ACS PHYSICAL CHEMISTRY Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

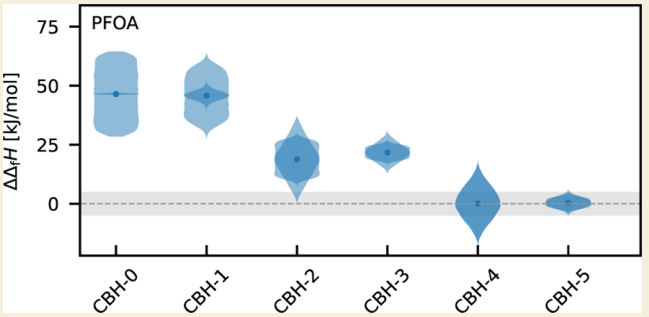
pubs.acs.org/physchemau
Article

Accurate Enthalpies of Formation for PFAS from First-Principles: Combining Different Levels of Theory in a Generalized Thermochemical Hierarchy

Kento Abeywardane and C. Franklin Goldsmith*

Cite This: ACS Phys. Chem Au 2024, 4, 247−258
Read Online

ABSTRACT: The enthalpies of formation are computed for a large number of per- and poly fluoroalkyl substances (PFAS) using a connectivity-based hierarchy (CBH) approach. A combination of different electronic structure methods are used to provide the reference data in a hierarchical manner. The ANL0 method, in conjunction with the active thermochemical tables, provides enthalpies of formation for smaller species with subchemical accuracy. Coupled-cluster theory with explicit correlations are used to compute enthalpies of formation for intermediate species, based upon the ANL0 results. For the largest PFAS, including perfluorooctanoic acid (PFOA) and heptafluoropropylene oxide dimer acid (GenX), coupled-cluster theory with local correlations is used. The sequence of homodesmotic reactions proposed by the CBH are determined automatically by a new open-source code, AutoCBH. The results are the first reported enthalpies of formation for the majority of the species. A convergence analysis and global uncertainty quantification confirm that the enthalpies of formation at 0 K should be accurate to within ±5 kJ/mol. This new approach is not limited to PFAS, but can be applied to many chemical systems.



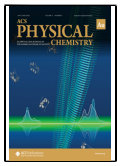
KEYWORDS: *PFAS, generalized hierarchy, CBH, PFOA, GenX, thermochemistry*

1. INTRODUCTION

Per- and poly fluoroalkyl substances (PFAS) are a major environmental hazard. PFAS have been used for decades as aqueous film-forming foams for fire suppression, surface coatings in textiles and apparel, and for food packaging, which has led to extensive groundwater contamination.[1−3] Unfortunately, PFAS are resistant to most remediation technologies, even those that work for other halogenated pollutants.[4] Among the most common PFAS found in the environment is perfluorooctanoic acid (PFOA). In response to its prevalence and toxicity, subsequent replacements for PFOA included heptafluoropropylene oxide dimer acid (HFPO−DA, $C_6HF_{11}O_3$); the ammonium salt of HFPO−DA is commercially are known as GenX.[5,6] At present, the most effective means of eliminating PFAS is via thermal destruction, or incineration.[7−9]

The removal and ultimate destruction of PFAS has led to recent interest in using computational chemistry to describe the reaction pathways for thermal treatment of PFAS. Several research groups have developed detailed chemical kinetic mechanisms for PFOA and other PFAS. Khan et al. used perfluoromethanesulfonic acid and perfluoroethanesulfonic acid as surrogates for perfluorooctanesulfonic acid, using various density functional and compound methods.[10] Altarawneh et al. used perfluoropentanoic acid (PFPA, $C_4F_9C(O)$-OH) as a surrogate for PFOA and used density functional theory to obtain the energetics.[11] Similarly, Weber et al. used perfluorobutanoic acid (PFBA, $C_3F_7C(O)OH$) as a surrogate for PFOA,[12] but they used the more accurate G4 method.[13] In a separate study, Khan et al. provided a systematic study for the decomposition of perfluorocarboxylic acids, up to and including PFOA.[14] Lorpaiboon and Ho computed bond dissociation energies (BDE) for a broad range of PFAS containing up to 26 heavy atoms (e.g., PFPA), based upon DLPNO−CCSD(T) single-point energies from DFT geometries.[15] Raza et al. computed energetic properties using B3LYP + D3BJ/6-311+G(2d,2p), and then used various machine learning approaches to predict defluorination.[16] As with PFOA and other perfluoro carboxylic acids, several groups have investigated the thermal destruction of GenX using computational chemistry (following literature convention, we will use "GenX" to refer to HFPO−DA, rather than just the

Received: October 25, 2023
Revised: January 5, 2024
Accepted: January 8, 2024
Published: February 27, 2024



ammonium salt, as well). Adi and Altarawneh created a detailed kinetic model, based upon M06-2X/6-311+G(d,p) calculations.[17] Alinezhad et al. computed BDE using M06-2X/Def2-TZVP.[18] Similarly, Paultre et al. computed BDEs, but they included coupled cluster single-point calculations as well, CCSD(T)/6-311+G(d,p)//ωB97xD/6-311+G(d,p).[19] Blotevogel et al. computed BDE and unimolecular decomposition rates using (DLPNO)-CCSD(T)/aug-cc-pVTZ//ωB97X-D/6-311+G(2d,2p).[20] Additionally, Ding et al. computed decomposition kinetics for solvated GenX using M06-2X/Def2-SVP//B3LYP/6-31G(d) with the SMD solvent model.[21]

In a detailed chemical kinetic mechanism, the enthalpy of a chemical species is typically represented by the standard enthalpy of formation $\Delta_f H°$, or the change in enthalpy required to form the substance from its constituent elements in their reference state. Unfortunately, the $\Delta_f H^i$ (the superscript ° is dropped for clarity) is not a quantity that is measured directly; rather, it is obtained or inferred from other measurements (e.g., calorimetry), which require references species. Moreover, these measurements have never been performed for the vast majority of species i − and in the case of highly reactive intermediates, such as most radicals, they are unlikely ever to be performed. Consequently, the $\Delta_f H^i$ must be computed, either through an estimation method (such as group additivity) or using computational quantum chemistry (QM). For smaller molecules (here defined to be approximately no more than six non-hydrogen atoms), various compound methods have been developed−such as HEAT,[22] focal point,[23] W4,[24] and ANL0[25]—that can provide enthalpies of formation with uncertainty below 1−2 kJ/mol (here we consider uncertainty to be defined by the 95% confidence interval, which is approximately twice the standard deviation, σ, for a normal distribution). For larger molecules, however, these methods are at best impractical, and more computationally affordable methods are employed. These methods aspire to provide enthalpies of formation that are close to chemical accuracy (typically defined to be ±1 kcal/mol).[26] For example, Farina et al. used the compound method G4 to compute the enthalpy of formation for a large number of fluorinated compounds,[27] and Paulechka and Kazakov developed their own compound method to compute the enthalpy of formation and develop their own reference library.[28]

Many of the previously cited approaches compute the enthalpy of formation by first computing the theoretical atomization energy (TAE). The disadvantage to using the TAE is the profound difference in electronic structure between the target species and the isolated atoms. Alternatively, one can formulate a hypothetical reaction in which the target PFAS is formed from a set of molecular reactants that contain all the required elements. The reaction enthalpy at 0 K, $\Delta H_{rxn}(0K)$, is computed using an electronic structure method

$$\Delta H_{rxn}^{QM}(0K) = E_0^{PFAS} + \sum_{i \neq PFAS} \nu_i E_0^i \quad (1)$$

where E_0^i is the sum of electronic and zero-point energy for species i obtained by the QM calculations, and ν_i is the stoichiometric coefficient. The enthalpy of formation of the target PFAS at 0 K is then given by

$$\Delta_f H^{PFAS}(0K) = \Delta H_{rxn}^{QM}(0K) - \sum_{i \neq PFAS} \nu_i \Delta_f H^i(0K) \quad (2)$$

For some gas-phase molecules, the enthalpies of formation are known with astonishing accuracy. For example, many species in the active thermochemical tables (ATcT)[29,30] are known with sub kJ mol^{-1} accuracy at 95% confidence limits. When the ATcT is used to provide the reference values in eq 2, the uncertainty in $\Delta_f H^{PFAS}$ is dominated by the error in the computed reaction enthalpy, ΔH_{rxn}^{QM}. Unfortunately, methods like G4 are limited in the size of PFAS that can be considered (e.g., ref 27 only considered up to 12 heavy atoms in their study, $CF_3CF_2OCF_3$, and ref 12 used $C_3F_7C(O)OH$, PFBA, PFBA, as a surrogate for PFOA). Some compromise must be made in the selection of electronic structure method for PFOA and similarly sized PFAS.

This situation can be improved, however, by a more thoughtful choice of reference species. By selecting reactants that more closely resemble the product, the electronic structure calculations can benefit from error cancellation in ΔH_{rxn}^{QM}. Pople and co-workers introduced the concept of isodesmic reactions, which are designed to conserve bond type.[31,32] These ideas have been generalized into more elaborate methods that conserve various aspects of the molecular environment, as described by the homodesmotic reaction classifications of Wheeler et al.[33] Raghavachari and co-workers developed the connectivity-based hierarchy (CBH) as a systematic method for obtaining homodesmotic reactions.[34−36] The CBH approach provides a simple and efficient recipe for improving the reference reaction. The approach classifies reactions in terms of a series of levels (or rungs). The lowest rung, CBH-0, is an isogyric reaction in which the heavy (i.e., non-hydrogen) atoms are fully hydrogenated. The next level, CBH-1, is an isodesmic reaction that conserves bonds; CBH-2 conserves the immediate environment of each atom, and so forth. The CBH method alternates between atom-centered and bond-centered conservation, and each new level expands the size of the molecular fragments, thereby ensuring increasing structural similarity to the target species. One particular advantage of the CBH method is that it can be automated.[37,38] The CBH method has been adopted by many groups in the combustion community for hydrocarbon enthalpies of formation,[39−43] and was recently extended to include adsorbates on metals as well.[44]

In the application of the CBH method, the availability of reference data is likely to be the primary limitation. For example, if we consider PFOA, it is possible to formulate CBH levels all the way up to CBH-8

$$C_7F_{15}CH_3 + CH_3(CF_2)_6C(O)OH - CH_3(CF_2)_6CH_3 \xrightarrow{CBH\text{-}8} C_7F_{15}C(O)OH$$

However, the enthalpies of formation for the three species on the left-hand side are unknown. Indeed, the ATcT only contains reference species for the first two levels

$$8CH_4 + 15HF + 2H_2O - 25H_2 \xrightarrow{CBH\text{-}0} C_7F_{15}C(O)OH \quad (PFOA\text{-}0)$$
(PFOA - 0)

$$15CH_3F + 7C_2H_6 + CH_2O + CH_3OH - 23CH_4 \xrightarrow{CBH\text{-}1} C_7F_{15}C(O)OH \quad (PFOA\text{-}1)$$
(PFOA - 1)

Although the error cancellation of CBH-1 is expected to be much better than CBH-0, the reference species at CBH-1 are still quite dissimilar from a perfluorocarboxylic acid (e.g.,

https://doi.org/10.1021/acsphyschemau.3c00056
ACS Phys. Chem Au 2024, 4, 247−258

neither the carboxlic acid nor the $-CF_2-$ functional group is represented). The lack of independent reference data presents a challenge for the CBH approach: how can one take advantage of advanced error cancellation schema when one is limited by the availability of reference data?

In the present work, we propose a solution to this problem by creating a hierarchy of compound electronic structure methods that will be used to compute the missing reference data. As noted above, it is the size of the target PFAS that limits the selection of ΔH_{rxn}^{QM} in eq 2. Accordingly, the size of the reference species will determine which method is available. We combine ANL0 with ATcT to compute the enthalpy of formation for 14 different per- and polyfluoro species, include alkanes, alcohols, ethers, and carboxylic acids. These species were chosen so as to create a library of "building blocks" that contain all the functional groups needed for CBH-2 and CBH-3 for PFOA and other PFAS. The enthalpy of formation for these species are obtained using isodesmic reactions, CBH-1. This particular combination is effectively identical to the "CBH-ANL" approach introduced by Elliott et al., who applied it with great success for problems related to hydrocarbon oxidation.[43] Once the high-accuracy ANL0 library has been determined, an intermediate-accuracy method—based upon coupled cluster theory with explicit correlations—is used to create a complementary library of even larger building blocks. This intermediate set contains the functional groups needed for CBH-4 and CBH-5. Finally, the enthalpy of reaction for PFOA and other large PFAS is computed using a third coupled cluster method, one with local correlations. As will be demonstrated below, this approach of combining a hierarchy of homodesmotic reactions with a hierarchy of electronic structure methods enables the calculation of enthalpies of formation with sub chemical acccuracy, even for species as large as PFOA. The present work focuses exclusively on closed-shell spaces; future work will expand the method to include radicals.

2. METHODS

2.1. Electronic Structure Theory

Multiple different methods are used throughout this work. For the smallest molecules, the compound method ANL0 was used.[25] In ref 45, ANL0 was extended to include fluorine atoms. ANL0 begins with geometry relaxation and normal-mode analysis using UCCSD(T)/cc-pVTZ; it is followed by a series of single-point energy calculations to correct for basis set extrapolation, higher order excitation, core—valence interactions, relativitistic effects, and anharmonicity in the zero-point energy. For a more detailed discussion, the reader is referred to ref 25. The ANL0 method is expected to be accurate to 2σ = 2 kJ/mol, or well within chemical accuracy, even when using isogyric reactions (CBH-0).[25,45] Unfortunately, this accuracy comes at a formidable computational cost, and it typically is limited to species with fewer than 10 non-hydrogen atoms (indeed, if the species belongs to the C_1 point group, then the maximum number of heavy atoms is fewer, e.g. six). The list of species computed using ANL0 (in addition to those obtained in ref 45) are listed in Table 2. Using the ATcT as the reference library, the ANL0 enthalpies of formation were computed at the CBH-1 (and occasionally CBH-2) level.

For the intermediate level, geometry optimization and normal-mode analysis were performed using the double-hybrid functional B2PLYPD3/cc-pVTZ,[46−48] followed by single-point calculations using coupled cluster theory with explicit correlations, UCCSD(T)-F12b/cc-pVTZ-f12.[49,50] The largest species obtained using UCCSD-(T)-F12b/cc-pVTZ-f12//B2PLYPD3/cc-pVTZ (hereafter, "F12" for short) was PFBA, $C_3F_7C(O)OH$. The complete list of species computed using F12 are listed in Table 3. Using ATcT and CBH-ANL as reference libraries, the F12 enthalpies of formation were computed at the CBH-3 (and occasionally CBH-4) level. In ref 38, the UCCSD(T)-F12b/cc-pVTZ-f12//B2PLYPD3/cc-pVTZ enthalpies were compared to CBH-ANL results for hydrocarbons, and it was found that the mean absolute deviation dropped from 3.9 kJ/mol at CBH-0 to 0.1 kJ/mol at CBH-2.

The largest species, C_5-species and up, necessitated a different approach. Two density functionals were used, ωB97xd/cc-pVTZ[51] and M06-2X/cc-pVTZ,[52] for the optimization and frequency calculation. Single-point energies were obtained using coupled cluster theory with local correlations, DLPNO−CCSD(T)/aug-cc-pVQZ.[53,54] These two methods—"$DLPNO_\omega$" and "$DLPNO_M$", respectively—are used to compute the ΔH_{rxn}^{QM} in eq 2. The highest CBH level is used where all of the reactants—the $\sum_{i\neq PFAS}\nu_i\Delta_f H^i(0K)$ in 2 can be obtained from either ATcT, ANL0, or F12. The list of target PFAS computed using $DLPNO_\omega$ and $DLPNO_M$ are provided in Table 4.

For all species, the lowest energy conformer was first selected using M06-2X/cc-pVTZ. Once the global minimum energy conformer was obtained, torsional scans were performed for appropriate dihedral angles in 10° increments using M06-2X/cc-pVTZ. These scans capture the contribution of conformers to the partition function. The corresponding partition function for the isolated torsional motion was computed via summation over the energy levels for the corresponding 1D Schrödinger equation; all other internal degrees of freedom were

Table 1. Enthalpies of Formation at 0 K, $\Delta_f H^\circ_{ATcT}$, for the Reference Species, Taken from the ATcT (in kJ/mol)

name	$\Delta_f H^\circ_{ATcT}$	uncertainty
H_2	0.000	(exact)
HF	−272.680	±0.019
CH_4	−66.551	±0.048
CH_3F	−227.48	±0.23
CH_2F_2	−443.38	±0.33
CHF_3	−689.34	±0.38
CF_4	−927.77	±0.23
C_2H_6	−68.39	±0.12
C_2F_6	−1334.11	±0.82
C_3H_8	−82.72	±0.16
H_2O	−238.898	±0.025
CH_2O	−105.38	±0.096
CF_2O	−603.36	±0.40
CH_3OH	−190.04	±0.15
CH_3OCH_3	−166.51	±0.37
$CH_3C(O)OH$	−418.53	±0.36

Table 2. Enthalpies of Formation at 0 and 298 K for the First Tier of Computed Species, as Determined by the ANL0 Compound Method (in kJ/mol)

name	$\Delta_f H^\circ_{ANL0,0K}$	$\Delta_f H^\circ_{ANL0,298K}$	level
CH_3CF_3	−739.7 ± 1.5	−752.5 ± 2.0	CBH-2
$CH_3CF_2CH_3$	−533.4 ± 1.5	−553.1 ± 2.0	CBH-1.5-avg
$CH_3CF_2CF_3$	−1154.4 ± 2.0	−1169.9 ± 2.5	CBH-1.5-HF
C_3F_8	−1743.8 ± 2.0	−1754.7 ± 2.5	CBH-1.5-avg
$CH_3(CF_2)_2CH_3$	−969.3 ± 2.0	−991.4 ± 2.5	CBH-1.5-avg
CF_3OH	−901.6 ± 1.5	−902.0 ± 2.0	CBH-1-avg
CH_3CF_2OH	−712.3 ± 1.5	−727.8 ± 2.0	CBH-1.5-HF
C_2F_5OH	−1315.4 ± 1.5	−1325.7 ± 2.0	CBH-1.5-avg
$CH_3CF(OH)CH_3$	−496.9 ± 2.0	−519.7 ± 2.5	CBH-1.5-HF
CH_3OCF_3	−874.7 ± 1.0	−889.0 ± 1.5	CBH-1.5-HF
CF_3OCF_3	−1531.3 ± 1.5	−1542.7 ± 2.0	CBH-1.5-avg
$CH_3CF_2OCH_3$	−682.8 ± 1.5	−704.5 ± 2.0	CBH-1.5-avg
$CF_3C(O)OH$	−1017.3 ± 2.0	−1028.1 ± 2.5	CBH-1.5-avg
$CH_3CF_2C(O)OH$	−833.7 ± 2.0	−850.0 ± 2.5	CBH-1.5-avg

https://doi.org/10.1021/acsphyschemau.3c00056
ACS Phys. Chem Au 2024, 4, 247−258

Table 3. Enthalpies of Formation at 0 and 298 K for the Second Tier of Computed Species, as Determined by UCCSD(T)-F12b/cc-pVTZ-f12//B2PLYPD3/cc-pVTZ (in kJ/mol)

name	$\Delta_f H_{F12,0K}^\circ$	$\Delta_f H_{F12,298K}^\circ$	level
$C_3F_7CH_3$	−1564.4 ± 3.0	−1581.6 ± 3.5	CBH-2.5-avg
C_4F_{10}	−2150.6 ± 3.0	−2163.8 ± 3.5	CBH-3-avg
$CH_3(CF_2)_3CH_3$	−1378.0 ± 3.0	−1401.2 ± 3.5	CBH-3-avg
$C_4F_9CH_3$	−1973.7 ± 3.0	−1993.9 ± 3.5	CBH-4-HF
$CH_3(CF_2)_4CH_3$	−1792.8 ± 4.0	−1817.4 ± 4.5	CBH-3-avg
$CH_3(CF_2)_2OH$	−1133.7 ± 3.0	−1151.2 ± 3.5	CBH-2-avg
$CH_3CF(OH)CF_3$	−1133.9 ± 2.0	−1152.7 ± 2.5	CBH-2-avg
C_3F_7OH	−1725.0 ± 2.0	−1737.5 ± 2.5	CBH-2-avg
$CF_3CF(OH)CF_3$	−1729.7 ± 3.0	−1744.0 ± 3.5	CBH-2.5-avg
$CF_3CF_2OCH_3$	−1289.3 ± 2.0	−1305.6 ± 2.5	CBH-3-HF
$CH_3CF_2OCF_3$	−1350.8 ± 2.0	−1367.6 ± 2.5	CBH-2.5-avg
$CF_3CF_2OCF_3$	−1944.6 ± 2.0	−1961.7 ± 2.5	CBH-2.5-avg
$CH_3CF(CH_3)OCH_3$	−466.0 ± 2.0	−494.2 ± 2.5	CBH-2-avg
$CH_3CF(CH_3)OCF_3$	−1153.6 ± 3.0	−1177.3 ± 3.5	CBH-2.5-avg
$CF_3CF(CF_3)OCH_3$	−1696.8 ± 3.0	−1715.0 ± 3.5	CBH-3-HF
$CF_3CF(CF_3)OCF_3$	−2353.1 ± 3.0	−2366.6 ± 3.5	CBH-2.5-avg
$C_2F_5C(O)OH$	−1432.9 ± 2.0	−1445.9 ± 2.5	CBH-2.5-avg
$CH_3CF_2CF_2C(O)OH$	−1255.2 ± 4.0	−1273.8 ± 4.5	CBH-3-avg
$CH_3CF(OH)C(O)OH$	−822.0 ± 4.0	−842.3 ± 4.5	CBH-2-avg
$CF_3CF(OH)C(O)OH$	−1430.5 ± 3.0	−1447.1 ± 3.5	CBH-2-avg
$C_3F_7C(O)OH$	−1845.4 ± 3.0	−1859.7 ± 3.5	CBH-2-avg

assumed to behave as harmonic oscillators. The thermophysical properties—including the enthalpy increment, entropy, and temperature-dependent heat capacity at constant pressure—were computed and then converted into NASA7 polynomials. The calculation of thermophysical properties was performed using subroutines from AutoMech,[55−57] which is part of the Pasr family of computational kinetics software from Argonne National Laboratory. All density functional theory calculations were performed using Gaussian09.[58] All wave function methods were performed using Molpro,[59] with the exception of the DLPNO calculations, which were performed in Orca.[60]

2.2. CF_4-Based CBH Schema

In the original CBH methodology, the reference species for the fluorine atoms is HF at CBH-0, and it is CH_3F at CBH-1. In ref 45, we proposed using CF_4 instead of HF for the isogyric reaction in the ANL0 method. In the present work, we build off this idea and consider an alternative CBH schema in which CF_4 replaces HF at the CBH-0 level. The implicit expectation is that a reaction sequence that begins with a fully fluorinated carbon atom at the lowest rung will outperform a reaction sequence that begins with HF. We refer to the original method as "CBH-*n*-HF", and the new approach as "CBH-*n*-CF_4", where *n* is the CBH level. As will be demonstrated below, there are advantages and disadvantages to both approaches. The CBH-*n*-CF_4 does indeed converge more quickly; however, the CBH-*n*-HF method is able to go to higher rungs.

Because of the introduction of CBH-*n*-CF_4, it is now possible to have two different reactions at the same CBH-*n*. This raises another issue, which is how best to average them. A consequence of the CBH methodology is that the enthalpy of reaction will limit toward zero as $n \to \infty$. Intuitively, this observation makes sense: the greater the structural similarity between the reactants and the product, the less energy is required/liberated in the reaction to form the product. Consequently, one could, in principle, use the enthalpy of reaction as a metric for convergence: at a given level *n*, the CBH-*n*-X with the lower $|\Delta H_{rxn}|$ is likely to be more accurate. Note, we are not suggesting that the enthalpy of reaction, by itself, is necessarily a reliable metric for error cancellation; rather, we restrict our focus to comparisons of different homodesmotic reactions. Accordingly, we propose a weighting scheme that is biased toward the reaction with the lower (absolute) enthalpy of reaction

$$w^{\mathrm{X}} = \frac{|\Delta H_{rxn}^{\text{CBH-}n\text{-X}}|^{-1}}{\sum_k |\Delta H_{rxn}^{\text{CBH-}n\text{-}k}|^{-1}} \tag{3}$$

An obvious objection to eq 3 is the issue of a singularity at $\Delta H_{rxn} = 0$. In reality, however, we are unlikely to reach a rung at which this singularity is possible, and the enthalpies of reaction will be $|\Delta H_{rxn}| \gg 0$.

2.3. AutoCBH

AutoCBH is a Python package developed to automate the tedious nature of (i) deriving CBH schema and (ii) computing enthalpies of formation by optimally laddering different methods of electronic structure. For the former, a user can obtain the entire designated CBH schema (e.g., CBH-*n*-HF or CBH-*n*-CF_4) for any target molecule given its SMILES string. The CBH method derives its reference species by breaking the target molecule into differently sized chemical fragments depending on the rung *n*. This operation becomes intuitive once the representation of the molecules is converted to simple chemical graphs where nodes are atoms and bonds are edges. The fragments are thus determined by selecting *n*-hop neighbors of atoms or bonds and then saturating them with either hydrogen or fluorine to return valid molecules (i.e., satisfy expected valence).

After a CBH scheme is defined, he criteria for choosing the optimal reference reaction consists of using the best available electronic structure method while also ensuring that the reaction is a linearly independent combination of reference species (see Section 3.1). The user provides (i) their database containing any QM-derived energies for all species and (ii) a rank ordering of those QM method levels of theory. For a given target molecule, a series of operations are run on the derived CBH schema including: validation of the reference species, selection of the best possible QM method, and decomposition of reactions to check for linear independence. If both HF and CF_4 CBH schema are utilized, or multiple QM methods reside at the same level of theory (e.g., $DLPNO_\omega$ and $DLPNO_M$), the enthalpies of formation are weighted and combined using eq 3. This process

Table 4. Enthalpies of Formation at 0 and 298 K for Larger PFAS, as Determined by UCCSD(T)-DLPNO/aug-cc-pVQZ//ωb97xd/cc-pVTZ and UCCSD(T)-DLPNO/aug-cc-pVQZ//M062X/cc-pVTZ (in kJ/mol)

name	$\Delta_f H_{DLPNO\omega,0K}^\circ$	$\Delta_f H_{DLPNO\omega,298K}^\circ$	$\Delta_f H_{DLPNOM,0K}^\circ$	level
C_5F_{12}	−2563.2 ± 4.0	−2577.6 ± 4.5	−2563.0 ± 4.0	CBH-4.5-avg
C_6F_{14}	−2973.8 ± 4.0	−2988.8 ± 4.5	−2974.0 ± 4.0	CBH-4.5-avg
C_7F_{16}	−3384.8 ± 4.0	−3401.3 ± 4.5	−3384.6 ± 4.0	CBH-4.5-avg
C_8F_{18}	−3795.8 ± 4.0	−3401.3 ± 4.5	−3795.5 ± 4.0	CBH-4.5-avg
$C_4F_9C(O)OH$	−2256.5 ± 4.0	−2271.5 ± 4.5	−2256.5 ± 4.0	CBH-4.5-avg
$C_5F_{11}C(O)OH$	−2667.4 ± 4.0	−2682.7 ± 4.5	−2667.4 ± 4.0	CBH-4.5-avg
$C_6F_{13}C(O)OH$	−3078.3 ± 4.0	−3094.6 ± 4.5	−3078.2 ± 4.0	CBH-4.5-avg
$C_7F_{15}C(O)OH$	−3489.1 ± 4.0	−3507.4 ± 4.5	−3488.8 ± 4.0	CBH-4.5-avg
FRD-903 (GenX)	−2883.5 ± 4.0	−2901.9 ± 4.5	−2886.1 ± 4.0	CBH-3.5-avg

https://doi.org/10.1021/acsphyschemau.3c00056
ACS Phys. Chem Au 2024, 4, 247−258

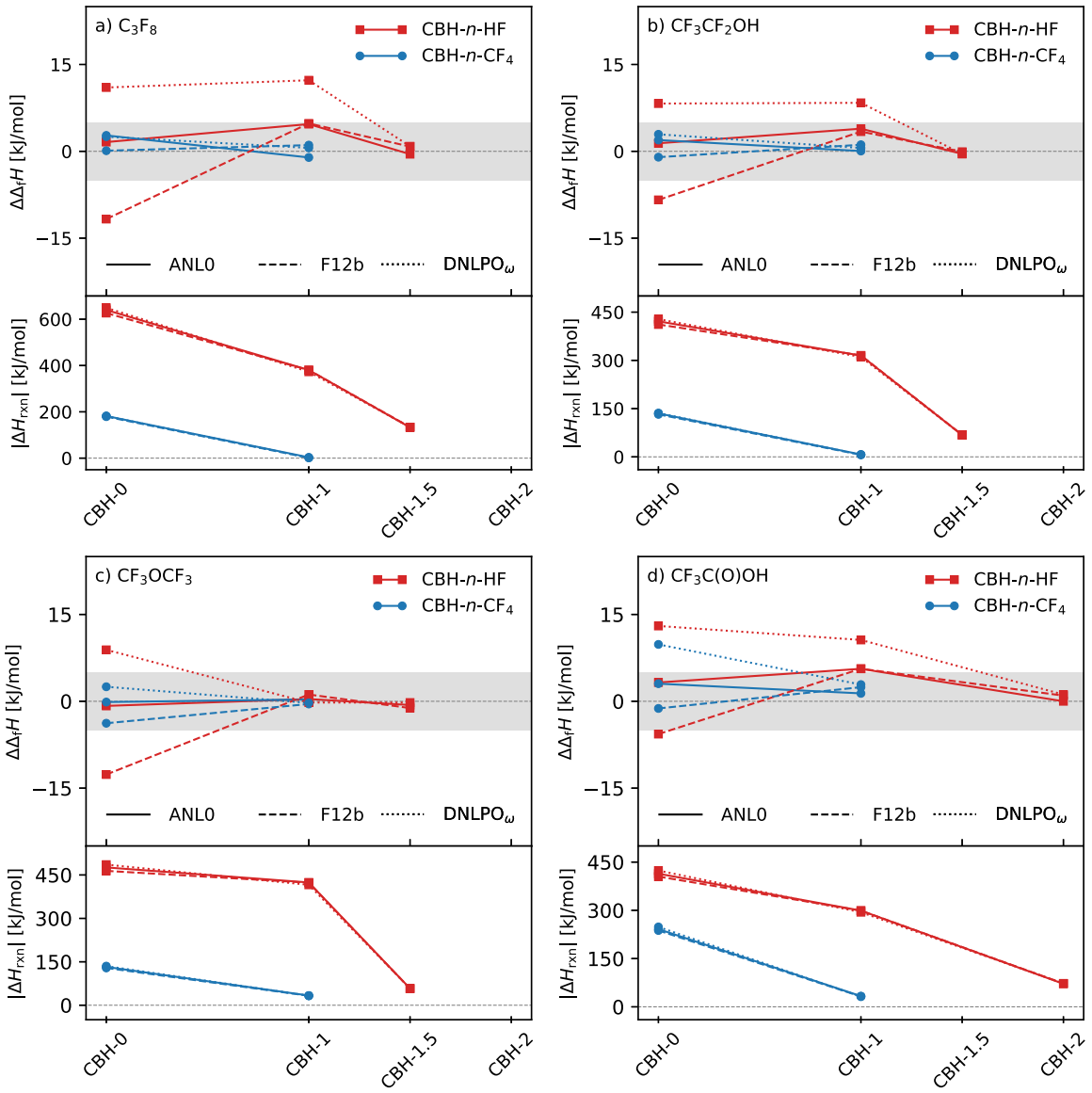


Figure 1. Comparison of the three different computational methods for the enthalpies of formation of (a) C_3F_8, (b) CF_3CF_2OH, (c) CF_3OCF_3, and (d) $CF_3C(O)OH$. In each subplot, the top pane is difference in the heat of formation from the final value, and the lower plot is the absolute value of the heat of reaction. The shaded region denotes ±4 kJ/mol.

iterates through all species, from small to large, because smaller species use more accurate QM methods and larger molecules can reference the smaller ones. AutoCBH finally reports the calculated enthalpies of formation and the attributed CBH reaction used for each molecule.

A vast and complicated network of molecules is the result of the hierarchical laddering of the CBH method. In this network, certain small molecules can be disproportionately referenced by larger ones, directly or indirectly. Consequently, the error of those small molecules will have a significant impact on the accuracy of many of its "descendent" large molecules. Therefore, it is important to focus efforts on improving the accuracy of values for those molecules. To help identify these critical molecules, AutoCBH also provides helper modules to visualize these thermochemical networks and quantify the impact of uncertainty propagation.

2.4. Reference Data

The ANL0 approach (as well as the F12) still require reference data. Table 1 lists 16 species that were taken from the Active Thermochemical Database.

These reference species contain all the bond types needed to form isodesmic (CBH-1) reactions. Unfortunately, the ATcT is still missing a few key species that would enable CBH-2 for PFAS.

3. RESULTS AND DISCUSSION

3.1. Construction of the Reference Libraries Using ANL0 and F12

Table 2 lists the 14 per- and polyfluoro species computed using the ANL0 method; Table 3 lists the 21 species computed using the F12 method. Also contained in these tables are the enthalpies of formation at 298.15 K, $\Delta_f H_{298\ K}^{\circ}$. Both tables list the CBH rung or average, given the availability of reference data. In terms of selecting the CBH average, we used the following rubric. If the highest available rung n is the same for both CBH-n-HF and CBH-n-CF_4, then the results are averaged according to eq 3. However, in many cases, the HF reference expression will permit a higher rung, CBH-n+1-HF vs CBH-n-CF_4. In those cases, we compare the corresponding enthalpies

https://doi.org/10.1021/acsphyschemau.3c00056
ACS Phys. Chem Au 2024, 4, 247−258

of reaction. If $|\Delta H_{rxn}^{CBH\text{-}n+1\text{-}HF}| < |\Delta H_{rxn}^{CBH\text{-}n\text{-}CF_4}|$ then we assume that the higher rung is superior and select CBH-n+1-HF accordingly. If, on the other hand, $|\Delta H_{rxn}^{CBH\text{-}n\text{-}CF_4}| < |\Delta H_{rxn}^{CBH\text{-}n+1\text{-}HF}|$ then we use the weighted average of the two results.

For many species in the ANL0 set, it was possible to find alternative reactions that were arguably superior to those proposed by AutoCBH. For example, for 1,1,1-trifluoroethane, CH_3CF_3, the CBH algorithm stopped at CBH-1-HF and CBH-1-CF_4, but a CBH-2 reaction is possible

$$C_2H_6 + 3CH_3F - 3CH_4 \xrightarrow{\text{CBH-1-HF}} CH_3CF_3, \quad \Delta H_{rxn} = -186.3[\text{kJ/mol}]$$

$$C_2F_6 + 3CHF_3 - 3CF_4 \xrightarrow{\text{CBH-1-CF}_4} CH_3CF_3, \quad \Delta H_{rxn} = -107.4[\text{kJ/mol}]$$

$$\frac{1}{2}C_2H_6 + \frac{1}{2}C_2F_6 \xrightarrow{\text{CBH-2}} CH_3CF_3, \quad \Delta H_{rxn} = -38.49[\text{kJ/mol}]$$

In other cases, we found reactions that were between CBH-1 and CBH-2 in the sense that they conserved some, but not all, of the bond-centered environments of the CBH-2 level. For example, $CH_3CF_2CH_3$ is required for CBH-2-HF calculations for larger PFAS. Unfortunately, this species is not in the ATcT. Computing the enthalpy of formation for $CH_3CF_2CH_3$ at CBH-2-HF would require both $CH_3{-}C$ and $C{-}CF_2{-}C$ subunits, but the latter cannot be found in the ATcT (indeed, $CH_3CF_2CH_3$ is itself the simplest saturated species to contain that subunit). At the CBH-1 level, AutoCBH provides

$$2CH_3F + 2C_2H_6 - 3CH_4 \xrightarrow{\text{CBH-1-HF}} CH_3CF_2CH_3, \quad \Delta H_{rxn} = -138.1[\text{kJ/mol}]$$

$$6CHF_3 + 2C_2F_6 - 7CF_4 \xrightarrow{\text{CBH-1-CF}_4} CH_3CF_2CH_3, \quad \Delta H_{rxn} = -224.1[\text{kJ/mol}]$$

Alternatively, we consider two different reactions

$$C_3H_8 - CH_4 + CH_2F_2 \xrightarrow{\text{CBH-1.5-HF}} CH_3CF_2CH_3, \quad \Delta H_{rxn} = -73.8[\text{kJ/mol}]$$

$$C_3H_8 + CF_4 - CH_2F_2 \xrightarrow{\text{CBH-1.5-CF}_4} CH_3CF_2CH_3, \quad \Delta H_{rxn} = 33.7[\text{kJ/mol}]$$

CBH-1.5-HF yields $\Delta_f H_{0K}^{CH_3CF_2CH_3} = -533.3$ kJ/mol, and CBH-1.5-CF_4 yields $\Delta_f H_{0K}^{CH_3CF_2CH_3} = -533.4$ kJ/mol. The advantage of these two equations, in addition to having a substantially lower enthalpies of reaction, is that they contain propane. We label these equations "CBH-1.5", because they are effectively halfway between CBH-1 and CBH-2. We generalize this observation and use the n + 1/2 description when a reaction has half the reference species needed for n + 1 level and is expected to be slightly more accurate or reliable than n. These exceptions are generally only relevant for smaller molecules, however, and are not used for the larger PFAS in this study.

For C_3F_8, the highest available rung hypothetically would be

$$2C_2F_5CH_3 - CH_3CF_2CH_3 \xrightarrow{\text{CBH-3-HF}} C_3F_8, \quad \Delta H_{rxn} = 32.6[\text{kJ/mol}]$$

Strictly speaking, though, it is not possible to compute perfluoropropane at the CBH-3 level using ANL0. As documented above, $CH_3CF_2CH_3$ is limited to CBH-1-HF or CBH-1.5-HF (as is $C_2F_5CH_3$, for precisely the same reason). Since both species also were obtained using ANL0, the final enthalpy of formation for C_3F_8 would in fact collapse back to a linear combination of the same reference species used for $CH_3CF_2CH_3$ and $C_2F_5CH_3$ (i.e., a combination of CH_4, CH_2F_2, CF_4, C_2H_6, C_2F_6, and C_3H_8). Accordingly, it is not possible to construct a linearly independent CBH-2 reaction (let alone CBH-3) from these species. Indeed, as is evident in Table 2, the absence of larger fluorinated compounds from the ATcT limits most of the ANL0 species to no higher than CBH-1.5. Nonetheless, as demonstrated in ref 38, the ANL0 method is capable of providing subchemical accuracy at the CBH-0 level, and the CBH-1 results will be even more reliable.

Figure 1 presents the results for four different perfluorinated species: C_3F_8 (Figure 1a), CF_3CF_2OH (Figure 1b), CF_3OCF_3 (Figure 1c), and $CF_3C(O)OH$ (Figure 1d). These four species were the largest perfluoro-alkane, alcohol, ether, and carboxylic acid, respectively, that could be obtained via ANL0. The red lines are the results obtained from the conventional CBH-n-HF, and the blue lines are from the alternative CBH-n-CF_4 approach. The solid lines are the ANL0 results, the dashed lines are the F12 results, and the dotted lines are the DLPNO$_\omega$ results. For each species, the top pane is the difference in the enthalpy of formation, relative to the final value, here assumed to be CBH-1.5-HF (or CBH-2-HF in the case of $CF_3C(O)$-OH); the lower pane is the enthalpy of reaction. As can be seen in all four subplots, the ANL0 maintains chemical accuracy (the shaded region), even at the CBH-0 level. The F12 approach obtains chemical accuracy for CBH-0-CF_4 and CBH-1-HF. Similarly, the DLPNO$_\omega$ is converged by CBH-1-CF_4 and CBH-1.5-HF.

The uncertainty values in Tables 2 and 3 are estimated based upon a combination of the CBH rung and the convergence with respect to lower rungs and are intended to be 2σ values. The uncertainty for $\Delta_f H_{298\ K}^\circ$ is increased by 0.5 kJ/mol to account for additional error in the enthalpy increment. The results in Figure 1 help to illustrate the uncertainty estimates in Table 2. For example, in the case of C_3F_8, the two highest rungs were CBH-1-CF_4 ($\Delta_f H_{0K}^\circ = -1743.9$) and CBH-1.5-HF ($\Delta_f H_{0K}^\circ = -1742.9$), with a standard deviation of 0.53 kJ/mol (the final weighted average is $\Delta_f H_{0K}^\circ = -1743.8$ kJ/mol). Similarly, for CF_3OH, the highest rungs were CBH-1-HF ($\Delta_f H_{0K}^\circ = -900.9$) and CBH-1-$CF_4$ ($\Delta_f H_{0K}^\circ = -901.7$), with a standard deviation of 0.40 kJ/mol and a final weighted average of $\Delta_f H_{0K}^\circ = -901.6$ kJ/mol. At this point, we lack sufficient data or justification to provide the uncertainty values with greater precision. Instead, we consider the standard deviation, σ, between the final value and the lower rungs, round it up to the nearest 0.5 kJ/mol, and include the contribution from the ATcT in the final uncertainty estimate for 2σ. Additional details are provided in Section 3.3 below.

There is an important caveat to the results in Figure 1: these results should not be taken to imply that DLPNO$_\omega$ always will be converged by CBH-2. As indicated by the lower pane in each subplot, the convergence depends strongly on the enthalpy of reaction. As the target PFAS increases in size,

https://doi.org/10.1021/acsphyschemau.3c00056
ACS Phys. Chem Au 2024, 4, 247−258

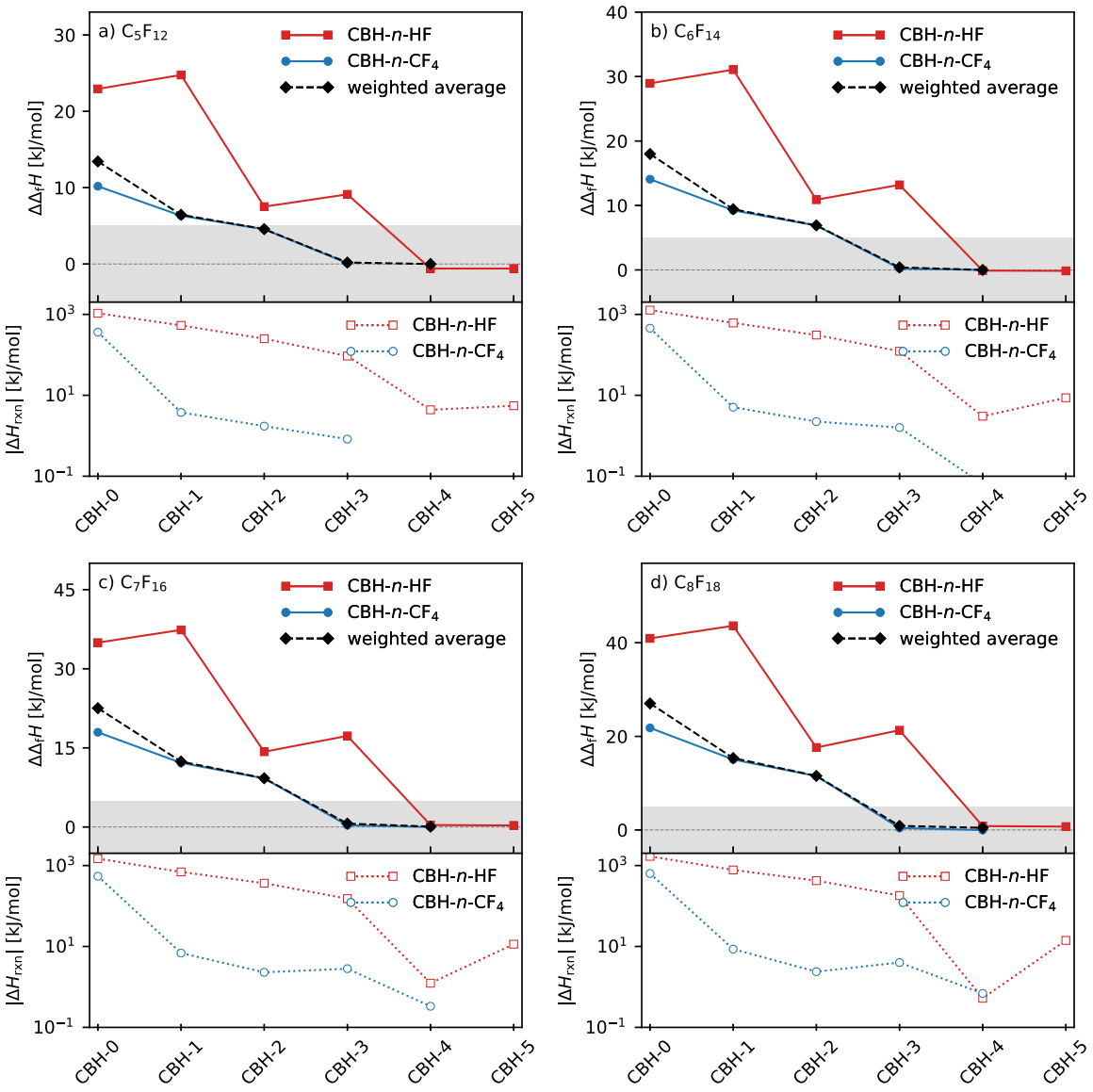


Figure 2. Convergence of the enthalpies of formation of perfluoroalkanes: (a) C_5F_{12}, (b) C_6F_{14}, (c) C_7F_{16}, and (d) C_8F_{18}.

the $|\Delta H_{rxn}|$ at a given CBH-n will increase. In other words, although the results in Figure 1 suggest that DLPNO$_\omega$ has converged for $n = 2$, in fact larger values of n will be necessary for larger PFAS. Instead, we suggest that the CBH enthalpy of formation has converged for a value n when the CBH-$(n-1)$-CF_4, CBH-$(n-1)$-HF, and CBH-n-HF agree to within ±2 kJ/mol.

3.2. Larger PFAS Enthalpies of Formation Obtained via DLPNO

Table 4 lists the enthalpies of formation for the largest perfluoroalkanes and perfluorocarboxlylic acids, as well as for GenX. Both DLPNO$_\omega$ and DLPNO$_M$ results are listed, as well as the highest available CBH rung.

Figure 2 presents the results for the four largest perfluoroalkanes. The final values for perfluoroalkanes are listed in Table 4. All four species were computed using DLPNO$_\omega$ (the results for DLPNO$_M$ were effectively indistinguishable at the higher levels). As with Figure 1, the results in Figure 2 suggest that the CBH-n-CF_4 converge more quickly than CBH-n-HF for a given value of n, but that CBH-n-HF is able to go to a higher values of n. Also shown in Figure 2 are the results obtained using the weighted average from eq 3 (dashed black lines). Given that the CBH-n-CF_4 consistently has a lower enthalpy of reation at each n, it is not surprisingly that the weighted average closely tracks the CBH-n-CF_4 results. For all the alkanes, $|\Delta H_{rxn}^{\text{CBH-4-CF}_4}| < |\Delta H_{rxn}^{\text{CBH-5-HF}}|$, and thus the final value is the weighted average of the two. The standard deviation for CBH-3-CF_4, CBH-4-HF, CBH-5-HF was 0.4 kJ/mol. If the DLPNO$_M$ results are included, the standard deviation increases to 0.7 kJ/mol.

Figure 3 presents the DLPNO$_\omega$ results for the four largest perfluorocarboxylic acids. The final values for carboxylic acids are listed in Table 4. The results are consistent with the perfluoroalkanes in Figure 2, with (i) CBH-n-CF_4 converge more quickly than CBH-n-HF, (ii) the weighted average favoring CBH-n-CF_4, and (iii) CBH-n-HF going to a higher rung. For all the carboxylic acids, $|\Delta H_{rxn}^{\text{CBH-4-CF}_4}| < |\Delta H_{rxn}^{\text{CBH-5-HF}}|$, and thus the final value is the weighted average of the two. The standard deviation for CBH-3-CF_4, CBH-4-HF, CBH-5-HF was 0.3 kJ/mol. If the

https://doi.org/10.1021/acsphyschemau.3c00056
ACS Phys. Chem Au 2024, 4, 247−258

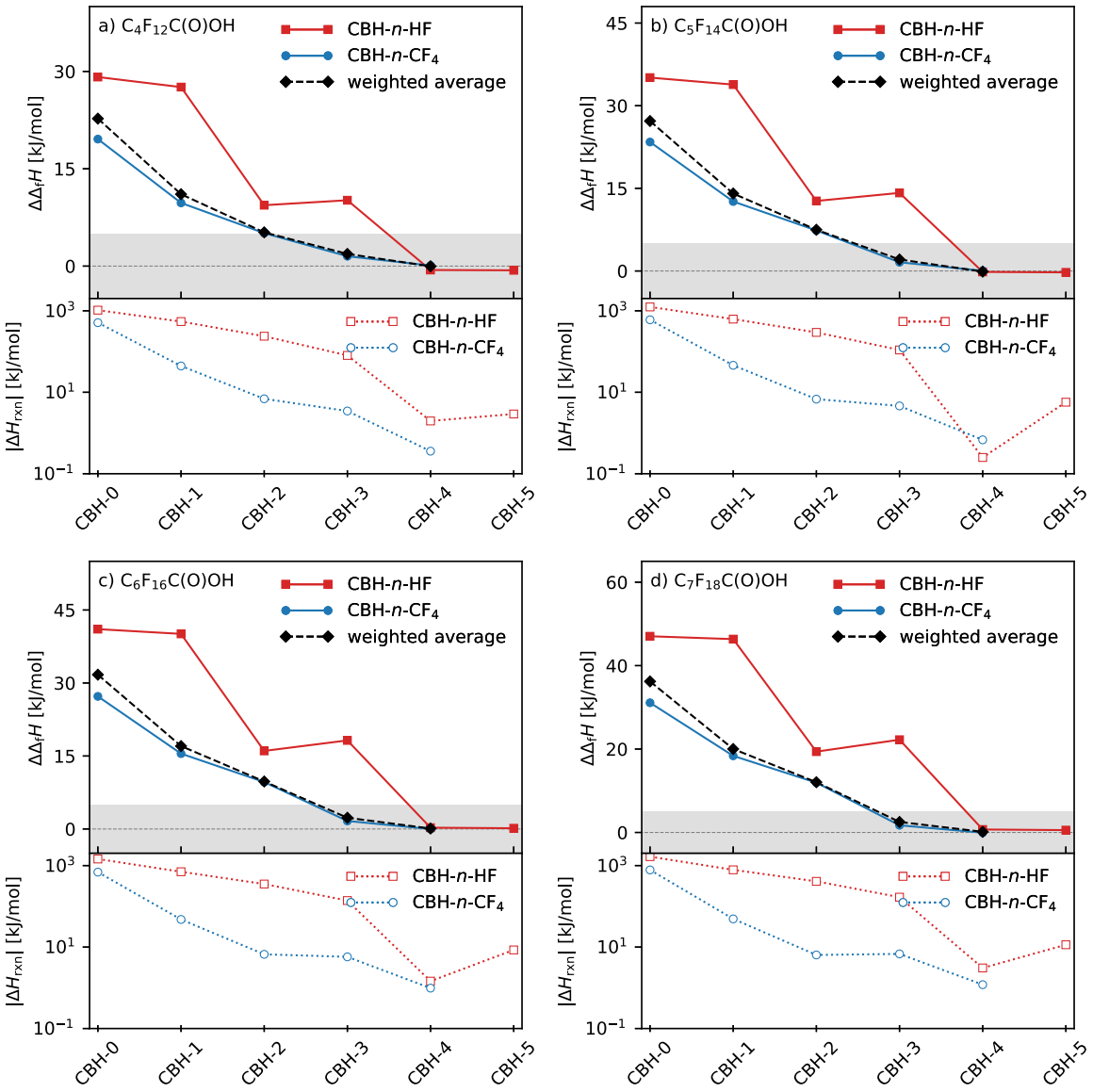


Figure 3. Convergence of the enthalpies of formation of perfluorocarboxylic acids: (a) $C_4F_9C(O)OH$, (b) $C_5F_{11}C(O)OH$, (c) $C_6F_{13}C(O)OH$, and (d) $C_7F_{15}C(O)OH$.

$DLPNO_M$ results are included, the standard deviation increases to 0.6 kJ/mol.

Table 4 contains the computed enthalpy of formation at 0 K for GenX, and Figure 4 presents the convergence study. To the best of our knowledge, these data are the first published values for this species. The more complicated structure of GenX necessitated multiple different functional groups and reference species. Indeed, more than a third of the species in Table 2 and more than half the species in Table 3 were computed just to provide the necessary building blocks for GenX. Despite the additional computational cost, this example highlights the power of the CBH approach in general and the AutoCBH package in particular.

3.3. Uncertainty Quantification for PFOA

The uncertainty in the enthalpy of formation for a given PFAS depends upon the uncertainties in both terms on the right-hand side of eq 2: the enthalpy of reaction, ΔH_{rxn}^{QM}, and the sum of the enthalpies of formation of the reactants, $\sum_i \nu_i \Delta_f H^i$. At lower CBH levels, the hypothetical reaction will consist of a large numbers of small molecules. For example, CBH-0 and CBH-1 for PFOA require 50 and 47 reactants, respectively, [see equations (PFOA-0) PFOA-0 and (PFOA-1) PFOA-1]. The enthalpies of formation for these small molecules are known with exceedingly high accuracy; e.g. the largest uncertainty for CBH-0 is for CH_4, ±0.048 kJ/mol, and for CBH-1 it is CH_3F, ±0.38 kJ/mol (see Table 1). On the other hand, the enthalpy of reaction at these lowest rungs is quite large (1658 kJ/mol and −787 kJ/mol, respectively). Given the lack of structural similarly between the reactants and product at CBH-0-HF, there is no reason to expect any error cancellation, and the situation with CBH-1-HF is only slightly better. Consequently, at the lowest rungs, the uncertainty in the derived enthalpy of formation is entirely dominated by the uncertainty in the computed enthalpy of reaction.

At higher rungs, the number of reactants will decrease, and the size of the reactants will increase. If the smaller molecules are known with high accuracy, a corollary is that larger molecules have greater uncertainty. In other words, we typically expect the uncertainty in the enthalpy of formation to increase with molecular size, and a cursory glance at the

https://doi.org/10.1021/acsphyschemau.3c00056
ACS Phys. Chem Au 2024, 4, 247−258

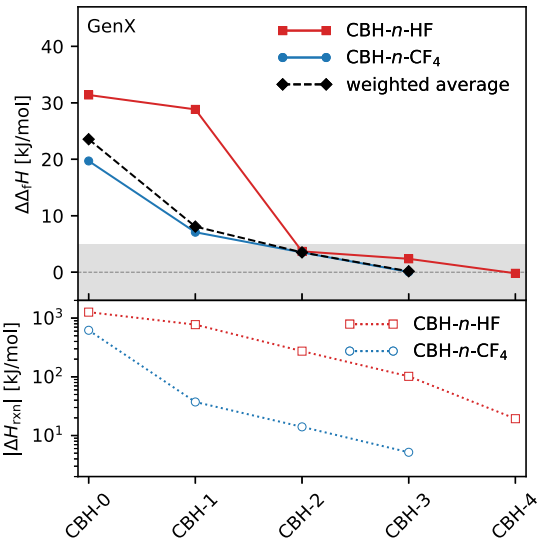


Figure 4. Convergence of the enthalpies of formation of GenX.

species in Table 1 confirms this expectation (e.g., CH_4 vs C_2H_6, CF_4 vs C_2F_6, CH_2O vs $CH_3C(O)OH$, etc.). This rather benign observation presents a challenge to the error cancellation methodology. As we climb the CBH ladder, the uncertainty due to the computed enthalpy of reaction will decrease, but the uncertainty due to the reference species will increase. If the methodology is to serve its desired purpose, we must ensure that the *overall* uncertainty decreases.

To further explore this problem, we consider a more rigorous uncertainty quantification for PFOA. We make the following assumptions. First, we assume that the ATcT uncertainties are normally distributed (with a mean of zero). This assumption is consistent with the statistical intent of the ATcT, in which the stated uncertainties correspond to the 95% confidence intervals. The width of the distribution was set so that 2σ was equal to the uncertainty in Table 2. Furthermore, the 16 species in Table 2 are assumed to be uncorrelated. This assumption is arguably strong; many of the fluorinated species in the ATcT share common sources, and we expect some degree of correlation is possible. Assuming that these species are uncorrelated and independent will overestimate their contribution. As will be demonstrated, the uncertainty in the ATcT is so low that this assumption is of little consequence.

Next, the enthalpy of formation for each species in the reference libraries (Tables 2 and 3) is assigned an uncertainty. This uncertainty is based upon (i) the method that was used (ANL0 vs F12), (ii) the CBH level that was used, and (iii) the standard deviation compared to lower levels, and/or CBH-*n*-HF versus CBH-*n*-CF_4. Whereas the uncertainties in the ATcT are assumed to be normal, the uncertainties in the derived reference species are assumed to be uniform (with a mean of zero). Additionally, whereas the uncertainties in the ATcT are assumed to be uncorrelated, we assume that the uncertainties in the derived reference species are correlated, due to the common method in ΔH_{rxn}^{QM}. In other words, if two reference species were computed using ANL0, then we expect their respective $\delta\Delta H_{rxn}^{ANL0}$ to have the same sign. Thus, all species in Table 2 exhibit some degree of correlation, as do all species in Table 3 and all species in Table 4.

Finally, we assign uncertainties to the enthalpies of reaction for DLPNO$_\omega$. We assume that the uncertainty in $\Delta H_{rxn}^{DLPNO_\omega}(n)$ scales inversely with the CBH level n, corresponding with the assumption of increasing error cancellation. For first six CBH levels, CBH-0 to CBH-5, we assign uncertainties of $\delta\Delta H_{rxn}^{DLPNO_\omega}(n) = \pm[20, 15, 10, 5, 3, 2]$ kJ/mol, respectively. These values, along with the uncertainties in Tables 2 and 3, are intended to represent the 2σ bounds. We readily concede that the assigned uncertainties—whether in Tables 2 and 3 or in the $\delta\Delta H_{rxn}^{DLPNO_\omega}(n)$—are debatable. The underlying assumptions, however—that (i) ANL0 is more accurate than F12, and F12 is more accurate than DLPNO$_\omega$, and (ii) $\delta\Delta H_{rxn}^{DLPNO_\omega}(n) < \delta\Delta H_{rxn}^{DLPNO_\omega}(n-1)$—should not be controversial.

The total number of perturbation variables is $N = 19{:}16$ for the independent ATcT species (normally distributed) and 3 for the electronic structure methods (uniformly distributed). These 19 perturbation variables were randomly sampled 20,000 times. The results are plotted in Figure 5 for PFOA.

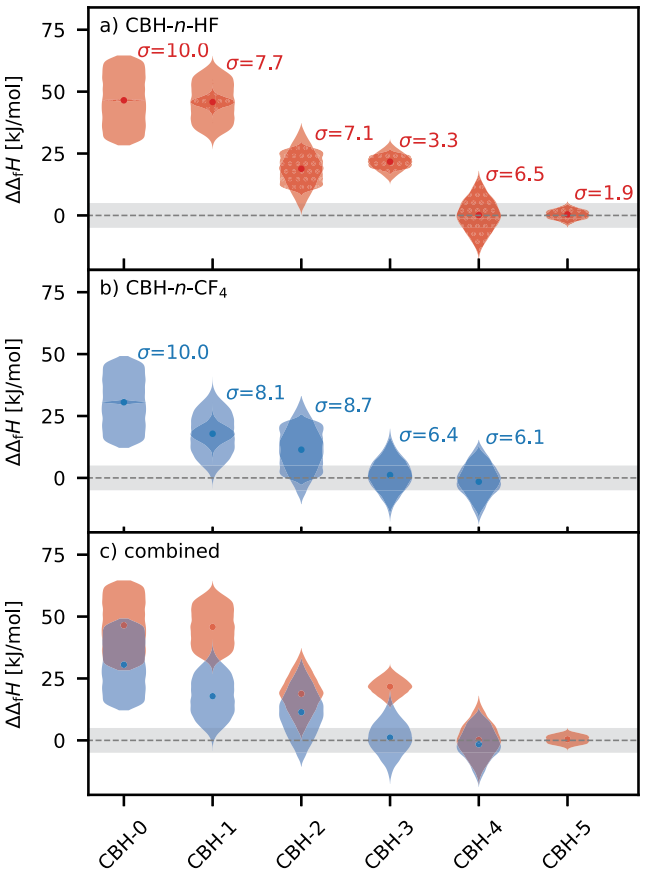


Figure 5. Uncertainty analysis for PFOA, $C_7F_{15}C(O)OH$. (a) CBH using HF has the reference species, (b) CBH using CF_4 as the reference species, and (c) comparison on the total uncertainty from both (a,b). In (a,b) the darker region in the middle is due to uncertainty in the $\Delta_f H^i$ of the reference species, and the larger region also includes the uncertainty in the $\Delta H_{rxn}^{DLPNO_\omega}$.

In Figure 5, the lightly shaded region is the total uncertainty. The darker core in Figure 5a,b is the uncertainty due to the reference species. As expected, the contribution of the reference species to the total uncertainty is negligible at CBH-0, and the total uncertainty is dominated by the computed enthalpy of reaction. By CBH-4, however, the situation is reversed: the uncertainty in the reference species account for virtually all of the total uncertainty. Importantly,

https://doi.org/10.1021/acsphyschemau.3c00056
ACS Phys. Chem Au 2024, 4, 247−258

though, the total variance decreases with increasing n. The one exception is for CBH-4-HF. The increase in the standard deviation (from σ = 3.3 at CBH-3-HF to σ = 6.5 at CBH-4-HF) is due to the fact that the reference species in CBH-4-HF had comparatively larger uncertainties of ±4 kJ/mol (see Supporting Information for details). Figure 5c illustrates the convergence of the different CBH methods for PFOA. If we take the two highest rungs for each CBH method (CBH-3-CF_4, CBH-4-HF, CBH-4-CF_4, and CBH-5-HF) and take their respective means, the standard deviation of those means is 0.9 kJ/mol. The standard deviation for CBH-5-HF is 1.9 kJ/mol, Figure 5. Consequently, based upon both the convergence study and uncertainty analysis, we believe that the enthalpy of formation for PFOA at 0 K is −3488 ± 4 kJ/mol at 2σ uncertainty. We assume that the other species in Table 4 have similarly uncertainties of ±4 kJ/mol, given that they are smaller in size.

4. CONCLUSIONS

A hierarchy of electronic structure methods was combined with the connectivity based hierarchy scheme to compute the standard enthalpies of formation for a large number of per- and poly fluoroalkyl substances. These values include the first reported $\Delta_f H°$ for many PFAS. Two distinct CBH schema are discussed, a conventional approach based on HF, and alternative approach based upon CF_4 at the lowest rung. The CF_4-based approach converges more quickly, but the HF-based approach is able to go to higher levels of error cancellation. A new open-source code, AutoCBH, is introduced, which automatically computes the CBH levels, with flexibility to include alternative schema. The combined approach provides enthalpies of formation at 0K that should maintain chemical accuracy. This assertion was tested for PFOA using a global uncertainty quantification approach, and the results confirmed convergence to within ±4 kJ/mol at 2σ uncertainty. This new approach is not limited to PFAS, but could be applied to any chemical system, subject to the availability of reference data and suitably accurate electronic structure methods.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsphyschemau.3c00056.

List of the specific reactions used to compute the enthalpies of formation at the different CBH levels. NASA7 polynomials for the species (PDF)

Additional SI file (TXT)

AUTHOR INFORMATION

Corresponding Author

C. Franklin Goldsmith − *Chemical Engineering Group, School of Engineering, Brown University, Providence, Rhode Island 02912, United States;* orcid.org/0000-0002-2212-0172; Email: franklin_goldsmith@brown.edu

Author

Kento Abeywardane − *Chemical Engineering Group, School of Engineering, Brown University, Providence, Rhode Island 02912, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsphyschemau.3c00056

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Bjarne Kreitz for many helpful discussions on the CBH methodology. This work was supported by the U.S. Army Corps of Engineers, Strategic Environmental Research and Development Program (SERDP), through award number 008305, with Dr. Andrea Leeson as the program manager, and by the US Department of Energy, Basic Energy Science, through grant number DE-SC0019489, with Dr. Wade Sisk as the program manager.

REFERENCES

(1) Moody, C. A.; Field, J. A. Perfluorinated Surfactants and the Environmental Implications of Their Use in Fire-Fighting Foams. *Environ. Sci. Technol.* **2000**, *34*, 3864−3870.

(2) Schultz, M. M.; Barofsky, D. F.; Field, J. A. Fluorinated Alkyl Surfactants. *Environ. Eng. Sci.* **2003**, *20*, 487−501.

(3) Sunderland, E. M.; Hu, X. C.; Dassuncao, C.; Tokranov, A. K.; Wagner, C. C.; Allen, J. G. A review of the pathways of human exposure to poly- and perfluoroalkyl substances (PFASs) and present understanding of health effects. *J. Expo. Sci. Environ. Epidemiol.* **2019**, *29*, 131−147.

(4) Stroo, H. F.; Leeson, A.; Marqusee, J. A.; Johnson, P. C.; Ward, C. H.; Kavanaugh, M. C.; Sale, T. C.; Newell, C. J.; Pennell, K. D.; Lebrón, C. A.; Unger, M. Chlorinated Ethene Source Remediation: Lessons Learned. *Environ. Sci. Technol.* **2012**, *46*, 6438−6447.

(5) Strynar, M.; Dagnino, S.; McMahen, R.; Liang, S.; Lindstrom, A.; Andersen, E.; McMillan, L.; Thurman, M.; Ferrer, I.; Ball, C. Identification of Novel Perfluoroalkyl Ether Carboxylic Acids (PFECAs) and Sulfonic Acids (PFESAs) in Natural Waters Using Accurate Mass Time-of-Flight Mass Spectrometry (TOFMS). *Environ. Sci. Technol.* **2015**, *49*, 11622−11630.

(6) Sun, M.; Arevalo, E.; Strynar, M.; Lindstrom, A.; Richardson, M.; Kearns, B.; Pickett, A.; Smith, C.; Knappe, D. R. U. Legacy and Emerging Perfluoroalkyl Substances Are Important Drinking Water Contaminants in the Cape Fear River Watershed of North Carolina. *Environ. Sci. Technol. Lett.* **2016**, *3*, 415−419.

(7) Kucharzyk, K. H.; Darlington, R.; Benotti, M.; Deeb, R.; Hawley, E. Novel treatment technologies for PFAS compounds: A critical review. *J. Environ. Manage.* **2017**, *204*, 757−764. Global Trends in the Environmental Remediation Industry

(8) Winchell, L. J.; Ross, J. J.; Wells, M. J. M.; Fonoll, X.; Norton, J. W.; Bell, K. Y. Per- and polyfluoroalkyl substances thermal destruction at water resource recovery facilities: A state of the science review. *Water Environ. Res.* **2021**, *93*, 826−843.

(9) Wang, J.; Lin, Z.; He, X.; Song, M.; Westerhoff, P.; Doudrick, K.; Hanigan, D. Critical Review of Thermal Decomposition of Per- and Polyfluoroalkyl Substances: Mechanisms and Implications for Thermal Treatment Processes. *Environ. Sci. Technol.* **2022**, *56*, 5355−5370.

(10) Khan, M.; So, S.; da Silva, G. Decomposition kinetics of perfluorinated sulfonic acids. *Chemosphere* **2020**, *238*, 124615.

(11) Altarawneh, M.; Almatarneh, M. H.; Dlugogorski, B. Z. Thermal decomposition of perfluorinated carboxylic acids: Kinetic model and theoretical requirements for PFAS incineration. *Chemosphere* **2022**, *286*, 131685.

(12) Weber, N. H.; Dixon, L. J.; Stockenhuber, S. P.; Grimison, C. C.; Lucas, J. A.; Mackie, J. C.; Stockenhuber, M.; Kennedy, E. M. Thermal decomposition of PFOA: Influence of reactor and reaction conditions on product formation. *Chem. Eng. Sci.* **2023**, *278*, 118924.

(13) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 theory. *J. Chem. Phys.* **2007**, *126*..

(14) Khan, M. Y.; Song, J.; Narimani, M.; da Silva, G. Thermal decomposition mechanism and kinetics of perfluorooctanoic acid

(PFOA) and other perfluorinated carboxylic acids: a theoretical study. *Environ. Sci.: Processes Impacts* **2022**, *24*, 2475−2487.
(15) Lorpaiboon, W.; Ho, J. High-Level Quantum Chemical Prediction of C−F Bond Dissociation Energies of Perfluoroalkyl Substances. *J. Phys. Chem. A* **2023**, *127*, 7943−7953.
(16) Raza, A.; Bardhan, S.; Xu, L.; Yamijala, S. S. R. K. C.; Lian, C.; Kwon, H.; Wong, B. M. A Machine Learning Approach for Predicting Defluorination of Per- and Polyfluoroalkyl Substances (PFAS) for Their Efficient Treatment and Removal. *Environ. Sci. Technol. Lett.* **2019**, *6*, 624−629.
(17) Adi, M. A.; Altarawneh, M. Thermal decomposition of heptafluoropropylene-oxide-dimer acid (GenX). *Chemosphere* **2022**, *289*, 133118.
(18) Alinezhad, A.; Shao, H.; Litvanova, K.; Sun, R.; Kubatova, A.; Zhang, W.; Li, Y.; Xiao, F. Mechanistic Investigations of Thermal Decomposition of Perfluoroalkyl Ether Carboxylic Acids and Short-Chain Perfluoroalkyl Carboxylic Acids. *Environ. Sci. Technol.* **2023**, *57*, 8796−8807. PMID: 37195265
(19) Paultre, C.-B.; Mebel, A. M.; O'Shea, K. E. Computational Study of the Gas-Phase Thermal Degradation of Perfluoroalkyl Carboxylic Acids. *J. Phys. Chem. A* **2022**, *126*, 8753−8760. PMID: 36374611
(20) Blotevogel, J.; Giraud, R. J.; Rappé, A. K. Incinerability of PFOA and HFPO-DA: Mechanisms, kinetics, and thermal stability ranking. *Chem. Eng. J.* **2023**, *457*, 141235.
(21) Ding, X.; Song, X.; Chen, X.; Ding, D.; Xu, C.; Chen, H. Degradation and mechanism of hexafluoropropylene oxide dimer acid by thermally activated persulfate in aqueous solutions. *Chemosphere* **2022**, *286*, 131720.
(22) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. High-accuracy extrapolated ab initio thermochemistry. III. Additional improvements and overview. *J. Chem. Phys.* **2008**, *128*, 114111.
(23) Császár, A. G.; Allen, W. D.; Schaefer, H. F.; Henry, F. In pursuit of the ab initio limit for conformational energy prototypes. *J. Chem. Phys.* **1998**, *108*, 9751−9764.
(24) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions. *J. Chem. Phys.* **2006**, *125*, 144108.
(25) Klippenstein, S. J.; Harding, L. B.; Ruscic, B. Ab initio Computations and Active Thermochemical Tables Hand in Hand: Heats of Formation of Core Combustion Species. *J. Phys. Chem. A* **2017**, *121*, 6580−6602.
(26) Somers, K. P.; Simmie, J. M. Benchmarking Compound Methods (CBS-QB3, CBS-APNO, G3, G4, W1BD) against the Active Thermochemical Tables: Formation Enthalpies of Radicals. *J. Phys. Chem. A* **2015**, *119*, 8922−8933.
(27) Farina, D. S.; Sirumalla, S. K.; Mazeau, E. J.; West, R. H. Extensive High-Accuracy Thermochemistry and Group Additivity Values for Halocarbon Combustion Modeling. *Ind. Eng. Chem. Res.* **2021**, *60*, 15492−15501.
(28) Paulechka, E.; Kazakov, A. Critical Evaluation of the Enthalpies of Formation for Fluorinated Compounds Using Experimental Data and High-Level Ab Initio Calculations. *J. Chem. Eng. Data* **2019**, *64*, 4863−4874.
(29) Ruscic, B.; Bross, D. H. Active Thermochemical Tables (ATcT) values based on ver. 1.122 of the Thermochemical Network, 2023. https://atct.anl.gov/ (accessed on Aug 01, 2023).
(30) Ruscic, B. Uncertainty quantification in thermochemistry, benchmarking electronic structure computations, and Active Thermochemical Tables. *Int. J. Quantum Chem.* **2014**, *114*, 1097−1101.
(31) Hehre, W. J.; Ditchfield, R.; Radom, L.; Pople, J. A. Molecular Orbital Theory of the Electronic Structure of Organic Compounds. V. Molecular Theory of Bond Separation. *J. Am. Chem. Soc.* **1970**, *92*, 4796−4801.
(32) Pople, J. A.; Radom, L.; Hehre, W. J. Molecular Orbital Theory of the Electronic Structure of Organic Compounds. VII. Systematic Study of Energies, Conformations, and Bond Interactions. *J. Am. Chem. Soc.* **1971**, *93*, 289−300.
(33) Wheeler, S. E.; Houk, K. N.; Schleyer, P. v. R.; Allen, W. D. A Hierarchy of Homodesmotic Reactions for Thermochemistry. *J. Am. Chem. Soc.* **2009**, *131*, 2547−2560.
(34) Ramabhadran, R. O.; Raghavachari, K. Theoretical Thermochemistry for Organic Molecules: Development of the Generalized Connectivity-Based Hierarchy. *J. Chem. Theory Comput.* **2011**, *7*, 2094−2103.
(35) Ramabhadran, R. O.; Raghavachari, K. Extrapolation to the Gold-Standard in Quantum Chemistry: Computationally Efficient and Accurate CCSD(T) Energies for Large Molecules Using an Automated Thermochemical Hierarchy. *J. Chem. Theory Comput.* **2013**, *9*, 3986−3994.
(36) Sengupta, A.; Raghavachari, K. Prediction of Accurate Thermochemistry of Medium and Large Sized Radicals Using Connectivity-Based Hierarchy (CBH). *J. Chem. Theory Comput.* **2014**, *10*, 4342−4350.
(37) Xu, H.; Xu, Z.; Liu, L.; Li, Z.; Zhu, Q.; Ren, H. Method and automatic program for accurate thermodynamic data of reaction mechanisms for combustion modeling. *Fuel* **2022**, *329*, 125431.
(38) Elliott, S. N.; Keçeli, M.; Ghosh, M. K.; Somers, K. P.; Curran, H. J.; Klippenstein, S. J. High-Accuracy Heats of Formation for Alkane Oxidation: From Small to Large via the Automated CBH-ANL Method. *J. Phys. Chem. A* **2023**, *127*, 1512−1531.
(39) Klippenstein, S. J.; Cavallotti, C. *Mathematical Modelling of Gas-Phase Complex Reaction Systems: Pyrolysis and Combustion*; Computer Aided Chemical Engineering; Faravelli, T., Manenti, F., Ranzi, E., Eds.; Elsevier, 2019; Vol. *45*, pp 115−167.
(40) Mulvihill, C. R.; Danilack, A. D.; Goldsmith, C. F.; Demireva, M.; Sheps, L.; Georgievskii, Y.; Elliott, S. N.; Klippenstein, S. J. Non-Boltzmann Effects in Chain Branching and Pathway Branching for Diethyl Ether Oxidation. *Energy Fuels* **2021**, *35*, 17890−17908.
(41) Klippenstein, S. J. Spiers Memorial Lecture: Theory of unimolecular reactions. *Faraday Discuss.* **2022**, *238*, 11−67.
(42) Ghosh, M. K.; Elliott, S. N.; Somers, K. P.; Klippenstein, S. J.; Curran, H. J. Group additivity values for the heat of formation of C_2−C_8 alkanes, alkyl hydroperoxides, and their radicals. *Combust. Flame* **2023**, *257*, 112492.
(43) Elliott, S. N.; Moore, K. B.; Copan, A. V.; Georgievskii, Y.; Keçeli, M.; Somers, K. P.; Ghosh, M. K.; Curran, H. J.; Klippenstein, S. J. Systematically derived thermodynamic properties for alkane oxidation. *Combust. Flame* **2023**, *257*, 112487.
(44) Kreitz, B.; Abeywardane, K.; Goldsmith, C. F. Linking Experimental and Ab Initio Thermochemistry of Adsorbates with a Generalized Thermochemical Hierarchy. *J. Chem. Theory Comput.* **2023**, *19*, 4149−4162.
(45) Sharma, S.; Abeywardane, K.; Goldsmith, C. F. Theory-Based Mechanism for Fluoromethane Combustion I: Thermochemistry and Abstraction Reactions. *J. Phys. Chem. A* **2023**, *127*, 1499−1511.
(46) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.
(47) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(48) Goerigk, L.; Grimme, S. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670.
(49) Adler, T. B.; Knizia, G.; Werner, H.-J. A simple and efficient CCSD(T)-F12 approximation. *J. Chem. Phys.* **2007**, *127*, 221106.
(50) Adler, T. B.; Werner, H.-J.; Manby, F. R. Local explicitly correlated second-order perturbation theory for the accurate treatment of large molecules. *J. Chem. Phys.* **2009**, *130*, 054106.
(51) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom−atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.

(52) Zhao, Y.; Truhlar, D. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(53) Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. Exploring the Accuracy Limits of Local Pair Natural Orbital Coupled-Cluster Theory. *J. Chem. Theory Comput.* **2015**, *11*, 1525−1539.
(54) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse maps: A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory. *J. Chem. Phys.* **2016**, *144*, 024109.
(55) Elliott, S. N.; Moore, K. B.; Moberg, D.; Mulvihill, C.; Harding, L. B.; Georgievskii, Y.; Jasper, A. W.; Copan, A. V.; Maffei, L. P.; Cavallotti, C.; Burke, M. P.; Goldsmith, C. F.; Klippenstein, S. J. AutoMech: Large-Scale Automated Mechanism Modeling, v0.1.2, 2022. http://tcg.cse.anl.gov/papr/codes/automech.html/ (accessed on Jan 01, 2022).
(56) Elliott, S. N.; Moore, K. B.; Copan, A. V.; Keçeli, M.; Cavallotti, C.; Georgievskii, Y.; Schaefer, H. F.; Klippenstein, S. J. Automated theoretical chemical kinetics: Predicting the kinetics for the initial stages of pyrolysis. *Proc. Combust. Inst.* **2021**, *38*, 375−384.
(57) Zaleski, D. P.; Sivaramakrishnan, R.; Weller, H. R.; Seifert, N. A.; Bross, D. H.; Ruscic, B.; Moore, K. B.; Elliott, S. N.; Copan, A. V.; Harding, L. B.; Klippenstein, S. J.; Field, R. W.; Prozument, K. Substitution Reactions in the Pyrolysis of Acetone Revealed through a Modeling, Experiment, Theory Paradigm. *J. Am. Chem. Soc.* **2021**, *143*, 3124−3142.
(58) Frisch, M. J.; et al. *Gaussian 09*, Revision D.01.; Gaussian Inc: Wallingford CT, 2013.
(59) Werner, H.-J.; et al. *MOLPRO, Version 2022.1, a Package of Ab Initio Programs*, 2022.
(60) Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, No. e1327.

https://doi.org/10.1021/acsphyschemau.3c00056
ACS Phys. Chem Au 2024, 4, 247−258