

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 13683–13689

www.rsc.org/pccp

PAPER

# Validation of electronic structure methods for isomerization reactions of large organic molecules

Sijie Luo,<sup>a</sup> Yan Zhao<sup>b</sup> and Donald G. Truhlar<sup>\*a</sup>

Received 20th March 2011, Accepted 25th May 2011

DOI: 10.1039/c1cp20834a

In this work the ISOL24 database of isomerization energies of large organic molecules presented by Huenerbein *et al.* [*Phys. Chem. Chem. Phys.*, 2010, **12**, 6940] is updated, resulting in the new benchmark database called ISOL24/11, and this database is used to test 50 electronic model chemistries. To accomplish the update, the very expensive and highly accurate CCSD(T)-F12a/aug-cc-pVDZ method is first exploited to investigate a six-reaction subset of the 24 reactions, and by comparison of various methods with the benchmark, MCQCISD-MPW is confirmed to be of high accuracy. The final ISOL24/11 database is composed of six reaction energies calculated by CCSD(T)-F12a/aug-cc-pVDZ and 18 calculated by MCQCISD-MPW. We then tested 40 single-component density functionals (both local and hybrid), eight doubly hybrid functionals, and two other methods against ISOL24/11. It is found that the SCS-MP3/CBS method, which is used as benchmark for the original ISOL24, has an MUE of 1.68 kcal mol<sup>-1</sup>, which is close to or larger than some of the best tested DFT methods. Using the new benchmark, we find  $\omega$ B97X-D and MC3MPWB to be the best single-component and doubly hybrid functionals respectively, with PBE0-D3 and MC3MPW performing almost as well. The best single-component density functionals without molecular mechanics dispersion-like terms are M08-SO, M08-HX, M05-2X, and M06-2X. The best single-component density functionals without Hartree–Fock exchange are M06-L-D3 when MM terms are included and M06-L when they are not.

## 1. Introduction

It is important to test electronic structure model chemistries<sup>1</sup> against a wide variety of data. Many validation studies have been based on small molecules where accurate data is most readily available. But some theoretical methods that work well for small molecules are not so successful for large ones. Further tests are necessary, but extracting accurate electronic structure test data from experiment is problematic for large molecules because electronic structure calculations directly predict Born–Oppenheimer energies ( $E$ ), but experiment yields enthalpies ( $H$ ) or free energies ( $G$ ). Extracting  $E$  from  $H$  or  $G$  requires the removal of conformational-vibrational-rotational energy, which becomes progressively more complicated as molecular size increases.<sup>2</sup> Furthermore experimental  $H$  and  $G$  data becomes sparser as molecular size increases. In recent years though an alternative approach has become more affordable, namely the direct calculation of  $E$  by high-level wave function theory.<sup>3</sup> This also becomes harder as molecular size increases,

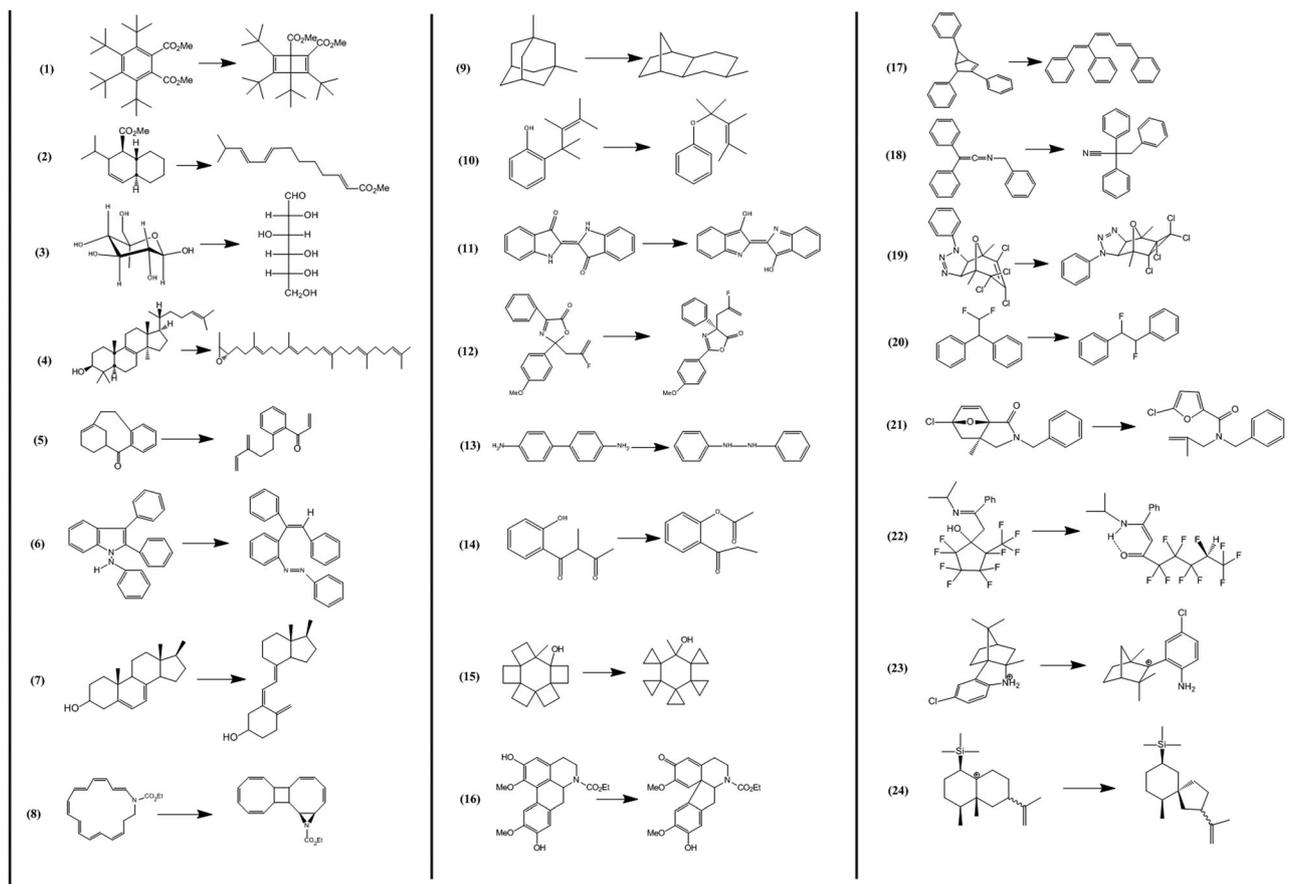
but progress is being made in extending the capabilities to larger systems.

It has been pointed out<sup>4</sup> that “a good place to start is the energetics of isomerization reactions since this allows direct comparison on the performance for differences in bonding, conjugation, and steric effects.” Recently, Huenerbein *et al.*<sup>5</sup> very constructively pushed the limit of what can be done in this regard by using spin-component-scaled third-order Møller–Plesset perturbation theory (SCS-MP3) to calculate reference values of the isomerization energy ( $\Delta E$ ) of 24 large molecules containing 24–81 atoms. They estimated uncertainties to be typically smaller than 2 kcal mol<sup>-1</sup> with maximum uncertainties of up to 4–5 kcal mol<sup>-1</sup>. The reference values were used to test 22 density functional methods that were found to have mean unsigned deviations from the reference data of 2.5–14.8 kcal mol<sup>-1</sup>. Their reference database is called ISOL24.<sup>5</sup> The ISOL24 reactions are shown in Fig. 1, which numbers each reaction (from 1 to 24) as in ref. 5, for convenience.

The objective of the present work is twofold: (1) to improve the accuracy of the reference data and (2) to test 50 additional electronic structure methods that are affordable for the entire benchmark suite. An electronic structure method is a short name for an electronic structure model chemistry,<sup>1</sup> which is a combination of a wave function level (*e.g.*, MP3 or Hartree–Fock) or

<sup>a</sup> Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, MN 55455-0431, USA

<sup>b</sup> Commercial Print Engine Lab., HP Laboratories, Hewlett-Packard Co., 1501 Page Mill Road, Palo Alto, CA 94304, USA



**Fig. 1** The 24 reactions of ISOL24 and ISOL24/11.

an approximate density functional with a one-electron basis set or a one-electron-basis-set extrapolation procedure. A method is denoted, as usual, as L/B where L denotes the wave function level or density functional, and B denotes the basis set or basis-set extrapolation procedure.

## 2. Methods

In the present work we directly use three basis sets: aDZ<sup>6</sup> (shorthand for aug-cc-pVDZ), MG3S,<sup>7</sup> and MG3SXP,<sup>8</sup> where aDZ is an augmented valence double zeta plus polarization set, MG3S is a minimally augmented valence triple zeta plus double polarization set, and MG3SXP is MG3S plus extra polarization functions. We also consider some calculations extrapolated to a complete basis set (CBS), and we consider some multi-coefficient correlation methods<sup>9–13</sup> (MCCMs) involving different basis sets for different components, as defined in the original references, which are given below.

All geometries were optimized by Huenerbein *et al.*<sup>5</sup> with the B97-D/TZVP method. Since the same geometries are used for the reference calculations as for the tested methods, the comparisons are a direct test of the ability of the tested methods to predict relative energies ( $\Delta E$ ) at pre-specified geometries. But since the geometries predicted by B97-D/TZVP are expected to be reasonably close to the accurate equilibrium geometries, we interpret the results as Born–Oppenheimer isomerization energies.

In order to validate affordable methods for improving the reference results, we first carried out CCSD(T)-F12a/aDZ wave function calculations on the six smallest reactions (reactions 3, 9, 10, 13, 14, and 20); these reactions have 24–35 atoms. Here CCSD denotes coupled cluster theory with single and double excitations,<sup>15</sup> (T) denotes a quasiperturbative treatment of connected triple excitations,<sup>16</sup> and F12a denotes a simplified version<sup>17</sup> of the F12<sup>18</sup> method that includes functions of interelectronic distance ( $r_{12}$ ) to accelerate convergence. The CCSD(T)-F12a/aDZ calculations are expected to be close to the CBS limit of CCSD(T).<sup>17,19</sup> We then compare these calculations to CCSD-F12a/aDZ, SCS-MP3/CBS, and seven MCCMs, which are explained next.

MCG3/3,<sup>11</sup> G3SX(MP3),<sup>10</sup> and BMC-CCSD<sup>12</sup> are MCCMs based entirely on wave function components. MCG3/3 and G3SX(MP3) each involve an expensive CCSD(T) step, but BMC-CCSD involves no step more expensive than CCSD and thus is more affordable for large systems.

MCG3-MPW<sup>13</sup> and MCG3-MPWB<sup>13</sup> are MCCMs that combine wave function and density functional components with the highest-order wave function component being CCSD(T). MCQCISD/MPW<sup>13</sup> and MCQCISD/MPWB<sup>13</sup> are similar but the highest-order wave function component is the less expensive quadratic configuration interaction with single and double excitations (QCISD). MC3MPW<sup>13</sup> and MC3MPWB<sup>13</sup> are even cheaper because the highest level of correlation included is based on the MP2 method. Since the Hartree–Fock

**Table 1** Density functional applied to all 24 reactions

Functional	Year	Ref.	Type <sup>a</sup>
BP86	1986–8	20, 21	GGA
B3LYP	1993–4	22, 23	Global-hybrid GGA
PBE	1996	24	GGA
PBE0	1996–8	25, 26	Global-hybrid GGA
$\tau$ HCTHhyb	2002	27	Global-hybrid <i>meta</i> -GGA
TPSS	2003	28	Meta-GGA
TPSSh	2003	29	Global-hybrid <i>meta</i> -GGA
BMK	2004	30	Global-hybrid <i>meta</i> -GGA
MPW3LYP	2004	31	Global-hybrid GGA
MC3MPW	2005	13	Doubly hybrid
MC3MPWB	2005	13	Doubly hybrid
MCQCISD-MPW	2005	13	Doubly hybrid
MCQCISD-MPWB	2005	13	Doubly hybrid
MOHLYP	2005	32	Global-hybrid GGA
PW6B95	2005	46	Global-hybrid <i>meta</i> -GGA
M05	2005	33	Global-hybrid <i>meta</i> -GGA
LC- $\omega$ PBE	2006	34	RS-hybrid GGA
B97-D	2006	35	GGA + MM
M05-2X	2006	36	Global-hybrid <i>meta</i> -GGA
M06-L	2006	37	Meta-GGA
M06-HF	2006	38	Global-hybrid <i>meta</i> -GGA
B2-PLYP	2006	39	Doubly hybrid
mPW2-PLYP	2006	40	Doubly hybrid
M06	2008	41	Global-hybrid <i>meta</i> -GGA
M06-2X	2008	41	Global-hybrid <i>meta</i> -GGA
M08-HX	2008	8	Global-hybrid <i>meta</i> -GGA
M08-SO	2008	8	Global-hybrid <i>meta</i> -GGA
$\omega$ B97	2008	42	RS-hybrid GGA
$\omega$ B97X	2008	42	RS-hybrid GGA
$\omega$ B97X-D	2008	43	RS-hybrid GGA + MM
B2GP-PLYP	2008	47	Doubly hybrid
M06-L-D	2009	37, 44	Global-hybrid <i>meta</i> -GGA + MM
M06-D	2009	41, 44	Global-hybrid <i>meta</i> -GGA + MM
BP86-D3	2010	20, 21, 45	GGA + MM
B3LYP-D3	2010	22, 23, 45	Global-hybrid GGA + MM
PBE-D3	2010	24, 45	GGA + MM
PBE0-D3	2010	25, 26, 45	Global-hybrid GGA + MM
TPSS-D3	2010	28, 45	Meta-GGA + MM
B2P-PLYP-D3	2010	39, 45	Doubly hybrid + MM
B97-D3	2010	45	GGA + MM
B2GP-PLYP-D3	2011	47, 48	Doubly hybrid + MM
TPSSh-D3	2011	29, 49	Global-hybrid <i>meta</i> -GGA + MM
PW6B95-D3	2011	46, 49	Global-hybrid <i>meta</i> -GGA + MM
M05-D3	2011	33, 49	Global-hybrid <i>meta</i> -GGA + MM
M05-2X-D3	2011	36, 49	Global-hybrid <i>meta</i> -GGA + MM
M06-L-D3	2011	37, 49	Meta-GGA + MM
M06-HF-D3	2011	38, 49	Global-hybrid <i>meta</i> -GGA + MM
M06-D3	2011	41, 49	Global-hybrid <i>meta</i> -GGA + MM
M06-2X-D3	2011	41, 49	Global-hybrid <i>meta</i> -GGA + MM

<sup>a</sup> “GGA” denotes generalized gradient approximation; “RS” denotes range-separated; “+ MM” denotes the addition of a post-SCF empirical molecular mechanics term.

orbitals on which CCSD(T), QCISD, and MP2 are built and the generalized Kohn–Sham orbitals on which some density functional components are built are both functionals of the electron density, these four methods are all doubly hybrid<sup>14</sup> density functional methods.

The other density functionals<sup>8,20–49</sup> considered in this work are all applied to all 24 reactions. For comparison we also include three other doubly hybrid density functional methods, namely, B2-PLYP,<sup>39</sup> mPW2-PLYP<sup>40</sup> and B2GP-PLYP,<sup>47</sup> in which the MP2-like correlation is combined with a DFT calculation. An empirical molecular-mechanics (MM) dispersion-like term can be added to these functionals in a post-SCF fashion (just as in the single-component functionals discussed below),

and two of such combined doubly-hybrid-functional-plus-MM methods are also included in the test. All density functionals considered in this article are listed and explained in Table 1. The table contains 40 single-component density functionals and nine doubly hybrid functionals. The 40 single-component functionals consist of 22 without MM terms and 18 with them. The nine multi-component functionals consist of seven without MM and two with MM.

### 3. Reference values

Table 2 compares the isomerization energies calculated by 11 methods for the six smallest systems in the ISOL24 database.

**Table 2** Isomerization energies (kcal mol<sup>-1</sup>) for a 6-reaction subset of the ISOL24 database

Reaction no.	CCSD(T)-F12a/aDZ	CCSD-F12/aDZ	MCG3-MPW	MCG3-MPW	MG3/3	MCQCISD-MPW	MCQCISD-MPW	SCS-MP3/CBS	BMC-CCSD	M06-L/MG3SXP	M06-L-DF <sup>a</sup>
3	9.77	10.98	11.05	10.67	10.57	11.30	11.02	11.70	11.16	4.58	4.58
9	21.76	21.34	21.21	20.93	22.44	20.90	20.33	22.30	20.20	19.05	19.03
10	6.82	5.93	6.34	6.25	6.92	6.30	6.10	7.90	8.57	2.72	2.70
13	33.52	32.75	33.52	33.48	34.57	33.47	33.30	35.10	34.75	32.44	32.42
14	5.30	4.19	5.22	5.04	5.19	5.43	5.18	5.20	5.33	3.46	3.45
20	4.66	4.70	4.86	4.76	4.69	4.63	4.47	4.50	4.75	3.14	3.15
MUE	0.00 <sup>b</sup>	0.74	0.43	0.45	0.46	0.52	0.66	0.90	1.01	2.74	2.75
Cost	3000	2600	1900	1900	1700	370	370	190	520	21	12

<sup>a</sup> Same basis but with density fitting. <sup>b</sup> By definition.

The last row of the table gives the “cost,” which is defined as the sum of the computer times (in processor hours) for single-point energy calculations on reactions 10 and 20 (the two largest systems in Table 2) divided by the sum of the computer times for single-point MP2/6-31G calculations on the same two reactions, run with the same software on the same computer. We recognize that such timings depend on the software, the computer and its load, and the parallelization, but nevertheless they are usually meaningful when they differ by more than a factor of about 1.5–2. All costs are rounded to two significant figures. All coupled cluster and QCISD results were calculated with *Molpro*<sup>50</sup> and all density functional calculations were run with *Gaussian 09*.<sup>51</sup>

The second last line of Table 1 is the mean unsigned error (MUE) of each method as compared to the CCSD(T)-F12a/aDZ results. The BMC-CCSD, SCS-MP3/CBS, and CCSD-F12/aDZ methods have MUEs of 1.01, 0.90, and 0.74 kcal mol<sup>-1</sup> respectively, and the four doubly hybrid methods have MUEs in the range 0.43–0.66 kcal mol<sup>-1</sup>. The most accurate method that is affordable for the entire database of 24 reactions is MCQCISD-MPW. This method was applied to the other 18 reactions, and a new reference database was created containing CCSD(T)-F12a/aDZ results for the six reactions of Table 2 and MCQCISD-MPW results for the other 18. This new set of reference data is called ISOL24/11 where the last two digits signify the update of 2011. The isomerization energies of the ISOL24/11 database are listed in Table 3.

#### 4. Testing density functionals

All 40 single-component density functionals were tested against ISOL24/11 with the MG3SXP basis set. We also tested

**Table 3** The ISOL24/11 database (kcal mol<sup>-1</sup>)

Reaction	$\Delta E$	Reaction	$\Delta E$
1	69.17	13	33.52
2	37.54	14	5.30
3	9.77	15	3.06
4	66.43	16	22.78
5	32.84	17	10.33
6	25.51	18	22.57
7	17.37	19	18.25
8	22.34	20	4.66
9	21.76	21	11.21
10	6.82	22	0.77
11	37.87	23	23.43
12	0.20	24	14.94

four of them with the MG3 basis set. As for the doubly hybrid functionals, B2-PLYP, B2GP-PLYP, and mPW2-PLYP with and without MM terms are also tested with the MG3SXP basis set, but MC3MPW and MC3MPWB have different basis sets for different level of calculations. Finally we tested MCQCISD-MPW, BMC-CCSD, and SCS-MP3/CBS. In every case we computed the mean unsigned error (MUE). These MUEs are given in Tables 4–6.

Table 4 compares the results of M06, M06-2X and M06-HF with MG3S and MG3SXP basis sets. MG3SXP (where XP denotes “extra polarization”) differs from MG3S in that it adds an extra polarization function to MG3S basis. In particular, the 2df polarization functions of MG3S on Li–Ne are replaced by a 3df set, and the 3d2f polarization functions on Al–Ar are replaced by a 4d2f set. It has been shown in previous research that the extra polarization functions of MG3SXP is necessary for obtaining accurate results in certain systems containing fluorine atoms, and considering that there are three reactions (reaction 12, 20 and 22) in the ISOL24/11 database involving fluorine, it is beneficial to clarify the necessity of extra polarization. The MUEs of M06 and M06-2X decrease by 1.27 and 1.15 kcal mol<sup>-1</sup> after adding the extra polarization, while the result of M06-HF with MG3SXP is unexpectedly worse than MG3S. In general we believe that MG3SXP provides validating results by properly treating the fluorine-containing systems with the extra polarization functions, and the quality of various density functionals can be appropriately compared by use of this basis set.

Table 5 compares the performance of various single-component and doubly hybrid density functionals against the ISOL24/11 database. Empirical dispersion-like MM terms (D-correction) can also be added to the DFT calculations in a post-SCF fashion to improve accuracy for weak interactions, and here we include three versions of them. The label “D” refers to the D-correction in ref. 44 except for two cases: in  $\omega$ B97X-D it refers to ref. 42 and in B97-D, it refers to ref. 35. On the other hand, “D3” refers to the D-correction in ref. 45 and 46.

**Table 4** Mean unsigned error (kcal mol<sup>-1</sup>) for density functionals tested with both MG3S and MG3SXP basis sets

Functionals	MUE	
	MG3S	MG3SXP
M06-2X	3.65	2.46
M06	4.29	3.06
M06-HF	3.56	4.47

**Table 5** Mean unsigned errors (kcal mol<sup>-1</sup>) for 48 density functionals. All calculations are performed with MG3SXP basis except for MCQCISD-MPW, MC3MPW, and MC3MPWB, which have method-specific basis sets for each of their components

Functional	MUE	Functional	MUE	Functional	MUE
MCQCISD-MPWB	0.47	PW6B95-D3	3.09	TPSS-D3	5.91
B2GP-PLYP-D3	1.36	ωB97	3.42	M06-L	5.94
MC3MPWB	1.65	PBE0	3.43	B3LYP-D3	6.02
MC3MPW	1.79	BMK	3.46	PBE	6.19
ωB97X-D	1.90	mPW2-PLYP	3.74	M05	6.49
PBE0-D3	1.91	BP86-D3	4.20	B97-D	6.71
M08-SO	2.19	B2-PLYP	4.39	TPSSh	6.78
M08-HX	2.27	PBE-D3	4.41	BP86	7.58
B2GP-PLYP	2.33	TPSSh-D3	4.41	TPSS	8.37
M06-2X-D3	2.34	PW6B95	4.44	MPW3LYP	8.48
M05-2X	2.35	M06-HF	4.47	B3LYP	8.98
M06-2X	2.46	LC-ωPBE	4.73	MOHLYP	13.49
M05-2X-D3	2.48	M06-HF-D3	4.77		
M06-D3	2.59	M06-L-D	4.82		
ωB97X	2.77	B97-D3	5.17		
M06-D	2.81	M05-D3	5.61		
B2-PLYP-D3	2.86	M06-L-D3	5.73		
M06	3.06	τHCTHhyb	5.87		

**Table 6** Mean unsigned errors (kcal mol<sup>-1</sup>) for other methods

Method	MUE
BMC-CCSD	1.46
SCS-MP3/CBS	1.68

Although Table 1 contains 49 density functionals, Table 5 contains only 48 because it omits MCQCISD-MPW, which is used for some of the benchmark values.

We first look at the “pure” density functionals without MM corrections. (These are the density functionals that do not have a suffix-D or -D3.) The popular B3LYP functional has an enormous MUE of 8.98 kcal mol<sup>-1</sup>, while the best single-component functional without D-correction, namely, M08-SO, achieves a MUE of 2.19 kcal mol<sup>-1</sup>. Several other Minnesota functionals, namely, M08-HX, M05-2X and M06-2X also perform quite well with MUEs of 2.27, 2.35 and 2.46 kcal mol<sup>-1</sup> respectively. Furthermore, we observe no improvement when comparing PBE with its *meta*-GGA counterpart TPSS, but the hybrid PBE0 almost halves the error of PBE. The range-separated functional ωB97X achieves an MUE of 2.77 kcal mol<sup>-1</sup>, but the other two RS-hybrid functionals, namely, ωB97 and LC-ωPBE, do not perform as well. This is well understood; the advantage of ωB97X is that it is not constrained to have zero Hartree–Fock exchange at small inter-electronic distances.

On the other hand, the best MM-corrected functional, which is also the best single-component functional overall, is ωB97X-D with an MUE of 1.90 kcal mol<sup>-1</sup>. Also surprising is the good performance of PBE0-D3, which is almost equally good as ωB97X-D and also outperforms all the other single-component functionals in our test. Moreover, most of the Minnesota functionals do not significantly improve by adding the MM corrections, while other MM-corrected functionals in general significantly improve over their non-corrected counterparts.

The only exception is that M06-L-D has a MUE of approximately 1 kcal mol<sup>-1</sup> smaller than M06-L. We attribute the above observations to the fact that the whole family of Minnesota functionals has incorporated reasonably accurate medium-range correlation energy by virtue of their functional forms and their parameterization process and adding post-SCF dispersion-like corrections thus does not improve their performance significantly. If we restrict ourselves to functionals without Hartree–Fock exchange (which is an important consideration from the point of view of cost, especially for extended systems), the best functional without MM terms is M06-L and the best with MM terms is M06-L-D3.

We then proceed to consider the more expensive doubly hybrid functionals, in which an MP2-like or QCISD-like term is added as a post-SCF term to the DFT or weighted DFT result.

Because MP2 and QCISD are more expensive than DFT for a given basis set for large systems, these methods-when used with a single basis set-have the computational cost of MP2 or QCISD rather than that of DFT; however, this does not affect MCQCISD-type and MC3M-type methods because they involve two basis sets. It should be noticed that the cost of MP2 steps can be reduced by using RI-MP2 with appropriate auxiliary basis sets, but this is not adopted in this paper. For all the doubly hybrid functionals based on MP2 as the WFT component, which excludes those in Table 2, but includes B2-PLYP, B2GP-PLYP, mPW2-PLYP, MC3MPW, and MC3MPWB, the average cost is approximately one to two orders of magnitude higher than for single-component functionals. However, these methods are still quite affordable with presently available computational resources.

Although the MUE of MCQCISD-MPWB is only 0.46 kcal mol<sup>-1</sup>, which is the smallest of all the tested methods, this is partly due to the fact that it is very similar to MCQCISD-MPW, which is the method we use to calculate 18 of the 24 reactions. On the other hand, the value is probably reasonable since it is close to the values in Table 2, which are based on comparison to the very accurate results. MC3MPWB and MC3MPW both achieve good performance without MM terms, achieving MUEs of 1.79 and 1.65 kcal mol<sup>-1</sup>, which outperform all the single-component functionals. B2GP-PLYP also performs well, and its accuracy is further improved by adding the MM term.

Finally in Table 6 we compare the results for the methods whose costs scale as N<sup>6</sup>, namely, SCS-MP3/CBS and BMC-CCSD. SCS-MP3/CBS is the method used in ref. 5 to construct the original ISOL24 database, and we find its MUE against the updated ISOL24/11 to be 1.68 kcal mol<sup>-1</sup>. This result is consistent with the error estimation of 2 kcal mol<sup>-1</sup> in ref. 5, and also implies that the original ISOL24 is not accurate enough for evaluation of methods such as B2GP-PLYP-D3 or MC3MPW, for which the MUE is even smaller than SCS-MP3/CBS. Second, BMC-CCSD is found to have an MUE of 1.46 kcal mol<sup>-1</sup>, which is only slightly worse than the best method in Table 5 if we exclude MCQCISD-MPW. In spite of the fact that it scales as N<sup>6</sup>, its most expensive step only requires a CCSD calculation with a polarized valence double-zeta basis set, which should be affordable for a large number of applications.

**Table 7** Mean unsigned errors (kcal mol<sup>-1</sup>) of the 6-reaction subset calculated by all the methods used in this paper with the results obtained by CCSD(T)-F12a/aDZ taken as reference data

Functional	MUE	Functional	MUE	Functional	MUE
MCG3-MPW	0.43	B2-PLYP	1.28	M05-D3	2.43
MCG3-MPWb	0.45	M06-D3	1.36	M06-HF-D3	2.43
MG3/3	0.46	ωB97	1.48	MPW3LYP	2.46
MCQCISD-MPW	0.52	PBE0	1.50	B97-D3	2.52
M08-HX	0.60	M06-2X-D3	1.53	B3LYP	2.63
B2G-PLYP-D3	0.64	M06-2X	1.53	M06-L-D3	2.73
MCQCISD-MPWb	0.66	ωB97X	1.55	M06-L	2.74
B2G-PLYP	0.67	M06-D	1.67	M05	2.75
CCSD-F12a/aDZ	0.74	B97D	1.75	TPSSh-D3	2.76
SCS-MP3/CBS	0.90	PW6B95-D3	1.75	TPSSh	3.12
B2-PLYP-D3	1.00	BMK	1.81	TPSS-D3	3.34
BMC-CCSD	1.01	τHCTHhyb	1.82	TPSS	3.67
MC3MPW	1.03	PBE-D3	1.89	MOHLYP	4.32
mPW2-PLYP	1.11	LC-ωPBE	1.91		
MC3MPWB	1.14	BP86-D3	1.99		
PBE0-D3	1.14	PW6B95	2.00		
ωB97X-D	1.14	PBE	2.02		
M08-SO	1.17	M06-L-D	2.09		
M05-2X	1.23	B3LYP-D3	2.14		
M06	1.26	BP86	2.30		
M05-2X-D3	1.27	M06-HF	2.42		

## 5. Concluding remarks

We have updated the ISOL24 database constructed in ref. 5 by methods of higher accuracy, and the new ISOL24/11 database is composed of six reactions calculated by CCSD(T)/aDZ and 18 reactions by MCQCISD-MPW. This is validated by a careful investigation of a subset of the smallest six reactions by the very accurate CCSD(T)-F12a/aDZ method, which should be very close to the CCSD(T) results with a complete basis set, and MCQCISD-MPW is then found to be the most accurate method we can afford for the remaining 18 reactions. The new ISOL24/11 database is recommended for future tests and parameterization of methods aimed at treatment of chemical reactions of large size.

We have tested 48 single-component and doubly hybrid functionals with and without MM terms, together with 2 other methods, against the ISOL24/11 database, and we conclude this paper with recommendations as follows.

(a) ωB97X-D and PBE0-D3 are the most highly recommended single-component functionals with MM corrections, while Minnesota functionals with high HF exchange, namely, M08-SO, M08-HX, M05-2X and M06-2X perform the best even without MM terms (D-correction).

(b) For the more expensive doubly hybrid functionals, whose most expensive component is an MP2 step, B2GP-PLYP-D3 and MC3MPWB are recommended.

(c) The even more expensive MCG3-MPW, MCQCISD-MPW, and BMC-CCSD methods are also recommended for benchmark of large reactions involving weak-interactions, whenever they are affordable.

## Appendix

Table 7 shows a comparison of the methods based only on the six reactions for which we have the most accurate results.

The results differ from those in Tables 5 and 6 primarily because this subset of reactions is smaller and less diverse.

## Acknowledgements

This work was supported in part by the National Science Foundation under Grant No. CHE03-49122. This research was performed in part using the MSCF in EMSL, a national scientific user facility sponsored by the US Department of Energy Office of Basic Energy Research and located at PNNL and also using a computer time grant from Minnesota Supercomputing Institute.

## References

- 1 J. A. Pople, *Rev. Mod. Phys.*, 1999, **71**, 1267.
- 2 J. Zheng, T. Yu, E. Papajak, I. M. Alecu, S. L. Mielke and D. G. Truhlar, *Phys. Chem. Chem. Phys.*, 2011, **13**, 10885.
- 3 See, e.g., A. Karton, E. Rabinovich, J. M. L. Martin and B. Ruscic, *J. Chem. Phys.*, 2006, **125**, 144108.
- 4 K. W. Sattelmeyer, J. Tirado-Rives and W. L. Jorgensen, *J. Phys. Chem. A*, 2006, **110**, 13551.
- 5 R. Huenerbein, B. Schirmer, J. Moellmann and S. Grimme, *Phys. Chem. Chem. Phys.*, 2010, **12**, 6940.
- 6 R. A. Kendall, T. H. Dunning Jr. and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796.
- 7 (a) R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1980, **72**, 650; (b) T. Clark, J. Chandrasekhar, G. W. Spitznagel and P. v. R. Schleyer, *J. Comput. Chem.*, 1983, **4**, 294; (c) L. A. Curtiss, K. Raghavachari, C. Redfern, V. Rassolov and J. A. Pople, *J. Chem. Phys.*, 1998, **109**, 7764; (d) P. L. Fast, M. L. Sanchez and D. G. Truhlar, *Chem. Phys. Lett.*, 1999, **306**, 407; (e) B. J. Lynch, Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2003, **107**, 1384.
- 8 Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.*, 2008, **4**, 1849.
- 9 P. L. Fast, J. C. Corchado, M. L. Sanchez and D. G. Truhlar, *J. Phys. Chem. A*, 1999, **103**, 5129.
- 10 L. A. Curtiss, P. C. Redfern, K. Raghavachari and J. A. Pople, *J. Chem. Phys.*, 2001, **114**, 108.
- 11 B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2003, **107**, 3898.
- 12 B. J. Lynch, Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2005, **109**, 1643.
- 13 Y. Zhao, B. J. Lynch and D. G. Truhlar, *Phys. Chem. Chem. Phys.*, 2005, **7**, 43.
- 14 Y. Zhao, B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 4786.
- 15 G. D. Purvis III and R. J. Bartlett, *J. Chem. Phys.*, 1982, **76**, 1910.
- 16 K. Raghavachari, G. W. Trucks, J. A. Pople and M. Head-Gordon, *Chem. Phys. Lett.*, 1989, **157**, 479.
- 17 G. Knizia, T. B. Adler and H.-J. Werner, *J. Chem. Phys.*, 2009, **130**, 054104.
- 18 J. Noga, S. Kedzuch, J. Simunek and S. Ten-no, *J. Chem. Phys.*, 2008, **128**, 174103.
- 19 E. Papajak and D. G. Truhlar, *J. Chem. Theory Comput.*, 2011, **7**, 10.
- 20 J. P. Perdew, *Phys. Rev.*, 1986, **33**, 8862.
- 21 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098.
- 22 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.
- 23 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623.
- 24 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.
- 25 J. P. Perdew, M. Ernzerhof and K. Burke, *J. Chem. Phys.*, 1996, **105**, 9982.
- 26 C. Adamo and V. Barone, *Chem. Phys. Lett.*, 1998, **298**, 113.
- 27 A. D. Boese and N. C. Handy, *J. Chem. Phys.*, 2002, **116**, 9559.
- 28 J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.
- 29 V. N. Staroverov, G. E. Scuseria, J. Tao and J. P. Perdew, *J. Chem. Phys.*, 2003, **119**, 12129.

- 30 A. D. Boese and J. M. L. Martin, *J. Chem. Phys.*, 2004, **121**, 3405.
- 31 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 6908.
- 32 N. Schultz, Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2005, **109**, 11127.
- 33 Y. Zhao, N. E. Schultz and D. G. Truhlar, *J. Chem. Phys.*, 2005, **123**, 161103.
- 34 O. A. Vydrov, J. Heyd, A. V. Krukau and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 074106.
- 35 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787.
- 36 Y. Zhao, N. Schultz and D. G. Truhlar, *J. Chem. Theory Comput.*, 2006, **2**, 364.
- 37 Y. Zhao and D. G. Truhlar, *J. Chem. Phys.*, 2006, **125**, 194101.
- 38 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2006, **110**, 13126.
- 39 S. Grimme, *J. Chem. Phys.*, 2006, **124**, 034108.
- 40 T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys.*, 2006, **8**, 4398.
- 41 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215.
- 42 J.-D. Chai and M. Head-Gordon, *J. Chem. Phys.*, 2008, **128**, 084106.
- 43 J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615.
- 44 A. Karton, D. Gruzman and J. M. L. Martin, *J. Phys. Chem. A*, 2009, **113**, 8434.
- 45 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
- 46 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2005, **109**, 5656.
- 47 A. Karton, A. Tarnopolsky, J.-F. Lamère, G. C. Schatz and J. M. L. Martin, *J. Phys. Chem. A*, 2008, **112**, 12868.
- 48 L. Goerigk and S. Grimme, *J. Chem. Theory Comput.*, 2011, **7**, 291.
- 49 L. Goerigk and S. Grimme, <http://toc.uni-muenster.de/DFTD3/functionals.html> (accessed Jan. 15, 2011).
- 50 H.-J. Werner, P. J. Knowles, F. R. Manby, M. Schütz, P. Celani, G. Knizia, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklaß, P. Palmieri, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang and A. Wolf, *Molpro*, version 2010, University College Cardiff, Cardiff, 2010.
- 51 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *GAUSSIAN 09 (Revision A.02)*, Gaussian, Inc., Wallingford CT, 2009.