Computational and Theoretical Chemistry 1072 (2015) 58–62

Contents lists available at ScienceDirect

Computational and Theoretical Chemistry

journal homepage: www.elsevier.com/locate/comptc

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, TX 76203-5017, United States

ARTICLE INFO

Article history:
Received 24 June 2015
Received in revised form 20 August 2015
Accepted 22 August 2015
Available online 2 September 2015

Keywords:
Composite methods
ccCA
Coupled cluster
Extrapolation
Complete basis set limit
Convergence behavior

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 0.87 kcal mol^{-1}. The MAD of the ccCA enthalpies from the CCSD(T,FC1)/aug-cc-pCV∞Z-DK enthalpies is 0.84 kcal mol^{-1}.

© 2015 Published by Elsevier B.V.

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 N_2, 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 N_2 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 1*s* 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-pV*n*Z, 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 2*d* and 1*f* functions of the cc-pVTZ sets for boron through fluorine recover approximately the same amount of correlation energy, as do the 3*d*, 2*f*, and 1*g* 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

---

* Corresponding author. Tel.: +1 940 565 4372; fax: +1 940 565 4318.
E-mail address: akwilson@unt.edu (A.K. Wilson).

http://dx.doi.org/10.1016/j.comptc.2015.08.015
2210-271X/© 2015 Published by Elsevier B.V.

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* (G*n*) methods [15–19], Weizmann-*n* (W*n*) 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 G*n* 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 C CSD(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 for a total of 147 energies. (It should be noted that COF_2 has not been included in this data set, as it has been recommended for removal due to uncertainties associated with the experimental enthalpy of formation [39,52–54].)

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 [55–60]. The second order spin-free Douglas–Kroll Hamiltonian was utilized to account for relativistic effects [61–63].

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. (Eq. (1)) [10,64]:

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

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 (Eq. (2)) [11]:

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

In Eq. (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 nth ζ-level, and B and C are fitting parameters. The S3 scheme was developed by Schwartz [65] and Helgaker and co-workers [66], which is based on the cubic inverse power of the highest angular momentum included within the basis set (Eq. (3)):

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

In Eq. (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. Eq. (2) has been shown to converge more rapidly and therefore underestimates the CBS limit, while Eq. (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 Balabanov [67]. This approach has been utilized many times [see, e.g. 4,28,68–72] The same extrapolations have been used for the CCSD(T) energies.

Download English Version:

https://daneshyari.com/en/article/5393121

Download Persian Version:

https://daneshyari.com/article/5393121

Daneshyari.com