

ARTICLE

Assessment of Contemporary Theoretical Methods for Bond Dissociation Enthalpies

Lu Li^{a,b}, Hong-jun Fan^{b*}, Hao-quan Hu^{a*}

a. State Key Laboratory of Fine Chemicals, Institute of Coal Chemical Engineering, School of Chemical Engineering, Dalian University of Technology, Dalian 116024, China

b. State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China

(Dated: Received on December 30, 2015; Accepted on January 25, 2016)

The density functional theory (DFT) is the most popular method for evaluating bond dissociation enthalpies (BDEs) of most molecules. Thus, we are committed to looking for alternative methods that can balance the computational cost and higher precision to the best for large systems. The performance of DFT, double-hybrid DFT, and high-level composite methods are examined. The tested sets contain monocyclic and polycyclic aromatic molecules, branched hydrocarbons, small inorganic molecules, *etc.* The results show that the mPW2PLYP and G4MP2 methods achieve reasonable agreement with the benchmark values for most tested molecules, and the mean absolute deviations are 2.43 and 1.96 kcal/mol after excluding the BDEs of branched hydrocarbons. We recommend the G4MP2 is the most appropriate method for small systems (atoms number ≤ 20); the double-hybrid DFT methods are advised for large aromatic molecules in medium size ($20 \leq$ atoms number ≤ 50), and the double-hybrid DFT methods with empirical dispersion correction are recommended for long-chain and branched hydrocarbons in the same size scope; the DFT methods are advised to apply for large systems (atoms number ≥ 50), and the M06-2X and B3P86 methods are also favorable. Moreover, the differences of optimized geometry of different methods are discussed and the effects of basis sets for various methods are investigated.

Key words: Bond dissociation enthalpies, Density functional theory, Double-hybrid density functional theory, High-level composite methods

I. INTRODUCTION

The generating and breaking of bonds is the basis of all chemical reactions. The bond dissociation enthalpies (BDEs) of a chemical bond, which measures the bonding strength plays an important role in determining the reactivity. Therefore, it is desirable to make a responsible knowledge of the energies required to break bonds and the energies released upon their formation. Unfortunately, it is too difficult to obtain the formation enthalpies of some large compounds and radicals [1–3], the number of experimental BDEs is quite limited.

Theoretical computation offers an alternative approach to obtain BDEs [4–13]. Indeed, since the early studies in computational chemistry, a lot of researchers in various fields have used the theoretical methods either to support their experimental results or to estimate unknown BDEs value. Nevertheless, different levels of theories are very different from each other. Schwabe and Grimme [14] compared the per-

formance of the BLYP, TPSS, B3LYP, B2PLYP and mPW2PLYP methods for the heats of formation (HOF) in the G3/05 set [15], they presented that the B2PLYP and the mPW2PLYP gave by far the lowest mean absolute deviation (MAD) over the whole G3/05 set (2.5 and 2.1 kcal/mol, respectively). Accordingly, they expanded their initial study on the G2 set [16] by 271 HOF, 105 ionization potentials, 63 electron affinities, 10 proton affinities, and 6 binding energies of hydrogen-bridged complexes, applying the B3LYP, B2PLYP, and mPW2PLYP methods to the full G3/05 test set for further validation of their performance. Notably, the test set contained many large molecules and heavy atoms up to Kr. Their analyses also revealed that the best performance of mPW2PLYP compared to other studied methods.

Chan and Radom did a comprehensive investigation to search the theoretical procedures that are both adequately accurate but less demanding on computational resources [17]. They concluded that the W1w BDEs generally showed very good agreement with experimental values, but also revealed large discrepancies in a number of cases. Then they had further refined their theoretical values at the W1w+T(Q), and W2w level.

* Authors to whom correspondence should be addressed. E-mail: fanhj@dicp.ac.cn, hhu@dlut.edu.cn

These higher-level calculations yielded BDEs that were consistent with W1w in all cases. They also found that double-hybrid DFT procedures generally give smaller overall derivations for absolute BDEs than those obtained from typical DFT procedures. Comparison of the performance of hybrid DFTs, the M06-2X method emerges as the overall best performer.

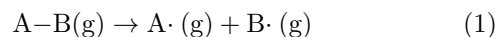
According to the findings of other researchers, HF and MP2 are often less reliable because of the spin contamination in dealing with radical species [18–20]. By comparison, high-level precision methods such as the W1w, CCSD(T), CBS series, and Gaussian-n series. Provide an excellent estimation of BDE values [15, 17, 21–26], however, these methods are expensive and limited to applying on very small systems. For a lot applications in the field of energy such as coal, petroleum, and biomass *etc.*, the thermal reactions in these processes generally involve larger systems. The large monocyclic aromatic molecules and larger polyaromatic molecules are of high interest in these larger systems. The monocyclic aromatic molecules are representative of the functionalities existing in coal, and understanding the thermochemistry and reactivity knowledge of these monocyclic aromatic molecules is useful to better understand the reactive behavior of complex molecules in the coal, which may lead to advances in coal processing. The polycyclic aromatic hydrocarbons (PAHs) have attracted increasing attention in recent years [27–29]. They can be used as model representatives to examine the elementary reactions for the growth of coke layers in coal and petroleum processing [30–32]. The PAHs are also the key elements within incomplete combustion processes, and found to form the largest class of known carcinogens and mutagens [33]. It is still challenging to understand the thermodynamic characteristics of PAHs. Previous work on BDEs of C–H and C–C bonds of PAHs and the effect of polyaromatic environment on the BDEs generally used DFT methods [34–37]. Therefore, it is necessary to find a method that can balance accurate results with computational economy to the greatest extent, especially for relatively large systems including the monocyclic aromatic molecules and PAHs which are the most common compounds appeared in the processing of coal, petroleum and biomass.

Based on these several suitable methods, in the present work, we screened a number of methods as potential candidates and singled out the B3LYP [38], M06-2X [39], mPW2PLYP [12], mPW2PLYPD [40], B2PLYP [12], B2PLYPD [40], G4MP2 [41], and CCSD(T) [42, 43] to do systematic investigation.

II. COMPUTATIONAL DETAILS

The BDEs are defined as the enthalpy of the following reaction required to break the bond A–B to form two

radicals at 298.15 K and 1 atm in the gas phase:



The BDE value can be estimated from Eq.(2) [34]:

$$BDE(A-B) = [H_{298}(A\cdot) + H_{298}(B\cdot)] - H_{298}(A-B) \quad (2)$$

The enthalpy of each species can be calculated from the following equation [44]:

$$H(T) = E + ZPE + H_{\text{trans}} + H_{\text{rot}} + H_{\text{vib}} + RT \quad (3)$$

where E is electronic energy, ZPE is the zero point energy, H_{trans} , H_{rot} , and H_{vib} are the standard temperature correction terms calculated with the equilibrium statistical mechanics with harmonic oscillator and rigid rotor approximations.







In our work, calculations are all carried out with GAUSSIAN 09 [45] packages. The geometries of reactants and resultant radical species are optimized on the X/cc-pVDZ [46] level, where the X is the selected studied methods, including B3LYP, M06-2X, mPW2PLYP, and B2PLYP. The double hybrid methods combine exact HF exchange with an MP2-like correlation to a DFT calculation, which have the same computational cost as MP2 and good accuracy. The minimum energy structure can be verified and the thermal contributions can be obtained by frequency calculations at the same level. Single point energy calculations are conducted at the X/cc-pVTZ [47] level (X represents the corresponding method as introduced above). In addition, single point energies calculations by mPW2PLYPD at cc-pVDZ, cc-pVTZ level and CCSD(T) at cc-pVDZ, cc-pVTZ, and cc-pVQZ [48] level all start with mPW2PLYP/cc-pVDZ geometry optimization and then are corrected the thermochemical data by using mPW2PLYP/cc-pVDZ frequency calculation. Similarly, the single point energies of B2PLYPD method are calculated by using the optimized geometry at B2PLYP/cc-pVDZ level and the thermochemical data are also corrected.

III. RESULTS AND DISCUSSION

A. Evaluation of different methods for the small molecules

In Chan and Radom's work, they focused on the nonaromatic compounds with fewer heavy atoms [17], and our previous work contained several calculated BDEs showing that the mPW2PLYP gave excellent performance on evaluating the BDEs of monocyclic aromatic molecules compared to other methods [49], we now examine the performance of various methods for the evaluation of BDEs on more monocyclic aromatic molecules. Several common non-aromatic organic molecules are included to compare together to form a preliminary assessment for the performance of different

TABLE I Experimental BDEs and calculated BDEs by different methods for small molecules (kcal/mol).

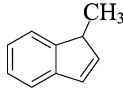
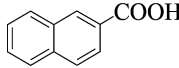
Bonds ^a	B3 LYP	M06- 2X	mPW2 PLYP	mPW2 PLYPD	B2 PLYP	B2 PLYPD	G4 MP2	CCSD (T)	Expt.
Ph-H	110.6	111.2	111.7	112.0	111.5	111.9	113.0	113.7	112.9 [2]
<i>o</i> -NC ₅ H ₄ -H	104.9	106.3	106.0	106.2	105.6	105.9	107.7	108.7	105 [36]
<i>p</i> -NC ₅ H ₄ -H	111.0	112.3	112.5	112.7	112.1	112.5	114.1	114.6	112 [36]
<i>m</i> -NC ₅ H ₄ -H	109.7	110.6	111.0	111.3	110.7	111.1	112.4	113.2	112 [36]
	104.9	107.0	106.3	106.6	105.9	106.2	109.3	107.1	103 [36]
	106.4	107.3	107.4	107.6	107.1	107.3	108.5	110.0	98 [36]
	104.0	105.7	105.3	105.5	104.8	105.1	106.3	107.9	103 [36]
	112.0	113.6	113.4	113.7	113.0	113.3	114.7	115.6	112 [36]
	118.3	118.2	119.1	119.3	118.8	119.2	118.2	119.0	112.3 [36]
	117.8	117.8	118.6	118.8	118.3	118.6	117.7	118.7	112.3 [36]
PhCH ₂ -H	87.4	90.6	90.2	90.4	90.3	90.6	90.8	91.7	89.8 [2]
Ph-CH ₃	96.7	103.1	101.2	102.2	100.7	102.2	104.3	104.5	101.8 [35]
Ph(CH ₃)CH-H	83.8	86.9	87.0	87.4	87.1	87.6	87.7	89.3	86.9 [56]
PhCH ₂ -CH ₃	69.4	78.2	75.7	76.8	75.5	77.1	78.2	78.8	77.6 [2]
Ph-CH ₂ CH ₃	92.5	100.3	97.5	100.4	98.0	100.5	102.8	103.3	101.5 [63]
Ph-OCH ₃	92.9	102.8	99.1	99.6	98.8	100.8	101.6	103.0	101.0 [2]
PhO-CH ₃	57.9	69.2	64.7	65.7	64.4	65.9	65.8	66.1	63.8 [7]
PhO-CH ₂ CH ₃	58.3	70.6	66.3	67.8	66.0	68.0	67.9	68.7	63.1 [7]
<i>o</i> -C ₅ H ₄ NCH ₂ -H	101.8	103.5	102.3	102.7	102.2	102.6	101.6	104.7	96.0 [36]
PhCH=CHCH ₂ -CH ₃	62.0	71.1	68.5	69.4	68.4	69.7	70.9	72.6	67.5 [7]
CH ₂ =CHPhCH ₂ -CH ₃	67.5	77.0	74.9	76.0	74.8	76.4	76.8	78.4	71.4 [7]
PhCH ₂ -N(CH ₃) ₂	55.0	69.5	64.5	67.0	64.2	67.7	68.2	69.8	61.7 [7]
PhCH ₂ -NHCH ₃	60.4	71.9	68.0	69.6	67.9	70.0	70.7	71.6	68.7 [7]
Ph(CH ₃)CH-COOCH ₃	68.1	80.1	77.2	79.3	76.3	79.2	79.4		77.1 [44]
CH ₃ -H	103.5	104.1	103.3	103.4	103.2	103.3	104.6	103.3	104.8 [63]
CH ₃ -CH ₃	85.0	90.7	87.4	87.8	87.1	87.7	89.5	88.3	90.1 [2]
CH ₃ CH ₂ -H	98.8	100.2	99.3	99.5	99.2	99.5	100.7	100.1	100.5 [63]
CH ₃ -OH	85.8	93.1	89.0	89.2	89.1	89.4	91.3	88.9	92.1 [2]
CH ₃ -OCH ₃	76.0	85.1	80.2	81.0	80.1	81.1	83.0	81.8	83.2 [2]
CH ₃ -CH ₂ CH ₃	88.8	88.6	85.6	86.5	85.2	86.5	88.4	87.6	88.2 [63]
CH ₃ -CH=CH ₂	95.3	101.0	98.8	99.5	98.4	99.3	100.6	99.8	101.4 [2]
CH ₃ CH ₂ -CH=CH ₂	91.8	98.7	96.7	97.8	96.2	97.7	99.1	98.8	100.2 [63]
CH ₃ CH ₂ -CH ₂ CH=CH ₂	63.8	73.8	69.9	71.4	69.5	71.6	74.0	73.7	74.7 [63]
CH ₂ CH=CH ₂ -CH ₂ CH=CH ₂	48.8	60.7	56.0	57.4	55.9	57.9	60.6	60.5	61.4 [63]
MAD	4.98	2.49	2.47	2.45	2.49	2.41	2.23	3.11	
LD	-12.6	7.8	6.8	7.0	6.5	6.9	6.5	8.7	

^a The bond homolysis site is marked with a dash “-”.

methods. Most of those molecules are not only important chemical raw material, but they are also key factors within petroleum distillate catalytic cracking process. The test set contains 26 parent compounds with small size including monocyclic aromatic molecules and non-aromatic organic molecules, which are no more than 10 heavy atoms. The produced 34 BDEs after related homolytic bond cleavage are summarized in Table I.

An exception can be found for pyrimidine of C2-H homolytic bond cleavage. It's clear to see that not only B3LYP but also high level calculation G4MP2 and even CCSD(T) all can't achieve agreeable value. Therefore, the experimental value should be problematic. Obviously, all of the methods yield reasonable values except B3LYP, after excluding the pyrimidine of C2-H. The B3LYP yields the largest MAD of 4.98 kcal/mol. In

TABLE II Experimental BDEs and calculated BDEs by different methods for large aromatic hydrocarbon compounds (kcal/mol).

Bonds	B3 LYP	M06- 2X	mPW2 PLYP	mPW2 PLYPD	B2 PLYP	B2 PLYPD	G4 MP2	B3 LYP[r]	B3 P86[r]	Expt.
Ph-Ph	106.5	113.7	115.4	117.6	114.8	117.8	119.8	108.0 [35]	112.9 [35]	113.7 [35]
PhCH ₂ -Ph	79.3	90.0	88.4	90.6	88.1	91.1	92.8	79.6 [35]	84.8 [35]	89.6 [35]
PhCH ₂ -CH ₂ Ph	52.9	67.5	64.2	66.9	64.2	67.8	68.6	54.1 [35]	59.5 [35]	61.4 [35]
PhCH ₂ CH ₂ -CH ₂ Ph	66.1	77.7	74.6	77.1	74.3	77.7	76.7	66.8 [35]	71.6 [35]	73.9 [35]
Ph(CH ₃)CHCH(Ph)-CH ₃	63.9	76.5	72.5	75.6	72.1	76.2	75.3	60.2 [35]	65.4 [35]	67.6 [35]
Ph ₂ CH-CH ₂ Ph	42.3	59.0	56.4	60.0	56.4	61.3	63.2	43.4 [35]	49.3 [35]	57.8 [35]
Ph ₂ CH-CH ₃	59.8	71.9	68.8	70.7	67.7	70.4	73.6			65.7 [64]
Ph ₂ C(CH ₃)-CH ₃	58.1	71.8	67.8	70.5	67.4	71.0	73.2			67.6 [64]
	59.7	69.4	65.5	67.0	65.1	67.2	67.8			70.0 [64]
Ph ₂ CH-H	78.5	83.2	82.9	83.2	83.0	83.4	84.5	78.7 [7]	81.6 [7]	82.7 [7]
Ph ₂ CH ₃ C-H	78.5	82.5	82.7	83.1	81.8	82.4	83.9	78.7 [7]	81.0 [7]	81.0 [7]
	99.6	107.7	105.6	106.8	104.5	106.2	102.6			104.8 [44]
MAD	7.55	3.03	1.93	3.28	1.68	3.53	4.58	7.28	2.70	
LD	-15.5	8.9	4.9	8.0	-4.9	8.6	7.9	-14.4	-8.5	

addition, it also produces the largest deviation (LD) of -12.6 kcal/mol in the set. The rest methods give more reasonable results with the MADs range from 2 kcal/mol to 4 kcal/mol. The G4MP2 achieves the smallest MAD of 2.23 kcal/mol, and it is slightly better than the double-hybrid methods mPW2PLYP and B2PLYP, with the value of MADs are 2.47 and 2.49 kcal/mol. Adding the empirical dispersion correction (specified by "D") to mPW2PLYP and B2PLYP give insignificant change, which slightly improved their performance, with MADs of 2.45, 2.41 kcal/mol, respectively. The performance of M06-2X is comparable to the double-hybrid methods, with the MAD of 2.49 kcal/mol. The CCSD(T) shows little larger but still acceptable MAD of 3.11 kcal/mol.

B. Evaluation of different methods for larger aromatic hydrocarbon compounds

In addition to these results, the performance of different methods for more aromatic hydrocarbon compounds at larger size are investigated and shown in Table II, including compounds containing no less than two benzene rings or condensed aromatic rings, and these molecules are generally considered as coal model compounds to represent the specific units in coal structure. The referred results are listed to assess various methods. The computed data of B3LYP[r] and B3P86[r] from previous studies of Li *et al.* [7, 35], proposed that B3LYP gave the largest MAD, and the B3P86 method emerges as the overall best performer. Our results show that the double-hybrid methods mPW2PLYP and B2PLYP give smaller MAD of

1.93 and 1.68 kcal/mol for the studied compounds which have relatively large size with heavy atom number ≥ 10 . The mPW2PLYP and B2PLYP with empirical dispersion correction and G4MP2 methods are generally associated with overestimation of these large compounds, with MADs of 3.28, 3.53, 4.58 kcal/mol, respectively. The tested DFT methods perform worse for the evaluation of BDEs than the double-hybrid methods as expected. The B3LYP method gives the largest MAD of 7.55 kcal/mol, and is consistent with the literature results with MAD of 7.28 kcal/mol [7, 35]. The value of MAD for M06-2X is 3.03 kcal/mol. The referenced results of B3P86 show little larger dispersions but still acceptable MAD of 2.70 kcal/mol [7, 35].

C. BDEs of particular methods for branched hydrocarbons

On the basis of the results in Table I and Table II, the double-hybrid methods and the G4MP2 give relatively better performance. The mPW2PLYP, mPW2PLYPD, and the G4MP2 are tested to investigate their performance on evaluating the BDEs of other kinds of compounds. The branched hydrocarbons are common components in the processing of coal and petroleum, which are examined here since they have been extensively interested [50, 51], and the results are summarized in Table III. In addition, due to the favorable performance of B3P86 method for evaluating BDEs of large aromatic compounds, we also calculate B3P86 BDEs for comparison. The BDEs of B3P86 are obtained using the same equations with other methods (*e.g.* B3LYP) as introduced in computational details. It is worthy to note that G4MP2 provides an excellent approximation

TABLE III Experimental BDEs and calculated BDEs by different methods for the branched hydrocarbons (kcal/mol).

Bonds	mPW2PLYP	mPW2PLYPD	G4MP2	B3P86	Expt.
(CH ₃) ₂ CH–CH=CH ₂	94.8	96.5	98.1	92.8	99.9 [63]
(CH ₃) ₂ CH–CH ₂ CH=CH ₂	67.9	70.3	73.3	65.0	75.2 [2]
(CH ₃) ₂ CH–CH(CH ₃) ₂	77.6	81.2	83.9	74.4	85.6 [2]
(CH ₃) ₃ C–CH=CH ₂	92.2	94.8	96.9	89.2	97.8 [2]
(CH ₃) ₃ C–CH ₂ CH=CH ₂	65.4	68.8	72.2	61.3	73.2 [2]
(CH ₃) ₃ C–CH(CH ₃) ₂	73.9	78.7	81.9	69.6	82.7 [2]
(CH ₃) ₃ C–C(CH ₃) ₃	69.0	75.2	78.7	63.5	78.6 [2]
MAD	7.46	3.93	1.17	11.03	
LD	–9.6	–4.9	–1.9	–15.1	

to the experimental BDEs for branched hydrocarbons, with an MAD of only 1.17 kcal/mol. The mPW2PLYP gives large discrepancies between experiment and the theoretical values with an MAD of 7.46 kcal/mol. Using a dispersion correction lead to smaller MAD, of 3.93 kcal/mol. The hybrid DFT procedure B3P86 produces the largest MAD of 11.03 kcal/mol, and the maximum discrepancy is up to –15.1 kcal/mol.

D. Structural optimized geometry comparison of various methods

The comparison of structural optimized geometry at different levels is summarized in Table IV. The B2PLYP method shows the smallest deviation compared with the mPW2PLYP method. The G4MP2 gives the largest deviation of bond lengths, and the M06-2X gives the largest deviation of bond angle.

E. Effect of basis sets on mPW2PLYP and CCSD(T) calculation

The effects of different basis sets on BDEs calculations are also investigated, the results are shown in Table V. Due to the size limitation of computing system at CCSD(T)/cc-pVQZ level, the values we don't have are omitted. From Table V, it is clear that the BDEs of mPW2PLYP change from a relatively small basis set (cc-pVDZ) to a extended one (cc-pVTZ), and the medium basis set (cc-pVTZ) to a larger one (cc-pVQZ) cause variations of 2.10 and 0.46 kcal/mol, respectively. In comparison, the same change of basis set results in variations of as large as 4.35 and 2.46 kcal/mol for the CCSD(T) BDEs calculations. Therefore, the CCSD(T) method is fairly sensitive toward the basis sets in the calculation of BDEs, which means that it is necessary to employ large enough basis set (*e.g.* cc-pVQZ) when using CCSD(T) BDEs calculations as the benchmark. O'Reilly *et al.* [52] investigated the BDEs of 31 N–H and 31 N–Cl bonds by a large variety of contemporary methods, their results show that changing the size of the basis sets cause different energy variations of different DFT methods, the BDEs of DFT methods change from A'VDZ to A'VTZ cause energy variations

TABLE IV Structural optimized geometry comparison with selected methods for bond lengths and bond angle (MADs relative to mPW2PLYP values).

B2PLYP	Length/Å	Angle/(°)
B3LYP	0.005	0.20
M06-2X	0.006	0.25
B2PLYP	0.003	0.03
G4MP2	0.012	0.15







of ~1 kcal/mol, and the A'VTZ to A'VQZ cause energy changes within 0.5 kcal/mol. Thus, it can be concluded that the cc-pVTZ basis set can be used to reliably predict the BDEs for most methods.

F. BDEs of particular methods for the extended compounds

Despite the importance of the BDEs and lots of researches on it. Most of these studies focused on hydrocarbons which produced carbon-centered radicals even if several heteroatoms can be found in the molecule, and numerous C–H BDEs and C–C BDEs have been reported [53–64]. In this work, we are interested to verify the credibility of selected computing methods on a broader range. Various typical inorganic compounds have been chosen and the BDEs of non single bonds are calculated. Due to lacking of enough experiment values, in our further studies we will use CCSD(T) as the benchmark theoretical method which is generally considered to be the most accurate method and cc-pVQZ is selected because it is affected obviously by the size of the basis set. The results are summarized in Table VI. The G4MP2 and mPW2PLYPD can not be used for some certain elements, therefore the BDEs of some molecules cannot be computed.

Overall, the G4MP2 show the best performance, which is consistent with our former study on the small molecule compounds, showing an MAD of 1.51 kcal/mol to the selected benchmark. The results using other two methods mPW2PLYP and mPW2PLYPD are reasonably close to the benchmark values for the majority of investigated species, with MADs being 2.67,

TABLE V Comparison of basis sets for mPW2PLYP and CCSD(T) calculation of BDEs (kcal/mol).

Bonds	mPW2PLYP/ cc-pVDZ	mPW2PLYP/ cc-pVTZ	mPW2PLYP/ cc-pVQZ	CCSD(T)/ cc-pVDZ	CCSD(T)/ cc-pVTZ	CCSD(T)/ cc-pVQZ
Ph-H	110.0	111.7	112.0	109.9	113.7	
<i>o</i> -NC ₄ H ₄ -H	104.3	106.0	106.2	105.0	108.7	
<i>p</i> -NC ₄ H ₄ -H	110.7	112.5	112.7	110.7	114.6	
<i>m</i> -NC ₄ H ₄ -H	109.2	111.0	111.3	109.4	113.2	
	104.7	106.3	106.6	103.7	107.1	
	105.5	107.4	107.7	106.1	110.0	
	103.5	105.3	105.5	104.2	107.9	
	111.6	113.4	113.7	111.7	115.6	
	116.9	119.1	119.4	114.9	119.0	
	116.5	118.6	118.9	114.6	118.7	
PhCH ₂ -H	88.9	90.2	90.3	89.0	91.7	
Ph(CH ₃)CH-H	85.8	87.0	87.2	86.8	89.3	
PhCH ₂ -CH ₃	77.2	75.7	75.7	78.0	78.8	
<i>o</i> -C ₅ H ₄ NCH ₂ -H	100.7	102.3	102.5	101.1	104.7	
CH ₃ -CH ₃	88.7	87.4	87.4	87.0	88.3	89.1
CH ₃ -OH	88.3	89.0	89.5	84.6	88.9	90.6
CH ₃ -OCH ₃	80.0	80.2	80.5	78.3	81.8	83.1
FHN-F	58.6	62.4	63.3	53.7	62.1	64.5
F ₂ N-Cl	26.8	30.5	31.6	24.8	31.7	34.4
MeO-SH	47.4	52.4	53.7	44.9	53.2	56.0
H ₃ Si-PH ₂	64.7	67.1	67.9	62.0	67.6	69.8
MeHP-H	76.7	78.8	79.4	75.3	79.1	80.4
H ₂ P-PH ₂	50.8	53.1	53.9	47.5	53.2	55.6
MeS-Cl	55.1	60.3	61.9	50.9	59.7	63.3

2.55 kcal/mol, respectively. The LDs for mPW2PLYP, mPW2PLYPD and G4MP2 are very close, with the values of 5.8, 5.8, and 5.2 kcal/mol, respectively. Notably, the largest discrepancies are all found for BDEs of HC≡N. In addition, G4MP2 method mainly overestimate the BDEs and large deviations are observed for CH₂=O, HC≡N, and CH₂=S, which are all non single bond species. It can also be noted that the dispersion-corrected procedure mPW2PLYPD give slightly better performance than the corresponding method mPW2PLYP, but the effect is almost negligible.

IV. CONCLUSION

Systematic assessment of the accuracy of quantum chemistry methods is an essential prerequisite for their routine use on predicting molecule thermochemistry. In this work, the performance of a variety of contemporary theoretical procedures on calculating the BDEs of the

selected species is assessed. The final MAD values for all examined compounds by various tested methods are summarized in Table VII. The following key observations emerge from the present study:

(i) G4MP2 generally gives the best agreement with the experiment values for small molecule compounds, especially for the branched hydrocarbons. Among the large aromatic compounds examined in this study, G4MP2 procedure perform less well than the double-hybrid DFT methods mPW2PLYP and B2PLYP. In addition, G4MP2 is generally associated with overestimation of these large aromatic compounds.

(ii) Double-hybrid DFT methods including mPW2PLYP and B2PLYP give reasonably close to the experimental values or the benchmark for the major investigated species (whether for small systems, or relatively large systems), except for the branched hydrocarbons. The mPW2PLYP performs comparably to B2PLYP for the test sets under study. The basis

TABLE VI Calculated BDEs of the molecules at different levels (kcal/mol).

Bonds	mPW2 PLYP	mPW2D PLYP	G4MP2	CCSD(T)/ cc-pVQZ	Bonds	mPW2 PLYP	mPW2D PLYP	G4MP2	CCSD(T)/ cc-pVQZ
Na-Cl	93.5	93.6	97.9	97.2	CH ₃ -F	107.4	107.5	110.0	108.9
H-Cl	100.5	100.5	103.4	102.5	FHN-F	62.4	62.5	65.2	64.5
H-F	131.4	131.4	136.6	134.9	CH ₃ -Cl	80.3	80.4	83.8	82.5
H-OH	114.7	114.7	118.2	117.4	CH ₃ S-Cl	60.3	60.7	66.4	63.3
HS-H	88.5	88.6	91.4	90.3	F ₂ N-Cl	30.5	30.6	35.1	34.4
HSe-H	79.9	80.0	81.4	82.4	C ₂ H ₃ -F	122.4	122.6	123.2	123.2
NH ₂ -H	104.6	104.6	106.8	106.1	C ₂ H ₃ -Cl	91.7	91.9	94.8	94.0
PH ₂ -H	80.2	80.4	82.2	81.5	CH ₂ =CH ₂	172.0	172.2	175.3	172.2
Cl-Cl	53.1	53.1	58.5	55.7	HC≡CH	228.7	228.9	228.0	226.3
F-F	33.0	33.0	36.8	36.1	CH ₂ =O	175.9	176.0	179.9	176.4
HO-OH	46.9	47.0	48.1	48.3	HC≡N	242.4	242.4	241.8	236.6
NH ₂ -NH ₂	62.8	63.2	64.8	64.1	CH ₂ =S	126.5	126.6	133.7	128.6
NH ₂ NH-H	79.4	79.5	82.3	82.0	CH ₃ -I	59.7	59.9		64.0
CH ₃ NH-H	96.8	97.0	99.0	98.8	Na-I	75.3	75.7		80.7
CH ₃ -NH ₂	81.5	81.8	83.8	83.1	I-I	43.6	43.6		48.3
CH ₃ -NO ₂	56.3	56.8	59.8	60.0	PbH-H	58.5	58.6		62.2
PH ₂ -PH ₂	53.1	53.7	57.1	55.7	H-BH ₂ NH ₃	100.8	101.2	100.8	101.1
PH ₂ PH-H	76.2	76.5	78.3	77.8	BH ₃ NH ₂ -H	91.5	91.8	94.5	95.0
CH ₃ PH-H	78.8	79.1	80.5	80.4	CaCl-Cl	113.6	113.7	115.2	116.0
CH ₃ -PH ₂	67.9	68.4	71.2	70.2	AlCl ₂ -Cl	111.6	112.0	115.3	115.4
H ₃ Si-PH ₂	67.1	67.7	70.3	69.8	MAD	2.54	2.35	0.99	
CH ₃ O-SH	52.4	52.9	57.6	56.0	LD	5.8	5.8	5.2	

TABLE VII The final MAD and LD values for all examined compounds by various tested methods (kcal/mol).

	Branched hydrocarbons		Other compounds	
	MAD	LD	MAD	LD
B3LYP			5.67	-15.5
M06-2X			2.64	8.9
B3P86	11.03	-15.1		
mPW2PLYP	7.46	-9.6	2.43	6.8
mPW2PLYPD	3.93	-4.9	2.52	8.0
B2PLYP			2.52	6.5
B2PLYPD			2.27	8.6
G4MP2	1.17	-1.9	1.96	7.9
CCSD(T)			3.11	8.7

set effects on mPW2PLYP calculation of BDEs are not significant, especially changing the medium basis set (cc-pVTZ) to a extended one (cc-pVQZ) leads to negligible differences.

(iii) The mPW2PLYPD and B2PLYPD with empirical dispersion correction greatly improve the performance of predicting BDEs of branched hydrocarbons compared to the mPW2PLYP and B2PLYP, but show slightly even insignificant change to the small compounds in our study. The larger derivations can be found in calculating BDEs of large aromatic compounds, comparable to those obtained by the corresponding methods without the empirical dispersion correction.

(iv) The CCSD(T) method is fairly sensitive toward

the basis sets in the calculation of BDEs. Thus, it is necessary to choose enough large basis set using the CCSD(T) method to evaluate BDEs.

(v) Among the studied hybrid DFT methods, namely, B3LYP, M06-2X, B3P86, for the evaluation of BDEs, the M06-2X and B3P86 methods provide acceptable performance for the majority studied systems, even though they are not as good as double-hybrid DFTs methods. Large discrepancies can be found in the calculating of B3P86 for branched hydrocarbons.

Taken together with the results of our work, the G4MP2 and double-hybrid DFT methods give satisfactory performance on evaluating BDEs for majority examined compounds. For small systems (atoms number ≤ 20), including monocyclic aromatic molecules, non-aromatic organic molecules and inorganic molecules, the G4MP2 and double-hybrid DFT methods all give reasonable BDEs, and the G4MP2 performs a little better than double-hybrid DFT methods. Thus, we recommend choosing the G4MP2 method for small molecules. For medium systems ($20 \leq$ atoms number ≤ 50), that the scope include the most common compounds which involved in the processing of various raw materials of fuel. The double-hybrid DFT methods mPW2PLYP and B2PLYP are advised for large aromatic molecules, and the thermodynamic characteristics of these PAHs is significant for understanding the processing of the coal, petroleum and biomass. The mPW2PLYPD and B2PLYPD with empirical dispersion correction are recommended for long-chain and branched hydrocarbons. For large systems (atoms

number ≥ 50), DFT methods are the most appropriate solution. The M06-2X and B3P86 methods are suggested to apply for the calculation of large molecules.

V. ACKNOWLEDGMENTS

This work was supported by the National Basic Research Program of China (No.2011CB201301), the Key Program Project of Joint Fund of Coal Research, the National Natural Science Foundation of China and Shenhua Group (No.51134014), DICP DMTO201404, and Key International ST Cooperation and Exchange Projects (No.2013DFG60060).

- [1] O. Dorofeeva, V. P. Novikov, and D. B. Neumann, *J. Phys. Chem. Ref. Data* **30**, 475 (2001).
- [2] S. J. Blanksby and G. B. Ellison, *Acc. Chem. Res.* **36**, 255 (2003).
- [3] J. Berkowitz, G. B. Ellison, and D. Gutman, *J. Phys. Chem.* **98**, 2744 (1994).
- [4] W. Zheng, W. Xu, Y. Wang, and Z. Chen, *Comp. Theor. Chem.* **1027**, 116 (2014).
- [5] A. Vagánek, J. Rimarčík, M. Ilčín, P. Škorňa, V. Lukeš, and E. Klein, *Comp. Theor. Chem.* **1014**, 60 (2013).
- [6] E. R. Johnson, O. J. Clarkin, and G. A. DiLabio, *J. Phys. Chem. A* **107**, 9953 (2003).
- [7] X. Q. Yao, X. J. Hou, H. Jiao, H. W. Xiang, and Y. W. Li, *J. Phys. Chem. A* **107**, 9991 (2003).
- [8] I. I. Marochkin and O. V. Dorofeeva, *Comp. Theor. Chem.* **991**, 182 (2012).
- [9] G. P. F. Wood, D. Moran, R. Jacob, and L. Radom, *J. Phys. Chem. A* **109**, 6318 (2005).
- [10] H. Z. Yu, F. Fu, L. Zhang, Y. Fu, Z. M. Dang, and J. Shi, *Phys. Chem. Chem. Phys.* **16**, 20964 (2014).
- [11] W. R. Zheng, J. L. Xu, T. Huang, Z. C. Chen, and Q. Yang, *Comp. Theor. Chem.* **968**, 1 (2011).
- [12] W. R. Zheng, Z. C. Chen, and W. X. Xu, *Chin. J. Chem. Phys.* **26**, 541 (2013).
- [13] B. Wang, Y. Fu, H. Z. Yu, and J. Shi, *Chin. J. Chem. Phys.* **27**, 640 (2014).
- [14] T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys.* **8**, 4398 (2006).
- [15] L. A. Curtiss, P. C. Redfern, and K. Raghavachari, *J. Chem. Phys.* **123**, 124107 (2005).
- [16] S. Grimme, *J. Chem. Phys.* **124**, 034108 (2006).
- [17] B. Chan and L. Radom, *J. Phys. Chem. A* **116**, 4975 (2012).
- [18] D. J. Henry, C. J. Parkinson, P. M. Mayer, and L. Radom, *J. Phys. Chem. A* **105**, 6750 (2001).
- [19] K. S. Song, Y. H. Cheng, Y. Fu, L. Liu, X. S. Li, and Q. X. Guo, *J. Phys. Chem. A* **106**, 6651 (2002).
- [20] J. Cioslowski, G. Liu, M. Martinov, P. Piskorz, and D. Moncrieff, *J. Am. Chem. Soc.* **118**, 5261 (1996).
- [21] L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Ras-solov, and J. A. Pople, *J. Chem. Phys.* **109**, 7764 (1998).
- [22] J. W. Ochterski, G. A. Petersson, and K. B. Wiberg, *J. Am. Chem. Soc.* **117**, 11299 (1995).
- [23] A. G. Baboul, L. A. Curtiss, P. C. Redfern, and K. Raghavachari, *J. Chem. Phys.* **110**, 7650 (1999).
- [24] S. W. Zhao, L. Liu, Y. Fu, and Q. X. Guo, *J. Phys. Org. Chem.* **18**, 353 (2005).
- [25] L. A. Curtiss, P. C. Redfern, and K. Raghavachari, *J. Chem. Phys.* **126**, 084108 (2007).
- [26] A. S. Menon, G. P. F. Wood, D. Moran, and L. Radom, *J. Phys. Chem. A* **111**, 13638 (2007).
- [27] A. K. Haritash and C. P. Kaushik, *J. Hazard. Mater* **169**, 1 (2009).
- [28] S. K. Samanta, O. V. Singh, and R. K. Jain, *Trends Biotechnol.* **20**, 243 (2002).
- [29] W. Wilcke, *J. Soil Sci. Plant Nut.* **163**, 229 (2000).
- [30] S. Wauters and G. B. Marin, *Chem. Eng. J.* **82**, 267 (2001).
- [31] K. Hemelsoet, V. Van Speybroeck, D. Moran, G. B. Marin, L. Radom, and M. Waroquier, *J. Phys. Chem. A* **110**, 13624 (2006).
- [32] V. Van Speybroeck, K. Hemelsoet, B. Minner, G. B. Marin, and M. Waroquier, *Mol. Simul.* **33**, 879 (2007).
- [33] M. F. Denissenko, A. Pao, M. S. Tang, and G. P. Pfeifer, *Science* **274**, 430 (1996).
- [34] K. Hemelsoet, V. V. Speybroeck, and M. Waroquier, *J. Phys. Chem. A* **112**, 13566 (2008).
- [35] X. Q. Yao, X. J. Hou, G. S. Wu, Y. Y. Xu, H. W. Xiang, H. Jiao, and Y. W. Li, *J. Phys. Chem. A* **106**, 7184 (2002).
- [36] C. Barckholtz, T. A. Barckholtz, and C. M. Hadad, *J. Am. Chem. Soc.* **121**, 491 (1999).
- [37] V. Van Speybroeck, G. B. Marin, and M. Waroquier, *ChemPhysChem* **7**, 2205 (2006).
- [38] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).
- [39] Y. Zhao and D. Truhlar, *Theor. Chem. Account.* **120**, 215 (2008).
- [40] T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys.* **9**, 3397 (2007).
- [41] L. A. Curtiss, P. C. Redfern, and K. Raghavachari, *J. Chem. Phys.* **127**, 124105 (2007).
- [42] G. E. Scuseria, C. L. Janssen, and H. F. Schaefer, *J. Chem. Phys.* **89**, 7382 (1988).
- [43] J. A. Pople, M. Headgordon, and K. Raghavachari, *J. Chem. Phys.* **87**, 5968 (1987).
- [44] J. Shi, X. Y. Huang, J. P. Wang, and R. Li, *J. Phys. Chem. A* **114**, 6263 (2010).
- [45] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J.

- Cioslowski, and D. J. Fox, *Gaussian 09, Revision A.02*, Wallingford CT: Gaussian Inc., (2009).
- [46] J. T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- [47] R. A. Kendall, J. T. H. Dunning, and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- [48] D. E. Woon and J. T. H. Dunning, *J. Chem. Phys.* **98**, 1358 (1993).
- [49] L. Li, H. J. Fan, and H. Q. Hu, *Fuel* **153**, 70 (2015).
- [50] W. C. McKee and P. V. Schleyer, *J. Am. Chem. Soc.* **135**, 13008 (2013).
- [51] M. D. Wodrich, W. C. McKee, and P. V. Schleyer, *J. Org. Chem.* **76**, 2439 (2011).
- [52] R. J. O'Reilly, A. Karton, and L. Radom, *Int. J. Quantum. Chem.* **112**, 1862 (2012).
- [53] Y. Feng, L. Liu, J. T. Wang, S. W. Zhao, and Q. X. Guo, *J. Org. Chem.* **69**, 3129 (2004).
- [54] Z. Tian, A. Fattahi, L. Lis, and S. R. Kass, *J. Am. Chem. Soc.* **128**, 17087 (2006).
- [55] E. I. Izgorodina, M. L. Coote, and L. Radom, *J. Phys. Chem. A* **109**, 7558 (2005).
- [56] C. J. Hayes and C. M. Hadad, *J. Phys. Chem. A* **113**, 12370 (2009).
- [57] M. L. Coote, A. Pross, and L. Radom, *Org. Lett.* **5**, 4689 (2003).
- [58] J. P. Senosiain, J. H. Han, C. B. Musgrave, and D. M. Golden, *Faraday Disc.* **119**, 173 (2002).
- [59] D. Kaur, R. P. Kaur, and R. Kaur, *J. Mol. Struct.: THEOCHEM* **803**, 95 (2007).
- [60] K. Hemelsoet, F. Van Durme, V. Van Speybroeck, M. F. Reyniers, and M. Waroquier, *J. Phys. Chem. A* **114**, 2864 (2010).
- [61] P. C. Nam, M. T. Nguyen, and A. K. Chandra, *J. Phys. Chem. A* **109**, 10342 (2005).
- [62] Y. Feng, L. Liu, J. T. Wang, H. Huang, and Q. X. Guo, *J. Chem. Inf. Comput. Sci.* **43**, 2005 (2003).
- [63] A. A. Zavitsas, D. W. Rogers, and N. Matsunaga, *J. Org. Chem.* **75**, 5697 (2010).
- [64] D. A. Robaugh and S. E. Stein, *J. Am. Chem. Soc.* **108**, 3224 (1986).
- [65] K. A. Peterson, *J. Chem. Phys.* **119**, 11099 (2003).
- [66] B. Metz, H. Stoll, and M. Dolg, *J. Chem. Phys.* **113**, 2563 (2000).
- [67] K. A. Peterson, B. C. Shepler, D. Figgen, and H. Stoll, *J. Phys. Chem. A* **110**, 13877 (2006).