Polycyclic aromatic hydrocarbons (PAHs) are key reference materials for the validation and parameterization of computationally cost-effective procedures such as density functional theory (DFT), semiempirical molecular orbital theory, and molecular mechanics. We obtain accurate heats of formation ($\Delta H_{f,298}$) for 20 PAHs with up to 18 carbon atoms by means of the explicitly correlated W1-F12 thermochemical procedure. The heats of formation are obtained via atomization reactions and quasiisodesmic reactions involving CH$_4$, C$_2$H$_4$, and C$_6$H$_6$ for which accurate experimental $\Delta H_{f,298}$ values are available from the ATcT thermochemical network. We show that for large PAHs the differences between W1-F12 heats of formation obtained from atomization and quasiisodesmic reactions increase with the size of the system and range between 1.7 (C$_7$H$_8$) and 8.9 (Chrysene, C$_{18}$H$_{12}$) kJ mol$^{-1}$. This suggests that atomization reactions should be used with caution for obtaining heats of formation for medium-sized systems even when highly accurate thermochemical procedures (such as W1-F12 theory) are used. For eight PAH compounds (toluene, indene, acenaphthylene, biphenyl, diphenylmethane, anthracene, pyrene, and chrysene) our best theoretical values agree with the best experimental values to within ~1 kJ mol$^{-1}$; for six additional systems (indane, naphthalene, acenaphthene, phenanthrene, and m-terphenyl) agreement between theory and experiment is good with deviations ranging between 2–4 kJ mol$^{-1}$. However, for four systems (p-terphenyl, fluorene, pyracene, and pyracyclene) our best W1-F12 values suggest that the experimental $\Delta H_{f,298}$ values should be revised by significant amounts ranging from 6.5 and 24.4 kJ mol$^{-1}$.

Keywords: Computational thermochemistry, CCSD(T), W1-F12 theory, atomization reactions, isodesmic reactions, Polycyclic aromatic hydrocarbon.
1. Introduction

Polycyclic aromatic hydrocarbons (PAHs) are a major source of pollution with harmful health impacts generated through processes of incomplete combustion of organic materials.\textsuperscript{1} Such processes include natural and human activities such as emission from gas and diesel vehicles, coal-based power plants, and bush fires. PAHs are present in soils, oceans, the atmosphere, and possibly in the interstellar medium.\textsuperscript{1} The smaller PAHs with say 2–5 rings are of particular interest since they play critical roles in the chemistry of these environments.\textsuperscript{2,3,4,5,6,7,8,9,10,11} In addition, many of the PAHs listed as toxic by the United States Environmental Protection Agency (US EPA)\textsuperscript{12} are relatively small PAHs with 2–4 rings (e.g., derivatives of benzene, naphthalene, fluorene).

The heat of formation is the most fundamental thermodynamic property needed for calculating reaction enthalpies. Recent years have witnessed increasingly important roles played by accurate quantum chemical composite procedures in predicting these properties, in particular (i) when accurate experimental values are not available,\textsuperscript{13,14,15,16,17,18,19,20,21,22,23,24} or (ii) combining them with the available experimental data in thermochemical networks.\textsuperscript{25,26,27,28} In this work we will evaluate the experimental heats of formation for 20 medium-sized PAHs by means of the high-level, ab initio W1-F12 thermochemical protocol.\textsuperscript{29} W1-F12 theory represents a layered extrapolations to the all-electron, relativistic CCSD(T)/CBS energy (complete basis-set limit coupled cluster with singles, doubles, and quasiperturbative triple excitations). It is important to note that, as customary in experimental (and high-level computational) thermochemistry, the term ‘accuracy’ here refers to 95% confidence intervals rather than root-mean-square deviations (RMSDs) or mean-absolute deviations (MADs).\textsuperscript{13,30}

Here, we obtain the W1-F12 heats of formation ($\Delta H_{f,298}$) for the 20 PAHs via atomization reactions and reactions involving CH$_4$, C$_2$H$_4$, and C$_6$H$_6$ for which accurate experimental $\Delta H_{f,298}$ values are available from the active thermochemical (ATcT) network.\textsuperscript{31,32,33} In this context, it is important to
mention the works of Wilson et al.\textsuperscript{23} and Bakowies,\textsuperscript{24} which calculated the heats of formation of some of the PAHs considered here using high-level thermochemical protocols. We show that heats of formation obtained from atomization reactions or from reactions involving the above species can differ by amounts ranging from 1.7 kJ mol$^{-1}$ for small PAHs (e.g., C$_7$H$_8$) and up to 8.9 kJ mol$^{-1}$ for medium-sized PAHs (e.g., Chrysene, C$_{18}$H$_{12}$). These results suggest that atomization reactions should be used with caution for obtaining heats of formation for medium-sized systems even when high-level CCSD(T)/CBS composite ab initio methods are used.

2. Computational details

In order to obtain accurate thermochemical properties for the PAHs, calculations have been carried out using the high-level, ab initio, W1-F12 procedure with the Molpro 2012.1 program suite.\textsuperscript{34,35} W1-F12 theory combines explicitly correlated F12 techniques\textsuperscript{36} with basis-set extrapolations in order to approximate the CCSD(T) basis-set-limit energy. Due to the drastically accelerated basis-set convergence of the F12 methods,\textsuperscript{37,38} W1-F12 is superior to the original W1 method\textsuperscript{39} in terms of computational cost.\textsuperscript{29} For the sake of making the article self-contained, we will briefly outline the various steps in W1-F12 theory (for further details see refs. 29 and 40). The Hartree–Fock component is extrapolated from the VDZ-F12 and VTZ-F12 basis sets, using the $E(L) = E_\infty + A/L^\alpha$ two-point extrapolation formula, with $\alpha = 5$ (where V$nZ$-F12 denotes the cc-pV$nZ$-F12 basis sets of Peterson et al.,\textsuperscript{37} which were specifically developed for explicitly correlated calculations). Note that the complementary auxiliary basis set (CABS) singles correction is included in the SCF energy.\textsuperscript{41,42,43} The valence CCSD-F12 correlation energy is extrapolated from the same basis sets, using the above two-point extrapolation with $\alpha = 3.38$. Optimal values for the geminal Slater exponents ($\beta$) used in conjunction with the V$nZ$-F12 basis sets were taken from ref. 38. The (T) valence correlation energy is
obtained from standard CCSD(T) calculations,\textsuperscript{39} namely, extrapolated from the A'VDZ and A'VTZ basis sets using the above two-point extrapolation formula with $\alpha = 3.22$ (where A'V$n$Z indicates the combination of the standard correlation-consistent cc-pV$n$Z basis sets on H and the aug-cc-pV$n$Z basis sets on C).\textsuperscript{44,45} In all of the explicitly correlated coupled cluster calculations the diagonal, fixed-amplitude 3C(FIX) ansatz,\textsuperscript{42,46,47,48} and the CCSD-F12b approximation are employed.\textsuperscript{43,49} The CCSD inner-shell contribution is calculated with the core-valence weighted correlation-consistent cc-pwCVTZ basis set of Peterson and Dunning,\textsuperscript{50} whilst the (T) inner-shell contribution is calculated with the cc-pwCVTZ(no f) basis set (where cc-pwCVTZ(no f) indicates the cc-pwCVTZ basis set without the f functions). The scalar relativistic contribution (in the second-order Douglas–Kroll–Hess approximation)\textsuperscript{51,52} is obtained as the difference between non-relativistic CCSD(T)/A'VDZ and relativistic CCSD(T)/A'VDZ-DK calculations.\textsuperscript{53} The diagonal Born–Oppenheimer corrections are calculated at the HF/cc-pVTZ and CCSD/cc-pVDZ levels of theory using the CFOUR program suite.\textsuperscript{54}

The geometries of all structures have been obtained at the B3LYP-D3/A'VTZ level of theory.\textsuperscript{55,56,57,58,59} We note that the D3 dispersion correction is expected to have a minor effect on the final geometries and a negligible effect on the final W1-F12 energies.\textsuperscript{60} For example, for C$_6$H$_6$ and C$_{10}$H$_8$, the MADs between the B3LYP/A'VTZ and B3LYP-D3/A'VTZ bond distances are 0.00018 Å (over the five unique C=C bonds) and 0.00005 Å (over the three unique C–H bonds). Moreover, the W1-F12 energies calculated using the two geometries differ by 0.000 kJ mol$^{-1}$ (C$_6$H$_6$) and 0.006 kJ mol$^{-1}$ (C$_{10}$H$_8$) (similar results have been obtained for m-terphenyl and p-terphenyl, \textit{vide infra}). Harmonic vibrational frequency analyses have been performed to confirm that all stationary points are equilibrium structures (i.e., they have all real frequencies). Zero-point vibrational energy (ZPVE) and enthalpic temperature corrections ($H_{298}-H_0$) have been obtained from such calculations. The ZPVEs have been scaled by 0.990 as recommended in ref. 40 for the B3LYP/A'VTZ level of theory. We note
however, that using the very similar scaling factor of 0.9896 from ref. 61 changes the scaled ZPVEs by small amounts ranging from 0.1–0.3 kJ mol⁻¹. All geometry optimizations and frequency calculations were performed using the Gaussian 09 program suite.⁶²

3. Results and discussion

3.1. W1-F12 heats of formation from atomization reactions. Table 1 and Figure 1 give an overview of the 20 PAH systems that are considered in the present work. We consider aromatic systems containing up to 18 carbon atoms and up to four rings. Benzene and toluene are not PAHs, but are considered here for the sake of completeness. Some of the systems contain fused aromatic rings (e.g., naphthalene, anthracene, and chrysene) and some systems contain aromatic rings that are not fused (e.g., diphenylmethane, m-terphenyl, and p-terphenyl). We also consider systems that contain nonaromatic rings (e.g., indene, indane, and 11H-benzo[b]fluorene).
Table 1. Overview of the 20 PAH systems which are considered in the present work. The structures are shown in Figure 1.

<table>
<thead>
<tr>
<th>Mol. Formula</th>
<th>Name</th>
<th>Symmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ( \text{C}_6\text{H}_6 )</td>
<td>Benzene</td>
<td>( D_{6h} )</td>
</tr>
<tr>
<td>2 ( \text{C}_7\text{H}_8 )</td>
<td>Toluene</td>
<td>( C_s )</td>
</tr>
<tr>
<td>3 ( \text{C}_9\text{H}_8 )</td>
<td>Indene</td>
<td>( C_s )</td>
</tr>
<tr>
<td>4 ( \text{C}<em>9\text{H}</em>{10} )</td>
<td>Indane</td>
<td>( C_s )</td>
</tr>
<tr>
<td>5 ( \text{C}_{10}\text{H}_8 )</td>
<td>Naphthalene</td>
<td>( D_{2h} )</td>
</tr>
<tr>
<td>6 ( \text{C}_{12}\text{H}_8 )</td>
<td>Acenaphthylene</td>
<td>( C_{2v} )</td>
</tr>
<tr>
<td>7 ( \text{C}_{12}\text{H}_8 )</td>
<td>Biphenylene</td>
<td>( D_{2h} )</td>
</tr>
<tr>
<td>8 ( \text{C}<em>{12}\text{H}</em>{10} )</td>
<td>Acenaphthene</td>
<td>( C_{2v} )</td>
</tr>
<tr>
<td>9 ( \text{C}<em>{12}\text{H}</em>{10} )</td>
<td>Biphenyl</td>
<td>( D_2 )</td>
</tr>
<tr>
<td>10 ( \text{C}_{13}\text{H}_8 )</td>
<td>Fluorene</td>
<td>( C_{2v} )</td>
</tr>
<tr>
<td>11 ( \text{C}<em>{13}\text{H}</em>{12} )</td>
<td>Diphenylmethane</td>
<td>( C_2 )</td>
</tr>
<tr>
<td>12 ( \text{C}_{14}\text{H}_8 )</td>
<td>Pyracyclene</td>
<td>( D_{2h} )</td>
</tr>
<tr>
<td>13 ( \text{C}<em>{14}\text{H}</em>{10} )</td>
<td>Anthracene</td>
<td>( D_{2h} )</td>
</tr>
<tr>
<td>14 ( \text{C}<em>{14}\text{H}</em>{10} )</td>
<td>Phenanthrene</td>
<td>( C_{2v} )</td>
</tr>
<tr>
<td>15 ( \text{C}<em>{14}\text{H}</em>{12} )</td>
<td>Pyracene</td>
<td>( D_{2h} )</td>
</tr>
<tr>
<td>16 ( \text{C}<em>{16}\text{H}</em>{10} )</td>
<td>Pyrene</td>
<td>( D_{2h} )</td>
</tr>
<tr>
<td>17 ( \text{C}<em>{17}\text{H}</em>{12} )</td>
<td>11H-Benzo[b]fluorene</td>
<td>( C_s )</td>
</tr>
<tr>
<td>18 ( \text{C}<em>{18}\text{H}</em>{12} )</td>
<td>Chrysene</td>
<td>( C_{2h} )</td>
</tr>
<tr>
<td>19 ( \text{C}<em>{18}\text{H}</em>{14} )</td>
<td>( m )-Terphenyl</td>
<td>( C_2 )</td>
</tr>
<tr>
<td>20 ( \text{C}<em>{18}\text{H}</em>{14} )</td>
<td>( p )-Terphenyl</td>
<td>( D_2 )</td>
</tr>
</tbody>
</table>
Figure 1. B3LYP-D3/A’VTZ optimized structures for the 20 PAHs considered in the present work. Atomic color scheme: H, white; C, gray.

Since W1-F12 theory represents a layered extrapolation to the all-electron CCSD(T) basis-set-limit energy, it is of interest to estimate whether the contributions from post-CCSD(T) excitations are likely to be significant for the PAHs considered in this work. The percentage of the total atomization
energy accounted for by parenthetical connected triple excitations, %TAE[(T)],\textsuperscript{13,63,64} has been shown to be a reliable energy-based diagnostic for the importance of post-CCSD(T) contributions to the total atomization energies. Table S1 of the Supporting Information gathers the %TAE[(T)] values for the species considered in the present work. The %TAE[(T)] values for these species lie in a very narrow range of 1.9–2.5%. It should be noted that for benzene post-CCSD(T) contributions from W4-F12 theory were found to reduce the atomization energy by 2.1 kJ mol\textsuperscript{-1} (\textit{vide infra}).

Let us proceed to calculating the total atomization energies (TAEs) for the set of 20 PAHs by means of W1-F12 theory. Table 2 gives the component breakdown of the W1-F12 atomization energies as well as the final TAEs at the bottom of the well (TAE\textsubscript{e}). The HF component is very large ranging from 4373.5 (benzene) to 12276.0 (p-terphenyl) kJ mol\textsuperscript{-1}. Despite the magnitude of the HF component, these results are expected to be close to the basis-set limit results. For example, for hydrocarbon cages with up to eight carbon atoms (C\textsubscript{4}H\textsubscript{4}, C\textsubscript{6}H\textsubscript{6}, and C\textsubscript{8}H\textsubscript{8}) the HF component from W1-F12 theory is less than 0.1 kJ mol\textsuperscript{-1} away from results obtained from W2-F12 theory.\textsuperscript{65}
Table 2. Component breakdown of the W1-F12 atomization energies for the 20 PAHs considered in the present work and predicted theoretical total atomization energies (kJ mol⁻¹).

<table>
<thead>
<tr>
<th></th>
<th>HF¹</th>
<th>CCSD-F12²</th>
<th>(T)³</th>
<th>CV⁴</th>
<th>Rel⁵</th>
<th>SO⁶</th>
<th>DBOC⁷</th>
<th>TAE,⁸</th>
<th>ΔZPVE⁹</th>
<th>TAE¹⁰</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene</td>
<td>4373.5</td>
<td>1216.9</td>
<td>111.7</td>
<td>29.9</td>
<td>-4.2</td>
<td>-2.1</td>
<td>0.5</td>
<td>5726.3</td>
<td>260.9</td>
<td>5465.3</td>
</tr>
<tr>
<td>Toluene</td>
<td>5328.8</td>
<td>1482.4</td>
<td>128.6</td>
<td>35.0</td>
<td>-5.0</td>
<td>-2.5</td>
<td>0.6</td>
<td>6968.1</td>
<td>331.7</td>
<td>6636.4</td>
</tr>
<tr>
<td>Indene</td>
<td>6307.0</td>
<td>1799.6</td>
<td>171.0</td>
<td>45.0</td>
<td>-6.3</td>
<td>-3.2</td>
<td>0.8</td>
<td>8313.9</td>
<td>364.3</td>
<td>7949.5</td>
</tr>
<tr>
<td>Indane</td>
<td>6789.3</td>
<td>1912.1</td>
<td>164.7</td>
<td>44.8</td>
<td>-6.5</td>
<td>-3.2</td>
<td>0.8</td>
<td>8902.0</td>
<td>426.5</td>
<td>8475.5</td>
</tr>
<tr>
<td>Naphthalene</td>
<td>6875.3</td>
<td>1944.9</td>
<td>195.4</td>
<td>50.2</td>
<td>-6.9</td>
<td>-3.5</td>
<td>0.8</td>
<td>9056.2</td>
<td>382.9</td>
<td>8673.2</td>
</tr>
<tr>
<td>Acenaphthylene</td>
<td>7839.4</td>
<td>2268.0</td>
<td>241.6</td>
<td>60.2</td>
<td>-8.3</td>
<td>-4.2</td>
<td>1.0</td>
<td>10397.6</td>
<td>413.8</td>
<td>9983.9</td>
</tr>
<tr>
<td>Biphenylene</td>
<td>7683.7</td>
<td>2266.9</td>
<td>240.7</td>
<td>59.5</td>
<td>-8.2</td>
<td>-4.2</td>
<td>1.0</td>
<td>10239.3</td>
<td>410.7</td>
<td>9828.6</td>
</tr>
<tr>
<td>Acenaphthene</td>
<td>8330.7</td>
<td>2377.7</td>
<td>232.8</td>
<td>60.2</td>
<td>-8.5</td>
<td>-4.2</td>
<td>1.0</td>
<td>10989.6</td>
<td>473.5</td>
<td>10516.1</td>
</tr>
<tr>
<td>Biphenyl</td>
<td>8324.6</td>
<td>2358.1</td>
<td>233.0</td>
<td>60.0</td>
<td>-8.3</td>
<td>-4.2</td>
<td>1.0</td>
<td>10964.2</td>
<td>471.7</td>
<td>10492.4</td>
</tr>
<tr>
<td>Fluorene</td>
<td>8854.4</td>
<td>2523.6</td>
<td>253.8</td>
<td>65.4</td>
<td>-9.1</td>
<td>-4.6</td>
<td>1.1</td>
<td>11684.6</td>
<td>488.0</td>
<td>11196.6</td>
</tr>
<tr>
<td>Diphenylnmethane</td>
<td>9266.5</td>
<td>2621.8</td>
<td>248.3</td>
<td>65.0</td>
<td>-9.1</td>
<td>-4.6</td>
<td>1.1</td>
<td>12189.0</td>
<td>544.6</td>
<td>11644.4</td>
</tr>
<tr>
<td>Pyracyclene</td>
<td>8740.9</td>
<td>2596.9</td>
<td>288.1</td>
<td>69.9</td>
<td>-9.6</td>
<td>-5.0</td>
<td>1.2</td>
<td>11682.3</td>
<td>443.3</td>
<td>11239.0</td>
</tr>
<tr>
<td>Anthracene</td>
<td>9355.1</td>
<td>2677.3</td>
<td>281.2</td>
<td>70.3</td>
<td>-9.7</td>
<td>-5.0</td>
<td>1.1</td>
<td>12370.4</td>
<td>503.8</td>
<td>11866.6</td>
</tr>
<tr>
<td>Phenanthrene</td>
<td>9384.1</td>
<td>2675.4</td>
<td>279.3</td>
<td>70.4</td>
<td>-9.7</td>
<td>-5.0</td>
<td>1.1</td>
<td>12395.6</td>
<td>505.0</td>
<td>11890.6</td>
</tr>
<tr>
<td>Pyracylene</td>
<td>9750.6</td>
<td>2814.1</td>
<td>271.0</td>
<td>70.0</td>
<td>-10.1</td>
<td>-5.0</td>
<td>1.2</td>
<td>12891.9</td>
<td>563.4</td>
<td>12328.5</td>
</tr>
<tr>
<td>Pyrene</td>
<td>10448.7</td>
<td>2992.5</td>
<td>326.2</td>
<td>80.7</td>
<td>-11.1</td>
<td>-5.7</td>
<td>1.3</td>
<td>13832.8</td>
<td>538.3</td>
<td>13294.4</td>
</tr>
<tr>
<td>Benzo[b]fluorene</td>
<td>11356.6</td>
<td>3252.3</td>
<td>338.0</td>
<td>85.6</td>
<td>-11.9</td>
<td>-6.0</td>
<td>1.2</td>
<td>15015.8</td>
<td>609.4</td>
<td>14406.4</td>
</tr>
<tr>
<td>Chrysene</td>
<td>11883.8</td>
<td>3407.9</td>
<td>364.2</td>
<td>90.6</td>
<td>-12.5</td>
<td>-6.4</td>
<td>1.4</td>
<td>15729.1</td>
<td>627.0</td>
<td>15102.1</td>
</tr>
<tr>
<td>m-Terphenyl</td>
<td>12275.7</td>
<td>3500.1</td>
<td>354.5</td>
<td>90.2</td>
<td>-12.5</td>
<td>-6.4</td>
<td>1.4</td>
<td>16203.0</td>
<td>682.0</td>
<td>15521.1</td>
</tr>
<tr>
<td>p-Terphenyl</td>
<td>12276.0</td>
<td>3499.7</td>
<td>354.5</td>
<td>90.2</td>
<td>-12.5</td>
<td>-6.4</td>
<td>1.5</td>
<td>16203.1</td>
<td>682.0</td>
<td>15521.1</td>
</tr>
</tbody>
</table>

¹Extrapolated from the cc-pVDZ-F12 and cc-pVTZ-F12 basis sets. ²Extrapolated from the A'VDZ and A'VTZ basis sets. ³CCSD(T) core-valence correction obtained as: CCSD/cc-pCVTZ + (T)/cc-pCVTZ(no f). ⁴CCSD(T)/cc-pVDZ-DK scalar relativistic correction. ⁵First-order atomic spin-orbit correction. ⁶CCSD diagonal Born–Oppenheimer correction. ⁷Vibrationless, relativistic, all-electron CCSD(T)/CBS total atomization energies. ⁸Zero-point vibrational energy corrections from B3LYP-D3/A'VTZ harmonic frequencies scaled by 0.990. ⁹ZPVE-inclusive, relativistic, all-electron CCSD(T)/CBS total atomization energies.

The valence CCSD-F12 correlation contribution increases the TAEs by amounts ranging from 1216.9 (benzene) to 3500.1 (p-terphenyl) kJ mol⁻¹. These values are expected to overestimate the CCSD/CBS values by chemically significant amounts for the larger PAHs considered in the present work. For example, for smaller hydrocarbons extrapolating the CCSD-F12 energy from the V[D,T]Z-F12 basis set pair overestimates the CCSD-F12/V[T,Q]Z-F12 values from W2-F12 theory by 0.13 (tetrahedrane), 0.33 (triprismane), and 0.52 (cubane) kJ mol⁻¹.² We expect that the deviations for the larger PAHs considered here would be even larger. The valence (T) correlation contributions still amount to hundreds of kJ mol⁻¹, namely they range between 111.7 (benzene) and 364.2 (chrysene) kJ.
The core-valence (CV) correlation contributions are also large, ranging from 29.9 (benzene) to 90.6 (chrysene) kJ mol$^{-1}$. The scalar relativistic contributions reduce the atomization energies by relatively large amounts ranging from 4.2 (benzene) to 12.5 (chrysene, $m$-terphenyl, and $p$-terphenyl) kJ mol$^{-1}$. Similarly, the first-order, atomic spin-orbit coupling contributions systematically reduce the atomization energies by up to 6.4 kJ mol$^{-1}$ (chrysene, $m$-terphenyl, and $p$-terphenyl).

The DBOC contributions at the HF/cc-pVTZ level of theory increase the TAEs by amounts of up to 2.7 kJ mol$^{-1}$ ($m$-terphenyl and $p$-terphenyl). However, the CCSD correlation contribution reduces the DBOCs by about 50% (see Table S2 of the Supporting Information). Thus, the overall DBOC contributions range between 0.5 (benzene) to 1.5 ($p$-terphenyl) kJ mol$^{-1}$. These CCSD correlation corrections to the DBOCs are consistent with previous results for systems with many hydrogen atoms.$^{13,40,65,66}$

Summing up the SCF, CCSD, (T), CV, Rel, fist-order SO, and DBOC contributions to the TAEs we obtain TAEs at the bottom-of-the-well (TAE$_e$) which are summarized in Table 2. The TAE$_e$ values are converted to TAEs at 0 K (TAE$_0$) by including scaled ZPVE corrections obtained from B3LYP-D3/A'VTZ harmonic frequencies. We convert the TAE$_0$ values to heats of formation at 0 K using ATcT atomic heats of formation at 0 K.$^{31,32,33}$ Subsequently, these $\Delta H_0$ values are converted to heats of formation at 298 K ($\Delta H_{298}^0$) using the CODATA$^{67}$ enthalpy functions ($H_{298}$–$H_0$) for the elemental reference states and the enthalpy functions for the PAHs are obtained (within the rigid-rotor harmonic oscillator approximation) from the B3LYP-D3/A'VTZ geometries and harmonic frequencies. Our final W1-F12 $\Delta H_0$ and $\Delta H_{298}^0$ values, which are obtained from atomization reactions, are listed in Table 3. For benzene, naphthalene, and anthracene $\Delta H_{298}^0$ values have been previously calculated at the W1-F12 level via atomization reactions.$^{29}$ These $\Delta H_{298}^0$ values differ from the ones obtained in the present work by 0.8 (benzene), 1.4 (naphthalene), and 2.2 (anthracene) kJ mol$^{-1}$ (Table 3). The
differences between the two sets of values are attributed, for the most part, to the different treatments of the DBOC and ZPVE terms.

Table 3. W1-F12 heats of formation at 298 K (ΔfH°298, kJ mol⁻¹) derived from atomization and quasiisodesmic reactions compared to experimental and theoretical data.

![Table](image)

3.2. Potential errors in the W1-F12 total atomization energies. An inspection of Table 2 reveals that for PAHs with more than 16 carbons the SCF component can exceed 10,000 kJ mol⁻¹, the CCSD component can exceed 3,000 kJ mol⁻¹, the ZPVE 600 kJ mol⁻¹, and the final TAE₀ values can exceed as much as 15,000 kJ mol⁻¹. Thus, even an error of a fraction of a percent in these components due to basis-set incompleteness and/or other deficiencies would lead to chemically significant errors in heats of formation derived from these TAEs. For example, a
0.01% error in the HF/V{D,T}Z-F12 component would lead to errors ranging between 0.4 (benzene) and 1.2 (terphenyl) kJ mol\(^{-1}\). An error of 0.1% in the correlation CCSD-F12/V{D,T}Z-F12 component would lead to errors ranging between 1.2 (benzene) and 3.5 (terphenyl) kJ mol\(^{-1}\). An error of 0.5% in the (T)/A’V{D,T}Z component would lead to errors ranging between 0.6 (benzene) and 1.8 (chrysene) kJ mol\(^{-1}\). Finally, an error of 3% in the core-valence (CV) correlation component would translate to errors ranging between 0.9 (benzene) and 2.7 (chrysene) kJ mol\(^{-1}\). Such errors are not outside the realm of possibility even for a high-level thermochemical protocol such as W1-F12. This is demonstrated in Table 4, which lists the SCF, CCSD, (T), and CV components obtained from W1-F12 and higher-level \(W_n\)-F12 theories for three illustrative medium-sized hydrocarbons (namely, benzene, cubane, and naphthalene). For the SCF component we obtain percentage errors ranging from 0.01% (cubane) to 0.03% (benzene). For the CCSD correlation component we obtain errors ranging from 0.04% (naphthalene) up to 0.13% (cubane). As for the (T) correlation component, the W1-F12 and W4-F12 values for benzene are practically identical, which is probably fortuitous. However, for naphthalene and cubane we obtain percentage errors of 0.2% and 0.6%, respectively. Finally, for the CV component we obtain errors of about 3% for all three compounds. As mentioned above, for systems like chrysene and terphenyl, these percentage errors could translate into absolute errors of up to 4 kJ mol\(^{-1}\) in the final heats of formation obtained from TAEs calculated at the W1-F12 level.
Table 4. Comparison between the SCF, CCSD, (T), and CV components of the TAE obtained from W1-F12 theory and higher-level \( W_n \) theories for three medium-sized hydrocarbons (in kJ mol\(^{-1}\)).

<table>
<thead>
<tr>
<th></th>
<th>SCF</th>
<th>( W_n^a )</th>
<th>% Error(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene (C(_6)H(_6))</td>
<td>4373.5</td>
<td>4372.3(^c)</td>
<td>0.03</td>
</tr>
<tr>
<td>SCF</td>
<td>1216.9</td>
<td>1216.3(^c)</td>
<td>0.05</td>
</tr>
<tr>
<td>(T)</td>
<td>111.7</td>
<td>111.7(^c)</td>
<td>0.01</td>
</tr>
<tr>
<td>CV</td>
<td>29.9</td>
<td>30.8(^c)</td>
<td>3.02</td>
</tr>
<tr>
<td>Cubane (C(_8)H(_8))</td>
<td>5246.2</td>
<td>5246.7(^d)</td>
<td>0.01</td>
</tr>
<tr>
<td>SCF</td>
<td>1714.6</td>
<td>1712.4(^d)</td>
<td>0.13</td>
</tr>
<tr>
<td>(T)</td>
<td>144.1</td>
<td>143.3(^d)</td>
<td>0.61</td>
</tr>
<tr>
<td>CV</td>
<td>37.1</td>
<td>38.1(^d)</td>
<td>2.64</td>
</tr>
<tr>
<td>Naphthalene (C(_{10})H(_8))</td>
<td>6875.3</td>
<td>6875.4(^d)</td>
<td>0.00</td>
</tr>
<tr>
<td>SCF</td>
<td>1944.9</td>
<td>1944.2(^d)</td>
<td>0.04</td>
</tr>
<tr>
<td>(T)</td>
<td>195.4</td>
<td>195.7(^d)</td>
<td>0.16</td>
</tr>
<tr>
<td>CV</td>
<td>50.2</td>
<td>51.8(^d)</td>
<td>3.01</td>
</tr>
</tbody>
</table>

\(^a\) Calculated at the following levels: \( W_4 \)-F12 (benzene), \( W_2 \)-F12 (cubane), and \( W_2h \) (naphthalene). \(^b\) Percentage error in W1-F12 relative to the higher-level \( W_n \) results. \(^c\) Taken from ref. 72. \(^d\) Taken from ref. 65.

For benzene we can also estimate the error due to the neglect of post-CCSD(T) correlation contributions. A computational foray into the theoretical TAE of benzene found that the CCSDT–CCSD(T) (T–(T)) component extrapolated from cc-pV\{T,Q\}Z basis sets amounts to \(-11.2\) kJ mol\(^{-1}\) and the CCSDT(Q)–CCSDT ((Q)) component calculated with the cc-pVDZ basis set amounts to \(+6.8\) kJ mol\(^{-1}\). A more recent investigation at the W4-F12 level,\(^72\) found that the T–(T)/cc-pV\{D,T\}Z component of \(-11.0\) kJ mol\(^{-1}\) is in excellent agreement with the above T–(T)/cc-pV\{T,Q\}Z value. However, the (Q)/cc-pV\{D,T\}Z component of \(+8.9\) kJ mol\(^{-1}\) was found to be substantially larger than the above (Q)/cc-pVDZ value. Using the values from ref. 72 we obtain a post-CCSD(T) contribution of \(-2.1\) kJ mol\(^{-1}\) for benzene.

Last but not least, another potential source of error in our W1-F12 atomization energies is the ZPVE component. The ZPVEs in the considered PAHs range between 260.9 (benzene) and 682.0 (terphenyl) kJ mol\(^{-1}\), thus even an error of a fraction of a percent, due to neglect of explicit
anharmonicity and inaccuracies in the harmonic frequencies, will translate into chemically significant errors in the final TAEs. It has been previously noted that for hydrocarbons the factor limiting accuracy of W1-F12 theory and similar thermochemical protocols will increasingly be the quality of the ZPVE.\textsuperscript{20,73}

3.3 W1-F12 heats of formation from quasiisodesmic reactions. The reason for the large components of the TAEs, and indeed the large TAEs themselves, is that atomization reactions do not conserve any of the bonding environment between reactants and products. It is well established that, for a given level of theory, the reaction energy becomes more accurate as larger molecular fragments are conserved on the two sides of the reaction due to an increasing degree of error cancellation between reactants and products.\textsuperscript{17,18,20,23,65,74,75,76,77,78,79,80,81,82,83,84,85} Nevertheless, obtaining the heat of formation for a given molecule using alternative reaction is only possible if highly accurate heats of formation are available for all the molecular fragments involved in the reaction (apart from the parent molecule). The success of this approach relies on three key factors: (i) the accuracy of the available experimental or theoretical heats of formation used for the molecular fragments, (ii) the level of theory at which the reaction enthalpy is calculated, and (iii) the degree to which the chosen reaction conserves the chemical environments on the two sides of the reaction (see also additional discussion in ref. 79). In this section we obtain heats of formation for the set of PAHs via two quasiisodesmic reactions that consider the above criteria:

\begin{align*}
C_nH_m & \rightarrow a \text{C}_6\text{H}_6 + b \text{CH}_4 \quad (1) \\
C_nH_m & \rightarrow a \text{C}_6\text{H}_5 + b \text{C}_2\text{H}_4 \quad (2)
\end{align*}
We note that with few exceptions (e.g., reaction (2) for C_{10}H_{8}), reactions (1) and (2) are not isodesmic, that is they do not conserve the numbers of each formal bond type. Nevertheless, they do preserve the chemical environments of the reactants and products to a large extent since they involve aromatic molecules on both sides of the reaction. Thus, we will collectively refer to these reactions as quasiisodesmic. Table 3 lists the W1-F12 heats of formation at 298 K ($\Delta fH^{\circ}_{298}$) obtained from reactions (1) and (2) (the stoichiometric coefficients are given in Table S3 of the Supporting information). Inspection of Table 3 reveals that in most cases the W1-F12 $\Delta fH^{\circ}_{298}$ values obtained from reactions (1) and (2) differ by less than 1 kJ mol\(^{-1}\) from each other and the maximum differences do not exceed 2.3 kJ mol\(^{-1}\). These relatively small differences between the heats of formation obtained from reactions (1) and (2) indicate that the degree to which these reactions conserve the chemical environments on the two sides of the reaction are sufficiently high at the W1-F12 level of theory. Or in other words, an indication that the reaction enthalpy is calculated using sufficiently high levels of theory (for both the electronic and secondary energy contributions) is provided by the resultant heat of formation showing relatively little variation with respect to the choice of the reaction. It is important to note that accurate experimental ATcT heats of formation, associated with well-defined error bars, are available for the auxiliary species considered in reactions (1) and (2).\(^{31,32,33}\) Namely, we use the following ATcT heats of formation:

$\Delta fH^{\circ}_{298}[\text{CH}_4(\text{g})] = -74.53 \pm 0.06$, $\Delta fH^{\circ}_{298}[\text{C}_2\text{H}_4(\text{g})] = 52.35 \pm 0.12$, and $\Delta fH^{\circ}_{298}[\text{C}_6\text{H}_6(\text{g})] = 83.10 \pm 0.23$ kJ mol\(^{-1}\).\(^{69}\) We can estimate an overall uncertainty to the W1-F12 heat of formation for the parent molecule calculated form the reactions (1) and (2) using the above experimental uncertainties for methane, ethylene, and benzene and an estimated uncertainty of 4.6 kJ mol\(^{-1}\) assigned to the W1-F12 reaction enthalpy. The uncertainty of 4.6 kJ mol\(^{-1}\) assigned to the W1-F12 reaction enthalpy is taken as twice the RMSD reported in ref. 29 for a set of 97 first-row
total atomization energies and a residual uncertainty for the scaled DFT ZPVE component taken as twice the RMSD reported in ref. 61. We note that since the experimental ATcT uncertainties are very small the final estimated uncertainties for the parent molecules are dominated by the uncertainties for the W1-F12 reaction enthalpies.

3.4. Comparison between the heats of formation obtained from reactions (1) and (2) and atomization reactions. Overall, the choice between which reaction to use (i.e., (1) or (2)) is somewhat arbitrary due to the fairly small variation in the predicted heats of formation (Table 3). The average deviation between the $\Delta H_{298}^\circ$ values predicted from reactions (1) and (2) is merely 1.2 kJ mol$^{-1}$ and the maximum difference does not exceed 2.3 kJ mol$^{-1}$. These differences are below the intrinsic accuracy of the W1-F12 method. Having said that, it can be argued that reaction (1) is better suited for treating PAHs which contain both sp$^2$ and sp$^3$ carbons and reaction (2) is better suited for treating PAHs which contain only sp$^3$ carbons. These $\Delta H_{298}^\circ$ values are chosen as our best theoretical predictions and are listed in a separate column in Table 3. Comparison of our best $\Delta H_{298}^\circ$ values obtained from reactions (1) and (2) with those obtained from atomization reaction reveal very significant differences reaching up to 8.9 kJ mol$^{-1}$ (!!). Figure 2 shows these differences for the PAHs considered in the present work. Inspection of this figure reveals a general correlation between the number of carbon atoms in the PAHs and the deviation between the $\Delta H_{298}^\circ$ values obtained from atomization reactions and reactions (1) or (2). Namely, for PAHs with less than 10 carbons (toluene, indene, and indane) the deviations range between 1.7–2.8 kJ mol$^{-1}$; for PAHs with 10–13 carbons (naphthalene, acenaphthylene, biphenylene, acenaphthene, biphenyl, fluorene, and diphenylmethane) the deviations range between 3.9–5.9 kJ mol$^{-1}$; for PAHs with 14–18 carbons (pyracyclene, anthracene, phenanthrene,
pyrene, benzo[b]fluorene chrysene, and terphenyl) the deviations range between 6.0–8.9 kJ mol⁻¹.

Figure 2. Dependence of the difference between $\Delta H_{298}^\circ$ values obtained from atomization and quasiisodesmic reactions ($\Delta E = \Delta H_{298}^\circ$ (atomization) – $\Delta H_{298}^\circ$ (quasiisodesmic), in kJ mol⁻¹) on the number of carbon atoms in the PAH.

The general increase in discrepancy between $\Delta H_{298}^\circ$ values obtained from atomization and quasiisodesmic reactions with the number of carbons in the PAH suggests that the errors in the atomization energies tend to increase with the size of the system and illustrates the potential issues of using atomization reactions for obtaining heats of formations for medium-sized systems even when highly accurate CCSD(T)/CBS composite ab initio methods are used. In this context, it should be pointed out that it has been shown that post-CCSD(T) contributions to the total atomization energies also tend to increase with the size of the system.⁸⁶,⁸⁷

3.5. Comparison between the heats of formation obtained from reactions (1) and (2) with available experimental data. Table 3 lists the available experimental heats of formations for the PAHs considered in the present work from the compilation by Roux, Temprado, Chickos, and
Nagano (RTCN).\textsuperscript{68} The $\Delta H^\circ_{298}$ value for toluene $50.1 \pm 1.1$ (RTCN) is in excellent agreement with the ATcT value of $50.01 \pm 0.34$ kJ mol$^{-1}$. The two values agree to within 0.1 kJ mol$^{-1}$, albeit the ATcT value is associated with a significantly smaller uncertainty. At the W1-F12 level we obtain $\Delta H^\circ_{298} = 50.8$ kJ mol$^{-1}$ which is less than 1 kJ mol$^{-1}$ away from experiment. This small discrepancy between W1-F12 increases our confidence in the combination of the chosen quasiisodesmic reaction and W1-F12 level of theory. The deviations between our best theoretical $\Delta H^\circ_{298}$ values and the experimental RTCN values for the other PAHs are listed in Table 3. For five systems the theoretical $\Delta H^\circ_{298}$ values are practically identical to the RTCN values with deviations (given in parentheses) smaller than 0.5 kJ mol$^{-1}$, namely they are: chrysene ($-0.1$), acenaphthylene ($-0.3$), biphenyl ($-0.5$), anthracene (0.3), and pyrene (0.4 kJ mol$^{-1}$). For three additional PAHs there is excellent agreement between theory and experiment with deviations below 2 kJ mol$^{-1}$, namely: naphthalene ($-1.7$), indene ($-1.4$), and diphenylmethane (1.2 kJ mol$^{-1}$). For five compounds the deviations between theory and experiment are somewhat larger, but still below the uncertainties associated with the W1-F12 values, namely the deviations are: indane ($-2.7$), acenaphthene ($-2.8$), biphenylene (2.2), phenanthrene (3.5), and m-terphenyl (4.5 kJ mol$^{-1}$).

Overall, for 14 of the 18 systems there is good-to-excellent agreement between theory and experiment with deviations ranging between 0.3 and 4.5 kJ mol$^{-1}$. However, for four systems our W1-F12 heats of formations obtained via quasiisodesmic reactions indicate that the experimental RTCN values should be substantially revised. Namely, the deviations between theory and experiment are $-8.9$ (p-terphenyl), 9.2 (fluorene), 16.4 (pyracene), and 24.4 (pyracyclene) kJ mol$^{-1}$. For comparison, the sum of the respective uncertainties is 8.5 (p-terphenyl), 7.7 (fluorene), and 9.9 (pyracene) kJ mol$^{-1}$. These results suggest that the above
experimental values (especially for pyracene and pyracyclene) should be reexamined. We note that our best theoretical heat of formation for fluorene deviates by 6.5 kJ mol\(^{-1}\) from the more recent experimental value of Monte et al.\(^{70}\) i.e., the two values agree to within overlapping uncertainties. Our best theoretical value for fluorene is also in good agreement with the recent theoretical value of Bakowies obtained with the ATOMIC(hc) protocol,\(^{24}\) namely the two values deviate by 2.4 kJ mol\(^{-1}\) (Table 3). Our best theoretical value for pyracyclene is also in excellent agreement with the ATOMIC(hc) value, i.e., the two values deviate by merely 0.4 kJ mol\(^{-1}\) (Table 3).

Finally, it should be pointed out that whilst the experimental heats of formation for m- and p-terphenyl suggest that the former isomer is more stable by 4.4 kJ mol\(^{-1}\) (Table 3), our theoretical heats of formation indicate that these isomers are isoenergetic. In this context, it should be pointed out that due to the structural similarity of these isomers, the W1-F12 isomerization energy should be associated with a much smaller uncertainty than the heats of formation. We also note that, as discussed in the Computational details Section, the D3 dispersion correction has a minor effect on the geometries of the isomers. For example, for the m-terphenyl isomer the MAD between the B3LYP/A'VTZ and B3LYP-D3/A'VTZ bond distances is 0.00035 Å (over the unique C=C and C–C bonds) and 0.00004 Å (over the unique C–H bonds). In addition, the G4(MP2) energies calculated using the two geometries differ by small amounts of 0.04 (m-terphenyl) and 0.09 (p-terphenyl) kJ mol\(^{-1}\).
4. Conclusions

We obtain accurate heats of formation ($\Delta H_{f298}$) for a set of 20 PAHs with up to 18 carbon atoms at the CCSD(T)/CBS level of theory by means of W1-F12 theory. We obtain the heats of formation via atomization reactions and quasiisodesmic reactions involving CH$_4$, C$_2$H$_4$, and C$_6$H$_6$ for which accurate experimental $\Delta H_{f298}$ values are available from the ATcT thermochemical network. We show that for the larger PAHs with 14–18 carbon atoms the deviations between the heats of formation obtained from atomization and quasiisodesmic reactions can differ by chemically significant amounts ranging between 6.0 (benzo[b]fluorene) and 8.9 (chrysene) kJ mol$^{-1}$. For smaller PAHs with 10–13 the deviations between the heats of formation obtained from atomization and quasiisodesmic reaction still differ by chemically significant amounts ranging from 3.9 (acenaphthene and diphenylmethane) and 5.9 (acenaphthylene and biphenylene) kJ mol$^{-1}$. These discrepancies between the two approaches for obtaining heats of formation, demonstrate that errors in total atomization energies can become chemically significant for medium-sized systems even when high-level composite ab initio procedures such as W1-F12 theory are used.

Our best W1-F12 heats of formation ($\Delta H_{f298}$) values obtained from quasiisodesmic reactions are: 50.8 (toluene), 159.8 (indene), 58.2 (indane), 148.9 (naphthalene), 262.9 (acenaphthylene), 419.4 (biphenylene), 154.0 (acenaphthene), 179.8 (biphenyl), 185.9 (fluorene), 164.9 (diphenylmethane), 433.0 (pyracyclene), 229.7 (anthracene), 205.7 (phenanthrene), 190.7 (pyracene), 225.9 (pyrene), 249.6 (benzo[b]fluorene), 268.6 (chrysene), 275.5 (m-terphenyl and p-terphenyl) kJ mol$^{-1}$. The W1-F12 heats of formation are all associated with similar uncertainties of $\pm$ 4.6 (or $\pm$ 4.7) kJ mol$^{-1}$. We note that no experimental gas-phase enthalpy of
formation has been reported for benzo[b]fluorene, thus the above theoretical value is recommended for this system.

For six systems (toluene, chrysene, acenaphthylene, biphenyl, anthracene, and pyrene) our best theoretical heats of formation are less than 1 kJ mol\(^{-1}\) away from the experimental values compiled by RTCN. For eight additional systems (naphthalene, indene, diphenylmethane, indane, acenaphthene, biphenylene, phenanthrene, and m-terphenyl) there is good-to-excellent agreement between theory and experiment with deviations ranging between 1.2–4.5 kJ mol\(^{-1}\). However, for four systems our accurate W1-F12 heats of formations indicate that the experimental RTCN values should be substantially revised. For p-terphenyl and fluorene the deviations between theory and experiment are nearly 10 kJ mol\(^{-1}\), and for pyracene and pyracyclene they are even larger, namely 16.4 and 24.4 kJ mol\(^{-1}\), respectively. (We note that the W1-F12 heat of formation for fluorene is in better agreement with the more recent experimental value of Monte \textit{et al.},\textsuperscript{70} i.e., the two values deviate by 6.5 kJ mol\(^{-1}\).) These large discrepancies between theory and experiment are larger than any potential deficiencies in our theoretical methodology (e.g., neglect of post-CCSD(T) excitations and explicit anharmonicity in the ZPVE) the effect of which should be diminished by the use of quasisodesmic reactions that conserve large chemical environments on the two sides of the reaction. We also note that whilst the experimental heats of formation for m- and p-terphenyl suggest that the former isomer is more stable by 4.4 kJ mol\(^{-1}\), our theoretical heats of formation suggest that these two isomers are isoenergetic.
**Supplementary data**

Diagnostics indicating the importance of post-CCSD(T) correlation effects for the species considered in this work (Table S1); diagonal Born–Oppenheimer corrections for the species considered in this work (Table S2); stoichiometric coefficients for reactions (1) and (2) for all species considered in this work (Table S3); CCSD(T)/CBS heats of formation at 0 K (Table S4); W1-F12 thermochemical data for CH₄ and C₂H₄ (Table S5); B3LYP-D3/A'VTZ optimized geometries for the species considered in this work (Table S6); B3LYP-D3/A'VTZ harmonic frequencies for the species considered in this work (Table S7); chemical compounds and models considered in the present work (Table S8); and full references for Molpro 2012.1, CFOUR, and Gaussian 09 (Table S9).

**Corresponding author**

*E-Mail: amir.karton@uwa.edu.au*

**Acknowledgments**

We gratefully acknowledge the generous allocation of computing time from the National Computational Infrastructure (NCI) National Facility, and system administration support provided by the Faculty of Science at UWA to the Linux cluster of the Karton group. AK is the recipient of an Australian Research Council (ARC) Future Fellowship (Project No. FT170100373). The authors would like to thank the referees of this manuscript for their valuable comments and suggestions.
References


81 Wan, W.; Karton, A. Heat of formation for C\textsubscript{60} by means of the G4(MP2) thermochemical protocol through reactions in which C\textsubscript{60} is broken down into corannulene and sumanene. Chem. Phys. Lett. 2016, 643, 34.


