Can DFT and ab initio methods describe all aspects of the potential energy surface of cycloreversion reactions?

Li-Juan Yu, Farzaneh Sarrami, Robert J. O'Reilly & Amir Karton

⊞ View supplementary material ⧉

▦ Published online: 15 Sep 2015.

✎ Submit your article to this journal ⧉

🔍 View related articles ⧉

▣ View Crossmark data ⧉

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Can DFT and ab initio methods describe all aspects of the potential energy surface of cycloreversion reactions?

Li-Juan Yu[a], Farzaneh Sarrami[a], Robert J. O'Reilly[b] and Amir Karton [a]

[a]School of Chemistry and Biochemistry, The University of Western Australia, Perth, Australia; [b]Department of Chemistry, School of Science and Technology, Nazarbayev University, Astana, Republic of Kazakhstan
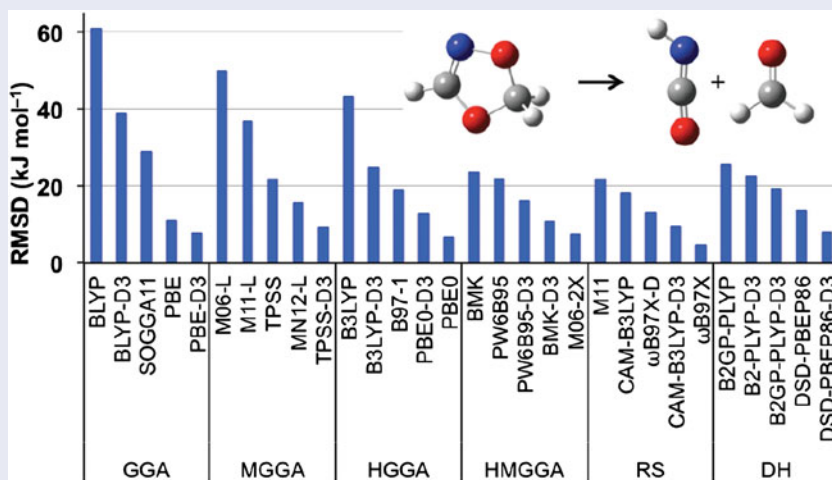
ABSTRACT

We introduce a representative benchmark database of 20 cycloreversion reaction energies obtained by means of the high-level W1 thermochemical protocol. We use these benchmark values to assess the performance of a variety of contemporary DFT, double-hybrid DFT (DHDFT), standard ab initio, and compound thermochemistry methods. We show that this set of reaction energies provides an extremely challenging test for nearly all of the considered DFT and DHDFT methods. For example, about 80% of the considered functionals result in root-mean-square deviations (RMSDs) above 10 kJ mol^{-1}. The best DFT and DHDFT procedures are ωB97X and DSD-PBEP86-D3, with RMSDs of 4.7 and 7.9 kJ mol^{-1}, respectively. Coupled with the fact that the barrier heights for these reactions also pose a significant challenge for many DFT methods, this work shows that only a handful of functionals can quantitatively describe all aspects of the potential energy surface of this important class of reactions. In addition, this work shows that London dispersion effects are particularly large for this class of reactions. For example, empirical D3 dispersion corrections reduce the RMSDs for the DFT and DHDFT procedures by amounts ranging from 3.5 (PBE and B2K-PLYP) to 22.0 (BLYP) kJ mol^{-1}.

1. Introduction

Over the past two decades density functional theory (DFT) has become the most applied quantum chemical method due to its attractive accuracy-to-computational cost ratio. The approximations for the exchange-correlation functional can be classified according to their rung on Perdew's 'Jacob's Ladder' [1]: (1) the local density approximation; (2) pure generalised gradient approximation (GGA) employing both the local density and the reduced density gradient; (3) the meta-GGAs which additionally employ the kinetic energy density; (4) the hybrid-GGAs and hybrid-meta-GGAs which in addition involve the occupied orbitals; and (5) the double-hybrid functionals which additionally employ the virtual orbitals. One weakness of the pure and hybrid-GGAs is that there is often a trade-off between their performance for thermochemistry and thermochemical kinetics [2–6]. For example, for hybrid-GGAs, the optimal percentage of Hartree–Fock exchange for thermochemical properties tends to be around the $\sim$ 20% mark, while for

thermochemical kinetics it tends to be at least twice as much [3–5, 7–9]. Hybrid-meta GGA and double-hybrid procedures can overcome this deficiency to a considerable extent [5, 9–12]. However, it is important to identify reactions for which these more sophisticated functionals fail to offer a quantitative description of both the reaction energies and barrier heights. Identifying these problematic reactions is important since (1) DFT functionals should be applied with caution for modelling these reactions, and (2) it may assist in the development of next-generation functionals with improved performance.

Cycloreversion of heteroatom rings, in which a cyclic structure is fragmented into simple unsaturated building blocks, is an important class of organic reactions which are widely used in synthetic and medicinal chemistry [13–16]. We have recently evaluated the performance of a range of contemporary DFT methods for the reaction barrier heights of 20 cycloreversion reactions of 5-membered heterocyclic rings (known as the CRBH20 dataset) [17]. These reaction barrier heights provide a challenging test for DFT and double-hybrid DFT (DHDFT) methods. For example, of the 65 tested DFT and DHDFT functionals, only two DHDFT functionals (B2K-PLYP and DSD-PBEP86) result in root-mean-square deviations (RMSDs) below the 'chemical accuracy' threshold ('chemical accuracy' is arbitrarily defined here as RMSDs ≤ 4.2 kJ mol^{-1} from accurate reference data).

The present study aims to answer an important question: can DFT methods quantitatively describe all aspects of the potential energy surfaces of cycloreversion reactions? (i.e., exhibit reasonably good performance for both reaction energies and barrier heights). To this end, we use a more lenient cut-off of twice the threshold of chemical accuracy (i.e., RMSD ≤ 8.4 kJ mol^{-1}) to indicate that a functional is performing reasonably well for these challenging reactions. We show that only a handful of DFT and DHDFT functionals exhibit reasonable performance for the calculation of both the energies and barrier heights of cycloreversion reactions. For example, only 6 of the 65 considered DFT functionals exhibit RMSDs ≤ 8.4 kJ mol^{-1} for both the reaction energies and barrier heights (namely, 3 hybrid GGA, 2 hybrid-meta GGA, and 1 double-hybrid functional). This result is significant because it demonstrates that even sophisticated functionals from rungs four and five of Jacob's Ladder cannot quantitatively describe all aspects of the potential surface of this important class of reactions.

2. Computational methods

We evaluate the performance of a large number of DFT and ab initio procedures for a set of 20 cycloreversion reaction energies (to be known as the CR20 dataset). In order to obtain reliable reference reaction energies, calculations have been carried out using the high-level, ab initio W1w theory [18, 19] or its explicitly correlated version (W1-F12 theory) [20] with the Molpro 2012.1 program suite [21, 22]. These theories represent layered extrapolations to the all-electron CCSD(T)/CBS limit [23, 24], and can achieve 'sub-chemical accuracy' for atomisation reactions [25]. For example, W1w and W1-F12 theories are associated with RMSDs of 2.6 and 3.1 kJ mol^{-1} for a set of 140 very accurate atomisation energies obtained at the full configuration interaction (FCI) infinite basis-set limit [18, 20, 25]. Nevertheless, we point out that for systems containing only first-row elements (and H) W1-F12 shows better performance. Specifically, for the 97 first-row atomisation energies in the W4-11 dataset [25], W1-F12 attains an RMSD of 1.9 kJ mol^{-1} relative to reference atomisation energies at the FCI infinite basis-set limit [20]. Therefore, in the present work, we obtain reaction energies with W1-F12 theory for systems containing only first-row elements and with W1w theory for systems containing second-row elements. For the sake of making the article self-contained, we will briefly outline the various steps in W1w theory (for further details see Refs. [18] and [19]). The Hartree–Fock (HF) component is extrapolated from the A′VTZ and A′VQZ basis sets, using the $E(L) = E_\infty + A/L^\alpha$ two-point extrapolation formula with $\alpha = 5$. The notation A′VnZ indicates the combination of the standard correlation-consistent cc-pVnZ basis sets on H [26], the aug-cc-pVnZ basis sets on first-row atoms [27], and the aug-cc-pV(n+d)Z basis sets on second-row atoms [28]. The valence CCSD correlation energy is extrapolated from the same basis sets with $\alpha = 3.22$. The quasiperturbative triples, (T), correction is extrapolated from the A′VDZ and A′VTZ basis sets with $\alpha = 3.22$. The CCSD(T) inner-shell contribution is calculated with the MTsmall basis set, which is a completely decontracted cc-pVTZ basis set with tight 2d1f functions added [18]. W1-F12 theory is an explicitly correlated version of W1w, which combines explicitly correlated F12 techniques with basis-set extrapolations in order to approximate the CCSD(T)/CBS energy. The computational protocol of W1-F12 theory has been specified and rationalised in detail in Ref. [20] (for a concise summary of the various steps in W1-F12 theory see Ref. [29]).

The geometries of all structures have been obtained at the B3LYP/A′VTZ level of theory [30–32]. Harmonic vibrational analyses have been performed to confirm each stationary point as an equilibrium structure (i.e., consisting of all real frequencies). All geometry optimisations and frequency calculations were performed using the Gaussian 09 program suite [33].

Finally, we note that in order to facilitate a direct comparison between the approximate theoretical procedures and the Wn results, the B3LYP/A′VTZ geometries are used for all calculations throughout this work.

The DFT exchange-correlation functionals considered in the present study (ordered by their rung on Jacob's Ladder)[1] are the pure generalised gradient approximation (GGA) functionals: BLYP [30, 34], B97-D [35], HCTH407 [36], PBE [37], BP86 [34, 38], BPW91 [34, 39], SOGGA11 [40], N12 [41]; the meta-GGAs (MGGAs): M06-L [42], TPSS [43], τ-HCTH [44], VSXC [45], BB95 [46], M11-L [47], MN12-L [48]; the hybrid-GGAs (HGGAs): BH&HLYP [49], B3LYP [30–32], B3P86 [31, 38], B3PW91 [31, 39], PBE0 [50], B97-1 [51], B98 [52], X3LYP [53], SOGGA11-X [54]; the hybrid-meta-GGAs (HMGGAs): M05 [55], M05-2X [56], M06 [9], M06-2X [9], M06-HF [9], BMK [3], B1B95 [34, 46], TPSSh [57], τ-HCTHh [44], PW6B95 [58], and the DHDFT procedures: [11] B2-PLYP [59], B2GP-PLYP [10], B2K-PLYP [60], B2T-PLYP [60], DSD-BLYP [61], DSD-PBEP86 [11, 62], and PWPB95 [63]. In addition to the global HGGAs and HMGGAs, we also consider the following range-separated (RS) functionals: CAM-B3LYP [64], LC-ωPBE [65], ωB97 [66], ωB97X [66], ωB97X-D [67], and M11 [68]. Empirical D3 dispersion corrections [69, 70, 71] are included in some cases using the Becke−Johnson [72] damping potential as recommended in Ref. [69] (denoted by the suffix -D3). We note that the suffix -D in B97-D and ωB97X-D indicates the original dispersion correction rather than the D3 correction. The standard DFT calculations were carried out in conjunction with the A′VTZ basis set, while the DHDFT calculations, which exhibit slower basis set convergence, were carried out in conjunction with the A′VQZ basis set. All the DFT and DHDFT were performed using the Gaussian 09 [33] and ORCA [73, 74] program suites.

In addition, the performance of composite thermochemical procedures and standard ab initio methods are also assessed. We consider the following composite procedures: G4 [75], G4(MP2) [76], G4(MP2)-6X [77], CBS-QB3 [78], and CBS-APNO [79], and the following ab initio methods: MP2, SCS-MP2 [80], MP2.5 [81], MP3, MP3.5 [82], SCS-MP3 [83], MP4, MP4(SDQ), MP4$_{av}$ [17], CCSD, SCS-CCSD [84], SCS(MI)CCSD [85], and CCSD(T). The performance of the MP2- and MP3-based procedures is evaluated in conjunction with the A′VQZ basis set whereas the performance of the computationally more expensive MP4-based and coupled cluster procedures is evaluated in conjunction with the A′VTZ basis set.
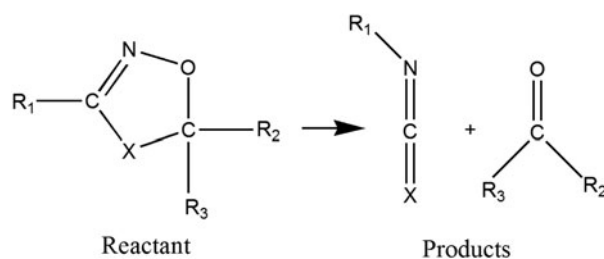

Figure 1. Schematic representation of the reactions in the CR20 database. The X center and the R_1, R_2, and R_3 functional groups are listed in Table 1.

3. Results and discussion

3.1. Overview of the benchmark reaction energies in the CR20 database

A schematic representation of the 20 cycloreversion reactions in the CR20 database is given in Figure 1, while the heteroatom (X) and substituents (R_1–R_3) are listed in Table 1. All the cycloreversion reactions involve cleavage of a heterocyclic ring into two unsaturated fragments. In addition, these reactions involve a migration of the R_1 substituent across the C=N bond (Figure 1). There are two types of reactions in the CR20 dataset (Table 1):

(1) Reactions **1–10** (X = O) are cycloreversions of dioxazoles forming carbonyls and isocyanates
(2) Reactions **11–20** (X = S) are cycloreversions of oxathiazoles forming carbonyls and isothiocyanates

The species in the CR20 database cover a broad spectrum of substituents and bonding situations. We consider

Table 1. Cycloreversion reactions in the CR20 database (see also Figure 1).

Reaction	X	R_1	R_2	R_3
1	O	H	H	H
2	O	Me	H	H
3	O	Et	H	H
4	O	CH$_2$F	H	H
5	O	Me	Me	H
6	O	Me	OH	H
7	O	Me	NH$_2$	H
8	O	Me	F	H
9	O	Me	Me	Me
10	O	Me	a	
11	S	H	H	H
12	S	Me	H	H
13	S	Et	H	H
14	S	CH$_2$F	H	H
15	S	Me	Me	H
16	S	Me	OH	H
17	S	Me	NH$_2$	H
18	S	Me	F	H
19	S	Me	Me	Me
20	S	Me	a	

aThe CR$_2$R$_3$ moiety is replaced with a C=O group (see Figure 1).

two parent structures in which $R_1 = R_2 = R_3 = H$ (and Me). In addition, the R_1, R_2, and R_3 substituents are varied such that: (1) $R_2 = H$, Me, NH_2, OH, and F, while maintaining a constant migrating group (i.e., $R_1 = $ Me), and (2) the migrating group is varied ($R_1 = H$, Me, Et, and CH_2F), while maintaining R_2 and R_3 constant (i.e., $R_2 = R_3 = H$). In addition, we consider the possibility of replacing the CR_2R_3 moiety with a carbonyl group.

Benchmark reference data have been obtained by means of the high-level W1-F12 and W1w procedures [18–20]. Since these procedures represent layered extrapolations to the CCSD(T) basis-set-limit energy, it is of interest to estimate whether the contributions from post-CCSD(T) excitations are likely to be significant. The percentage of the total atomisation energy accounted for by parenthetical connected triple excitations, $\%TAE_e[(T)]$, has been shown to be a reliable energy-based diagnostic for the importance of post-CCSD(T) contributions. It has been suggested that $\%TAE_e[(T)] < 2$ indicates systems that are dominated by dynamical correlation, while $2 < \%TAE_e[(T)] < 5$ indicates systems that include mild non-dynamical correlation [19, 25, 86]. Table S1 (Supplementary data, Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/00268976.2015.1081418) gathers the $\%TAE_e[(T)]$ values for the reactants and products involved in the CR20 dataset. The $\%TAE_e[(T)]$ values for these species lie in the range 1.6–3.9 (where in 87% of cases the $\%TAE_e[(T)] \leq 3.0$). These values suggest that the CCSD(T) method is adequate for the description of the reaction energies in the CR20 database and that our bottom-of-the-well CCSD(T)/CBS benchmark values should be well below 1 kcal mol^{-1} from the full configuration interaction (FCI) basis-set limit [20, 25, 87].

The component breakdown of the W1-F12 and W1w reaction energies are gathered in Table 2. All the reactions in the CR20 database are exothermic with reaction energies ranging from −32.0 (**12**) to −149.4 (**10**) kJ mol^{-1}. The reactions of the dioxazole systems are systematically more exothermic than those of the oxathiazoles (by amounts ranging from 8.7–56.9 kJ mol^{-1}). For both the dioxazoles and oxathiazoles the most exothermic reaction is obtained when R_1 is a methyl group and CR_2R_3 is a C=O group, namely $\Delta E_e = -149.4$ (dioxazole, **10**) and −133.1 (oxathiazole, **20**) kJ mol^{-1}. Similarly, for both ring types the least exothermic reaction is obtained for $R_1 = $ Me and $R_2 = R_3 = H$, namely $\Delta E_e = -67.6$ (dioxazole, **2**) and –32.0 (oxathiazole, **12**) kJ mol^{-1}.

Inspection of Table 2 reveals that the HF/CBS level of theory systematically underestimates the reaction energies by amounts ranging from 51.1 (**8**) to 86.3 (**19**) kJ mol^{-1} relative to the CCSD(T)/CBS reference values.

Table 2. Component breakdown of the benchmark W1-F12 and W1w reaction energies for the reactions in the CR20 database (in kJ mol^{-1}, see Table 1 and Figure 1 for the reaction definitions).[a]

Reaction	ΔSCF	ΔCCSD	Δ(T)	ΔCV[b]	ΔE_e[c]
1	−182.6	49.6	3.7	−1.9	−131.2
2	−122.6	52.9	3.8	−1.7	−67.6
3	−133.1	54.5	4.1	−1.8	−76.3
4	−158.2	55.6	4.3	−1.6	−99.9
5	−135.2	58.1	4.5	−1.8	−74.4
6	−177.3	55.7	3.5	−2.1	−120.3
7	−189.5	59.3	3.9	−2.5	−128.8
8	−141.8	50.6	2.5	−2.0	−90.7
9	−147.8	64.2	5.8	−1.9	−79.7
10	−206.4	55.5	3.7	−2.1	−149.4
11	−131.5	54.7	4.0	−1.5	−74.3
12	−99.9	63.2	6.2	−1.5	−32.0
13	−111.2	65.7	6.8	−1.5	−40.2
14	−117.0	61.7	4.9	−1.3	−51.6
15	−124.0	71.0	8.0	−1.7	−46.8
16	−181.2	69.2	7.2	−1.9	−106.7
17	−182.2	72.6	7.4	−2.4	−104.5
18	−149.1	63.2	5.8	−1.9	−82.0
19	−144.0	78.4	9.7	−1.9	−57.7
20	−210.1	70.3	8.7	−2.1	−133.1

[a]The reference values for reactions **1–10** are from W1-F12 theory, and for reactions **11–20** from W1w theory. [b]Core-valence correction. [c]All-electron, vibrationless CCSD(T) basis-set-limit reaction energies.

The valence CCSD correlation contribution to the reaction energies increases the reaction energies by amounts ranging between 49.6 (**1**) and 78.4 (**19**) kJ mol^{-1}. The quasiperturbative triples, (T), contribution further increases the reaction energies by chemically significant amounts ranging from 2.5 (**8**) to 9.7 (**19**) kJ mol^{-1}. Finally, we note that the core-valence correlation contribution systematically reduces the reaction energies by relatively modest amounts ranging from 1.3 (**14**) to 2.5 (**7**) kJ mol^{-1}.

3.2. Performance of standard and composite ab initio procedures for the reaction energies in the CR20 database

Table 3 gives an overview of the performance of the composite G3, G3(MP2), G4, G4(MP2), G4(MP2)-6X, CBS-QB3, and CBS-APNO procedures, as well as several ab initio methods (e.g., MP2, MP2.5, MP3, MP3.5, MP4, SCS-MP2, SCS-MP3, CCSD, and CCSD(T)). Inspection of Table 3 reveals that most of the composite procedures do not have difficulties with the reactions in the CR20 dataset. The CBS-type procedures show excellent performance with RMSDs of 2.9 (CBS-QB3) and 3.4 (CBS-APNO) kJ mol^{-1}. Three of the Gn-type procedures result in RMSDs below the 'chemical accuracy' threshold, namely, G3 and G4 with an RMSD of 4.1 kJ mol^{-1} and G4(MP2)-6X with a slightly lower RMSD of 3.8 kJ mol^{-1}. The rest of the Gn procedures result in RMSDs

Table 3. Statistical analysis for the performance of composite and standard ab initio methods for the calculation of the reaction energies in the CR20 database (in kJ mol^{-1}).

Basis set	Methods	RMSD[a]	MAD[a]	MSD[a]	LD[a,b]
	G3	4.1	3.5	−2.9	7.9 (**13**)
	G3(MP2)	5.0	4.4	−4.1	9.2 (**13**)
	G3B3	6.7	6.2	−6.1	10.6 (**12**)
	G3(MP2)B3	7.4	6.9	−6.9	11.0 (**12**)
	G4	4.1	3.8	−3.6	6.5 (**12**)
	G4(MP2)	4.7	4.3	−4.3	7.4 (**14**)
	G4(MP2)-6X	3.8	3.4	−3.1	6.9 (**14**)
	CBS-QB3	2.9	2.3	−1.5	6.0 (**12**)
	CBS-APNO[c]	3.4	1.6	3.2	5.8 (**9**)
A′VQZ	HF	65.1	64.4	−64.4	85.8 (**19**)
	MP2	14.8	14.0	−14.0	25.8 (**10**)
	SCS-MP2	23.4	23.1	−23.1	32.3 (**10**)
	MP2.5	3.6	3.2	−3.2	8.5 (**20**)
	MP3	8.6	7.6	7.6	13.8 (**8**)
	SCS-MP3	20.7	20.5	−20.5	27.6 (**10**)
A′VTZ	MP3.5	2.4	2.0	1.6	4.4 (**9**)
	MP4(SDQ)	4.8	4.6	−4.6	6.8 (**20**)
	MP4(SDTQ)	6.9	6.5	−6.5	12.4 (**10**)
	MP4$_{av}$[d]	5.7	5.6	−5.6	8.3 (**10**)
	CCSD	2.4	2.0	−0.6	4.2 (**20**)
	SCS-CCSD	12.4	12.2	12.2	16.0 (**7**)
	SCS(MI)CCSD	10.9	10.7	10.7	13.6 (**7**)
	CCSD(T)	4.7	4.4	4.4	6.8 (**9**)
	CCSD(T)/CBS(MP2)[e]	4.5	4.1	4.1	7.2 (**17**)

[a]RMSD = root-mean-square deviation, MAD = mean-absolute deviation, MSD = mean-signed deviation, LD = largest deviation (in absolute value). [b]The reaction numbers are given in parenthesis (see also Table 1). [c]Error statistics over the dioxazole systems only. [d]MP4$_{av}$ is defined as the average of MP4(SDQ) and MP4(SDTQ), see Ref. [17] for further details. [e]CCSD(T)/CBS(MP2) ≈ CCSD(T)/A′VDZ + MP2/A′V{T,Q}Z − MP2/A′VDZ.

ranging between 4.7 (G3(MP2)) and 7.4 (G3(MP2)B3) kJ mol^{-1}. Finally, we note that, apart from CBS-APNO, all the composite procedures tend to systematically underestimate the reaction energies.

We now turn our attention to the performance of the standard wavefunction methods. Table 3 gives these results in conjunction with the A′VQZ for the computationally more economical MP2- and MP3-based methods, while the performance of the MP4-based and coupled cluster methods is evaluated in conjunction with the A′VTZ basis set. Table S2 of the Supplementary data (Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/00268976.2015.1081418.) gives results for all the standard ab initio methods in conjunction with the A′VnZ basis sets (n = D and T). We start by noting that for the methods for which we have both A′VTZ and A′VQZ results (HF, MP2, SCS-MP2, MP2.5, MP3, and SCS-MP3) the difference in the overall RMSDs between the A′VTZ and A′VQZ basis sets are smaller than 2.1 kJ mol^{-1}. However, in nearly all cases the A′VTZ basis set performs better than the A′VQZ basis set (relative to the W1-F12 and W1w reference values). For the most part, this results from error compensation between basis set incompleteness and the neglect of higher level correlation effects.

Inspection of Table 3 reveals that the CR20 database is a challenging target for nearly all of the standard ab initio methods. Only 3 of the 15 examined procedures result in RMSDs below the 'chemical accuracy' threshold (MP2.5, MP3.5, and CCSD). As mentioned in Section 3.1, Hartree–Fock theory systematically and severely underestimates the W1-F12 and W1w reaction energies. For the HF/A′VQZ level of theory we obtain an RMSD of 65.1 kJ mol^{-1} and an MSD = −1 × MAD = −64.4 kJ mol^{-1} (we note that these values remain essentially unchanged for the HF/CBS level of theory from W1-F12 and W1w theories). Second-order Møller–Plesset perturbation theory (MP2) systematically underestimates the reaction energies and results in an RMSD of 14.8 kJ mol^{-1} and an MSD = −1 × MAD = −14.0 kJ mol^{-1}. Nevertheless, it should be pointed out that MP2 still results in better performance compared to most of the DHDFT functionals (*vide infra*). Inclusion of higher order excitations in the MP3 procedure improves upon the performance of MP2 and leads to an RMSD of 8.6 kJ mol^{-1}. However, MP3 seems to overcorrect the deficiencies of MP2, and leads to systematic overestimation of the reaction energies (as is evident from MSD = MAD = 7.6 kJ mol^{-1}). In this situation, MP2.5, which is an average of MP2 and MP3, results in very good performance with an RMSD of 3.6 kJ mol^{-1} for the CR20 dataset. Fourth-order Møller–Plesset perturbation theory (MP4) provides a slight improvement over MP3, with an RMSD of 6.9 kJ mol^{-1}. However, in contrast to MP3, MP4 systematically underestimates the reaction energies with MSD = −1 × MAD = −6.5 kJ mol^{-1}. We note in passing, that similar oscillatory behaviour of the MPn (n = 2, 3, and 4) series has been previously observed for both thermochemical [81, 83, 88–90] and kinetic properties [82, 91]. Since MP3 systematically overestimates the reaction energies and MP4 systematically underestimates them by similar amounts, MP3.5, which is an average of MP3 and MP4, results in excellent performance with an RMSD of only 2.4 kJ mol^{-1}. In fact, this method outperforms all of the considered DHDFT and MPn-based procedures. The excellent performance of MP3.5 is also demonstrated by a relatively small MSD of −1.6 kJ mol^{-1}. Finally, it should be pointed out that scaling the same-spin and opposite-spin components of the MP2 and MP3 correlation energies, as in the SCS-MP2 and SCS-MP3 procedures, worsens the performance of both procedures and leads to RMSDs of over 20 kJ mol^{-1} (Table 3). This deterioration in performance upon scaling of the same-spin and opposite-spin components is rather unusual for reaction energies [92]. For comparison, for the set of 140 atomisation energies in the W4-11 dataset we obtain the following RMSDs for the MPn and SCS-MPn procedures: 45.8 (MP2), 28.1 (SCS-MP2), 60.5 (MP3), and 20.0 kJ mol^{-1} (SCS-MP3) [25].

The CCSD/A′VTZ level of theory shows excellent performance with an RMSD of 2.4 kJ mol^{-1}. It should be noted, however, that at the CCSD/CBS level of theory (taken from W1-F12 and W1w theories) this RMSD increases to 4.1 kJ mol^{-1}. The CCSD(T) method attains RMSDs of 8.7 and 4.7 kJ mol^{-1} in conjunction with the A′VDZ and A′VTZ basis sets, respectively. Therefore, it seems that the CCSD/A′VTZ level of theory benefits from some error compensation between basis-set incompleteness and the neglect of quasiperturbative triples, (T), corrections. It is of interest to assess the performance of the CCSD(T) method using an additivity-based approach in which the CCSD(T)/CBS energy is estimated from the CCSD(T)/A′VDZ energy and an MP2-based basis-set-correction term (ΔMP2 = MP2/A′V{T,Q}Z − MP2/A′VDZ, where the MP2/A′V{T,Q}Z energy is extrapolated to the basis-set limit with an extrapolation exponent of 3) [93]. This cost-effective approach, which is denoted here by CCSD(T)/CBS(MP2), has been widely used for obtaining non-covalent interaction energies close to the CCSD(T)/CBS limit [94–98]. More recently, this method has also been found to give good performance for reaction energies [89, 90, 99] and barrier heights [17, 82]. For the reactions in the CR20 database the CCSD(T)/MP2(CBS) method gives an RMSD and MSD of 4.5 kJ mol^{-1}, just above the threshold of chemical accuracy.

3.3. Performance of DFT and DHDFT procedures for the reaction energies in the CR20 database

Table 4 gives an overview of the performance of a large number of contemporary DFT and DHDFT functionals (with and without empirical D3 dispersion corrections). We start by making the following general observations:

- The CR20 database proves to be an extremely challenging test for DFT methods. None of the 65 DFT and DHDFT functionals attain an RMSD below the threshold of chemical accuracy. Only 10 functionals attain RMSDs below 8.4 kJ mol^{-1} (twice the threshold of chemical accuracy). The rest of the functionals attain RMSDs ranging between 8.6 (ωB97) and 60.9 (BLYP) kJ mol^{-1}.
- Of the 53 conventional DFT methods (rungs 2–4 of Jacob's Ladder), nine functionals give RMSDs < 8.4 kJ mol^{-1}. Of these, the range-separated hybrid GGA ωB97X functional of Chai and Head-Gordon gives the best performance with an RMSD of 4.7 kJ mol^{-1}, while the hybrid GGA (B3PW91-D3) and the hybrid-meta GGA (M05-2X) tie for the second place with RMSDs of 5.3 and 5.4 kJ mol^{-1}, respectively.

- Of the 12 considered double-hybrid functionals, only one functional (DSD-PBEP86-D3) gives an RMSD below 8.4 kJ mol^{-1}. The other DHDFT functionals attain RMSDs between 13.6 (DSD-PBEP86) and 33.6 (B2-PLYP) kJ mol^{-1}.
- None of the considered DFT and DHDFT functionals result in largest deviations that are below 8.4 kJ mol^{-1}. In particular, the largest deviations range between 9.8 (ωB97X) and 79.2 (BLYP) kJ mol^{-1}. We note that reactions **7** and **19** seem to be particularly problematic for most of the DFT procedures.
- With few exceptions all the conventional and double-hybrid functionals tend to systematically underestimate the reaction energies, as evident from MSD $\approx$ $-1 \times$ MAD (notable exceptions include ωB97, PBE0, LC-ωPBE, and M06-HF).
- Dispersion corrections tend to dramatically improve agreement with the W1-F12 and W1w results. For most of the functionals the RMSDs are reduced by 15–74% upon inclusion of the D3 dispersion corrections.

Two of the 11 considered GGA functionals give relatively good performance with RMSDs below the 8.4 kJ mol^{-1} threshold (RMSD = 7.6 (PBE-D3) and 7.7 (BP86-D3) kJ mol^{-1}). We note that removing the D3 dispersion correction significantly increases the RMSDs for these two functionals, namely to 11.1 (PBE) and 21.7 (BP86) kJ mol^{-1}. The rest of the GGAs give very poor performance with RMSDs between 22.8 (N12) and 60.9 (BLYP) kJ mol^{-1}.

Inclusion of the kinetic energy density in the MGGA procedures does not improve the performance. The best performing MGGA functional attains an RMSD of 9.2 kJ mol^{-1} (TPSS-D3); we note that removing the D3 dispersion correction significantly increases the RMSD by 12.4 kJ mol^{-1}. With the exception of PBE-D3, all the GGA and MGGA functionals systematically underestimate the reaction energies as evident from MSD $\approx$ $-1 \times$ MAD. PBE-D3 attains a slightly positive MSD of +2.4 kJ mol^{-1} (cf. with an MSD of -8.0 kJ mol^{-1} for PBE). These results suggest that, apart from PBE-D3 and BP86-D3, the GGA and MGGA functionals should be applied with caution when calculating cycloreversion reaction energies.

The RMSDs for the HGGA functionals span a wide range, from 5.3 (B3PW91-D3) to 43.2 (B3LYP) kJ mol^{-1}. Of the 13 considered hybrid GGAs, three functionals result in RMSDs < 8.4 kJ mol^{-1}, these are: B3PW91-D3 (5.3), SOGGA11-X (6.1), and PBE0 (6.7 kJ mol^{-1}). We note, however, that the RMSD for B3PW91-D3 increases to as much as 20.2 kJ mol^{-1} upon removal of the D3 dispersion correction. We also note that the popular B3LYP functional shows very poor performance with RMSDs

Table 4. Statistical analysis for the performance of DFT and DHDFT procedures for the calculation of the reaction energies in the CR20 database (in kJ mol^{-1}).[a,b]

Type[c]	Method	RMSD	MAD	MSD	LD
GGA	BLYP	60.9	60.3	−60.3	79.2 (**19**)
	BLYP-D3	38.9	38.4	−38.4	48.7 (**7**)
	B97-D	52.4	52.2	−52.2	59.6 (**17**)
	HCTH407	54.4	53.6	−53.6	69.5 (**19**)
	PBE	11.1	9.0	−8.0	21.4 (**9**)
	PBE-D3	7.6	6.4	2.4	18.6 (**20**)
	BP86	21.7	20.3	−20.3	34.4 (**19**)
	BP86-D3	7.7	5.8	−3.5	15.3 (**7**)
	BPW91	31.1	30.0	−30.0	45.2 (**9**)
	SOGGA11	29.1	26.8	−26.8	48.3 (**7**)
	N12	22.8	20.7	−20.7	37.7 (**19**)
MGGA	M06-L	49.9	49.6	−49.6	58.4 (**7**)
	TPSS	21.6	19.7	−19.6	36.6 (**19**)
	TPSS-D3	9.2	8.2	−6.4	15.5 (**7**)
	τ-HCTH	49.8	48.5	−48.5	69.5 (**9**)
	VSXC	47.5	47.2	−47.2	59.9 (**7**)
	BB95	36.8	35.8	−35.8	50.1 (**19**)
	M11-L	36.9	36.4	−36.4	46.4 (**11**)
	MN12-L	15.7	12.7	−10.9	30.4 (**17**)
HGGA	BH&HLYP	28.9	27.6	−27.6	47.1 (**19**)
	BH&HLYP-D3	14.3	13.6	−13.6	22.7 (**19**)
	B3LYP	43.2	42.5	−42.5	61.3 (**19**)
	B3LYP-D3	24.9	24.5	−24.5	31.6 (**7**)
	B3P86	9.7	7.4	−7.0	21.8 (**19**)
	B3PW91	20.2	18.9	−18.9	34.3 (**19**)
	B3PW91-D3	5.3	4.2	−0.6	10.7 (**7**)
	PBE0	6.7	5.9	2.9	10.7 (**11**)
	PBE0-D3	13.0	12.1	12.1	18.9 (**20**)
	B97-1	18.8	17.7	−17.7	31.9 (**19**)
	B98	24.0	22.9	−22.9	39.0 (**19**)
	X3LYP	37.3	36.5	−36.5	54.9 (**19**)
	SOGGA11-X	6.1	4.3	−3.4	15.9 (**19**)
HMGGA	M05	57.3	56.8	−56.8	66.8 (**19**)
	M05-2X	5.4	4.4	2.7	11.1 (**1**)
	M06	45.4	45.2	−45.2	51.8 (**7**)
	M06-2X	7.4	6.0	−4.6	15.3 (**19**)
	M06-HF	33.9	32.3	32.3	54.8 (**1**)
	BMK	23.6	22.6	−22.6	35.1 (**19**)
	BMK-D3	10.8	10.1	−10.1	17.5 (**17**)
	B1B95	19.6	18.4	−18.4	33.3 (**19**)
	B1B95-D3	6.7	5.3	−5.0	13.4 (**7**)
	TPSSh	17.2	15.4	−14.9	31.9 (**19**)
	τ-HCTHh	21.4	20.3	−20.3	34.5 (**19**)
	PW6B95	21.8	20.8	−20.8	35.3 (**19**)
	PW6B95-D3	16.2	15.3	−15.3	25.2 (**19**)
RS	CAM-B3LYP	18.2	16.7	−16.7	33.8 (**19**)
	CAM-B3LYP-D3	9.6	8.3	−8.3	18.4 (**19**)
	LC-ωPBE	12.8	11.5	11.5	22.1 (**11**)
	LC-ωPBE-D3	22.2	21.6	21.6	30.4 (**11**)
	ωB97	8.6	7.3	7.1	15.7 (**11**)
	ωB97X	4.7	3.8	−1.9	9.8 (**9**)
	ωB97X-D	13.1	12.5	−12.5	20.2 (**19**)
	M11	21.7	21.2	−21.2	29.2 (**17**)
DH	B2-PLYP	33.6	33.4	−33.4	40.4 (**19**)
	B2-PLYP-D3	22.6	22.4	−22.4	28.6 (**7**)
	B2GP-PLYP	25.6	25.5	−25.5	30.5 (**7**)
	B2GP-PLYP-D3	19.2	18.9	−18.9	25.5 (**7**)
	B2K-PLYP	22.8	22.6	−22.6	27.0 (**17**)
	B2K-PLYP-D3	19.3	19.2	−19.2	23.4 (**7**)
	B2T-PLYP	29.3	29.1	−29.1	35.2 (**19**)
	DSD-BLYP	25.1	25.0	−25.0	29.0 (**7**)
	DSD-PBEP86	13.6	13.3	−13.3	17.6 (**7**)
	DSD-PBEP86-D3	7.9	7.2	−7.2	12.7 (**7**)
	PWPB95	18.2	17.8	−17.8	25.0 (**19**)
	PWPB95-D3	13.9	13.5	−13.5	19.0 (**7**)

[a]Footnotes a and b to Table 3 apply here. [b]The standard DFT calculations are carried out in conjunction with the A′VTZ basis set while the DHDFT calculations are carried out in conjunction with the A′VQZ basis set. [c]GGA = generalised gradient approximation, HGGA = hybrid-GGA, MGGA = meta-GGA, RS = range-separated, HMGGA = hybrid-meta-GGA, DH = double hybrid.

of 24.9 and 43.2 and kJ mol^{-1} with and without the D3 dispersion correction, respectively. It is instructive to compare the performance of the hybrid functionals that combine Becke's three-parameter exchange functional with different gradient-corrected correlation functionals. These functionals give RMSDs of 9.7 (B3P86), 20.2 (B3PW91), and 43.2 (B3LYP) kJ mol^{-1}. This reflects the poor performance of the LYP and PW91 correlation functionals for these reactions (see also the discussion of the performance of the DHDFT functionals below). Similar observations have also been noted for other types of reactions, such as bond dissociation energies of N–H and N–Cl bonds [100].

In general, the HMGGA functionals give similar performance to the HGGAs, with a wide spread of RMSDs. Interestingly, both the best and worst performing HMGGA functionals belong to the Minnesota-05 family, with RMSDs of 5.4 (M05-2X) to 57.3 (M05) kJ mol^{-1}. Three out of the 13 considered HMGGA functionals result in RMSDs < 8.4 kJ mol^{-1}, namely M05-2X (5.4), B1B95-D3 (6.7), and M06-2X (7.4 kJ mol^{-1}). We note that the RMSD for B1B95-D3 increases to as much as 19.6 kJ mol^{-1} upon removal of the D3 dispersion correction.

Inspection of Table 4 reveals that the optimal percentage of HF-like exchange for the HGGA and HMGGA functionals spans over a wide range of 20–55%. For example, the best performing functionals with 20–28% of HF-like exchange are (RMSDs are given in parenthesis): B3P86 (7.3), PBE0-D3 (5.8), PBE0 (6.2), and B1B95-D3 (7.7 kJ mol^{-1}). At the other end, functionals with 40–55% of HF-like exchange that perform well are: SOGGA11-X (5.6), BH&HLYP-D3 (7.2), and M06-2X (6.8 kJ mol^{-1}). We note that similar observations have been made for the barrier heights of the reactions in the CRBH20 dataset [17].

The RMSDs for the range-separated hybrid- and hybrid-meta-GGAs range between 4.7 (ωB97X) and 22.2 (LC-ωPBE-D3) kJ mol^{-1}, where ωB97X clearly outperforms the rest of the functionals. The good performance of ωB97X is followed by ωB97 (RMSD = 8.6) and CAM-B3LYP-D3 (RMSD = 9.6 kJ mol^{-1}).

All of the considered double-hybrid functionals lead to disappointing RMSDs considering their high computational cost. In particular, the RMSDs for the DHDFT procedures range between 7.9 (DSD-PBEP86-D3) and 33.6 (B2-PLYP) kJ mol^{-1}. We note that the DHDFT functionals that employ the LYP correlation functional give poor performance with RMSDs ranging between 19.2 (B2GP-PLYP-D3) and 33.6 (B2-PLYP) kJ mol^{-1}. Notably, these DHDFT procedures perform significantly worse than MP2 (which attains an RMSD of 14.8 kJ mol^{-1},

Table 3). The spin-component-scaled DHDFT procedures (PWPB95 and DSD-PBEP86) give better performance, but still perform on par or worse than MP2, with RMSDs of 18.2 and 13.6 kJ mol^{-1}, respectively. Finally, we note that dispersion corrections reduce the RMSDs of the DHDFT functionals by considerable amounts; however, DSD-PBEP86-D3 (RMSD = 7.8 kJ mol^{-1}) is the only DHDFT functional that manages to slightly surpass the RMSD cut-off value of 8.4 kJ mol^{-1}.

As mentioned above, without dispersion corrections nearly all the conventional and double-hybrid functionals tend to systematically underestimate the reaction energies, as evident from MSD $\approx$ −1 × MAD. Dispersion effects make the reaction energies less exothermic since they stabilise the heterocyclic rings to a larger extent than the products of the cycloreversion reaction (Figure 1). Therefore, dispersion D3 corrections tend to systematically improve the performance of the DFT and DHDFT methods. For the few functionals that tend to overestimate the reaction energies (e.g., PBE0 and LC-ωPBE) there is no point in including D3 corrections since they can only increase the errors. Table 5 gathers the differences in RMSD between the dispersion-corrected and uncorrected DFT functionals (ΔRMSD = RMSD(DFT) − RMSD(DFT-D3)). A positive ΔRMSD value indicates that the dispersion correction improves the performance of the functional, whereas a negative ΔRMSD value indicates deterioration in performance. Inspection of Table 5 reveals that, with few exceptions, the dispersion D3 corrections significantly improve the agreement with the

Table 5. Overview of the performance of various DFT functionals with and without empirical D3 dispersion corrections. The tabulated values are ΔRMSD = RMSD(DFT) − RMSD(DFT-D3) (in kJ mol^{-1}).[a]

Type	Method	ΔRMSD
GGA	BLYP	22.0
	PBE	3.5
	BP86	14.0
MGGA	TPSS	12.4
HGGA	BH&HLYP	14.6
	B3LYP	18.3
	B3PW91	14.9
	PBE0	− 6.3
HMGGA	BMK	12.8
	B1B95	12.9
	PW6B95	5.6
RS	CAM-B3LYP	8.6
	LC-ωPBE	− 9.4
	ωB97X[b]	− 8.4
DH	B2-PLYP	11.0
	B2GP-PLYP	6.4
	B2K-PLYP	3.5
	DSD-PBEP86	5.7
	PWPB95	4.3

[a]Footnotes a–c to Table 4 apply here. [b]The original dispersion correction rather than the D3 correction is used for this functional.

W1-F12 and W1w results. In particular, the RMSDs are reduced by amounts ranging from 3.5 (PBE) to 22.0 (BLYP) kJ mol^{-1} upon inclusion of the dispersion correction. Cases where the RMSD is reduced by more than 50% upon inclusion of the D3 dispersion correction include: BH&HLYP, BMK, TPSS, BP86, B1B95, and B3PW91. Finally, it is worth pointing out that sizable dispersion corrections are still obtained for the DHDFT functionals; in particular, the RMSDs are reduced by amounts ranging between 3.5 (B2K-PLYP) and 11.0 (B2-PLYP) kJ mol^{-1}.

It is generally accepted that the accuracy of DFT methods should increase as one climbs the rungs of Jacob's Ladder. One way of demonstrating this is to take the average of the RMSDs obtained for the DFT functionals from each rung of Jacob's Ladder. Let us consider, for example, the performance of DFT for the atomisation energies in the W4-11 dataset. Reference [25] reports the RMSDs for a wide range of functionals from rungs 2–5 of Jacob's Ladder relative to highly accurate reference values from W4 (or higher) theory [19]. Taking the average of the RMSDs obtained for the functionals from each rung of Jacob's Ladder we obtain: 37.8 (pure GGAs, rung two), 21.9 (MGGAs, rung three), 16.1 (HMGGAs, rung four), and 9.1 (DHDFT, rung five) kJ mol^{-1}. Thus, the averaged RMSD is reduced as one climbs the rungs of Jacob's Ladder. In contrast, for the reactions in the CR20 dataset we obtain the following averaged RMSDs: 30.7 (over 11 pure GGAs), 33.4 (over 8 MGGAs), 22.1 (over 13 HMGGAs), and 20.9 (over 12 DHDFT) kJ mol^{-1}. Thus, there is little or no improvement when going from rungs 2 → 3 and 4 → 5. Having said that, we note that taking the average RMSD over a set of ~10 DFT functionals from each rung of Jacob's Ladder is a crude approach for differentiating between the performance of functionals from different rungs. Thus, it is instructive to also compare the RMSD of the best performer from each rung rather than the averaged RMSD. The RMSD of the best performers are: 7.6 (GGA), 9.2 (MGGA), 5.4 (HMGGA), and 7.9 (DHDFT) kJ mol^{-1}. Again, we see that there is no improvement in performance when going from rungs 2 → 3 and 4 → 5.

3.4. Performance of theoretical procedures for both reaction energies and barrier heights

In order to carry out computational investigations of the mechanisms of cycloreversion reactions we need to identify theoretical procedures that can accurately describe all aspects of the potential energy surface (i.e., give both reliable reaction energies and barrier heights). This is particularly important for investigations, such as Ref. [14], that consider a number of competing reaction pathways which involve reaction energies and barrier heights that

are separated by only a few kJ mol^{-1}. The results presented here for reaction energies and in Ref. [17] for reaction barrier heights reveal a frustrating situation in which it is hard to find theoretical procedures that perform reasonably well for both reaction energies and barrier heights. This section identifies the best DFT and ab initio methods that are recommended for investigating the entire potential energy surface of cycloreversion reactions.

First, let us consider the performance of the DFT methods. None of the considered DFT and DHDFT methods are able to calculate the reaction energies with RMSDs below the chemical accuracy threshold (Table 4). Thus, we will point out the procedures that result in RMSDs below the threshold of twice that of chemical accuracy (i.e., RMSD ≤ 8.4 kJ mol^{-1}). In Section 3.3 we identified 10 functionals that can achieve this goal. Namely, the GGAs: PBE-D3 and BP86-D3; the HGGAs B3PW91-D3, SOGGA11-X, and PBE0; the range-separated HGGA ωB97X; the HMGGAs M05-2X, B1B95-D3, and M06-2X; and the DHDFT DSD-PBEP86-D3. As may be expected, both GGA functionals significantly underestimate the reaction barrier heights (with RMSDs of over 30 kJ mol^{-1} for the CRBH20 dataset) [17]. The HGGA B3PW91-D3 and the HMGGA M05-2X also result in relatively poor performance for the reaction barrier heights (with RMSDs of over 10 kJ mol^{-1}) [17]. We therefore do not recommend the use of these four functionals for the calculation of both the energetics and kinetics of cycloreversion reaction profiles. Table 6 gathers the RMSDs for the CR20 and CRBH20 databases for the rest of the best performing functionals. Looking at the average of the RMSDs for both

Table 6. Performance of DFT and ab initio methods for both the reaction energies (CR20 dataset) and barrier heights (CRBH20 dataset) of cycloreversion reactions (in kJ mol^{-1}).

Methods	CR20[a]	CRBH20[a]	Both datasets[b]
DFT and DHDFT methods[c]			
SOGGA11-X	6.1	5.6	5.9
PBE0	6.7	6.2	6.5
B1B95-D3	6.7	7.7	7.2
M06-2X	7.4	6.8	7.1
ωB97X	4.7	7.3	6.0
DSD-PBEP86-D3	7.9	2.5	5.2
Composite ab initio methods[d]			
CBS-QB3	2.9	3.1	3.0
CBS-APNO[e]	3.4	1.5	2.5
Standard ab initio methods[f]			
CCSD(T)/A′VTZ	4.7	5.0	4.9
CCSD(T)/CBS(MP2)[g]	4.5	1.7	3.1

[a]RMSDs for the CR20 and CRBH20 datasets. [b]Average of the RMSD for the CR20 and CRBH20 datasets. [c]DFT methods with RMSD < 8.4 kJ mol^{-1} for both the CR20 and CRBH20 datasets. [d]Composite ab initio methods with RMSD < 4.2 kJ mol^{-1} for both the CR20 and CRBH20 datasets. [e]Error statistics over the dioxazole systems only. [f]Standard ab initio methods with RMSD ≤ 5.0 kJ mol^{-1} for both the CR20 and CRBH20 datasets. [g]CCSD(T)/CBS(MP2) ≈ CCSD(T)/A′VDZ + MP2/A′V{T,Q}Z − MP2/A′VDZ.

the CR20 and CRBH20 databases, the DHDFT DSD-PBEP86-D3 method emerges as the best performer with an averaged RMSD of 5.2 kJ mol^{-1}. However, while this method gives superb performance for the reaction barrier heights (RMSD = 2.5 kJ mol^{-1}), it shows relatively poor performance for the reaction energies (RMSD = 7.9 kJ mol^{-1}). The HGGA SOGGA11-X and the range-separated HGGA ωB97X emerge as the second best performers with averaged RMSDs of 5.9–6.0 kJ mol^{-1}. Of these functionals, SOGGA11-X gives slightly better performance for the reaction barrier heights, whereas ωB97X gives slightly better performance for the reaction energies. Overall, both of these functionals are equally recommended for investigations of the mechanisms of cycloreversion reactions.

Let us now move to the performance of the composite ab initio methods. In Section 3.2 we identified five methods that are able to calculate the reaction energies with RMSDs below the chemical accuracy threshold (Table 3). Namely, CBS-QB3, CBS-APNO, G3, G4, and G4(MP2)-6X. Of these, the Gn methods result in RMSDs above the threshold of chemical accuracy for the reaction barrier heights (namely, RMSD = 4.9–6.4 kJ mol^{-1}). However, the CBS-type methods result in excellent performance for both the reaction energies and barrier heights, with RMSDs well below the threshold of chemical accuracy. Whereas CBS-APNO gives significantly better performance for the reaction barrier heights, CBS-QB3 gives better performance for the reaction energies. Overall, both methods are recommended for investigations of the potential energy surfaces of cycloreversion reactions.

Finally, let us consider the performance of the standard ab initio methods. In Section 3.2 we identified three methods with RMSDs below the chemical accuracy threshold for the CR20 dataset (namely, MP2.5/A$'$VQZ, MP3.5/A$'$VTZ, and CCSD/A$'$VTZ). Of these, MP2.5 and MP3.5 exhibit particularly poor performance for the barrier heights, with RMSDs of 26.4 and 11.3 kJ mol^{-1}, respectively. The CCSD/A$'$VTZ level of theory also results in an unsatisfactory RMSD of 6.3 kJ mol^{-1} for the CRBH20 database. Table 6 lists the standard ab initio methods that result in RMSDs $\leq$ 5.0 kJ mol^{-1} for both the CR20 and CRBH20 datasets. The CCSD(T)/A$'$VTZ level of theory just achieves this goal with RMSDs of 4.7 and 5.0 kJ mol^{-1}, respectively. The MP2-based additivity approach for estimating the CCSD(T)/CBS energy (CCSD(T)/CBS(MP2)) shows similar performance for the reaction energies, but significantly improved performance for the reaction barrier heights. We therefore recommend the CCSD(T)/CBS(MP2) method for investigating the potential energy surface of cycloreversion reactions.

4. Conclusions

We introduce a representative benchmark database of 20 cycloreversion reaction energies (to be known as the CR20 dataset). The reference reaction energies are obtained at the CCSD(T)/CBS level by means of the high-level W1-F12 and W1w thermochemical protocols. These benchmark values allow us to assess the performance of more approximate theoretical procedures. Specifically, we examine the performance of a variety of contemporary DFT and DHDFT procedures, as well as a number of standard and composite ab initio methods. With regard to the performance of the DFT and DHDFT procedures we make the following observations:

- The CR20 dataset proves to be an extremely challenging test for all of the considered DFT and DHDFT procedures (rungs 2–5 of Jacob's Ladder). With no exceptions, all the DFT and DHDFT methods result in RMSDs above the 'chemical accuracy' threshold (i.e., RMSDs > 4.2 kJ mol^{-1}). Increasing the threshold of an acceptable RMSD to twice that of chemical accuracy (i.e., RMSD = 8.4 kJ mol^{-1}), results in 10 functionals with acceptable performance.
- The best performing DFT functionals (RMSDs are given in parenthesis) are the GGAs: PBE-D3 (7.6) and BP86-D3 (7.7); the HGGAs B3PW91-D3 (5.3), SOGGA11-X (6.1), and PBE0 (6.7); the RS HGGA ωB97X (4.7); the HMGGAs M05-2X (5.4), B1B95-D3 (6.7), and M06-2X (7.4); and the DHDFT DSD-PBEP86-D3 (7.9 kJ mol^{-1}).
- The rest of the functionals attain RMSDs ranging between 8.6 (ωB97) and 60.9 (BLYP) kJ mol^{-1}.
- Empirical D3 dispersion corrections significantly improve the performance of most of the DFT functionals and are essential for obtaining reasonable performance for the CR20 dataset. In particular, the RMSDs for most of the functionals are reduced by 15–74% upon inclusion of the D3 dispersion corrections.

With regard to the performance of the standard and composite ab initio procedures, we draw the following conclusions:

- The composite ab initio procedures show good-to-excellent performance for the CR20 dataset. The best performing composite procedures with RMSDs below the threshold of chemical accuracy are: CBS-QB3 (2.9), CBS-APNO (3.4), G4(MP2)-6X (3.8), G3 and G4 (4.1 kJ mol^{-1}).
- The best performing MPn-based procedures with RMSDs below the threshold of chemical accuracy are: MP2.5 (3.6) and MP3.5 (2.4 kJ mol^{-1}).

The rest of the MPn-based procedures show relatively poor performance with RMSDs between 4.8 (MP4(SDQ)) and 23.4 (SCS-MP2) kJ mol^{-1}.

- The CCSD(T) method in conjunction with the A′VDZ and A′VTZ basis sets gives RMSDs of 8.7 and 4.7 kJ mol^{-1}, respectively. Estimating the CCSD(T)/CBS energy from the CCSD(T)/A′VDZ energy and adding an MP2-based basis-set-correction term (i.e., the CCSD(T)/CBS(MP2) method) results in an RMSD of 4.5 kJ mol^{-1}.

Based on the results presented here for reaction energies and in Ref. [17] for reaction barrier heights we find that only a handful of DFT and ab initio methods are suitable for investigating the entire potential energy surface of cycloreversion reactions. In particular, we recommend the following methods for such investigations:

- Of the DFT methods, the HGGA SOGGA11-X and the range-separated HGGA ωB97X emerge as the best performers with RMSDs between 4.7–7.3 kJ mol^{-1}, for the reaction energies and barrier heights.
- Of the DHDFT methods, DSD-PBEP86-D3 shows the best performance with RMSDs of 2.5 and 7.9 kJ mol^{-1} for the CRBH20 and CR20 databases, respectively.
- Of the standard ab initio methods, the CCSD(T)/CBS(MP2) method emerges as the best performer with RMSDs of 1.7 and 4.5 kJ mol^{-1} for the CRBH20 and CR20 databases, respectively.
- Of the composite methods, CBS-APNO and CBS-QB3 give excellent performance for both the reaction energies and barrier heights, with RMSDs between 1.5 and 3.4 kJ mol^{-1}. Overall, the CBS-type procedures are the methods of choice for investigating the entire potential energy surface of cycloreversion reactions.

Acknowledgments

Disclosure statement

Funding

Supplemental data

Supplementary material can be found at http://dx.doi.org/10.1080/00268976.2015.1081418.

ORCID

Amir Karton http://orcid.org/0000-0002-7981-508X

References

[1] J.P. Perdew and K. Schmidt, AIP Conf. Proc. **577**, 1 (2001).

[2] J. Baker, J. Andzelm, M. Muir, and P.R. Taylor, Chem. Phys. Lett. **237**, 53 (1995).

[3] J.L. Durant, Chem. Phys. Lett. **256**, 595 (1996).

[4] A.D. Boese, J.M.L. Martin, and N.C. Handy, J. Chem. Phys. **119**, 3005 (2003).

[5] A.D. Boese and J.M.L. Martin, J. Chem. Phys. **121**, 3405 (2004).

[6] A.J. Cohen, P. Mori-Sánchez, and W. Yang, Chem. Rev. **112**, 289 (2012).

[7] B.J. Lynch, P.L. Fast, M. Harris, and D.G. Truhlar, J. Phys. Chem. A **104**, 4811 (2000).

[8] J.K. Kang and C.B. Musgrave, J. Chem. Phys. **115**, 11040 (2001).

[9] Y. Zhao and D.G. Truhlar, Theor. Chem. Acc. **120**, 215 (2008).

[10] A. Karton, A. Tarnopolsky, J.–F. Lamere, G.C. Schatz, and J.M.L. Martin, J. Phys. Chem. A **112**, 12868 (2008).

[11] S. Kozuch and J.M.L. Martin, J. Comp. Chem. **34**, 2327 (2013).

[12] L. Goerigk and S. Grimme, WIREs Comput. Mol. Sci. **4**, 576 (2014).

[13] B.A. Burkett, J.M. Kane-Barber, R.J. O'Reilly, and L. Shi, Tetrahedron Lett. **48**, 5355 (2007).

[14] R.J. O'Reilly and L. Radom, Org. Lett. **11**, 1325 (2009).

[15] J.N. Kim, K.S. Jung, H.J. Lee, and J.S. Son, Tetrahedron Lett. **38**, 1596 (1997).

[16] Y. Hu, C.-Y. Li, X.-M. Wang, Y.-H. Yang, and H.-L. Zhu, Chem. Rev. **114**, 5572 (2014).

[17] L.-J. Yu, F. Sarrami, R.J. O'Reilly, and A. Karton, Chem. Phys. **458**, 1 (2015).

[18] J.M.L. Martin and G. Oliveira, J. Chem. Phys. **111**, 1843 (1999).

[19] A. Karton, E. Rabinovich, J.M.L. Martin, and B. Ruscic, J. Chem. Phys. **125**, 144108 (2006).

[20] A. Karton and J.M.L. Martin, J. Chem. Phys. **136**, 124114 (2012).

[21] MOLPRO is a package of ab initio programs written by H.-J. Werner, P.J. Knowles, G. Knizia, F.R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, et al. See: http:www.molpro.net.

[22] H.-J. Werner, P.J. Knowles, G. Knizia, F.R. Manby, and M. Schütz, WIREs Comput. Mol. Sci. **2**, 242 (2012).

[23] K.A. Peterson, D. Feller, and D.A. Dixon, Theor. Chem. Acc. **131**, 1079 (2012).

[24] T. Helgaker, W. Klopper, and D.P. Tew, Mol. Phys. **106**, 2107 (2008).

[25] A. Karton, S. Daon, and J.M.L. Martin, Chem. Phys. Lett. **510**, 165 (2011).

[26] T.H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[27] R.A. Kendall, T.H. Dunning Jr., and R.J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[28] T.H. Dunning, K.A. Peterson, and A.K. Wilson, J. Chem. Phys. **114**, 9244 (2001).

[29] A. Karton, Chem. Phys. Lett. **585**, 330 (2014).

[30] C. Lee, W. Yang, and R.G. Parr, Phys. Rev. B **37**, 785 (1988).

[31] A.D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[32] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, and M.J. Frisch, J. Phys. Chem. **98**, 11623 (1994).

[33] 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, et al. *Gaussian 09, Revision D.01* (Gaussian, Inc.: Wallingford CT, 2009).

[34] A.D. Becke, Phys. Rev. A **38**, 3098 (1988).

[35] S. Grimme, J. Comp. Chem. **27**, 1787 (2006).

[36] A.D. Boese and N.C. Handy, J. Chem. Phys. **114**, 5497 (2001).

[37] J.P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); ibid Phys. Rev. Lett. **78**, 1396 (1997).

[38] J.P. Perdew, Phys. Rev. B **33**, 8822 (1986).

[39] J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).

[40] R. Peverati, Y. Zhao, and D.G. Truhlar, J. Phys. Chem. Lett. **2**, 1991 (2011).

[41] R. Peverati and D.G. Truhlar, J. Chem. Theory Comput. **8**, 2310 (2012).

[42] Y. Zhao and D.G. Truhlar, J. Chem. Phys. **125**, 194101 (2006).

[43] J.M. Tao, J.P. Perdew, V.N. Staroverov, and G.E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).

[44] A.D. Boese and N.C. Handy, J. Chem. Phys. **116**, 9559 (2002).

[45] T. van Voorhis and G.E. Scuseria, J. Chem. Phys. **109**, 400 (1998).

[46] A.D. Becke, J. Chem. Phys. **104**, 1040 (1996).

[47] R. Peverati and D.G. Truhlar, J. Phys. Chem. Lett. **3**, 117 (2012).

[48] R. Peverati and D.G. Truhlar, Phys. Chem. Chem. Phys. **10**, 13171 (2012).

[49] A.D. Becke, J. Chem. Phys. **98**, 1372 (1993).

[50] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).

[51] F.A. Hamprecht, A.J. Cohen, D.J. Tozer, and N.C. Handy, J. Chem. Phys. **109**, 6264 (1998).

[52] H.L. Schmider and A.D. Becke, J. Chem. Phys. **108**, 9624 (1998).

[53] X. Xu, Q. Zhang, R.P. Muller, and W.A. Goddard, J. Chem. Phys. **122**, 014105 (2005).

[54] R. Peverati and D.G. Truhlar, J. Chem. Phys. **135**, 191102 (2011).

[55] Y. Zhao, N.E. Schultz, and D.G. Truhlar, J. Chem. Phys. **123**, 161103 (2005).

[56] Y. Zhao, N.E. Schultz, and D.G. Truhlar, J. Chem. Theory Comput. **2**, 364 (2006).

[57] V.N. Staroverov, G.E. Scuseria, J. Tao, and J.P. Perdew, J. Chem. Phys. **119**, 12129 (2003).

[58] Y. Zhao and D.G. Truhlar, J. Phys. Chem. A **109**, 5656 (2005).

[59] S. Grimme, J. Chem. Phys. **124**, 034108 (2006).

[60] A. Tarnopolsky, A. Karton, R. Sertchook, D. Vuzman, and J.M.L. Martin, J. Phys. Chem. A **112**, 3 (2008).

[61] S. Kozuch, D. Gruzman, and J.M.L. Martin, J. Phys. Chem. C **114**, 20801 (2010).

[62] S. Kozuch and J.M.L. Martin, Phys. Chem. Chem. Phys. **13**, 20104 (2011).

[63] L. Goerigk and S. Grimme, J. Chem. Theory Comput. **7**, 291 (2011).

[64] T. Yanai, D. Tew, and N. Handy, Chem. Phys. Lett. **393**, 51 (2004).

[65] O.A. Vydrov and G.E. Scuseria, J. Chem. Phys. **125**, 34109 (2006).

[66] J.–D. Chai and M. Head-Gordon, J. Chem. Phys. **128**, 084106 (2008).

[67] J.–D. Chai and M. Head-Gordon, Phys. Chem. Chem. Phys. **10**, 6615 (2008).

[68] R. Peverati and D.G. Truhlar, J. Phys. Chem. Lett. **2**, 2810 (2011).

[69] S. Grimme, S. Ehrlich, and L. Goerigk, J. Comput. Chem. **32**, 1456 (2011).

[70] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).

[71] S. Grimme, WIREs Comput. Mol. Sci. **1**, 211 (2011).

[72] A.D. Becke and E.R. Johnson, J. Chem. Phys. **123**, 154101 (2005).

[73] ORCA is a modern electronic structure program package written by F. Neese, with contributions from Ute Becker, Dmytro Bykov, Dmitry Ganyushin, Andreas Hansen, Robert Izsak, Dimitrios G. Liakos, Christian Kollmar, Simone Kossmann, Dimitrios A. Pantazis, Taras Petrenko, et al. Max Planck Institute for Chemical Energy Conversion, 2014.

[74] F. Neese, WIREs Comput. Mol. Sci. **2**, 73 (2012).

[75] L.A. Curtiss, P.C. Redfern, and K. Raghavachari, J. Chem. Phys. **126**, 084108 (2007).

[76] L.A. Curtiss, P.C. Redfern, and K. Raghavachari, J. Chem. Phys. **127**, 124105 (2007).

[77] B. Chan, J. Deng, and L. Radom, J. Chem. Theory Comput. **7**, 112 (2011).

[78] J.A. Montgomery Jr., M.J. Frisch, J.W. Ochterski, and G.A. Petersson, J. Chem. Phys. **110**, 2822 (1999); ibid **112**, 6532 (2000).

[79] J.W. Ochterski, G.A. Petersson, and J.A. Montgomery Jr., J. Chem. Phys. **104**, 2598 (1996).

[80] S. Grimme, J. Chem. Phys. **118**, 9095 (2003).

[81] M. Pitonak, P. Neogrady, J. Cerny, S. Grimme, and P. Hobza, Chem. Phys. Chem. **10**, 282 (2009).

[82] A. Karton and L. Goerigk, J. Comp. Chem. **36**, 622 (2015).

[83] S. Grimme, J. Comp. Chem. **24**, 1529 (2003).

[84] T. Takatani, E.E. Hohenstein, and C.D. Sherrill, J. Chem. Phys. **128**, 124111 (2008).

[85] M. Pitonak, J. Rezac, and P. Hobza, Phys. Chem. Chem. Phys. **12**, 9611 (2011).

[86] A. Karton, and J.M.L. Martin, J. Chem. Phys. **133**, 144102 (2010).

[87] A. Karton, I. Kaminker, and J.M.L. Martin, J. Phys. Chem. A **113**, 7610 (2009).

[88] M.L. Leininger, W.D. Allen, H.F. Schaefer, and C.D. Sherrill, J. Chem. Phys. **112**, 9213 (2000).

[89] L.-J. Yu and A. Karton, Chem. Phys. **441**, 166 (2014).

[90] L.-J. Yu, F. Sarrami, A. Karton, and R.J. O'Reilly, Mol. Phys. **113**, 1284 (2015).

[91] A. Karton, R.J. O'Reilly, B. Chan, and L. Radom, J. Chem. Theory Comput. **8**, 3128 (2012).

[92] S. Grimme, L. Goerigk, and R.F. Fink, WIREs Comput. Mol. Sci. **2**, 886 (2012).

[93] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A.K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).

[94] B. Brauer, M.K. Kesharwani, and J.M.L. Martin, J. Chem. Theory Comput. **10**, 3791 (2014).

[95] L. Goerigk, A. Karton, J.M.L. Martin, and L. Radom, Phys. Chem. Chem. Phys. **15**, 7028 (2013).

[96] D.G. Liakos and F. Neese, J. Phys. Chem. A **116**, 4801 (2012).

[97] P. Jurecka, J. Sponer, J. Cerny, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).

[98] P. Jurečka and P. Hobza, Chem. Phys. Lett. **365**, 89 (2002).

[99] X. He, L. Fusti-Molnar, and K.M. Merz, Jr., J. Phys. Chem. A **113**, 10096 (2009).

[100] R.J. O'Reilly, A. Karton, and L. Radom, Int. J. Quantum Chem. **116**, 1862 (2012).

Appendix A. Supplementary data

Diagnostics indicating the importance of post-CCSD(T) contributions for the species involved in the CR20 database (Table S1); overview of the basis-set convergence of the standard and modified ab initio procedures (Table S2); overview of the basis-set convergence of conventional DFT methods (Table S3); and full references for Ref. [21] (Molpro 2012), Ref. [33] (Gaussian 09), and Ref. [73] (ORCA). For the convenience of the reader the Supplementary data also includes a directory (CR20_input_files_and_script.zip) with the Gaussian 09 input files for the species involved in the CR20 database and a perl script that calculates the reaction energies and the error statistics from our W1-F12 and W1w reference values.