

Economical Post-CCSD(T) Computational Thermochemistry Protocol and Applications to Some Aromatic Compounds[†]

Amir Karton, Ilya Kaminker, and Jan M. L. Martin*

Department of Organic Chemistry, Weizmann Institute of Science, IL-76100 Rehovot, Israel

Received: January 4, 2009; Revised Manuscript Received: January 25, 2009

To achieve a kilojoules-per-mole level of accuracy consistently in computational thermochemistry, the inclusion of post-CCSD(T) correlation effects cannot be avoided. Such effects are included in the W4 and HEAT computational thermochemistry protocols. The principal bottleneck in carrying out such calculations for larger systems is the evaluation of the \hat{T}_3 -(T) term. We propose a cost-effective empirical approximation for this term that does not entail any reliance on experimental data. For first-row molecules, our W3.2lite protocol yields atomization energies with a 95% confidence interval of ~ 0.4 kcal/mol at the expense of introducing two such parameters. W3.2lite has been successfully applied to aromatic and aliphatic hydrocarbons such as benzene, fulvene, phenyl radical, pyridine, furan, benzyne isomers, *trans*-butadiene, cyclobutene, [1.1.1]propellane, and bicyclo[1.1.1]pentane. The W3.2lite predictions for fulvene, phenyl radical, cyclobutene, and [1.1.1]propellane are impossible to reconcile with experiment and suggest that remeasurement may be in order.

I. Introduction

Recently, computational thermochemistry in the sub-kilojoule-per-mole accuracy range has become possible on a near-“black box” basis for small molecules thanks to two computational thermochemistry approaches: HEAT (highly accurate extrapolated ab initio thermochemistry) developed by an international group of researchers^{1–3} and W4 (Weizmann-4) by the Martin group at Weizmann.⁴ (We note in passing that validation against experiment in this accuracy range would have been impossible without the development of the active thermochemical tables approach^{5–7} of Ruscic and coworkers.)

The most time-consuming steps in both W4 and HEAT are post-CCSD(T) correlation calculations, which are essential for consistently reaching this level of accuracy: if one neglects these steps (as is done in, e.g., the older W1 and W2 methods),^{8,9} the rms deviation for a representative sample of molecules will increase by about an order of magnitude, from about 0.3 to 3 kJ/mol, unless one limits oneself to molecules dominated by dynamical correlation.

The greatest weakness, at present, of both HEAT and W4 are the strenuous demands on computational resources places by the post-CCSD(T) calculations. Fully iterative CCSDT (coupled cluster with all single, double, and triple excitations) calculations exhibit asymptotic CPU time scaling $\propto n^3N^5$ (where n is the number of correlated electrons and N is the number of orbitals) that reaches $\propto n^4N^5$ for CCSDT(Q) (i.e., CCSDT with quasiperturbative connected quadruples), $\propto n^4N^6$ for fully iterative CCSDTQ, and $\propto n^5N^7$ for CCSDTQ5.

In practice, W4 is limited to at most four heavy atoms with some symmetry (BF₃, SO₃, and very recently FOOF),¹⁰ whereas the reduced-cost variant W4lite⁴ was successfully applied to the SiF₄ molecule.¹¹ Even the hardware available to us (8-core AMD Opteron and Intel Cloverton systems, each with 32 GB of RAM and a 2 TB striped disk array) was stretched to the limit by these calculations.

Basis set convergence of these contributions was recently studied in great detail:¹² it accelerates in the sequence CCSDT–CCSD(T) < CCSDT(Q)–CCSDT < CCSDTQ–CCSDT < CCSDTQ5–CCSDTQ. As a result, even in W4lite and W3.2, the \hat{T}_3 -(T) contribution (that is, the CCSDT–CCSD(T) difference) is obtained from relatively large basis sets of up to spdf quality (Dunning’s correlation consistent cc-pVTZ, to be precise), whereas the \hat{T}_4 and higher contributions in W4lite, W3.2, and HEAT345(Q) are actually obtained from a single CCSDT(Q) calculation in the fairly small cc-pVDZ basis set. We note here that whereas the CCSDT(Q) calculation is computationally tedious it requires only fairly modest amounts of memory and lends itself extremely well to parallelism. This leaves the iterative \hat{T}_3 -(T) step as the main computational bottleneck, not just in terms of CPU time but also of memory requirements.

In the present contribution, we propose an approximation that at the expense of introducing one to two empirical scaling parameters, none of which are derived from experiment, allows post-CCSD(T) computational thermochemistry on considerably larger systems. Applications to molecules such as benzene, fulvene, phenyl radical, pyridine, furan, benzyne isomers, *trans*-butadiene, cyclobutene, [1.1.1]propellane, and bicyclo[1.1.1]pentane will be presented as proofs of concept.

II. Computational Methods

All calculations were carried out on the Linux cluster of the Martin group at Weizmann. Geometry optimizations were carried out using Gaussian 03 revision C.01.¹³ The B3LYP^{14,15} DFT hybrid exchange correlation functional was used in conjunction with the pc-2^{16–18} polarization consistent basis set of Jensen. For the molecules in sets I and II as well as for the benzyne isomers, cyclobutene, and *trans*-butadiene, reference geometries were optimized at the CCSD(T)/cc-pV(Q+d)Z level of theory with only valence electrons correlated.

Unless otherwise stated, all post-CCSD(T) calculations were carried out using an OpenMP-parallel version of Mihály Kállay’s

[†] Part of the special issue “Robert Benny Gerber Festschrift”.

* Corresponding author. E-mail: gershom@weizmann.ac.il.

general coupled cluster code MRCC¹⁹ interfaced to the Austin–Mainz–Budapest version of the ACES II program system.²⁰ All large-scale self-consistent field (SCF), CCSD, and CCSD(T) calculations^{21,22} were carried out using version 2006.1 of the MOLPRO²³ program system. The diagonal Born–Oppenheimer correction (DBOC) calculations and some CCSDT calculations were carried out using the Austin–Mainz–Budapest version of the ACES II program system.²⁰

Unless otherwise indicated, all basis sets employed belong to the correlation-consistent family of Dunning and coworkers.^{24–28} For the large scale SCF, CCSD, and CCSD(T) single-point calculations, we combined the regular cc-pVnZ basis sets²⁴ on H with aug-cc-pVnZ basis sets²⁵ on Be–F and aug-cc-pV(n+d)Z basis sets²⁶ on Al–Cl. For convenience, we will denote this combination by AVnZ and regular cc-pVnZ on all atoms by PVnZ. In core-valence correlation calculations, the augmented version of the core-valence weighted correlation consistent basis sets of Peterson and Dunning were employed.²⁷ Scalar relativistic calculations were carried out using the Pacific Northwest National Laboratory (PNNL) Douglas–Kroll–Hess relativistically contracted correlation basis sets.²⁸ For the valence correlation post-CCSD(T) calculations, regular cc-pVnZ basis sets were employed as well as the atomic natural orbital (ANO) [4s3p1d/3s1p] basis set of Roos and coworkers,^{29,30} denoted by ANO431.

All basis set extrapolations, unless otherwise indicated, are carried out using the two-point extrapolation $A + B/L^3$ (where L is the highest angular momentum present in the basis set). In W3.2lite theory, we also consider extrapolations of the form of a simple linear combination of two basis sets for the valence CCSDT–CCSD(T) contribution.

The computational protocols of W4 theory, of the simplified variants W4lite, W3.2, and W2.2, and of the post-W4 methods W4.2, W4.3, and W4.4 have been specified and rationalized in great detail elsewhere.^{4,12} For the sake of making the article self-contained, we will briefly outline the various steps in W3.2 theory: (1) The reference geometry is obtained at the CCSD(T)/cc-pV(Q+d)Z level (frozen core). (2) The ROHF-SCF contribution is extrapolated from AVQZ and AV5Z basis sets using the Karton–Martin modification³¹ of Jensen’s extrapolation formula³²

$$E_{\text{HF},L} = E_{\text{HF},\infty} + A(L + 1) \exp(-9\sqrt{L}) \quad (1)$$

(3) The RCCSD valence correlation energy is calculated using AVQZ and AV5Z basis sets, using the Watts–Gauss–Bartlett (also known as ACES) definition²² for open-shell systems. Following the suggestion of Klopper,³³ it is partitioned in singlet-coupled pair energies, triplet-coupled pair energies, and \hat{T}_1 terms. The \hat{T}_1 term (which exhibits very weak basis set dependence) is simply set to be equal to that in the largest basis set, whereas the singlet-coupled and triplet-coupled pair energies are extrapolated using $A + B/L^\alpha$ with $\alpha_s = 3$ and $\alpha_T = 5$. These expressions are physically motivated by the partial-wave expansion for pair correlation energies in helium-like atoms^{34–36} as well as by empirical observations.^{8,37} (4) The (T) valence correlation energy is extrapolated from AVTZ and AVQZ calculations. For open-shell systems, the Werner–Knowles–Hampel (also known as MOLPRO) definition³⁸ of the restricted open-shell CCSD(T) energy is employed throughout rather than the original Watts–Gauss–Bartlett²² (also known as ACES II) definition, unless otherwise indicated. (5) The CCSDT–CCSD(T) difference is extrapolated from PVDZ and PVTZ basis sets.

(6) The \hat{T}_4 difference is estimated at the UCCSDT(Q)/PVDZ level of theory. (7) The difference between ACES II and MOLPRO definitions of the valence RCCSD(T) definition is extrapolated from PVDZ and PVTZ basis sets. One half of this contribution is added to the final result. (8) The inner-shell correlation contribution is extrapolated from RCCSD(T)/aug-cc-pwCVTZ and RCCSD(T)/aug-cc-pwCVQZ calculations. (9) The scalar relativistic contribution is obtained from the difference between nonrelativistic RCCSD(T)/AVQZ and second-order Douglas–Kroll RCCSD(T)/DK-aug-cc-pV(Q+d)Z calculations. (10) Atomic spin–orbit coupling terms are taken from the experimental fine structure. (11) Finally, diagonal Born–Oppenheimer corrections are obtained from RHF/AVTZ calculations.

III. Results and Discussion

For systems consisting of more than four non-hydrogen atoms, full W4 is not a viable option with current mainstream technology. If one wants to treat such a system by means of post-W2 methods, then W3.2 would probably be the only feasible starting point. The most strenuous calculations in W3.2, which might be considered to be the bottleneck of the method, are the CCSDT/PVTZ and CCSDT(Q)/PVDZ single-point calculations. For example, for the phenyl radical (C_{2v} symmetry), the quasiperturbative CCSDT(Q)/PVDZ calculation involves 67×10^9 determinants in the (Q) stage, whereas a fully iterative CCSDT/PVTZ calculation, if it were feasible, would involve 4.7×10^9 amplitudes. (In practice, the memory requirements are too forbidding, even for a machine with 32 GB of main memory.)

A. Higher-Order Connected Triples Contributions to the Total Atomization Energies. Because these are clearly the narrowest of the two bottlenecks, let us first consider the higher-order connected triples contributions (also known as “ \hat{T}_3 –(T)”). We were able to obtain reference values at the W4.4 theory level¹² for a chemically diverse training set of 51 first- and second-row small molecules, hereinafter referred to as “set I”. They are: BH, BH₃, CH, CH₂, CH₃, CH₄, C₂H, C₂H₂, C₂H₄, NH, NH₂, NH₃, N₂H, N₂H₂, OH, H₂O, HF, HCN, HNO, H₂CO, HOF, N₂, O₂, F₂, BF, CO, CN, CF, NO, OF, BeF₂, CO₂, AlH, AlH₃, SiH, SiH₄, PH₃, HS, H₂S, HCl, AlF, SiO, SiF, CS, SO, ClO, ClF, P₂, S₂, Cl₂, and AlCl. The chosen set, which includes radicals, polar systems, hydrides, and nonhydrides with single and multiple bonds, evidently spans the gamut from systems dominated by a single reference configuration to systems dominated by nondynamical correlation. Pathological cases such as BeO, MgO, Be₂, B₂, C₂, and BN ($a^1\Sigma^+$), for which the subsequent approximations proved to be too crude to hold, are omitted from consideration.

It has recently been shown¹² that higher-order triples contributions to the total atomization energy (TAE) extrapolated from the PV{T,Q}Z basis set pair are very close to the basis set limit and that convergence is slower than $\alpha = 3$. Fitting against PV{5,6}Z results suggested an effective exponent of $\alpha = 2.5$. Hereinafter, we consider the PV{T,Q}Z results extrapolated with exponent $\alpha = 2.5$ as the basis set limit for the \hat{T}_3 –(T) contribution. In general, this contribution at the basis set limit systematically reduces the TAE (notable exceptions are the CH and OF radicals with increases of 0.10 and 0.28 kcal/mol, respectively). Considerable basis set sensitivity is seen, especially for species containing multiple and very polar bonds. The PVDZ basis set is too limited to describe higher-order triples contributions; in particular, \hat{T}_3 –(T) contributions are generally positive with the PVDZ basis set and change sign as the basis set is expanded. The PV{D,T}Z pair, however, affords quite

TABLE 1: \hat{T}_3 -(T) Error Statistics with Respect to Our Best Results for Set I (in kcal/mol)

basis sets	1st row (32 species)				1st/2nd row (19 species) ^a				all (51 species)				
	<i>a</i>	<i>R</i> ²	MAD	rmsd	<i>a</i>	<i>R</i> ²	MAD	rmsd	<i>a</i>	<i>R</i> ²	MAD	rmsd	
PVTZ			0.12	0.14			0.11	0.14			0.12	0.14	
PVTZ(nof2d)			0.16	0.19			0.20	0.25					
PVTZ(nof1d)			0.23	0.28			0.29	0.37					
ANO431			0.25	0.29			0.26	0.35					
<i>a</i> × PVTZ	1.40	0.97	0.06	0.07	1.42	0.98	0.04	0.05	1.41	0.97	0.05	0.06	
<i>a</i> × PVTZ(nof2d)	1.54 ^b	0.95 ^b	0.08 ^b	0.10 ^b	1.94 ⁱ	0.90 ⁱ	0.12 ⁱ	0.15 ⁱ	1.64	0.91	0.10	0.13	
<i>a</i> × PVTZ(nof1d)	1.94 ^c	0.87 ^c	0.14 ^c	0.19 ^c	2.21	0.65	0.28	0.34					
<i>a</i> × ANO431	1.89 ^d	0.84 ^d	0.18 ^d	0.23 ^d	2.14	0.61	0.22	0.30					
(1 - <i>a</i>) × PVDZ + <i>a</i> × PVTZ(nof1d)	2.39 ^e	0.93 ^e	0.07 ^e	0.08 ^e									
(1 - <i>a</i>) × PVDZ + <i>a</i> × PVTZ(nof1d)(nophonH)	2.40 ^f	0.93 ^f	0.06 ^f	0.08 ^f									
(1 - <i>a</i>) × PVDZ + <i>a</i> × PVTZ(nof2d)	1.62 ^g	0.96 ^g	0.05 ^g	0.07 ^g	2.08 ^j	0.93 ^j	0.07 ^j	0.09 ^j	1.74 ^l	0.92 ^l	0.06 ^l	0.09 ^l	
(1 - <i>a</i>) × PVDZ + <i>a</i> × PVTZ(nof2d)(1ponH)	1.62 ^h	0.96 ^h	0.05 ^h	0.07 ^h	2.08 ^k	0.93 ^k	0.06 ^k	0.08 ^k	1.74 ^m	0.92 ^m	0.06 ^m	0.09 ^m	
(1 - <i>a</i>) × PVDZ + <i>a</i> × PVTZ	1.43	0.98	0.03	0.05	1.42	0.99	0.03	0.04					
PV{D,T}Z			0.98	0.03			0.99	0.03			0.98	0.03	0.04

basis sets	considering only the 1st row hydrides of set I (16 species)			
	<i>a</i>	<i>R</i> ²	MAD	rmsd
<i>a</i> × PVTZ (nof1d)	2.32	0.91	0.13	0.15
<i>a</i> × PVTZ(nof1d)(nophonH)	2.44	0.91	0.12	0.14
<i>a</i> × PVTZ(nof2d)	1.74	0.97	0.07	0.08
<i>a</i> × PVTZ(nof2d)(1ponH)	1.76	0.97	0.07	0.08
(1 - <i>a</i>) × PVDZ + <i>a</i> × PVTZ(nof1d)	2.81	0.99	0.03	0.04
(1 - <i>a</i>) × PVDZ + <i>a</i> × PVTZ(nof1d)(nophonH)	2.80	0.99	0.03	0.03
(1 - <i>a</i>) × PVDZ + <i>a</i> × PVTZ(nof2d)	1.83	1.00	0.02	0.03
(1 - <i>a</i>) × PVDZ + <i>a</i> × PVTZ(nof2d)(1ponH)	1.83	1.00	0.01	0.02

^a Here and in all subsequent tables, this notation refers to systems containing at least one second-row element. ^b Eliminating the main outliers CN, OF, and HNO results in $a = 1.58$, $R^2 = 0.97$, $MAD = 0.07$, and $rmsd = 0.08$ kcal/mol. ^c Eliminating the main outliers CN and OF results in $a = 2.12$, $R^2 = 0.92$, $MAD = 0.12$, and $rmsd = 0.14$ kcal/mol. ^d Eliminating the main outliers CN and OF results in $a = 2.13$, $R^2 = 0.89$, $MAD = 0.15$, and $rmsd = 0.18$ kcal/mol. ^e Eliminating the main outliers HNO, O₂, F₂, and OF results in $a = 2.58$, $R^2 = 0.98$, $MAD = 0.04$, and $rmsd = 0.05$ kcal/mol. ^f Eliminating the main outliers HNO, O₂, F₂, and OF results in $a = 2.59$, $R^2 = 0.98$, $MAD = 0.03$, and $rmsd = 0.05$ kcal/mol. ^g Eliminating the main outliers HNO, O₂, F₂, and OF results in $a = 1.72$, $R^2 = 0.99$, $MAD = 0.03$, and $rmsd = 0.03$ kcal/mol. ^h Eliminating the main outliers HNO, O₂, F₂, and OF results in $a = 1.72$, $R^2 = 0.99$, $MAD = 0.02$, and $rmsd = 0.03$ kcal/mol. ⁱ Eliminating the main outliers ClO, PH₃, and SO results in $a = 2.25$, $R^2 = 0.93$, $MAD = 0.08$, and $rmsd = 0.09$ kcal/mol. ^j Eliminating the main outlier P₂ results in $a = 1.98$, $R^2 = 0.96$, $MAD = 0.06$, and $rmsd = 0.07$ kcal/mol. ^k Eliminating the main outlier P₂ results in $a = 1.98$, $R^2 = 0.96$, $MAD = 0.05$, and $rmsd = 0.06$ kcal/mol. ^l Eliminating the main outliers HNO, OF, and P₂ results in $a = 1.76$, $R^2 = 0.96$, $MAD = 0.05$, and $rmsd = 0.07$ kcal/mol. ^m Eliminating the main outliers HNO, OF, and P₂ results in $a = 1.76$, $R^2 = 0.96$, $MAD = 0.04$, and $rmsd = 0.06$ kcal/mol.

reasonable predictions (e.g., the rmsd for set I with respect to the basis set limit is 0.04 kcal/mol) and in fact is used in W3.2, W4lite, W4, and W4.2 theories.⁴

In an attempt to lower the computational cost of the CCSDT/PVTZ calculation in W3.2 theory, we consider five smaller basis sets: (i) the ANO431 basis set; (ii) the spd and sp parts of the PVTZ basis set on heavy atoms and hydrogen, respectively, denoted by PVTZ(nof2d); (iii) the PVTZ(nof2d) basis set with the *p* functions on hydrogen replaced by the *p* function of the PVDZ basis set, denoted PVTZ(nof2d)(1ponH); (iv) the sp part of the PVTZ basis set combined with the *d* function from the PVDZ basis set on heavy atoms and the *s* part of the PVTZ basis set combined with the *p* function from the PVDZ basis set on hydrogen, denoted by PVTZ(nof1d); and (v) the PVTZ(nof1d) basis set without the *p* function on hydrogen, denoted by PVTZ(nof2d)(nophonH). These five basis sets consist of [4s3p1d/3s1p], [4s3p2d/3s2p], [4s3p2d/3s1p], [4s3p1d/3s1p], and [4s3p1d/3s] contracted Gaussian-type orbitals (GTO), respectively. The basis set dependence is shown in Table S1 of the Supporting Information, and the error statistics for set I with respect to the basis set limit are summarized in Table 1. Here we adopt the same strategy that was used in W3 and W4 theories, namely, using a scaling factor determined from least-squares fitting to the best available basis set limits over our training set of molecules.

Let us first consider the first-row species of set I (Table 1). The root-mean-square deviation (rmsd) of the unscaled PVTZ, PVTZ(nof2d), PVTZ(nof1d), and ANO431 basis sets is 0.14, 0.19, 0.28, and 0.29 kcal/mol, respectively. Simple scaling of these basis sets reduces the rmsd by 20–50% to 0.07, 0.10, 0.19, and 0.23 kcal/mol, respectively. In reality, such scaling becomes less effective because the basis set is less complete. We then considered the possibility of using a “multicoefficient” or “extrapolation” approach of the type $E_{T-(T)}(\infty) = \alpha \times E_{T-(T)}(\text{PVDZ}) + \beta \times E_{T-(T)}(\text{basis set}_2)$, where basis set₂ is one of the aforementioned truncated basis sets. (It should be noted that this is not an extrapolation in terms of *L*.) Like Schwenke³⁹ in a different context, we believe it is useful to impose the condition $\alpha + \beta = 1$, which leads to an expression with only a single adjustable parameter

$$E_{T-(T)}(\infty) = E_{T-(T)}(\text{PVDZ}) + a[E_{T-(T)}(\text{basis set}_2) - E_{T-(T)}(\text{PVDZ})] \quad (2)$$

Whereas this expression has no theoretical rationale, in practice, we find that it considerably improves performance over simple scaling: For example, using the PVDZ and PVTZ(nof2d) basis sets, we find $a = 1.62$, $R^2 = 0.96$, and $rmsd = 0.07$ kcal/mol. Upon exclusion of the four main outliers (HNO, O₂, F₂, and

TABLE 2: Error Statistics \hat{T}_4 Approximation versus (Q)/PV{T,Q}Z + [Q-(Q)]/PVTZ + \hat{T}_5 /PVDZ Results for Set II (in kcal/mol)

	1 st row (24 species)				1 st /2 nd row (19 species)				all (43 species)			
	<i>a</i>	<i>R</i> ²	MAD	rmsd	<i>a</i>	<i>R</i> ²	MAD	rmsd	<i>a</i>	<i>R</i> ²	MAD	rmsd
(Q)/PVDZ			0.06	0.09			0.16	0.24			0.10	0.17
<i>a</i> × (Q)/PVDZ	1.08 ^b	0.97 ^b	0.05 ^b	0.07 ^b	1.34 ^h	0.85 ^h	0.13 ^h	0.18 ^h	1.17 ⁿ	0.89 ⁿ	0.10 ⁿ	0.14 ⁿ
(Q)/PVDZ(noponH)			0.06	0.09			0.17	0.24			0.11	0.17
<i>a</i> × (Q)/PVDZ(noponH)	1.08 ^c	0.97 ^c	0.06 ^c	0.08 ^c	1.34 ⁱ	0.85 ⁱ	0.14 ⁱ	0.18 ⁱ	1.17 ^o	0.89 ^o	0.10 ^o	0.15 ^o
(Q)/ANO431			0.06	0.09			0.15	0.22			0.10	0.16
<i>a</i> × (Q)/ANO431	1.11 ^d	0.98 ^d	0.04 ^d	0.06 ^d	1.40 ^j	0.94 ^j	0.08 ^j	0.11 ^j	1.21 ^p	0.94 ^p	0.08 ^p	0.11 ^p
(Q)/PVTZ(nof1d)			0.08	0.11			0.17	0.24			0.12	0.18
<i>a</i> × (Q)/PVTZ(nof1d)	1.18 ^e	0.98 ^e	0.04 ^e	0.06 ^e	1.47 ^k	0.93 ^k	0.09 ^k	0.13 ^k	1.28 ^q	0.94 ^q	0.08 ^q	0.12 ^q
<i>a</i> × (Q)/PVDZ + <i>b</i> × (Q)/PVTZ(nof1d)	(0.32, 0.83) ^f	0.98 ^f	0.04 ^f	0.06 ^f	(-1.31, 2.87) ^l	0.96 ^l	0.08 ^l	0.10 ^l	(-0.55, 1.87) ^r	0.94 ^r	0.07 ^r	0.11 ^r
<i>a</i> × [\hat{T}_4 -(Q)/PVDZ(nod) + (Q)/PVTZ(nof1d)] ^a	1.17 ^g	0.97 ^g	0.05 ^g	0.08 ^g	1.51 ^m	0.96 ^m	0.06 ^m	0.08 ^m	1.38	0.94	0.07	0.11

considering only the 1st row hydrides of set II				
	<i>a</i>	<i>R</i> ²	MAD	rmsd
(Q)/PVDZ			0.06	0.10
<i>a</i> × (Q)/PVDZ	1.23	0.97	0.04	0.06
(Q)/PVDZ(noponH)			0.07	0.11
<i>a</i> × (Q)/PVDZ(noponH)	1.23	0.96	0.05	0.07
(Q)/ANO431			0.05	0.09
<i>a</i> × (Q)/ANO431	1.26	0.99	0.03	0.04

^a Parameterized against our best available \hat{T}_4 results. ^b Eliminating the main outliers C₂H₂ and C₂H results in *a* = 1.06, *R*² = 0.99, MAD = 0.03, and rmsd = 0.05 kcal/mol. ^c Eliminating the main outliers C₂H₂ and C₂H results in *a* = 1.05, *R*² = 0.99, MAD = 0.04, and rmsd = 0.05 kcal/mol. ^d Eliminating the main outliers C₂H₂ and C₂H results in *a* = 1.09, *R*² = 0.99, MAD = 0.03, and rmsd = 0.04 kcal/mol. ^e Eliminating the main outlier CN results in *a* = 1.23, *R*² = 0.99, MAD = 0.02, and rmsd = 0.04 kcal/mol. ^f After eliminating the main outliers C₂H₂ and C₂H, the scaling factors become (0.53, 0.57), *R*² = 0.99, MAD = 0.03, and rmsd = 0.04 kcal/mol. ^g Eliminating the main outlier CN results in *a* = 1.25, *R*² = 0.99, MAD = 0.02, and rmsd = 0.03 kcal/mol. ^h Eliminating the main outlier SiO results in *a* = 1.50, *R*² = 0.91, MAD = 0.10, and rmsd = 0.13 kcal/mol. ⁱ Eliminating the main outlier SiO results in *a* = 1.50, *R*² = 0.91, MAD = 0.11, and rmsd = 0.14 kcal/mol. ^j Eliminating the main outlier SiO results in *a* = 1.49, *R*² = 0.97, MAD = 0.07, and rmsd = 0.08 kcal/mol. ^k Eliminating the main outlier SiO results in *a* = 1.59, *R*² = 0.97, MAD = 0.06, and rmsd = 0.09 kcal/mol. ^l After eliminating the main outlier, SiO, the scaling factors become (-0.65, 2.26), *R*² = 0.98, MAD = 0.06, and rmsd = 0.08 kcal/mol. ^m Eliminating the main outlier SiO results in *a* = 1.58, *R*² = 0.98, MAD = 0.04, and rmsd = 0.06 kcal/mol. ⁿ Eliminating the main outliers P₂, S₂, and CS results in *a* = 1.09, *R*² = 0.95, MAD = 0.07, and rmsd = 0.09 kcal/mol. ^o Eliminating the main outliers P₂, S₂, and CS results in *a* = 1.09, *R*² = 0.95, MAD = 0.07, and rmsd = 0.09 kcal/mol. ^p Eliminating the main outliers P₂, S₂, and CS results in *a* = 1.14, *R*² = 0.97, MAD = 0.05, and rmsd = 0.07 kcal/mol. ^q Eliminating the main outliers P₂, S₂, and CS results in *a* = 1.20, *R*² = 0.97, MAD = 0.05, and rmsd = 0.08 kcal/mol. ^r After eliminating the main outliers P₂, S₂, and CS, the scaling factors become (0.26, 0.91), *R*² = 0.97, MAD = 0.05, and rmsd = 0.08 kcal/mol.

OF) an even better fit is obtained: *a* = 1.72, *R*² = 0.99, and rmsd = 0.03 kcal/mol. Using the PVDZ and PVTZ(nof1d) basis sets, we find *a* = 2.39, *R*² = 0.93, and rmsd = 0.08 kcal/mol. (Again, elimination of the four main outliers, HNO, O₂, F₂, and OF, results in: *a* = 2.58, *R*² = 0.98, and rmsd = 0.05 kcal/mol.) (For the sake of completeness, we note that we also attempted to relax the $\alpha + \beta = 1$ condition, but found no visible improvement; in addition, $\alpha + \beta$ was very close to 1 for the two-parameter optimized solutions.) As a numerical experiment, we considered a linear combination of the PVDZ and PVTZ basis sets and obtained *a* = 1.43, which is practically equivalent to the regular two-point *A* + *B*/*L*³ extrapolation. Finally, we note that taking off one *p* function on H in the PVTZ(nof2d) or PVTZ(nof1d) basis sets is of very little consequence for the subset of 16 first-row hydrides. (See Table 1.)

Turning now to the species containing second-row elements in set I, the PVDZ, ANO431 and PVTZ(nof1d) basis sets are found to be simply too small to give any useful predictions. The smallest basis set for which a useful correlation with the basis set limit results can be found is PVTZ(nof2d). For the subset of 19 molecules containing second-row atoms of set I, we obtain a scaling factor of 1.94, *R*² = 0.90, and rmsd = 0.15 kcal/mol. Excluding the three main outliers (SO, ClO, and PH₃) results in a scaling coefficient of 2.25, *R*² = 0.93, and rmsd = 0.09 kcal/mol. If the PVDZ results are included via eq 2, then we obtain *a* = 1.98, *R*² = 0.96, and rmsd = 0.07 kcal/mol upon exclusion of the main outlier, P₂.

B. Post-CCSDT Contributions to the Total Atomization

Energies. We will now consider the post-CCSDT contributions to the molecular TAE. For a set of 43 first- and second-row small molecules (hereinafter referred to as “set II”), we were able to obtain $\hat{T}_4 + \hat{T}_5$ contributions at the W4.3 theory level; that is, parenthetical connected quadruples (CCSDT(Q)-CCSDT) are extrapolated from PVTZ and PVQZ basis sets, higher-order quadruple contributions (\hat{T}_4 -CCSDT(Q)) are calculated with the PVTZ basis set, and connected quintuple contributions (CCSDTQ5-CCSDTQ) are calculated with the PVDZ basis set. Set II includes the following molecules: BH, BH₃, CH, CH₂, CH₃, CH₄, C₂H, C₂H₂, NH, NH₂, NH₃, OH, H₂O, HF, HCN, N₂, O₂, F₂, BF, CO, CN, CF, NO, OF, AlH, AlH₃, SiH, SiH₄, PH₃, HS, H₂S, HCl, AlF, SiO, SiF, CS, SO, ClO, ClF, P₂, S₂, Cl₂, and AlCl. We first consider scaling of CCSDT(Q)-CCSDT contributions calculated with the PVDZ and ANO431 basis sets against our best available $\hat{T}_4 + \hat{T}_5$ results. Table 2 depicts the overall error statistics for set II. The basis set dependence is shown in Table S2 of the Supporting Information.

For all systems considered, the overall $\hat{T}_4 + \hat{T}_5$ contributions are positive, ranging from 0.03 (CH₂) to 1.6 kcal/mol (P₂). Without exception, the CCSDT(Q)-CCSDT contributions to the TAE are likewise positive. As observed earlier,⁴ CCSDT(Q)-CCSDT contributions generally increase monotonically with the basis set size, except for very polar systems where the PVDZ results are overestimates. (For example, for OH, H₂O, SiF, AlF, and HF, the PVDZ results are higher than the PV{T,Q}Z limits

TABLE 3: Error Statistics W_n Variants versus ATcT (in kcal/mol) for the Set of Molecules Considered in Reference 4

method		$\hat{T}_3-(T)$	(Q)	MSD	MAD	rmsd
1st Row (19 Species) ^a						
W2.2	N/A		N/A	-0.49	0.53	0.95
W3.2	PV{D,T}Z		PVDZ	-0.10	0.17	0.23
W3.2	PV{D,T}Z		1.1 × PVDZ	-0.01	0.11	0.13
W3.2lite(a)	-0.7 × PVDZ + 1.7 × PVTZ(nof2d)		1.1 × PVDZ	-0.05	0.14	0.19
W3.2lite(b)	-1.6 × PVDZ + 2.6 × PVTZ(nof1d)		1.1 × PVDZ	-0.07	0.17	0.21
W3.2lite(c)	-1.6 × PVDZ + 2.6 × PVTZ(nof1d)(no _{pon} H)		1.1 × PVDZ(no _{pon} H)	-0.08	0.16	0.21
W4lite	PV{D,T}Z		PVDZ	-0.03	0.09	0.12
W4lite	PV{D,T}Z		1.1 × PVDZ	0.06	0.08	0.11
W4	PV{D,T}Z		^b	-0.01	0.07	0.09
1st/2nd Row (6 Species) ^c						
W2.2	N/A		N/A	-0.18	0.22	0.34
W3.2	PV{D,T}Z		PVDZ	-0.09	0.13	0.16
W3.2	PV{D,T}Z		1.1 × PVDZ	-0.03	0.10	0.12
W3.2lite(d)	-1.0 × PVDZ + 2.0 × PVTZ(nof2d)		1.5 × ANO431	0.00	0.11	0.13
W4lite	PV{D,T}Z		PVDZ	-0.08	0.11	0.16
W4lite	PV{D,T}Z		1.1 × PVDZ	-0.02	0.13	0.17
W4	PV{D,T}Z		^b	-0.02	0.07	0.08
All (25 Species)						
W2.2	N/A		N/A	-0.41	0.46	0.85
W3.2	PV{D,T}Z		PVDZ	-0.10	0.16	0.22
W3.2	PV{D,T}Z		1.1 × PVDZ	-0.02	0.11	0.13
W3.2lite(a)	-0.7 × PVDZ + 1.7 × PVTZ(nof2d)		1.1 × PVDZ	-0.04	0.14	0.17
W3.2lite(b)	-1.6 × PVDZ + 2.6 × PVTZ(nof1d)		1.1 × PVDZ	-0.03	0.16	0.20
W3.2lite(c)	-1.6 × PVDZ + 2.6 × PVTZ(nof2d)(no _{pon} H)		1.1 × PVDZ(no _{pon} H)	-0.04	0.16	0.20
W3.2lite(d)	-1.0 × PVDZ + 2.0 × PVTZ(nof2d)		1.5 × ANO431	0.01	0.13	0.15
W4lite	PV{D,T}Z		PVDZ	-0.04	0.10	0.13
W4lite	PV{D,T}Z		1.1 × PVDZ	0.04	0.09	0.13
W4	PV{D,T}Z		^b	-0.01	0.07	0.08

^a H₂O, C₂H₂, CH₃, CH₄, CH, CO₂, CO, F₂, HF, N₂, NH₃, N₂O, NO, O₂, O₃, NO₂, C₂H₄, H₂CO, and HNO. ^b Here $\hat{T}_4 = 1.1(E[\text{CCSDT}(\text{Q})/\text{PVTZ}] - E[\text{CCSDT}/\text{PVTZ}] + E[\text{CCSDTQ}/\text{PVDZ}] - E[\text{CCSDT}(\text{Q})/\text{PVDZ}])$. ^c H₂S, SO, SO₂, HCl, HOCl, and Cl₂.

by 0.02, 0.03, 0.03, 0.04, and 0.06 kcal/mol, respectively). In general, (Q)/PVDZ and (Q)/ANO431 underestimate the overall $\hat{T}_4 + \hat{T}_5$ contribution. Again, very polar cases may be an exception to this rule. For the ANO431 basis set, the only notable exception is CN, whereas for the PVDZ basis set, the (Q) contribution overestimates the overall post-CCSDT contribution for quite a few polar systems, most notably for BF, CF, OF, H₂O, HF, AlF, and SiO by 0.04, 0.04, 0.05, 0.05, 0.08, 0.08, and 0.10 kcal/mol, respectively.

For the first-row subset of set II, there is a rather high statistical correlation between the CCSDT(Q)–CCSDT values and the reference $\hat{T}_4 + \hat{T}_5$ numbers; $R^2 = 0.974$ and 0.978 for the PVDZ and ANO431 basis sets, respectively. Eliminating the two main outliers, C₂H and C₂H₂, results in scaling factors of 1.06 and 1.09, $R^2 = 0.988$ and 0.990 , and rmsd = 0.05 and 0.04 kcal/mol, respectively, for the PVDZ and ANO431 basis sets. Again, we note that removing the *p* function on H from the PVDZ basis set (denoted, PVDZ(no_{pon}H)) is of very little consequence on the scaling factor and error statistics for the subset of 15 first-row hydrides. (See Table 2.)

We also considered more elaborate approximations to the \hat{T}_4 contribution, namely, $a \times (\text{Q})/\text{PVTZ}(\text{nof}1d)$, $a \times (\text{Q})/\text{PVDZ} + b \times (\text{Q})/\text{PVTZ}(\text{nof}1d)$, and $a \times [\hat{T}_4-(\text{Q})/\text{PVDZ}(\text{nod}) + (\text{Q})/\text{PVTZ}(\text{nof}1d)]$, where PVDZ(nod) denotes the sp part of the PVDZ basis set, but found no appreciable improvement with their use.

For the subset of molecules containing second-row atoms of set II, the situation is less encouraging. For the PVDZ basis set, we obtain a scaling factor of 1.50, $R^2 = 0.91$, MAD = 0.10, and rmsd = 0.13 kcal/mol, after eliminating the main outlier, SiO. The ANO431 basis set, however, performs more consistently, giving a scaling factor of 1.49, $R^2 = 0.97$, MAD

= 0.07, and rmsd = 0.08 kcal/mol after SiO is eliminated. The $a \times [\hat{T}_4-(\text{Q})/\text{PVDZ}(\text{nod}) + (\text{Q})/\text{PVTZ}(\text{nof}1d)]$ formula admits some improvement: $a = 1.58$, $R^2 = 0.98$, MAD = 0.04, and rmsd = 0.06 kcal/mol after SiO is eliminated.

C. W3.2lite Theory. The results of the previous sections can be incorporated in a computationally more economical version of W3.2 theory, to be known as W3.2lite. Table 3 gives the error statistics for the set of ATcT species that were considered in our original W4 paper.⁴ (The individual errors can be found in Table S3 of the Supporting Information.)

Let us first consider the first-row systems. We propose three alternatives for calculating the $\hat{T}_3-(T)$ contribution: (variant a) from eq 2 with $a = 1.7$ and basis set₂ = PVTZ(nof2d); (variant b) from eq 2 with $a = 2.6$ and basis set₂ = PVTZ(nof1d); or (variant c) from eq 2 with $a = 2.6$ and basis set₂ = PVTZ(nof1d)(no_{pon}H). As for post-CCSDT contributions, the ANO431 and PVDZ basis sets show similar performance. We therefore suggest use of the smaller basis set, that is, UCCSDT(Q)/PVDZ–UCCSDT/PVDZ scaled by 1.1 for variants (a) and (b), and UCCSDT(Q)/PVDZ(no_{pon}H)–UCCSDT/PVDZ(no_{pon}H) scaled by 1.1 for the computationally less expensive variant (c). The greatest improvement of variant (a) over variant (b) is seen for N₂, O₂, F₂, NO, HNO, and N₂O, for which the error (in absolute value) drops by about 0.07 kcal/mol. All in all, variant (a) offers little improvement over variant (b). For example, the MAD and rmsd of the former are lower by only 0.03 and 0.02 kcal/mol. Therefore, the extra computational cost of W3.2lite(a) does not seem to warrant its use, although it should be noted (Table S3 of the Supporting Information) that the errors of W3.2lite(a) are systematically lower than those of W3.2lite(b) at the price of adding only five basis functions per heavy atom.

TABLE 4: Diagnostics for Importance of Nondynamical Correlation

	% TAE _e				CCSD/cc-pVTZ		
	[SCF] ^a	[(T)] ^a	[post-CCSD(T)] ^a	[(Q)] ^a	\mathcal{T}_1 diagnostic	\mathcal{D}_1 diagnostic	largest T_2 amplitudes
benzene	76.8	2.0	-0.04	0.13	0.010	0.028	0.080 (×2)
fulvene	76.0	2.0	-0.01	0.14	0.011	0.035	0.098
phenyl radical	76.3	2.1	-0.001	0.16	0.013	0.042	0.081
pyridine	74.0	2.3	-0.04	0.16	0.011	0.029	0.085
furane	73.5	2.3	-0.03	0.17	0.013	0.043	0.077
<i>o</i> -benzyne(¹ A ₁)	73.1	2.8	-0.01	0.25	0.013	0.039	0.153
<i>m</i> -benzyne(¹ A ₁)	71.8	3.1	0.02	0.36	0.015	0.046	0.188
<i>p</i> -benzyne(¹ A _g)	68.1	4.3	0.16	0.68	0.019	0.063	0.353
<i>p</i> -benzyne(³ B _{1u})	75.7	2.3	0.01	0.16	0.015	0.048	0.083
<i>trans</i> -butadiene	76.9	1.6	-0.002	0.09	0.011	0.034	0.108
cyclobutene	76.2	1.6	-0.02	0.09	0.010	0.030	0.104
[1.1.1]propellane(¹ A ₁)	74.2	2.0	-0.04	0.14	0.010	0.027	0.100
[1.1.1]propellane(³ A ₂)	74.4	1.9	0.00	0.11	0.011	0.027	0.026
bicyclo[1.1.1]pentane(¹ A ₁)	75.7	1.6	-0.04	0.08	0.008	0.017	0.024

^a Percentage of the total atomization energy related to nonrelativistic, clamped-nuclei values with inner shell electrons constrained to be doubly occupied.

There is no visible difference in performance between W3.2lite(b) and W3.2lite(c). For the set of 19 ATcT species that were considered in our original W4 paper,⁴ both variants attain a MSD, MAD, and rmsd of -0.08, 0.17, and 0.21 kcal/mol, which implies a 95% confidence interval of ~0.4 kcal/mol.

Turning to the species containing second-row atoms, for the \hat{T}_3 -(T) contribution (variant d), we find an empirical expression in the form of eq 2 with $a = 2.0$ and basis set₂ = PVTZ(nof2d). As for post-CCSDT contributions, the results of the previous section suggest using the ANO431 basis set scaled by 1.5. For the set of six species involving second-row elements for which we have ATcT values, we obtain a MSD, MAD, and rmsd of 0.00, 0.1, and 0.13 kcal/mol, respectively. (Although we note that a subset of six molecules is too small to provide statistically meaningful data.)

In summary, the changes in W3.2lite relative to W3.2 theory (Computational Methods section) are as follows: (1) The reference geometry is obtained at the B3LYP/PVTZ level of theory. (2) The \hat{T}_3 -(T) difference for first-row systems is obtained from the empirical expressions $-1.6 \times \text{PVDZ} + 2.6 \times \text{PVTZ}(\text{nof}1d)$ (variant b) and $-1.6 \times \text{PVDZ} + 2.6 \times \text{PVTZ}(\text{nof}1d)(\text{noponH})$ (variant c) and for second-row systems from $-1.0 \times \text{PVDZ} + 2.0 \times \text{PVTZ}(\text{nof}2d)$ (variant d). (3) Post-CCSDT contributions for first-row systems are estimated from UCCSDT(Q)/PVDZ-UCCSDT(PVDZ) scaled by 1.1 (variant b) and UCCSDT(Q)/PVDZ(noponH)-UCCSDT(PVDZ(noponH)) scaled by 1.1 (variant c) and for second-row systems from UCCSDT(Q)/ANO431-UCCSDT/ANO431 scaled by 1.5 (variant d).

To summarize, we assert that W3.2lite is more suitable for first-row systems and recommend the use of variant c for routine application because it is the cheapest and shows very similar performance to variants a and b. For systems that contain a small minority of second-row atoms, we recommend treating them as first-row atoms; for systems where second-row atoms predominate, we recommend variant d, that is, the second-row variant of the protocol.

Finally, we note that scaling the (Q)/PVDZ contribution in W3.2 by 1.1 cuts the rmsd by 0.10 and 0.04 kcal/mol for first- and second-row systems, respectively. Interestingly, such scaling does not improve the overall error statistics of W4lite theory. (See Table 3.)

D. Some Illustrative Examples. Aromatic systems, in general, have not been widely investigated by means of

benchmark thermochemical protocols. In this section, we apply W3.2lite to seven aromatic compounds, namely, benzene, phenyl radical, *o*-benzyne, *m*-benzyne, *p*-benzyne, pyridine, and furan as well as to *trans*-butadiene, cyclobutene, fulvene, [1.1.1]propellane, and bicyclo[1.1.1]pentane. For the sake of comparison, we will also include W2.2 results, which only differ from W3.2lite in the complete neglect of post-CCSD(T) correlation. For the smaller systems, *trans*-butadiene and cyclobutene, we were also able to obtain results at the W3.2 level from CCSD(T)/PVQZ reference geometries.

The ZPVEs for all of the systems except benzene and *p*-benzyne were taken to be one-half the sum of the RB3LYP/pc-2 harmonic frequencies scaled by 0.985. An uncertainty of 0.2 kcal/mol was assigned to the scaled DFT ZPVEs (taken to be twice the rmsd reported in ref 40 for a set of 39 molecules). For benzene, the anharmonic ZPVE from ref 41 was used; this value differs from the scaled DFT value by merely 0.03 kcal/mol. For singlet *p*-benzyne, the UB3LYP/pc-2 harmonic frequencies were used because a single-determinantal restricted description leads to a closed-shell singlet rather than a singlet diradical and to unrealistic vibrational frequencies. (See ref 42 for a detailed discussion.) Unless otherwise noted, experimental data for the heats of formation at 0 K were taken from the NIST Computational Chemistry Comparison and Benchmark Database (CCCBDB).⁴³ The atomization energies quoted in CCCBDB assume CODATA⁴⁴ values for the atomic heats of formation; however, particularly for carbon atom, the ATcT value⁴⁵ is significantly higher than the CODATA value, and these results are substituted in experimental atomization energies for benzene, pyridine, and so on that are increased by 0.3 to 0.5 kcal/mol over the ATcT values. Therefore, throughout this section, the experimental TAE₀ values were obtained from heats of formation at 0 K using ATcT atomic heats of formation. In cases where only experimental heats of formation at 298 K were available, they were first converted to 0 K using $H_{298}-H_0$ for H₂(g) 2.024 ± 0.000, C(cr, graphite) 0.251 ± 0.005, and N₂(g) 2.072 ± 0.000 kcal/mol from CODATA⁴⁴ and molecular $H_{298}-H_0$ calculated at the B3LYP/pc-2 level of theory within the RRHO approximation. In addition, for benzene, an ATcT⁴⁵ atomization energy is available.

Diagnostics for the importance of nondynamical correlation can be found in Table 4. These include the \mathcal{T}_1 and \mathcal{D}_1 diagnostics,^{46,47} the largest T_2 amplitudes at the CCSD/PVTZ level, and percentages of the total atomization energy accounted for at the SCF level, by parenthetical triples (shown elsewhere⁴

to be a more reliable predictor of the importance of post-CCSD(T) correlation effects than any other diagnostic), and overall by post-CCSD(T) correlation effects. A component breakdown of the W3.2lite data is given in Table 5, and the theoretical TAEs, heats of formation, some bond dissociation energies, and adiabatic singlet–triplet splittings of *p*-benzynes and [1.1.1]propellane are compared with available experimental values in Table 6.

After perusing Table 5, it becomes evident that the \hat{T}_3 –(T) contribution from W3.2lite(b) is slightly more negative and the (Q) is slightly more positive than the respective contributions from W3.2lite(c) theory, in effect leaving the post-CCSD(T) contributions of both variants roughly equal. Singlet *p*-benzynes, however, is an exception because the (Q) contribution of the two variants differs by as much as 0.20 kcal/mol. Throughout the remainder of this section, we will consider W3.2lite(c) theory for [1.1.1]propellane and bicyclo[1.1.1]pentane and W3.2lite(b) theory for the rest of the systems.

For benzene, our best \hat{T}_3 –(T) reduces the atomization energy by 2.2 kcal/mol, which is mostly but not wholly compensated by the (Q) contribution, 1.8 kcal/mol. Overall, the W3.2lite TAE₀[benzene] is 0.33 kcal/mol lower than W2.2. The said value, 1306.17 ± 0.4 kcal/mol, is in embarrassingly close agreement with ATcT (1306.10 ± 0.09 kcal/mol). The CCCBDB value (1306.13 ± 0.2) agrees well with ATcT and our calculations if the revised atomic heat of formation of carbon is employed.

For fulvene, there is near-perfect cancellation between the \hat{T}_3 –(T) contribution (−2.1 kcal/mol) and the (Q) contribution (+2.0 kcal/mol). Overall, the W3.2lite TAE₀, 1275.0 ± 0.4, is identical to the W2.2 value. The CCCBDB⁴³ value, however, is 2.3 kcal lower than the W3.2lite value. This discrepancy between theory and experiment is about 3 times the sum of the respective uncertainties and calls for a reexamination of the experimental value. Diagnostics for the importance of nondynamical correlation effects (given in Table 4) indicate that this system is mostly dominated by dynamical correlation, and thus imperfections in our calculated post-CCSD(T) contributions are not expected to account for the discrepancy between theory and experiment. As shown in ref 48 and references therein, more approximate compound thermochemistry methods such as G2MP2, HL2, G3B3, and CBS-APNO also suggest that the experimental heat of formation is too high by ~2 kcal/mol. From another perspective, W3.2lite reaction energies for the benzene → fulvene isomerization reaction (31.2 ± 0.6 kcal/mol) are 2.3 kcal/mol lower than the CCCBDB isomerization energy (33.5 ± 0.3 kcal/mol).

Two theoretical studies on the phenyl radical have been recently reported. Lau and Ng⁴⁹ calculated the TAE of C₆H₅ at the CCSD(T)/CBS level of theory with basis sets up to spdfigh quality and including ZPVE, core-valence, and scalar relativistic corrections; they arrived at Δ*H*_{f,0K} = 84.0 kcal/mol, which is in good agreement with the experimental value (84.3 ± 0.6 kcal/mol).⁵⁰ Montgomery and coworkers⁵¹ used their recently proposed ROCBS-QB3 method and arrived at *D*₂₉₈(C₆H₅–H) = 113.1 kcal/mol, which is, again, in good agreement with the experimental value⁵⁰ (113.5 ± 0.5 kcal/mol).

The UHF wave function for phenyl radical suffers from severe spin contamination (<*S*²> = 1.33), and even the UCCSD wave function still has <*S*²> = 0.86 compared with 3/4 for a pure doublet state. What is more, the RCCSDT/PVDZ total energy, −230.898134, is significantly lower than the UCCSDT/PVDZ value, −230.897692 kcal/mol. Whereas the \hat{T}_3 –(T) calculations were carried out from ROHF wave functions and will be

immune to spin contamination, the CCSDT(Q)/PVDZ of necessity employs a UHF reference, and we may assume that the (Q) contribution of 2.01 kcal/mol is an underestimate. Consequently, the perfect cancellation between \hat{T}_3 –(T) and (Q) seen in Table 5 is almost certainly deceptive, and the W3.2lite TAE₀ = 1195.15 ± 0.4 kcal/mol is biased downward, if anything. This makes it even harder to reconcile with the experimental value⁵⁰ of 1194.2 ± 0.6 kcal/mol (using ATcT atoms), let alone 1193.7 kcal/mol using CODATA atoms. The C₆H₅–H bond dissociation energy was measured to be 112.0 ± 0.6 kcal/mol in ref 50; our calculation is 1 kcal/mol lower, yet the uncertainty in the calculations should be much lower than that for the total atomization energy. We suggest that redetermination of the experimental value may be in order.

For pyridine, we again see imperfect cancellation between \hat{T}_3 –(T), −2.4 kcal/mol, and (Q), +2.0 kcal/mol. Overall, the W3.2lite TAE₀ is 0.4 kcal/mol lower than its W2.2 counterpart. Our W3.2lite TAE₀, 1183.35 ± 0.4 kcal/mol, again agrees embarrassingly well with CCCBDB, 1183.3 ± 0.15 kcal/mol, if ATcT atomic heats of formation are used. We note that the uncertainties introduced by the DFT geometries and zero-point vibrational energy are at least one order of magnitude larger than the discrepancy between W3.2lite and CCCBDB; such good agreement is fortuitous.

The B3LYP/pc-2 optimized geometry for furan is in good agreement with the CCSD(T)/AVQZ geometry of Feller et al.⁵² (bond lengths agree to 0.005 Å or better). The W2.2 TAE₀ is 951.09 ± 1.7 kcal/mol. The valence \hat{T}_3 –(T) contribution to the TAE is −1.87 kcal/mol, and the (Q) contribution is 1.69 kcal/mol. The W3.2lite TAE₀ (951.0 ± 0.4 kcal/mol) is merely 0.07 kcal/mol higher than the CCCBDB value (with ATcT atoms). Feller et al. calculated the TAE₀ of furan at the CCSD(T)/CBS level of theory with basis sets up to spdfigh quality, and including ZPVE, core-valence, scalar relativistic, spin–orbit, and higher-order-correlation corrections arrived at TAE₀ = 950.0 ± 0.5 kcal/mol.⁵²

The thermochemistry of singlet benzyne has been widely studied in recent years both theoretically^{42,53–55} and experimentally^{50,56,57} because of their interesting electronic structures and chemical properties. *p*-Benzynes, a key intermediate in the Bergman cyclization of enediyne,⁵⁸ is of special interest because of its role in the cleavage of double-stranded DNA and its potential antitumor activity. As the distance between the two unpaired electrons increases upon going from *o*-benzynes through *m*-benzynes to *p*-benzynes, some systematic trends can be established: (a) singlet diradical character is increased; (b) thermodynamic stability is decreased; (c) singlet–triplet splitting is decreased; and (d) nondynamical correlation effects become more pronounced. The percentage of the TAE_e accounted for by (T) triples of 2.8, 3.1, and 4.3, for *o*-, *m*-, and *p*-benzynes, respectively, indicates that nondynamical correlation effects become increasingly important along this series. This is also reflected in the magnitude of the largest \hat{T}_2 amplitudes: 0.15, 0.19, and 0.35, respectively (Table 4). For the benzynes systems, the CCSD(T)/PVQZ reference geometries substantially differ from the DFT optimized geometries. Figure 1 compares the CCSD(T)/PVQZ and B3LYP/pc-2 C–C bond lengths and diradical separations. The C–H bond lengths predicted by the two levels of theory differ by up to 0.005 Å, whereas the C–C bond lengths differ by as much as 0.02, 0.03, and 0.05 Å for *o*-, *m*-, and *p*-benzynes, respectively. Furthermore, the DFT-optimized diradical separations are shorter than the coupled cluster ones by 0.02, 0.44 (!), and 0.02 Å, respectively, for *o*-, *m*- and *p*-benzynes. Therefore, as previously noted by Cramer

TABLE 5: Component Breakdown of the W2.2 and W3.2lite Atomization Energies (in kcal/mol)^a

	SCF		CCSD		\hat{T}_3 -T		(Q)		inner shell		scalar rel.		spin-orbit		DBOC		M-A ^b		TAE _c		ZPVE	
	(A)	(A)	(B)	(A)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)	(N)	(O)	(P)	(Q)	(R)	(S)	(T)
benzene	1045.36	290.36	26.76	-2.11	-2.18	-2.11	1.78	1.72	7.09	7.37	-0.99	-0.51	0.23	0.13	1368.58	1368.252	1368.247	1368.247	1368.247	1368.247	62.08	62.08
fulvene	1010.08	292.50	26.83	-2.00	-2.05	-2.00	1.99	1.92	6.97	7.23	-0.97	-0.51	0.26	0.13	1335.42	1335.40	1335.42	1335.42	1335.42	1335.42	60.42	60.42
phenyl radical	948.18	268.52	26.35	-1.91	-1.95	-1.91	2.01	1.97	6.89	7.16	-0.96	-0.51	0.21	0.12	1248.94	1249.067	1249.068	1249.068	1249.068	1249.068	53.92	53.92
pyridine	911.56	293.55	27.98	-2.38	-2.38	-2.32	1.98	1.92	6.35	6.59	-1.02	-0.42	0.21	0.16	1238.45	1238.127	1238.128	1238.128	1238.128	1238.128	54.78	54.78
furane	727.95	240.02	22.46	-1.87	-1.87	-1.84	1.69	1.66	4.98	5.16	-0.91	-0.56	0.21	0.11	994.33	994.20	994.21	994.21	994.21	994.21	43.24	43.24
<i>o</i> -benzynes(A ₁) ^c	844.99	278.45	31.90	-2.92	-2.92	-2.86	2.93	2.87	6.83	7.10	-0.96	-0.51	0.18	0.13	1161.16	1161.233	1161.227	1161.227	1161.227	1161.227	46.43	46.43
<i>m</i> -benzynes(A ₁) ^c	819.44	286.18	35.80	-3.78	-3.78	-3.68	4.15	4.08	6.65	6.63	-0.91	-0.51	0.15	0.13	1146.78	1147.24	1147.22	1147.22	1147.22	1147.22	45.42	45.42
<i>p</i> -benzynes(A _g) ^c	768.01	310.12	48.01	-5.58	-5.58	-5.56	7.90	7.70	6.06	6.26	-0.93	-0.51	0.10 ^d	0.13	1131.07	1133.26	1133.44	1133.44	1133.44	1133.44	45.62	45.62
<i>p</i> -benzynes(B _{1u}) ^c	850.24	246.39	26.01	-1.71	-1.71	-1.68	1.84	1.80	6.57	6.81	-0.95	-0.51	0.19	0.12	1128.18	1128.37	1128.38	1128.38	1128.38	1128.38	45.91	45.91
<i>trans</i> -butadiene	775.57	216.59	16.26	-0.98	-0.98	-0.94	1.00	0.93	4.64	4.85	-0.65	-0.34	0.20	0.08	1012.47	1012.51	1012.54	1012.54	1012.54	1012.54	52.51	52.51
cyclobutene	759.96	221.41	16.00	-1.14	-1.14	-1.12	0.96	0.92	4.53	4.68	-0.70	-0.34	0.19	0.08	1001.21	1001.05	1001.07	1001.07	1001.07	1001.07	53.30	53.30
[1.1.1]propellane(A ₁)	832.83	266.98	22.83	-2.01	-2.01	-2.01	1.62	1.62	5.28	5.43	-0.83	-0.42	0.24	0.10	1127.05	1126.71	1126.71	1126.71	1126.71	1126.71	57.52	57.52
[1.1.1]propellane(A ₂)	762.90	243.08	19.06	-1.13	-1.13	-1.13	1.11	1.11	5.02	5.15	-0.89	-0.42	0.21	0.10	1029.09	1029.13	1029.13	1029.13	1029.13	1029.13	56.35	56.35
bicyclo[1.1.1]pentane(A ₁)	966.76	290.26	20.28	-1.50	-1.50	-1.50	1.08	1.08	5.43	5.57	-0.92	-0.42	0.25	0.10	1281.79	1281.41	1281.41	1281.41	1281.41	1281.41	72.12	72.12
Reactions																						
benzene → fulvene	35.28	-2.14	-0.07	-0.13	-0.13	-0.11	-0.21	-0.20	0.13	0.13	-0.01	0	-0.03	0.00	33.16	32.85	32.83	32.83	32.83	32.83	-1.66	-1.66
benzene → H + C ₆ H ₅	97.18	21.84	0.41	-0.23	-0.23	-0.21	-0.24	-0.25	0.21	0.21	-0.02	0	0.02	0.00	119.64	119.185	119.179	119.179	119.179	119.179	-8.16	-8.16
C ₆ H ₅ → H + <i>o</i> -benzynes	103.19	-9.94	-5.55	0.97	0.97	0.95	-0.91	-0.90	0.05	0.05	0.00	0	0.03	0.00	87.79	87.83	87.84	87.84	87.84	87.84	-7.49	-7.49
C ₆ H ₅ → H + <i>m</i> -benzynes	128.75	-17.66	-9.45	1.82	1.82	1.78	-2.14	-2.11	0.52	0.52	-0.05	0	0.06	0.00	102.16	101.83	101.85	101.85	101.85	101.85	-8.50	-8.50
C ₆ H ₅ → H + <i>p</i> -benzynes(A _g)	180.17	-41.61	-21.66	3.63	3.63	3.66	-5.88	-5.73	0.89	0.89	-0.03	0	0.11	0.00	117.88	115.81	115.62	115.62	115.62	115.62	-8.30	-8.30
<i>trans</i> -butadiene → cyclobutene	15.61	-4.82	0.26	0.17	0.17	0.18	0.04	0.01	0.17	0.17	0.05	0	0.01	0.00	11.27	11.46	11.47	11.47	11.47	11.47	0.79	0.79
[1.1.1]propellane(A ₁) + H ₂ → bicyclo[1.1.1]pentane(A ₁)	-50.10	2.35	2.55	-0.50	-0.50	-0.50	0.04	0.01	-0.14	0.08	0.08	0	0.05	0.00	-45.22	-45.19	-45.19	-45.19	-45.19	-45.19	8.38	8.38
Adiabatic Singlet-Triplet Splittings																						
<i>p</i> -benzynes	-82.23	63.73	22.00	-3.877	-3.877	-3.884	6.06	5.89	-0.54	0.02	0.02	0	-0.09	0.00	2.88	4.89	5.07	5.07	5.07	5.07	0.29	0.29
[1.1.1]propellane	69.93	23.90	3.77	-0.95	-0.95	-0.95	0.50	0.50	0.27	0.06	0.06	0	0.03	0.00	97.96	97.59	97.59	97.59	97.59	97.59	-1.17	-1.17

^a(A) AV{Q,5}Z; (B) AV{T,Q}Z; (C) -1.6 × PVDZ + 2.6 × PVTZ(nof1d) from W3.2lite(b) theory; (D) -1.6 × PVDZ + 2.6 × PVTZ(nof1d)(noptonH) from W3.2lite(c) theory; (E) PVDZ from W3.2lite(b) theory; (F) PVDZ(noponH) from W3.2lite(c) theory; (G) MTsmall; (H) aug-cc-pwCV{T,Q}Z; (I) DK-aug-cc-pVQZ; (J) aug-cc-pVTZ; (K) PV{D,T}Z; (L) C₆H₆ from ref 41; all of the rest: B3LYP/pc-2 harmonic frequencies scaled by 0.985. ^b Difference between the ACES II and MOLPRO definitions of the valence RCCSD(T). ^c CCSD(T)/PVQZ reference geometries. (See the text.) ^d HF/AVTZ//HF/AVTZ level of theory. (See the text.)

TABLE 6: Total Atomization Energies, Heats of Formation, Reaction Energies, and Singlet–Triplet Splittings at 0 K (kcal/mol)

	TAE ₀ ^a				ΔH _{f,0} ^c		
	W2.2	W3.2lite ^b	W3.2	exptl ^c	W3.2lite ^{b,d}	W3.2	exptl ^c
benzene	1306.5	1306.17 ± 0.4		1306.13 ± 0.20 1306.10 ± 0.09 ^f	23.96 ± 0.4		24.00 ± 0.12 24.03 ± 0.18 ^f
fulvene	1275.0	1275.01 ± 0.4		1272.66 ± 0.27	55.12 ± 0.5		57.47 ± 0.22 ^g
phenyl radical	1195.0	1195.15 ± 0.4		1194.20 ± 0.62	83.35 ± 0.5		84.3 ± 0.6
pyridine	1183.7	1183.35 ± 0.4		1183.29 ± 0.15	37.56 ± 0.5		37.62 ± 0.07 ^g
furane	951.1	950.97 ± 0.4		950.90 ± 0.15	-5.23 ± 0.5		-5.16 ± 0.10
<i>o</i> -benzyne(¹ A ₁)	1114.7	1114.80 ± 0.4		1118.1 ± 3	112.06 ± 0.5		108.8 ± 3 ^{g,h}
<i>m</i> -benzyne(¹ A ₁)	1101.4	1101.80 ± 0.4		1102.8 ± 3.1	125.06 ± 0.5		124.1 ± 3.1 ^{g,i}
<i>p</i> -benzyne(¹ A _g)	1085.4	1087.83 ± 0.4		1087.3 ± 3.3 1086.8 ± 2.9 1086.6 ± 1.0	139.03 ± 0.5		139.5 ± 3.3 ^{g,j} 140.0 ± 2.9 ^{g,k} 140.2 ± 1.0 ^{g,l}
<i>p</i> -benzyne(³ B _{1u})	1082.3	1082.47 ± 0.4		1083.1 ± 2.9	144.39 ± 0.5		143.8 ± 2.9 ^{g,m}
<i>trans</i> -butadiene	960.0	960.03 ± 0.4	959.77 ± 0.3	960.26 ± 0.17	29.99 ± 0.5	30.25 ± 0.4	29.76 ± 0.14
cyclobutene	947.9	947.77 ± 0.4	947.45 ± 0.3	948.48 ± 0.37	42.25 ± 0.5	42.57 ± 0.4	41.54 ± 0.36
propellane(¹ A ₁)	1069.5	1069.20 ± 0.4 ⁿ		1072 ± 1	90.87 ± 0.5		88 ± 1 ^{g,o}
propellane(³ A ₂)	972.7	972.78 ± 0.4 ⁿ			187.30 ± 0.5		
bicyclo[1.1.1]pentane	1209.7	1209.29 ± 0.4 ⁿ			54.04 ± 0.5		
Reactions							
benzene → fulvene		31.16 ± 0.6					
benzene → C ₆ H ₅ + H		111.02 ± 0.6		112.0 ± 0.6 ^p			
C ₆ H ₅ → H + <i>o</i> -benzyne		80.35 ± 0.6					
C ₆ H ₅ → H + <i>m</i> -benzyne		93.35 ± 0.6					
C ₆ H ₅ → H + <i>p</i> -benzyne(¹ A _g)		107.32 ± 0.6					
<i>trans</i> -butadiene → cyclobutene		12.26 ± 0.6	12.33 ± 0.4				
[1.1.1]propellane(¹ A ₁) + H ₂ → bicyclo[1.1.1]pentane(¹ A ₁)		-36.80 ± 0.7					
Adiabatic Singlet–Triplet Splittings							
<i>p</i> -benzyne	3.2	5.36 ± 0.6		3.8 ± 0.5 ^q			
[1.1.1]propellane	96.8	96.42 ± 0.6		94.3 ^r			

^a Theoretical uncertainties correspond to 95% confidence intervals taken as twice the rmsd (1.7, 0.4, and 0.3 kcal/mol for W2.2, W3.2lite, and W3.2, respectively). ^b Variant W3.2lite(b) unless otherwise noted. ^c Unless otherwise noted, converted from the experimental ΔH_{f,0} using the ATcT⁷¹ atomic heats of formation (C 170.055 ± 0.026, H 51.633 ± 0.000, N 112.469 ± 0.006 kcal/mol). ^d Converted from TAE₀ using the ATcT⁷¹ atomic heats of formation, cfr. (c). ^e Unless otherwise noted, experimental values from CCCBDB.⁷² ^f From ATcT.⁴⁵ ^g Experimental ΔH_{f,298} was converted to 0 K using H₂₉₈ - H₀ for H₂(g) 2.024 ± 0.000, C(cr,graphite) 0.251 ± 0.005, and N₂(g) 2.072 ± 0.000 kcal/mol from CODATA;⁴⁴ molecular H₂₉₈ - H₀ calculated at the B3LYP/pc-2 level of theory within the RRHO approximation (3.69, 3.27, 3.38, 3.47, 3.33, 3.34, 3.09 kcal/mol for fulvene, pyridine, *o*-benzyne, *m*-benzyne, singlet *p*-benzyne, triplet *p*-benzyne, [1.1.1]propellane). ^h Converted to 0 K from ΔH_{f,298} = 106.6 ± 3 given in ref 56. ⁱ Converted to 0 K from ΔH_{f,298} = 122.0 ± 3.1 given in ref 56. ^j Converted to 0 K from ΔH_{f,298} = 137.3 ± 3.3 given in ref 56. ^k Converted to 0 K from ΔH_{f,298} = 137.8 ± 2.9 given in ref. 50. ^l Converted to 0 K from ΔH_{f,298} = 138.0 ± 1.0 given in ref. 57. ^m Converted to 0 K from ΔH_{f,298} = 141.6 ± 2.9 given in ref. 60. ⁿ Variant W3.2lite(c). ^o Converted to 0 K from estimated ΔH_{f,298} = 84 ± 1 given in ref 67. ^p From ref 50. ^q From ref 60. ^r From ref 69.

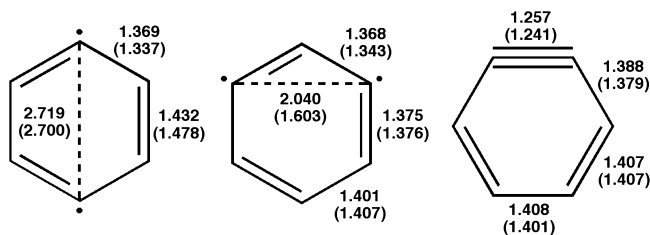


Figure 1. CCSD(T)/PVQZ and B3LYP/pc-2 (in parentheses) optimized C–C bond lengths and diradical separations (in angstroms) for the *p*-, *m*-, and *o*-benzyne isomers.

et al.⁵³ and Cremer and coworkers,⁵⁵ at the DFT level, *m*-benzyne has more of a “bicyclic” structure with a 1.6 Å distance between the two dehydrocarbons. This artifact, associated with the single-configurational representation provided by the DFT method, is clearly absent at the CCSD(T)/PVQZ level with a diradical separation of 2.04 Å.⁵⁹ Therefore, for the W2.2 and W3.2lite single-point calculations, CCSD(T)/PVQZ reference geometries were used.

Post-CCSD(T) contributions substantially increase along the series *o*-, *m*-, and *p*-benzyne. The higher-order connected triples, \hat{T}_3 –(T), reduce the binding energies by 2.92, 3.78, and 5.58

kcal/mol, respectively, and the parenthetical quadruples (Q) contribution increases the binding energies by 2.93, 4.15, and 7.90 kcal/mol, respectively. Overall, our best TAE₀ are 1114.80 ± 0.4, 1101.80 ± 0.4, and 1087.83 ± 0.4 kcal/mol. These correspond to ΔH_{f,0}^c of 112.06 ± 0.5, 125.06 ± 0.5, and 139.03 ± 0.5 kcal/mol. Squires et al.⁵⁶ determined the heats of formation at 298 K from collision-induced dissociation threshold energy measurements to be 106.6 ± 3.0, 122.0 ± 3.1, and 137.3 ± 3.3 kcal/mol, respectively. Converting these to 0 K, we obtain ΔH_{f,0}^c of 108.8 ± 3, 124.1 ± 3.1, and 139.5 ± 3.3 kcal/mol for *o*-, *m*-, and *p*-benzyne, respectively. The W3.2lite heat of formation is higher than the experimental value by 3.3 kcal/mol for *o*-benzyne (slightly below the sum of the respective uncertainties) and by 1.0 kcal/mol for *m*-benzyne. For *p*-benzyne, two more independent experimental determinations are available (after conversion to 0 K): 140.2 ± 1.0⁵⁷ and 140.0 ± 2.9 kcal/mol.⁵⁰ Our best theoretical value is thus 0.5 to 1.2 kcal/mol lower than the experimental value.

Post-CCSD(T) correlation effects are much less pronounced for the low-lying triplet excited state of *p*-benzyne than for the singlet ground state (e.g., the %TAE_e [(T)] values are 2.3 and 4.3, and the largest \hat{T}_2 amplitudes are 0.08 and 0.35 for the two

states, respectively). Or, to put it differently, there is near-perfect cancellation between the \hat{T}_3 -(T) contribution (-1.71 kcal/mol) and the (Q) contribution ($+1.84$ kcal/mol). At the W3.2lite level, we obtain a TAE_0 of 1082.5 ± 0.4 kcal/mol, resulting in an adiabatic singlet–triplet splitting at 0 K of 5.4 ± 0.6 kcal/mol. Wenthold et al.⁶⁰ obtained an adiabatic singlet–triplet splitting of 3.8 ± 0.5 kcal/mol from ultraviolet photoelectron spectroscopy (UPS) measurements, which is 1.6 kcal/mol lower than our W3.2lite value. For the vibrationless transition, we obtain 5.07 ± 0.6 kcal/mol, which is in good agreement with the ASPT2(12,12)/ANO-L value of 5.5 kcal/mol reported in ref 61.

For *trans*-butadiene, the W2.2 and W3.2lite TAE_0 agree to within 0.03 kcal/mol. The W3.2 TAE_0 (959.77 ± 0.3 kcal/mol) is 0.3 kcal/mol lower than the W3.2lite value, where the difference comes from the improved extrapolation of the higher-order triples, \hat{T}_3 -(T), contribution. The said W3.2 TAE_0 is 0.5 kcal/mol below the CCCBDB (using ATcT atoms). For cyclobutene, the W3.2 TAE_0 (947.45 ± 0.4 kcal/mol) is 0.3 and 0.5 kcal/mol lower than the W3.2lite and W2.2 values, respectively. The CCCBDB TAE_0 is 1.0 kcal/mol higher than the W3.2 value.

The cyclobutene \rightarrow butadiene electrocyclic ring-opening reaction is a classic example of a pericyclic rearrangement whose stereochemical selectivity follows the Woodward–Hoffmann rules. At both W3.2lite and W3.2 levels, we obtain a reaction energy of -12.3 kcal/mol, which is 0.3 kcal/mol below the CBS-QB3 value of Houk and coworkers⁶² and 0.5 kcal/mol above the value obtained using the CCCBDB TAE_0 s.

As a last example, we consider a highly strained hydrocarbon, namely, [1.1.1]propellane and the related bicyclo[1.1.1]pentane. (On a historical note, the name and first synthesis of [1.1.1]propellane were reported by the late Ginsburg and coworkers⁶³ from the Technion.) Small-ring propellanes are an intriguing class of strained alkanes that have attracted considerable theoretical interest (see, e.g., refs 64–67 and references therein) because of their inverted-configuration bridgehead carbons. In particular, the nature of the central bond between the two “inverted” bridgehead carbons has been widely debated. (See refs 65 and 66 for two recent studies.) Our B3LYP/pc-2 reference geometry is in reasonable agreement with the gas-phase electron diffraction structure reported by Hedberg and Hedberg:⁶⁸ the C–H and C–C bond lengths agree to within 0.02 and 0.01 Å, respectively, and the bond angles agree to within 1° . As expected, the r_e distances underestimate the experimental r_g distances because of anharmonicity of the potential energy surface around the r_e positions. For the purpose of illustration, the calculated r_z – r_e differences from a B3LYP/pc-1 anharmonic force-field calculation are 0.006 and 0.009 Å for C–H and C–C bonds, respectively. The theoretical interbridgehead separation is 0.02 Å shorter than the experimental r_g distance, even after the said r_z – r_e correction of 0.01 Å is applied. The TAE_0 of [1.1.1]propellane is 1069.54 kcal/mol at the W2.2 level. The valence \hat{T}_3 -(T) contribution to the TAE (-2.01 kcal/mol) is only partially compensated by the (Q) contribution of 1.62 kcal/mol, resulting in a W3.2lite TAE_0 of 1069.20 ± 0.4 kcal/mol. Using ATcT atomic heats of formation at 0 K, we obtain $\Delta H_{f,0}^\circ = 90.87 \pm 0.5$ kcal/mol, which is 2.9 kcal/mol higher than the “semiexperimental” estimate of Wiberg et al.,⁶⁷ $\Delta H_{f,0}^\circ = 88 \pm 1$ kcal/mol, which was derived from a calorimetric measurement of the heat of reaction of [1.1.1]propellane with acetic acid. At the W3.2lite level, we obtain an adiabatic singlet–triplet gap of 96.42 ± 0.6 kcal/mol, which is 2.1 kcal/mol higher than the experimental value of 94.3 kcal/mol, estimated by Schafer et al.⁶⁹ from low-energy electron-impact spectroscopy.

In an attempt to eliminate the significant computational cost of the inner-shell CCSD(T)/aug-cc-pwCVQZ calculation, we considered this contribution with the considerably smaller MTsmall basis set.⁸ In general, the MTsmall basis sets underestimate the aug-cc-pwCV{T,Q}Z results by 0.1 to 0.3 kcal/mol (Table 5).

IV. Conclusions

The principal bottleneck in carrying out W4- or HEAT-type computational thermochemistry for larger systems is the evaluation of the \hat{T}_3 -(T) term. We propose a cost-effective empirical approximation for this term that does not entail any reliance on experimental data. For first-row molecules, our W3.2lite protocol, at the expense of introducing two such parameters, yields atomization energies with a 95% confidence interval of about 0.4 kcal/mol. W3.2lite has been successfully applied to aromatic and aliphatic hydrocarbons such as benzene, fulvene, phenyl radical, pyridine, furan, benzyne isomers, *trans*-butadiene, cyclobutene, [1.1.1]propellane, and bicyclo[1.1.1]pentane. The W3.2lite predictions for fulvene, phenyl radical, cyclobutene, and [1.1.1]propellane are impossible to reconcile with experiment and suggest that remeasurement may be in order.

Acknowledgment. We thank Dr. Branko Ruscic (Argonne National Laboratory) for early communication of ref 45. Research at Weizmann was funded by the Israel Science Foundation (grant 709/05), the Minerva Foundation (Munich, Germany), and the Helen and Martin Kimmel Center for Molecular Design. J.M.L.M. is the incumbent of the Baroness Thatcher Professorial Chair of Chemistry and a member *ad personam* of the Lise Meitner-Minerva Center for Computational Quantum Chemistry. The research presented in this article is part of ongoing work in the framework of a Task Group of the International Union of Pure and Applied Chemistry (IUPAC) on “Selected free radicals and critical intermediates: thermodynamic properties from theory and experiment” (2000-013-2-100, renewal 2003-024-1-100). See ref 70 for further details.

Supporting Information Available: Convergence of the contribution of valence higher-order triples, CCSDT–CCSD(T), and valence parenthetical quadruples, CCSDT(Q)–CCSDT, to the total atomization energy, W_n and ATcT total atomization energies at 0 K, and equilibrium geometries. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599.
- (2) Bomble, Y. J.; Vázquez, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2006**, *125*, 064108.
- (3) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2008**, *128*, 114111.
- (4) Karton, A.; Rabinovich, E.; Martin, J. M. L. *J. Chem. Phys.* **2006**, *125*, 144108.
- (5) Ruscic, B.; Pinzon, R. E.; Morton, M. L.; von Laszewski, G.; Bittner, S.; Nijsure, S. G.; Amin, K. A.; Minkoff, M.; Wagner, A. F. *J. Phys. Chem. A* **2004**, *108*, 9979.
- (6) Ruscic, B. Active Thermochemical Tables. In *McGraw-Hill 2005 Yearbook of Science and Technology*; McGraw-Hill: New York, 2004; pp 3–7.
- (7) 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*, 561–570.
- (8) Martin, J. M. L.; Oliveira, G. *J. Chem. Phys.* **1999**, *111*, 1843.
- (9) Parthiban, S.; Martin, J. M. L. *J. Chem. Phys.* **2001**, *114*, 6014.
- (10) Karton, A.; Parthiban, S.; Martin, J. M. L. *J. Phys. Chem. A*, published online Jan 21, <http://dx.doi.org/10.1021/jp8087435>.

- (11) Karton, A.; Martin, J. M. L. *J. Phys. Chem. A* **2007**, *111*, 5936.
- (12) Karton, A.; Martin, J. M. L.; Taylor, P. R. *J. Chem. Phys.* **2007**, *127*, 064104.
- (13) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.01; Gaussian, Inc.: Wallingford, CT.
- (14) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (15) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (16) Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113; Erratum: **2002**, *116*, 3502.
- (17) Jensen, F. *J. Chem. Phys.* **2002**, *116*, 7372.
- (18) Jensen, F.; Helgaker, T. *J. Chem. Phys.* **2004**, *121*, 3463.
- (19) MRCC is a string-based general coupled cluster program suite written by M. Kállay. See also: Kállay, M.; Surján, P. R. *J. Chem. Phys.* **2001**, *115*, 2945, as well as <http://www.mrcc.hu>.
- (20) ACES II (Austin–Mainz–Budapest version) is an electronic structure program system written by Stanton, F. J.; Gauss, J.; Watts, J. D.; Szalay, P. G.; Bartlett, R. J.; with contributions from Auer, A. A.; Bernholdt, D. B.; Christiansen, O.; Harding, M. E.; Heckert, M.; Heun, O.; Huber, C.; Jonsson, D.; Jusélius, J.; Lauderdale, W. J.; Metzroth, T.; Ruud, K.
- (21) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.
- (22) Watts, J. D.; Gauss, J.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 8718.
- (23) MOLPRO is a package of ab initio programs written by Werner, H.-J.; Knowles, P. J.; Schütz, M.; Rauhut, G.; Manby, F. R.; Amos, R. D.; Bernholdt, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Heter, G.; Lloyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklaß, A.; Palmieri, P.; Pitzer, R.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.
- (24) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
- (25) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
- (26) Dunning, T. H.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244.
- (27) Peterson, K. A.; Dunning, T. H. *J. Chem. Phys.* **2002**, *117*, 10548.
- (28) de Jong, W. A.; Harrison, R. J.; Dixon, D. A. *J. Chem. Phys.* **2001**, *114*, 48.
- (29) Widmark, P.-O.; Malmqvist, P.-A.; Roos, B. O. *Theor. Chim. Acta.* **1990**, *77*, 291.
- (30) Widmark, P.-O.; Persson, B. J.; Roos, B. O. *Theor. Chim. Acta.* **1991**, *79*, 419.
- (31) Karton, A.; Martin, J. M. L. *Theor. Chem. Acc.* **2006**, *115*, 330.
- (32) Jensen, F. *Theor. Chem. Acc.* **2005**, *113*, 267.
- (33) Klopper, W. *Mol. Phys.* **2001**, *99*, 481.
- (34) Schwartz, C. In *Methods in Computational Chemistry*; Alder, B. J., Ed.; Academic: New York, 1963; Vol. 2.
- (35) Hill, R. N. *J. Chem. Phys.* **1985**, *83*, 1173.
- (36) Kutzelnigg, W.; Morgan, J. D., III *J. Chem. Phys.* **1992**, *96*, 4484; Erratum: **1992**, *97*, 8821.
- (37) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.
- (38) Knowles, P. J.; Hampel, C.; Werner, H.-J. *J. Chem. Phys.* **2000**, *99*, 5219; Erratum: **2000**, *112*, 3106.
- (39) Schwenke, D. W. *J. Chem. Phys.* **2005**, *122*, 014107.
- (40) Merrick, J. P.; Moran, D.; Radom, L. *J. Phys. Chem. A* **2007**, *111*, 11683.
- (41) Parthiban, S.; Martin, J. M. L. *J. Chem. Phys.* **2001**, *115*, 2051.
- (42) Crawford, T. D.; Kraka, E.; Stanton, J. F.; Cremer, D. *J. Chem. Phys.* **2001**, *114*, 10638.
- (43) NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101, Release 12; Johnson, R. D., III, Ed.; Aug 2005; <http://srdata.nist.gov/cccbdb>.
- (44) Cox, J. D.; Wagman, D. D.; Medvedev, V. A. *CODATA Key Values for Thermodynamics*; Hemisphere Publishing Corp.: New York, 1989. See also <http://www.codata.org/resources/databases/key1.html>.
- (45) Ruscic, B. Unpublished results obtained from Active Thermochemical Tables (ATcT) version 1.25 using the Core Argonne Thermochemical Network version 1.056 (2006).
- (46) (a) Lee, T. J.; Rice, J. E.; Scuseria, G. E.; Schaefer, H. F., III. *Theor. Chim. Acta* **1989**, *75*, 81. (b) Lee, T. J.; Taylor, P. R. *Theor. Chim. Acta* **1989**, *23*, 199.
- (47) (a) Leininger, M. L.; Nielsen, I. M. B.; Crawford, T. D.; Janssen, C. L. *Chem. Phys. Lett.* **2000**, *328*, 431. (b) Nielsen, I. M. B.; Janssen, C. L. *Chem. Phys. Lett.* **1999**, *310*, 568. (c) Janssen, C. L.; Nielsen, I. M. B. *Chem. Phys. Lett.* **1998**, *290*, 423.
- (48) Lories, X.; Vandooren, J.; Peeters, D. *Chem. Phys. Lett.* **2008**, *452*, 29.
- (49) Lau, K.-C.; Ng, C. Y. *J. Chem. Phys.* **2006**, *124*, 044323.
- (50) Davico, G. E.; Bierbaum, V. M.; DePuy, C. H.; Ellison, G. B.; Squires, R. R. *J. Am. Chem. Soc.* **1995**, *117*, 2590.
- (51) Geoffrey, P. F.; Radom, L.; Petersson, G. A.; Barnes, E. C.; Frisch, M. J.; Montgomery, J. A., Jr. *J. Chem. Phys.* **2006**, *125*, 094106.
- (52) Feller, D.; Franz, J. A. *J. Phys. Chem. A* **2000**, *104*, 9017.
- (53) Cramer, C. J.; Nash, J. J.; Squires, R. R. *Chem. Phys. Lett.* **1997**, *277*, 311.
- (54) Lindh, R.; Bernhardsson, A.; Schütz, M. *J. Chem. Phys. A* **1999**, *103*, 9913.
- (55) Smith, C. E.; Crawford, T. D.; Cremer, D. *J. Chem. Phys.* **2005**, *122*, 174309.
- (56) Wenthold, P. G.; Squires, R. R. *J. Am. Chem. Soc.* **1994**, *116*, 6401.
- (57) Roth, W. R.; Hopf, H.; Horn, C. *Chem. Ber.* **1994**, *127*, 1765.
- (58) Jones, R. R.; Bergman, R. G. *J. Am. Chem. Soc.* **1972**, *94*, 660.
- (59) It is also noted that in this case, higher-order connected triple excitations, \hat{T}_3 -(T), have a rather moderate effect on the geometry optimization; with the 6-31G(d,p) basis set, the CCSD(T) and CCSDT optimized bond lengths and diradical separation differ by no more than 0.001 and 0.01 Å, respectively. (See ref 55).
- (60) Wenthold, P. G.; Squires, R. R.; Lineberger, W. C. *J. Am. Chem. Soc.* **1998**, *120*, 5279.
- (61) Li, H.; Yu, S.-Y.; Huang, M.-B.; Wang, Z.-X. *Chem. Phys. Lett.* **2007**, *450*, 12.
- (62) Guner, V.; Khuong, K. S.; Leach, A. G.; Lee, P. S.; Bartberger, M. D.; Houk, K. N. *J. Phys. Chem. A* **2003**, *107*, 11445.
- (63) Ginsburg, D. *Acc. Chem. Res.* **1969**, *2*, 121.
- (64) Feller, D.; Davidson, E. R. *J. Am. Chem. Soc.* **1987**, *109*, 4133.
- (65) Polo, V.; Andres, J.; Silvi, B. *J. Comput. Chem.* **2007**, *28*, 857.
- (66) Wu, W.; Gu, J.; Song, J.; Shaik, S.; Hiberty, P. C. *Angew. Chem., Int. Ed.* **2008**, *48*, 1407.
- (67) Wiberg, K. B.; Dailey, W. P.; Walker, F. H.; Waddell, S. T.; Crocker, L. S.; Newton, M. J. *J. Am. Chem. Soc.* **1985**, *107*, 7247.
- (68) Hedberg, L.; Hedberg, K. J. *J. Am. Chem. Soc.* **1985**, *107*, 7257.
- (69) Schafer, O.; Allen, M.; Szeimies, G.; Sanktjohanser, M. *J. Am. Chem. Soc.* **1992**, *114*, 8180.
- (70) Ruscic, B.; Boggs, J. E.; Bortat, A.; Csaszar, A. G.; Demaison, J.; Janoschek, R.; Martin, J. M. L.; Morton, M. L.; Rossi, M. J.; Stanton, J. F.; Szalay, P. G.; Westmoreland, P. R.; Zabel, F.; Berces, T. *J. Phys. Chem. Ref. Data* **2005**, *34*, 573. See also <http://www.iupac.org/projects/2003/2003-024-1-100.html>.
- (71) Tasi, G.; Izsák, R.; Matisz, G.; Császár, A. G.; Kállay, M.; Ruscic, B.; Stanton, J. F. *Chem. Phys. Chem.* **2006**, *7*, 1664.
- (72) NIST Computational Chemistry Comparison and Benchmark Database, Release 12; Johnson, R. D., Ed.; Aug 2005; <http://srdata.nist.gov/cccbdb/>.