Benchmarking Density-Functionals on Structural Parameters of Small/Medium-Sized Organic Molecules

Éric Brémond,[†] Marika Savarese,[†] Neil Qiang Su,[‡] Ángel José Pérez-Jiménez,[¶]

Xin Xu,[‡] Juan Carlos Sancho-García,[¶] and Carlo Adamo[*,§,‖,†]

CompuNet, Istituto Italiano di Tecnologia, via Morego 30, I-16163 Genoa, Italy, Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, MOE Laboratory for Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200433, China, Departamento de Química Física, Universidad de Alicante, E-03080 Alicante, Spain, Chimie ParisTech, PSL Research University, CNRS, Institut de Recherche de Chimie Paris IRCP, F-75005 Paris, France, and Institut Universitaire de France, 103 Boulevard Saint Michel, F-75005 Paris, France

E-mail: carlo.adamo@chimie-paristech.fr

[*]To whom correspondence should be addressed
[†]Istituto Italiano di Tecnologia
[‡]Fudan University
[¶]Universidad de Alicante
[§]Chimie ParisTech
[‖]Institut Universitaire de France

Abstract

In this letter we report the error analysis of 59 exchange-correlation functionals in evaluating the structural parameters of small- and medium-sized organic molecules. From this analysis, recently developed double-hybrids, such as xDH-PBE0, emerge as the most reliable methods, while global-hybrids confirm their robustness in reproducing molecular structures. Notably the M06-L density-functional is the only semilocal method reaching an accuracy comparable to hybrids'. A comparison with errors obtained on energetic databases (including thermochemistry, reaction barriers and interaction energies) indicate that most of the functionals have a coherent behavior, showing low (or high) deviations on both energy and structure datasets. Only a few of them are more prone toward one of these two properties.

The quality of any method rooted in density functional theory (DFT) is (strongly) affected by the choice of the exchange-correlation functional (ECF), which gives the unknown term of the Kohn-Sham energy. If from one side the spreading of DFT in chemistry and physics has encouraged the research of new and better-performing density-functionals,[1] from the other side their validation has become a due step before any routine application. Such a benchmark passes through a careful evaluation (and consequent statistical analysis) of the errors on defined properties and systems sets.

Starting from the nineties, a large effort has been made in order to define standard benchmark sets allowing for a meaningful and fair comparison between different ECFs.[2–8] Among the properties firstly targeted, atomization energies, ionization potentials and electron affinities[2–4] as well as bond lengths and angles of (mostly) small organic systems received a particular attention.[5,6] Later, several other key properties were added to these, such as different spectroscopic observables,[9–13] gaps in solids,[8,14] lattice constants,[8,15,16] and reaction energies,[17] just to mention some.

Since performances on properties and structural parameters are generally believed to be disconnected, it is a commonly-accepted practice to carry out such benchmarks at given molecular structures. However some exceptions can be found in the literature,[18–20] mainly concerning specific cases, like transition state structures,[21–23] weakly bound interacting complexes,[24,25] conjugated polymers[26] or H-bonds.[27] Most of these latter studies showed that ECFs performing well on a given non-structural property are not necessarily the best candidates for an accurate determination of molecular geometries. Nevertheless, properties and structures are often evaluated using the same ECF, which prompts for more systematic studies.

One of the main reasons concerning the recent deficiency of benchmarks on molecular structures is the lack of accurate reference values to perform these systematic investigations. Within this framework, Barone and coworkers have recently developed two reference databases of semiexperimental equilibrium structures of semirigid organic molecules named

here **CCse21** and **B3se47**.[28–30] Whereas the former gathers a collection of 21 small organic molecules ranging from tri- to octo-atomic systems, the latter is the subsequent extension including 26 additional medium-sized organic systems dealing with various types of covalent bonds and different molecular skeletons (see Figure S1 and S2 in the Supporting Information). Both databases are an excellent diagnostic test to discriminate density-functionals in modeling structural parameters of organic systems.

In this Letter, we use these two datasets to thoroughly benchmark the accuracy of 59 ECFs (reported in Table 1), and 3 post-Hartree-Fock (post-HF) approaches derived from the second-order Møller-Plesset theory in its canonical (MP2) or spin-scaled versions (SCS- and SOS-MP2). For the sake of completeness, the Hartree-Fock (HF) values are also reported. The references and further details of all the considered computational methods involved in this Letter are given in Table S1 of the Supporting Information.

In order to discriminate the accuracy of the selected approaches, we define a criterion based on the matrix containing all the interatomic distances. For each system, we compute the mean absolute deviation (MAD) over the distance matrix of the probed and the reference geometries, and calculate the averaged deviation over the set. Figure 1 reports these statistics for the 63 computational approaches considered in this Letter (see Table S2 and S3 in the Supporting Information for more details).

For the **CCse21** dataset, the deviations span from 0.002 to 0.016 Å for xDH-PBE0 and HF methods, respectively. Within this interval, a smooth transition from high to low accuracy is observed. Apart from the worst performing ECFs like BLYP, B97D, B97D3 or TPSS, most of the methods give a slight increase of the distance matrix deviation ($\sim 1 \cdot 10^{-3}$ Å) going from the **CCse21** to the **B3se47** database. In other words, most of the methods show a coherent behavior for both small- and medium-sized semirigid organic compounds.

Going through the details, the top rank performing density-functionals is ruled by double-hybrids, and more specially by the xDH-PBE0 double-hybrid and some reparameterization variants of the B2-PLYP density-functional with deviations lower than 0.003 Å on the dis-

tance matrix criterion. These ECFs (containing a fraction of nonlocal correlation) out-perform the other considered methods for the description of minimum energy structures. They are followed by modern and highly parameterized range-separated hybrids belonging to the ωB97 family of ECFs, which are often underlined in the literature as promising approaches to model several other chemical properties,[11,27,31] and by the large panel of global-hybrids chosen to perform this investigation. Among these global-hybrids, SOGGA11-X, an exchange-hybridized variant of the semilocal SOGGA11 already highlighted for its accuracy in modeling bond lengths,[18–20] and the parameter-free PBE0 density-functionals, are part of the best performing approaches with an error of 0.004 Å for both the **CCse21** and **B3se47** datasets. The popular B3LYP global-hybrid is on the line with an error of 0.005 Å on the **CCse21** dataset, and of 0.007 Å on the **B3se47** database. Distance matrix deviations are larger than 0.007 Å for semilocal ECFs (**CCse21** database), noting that mGGAs are generally more accurate than GGAs. At this point, two remarkable results should be underlined. From one side GGA density-functionals casting empirical dispersion corrections such as B97D and B97D3 are worse performing than their parent approaches, thus suggesting problems in their parameterization procedure. On the other side M06-L is the only semilocal ECF giving results comparable to those obtained with the best performing (and more computationally expensive) hybrid density-functionals.

Another interesting issue arising from Figure 1 concerns the performance of the Minnesota density-functionals[8] excluding the already mentioned SOGGA11-X. First generation density-functionals (*e.g.*, M05, M06) are more accurate than those belonging to the second one (*e.g*, M11, N12-SX, MN12-SX). This effect could be related to the transition from global- to range-separated hybridization scheme. Note that the behavior of semilocal approaches is more difficult to rationalize even if some of the most recent ones are mainly parameterized for thermochemistry (M11-L).

Figure 2 gives some insights on the accuracy of the computational approaches to model the length of selected CH, CC and CO bonds extracted from the **CCse21** database (see Table

S4 in the Supporting Information for more details). As a general trend, CC and CO bond lengths are underestimated (negative mean deviations) while CH ones are often overestimated (positive mean deviations). Semilocal density-functionals such as BLYP, PBE, B97D, B97D3, and TPSS, or post-HF methods infringe this last assertion and show a general overestimation of all the analyzed bonds. Here again Figure 2 illustrates the excellent performances of the top rank double-hybrids, which remarks the effect of a nonlocal correlation correction to accurately model these three types of covalent bond lengths. Hybrids such as SOGGA11-X and ωB97 are close to double-hybrids, performing well for CH and CC bond types, but are inferior to double-hybrids in the case of CO. However the results on the CO bond type are less statistically meaningful due to the small number of cases.

To summarize our results, Figure 3 reports a diagram ranking the computational approaches according to their normalized distance matrix deviations obtained on the **CCse21** database. Within this diagram, the scores of the best and least performing methods (xDH-PBE0 and HF, respectively) are set to 100 and 0%, respectively. Each column gathers the methods by a 10% decrease of the score of xDH-PBE0, thus indicating an increase of the error with respect to the most accurate functional. In particular, the first left column contains 8 methods having scores ranging between 100 and 90% of the xDH-PBE0 error, and the second left column contains 18 methods scoring between 90 and 80%.

Among the 53 computational approaches investigated in this Letter, double-hybrids appear to be the best candidates. Seven of them are part of the first left column, or in other words are part of the computational methods having a normalized error smaller than 10% with respect to the xDH-PBE0 double-hybrid. Double-hybrids are still computationally too demanding to easily afford investigations on larger systems but, fortunately, non-hybrid and singly-hybridized ECFs can afford them, of course taking into account the trade-off between accuracy and computational cost. The 26 methods clustered in the two first left columns (more than 40% of the computational methods investigated in this Letter) perform worse than 20% of the xDH-PBE0 score. Other 19 methods (ranging from X3LYP to M11, third

6

left column) have scores between 80 and 70% of it, while the remaining part (18 methods) perform worse than 70% of the score of the best candidate. The first two left columns include the best density-functionals of each class of hybrids defined in Table 1. Only one, out of 26 (M06-L), is non-hybrid, thus underlining the role of nonlocal exchange to obtain accurate geometrical values. Nevertheless, this contribution does not have to excess 40% in order to get the best improvement. The best performing hybrid ECFs are PBE0 (among those casting less than 40% of exact-like exchange), and SOGGA11-X (among those casting more than 40% of exact-like exchange), and ωB97X (for range-separated hybrids).

In order to verify the existence of a relationship between the performances obtained for structural parameters and energetic properties, we plot in Figure 4 the errors for the GMTKN30[7] and CE345[8] databases with respect to the errors for molecular structures for a selection of ECFs. Energetic data are taken from the literature.[7,8,32,33] The GMTKN30 and CE345 databases are two extended sets containing 841 and 345 reference relative energies, respectively, specifically built to analyze the performance of computational methods for main group thermochemistry, kinetics and weak-interactions. In both cases, most of the selected density-functionals cluster on the left down corner of the correlation graphs, and present a deviation of 1.5 to 6.0 kcal mol^{-1} for energetic, and of 0.003 to 0.008 Å for structural properties. Qualitatively speaking, density-functionals close to the diagonal, *i.e.* small or large deviations with respect to both energetic and structural criteria, provide coherent performances on the two properties. In contrast, ECFs far from it are more specialized for one property. This is the case for M11-L which is known to give small errors on energy and B3LYP which performs better for structures than for energies. We have also to notice the significantly different behaviors on the two energetic sets observed for a few density-functionals such as BLYP and B3LYP. Their differences on predicted errors are as larger as 3 kcal mol^{-1} (BLYP), and could be related to the reduced number and different types of reactions considered in the CE345 database.

In summary, in this Letter we report a careful investigation of the performances of 59

density-functionals in computing structural parameters of small- to medium-sized semirigid organic compounds belonging to the **CCse21** and **B3se47** databases. Our results indicate that the xDH-PBE0 double-hybrid provides the lowest deviations, while PBE0 and SOGGA11-X are the best global-hybrids. The performances of the M06-L semilocal density-functional are also remarkable, while ωB97X is the most reliable range-separated hybrid. Nevertheless 23 density-functionals are particularly accurate on structural parameters, providing deviations not larger than 20% from that given by xDH-PBE0. A comparison with errors obtained for two energy databases (*i.e.* GMTKN30 and CE345) shows a global coherence of most of the considered functionals, while few of them (*i.e.* BLYP and M11-L) appear to be more prone toward one of the considered properties (structure or energy). More generally, our results indicate that structural databases should be considered in the validation step of any exchange-correlation density-functional, or even better, should be included into the training set of parameterized approaches.

Computational Methods Except when mentioned, all the computations were performed with a development version of the Gaussian suite of programs[34] using an ultrafine grid and a tight criterion for energy and geometry optimization convergence. Double-hybrid density-functionals such as B2GP-PLYP, B2T-PLYP, B2K-PLYP, B2π-PLYP, PBE0-DH, PBE0-2, PBE-QIDH and TPSS-QIDH, and post-HF methods such as SCS- and SOS-MP2 were fully implemented within the computational Gaussian code. XYG3 and xDH-PBE0 double-hybrids were implemented in the NorthWest computational Chemistry (NWChem) software package,[35] and computations with these two density-functionals were performed with similar convergence criteria. We direct the reader to Table S1 in the Supporting Information to get the original references and descriptions of the builtin and implemented density-functionals included into the softwares. The aug-cc-pVTZ Dunning-augmented triple-ζ basis set was set for all the computations.

Acknowledgement

Table 1: List of the exchange-correlation density-functionals considered in this Letter ordered according to their hybridization scheme.

non-hybrid	global-hybrid ($< 40\%$ of EXX[b])	global-hybrid ($> 40\%$ of EXX[b])	range-separated hybrid	double-hybrid
SVWN	APF	BH&HLYP	ωB97	B2-PLYP
BLYP	APFD	BMK[a]	ωB97X	B2π-PLYP
PBE	B3LYP	M05-2X[a]	ωB97XD	B2GP-PLYP
B97D	B3PW91	M06-2X[a]	CAM-B3LYP	B2K-PLYP
B97D3	B97-1	M06-HF[a]	HISSb	B2T-PLYP
SOGGA11	B97-2	SOGGA11-X	HSE06	PBE-QIDH
N12	B98		LC-ωPBE	PBE0-2
HTCH407	M05[a]		LC-BLYP	PBE0-DH
TPSS[a]	M06[a]		LC-PBE	TPSS-QIDH[a]
M06-L[a]	O3LYP		M11[a]	XYG3
M11-L[a]	PBE0		MN12-SX[a]	mPW2-PLYP
MN12-L[a]	TPSSh[a]		N12-SX	xDH-PBE0
	X3LYP			
	mPW1LYP			
	mPW1PBE			
	mPW1PW91			
	mPW3PBE			

[a]These density-functionals also depend on kinetic energy density (τ).
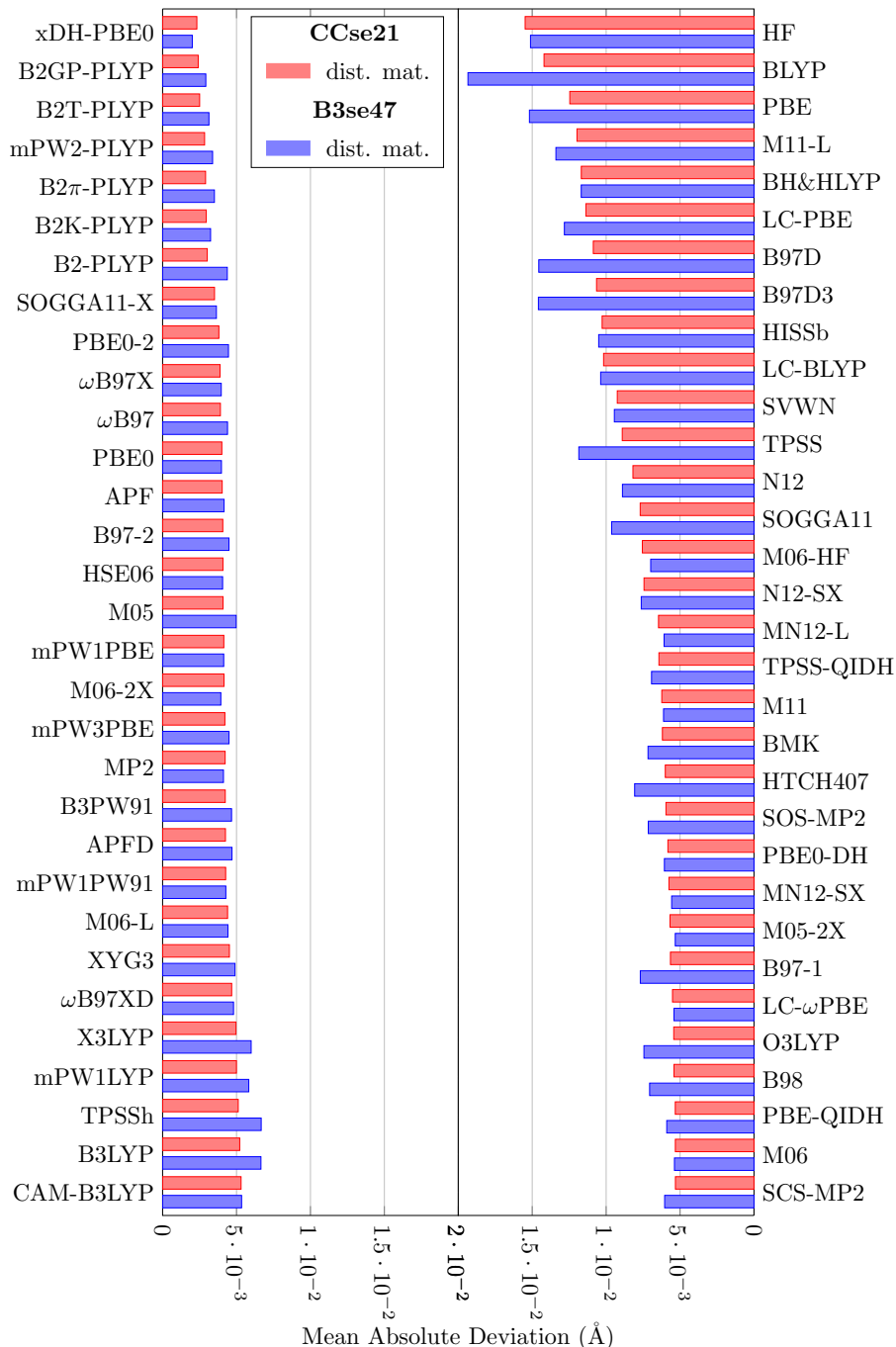
[b]EXX denotes the fraction of exact-like exchange.

Figure 1: Mean absolute deviations (in Å) of the distance matrices for the **CCse21** (red, 21 items) and **B3se47** (blue, 47 items) test sets. All the computational approaches are ranked according to their performance on the **CCse21** dataset. All the geometries are fully optimized with the aug-cc-pVTZ basis set.
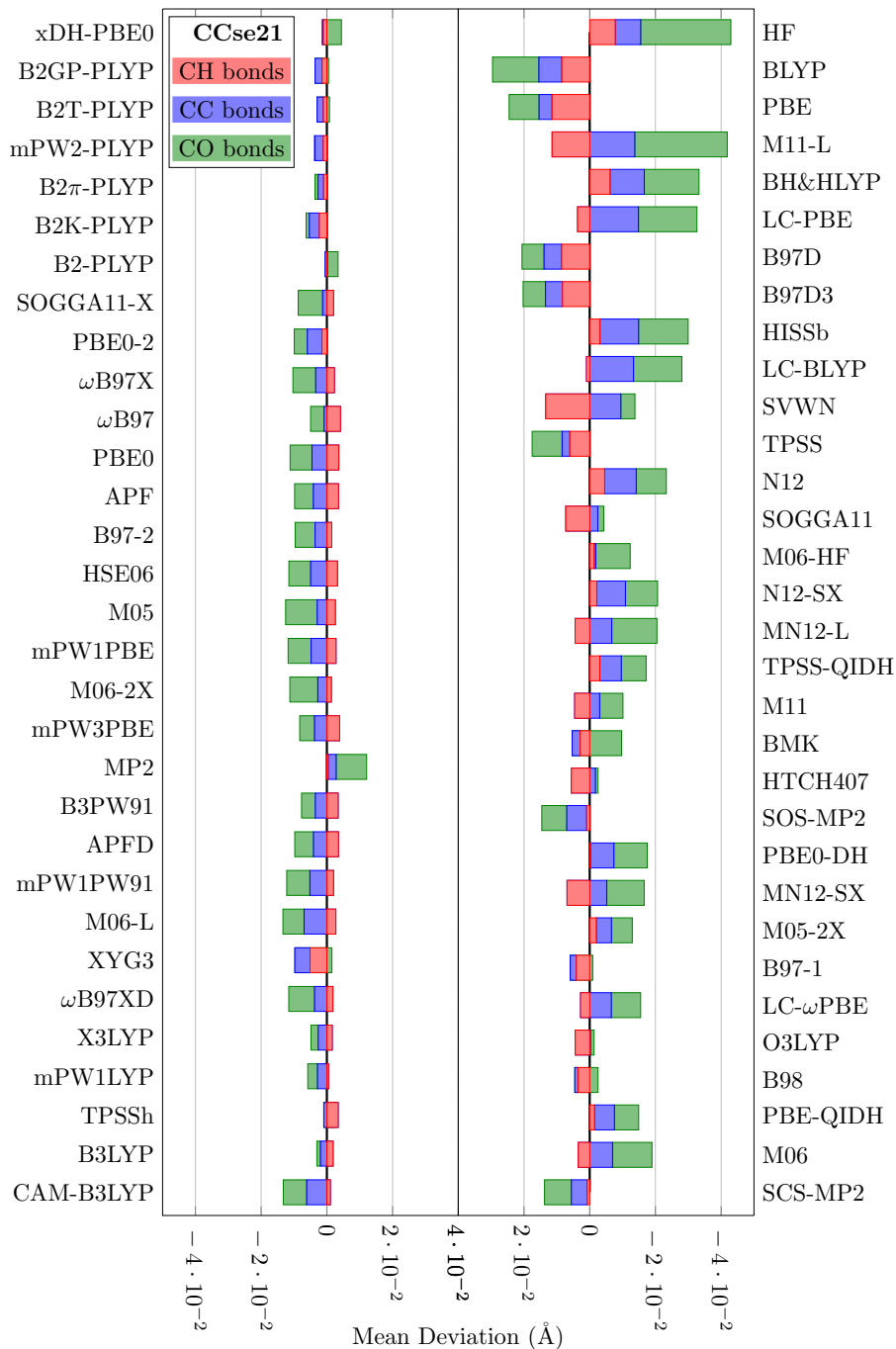
Figure 2: Cumulative mean deviations (in Å) of the CH (red, 27 items), CC (blue, 18 items) and CO (green, 7 items) bonds for the **CCse21** test set. All the geometries are fully optimized with the aug-cc-pVTZ basis set.
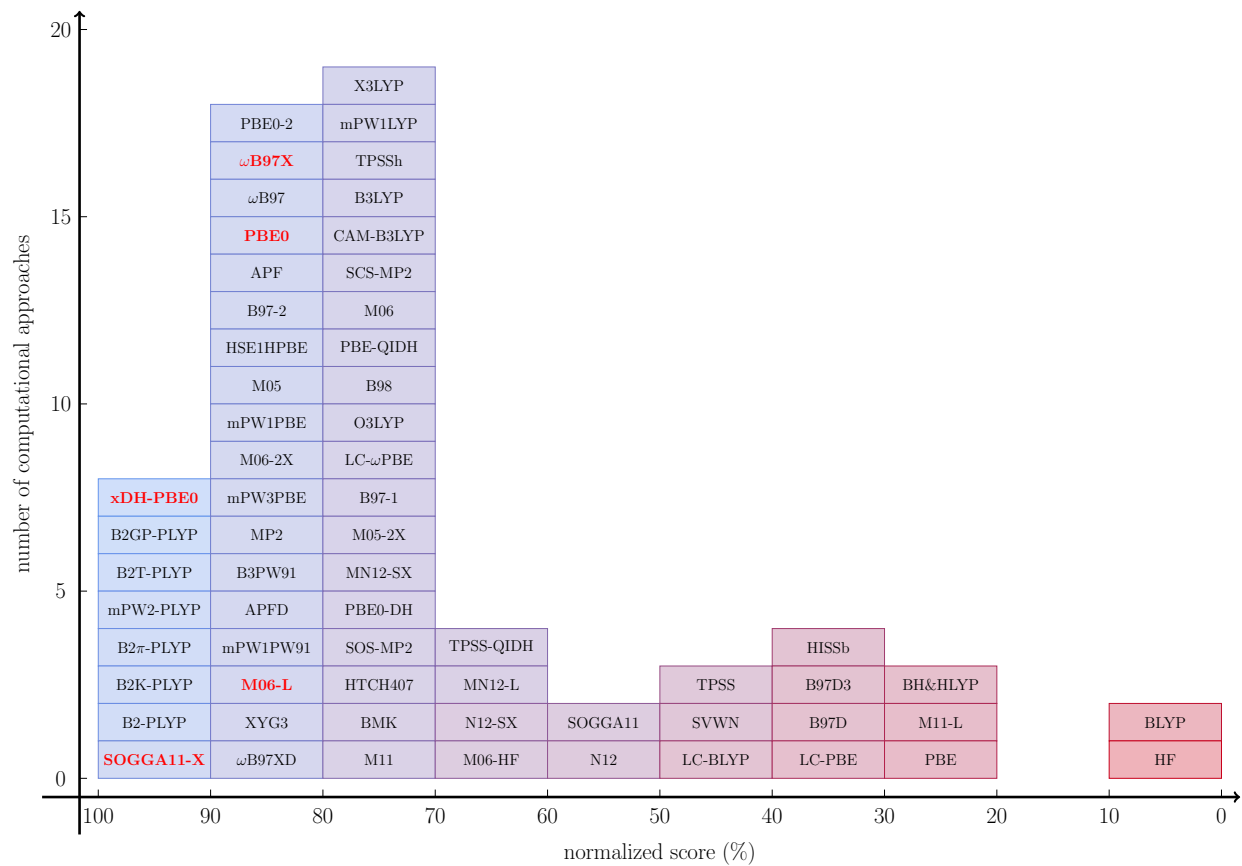
12

Figure 3: Ranking diagram of normalized performances (in %) built according to the distance matrix deviations included into the **CCse21** test set. The red to blue color shade illustrates the worst to the best performing computational approach. A red-colored bold face font denotes the best performing density-functional of each family as reported in Table 1.
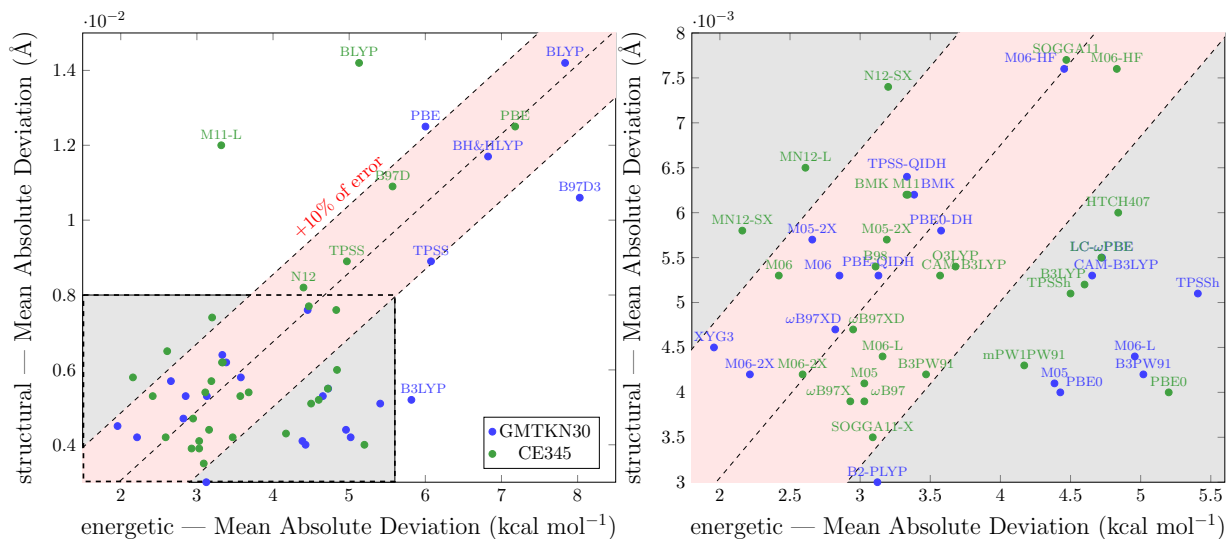
13

Figure 4: Structural (in Å) versus energetic (in kcal mol^{-1}) properties mean absolute deviations. Structure parameter deviations are computed on the **CCse21** test set, while energetic weighted total mean absolute deviations for the GMTKN30 (blue) and CE345 (green) databases are taken from references 7 and 8. The diagonal line connects the lowest and the highest errors on the **CCse21** and GMTKN30 sets, and the red area delimits the 10% error zone around this diagonal. Right graph is a zoom of the back-dashed area drawn on the left graph.

Supporting Information Available

References and descriptions of the computational approaches involved in this Letter are reported in Table S1. Representations of the molecules included into the **CCse21** and **B3se47** dataset are given in Figure S1 and S2, respectively. Individual distance matrix deviations on the **CCse21** and **B3se47** are reported in Table S2 and S3, respectively. Table S4 reports the error on the bond length for a selection of bonds derived from the **CCse21** database. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Perdew, J. P. In *Density Functional Theory: A Bridge Between Chemistry and Physics*; Geerlings, P., DeProft, F., Langenaeker, W., Eds.; Vrije Universiteit Brussel Press: Brussel, 1999.

(2) Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622–5629.

(3) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221–7230.

(4) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063–1079.

(5) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 5612–5626.

(6) Laming, G. J.; Termath, V.; Handy, N. C. *J. Chem. Phys.* **1993**, *99*, 8765–8773.

(7) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 291–309.

(8) Peverati, R.; Truhlar, D. G. *Phil. Trans. R. Soc. A* **2014**, *372*, 20120476.

(9) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(10) Jacquemin, D.; Wathelet, V.; Perpète, E. A.; Adamo, C. *J. Chem. Theory Comput.* **2009**, *5*, 2420–2435.

(11) Jacquemin, D.; Brémond, E.; Planchat, A.; Ciofini, I.; Adamo, C. *J. Chem. Theory Comput.* **2011**, *7*, 1882–1892.

(12) Isegawa, M.; Truhlar, D. G. *J. Chem. Phys.* **2013**, *138*, 134111.

(13) Muniz-Miranda, F.; Pedone, A.; Battistelli, G.; Montalti, M.; Bloino, J.; Barone, V. *J. Chem. Theory Comput.* **2015**, *11*, 5371–5384.

(14) Civalleri, B.; Presti, D.; Dovesi, R.; Savin, A. *Chemical Modelling: Applications and Theory Volume 9*; The Royal Society of Chemistry, 2012; Vol. 9; pp 168–185.

(15) Janthon, P.; Luo, S. A.; Kozlov, S. M.; Viñes, F.; Limtrakul, J.; Truhlar, D. G.; Illas, F. *J. Chem. Theory Comput.* **2014**, *10*, 3832–3839.

(16) Labat, F.; Brémond, E.; Cortona, P.; Adamo, C. *J. Mol. Mod.* **2013**, *19*, 2791–2796.

(17) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.

(18) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2008**, *128*, 184109.

(19) Peverati, R.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. Lett.* **2011**, *2*, 1991–1997.

(20) Peverati, R.; Truhlar, D. G. *J. Chem. Phys.* **2011**, *135*, 191102.

(21) Su, N. Q.; Adamo, C.; Xu, X. *J. Chem. Phys.* **2013**, *139*, 174106.

(22) Xu, X.; Alecu, I. M.; Truhlar, D. G. *J. Chem. Theory Comput.* **2011**, *7*, 1667–1676.

(23) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 2936–2941.

(24) Grimme, S.; Steinmetz, M. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031–16042.

(25) Steinmann, S. N.; Corminboeuf, C. *J. Chem. Theory Comput.* **2011**, *7*, 3567–3577.

(26) Wykes, M.; Su, N. Q.; Xu, X.; Adamo, C.; Sancho-García, J.-C. *J. Chem. Theory Comput.* **2015**, *11*, 832–838.

(27) Mangiatordi, G. F.; Brémond, E.; Adamo, C. *J. Chem. Theory. Comput.* **2012**, *8*, 3082–3088.

(28) Piccardo, M.; Penocchio, E.; Puzzarini, C.; Biczysko, M.; Barone, V. *J. Phys. Chem. A* **2015**, *119*, 2058–2082.

(29) Penocchio, E.; Piccardo, M.; Barone, V. *J. Chem. Theory Comput.* **2015**, *11*, 4689–4707.

(30) Reference structures were downloaded from http://dreamslab.sns.it/molecules/index.html the April, 20th 2015.

(31) Steinmann, S. N.; Corminboeuf, C. *J. Chem. Theory. Comput.* **2012**, *8*, 4305–4316.

(32) Bousquet, D.; Brémond, E.; Sancho-García, J. C.; Ciofini, I.; Adamo, C. *J. Chem. Theory Comput.* **2013**, *9*, 3444–3452.

(33) Brémond, E.; Sancho-García, J. C.; Pérez-Jiménez, A. J.; Adamo, C. *J. Chem. Phys.* **2014**, *141*, 031101–031104.

(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.;

Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian Development Version Revision H.11. 2010; Gaussian Inc. Wallingford CT.

(35) Valiev, M.; Bylaska, E.; Govind, N.; Kowalski, K.; Straatsma, T.; Dam, H. V.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T.; de Jong, W. *Comput. Phys. Commun.* **2010**, *181*, 1477–1489.

Graphical TOC Entry