Received: 1 May 2020 | Revised: 15 July 2020 | Accepted: 18 July 2020
DOI: 10.1002/wcms.1501

FOCUS ARTICLE

WIREs COMPUTATIONAL MOLECULAR SCIENCE WILEY

Applications of isodesmic-type reactions for computational thermochemistry

Bun Chan[1] | **Eric Collins**[2] | **Krishnan Raghavachari**[2]

[1]Graduate School of Engineering, Nagasaki University, Nagasaki, Japan
[2]Department of Chemistry, Indiana University, Bloomington, Indiana

Correspondence
Bun Chan, Graduate School of Engineering, Nagasaki University, Bunkyo 1-14, Nagasaki-shi, Nagasaki 852-8521, Japan.
Email: bun.chan@nagasaki-u.ac.jp

Abstract

In computational thermochemistry, "isodesmic-type" reactions play a significant role for obtaining accurate thermochemical quantities using low-cost methods that can be applied to large systems. This review touches on some of the examples. For instance, a series of relative bond dissociation energies (BDEs) have been devised to calculate absolute BDEs with near-chemical-accuracy (~5 kJ mol^{-1}) using density functional theory (DFT) methods. To facilitate the applicability of isodesmic-type reactions, the connectivity-based hierarchy (CBH) has been developed to automate the systematic generation of isodesmic-type reaction schemes, and applied to large organic and biomolecular systems. The related netCBH scheme yields accurate reaction energies in complex organic reactions, achieving coupled-cluster quality results at DFT cost. Isodesmic-type reactions have been used to obtain heats of formation for medium-sized fullerenes, with uncertainties of ~20 kJ mol^{-1} up to C_{180}. In comparison, the literature C_{60} heat of formation has an uncertainty of 100 kJ mol^{-1}. Importantly, it fills the gap in which heats of formation for those larger fullerenes are not available. These studies showcase how isodesmic-type reactions propel the accuracy of quantum chemistry computations to a level that rivals or even betters modern experimental determinations, particularly for systems that are difficult to study experimentally.

This article is categorized under:
Structure and Mechanism > Reaction Mechanisms and Catalysis
Electronic Structure Theory > Combined QM/MM Methods
Theoretical and Physical Chemistry > Thermochemistry

KEYWORDS
computational chemistry, connectivity-based hierarchy, fullerene, isodesmic reaction, thermochemistry

1 | INTRODUCTION

Thermochemistry concerns the energy change of chemical transformations. It covers reaction energies, which include fundamental quantities of a chemical species such as its heat of formation and ionization energy. Many chemical reactivities can thus be described by thermochemistry, and therefore, the accurate determination of these quantities is important to the understanding of existing chemistry and the prediction of new ones. Let us consider a generic reaction

https://doi.org/10.1002/wcms.1501

A → B, accurate determination of the reaction energy can be accomplished by two means. The first being the determination of accurate energies for both A and B, and the second being the reliance on a high degree of error cancellation between the two energies.

The direct and accurate computation of total energies for small chemical species has become possible with the development of supercomputer technologies, advanced software algorithms, and efficient approximate protocols such as black-box quantum chemistry composite procedures.[1,2] These include, notably, the G*n*-type,[3–5] CBS,[6] and ccCA[7] protocols, the W*n*-type methods,[8–10] and the various HEAT procedures.[11] Nonetheless, highly accurate quantum chemistry computations are, by their fundamental nature, associated with high-order scaling with system size, thus rendering them prohibitive even for medium-sized systems and even with the most advanced computer systems.[12,13]

In situations where the systems of interest are sizable, the exploitation of systematic cancellation of errors provides a viable (and often the only) alternative for quantum-chemistry calculations with adequate accuracy. This article illustrates how one can put this into practice with various isodesmic-type reaction schemes, which are loosely defined as reactions in which the types of bonds in the products are the same as those in the reactants.[14] In particular, we will illustrate the application of such an approach to some large molecular systems.

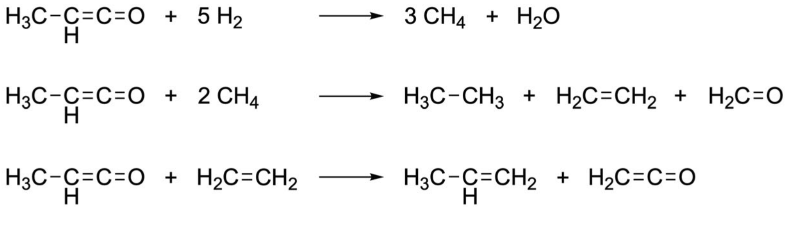
2 | PIONEERING DEVELOPMENTS

In the early days of computational quantum chemistry, the calculation of even small molecules (with just a few non-hydrogen atoms) was feasible only with low-level protocols such as Hartree–Fock (HF) with a small basis set. The resulting calculated energies, as judged by the so-called "chemical accuracy",[15] were often inadequate for chemical applications. Thus, the concept of isodesmic reactions was proposed as a means to obtain reasonable thermochemical quantities within the limitation of existing computational resources of the day.[16]

Isodesmic reactions are actual or hypothetical "bond separation" processes in which the number of each type of formal bonds is retained. It can be illustrated by reaction 2 in Figure 1. On the reactant side, the main substrate (methylketene) has one C—C single bond, one C═C double bond, and one C═O double bond. On the product side of the equation, these bonds are "separated" from one another in the form of isolated ethane, ethene, and formaldehyde. The equation is balanced with two methane molecules, which provide not only the correct number of carbon atoms but also the same number (12) of C—H bonds on both sides of the equation. In a landmark study, Pople and co-workers demonstrated that some prototypical isodemsic reactions can be computed with reasonable accuracy of a few kcal mol^{-1} with even HF in conjunction with the STO-3G minimal basis set.[16]

In addition to isodesmic reactions, the separation of conceptual chemical building blocks can take other forms. Isogyric reactions (Figure 1, reaction 1)[17] require retention of the total number of formal electron pairs but there is no requirement for the conservation of the different types of bonds, which is mandatory for isodesmic reactions. Homodesmotic reactions (Figure 1, reaction 3),[18] in their simplest form, essentially retain the numbers and types of atoms around each atom of a molecule. Such a definition necessitates the satisfaction of the requirements for isodesmic reactions. Isogyric, isodesmic, and homodesmotic reactions represent a series of reaction types in which an increasing amount of bonding information of the reactants is preserved in the products. In general, the accuracy for computed reaction energies also increase in the same order.[19]

As mentioned earlier, one can accurately compute reaction energies for isodesmic reactions, often with low-level quantum chemistry methods. In general, the accuracy for computed reaction energies for homodesmotic reactions is even better (but it works less well for isogyric reactions, as expected). If accurate thermochemical quantities are available for all species in the reaction except the one for the target species, for example, methylketene in Figure 1, then one can obtain the target value with good accuracy (see Box 1 for a working example). In simple cases, the requirement for



(1) $H_3C{-}\underset{H}{C}{=}C{=}O + 5\,H_2 \longrightarrow 3\,CH_4 + H_2O$

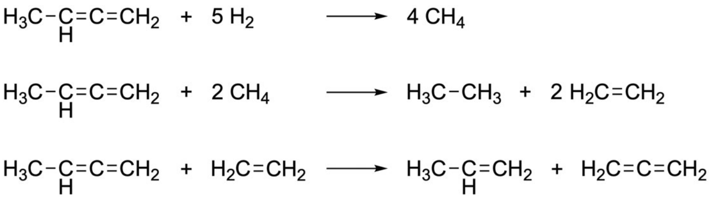
(2) $H_3C{-}\underset{H}{C}{=}C{=}O + 2\,CH_4 \longrightarrow H_3C{-}CH_3 + H_2C{=}CH_2 + H_2C{=}O$

(3) $H_3C{-}\underset{H}{C}{=}C{=}O + H_2C{=}CH_2 \longrightarrow H_3C{-}\underset{H}{C}{=}CH_2 + H_2C{=}C{=}O$

FIGURE 1 Examples of (1) isogyric, (2) isodesmic, and (3) homodesmotic reactions

BOX 1 A working example of using isodesmic-type reactions to obtain accurate thermochemical quantity

In this example, we consider the heat of formation (HoF) of 1,2-butadiene as the target quantity to be calculated using isodesmic-type reactions. It is analogous to methylketene in reactions 1–3. However, while the HoF for methylketene is yet to be experimentally determined, the experimental HoF for 1,2-butadiene is known (162.2 kJ mol^{-1}), thus allowing assessment of our computed results. We will use reactions A–C that are analogous to reactions 1–3 to compute the target HoF.



(A) $H_3C-\underset{H}{C}=C=CH_2 + 5\,H_2 \longrightarrow 4\,CH_4$

(B) $H_3C-\underset{H}{C}=C=CH_2 + 2\,CH_4 \longrightarrow H_3C-CH_3 + 2\,H_2C=CH_2$

(C) $H_3C-\underset{H}{C}=C=CH_2 + H_2C=CH_2 \longrightarrow H_3C-\underset{H}{C}=CH_2 + H_2C=C=CH_2$

The experimental HoFs for the complementary species in A–C are all available with chemical accuracy [0.0 (H_2), −74.5 (CH_4), −84.0 (C_2H_6), 52.4 (C_2H_4), 19.9 ($CH_3CH{=}CH_2$), and 189.9 ($CH_2{=}C{=}CH_2$) kJ mol^{-1}]. To obtain the 298 K reaction enthalpies, we use the HF/STO-3G//AM1 protocol, that is, AM1 geometries, zero-point vibrational energies [ZPVE], and thermal corrections for 298 K enthalpies [Tc(H_{298})], and HF/STO-3G single-point energies [E(HF)] on AM1 geometries. These values are given in the following table.

Species	E(HF) Hartree	ZPVE kJ mol^{-1}	Tc (H_{298}) kJ mol^{-1}
H_2	−1.11611	26.0	8.7
CH_4	−39.72421	116.6	10.0
C_2H_6	−78.30042	195.0	11.9
C_2H_4	−77.07232	133.9	10.5
$CH_3CH{=}CH_2$	−115.65469	209.8	13.6
$CH_2{=}C{=}CH_2$	−114.42037	147.4	12.4
$CH_3CH{=}C{=}CH_2$	−153.00141	223.1	15.9

The total 298 K enthalpy for a given species is simply E(HF) + ZPVE + Tc(H_{298}). Using the resulting 298 K enthalpies, the reaction energies for A–C are calculated simply as Σ [E(HF),products] – Σ [E(HF),reactants]. The values are −732.7 (A), 16.1 (B), and −3.8 (C) kJ Mol^{-1}. For each reaction, the enthalpy is also equal to Σ [HoF,products] – Σ [HoF,reactants]. For reaction A, we thus have −732.7 [reaction energy] = 4 × −74.5 [HoF (CH_4)] – {5 × 0.0 [HoF(H_2)] + HoF($CH_3CH{=}C{=}CH_2$)}, which leads to an HoF($CH_3CH{=}C{=}CH_2$) of 434.6 kJ Mol^{-1}. This value deviates from the experimental value of 162.2 kJ mol^{-1} by 272.4 kJ mol^{-1}. In comparison, the corresponding target HoF values obtained with reactions B and C are 153.7 and 161.2 kJ mol^{-1}, which are in good agreement with experiment. These results thus illustrate that accurate thermochemical quantities can be obtained using economical computational chemistry methods in conjunction with reasonably well-balanced isodesmic-type reactions.

accurate thermochemical values for the accompanying species is often satisfied. This is because these species are usually small molecules (e.g., H_2, CH_4, and H_2O in reaction 1) for which highly accurate experimental values are available, or they can be computed with higher-accuracy methods due to their modest sizes. There are additional considerations in relation to error propagation when one attempts to maximize accuracy, which will be briefly discussed in the following sections.

3 | CONTEMPORARY ISODESMIC-TYPE SCHEMES

3.1 | Systematic hierarchy of homodesmotic reactions

The aforementioned isodemic-type reactions are seemingly well defined for the simple cases in Figure 1. However, the definitions are not always applied rigorously; isodesmic reactions are sometimes mistakenly quoted as homodesmotic.[20–22] In addition, different variants of homodesmotic and related reaction types have been proposed, which include hyper-homodesmotic,[23] semi-homodesmotic,[24] isogeitonic,[25] and isoplesiotic[26] reactions to name but a few.

In an attempt to resolve the confusion, Wheeler et al. have defined a systematic series of isodesmic-type reactions.[19] Their highest level of reaction type (hyper-homodesmotic reaction) preserves the first layer of neighboring groups surrounding each bond. They show that, for a range of hydrocarbons, the heats for formation can be obtained with an accuracy of ~1 kJ mol^{-1} using a range of modest-level quantum chemistry methods in conjunction with this highest level of hierarchy. While their systematic scheme has originally been formulated for closed-shell species such as hydrocarbons, it has subsequently been extended to cover radicals.[27] For more details on the schemes of Wheeler and coworkers, we refer readers to a previous review article in this journal.[28]

3.2 | Relative bond dissociation energies

We note that isodesmic-type reactions have often been used for the calculation of heats of formation, which is a "global" property for a given molecule. Chemical reactions, however, usually occur at a particular site of a molecule. In this regard, accurate determination of such "local" thermochemical properties is also of practical relevance. An important local molecular property is the bond dissociation energy (BDE), which is a key reactivity indicator. In this regard, Chan and Radom have explored the use of (loosely speaking) isodesmic-type reactions to obtain accurate BDEs with a combination of high-level and moderate quantum-chemistry methods. Specifically, they have devised a range of *relative* BDEs [Figure 2, RBDE(a)–(c)] for a set of 261 BDEs comprising systematically of bonds between the elements H, C, N, O, F, Si, P, S, and Cl (termed the BDE261 set).[29]

When one computes a BDE directly, the reaction typically involves a closed-shell species on one side of the equation and two radicals on the other side. Because the treatment of radical species may be subject to spin contamination effects that are sometimes problematic,[30] the calculation of BDEs using modest-level of methods often does not yield a high-level of accuracy.[29] For the BDE261 set, the ubiquitous B3-LYP[31] method yields a mean absolute deviation (MAD) of 26.7 kJ mol^{-1} against reference values obtained using the high-level W1[32] method (Table 1). The use of the double-hybrid density-functional-theory (DH-DFT) method DuT-D3[33] provides a vastly improved description of the BDEs, having an MAD of 6.4 kJ mol^{-1} that falls just outside of "chemical accuracy" of 1 kcal mol^{-1} (4.2 kJ mol^{-1}); the accuracy is further improved with the higher-level G4(MP2)-6X[4] method (4.8 kJ mol^{-1}). Note that the terms "mean absolute

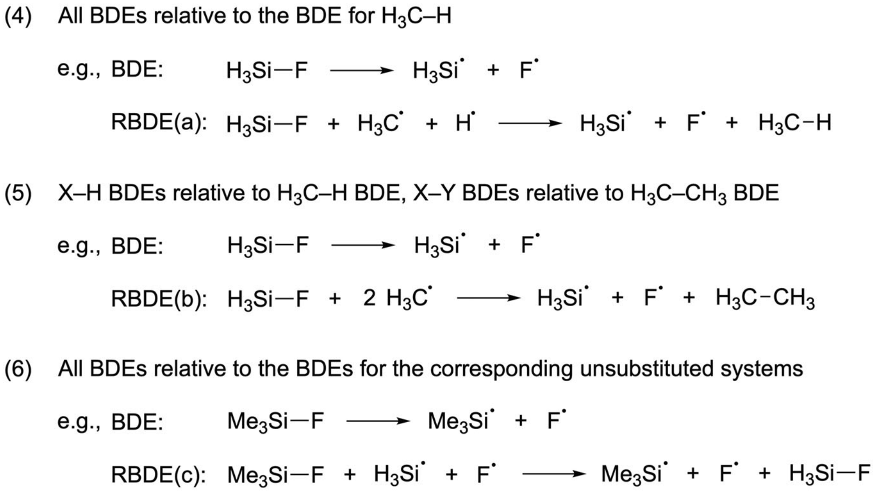


FIGURE 2 Examples of relative bond dissociation energies (RBDEs) for the estimation of absolute BDEs by exploiting cancellations of systematic deviations

deviation" and "mean absolute error" are often used interchangeably in the literature. To avoid confusion, in this article, MAD refers to the average of the absolute differences between values obtained with two different methods.

In comparison, in the calculation of an RBDE, the overall equation balances the number of radical species on both sides, thus facilitating cancellation of deviations associated with treatments of radicals. An accurately computed RBDE can then be combined with an accurate reference BDE to yield an accurate absolute BDE for the target system. The RBDE(a) scheme [Figure 2, (4)] represents a simple case where the target BDE can be computed using RBDE with the H_3C–H BDE employed as a reference. It can be seen that the accuracy uniformly improved with RBDE(a), although the MAD for B3-LYP is still substantial (20.1 kJ mol^{-1}).

It is noteworthy that, when the target BDE is associated with breaking a C–H bond, RBDE(a) is equivalent to the commonly used quantity "radical stabilization energy" (RSE) for carbon-centered radicals ($R^\bullet$), which is the energy of the reaction $R^\bullet + H_3C\text{–}H \rightarrow R\text{–}H + H_3C^\bullet$. In comparison with the results for BDE261, the accuracy of B3-LYP for RSEs is considerably better (MAD ~5 kJ mol^{-1}) for a prototypical set of carbon radicals.[34] Such a difference highlights the utility of balancing the chemical features in isodesmic-type reactions, as in the case of RSE calculations, that is, balancing C—H bond with C—H bond, carbon-centered radical with carbon-centered radical.

In comparison with RBDE(a), the RBDE(b) scheme represents a better balance of chemical features by providing a better reference, namely the $H_3C—CH_3$ bond, for BDEs that involving breaking a bond between two heavy atoms. This slight increase in the complexity of the isodesmic-type scheme leads to better accuracy for all three methods in Table 1; in particular, the MAD for B3-LYP is significantly lowered to 7.4 kJ mol^{-1}. The improvements gained by going from RBDE(b) to RBDE(c) are slight. We note that, for the BDE261 set of systems, RBDE(c) requires 45 reference BDEs, as opposed to just two for RBDE(b). This illustrates a consideration when one applies isodesmic-type reaction schemes. Specifically, when one devises "better" but more complex protocols, one may at some point encounter diminishing returns. Where the "sweet spot" lies, however, might not be entirely clear nor uniquely defined; it may vary depending on the system and property of interest.

3.3 | Hierarchy of relative bond dissociation energies

The BDE261 set that we have discussed in the previous section involves a diverse set of bond types. The substituents on the substrates (and the corresponding radicals), however, are relatively limited (H_3C– and F–). Given the structural diversity of species (e.g., peptides and DNAs in biological systems) for which their BDEs are of practical relevance, it is of interest to accurately evaluate BDEs for species with a wide range of substitution patterns. In this regard, Radom and co-workers have examined the accuracy of a variety of quantum-chemistry methods for the calculation of RSEs for a range of substituted carbon-centered radicals.[35]

Consider a generic radical $(X)(Y)(Z)C^\bullet$, its RSE is defined relative to $H_3C^\bullet$ (Figure 3, reaction 7). Thus, it quantifies the effect of simultaneous substitution of all three hydrogen atoms with the groups X, Y, and Z. One can also intuitively consider the total effect of the three functional groups as the sum of the effects of each group acting independently on the methyl radical. Such an additive scheme would be an accurate approximation to the actual RSE if the effect of one substituent on the radical does not affect how the radical interacts with another substituent. Conversely, the deviation of this additivity from the actual RSE (i.e., DARSE, Equation 11) would be an indicator for the extent of an overall cooperative or hindering interactions between the substituents. For a tri-substituted carbon-centered radical, that is, $X \neq H$ and $Y \neq H$ and $Z \neq H$, one can further extend this concept of deviation from additivity by including di-substituted species $[(X)(Y)HC^\bullet, (Y)(Z)HC^\bullet$, and $(Z)(X)HC^\bullet]$ in the set of reference systems, leading to the quantity DPARSE.

While Reference 35 focuses on the chemical aspect of RSE, DARSE, and DPARSE, it is perhaps not difficult to see that they form a hierarchy of isodesmic-type reactions with increasing sophistication in the order RSE → DARSE →

TABLE 1 Mean absolute deviations (kJ mol^{-1}) from reference W1 values for the various quantum-chemistry methods for the calculation of the BDE261 set of bond dissociation energies (BDEs) and relative BDEs using isodesmic-type schemes listed in Figure 2

Method	BDE	RBDE(a)	RBDE(b)	RBDE(c)
B3-LYP	26.7	20.1	7.4	6.1
DuT-D3	6.4	4.5	2.7	1.3
G4(MP2)-6X	4.8	3.6	2.3	1.2

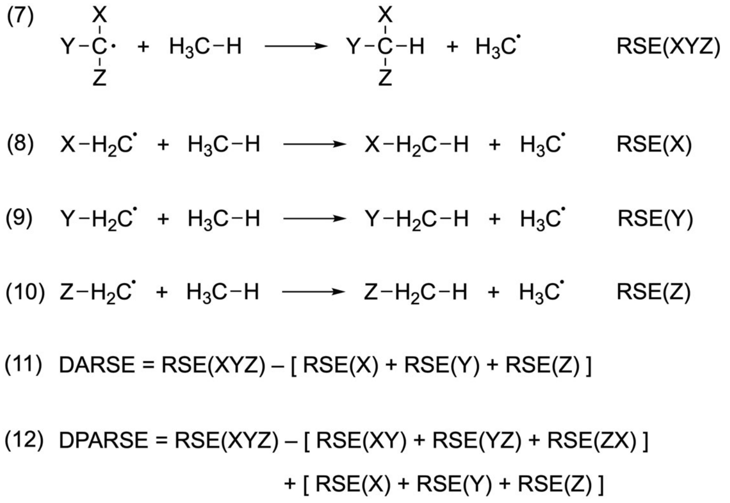
DPARSE, which one may be able to exploit for obtaining accurate thermochemical quantities for a (relatively large) target system. At this point, we note again that RSEs are related to relative bond dissociation energies (RBDEs) when the reference system is the $H_3C—H$ BDE. In a subsequent study,[36] the concept of a systematic sequence of RSEs has been adopted to the calculation of corresponding BDE-type quantities, namely BDE, RBDE, deviation from additivity of the RBDE (DARBDE) and deviation from pair-wise additivity of the RBDE (DPARBDE).

For a set of ~50 substituted carbon-centered radicals, the associated C–H BDEs and related relative quantities have been obtained with the G4(MP2)-6X method, which itself has been selected on the basis of benchmark for a subset using values computed with the even-higher-level W1X-2[37] method. The systems include σ and π electron-withdrawing and electron-donating groups such as CH_3, CF_3, BH_2, and NH_2. Some key results from the assessment are shown in Table 2.

While M06-2X[38] and DuT-D3 are quite accurate for the direct computation of the C–H BDEs, this is not the case for the other three methods in Table 2. In comparison to direct calculations of BDEs, the accuracy generally improves with the hierarchy of relative BDEs in the order BDE → RBDE → DARBDE → DPARBDE. Notably, for B-LYP,[39] M06-L[40] and B3-LYP, large improvements can be seen along the way, such that their accuracies for DARBDE and DPARBDE are not far behind those for M06-2X and DuT-D3. In Reference 36, the concept of the systematic hierarchy of relative BDEs has also been applied to C–CH_3 and C–F BDEs, with similar findings of improving accuracies along with the series.

Among DARBDE and DPARBDE, the former is more straightforward to apply because it only involves, along with the target species, just the mono-substituted systems (analogous to Equation 11), whereas the calculation of DPARBDE also requires di-substituted species. Thus, one may exploit accurately calculated DARBDE using low-cost methods in combination with higher-level RBDEs for mono-substituted systems to adequately and efficiently obtain RBDEs for multiply substituted species. Table 3 demonstrates how this can be put into practice.

The M06-2X method, as a hybrid-DFT, provides a good description of the RBDEs at a reasonable computational cost. Nonetheless, its accuracy for C—CH_3 RBDEs is somewhat worse than those for C—H and C—F RBDEs. This aspect



$$\text{(7)}\quad Y\!-\!\overset{X}{\underset{Z}{C}}\cdot + H_3C\!-\!H \longrightarrow Y\!-\!\overset{X}{\underset{Z}{C}}\!-\!H + H_3C^{\bullet} \qquad \text{RSE(XYZ)}$$

$$\text{(8)}\quad X\!-\!H_2C^{\bullet} + H_3C\!-\!H \longrightarrow X\!-\!H_2C\!-\!H + H_3C^{\bullet} \qquad \text{RSE(X)}$$

$$\text{(9)}\quad Y\!-\!H_2C^{\bullet} + H_3C\!-\!H \longrightarrow Y\!-\!H_2C\!-\!H + H_3C^{\bullet} \qquad \text{RSE(Y)}$$

$$\text{(10)}\quad Z\!-\!H_2C^{\bullet} + H_3C\!-\!H \longrightarrow Z\!-\!H_2C\!-\!H + H_3C^{\bullet} \qquad \text{RSE(Z)}$$

$$\text{(11)}\quad \text{DARSE} = \text{RSE(XYZ)} - [\,\text{RSE(X)} + \text{RSE(Y)} + \text{RSE(Z)}\,]$$

$$\text{(12)}\quad \text{DPARSE} = \text{RSE(XYZ)} - [\,\text{RSE(XY)} + \text{RSE(YZ)} + \text{RSE(ZX)}\,] + [\,\text{RSE(X)} + \text{RSE(Y)} + \text{RSE(Z)}\,]$$

FIGURE 3 Radical stabilization energy (RSE, 7–10), deviation from additivity of the RSE (DARSE, 11) devised from the RSEs, and deviation from pair-wise additivity of the RSE (DPARSE, 12) devised in similar fashion by including di-substituted systems

TABLE 2 Mean absolute deviations (kJ mol^{-1}) from reference G4(MP2)-6X Values for the various quantum-chemistry methods for the calculation of a set of ~50 C—H bond dissociation energies (BDEs), relative BDEs (RBDEs), deviation from the additivity of the BDEs (DARBDEs), and deviation from the pair-wise additivity of the BDEs (DPARBDEs)

Method	BDE	RBDE	DARBDE	DPARBDE
B-LYP	31.1	21.2	5.9	3.9
M06-L	22.2	17.2	4.4	2.3
B3-LYP	19.3	14.0	4.0	2.6
M06-2X	4.6	3.5	3.7	1.1
DuT-D3	4.6	4.6	2.7	0.9

TABLE 3 Mean absolute deviations (kJ mol^{-1}) with respect to benchmark G4(MP2)-6X values for C—H, C—CH_3, and C—F relative bond dissociation energies (RBDEs) calculated using lower-cost procedures and estimated by the various approximation schemes using a combination of RBDEs and deviation from the additivity of the RBDEs (DARBDEs)

Scheme	A	B	C	D
RBDE	M06-2X	G4(MP2)-6X	M06-L	M06-2X
DARBDE	–	M06-2X	–	M06-L
C—H	3.5	3.7	17.2	5.5
C—CH_3	5.3	3.5	22.9	6.2
C—F	3.4	3.6	10.8	5.2

can be improved by using the higher-level G4(MP2)-6X method to obtain reference RBDEs for mono-substituted systems, and combine them with DARBDEs calculated with M06-2X. Conversely, if one were to reduce the cost for computing RBDEs using M06-2X, one may typically use a lower-cost method such as the non-hybrid-DFT method M06-L to directly calculate the RBDE. However, the resulting values are associated with large deviations. These deviations can be substantially lowered when one combines M06-L DARBDEs with reference M06-2X RBDEs, with the resulting MADs being just slightly larger than those for M06-2X.

4 | CONNECTIVITY-BASED HIERARCHY

4.1 | A generalized isodesmic hierachy

Although the idea of the isodesmic scheme dates back to 1970s, most of the more sophisticated schemes in the literature, for example, the homodesmotic hierarchy, have traditionally been somewhat restricted in their applications to a relatively limited range of organic molecules such as hydrocarbons.[19] This has historically been due to the difficulties in extending and generalizing such schemes to cover the vast structural variety present in organic chemistry, owing to their dependence on predefined reactant and product molecules that are based on complex definitions. Overall, there was a clear need to develop a generalizable hierarchy of isodesmic-based schemes that are applicable to all functional groups along with an easy, automated way to generate the associated reactions.

To overcome such limitations present in many isodesmic schemes, Raghavachari and co-workers developed the generalized Connectivity-Based Hierarchy (CBH) of error-cancellation schemes.[41] This thermochemical hierarchy provides a systematic approach based solely on the connectivity of the atoms in a molecule, bypassing any manual balancing of bond and hybridization types. Different levels of CBH can be visualized as the rungs of a ladder, such that ascending the rungs of the hierarchy increasingly preserves chemical environments, and a better matching of the bond-types and hybridization-types is automatically achieved. Much like other isodesmic-based hierarchies, higher rungs of CBH follow the preservation rules of all lower rungs, that is, homodesmotic reactions are also isodesmic and isogyric. As an illustration, the lower rungs (referred to as CBH-n) up to CBH-3 are given for propyl-pent-4-enoate ($C_8H_{14}O_2$) in Figure 4.

These rungs (referred to as CBH-n) alternate between atom- and bond-centered reactions, starting by preserving either the number of each atom or the number of each bond. For example, the lowest atom-centric rung is constructed by extracting all the non-hydrogen atoms and terminating the open valences with hydrogen atoms to create CBH-0 product molecules (called fragments, reaction 13). Similarly, the lowest bond-centric rung (CBH-1, reaction 14) is generated by extracting all the non-hydrogen bonds and terminating them with hydrogen atoms. Subsequent CBH-n rungs, include immediate connectivity to all CBH-($n-2$) type product molecules (fragments), for example, CBH-2 (reaction 15) includes all immediate connectivity to each non-hydrogen atom (CBH-0 fragments), and CBH-3 (reaction 16) includes all immediate connectivity to all heavy-atom bonds (CBH-1 fragments). Additionally, the reactant side of a given CBH rung is defined as overlapping fragments between product molecules and will be found among the products of the previous rung, establishing an extended relationship between the rungs.[41]

As a result, the rungs of CBH align directly with other previously defined schemes in the following order for CBH-0 to CBH-3: isogyric → isodesmic → hypohomodesmotic → hyperhomodesmotic. Higher rungs (CBH-4 and beyond) are defined by the natural progression of the hierarchy, however, in practice, acceptable performance (within chemical accuracy) is typically provided by the CBH-2 and CBH-3 schemes, as shown in the next section. Overall, the underlying

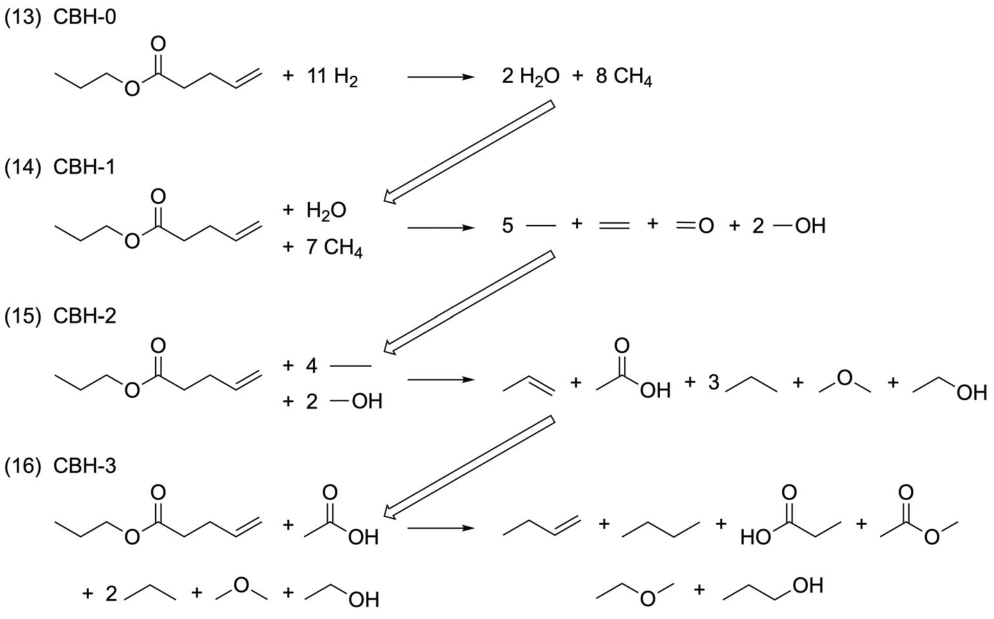


FIGURE 4 Traditional CBH rungs for propyl-pent-4-enoate ($C_8H_{14}O_2$)

logic behind the n-homodesmotic hierarchy of reaction schemes is similar to CBH. However, under the definitions of some of these other related hierarchies, multiple reactions can be written for a given molecule which follow the same preservation rules and fall under the same classification.[41,42] The systematic nature of the CBH protocol along with its dependence solely on the connectivity and chemical bonding between the atoms eliminates this ambiguity, allowing for only one CBH reaction to exist for a given rung and facilitating the automated construction of the reactions.[43]

4.2 | Performance of CBH for thermochemistry

Along with the generation of the reaction schemes, CBH introduces a slightly modified formalism with analogy to fragmentation methods.[44] This arises from the problems associated with traditional protocols for the calculation of enthalpies of formation, particularly when applying higher levels of preservation, that is, higher rungs of the hierarchy beyond the isodesmic (CBH-1) scheme. When ascending the hierarchy, larger product molecules are formed to satisfy the corresponding preservation rules, for example, up to 8 non-hydrogen atoms for CBH-3. In the traditional protocol for the isodesmic-based schemes, reaction enthalpies are computed via electronic structure methods then used in conjunction with accurate reference enthalpies (typically experimentally derived) on the fragments to solve for the heat of formation of a (larger) molecule of interest through Hess' Law. Typically, the uncertainty of experimental values scales with the size of the molecule, leading to unreliable experimental values for larger molecules. Consequently, the traditional protocol cannot be used easily in conjunction with higher rungs.

Instead, more expensive levels of theory, for example, CCSD(T) or G4, can take the place of experiments and be utilized to bypass the problem altogether. In this way, the application of the CBH correction schemes are analogous to fragmentation-based methods, in which the energy, enthalpy, or free energy of the full molecule is calculated at a lower level of theory (typically DFT), then smaller calculations on the reactant and products fragments of the CBH reaction are performed with a more expensive level of theory to construct local corrections to the lower level of theory. The final corrected thermochemical property is then approximately equal to the full molecule calculated at the higher level of theory within a margin of error depending on the fragment size.

To date, the standard CBH correction scheme has been benchmarked on closed- and open-shell, H,C,N,O,F,S,Cl-containing organic and biomolecules, both neutral and cationic with a wide variety of functional groups.[41,45–49] Previous performance of CBH on the calculation of the heat of formation of a set of neutral organic molecules from Reference 41 is showcased in Table 4. Using the CBH-3 correction scheme, mean absolute deviations within chemical accuracy are achieved for many different low levels of theory.

TABLE 4 Mean absolute deviations (kJ mol^{-1}) for 298 K heats of formational obtained with traditional CBH

Method	CBH-1	CBH-2	CBH-3
B-PW91	30.2	6.7	7.0
BMK	20.6	5.4	4.2
B3-LYP	31.7	6.5	4.9
M05-2X	8.7	5.7	3.1
M06-2X	11.4	7.2	3.6
TPSSh	32.8	5.4	5.4
B2-PLYP	17.6	4.8	3.7

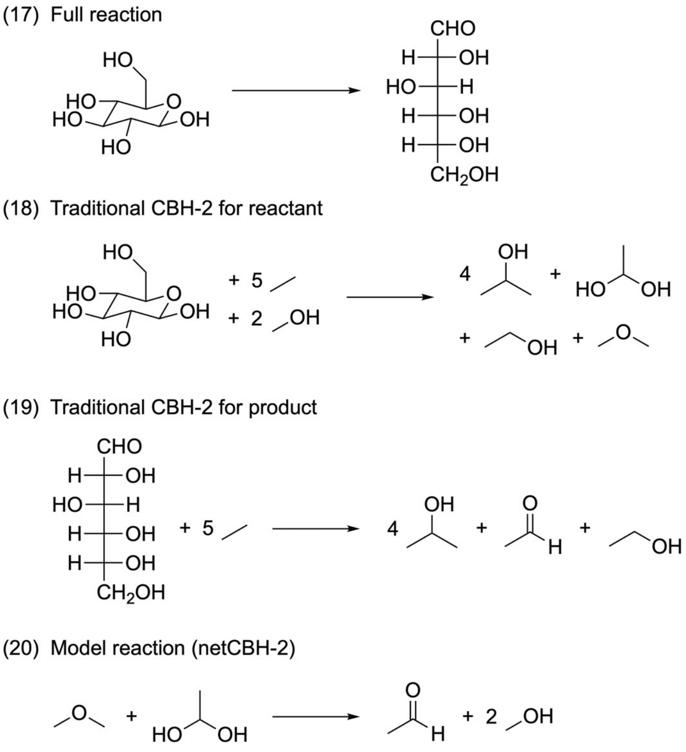


FIGURE 5 NetCBH reaction for carbohydrate ring opening isomerization reaction

5 | THE USE OF CBH FOR COMPLEX ORGANIC REACTIONS

5.1 | The netCBH protocol for reaction energies

Recently the traditional CBH protocol has been extended to be applicable to complex organic reactions.[50] In this scheme, a correction is applied to reaction enthalpies calculated at a low level of theory through the "netCBH" protocol. As stated in the previous section, these chemical reactions are a more "local" property compared to the "global" heat of formation. Thus, for reactions in which the product and reactant parent molecules are similar, the corresponding CBH fragments mostly cancel out on both sides of the reaction with the only fragments involving the transformation are needed to construct a correction.

One such reaction, the ring-opening isomerization reaction of carbohydrate, is shown in Figure 5 (reaction 17). To construct the netCBH scheme, first, the traditional CBH-2 schemes are created for both the reactant (reaction 18) and product (reaction 19) molecules. The two reactions are then subtracted from one another giving the net reaction fragments (reaction 20). Since this transformation only involves the conversion of the hydroxyl group to its linear carbonyl form, the model netCBH reaction in this case only involves the fragments around this functional group. The netCBH-2 scheme can then be used to calculate a correction to the original reaction in a similar way to the traditional CBH protocol.

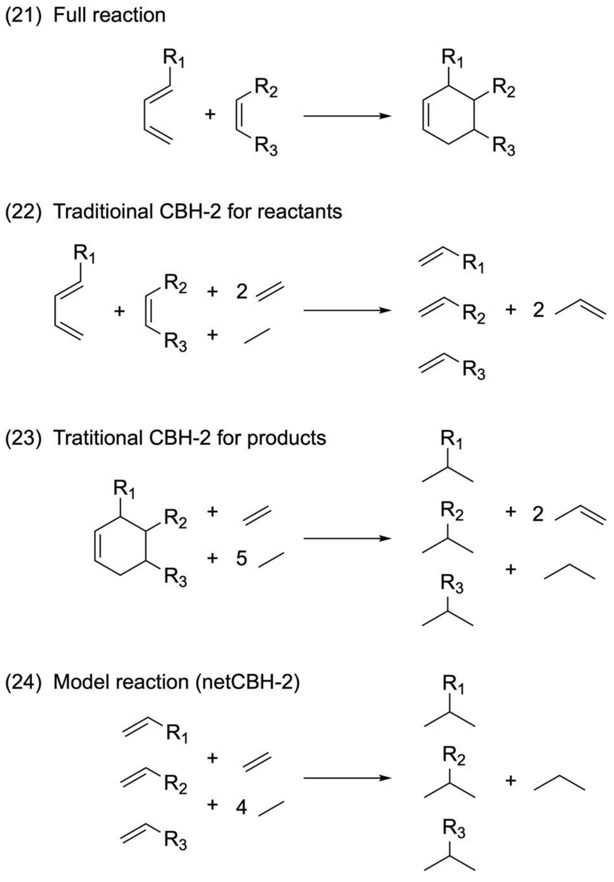
This protocol also works for multiple reactant or product molecules, such as in Diels-Alder reactions shown in Figure 6 (reactions 21–24), in which the CBH-n scheme of each species are constructed separately then added together. Here, the model reaction captures the direct change of bonding environment around the R-groups involved in the carbon–carbon bond formation. More generally, the netCBH protocol was applied to common, complex organic reactions, including Diels–Alder, aldol condensation, Pauson–Khand reaction, aminoxylation, and isomerization reactions.

Many of these transformations involve interactions that are difficult for DFT to model to an acceptable accuracy. For example, B3-LYP and many other density functionals do not accurately model Diels–Alder reactions, underestimating the reaction enthalpies due to an inadequate description of $\sigma \rightarrow \pi$ bond transformations, along with hyperconjugation effects present in cyclic and bicyclic products.[47] The application of the ΔCBH-1 scheme only slightly improves the performance of raw B3-LYP on Diels–Alder reaction enthalpies, reducing the MAD from 63.5 to 45.3 kJ mol^{-1}. This highlights the inability of the isodesmic scheme to appropriately capture the chemical transformation. Drastically better accuracies are achieved with the ΔCBH-2 scheme, further reducing the MAD to 3.7 kJ mol^{-1}, well within the chemical accuracy threshold.

5.2 | Towards chemical accuracy for complex reactions

In addition to the examples given above, impressive improvements are similarly seen for a larger set of reactions for a wide range of density functional and wave-function based methods, highlighted in Table 5 from Reference 50. If greater accuracy is required, the remaining errors in these corrected reaction energies could also be mitigated with the use of rungs beyond CBH-2, since the local corrections will capture a larger portion of the chemical environment. Many of the CBH-2 errors for these complex reactions line up with the previous CBH-2 results shown in Table 4, indicating these reactions could potentially benefit further from the CBH-3 correction scheme.

It is important to note that these results are almost completely *density-functional independent*. One of the major problems with the calculation of thermochemical properties with DFT is the substantial disagreement between different levels of theory, leading to a high dependence on the density functional employed. The same is true for reaction


(21) Full reaction

(22) Traditioinal CBH-2 for reactants

(23) Tratitional CBH-2 for products

(24) Model reaction (netCBH-2)

FIGURE 6 NetCBH reaction for generic Diels–Alder reactions

TABLE 5 Mean absolute deviations (kJ mol^{-1}) for netCBH reaction for complex organic reactions

Method	CBH-0	CBH-1	CBH-2
	DFT methods[a]		
B3-PW91	31.8	36.4	6.7
B3-LYP	54.0	40.2	7.1
BMK	16.7	23.0	5.9
CAM-B3-LYP	29.7	33.1	8.0
M06-2X	10.5	19.3	9.2
TPSSh	37.7	35.2	7.1
*ω*B97XD	9.2	20.1	6.7
B3-LYP-D3BJ	26.4	22.2	5.4
CAM-B3-LYP-D3BJ	16.3	22.2	6.7
B3-LYP-NL	19.3	20.1	6.3
B2-PLYP	28.9	23.4	4.6
B2-PLYP-D3BJ	15.5	14.2	4.2
	Wavefunction methods[a]		
HF	55.7	39.8	8.8
MP2	16.7	10.0	5.4
SCS-MP2	7.5	4.6	4.6
CCSD(T)[b]	3.8	2.5	2.1

[a] 6–311++G(3df,2p) basis set employed unless otherwise stated.
[b] Extrapolated using the formula CCSD(T)/6–311++G(3df,2p) = MP2/6–311++G(3df,2p) + CCSD(T)/6–31+G(d,p) – MP2/6–31+G(d,p).

energies calculated here at the DFT/TZV level of theory. Once corrected with the ΔCBH-2 protocol, reaction energies across various methods agree with each other, and virtually any density functional can be used as the low level of theory with an acceptable accuracy. Furthermore, the overall computational cost is determined by the cost of the low level of theory on the full molecule (10–30 non-hydrogen atoms), since the more expensive level of theory is only performed on the CBH-2 fragments (3–5 non-hydrogen atoms). Molecules involved in these transformations are much larger in size than the fragments generated in the CBH-2 reaction scheme, allowing for coupled-cluster quality reaction energies to be obtained at the cost of DFT.

As seen in DFT, even more cost-effective methods, such as minimal basis set Hartree–Fock (HF) or semi-empirical methods can also benefit greatly from the ΔCBH-2 corrections.[51] In many cases, these methods outperform DFT for reaction energies demonstrating that they also benefit significantly from error cancellation. Although these methods do not fall within the chemical accuracy threshold, their performance highlights the importance of selecting an appropriate low-level theory, since they are solely responsible for the modeling of long-range physics for these reactions along with providing errors which are systematic. Overall, the correction schemes established with CBH and netCBH protocols allow the study of both absolute heats of formation as well as complex organic reaction within the realm of chemical accuracy at a reasonable cost.

6 | APPLICATION TO FULLERENE THERMOCHEMISTRY

6.1 | Background on fullerene chemistry

We now present the last set of case studies in this review. Recently, isodesmic-type reactions have been applied to fullerenes with the aim of obtaining computational heats of formation values with the best achievable accuracy.[52,53] These works were motivated by the importance of fullerenes in chemical sciences and yet their experimentally determined heats of formation either carry substantial uncertainties (e.g., 2,560 ± 100 kJ mol^{-1} for the buckyball C_{60})[54] or do not exist at all. The combination of the strength in applying isodesmic-type reactions and advances in computational chemistry technologies has enabled the theoretical approach for tackling this issue.

Fullerene molecules are hollow carbon clusters consisting of a single sheet of *sp*^{2} carbon atoms. In this regard, they are similar to typical polycyclic aromatic hydrocarbon (PAH) molecules, which also comprise mainly *sp*^{2} carbon atoms. A key distinction between fullerenes and most PAHs is that, while PAHs usually have all *sp*^{2} carbon atoms in their natural planar form, the spherical/ellipsoidal shapes of fullerenes require distortion of their constituent atoms from planarity. The non-planar structures of fullerenes are a consequence of them having pentagon rings in addition to hexagons, whereas PAHs generally contain hexagon rings. A specific class of fullerene isomers are those that satisfy the "isolated pentagon rule" (IPR).[55] These isomers, by definition, do not have pentagon–pentagon junctions, and they are often lower in energies than those (non-IPR) isomers that possess such junctions. A prime example of IPR fullerene is the icosahedron (I_h), C_{60}, that is, the famed buckyball fullerene.

With the understanding of these key chemical characteristics of fullerenes, one can then construct isodesmic-type reactions that conserve these very characteristics to maximize cancellation of errors. In Reference 52 that focuses on I_h C_{60}, a key ingredient in the equations is corannulene ($C_{20}H_{10}$), which is an atypical PAH with a bowl-shaped structure. It consists of a pentagon ring surrounded by five interconnected hexagons, which is a key structural motif for I_h C_{60} (Figure 7). Importantly, this bowl-shape molecule serves the purpose of partially conserving the curvature and the associated strain in I_h C_{60}.

6.2 | Heat of formation for corannulene

To successfully obtain accurate thermochemical properties of a target molecule with isodemic-type reactions, one relies not only on error cancellation for the reaction, but also on having accompanying species with small uncertainties. In relation to the application to I_h C_{60}, the established experimental heat of formation for corannulene has an uncertainty of ±9.2 kJ mol^{-1}.[54] This is deemed inadequate as it would amount to tens of kJ mol^{-1} in the total uncertainty for I_h C_{60}. This shortcoming has been overcome in Reference 52 by applying isodesmic-type reaction schemes (Figure 8) and the accurate quantum chemistry method W1h[56] for the calculation of the corannulene's heat of formation.

A series of similar reactions for other PAHs have been used to assess the total uncertainties in these reaction schemes, with the conclusion being that reaction 26 provides the best overall accuracy. This leads to the final recommended heat of formation for corannulene of 485.2 ± 7.9 kJ mol^{-1}. While the uncertainty of ±7.9 kJ mol^{-1}, in the first instance, does not appear to be a significant improvement over that (±9.2 kJ mol^{-1}) for the experimental value, a subsequent study[53] has more rigorously re-examined the uncertainty in the theoretical value. As a result, a smaller (but still conservative) value of ±4.4 kJ mol^{-1} has been proposed. Significantly, the results in References 52,53 reveal that the existing experimental heat of formation for corannulene (458.5 kJ mol^{-1}) is likely to deviate from the actual value

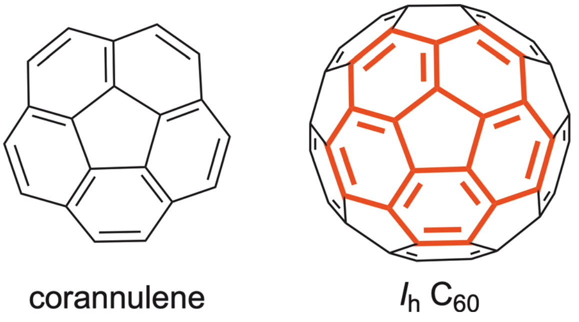


FIGURE 7 Corannulene and its connection with I_h C_{60} as a structural motif

$$\text{(25)}\quad \text{Ar} + m_1\,\text{CH}_4 \longrightarrow n_1\,\text{C}_2\text{H}_6 + o_1\,\text{C}_2\text{H}_4$$

$$\text{(26)}\quad \text{Ar} + m_2\,\text{C}_2\text{H}_4 \longrightarrow n_2\,\text{C}_6\text{H}_6 \quad \text{(benzene)}$$

$$\text{(27)}\quad \text{Ar} + m_3\,\text{C}_6\text{H}_6 \longrightarrow n_3\,\text{C}_{10}\text{H}_8 \quad \text{(naphthalene)}$$

$$\text{(28)}\quad \text{Ar} + m_4\,\text{C}_{10}\text{H}_8 \longrightarrow n_4\,\text{C}_{14}\text{H}_{10} \quad \text{(phenanthrene)}$$

FIGURE 8 Isodesmic-type reactions used for obtaining the heats of formation for aromatic systems [Ar = naphthalene ($C_{10}H_8$), phenanthrene ($C_{14}H_{10}$), triphenylene ($C_{18}H_{12}$), and corannulene ($C_{20}H_{10}$)]

by almost 30 kJ mol^{-1}. This revelation has prevented a significant error from being incurred in the subsequently calculated heat of formation for I_h C_{60}.

6.3 | Heat of formation for I_h C_{60}

The isodemistic-type reactions used in Reference 52 for obtaining I_h-C_{60} heat of formation involve just corannulene and planar PAHs (Figure 9, reactions 29–32). A robust double-hybrid (DH) DFT method (DSD-PBE-P86)[57] has been used to obtain the reaction energies. The four reactions yield heats of formation values ranging from 2,490.2 to 2,554.1 kJ mol^{-1}. Among the species in these reactions, benzene and naphthalene have fairly accurate reference heats of formation values, with uncertainties of ±0.3, and ±1.5 mol^{-1}.[58,59] The quoted uncertainties for the literature heats of formation for phenanthrene and triphenylene are ±2.3 and ±4.4 kJ mol^{-1},[59] which would contribute ±23 and ±44 kJ mol^{-1} to the calculated heats of formation values for I_h C_{60}.

Reactions that involve larger planar PAHs are also associated with a larger number of corannulene, which also has a relatively large uncertainty for its heat of formation. On the other hand, one can expect that the use of larger PAHs would provide better conservation of the chemical characteristics with corannulene and thus smaller errors in the calculated reaction energies. As a compromise of these contrasting factors, in Reference 52, the recommended heat of formation for I_h C_{60} (2,521.6 kJ mol^{-1}) has been taken as the average value of the ones obtained with reaction 29 (benzene) and reaction 30 (naphthalene).

In Reference 53, a somewhat different approach is used to obtain the I_h-C_{60} heat of formation. It employs a slightly different DH-DFT method (DSD-PBE-PBE)[57] that, in comparison with DSD-PBE-P86, yields more accurate reaction energies for isodesmic-type reactions related to those shown in 25–28 (Figure 8). In addition, rigorous error analysis has been carried out to obtain uncertainties for each of the reactions 33–36. The final heat of formation value for I_h C_{60} and its associated uncertainties is taken from the overlapping region spanned by the error bars of all four reactions. This value (2,520.0 ± 20.7 kJ mol^{-1}) is in good agreement with the one in Reference 52. In other subsequent studies,[60,61] similar values ranged in ~2,510–2,530 kJ mol^{-1} have been proposed. They all suggest that the existing literature value of 2,560 kJ mol^{-1} is somewhat too positive, by ~40 kJ mol^{-1}.

6.4 | Heats of formation for larger fullerenes

The provision of adequate I_h-C_{60} heat of formation in References 52,53 paves the way for obtaining heats of formation for larger fullerenes. Specifically, we can intuitively expect that a fullerene that is slightly larger than I_h C_{60} shares similar chemical characteristics with I_h C_{60}. Thus, a well-constructed isodesmic-type reaction with the two similar-sized fullerenes might be associated with a large degree of error cancellation.

In Reference 53, the smallest fullerene examined that is larger than I_h C_{60} is C_{70}. The difference of 10 carbon atoms is balanced with unsaturated hydrocarbon molecules, specifically the *trans*-butadiene/benzene and the styrene/naphthalene pairs (Figure 10). These reactions not only preserve the chemistry of the two similar-sized fullerenes, the small number of excess atoms in C_{70} also necessitate the inclusion of just a small number of spectator species to balance the equations. This in turn reduces the total uncertainty in the calculated heat of formation for C_{70}.

The two reactions (33 and 34) lead to different heat of formation values for C_{70} together with different error bars. The recommended heat of formation and the associated uncertainty is taken from the overlapping region of the two ranges. This value is then used to obtain the heat of formation for C_{76}, which is subsequently applied in the isodesmic-type reactions for C_{78}, and so on. In total, heats of formation for ten medium-sized fullerenes (up to C_{320}) have been obtained in Reference 53. They are 2,520.0 ± 20.7 (C_{60}), 2,683.4 ± 17.7 (C_{70}), 2,862.0 ± 18.5 (C_{76}), 2,878.8 ± 13.3 (C_{78}),

$$(29)\quad C_{60} + 10\,C_6H_6 \longrightarrow 6\,C_{20}H_{10}$$

$$(30)\quad C_{60} + 10\,C_{10}H_8 \longrightarrow 8\,C_{20}H_{10}$$

$$(31)\quad C_{60} + 10\,C_{14}H_{10} \longrightarrow 10\,C_{20}H_{10}$$

$$(32)\quad C_{60} + 10\,C_{18}H_{12} \longrightarrow 12\,C_{20}H_{10}$$

FIGURE 9 Isodesmic-type reactions used for obtaining the heats of formation for I_h C_{60} using corannulene ($C_{20}H_{10}$) and planar aromatic hydrocarbons with increasing sizes [naphthalene ($C_{10}H_8$), phenanthrene ($C_{14}H_{10}$) and triphenylene ($C_{18}H_{12}$)]

$$\text{(33)} \quad C_{70} + 5\,C_4H_6 \longrightarrow C_{60} + 5\,C_6H_6$$

$$\text{(34)} \quad C_{70} + 5\,C_8H_8 \longrightarrow C_{60} + 5\,C_{10}H_8$$

$$\text{(35)} \quad C_{76} + C_{14}H_{10} \longrightarrow C_{70} + C_{20}H_{10}$$

$$\text{(36)} \quad C_{78} + C_4H_6 \longrightarrow C_{76} + C_6H_6$$

$$\text{(37)} \quad C_{78} + C_8H_8 \longrightarrow C_{76} + C_{10}H_8$$

$$\text{(38)} \quad C_{180} \longrightarrow 2\,C_{90}$$

$$\text{(39)} \quad C_{180} \longrightarrow 3\,C_{60}$$

FIGURE 10 Examples of isodesmic-type reactions used for obtaining the heats of formation for medium-sized fullerenes using similar-sized fullerenes together with accompanying species *trans*-butadiene (C_4H_6), benzene (C_6H_6), styrene (C_8H_{10}), naphthalene ($C_{10}H_8$), phenanthrene ($C_{14}H_{10}$), and corannulene ($C_{20}H_{10}$)

2,946.4 ± 14.5 (C_{84}), 3,067.3 ± 15.4 (C_{90}), 3,156.6 ± 16.2 (C_{96}), 3,967.7 ± 33.4 (C_{180}), 4364 (C_{240}) and 5415 (C_{320}) kJ mol^{-1}. For many of these, bracketing with values and their associated error bars obtained with multiple isodesmic-type reactions yield contained uncertainties of ~20 kJ mol^{-1}.

In subsequent studies,[62,63] the quantum-chemistry and isodesmic-type-reaction protocols established in Reference 53 have been applied to obtain heats of formation for an even larger number of small to medium-sized fullerenes. Relative stabilities for different isomers of a given sized fullerene have also been determined. The large number of accurate stability data has been applied in Reference 63 to build a chemically intuitive model for estimating the heats of formation for all fullerenes. In Reference 62, the calculated heats for formation for C_{84} isomers are used to resolve identities of some experimentally observed isomers. These studies have showcased the utility of accurate thermochemical properties in answering experimental questions. In turn, they illustrate the use of isodesmic-type reactions in conjunction with quantum chemistry computations for advancing chemical sciences.

7 | OTHER APPLICATIONS AND FUTURE PROSPECTS

The above discussion and case studies mainly concern the use of isodesmic-type reactions for obtaining accurate thermochemical quantities, and we have focused primarily on works of our own. Nevertheless, before we conclude, we now briefly touch on selected results of others to illustrate the wider range of applications of isodesmic-type reactions. We do note, however, that applications involving similar ideas are numerous in the literature and discussing them comprehensively is well beyond the scope of this article.

In addition to accurate computations related to thermochemistry as discussed in this article, a complementary, yet powerful, utility of isodesmic-type reactions is to provide a deeper understanding of the fundamentals in chemistry. For example, we have earlier touched on the concept of "deviation from additivity of the radical stabilization energies" (DARSE) and related quantities, which can be considered as effective measures of how substituents surrounding a carbon-centered radical interact with one another.

Many other useful concepts can be derived from an isodesmic-based analysis. In one study, a series of isodesmic-type reactions of cyclic hydrocarbons have been used to quantify bond strains in those systems.[64] In another study, aromaticity for a wide range of aromatic and anti-aromatic compounds have been determined.[65] For systems where both bond strain and (anti-)aromaticity are relevant, isodesmic-type reactions can be used to separate and independently evaluate these two factors.[66,67] Other chemical phenomena that can be examined using isodesmic-type reactions include, for example, intramolecular hydrogen bonds[68] and the trans effect.[69]

We also reiterate another feature of isodesmic-type reactions, namely its analogy to fragment-based methods (see sections 4 and 5). The latter include, for instance, the group additivity schemes of Benson.[70] Brief comparisons between the two types of methods have been given in References 19, 28, and 42. Here we simply note that isodesmic-type and fragment-based methods can be used in a complementary manner. As an example,[71] high-level wave-function composite schemes, in conjunction with isodesmic-type reactions, have been applied to the calculation of heats of formation for small to medium-sized hydrocarbons; the "isodesmic" values are then further improved using a statistically derived group-additivity scheme to yield the eventual best estimates.

Finally, we note that the accuracy of isodesmic-type reactions obtained using low-level methods often depends on how well the chemical features in the reactants are preserved in the products. In principle, this would exclude extended systems for which small structural motifs are deemed inadequate for capturing the characteristics that might be unique in the extended structure. An example for this is graphene and related systems, in which the π bond networks are often fully delocalized. There are, however, reports of specifically constructed isodesmic-type schemes that are applicable to such species.[72,73] On that note, we are hopeful that, with further educated imaginations and careful developments, isodesmic-type reactions can be utilized in an even wider range of chemical applications.

8 | CONCLUDING REMARKS

Computational quantum chemistry has become a powerful tool for obtaining thermochemical quantities with an accuracy that, at its best, surpasses precise experimental determinations. The direct approach for accurate quantum chemistry calculations of thermochemical properties, however, necessitates the use of high-level methods that are still limited to relatively small molecules.

The use of "isodesmic-type" reactions provides an indirect means for obtaining adequately accurate thermochemical quantities using lower-methods, which can be applied to substantially larger systems. In essence, these reactions are constructed in an attempt to preserve the chemical features between reactants and products. In this manner, errors incurred on the reactant side would ideally be similar to those on the product side regardless of the method used to obtain the reaction energy. This would then lead to superior cancellation of errors, and thus accurate reaction energies even when a moderate method is employed in the computations.

Over decades of development, many isodesmic-type reaction schemes have been proposed, and they have been demonstrated to yield good accuracy for a diverse range of thermochemical quantities over a wide range of systems. Some examples given in this review include the use of a hierarchy of relative bond dissociation energies (BDEs) to calculate absolute BDEs. In these cases, it has been shown that near-chemical-accuracy ($\sim$5 kJ mol^{-1}) can be achieved using a combination of low-cost hybrid and non-hybrid density functional theory (DFT) methods. Notably, this outcome represents a substantial improvement over the direct application of non-hybrid DFT methods (deviations $\sim$10–20 kJ mol^{-1}), with little requirement for additional computational resources.

To enable straightforward application of isodesmic-type reactions, the connectivity-based hierarchy (CBH) has been defined, which provides a general and automated procedure to generate an increasingly sophisticated hierarchy of isodesmic-type reaction schemes. It leads to a well-defined protocol to obtain accurate thermochemical information on larger molecules containing most functional groups present in organic and biomolecular systems. It also eliminates the need for accurate experimental data on any of the smaller reactant or product molecules by using a higher level of theory to compute the needed reference energies. The resulting netCBH scheme provides an efficient protocol to obtain accurate reaction energies in complex organic reactions, effectively providing coupled-cluster quality results at DFT cost.

Isodesmic-type reactions, in conjunction with moderate quantum chemistry methods, have been used to obtain accurate heats of formation for a series of medium-sized fullerene molecules from C_{60} to C_{320}. In these cases, the isodesmic-type reactions are constructed to preserve the fundamental nature that fullerenes are deformed polycyclic aromatic hydrocarbons. For C_{60}, it involves the bowl-shaped corannulene as a "structural motif". For each of the larger fullerenes, a similar-sized fullerene is employed to provide a close structural resemblance.

For most of the fullerenes, rigorous analysis of the sources of error indicate that these values are generally associated with uncertainties of $\sim$20 kJ mol^{-1} up to C_{180}. To put this in perspective, the literature value for C_{60} heat of formation has an uncertainty of 100 kJ mol^{-1}. Furthermore, heats of formation value for larger fullerenes are typically not available, which, presumably, is in part related to the difficulty for experimental determination. These studies demonstrated the utility of isodesmic-type reactions in propelling the accuracy of quantum chemistry computations to a level that rivals or even betters modern experimental determinations, particularly for systems that are difficult to study experimentally.

ACKNOWLEDGMENTS

Bun Chan acknowledges generous grants of computer time from the RIKEN Information Systems and Cybersecurity (ISC, project Q20266), Japan. Krishnan Raghavachari acknowledges financial support from the National Science Foundation (NSF, grant number CHE-1665427).

CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

AUTHOR CONTRIBUTIONS

Bun Chan: Conceptualization; writing-original draft; writing-review and editing. **Krishnan Raghvachari:** Conceptualization; writing-original draft; writing-review and editing. **Eric Collins:** Conceptualization; writing-original draft; writing-review and editing.

ORCID

Bun Chan https://orcid.org/0000-0002-0082-5497

RELATED WIREs ARTICLES

Determination of enthalpies (Heats) of formation
Homodesmotic reactions for thermochemistry

REFERENCES

1. Cioslowski J, editor. *Quantum-mechanical prediction of thermochemical data*. New York: Kluwer Academic, 2001.
2. Martin JML. Computational thermochemistry: A brief overview of quantum mechanical approaches. *Annu Rep Comput Chem*. 2005;*1*:31–43.
3. Curtiss LA, Redfern PC, Raghavachari K. G*n* theory. *WIREs Comput Mol Sci*. 2011;*1*:810–825.
4. Chan B, Deng J, Radom L. G4(MP2)-6X: A cost-effective improvement to G4(MP2). *J Chem Theory Comput*. 2011;*7*:112–120.
5. Chan B, Karton A, Raghavachari K. G4(MP2)-XK: A variant of the G4(MP2)-6X composite method with expanded applicability for main-group elements up to radon. *J Chem Theory Comput*. 2019;*15*:4478–4484.
6. Petersson GA. Complete basis set models for chemical reactivity: From the helium atom to enzyme kinetics. In: Cioslowski J, editor. *Quantum-mechanical prediction of thermochemical data*. New York: Kluwer Academic, 2001; p. 99–130.
7. Deyonker NJ, Cundari TR, Wilson AK. The correlation consistent composite approach (ccCA): An alternative to the Gaussian-*n* methods. *J Chem Phys*. 2006;*124*:114104-1-17.
8. Karton A. A computational chemist's guide to accurate thermochemistry for organic molecules. *WIREs Comput Mol Sci*. 2016;*6*:292–310.
9. Chan B. How to computationally calculate thermochemical properties objectively, accurately, and as economically as possible. *Pure Appl Chem*. 2017;*89*:699–713.
10. Chan B. Unification of the W1X and G4(MP2)-6X composite protocols. *J Chem Theory Comput*. 2017;*13*:2642–2649.
11. Tajti A, Szalay PG, Császár AG, et al. HEAT: High accuracy extrapolated ab initio thermochemistry. *J Chem Phys*. 2004;*121*:11599–11613.
12. Bartlett RJ. Coupled-cluster theory and its equation-of-motion extensions. *WIREs Comput Mol Sci*. 2011;*2*:126–138.
13. Katouda M, Nakajima T. MPI/OpenMP hybrid parallel algorithm of resolution of identity second-order Møller–Plesset perturbation calculation for massively parallel multicore supercomputers. *J Chem Theory Comput*. 2013;*9*:5373–5380.
14. Muller P. Glossary of terms used in physical organic chemistry. *Pure Appl Chem*. 1994;*66*:1077–1184.
15. Pople JA. Nobel lecture: quantum chemical models. *Rev Mod Phys*. 1999;*71*:1267–1274.
16. Hehre WJ, Ditchfield R, Radom L, Pople JA. Molecular orbital theory of the electronic structure of organic compounds. V. Molecular theory of bond separation. *J Am Chem Soc*. 1970;*92*:4796–4801.
17. Hehre WJ, Radom L, Schleyer PR, Pople JA. *Ab initio molecular orbital theory*. New York: Wiley-Interscience, 1986.
18. George P, Trachtman M, Bock CW, Brett AM. An alternative approach to the problem of assessing stabilization energies in cyclic conjugated hydrocarbons. *Theor Chem Acc*. 1975;*38*:121–129.
19. Wheeler SE, Houk KN, Schleyer PR, Allen WD. A hierarchy of homodesmotic reactions for thermochemistry. *J Am Chem Soc*. 2009;*131*:2547–2560.
20. Ma BY, Sulzbach HM, Xie YM, Schaefer HF. π-Electron delocalization and compression in acyclic acetylenic precursors to multidimensional carbon networks: Comparison with experiment for the recently synthesized tris(trimethylsilyl)-substituted tetraethynylmethane. Structures, thermochemistry, infrared spectra, polarizabilities, and hyperpolarizabilities. *J Am Chem Soc*. 1994;*116*:3529–3538.
21. Minkin VI, Minyaev RM. Cyclic aromatic systems with hypervalent centers. *Chem Rev*. 2001;*101*:1247–1265.
22. Caramori GF, de Oliveira KT, Galembeck SE, Bultinck P, Constantino MG. Aromaticity and homoaromaticity in methano[10]annulenes. *J Org Chem*. 2007;*72*:2698.
23. Hess BA Jr, Schaad LJ. Ab initio calculation of resonance energies. Benzene and cyclobutadiene. *J Am Chem Soc*. 1983;*105*:7500–7505.
24. Nyulaszi L, Varnai P, Veszpremi T. About the aromaticity of five-membered heterocycles. *J Mol Struct (THEOCHEM)*. 1995;*358*:55–61.
25. El-Nahas AM, Bozzelli JW, Simmie JM, Navarro MV, Black G, Curran HJ. Thermochemistry of acetonyl and related radicals. *J Phys Chem A*. 2006;*110*:13618–13623.

26. George P, Bock CW, Trachtman M. The matching of structural elements in reactions for evaluating stabilization energies for benzene and monosilabenzene. *Theor Chim Acta*. 1987;*71*:289–298.
27. Wodrich MD, Corminboeuf C, Wheeler SE. Accurate thermochemistry of hydrocarbon radicals via an extended generalized bond separation reaction scheme. *J Phys Chem A*. 2012;*116*:3436–3447.
28. Wheeler SE. Homodesmotic reactions for thermochemistry. *WIREs Comput Mol Sci*. 2012;*2*:204–220.
29. Chan B, Radom L. BDE261: A comprehensive set of high-level theoretical bond dissociation enthalpies. *J Phys Chem A*. 2012;*116*: 4975–4986.
30. Menon AS, Radom L. Consequences of spin contamination in unrestricted calculations on open-shell species: Effect of Hartree-Fock and Møller-Plesset contributions in hybrid and double-hybrid density functional theory approaches. *J Phys Chem A*. 2008;*112*: 13225–13230.
31. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J Phys Chem*. 1994;*98*:11623–11627.
32. Martin JML, de Oliveira G. Towards standard methods for benchmark quality ab initio thermochemistry—W1 and W2 theory. *J Chem Phys*. 1999;*111*:1843–1856.
33. Chan B, Radom L. Obtaining good performance with triple-ζ-type basis sets in double-hybrid density functional theory procedures. *J Chem Theory Comput*. 2011;*7*:2852–2863.
34. Henry DJ, Parkinson CJ, Mayer PM, Radom L. Bond dissociation energies and radical stabilization energies associated with substituted methyl radicals. *J Phys Chem A*. 2001;*105*:6750–6756.
35. Menon AS, Henry DJ, Bally T, Radom L. Effect of substituents on the stabilities of multiply-substituted carbon-centered radicals. *Org Biomol Chem*. 2011;*9*:3636–3657.
36. Chan B, Radom L. Hierarchy of relative bond dissociation enthalpies and their use to efficiently compute accurate absolute bond dissociation enthalpies for C−H, C−C, and C−F bonds. *J Phys Chem A*. 2013;*117*:3666–3675.
37. Chan B, Radom L. W1X-1 and W1X-2: W1-quality accuracy with an order of magnitude reduction in computational cost. *J Chem Theory Comput*. 2012;*8*:4259–4269.
38. Zhao Y, Truhlar DG. 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*. 2006;*120*:215–241.
39. Johnson BG, Gill PMW, Pople JA. Preliminary results on the performance of a family of density functional methods. *J Chem Phys*. 1992; *97*:7846–7848.
40. Zhao Y, Truhlar DG. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J Chem Phys*. 2006;*125*:194101–1–18.
41. Ramabhadran RO, Raghavachari K. Theoretical thermochemistry for organic molecules: Development of the generalized connectivity-based hierarchy. *J Chem Theory Comput*. 2011;*7*:2094–2103.
42. Ramabhadran RO, Raghavachari K. The successful merger of theoretical thermochemistry with fragment-based methods in quantum chemistry. *Acc Chem Res*. 2014;*47*:3596–3604.
43. The CBH hierarchy depends only on the connectivity and the valence-bond structure of a given molecule. If multiple resonance structures can be written for a molecule with different bond orders between the connected atoms, each one will define a set of unique CBH reaction schemes. See references 45 and 48 for further discussions in the case of aromatic systems and delocalized carbocations.
44. Raghavachari K, Saha A. Accurate composite and fragment-based quantum chemical models for large molecules. *Chem Rev*. 2015;*115*: 5643–5677. and references therein.
45. Ramabhadran RO, Raghavachari K. Connectivity-based hierarchy for theoretical thermochemistry: Assessment using wave function-based methods. *J Phys Chem A*. 2012;*116*:7531–7537.
46. Ramabhadran RO, Sengupta A, Raghavachari K. Application of the generalized connectivity-based hierarchy to biomonomers: Enthalpies of formation of cysteine and methionine. *J Phys Chem A*. 2013;*117*:4973–4980.
47. 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.
48. Collins EM, Sengupta A, AbuSalim DI, Raghavachari K. Accurate thermochemistry for organic cations via error cancellation using connectivity-based hierarchy. *J Phys Chem A*. 2018;*122*:1807–1812.
49. Sengupta A, Ramabhadran RO, Raghavachari K. Accurate and computationally efficient prediction of thermochemical properties of biomolecules using the generalized connectivity-based hierarchy. *J Phys Chem B*. 2014;*118*:9631–9643.
50. Sengupta A, Raghavachari K. Solving the density functional conundrum: Elimination of systematic errors to derive accurate reaction enthalpies of complex organic reactions. *Org Lett*. 2017;*19*:2576–2579.
51. Kromann JC, Welford A, Christensen AS, Jensen JH. Random versus systematic errors in reaction enthalpies computed using semiempirical and minimal basis set methods. *ACS Omega*. 2018;*3*:4372–4377.
52. Karton A, Chan B, Raghavachari K, Radom L. Evaluation of the heats of formation of corannulene and C_{60} by means of high-level theoretical procedures. *J Phys Chem A*. 2013;*117*:1834–1842.
53. Chan B, Kawashima Y, Katouda M, Nakajima T, Hirao K. From C_{60} to infinity: Large-scale quantum chemistry calculations of the heats of formation of higher fullerenes. *J Chem Theory Comput*. 2016;*138*:1420–1429.

54. Linstrom PJ, Mallard WG, editors. *NIST chemistry WebBook, NIST standard reference database number 69*. Gaithersburg, MD: National Institute of Standards and Technology, 2020. https://doi.org/10.18434/T4D303.
55. Kroto HW. The stability of the fullerenes C_n, with n = 24, 28, 32, 36, 50, 60 and 70. *Nature*. 1987;*329*:529–531.
56. Martin JML, Parthiban S. Understanding chemical reactivity. In: Cioslowski J, editor. *Quantum-mechanical prediction of thermochemical data*. New York: Kluwer Academic, 2001; p. 31–65.
57. Kozuch S, Martin JML. Spin-component-scaled double hybrids: An extensive search for the best fifth-rung functionals blending DFT and perturbation theory. *J Comput Chem*. 2013;*34*:2327–2344.
58. Ruscic B, Pinzon RE, Morton ML, et al. Introduction to active thermochemical tables: Several "key" enthalpies of formation revisited. *J Phys Chem A*. 2004;*108*:9979–9997.
59. Roux MV, Temprado M, Chickos JS, Nagano YJ. Critically evaluated thermochemical properties of polycyclic aromatic hydrocarbons. *Phys Chem Ref Data*. 2008;*37*:1855–1996.
60. Wan W, Karton A. Heat of formation for C_{60} by means of the G4(MP2) thermochemicalprotocol through reactions in which C_{60} is broken down intocorannulene and sumanene. *Chem Phys Lett*. 2016;*643*:34–38.
61. Bumpus JA. Gas-phase heat of formation values for buckminsterfullerene (C_{60}), C_{70} fullerene (C_{70}), corannulene, coronene, sumanene, and other polycyclic aromatic hydrocarbons calculated using density functional theory (M06-2X) coupled with a versatile inexpensive group-equivalent approach. *J Phys Chem A*. 2018;*122*:6615–6632.
62. Waite SL, Chan B, Karton A, Page AJ. Accurate thermochemical and kinetic stabilities of C_{84} isomers. *J Phys Chem A*. 2018;*112*: 4768–4777.
63. Chan B, Kawashima Y, Dawson W, Katouda M, Nakajima T, Hirao K. A simple model for relative energies of all fullerenes reveals the interplay between intrinsic resonance and structural deformation effects in medium-sized fullerenes. *J Chem Theory Comput*. 2019;*15*: 1255–1264.
64. Daoust KJ, Hernandez SM, Konrad KM, Mackie ID, Winstanley J Jr, Johnson RP. Strain estimates for small-ring cyclic allenes and butatrienes. *J Org Chem*. 2006;*71*:5708–5714.
65. Suresh CH, Koga N. An isodesmic reaction based approach to aromaticity of a large spectrum of molecules. *Chem Phys Lett*. 2006;*419*: 550–556.
66. Hohlneicher G, Packschies L, Weber J. On the σ,π-energy separation of the aromatic stabilization energy of cyclobutadiene. *Phys Chem Chem Phys*. 2007;*9*:2517–2530.
67. Suresh CH, Lincy TL, Mohan N, Rakhi R. Aromatization energy and strain energy of buckminsterfullerene from homodesmotic reactions. *J Phys Chem A*. 2015;*119*:6683–6688.
68. Deshmukh MM, Suresh CH, Gadre SR. Intramolecular hydrogen bond energy in polyhydroxy systems: A critical comparison of molecular tailoring and isodesmic approaches. *J Phys Chem A*. 2007;*111*:6472–6480.
69. Sajith PK, Suresh CH. Quantification of mutual *trans* influence of ligands in Pd(II) complexes: A combined approach using isodesmic reactions and AIM analysis. *Dalton Trans*. 2010;*39*:815–822.
70. Cohen N, Benson SW. Estimation of heats of formation of organic compounds by additivity methods. *Chem Rev*. 1993;*93*:2419–2438.
71. Bakowies D. Estimating systematic error and uncertainty in ab initio thermochemistry: II. ATOMIC(hc) enthalpies of formation for a large set of hydrocarbons. *J Chem Theory Comput*. 2020;*16*:399–426.
72. Sivaramakrishnan R, Tranter RS, Brezinsky K. Ring conserved isodesmic reactions: A new method for estimating the heats of formation of aromatics and PAHs. *J Phys Chem A*. 2005;*109*:1621–1628.
73. Huang L, Bohorquez HJ, Matta CF, Massa L. The kernel energy method: Application to graphene and extended aromatics. *Int J Quant Chem*. 2011;*111*:4150–4157.

How to cite this article: Chan B, Collins E, Raghavachari K. Applications of isodesmic-type reactions for computational thermochemistry. *WIREs Comput Mol Sci*. 2021;11:e1501. https://doi.org/10.1002/wcms.1501