

CHEMISTRY

A European Journal

A Journal of



Accepted Article

Title: Assessing how correlated MO calculations can perform vs Kohn-Sham DFT: barrier heights/isomerizations

Authors: Antonio Varandas

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *Chem. Eur. J.* 10.1002/chem.201700928

Link to VoR: <http://dx.doi.org/10.1002/chem.201700928>

Supported by
ACES

WILEY-VCH

Assessing how correlated MO calculations can perform vs Kohn-Sham DFT: barrier heights/isomerizations

A. J. C. Varandas,^{*,†} M. Martínez González,^{†,‡} L. A. Montero-Cabrera,[‡] and J. M. Garcia de la Vega^{†,¶}

[†]*Departamento de Química, and Centro de Química, Universidade de Coimbra, 3004-535 Coimbra, Portugal*

[‡]*Universidad de La Habana, Facultad de Química, calle San Lázaro sn., 10400 La Habana, Cuba*

[¶]*Departamento de Química-Física Aplicada, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain*

E-mail: varandas@uc.pt

Abstract

To assess the title issue, 38 hydrogen transfer barrier heights and 38 non-hydrogen transfer barrier heights/isomerizations extracted from extensive databases have been considered, in addition to four 2p isomerization reactions and six others for large organic molecules. All Kohn-Sham DFT calculations have employed the popular M06-2X functional, while the correlated MO-based ones are from single-reference MP2 and CCSD(T) methods. They have all utilized the same basis sets, with the raw MO energies subsequently extrapolated to the complete basis set limit without additional cost. MP2 calculations are found to be as cost effective as DFT ones and often slightly

more, while showing a satisfactory accuracy when compared with the reference data. Although the focus is on barrier heights, the results may bear broader implications in that one may see the successes and difficulties of DFT when compared with traditional MO theories for the same data.

1 Introduction

Density functional theory^{1,2} (DFT) is the leading first-principles method in computing electronic structures and properties in molecular science. Indeed, DFT has enabled electronic structure calculations on complex chemical problems with an accuracy often unobtainable by other approaches, which makes it useful for modelling, prediction, design and understanding.³ Mainstream DFT applications are based on Kohn-Sham (KS) DFT,^{4,5} which is formally exact. Yet, despite offering in principle an exact formulation of quantum mechanical electronic structure theory, KS DFT relies in practice on approximate (often empirical) exchange-correlation (XC) functionals. In fact, the exact form of the XC functional is unknown and essentially unknowable:⁴ were it known, and DFT would take into full account all complex many-body effects at a computation cost characteristic of mean field approximations. Amongst the many reported functionals in the literature, the Minnesota global-hybrid meta functionals^{3,6,7} have shown to have a particularly good performance for a variety of systems which served either as test set for their calibration or as ground for evaluating their predictive performance.⁸ It turns out that the best available functional for one application is often not the best for another. Although unquestionable progress (of no concern here) has been made, this situation has been standing for years. This justifies our quest for testing how traditional correlated electronic structure methods based on molecular orbital (MO) theory can perform when compared with DFT for the predicted properties and cost effectiveness.

As the title shows, the focus will be here on hydrogen transfer barrier heights and non-hydrogen transfer barrier heights/isomerizations since they are a key goal of chemistry, and often an attribute where correlated single-reference MO theories such as second-order Möller-

Plesset perturbation theory (MP2) and coupled-cluster single and double electronic excitations with a perturbative correction for triples [CCSD(T)] are believed to have modest performances as they involve species far from their equilibrium geometries.^{9,10} We will further concentrate on comparing the performance of such MO methods with DFT employing the Minnesota density functionals, specifically M06-2X which has been extensively compared in the literature with other functionals.³ Indeed, as an overall guide, it has most recently been considered from an exhaustive study⁸ of such functionals as "one of the two hybrid functionals most broadly useful", although as also remarked "each has different strengths and weaknesses". Thence, it would be interesting to extend the discussion to other functionals, primarily dispersion-corrected ones such as B3LYP-Dn.^{11,12} These mix the conventional B3LYP^{13,14} functional with an add-on energy term which is a simple function of interatomic distances containing adjustable parameters fitted to conformational and interaction energies computed via CCSD(T) calculations at the complete-basis-set extrapolated limit, CCSD(T)/CBS. The fitting is then functional-specific, with¹¹ D2 implying energy corrections for all pairs of atoms while¹² D3 considers also triplets of atoms to account for three-body effects. However, for a more focused discussion, only the M06-2X functional will be examined. Indeed, it is expected to provide the best performance since the systems involved on the test databases here utilized have been employed for the calibration of M06-2X.^{3,6,7}

In this work, we will then hope to show that MP2 in its restricted open-shell variant (RMP2, denoted simply here as MP2) can perform well both in its predictive capabilities and particularly in cost-effectiveness when compared with DFT/M06-2X. Note that both MO methods here discussed are size-consistent, a primary requirement for accuracy in studying intermolecular forces and reaction dynamics.^{9,10} Thence, the present results may have broader implications in that they may at least call for testing the successes and failures of many kinds of other density functionals when compared to MP2 for the same data. The present work provides therefore a first status report on the quest of how the two most popular *ab initio* MO methods, namely MP2 and CCSD(T) - this commonly considered the

golden approach of quantum chemists -, perform versus the state-of-the-art DFT/M06-2X approach. Being a fairly obvious goal to pursue, the question may then arise of why it has only now been put. As it will become clear later, the major motivation can be attributed to the quite good performance of extrapolation techniques to yield CBS results. Although the success of multilevel techniques such as CBS is well recognized,¹⁵ the novelty is on its demonstrated^{16–18} reliability even when using the most cost-effective MO correlated single-reference raw energies, thus allowing a useful comparison with the experimental values.

2 Theory

A general and most convenient partition of the molecular energy is

$$E = E_{\text{HF}} + E_{\text{cor}} \quad (1)$$

where HF stands for Hartree-Fock, and “cor” for correlation energy. Calculations of the HF and correlation energy contributions have then been carried out for all systems and reference geometries given elsewhere.¹⁹ In all cases, an atomic basis set of the cc-pVXZ Dunning’s family^{20,21} has been used. For proper comparison, all KS DFT results have here been obtained with the same basis set.

The present study involved therefore two basic steps. First, the reactants and products were characterized using DFT and high level methods of *ab initio* correlated electronic structure theory. Specifically, both MP2 and coupled-cluster methods were employed, this in its popular CCSD(T) version. Core correlation effects have been ignored.¹⁰ The second step involves single-point calculations for the supermolecular (intermediary) species. No attempt has been done to obtain optimized structures (these were taken as used for calibrating/testing the M06-2X functional³), although it should be recommended for more genuine estimates at those levels of theory.¹⁹

Unrestricted-Hartree-Fock (UHF) wave functions that allow the bonding electrons to

localize on the separated fragments are popular but they suffer of spin contamination, which is particularly important in the intermediate region where the bonds are elongated but not completely broken.^{9,10} On the other hand, it is well known to facilitate derivation of an UMP method, while addition of electron correlation tends to decrease spin contamination, albeit at a small level for low orders of the MP perturbation expansion.^{9,10} Indeed, despite recovering a significant fraction of the electron correlation (both static and dynamical), the net effect tends to be an increase in energy due to spin contamination. Although this may have no consequences at dissociation, when the different spin states have equal energies, it tends to make the RMP energy lower than the UMP one in the intermediate region, yet with the RHF (denoted simply as HF) energy higher than the UHF one. Indeed, spin contamination in the UHF wave function tends to make the UMP expansion converge slower than RMP.²² For open-shell systems where RHF cannot be used, this may recommend that the reference wave function should be of the restricted open-shell HF type, as the results here reported seem to corroborate. Note that the CBS extrapolation schemes here employed have been calibrated for ROHF and ROMP2 (thence RHF and RMP2 or just HF and MP2) methods, and hence are here used for UMP2 in an ad hoc manner.

For the purpose of generality and having in mind relatively large species, the simplest Dunning's^{20,21} correlation consistent bases have been utilized: cc-pVXZ, denoted shortly as VXZ. For affordability, a major asset of the present proposal, all calculations have employed mostly basis with the two lowest cardinal numbers, $X = D$ and T , although a higher member of the hierarchy ($X = Q$) has occasionally been utilized. Strictly speaking, we should emphasize that a composite procedure is here utilized, namely MP2/VTZ//QCISD and CCSD(T)/VTZ//QCISD; as usual, the notation implies MP2 or CCSD(T) single-point calculations at a QCISD (or another high level of theory) geometry;¹⁹ due to irrelevance for the discussion, no indication is given in the text (see Tables 1 and 2) of the specific basis set used in QCISD. [Parenthetically, suffice it to say that QCISD/MG3 was utilized to obtain the optimized geometries for the reactions of the HTBH38/08 and NHTBH38/08 databases,³

PBE0-D3/def2-TZVPP^{23,24} for 2pIsoE4 (2p isomerization energies), and B97-D/TZVP^{25,26} for IsoL6/11 (six isomerization energies of large organic molecules). Regarding the reference values for the HTBH38/08 and NHTBH38/08 energies, they come from a myriad of theoretical and experimental sources as shown in Tables 1 and 2. In turn, for the 2pIsoE4 database, the reported values originate from spin-component-scaling [SCS-MP2/cc-p(wC)V5Z] calculations with further corrections for higher correlation effects based on the CCSD(T)²⁷ method. Similarly, for the IsoL6-11 database, calculations at the CCSD(T)-F12a/aug-cc-pVDZ level of theory were utilized.²⁸ Although the above composite schemes are popular,¹⁰ reasons were given^{29,30} for caution in practicing the approach. One should recall that similar MO calculations of aromatic hydrocarbons, often generate³¹ anomalous results when employing popular basis sets of the Pople type (*e.g.*, 6-311G,^{32,33} etc). To our knowledge, no such stark warnings for black-box ab initio studies have been reported when using correlation consistent basis sets even belonging to the lowest hierarchical ranks.

For enhanced accuracy, the electronic energy along the reaction path is split as in Eq. (1), with each energy component separately extrapolated to the CBS limit. The procedure has been described in detail elsewhere,^{16,34,35} and hence is here only briefly addressed. For consistency, the same pair of cardinal/hierarchical numbers will be used for the extrapolations. We employ our own HF-CBS protocol³⁶ for HF extrapolations from *DZ* and *TZ* or *TZ* and *QZ* bases, denoted (d, t) and (t, q) ; $X = D, T, \dots$ are cardinal numbers, while $x = d, t, \dots$ are the corresponding hierarchical numbers. These dual-level extrapolations are of the form $A + B \exp(-\gamma x)$, with the CBS energy obtained from a protocol³⁷ (see the Supplementary Information, SI) believed to yield (d, t) absolute energies converged to a root mean square error $< 0.5 \text{ kJ mol}^{-1}$, thence satisfactory for the present purposes. The expectations are that convergence gets tighter and tighter with increasing pairs of the hierarchical numbers.

Extrapolations of the correlation energy find a basis on studies of its dependence on the partial wave quantum number for two-electron atomic systems and second-order pair energies in many-electron atoms.^{38,39} The CBS protocol employs the inverse power form^{17,40}

$E_X^{\text{cor}} = E_\infty^{\text{cor}} + A_3/x^3$, where E_X^{cor} is the correlation energy for cardinal number X , and E_∞^{cor} and A_3 are parameters determined from calculations for two affordable hierarchical numbers (x or X); note that the hierarchical x values are theory specific.¹⁷ Suffice it to add that traditional extrapolation protocols are reliable only when based on X values typically larger than Q , which makes re-hierarchization^{16,17} crucial for affordability (use of D and T bases or just D or T ;¹⁸ note that the number of basis functions scales as X^3). Thus, the protocol utilized assumes here too a popular form (see the SI) which has been shown^{16,17} to yield accurate results for a variety of systems and methods.^{34,41–43}

The basis set superposition error⁴⁵ has been ignored in calculating the energetics. Indeed, use of the counterpoise technique⁴⁵ is deemed unjustified^{16,46,47} due to employing CBS extrapolated energies. In fact, counterpoise is not warranted with small basis functions⁴⁸ such as DZ and TZ , with CBS being the only logical and reliable alternative.^{16,46,47}

All electronic structure calculations [M06-2X, CCSD(T), and MP2] both for the direct and reverse reactions employed MOLPRO.⁴⁴ Note that the CPU times refer always to the total spent with the two basis sets plus CBS extrapolation. They were performed by storing the 2e integrals on the hard drive, and no approximations such as Resolution of the Identity were used: the standard MP2 and CCSD(T) methods were always utilized. As a measure of computational cost for MP2/CBS and CCSD(T), the CPU time ratio against the M06-2X functional has always been used.

3 Results and discussion

The two major databases of barrier heights used in the present work are as follows. The database of hydrogen transfer barrier heights (HTBH38/08) includes 38 transition state barrier heights for 19 hydrogen transfer reactions, 18 of which involve radicals as reactants and products. All reactions in HTBH38/08 are isodesmic, with 6 revised in 2008.⁶³ All geometries have been obtained at QCISD-MG3 level of theory.⁶⁴ Note that the MG3 basis

Table 1: Calculated barrier heights, in kcal mol⁻¹, for the hydrogen transfer barrier heights of the HTBH data base.^{49–51}

#	reaction	Direct					Reverse					Energy reference
		Ref. ^{a)}	MP2 ^{b)}	UMP2 ^{b)}	CCSD(T) ^{b)}	M06-2X ^{c)}	Ref. ^{a)}	MP2 ^{b)}	UMP2 ^{b)}	CCSD(T) ^{b)}	M06-2X ^{c)}	
1	H + HCl → H ₂ + Cl	5.7	8.45	9.96	4.90	3.82	8.7	7.73	8.10	8.86	8.37	exp. ^{52,53}
2	OH + H ₂ → H ₂ O + H	5.1	7.77	7.04	5.80	5.37	21.2	30.96	30.78	20.86	20.01	SAE ^{e)} 54,55
3	CH ₃ + H ₂ → CH ₄ + H	12.1	11.76	12.95	12.23	12.27	15.3	17.84	19.51	14.72	15.46	exp. ⁵²
4	OH + CH ₄ → H ₂ O + CH ₃	6.7	6.52	7.20	6.69	5.11	19.6	23.63	24.38	19.26	16.56	exp. ⁵²
5	H + H ₂ → H ₂ + H	9.6	12.26	13.08	9.79	15.08	9.6	12.26	13.08	9.79	15.08	QM ^{f)} 56
6	OH + NH ₃ → H ₂ O + NH ₂	3.2	8.25		5.20	1.70	12.7	20.07		15.16	10.76	exp. ⁵²
7	HCl + CH ₃ → CH ₄ + Cl	1.7	1.20	2.49	2.15	-0.22	7.9	6.57	7.19	8.61	7.52	exp. ⁵²
8	OH + C ₂ H ₆ → H ₂ O + C ₂ H ₅	3.4	4.77	4.88	3.95	2.41	19.9	25.13	25.17	19.91	18.00	exp. ⁵²
9	F + H ₂ → HF + H	1.8	5.32	4.58	2.48	2.57	33.4	45.41	45.21	33.38	31.08	MRCI(Q) 57
10	O + CH ₄ → OH + CH ₃	13.7	12.28	16.79	15.03	13.45	8.1	7.88	11.86	8.82	5.35	exp. ⁵²
11	H + PH ₃ → H ₂ + PH ₂	3.1	5.48	5.24	2.99	3.42	23.2	26.45	25.18	26.20	26.46	exp. ^{52,53}
12	H + HO → H ₂ + O	10.7	15.82	17.31	10.37	9.41	13.1	14.14	15.67	14.09	14.32	MRCI(Q)/CBS ⁵⁸
13	H + H ₂ S → H ₂ + HS	3.5	6.23	6.32	3.94	4.26	17.3	17.92	16.90	19.07	19.15	exp. ^{52,53}
14	O + HCl → OH + Cl	9.8	18.67	15.16	12.86	8.04	10.4	19.63	14.93	13.10	7.68	exp. ⁵²
15	CH ₃ + NH ₂ → CH ₄ + NH	8	7.53	10.83	8.83	5.80	22.4	20.09	23.79	22.47	20.74	exp. ^{52,53}
16	C ₂ H ₅ + NH ₂ → C ₂ H ₆ + NH	7.5	7.49	11.35	9.55	7.39	18.3	16.80	21.20	19.80	18.18	exp. ^{52,53}
17	NH ₂ + C ₂ H ₆ → NH ₃ + C ₂ H ₅	10.4	7.37	11.23	11.36	9.88	17.4	15.92	19.73	17.37	16.41	exp. ⁵²
18	NH ₂ + CH ₄ → NH ₃ + CH ₃	14.5	9.75	13.82	13.97	12.37	17.8	15.04	19.20	16.59	14.75	exp. ⁵⁹
19	s - t c - C ₅ H ₈ → s - t c - C ₅ H ₈ ^{d)}	38.4	33.23	33.24	39.67	39.27	38.4	33.23	33.24	39.67	39.27	SAE ^{e)} 59
MAE			2.79	2.75	0.86	1.30		3.87	3.71	0.95	1.85	

^{a)}Reference data, reported as best know estimates,³ obtained as indicated on the last column. ^{b)}After CBS(*d, t*) extrapolation. ^{c)}Calculated in the present work with the cc-pVTZ basis. ^{d)}t stands for trans and c for cis. ^{e)}SAE stands for semiempirical adjustment to experiment. ^{f)}QM stands for Quantum Monte Carlo.

Table 2: Calculated barrier heights, in kcal mol⁻¹, for the non-hydrogen transfer barrier heights of the NHTBH data base.^{49–51}

#	reaction	Direct					Reverse					Energy reference
		Ref. ^{a)}	MP2 ^{b)}	UMP2 ^{b)}	CCSD(T) ^{b)}	M06-2X ^{c)}	Ref. ^{a)}	MP2 ^{b)}	UMP2 ^{b)}	CCSD(T) ^{b)}	M06-2X ^{c)}	
1	H + N ₂ O → OH + N ₂	17.13	21.46	35.77	18.56	17.82	82.47	75.72	89.48	84.96	81.18	W4 ⁶⁰
2	H + FH → HF + H	42.18	44.44	46.98	43.12	39.85	42.18	44.44	46.98	43.12	39.85	W1 ⁶¹
3	H + ClH → HCl + H	18	18.55	22.80	18.73	19.13	18	18.55	22.80	18.73	19.13	MRCI(Q)/CBS ⁵⁸
4	H + FCH ₃ → HF + CH ₃	30.38	36.35	36.23	31.89	31.48	57.02	61.12	60.52	58.92	54.04	W1 ⁶¹
5	H + F ₂ → HF + F	2.27	4.74	30.54	3.04	2.26	106.18	106.76	132.02	106.29	107.46	W1 ⁶¹
6	CH ₃ + FCl → CH ₃ F + Cl	6.73	8.31	21.14	8.37	4.30	60	63.58	75.74	62.66	60.15	W3.2 ⁶⁰
7	F ⁻ + CH ₃ F → FCH ₃ + F ⁻	-0.34	-7.85	-7.84	-7.90	-9.46	-0.34	-7.85	-7.84	-7.90	-9.46	W2h ⁶²
8	F ⁻ CH ₃ F → FCH ₃ F ⁻	13.38	11.99		11.86	13.26	13.38	11.99		11.86	13.26	W2h ⁶²
9	Cl ⁻ + CH ₃ Cl → ClCH ₃ + Cl ⁻	3.1	1.09	1.09	0.13	-0.86	3.1	1.09	1.09	0.13	-0.86	SAE ^{d)} 61
10	Cl ⁻ CH ₃ Cl → ClCH ₃ Cl ⁻	13.41	13.23		12.02	12.23	13.41	13.23		12.02	12.23	W3.2 ⁶⁰
11	F ⁻ + CH ₃ Cl → FCH ₃ + Cl ⁻	-12.54	-20.63	-20.63	-22.03	-25.28	20.11	22.10	22.10	23.04	24.69	W1 ⁶²
12	F ⁻ CH ₃ Cl → FCH ₃ Cl ⁻	3.44	1.46		0.26	0.55	29.42	32.93		33.60	36.30	W3.2 ⁶⁰
13	OH ⁻ + CH ₃ F → HOCH ₃ + F ⁻	-2.44	-12.03	-12.03	-11.50	-12.65	17.66	9.11	9.11	8.98	8.74	W3.2 ⁶⁰
14	OH ⁻ CH ₃ F → HOCH ₃ F ⁻	10.96	7.29		7.40	8.46	47.2	49.84		49.42	52.45	W1 ⁶¹
15	H + N ₂ → HN ₂	14.36	16.27	27.46	14.99	13.90	10.61	5.70	8.61	11.44	10.93	W4 ⁶⁰
16	H + CO → HCO	3.17	5.49	5.43	3.48	3.49	22.68	24.35	23.24	23.29	22.44	W1 ⁶¹
17	H + C ₂ H ₄ → CH ₃ CH ₂	1.72	3.79	9.14	2.14	3.00	41.75	42.52	47.25	42.93	43.64	VSEC ⁶⁰
18	CH ₃ + C ₂ H ₄ → CH ₃ CH ₂ CH ₂	6.85	7.15	12.73	6.85	5.93	32.97	35.28	40.69	32.96	34.08	W1 ⁶¹
19	HCN → HNC	48.07	52.23	52.23	47.95	45.96	32.82	34.75	34.75	33.28	32.98	W4 ⁶⁰
MAE			3.28	9.12	2.49	2.93		3.01	6.63	2.28	2.78	

^{a)}Reference data, reported as best know estimates,³ obtained as indicated on the last column. ^{b)}After CBS(*d, t*) extrapolation. ^{c)}Calculated in the present work with the cc-pVTZ basis. ^{d)}SAE stands for semiempirical adjustment to experiment.

set, also known as G3LargeMP2, is identical to 6-311++G(3d2f,2df,2p) for H through Si and very similar to 6-311+G(3d2f) for P, S, and Cl. For this database, the DFT calculations reported by Truhlar and coworkers used the MG3S basis set,⁶⁵ which is identical to MG3 except that diffuse functions on hydrogen have been removed.⁶⁶ In turn, the original version of the database for non-hydrogen transfer barrier heights (NHTBH38/08) dates back to 2004⁵⁰ by joining three older databases containing 38 transition state barrier heights. It contains 12 barrier heights for heavy-atom transfer reactions, 16 for nucleophilic substitution (NS) reactions, and 10 for non-NS unimolecular and association reactions. Likewise the other database, geometries are obtained at the QCISD-MG3 level of theory,⁶⁴ with 18 reference data in the NHTBH38 database revised in 2008.⁶³ For this database, the DFT calculations reported by Truhlar and coworkers used the MG3S basis set.⁶⁵ Since there is no risk of confusion, the above two databases will be briefly referred as HTBH and NHTBH in an obvious correspondence. For consistency, novel DFT/M06-2X calculations are performed but with the VTZ basis set. These are followed by (restricted open-shell type) MP2 and CCSD(T) calculations employing the same basis sets, *i.e.*, with $X = D$ and T . Because the MP2/VTZ calculations turn often to be more cost-effective than the M06-2X ones, additional MP2/VQZ calculations have been performed. In turn, extra CBS extrapolations are done, namely MP2/CBS(d, q) and MP2/CBS(t, q), thence allowing a test of convergence of the CBS extrapolations. Furthermore, UMP2 calculations have been carried out, and subsequently CBS(d, t) extrapolated assuming that the extrapolation scheme developed for HF and MP2 energies still apply. All calculated raw energies are gathered as SI.

Figure 1 top panel compares M06-2X/DFT vs MP2/CBS(d, t) and CCSD(T)/CBS(d, t) errors for the 19 forward reactions of the HTBH database, while the errors for the backward reactions are on the bottom panel. Corresponding comparisons but for the reactions of the NHTBH database are in Figure 2. Tables 1 and 2 gather the basic energetic data, while the SI collects the full calculated energetics and timings for all 76 reactions here studied, including the data used for the graphical and tabular material utilized in the main

text. The general trend indicates that MP2/CBS(d, t) [this extends to MP2/CBS(d, q) and MP2/CBS(t, q)] tends to overestimate the reference while DFT/M06-2X underestimates it. An exception, where both underestimate the reference value, is *e.g.* the HTBH reaction $\text{HCl} + \text{CH}_3 \rightarrow \text{CH}_4 + \text{Cl}$ (reaction # 7) with a barrier height of $1.7 \text{ kcal mol}^{-1}$. In this case, the MP2/CBS(d, t) value is $1.2 \text{ kcal mol}^{-1}$ compared to $-0.2 \text{ kcal mol}^{-1}$ for DFT/M06-2X. Another interesting case may be the isomerization reaction $\text{F}^- \text{CH}_3 \text{Cl} \rightarrow \text{FCH}_3 \text{Cl}^-$ from the NHTBH database, where MP2/CBS(d, t) outperforms DFT/M06-2X both for the direct and reverse processes. Indeed, a similar observation is valid for many isomerization reactions of the type $\text{X}^- \text{CH}_3 \text{Y} \rightarrow \text{XCH}_3 \text{Y}^-$ as well as bimolecular reactions (X and Y are electronegative atoms or groups) where MP2/CBS(d, t) is found to perform quite satisfactorily. Interestingly, the mean absolute errors⁶⁷ (MAE) using MP2 are larger (HTBH: $4.14 \text{ kcal mol}^{-1}$ NHTBH: $5.58 \text{ kcal mol}^{-1}$) than the corresponding values here reported (HTBH: $3.33 \text{ kcal mol}^{-1}$; NHTBH: $3.14 \text{ kcal mol}^{-1}$). This is probably to be expected, with one of the reasons being the use by Trulhar's group of the MG3S basis set while, for consistency, cc-pVXZ has always been here employed.

Time performances are compared in Figures 3 and 4. Because MP2/CBS(d, t) times tend to be even smaller than the DFT/M06-2X ones, it is appropriate to consider the next levels of extrapolation: MP2/CBS(d, q) and MP2/CBS(t, q). The range of results covered by these values is indicated by the error bars in Figures 1 to 4. Note that rather than times, we report time-ratios obtained by taking the M06-2X values as reference, *i.e.* $t_Y/t_{\text{M06-2X}}$ where Y stands for the MO-based method.

Regarding the energetics of the above 76 reactions, the following observations are in order: (a) The CCSD(T) method is clearly the correlated single-reference method that yields results in best overall agreement with the reference data. In fact, this may make the present CCSD(T) calculations the most reliable ones. Naturally, the price to pay reflects on the cost that may reach from several times to one order of magnitude or even more in relation to MP2/CBS(d, t) and DFT/M06-2X. (b) The DFT/M06-2X calculations employing the VTZ

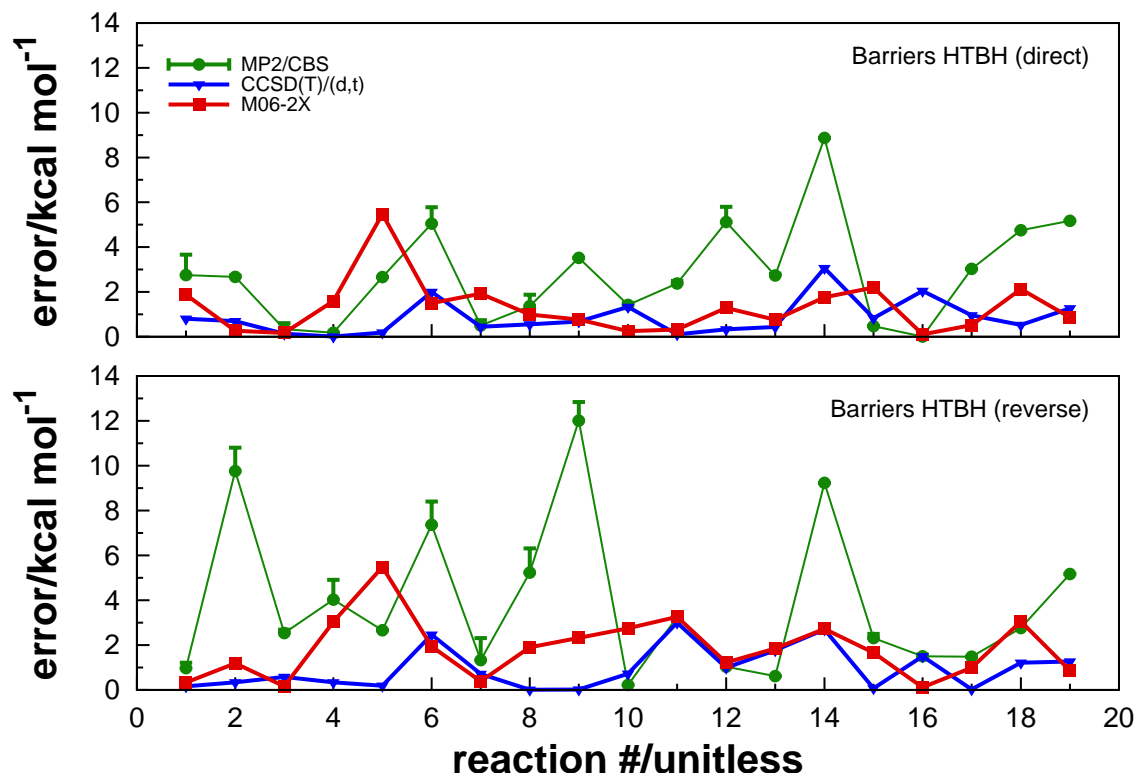


Figure 1: Errors in the predicted barrier heights/isomerization for the 19 HTBH reactions studied in the present work: top panel, forward; bottom panel, reverse. Indicated by solid circles are the MP2/CBS(d,t) results while the error-bar-like symbols indicate the range covered by all [(d,t), (d,q), and (t,q)] CBS/extrapolated values. If the symbol is not visible, it simply implies that all values are indistinguishable within the scale of the plot. The corresponding DFT/M06-2X and CCSDT(T)/CBS(d,t) values are indicated by the solid squares and down-triangles, respectively. The line connecting the points of a given family is meant only to guide the eye.

basis set show the second overall best performance as measured from the MAE relative to the reference data (these, except for some values that were updated since the creation of the functional, were used to calibrate M06-2X using the MG3S basis set). This may not be totally surprising because the VTZ basis set is not expected to yield drastically distinct results⁶⁸ when compared with MG3S that was employed to calibrate the M06-2X functional. However, if one counts the successful cases with respect to the reference data, the results favor MP2/CBS(d,t) in 8 out of the 19 cases both for the direct and reverse HTBH reactions. The corresponding number for both the direct and reverse NHTBH reactions is 10. Thus, one has overall $\sim 50\%$ of successes for both DFT/M06-2X and

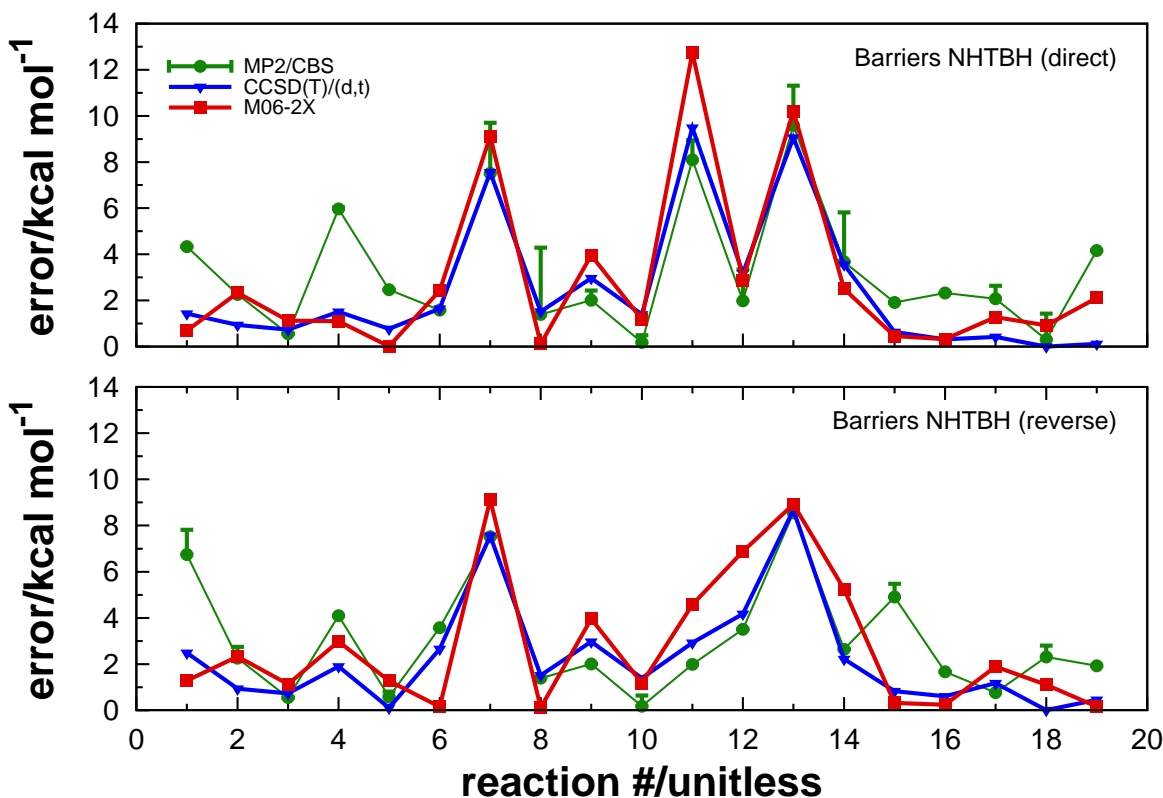


Figure 2: As in Figure 1 but for the NHTBH reactions.

MP2/CBS(d, t). The statistics is maintained if the reference data is taken to be the CCSD(T) results here reported. (c) Reaction barriers are prone to error cancellations as they result from subtracting the absolute energy at the asymptote from the absolute energy of the intermediary at the geometry where the barrier occurs. Besides the reasons associated to the different electronic structures that are likely to be involved, such a difference may also explain why the various theories describe the barrier for the forward and reverse reactions with distinct performances. A poor performance of such cancellations, partly due to spin contamination that affects distinctly the intermediate and the asymptotes, may be the reason for the poor performance of the UMP2/CBS(d, t) method in the case of the NHTBH reactions. This is not apparent for the HTBH reactions, where the UMP2 performance is even slightly better than MP2 when judged from the overall MAE. (d) For the 19 forward and reverse NHTBH reactions, MP2/CBS(d, t) results show a MAE that underperform DFT/M06-2X by only $\sim 0.5 \text{ kcal mol}^{-1}$ ($0.35 \text{ kcal mol}^{-1}$ for the forward reactions, and $0.23 \text{ kcal mol}^{-1}$ for

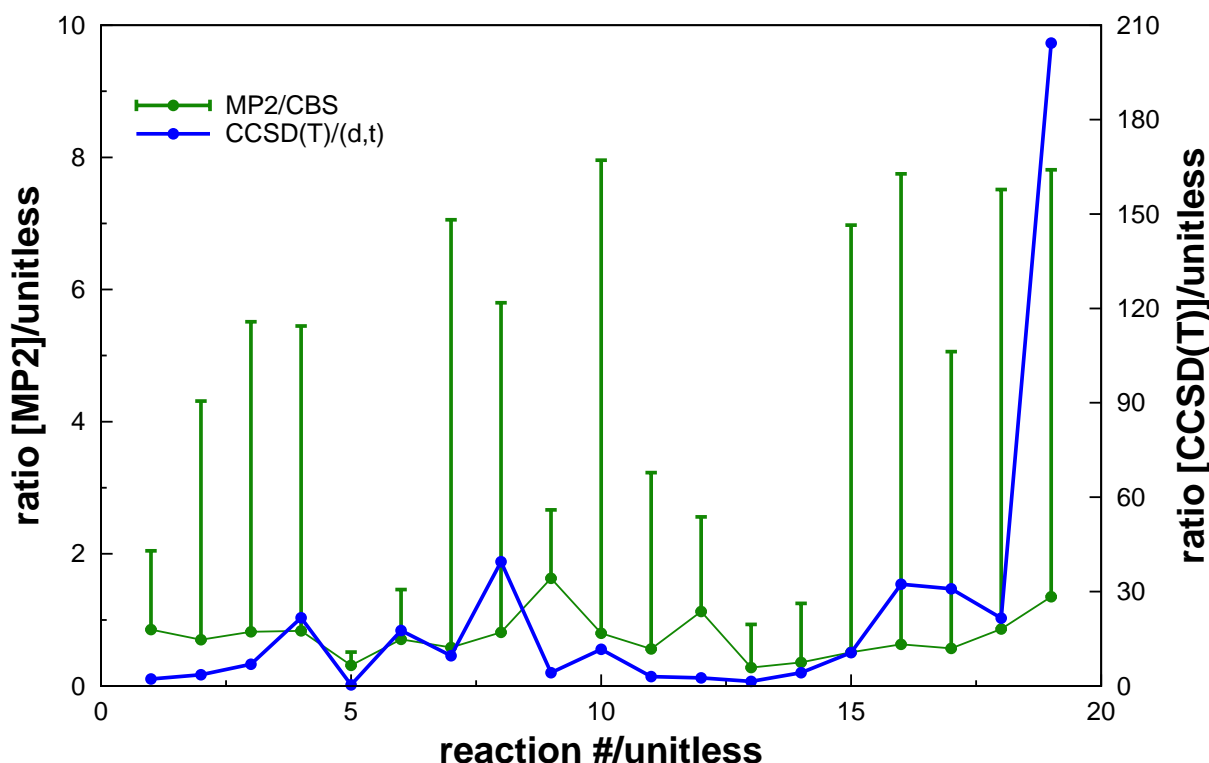


Figure 3: Ratio between the total times spent in the calculations of the heights/isomerizations with the various correlated MO-methods here employed for the HTBH reactions with respect to the corresponding DFT/M06-2X ones. Note that the left-hand-side Y-axis applies to MP2/CBS, while the rhs refers to CCSD(T)/CBS. Naturally, MP2/CBS(d, t) corresponds to the minimum time-ratio, and the maximum to MP2/CBS(t, q).

the reverse ones). Such values are slightly increased for the HTBH reactions, being the MAE larger by $1.49 \text{ kcal mol}^{-1}$ when compared with the DFT/M06-2X value for the direct processes, and $2.02 \text{ kcal mol}^{-1}$ for the reverse ones. Interestingly, the results appear to be fairly well converged at this level of theory when compared with the somewhat more expensive MP2/CBS(d, q) and MP2/CBS(t, q) results, with an exception being the forward HTBH reaction #13 ($\text{H} + \text{H}_2\text{S} \rightarrow \text{H}_2 + \text{HS}$). An explanation can only be traced to the performance of the MP2 method with respect to the reference values. Regarding DFT, it is interesting to note that the larger differences between the reference and M06-2X barrier energies are 5.48 and $9.12 \text{ kcal mol}^{-1}$ for reactions HTBH #5, $\text{H} + \text{H}_2 \rightarrow \text{H}_2 + \text{H}$, and NHTBH #7, $\text{F}^- + \text{CH}_3\text{F} \rightarrow \text{FCH}_3 + \text{F}^-$. The most underperforming reaction for DFT calculations is by far the NHTBH

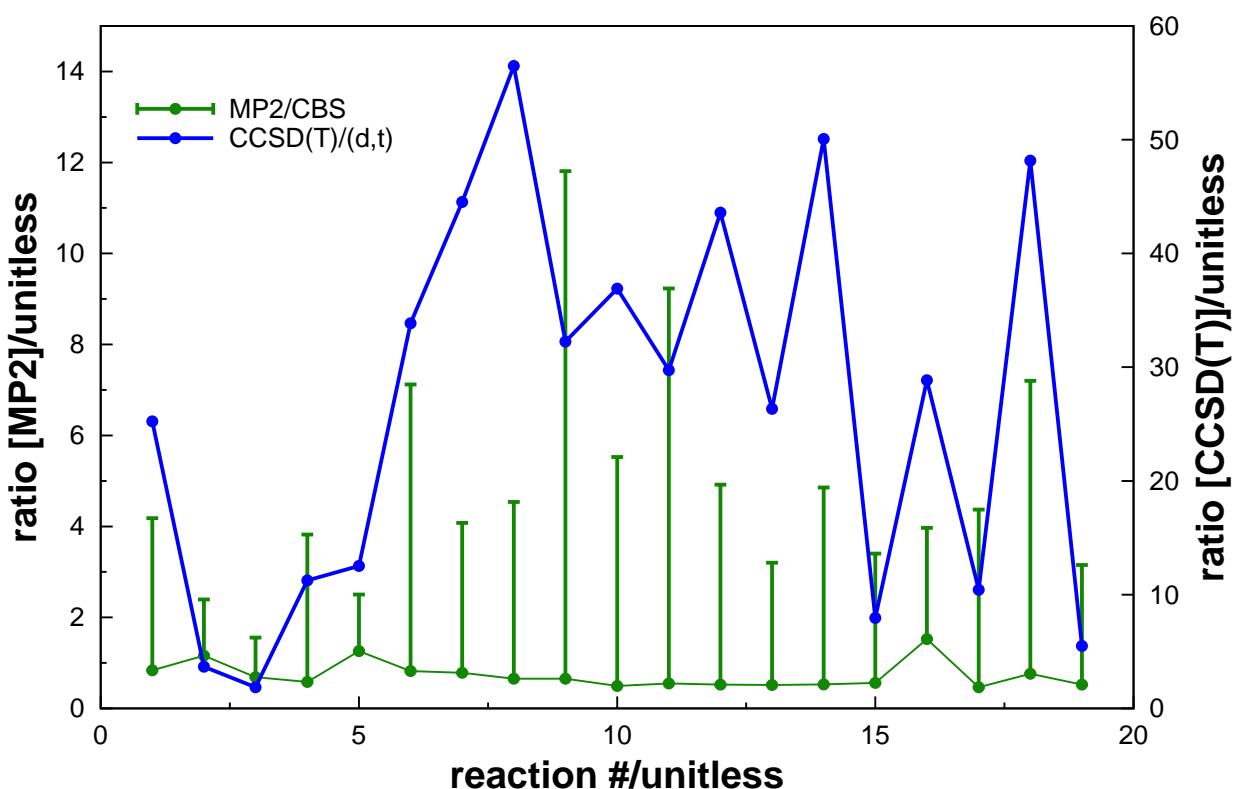


Figure 4: As in Figure 3 but for the NHTBH reactions.

direct reaction #11, $\text{F}^- + \text{CH}_3\text{Cl} \rightarrow \text{FCH}_3 + \text{Cl}^-$, with $12.74 \text{ kcal mol}^{-1}$. In turn, the larger MP2 barrier-height differences refer to the reverse HTBH reactions #9 ($12.01 \text{ kcal mol}^{-1}$), and #14 ($9.23 \text{ kcal mol}^{-1}$). Moreover, two NHTBH reactions display significant differences, #11 direct ($8.09 \text{ kcal mol}^{-1}$) and #13 direct and reverse (9.59 and $8.55 \text{ kcal mol}^{-1}$) reactions. Regarding CCSD(T)/CBS(*d, t*), reasonably small (≤ 1 and $2.5 \text{ kcal mol}^{-1}$) MAE are obtained for the HTBH and NHTBH reactions (respectively). The ones that mostly underperform are the NHTBH reactions $\text{X}^- + \text{CH}_3\text{Y} \rightarrow \text{XCH}_3 + \text{Y}^-$, which show barrier-height energy differences of 7.54 , 9.49 , and $8.06 \text{ kcal mol}^{-1}$, respectively for direct reactions #7, #11 and #13, all exhibiting negative barrier-heights [both at the MP2 and CCSD(T) levels of theory, as well as the reference itself]. Of course, a full set of experimental values might offer a more stringent ground for critical assessment but they are largely unavailable, and hence not considered here or in the databases themselves.

Figure 5 shows the errors and time ratios for the reactions of the 2pIsoE4 database. The trends found in previous cases are here too observed: time ratios close to unit, and errors

of MP2/CBS(d, t) comparable to M06-2X. Similarly, Figure 6 illustrates the corresponding errors and time ratios for the isomerization reactions of the IsoL6/11 database. Once again, the trends reported above are maintained: an almost equal performance for the energies, and a similar cost in MP2/CBS(d, t) vs M06-2X. Note that the calculated errors for the ten

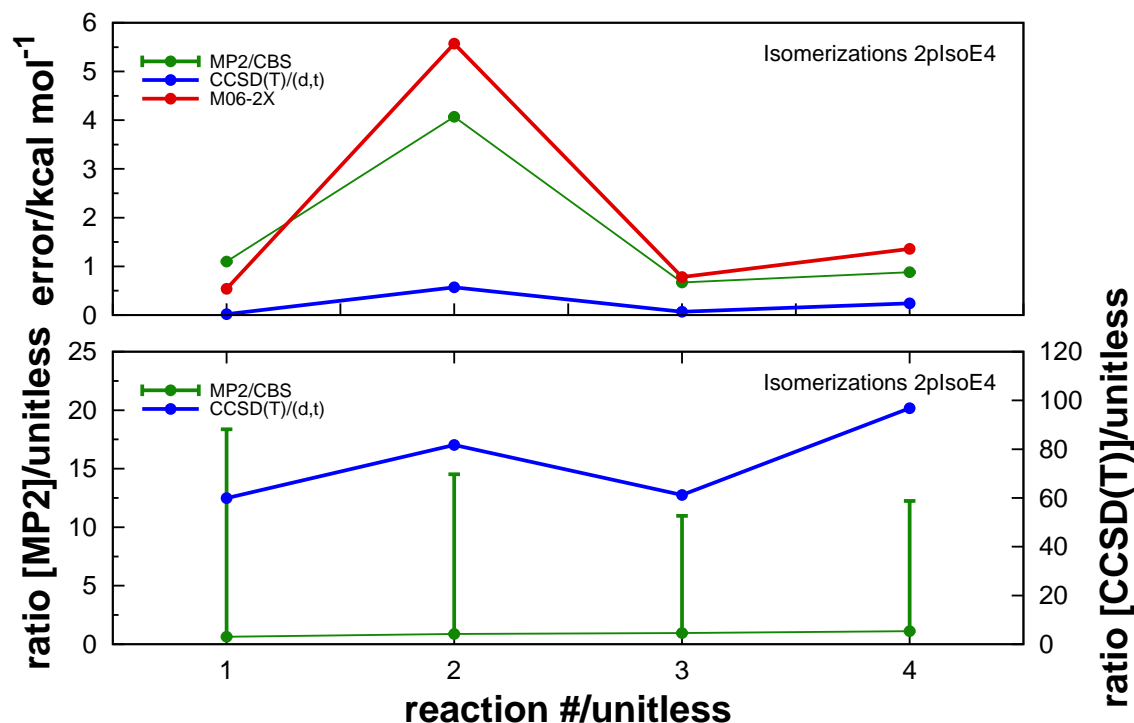


Figure 5: Errors and total time ratios for the 2p isomerization energies database. Where applicable, the left-hand-side Y-axis applies to MP2/CBS, while the rhs refers to CCSD(T)/CBS. As in previous plots, MP2/CBS(d, t) corresponds to the minimum time-ratio, and the maximum to MP2/CBS(t, q).

reactions of the databases with large molecules are basically the same using MP2/CBS(d, t) and MP2/CBS(t, q), although their cost differs by up to an order of magnitude (Figures 5 and 6). This emphasizes MP2/CBS(d, t) as one of most cost-effective tools for barrier heights of medium and large-sized systems. Naturally, standard CCSD(T) tends to be prohibitive.

4 Concluding remarks

Timewise, the most rewarding fact concerns MP2/VDZ, MP2/VTZ and MP2/CBS(d, t)

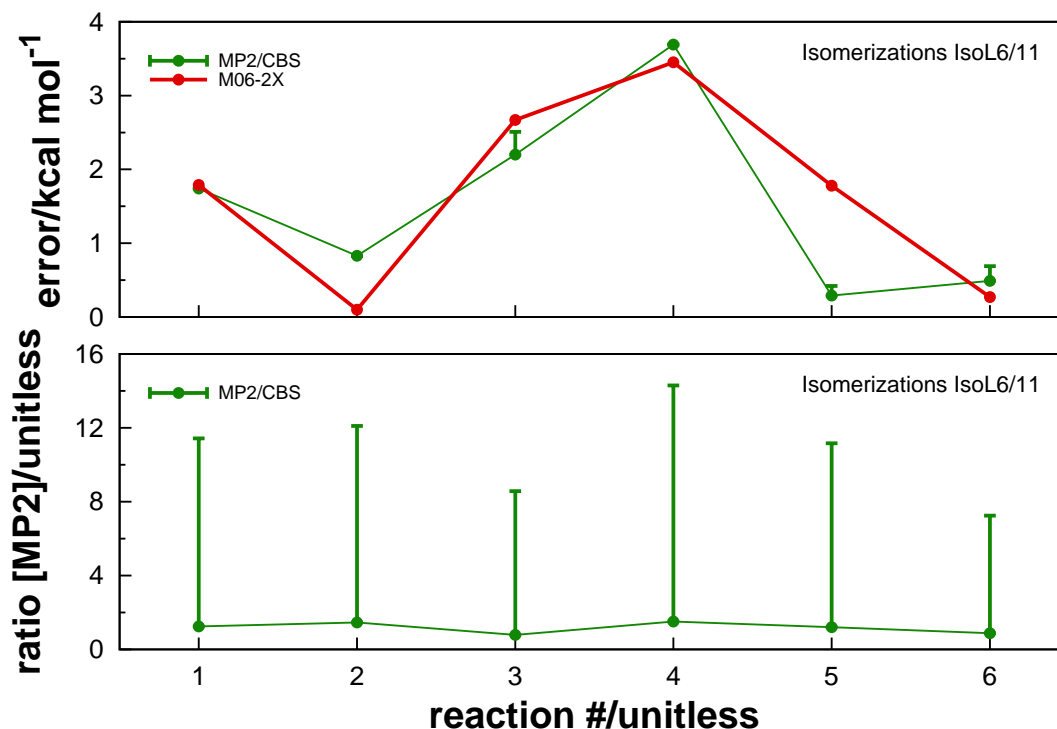


Figure 6: As in Figure 5 but for the reactions included in the isomerization database for large molecules.

which show to rival in cost with DFT/M06-2X while performing on average even cheaper than the latter irrespective of the reaction considered. Not the least important is the observation that convergence of the MP2 results can be well assessed by varying the hierarchy of the basis set. Although CBS extrapolation is also possible with DFT,⁶⁹ the final result depends critically on the selected functional since DFT appears to be less sensitive to the size of the basis. Times can too be expected to vary with the choice of functional, an issue out of the scope of the present work. Indeed, the focus has been on assessing the performance of MP2/CBS vs DFT/M06-2X results using one of the most popular functionals available for barrier heights/isomerizations. Of course, M06-2X may miss relevant dispersion interactions,⁷⁰ but this should not alter the general findings when the intermediary lies far away from the asymptotes.⁷¹ In fact, even if the calculated energetics may change, no significant differences are expected in computational cost. In this regard, calculations are being performed with other popular density functionals, the results of which will hopefully be reported

elsewhere. Suffice it to advance that MP2/CBS(d, t) clearly outperforms DFT/B3LYP-D3 using the same cc-pVXZ basis set by showing energy errors a factor of 2-3 smaller for the test set here utilized. Similarly, the DFT/B3LYP-D3 calculations are timewise more or less costly as M06-2X. Thus, the work in progress appears to confirm the trends here reported even for other functionals.

Of course, the geometries here utilized have been computed at the QCISD level which is well known to give results roughly equal to CCSD. Since the big jump in comparing the performances of MP2, CCSD and CCSD(T) is expected to lie between CCSD and CCSD(T), not between MP2 and CCSD, this offers a bias toward MP2/CBS(d, t)/QCISD/MG3. It turns out that the composite scheme M06-2X/MG3S//QCISD/MG3 has been utilized to calibrate the functional, thence suggesting a contrary conclusion although largely relying on a multiparametric fit. In fact, we should emphasize to have here redone all calculations at CCSD(T)/cc-pVXZ ($X = D, T$) and CCSD(T)/CBS(d, t) levels of theory, with the same conclusions holding as extracted from the data actually utilized. Indeed, an impartial view cannot claim M06-2X to be an unbiased predictor for the data here considered since it has been utilized for its calibration. Although no other computer packages but MOLPRO were employed for the electronic structure calculations, the main conclusions should remain valid when optimized packages for each specific method are utilized. To summarize, if to its well established reliability, one adds the reported good performance, MP2/CBS(d, t) suggests itself as the lowest-cost correlated MO-based theory capable of rivaling with KS DFT for reactions involving small, medium, and even large molecular species. Conversely to DFT, where the results depend on the chosen functional with specific functionals being optimal for specific properties, MP2 is an ab initio theory universally applicable to any molecular property. Indeed, MP2 has the advantage of describing accurately both the energy and density, a conundrum that has recently emerged for DFT.⁷² Obviously, any known weaknesses of MP2 (*e.g.*, the one due to its single-reference nature) will remain but this too must be cautioned for KS DFT. A final remark to note the work performed towards

conveying MP2 or slightly more sophisticated versions of MP theory^{73–76} the quality of CCSD(T) via spin-component scaling approaches.^{77–80} Eventually, this may lead to a linear scaling algorithm to efficiently compute intermolecular interactions,^{79,81} and hence reduce even further the cost of MP2. Definitely, both such issues jointly with the possibility of using MP2/CBS(d, t) to calculate molecular properties deserve further attention.

Acknowledgements

A.J.C.V. thanks Professor Donald Truhlar for stimulating discussions that led to perform this work, and for sending a copy of the Minnesota databases here employed. He also thanks the support of COST Action CM1405. M.M.G. thanks the financial support from the Coimbra Chemistry Centre during his stay at Coimbra for PhD studies. J.M.G.V. acknowledges the Ministerio de Educación, Cultura y Deporte of Spain for the grant (PRX2016-00300), which has permitted his stay at the University of Coimbra. This work has the support of Fundação para a Ciência e a Tecnologia, Portugal, as well as the support to the Coimbra Chemistry Centre through the project PEst-OE/QUI/UI0313/2014.

References

- (1) Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864–871.
- (2) Levy, M. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 6062–6065.
- (3) Peverati, R.; Truhlar, D. G. *Phil. Trans. R. Soc. A* **2014**, *372*, 20120476.
- (4) Kohn, W.; Sham, L. *Phys. Rev. A* **1965**, *140*, A1133–A1138.
- (5) Kohn, W. *Rev. Mod. Phys.* **1998**, *71*, 1253–1266.
- (6) Yu, H. S.; Li, S. L.; Truhlar, D. G. *J. Chem. Phys.* **2016**, *145*, 130901.

- (7) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. *Chem. Sci.* **2016**, *7*, 5032–5051.
- (8) Mardirossian, N.; Head-Gordon, M. *J. Chem. Theory Comput.* **2016**, *12*, 4303–4325.
- (9) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley: Chichester, 2000.
- (10) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; Wiley: Chichester, 2006.
- (11) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.
- (12) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (13) Becke, A. D. *J. Chem. Phys.* **1993**, *137*, 064113.
- (14) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (15) Fast, P. L.; Truhlar, D. G. *J. Chem. Phys.* **2000**, *104*, 6111–6116.
- (16) Varandas, A. J. C. *J. Phys. Chem. A* **2010**, *114*, 8505–8516.
- (17) Varandas, A. J. C.; Pansini, F. N. N. *J. Chem. Phys.* **2014**, *141*, 224113.
- (18) Pansini, F. N. N.; Varandas, A. J. C. *Chem. Phys. Lett.* **2015**, *631-632*, 70–77.
- (19) Yu, H. S.; Zhang, W.; Verma, P.; He, X.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2015**, *17*, 12146–12160.
- (20) Dunning Jr., T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (21) Dunning, Jr, T. H.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244–9253.
- (22) Handy, N. C.; Knowles, P. J.; Somasundram, K. *Theor. Chim. Acta* **1985**, *68*, 87–100.
- (23) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

- (24) Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829–5835.
- (25) Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554–8560.
- (26) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (27) Schwabe, T. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14559.
- (28) Luo, S.; Zhao, Y.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13683.
- (29) Varandas, A. J. C. *J. Phys. Chem. A* **2013**, *117*, 7393–7407.
- (30) Viegas, L. P.; Varandas, A. J. C. *J. Comput. Chem.* **2014**, *35*, 507–517.
- (31) Moran, D.; Simmonett, A. C.; Leach III, F. E.; Allen, W. D.; v. R. Schleyer, P.; Schaefer III, H. F. *J. Am. Chem. Soc.* **2006**, *128*, 9342–9343.
- (32) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265–3269.
- (33) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *Theor. Chem. Acc.* **1980**, *72*, 650–654.
- (34) Varandas, A. J. C. *J. Chem. Phys.* **2007**, *126*, 244105.
- (35) Varandas, A. J. C. *Int. J. Quantum Chem.* **2010**, *111*, 416–429.
- (36) Karton, A.; Martin, J. M. L. *Theor. Chem. Acc.* **2006**, *115*, 330–333.
- (37) Pansini, F. N. N.; Neto, A. C.; Varandas, A. J. C. *Theor Chem Acc.* **2016**, *135*, 261.
- (38) Schwartz, C. *Phys. Rev.* **1962**, *126*, 1015.
- (39) Kutzelnigg, W.; Morgan III, J. D. *J. Chem. Phys.* **1992**, *96*, 4484–4508.
- (40) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.
- (41) Varandas, A. J. C. *Physica Scripta; Comm. At. Opt. Mol. Phys.* **2007**, *76*, C28–C35.

- (42) Varandas, A. J. C. *Chem. Phys. Lett.* **2007**, *443*, 398–407.
- (43) Varandas, A. J. C. *J. Chem. Phys.* **2007**, *127*, 114316.
- (44) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; and others, M. S. MOLPRO, version 2012.1, a package of *ab initio* programs. 2012; see <http://www.molpro.net>.
- (45) Boys, F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.
- (46) Varandas, A. J. C. *Theor. Chem. Acc.* **2008**, *119*, 511–521.
- (47) Varandas, A. J. C. *J. Phys. Chem. A* **2011**, *115*, 2668–2668.
- (48) Schwenke, D. W.; Truhlar, D. G. *J. Chem. Phys.* **1985**, *82*, 2418–2426.
- (49) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *2*, 364–382.
- (50) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 43–52.
- (51) Zhao, Y.; Gonzalez-Garcia, N.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 2012–2018.
- (52) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 4811.
- (53) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2002**, *106*, 842.
- (54) Troya, D.; Lakin, M. J.; Schatz, G. C.; González, M. *J. Chem. Phys.* **2001**, *115*, 1828.
- (55) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 2715.
- (56) Diedrich, D. L.; Anderson, J. B. *J. Chem. Phys.* **1994**, *100*, 8089.
- (57) Stark, K.; Werner, H. J. *J. Chem. Phys.* **1996**, *104*, 6515.
- (58) Peterson, K. A.; Dunning, T. H. *J. Phys. Chem. A* **1997**, *101*, 6280.
- (59) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 2936.
- (60) Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput* **2009**, *5*, 808.

- (61) Zhao, Y.; González-García, N.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 2012.
- (62) Parthiban, S.; de Oliveira, G.; Martin, J. M. L. *J. Phys. Chem. A* **2001**, *105*, 895.
- (63) Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 808–821.
- (64) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968–5975.
- (65) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384–1388.
- (66) Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 569–582.
- (67) Peverati, R.; Truhlar, D. G. *Phil. Trans. R. Soc. A* **2012**, *372*, 0476.
- (68) Boese, A. D.; Martin, J. M. L.; Handy, N. C. *J. Chem. Phys.* **2003**, *119*, 3005–3014.
- (69) Chuang, Y.-Y.; Chen, S.-M. *J. Comput. Chem.* **2011**, *32*, 1671–1679.
- (70) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
- (71) Torrent-Sucarrat, M.; Varandas, A. J. C. *Chem. Eur. J.* **2016**, *22*, 14056–14063.
- (72) Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. *Science* **2017**, *355*, 49–52.
- (73) Pitoňák, M.; Neogrády, P.; Černý, J.; Grimme, S.; Hobza, P. *ChemPhysChem* **2009**, *10*, 282–289.
- (74) Riley, K. E.; Řezáč, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2011**, *13*, 21121–21125.
- (75) Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 3364–3374.
- (76) Hostaš, J.; Jakubec, D.; Laskowski, R. A.; Gnanasekaran, R.; Řezáč, J.; Vondrášek, J.; Hobza, P. *J. Chem. Theory Comput.* **2015**, *11*, 4086–4092.
- (77) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

- (78) Lochan, R. C.; Yung, Y. S.; Head-Gordon, M. *J. Phys. Chem. A* **2005**, *109*, 7598–7605.
- (79) Jung, Y.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2831–2840.
- (80) Varandas, A. J. C. *J. Chem. Phys.* **2010**, *133*, 064104.
- (81) DiStasio, R. A.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 1073–1083.