Benchmark Accuracy in Thermochemistry, Kinetics, and Noncovalent Interactions

Amir Karton*

School of Science and Technology, University of New England, Armidale, NSW 2351, Australia.

Abstract

Composite ab initio methods are amongst the most accurate quantum chemical theories used for exploring energetic and spectroscopic properties at the molecular and atomic levels. This field of quantum chemistry began with the development of the Gaussian-*n* methods by John Pople, Larry Curtiss, Krishnan Raghavachari, and co-workers in the late 1980s. Due to major advances in quantum chemical theory, composite methods have advanced to the level where they can predict challenging thermochemical and kinetic properties with confident sub-kcal/mol accuracy. In addition, due to advances in high-performance supercomputer technology, these theories are routinely used for examining chemical processes involving molecules with dozens of atoms. Consequently, composite ab initio methods are being widely used in many branches of chemistry. Here we will review the top end of these procedures, namely composite ab initio methods that are capable of achieving benchmark accuracy (i.e., confident sub-kJ/mol predictions) in thermochemistry, kinetics, and noncovalent interactions. This chapter reviews the accuracy and applicability of high-level composite ab initio methods. General recommendations are given for selecting suitable methods for a given property and levels of accuracy.

Keywords

Quantum chemistry, Computational thermochemistry, Composite ab initio methods, Coupled-cluster theory, CCSD(T), CCSDT(Q), CCSDTQ, Post-CCSD(T), Infinite basis set limit, Correlation-consistent basis sets, Core-valance, Scalar relativistic, Spin-orbit, Born–Oppenheimer, Chemical accuracy, Benchmark accuracy, Pople diagram.

*E-mail: amir.karton@une.edu.au

Cite as:
A. Karton. Benchmark Accuracy in Thermochemistry, Kinetics, and Noncovalent Interactions. Comprehensive Computational Chemistry, 1st Edition, Eds: R. J. Boyd and M. Yanez, Elsevier (2023).

Key points

- Composite ab initio methods are multistep theoretical procedures capable of confident sub-kcal/mol accuracy
- High-level composite ab initio methods are a subset of these procedures capable of confident sub-kJ/mol accuracy
- High-level composite methods explicitly include all energetic terms that contribute to the molecular binding energies at the sub-kJ/mol level
- These terms include post-CCSD(T), core-valence, relativistic, spin-orbit, Born–Oppenheimer, and zero-point vibrational energy corrections
- Additional terms may include conformational corrections for floppy molecules and tunneling corrections for reaction barrier heights

Glossary

- HF, Hartree–Fock
- SCF, self-consistent field
- CC, coupled cluster
- CCSD, CC with single and double excitations
- CCSD(T), CC with singles, doubles, and quasiperturbative triple excitations
- CCSDT, CC with single, double, and triple excitations
- CCSDT(Q), CC with singles, doubles, triples, and quasiperturbative quadruple excitations
- CCSDTQ, CC with single, double, triple, and quadrupole excitations
- CCSDTQ(5), CC with singles, doubles, triples, quadruples, and quasiperturbative quintuple excitations
- CCSDTQ5, CC with single, double, triple, quadrupole, and quintuple excitations
- ΔCCSD, CCSD – HF difference
- Δ(T), CCSD(T) – CCSD difference
- ΔT–(T), CCSDT – CCSD(T) difference
- Δ(Q), CCSDT(Q) – CCSDT difference
- ΔQ–(Q), CCSDTQ – CCSDT(Q) difference
- ΔQ, CCSDTQ – CCSDT difference
- Δ(5), CCSDTQ(5) – CCSDTQ difference
- Δ5–(5), CCSDTQ5 – CCSDTQ(5) difference
- Δ5, CCSDTQ5 – CCSDTQ difference
- Δ(6), CCSDTQ5(6) – CCSDTQ5 difference
- Δ6–(6), CCSDTQ56 – CCSDTQ5(6) difference
- Δ6, CCSDTQ56 – CCSDTQ5 difference
- FCI, full configuration interaction
- DFT, density functional theory
- DHDFT, double-hybrid density functional theory
- DBOC, diagonal Born–Oppenheimer correction
- ZPVE, zero-point vibrational energy
- ATcT, Active Thermochemical Tables
- MAD, mean-absolute deviation
- RMSD, root-mean-square deviation
- 95% CI, 95% confidence interval

Contents

1. Introduction

One of the primary and most important aims of quantum chemistry was to develop computational schemes to narrow the gap between theoretical predictions and experimental observations. A great deal of progress has been achieved over the past 25 years since the award of the chemistry Nobel Prize to John Pople for the development of quantum chemical computational methods. Today, quantum chemical composite ab initio methods can predict gas-phase thermochemical and kinetic properties with confident sub-kJ/mol accuracy relative to highly accurate experimental data. These chemical properties include heats of formation, ionization potentials, electron affinities, reaction energies, complexation energies, reaction barrier heights, and reaction rates. High-level composite ab initio methods are broadly defined here as methods that can obtain challenging thermochemical properties such as total atomization energies (TAEs) with sub-kJ/mol accuracy on average (and worst-case errors for problematic molecules < 1 kJ/mol). We note that this level of accuracy rivals or even surpasses that of many experiments.

In several areas of chemistry, it is essential to achieve this level of accuracy. One example is interstellar medium (ISM) chemistry, where the temperature of the gas can be as cold as 10 K. At this temperature, a change of 1 kJ mol^{-1} in a reaction barrier height corresponds to a change of five orders of magnitude in the reaction rate. Another example is the accurate modeling of kinetics and mechanisms of chemical reactions in complex chemical environments involving multiple competing reaction pathways. An Example of such gas-phase chemical environments includes atmospheric chemistry. Highly accurate theoretical chemical properties with confident sub-kJ/mol accuracy are also used for generating thermochemical and kinetic databases involving a diverse range of small-to-medium-sized molecules. Such databases are an important supplement to highly accurate experimental thermochemical databases such as the Active Thermochemical Tables

(ATcT) database of Ruscic and co-workers. For example, the W4-17 database is a set of 200 total atomization energies at 0 K (TAE_0), which have been obtained using near-exact theoretical electronic, relativistic, Born–Oppenheimer, and zero-point vibrational energies.[1,2,3,4] This database includes organic and inorganic species with up to eight non-hydrogen atoms, which cover a broad spectrum of bonding situations, electronic states, and multireference characters. The DBH24-08[5,6] and BH28[7] databases are examples of theoretical databases of highly accurate reaction barrier heights. These databases are essential for benchmarking the performance of lower-level quantum chemical methods, such as composite ab initio methods that approximate the CCSD(T) energy at various degrees of basis set completeness (e.g., CBS-QB3, G4, G4(MP2), W1, W1-F12, and ccCA) and perturbation theory methods such as double-hybrid DFT. Then these lower-level methods, which are capable of achieving chemical accuracy for a wide range of thermochemical and kinetic properties, are used for benchmarking computationally economical methods such as conventional DFT methods.[8,9]

A key difference between high-level composite ab initio methods (e.g., W4, W4-F12, and HEAT-456QP) and lower-level CCSD(T)-based composite ab initio methods, apart from their intrinsic accuracies and computational costs, is the high-level methods are more robust, i.e., applicable to any arbitrary first- and second-row system (including pathologically multireference molecules such as peroxides, halogen oxides, and carbon clusters).[1,2,3,10,11,12,13] Therefore, these methods can be considered as 'black box' composite ab initio methods. This point is particularly important for species that are not included in the ATcT database since it can be challenging to obtain highly accurate experimental data for new systems. Furthermore, high-level composite ab initio methods may be the only practical avenue to investigate rare, hypothetical, or transient

species (e.g., free radicals, reaction intermediates, and transition states) and properties that are not accessible by experiment.

2. Accuracy in Quantum Chemical Calculations

As mentioned above, a primary goal of composite ab initio methods is to obtain accurate thermochemical and kinetic properties with well-defined error bars that are directly comparable to experimental measurements. In this context, it is useful to define two arbitrary levels of accuracy that are commonly used in computational thermochemistry – 'chemical' and 'benchmark' accuracy.[3] Chemical accuracy refers to errors of up to 1.0 kcal mol^{-1} (or 4.2 kJ mol^{-1}), whereas benchmark accuracy refers to errors of up to 1.0 kJ mol^{-1} from sufficiently accurate experimental or theoretical data. As indicated here, chemical and benchmark accuracies refer only to cut-off errors but not to the way they are quantified, i.e., mean absolute deviations (MADs), root-mean-square deviations (RMSDs), or 95% confidence intervals (CIs).[14] Therefore, the terms chemical and benchmark accuracies should be used and interpreted with caution.

The performance of quantum chemical methods varies widely depending on the chemical property being considered (e.g., TAEs, reaction energies, reaction barrier heights, and nonbonded interactions) and chemical systems (e.g., multireference character, spin state, bond polarity, and formal oxidation state). In the context of high-level composite ab initio methods, chemical and benchmark accuracies are typically assessed against highly challenging chemical properties such as TAEs which involve breaking all chemical bonds in a molecule. In addition, a typical evaluation dataset would include systems dominated by a single reference configuration (e.g., CO_2, CH_3COOH, C_4H_{10}), systems that exhibit appreciable multireference character (e.g., H_2O_2, C_2N_2, N_2O_4), and pathologically multireference systems (e.g., C_2, O_3, F_2O_2). Evaluating performance in

this manner provides upper-bounds errors, and the performance for less challenging chemical systems and properties should be markedly improved. For example, higher levels of accuracy are expected for less challenging chemical properties that benefit from systematic error cancelation between reactants and products.[15,16] It is well established that the performance of quantum chemical methods improves along the following sequence of reactions:[2,15,17,18,19,20,21]

Atomization → Isogyric → Isodesmic → Hypohomodesmotic → Homodesmotic → Hyperhomodesmotic

The same trend is observed along a generalized formulation of this sequence, namely the connectivity-based hierarchical (CBH-*n*) reaction scheme:[16,18,22,23,24]

CBH-0 (atomization) → CBH-1 (atom-centric) → CBH-2 (bond-centric) → CBH-3 (higher bond-centric) → •••

Even with increasing levels of error cancellation between reactants and products, the above reactions are still challenging since they involve the breaking and forming of covalent bonds. Chemical transformations involving nonbonded interactions are less challenging for quantum chemical methods. These interactions may include hydrogen/halogen bonds and dispersion interactions. Examples of such properties are conformational isomerizations,[25,26,27,28] and complexation energies of nonbonded complexes.[29,30,31,32] For such properties, it is generally accepted that chemical accuracy refers to 0.1 kcal mol^{-1} and benchmark accuracy refers to 0.1 kJ mol^{-1}.[3,8,33]

3. Overview of Composite Ab Initio Methods

Single-point energy calculations carried out with standard ab initio methods such as DFT (B3LYP, PBE0, and ωB97X-D, and M06-2X), MP2, and CCSD(T) obtain the nonrelativistic electronic energy, which is not directly comparable to experimentally observable energies at finite temperatures. In contrast, composite ab initio methods are designed to obtain energetic and spectroscopic properties that are both directly comparable to experimental measurements and are associated with well-defined error bars. In addition, standard ab initio methods cannot be considered near-'black box' methods and may result in very large errors for challenging chemical properties (e.g., TAEs) and systems (e.g., molecules dominated by multireference effects). To understand how composite ab initio methods work, it is instructive to imagine a three-dimensional diagram, for which the axes are (i) the one-particle space, (ii) the *n*-particle space, and (iii) corrections to the nonrelativistic electronic energy. The one-particle and *n*-particle spaces represent the level of theory used for solving the nonrelativistic Schrödinger equation. Coupled-cluster (CC) theory provides a systematic road map for converging the *n*-particle space.[34,35] Similarly, the correlation consistent basis sets of Dunning and co-workers[36,37,38,39] in conjunction with basis set extrapolations, provide an effective approach for converging the *n*-particle space. Therefore, in practically all high-level composite ab initio methods, coupled-cluster theory and the correlation consistent basis sets are used for converging the one-particle and *n*-particle spaces, respectively. Figure 1 depicts a modified Pople diagram illustrating the convergence of these spaces. The third axis (blue arrow) represents any energetic contributions that are needed for converting the nonrelativistic electronic energy to experimentally observable energy. These corrections may include the core-valence, scalar relativistic, spin-orbit, Born–Oppenheimer, zero-point vibrational energy, enthalpic, and entropic corrections.

Different composite ab initio methods approximate different points on the Pople diagram. It is instructive to classify two types of composite ab initio methods:

- CCSD(T) methods
- Post-CCSD(T) methods (referred to here as high-level composite ab initio methods)

Since the CCSD(T) method provides a highly cost-effective approach for obtaining accurate electronic energies for systems that are dominated by a single-reference configuration, most composite ab initio methods belong to the first class of methods. Popular examples of CCSD(T)-based composite ab initio methods include the Gaussian-*n* (G*n*) methods[40,41,42,43,44] and variants thereof,[45,46,47,48,49,50,51,52] complete basis set (CBS) model chemistries,[53,54,55,56,57,58,59] lower-level Weizmann-*n* (W*n*)[60,61] and W*n*X methods,[62] and the correlation consistent composite approach (ccCA) methods.[63,64,65,66,67,68,69,70,71,72] The present review focuses on post-CCSD(T) composite ab initio methods that are capable of achieving benchmark accuracy. For in-depth overviews of composite ab initio methods, we refer the reader to several excellent reviews.[2,3,4,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91] For reviews of coupled-cluster theory, we refer the reader to references 92, 93, 94, and 95.

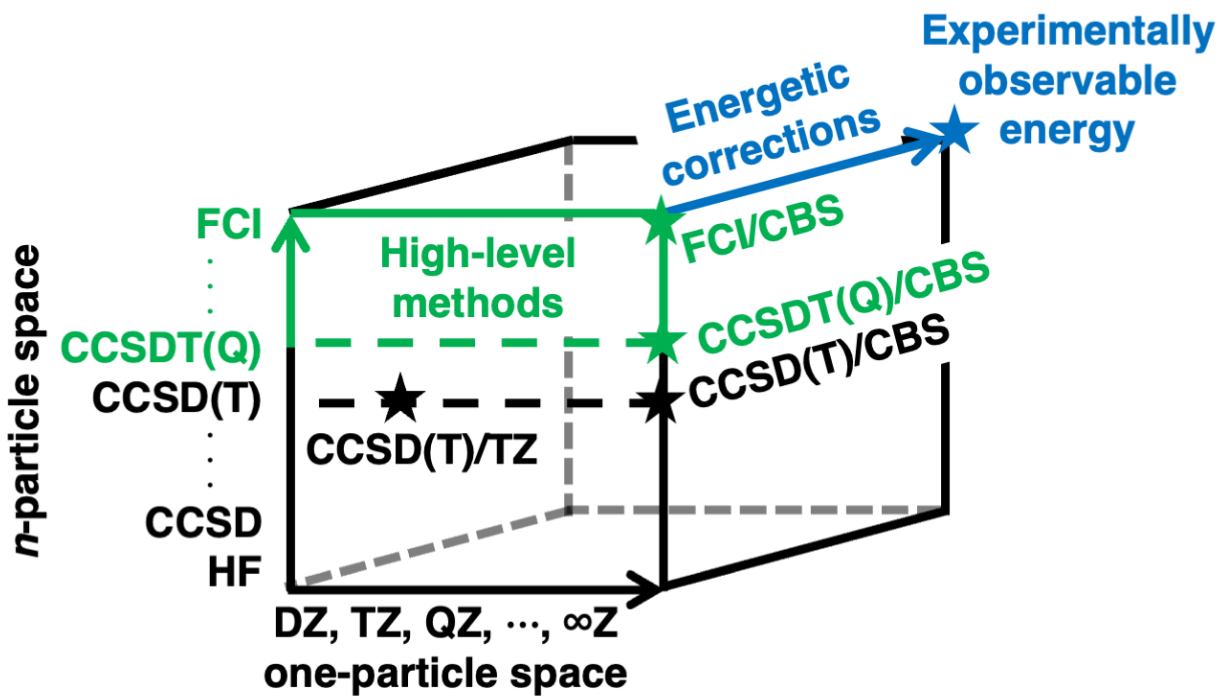

Figure 1. A three-dimensional Pople diagram illustrating the three axes defining composite ab initio methods, namely, the one-particle space, *n*-particle space, and any energetic corrections to the nonrelativistic electronic energy. The top part of the front pane of the Pople diagram (shown in green) represents the levels of theory approximated by high-level composite ab initio methods. The blue axis represents any additional energetic contributions that are needed for a meaningful comparison with experimentally observable properties (e.g., scalar relativistic, spin-orbit, Born–Oppenheimer, and zero-point vibrational energy corrections).

3.1. Computational Cost of Composite Ab Initio Methods

Before proceeding to an in-depth discussion of post-CCSD(T) composite ab initio methods capable of confident benchmark accuracy for challenging chemical systems and properties. It is important to emphasize that this level of accuracy and general applicability comes with a great cost in terms of computational resources and time. Table 1 gives an overview of the computational cost of coupled cluster methods. The computational cost increases by about an order of magnitude with each step along the series CCSD → CCSD(T) → CCSDT → CCSDT(Q) → CCSDTQ →

CCSDTQ(5) → CCSDTQ5. Therefore, whereas the CCSD(T) method is associated with formal scaling ~ N_{bas}^7, the CCSDT(Q) method, which is the lowest level of excitation involved in post-CCSD(T) methods, scales as ~ N_{bas}^9. Furthermore, the most accurate composite ab initio methods (such as W4.3 and W4.4 theories) include post-CCSD(T) contributions up to CCSDTQ56, which scales as ~ N_{bas}^{14}.

Table 1. Overview of formal computational scaling of coupled-cluster (CC) methods.

CC level	Cost[a]
CCSD	~ N_{bas}^6
CCSD(T)	~ N_{bas}^7
CCSDT	~ N_{bas}^8
CCSDT(Q)	~ N_{bas}^9
CCSDTQ	~ N_{bas}^{10}
CCSDTQ(5)	~ N_{bas}^{11}
CCSDTQ5	~ N_{bas}^{12}
CCSDTQ5(6)	~ N_{bas}^{13}
CCSDTQ56	~ N_{bas}^{14}

[a] N_{bas} is the number of basis functions. For the sake of simplicity, these scalings do not take into account the number of iterations needed to reach convergence of the coupled-cluster energy (for further details, see ref. 3).

When considering the computational cost of post-CCSD(T) composite ab initio methods, we must remember that successively higher-level coupled cluster methods converge increasingly faster with the basis set size.[2,3,75] Therefore, higher-order excitations require smaller basis sets to converge to the same level of accuracy as the lower-level excitations. For example, in the popular W4 theory, the largest basis set used for extrapolating the CCSD and CCSD(T) components are the aug-cc-pV(6+d)Z and aug-cc-pV(5+d)Z, respectively. However, it is sufficient to calculate the CCSDT(Q) and CCSDTQ energies with the cc-pVTZ and cc-pVDZ basis sets, respectively.[96,97] Note that the post-CCSD(T) contributions are also calculated without diffuse functions or high-exponent *d* functions on second-row elements.[98] To get a better idea of the computational cost of

these components, let us look at the computational cost of these methods for a representative small molecule. Table 2 gives the relative computational cost of the coupled-cluster methods in conjunction with the cc-pV*n*Z basis sets for *n*-pentane (C_5H_{12}). For simplicity, we consider the non-augmented cc-pV*n*Z basis sets, which for *n*-pentane involve 130 (cc-pVDZ), 318 (cc-pVTZ), 635 (cc-pVQZ), 1115 (cc-pV5Z), and 1792 (cc-pV6Z) basis functions. The largest calculations in W1 theory – a representative CCSD(T) composite ab initio method – are the CCSD/cc-pVQZ and CCSD(T)/cc-pVTZ calculations. These calculations have a similar relative computational scaling (of 10^4–10^5, Table 2). Indicating that W1 theory is balanced in terms of the computational cost of the CCSD and CCSD(T) components. Let us move on to a representative post-CCSD(T) composite ab initio method – W4 theory. The largest calculations in W4 theory are the CCSD/V6Z, CCSD(T)/V5Z, CCSDT(Q)/VTZ, and CCSDTQ/VDZ. Inspection of Table 2 shows that the CCSD(T) and post-CCSD(T) are associated with similar computational scaling once the size of the basis set used for each component is considered. Again, indicating that W4 theory is a balanced theory in terms of the computational cost of the CSD(T) and post-CCSD(T) steps. We note that since our estimated computational scalings here do not take into account the number of iterations needed to reach convergence of the coupled-cluster energy, the computational scalings of the full iterative methods (CCSD, CCSDT, CCSDTQ) are underestimated. In practice, the CCSDTQ/VDZ calculation turns out to be the computationally most demanding calculation in W4 theory.

Table 2. Relative formal computational scaling of coupled-cluster (CC) methods in conjunction with the cc-pV*n*Z (V*n*Z) basis sets for a representative small molecule – *n*-pentane (C_5H_{12}). The largest basis sets used for each component in a representative CCSD(T) composite ab initio method (W1 theory) are highlighted in green. The largest basis sets used for each component in a representative post-CCSD(T) composite ab initio method (W4 theory) are highlighted in yellow.[a]

CC/V*n*Z	VDZ	VTZ	VQZ	V5Z	V6Z
CCSD	1	2×10^{2}	1×10^{4}	4×10^{5}	7×10^{6}
CCSD(T)	1×10^{2}	7×10^{4}	9×10^{6}	4×10^{8}	1×10^{10}
CCSDT	2×10^{4}	2×10^{7}	5×10^{9}	5×10^{11}	
CCSDT(Q)	2×10^{6}	7×10^{9}	3×10^{12}		
CCSDTQ	3×10^{8}	2×10^{12}			
CCSDTQ(5)	4×10^{10}				

[a]Values are given relative to the estimated computational cost of the CCSD/VDZ calculation. The relative values are obtained using the estimated computational scalings in Table 1.

Table 2 demonstrates that the use of significantly smaller basis sets for the post-CCSD(T) calculations in high-level composite ab initio methods results in a relatively balanced computational cost across the CCSD(T) and post-CCSD(T) components. However, it also illustrates the exponential increase in computational cost when moving from a CCSD(T) composite method, such as W1 theory, to a post-CCSD(T) composite method, such as W4 theory. In practice, the largest molecules W4 theory has been applied to are small systems with up to ~5 non-hydrogen atoms (e.g., *n*-butane C_4H_{10}, acetic acid $HOClO_2$, cyanogen C_2N_2, tetrahedrane C_4H_4, CCl_4, SiF_4).[1,18] In contrast, W1 theory (or its successors W1w and W1-F12 theories) have been applied to systems as large as arginine ($C_6H_{14}N_4O_2$),[99] terphenyl ($C_{18}H_{14}$),[100] corannulene ($C_{20}H_{10}$),[101] sumanene ($C_{21}H_{12}$),[101] dodecahedrane ($C_{20}H_{20}$),[18] and carbon clusters (C_{20} and C_{24}).[102] In this context, it should be pointed out that W4 theory is a CCSDTQ5/CBS composite ab initio method that is applicable to small systems. However, post-CCSD(T) methods that calculate the CCSDT(Q)/CBS energy, such as W4lite, provide intermediate applicability between W1 and W4 theories. For example, CCSDT(Q)/CBS composite methods have been applied to systems with up

to ~10 non-hydrogen atoms, such as the SF_6^- anion,[103] C_2X_6 (X = F, Cl),[1] phosphorus sulfide cages (P_4S_4),[104] bullvalene ($C_{10}H_{10}$),[105] the octasulfur ring (S_8),[106] and even the benzene dimer.[107]

3.2. Accuracy of Composite Ab Initio Methods

A number of post-CCSD(T) composite ab initio methods have been developed over the past two decades.[2,3,75,78] These include the focal-point analysis (FPA),[108,109,110,111,112] Weizmann-*n* (W3, W3-F12, W4, W4-F12, W4.2, W4.3, and W4.4)[96,97,113,114] and W*n*X,[115,116] high-accuracy extrapolated ab initio thermochemistry (HEAT),[117,118,119,120,121] Feller–Peterson–Dixon (FPD),[90,91,122,123,124,125,126] and interference-corrected explicitly correlated second-order perturbation theory (INT-MP2-F12)[127] methods. As shown above, these methods are by far more computationally demanding than the CCSD(T) composite ab initio methods. However, they are capable of consistently better accuracy. This is clearly illustrated when looking at RMSDs obtained for TAEs. Figure 2 gives the RMSDs for CCSD(T) composite ab initio methods for a large and diverse set of 183 TAEs in the W4-17 database.[1] Most of the TAEs in the W4-17 database are calculated at the CCSDTQ5/CBS or CCSDTQ56/CBS levels of theory. Whereas for a small subset of larger molecules (e.g., C_5H_{12}, C_6H_6, CH_3COOH, $HClO_4$, and SF_6), the reference TAEs are calculated at the CCSDT(Q)/CBS level of theory. These reference values are associated with a 95% confidence interval well below 1 kJ mol^{-1} and are sufficiently accurate for benchmarking the performance of CCSD(T) composite ab initio methods.[3] In terms of chemical diversity, the W4-17 database includes organic and inorganic species involving single and multiple bonds with varying degrees of covalent and ionic characters, as well as radicals and triplet systems.[1] Figure 2 also gives the RMSDs for the post-CCSD(T) composite ab initio methods; however, due to the high accuracy of these methods, their performance cannot be assessed relative to the theoretical

TAEs in the W4-17 database. Instead, we use highly accurate experimental TAEs from the ATcT database associated with error bars ≤ 0.06 kcal mol^{-1} (for further discussion, see refs. 3 and 75).[128,129] Figure 2 shows that there is a clear improvement in performance when moving from the CCSD(T) to the post-CCSD(T) composite ab initio methods. In terms of RMSDs, the CCSD(T) methods achieve chemical accuracy; namely, the RMSDs range between 0.95 (G4) and 0.55 (W2-F12) kcal mol^{-1}. Whereas the post-CCSD(T) methods achieve benchmark accuracy; namely, the RMSDs range between 0.168 (W3-F12) and 0.060 (W4.*x*) kcal mol^{-1} (where W4.*x* refers to W4.2 theory or higher). It is important to note that with two exceptions (W3-F12 and W3.2 theories), all the post-CCSD(T) composite ab initio methods result in RMSDs ≤ 0.10 kcal mol^{-1}. These RMSDs translate to 95% confidence intervals lower than 1 kJ mol^{-1}.

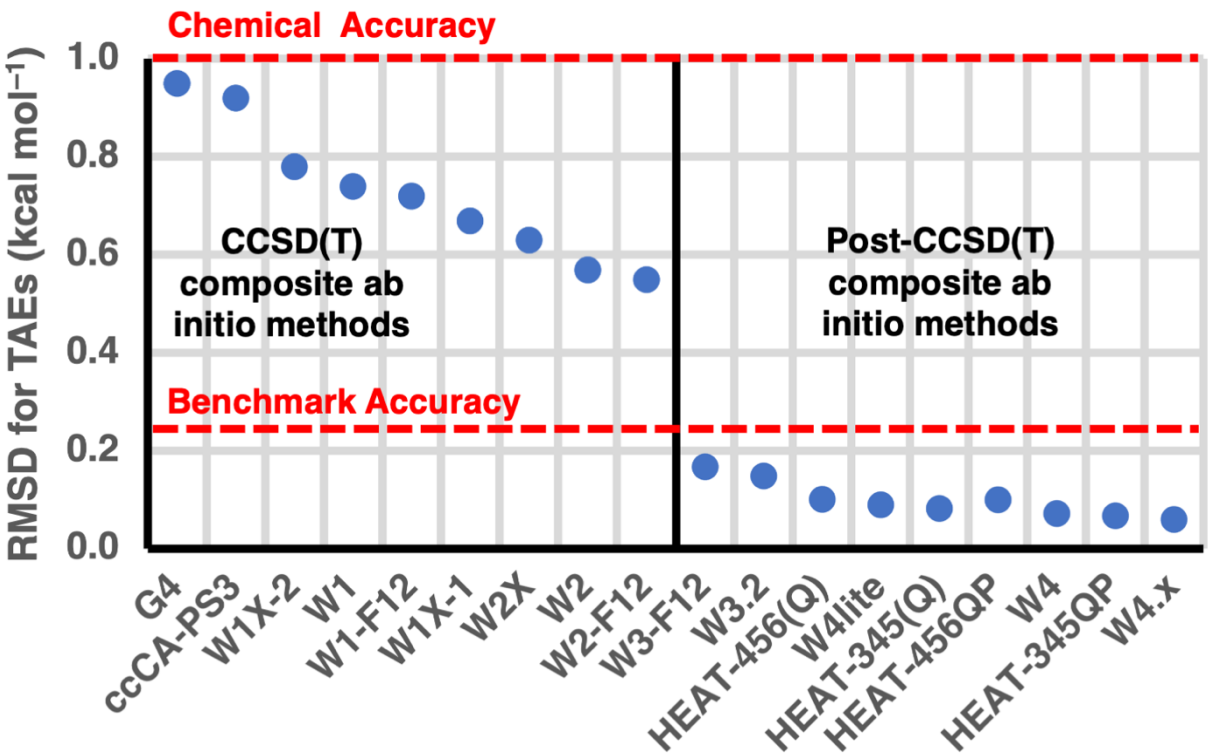

Figure 2. Root-mean-square deviations (RMSDs) in kcal mol^{-1} for a selected set of CCSD(T) and post-CCSD(T) composite ab initio methods for total atomization energies (TAEs). The performance of CCSD(T) composite ab initio methods is evaluated relative to the W4-17 database. The performance of post-CCSD(T) methods is evaluated relative to a smaller set of highly accurate experimental TAEs from ATcT (see main text and refs. 3 and 75 for further details).

4. Energy Components in High-Level Composite Ab Initio Methods

Composite ab initio methods rely on the following general equation to obtain energies ($E_{composite}$) that are directly comparable with experimentally observed energies

$$E_{composite} \approx E_{elect} + \Delta E_{CV} + \Delta E_{SR} + \Delta E_{SO} + \Delta E_{DBOC} + \Delta E_{ZPVE} + \Delta E_{HCF} + \cdots \quad (1)$$

Where E_{elect} is the nonrelativistic electronic energy, ΔE_{CV} is a core–valence correction, ΔE_{SR} is a scalar relativistic correction, ΔE_{SO} is a spin-orbit correction, ΔE_{BO} is a Born–Oppenheimer correction, ΔE_{ZPVE} is a zero-point vibrational energy correction, and ΔE_{HCF} is a heat content

function (or enthalpy function) correction. To relate this equation to the 3D Pople diagram in Figure 1, the nonrelativistic electronic energy is represented on the upper front face of the cube (green arrows), and all the secondary energetic correction terms (ΔE_{CV}, ΔE_{SR}, ΔE_{SO}, ΔE_{DBOC}, ΔE_{ZPVE}, ΔE_{HCF}, ...) are represented on the side face under the 'energetic corrections' arrow. In high-level composite ab initio methods, all these terms are calculated explicitly. Ideally, each term should be calculated to within benchmark accuracy so that the overall method does not rely on fortuitous error cancellation between the various components. The ellipsis at the end of Eq. 1 indicates any additional energetic component that affects the energetic property at hand to within benchmark accuracy. Depending on the chemical system and property, additional corrections may include conformational corrections to the enthalpy for floppy molecules,[99,130] and quantum mechanical tunneling (QMT) corrections for certain reaction barrier heights.[131] In the following sections, we will provide a comprehensive overview of the magnitude of the energetic terms in Eq. 1. For this purpose, we will use the W4-17 benchmark database,[1] which includes 200 TAEs at 0 K obtained via the W4lite, W4, and W4.*x* theories. In particular, for most systems in the W4-17 database (60%), the TAEs are obtained at the CCSDTQ5/CBS level via W4 and W4.2 theories, for a subset of smaller molecules (23%), the TAEs are obtained at the CCSDTQ56/CBS level via W4.3 and W4.4 theories, and for a subset of larger molecules (17%) such as benzene, SF_6, and C_2Cl_6 the TAEs are obtained at the CCSDT(Q)/CBS level via W4lite theory. Overall, the W4-17 dataset includes 200 organic and inorganic species with up to eight non-hydrogen atoms, which cover a broad spectrum of bonding situations, electronic states, and multireference characters.[1]

The nonrelativistic electronic energy (E_{elect}) is by far the largest energetic term, and therefore it serves as the base energy in Eq. 1. Typically, E_{elect} accounts for about 95% of the overall TAE at 0K. High-level composite ab initio methods can be divided into two subcategories:

- CCSDT(Q) methods
- Post-CCSDT(Q) methods

The post-CCSDT(Q) methods typically calculate the CCSDTQ5/CBS energy (e.g., W4 and HEAT-456QP theories). However, the CCSDTQ(5)/CBS and CCSDTQ5(6)/CBS energies are also frequently considered for slightly smaller or larger systems, respectively. As already mentioned, with current mainstream computer hardware, the CCSDT(Q)/CBS methods are applicable to systems with up to ~10 non-hydrogen atoms, whereas the post-CCSDT(Q) methods are typically applicable to systems with up to ~5 non-hydrogen atoms. The nonrelativistic electronic energy in these methods is calculated via the following equations:

$$E_{elect,CCSDT(Q)} = HF/CBS_{HF} + \Delta CCSD/CBS_{\Delta CCSD} + \Delta(T)/CBS_{\Delta(T)} + \Delta T\text{–}(T)/CBS_{\Delta T\text{–}(T)} + \Delta(Q)/CBS_{\Delta(Q)} \quad (2)$$

$$E_{elect,CCSDTQ5} = HF/CBS_{HF} + \Delta CCSD/CBS_{\Delta CCSD} + \Delta(T)/CBS_{\Delta(T)} + \Delta T\text{–}(T)/CBS_{\Delta T\text{–}(T)} + \Delta(Q)/CBS_{\Delta(Q)} + \Delta Q\text{–}(Q)/CBS_{\Delta Q\text{–}(Q)} + \Delta(5)/CBS_{\Delta(5)} + \Delta 5\text{–}(5)/CBS_{\Delta 5\text{–}(5)} \quad (3)$$

Here, all the energetic components are converged separately to the complete basis set limit and CBS_{comp} designates the basis set or basis set extrapolation used for each energetic component. The ΔCCSD, Δ(T), ΔT–(T), Δ(Q), ΔQ–(Q), Δ(5), and Δ5–(5) terms correspond to the correlation components; for example, ΔCCSD = CCSD – HF, Δ(T) = CCSD(T) – CCSD, ΔT–(T) = CCSDT –

CCSD(T), and so on (see Glossary for further details). The HF, ΔCCSD, Δ(T), and ΔT–(T) components are typically extrapolated to the CBS limit, whereas the Δ(Q) and higher-level terms are typically calculated with a single basis set. This partitioning of the nonrelativistic electronic energy is a highly efficient approach for obtaining the CCSDT(Q)/CBS and CCSDTQ5/CBS energies. The success of this approach relies on the significantly faster basis set convergence of successively higher-order coupled cluster expansion terms.[2,3,75]

4.1. Valence CCSD(T) Energy Components

Figure 3 gives an overview of the distribution of the HF component of the 200 TAEs in the W4-17 database. The HF component in the W4-17 database is extrapolated using the jul-cc-pV(5+d)Z and jul-cc-pV(6+d)Z basis sets.[132] For the HF component, the jul-cc-pV{5,6}Z basis set extrapolation provides near-exact results relative to numerical HF reference values.[96,133,134] For well-behaved systems that are not dominated by multireference effects, the HF energy is by far the largest contributor to the overall nonrelativistic electronic energy (Eqs. 2 and 3). The HF components in the W4-17 spread over a wide range of nearly 1300 kcal mol^{-1}, where the largest contribution of 1239.83 kcal mol^{-1} is observed for *n*-pentane. For homologous series, there will generally be a linear correlation between the magnitude of the HF component and the size of the system. This is illustrated for the homologous series of CH_n hydrocarbons (CH, CH_2, CH_3, and CH_4) and straight-chain alkanes (ethane, propane, *n*-butane, and *n*-pentane) (Figure 3b). However, this may not be true for more complex chemical systems where non-systematic multireference and bonding effects may play a role. For example, we obtain the following HF components for the oxoacids of chlorine 86.70 (HOCl), 59.07($HClO_2$), 67.58($HClO_3$), 80.52 ($HClO_4$) kcal mol^{-1}.

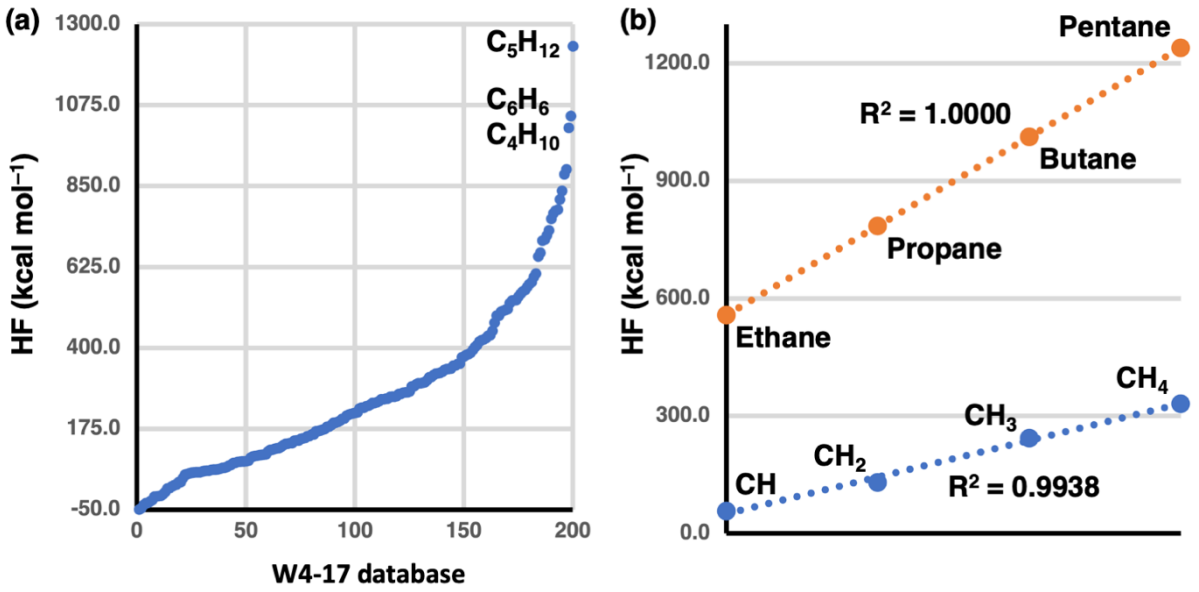

Figure 3. (a) Overview of the HF/CBS component for the 200 TAEs in the W4-17 database. (b) Overview of the HF/CBS component for the TAEs of two homologous series of hydrocarbons (in kcal mol^{-1}).

It is well known that pathologically multireference systems can be unbound at the SCF level.[2,3,75] Indeed, an inspection of Figure 3a reveals that a dozen systems in the W4-17 database are associated with a negative TAE at the Hatree–Fock level. These systems are (the HF component of the TAE is given in parentheses) FOOF (–48.79), O_3 (–45.09), ClF_5 (–36.36), F_2 (–31.08), FOF (–30.21), ClO_3 (–26.01), FOO (–22.48), ClOO (–13.00), OF (–12.23), BN $^1\Sigma^+$ (–11.35), OClO (–10.77), and ClF_3 (–5.50 kcal mol^{-1}). The magnitude of the negative HF component is related to both the multireference character of the system and to the strength of the bonds in the system. For example, F_2 and F_2O both have a similar TAE at the HF/CBS level (of about –30 kcal mol^{-1}); however, F_2 is more loosely bound, whereas F_2O has a stronger multireference character. We also note that the highly multireference system ClOOCl has a near-zero TAE of 0.19 kcal mol^{-1} at the HF/CBS level.

Figure 4 gives an overview of the distribution of the ΔCCSD correlation component for the 200 TAEs in the W4-17 database. The CCSD correlation component is extrapolated using the jul-cc-pV(5+d)Z and jul-cc-pV(6+d)Z basis sets.[96] For recent overviews of the basis set extrapolations used in high-level composite ab initio theories, see refs. 3 and 4. Whilst the ΔCCSD correlation energy exhibits a significantly slower basis set convergence than the HF component, the jul-cc-pV{5,6}Z basis set extrapolations are expected to result in errors well below the threshold of benchmark accuracy.[96,114,135,136,137] The ΔCCSD/CBS correlation components in the W4-17 still spread over a wide range of 316 kcal mol^{-1}, between 18.52 (AlH) and 334.44 (*n*-pentane) kcal mol^{-1}. We note that a very large ΔCCSD contribution of 313.27 kcal mol^{-1} is also obtained for dinitrogen tetroxide (N_2O_4). Again, it is interesting to examine the magnitude of the ΔCCSD/CBS correlation contribution for several homologous series that are included in the W4-17 database (Figure 4b). Here, we find a linear correlation with the size of the system for the homologous series of CH_n hydrocarbons (CH, CH_2, CH_3, and CH_4), straight-chain alkanes (ethane, propane, *n*-butane, and *n*-pentane), and oxoacids of chlorine (HOCl, $HClO_2$, $HClO_3$, $HClO_4$) (Figure 4b). Inspection of Figures 3 and 4 demonstrates that even for relatively small systems, the HF and ΔCCSD contributions to the TAEs typically amount to hundreds of kcal mol^{-1}. The mean HF and ΔCCSD contributions for the W4-17 dataset are 270.63 and 123.73 kcal mol^{-1}, respectively. For both components, the largest contribution is obtained for *n*-pentane, namely 1239.83 (HF) and 334.44 (ΔCCSD) kcal mol^{-1}. It is important to note that even a 0.1% error due to basis set incompleteness or the use of a nonoptimal basis set extrapolation scheme will translate into errors larger than 1 kJ mol^{-1} in the HF and ΔCCSD contributions. Therefore, in high-level composite ab initio methods, these two components have to be calculated with errors smaller than 0.01% in order to be

applicable to systems that most chemists would consider small molecules (e.g., *n*-butane, *n*-pentane, acetic acid, and benzene). This level of accuracy would have been unimaginable when composite ab initio methods were starting to be developed and illustrates how far this field has progressed over the past 35 years. Today, high-level composite ab initio theories, such as HEAT-345(Q), W4lite, and W4-F12, can reproduce the ATcT total atomization energy of benzene with sub-kJ/mol accuracy.[107,138] For a system with a TAE of nearly 5500 kJ mol^{-1}, this level of accuracy indeed corresponds to an error of about 0.01%.

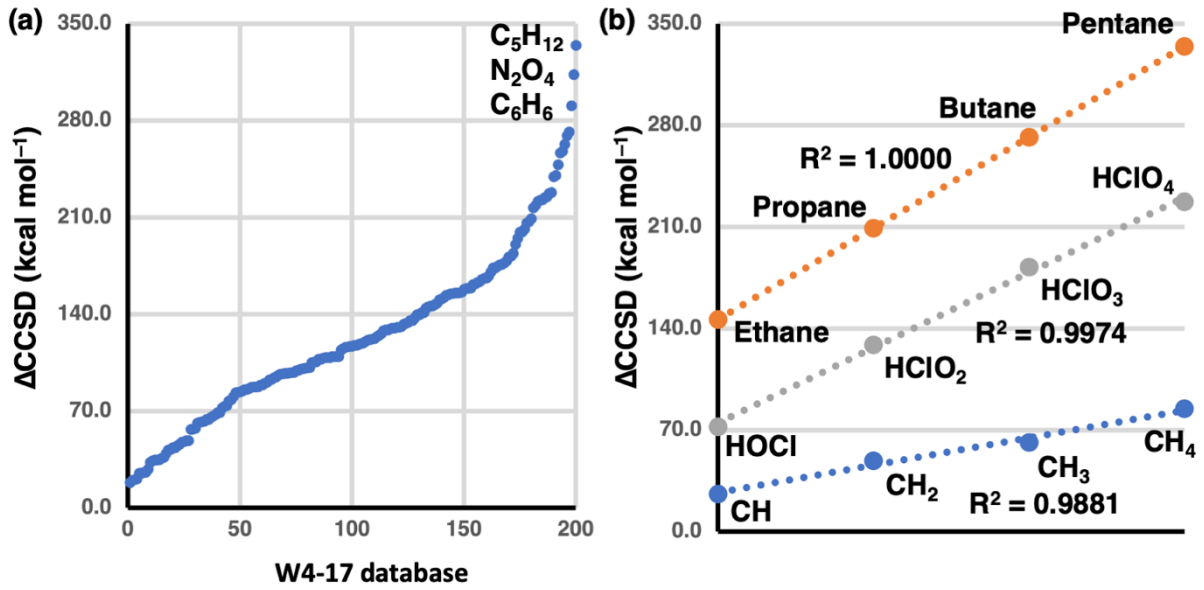

Figure 4. (a) Overview of the ΔCCSD/CBS correlation contribution for the 200 TAEs in the W4-17 database. (b) Overview of the ΔCCSD/CBS correlation contribution for the TAEs of three homologous series CH_n hydrocarbons, straight-chain alkanes, and oxoacids of chlorine (in kcal mol^{-1}).

Figure 5 gives an overview of the distribution of the Δ(T) correlation contribution for the 200 TAEs in the W4-17 database. The Δ(T) correlation energy exhibits a faster basis set convergence than the ΔCCSD correlation energy,[96,97,136] and in the W4-17 database it is

extrapolated from the jul-cc-pV(Q+d)Z and jul-cc-pV(5+d)Z basis sets. In a recent comprehensive study of the Δ(T) correlation component for a subset of the W4-17 database (namely the W4-08 subset), Martin has shown that the jul-cc-pV{5,6}Z basis set extrapolation convergences to within about 0.01 kcal mol^{-1} relative to near-infinite basis set limit reference values.[4] Generally speaking, the Δ(T) contributions to the TAEs are smaller by about an order of magnitude relative to the ΔCCSD correlation contributions. With the exception of dinitrogen tetroxide (N_2O_4), the Δ(T) contributions to the TAEs in the W4-17 database range between 0.07 (AlH) and 28.47 (S_4) kcal mol^{-1}. For N_2O_4, we obtain an unusually large Δ(T) contribution of 42.85 kcal mol^{-1} due to considerable nondynamical correlation effects and the size of the system. For the homologous series in the W4-F12 database, there is a general linear correlation between the magnitude of the Δ(T) component and the size of the system (Figure 5b). However, as we move to higher-level correlation effects, the degree of linear correlation becomes less pronounced. This is illustrated, for example, by the Δ(T) contribution to the TAE of CH_2 and CH_3 being practically the same, namely 1.9 kcal mol^{-1}, as well as by smaller squared correlation coefficients (R^2) values in Figure 5b for the CH_n hydrocarbons, alkanes, and oxoacids homologous series compared to Figure 4b.

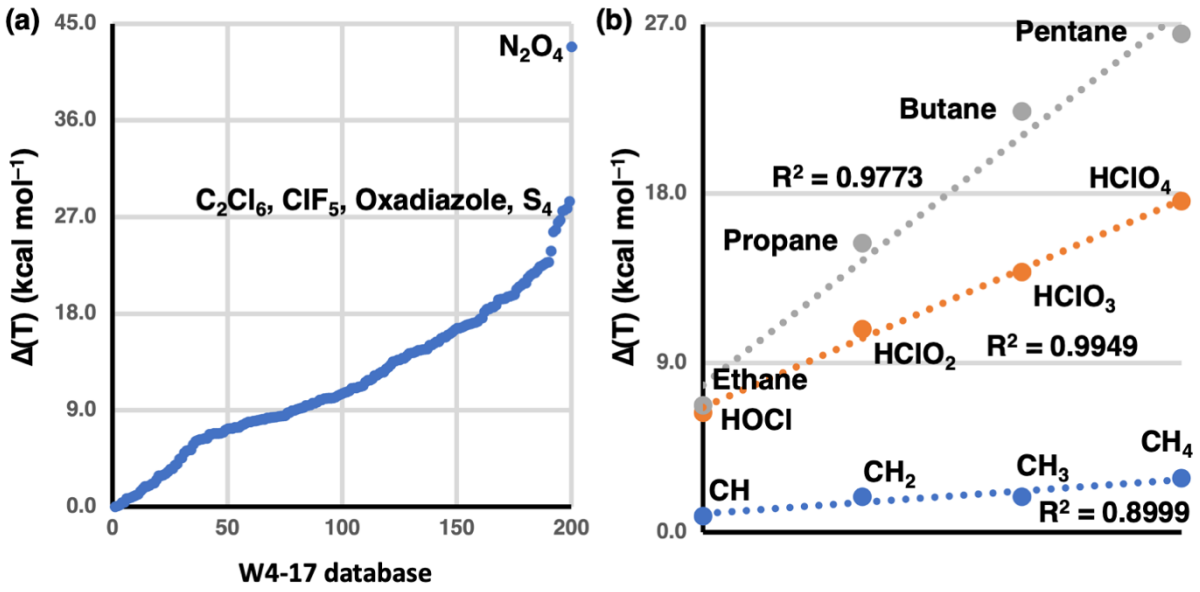

Figure 5. (a) Overview of the Δ(T)/CBS correlation contribution for the 200 TAEs in the W4-17 database. (b) Overview of the Δ(T)/CBS correlation contribution for the TAEs of three homologous series CH_n hydrocarbons, straight-chain alkanes, and oxoacids of chlorine (in kcal mol^{-1}).

4.2. Post-CCSD(T) Energy Components

Before proceeding to a discussion of the post-CCSD(T) contributions to the TAEs, we refer the reader to a comprehensive discussion of this topic in reference 3. Figure 6 gives an overview of the magnitude of the higher-order triples (ΔT–(T) = CCSDT – CCSD(T)) and connected quadruple (ΔQ = CCSDTQ – CCSDT) contributions to the TAEs in the W4-17 database. It is well established that the ΔT–(T) contribution tends to systematically reduce the TAEs, i.e., they are predominantly negative. Figure 6a shows that 94% of the ΔT–(T) contributions in the W4-17 are negative. We note that the positive contributions are very small; namely, they range between +0.02 (B_2H_6) and +0.28 (OF) kcal mol^{-1}. The negative ΔT–(T) contributions range between –0.01 (CH_2, NH_2, and HS) and –3.28 (P_4) kcal mol^{-1}, where contributions larger than 2.0 kcal mol^{-1} in absolute value are obtained for 7% of the systems. Particularly large negative contributions are obtained for

1,2,5-oxadiazole (–2.54), benzene (–2.62), BN (–2.63), S_4 (–2.78), N_2O_4 (–2.92), C_2Cl_6 (–3.07), and P_4 (–3.28 kcal mol^{-1}). The connected quadruple contributions, on the other hand, are all positive and range between +0.03 (CH_2 $^3B_1$) and +4.84 (S_4) kcal mol^{-1}. Contributions larger than 2.0 kcal mol^{-1} are obtained for 9% of the systems, where particularly large contributions are obtained for 1,2,5-oxadiazole (+2.50), P_4 (+2.57), FO_2 (+2.91), ClO_2 (+3.13), O_3 (+3.81), N_2O_4 (+4.21), and S_4 (+4.84 kcal mol^{-1}).

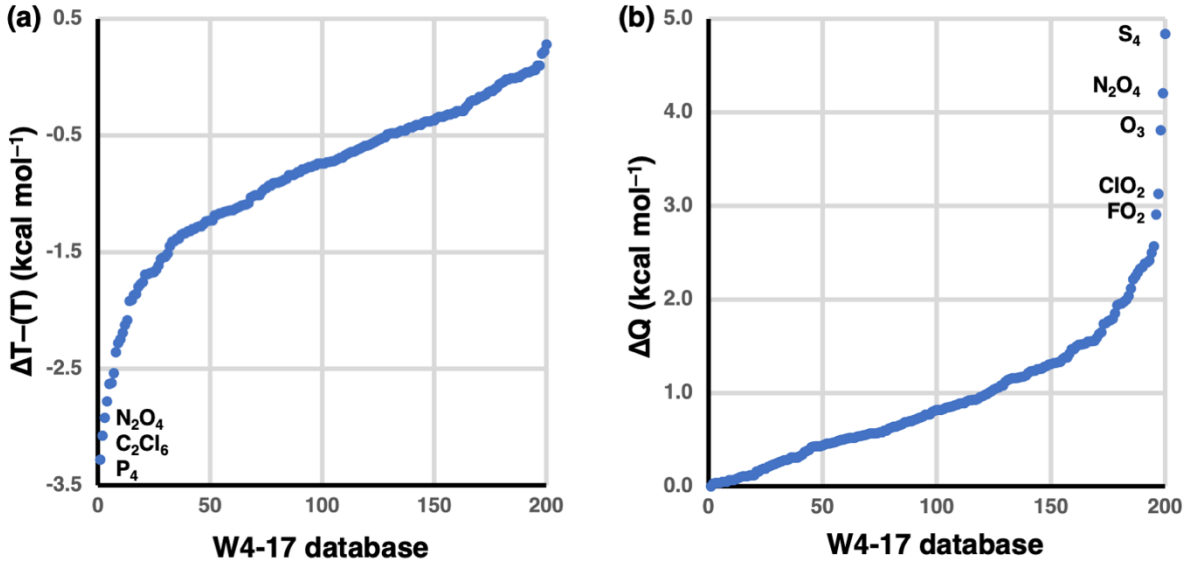

Figure 6. Overview of the (a) ΔT–(T) and (b) ΔQ correlation contributions for the 200 TAEs in the W4-17 database (in kcal mol^{-1}).

The effective cancellation between the negative ΔT–(T) contributions and positive quadruple ΔQ contributions leads to the excellent performance (relative to computational cost) of the CCSD(T)-based composite ab initio methods.[3,75,82,87,88,113,96,97,117,118,119,139,140,141,142,143] Indeed, the CCSD(T) method is one of the most successful yet computationally economical approaches in quantum chemistry and therefore is commonly known as the 'gold standard' of quantum

chemistry. Accordingly, as illustrated in Figure 2, many CCSD(T)-based composite ab initio methods achieve chemical accuracy in terms of RMSDs.

However, it is critical to emphasize that whilst the cancellation between the ΔT–(T) and ΔQ components is qualitative, it is not quantitative. Figure 7 depicts the ΔT–(T) and ΔQ contributions to the TAEs in the W4-17 database, where the ΔQ contributions are ordered by increasing value, and the ΔT–(T) contributions are ordered in the same order as the ΔQ contributions. Inspection of this figure illustrates that, overall, there is a relatively poor correlation between the magnitudes of the ΔT–(T) and ΔQ contributions. This is demonstrated by a squared correlation coefficient of $R^2 = 0.4688$ between the magnitude of the ΔT–(T) and ΔQ contributions (in absolute value) for the 200 systems in the W4-17 database.

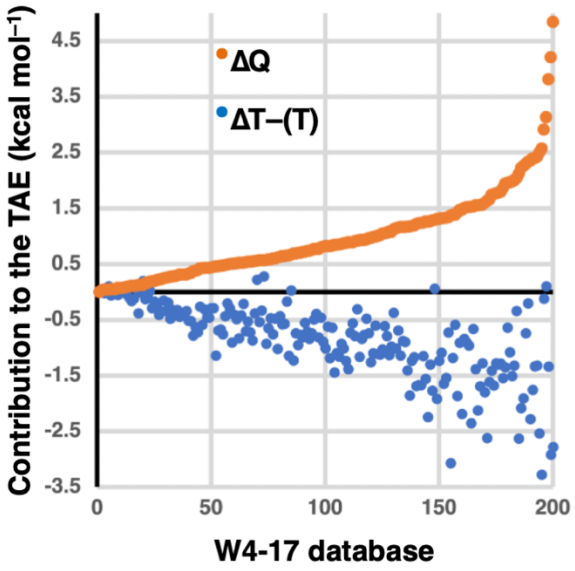

Figure 7. Overview of the (a) ΔT–(T) and (b) ΔQ correlation contributions to the 200 TAEs in the W4-17 database (in kcal mol^{-1}), where the ΔQ contributions (orange dots) are ordered by increasing value, and the ΔT–(T) contributions (blue dots) are ordered in the same order as the ΔQ contributions.

To illustrate the strength of CCSD(T)-based composite ab initio methods, Table 3 lists 23 selected systems for which both the ΔT–(T) and ΔQ contributions exceed 1 kcal mol^{-1} (in absolute value), yet the cancellation between these two contributions results in an overall contribution below 1 kJ mol^{-1}. In other words, Table 3 lists selected systems with large ΔT–(T) and ΔQ contributions for which the CCSD(T) approximation works remarkably well. It is interesting to see how diverse the systems in Table 3 are. This set includes organic systems with multiple polar bonds (e.g., acetic acid, glyoxal, and C_2F_4), highly strained organic systems (e.g., oxirene, dioxirane, 1,3-dioxetan-2-one, and cyclobutadiene), non-strained cyclic systems (e.g., furazan), inorganic systems with highly polar bonds where the central atom is in a high formal oxidation state (e.g., PF_5, SF_6, and $HClO_4$), and even systems with a considerable degree of multireference character (e.g., C_2 ($^1\sum_g^+$) and SO_3).

Table 3. Overview of the magnitude of the ΔT–(T) and ΔQ contributions to the TAEs for selected systems in the W4-17 database for which the CCSD(T) approximation works remarkably well (in kcal mol^{-1}).

Component	ΔT–(T)	ΔQ	Sum[a]
PF_5	–1.41	1.17	–0.24
C_2Cl_2	–1.54	1.33	–0.21
1,3-Dioxetan-2-one	–1.80	1.59	–0.20
SF_6	–1.69	1.56	–0.13
HC_2Cl	–1.11	1.01	–0.10
C_2F_2	–1.12	1.02	–0.10
Cl_2CO	–1.23	1.18	–0.05
Acetic acid	–1.13	1.08	–0.05
C_2F_4	–1.30	1.26	–0.04
Furazan	–2.54	2.50	–0.04
CO_2	–1.03	1.04	0.01
Oxirene	–1.17	1.19	0.02
HOCN	–1.14	1.16	0.02
Formic anhydride	–1.61	1.65	0.04
SO_3	–1.68	1.75	0.07
C_2 ($^1\Sigma_g^+$)	–2.28	2.35	0.07
Glyoxal	–1.23	1.31	0.08
ClCN	–1.15	1.24	0.09
HONC	–1.01	1.14	0.13
$HClO_4$	–2.08	2.22	0.14
Cyclobutadiene	–1.38	1.55	0.17
Dioxirane	–1.18	1.38	0.20
OCS	–1.09	1.32	0.23

[a] Sum of the ΔT–(T) and ΔQ contributions.

On the other hand, Table 4 lists selected systems for which there is poor (or no) cancellation between the ΔT–(T) and ΔQ contributions. For all the systems in Table 4, the sum of the ΔT–(T) and ΔQ contributions to the TAEs ranges between 0.8 and 3.2 kcal mol^{-1} (in absolute value). As expected, many of the systems in Table 4 are highly multireference, e.g., B_2, HO_3, O_3, FO_2, F_2O_2, N_2O_4, ClO_2, and S_4. Nevertheless, this subset includes systems that are not multireference by any stretch of the imagination, e.g., organic systems with multiple polar bonds (e.g., C_2Cl_6, C_2Cl_4), non-strained cyclic systems (e.g., benzene and thiophene), and ClF_5, which is an inorganic system

with highly polar bonds and a central atom in a formal high oxidation state. We also note that these systems are similar to those in Table 3, which makes it hard to a priori predict whether a CCSD(T) composite ab initio method would give a reliable TAE without explicitly calculating the post-CCSD(T) contributions.

Table 4. Overview of the magnitude of the ΔT–(T) and ΔQ contributions to the TAEs for selected systems in the W4-17 database for which the CCSD(T) approximation breaks (in kcal mol^{-1}).

Component	ΔT–(T)	ΔQ	Sum[a]
C_2Cl_6	–3.07	1.36	–1.71
Benzene	–2.62	1.63	–0.99
C_2Cl_4	–2.24	1.26	–0.99
Thiophene	–2.36	1.52	–0.83
OF	0.28	0.57	0.85
FNO	–0.66	1.53	0.87
NO_2	–1.08	1.96	0.88
ClF_5	–1.33	2.40	1.06
S_3	–1.32	2.42	1.10
N_2O_4	–2.92	4.21	1.29
ClNO	–0.64	1.95	1.32
B_2	0.06	1.29	1.35
FOOF	–0.74	2.29	1.55
cis-HO_3	–0.34	2.04	1.70
S_4	–2.78	4.84	2.06
trans-HO_3	–0.20	2.33	2.13
O_3	–1.34	3.81	2.47
FO_2	–0.12	2.91	2.79
ClOO	0.10	3.13	3.23

[a] Sum of the ΔT–(T) and ΔQ contributions.

The results in Tables 3 and 4 illustrate the challenges associated with choosing an appropriate composite ab initio method, i.e., a CCSD(T) or post-CCSD(T) method, without an a priori indicator for the importance of post-CCSD(T) contributions to the TAEs. Nevertheless, it has been found that the percentage of the TAE accounted for by the Δ(T) contribution relative to the CCSD(T) TAE provides a reliable diagnostic for the importance of post-CCSD(T)

contributions to the TAEs.[2,75,96] This useful diagnostic, which is often referred to as the %TAE[(T)] diagnostic, can be conveniently obtained from a CCSD(T) calculation in conjunction with a relatively small double-ζ basis set.[2] Therefore, it should be considered before running the extensive post-CCSD(T) calculations involved in high-level composite ab initio methods. The %TAE[(T)] diagnostic is defined as the percentage of the TAE accounted for by the Δ(T) contribution to the TAE, namely

$$\%\mathrm{TAE}[(\mathrm{T})] = \frac{\mathrm{TAE}[\mathrm{CCSD(T)}] - \mathrm{TAE}[\mathrm{CCSD}]}{\mathrm{TAE}[\mathrm{CCSD(T)}]} \tag{4}$$

where TAE[CCSD] and TAE[CCSD(T)] are the TAEs calculated at the CCSD and CCSD(T) levels, respectively. It has been found that there is a high statistical correlation between the %TAE[(T)] values and the percentage of the TAE accounted for by the ΔQ contribution (%TAE[ΔQ]). In particular, for the diverse set of 200 molecules in the W4-17 database, we obtain a squared correlation coefficient of $R^2 = 0.9300$ between the %TAE[(T)] and %TAE[ΔQ] values. An extensive analysis done for the W4-11 database of 140 TAEs suggests the following cut-off ranges for the %TAE[(T)] values[2,75]

- %TAE[(T)] below 5% suggests post-CCSD(T) contributions to the TAE should be below about 0.5 kcal mol^{-1}
- %TAE[(T)] between 5–10% suggests post-CCSD(T) contributions to the TAE should be below about 1 kcal mol^{-1}
- %TAE[(T)] above 10% suggests that post-CCSD(T) contributions to the TAE can certainly exceed 1 kcal mol^{-1}

Let us move to the Δ5 contributions to the TAEs. Most of the TAEs in the W4-17 database (167 out of 200) were obtained at the CCSDTQ5/CBS (or higher) level of theory via W4 and W4.*x* theories. The rest of the TAEs (33 out of 200) were obtained at the CCSDT(Q)/CBS level via W4lite theory. Figure 8 depicts the Δ5 correlation contributions to the 167 TAEs in the W4-17 database for which these contributions are available. The Δ5 contributions are practically all positive. We note however, that for a small number of halogenated systems (BF, CF, CHF_3, AlF, PF_3, CCl_2, CF_2Cl_2, $HClO_3$, and $AlCl_3$), contributions between 0.00 and –0.03 kcal mol^{-1} are obtained. The largest Δ5 contributions are obtained for highly multireference systems, namely 0.41 (O_3), 0.32 (C_2 ($^1\Sigma_g^+$)), 0.30 (S_4), 0.27 (F_2O_2), 0.24 (C_2N_2), 0.23 (FO_2), 0.22 (ClOO), and 0.21 (SO_3) kcal mol^{-1}. Yet, sizeable contributions between 0.1–0.2 kcal mol^{-1} are obtained for 25 molecules that span the gamut from systems dominated by moderate nondynamical correlation effects (e.g., CN, N_2, NO_2, P_2, and SO_2) to strong nondynamical correlation effects (e.g., BN, HO_3, P_4, S_3, and OClO). For all of these systems, explicitly calculating the Δ5 contributions is necessary in order to obtain confident benchmark accuracy.

Figure 8 depicts the Δ6 contributions to the TAEs for the 14 systems for which we obtain Δ6 contributions larger than 0.004 kcal mol^{-1}. As expected, these contributions are very small and can generally be neglected even for benchmark accuracy. The largest Δ6 contributions are obtained for the highly multireference systems BN (0.04) and C_2 ($^1\Sigma_g^+$) (0.07 kcal mol^{-1}).[13,144] Interestingly, a relatively high contribution of 0.03 kcal mol^{-1} is also obtained for P_2. Contributions beyond Δ6 can certainly be neglected. For example, the Δ7 contributions to the TAE of C_2 ($^1\Sigma_g^+$) is merely 0.003 kcal mol^{-1}.[13]

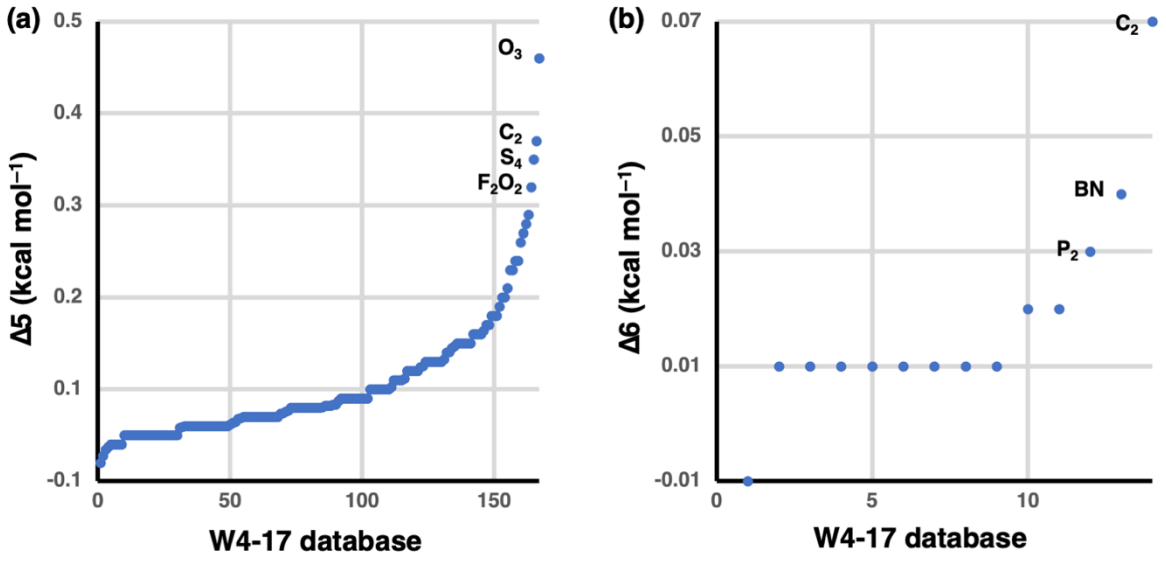

Figure 8. Overview of the (a) Δ5 and (b) Δ6 correlation contributions for the TAEs for the systems in the W4-17 database for which these contributions are available (in kcal mol^{-1}), Δ6 contributions are shown for the 14 systems for which these contributions are larger than 0.004 kcal mol^{-1}.

4.3. Secondary Energy Corrections

Let us turn our attention to the secondary energy corrections in high-level composite ab initio methods. In nearly all composite ab initio methods, the post-CCSD(T) calculations (e.g., CCSDT, CCSDT(Q), CCSDTQ, CCSDTQ(5), and CCSDTQ5) are carried out within the frozen-core approximation, i.e., with only the valence electrons being correlated. Composite ab initio methods that are generally applicable to second-row systems, such as the Weizmann-*n* theories, also employ the frozen-core approximation in the ΔCCSD and Δ(T) calculations. In contrast, in methods that were designed for treating mainly first-row systems, the CCSD and CCSD(T) calculations are carried out with the valence and core-valence electrons correlated. This distinction makes the W*n* theories applicable to relatively large second-row systems at a realistic computational cost. Notable examples include the calculation of the TAEs of P_4S_4 and S_8 at the CCSDT(Q)/CBS level of theory via W4lite theory,[104,106] and the TAE of P_4S_{10} at the CCSD(T)/CBS

level via W1-F12 theory.[104] The Weizmann-*n* methods include a core-valence (CV) correction to account for inner-shell correlation. The CV correction is calculated at the CCSD(T) level via the following equation

$$\Delta CV = CCSD(T)_{all}/CBS - CCSD(T)_{val}/CBS \quad (5)$$

In W*n* theories (*n* = 2–4) the ΔCV correction is extrapolated from the aug-cc-pWCV{T,Q}Z basis set pair. Calculating the CV correction at the CCSD(T)/aug-cc-pWCV{T,Q}Z level of theory provides confident sub-kJ/mol accuracy with an RMSD of merely 0.03 kcal mol^{-1} for the 200 molecules in the W4-17 dataset.[114,145] Figure 9a provides an overview of the CV correction to the TAEs in the W4-17 database. The ΔCV correction spans over a wide range from –0.67 (AlH_3) to +7.37 (benzene) kcal mol^{-1}. Negative ΔCV corrections are obtained for aluminum and silicon systems (e.g., AlH, AlH_3, $AlCl_3$, SiH_4, and Si_2H_6) as well as for fluorine systems (e.g., F_2, F_2O, F_2O_2, ClF_3, ClF_5, and SF_6). However, for most systems in the W4-17 database, the ΔCV correction is positive. Particularly large ΔCV corrections exceeding 5 kcal mol^{-1} are obtained for medium-sized organic systems such as tetrahedrane, furan, thiophene, pyrrole, borole, *n*-pentane, cyclopentadiene, and benzene. Notably, most of these systems do not contain second-row atoms, illustrating the importance of ΔCV corrections for first-row organic systems. Clearly, the ΔCV correction cannot be neglected for achieving chemical accuracy. Furthermore, to achieve benchmark accuracy, the ΔCV correction has to be treated at a sufficiently high level of theory (*vide supra*).[96,114,145,146]

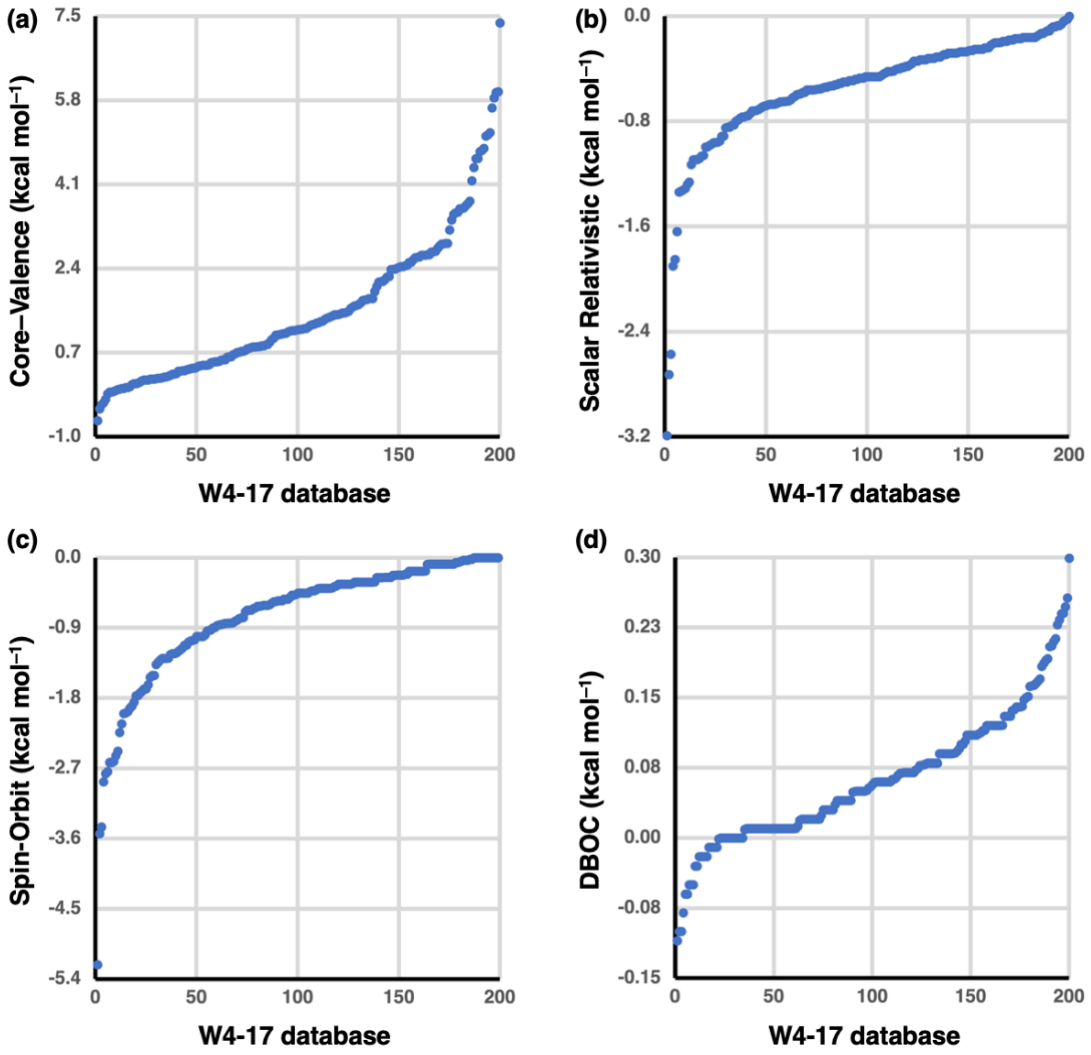

Figure 9. Overview of the (a) core-valence, (b) scalar relativistic, (c) spin-orbit, and (d) diagonal Born–Oppenheimer corrections for the 200 TAEs in the W4-17 database.

It is well established that relativistic effects can play a significant role in first-and second-row systems. Figure 9b depicts the scalar relativistic (ΔSR) corrections for the 200 molecules in the W4-17 database. These ΔSR corrections are calculated at the CCSD(T)/aug-cc-pVQZ-DK level using the second-order Douglas–Kroll–Hess (DKH) approximation which has been shown to yield results in close agreement with the full relativistic treatment for first- and second-row systems.[147,148,149,150] The ΔSR corrections to the TAEs range between –0.02 (BH) and –3.19 (SF_6)

kcal mol^{-1}. Particularly large ΔSR corrections ranging between 1.1 and 3.2 kcal mol^{-1} are obtained for halogenated hydrocarbons (e.g., C_2F_4, C_2F_6, $C_2F_2Cl_4$, C_2Cl_4, and C_2Cl_6), oxygenated and fluorinated systems with a central second-row atom in a high formal oxidation state (e.g., SO_3, ClO_3, $HClO_3$, $HClO_4$, AlF_3, SiF_4, PF_5, and SF_6), cyclic heteroatom organic systems (e.g., silole and thiophane), and for Si_2H_6. Figure 9c also provides an overview of the spin-orbit (ΔSO) corrections to the TAEs in the W4-17 database. High-level composite ab initio methods designed to treat mainly first- and second-row molecules typically include only first-order atomic and molecular ΔSO corrections. These corrections are universally negative and tend to be large for compounds with multiple oxygen, fluorine, and second-row atoms. For example, ΔSO corrections ranging between –2.0 and –3.0 kcal mol^{-1} are obtained for CF_2Cl_2, CCl_3H, C_2F_6, $C_2F_2Cl_2$, Cl_2O_2, $AlCl_3$, SF_6, ClF_5, and S_4. Yet, notably larger ΔSO corrections are obtained for perchlorinated organic compounds, namely –3.45 (CCl_4), –3.53 (C_2Cl_4), and –5.21 (C_2Cl_6) kcal mol^{-1}.

Figure 9d provides an overview of the deviations from the Born–Oppenheimer approximation, which are obtained via diagonal Born–Oppenheimer corrections (ΔDBOC). As can be seen, these corrections are much smaller than the ΔCV, ΔSR, and ΔSO corrections. For the 200 systems in the W4-17 database, the ΔDBOC contributions to the TAEs are calculated at the HF/aug-cc-pVTZ level of theory. However, it has been found that DBOC contributions obtained at the SCF level represent overestimations and that correlation contributions can reduce the HF DBOC contribution by up to 50%.[3,18,75,97,99,100,101,119,151,152] Therefore, it is recommended that ΔDBOC contributions calculated at the SCF level should be scaled by a factor of 0.5.[3] For the relatively small systems in the W4-17 database the ΔDBOC contributions range between –0.11 (CH) and +0.30 (*n*-pentane) at the HF/aug-cc-pVTZ level of theory. Therefore, these corrections cannot be neglected for achieving benchmark accuracy. However, we note that for larger hydrocarbons the

ΔDBOC contribution cannot be neglected even for achieving confident chemical accuracy for example, at the CCSD level, the following ΔDBOC contributions are obtained 0.32 (octahedrane, $(CH)_{12}$),[18] 0.33 (chrysene, $C_{18}H_{12}$),[100] 0.33 (meta-terphenyl, $C_{18}H_{14}$),[100] 0.36 (para-terphenyl, $C_{18}H_{14}$),[100] and 0.42 (dodecahedrane, $(CH)_{20}$) kcal mol^{-1}.[18] Since these ΔDBOC contributions are calculated at the CCSD level they represent highly accurate estimates and do not need to be scaled down.[3]

Summing up the CCSDT(Q)/CBS or CCSDTQ5/CBS nonrelativistic electronic TAE with the ΔCV, ΔSR, ΔSO, and ΔDBOC corrections to the TAE gives the all-electron, relativistic DBOC-inclusive TAE at the bottom of the well (TAE_e). The zero-point vibrational energy (ZPVE) is required to convert the TAE_e to a total atomization energy at 0 K (TAE_0) that can be compared with experiment. With the exception of a handful of diatomics in the W4-17 database, the ZPVE is the largest secondary energetic correction. Figure 10 provides an overview of the ZPVEs for the molecules in the W4-17 database. The ZPVEs span a wide range from 0.69 (AlCl) to 99.45 (*n*-pentane) kcal mol^{-1}. Calculating the ZPVE of the larger molecules in the W4-17 database (e.g., propane, *n*-butane, *n*-pentane, and benzene) with sub-kJ mol^{-1} accuracy can be one of the most challenging steps in high-level composite ab initio methods. For such systems, the calculation of harmonic frequencies at the CCSD(T)/cc-pVQZ level with cubic and quartic anharmonic corrections (e.g., at the MP2 or CCSD(T) levels) becomes prohibitively expansive.[25,75,99,151] In such cases, a computationally feasible, yet fairly reliable, approach is to scale the CCSD(T) or DHDFT harmonic ZPVE.[153,154,155,156,157,158,159,160,161,162,163] It has been shown that scaling of CCSD(T) or double-hybrid DFT harmonic frequencies can yield anharmonic ZPVEs with accuracies approaching those obtained from quartic force fields.[162]

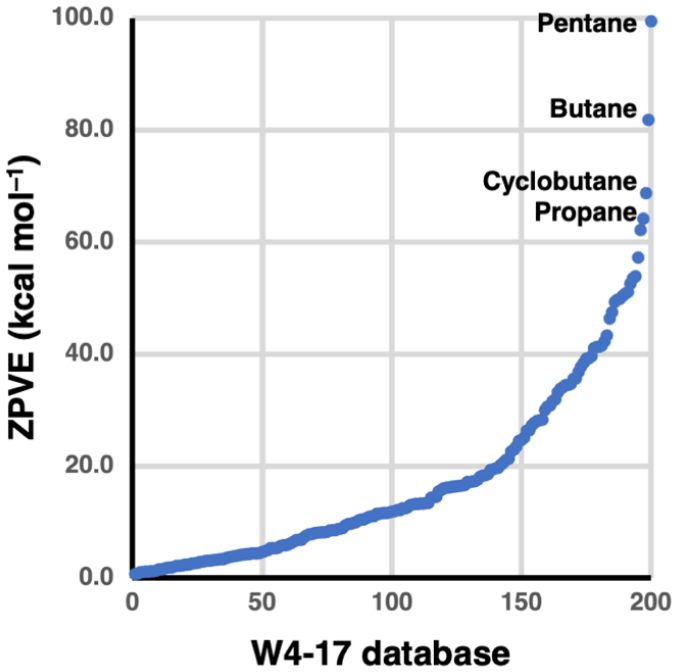

Figure 10. Overview of the zero-point vibrational energy (ZPVE) corrections for the 200 TAEs in the W4-17 database (in kcal mol^{-1}).

5. Putting It All Together for Thermochemistry, Kinetics, and Noncovalent Interactions

To summarize this chapter, it is useful to examine specific examples in which high-level composite ab initio methods have been used to obtain TAEs, reaction barrier heights, and binding energies of weakly bound complexes. In particular, we will consider the (i) TAEs of $C_2(^1\Sigma_g^+)$,[12,13] S_4,[1,2] and C_6H_6,[114] (ii) the reaction barrier heights for the Cope rearrangement in bullvalene[105,131] and for a prototypical cycloreversion reaction,[7] and (iii) the binding energies of the NO_2 and benzene dimers.[1,107]

Let us begin with dicarbon, which is among the most abundant molecules observed in stars, circumstellar shells, nebulae, comets, and the ISM.[164] For this small molecule, we can truly approach the exact relativistic solution to the Schrödinger equation. For the highly multireference

singlet ground $^1\Sigma_g^+$ state of C_2, high-level calculations of up to CCSDTQ567 are available in the literature.[12,13,82,125,165] The results from W4.3 (and higher) theories are summarized in Table 5. At the HF/aug-cc-pV{5,6}Z level of theory, the two carbon atoms are weakly bound with a dissociation energy of merely 18.38 kcal mol^{-1}. The ΔCCSD/aug-cc-pV{5,6}Z correlation component is nearly six times as large and amounts to 107.60 kcal mol^{-1}. The Δ(T)/aug-cc-pV{Q,5}Z correlation component further increases the dissociation energy by 19.37 kcal mol^{-1}. Overall, at the CCSD(T)/CBS level of theory, we obtain a dissociation energy of 145.35 kcal mol^{-1}. We note that this value is in excellent agreement with the infinite basis set limit CCSD(T)/aug-cc-pV{8,9}Z energy of 145.37 kcal mol^{-1} from Feller, Peterson, and Ruscic.[125] The higher-order triples, ΔT–(T), component, has been extrapolated to the infinite basis set limit from the cc-pV{6,7}Z basis sets and reduces the dissociation energy by as much as –2.27 kcal mol^{-1}.[13] In this case, we obtain the expected pattern of the Δ(T), Δ(Q), and Δ(5) components, all increasing the dissociation energy. Whereas the higher-order triples, quadruples, and quintuples components (ΔT–(T), ΔQ–(Q), and Δ5–(5)) all decrease the dissociation energy. Interestingly, the ΔT–(T), ΔQ–(Q), and Δ5–(5) components of –2.268, –1.151, and –0.053 kcal mol^{-1}, respectively, lie on a perfectly straight line with a squared correlation coefficient of R^2 = 0.99998. The sextuple contributions, Δ6, are still relevant for benchmark accuracy and amount to 0.063 kcal mol^{-1}, whereas the septuple contributions are an order of magnitude smaller and can be safely neglected (Table 5). Overall, the post-CCSDT contributions up to CCSDTQ567 amount to 2.69 kcal mol^{-1}. Therefore, they outweigh the ΔT–(T) component of –2.27 kcal mol^{-1}. The nonrelativistic electronic CCSDTQ567/CBS energy amounts to 145.78 kcal mol^{-1}. Moving to the secondary energetic corrections, at the W4.3 level, we obtain ΔCV = 1.25, ΔSR = –0.17, ΔSO = –0.17, ΔDBOC = 0.03, ZPVE = 2.64 kcal mol^{-1}. Summing up we obtain a CCSDTQ567/CBS dissociation energy at 0 K

(D_0) of 144.08 kcal mol^{-1}.[13] This value agrees within overlapping uncertainties with the experimental ATcT value of 144.01 ± 0.07 kcal mol^{-1}.[125]

Table 5. Overview of the energy components involved in high-level composite ab initio methods for the total atomization energies of three representative molecules C_2, S_4, and C_6H_6 (in kcal mol^{-1}).

Component	C_2 ($^1\Sigma_g^+$) CCSDTQ567/CBS[a]	S_4 CCSDTQ(5)/CBS[b]	C_6H_6 CCSDT(Q)/CBS[c]
HF	18.38	73.293	1045.01
ΔCCSD	107.60	129.531	290.71
Δ(T)	19.37	28.465	26.70
ΔT–(T)	–2.268	–2.782	–2.62
Δ(Q)	3.420	4.839[d]	2.13
ΔQ–(Q)	–1.151		N/A
Δ(5)	0.412	0.297	N/A
Δ5–(5)	–0.053	N/A	N/A
Δ6	0.063	N/A	N/A
Δ7	0.003	N/A	N/A
ΔCV	1.25	0.680	7.37
ΔSR	–0.17	–0.671	–0.99
ΔSO	–0.17	–2.239	–0.51
ΔDBOC	0.03	0.006	0.23
ZPVE	2.64	3.292	62.16
TAE_e	146.716	231.444	1368.10
TAE_0	144.076	228.152	1305.94
Expt. (ATcT)	144.01 ± 0.07	N/A	1305.9 ± 0.1
Theor.–Expt.	0.07	N/A	0.04

[a]The ΔT–(T), Δ(Q), ΔQ–(Q), Δ(5), Δ5–(5), Δ6, Δ7 components are taken from ref. 13, all other components are calculated via W4.3 theory and taken from ref. 12. [b]All components are calculated via W4 theory and are taken from ref. 2. [c]All components are calculated via W4-F12 theory and are taken from ref. 114. [d]This is the ΔQ = CCSDTQ – CCSDT contribution.

Let us move on to a larger system composed only of second-row atoms – tetrasulfur (S_4). The TAE of this highly multireference system has been obtained via W4 theory at the CCSDTQ(5)/CBS level of theory.[1,2] The HF/aug-cc-pV{5,6}Z, ΔCCSD/aug-cc-pV{5,6}Z, and Δ(T)/aug-cc-pV{Q,5}Z components to the TAE are 73.293, 129.531, and 28.465 kcal mol^{-1},

respectively. Resulting in a CCSD(T)/CBS TAE of 231.289 kcal mol^{-1}. The ΔT–(T) component amounts to as much as –2.782 kcal mol^{-1}. However, it is significantly outweighed by the ΔQ component of 4.839 kcal mol^{-1}. The Δ(5) component is still very significant and, on its own, exceeds 1 kJ mol^{-1}; namely, it amounts to 0.297 kcal mol^{-1}. Overall, the post-CCSD(T) contributions to the TAE are quite hefty, reaching 2.354 kcal mol^{-1}. We note in passing that other small molecules for which the post-CCSD(T) contributions to the TAE exceed 2 kcal mol^{-1} (or even 3 kcal mol^{-1}) include HO_3, O_3, FO_2, and ClO_2 (see ref. 3 for a comprehensive discussion). Overall, the nonrelativistic CCSDTQ(5)/CBS TAE amounts to 233.643 kcal mol^{-1}. For this system, the ΔCV, ΔSR, and ΔSO corrections are also quite hefty, albeit not as large as the post-CCSD(T) corrections. Namely they amount to ΔCV = 0.680, ΔSR = –0.671, ΔSO = –2.239 kcal mol^{-1}. Adding all the secondary energetic corrections, including the ΔDBOC = 0.03 and ZPVE = 2.64 kcal mol^{-1} contributions, we obtain a TAE_0 of 228.152 kcal mol^{-1} from W4 theory.

Let us move to a larger first-row organic system – benzene, for which TAEs have been obtained at the CCSDT(Q)/CBS level.[114,138] Table 5 shows the component breakdown of the W4-F12 TAE, however, we note that overall this value is practically identical to the earlier TAE obtained by Harding, Vázquez, Gauss, Stanton, and Kállay (HVGSK),[138] see ref. 114 for a comprehensive discussion. Benzene is not a highly multireference system, yet high-level composite ab initio methods are required for calculating its TAE within chemical accuracy (and certainly benchmark accuracy). At the valence CCSD(T)/CBS limit, we obtain a TAE of 1362.42 kcal mol^{-1} from W4-F12 theory.[114] The higher-order triples contribution amounts to –2.62 kcal mol^{-1} at the ΔT–(T)/cc-pV{D,T}Z level. This value is expected to be sufficiently accurate at the benchmark level, however, some caution is needed when treating the Δ(Q) component. In W4lite theory the Δ(Q) component is calculated in conjunction with the cc-pVDZ basis set and amounts

to 1.63 kcal mol^{-1}. However, as shown in ref. 114, when extrapolated from the cc-pV{D,T}Z basis sets, the magnitude of the Δ(Q) component increases to as much as 2.13 kcal mol^{-1}. Overall, the best post-CCSD(T) contribution amounts to –0.49 kcal mol^{-1}. We note that in contrast to C_2 and S_4 where post-CCSD(T) contributions increase the TAE, for benzene they decrease the TAE. The ΔCV correction amounts to as much as 7.37 kcal mol^{-1} at the CCSD(T)/aug-cc-pwCV{T,Q}Z level of theory. For comparison, calculating the ΔCV correction at the CCSD(T)/aug-cc-pwCVTZ level of theory (in a similar manner to W1 theory) results in ΔCV = 6.67 kcal mol^{-1}, which is lower by as much as 0.70 kcal mol^{-1} than the CBS value. Therefore, some caution is required when attempting to achieve chemical accuracy for the ΔCV component. The secondary energetic components from W4-F12 theory are ΔSR = –0.99, ΔSO = –0.51, ΔDBOC = 0.23, and ZPVE = 62.16 kcal mol^{-1}. Adding the best CCSD(T), post-CCSD(T), and secondary energetic corrections results in a TAE_0 of 1305.94 kcal mol^{-1}, which is practically spot on the ATcT value of 1305.9 ± 0.1 kcal mol^{-1}.

Let us move on to the calculation of reaction barrier heights. Table 6 gives the component breakdown of the CCSDT(Q)/CBS barrier heights for two challenging chemical reactions (i) the Cope rearrangement in bullvalene and (ii) the cycloreversion of the $C_2H_3NO_2$ heterocycle. Bullvalene ($C_{10}H_{10}$) is one of the most fascinating hydrocarbon cages that were ever synthesized since it spontaneously undergoes a series of rapid degenerate Cope rearrangements.[166,167,168,169] Consequently, at room temperature, the ^{1}H NMR spectrum of bullvalene exhibits a single broad peak, which becomes a single sharp peak at higher temperatures. Table 6 gives the component breakdown of the reaction barrier height for the Cope rearrangement in bullvalene from W3lite-F12 theory.[105,131] At the HF/CBS level, we obtain a strongly overestimated reaction barrier height of 25.84 kcal mol^{-1}. The ΔCCSD/CBS and Δ(T)/CBS contributions reduce the reaction barrier

height by –6.38 and –4.66 kcal mol^{-1}, respectively. Such that the valence CCSD(T)/CBS barrier height amounts to 14.80 kcal mol^{-1}. The ΔT–(T) and Δ(Q) contributions amount to +0.42 and –0.27 kcal mol^{-1}, respectively. Thus, overall the post-CCSD(T) contributions are relatively modest and increase the barrier height by 0.15 kcal mol^{-1}. The secondary energetic corrections amount to ΔCV = 0.32, ΔSR = –0.05, ΔDBOC = 0.02, ZPVE = –1.04, and H_{298}–H_0 = –0.14 kcal mol^{-1}. Overall, the classical relativistic CCSDT(Q)/CBS reaction barrier height at 298 K amounts to 14.08 kcal mol^{-1}. Importantly, for this rearrangement, a heavy atom tunneling correction amounts to as much as 0.22 kcal mol^{-1} at room temperature and has to be considered in the kJ/mol regime.[131] Overall, the best relativistic CCSDT(Q)/CBS reaction barrier height, including a QMT correction at 298 K, amounts to 13.85 kcal mol^{-1}. This highly accurate theoretical reaction barrier height suggests that the gas-phase experimental value of 13.2 kcal mol^{-1} should be revised upward. We note that these ^{1}H NMR measurements were conducted over 30 years ago over a relatively small temperature range, and therefore more elaborate NMR measurements may resolve the discrepancy between theory and experiment.

Table 6. Overview of the energy components involved in high-level composite ab initio methods for two representative reaction barrier heights (in kcal mol^{-1}).

Component	Cope rearrangement of bullvalene ($C_{10}H_{10}$) CCSDT(Q)/CBS[a]	Cycloreversion of $C_2H_3NO_2$[b] CCSDT(Q)/CBS[c]
HF/CBS	25.84	44.12
ΔCCSD	–6.38	5.40
Δ(T)	–4.66	–2.53
ΔT–(T)	0.42	–0.69
Δ(Q)	–0.27	–0.41
ΔCV	0.32	N/A[g]
ΔSR	–0.05	N/A[g]
ΔDBOC	0.02	N/A[g]
ZPVE	–1.04	N/A[g]
H_{298}–H_0	–0.14	N/A[g]
QMT[d]	0.22	N/A[g]
$\Delta E^{\ddagger}_{e,classical}$	15.26	45.91
$\Delta H^{\ddagger}_{298,classical}$	14.08	
$\Delta H^{\ddagger}_{298,tunneling}$	13.85	
Expt.	13.2[e]	
Theor.–Expt.	0.66[f]	

[a]All components are calculated via W*n* theories and are taken from refs. 105 and 131. [b]All components are calculated via W3lite-F12 theory and are taken from ref. 7. [c]See cycloreversion barrier height (CRBH) reaction number **1** in ref. 7. [d]Quantum mechanical tunneling (QMT) correction taken from ref. 131. [e]Taken from ref. 170. [f]Note that this experimental value is not sufficiently accurate for benchmarking high-level composite ab initio methods (see further discussion in refs. 105 and 131). [g]This component was not calculated in the original work.

Let us move to examine a reaction barrier height in which post-CCSD(T) contributions play a more significant role. Cycloreversion reactions (or, in the reverse direction, cycloaddition reactions) play an important role in organic synthesis since they provide a synthetic route for breaking down heterocyclic rings into small unsaturated building blocks. Table 6 gives the components breakdown of the CCSDT(Q)/CBS reaction barrier height for a typical cycloreversion reaction in which a $C_2H_3NO_2$ heterocycle is broken down into formaldehyde and H–N=C=O

fragments (for further details, see reaction CRBH **1** in ref. 7, which is the same as reaction **1** in ref. 171). Here, we only examine the nonrelativistic, valence CCSDT(Q)/CBS reaction barrier height (Table 6). At the HF/CBS level, we obtain a reaction barrier height of 44.12 kcal mol^{-1}. The ΔCCSD/CBS and Δ(T)/CBS correlation components amount to 5.40 and –2.53 kcal mol^{-1}, respectively. Resulting in a valence CCSD(T)/CBS reaction barrier height of 46.99 kcal mol^{-1}. In a similar manner to TAEs, for most reaction barrier heights, the ΔT–(T) and Δ(Q) components have opposite signs and tend to cancel each other out. However, it has been found in ref. 7 that for cycloreversion reactions, both the ΔT–(T) and Δ(Q) components systematically reduce the reaction barrier heights. For the reaction at hand, we obtain T–(T) = –0.69 and Δ(Q) = –0.41 kcal mol^{-1}. Thus, overall the post-CCSD(T) contributions reduce the reaction barrier height by 1.1 kcal mol^{-1}. This demonstrates that post-CCSD(T) contributions are required even for chemical accuracy.

As a final example, let us consider two weakly interacting dimers (i) the nitrogen dioxide dimer (O_2N•••NO_2),[1] and (ii) the benzene parallel displaced dimer (C_6H_6•••C_6H_6).[107] Table 7 gives the component breakdown of these two binding energies. The O_2N•••NO_2 binding energy has been obtained at the CCSDT(Q)/CBS level of theory via W4lite theory. Dinitrogen tetroxide has an unusually long N–N bond calculated to be 1.7346 Å at the CCSD(T)/cc-pVQZ level of theory. For comparison, the length of the single N–N bond in hydrazine is calculated to be 1.4383 Å at the same level of theory. Here, we will examine the dissociation energy $N_2O_4 \rightarrow 2NO_2$. At the HF/aug-cc-pV{5,6}Z level, the two NO_2 molecules are unbound with an interaction energy of –6.24 kcal mol^{-1}. The ΔCCSD/aug-cc-pV{5,6}Z component increases the binding energy by 18.95 kcal mol^{-1}, and the Δ(T)/aug-cc-pV{Q,5}Z component further increases the binding energy by 4.12 kcal mol^{-1}. At the valence CCSD(T)/CBS level of theory, we obtain a binding energy of 16.83 kcal mol^{-1}. The ΔT–(T)/cc-pV{D,T}Z component reduces the binding energy by –0.75 kcal mol^{-1} and

the Δ(Q)/cc-pVDZ component increases it by a relatively small amount of 0.10 kcal mol^{-1}. Overall, post-CCSD(T) contributions to the binding energy are fairly significant and amount to –0.65 kcal mol^{-1}. The ΔCV, ΔSR, and ΔDBOC corrections tend to cancel each other out and collectively amount to –0.01 kcal mol^{-1}. The ZPVE contribution reduces the dissociation energy by as much as –3.74 kcal mol^{-1}. Overall, we obtain a relativistic CCSDT(Q)/CBS binding energy of 12.42 kcal mol^{-1}, which is lower by 0.38 kcal mol^{-1} than the ATcT value of 12.80 ± 0.05 kcal mol^{-1} (Table 7). Given that N_2O_4 is a highly multireference system (*vide supra*), the discrepancy between theory and experiment is mainly attributed to the basis set incompleteness of the Δ(Q)/cc-pVDZ component and the complete neglect of the higher-order quadruple excitations (ΔQ–(Q)).

Table 7. Overview of the energy components involved in high-level composite ab initio methods for the binding energies (BEs) of two weakly bound dimers (in kcal mol^{-1}).

Component	O_2N•••NO_2 CCSDT(Q)/CBS[a]	C_6H_6•••C_6H_6[b] CCSDT(Q)/CBS[c]
HF/CBS	–6.24	–4.510
ΔCCSD	18.95	5.797
Δ(T)	4.12	1.303
ΔT–(T)	–0.75	–0.246
Δ(Q)	0.10	0.187
ΔCV	0.07	0.027
ΔSR	–0.13	0.003
ΔDBOC	0.05	–0.001
ZPVE	–3.74	–0.240
BE_e	16.16	2.550
BE_0	12.42	2.310
Expt.	12.80 ± 0.05[d]	2.4 ± 0.4[e]
Theor.–Expt.	–0.38	–0.09

[a]All components are calculated via W4lite theory and are taken from ref. 1. [b]All components are taken from ref. 107. [c]This component was not calculated in the original work. [d]Taken from ATcT, ref. 172. [e]Taken from ref. 173.

Finally, let us consider the benzene dimer.[107] Similar to the O_2N•••NO_2 dimer, at the HF/jul-cc-pV{5,6}Z level, the two benzenes are unbound with an interaction energy of –4.510 kcal mol^{-}

[1]. The ΔCCSD/jul-cc-pV{Q,5}Z component increases the binding energy by 5.797 kcal mol^{-1}, and the Δ(T)/jul-cc-pV{T,Q}Z component further increases the binding energy by 1.303 kcal mol^{-1}. At the valence CCSD(T)/CBS level of theory, we obtain a binding energy of 2.580 kcal mol^{-1}.[107] The ΔT–(T)/cc-pVTZ(no f 1d) component reduces the binding energy by –0.246 kcal mol^{-1}. Where cc-pVTZ(no f 1d) denotes a truncated version of the cc-pVTZ basis set that combines the sp part of the cc-pVTZ basis set with the d function from the cc-pVDZ basis set on carbon and the s part of the cc-pVTZ basis set with the p function from the cc-pVDZ basis set on hydrogen. The Δ(Q)/cc-pVDZ component increases the binding energy by a relatively significant amount of 0.187 kcal mol^{-1}. Overall, due to the cancellation between the ΔT–(T) and Δ(Q) components, post-CCSD(T) contributions to the binding energy are fairly small and amount to –0.059 kcal mol^{-1}. The ΔCV, ΔSR, and ΔDBOC corrections are also relatively small and collectively amount to 0.029 kcal mol^{-1}. The ZPVE contribution reduces the binding energy by as much as –0.240 kcal mol^{-1}. Overall, we obtain a relativistic CCSDT(Q)/CBS binding energy of 2.310 kcal mol^{-1}, which is lower by 0.09 kcal mol^{-1} than the experimental value of 2.4 ± 0.4 kcal mol^{-1} (Table 7). Thus, the discrepancy between theory and experiment is well below the experimental uncertainty.

Acknowledgments

I gratefully thank Professor Gershom (Jan M. L.) Martin (Weizmann Institute of Science), Professor Leo Radom (University of Sydney), and Professor Krishnan Raghavachari (Indiana University, Bloomington) for many fruitful discussions and insights on composite ab initio methods over the past two decades.

References

[1] Karton, A.; Sylvetsky, N.; Martin, J. M. L. W4-17: A diverse and high-confidence dataset of atomization energies for benchmarking high-level electronic structure methods. *J. Comput. Chem.* **2017**, 38, 2063.

[2] Karton, A.; Daon, S.; Martin, J. M. L. W4-11: A high-confidence benchmark dataset for computational thermochemistry derived from first-principles W4 data. *Chem. Phys. Lett.* **2011**, 510, 165.

[3] Karton, A. Quantum mechanical thermochemical predictions 100 years after the Schrodinger equation. *Annu. Rep. Comput. Chem.* **2022**, 18, 123.

[4] Martin, J. M. L. Basis Set Convergence and Extrapolation of Connected Triple Excitation Contributions (T) in Computational Thermochemistry: The W4-17 Benchmark with Up to k Functions. Quantum Science — Frontier of Chemistry and Physics, edited by Taku Onishi, Springer Nature, 2022. DOI: 10.1007/978-981-19-4421-5_8.

[5] Karton, A.; Tarnopolsky, A.; Lamere, J. F.; Schatz, G. C.; Martin, J. M. L. Highly accurate first-principles benchmark datasets for the parametrization and validation of density functional and other approximate methods. Derivation of a robust, generally applicable, double-hybrid functional for thermochemistry and thermochemical kinetics. *J. Phys. Chem. A* **2008**, 112, 12868.

[6] Zheng, J.; Zhao, Y.; Truhlar, D. G. The DBH24/08 Database and Its Use to Assess Electronic Structure Model Chemistries for Chemical Reaction Barrier Heights. *J. Chem. Theory Comput.* **2009**, 5, 808.

[7] A. Karton, Highly Accurate CCSDT(Q)/CBS Reaction Barrier Heights for a Diverse Set of Transition Structures: Basis Set Convergence and Cost-Effective Approaches for Estimating Post-CCSD(T) Contributions. *J. Phys. Chem. A* **2019**, 123, 6720.

[8] Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, 115, 2315.

[9] Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, 19, 32184.

[10] Karton, A.; Parthiban, S.; Martin J. M. L. Post-CCSD(T) ab Initio Thermochemistry of Halogen Oxides and Related Hydrides XOX, XOOX, HOX, XO_n, and HXO_n (X = F, Cl), and Evaluation of DFT Methods for These Systems. *J. Phys. Chem. A* **2009**,113, 4802.

[11] Feller, D.; Peterson K. A.; Dixon D. A. Refined theoretical estimates of the atomization energies and molecular structures of selected small oxygen fluorides. *J. Phys. Chem. A* **2010**, 114, 613.

[12] Karton, A.; Tarnopolsky, A.; Martin, J. M. L. Atomization energies of the carbon clusters C*n* (*n* = 2–10) revisited by means of W4 theory as well as density functional, G*n*, and CBS methods. Mol. Phys. 2009, 107, 977.

[13] Karton, A. Basis set convergence of high-order coupled cluster methods up to CCSDTQ567 for a highly multireference molecule; *Chem. Phys. Lett.* **2019**, 737, 136810.

[14] Ruscic, B. Uncertainty Quantification in Thermochemistry, Benchmarking Electronic Structure Computations, and Active Thermochemical Tables. *Int. J. Quantum Chem.* **2014**, 114, 1097.

[15] Wheeler, S. E. Homodesmotic reactions for thermochemistry, *WIREs Comput. Mol. Sci.* **2012**, 2, 204.

[16] Chan, B.; Collins, E.; Raghavachari, K. Applications of isodesmic-type reactions for computational thermochemistry. *WIREs Comput. Mol. Sci.* **2021**, 11, e1501.

[17] Wheeler, S. E.; Houk, K. N.; Schleyer, P. v. R.; Allen, W. D. A Hierarchy of Homodesmotic Reactions for Thermochemistry. *J. Am. Chem. Soc.* **2009**, 131, 2547.

[18] Karton, A.; Schreiner, P. R.; Martin, J. M. L. Heats of formation of platonic hydrocarbon cages by means of high-level thermochemical procedures. *J. Comput. Chem.* **2016**, 37, 49.

[19] Yu, L.-J.; Karton, A. Assessment of theoretical procedures for a diverse set of isomerization reactions involving double-bond migration in conjugated dienes. *Chem. Phys.* **2014**, 441, 166.

[20] Karton, A. How Reliable Is DFT in Predicting Relative Energies of Polycyclic Aromatic Hydrocarbon Isomers? Comparison of Functionals from Different Rungs of Jacob's Ladder. *Chem. Phys.* **2014**, 441, 166.

[21] Chan, B.; Karton, A. PAH335 – A diverse database of highly accurate CCSD(T) isomerization energies of 335 polycyclic aromatic hydrocarbons. *Chem. Phys. Lett.* **2023**, vol, page.

[22] Ramabhadran, R. O.; Raghavachari, K. Theoretical Thermochemistry for Organic Molecules: Development of the Generalized Connectivity-Based Hierarchy. *J. Chem. Theory Comput.* **2011**, 7, 2094.

[23] Ramabhadran, R. O.; Raghavachari, K. Connectivity-Based Hierarchy for theoretical thermochemistry: assessment using wave function-based methods. *J. Phys. Chem. A* **2012**, 116, 7531.

[24] Ramabhadran, R. O.; Raghavachari, K. The successful merger of theoretical thermochemistry with fragment-based methods in quantum chemistry. *Acc. Chem. Res.* **2014**, 47, 3596.

[25] Karton, A.; Gruzman, D.; Martin, J. M. L. Benchmark thermochemistry of the C_nH_{2n+2} alkane isomers (n = 2–8) and performance of DFT and composite ab initio methods for dispersion-driven isomeric equilibria. *J. Phys. Chem. A* **2009**, 113, 8434.

[26] Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. L. The melatonin conformer space: benchmark and assessment of wavefunction and DFT methods for a paradigmatic biological and pharmacological molecule. *J. Phys. Chem. A* **2013**, 117, 2269.

[27] Kesharwani, M. K.; Karton, A.; Martin, J. M. L. Benchmark ab initio conformational energies for the proteinogenic amino acids through explicitly correlated methods. Assessment of density functional methods. *J. Chem. Theory Comput.* **2016**, 12, 444.

[28] Goerigk, L.; Karton, A.; Martin, J. M. L.; Radom, L. Accurate quantum chemical energies for tetrapeptide conformations: Why MP2 data with an insufficient basis set should be handled with caution. *Phys. Chem. Chem. Phys.* **2013**, 15, 7028.

[29] Sedlak, R.; Janowski, T.; Pitonák, M.; Rezác, J.; Pulay, P.; Hobza, P. Accuracy of Quantum Chemical Methods for Large Noncovalent Complexes. *J. Chem. Theory Comput.* **2013**, 9, 3364.

[30] P. Jurečka, J. Šponer, J. Černýa, P. Hobza, Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, 8, 1985.

[31] Řezáč, J.; Riley, K. E.; Hobza, P. Extensions of the S66 Data Set: More Accurate Interaction Energies and Angular-Displaced Nonequilibrium Geometries. *J. Chem. Theory Comput.* **2011**, 7, 3466.

[32] Santra, G.; Semidalas, E.; Mehta, N.; Karton, A.; Martin, J. M. L. S66x8 noncovalent interactions revisited: new benchmark and performance of composite localized coupled-cluster methods. *Phys. Chem. Chem. Phys.* **2022**, 24, 25555.

[33] Mehta, N.; Fellowes, T.; White, J. M.; Goerigk, L. CHAL336 Benchmark Set: How Well Do Quantum-Chemical Methods Describe Chalcogen-Bonding Interactions? *J. Chem. Theory Comput.* **2021**, 17, 2783.

[34] I. Shavitt and R. J. Bartlett, in Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory, Cambridge Molecular Science Series (Cambridge University Press, Cambridge, 2009).

[35] Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, 79, 291.

[36] Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* 1989, **90**, 1007.

[37] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, 96, 6796.

[38] A. K. Wilson, T. van Mourik, and T. H. Dunning, Gaussian basis sets for use in correlated molecular calculations. VI Sextuple zeta correlation consistent basis sets for boron through neon. *J. Mol. Struct.: THEOCHEM* **1996**, 388, 339.

[39] van Mourik, T.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. Sextuple zeta and augmented sextuple zeta correlation consistent basis sets for aluminum through argon. *Int. J. Quantum Chem.* **2000**, **76**, 205.

[40] Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. Gaussian-1 theory: a general procedure for prediction of molecular-energies. *J. Chem. Phys*. **1989**, 90, 5622.

[41] Curtiss, L. A.; Jones, C.; Trucks, G. W.; Raghavachari, K.; Pople, J. A. Gaussian-1 theory of molecular-energies for 2nd-row compounds. *J. Chem. Phys.* **1990**, 93, 2537.

[42] Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. Gaussian-2 theory for molecular energies of first- and second-row compounds. *J. Chem. Phys.* **1991**, 94, 7221.

[43] Curtiss L. A.; Raghavachari K.; Redfern P. C.; Pople J. A. Gaussian-3 (G3) theory for molecules containing first and second-row atoms. *J. Chem. Phys.* **1998**, 109, 7764.

[44] Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 theory. *J. Chem. Phys.* **2007**, 126, 084108.

[45]Henry, D. J.; Sullivan, M. B.; Radom, L. G3-RAD and G3X-RAD: Modified Gaussian-3 (G3) and Gaussian-3X (G3X) procedures for radical thermochemistry. *J. Chem. Phys.* **2003**, 118, 4849.

[46] Chan, B.; Coote, M. L.; Radom. L. G4-SP, G4(MP2)-SP, G4-sc, and G4(MP2)-sc: Modifications to G4 and G4(MP2) for the Treatment of Medium-Sized Radicals. *J. Chem. Theory Comput.* **2010**, 6, 2647.

[47] Chan, B.; Deng, J.; Radom, L. G4(MP2)-6X: a cost-effective improvement to G4(MP2). *J. Chem. Theory Comput.* **2011**, 7, 112.

[48] Karton, A.; O'Reilly, R. J.; Chan, B.; Radom, L. Determination of barrier heights for proton exchange in small water, ammonia, and hydrogen fluoride clusters with G4(MP2)-type, MPn, and SCS-MPn procedures–a caveat. *J. Chem. Theory Comput.* **2012**, 8, 3128.

[49] da Silva, G. G3X-K theory: A composite theoretical method for thermochemical kinetics. *Chem. Phys. Lett.* **2013**, 558, 109.

[50] Chan, B.; Karton, A.; Raghavachari, K.; Radom, L. Restricted Open-Shell G4(MP2)-Type Procedures. *J. Phys. Chem. A* **2016**, 120, 9299.

[51] Chan, B.; Karton, A.; Raghavachari. K. G4(MP2)-XK: A Variant of the G4(MP2)-6X Composite Method with Expanded Applicability for Main Group Elements up to Radon. *J. Chem. Theory Comput.* **2019**, 15, 4478.

[52] Semidalas, E.; Martin, J. M. L. Canonical and DLPNO-based G4(MP2)-XK-inspired composite wavefunction methods parametrized against the GMTKN55 training set: Are they more accurate and/or robust than double-hybrid DFT? *J. Chem. Theory Comput.* **2020**, 16, 4238.

[53] Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J.

A complete basis set model chemistry. I. The total energies of closed-shell atoms and hydrides of the first-row elements. *J. Chem. Phys.* **1988**, 89, 2193.

[54] Petersson, G. A.; Al-Laham, M. A. A complete basis set model chemistry. II. Open-shell systems and the total energies of the first-row atoms. *J. Chem. Phys.* **1991**, 94, 6081.

[55] Petersson, G. A.; Bennett, A.; Tensfeldt, T. G. A complete basis set model chemistry. III. The complete basis set-quadratic configuration interaction family of methods. *J. Chem. Phys.* **1991**, 94, 6091.

[56] Montgomery, J. A. Jr.; Ochterski, J. W.; Petersson, G. A. A complete basis set model chemistry. IV. An improved atomic pair natural orbital method. *J. Chem. Phys.* **1994**, 101, 5900.

[57] Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A. Jr. A complete basis set model chemistry. V. Extensions to six or more heavy atoms. *J. Chem. Phys.* **1996**, 104, 2598.

[58] Montgomery, J. A. Jr.; Frisch M. J.; Ochterski, J. W.; Petersson, G. A. A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. *J. Chem. Phys.* **1999**, 110, 2822.

[59] Wood, G. P. F.; Radom, L.; Petersson, G. A.; Barnes, E. C.; Frisch, M. J.; Montgomery, J. A. Jr A restricted-open-shell complete-basis-set model chemistry. *J. Chem. Phys.* **2006**, 125, 094106.

[60] Martin, J. M. L.; de Oliveira G. Towards standard methods for benchmark quality ab initio thermochemistry—W1 and W2 theory. *J. Chem. Phys.* **1999**, 111, 1843.

[61] Karton, A.; Martin, J. M. L. Explicitly correlated Wn theory: W1-F12 and W2-F12. *J. Chem. Phys.* **2012**, 136, 124114.

[62] Chan, B.; Radom, L. W1X-1 and W1X-2: W1-Quality Accuracy with an Order of Magnitude Reduction in Computational Cost. *J. Chem. Theory Comput.* **2012**, 8, 4259.

[63] DeYonker, N. J.; Cundari, T. R.; Wilson, A. K. The correlation consistent composite approach (ccCA): an alternative to the Gaussian-*n* methods. *J. Chem. Phys.* **2006**, 124, 114104.

[64] DeYonker N. J.; Peterson K. A.; Steyl G.; Wilson A. K.; Cundari T. R. Quantitative computational thermochemistry of transition metal species. *J. Phys. Chem. A* **2007**, 111, 11269.

[65] DeYonker, N. J.; Williams, T. G.; Imel, A. E.; Cundari, T. R.; Wilson, A. K. Accurate thermochemistry for transition metal complexes from first-principles calculations. *J. Chem. Phys.* **2009**, 131, 024106.

[66] Mintz, B.; Williams, T. G.; Howard, L.; Wilson, A. K. Computation of potential energy surfaces with the multireference correlation consistent composite approach. *J. Chem. Phys.* **2009**, 130, 234104.

[67] DeYonker, N. J.; Wilson, B. R.; Pierpont, A. W.; Cundari, T. R.; Wilson, A. K. Towards the intrinsic error of the correlation consistent Composite Approach (ccCA). *Mol. Phys.* **2009**, 107, 1107.

[68] Prascher, B. P.; Lai, J. D.; Wilson, A. K. The resolution of the identity approximation applied to the correlation consistent composite approach. *J. Chem. Phys.* **2009**, 131, 044130.

[69] Laury, M. L.; DeYonker, N. J.; Jiang, W.; Wilson, A. K. A pseudopotential-based composite method: The relativistic pseudopotential correlation consistent composite approach for molecules containing 4d transition metals (Y–Cd). *J. Chem. Phys.* **2011**, 135, 214103.

[70] Laury, M. L.; Wilson, A. K. Examining the heavy p-block with a pseudopotential-based composite method: Atomic and molecular applications of rp-ccCA. *J. Chem. Phys.* **2012**, 137, 214111.

[71] Mahler, A.; Wilson, A. K. Explicitly Correlated Methods within the ccCA Methodology. *J. Chem. Theory Comput.* **2013**, 9, 1402.

[72] Welch, B. K.; Almeida, N. M. S.; Wilson, A. K. Super ccCA (s-ccCA): an approach for accurate transition metal thermochemistry. *Mol. Phys.* **2021**, 119, e1963001.

[73] Patel, P.; Melin, T. R. L.; North, S. C.; Wilson, A. K. Ab initio composite methodologies: Their significance for the chemistry community. *Annu. Rep. Comput. Chem.* **2021**, 17, 113.

[74] Chan, B. How to Computationally Calculate Thermochemical Properties Objectively, Accurately, and as Economically as Possible. *Pure Appl. Chem.* **2017**, 89, 699.

[75] Karton, A. A computational chemist's guide to accurate thermochemistry for organic molecules. *WIREs Comput. Mol. Sci.* **2016**, 6, 292.

[76] Peterson, C.; Penchoff, D. A.; Wilson, A. K. Prediction of thermochemical properties across the periodic table: A review of the correlation consistent composite approach (ccCA) strategies and applications. *Annu. Rep. Comput. Chem.* **2016**, 12, 3.

[77] Jiang, W.; Wilson, A. K. Ab initio composite approaches: potential energy surfaces and excited electronic states. *Annu. Rep. Comput. Chem.* **2012**, 8, 29.

[78] Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical accuracy in ab initio thermochemistry and spectroscopy: current strategies and future challenges. *Theor. Chem. Acc.* **2012**, 131, 1079.

[79] Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. G*n* theory. *WIREs Comput. Mol. Sci.* **2011**, 1, 810.

[80] Klopper, W.; Bachorz, R. A.; Hättig, C.; Tew, D. P. Accurate computational thermochemistry from explicitly correlated coupled-cluster theory. *Theor. Chem. Acc.* **2010**, 126, 289.

[81] DeYonker, N.; Cundari, T. R.; Wilson, A. K. in: Piecuch, P.; Maruani, J.; Delgado-Barrio, G.; Wilson, S. (Eds.), Advances in the Theory of Atomic and Molecular Systems (Progress in Theoretical Chemistry and Physics, Vol. 19), Springer Netherlands, Dordrecht, 2009 pp. 197–224.

[82] Feller, D.; Peterson, K. A.; Dixon, D. A. A survey of factors contributing to accurate theoretical predictions of atomization energies and molecular structures. *J. Chem. Phys.* **2008**, 129, 204105.

[83] Helgaker, T.; Klopper, W.; Tew, D. P. Quantitative quantum chemistry. *Mol. Phys.* **2008**, 106, 2107.

[84] Fabian, W. M. F. Accurate thermochemistry from quantum chemical calculations? *Monatsch. Chem.* **2008**, 139, 309.

[85] Martin, J. M. L. Computational Thermochemistry: A Brief Overview of Quantum Mechanical Approaches. *Annu. Rep. Comput. Chem.* **2005**, 1, 31.

[86] Friesner, R. A. Ab Initio Quantum Chemistry: Methodology and Applications. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, 102, 6648.

[87] Helgaker, T.; Klopper, W.; Bak, K. L.; Halkier, A.; Jørgensen, P.; Olsen, J. Highly accurate ab initio computation of thermochemical data. Understanding Chemical Reactivity, Vol. 22: Quantum-Mechanical Prediction of Thermochemical Data, Cioslowski, J. ed. (Kluwer, Dordrecht, 2001), pp. 1–30.

[88] Martin, J. M. L.; Parthiban, S. W1 and W2 theory and their variants: thermochemistry in the kJ/mol accuracy range. Understanding Chemical Reactivity, Vol. 22: Quantum-Mechanical Prediction of Thermochemical Data, Cioslowski, J. ed. (Kluwer, Dordrecht, 2001), pp. 31–65.

[89] Dunning, T. H. A Road Map for the Calculation of Molecular Binding Energies. *J. Phys. Chem. A* **2000**, 104, 9062.

[90] Dixon, D. A.; Feller D.; Peterson K. A. A practical guide to reliable first principles computational thermochemistry predictions across the periodic table. *Annu. Rep. Comput. Chem.* **2012**, 8, 1.

[91] Feller D.; Peterson K. A.; Dixon D. The Impact of Larger Basis Sets and Explicitly Correlated Coupled Cluster Theory on the Feller–Peterson–Dixon Composite Method. *Annu. Rep. Comput. Chem.* **2016**, 12, 47.

[92] Crawford, T. D.; Kumar, A.; Bazanté, A. P.; Remigio, R. D.; Reduced-scaling coupled cluster response theory: Challenges and opportunities. *WIREs Comput. Mol. Sci.* **2019**, 9, e1406.

[93] Bartlett, R. J. The coupled-cluster revolution. *Mol. Phys.* **2010**, 108, 2905.

[94] Shavitt, I.; Bartlett, R. J. Many-Body Methods in Chemistry and Physics: Many-Body Perturbation Theory and Coupled-Cluster Theory. Cambridge: Cambridge University Press; 2009.

[95] Bartlett, R. J.; Musial, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, 79, 291.

[96] Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. W4 theory for computational thermochemistry: in pursuit of confident sub-kJ/mol predictions. *J. Chem. Phys.* **2006**, 125, 144108.

[97] Karton, A.; Taylor, P. R.; Martin, J. M. L. Basis set convergence of post-CCSD contributions to molecular atomization energies. *J. Chem. Phys.* **2007**, 127, 064104.

[98] Karton, A. Tightening the Screws: The Importance of Tight d Functions in Coupled-Cluster Calculations up to the CCSDT(Q) Level. *J. Phys. Chem. A* **2022**, 126, 8544.

[99] Karton, A.; Yu, L.-J.; Kesharwani, M. K.; Martin, J. M. L. Heats of formation of the amino acids re-examined by means of W1-F12 and W2-F12 theories. *Theor. Chem. Acc.* **2014**, 133, 1483.

[100] Karton, A.; Chan, B. Accurate Heats of Formation for Polycyclic Aromatic Hydrocarbons: A High-Level Ab Initio Perspective. *J. Chem. Eng. Data* **2021**, 66, 3453.

[101] Karton, A.; Chan, B.; Raghavachari, K.; Radom L. Evaluation of the heats of formation of corannulene and C_{60} by means of high-level theoretical procedures. *J. Phys. Chem. A* **2013**, 117, 1834.

[102] Manna, D.; Martin, J. M. L. What Are the Ground State Structures of C_{20} and C_{24}? An Explicitly Correlated Ab Initio Approach; *J. Phys. Chem. A* **2016**, 120, 153.

[103] Karton, A.; Martin, J. M. L. Comment on "Revised electron Affinity of SF_6 from Kinetic Data" [J. Chem. Phys. 136, 121102 2012]. *J. Chem. Phys.* 2012, 136, 197101.

[104] Kroeger, A. A.; Karton, A. Thermochemistry of phosphorus sulfide cages: an extreme challenge for high-level ab initio methods. *Struct. Chem.* **2019**, 30, 1665.

[105] Karton, A. Cope rearrangements in shapeshifting molecules re-examined by means of high-level CCSDT(Q) composite ab initio methods. *Chem. Phys. Lett.* **2020**, 759, 138018.

[106] Karton, A. High-level thermochemistry for the octasulfur ring: A converged coupled cluster perspective for a challenging second-row system. *Chem. Phys. Impact* **2021**, 3, 100047.

[107] Karton, A.; Martin, J. M. L. Prototypical π–π dimers re-examined by means of high-level CCSDT(Q) composite ab initio methods. *J. Chem. Phys.* **2021**, 154, 124117.

[108] Allen, W. D.; East, A. L. L.; Császár, A. G. Ab initio anharmonic vibrational analyses of non-rigid molecules. In: Laane, J.; Dakkouri, M.; van der Veken, B.; Oberhammer, H. eds. Structures and Conformations of Non-Rigid Molecules. Dordrecht: Kluwer; 1993, 343.

[109] East, A. L. L.; Allen, W. D. The heat of formation of NCO. *J. Chem. Phys.* **1993**, 99, 4638.

[110] Klippenstein, S. J.; East, A. L. L.; Allen, W. D. A high level ab initio map and direct statistical treatment of the fragmentation of singlet ketene. *J. Chem. Phys.* **1996**, 105, 118.

[111] Császár, A. G.; Allen, W. D.; Schaefer, H. F. In pursuit of the ab initio limit for conformational energy prototypes. *J. Chem. Phys.* **1998**, 108, 9751.

[112] Schuurman, M. S.; Muir, S. R.; Allen, W. D.; Schaefer, H. F. Toward subchemical accuracy in computational thermochemistry: Focal point analysis of the heat of formation of NCO and [H,N,C,O] isomers. *J. Chem. Phys.* **2004**, 120, 11586.

[113] Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kallay, M.; Gauss, J. W3 theory: Robust computational thermochemistry in the kJ/mol accuracy range. *J. Chem. Phys.* **2004**, 120, 4129.

[114] Sylvetsky, N.; Peterson, K. A.; Karton, A.; Martin, J. M. L. Toward a W4-F12 approach: Can explicitly correlated and orbital-based ab initio CCSD(T) limits be reconciled? *J. Chem. Phys.* **2016**, 144, 214101.

[115] Chan, B.; Radom, L. W3X: a cost-effective post-CCSD(T) composite procedure. *J. Chem. Theory Comput.* **2013**, 9, 4769.

[116] Chan, B.; Radom, L. W2X and W3X-L: cost-effective approximations to W2 and W4 with kJ mol^{-1} accuracy. *J. Chem. Theory Comput.* **2015**, 11, 2109.

[117] Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. HEAT: High accuracy extrapolated ab initio thermochemistry. *J. Chem. Phys.* **2004**, 121, 11599.

[118] Bomble, Y. J.; Vázquez, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. High-accuracy extrapolated ab initio thermochemistry. II. Minor improvements to the protocol and a vital simplification. *J. Chem. Phys.* **2006**, 125, 064108.

[119] Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. High-accuracy extrapolated ab initio thermochemistry. III. Additional improvements and overview. *J. Chem. Phys.* **2008**, 128, 114111.

[120] Thorpei, J. H.; Lopez, C. A.; Nguyen, T. L.; Baraban, J. H.; Bross, D. H.; Ruscic, B.; Stanton, J. F. High-accuracy extrapolated ab initio thermochemistry. IV. A modified recipe for computational efficiency. *J. Chem. Phys.* **2019**, 150, 224102.

[121] Thorpe, J. H.; Kilburn, J. L.; Feller, D.; Changala, P. B.; Bross, D. H.; Ruscic, B.; Stanton, J. F. Elaborated thermochemical treatment of HF, CO, N_2, and H_2O: Insight into HEAT and its extensions. *J. Chem. Phys.* **2021**, 155, 184109.

[122] Feller D.; Peterson K. A.; Dixon D. A. Refined theoretical estimates of the atomization energies and molecular structures of selected small oxygen fluorides. *J. Phys. Chem. A* **2010**, 114, 613.

[123] Feller D.; Peterson K. A.; Dixon D. A. Ab Initio Coupled Cluster Determination of the Heats of Formation of $C_2H_2F_2$, C_2F_2, and C_2F_4. *J. Phys. Chem. A* **2011**, 115, 1440.

[124] Feller D.; Peterson K. A.; Dixon D. A. Further benchmarks of a composite, convergent, statisticallycalibrated coupled cluster-based approach for thermochemical and spectroscopic studies. *Mol Phys.* **2012**, 110, 2381.

[125] Feller D.; Peterson K. A.; Ruscic B. Improved accuracy benchmarks of small molecules using correlation consistent basis sets. *Theor. Chem. Acc.* **2014**, 133, 1407.

[126] Feller D. Estimating the intrinsic limit of the Feller-Peterson-Dixon composite approach when applied to adiabatic ionization potentials in atoms and small molecules. *J. Chem. Phys.* **2017**, 147, 034103.

[127] Vogiatzis, K. D.; Haunschild, R.; Klopper, W. Accurate atomization energies from combining coupled-cluster computations with interference-corrected explicitly correlated second-order perturbation theory. *Theor. Chem. Acc.* **2014**, 133, 1446.

[128] Ruscic, B.; Pinzon, R. E.; Morton, M. L.; von Laszewski, G.; Bittner, S. J.; Nijsure S. G.; Amin, K. A.; Minkoff, M.; Wagner, A. F. Introduction to active thermochemical tables: several "key" enthalpies of formation revisited. *J. Phys. Chem. A* **2004**, 108, 9979.

[129] Ruscic, B.; Pinzon, R, E.; von Laszewski, G.; Kodeboyina, D.; Burcat, A.; Leahy, D.; Montoya, D.; Wagner, A. F. Active thermochemical tables: thermochemistry for the 21st century. *J. Phys. Conf. Ser.* **2005**, 16, 561.

[130] Ramabhadran, R. O.; Sengupta, A.; Raghavachari, K. Application of the Generalized Connectivity-Based Hierarchy to Biomonomers: Enthalpies of Formation of Cysteine and Methionine. *J. Phys. Chem. A* **2013**, 117, 4973.

[131] Kozuch, S.; Schleif, T.; Karton, A. Quantum mechanical tunnelling: the missing term to achieve sub-kJ mol^{-1} barrier heights. *Phys. Chem. Chem. Phys.* **2021**, 23, 10888.

[132] E. Papajak and D. G. Truhlar, Convergent partially augmented basis sets for post-Hartree–Fock calculations of molecular properties and reaction barrier heights. *J. Chem. Theory Comput.* **2011**, 7, 10.

[133] Karton, A.; Martin, J. M. L. "Comment on: 'Estimating the Hartree–Fock limit from finite basis set calculations' [Jensen F (2005) theor Chem Acc 113:267]," *Theor. Chem. Acc.* **2005**, 115, 330.

[134] Jensen, S. R.; Saha, S.; Flores-Livas, J. A.; Huhn, W.; Blum, V.; Goedecker, S.; Frediani, L. The elephant in the room of density functional theory calculations. *J. Phys. Chem. Lett.* **2017**, 8, 1449.

[135] Schwenke, D. W. The extrapolation of one-electron basis sets in electronic structure calculations: how it should work and how it can be made to work. *J. Chem. Phys.* **2005**, 122, 014107.

[136] Ranasinghe, D. S.; Petersson, G. A. CCSD(T)/CBS atomic and molecular benchmarks for H through Ar. *J. Chem. Phys.* **2013**, 138, 144104.

[137] Kesharwani, M. K.; Sylvetsky, N.; Köhn A., Tew, D. P.; Martin, J. M. L. Do CCSD and approximate CCSD-F12 variants converge to the same basis set limits? The case of atomization energies. *J. Chem. Phys.* **2018**, 149, 154109.

[138] Harding, M. E.; Vázquez, J.; Gauss, J.; Stanton, J. F.; Kállay, M. Towards highly accurate ab initio thermochemistry of larger systems: Benzene. *J. Chem. Phys.* **2011**, 135, 044513.

[139] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett.* **1989**, 157, 479.

[140] Stanton, J. F. Why CCSD(T) works: a different perspective. *Chem. Phys. Lett.* **1997**, 281, 130.

[141] Bak, K. L.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Gauss, J. Coupled-cluster singles, doubles and triples (CCSDT) calculations of atomization energies. *Chem. Phys. Lett.* **2000**, 317, 116.

[142] Feller, D.; Peterson, K. A. Probing the Limits of Accuracy in Electronic Structure Calculations: Is Theory Capable of Results Uniformly Better Than "Chemical Accuracy". *J. Chem. Phys.* **2007**, 126, 114105.

[143] Raghavachari K. Historical perspective on: a fifthorder perturbation comparison of electron correlation theories [Volume 157, Issue 6, 26 May 1989, Pages 479–483]. *Chem. Phys. Lett.* **2013**, 589, 35.

[144] Karton, A.; Martin, J. M. L. The lowest singlet-triplet excitation energy of BN: A converged coupled cluster perspective. *J. Chem. Phys.* **2006**, 125, 144313.

[145] Sylvetsky, N.; Martin, J. M. L. Probing the basis set limit for thermochemical contributions of inner-shell correlation: balance of core-core and core-valence contributions. *Mol. Phys.* **2019**, 117, 1078.

[146] Ranasinghe, D. S.; Frisch, M. J.; Petersson, G. A. A density functional for core-valence correlation energy. *J. Chem. Phys.* **2015**, 143, 214111.

[147] Douglas, M.; Kroll, N. M. Quantum electrodynamical corrections to the fine structure of helium. *Ann. Phys.* **1974**, 82, 89.

[148] Heß, B. A. Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators. *Phys. Rev. A* **1986**, 33, 3742.

[149] Collins, C. L.; Dyall, K. G.; Schaefer, H. F. Relativistic and correlation effects in CuH, AgH, and AuH: comparison of various relativistic methods. *J. Chem. Phys.* **1995**, 102, 2024.

[150] Barysz, M.; Sadlej, A. J. Two-component methods of relativistic quantum chemistry: from the Douglas–Kroll approximation to the exact two-component formalism. *J. Mol. Struct. Theochem* **2001**, 573, 181.

[151] Karton, A.; Ruscic, B.; Martin, J. M. L. Benchmark atomization energy of ethane: importance of accurate zero-point vibrational energies and diagonal Born–Oppenheimer corrections for a 'simple' organic molecule. *J. Mol. Struct. Theochem* **2007**, 811, 345.

[152] Tajti, A.; Szalay, P. G.; Gauss, J. Perturbative treatment of the electron-correlation contribution to the diagonal Born–Oppenheimer correction. *J. Chem. Phys.* **2007**, 127, 014102.

[153] Grev, R. S.; Jenssen, C. L.; Schaefer, H. F. Concerning zero-point vibrational energy corrections to electronic energies. *J. Chem. Phys.* **1991**, 95, 5128.

[154] Pople, J. A.; Scott, A. P.; Wong, M. W.; Radom L. Scaling Factors for Obtaining Fundamental Vibrational Frequencies and Zero-Point Energies from HF/6-31G* and MP2/6-31G* Harmonic Frequencies. *Isr. J. Chem.* **1993**, 33, 345.

[155] Scott, A. P.; Radom, L. Harmonic Vibrational Frequencies: An Evaluation of Hartree–Fock, Møller–Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors. *J. Phys. Chem.* **1996**, 100, 16502.

[156] Sinha, P.; Boesch, S. E.; Gu, C.; Wheeler, R. A.; Wilson, A. K.; Harmonic Vibrational Frequencies: Scaling Factors for HF, B3LYP, and MP2 Methods in Combination with Correlation Consistent Basis Sets. *J Phys Chem A* **2004**, 108, 9213.

[157] Merrick, J. P.; Moran, D.; Radom, L. An Evaluation of Harmonic Vibrational Frequency Scale Factors. *J Phys Chem A* **2007**, 111, 11683.

[158] Irikura, K. K.; Johnson III, R. D.; Kacker R. N.; Kessel, R. J. Uncertainties in scaling factors for ab initio vibrational zero-point energies. *J Chem Phys* **2009**, 130, 114102.

[159] Alecu, I.; Zheng, J.; Zhao, Y.; Truhlar, D. G. Computational Thermochemistry: Scale Factor Databases and Scale Factors for Vibrational Frequencies Obtained from Electronic Model Chemistries. *J. Chem. Theory Comput.* **2010**, 6, 2872.

[160] Laury, M. L.; Carlson, M. J.; Wilson, A. K. Vibrational Frequency Scale Factors for Density Functional Theory and the Polarization Consistent Basis Sets. *J. Comput. Chem.* **2012**, 33, 2380.

[161] Martin, J. M. L.; Kesharwani, M. K. Assessment of CCSD(T)-F12 approximations and basis sets for harmonic vibrational frequencies. *J. Chem. Theory Comput.* **2014**, 10, 2085.

[162] Kesharwani, M. K.; Brauer, B.; Martin, J. M. L. Frequency and zero-point vibrational energy scale factors for double hybrid density functionals (and selected other methods): can anharmonic force fields be avoided? *J. Phys. Chem. A* **2015**, 119, 1701.

[163] Trujillo, J. C. Z.; McKemmish, L. K. Model Chemistry Recommendations for Scaled Harmonic Frequency Calculations: A Benchmark Study. *J. Phys. Chem. A* **2023**, 127, 1715.

[164] Schmidt. T. W. The Spectroscopy of C_2: A Cosmic Beacon. *Acc. Chem. Res.* **2021**, 54, 481.

[165] Borsovszky, J.; Nauta, K.; Jiang, J.; Hansen, C. S.; McKemmish, L. K.; Field, R. W.; Stanton, J. F.; Kable, S. H.; Schmidt. T. W. Photodissociation of dicarbon: How nature breaks an unusual multiple bond. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, 118, e2113315118.

[166] Bismillah, A. N.; Chapin, B. M.; Hussein, B. A.; McGonigal, P. R. Shapeshifting molecules: the story so far and the shape of things to come. *Chem. Sci.* **2020**, 11, 324.

[167] Csaszar, A. G.; Fabri, C.; Sarka, J. Quasistructural molecules. *WIREs Comput. Mol. Sci.* **2020**, 10, e1432.

[168] Graulich, N. The Cope rearrangement—the first born of a great family. *WIREs Comput. Mol. Sci.* **2011**, 1, 172.

[169] McKee, M. L. Fluctional molecules. *WIREs Comput. Mol. Sci.* **2011**, 1, 943.

[170] Moreno, P. O.; Suarez, C.; Tafazzoli, M.; True, N. S.; LeMaster, C. B. Gas phase NMR study of the degenerate Cope rearrangement of bullvalene. *J. Phys. Chem.* **1992**, 96, 10206.

[171] Yu, L.-J.; Sarrami, F.; O'Reilly, R. J.; Karton, A. Reaction barrier heights for cycloreversion of heterocyclic rings: An Achilles' heel for DFT and standard ab initio procedures. *Chem. Phys.* **2015**, 458, 1.

[172] Ruscic, B. Bross, D. H. Active Thermochemical Tables (ATcT) values based on ver. 1.124 of the Thermochemical Network, Argonne National Laboratory, Lemont, Illinois 2022; available at ATcT.anl.gov (retrieved on March 18, 2023).

[173] Grover, J. R.; Walters, E. A.; Hui, E. T. Dissociation energies of the benzene dimer and dimer cation. *J. Chem. Phys.* **1987**, 91, 3233.