



HAL
open science

When does a functional correctly describe both the structure and the energy of the transition state?

Neil Qiang Su, Pascal Pernot, Xin Xu, Andreas Savin

► To cite this version:

Neil Qiang Su, Pascal Pernot, Xin Xu, Andreas Savin. When does a functional correctly describe both the structure and the energy of the transition state?. *Journal of Molecular Modeling*, 2017, 23 (2), pp.65. 10.1007/s00894-017-3229-8 . hal-01472132

HAL Id: hal-01472132

<https://hal.sorbonne-universite.fr/hal-01472132>

Submitted on 20 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

When does a functional correctly describe both the structure, and the energy of the transition state?

Neil Qiang Su¹, Pascal Pernot^{2,3}, Xin Xu¹, and Andreas Savin^{4,5}

¹Collaborative Innovation Center of Chemistry for Energy Materials, Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, MOE Key Laboratory of Computational Physical Sciences, Department of Chemistry, Fudan University, Shanghai 200433, China

²CNRS, UMR8000, Laboratoire de Chimie Physique, F-91405 Orsay, France

³Univ. Paris-Sud, UMR8000, Laboratoire de Chimie Physique, F-91405 Orsay, France

⁴CNRS, UMR7616, Laboratoire de Chimie Théorique, F-75005 Paris, France

⁵UPMC Univ Paris 06, UMR7616, Laboratoire de Chimie Théorique, F-75005, Paris, France

October 28, 2016

Abstract

Requiring that several properties are well reproduced is a severe test on density functional approximations. This can be assessed through the estimation of joint and conditional success probabilities. An example is provided for a small set of molecules, for properties characterizing the transition states (geometries and energies).

Keywords benchmarks, density functional theory, method selection, uncertainty quantification

Introduction

Very often, an approximate calculation method is used to describe more than a single property. One may need two properties well described, e.g., the ionization potential and the electron affinity in order to describe the fundamental gap that is the difference between the two [1-4]. In another typical situation, some prior reliable information about one of the properties is verifiable, but another property is of interest. In such a case, the approximate calculation is performed for both properties, and the approximation used is considered unreliable (and is rejected), if the agreement between the calculated and verifiable property is unsatisfactory. For example, if an approximate calculation provides a lattice constant too far from a measured one, the confidence in the (not known) bulk modulus of the same system may be shattered, and the result of the calculation may not be published. In a previous paper [5], some of us have pointed out a simple logic on the use of benchmarks for multiple properties, which states that a given approximation works best for each of the properties separately does not necessarily work best in the study cases mentioned above. The reason for this behavior is that the errors made by an approximate method for one of the properties is not necessarily positively correlated with those for another property.

In the present paper, we choose to explore two important properties for reactions, through a data set characterizing molecular transition states, by both geometries and energies. The aim of our paper is to show that the requirements that

1. two properties are simultaneously well reproduced for a given system, or

2. one of the properties is well reproduced, knowing that the other is well reproduced

are strong, and selective, which are more severe than those usually considered.

Our paper should not be seen as recommending or banning using one of the 36 mentioned approximate methods showing up in the results section. The names of the studied functionals are presented only for illustration purpose only, and not for ranking. One reason is that the test set is small. Hence, the measures produced out of it are prone to large statistical errors. And another is that the reliability of the reference data is often not well quantified. Nevertheless, we believe that the present data set can be used as a plausible exemplification to show (i) that using the stronger requirements may be a more selective way to choose approximate methods, and (ii) that larger samples may be needed for the tests we propose.

Technical details

Reference data

We have considered the transition states for the following 12 reactions [6] (see Figure 1), which include three hydrogen transfer (HT) reactions, three heavy-atom transfer (HAT) reactions, three nucleophilic substitution (NS) reactions of anions, and three unimolecular and association (UA) reactions. These are subsets that belong to the widely used data sets developed by Zhao and Truhlar [7,8], NHTBH38 and HTBH38, each of which consists of 38 data points for non-hydrogen transfer barrier heights and hydrogen transfer barrier heights.

For these 12 reactions used in the present work, the geometric structures for the transition states have been used to set up a new dataset called TSG36 [9]. For the transition states (A . . . B . . . C, as shown in Figure 1), each is characterized by three distances (A . . . B, B . . . C, and C . . . A), involving the breaking and the forming bonds that make a total of 36 data points for the transition state geometries. As recommended [9], the reference geometric data are generally the BMC-CCSD [10] values.

In consonance with the TSG36 set, we coin the TSE36 set, where each reaction is characterized by three energies (forward barrier, backward barrier, and difference between the two). For the bimolecular nucleophilic substitution (S_N2) reaction between $Cl^- + CH_3Cl$, there exists a pre-complexation $[Cl^- \dots CH_3Cl]$ before forming the transition state $[Cl \dots CH_3 \dots Cl]^-$. The two barriers are with respect to the free reactants $Cl^- + CH_3Cl$ and the complex $[Cl^- \dots CH_3Cl]$, respectively. Hence the difference between the two barriers corresponds to the complexation energy. As recommended [7,8], the reference energy data are generally the W1 [11] values calculated at the optimized geometries at the QCISD/MG3 [12,13] level.

Calculation results

There is a plethora of approximate density functional methods. The acronyms for 33 representatives are given in the Appendix, which include doubly hybrid functionals [14-18], hybrid functionals [19-33], range-separated [34-37] or long range corrected [38,39] functionals, meta-GGAs [40,41] and GGAs (generalized gradient approximations) [42-47]. For some functionals, the dispersion corrections are applied [18,21,36,42,43]. The

results of LDAs (local density approximations) are not considered, as it is well-known that LDAs always significantly underestimate the barrier heights [7,8,14-16] and often fail to locate the desired transition states [48-50]. In fact, GGAs, and meta-GGAs can also lead to erroneous transition state structures [9, 48-50], while even hybrid functionals sometimes give unacceptable geometries [51,52].

For comparison, calculation results from some wavefunction methods, QCISD, MP2 and Hartree-Fock (HF), are also included.

Some of the results may be found in the literature [52-54], while, for the completeness, we have summarized all geometric and energetic results here and provided them as the supplementary material. For single point energies, the basis sets used are G3Large [55]. For geometry optimizations, the basis sets used are 6-311+G(3df,2p) [56]. The xDH-PBE0 data are taken from Ref. 53, where the basis sets used were 6-311+G(3df,2p).

Probabilities

We consider such properties as geometrical parameters (noted g), and energy differences (noted e). For each of the properties, we consider an approximation successful, if it reproduces all reference data to a given accuracy, e.g., errors ≤ 3 pm for interatomic distances, and ≤ 3 kcal/mol for energy differences.

Let $N(p)$ be the number of systems for which property $p \in \{g, e\}$ is successfully reproduced by the approximation under consideration. Furthermore, let $N(g \cap e)$ be the number of systems for which both properties are correctly reproduced. Finally, let N be

the number of systems in the benchmark set. For each of the functionals we calculate the probabilities

- to have success for a given property, $P(p) = N(p)/N$;
- to have success for both properties, $P(g \cap e) = N(g \cap e)/N$;
- to have success for the distances, once the energy differences has been checked

to be acceptable, $P(g|e) = N(g \cap e)/N(e)$; and

- to have success for the energy differences, once the geometrical parameters have been checked to be acceptable, $P(e|g) = N(g \cap e)/N(g)$.

The approximation is considered successful, for any of the $N=12$ reactions, when all three distances and/or all three energies are in error by less than the acceptance threshold.

Results and discussion

Some standard statistical indicators of the quality of an approximation (maximum absolute error, MAX; mean signed error, MSE; mean absolute deviation, MAD; root mean square deviation, RMSD) can be found in the supporting information. We notice that some methods can have very large maximal errors (MAX, up to 150 pm for geometries, or 40 kcal/mol for the energies), and that these will not disappear when enlarging the data set. Also sometimes a large bias (MSE) can be observed, and an important dispersion of the data (RMSD).

The correlation between the MAD and RMSD for each property and the corresponding success probabilities $P(g)$ and $P(e)$ are shown in Figure 2. Although one

observes a general tendency to a negative correlation, it is very lax. The best correlation is observed for energies, between $P(e)$ and the MAD. The Spearman correlation coefficient in this case is -0.9, and falls to -0.7 for the geometries. It appears therefore that the usual statistics cannot generally be used as proxies for the success probabilities, which encompass different information. In particular, the probabilities allow producing measures for simultaneous, or conditional success that forms the object of our paper.

Figure 3 shows the estimated probabilities for the following acceptance thresholds: 3 pm for distances, and 3 kcal/mol for energies. Despite these quite generous thresholds, we notice that only a few approximations reach the (modest) value of 0.5 or larger for all probabilities (XYG3, XYGJ-OS, xDH-PBE0). Even though some approximate methods can reach larger values for some of the probabilities, only these three methods get above 0.50 for $P(g \cap e)$. The limited accuracy is not a feature of density functional approximations only, but also of wave function methods: QCISD produces $P(g)$, and $P(e)$ larger than 0.5, but both $P(g \cap e)$ and $P(e|g)$ are smaller than 0.5. This method provides an example where both $P(g)$, and $P(e)$ are considered acceptable, but not $P(g \cap e)$.

If one considers conditional probabilities, many methods achieve values above 0.5. For instance, M06-L works apparently extremely well: both conditional probabilities equal to one, i.e., success for one property guarantees success for the other. However, we notice that both $P(g)$ and $P(e)$ are very small (1/6). Thus, it cannot be argued that the result is statistically significant with $N(g)=N(e)=2$. The conditional probabilities should not be considered when the success probabilities for individual properties are

low.

If we choose tighter, i.e., more reasonable in the sense that they are closer to chemical accuracy, acceptance thresholds (1.5 pm and 1.5 kcal/mol), none of the methods pass even the loose criterion of “rather acceptable than unacceptable” result for both properties, $P(g \cap e) \geq 0.5$, cf. Figure 4. Of course, the situation of the conditional probabilities cannot be considered as being improved, as $N(g)$ and $N(e)$ cannot be larger than those of the looser thresholds.

Let us now lower our requirements, and put the thresholds to 4.5 pm and 4.5 kcal/mol (Figure 5). Now, while many functionals still fail to give $P(g \cap e) \geq 0.5$, we can remark that the situation is improved for M06-2X, the range-separated functionals ω B97X, ω B97X-D, LC- ω PBE, as well as QCISD. Notice that although both $P(g)$ and $P(e)$ are better for M06-2X than for ω B97X-D, $P(g \cap e)$ has the same value for both methods.

Finally, let us consider a more satisfying selection probability for "good" methods, i.e., we require success probabilities larger than 0.68 (as for staying within one standard deviation for a normal distribution). Going back to the initially discussed “generous” thresholds of 3 pm and 3 kcal/mol (Figure 3), we see that none of the method yields a high enough probability for $P(g \cap e)$.

Conclusion

Transition state geometries and energy differences obtained using 36 approximate methods have been compared with reference data. We do not claim having established

a ranking of the approximations: there are many other functionals, e.g. Refs. 57-59, that could have been tested, and there are issues related to the size of the benchmark set, and of technical nature, e.g. the basis sets used. Hence, the objective of the present paper is different from some previous work, e.g. Ref. 60, where benchmarking study has been carried out to examine the functional performances on structural parameters and energies separately.

However, our present data are sufficient to illustrate that situations can arise when

- 1) different properties are required for a given system, or
- 2) a priori information can be used to accept or reject an approximation.

These requirements are stronger than those usually used, and could be more effectively used as selection criteria. For example, for a very generous acceptance of errors in distances not larger than 3 pm, and in energy differences not larger than 3 kcal/mol, we found that less than ten percent of the studied methods were able to reproduce the reference results with a probability larger than $\frac{1}{2}$.

Acknowledgment Henry Chermette was an early fighter on the path of density functional approximations. Applying them carefully was one of his main concerns. It is thus our pleasure to dedicate him our paper. This work was supported by the Ministry of Science and Technology (2013CB834606, 2011CB808505), and National Natural Science Foundation of China (91027044, 21133004).

Appendix

The acronyms for 33 representative density functional approximations.

Acronym	Type	References
XYG3	Double hybrid	[14]
XYGJ-OS	Double hybrid	[15]
xDH-PBE0	Double hybrid	[16]
B2PLYP	Double hybrid	[17]
B2PLYP-D	Double hybrid plus dispersion correction	[18]
B3LYP	Hybrid GGA	[19,20]
B3LYP-D3	Hybrid GGA plus dispersion correction	[19-21]
B3PW91	Hybrid GGA	[19]
PBE0	Hybrid GGA	[22,23]
PBE0-D3	Hybrid GGA plus dispersion correction	[21-23]
O3LYP	Hybrid GGA	[24]
X3LYP	Hybrid GGA	[25,26]
BHhLYP	Hybrid GGA	[27]
M06-2X	Hybrid meta-GGA	[28]
M06	Hybrid meta-GGA	[28]
TPSSh	Hybrid meta-GGA	[29]
BMK	Hybrid meta-GGA	[30]
B97-1	Hybrid GGA	[31]
B97-2	Hybrid GGA	[32]

B98	Hybrid GGA	[33]
CAM-B3LYP	Range-separated hybrid GGA	[34]
ω B97X	Range-separated hybrid GGA	[35]
ω B97X-D	Range-separated hybrid GGA plus dispersion correction	[36]
HSE06	Range-separated hybrid GGA	[37]
LC- ω PBE	Long range corrected hybrid GGA	[38]
LC-PBE	Long range corrected GGA	[39]
TPSS	meta-GGA	[40]
M06-L	meta-GGA	[41]
B97D	GGA plus dispersion correction	[42]
B97D3	GGA plus dispersion correction	[43]
HCTH407	GGA	[44]
BLYP	GGA	[45,46]
PBE	GGA	[47]

References

1. Su NQ, Zhang IY, Wu JM, Xu X (2011) *Front Chem China* 6:269
2. Su NQ, Yang WT, Mori-Sánchez P, Xu X (2014) *J Phys Chem A* 118:9201
3. Su NQ, Xu X (2015) *J Chem Theory Comput* 11:4677
4. Su NQ, Xu X (2016) *J Chem Theory Comput* 12:2285
5. Civalleri B, Dovesi R, Pernot P, Presti D, Savin A (2016) *Computation* 4:20

6. Zheng J, Zhao Y, Truhlar DG (2009) *J Chem Theory Comput* 5:808
7. Zhao Y, Gonzales-Garcia N, Truhlar DG (2005) *J Phys Chem A* 109:2012
8. Zhao Y, Lynch BJ, Truhlar DG (2005) *Phys Chem Chem Phys* 7:43
9. Xu X, Alecu IM, Truhlar DG (2011) *J Chem Theory Comput* 7:1667
10. Lynch BJ, Zhao Y, Truhlar DG (2005) *J Phys Chem A* 109:1643
11. Martin JML, de Oliveira G (1999) *J Chem Phys* 111:1843
12. Pople JA, Head-Gordon M, Raghavachari K (1987) *J Chem Phys* 87:5968
13. Lynch BJ, Zhao Y, Truhlar DG (2003) *J Phys Chem A* 107:1384
14. Zhang Y, Xu X, Goddard III WA (2009) *Proc Natl Acad Sci USA* 106:4963
15. Zhang IY, Xu X, Jung Y, Goddard III WA (2011) *Proc Natl Acad Sci USA* 108:19896
16. Zhang IY, Su NQ, Brémond É, Adamo C, Xu X (2012) *J Chem Phys* 136:174103
17. Grimme S (2006) *J Chem Phys* 124:034108
18. Schwabe T, Grimme S (2007) *Phys Chem Chem Phys* 9:3397
19. Becke AD (1993) *J Chem Phys* 98:5648
20. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) *J Phys Chem* 98:11623
- 21 Grimme S, Antony J, Ehrlich S, Krieg H (2010) *J Chem Phys* 132:154104
22. Adamo C, Barone V (1999) *J Chem Phys* 110:6158
23. Ernzerhof M, Scuseria GE (1999) *J Chem Phys* 110:5029
24. Cohen AJ, Handy NC (2001) *Mol Phys* 99:607
25. Xu X, Goddard III WA (2004) *Proc Natl Acad Sci USA* 101:2673
26. Xu X, Zhang Q, Muller RP, Goddard III WA. (2005) *J Chem Phys* 122:014105

27. Becke AD (1993) J Chem Phys 98:1372
28. Zhao Y, Truhlar DG (2008) Theor Chem Acc 120:215
29. Staroverov VN, Scuseria GE, Tao J, Perdew JP (2003) J Chem Phys 119:12129
30. Boese AD, Martin JML (2004) J Chem Phys 121:3405
31. Hamprecht FA, Cohen AJ, Tozer DJ, Handy NC (1998) J Chem Phys 109:6264
32. Wilson PJ, Bradley TJ, Tozer DJ (2001) J Chem Phys 115:9233
33. Schmider HL, Becke AD (1998) J Chem Phys 108:9624
34. Yanai T, Tew DP, Handy NC (2004) Chem Phys Lett 393:51
35. Chai J-D, Head-Gordon M (2008) J Chem Phys 128:084106
36. Chai J-D, Head-Gordon M (2008) Phys Chem Chem Phys 10:6615
37. Krukau AV, Vydrov OA, Izmaylov AF, Scuseria GE (2006) J Chem Phys 125:224106
38. Vydrov OA, Scuseria GE (2006) J Chem Phys 125:234109
39. Iikura H, Tsuneda T, Yanai T, Hirao K (2001) J Chem Phys 115:3540
40. Tao J, Perdew JP, Staroverov VN, Scuseria GE (2003) Phys Rev Lett 91:146401
41. Zhao Y, Truhlar DG (2006) J Chem Phys 125:194101
42. Grimme S (2006) J Comput Chem 27:1787
43. Grimme S, Ehrlich S, Goerigk L (2011) J Comput Chem 32:1456
44. Boese AD, Handy NC (2001) J Chem Phys 114:5497
45. Becke AD (1988) Phys Rev A 38:3098
46. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785
47. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865

48. Su NQ, Chen J, Su Z, Zhang DH, Xu X (2015) *J Chem Phys* 142:084107
49. Su NQ, Chen J, Xu X, Zhang DH (2016) *Acta Phys-Chim Sin* 32:119
50. Su NQ, Xu X (2016) *Mol Phys* 114:1207
51. Curtiss LA, Redfern PC, Raghavachari K (2010) *Chem Phys Lett* 499:168
52. Su NQ, Xu X, *Chem Comm* 2016, 52: 13840
53. Su NQ, Adamo C, Xu X (2013) *J Chem Phys* 139:74106
54. Su NQ, Xu X (2013) *Sci Sin Chim* 43:1761
55. Curtiss LA, Raghavachari K, Redfern PC, Rassolov V, Pople JA (1998) *J Chem Phys* 109:7764
56. Curtiss LA, Raghavachari K, Trucks GW, Pople JA (1991) *J Chem Phys* 94:7221
57. Brémond É, Sancho-García, JC, Pérez-Jiménez ÁJ, Adamo C (2014) *J Chem Phys* 141:031101
58. Yu HS, He X, Truhlar DG (2016) *J Chem Theory Comput* 12:1280
59. Mardirossian N, Head-Gordon M (2016) *J Chem Phys* 144:214110
60. Brémond É, Savarese M, Su NQ, Pérez-Jiménez ÁJ, Xu X, Sancho-García, JC, Adamo C (2016) *J Chem Theory Comput* 12:459

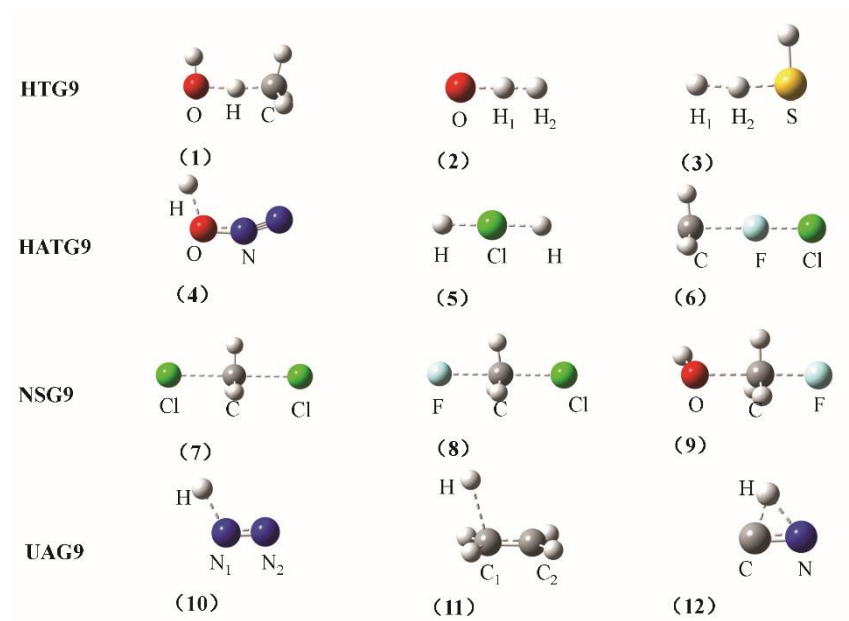


Fig. 1 Transition state structures in the TSG36 [9] set, which include three hydrogen transfer (HT) reactions, three heavy-atom transfer (HAT) reactions, three nucleophilic substitution (NS) reactions of anions, and three unimolecular and association (UA) reactions. Each transition state structure is characterized by three distances (A . . . B, B . . . C, and C . . . A), involving the breaking and the forming bonds that make a total of 9 data points (G9) for each type of reaction.

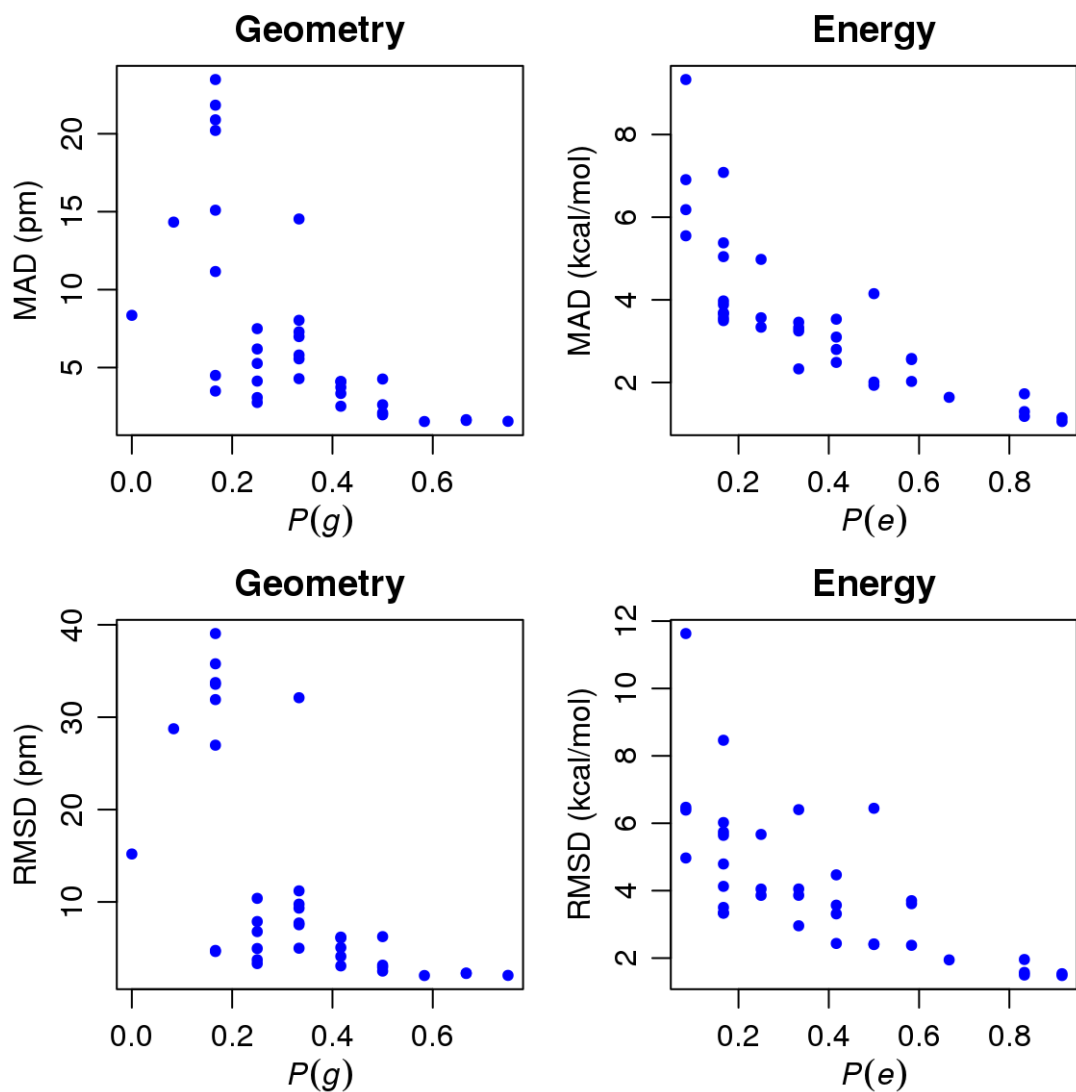


Fig. 2 Correlations between the MAD and RMSD for each property of energy and geometry and the corresponding success probabilities $P(g)$ and $P(e)$.

	$P(g)$	$P(e)$	$P(g \cap e)$	$P(g e)$	$P(e g)$	
XYG3	0.67	0.83	0.58	0.70	0.88	XYG3
XYGJ-OS	0.67	0.92	0.58	0.64	0.88	XYGJ-OS
xDH-PBE0	0.58	0.83	0.58	0.70	1.00	xDH-PBE0
B2PLYP	0.42	0.83	0.33	0.40	0.80	B2PLYP
B2PLYP-D	0.50	0.33	0.25	0.75	0.50	B2PLYP-D
B3LYP	0.33	0.17	0.08	0.50	0.25	B3LYP
B3LYP-D3	0.33	0.17	0.08	0.50	0.25	B3LYP-D3
B3PW91	0.33	0.25	0.08	0.33	0.25	B3PW91
PBE0	0.42	0.33	0.25	0.75	0.60	PBE0
PBE0-D3	0.42	0.17	0.17	1.00	0.40	PBE0-D3
O3LYP	0.25	0.42	0.08	0.20	0.33	O3LYP
X3LYP	0.33	0.17	0.08	0.50	0.25	X3LYP
BHHLYP	0.17	0.42	0.17	0.40	1.00	BHHLYP
M06-2X	0.50	0.92	0.42	0.45	0.83	M06-2X
M06	0.25	0.58	0.17	0.29	0.67	M06
TPSSh	0.33	0.08	0.08	1.00	0.25	TPSSh
BMK	0.25	0.67	0.25	0.38	1.00	BMK
B97-1	0.33	0.33	0.17	0.50	0.50	B97-1
B97-2	0.25	0.58	0.17	0.29	0.67	B97-2
B98	0.25	0.42	0.17	0.40	0.67	B98
CAM-B3LYP	0.42	0.42	0.17	0.40	0.40	CAM-B3LYP
ω B97X	0.25	0.50	0.25	0.50	1.00	ω B97X
ω B97X-D	0.42	0.50	0.33	0.67	0.80	ω B97X-D
HSE06	0.50	0.17	0.17	1.00	0.33	HSE06
LC- ω PBE	0.42	0.33	0.25	0.75	0.60	LC- ω PBE
LC-PBE	0.33	0.17	0.08	0.50	0.25	LC-PBE
TPSS	0.17	0.08	0.08	1.00	0.50	TPSS
M06-L	0.17	0.17	0.17	1.00	1.00	M06-L
B97-D	0.17	0.17	0.08	0.50	0.50	B97-D
B97-D3	0.17	0.17	0.08	0.50	0.50	B97-D3
HCTH407	0.08	0.50	0.08	0.17	1.00	HCTH407
BLYP	0.17	0.08	0.08	1.00	0.50	BLYP
PBE	0.17	0.17	0.08	0.50	0.50	PBE
QCISD	0.75	0.58	0.33	0.57	0.44	QCISD
MP2	0.17	0.25	0.00	0.00	0.00	MP2
HF	0.00	0.08	0.00	0.00	0.00	HF
	$P(g)$	$P(e)$	$P(g \cap e)$	$P(g e)$	$P(e g)$	

Fig. 3 Probabilities of success (errors in distances ≤ 3 pm and in energies ≤ 3 kcal/mol), with different density functional approximations, to have success for geometries, $P(g)$,

to have success for energy differences, $P(e)$, to have success for both properties, $P(g \cap e)$, to have success for the distances, once the energy differences has been checked to be acceptable, $P(g|e)$, to have success for the energy differences, once the geometrical parameters have been checked to be acceptable, $P(e|g)$. The orange areas correspond to probabilities above 0.50.

	$P(g)$	$P(e)$	$P(g \cap e)$	$P(g e)$	$P(e g)$	
XYG3	0.33	0.50	0.25	0.50	0.75	XYG3
XYGJ-OS	0.42	0.67	0.42	0.63	1.00	XYGJ-OS
xDH-PBE0	0.25	0.42	0.17	0.40	0.67	xDH-PBE0
B2PLYP	0.25	0.25	0.08	0.33	0.33	B2PLYP
B2PLYP-D	0.25	0.17	0.08	0.50	0.33	B2PLYP-D
B3LYP	0.17	0.08	0.08	1.00	0.50	B3LYP
B3LYP-D3	0.25	0.08	0.08	1.00	0.33	B3LYP-D3
B3PW91	0.17	0.08	0.08	1.00	0.50	B3PW91
PBE0	0.25	0.08	0.00	0.00	0.00	PBE0
PBE0-D3	0.25	0.08	0.00	0.00	0.00	PBE0-D3
O3LYP	0.25	0.25	0.08	0.33	0.33	O3LYP
X3LYP	0.25	0.08	0.08	1.00	0.33	X3LYP
BHHLYP	0.17	0.17	0.08	0.50	0.50	BHHLYP
M06-2X	0.17	0.58	0.08	0.14	0.50	M06-2X
M06	0.08	0.17	0.08	0.50	1.00	M06
TPSSh	0.08	0.08	0.08	1.00	1.00	TPSSh
BMK	0.17	0.33	0.00	0.00	0.00	BMK
B97-1	0.08	0.00	0.00	0.00	0.00	B97-1
B97-2	0.08	0.25	0.00	0.00	0.00	B97-2
B98	0.08	0.00	0.00	0.00	0.00	B98
CAM-B3LYP	0.08	0.08	0.00	0.00	0.00	CAM-B3LYP
ω B97X	0.08	0.08	0.00	0.00	0.00	ω B97X
ω B97X-D	0.08	0.17	0.08	0.50	1.00	ω B97X-D
HSE06	0.17	0.08	0.00	0.00	0.00	HSE06
LC- ω PBE	0.33	0.17	0.08	0.50	0.25	LC- ω PBE
LC-PBE	0.00	0.08	0.00	0.00	0.00	LC-PBE
TPSS	0.08	0.08	0.08	1.00	1.00	TPSS
M06-L	0.17	0.17	0.17	1.00	1.00	M06-L
B97-D	0.08	0.17	0.08	0.50	1.00	B97-D
B97-D3	0.00	0.08	0.00	0.00	0.00	B97-D3
HCTH407	0.08	0.17	0.08	0.50	1.00	HCTH407
BLYP	0.00	0.08	0.00	0.00	0.00	BLYP
PBE	0.00	0.00	0.00	0.00	0.00	PBE
QCISD	0.42	0.17	0.17	1.00	0.40	QCISD
MP2	0.00	0.17	0.00	0.00	0.00	MP2
HF	0.00	0.00	0.00	0.00	0.00	HF
	$P(g)$	$P(e)$	$P(g \cap e)$	$P(g e)$	$P(e g)$	

Fig. 4 Same as Figure 3, but with smaller tolerances (errors in distances ≤ 1.5 pm and in energies ≤ 1.5 kcal/mol).

	$P(g)$	$P(e)$	$P(g \cap e)$	$P(g e)$	$P(e g)$	
XYG3	0.92	1.00	0.92	0.92	1.00	XYG3
XYGJ-OS	0.83	1.00	0.83	0.83	1.00	XYGJ-OS
xDH-PBE0	0.92	0.92	0.83	0.91	0.91	xDH-PBE0
B2PLYP	0.50	0.83	0.42	0.50	0.83	B2PLYP
B2PLYP-D	0.50	0.58	0.25	0.43	0.50	B2PLYP-D
B3LYP	0.33	0.50	0.08	0.17	0.25	B3LYP
B3LYP-D3	0.33	0.33	0.08	0.25	0.25	B3LYP-D3
B3PW91	0.50	0.50	0.25	0.50	0.50	B3PW91
PBE0	0.50	0.50	0.33	0.67	0.67	PBE0
PBE0-D3	0.50	0.50	0.33	0.67	0.67	PBE0-D3
O3LYP	0.25	0.58	0.08	0.14	0.33	O3LYP
X3LYP	0.33	0.50	0.08	0.17	0.25	X3LYP
BHHLYP	0.50	0.58	0.33	0.57	0.67	BHHLYP
M06-2X	0.67	0.92	0.58	0.64	0.88	M06-2X
M06	0.42	0.83	0.42	0.50	1.00	M06
TPSSh	0.33	0.08	0.08	1.00	0.25	TPSSh
BMK	0.50	1.00	0.50	0.50	1.00	BMK
B97-1	0.33	0.58	0.17	0.29	0.50	B97-1
B97-2	0.42	0.67	0.33	0.50	0.80	B97-2
B98	0.42	0.58	0.25	0.43	0.60	B98
CAM-B3LYP	0.50	0.67	0.42	0.63	0.83	CAM-B3LYP
ω B97X	0.58	0.92	0.58	0.64	1.00	ω B97X
ω B97X-D	0.58	0.83	0.58	0.70	1.00	ω B97X-D
HSE06	0.50	0.50	0.33	0.67	0.67	HSE06
LC- ω PBE	0.83	0.67	0.67	1.00	0.80	LC- ω PBE
LC-PBE	0.50	0.33	0.25	0.75	0.50	LC-PBE
TPSS	0.33	0.08	0.08	1.00	0.25	TPSS
M06-L	0.25	0.50	0.17	0.33	0.67	M06-L
B97-D	0.17	0.33	0.17	0.50	1.00	B97-D
B97-D3	0.17	0.25	0.08	0.33	0.50	B97-D3
HCTH407	0.08	0.58	0.08	0.14	1.00	HCTH407
BLYP	0.25	0.17	0.08	0.50	0.33	BLYP
PBE	0.17	0.17	0.08	0.50	0.50	PBE
QCISD	0.83	0.83	0.67	0.80	0.80	QCISD
MP2	0.50	0.33	0.17	0.50	0.33	MP2
HF	0.25	0.08	0.08	1.00	0.33	HF
	$P(g)$	$P(e)$	$P(g \cap e)$	$P(g e)$	$P(e g)$	

Fig. 5 Same as Figure 3, but with larger tolerances (errors in distances ≤ 4.5 pm and in energies ≤ 4.5 kcal/mol).