

Geometries and Vibrational Frequencies of Small Radicals: Performance of Coupled Cluster and More Approximate Methods

Peter R. Tentscher[†] and J. Samuel Arey^{*,†,‡}

[†]Environmental Chemistry Modeling Laboratory, EPFL, Lausanne, Switzerland

[‡]Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dübendorf, Switzerland

Supporting Information

ABSTRACT: We generated a new set of reference geometries of small radicals using experimental equilibrium structures, as well as a benchmark-quality coupled cluster additivity scheme including up to quadruples excitations (CCSDTQ). Using these geometries and a set of experimental vibrational frequencies of open shell diatomics, we evaluated the performance of various coupled cluster methods based mainly on unrestricted references, using Dunning basis sets both with and without core correlation. Contrary to previous results, we found that UCCSD(T) and ROCCSD(T) perform equally well for geometries, better than CCSD, and close to their performances for closed shell systems. No improvement over CCSD(T) was achieved by using a Brueckner reference (BD(T)) or full triples (CCSDT). For frequencies, ROCCSD(T), BD(T), and CCSDT improve upon UCCSD(T), especially for the troublesome NO and CO⁺ cases. EOMIP-CCSD yields geometries and harmonic frequencies similar to CCSD, and qualitatively correct anharmonic (VPT2) contributions in all cases, like the RO-CC methods. The double hybrid DFT functional B2PLYP-D yields geometries and frequencies of similar quality to that of CCSD but at a much reduced cost. The meta hybrid functionals M06-2X, M06-HF, and BMK perform worse than CCSD, and worse than B3LYP, on average.

1. INTRODUCTION

Quantum chemistry enables the study of molecular systems that are not easily accessed experimentally. Radical species constitute a broad category of such systems, and thus, their theoretical description is of special importance. Calculations of high-quality geometries and vibrational frequencies are an integral part of such studies, as results can be used either for comparison with experiment, for (sub)-kJ/mol thermochemistry, or as benchmarks for the testing of more approximate methods. If the goal is “spectroscopic accuracy,” then errors should be <0.1 pm for bond lengths and 1 cm⁻¹ for vibrational frequencies. Stringent composite methods (e.g., HEAT,¹ W4²) offer the needed accuracy for single point calculations of systems with several heavy atoms. So far, their use for obtaining geometries and vibrational frequencies has been limited to small systems such as diatomics. For bigger systems, one must select the “best” method affordable.

For geometries of larger closed shell systems, the CCSD(T)^{3,4} method with triple- or quadruple- ζ basis sets is often the method of choice, and corresponding analytical gradient calculations can now be carried out for systems as big as ferrocene in a core correlated triple- ζ basis (672 basis functions).⁵ Geometry optimizations with this method yield highly accurate results: on a set of 28 small, closed shell molecules, Helgaker et al.⁶ reported a mean unsigned error (MUE) in bond lengths vs experimental results of 0.23 pm using the CCSD(T)/cc-pVTZ model chemistry. In a subsequent study using improved reference geometries, Bak et al.⁷ reported MUEs of 0.2 and 0.1 pm for triple- and quadruple- ζ basis sets, respectively; they observed only minor differences in performance between the different Dunning basis set types^{8–10} (cc-pVXZ, aug-cc-pVXZ, cc-pCVXZ) of the same cardinal number. Hence, the highly accurate

thermochemical methods HEAT^{1,11} and W4,² the former originating from the need of accurate models for radicals, base their calculations on CCSD(T) geometries and frequencies.

For open shell systems, the situation is less clear. The only study of bond lengths of radicals using a statistically relevant data set was published in 2001 by Byrd et al.:¹² Applying frozen core (FC)-ROCCSD(T)¹³/cc-pVTZ to a set of small radicals, the authors reported a MUE in bond length of 0.58 pm, inferior to the performance for closed shell systems. More surprisingly, UCCSD(T) was found to perform worse (MUE of 1.25 pm) than ROCCSD(T) and also worse than CCSD.¹⁴ However, the Byrd study¹² employed basis sets of limited size and did not account for core correlation, and these data are discussed in more detail later on.

For harmonic vibrational frequencies of diatomic radicals, Beran et al.¹⁵ reported erratic performance of CCSD(T). The reasons for the mixed performance of CCSD(T) have been analyzed in the subsequent literature,¹⁶ and some systems have been studied in great detail.¹⁷ As reviewed by Stanton and Gauss,¹⁶ these systems pose problems for *ab initio* methods that are not (or are to a lesser extent) encountered for their closed shell counterparts: spin contamination, symmetry breaking, problems associated with near-instabilities of the reference wave function, pseudo Jahn–Teller effects, and multireference character. The above-mentioned publications^{12,15,16} are highly cited, frequently as a justification for not performing CCSD(T) computations at all when radicals are involved, but to use either CCSD or density functional (DFT) methods instead. This is in strong contrast with the literature on closed shell molecules:

Received: November 22, 2011

Published: May 16, 2012

CCSD(T) is still considered the “gold standard” benchmark for the majority of chemical problems.

These studies^{12,15,16} concerning the shortcomings of CCSD(T) on radicals propose two strategies to improve the performance of CCSD(T): (a) a change of the reference wave function, e.g., from UHF to ROHF or vice versa, to Brueckner¹⁸ or otherwise optimized orbitals,¹⁹ or possibly to Kohn–Sham orbitals,¹⁵ or (b) switching from CCSD(T) to a method less dependent on the quality of the reference, such as connected triples (CCSDT)²⁰ or the CCSD(2)¹⁵ method. Beran et al.¹⁵ applied orbital-optimized coupled cluster with perturbative triples (OD(T)) and CCSD(2) to frequency calculations of some diatomic doublet radicals. This seemed to eliminate the biggest outliers of the frequency test set. However, the authors chose the unorthodox strategy of computing harmonic frequencies on the experimental geometries instead of using properly minimized geometries for each model chemistry. Thus, it remains unclear whether OD(T) and CCSD(2) could be good choices for geometry optimizations, as well, and whether the true performance of those model chemistries was assessed properly by the use of nonminimized geometries. In order to avoid problems with an open shell reference, it is also possible to use closed shell electronic states as references, such as in the EOMIP-CCSD²¹ and EOMEA-CCSD²² methods.

In this study, we (a) compile a set of accurate reference geometries for radicals, (b) evaluate the performance of different coupled cluster and more approximate methods for predicting these geometries, and (c) evaluate a selected set of model chemistries for vibrational frequencies of radical diatomics, with a special focus on methods that promise an improved description of radicals. Ultimately, we aim to assess the magnitude of error to be expected from different methods for both properties. In order to address the need to describe larger systems, we included a few DFT methods that have been shown to perform well across a variety of chemistry assessments.²³

As we will show in section 3.1.7, the choice of appropriate reference geometries is crucial when assessing the methods for bond lengths in our study. One should use true equilibrium r_e experimental structures when comparing theory to experiments. These are derived from equilibrium rotational constants, B_e , that are free from vibrational effects. For many, especially larger, systems, only B_0 constants are available, and the resulting r_0 structures correspond to an average over zero point vibrational motion.²⁴ However, in cases where experimental rotational–vibrational coupling constants are lacking or are of limited quality, these can be computed using quantum chemistry. From experimental B_0 constants and calculated coupling constants, semiempirical equilibrium structures, r_{se} , can be derived.^{25,26} Bak et al.⁷ used r_{se} structures for the evaluation of CCSD(T) equilibrium structures of closed shell species, and good agreement between r_{se} bond lengths and computed bond lengths was achieved. The authors also concluded that their r_{se} structures are likely of higher quality than the corresponding experimental r_e structures used in a previous study,⁶ and only when using r_{se} structures were they able to show that CCSD(T)/cc-pVQZ outperforms CCSD(T)/cc-pVTZ for bond lengths (MUEs of 0.20 and 0.13 pm, respectively). However, this strategy has not always been successful: an r_{se} structure of HCC,²⁷ using vibrational corrections calculated at the all-electron CCSD(T)/aug-cc-pCVQZ level of theory, left a discrepancy of 0.15 pm with respect to the theoretical best estimate that accounted for correlation up to the full configuration interaction (FCI) level.

The authors suggested that this remaining discrepancy most likely originates from the shortcomings of the VPT2 scheme²⁶ used for the vibrational corrections. In light of these considerations, we decided to compute accurate geometries for the molecules in our test set and then choose on a case-by-case basis whether we prefer experimental, semiempirical, or computed equilibrium structures as a reference.

To achieve highly accurate geometries using coupled cluster theory, the quality of the model chemistry can be improved in systematic ways using composite methods.^{16,28} In short, the quality of the reference wave function becomes less important as higher levels of correlation are incorporated into the calculation,^{16,28} cumulating in the correct correlation treatment (in the given basis set) when the full configuration interaction (FCI) limit is reached. As an approximation to FCI, higher correlated treatments²⁹ (CCSDT, CCSDTQ, etc.) can be incorporated into single point energy or derivative calculations. This strategy is applicable to geometry optimizations and frequency calculations. Gradients determined using the W4² scheme yielded excellent geometries on first-row diatomics³⁰ including the OH and CH radicals. Similar studies by Heckert et al.^{31,32} also found excellent agreement between computed and experimental bond lengths of mostly closed shell diatomics. We chose one of Heckert’s methods that applies basis-set converged CCSD(T) and up to CCSDTQ contributions as a “best affordable” approach for our benchmark geometries.

Using experimental equilibrium structures, computed benchmark geometries, and a set of experimental vibrational frequencies, we test several conventional wave function methods (CCSD(T), CCSD, MP2,³³ and HF), both with and without core correlation. Besides these widely used models, our focus is on methods that promise to improve upon them: coupled cluster calculation with Brueckner orbitals as a reference (BD, BD(T)),³⁴ with connected triples (CCSDT),²⁰ and the EOMIP²¹ approach for the treatment of ionized species. EOMIP-CCSD reportedly gives qualitatively correct descriptions of very problematic systems such as the nitrate radical.^{35,36} Thus, EOMIP-CCSD is of special importance to verify results from other methods. However, to obtain properties comparable to CCSD(T) quality, the EOMIP approach typically needs larger basis sets and some kind of triples excitation.³⁷ For more approximate methods, we (re)evaluate the optimized orbitals-based MP2 (O2-MP2)³⁸ as well as a selection of the different families of DFT methods that can be used to treat much larger systems, including double hybrid (B2PLYP-D³⁹), meta hybrid (M06-2X,⁴⁰ M06-HF,⁴⁰ BMK⁴¹), hybrid (B3LYP⁴²), and GGA (B97-D⁴³) functionals. We decided to use an unrestricted reference function in most cases. For the present geometry reference set, RO-CCSD(T) did not perform significantly better than UCCSD(T). RO methods can resolve (near-)instability problems of the reference wave function,¹⁷ but these are especially prone to artifactual symmetry breaking,^{1,16} which often results in nonsensical vibrational frequencies. UHF-based coupled cluster methods are also favorable from a computational cost point of view: to our knowledge, analytical second derivatives are not available for coupled cluster RO methods in currently available quantum chemistry software packages, and hence we focused on UHF-based methods. For the geometry test set, we emphasize a statistical evaluation of the errors of the individual methods. For vibrational frequencies, collective statistics are less meaningful, because the errors are dominated by a handful of problematic systems.

We do not aim to further investigate the origin of the problems for the treatment of open shell species by wave function methods. To this end, we refer the reader to the excellent review by Stanton and Gauss¹⁶ as well as the detailed analysis of “triplet instabilities” by Szalay et al.¹⁷ We also refer the reader to numerous studies on some of the individual radicals (see Methodology section for a list). These studies report on the problems they pose to quantum chemistry, and what methods have so far been found to perform best for each particular case. A comprehensive literature review thereon is out of the scope of the present work.

2. METHODOLOGY

2.1. Geometry Optimizations. Geometry optimizations using the HF, MP2,³³ CC2,⁴⁴ CCSD,¹⁴ and CCSD(T)³ model chemistries were conducted with the CFOUR v1 program package.⁴⁵ CC3,⁴⁶ CCSDT,⁴⁷ and higher excited calculations were carried out with the MRCC⁴⁸ package linked to the CFOUR v1 package. Most calculations were based on UHF reference wave functions, and in the remainder of the manuscript, we refer to them without the prefix U. CCSD(T) calculations based on an ROHF reference¹³ are prefixed with RO.⁴⁹ EOMIP-CCSD²¹ calculations were conducted with the CFOUR v1 package, and the underlying CCSD calculation was always the restricted $n + 1$ -electron configuration for the n -electron EOMIP calculation. Our initial guess for the root was ionization from the HOMO, but this did not always lead to the ionized state of lowest energy and had to be changed appropriately in some cases. In tables, EOMIP-CCSD is abbreviated as EOMIP.

All CFOUR calculations, except CC3, used analytic gradients.^{26,50} The RMS forces were converged to at least 10^{-8} Hartree per Bohr in Z-matrix coordinates. Dunning’s correlation consistent basis sets^{8–10} were used throughout and are abbreviated PVXZ, APVXZ, PCVXZ, and APCVXZ for cc-pVXZ, aug-cc-pVXZ, cc-pCVXZ, and aug-cc-pCVXZ ($X = D, T, Q$) in the remainder of the article. The (A)PVXZ bases are designed for use with the frozen core (FC) approximation, and the (A)PCVXZ bases are designed for use with correlated core electrons (all electron, AE). However, due to the increased computational expense of the PCVXZ bases, AE computations with PVXZ bases are frequently found in the literature; these computations do not account for all core correlation, but the resulting bias in geometries is found improved compared to the corresponding FC calculations.⁵¹ Because of their practical relevance, we included such methods in the present study.

We applied three composite schemes, by using the CFOUR implementations of additivity of energies, gradients, and force constants, as well as basis set extrapolation.^{31,32,52} Our theoretical best estimate computations were carried out with one of the additivity schemes formerly used by Heckert et al.^{31,32} The HF-SCF gradient was evaluated with the PV6Z basis and refined by CCSD(T)/PV6Z, CCSDT/PVTZ, and CCSDTQ/PVDZ calculations, all using the frozen core (FC) approximation. Core correlation was treated with an AE calculation at the CCSD(T)/PCVQZ level. This scheme assumes additivity of energy and gradient, and analytical gradient calculations were used throughout. In the rest of the manuscript, this is referred to as CCSDTQex. A computationally less demanding additivity scheme with up to triples excitations (CCSDT) using basis set extrapolation was used in the following way: the HF-SCF gradient was evaluated with a three-point extrapolation (PVTZ-PVQZ-PV5Z), the FC-CCSD(T) contribution

with a PVTZ-PVQZ extrapolation, and the FC-CCSDT contribution with a PVDZ-PVTZ extrapolation, and the contribution of core correlation was evaluated at the CCSD(T)/PCVTZ level. This is referred to as CCSDT_{ex}. Finally, an additivity and extrapolation scheme using only up to CCSD(T) calculations was used, being identical to CCSDT_{ex}, but lacking the CCSDT contribution. This is referred to as CCSD(T)_{ex}. BD(T),³⁴ BD,³⁴ B2PLYP-D,³⁹ M06-2X,⁴⁰ M06-HF,⁴⁰ BMK,⁴¹ B3LYP,⁴² and B97-D⁴³ calculations were carried out with the Gaussian 09 vA.02⁵³ suite using tight (maximum force $<1.5 \times 10^{-5}$, RMS force $<10^{-5}$ Hartree per Bohr, maximum displacement $<6 \times 10^{-5}$, RMS displacement $<4 \times 10^{-5}$ Bohr) geometry convergence criteria. For BD and BD(T), we used the “Conver=10” keyword, which corresponds to an energy convergence of 10^{-10} Hartrees and convergence of the singles amplitudes of 10^{-8} . MP2 calculations employing the resolution-of-identity approximation (RI-MP2)⁵⁴ were carried out with the Q-Chem 3.2.0.2⁵⁵ program. We conducted Scaled Opposite Spin MP2 (SOS-MP2)⁵⁶ using a scaling factor of 1.0 and the “optimized second order opposite spin” (O2) methodology.³⁸

A consistent choice of the reference function was made for all UHF and ROHF based methods; broken symmetry references were not used for symmetric molecules. For compounds that were not part of the frequency test set, the structures found were confirmed to be true minima by geometry optimizations and frequency calculations at the FC-CCSD/PVTZ level of theory. For all methods, geometries were optimized in the same symmetry and with the same occupation vectors as used for these frequency calculations.

2.2. Frequency Calculations. Harmonic frequencies were calculated using analytical second derivatives for the MP2,⁵⁷ CC2,⁵⁸ CCSD,⁵⁹ CCSD(T),⁵⁹ CCSDT,⁵⁸ and CCSDTQ_{ex}⁵² models with the CFOUR and CFOUR/MRCC packages. For technical reasons, frozen core calculations including CCSDT and higher excitations were calculated using numerical second derivatives of analytical first derivatives for N_2^+ , O_2^+ , and F_2^+ . For EOMIP-CCSD and ROHF-based frequencies, numerical second derivatives of analytical first derivatives were used throughout. For CC3⁴⁶ frequencies, numerical derivatives were used. Fundamental⁶⁰ frequencies were approximated by the VPT2²⁶ method as implemented in CFOUR v1. For all numerical derivatives, the default parameters of CFOUR v1 were employed. The equilibrium geometries for VPT2 calculations were optimized to 10^{-10} Hartree per Bohr in Z-matrix coordinates. Harmonic frequencies with the BD(T), BD, B2PLYP-D, BMK, M06-2X, M06-HF, B3LYP, and B97-D methods were conducted with Gaussian 09, using standard settings throughout. The current implementation of VPT2 in the Gaussian 09 package does not allow for computations of linear molecules; therefore, no anharmonic calculations with the corresponding methods were conducted. It would be interesting to proceed with a test of B2PLYP-D in a future study, as B2PLYP anharmonicity constants for closed shell systems are very promising.⁶¹ Frequency calculations with the SOS-O2-MP2 method were not correctly enabled in the Q-Chem 3.2.0.2 version available to the authors. Hence, we simply cite SOS-O2-MP2 frequencies for the same test set obtained in a previous study³⁸ using a development version of Q-Chem.

2.3. Reference Data. Geometries. The geometry reference set is a subset of the molecules considered in the study by Byrd et al.;¹² we discarded the CH_3O and $C_4H_2^+$ systems because of their larger size and the unavailability of experimental r_e

structures. For CH_3O , the reported experimental geometry⁶² is interpreted in C_{3v} symmetry. However, the real geometry is slightly distorted to C_s ,⁶³ and a comparison of the corresponding bond lengths is not meaningful. Both systems are large compared to the rest of our test set, and we did not conduct calculations beyond the AE-CCSD(T)/PCVQZ level for them. For the methods tested in our study, the discrepancies compared to experimental results are much larger than for other systems in the test set, i.e., up to 2.5 pm for the HC bond in HCCCCCH^+ and up to 1 pm for the CO bond in CH_3O , which may or may not be explained by not using rotational–vibrational interaction constants in the derivation of those geometries. Performing benchmark calculations was out of the scope of the present study, so we decided to not include these two systems in our test set. Nevertheless, these systems warrant further examination in future theoretical and experimental studies.

For the remaining 27 systems (32 bond lengths), we compiled an updated set of experimental values from the literature, which resulted in values significantly different from the ones used by Byrd et al.¹² Experimental values were compiled from the NIST CCCBDB database,⁶⁴ the collections by Huber and Herzberg,^{65,66} the Landolt-Börnstein book series,^{67–70} the references therein, and often from newer literature not included in those collections.

For all diatomic molecules, we recalculated equilibrium bond lengths from experimental B_e constants using a consistent set of physical constants (CODATA 2010 recommended values),⁷¹ and the reported uncertainties correspond to the propagation of 2σ in B_e . The r_e structures of H_2O^+ , NH_2 , BH_2 , CH_3 , HCO , and HOO are reproduced from the cited literature without modification. For CNC , CO_2^+ , HCC , HCP^+ , HNF , and N_3 , only experimental r_0 (r_s for HCC) structures are available. For CNC and N_3 , we estimated the experimental uncertainty from the reported rotational B_0 constants, as no experimental uncertainty was given in the original literature.

The data set was subdivided into subsets A and B. This was necessary because we conducted some of the more expensive calculations only for the smallest molecules (set A). All experimental bond lengths, corresponding experimental uncertainty, semiempirical structures (r_{se}) if available, and CCSDTQex theoretical best estimates (r_{ai}) are tabulated for sets A (Table 1) and B (Table 2).

Frequencies. The frequency test set consisted of harmonic and anharmonic frequencies of several diatomics, identical to the set used by Beran et al.,¹⁵ but using more recent experimental data. Experimental values for the harmonic frequencies, ω_e , and first anharmonicity constants, $\omega_e x_e$, are given in Table 3. Different frequencies have been reported for FH^+ , and we chose the (identical) values of Gewurtz et al.⁷² and Hovde et al.⁷³ Both authors report an anharmonic frequency of 2912.5 cm^{-1} . In a more recent study, Yench a et al.⁷⁴ report lower values, i.e., a ω_e of 3061.8 cm^{-1} and a $\omega_e x_e$ of 80.8 cm^{-1} . Our best theoretical estimates (see below) are much closer to the values of Gewurtz et al.,⁷² and we did not find indications of FH^+ being a problematic system for coupled cluster theory. Therefore, we consider Gewurtz et al.'s experimental values likely to be more accurate than those reported by Yench a et al.⁷⁴ Similarly, for F_2^+ , Huber and Herzberg⁶⁶ reported a ω_e of 1073.3 cm^{-1} , and Cormack et al., a value of 1126 cm^{-1} , whereas Tuckett et al.⁷⁵ measured 1089.9 cm^{-1} , in agreement with Yang et al.⁷⁶ (1091.5 cm^{-1}). As for FH^+ , the photoelectron

Table 1. Experimental and Theoretical Best Estimate Geometry Data Set A [pm], Where r_e Is the Experimental Equilibrium Geometry, r_0 Is the Experimental Zero-Point Geometry, r_{se} Is the Semiempirical Geometry, r_{ai} Is the *ab Initio* Best Estimate, and r_{ref} Is the Chosen Reference Geometry^a

molecule	$r_e \{r_0\}$	ref	r_{se}	ref	r_{ai}^d	r_{ref}
set A						
OH	96.966(9) ^b	93			96.951	r_e
H_2O^+	99.92(6)	70			99.859	r_e
FH^+	100.105(14) ^b	72			100.078	r_e
NH_2	102.541(120)	94	102.39	79	102.379	r_{ai}
NH_2	102.45(37)	67				
NH^+	106.898(6) ^b	95			106.846	r_e
CO^+	111.522(2) ^b	93			111.555	r_e
O_2^+	111.687(6) ^b	93			111.640	r_e
N_2^+	111.641(5) ^b	93			111.634	r_e
CH	111.791(2) ^{b,c}	96			111.797	r_e
NO	115.0784(14) ^b	93			115.062	r_e
CN	117.18074(4) ^b	93			117.191	r_e
BH_2	{118.1}	97			118.489	r_{ai}
BO	120.475(2) ^b	93			120.530	r_e
BH^+	120.329(133) ^b	98			120.230	r_{ai}
C_2^-	126.829(23) ^b	93			126.852	r_e
CF	127.217(43) ^b	93			127.246	r_e
F_2^+	131.19(38) ^b	76			131.059	r_{ai}
OF	135.41078(1) ^b	93			135.359	r_e

^aNumbers in parentheses are experimental uncertainties with respect to the last digit, e.g., $\pm 0.009\text{ pm}$ for OH. Subscript numbers represent digits that are considered inaccurate. Numbers in braces refer to experimental r_0 structures. ^bRecalculated from B_e given in reference. ^c B_e for CD was used. ^dThe present work.

spectroscopy value of Cormack et al.⁷⁷ is higher than the recent results from other types of spectroscopy, and we used the result of Yang et al.⁷⁶ for ω_e and $\omega_e x_e$.

3. RESULTS AND DISCUSSION

3.1. Geometries. We conducted geometry optimizations for a set of small radicals using several unrestricted wave function methods (HF, MP2, O2-MP2, CC2, CCSD, CCSD(T), CC3, CCSDT, BD, BD(T)), *ab initio* additivity schemes (CCSDTQex, CCSDTex, CCSD(T)ex), density functional methods (B2PLYP-D, M06-2X, M06-HF, BMK, B3LYP, and B97-D), ROHF-based wave function methods (MP2, CCSD, and CCSD(T)), and restricted EOMIP-CCSD calculations. We used Dunning's basis sets, with and without core correlation. We compared the results to reference values—either experimental or theoretical best estimates, depending on the case, as discussed in section 3.1.1. We report the following statistical descriptors: mean signed errors (MSEs), standard deviations (STDs), mean unsigned errors (MUEs), and the maximum (MAX) and the minimum (MIN) deviations. All statistical data are summarized in Table 4. The relevance of accurate geometries for thermochemical studies is briefly discussed in the Supporting Information. For the interpretation of the results, we remind the reader that the convergence of the one-electron treatment (the basis set) and the N-electron treatment (the correlation method) are independent of each other: the quality of an electron correlation method should be evaluated at the CBS limit, including core–valence correlation, to avoid error cancellation between the incompleteness of

Table 2. Experimental and Theoretical Best Estimate Geometry Data Set B [pm]

molecule	bond	$r_e \{r_0\}^a$	ref	r_{se}	ref	r_{ai}^e	r_{ref}
set B							
CH ₃	CH	107.6(1)	68			107.618	r_{ai}
CNC	CN	{124.45(5)} ^{b,c}	99			124.347	r_{ai}
CO ₂ ⁺	CO	{117.68(74)} ^b	100			117.454	r_{ai}
HCC	HC	{121.652(2)} ^d	101	1.069	27	106.281	r_{ai}
HCC	CC	{104.653(1)} ^d	101	1.207	27	120.821	r_{ai}
HCO	HC	111.91(50)	68			111.736	r_{ai}
HCO	CC	117.54(15)	68			117.616	r_{ai}
HCP ⁺	HC	{107.7(2)}	102			108.095	r_{ai}
HCP ⁺	CP	{160.13(3)}	102			159.792	r_{ai}
HNF	HN	{103.5(3)}	70			102.961	r_{ai}
HNF	NF	{137.3(10)}	70			136.886	r_{ai}
HOO	HO	97.07(20)	103			97.027	r_{ai}
HOO	OO	133.054(85)	103			132.975	r_{ai}
N ₃	NN	{118.115}	104			117.804	r_{ai}

^aNumbers in parentheses are experimental uncertainties with respect to the last digit, e.g., ± 0.1 pm for CH₃. Numbers in braces refer to experimental r_0 structures. ^bRecalculated from B_0 given in reference. ^cUncertainty taken from the rotational constant of ¹³CNC. ^dExperimental r_s structure. ^eThe present work.

Table 3. Experimental Frequencies [cm⁻¹]

molecule	ω_e	$\omega_e x_e$	ref
OH	3737.8	84.9	93
FH ⁺	3090.5	89.0	72
CH	2860.8	64.4	93
N ₂ ⁺	2207.0	16.1	93
CO ⁺	2214.1	15.1	93
CN	2068.6	13.1	93
O ₂ ⁺	1906.0	16.5	93
NO	1904.1	14.1	93
BO	1885.3	11.7	93
CF	1307.9	11.1	93
F ₂ ⁺	1091.5	9.0	76
OF	1053.0	9.9	93

correlation treatment, incompleteness of basis set, and/or core correlation treatment. However, the usefulness of a method/basis set combination often benefits from such error cancellation. As many methods produce similar error spreads (STDs), it is of some practical importance to find, among those, the method with the smallest bias (MSE), resulting in a good performance (MUE). Good performance often depends on error cancellation, and this is not necessarily indicative of the quality of the electron correlation method used.

3.1.1. Choice of Reference Values. CCSDTQex Bond Lengths.

For most of the diatomics in set A, highly accurate experimental bond lengths are available. Calculated CCSDTQex r_{ai} values show no significant bias and are within ~ 0.05 pm of experimental values in all cases, except BH⁺ and F₂⁺ (deviations of -0.10 and -0.52 pm, respectively). For F₂⁺ and BH⁺, the experimental bond lengths have a significant uncertainty, and the computed values are within this error.⁷⁸ The r_e structure of H₂O⁺ is reproduced well by CCSDTQex, and the experimental (r_0) BH₂ geometry is rather inaccurate. There is a significant discrepancy between theory and experiments for NH₂ (-0.16 pm), discussed below. On the basis of these results, we assign an uncertainty of ~ 0.05 pm to the CCSDTQex values, which is sufficient accuracy to enable useful comparisons to more approximate methods.

For set B, experimental r_e structures are available for CH₃, HCO, and HOO. For these molecules, CCSDTQex values fall within the rather large experimental uncertainties. For comparison to previous reports: Heckert et al.³¹ used the CCSDTQex model chemistry to study a set of small molecules of similar size, including three radicals (HF, H₂O, NH₃, CH₄, N₂, F₂, CO, HCN, HNC, HCCH, CO₂, OH, NH₂, CH₂, CN, and NO). They reported an MUE of 0.027 pm, MSE of -0.002 pm, STD of 0.036 pm, and a maximum deviation of 0.094 pm, close to the errors we observed in cases where reliable experimental data were available.

Angles. For the systems where angles are present (HCO, HNF, HOO, H₂O⁺, NH₂, BH₂), the angles computed by CCSDTQex are within 2° of the experimental values in all cases. The complete Z matrices of the computed geometries are given in the Supporting Information.

Selection of Reference Geometries. For diatomics, the experimental r_e values were chosen as the reference, except for BH⁺ and F₂⁺, based on the considerations given above. For H₂O⁺, we chose the experimental structure as the reference. For BH₂ and CH₃, we found the experimental values insufficiently accurate, and we chose the CCSDTQex structures. For NH₂, a significant discrepancy is found between the experimental r_e and calculated r_{ai} structures, but the calculated structure agrees well with a semiempirical structure reported by Demaison et al.,⁷⁹ so we chose the r_{ai} structure as the reference. All experimental r_e structures in set B have uncertainties of at least 0.1 pm, and hence we chose the CCSDTQex values as references, all of which fell within the experimental uncertainties. For systems with an experimental r_0 structure, CCSDTQex structures were chosen as the reference. Another strategy would be to calculate r_{se} structures for these compounds, but we consider the CCSDTQex structures to be accurate enough for the purpose of the present study.

3.1.2. Coupled Cluster Including Iterative Triples. The CCSDTQex model shows good performance (MUEs of 0.17 and 0.10 pm for sets A and B, respectively) and small bias (MSEs of -0.14 and -0.10 pm), slightly worse than the closed shell CCSDT/PVQZ results of Halkier et al.⁵¹ Pure CCSDT calculations were only carried out for set A. CCSDT typically

Table 4. Error in Bond Length with Respect to r_{ref} (see Tables 1 and 2), for Set A (left) and Set B (right) [pm]

method	corr. ^a	basis	MUE		MSE		STD		MAX		MIN	
CCSDT _{ex}	CBS		0.17	0.10	−0.14	−0.10	0.22	0.09	−0.82	−0.28	0.12	0.01
CCSDT	AE	PCVTZ	0.24		0.22		0.19		−0.15		0.63	
CCSDT	AE	PVTZ	0.21		0.04		0.25		−0.40		0.49	
CCSDT	FC	PVTZ	0.44		0.44		0.26		0.01		0.96	
CC3	AE	PVQZ	0.16	0.10	−0.06	0.00	0.18	0.12	−0.37	−0.18	0.40	0.20
CC3	AE	PVTZ	0.25	0.31	0.08	0.13	0.31	0.39	−0.42	−0.44	0.73	0.98
CC3	FC	PVTZ	0.48	0.58	0.48	0.58	0.32	0.36	0.09	0.14	1.25	1.45
CCSD(T)	CBS		0.24	0.16	−0.21	−0.16	0.24	0.15	−0.71	−0.46	0.17	0.00
CCSD(T)	AE	PCVQZ	0.15	0.08	−0.15	−0.07	0.16	0.10	−0.45	−0.30	0.05	0.05
ROCCSD(T)	AE	PCVQZ	0.10	0.06	−0.08	−0.04	0.14	0.08	−0.50	−0.21	0.06	0.05
CCSD(T)	AE	APVQZ	0.18	0.12	−0.17	−0.11	0.16	0.10	−0.48	−0.28	0.06	0.09
CCSD(T)	AE	PVQZ	0.20	0.15	−0.20	−0.13	0.14	0.10	−0.44	−0.29	0.03	0.14
CCSD(T)	AE	APCVTZ	0.27	0.32	0.26	0.32	0.18	0.17	−0.05	0.10	0.72	0.65
CCSD(T)	AE	PCVTZ	0.15	0.22	0.13	0.22	0.16	0.16	−0.10	0.02	0.54	0.63
CCSD(T)	AE	APVTZ	0.22	0.26	−0.12	−0.02	0.30	0.39	−0.83	−0.67	0.38	0.88
CCSD(T)	AE	PVTZ	0.16	0.25	−0.05	0.01	0.20	0.36	−0.42	−0.47	0.35	0.93
ROCCSD(T)	AE	PVTZ	0.21	0.28	0.02	0.05	0.26	0.36	−0.42	−0.44	0.47	0.91
BD(T)	AE	PVTZ	0.19	0.25	−0.04	−0.01	0.24	0.34	−0.42	−0.45	0.33	0.84
CCSD(T)	FC	PVTZ	0.35	0.46	0.35	0.46	0.22	0.33	0.07	0.12	0.86	1.41
ROCCSD(T)	FC	PVTZ	0.43	0.50	0.43	0.50	0.26	0.33	0.05	0.12	0.94	1.39
CCSD(T)	FC	PVDZ	1.74	1.78	1.74	1.78	0.77	0.64	0.63	0.82	3.55	3.34
ROCCSD(T)	FC	PVDZ	1.82	1.82	1.82	1.82	0.75	0.67	0.63	0.82	3.48	3.33
CCSD	AE	PCVTZ	0.52	0.34	−0.48	−0.33	0.62	0.26	−2.46	−0.91	0.14	0.00
CCSD	AE	PVTZ	0.64	0.52	−0.64	−0.52	0.53	0.26	−2.48	−1.12	−0.20	−0.09
CCSD	FC	APVTZ	0.36	0.13	−0.17	−0.03	0.57	0.19	−2.05	−0.32	0.34	0.37
CCSD	FC	PVTZ	0.36	0.17	−0.24	−0.08	0.59	0.25	−2.23	−0.67	0.32	0.38
ROCCSD	FC	PVTZ	0.37	0.18	−0.26	−0.10	0.61	0.25	−2.32	−0.66	0.32	0.33
CCSD	AE	PVDZ	1.18	1.24	1.16	1.24	0.64	0.62	−0.14	0.29	2.30	2.25
CCSD	FC	PVDZ	1.23	1.32	1.22	1.32	0.66	0.64	−0.09	0.35	2.40	2.36
B-CCD	AE	PVTZ	0.75	0.64	−0.75	−0.64	0.61	0.33	−2.74	−1.36	−0.21	−0.25
EOMIP	AE	APCVQZ	0.71	0.70	−0.70	−0.68	0.63	0.59	−2.25	−2.03	0.16	0.14
EOMIP	AE	APCVTZ	0.45		−0.36		0.53		−1.69		0.48	
EOMIP	FC	APVTZ	0.37	0.27	−0.14	−0.13	0.51	0.32	−1.49	−0.65	0.76	0.46
EOMIP	FC	PVTZ	0.39	0.41	−0.35	−0.36	0.53	0.41	−2.18	−1.42	0.23	0.19
EOMIP	FC	PVDZ		1.03		0.93		0.69		−0.66		1.76
CC2	AE	PCVTZ	0.82	0.58	−0.35	−0.53	0.93	0.45	−1.93	−1.67	1.48	0.35
CC2	AE	PVTZ	0.92	0.76	−0.50	−0.68	0.98	0.52	−1.93	−1.68	1.35	0.33
CC2	FC	PVTZ	0.85	0.45	−0.12	−0.27	1.07	0.50	−2.22	−1.34	1.89	0.67
MP2	AE	PCVTZ	1.46	0.90	−0.47	−0.66	2.00	1.18	−4.77	−3.71	3.78	1.34
MP2	FC	PVTZ	1.43	0.84	−0.28	−0.43	2.05	1.22	−4.54	−3.47	4.07	2.02
ROMP2	FC	PVTZ	1.44	0.63	0.73	0.19	2.19	0.85	−2.39	−1.90	6.59	1.37
O2-MP2	AE	PVTZ	0.49	0.49	−0.40	−0.49	0.48	0.21	−1.50	−0.81	0.46	−0.14
HF		PVTZ	2.73	1.81	−2.73	−1.81	2.13	0.98	−9.78	−3.36	−0.12	−0.37
B2PLYPD	FC	PVTZ	0.30	0.18	0.16	0.03	0.33	0.23	−0.51	−0.53	0.75	0.37
BMK		PVTZ	0.95	0.62	−0.71	−0.38	1.14	0.68	−4.15	−1.36	0.93	0.74
M062X		PVTZ	1.18	0.69	−0.93	−0.55	1.46	0.75	−5.01	−2.10	0.82	0.40
M06HF		PVTZ	1.73	1.15	−1.60	−1.15	1.89	0.90	−6.62	−3.17	0.68	−0.02
B3LYP		PVTZ	0.68	0.34	−0.15	0.01	0.82	0.42	−1.62	−0.88	1.32	0.70
B97D		PVTZ	0.89	0.87	0.80	0.87	0.76	0.37	−0.53	0.24	2.30	1.38
M06HF		PVTZ	1.73	1.15	−1.60	−1.15	1.89	0.90	−6.62	−3.17	0.68	−0.02
B3LYP		PVTZ	0.68	0.34	−0.15	0.01	0.82	0.42	−1.62	−0.88	1.32	0.70
B97D		PVTZ	0.89	0.87	0.80	0.87	0.76	0.37	−0.53	0.24	2.30	1.38

^aAE, all electrons correlated; FC, frozen core electrons; CBS, additivity scheme, see text for details.

performed worse compared to the above extrapolation schemes, with MUEs ranging from 0.20 to 0.45 pm. As the

CCSDT contribution dominates the computational cost of the CCSDT_{ex} model, CCSDT_{ex} or a similar scheme is

recommended instead of pure CCSDT. The CC3 model performs only slightly worse than CCSDT using the same basis sets; the PVQZ basis improves significantly over the triple- ζ basis sets.

3.1.3. Coupled Cluster Including Perturbative Triples. CCSD(T)ex exhibits worse performance than CCSD(T) (MUEs of 0.21 and 0.16 pm), due to the larger bias (MSEs of -0.20 and -0.16 pm). CCSD(T)ex exhibits a larger bias for set A than for set B, due to errors for CO^+ , O_2^+ , NO, CN, F_2^+ , and OF, several of which are systems known to be challenging to CCSD(T).¹⁷

For regular CCSD(T) calculations, we find trends similar to those reported by Bak et al.⁷ for closed shell species, although with slightly worse overall performance: quadruple- ζ basis sets yield MUEs down to ~ 0.1 pm, and triple- ζ basis sets perform slightly worse when used in AE calculations and significantly worse in FC calculations, exhibiting a positive bias of ~ 0.4 pm. Using double- ζ basis sets produces a large positive bias. Similar to Bak et al.'s results,⁷ the use of the appropriate core correlation basis (PCVXZ) produces only small changes in bond lengths compared to AE calculations with the PVXZ basis sets.

AE-CCSD(T) calculations with quadruple- ζ basis sets performed similar to or better than the CCSD(T)ex additivity scheme, and triple- ζ calculations only slightly worse. Also, quadruple- ζ results are only slightly worse than the CCSD(T)ex scheme, so the former should only be used for specific systems that have already been diagnosed to benefit from a CCSDT contribution (e.g., NO, CN, CO^+).

UCCSD(T) and ROCCSD(T) results do not differ much: using the PVTZ basis, UCCSD(T) seems slightly more appropriate, whereas with the PCVQZ basis, ROCCSD(T) performs slightly better. The differences in bond length between the two methods are usually <0.07 pm, with the exception of CN, CO^+ , HCC, and OF. The use of a Brueckner reference function (AE-BD(T)/PVTZ) does not improve statistics compared to AE-CCSD(T)/PVTZ.

Using the PVQZ basis, AE-CCSD(T) performs slightly worse than CC3, and the latter may be considered as an alternative, but current implementations lack analytical gradients. With triple- ζ basis sets, there is no statistical benefit from using iterative triples. From a cost/benefit point of view, AE-CCSD(T) with PVTZ or PVQZ bases provide good results (MUEs between 0.15 and 0.25 pm) at a rather moderate cost, with the PVQZ basis being clearly better for larger systems. In comparison, Bak et al.⁷ reported results for closed shell geometries, yielding MUEs of 0.20 and 0.13 pm for AE-CCSD(T) with the PVTZ and PVQZ basis sets, respectively.

3.1.4. Coupled Cluster with Doubles Excitations. There is significant variation among the performances achieved by the different CCSD computations. CCSD is typically less accurate than CCSD(T) and also shows increased error spread (STD). FC-CCSD/APVTZ performs best (MUE of 0.36 and 0.13 pm for sets A and B), apparently due to error cancellation of basis set incompleteness, frozen core electrons, and limited correlation treatment. All CCSD calculations have similar error spreads, so the performance is dictated by the bias (all models are biased toward short bond lengths). As with CCSD(T), the use of double- ζ basis sets results in an overestimation of bond lengths by >1 pm and a slightly increased standard deviation. The differences between ROCCSD and UCCSD are small: using FC-CCSD/PVTZ, the differences are <0.07 pm, except for CN, CO_2^+ , N_3 , CNC, and F_2^+ . The use of Brueckner orbitals slightly worsens the performance compared to AE-CCSD/

PVTZ; we observed an increase in both systematic error and standard deviation for sets A and B.

FC-EOMIP-CCSD/APVTZ is almost as accurate as FC-CCSD/APVTZ and more accurate than the other CCSD models tested, exhibiting MUEs of 0.36 and 0.27 pm for sets A and B. The use of the corresponding FC-PVTZ or AE-APCVTZ methods results in slightly worse accuracy, the latter indicating error cancellation in the FC-EOMIP-CCSD/APVTZ results. Geometries of some systems (e.g., OF, CF, CN, NO) show a strong dependence on basis set and core correlation. For notoriously difficult cases such as the nitrate radical, only EOMIP coupled cluster and multireference methods give a qualitatively correct geometry.^{35,36,80} Hence, for the systems studied here, it is encouraging that the EOMIP-CC formalism gives very reasonable geometries with only doubles excitations in combination with triple- ζ basis sets. Finally, CC2 is much less accurate than the above methods, with errors about twice as large as for CCSD.

3.1.5. Perturbation Theory. Among the tested perturbation theory methods, O2-MP2 is the most accurate, with MUEs of 0.49 for both sets A and B, somewhat worse than the best CCSD results, but better than CC2 for set A. Conventional MP2, either with UHF or ROHF references, with and without core correlation, is much less accurate (MUEs between 0.6 and 1.5 pm), followed by HF-SCF (MUEs of 2.73 and 1.81 pm).

For conventional MP2, MUE and STD values for sets A and B are roughly 2- to 4-fold larger than those observed for CCSD when a comparable basis set and frozen core approximation are used. Like CCSD, MP2 shows a smaller bias for the frozen core calculation, indicating cancellation of errors.

Compared to conventional MP2, the O2 methodology significantly improves results for most of the systems in sets A and B (errors in CO_2^+ , the CC bond in HCC, HCP+, the OO bond in HOO, CO^+ , O_2^+ , N_2^+ , NO, CN, BO, F_2^+ , and OF are reduced to <1 pm), but it also worsens some results compared to conventional MP2 (errors in CH, NH^+ , HNF, and the HC bond in HCC are increased to >1 pm). Overall, using optimized orbitals decreases the largest errors and constitutes a significant improvement over conventional MP2; errors approach conventional MP2 results for closed shell geometries.⁷ The implementation of O2-MP2 that we used (RI-MP2 with scaled opposite spin and optimized orbitals) is formally a fourth-order scaling method, but with a large prefactor because a perturbative calculation is carried out at each SCF cycle. Therefore, it is cost-competitive with conventional MP2 only for large systems. Like MP2, the CC2 approach is a fifth-order scaling method. CC2 improves significantly over the MP2 results, but it still does not reach "closed shell MP2" performance.

3.1.6. Hybrid, Meta Hybrid, and Double Hybrid DFT. Calculations with the B2PLYP-D, BMK, M06-2X, M06-HF, B3LYP, and B97-D density functional methods were conducted. Although B2PLYP-D was originally calibrated in conjunction with quadruple- ζ basis sets,³⁸ it performs remarkably well with the PVTZ basis, yielding an accuracy *en par* with the best CCSD results, with MUEs of 0.30 and 0.18 pm for sets A and B and little bias. The meta hybrid functionals are less accurate, exhibiting MUEs between 0.6 and 1.7 pm and pronounced negative bias. B3LYP performs better than the meta hybrids but does not rival CCSD. B97-D yields a bias of ~ 0.9 pm, but the bond lengths have a relatively low standard deviation. The B97-D MUEs are *en par* with the meta hybrid

functionals for set A, but worse for set B (except M06-HF). We briefly discuss the appropriateness of DFT-derived geometries for high-accuracy thermochemical calculations in the Supporting Information.

3.1.7. Comparison to Results Reported by Byrd et al.¹² The results obtained in the present study, using our reference geometries, differ significantly from results reported previously for some methods and basis sets. In the present study, identical methods were found to perform better for predicting geometries, on average, compared to the Byrd statistics. The Byrd et al. geometry test set differs from ours not only by the use of different reference geometries for the same systems but also by incorporating two more systems which we did not include in our test set: HCCCCCH⁺ and CH₃O. With the Byrd reference set, our calculations using AE-CCSD(T)/PCVQZ would yield an MUE of 0.51 pm, whereas we found MUEs of 0.15 and 0.08 pm for our improved reference sets A and B, respectively. Hence, the improved reference set is crucial for showing that AE-CCSD(T) performs significantly better than CCSD methods.

In Table 8, we report statistics for our frozen core results, employing methods identical to those used by Byrd et al. and with respect to the reference geometries used by Byrd et al., including CH₃O and HCCCCCH⁺. The following discussion refers to the Byrd et al. reference values throughout. The statistical descriptors reported by us and Byrd et al., using identical methods, differ at least slightly in all cases. This can possibly be explained by convergence of the reference wave function, for some systems, to a different electronic state in the two studies. However, the striking differences between the two sets of results in Table 8 are the performances of UHF versus ROHF based methods, especially for CCSD(T) and MP2. Byrd et al. reported that UCCSD(T) and UMP2 perform significantly worse than ROCCSD(T) and ROMP2. In contrast, we found that the reference function has only a small influence on the MUE, and that UCCSD(T) does not perform worse than UCCSD or ROCCSD. A possible rationalization is that for some systems, Byrd et al. may have used reference wave functions corresponding to different electronic states for UHF versus ROHF, whereas we ensured the same occupation in both cases. Figure 1 and 2 in Byrd et al.¹² support this hypothesis: for some systems, e.g., HCO, HCP⁺, CO₂⁺, N₃, and BO, differences in bond length between UCCSD(T) and ROCCSD(T) are >1 pm, whereas we observed maximum differences of <0.5 pm using identical methods. However, following this argument, the resemblance of the ROCCSD and ROCCSD(T) statistics between the previous and present studies would have to be coincidental. Also, perturbative methods (MP2, CCSD(T)) are more sensitive to the choice of reference function, and the differences in properties resulting from restricted and unrestricted calculations can be expected to be more pronounced when using these methods. Finally, we cannot rule out a software bug and/or limitation for frozen core gradients in the unspecified revision of the Florida-ACESII version used by Byrd et al.

3.2. Harmonic Frequency Calculations. We calculated harmonic frequencies for the same set of diatomics that Beran et al.¹⁵ studied, and we compared these values to true harmonic frequencies extracted from experimental data. All results are shown in Table 5. The error distributions are heavily influenced by a few outliers. Hence, for each method, we report the errors for individual systems, as well as the mean signed error (MSE), standard deviation (STD), and

mean unsigned error (MUE). The impact of errors in vibrational frequencies on thermochemistry is briefly discussed in the Supporting Information.

3.2.1. Coupled Cluster Including Triples Excitations and Additivity Schemes. We applied the CCSDTQex additivity scheme as used for geometries to the frequency test set, using additivity of the force constant. This led to good agreement with experimental results: the largest discrepancies are FH⁺ (14 cm⁻¹), F₂⁺ (9 cm⁻¹), and O₂⁺ (12 cm⁻¹). The CCSDTQex method appears to obtain slightly more reliable results for closed shell systems: Heckert⁵² applied it to calculate frequencies of a set of closed-shell molecules (HF, H₂O, CH₂, N₂, F₂, CO, HCN, HNC) and reported a MSE of -3.8, a MUE of 3.8, a STD of 2.5, and a maximum error of 8.6 cm⁻¹ with respect to experimental fundamental frequencies, where the anharmonicity was estimated at the AE-CCSD(T)/PCVQZ level. Tew et al.⁸¹ found that the lack of true basis set convergence on harmonic frequencies at the CCSD(T)/PV6Z level can result in errors of up to 3.5 cm⁻¹, a magnitude beginning to be relevant at the level of accuracy achieved by CCSDTQex. For comparison, W4 harmonic frequencies yielded a MUE < 1 cm⁻¹ for a set including 26 closed shell diatomics, CH, and OH.³⁰

Among the iterative triples methods, CCSDT/PVTZ performs best, and CC3 describes NO, CN, and CO⁺ significantly worse. By comparison, unrestricted CCSD(T) encounters (known) difficulties for the NO, CO⁺, and CN systems,¹⁷ irrespective of the basis set and treatment of core correlation. These systems were described slightly better by CC3. CCSD(T) results are comparable to those reported by Beran et al.,¹⁵ who employed experimental geometries to obtain computed force constants and resulting frequencies.

Using a ROHF reference improves the treatment of problematic systems (NO, CO⁺, CN, OF), which can in all cases be attributed to artifacts of the UHF reference.¹⁷ Using a Brueckner reference (AE-BD(T)) also improves upon UCCSD(T) by eliminating the outliers (NO, CO⁺, CN) and produces an accuracy similar to ROCCSD(T), likely for the same reasons.

3.2.2. Coupled Cluster with Doubles Excitations and Perturbation Theory. Compared to the above methods, CCSD has problems with a greater number of systems, and the use of a Brueckner reference does not ameliorate any of these failures. For most systems, BD provides poorer frequencies than CCSD, with deviations from experimental results up to ~160 cm⁻¹. The typical performances of CCSD and BD are poor compared to CCSD(T), with the exception of NO. EOMIP-CCSD does not perform significantly better than CCSD for the diatomics studied here. For EOMIP-CCSD, the use of the APVTZ basis removes some of the problem cases encountered with the PVTZ basis (OH, FH⁺), but frequencies for N₂⁺, CO⁺, and OF remain poor. The use of a quadruple- ζ basis with core correlation worsens the bias and does not improve the accuracy (STD). MP2 results are significantly worse than those of other methods in this study. However, previously reported O2-MP2 frequencies³⁸ are much more accurate compared to conventional MP2. O2-MP2 produces a MUE *en par* with the best CCSD results, but with a larger error spread.

3.2.3. Density Functional Theory. Among the tested DFT methods, B2PLYP-D performs best for frequencies, exhibiting an MUE of only 19 cm⁻¹, little bias, and an error spread smaller than the CCSD methods. It clearly outperforms almost all methods considered in this study, with the exceptions of the

Table 5. Errors in Harmonic Frequencies with Respect to Experiment [cm^{-1}]

method	corr. ^a	basis	OH	FH ⁺	CH	FH ⁺	CH	N ₂ ⁺	CO ⁺	CN	BO	CF	F ₂ ⁺	OF	MSE	STD	MUE
CCSDTQex	CBS	CBS	6	14	2	7	3	4	12	6	-1	0	9	3	5	5	6
CC3	AE	PVQZ	16	27	7	-1	46	65	24	130	-24	-5	35	29	29	40	34
CCSDT	AE	PVTZ	22	64	7	15	-4	22	41	4	0	20	39	2	19	20	20
CC3	AE	PVTZ	24	63	10	-7	37	62	14	135	-30	5	36	30	32	42	38
CCSDT	FC	PVTZ	4	47	-24	-6	-22	2	25	-7	-20	8	37	-2	4	23	17
CC3	FC	PVTZ	6	47	-21	-28	21	43	-2	132	-50	-7	34	31	17	47	35
CCSD(T)	AE	PCVQZ	17	29	1	13	87	68	35	203	13	10	30	37	45	56	45
ROCCSD(T)	AE	PCVQZ	17	29	1	11	6	8	36	15	4	10	33	16	15	12	15
CCSD(T)	AE	PVQZ	18	31	7	16	89	70	36	203	16	15	32	37	48	55	48
CCSD(T)	AE	PCVTZ	5	48	-15	-3	70	56	12	210	2	19	32	39	40	60	43
CCSD(T)	AE	PVTZ	26	66	10	9	78	66	25	215	11	24	31	38	50	57	50
ROCCSD(T)	AE	PVTZ	25	66	9	7	-4	6	26	15	1	24	34	16	19	19	19
BD(T)	AE	PVTZ	28	69	11	16	14	14	37	26	11	27	46	25	27	17	27
CCSD(T)	FC	PVTZ	7	50	-21	-13	62	45	9	219	-8	13	29	37	36	63	43
CCSD(T)	FC	PVDZ	-34	11	-45	-23	46	22	23	154	-34	-18	-132	-43	-6	68	49
ROCCSD(T)	FC	PVDZ	-34	10	-45	-25	-36	-34	24	10	-47	-18	-129	-77	-33	41	41
CCSD	AE	PCVTZ	39	75	5	76	93	103	134	93	55	46	143	59	77	39	77
CCSD	FC	APVTZ	19	54	2	60	78	86	120	78	35	19	132	45	61	40	61
ROCCSD	FC	APVTZ	19	55	2	68	83	65	125	71	35	18	136	51	61	41	61
CCSD	FC	PVTZ	40	75	-3	64	80	89	127	85	44	39	136	52	69	39	69
B-CCD	AE	PVTZ	62	95	30	105	122	88	160	120	78	61	152	73	96	38	96
CCSD	FC	PVDZ	-13	23	-30	48	55	58	127	57	15	5	-20	-52	23	49	42
EOMIP	AE	APCVQZ	61	85	53	150	113	103	116	99	69	25	80	147	92	37	92
EOMIP	FC	APVTZ	32	84	34	123	79	77	80	69	39	17	70	137	70	36	70
EOMIP	FC	PVTZ	105	102	51	126	81	76	85	78	45	64	73	153	87	31	87
EOMIP	FC	PVDZ	68	54	40	118	65	-170	87	86	16	41	-94	84	33	83	77
MP2	FC	PVTZ	83	62	83	-110	660	827	-413	1434	195	35	-213	332	248	507	371
ROMP2	FC	PVTZ	80	64	83	-180	-112	-264	-532	7	211	37	-315	54	-72	213	162
O2-MP2	AE	6-311G (2df2pd)	135	90	94	-51	8	-22	-39	32	21	45	-48	35	25	60	52
B2PLYPD	FC	PVTZ	14	-30	14	-4	14	10	-70	-3	-10	5	-18	69	-1	33	22
BMK		PVTZ	61	-42	-24	27	190	139	227	116	88	46	276	137	103	97	115
M06-2X		PVTZ	13	-40	16	194	180	168	305	173	89	50	289	134	131	108	138
M06-HF		PVTZ	16	-64	42	233	288	281	459	282	-598	9	384	146	123	279	234
B3LYP		PVTZ	-40	-98	-42	125	77	85	144	72	33	-2	143	65	47	78	77
B97D		PVTZ	-118	-121	-121	74	-23	24	43	-4	-60	-53	62	25	-23	72	61

^aAE, all electrons correlated; FC, frozen core electrons; CBS, additivity scheme, see text for details.

much more expensive ROCCSD(T), CCSDT, and BD(T). Among the remaining DFT methods, the same trend as for the geometries is observed: the meta hybrid functionals perform poorly compared to CCSD, whereas B3LYP produces more reliable results, only slightly worse than CCSD methods. B97-D produces accuracy comparable to B3LYP, possibly due to the absence of exact exchange, the latter being probably more prone than GGA methods to (near-) instabilities of the orbitals.⁸²

3.2.4. Scaling Factors. Some of the frequency results show a considerable bias but a relatively small error spread. For these methods (BD(T), BD, CCSD, EOMIP-CCSD), practical applications can possibly benefit from empirical frequency scaling factors. With the advent of useful semiempirical methods, it was noted that computed vibrational frequencies benefit from empirical corrections,^{83,84} and a systematic procedure for frequency scaling was devised a few years later.⁸⁵ To date, empirical frequency scaling factors for closed shell systems are available for many methods, and these correct the computed harmonic frequency to reproduce either harmonic experimental⁸⁶ or anharmonic experimental⁸⁷ vibrations. Here, we propose such scaling factors, although they are preliminary and should be validated on a larger data set including polyatomic systems. We chose to report scaling factors to correct for the intrinsic error of the model chemistry only, i.e., to achieve agreement between computed and experimental harmonic frequencies. In other words, the scaling factors do not correct for anharmonicity. Factors were fitted to minimize the MUEs and rounded to three decimal places, resulting in MSEs $<1\text{ cm}^{-1}$. The resulting scaling factors and associated error statistics are reported in Table 6. Bearing in mind the limitations of a small

Table 6. Proposed Scaling Factors to Correct Discrepancies between Calculated and Experimental Harmonic Frequencies [cm^{-1}]

method	corr. ^a	basis	SF	STD	MUE	MAX ^b
BD(T)	AE	PVTZ	0.987	19	16	31
CCSD	AE	PCVTZ	0.965	56	40	99
CCSD	FC	APVTZ	0.972	52	40	98
B-CCD	AE	PVTZ	0.957	59	42	98
EOMIP-CCSD	AE	APCVQZ	0.958	54	43	97
EOMIP-CCSD	FC	APVTZ	0.968	49	36	98
EOMIP-CCSD	FC	PVTZ	0.961	44	30	106

^aAE, all electrons correlated; FC, frozen core electrons; CBS, additivity scheme, see text for details. ^bLargest errors after scaling.

diatomic data set, these results suggest that scaling factors will improve the practical performance of these methods for open shell systems.

3.3. Anharmonic Frequencies. We used the VPT2²⁶ method to calculate the first anharmonicity constant, $\omega_e x_e$, for the systems given in Table 3 with different coupled cluster methods. Errors with respect to experiment are given in Table 7.

3.3.1. CCSDT and CCSD(T). The $\omega_e x_e$ values of the AE-CCSDT/PVTZ model are in good agreement with experimental results, with the important exception of NO. The $\omega_e x_e$ of NO has the wrong sign; i.e., it is qualitatively wrong. The FC-CCSDT/PVTZ results are of similar quality as for the AE calculation. The UCCSD(T) models perform worse than CCSDT and produce nonsensical anharmonicity constants for NO and a qualitatively incorrect $\omega_e x_e$ value for OF. However, for the remaining systems, CCSD(T) yields good agreement

with experimental results. The inclusion of core correlation had only little influence on performance for the present test set. As with harmonic frequency results, using a ROHF reference fixes the problems encountered for anharmonicity constants of NO and OF.

3.3.2. CCSD, EOMIP-CCSD, and ROMP2. UCCSD produces somewhat more reliable anharmonicity constants than UCCSD(T). Results for OF are in better agreement with experimental results and are close to the CCSDT values. The results for NO, although qualitatively wrong, are improved upon in comparison to CCSD(T). For the remaining systems, CCSD produces errors typically $<2\text{ cm}^{-1}$, whereas CCSD(T) yielded errors $<1\text{ cm}^{-1}$. Again, FC and AE calculations yielded almost identical results. Only RO-CC and EOMIP-CC methods yielded qualitatively correct results for all systems considered, including NO. For the systems that are already well-described by CCSD, EOMIP-CCSD yields only slightly worse results. On the basis of the promising ROCCSD(T) results, we also tested ROMP2, which gave anharmonicity constants of low quality.

3.4. Identifying Problematic Cases for CCSD(T). Frequency calculations using perturbative methods are especially sensitive to the reference wave function,^{15,16} resulting in large errors for ω_e and $\omega_e x_e$. CCSD(T) may be the best affordable method for many systems. Here we summarize strategies for detecting possible problems with CCSD(T) applied to frequencies, for cases where computationally more demanding methods are not feasible. Problems can be caused by different shortcomings of the reference wave function:

- Instabilities or near-instabilities can be present in a UHF reference function, as discussed in detail by Crawford et al.⁸⁸ and Szalay et al.¹⁷ To diagnose near-instabilities of the reference, the smallest absolute eigenvalues of the molecular orbital (MO) Hessian can be used,^{16,17} as available, e.g., in the CFOUR and PSI3⁸⁹ packages. The diagnosis is somewhat delicate, as to our knowledge no “threshold” value has been devised below which a system is to be classified as problematic. Detailed analysis of the orbitals as a function of molecular geometry¹⁷ can provide more information but may be difficult for polyatomic systems. For the present frequency test set, the smallest absolute eigenvalues (UHF) observed were 0.025 (NO), 0.036 (CH), 0.038 (O_2^+), 0.041 (CF), and 0.047 (F_2^+) (atomic units, see Supporting Information for a complete list). Of these, NO is known to be a pathological case of “triplet instability”.¹⁷ The OF system also suffers from such problems, but this is only apparent when analyzing the MO Hessian eigenvalues as a function of bond length.¹⁷ These problems can be solved by using a ROHF or Brueckner reference instead, which removed the outliers at the UCCSD(T) level in our frequency test set. However, other problems, not dealt with in the present work, can also occur.
- The system may exhibit multireference character/non-dynamical correlation; that is, a single Slater determinant describes the system insufficiently. The quality of the (single) reference can be judged by the T1 diagnostic,⁹⁰ and a poor reference may or may not be caused by multireference character of the system. From UCCSD results, we find that the T1 values of CN (0.07) and CO^+ (0.05) are problematic, and there are several systems (OF, NO, BO) with values around the 0.02 threshold, indicating the onset of questionable quality of the

Table 7. Errors in $\omega_e x_e$ with Respect to Experimental Results [cm^{-1}]

method	corr. ^a	basis	OH	FH ⁺	CH	N ₂ ⁺	CO ⁺	CN	O ₂ ⁺	NO	BO	CF	F ₂ ⁺	OF	MSE	STD	MUE
CCSDT	AE	PVTZ	-0.6	6.7	-0.6	-0.3	-1.2	-0.5	-1.5	-27.3	-0.2	-0.3	-0.2	4.1	-1.8	8.4	3.6
CCSDT	FC	PVTZ	-0.6	5.3	-1.0	-0.3	-1.3	-0.5	-1.5	-10.4	-0.4	-0.3	-0.2	3.0	-0.7	3.6	2.1
CCSD(T)	AE	PCVQZ	-0.5	5.7	-0.6	-0.4	-0.8	0.0	-1.2	183.6	-0.9	-0.3	0.1	-21.3	13.6	53.9	17.9
ROCCSD(T)	AE	PCVQZ	-0.4	11.1	-0.7	-0.4	0.0	-0.2	-1.2	-1.0	-0.1	-0.3	0.0	0.8	0.6	3.3	1.4
CCSD(T)	AE	PCVTZ	-1.7	5.5	-0.9	-0.3	-0.6	-0.1	-1.2	432.6	-1.1	-0.2	0.3	-23.2	34.1	125.7	39.0
CCSD(T)	AE	PVTZ	-1.0	6.7	-0.9	-0.2	-0.5	-0.1	-1.0	309.0	-1.1	-0.4	0.3	-21.6	24.1	90.0	28.6
CCSD(T)	FC	PVTZ	-1.0	5.2	-1.4	-0.3	-0.2	-0.3	-1.1	-572.0	-1.3	-0.4	0.3	-24.2	-49.7	164.6	50.6
CCSD(T)	FC	PVDZ	2.6	-0.4	2.7	0.3	0.3	-1.0	0.9	259.4	-1.4	-1.0	-0.1	5.1	22.3	74.7	22.9
ROCCSD(T)	FC	PVDZ	2.6	-0.4	2.7	0.3	-0.2	-0.1	0.8	0.3	-0.1	-1.0	-0.2	1.5	0.5	1.2	0.8
CCSD	AE	PCVTZ	-3.1	5.5	-2.1	-1.2	-1.4	-0.3	-2.7	-67.8	-1.1	-0.4	-0.8	-4.5	-6.7	19.4	7.6
CCSD	FC	APVTZ	-4.3	4.8	-2.9	-1.2	-1.3	-0.4	-2.6	-60.4	-1.0	-0.3	-0.9	-5.7	-6.4	17.2	7.2
ROCCSD	FC	APVTZ	-4.4	4.8	-2.9	-1.2	-1.5	-0.9	-2.6	-1.2	-1.0	-0.3	-1.0	-1.3	-1.1	2.2	1.9
CCSD	FC	PVTZ	-2.3	5.1	-2.5	-1.2	-1.4	-0.5	-2.5	-62.4	-1.2	-0.5	-0.8	-4.2	-6.2	17.8	7.0
CCSD	FC	PVDZ	1.0	-0.5	1.5	-0.6	-1.4	-0.9	-0.9	-30.7	-1.1	-1.2	-1.0	-14.3	-4.2	9.3	4.6
ROCCSD	FC	PVDZ	-2.4	5.1	-2.6	-1.2	-1.6	-1.0	-2.6	-1.1	-1.1	-0.5	-0.8	-1.0	-0.9	2.0	1.7
EOMIP	AE	APCVQZ	-4.6	-2.2	-6.5	-2.9	-0.7	-1.5	-2.2	-0.5	-0.3	-1.4	-0.6	-4.3	-2.3	1.9	2.3
EOMIP	FC	APVTZ	-6.8	-2.8	-7.6	-2.8	-0.9	-1.6	-2.2	-1.2	-0.3	-1.3	-0.6	-4.5	-2.7	2.4	2.7
EOMIP	FC	PVTZ	-7.5	-3.3	-8.1	-2.8	-0.9	-1.6	-2.1	-1.7	-0.4	-1.1	-0.4	-4.1	-2.8	2.6	2.8
EOMIP	FC	PVDZ	-3.5	-9.8	-4.5	-2.4	-1.0	-1.9	-0.4	-1.0	-0.3	-1.6	-0.8	-4.2	-2.6	2.7	2.6
ROMP2	FC	PVTZ	-7.3	6.1	-7.9	2.6	0.7	14.2	29.2	-4.8	-1.5	-0.5	19.8	-8.8	3.5	11.9	8.6

^aAE, all electrons correlated; FC, frozen core electrons.

Table 8. Statistical Descriptors of Geometry Errors for the Combined Data Set (Set A, Set B, CH₃O, and C₄H₂⁺) with Respect to Experimental Values Used by Byrd et al.¹² [pm]^a

method	MSE	STD	MUE	MAX
UHF ^b	−1.97	1.92	2.34	−9.22
UHF ^c	−2.11	1.83	2.32	−9.22
ROHF ^b	−2.48	1.96	2.76	−9.44
ROHF ^c	−2.26	3.00	2.32	−9.44
UMP2 ^b	−0.41	1.82	1.40	4.63
UMP2 ^c	−0.45	2.69	1.76	8.68
ROMP2 ^b	0.38	1.77	1.21	7.15
ROMP2 ^c	0.46	1.71	1.15	7.13
UCCSD ^b	−0.19	0.80	0.59	2.66
UCCSD ^c	−0.20	0.80	0.56	2.67
ROCCSD ^b	−0.22	0.81	0.61	2.66
ROCCSD ^c	−0.22	0.81	0.59	2.67
UCCSD(T) ^b	0.33	0.72	0.58	2.85
UCCSD(T) ^c	1.01	1.50	1.25	7.07
ROCCSD(T) ^b	0.38	0.70	0.60	2.85
ROCCSD(T) ^c	0.38	0.70	0.58	2.86
UCCSD(T)/PCVQZ ^{b,d}	−0.15	0.75	0.51	2.78

^aUnless mentioned otherwise, frozen core calculations with the PVTZ basis are reported. ^bThe present work. ^cTaken from Byrd et al.¹² ^dAll-electron calculation.

reference. When using ROCCSD, the T1 values are improved for the most severe cases (CN, 0.04; CO⁺, 0.03) and remain the same or increase slightly for the remaining systems. This may partly contribute to the improved treatment of these systems by ROHF-based methods.

Karton et al.² found the energy-based %TAE(T) diagnostic⁹¹ useful for predicting the importance of post-CCSD(T) contributions to the atomization energy, indicative of nondynamical correlation. The %TAE descriptor is not defined for ions, but in the same spirit, we can define %TBE(T) = ((TBE_{CCSD(T)} − TBE_{CCSD})/TBE_{CCSD(T)}) · 100, where TBE is the total bond dissociation energy calculated using either CCSD or CCSD(T). For neutral diatomics, the %TBE(T) definition is equivalent to %TAE(T). For our frequency test set, %TBE(T) values, calculated using AE-CCSD(T)/PCVTZ, differ by <0.5% between ROHF and UHF references, and we report mean values hereafter. F₂⁺ (14.9%) and OF (13.3%) are predicted to have severe nondynamical correlation. O₂⁺ (7.6%), N₂⁺ (6.9%), NO (6.1%), CN (5.6%), and CO⁺ (4.4%) are predicted to have moderate nondynamical correlation, and CF (3.8%) and BO (3.5%) are predicted to have mild nondynamical correlation. From these data, it is not obvious to what extent the insufficient treatment of nondynamical correlation by CCSD(T) affects the resulting vibrational frequencies: for example, the ROCCSD(T)/PCVQZ frequency for F₂⁺ exhibits significant error, whereas the OF frequency is not worse than the OH frequency. However, strong effects of multireference character on geometries and frequencies cannot be excluded for other systems. In severe cases, for example, those having equally contributing configurations,

single-determinantal methods generally fail, and such systems should be treated more appropriately by EOMIP-CCSD.

- (c) Another issue with (restricted and unrestricted) HF references is the unphysical localization of the unpaired electron. This is commonly referred to as symmetry breaking, but it is not limited to symmetric systems. ROHF references are more prone^{1,16} to this phenomenon. The fitness of the reference can be judged qualitatively by comparing charge or spin distributions of CCSD(T) with those of EOMIP-CCSD. The latter uses a reference based on a closed shell electronic state and produces a spin-pure, qualitatively correct result. Also the use of a Brueckner reference solves the symmetry breaking problem in many (but not all) cases.⁹² We did not encounter this phenomenon in the present study, since the known symmetry breakers (C₂[−], N₂⁺, O₂⁺, F₂⁺) were treated using a D_{2h} symmetric reference function.

We saw that ROCCSD(T) is a robust method for the vibrational frequencies of our test set, which is even more true for EOMIP-CCSD and regular CCSD, albeit at reduced accuracy. As the latter methods are computationally less demanding than CCSD(T), they provide a heuristic way to verify CCSD(T) results, even at moderate basis set size. Depending on the specific problem, UCCSD(T) or ROCCSD(T) calculations may be preferred, and a comparison between UCCSD(T), ROCCSD(T), CCSD, and EOMIP-CCSD results, also using only a moderate basis set size, can help to identify problematic cases prior to a detailed analysis of the specific problem.

4. CONCLUSIONS AND RECOMMENDATIONS

We have evaluated the performance of different popular wave function and DFT methods for the geometries and normal modes of small molecular radicals. The most sophisticated method was an additivity scheme that accounts for CCSD(T) contributions close to the basis set limit and excitations up to CCSDTQ (“CCSDTQex”). CCSDTQex results compared favorably to highly accurate experimental geometries available for some diatomic radicals, and we assigned an uncertainty of ~0.05 pm to CCSDTQex, accurate enough to serve as a benchmark for the evaluation of more approximate methods. CCSDTQex calculations using analytical first and second derivatives are routinely feasible for systems with up to three centers (and potentially slightly larger systems) using current computational resources. However, systems of interest to most practitioners are larger, and thus we evaluated more approximate methods, with the focus on coupled cluster theory.

In contrast to previous results by Byrd et al.,¹² we found that CCSD(T), using both UHF and ROHF references, provides geometries of good quality, close to its performance for comparable closed shell systems. For harmonic and anharmonic frequencies, UCCSD(T) provides good results for most systems, and some severe outliers present in our test set are cured by using a ROHF reference.

CCSDT and BD(T) provide more reliable results for frequencies but at a much elevated cost: due to its N⁸ scaling, CCSDT is prohibitively expensive when using a triple- ζ basis set. BD(T), while formally an N⁷ method, has a larger prefactor than CCSD(T) and lacks analytical derivatives, which results in elevated cost due to the use of numerical first and second derivatives.

Additionally, we have shown that coupled cluster methods based on ROHF references and on the EOMIP formalism are more robust for harmonic and anharmonic frequency calculations. To date, only analytical first derivative calculations are available for both approaches. Also, ROCCSD(T) analytical gradient calculations cannot be parallelized in the current version of CFOUR, in contrast to UCCSD(T) first and second derivatives. (However, a “trivial” parallelization is possible through purely numerical computation of vibrational frequencies from single point energies, implemented in CFOUR.) As our results stress the usefulness of ROCCSD(T) and EOMIP-CCSD frequency calculations, this may motivate the development of analytical second derivatives for these methods, more efficient implementations, and/or analytical derivatives for BD(T) calculations. At present, UCCSD(T) can be a cost-competitive alternative for geometries as well as vibrational frequencies, provided that problematic systems can be identified beforehand, as outlined above.

When the systems size exceeds what can be handled by CCSD(T), we would recommend the B2PLYP-D method for both geometries and normal modes. B2PLYP-D is an apparently reliable alternative to CCSD and produces similar results at lower cost. By comparison, recent meta hybrid functionals produce results worse than B3LYP throughout. Like wave function methods, DFT methods can suffer from (near-)instabilities of the MOs,⁸² although such results have not been widely reported. The most robust method for frequencies tested here is EOMIP-CCSD, and we recommend using an EOMIP-CCSD calculation to confirm results obtained with wave function and DFT methods whenever affordable.

Highly accurate theoretically derived geometries, vibrational frequencies, and related properties are relevant to (a) the meaningful interpretation, confirmation, and prediction of spectroscopic data and (b) thermochemical calculations with kJ/mol or better accuracy. For these applications, highly accurate geometries and vibrational frequencies, beyond DFT quality, are needed. For calculations up to kcal/mol accuracy, our results show that B3LYP geometries and frequencies may be adequate for many open-shell systems (see the Supporting Information). Additionally, (c) accurate geometries and vibrational frequencies are pertinent to computational chemistry method evaluation, as diagnostics of the quality of the PES of more approximate (DFT) methods for a system of interest.

Finally, our results should not be interpreted as a general evaluation of methods for the treatment of open-shell species. For example, we have not evaluated method performance for energetics of reactions or for intra- and intermolecular through-space interactions.

■ ASSOCIATED CONTENT

■ Supporting Information

(a) A brief discussion on the thermochemical relevance of errors in both bond lengths and frequencies and (b) complete Z matrices for the calculated benchmark geometries. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: samuel.arey@epfl.ch.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation (SNSF) ProDoc TM Grant PDFMP2-123028. We thank Michael Harding and John Stanton for their technical support of the CFOUR code. We acknowledge the Swiss National Supercomputing Center (CSCS) and the EPFL centralized HPC for computational resources. We thank two anonymous reviewers for their valuable and constructive comments.

■ REFERENCES

- (1) Tajti, A.; Szalay, P. G.; Csaszar, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vazquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599–11613.
- (2) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108.
- (3) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- (4) Bartlett, R. J.; Watts, J.; Kucharski, S.; Noga, J. *Chem. Phys. Lett.* **1990**, *165*, 513–522.
- (5) Harding, M. E.; Metzroth, T.; Gauss, J.; Auer, A. A. *J. Chem. Theory Comput.* **2008**, *4*, 64–74.
- (6) Helgaker, T.; Gauss, J.; Jørgensen, P.; Olsen, J. *J. Chem. Phys.* **1997**, *106*, 6430–6440.
- (7) Bak, K. L.; Gauss, J.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. *J. Chem. Phys.* **2001**, *114*, 6548–6556.
- (8) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (9) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (10) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1995**, *103*, 4572–4585.
- (11) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2008**, *128*, 114111.
- (12) Byrd, E. F. C.; Sherrill, C. D.; Head-Gordon, M. *J. Phys. Chem. A* **2001**, *105*, 9736–9747.
- (13) Gauss, J.; Lauderdale, W. J.; Stanton, J. F.; Watts, J. D.; Bartlett, R. J. *J. Chem. Phys. Lett.* **1991**, *182*, 207–215.
- (14) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.
- (15) Beran, G. J. O.; Gwaltney, S. R.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2003**, *5*, 2488.
- (16) Stanton, J. F.; Gauss, J. *Adv. Chem. Phys.* **2003**, *125*, 101–146.
- (17) Szalay, P. G.; Vázquez, J.; Simmons, C.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 7624–7631.
- (18) Stanton, J. F.; Gauss, J.; Bartlett, R. J. *J. Chem. Phys.* **1992**, *97*, 5554–5559.
- (19) Scuseria, G. E.; Schaefer, H. F., III. *Chem. Phys. Lett.* **1987**, *142*, 354–358.
- (20) Watts, J. D.; Bartlett, R. J. *J. Chem. Phys.* **1990**, *93*, 6104–6105.
- (21) Stanton, J. F.; Gauss, J. *J. Chem. Phys.* **1994**, *101*, 8938.
- (22) Nooijen, M.; Bartlett, R. J. *J. Chem. Phys.* **1995**, *102*, 3629–3647.
- (23) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2010**, *6*, 107–126.
- (24) Other procedurally defined types of experimental structures exist, but they are of limited interest for comparison with quantum chemical equilibrium geometries.
- (25) Puzzarini, C.; Stanton, J. F.; Gauss, J. *Int. Rev. Phys. Chem.* **2010**, *29*, 273–367.
- (26) Stanton, J. F.; Gauss, J. *Int. Rev. Phys. Chem.* **2000**, *19*, 61–95.
- (27) Szalay, P. G.; Thøgersen, L. S.; Olsen, J.; Kállay, M.; Gauss, J. *J. Phys. Chem. A* **2004**, *108*, 3030–3034.
- (28) Helgaker, T.; Klopper, W.; Tew, D. P. *Mol. Phys.* **2008**, *106*, 2107–2143.
- (29) Kállay, M.; Surján, P. R. *J. Chem. Phys.* **2001**, *115*, 2945–2954.
- (30) Karton, A.; Martin, J. M. L. *J. Chem. Phys.* **2010**, *133*, 144102.
- (31) Heckert, M.; Kállay, M.; Gauss, J. *Mol. Phys.* **2005**, *103*, 2109–2115.

- (32) Heckert, M.; Kállay, M.; Tew, D. P.; Klopper, W.; Gauss, J. J. *Chem. Phys.* **2006**, *125*, 044108.
- (33) Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359–401.
- (34) Handy, N. C.; Pople, J. A.; Head-Gordon, M.; Raghavachari, K.; Trucks, G. W. *Chem. Phys. Lett.* **1989**, *164*, 185–192.
- (35) Stanton, J. F. *Mol. Phys.* **2009**, *107*, 1059.
- (36) Stanton, J. F. *J. Chem. Phys.* **2007**, *126*, 134309.
- (37) Bomble, Y. J.; Saeh, J. C.; Stanton, J. F.; Szalay, P. G.; Kállay, M.; Gauss, J. J. *Chem. Phys.* **2005**, *122*, 154107.
- (38) Lochan, R. C.; Head-Gordon, M. J. *Chem. Phys.* **2007**, *126*, 164101.
- (39) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.
- (40) Zhao, Y.; Truhlar, D. *Theor. Chim. Acta* **2008**, *120*, 215–241.
- (41) Boese, A. D.; Martin, J. M. L. *J. Chem. Phys.* **2004**, *121*, 3405–3416.
- (42) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (43) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (44) Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409–418.
- (45) CFOUR, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package by Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. with contributions from Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Bomble, Y. J.; Cheng, L.; Christiansen, O.; Heckert, M.; Heun, O.; Huber, C.; Jagau, T.-C.; Jonsson, D.; Jusélius, J.; Klein, K.; Lauderdale, W. J.; Matthews, D. A.; Metzroth, T.; O'Neill, D. P.; Price, D. R.; Prochnow, E.; Ruud, K.; Schiffmann, F.; Schwalbach, W.; Stopkowitz, S.; Tajti, A.; Vázquez, J.; Wang, F.; Watts, J. D. and the integral packages MOLECULE (Almlöf, J.; Taylor, P. R.), PROPS (Taylor, P. R.), ABACUS (Helgaker, T.; Jensen, H. J. Aa.; Jørgensen, P.; Olsen, J.), and ECP routines by Mitin, A. V.; van Wüllen, C. For the current version, see <http://www.cfour.de> (release date June 2010).
- (46) Koch, H.; Christiansen, O.; Jørgensen, P.; de Merás, A. M. S.; Helgaker, T. *J. Chem. Phys.* **1997**, *106*, 1808–1818.
- (47) Kállay, M.; Szalay, P. G.; Surján, P. R. *J. Chem. Phys.* **2002**, *117*, 980–990.
- (48) Kállay, M. MRCC, a string-based quantum chemical program suite. See also Kállay, M.; Surján, P. R. *J. Chem. Phys.* **2001**, *115*, 2945, and <http://tc03.fkt.bme.hu/index.php?n=Main.Author> (accessed October 2010).
- (49) The ROCCSD(T) scheme implemented in CFOUR, also known as the Watts–Gauss–Bartlett definition, uses a different scheme to freeze orbitals in frozen core calculations than the ROCCSD(T) implementation in MOLPRO (Werner–Knowles–Hampel definition). Frozen core calculations using the different definitions can be expected to lead to slightly different results.
- (50) Kállay, M.; Gauss, J.; Szalay, P. G. *J. Chem. Phys.* **2003**, *119*, 2991–3004.
- (51) Halkier, A.; Jørgensen, P.; Gauss, J.; Helgaker, T. *Chem. Phys. Lett.* **1997**, *274*, 235–241.
- (52) Heckert, M. Hochgenaue Berechnungen von Energien und molekularen Eigenschaften mit Hilfe der Coupled-Cluster-Theorie. PhD Thesis, University of Mainz: Mainz, Germany, 2006.
- (53) Frisch, M. J. et al. *Gaussian 09*, revision A.02; Gaussian Inc.: Wallingford, CT, 2009.
- (54) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359–363.
- (55) Shao, Y.; et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.
- (56) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793–9802.
- (57) Stanton, J. F.; Gauss, J.; Bartlett, R. J. *Chem. Phys. Lett.* **1992**, *195*, 194–199.
- (58) Kállay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 6841–6848.
- (59) Szalay, P. G.; Gauss, J.; Stanton, J. F. *Theor. Chim. Acta* **1998**, *100*, 5–11.
- (60) Some inconsistency is present in the nomenclature of frequencies throughout the literature. In this article, fundamental simply denotes the anharmonic frequency.
- (61) Biczysko, M.; Panek, P.; Scalmani, G.; Bloino, J.; Barone, V. *J. Chem. Theory Comput.* **2010**, *6*, 2115–2125.
- (62) Momose, T.; Endo, Y.; Hirota, E.; Shida, T. *J. Chem. Phys.* **1988**, *88*, 5338.
- (63) Ruscic, B. *J. Phys. Chem. Ref. Data* **2005**, *34*, 573.
- (64) NIST Computational Chemistry Comparison and Benchmark Database, Release 15a. <http://cccbdb.nist.gov> (Accessed April 2010).
- (65) Herzberg, G. *Molecular Spectra and Molecular Structure I. Spectra of Diatomic Molecules*; Van Nostrand: New York, 1950.
- (66) Huber, K. P.; Herzberg, G. *Constants of Diatomic Molecules; Molecular Spectra and Molecular Structure V*; Van Nostrand Reinhold: New York, 1979.
- (67) Kuchitsu, K. *Structure Data of Free Polyatomic Molecules; Landolt-Börnstein*; Springer: Berlin, 2006; Vol. 28A.
- (68) Kuchitsu, K. *Structure Data of Free Polyatomic Molecules; Landolt-Börnstein*; Springer: Berlin, 1999; Vol. 25B.
- (69) Kuchitsu, K. *Structure Data of Free Polyatomic Molecules; Landolt-Börnstein*; Springer: Berlin, 1999; Vol. 25C.
- (70) Kuchitsu, K. *Structure Data of Free Polyatomic Molecules; Landolt-Börnstein*; Springer: Berlin, 1998; Vol. 25A.
- (71) Mohr, P.; Taylor, B.; Newell, D. The 2010 CODATA Recommended Values of the Fundamental Physical Constants (Web Version 6.0). National Institute of Standards and Technology: Gaithersburg, MD, 2011. <http://physics.nist.gov/constants> (accessed January 2012).
- (72) Gewurtz, S.; Lew, H.; Flainek, P. *Can. J. Phys.* **1975**, *53*, 1097–1108.
- (73) Hovde, D.; Keim, E.; Saykally, R. *Mol. Phys.* **1989**, *68*, 599–607.
- (74) Yench, A. J.; Cormack, A. J.; Donovan, R. J.; Hopkirk, A.; King, G. C. *J. Phys. B: At., Mol., Opt. Phys.* **1999**, *32*, 2539.
- (75) Tuckett, R.; Dale, A.; Jaffey, D.; Jarrett, P.; Kelly, T. *Mol. Phys.* **1983**, *49*, 475–486.
- (76) Yang, J.; Hao, Y.; Li, J.; Zhou, C.; Mo, Y. *J. Chem. Phys.* **2005**, *122*, 134308.
- (77) Cormack, A. J.; Yench, A. J.; Donovan, R. J.; Lawley, K. P.; Hopkirk, A.; King, G. C. *Chem. Phys.* **1996**, *213*, 439–448.
- (78) We noted that for BH⁺, electronic contributions to the rotational constant were not considered in the experimental reports. For similar diatomics, CH/CD and BH, electronic contributions to B_e caused an adjustment in bond length of up to ~0.1 pm,^{25,96} so this effect would likely require consideration in the event of a future revision of the experimental value for BH⁺.
- (79) Demaison, J.; Margules, L.; Boggs, J. E. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3359–3363.
- (80) Eisfeld, W.; Morokuma, K. *J. Chem. Phys.* **2000**, *113*, 5587–5597.
- (81) Tew, D. P.; Klopper, W.; Heckert, M.; Gauss, J. *J. Phys. Chem. A* **2007**, *111*, 11242–11248.
- (82) Lutnæs, O. B.; Helgaker, T.; Jaszuński, M. *Mol. Phys.* **2010**, *108*, 2579.
- (83) Dewar, M. J.; Ford, G. P.; McKee, M. L.; Rzepa, H. S.; Thiel, W.; Yamaguchi, Y. *J. Mol. Struct.* **1978**, *43*, 135–138.
- (84) Dewar, M. J. S.; Ford, G. P. *J. Am. Chem. Soc.* **1977**, *99*, 1685–1691.
- (85) Pulay, P.; Fogarasi, G.; Pongor, G.; Boggs, J. E.; Vargha, A. *J. Am. Chem. Soc.* **1983**, *105*, 7037–7047.
- (86) Csonka, G. I.; Ruzsinszky, A.; Perdew, J. P. *J. Phys. Chem. A* **2005**, *109*, 6779–6789.
- (87) Merrick, J. P.; Moran, D.; Radom, L. *J. Phys. Chem. A* **2007**, *111*, 11683–11700.
- (88) Crawford, T. D.; Stanton, J. F.; Allen, W. D.; Schaefer, H. F., III. *J. Chem. Phys.* **1997**, *107*, 10626–10632.
- (89) Crawford, T. D.; Sherrill, C. D.; Valeev, E. F.; Fermann, J. T.; King, R. A.; Leininger, M. L.; Brown, S. T.; Janssen, C. L.; Seidl, E. T.; Kenny, J. P.; Allen, W. D. *J. Comput. Chem.* **2007**, *28*, 1610–1616.
- (90) Lee, T. J.; Taylor, P. R. *Int. J. Quantum Chem.* **1989**, *36*, 199–207.
- (91) %TAE refers to the percentage of the total atomization energy that arises from the (T) contribution of a CCSD(T) calculation.

- (92) Crawford, T. D.; Stanton, J. F. *J. Chem. Phys.* **2000**, *112*, 7873–7879.
- (93) Irikura, K. K. *J. Phys. Chem. Ref. Data* **2007**, *36*, 389–397.
- (94) Kobayashi, K.; Ozeki, H.; Saito, S.; Tonooka, M.; Yamamoto, S. *J. Chem. Phys.* **1997**, *107*, 9289–9296.
- (95) Colin, R. *J. Mol. Spectrosc.* **1989**, *136*, 387–401.
- (96) Morino, I.; Matsumura, K.; Kawaguchi, K. *J. Mol. Spectrosc.* **1995**, *174*, 123–131.
- (97) Herzberg, G.; Johns, J. W. C. *Proc. R. Soc. London, Ser. A* **1967**, *298*, 142–159.
- (98) Viteri, C. R.; Gilkison, A. T.; Rixon, S. J.; Grant, E. R. *J. Chem. Phys.* **2006**, *124*, 144312.
- (99) Merer, A. J.; Travis, D. N. *Can. J. Phys.* **1966**, *44*, 353–372.
- (100) Kawaguchi, K.; Yamada, C.; Hirota, E. *J. Chem. Phys.* **1985**, *82*, 1174–1177.
- (101) Bogey, M.; Demuynck, C.; Destombes, J. *Mol. Phys.* **1989**, *66*, 955–960.
- (102) Sunahori, F. X.; Zhang, X.; Clouthier, D. J. *J. Chem. Phys.* **2007**, *127*, 104312.
- (103) Lubic, K. G.; Amano, T.; Uehara, H.; Kawaguchi, K.; Hirota, E. *J. Chem. Phys.* **1984**, *81*, 4826–4831.
- (104) Brazier, C. R.; Bernath, P. F.; Burkholder, J. B.; Howard, C. J. *J. Chem. Phys.* **1988**, *89*, 1762–1767.