite Approach (ccCA) to assess the ability of this approach to achieve its presumed CCSD(T)-level accuracy. The calculated enthalpies of formation from the G2/97 data set were compared to enthalpies computed at the CCSD(T,FC1)/aug-cc-pCV ∞ Z-DK level of theory. Deviations of both ccCA and CCSD(T) enthalpies from experiment have been evaluated. ccCA results in a mean absolute deviation (MAD) of 0.80 kcal mol⁻¹ from experiment while CCSD(T) has an MAD of 0.90 kcal mol⁻¹. The MAD of the ccCA enthalpies from the CCSD(T,FC1)/aug-cc-pCV ∞ Z-DK enthalpies is 0.87 kcal mol⁻¹.

Keywords: composite methods, ccCA, coupled cluster, extrapolation, complete basis set limit, convergence behavior

* akwilson@unt.edu, phone 940-565-4372, fax 940-565-4318

1 INTRODUCTION

The coupled cluster with single, double, and perturbative triple excitations [CCSD(T)] method has been widely accepted as the “gold standard” of *ab initio* electronic structure methods.[1] In fact, CCSD(T) has been shown to achieve energetic properties such as enthalpies of formation, binding energies, ionization energies, and electron affinities to within chemical accuracy (1 kcal mol⁻¹ of reliable experiment), on average, for hundreds of main group species.[2-5] However, to predict energetic properties to this level of accuracy, large basis sets such as sextuple- ζ or septuple- ζ level may be needed.[5,6] For example, it has been shown that the dissociation energy (D_e) of N₂, even at the CCSD(T)/cc-pV6Z level of theory, correlating only the valence electrons, there is still a deviation from experiment of 2 kcal mol⁻¹. [7] In order to compute D_e for N₂ to within 1 kcal mol⁻¹ of experiment (specifically, achieving a deviation of 0.1 kcal mol⁻¹ from experiment), the authors determined it necessary to include core-valence electron interactions through correlation of the valence and subvalence electrons (which, for first row atoms, is the correlation of all electrons; for second row atoms, is the correlation of all but the 1s electrons, etc.), as well as extrapolation of the energies to the complete basis set (CBS) limit. In fact, a 2001 study by Feller and Dixon determined that, at least for hydrocarbons, including the valence and subvalence electrons within the correlation space can account for up to 1 kcal mol⁻¹ per carbon atom.[3]

CBS extrapolations can account for errors arising from basis set incompleteness through a series of systematically converging basis sets, such as the correlation consistent family of basis sets by Dunning and co-workers.[8,9] This family of basis sets, denoted cc-pVnZ, where n indicates the ζ -level of the basis set, were designed such that all basis functions that recover roughly the same amount of correlation energy are added in shells (i.e. the 2d and 1f functions of the cc-pVTZ sets for boron through fluorine recover approximately the same amount of correlation energy, as do the 3d, 2f, and 1g functions of the cc-pVQZ sets, etc.). This design results in a hierarchical family of basis sets that has the property that they systematically approach the CBS limit. It was found that this systematic increase in energy converged to the CBS limit.[10,11] Using energies obtained from a series of basis sets (e.g. double, triple, and quadruple- ζ quality sets), the energy of the system could be extrapolated to the CBS limit.

Using large basis sets and correlating the valence and subvalence electrons greatly increases the computational cost of using CCSD(T) in terms of memory, CPU, and time required. This is due to the high scaling of CCSD(T), as it scales as $O(n_{occ}^3 + n_{virt}^4)$, where n_{occ} is the number of occupied orbitals and n_{virt} is the number of virtual orbitals within a calculation. Even with a moderate-sized basis set, such as a triple- ζ quality set, the size of molecule that can be investigated can be restricted by this steep computational scaling. As such, there has been much work done to reduce the computational cost via developments such as explicitly correlated,[12] local,[13,14] and composite methods.

Composite methods, or model chemistries, combine modest levels of theory with larger basis sets and more robust theories with smaller basis sets to approximate the results obtained at a much higher level of theory, albeit one with a much larger computational cost.[15] Some commonly used composite methodologies include the Gaussian-*n* (*Gn*) methods,[15-19] Weizmann-*n* (*Wn*) methods,[20-22] the Feller-Peterson-Dixon (FPD) procedure,[3,23,24] HEAT method,[25] the multi-coefficient methods of Truhlar and coworkers,[26] and the correlation consistent Composite Approach (ccCA), developed by Wilson and coworkers.[27-36] ccCA includes extrapolation of energies to the CBS limit and the addition of core-valence electron interactions within the methodology to account for important effects that may not be computationally feasible to recover in a single point calculation.

The success of composite methods rests on the assumption that important electronic effects can be calculated and used as additive terms to a reference energy, which is the “additivity assumption” established and demonstrated by the authors of the *Gn* methods.[18,19,37] In principle, an additive approach approximates the energy that could be achieved by the best overall theoretical method and the largest basis set utilized in any of the steps of the composite approach. For example, within the G4 methodology, CCSD(T) is used in a single point calculation to recover dynamic correlation energy beyond the MP2 level of theory; all electrons are included within the correlation space in an MP2-level calculation to account for important core-valence electron interactions (indicated by Full); and the G3LargeXP basis set, which is a 6-311G(3*df*,2*p*) set with additional polarization functions, is used in a single point MP2(Full) calculation to help account for larger basis set effects not recovered when smaller basis sets are utilized. Because of this, the effective accuracy of the G4 methodology should be comparable to a single CCSD(T,Full)/G3LargeXP calculation. In fact, when used to determine enthalpies of formation (ΔH_f 's) for the 125 small main group molecules in the G2-1 molecule set, G4 theory achieves an absolute deviation of 0.65 kcal mol⁻¹ as compared to the 0.78 kcal mol⁻¹ absolute deviation achieved by the targeted method.[18] Other studies have also shown that G2 and G3 each perform as well as or better than the effective level of theory that the model chemistries are designed to emulate.[18,19,37]

It is important to verify the accuracy and utility of any method, especially a model chemistry that combines different levels of theory and basis sets. To gauge the utility of computational approaches, established sets of molecules/molecular properties such as the Gaussian molecule/property sets (G2-1, G2/97, G3/99, and G3/05)[16,38-40] have proven useful (see, e.g., Refs. 27, 28, 40). The sets include experimental energies (e.g., ionization energies, electron affinities, enthalpies of formation, and proton affinities) with experimental uncertainties of less than one kcal mol⁻¹ and represent a wide variety of elements and bonding types.[19]

These molecule sets have been used as a gauge for variants of ccCA in earlier work.[27-36,41] ccCA is routinely able to achieve a mean absolute deviation (MAD) of less than 1 kcal mol⁻¹, on average, in the calculation of energetic properties (i.e., ΔH_f 's, ionization potentials, and electron affinities) for the molecules in these sets. Further, ccCA has been utilized to study

nearly two thousand molecules, covering a large portion of the periodic table, and properties including pK_a 's,[36] reaction barriers,[42] and potential energy curves for ground and excited state species,[31,43-45] in addition to the energetic properties mentioned earlier.[46-50]

In principle the ccCA methodology was designed to emulate CCSD(T,FC1)/aug-cc-pCV ∞ Z-DK (where FC1 indicates that the subvalence shell of electrons are included within the correlation space) yet at a much reduced computational cost. As ccCA has been successful in energetic calculations, including the predictions of energetic properties that have yet to be measured experimentally,[48] the methodology is compared to the targeted accuracy to ensure that it is indeed comparable. In this study, ΔH_f 's from the G2/97 molecule set[38] were used to compare the performance of ccCA to energies produced at the target level of theory, CCSD(T,FC1)/aug-cc-pCV ∞ Z-DK. The G2/97 molecule set includes 29 radicals, 35 nonhydrogen systems, 22 hydrocarbons, 47 substituted hydrocarbons, and 15 inorganic hydrides, including atoms from H to Cl.

2 COMPUTATIONAL METHODS

All calculations were carried out using MOLPRO 2010.[51] The optimized geometries and frequencies of molecules in the G2/97 set were obtained at the B3LYP/cc-pVTZ level of theory. Frequencies were scaled by 0.989 to obtain the ZPE,[28] which was then used for both ccCA and CCSD(T)-calculated ΔH_f 's. The ccCA methodology was utilized as described in Ref. 28. The CCSD(T,FC1) energies were calculated at the aug-cc-pCV n Z-DK level of theory, with n = D, T, Q, and 5.[52-57] The second order spin-free Douglas-Kroll Hamiltonian was utilized to account for relativistic effects.[58-60]

Two of the extrapolation schemes that have been included within the ccCA methodology to extrapolate atomic and molecular energies to the CBS limit[28] are utilized in this study for ccCA and CCSD(T). As correlation energies converge at a slower rate than HF energies, the SCF and correlation energies are extrapolated separately. The two-point extrapolation of Feller has been determined to be effective in extrapolating HF energies, as suggested by Halkier, et al. (Eqn. 1):[10,61]

$$E(n) = E_{HF-CBS} + B \exp(-1.63n) \quad (1)$$

Correlation energies were extrapolated by a combination of two extrapolation schemes: the Peterson (P) and the Schwartz-3 (S3), referred to as PS3 (which has previously been shown to perform quite well within the ccCA methodology and is recommended for use within ccCA).[28] The scheme from Peterson and co-workers (P) is a mixed exponential/Gaussian formula (Eqn. 2):[11]

$$E(n) = E_{CBS} + B \exp[-(n-1)] + C \exp[-(n-1)^2] \quad (2)$$

In Eqn. 2, n indicates the ζ -level of the basis set, E_{CBS} is the electronic energy at the CBS limit, E_n is the electronic energy at the n th ζ -level, and B and C are fitting parameters. The S3 scheme was developed by Schwartz[62] and Helgaker and co-workers,[63] which is based on the cubic inverse power of the highest angular momentum included within the basis set (Eqn. 3):

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

In Eqn. 3, $l_{\max}$ is the highest angular momentum included in the basis set. E_{CBS} is the energy at the CBS limit and B is a fitting constant. Eqn. 2 has been shown to converge more rapidly and therefore underestimates the CBS limit, while Eqn. 3 converges more slowly and therefore overestimates the CBS limit.[28,34] As such, an average of these two schemes is utilized, as proposed by Peterson and coworkers.[64] This approach has been utilized many times [see, e.g. 4,28,65-69] The same extrapolations have been used for the CCSD(T) energies.

3 RESULTS

The ΔH_f 's for each molecule of the G2/97 set were calculated using ccCA and CCSD(T) and compared to experiment. The resulting mean absolute deviations (MADs), mean signed deviations (MSDs), and root-mean-squared deviations (RMSDs) are reported in Table 1 and shown graphically in Figure 1. In comparing the MADs obtained using CCSD(T) to the MADs obtained from ccCA, CCSD(T) results in an MAD of 0.90 kcal mol⁻¹, while ccCA results in an MAD of 0.80 kcal mol⁻¹. The MSD for CCSD(T) is 0.12 kcal mol⁻¹, indicating a slight bias towards overestimation of the calculated ΔH_f 's. The MSD for ccCA is -0.26 kcal mol⁻¹, indicating that ccCA slightly underestimates the ΔH_f 's, as has been shown in previous studies.[28,34]. It is possible that this underestimation of the experimental values results from use of the different extrapolation schemes. The Peterson extrapolation tends to converge more rapidly and therefore underestimates the CBS limit, while the Schwartz-3 scheme tends to converge more slowly and therefore overestimates the CBS limit. Using an average of these schemes cancels out these effects, though not entirely. In addition the ΔCC correction is the largest within the methodology and at the triple- ζ level, CCSD(T) underestimates the ΔH_f 's as has previously been shown.[3] Despite this, incorporation of CCSD(T) as the ΔCC term within the ccCA methodology does reduce the average absolute error.[70] ccCA improves upon the accuracy obtained by CCSD(T) by more than 1 kcal for 23 molecules within the G2/97 set. For example, the CCSD(T) ΔH_f for hydrogen peroxide (HOOH) is 5.09 kcal mol⁻¹ from experiment, yet the value computed from ccCA is only 0.60 kcal mol⁻¹ from experiment. A similar accuracy gain is observed for Si₂H₆, in which the CCSD(T) ΔH_f is 3.97 kcal mol⁻¹ from experiment, while the ccCA ΔH_f is 1.08 kcal mol⁻¹ from experiment. Conversely, there are only 12 molecules of the 147 molecules included within the G2/97 set for which CCSD(T) exceeds the accuracy obtained by ccCA by more than 1 kcal.

ccCA has a lower absolute deviation from experiment for nonhydrogens, inorganic hydrides, and radical species than does CCSD(T). However, for hydrocarbons and substituted hydrocarbons, CCSD(T) outperformed ccCA. The subset of the G2/97 set with the largest absolute deviation from experiment for ccCA was found to be the 22 hydrocarbons at 0.92 kcal mol⁻¹. (This finding corroborates a previous study that the calculated ΔH_f 's for hydrocarbons show the largest deviation from experiment when using ccCA.[50]) The smallest deviation of ccCA-calculated ΔH_f 's occurred for the subset of 29 radical molecules, with a deviation of 0.64 kcal mol⁻¹. The ccCA ΔH_f 's are underestimated for each subgroup of the molecule set, with the exception of the inorganic hydrides, for which the signed deviation is 0.55 kcal mol⁻¹. For the CCSD(T) ΔH_f 's, the subgroup with the largest deviation occurs for the 15 inorganic hydrides in the molecule set, with an absolute deviation of 1.26 kcal mol⁻¹, compared to 0.89 kcal mol⁻¹ resulting from ccCA. (Larger errors for the ΔH_f 's of inorganic molecules in the G2/97 molecule set when using CCSD(T) has been previously noted.[3]) Notably, the subgroup with the lowest absolute deviation for CCSD(T) occurs for the hydrocarbons at 0.68 kcal mol⁻¹. The absolute deviations for the substituted hydrocarbons is 0.75 and 0.76 kcal mol⁻¹ for CCSD(T) and ccCA, respectively.

To provide further comparison of ccCA with CCSD(T,FC1)/aug-cc-pCV ∞ Z-DK, the total time required to calculate the total energy for 1-chloropropane using ccCA is just 5% of the total time required for the CCSD(T) calculations. The percentage of the total calculation taken by each step is listed in Table 2. For this calculation, the computational bottleneck of ccCA is the MP2/aug-cc-pVQZ step, which represents 62% of the total time required for the calculation of the ccCA energy. The bottleneck for all of the CCSD(T) calculations is the aug-cc-pwCV5Z-DK calculation, which is 90% of the required calculation time needed to compute the steps to reach the CCSD(T,FC1)/aug-cc-pCV ∞ Z-DK energy. CCSD(T, FC1)/aug-cc-pCV n Z-DK calculations take so much more time because of the high computational scaling of the method and on the basis set, as previously discussed.. The CCSD(T)-computed enthalpy of formation is 1.21 kcal mol⁻¹ from experiment, while the ccCA-computed enthalpy of formation is within 0.36 kcal mol⁻¹ of experiment. Even for molecules such as AlF₃, furan, and pyridine, in which the CCSD(T) ΔH_f 's are up to 2 kcal mol⁻¹ closer to experiment than those calculated by ccCA, the ccCA calculations take 5%, 6%, and 1% of the time required for the single point CCSD(T) calculations, respectively.

4 CONCLUSIONS

Although CCSD(T) is able to predict energetic properties for molecules that are within 1 kcal mol⁻¹ of experiment, overall, for main group species, the high computational cost of the method can restrict its use to smaller molecules. The correlation consistent Composite Approach is, by construction, comparable in accuracy, on average, to CCSD(T,FC1)/aug-cc-pCV ∞ Z-DK, in which the valence and subvalence electrons are included within the correlation space and the

energies have been extrapolated to the CBS limit. ccCA was utilized to calculate the ΔH_f 's for the 147 molecules included within the G2/97 data set and compared to the target accuracy. Overall, ccCA-PS3 achieved an MAD of 0.80 kcal mol⁻¹ from experiment, while CCSD(T,FC1) with the PS3 extrapolation scheme achieved an MAD of 0.90 kcal mol⁻¹. The MAD between ccCA and CCSD(T,FC1)/aug-cc-pCV ∞ Z-DK is 0.87 kcal mol⁻¹, with an MSD of 0.37 kcal mol⁻¹, indicating that ccCA does result in energies that are near those of the targeted approach.

ACKNOWLEDGEMENTS

This material is based on work supported by the National Science Foundation under grant numbers CHE-1213874 and CHE-1362479. RW acknowledges support from the National Science Foundation GFRP under grant number DGE-1144248. Computing resources were provided by the Academic Computing Services at the University of North Texas. Support from the United States Department of Energy for the Center for Advanced Scientific Computing and Modeling (CASCaM) is acknowledged.

References

- [1] T.H. Dunning Jr. A Road Map for the Calculation of Molecular Binding Energies, *J. Phys. Chem. A*. 104 (2000) 9062-9080.
- [2] R.J. Bartlett, M. Musial. Coupled-cluster theory in quantum chemistry, *Rev. Mod. Phys.* 79 (2007) 291-352.
- [3] D. Feller, D.A. Dixon. Extended benchmark studies of coupled cluster theory through triple excitations, *J. Chem. Phys.* 115 (2001) 3484-3496.
- [4] D. Feller, K.A. Peterson. Probing the limits of accuracy in electronic structure calculations: Is theory capable of results uniformly better than "chemical accuracy"? *J. Chem. Phys.* 126 (2007) 114105.
- [5] K.A. Peterson, D. Feller, D.A. Dixon. Chemical accuracy in ab initio thermochemistry and spectroscopy: current strategies and future challenges, *Theor. Chem. Acc.* 131 (2012) 1079.
- [6] D. Feller, K.A. Peterson, B. Ruscic. Improved accuracy benchmarks of small molecules using correlation consistent basis sets, *Theor. Chem. Acc.* 133 (2014) 1407.
- [7] A.K. Wilson, T. van Mourik, T.H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. VI. Sextuple zeta correlation consistent basis sets for boron through neon, *J. Molec. Struct. (THEOCHEM)*. 388 (1996) 339-349.
- [8] R.A. Kendall, T.H. Dunning Jr., R.J. Harrison. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions, *J. Chem. Phys.* 96 (1992) 6796-6806.
- [9] A.K. Wilson, T.H. Dunning Jr. Benchmark calculations with correlated molecular wave functions. X. Comparison with "exact" MP2 calculations on Ne, HF, H₂O, and N₂, *J. Chem. Phys.* 106 (1997) 8718-8726.
- [10] D. Feller. Application of systematic sequences of wave functions to the water dimer, *J. Chem. Phys.* 96 (1992) 6104.
- [11] K.A. Peterson, D.E. Woon, T.H. Dunning Jr. Benchmark calculations with correlated molecular wavefunctions. IV. The classical barrier height of the H+H₂ ->H₂+H reaction, *J. Chem. Phys.* 100 (1994) 7410-7415.
- [12] T.B. Adler, G. Knizia, H. Werner. A simple and efficient CCSD(T)-F12 approximation, *J. Chem. Phys.* 127 (2007) 221106.
- [13] J. Almlöf. Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach, *Chem. Phys. Lett.* 181 (1991) 319-320.
- [14] A.K. Wilson, J. Almlöf. Møller-Plesset correlation energies in a localized orbital basis using a Laplace transform technique, *Theor. Chim. Acta.* 95 (1997) 49-62.
- [15] J.A. Pople, M. Head-Gordon, D.J. Fox. Gaussian-1 Theory: A general procedure for prediction of molecular energies, *J. Chem. Phys.* 90 (1989) 5622-5629.
- [16] L.A. Curtiss, K. Raghavachari, G.W. Trucks, J.A. Pople. Gaussian-2 theory for molecular energies of first- and second-row compounds, *J. Chem. Phys.* 94 (1991) 7221-7230.
- [17] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rassolov, J.A. Pople. Gaussian-3 (G3) theory for molecules containing first and second-row atoms, *J. Chem. Phys.* 109 (1998) 7764-7776.
- [18] L.A. Curtiss, P.C. Redfern, K. Raghavachari. Gaussian-4 theory, *J. Chem. Phys.* 126 (2007) 084108.
- [19] L.A. Curtiss, P.C. Redfern, K. Raghavachari. *Gn* theory, *WIREs Comput. Mol. Sci.* 1 (2011) 810-825.
- [20] J.M.L. Martin, G. de Oliveira. Towards standard methods for benchmark quality *ab initio* thermochemistry - W1 and W2 theory, *J. Chem. Phys.* 111 (1999) 1843-1856.

- [21] A.D. Boese, M. Oren, O. Atasoylu, J.M.L. Martin, M. Kallay, J. Gauss. W3 theory: Robust computational thermochemistry in the kJ/mol accuracy range, *J. Chem. Phys.* 120 (2004) 4129-4141.
- [22] A. Karton, E. Rabinovich, J.M.L. Martin, B. Ruscic. W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions, *J. Chem. Phys.* 125 (2006) 144108.
- [23] D.A. Dixon, D. Feller, G. Sandrone. Heats of Formation of Simple Perfluorinated Carbon Compounds, *J. Phys. Chem. A.* 103 (1999) 4744-4751.
- [24] D. Feller, D.A. Dixon. Predicting the Heats of Formation of Model Hydrocarbons up to Benzene, *J. Phys. Chem. A.* 104 (2000) 3048.
- [25] M.E. Harding, J. Vazquez, B. Ruscic, A.K. Wilson, J. Gauss, J.F. Stanton. High-accuracy extrapolated *ab initio* thermochemistry. III. Additional improvements and overview, *J. Chem. Phys.* 128 (2008) 114111.
- [26] Y. Zhao, B.J. Lynch, D.G. Truhlar. Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics, *Phys. Chem. Chem. Phys.* 7 (2005) 43-52.
- [27] N.J. DeYonker, T.R. Cundari, A.K. Wilson. The correlation consistent composite approach (ccCA): An alternative to the Gaussian-*n* methods, *J. Chem. Phys.* 124 (2006) 114104.
- [28] N.J. DeYonker, B.R. Wilson, A.W. Pierpont, T.R. Cundari, A.K. Wilson. Towards the intrinsic error of the correlation consistent Composite Approach (ccCA), *Mol. Phys.* 107 (2009) 1107-1121.
- [29] N.J. DeYonker, T.G. Williams, A.E. Imel, T.R. Cundari, A.K. Wilson. Accurate thermochemistry for transition metal complexes from first-principles calculations, *J. Chem. Phys.* 131 (2009) 024106.
- [30] B.P. Prascher, J.D. Lai, A.K. Wilson. The resolution of the identity approximation applied to the correlation consistent composite approach, *J. Chem. Phys.* 131 (2009) 044130.
- [31] B. Mintz, T.G. Williams, L. Howard, A.K. Wilson. Computation of potential energy surfaces with the multireference correlation consistent composite approach, *J. Chem. Phys.* 130 (2009) 234104.
- [32] S.R. Das, T.G. Williams, M.L. Drummond, A.K. Wilson. A QM/QM Multilayer Composite Methodology: The ONIOM Correlation Consistent Composite Approach (ONIOM-ccCA), *J. Phys. Chem. A.* 114 (2010) 9394-9397.
- [33] M.L. Laury, N.J. DeYonker, W. Jiang, A.K. Wilson. A pseudopotential-based composite method: The relativistic pseudopotential correlation consistent composite approach for molecules containing 4*d* transition metals (Y-Cd), *J. Chem. Phys.* 135 (2011) 214103.
- [34] T.G. Williams, N.J. DeYonker, B.S. Ho, A.K. Wilson. The correlation consistent Composite Approach: The spin contamination effect on an MP2-based composite methodology, *Chem. Phys. Lett.* 504 (2011) 88-94.
- [35] A. Mahler, A.K. Wilson. Explicitly Correlated Methods within the ccCA Methodology, *J. Chem. Theory Comput.* 9 (2013) 1402-1407.
- [36] A.G. Riojas, A.K. Wilson. Solv-ccCA: Implicit Solvation and the Correlation Consistent Composite Approach for the Determination of pK_a, *J. Chem. Theory Comput.* 10 (2014) 1500-1510.
- [37] L.A. Curtiss, J.E. Carpenter, K. Raghavachari, J.A. Pople. Validity of additivity approximations used in Gaussian-2 theory, *J. Chem. Phys.* 96 (1992) 9030-9034.

- [38] L.A. Curtiss, K. Raghavachari, P.C. Redfern, J.A. Pople. Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation, *J. Chem. Phys.* 106 (1997) 1063-1079.
- [39] L.A. Curtiss, K. Raghavachari, P.C. Redfern, J.A. Pople. Assessment of Gaussian-3 and density functional theories for a larger experimental test set, *J. Chem. Phys.* 112 (2000) 7374-7383.
- [40] L.A. Curtiss, P.C. Redfern, K. Raghavachari. Assessment of the Gaussian-3 and density-functional theories on the G3/05 test set of experimental energies, *J. Chem. Phys.* 123 (2005) 124107.
- [41] S.A. Nedd, N.J. DeYonker, A.K. Wilson, P. Piecuch, M.S. Gordon. Incorporating a completely renormalized coupled cluster approach into a composite method for thermodynamic properties and reaction paths, *J. Chem. Phys.* 136 (2012) 144109.
- [42] W. Jiang, C.C. Jeffrey, A.K. Wilson. Empirical Correction of Nondynamical Correlation Energy for Density Functionals, *J. Phys. Chem. A.* 116 (2012) 9969-9978.
- [43] G.A. Oyedepo, A.K. Wilson. Multireference Correlation Consistent Composite Approach [MR-ccCA]: Toward Accurate Prediction of the Energetics of Excited and Transition State Chemistry, *J. Phys. Chem. A.* 114 (2010) 8806-8816.
- [44] W. Jiang, A.K. Wilson. Multireference composite approaches for the accurate study of ground and excited electronic states: C₂, N₂, and O₂, *J. Chem. Phys.* 134 (2011) 034101.
- [45] C. South, G. Schoendorff, A.K. Wilson. MR-ccCA: A route for accurate ground and excited state potential energy curves and spectroscopic properties for third-row diatomic molecules, *Comput. Theor. Chem.* 1040-1041 (2014) 72-83.
- [46] T.G. Williams, A.K. Wilson. Performance of the correlation-consistent composite approach for sulfur species, *J. Sulfur Chem.* 29 (2008) 353-365.
- [47] Y. Gao, N.J. DeYonker, E.C. Garret III, A.K. Wilson, T.R. Cundari, P. Marshall. Enthalpy of Formation of the Cyclohexadienyl Radical and the C-H Bond Enthalpy of 1,4-Cyclohexadiene: An Experimental and Computational Re-evaluation, *J. Phys. Chem. A.* 113 (2009) 6955-6963.
- [48] K.R. Jorgensen, G.A. Oyedepo, A.K. Wilson. Highly energetic nitrogen species: Reliable energetics via the correlation consistent Composite Approach (ccCA), *J. Hazard. Mat.* 186 (2011) 583-589.
- [49] G.A. Oyedepo, C. Peterson, A.K. Wilson. Accurate predictions of the energetics of silicon compounds using the multireference correlation consistent composite approach, *J. Chem. Phys.* 135 (2011) 094103.
- [50] B.R. Wilson, N.J. DeYonker, A.K. Wilson. Prediction of Hydrocarbon Enthalpies of Formation by Various Thermochemical Schemes, *J. Comp. Chem.* 33 (2012) 2032-2042.
- [51] H.-J. Werner, P.J. Knowles, G. Knizia, F.R. Manby, M. Schutz, P. Celani, et al., *Molpro*, 2010.1.
- [52] T.H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, *J. Chem. Phys.* 90 (1989) 1007-1023.
- [53] D.E. Woon, T.H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon, *J. Chem. Phys.* 98 (1993) 1358-1371.
- [54] T.H. Dunning Jr., KA Peterson, AK Wilson. Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited, *J. Chem. Phys.* 114 (2001) 9244-9253.

- [55] K.A. Peterson, T.H. Dunning Jr. Accurate correlation consistent basis sets for molecular core-valence correlation effects: The second row atoms Al-Ar, and the first row atoms B-Ne revisited, *J. Chem. Phys.* 117 (2002) 10548-10560.
- [56] S. Yockel, A.K. Wilson. Core-valence correlation consistent basis sets for second-row atoms (Al-Ar) revisited, *Theor. Chem. Acc.* 120 (2008) 119-131.
- [57] W.A. de Jong, R.J. Harrison, D.A. Dixon. Parallel Douglas-Kroll energy and gradients in NWChem: Estimating scalar relativistic effects using Douglas-Kroll contracted basis sets, *J. Chem. Phys.* 114 (2001) 48-53.
- [58] M. Douglass, N.M. Kroll. Quantum electrodynamical corrections to the fine structure of helium, *Ann. Phys.* 82 (1974) 89.
- [59] B.A. Hess. Applicability of the no-pair equation with free-particle projection operators to atomic and molecular structure calculations, *Phys. Rev. A.* 32 (1985) 756-763.
- [60] B.A. Hess. Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators, *Phys. Rev. A.* 33 (1986) 3742-3748.
- [61] A. Halkier, T. Helgaker, P. Jørgenson, W. Klopper, J. Olsen. Basis-set convergence of the energy in molecular Hartree-Fock calculations, *Chem. Phys. Lett.* 302 (1999) 437-446.
- [62] C. Schwartz. Importance of Angular Correlations Between Atomic Electrons, *Phys. Rev.* 126 (1962) 1015-1019.
- [63] T. Helgaker, W. Klopper, H. Koch, J. Noga. Basis-set convergence of correlated calculations on water, *J. Chem. Phys.* 106 (1997) 9639-9646.
- [64] N.B. Balabanov, K.A. Peterson. Mercury and Reactive Halogens: The Thermochemistry of $\text{Hg} + \{\text{Cl}_2, \text{Br}_2, \text{BrCl}, \text{ClO}, \text{and BrO}\}$, *J. Phys. Chem. A.* 107 (2003) 7465-7470.
- [65] D. Feller, D.A. Dixon, J.S. Francisco. Coupled Cluster Theory Determination of the Heats of Formation of Combustion-Related Compounds: CO, HCO, CO₂, HCO₂, HOCO, HC(O)OH, and HC(O)OOH, *J. Phys. Chem. A.* 107 (2003) 1604-1617.
- [66] D. Feller, D.A. Dixon. Coupled Cluster Theory and Multireference Configuration Interaction Study of FO, F₂O, FO₂, and FOOH, *J. Phys. Chem. A.* 107 (2003) 9641-9651.
- [67] L. Pollack, T.L. Windus, W.A. de Jong, D.A. Dixon. Thermodynamic Properties of the C₅, C₆, and C₈ *n*-Alkanes from ab Initio Electronic Structure Theory, *J. Phys. Chem. A.* 109 (2005) 6934-6938.
- [68] D. Feller, K.A. Peterson, T.D. Crawford. Sources of error in electronic structure calculations on small chemical systems, *J. Chem. Phys.* 124 (2006) 054107.
- [69] T.G. Williams, N.J. DeYonker, A.K. Wilson. Hartree-Fock complete basis set limit properties for transition metal diatomics, *J. Chem. Phys.* 128 (2008) 044101.
- [70] N.J. DeYonker, T. Grimes, S. Yockel, A. Dinescu, B. Mintz, T.R. Cundari, A. K. Wilson. The correlation-consistent composite approach: Application to the G3/99 test set, *J. Chem. Phys.* 125 (2006) 104111.

Table 1 - Deviation from experiment for ccCA and CCSD(T), in kcal mol⁻¹, for the ΔH_f 's of the G2/97 molecule set. FC1 indicates that the valence and subvalence electrons were correlated. The CCSD(T)/CBS limit is estimated by extrapolation of the aug-cc-pCVnZ-DK basis sets, n = D, T, Q, 5.

	ccCA	CCSD(T, FC1)/CBS
<i>Mean signed deviation (MSD)</i>		
Enthalpies of formation (147)	-0.26	0.12
Nonhydrogens (34)	-0.31	0.07
Hydrocarbons (22)	-0.82	0.18
Substituted hydrocarbons (47)	-0.23	0.14
Inorganic hydrides (15)	0.55	0.52
Radicals (29)	-0.27	-0.08
<i>Mean absolute deviation (MAD)</i>		
Enthalpies of formation (147)	0.80	0.90
Nonhydrogens (34)	0.88	1.14
Hydrocarbons (22)	0.92	0.68
Substituted hydrocarbons (47)	0.76	0.75
Inorganic hydrides (15)	0.89	1.26
Radicals (29)	0.64	0.85
<i>Root-mean-square deviation (RMSD)</i>		
Enthalpies of formation (147)	1.08	1.30
Nonhydrogens (34)	1.17	1.48
Hydrocarbons (22)	1.24	0.90
Substituted hydrocarbons (47)	1.10	1.05
Inorganic hydrides (15)	0.96	2.12
Radicals (29)	0.83	1.08

Figure 1 - Mean absolute deviation (MAD) from experiment for the calculated ΔH_f° 's of the G2/97 molecule set, calculated with ccCA and CCSD(T,FC1)/CBS. FC1 indicates that the valence and subvalence electrons were included within the correlation space. The CBS limit for the CCSD(T,FC1) calculations is estimated by extrapolation of the aug-cc-pCV n Z-DK energies, where $n = D, T, Q, 5$.

(For black and white printing) Figure 1 - Mean absolute deviation (MAD) from experiment for the calculated ΔH_f 's of the G2/97 molecule set, calculated with ccCA and CCSD(T,FC1)/CBS. FC1 indicates that the valence and subvalence electrons were included within the correlation space. The CBS limit for the CCSD(T,FC1) calculations is estimated by extrapolation of the aug-cc-pCV n Z-DK energies, where $n = D, T, Q, 5$.

Table 2. Percentage of time required and number of basis functions for the energy calculation of 1-chloropropane, for each step of ccCA and CCSD(T,FC1)/CBS.

ccCA		
	%	Basis Fcns
MP2/aVDZ	0.2	164
MP2/aVTZ	3.3	354
MP2/aVQZ	72.0	651
MP2/VTZ	0.6	227
MP2/aCVTZ	7.6	418
MP2-DK/VTZ-DK	0.6	354
CCSD(T)/VTZ	15.7	227
CCSD(T)		
	%	Basis Fcns
CCSD(T,FC1)/aCVDZ-DK	0.03	180
CCSD(T,FC1)/aCVTZ-DK	0.8	418
CCSD(T,FC1)/aCVQZ-DK	9.3	788
CCSD(T,FC1)/aCV5Z-DK	89.8	1325