

Prepared for *J. Phys. Chem. B*  
Revised July 5, 2005

# Improved Density Functionals for Water

Erin E. Dahlke and Donald G. Truhlar\*

*Department of Chemistry and Supercomputing Institute, University of Minnesota,*

*Minneapolis, Minnesota 55455-0431*

## **Abstract**

The accuracy of existing density functional methods for describing the noncovalent interaction energies in small water clusters is investigated by testing twenty-five density functionals against a data set of 28 water dimers and 8 water trimers whose structures are taken from the literature and from simulations. The most accurate functionals are found to be PW6B95 with a mean unsigned error of 0.13 kcal/mol and MPWB1K and B98 with mean unsigned errors of 0.15 kcal/mol; the best functional with no Hartree-Fock exchange is *m*PWLYP, which is a GGA with a mean unsigned error of 0.28 kcal/mol. In comparison, the most popular GGA functionals, PBE and BLYP, have mean unsigned errors of 0.52 and 1.03 kcal/mol respectively. Since GGAs are very cost efficient for both condensed-phase simulations and electronic structure calculations on large systems, we optimized four new GGAs for water. The best of these, PBE1W and MPWLYP1W, have mean unsigned errors of 0.12 and 0.17 kcal/mol

respectively. These new functionals are well suited for use in condensed phase simulations of water and ice.

## 1. Introduction

There has been a great deal of theoretical work focused on describing the energetics of small to moderately sized water clusters using both high-level quantum mechanical methods and density functional theory (DFT).<sup>1-14</sup> The DFT methods are more economical, and the use of molecular dynamics and Monte Carlo methods to study bulk water<sup>15-23</sup> has placed an increased importance on finding a density functional method that can accurately describe these systems. While a recent advance<sup>25</sup> has allowed for the inclusion of exact Hartree-Fock exchange into the plane wave code VASP<sup>26</sup>, many of the simulation methods that employ plane waves<sup>15</sup> require the use of a density functional method that does not contain any Hartree-Fock exchange, that is a non-hybrid method. (A hybrid functional contains Hartree-Fock exchange. Hybrid meta functionals contain Hartree-Fock exchange and kinetic energy density, and GGAs and local spin density approximations (LSDAs) contain neither.) Additionally, since calculation of the exchange terms is the bottleneck in hybrid DFT methods,<sup>25,27-29</sup> if one is interested in studying large clusters a non-hybrid method will be more efficient. Moreover, if one is interested in a broad range of aqueous chemistry, such as processes at water-metal interfaces,<sup>30,31</sup> it may be desirable to use a non-hybrid method, as functionals with no Hartree-Fock exchange perform better than hybrid methods for many systems containing transition metals.<sup>32-35</sup>

Several years ago, Hall *et. al*<sup>36</sup> pointed out that most density functionals underestimate the binding energy of the water dimer. Recent assessments of density functional methods for noncovalently-bonded complexes,<sup>37,38</sup> including complexes formed by hydrogen bonding, show that hybrid meta functionals can be used to accurately describe noncovalent interactions such as hydrogen bonding, pi-pi stacking, and dipole and charge-transfer interactions. Furthermore, they

show that both hybrid and hybrid meta functionals perform substantially better for non-bonded interactions than generalized gradient approximations (GGAs) or meta functionals. (A meta functional is one that contains kinetic energy density. GGAs and meta functionals depend on the density gradient as well as the density; LSDAs do not.) With this in mind, we set out to determine the accuracy of a variety of density functional methods against a database of water dimer and trimer data and to investigate the usefulness of GGAs and meta functionals for studying water interactions.

## 2. Database

In order to create a database against which to test existing DFT methods, we selected a set of water dimers and trimers both from structures found in the literature and from simulations of liquid water and ice. The literature clusters consist of a collection of ten water dimers and six water trimers representing stationary points on their respective potential energy surfaces, as taken from the work of Tschumper *et al.*<sup>39,40</sup> The accurate binding energies for these clusters were reported in references 39 and 40 based on calculations using the focal point method. Additionally, a collection of ten water dimers from molecular dynamics simulations of ice, as well as eight water dimers and two water trimers from Monte Carlo simulations of liquid water were included in the data set. Accurate binding energies for all clusters taken from simulations were computed, relative to the unrelaxed monomers, at the Weizmann-1<sup>41, 42</sup> level of theory using the MOLPRO<sup>43</sup> quantum chemistry package.

The dimers from the molecular dynamics simulations of ice consist of two structures taken from simulations<sup>44</sup> of ice XI, as well as eight structures taken from simulations<sup>44</sup> of ice VIII at pressures ranging from 0 GPa to 70 GPa in increments of 10 GPa (one dimer from each simulation). The eight dimers and two trimers from Monte Carlo simulations were taken from three different ensembles at seven different temperatures. One dimer was taken from the liquid

box of a Gibbs ensemble simulation<sup>22</sup> at each of the following temperatures: 323, 373, 423, 473, and 523 K. Additionally, one trimer from the liquid box of the 323 K simulation, and one dimer from the vapor box of the 523 K simulation were used. The remaining dimers and trimer were taken from the  $NpT$  and  $NVT$  ensembles, with one dimer from an  $NpT$  ensemble simulation<sup>23</sup> carried out at 298 K and 1 atmosphere of pressure, and one dimer and one trimer taken from an  $NVT$  ensemble simulation<sup>24</sup> run at 315 K.

We define the binding energy of a dimer and trimer, respectively, as

$$E_{\text{bind}} = E_A + E_B - E_{AB} \quad (3)$$

$$E_{\text{bind}} = E_A + E_B + E_C - E_{ABC} \quad (4)$$

For the literature clusters the monomer is a relaxed gas-phase water molecule ( $A = B = C$ ). For the clusters taken from simulations we chose to use the unrelaxed monomers from each cluster. The molecular dynamics simulations from which some of the dimers were taken use rigid water molecules, making the two monomers within a given dimer equivalent. The structures from the Monte Carlo simulations, however, do not use rigid monomers, making the monomers for a given dimer or trimer different. The intramolecular O-H bond lengths in these structures vary from 0.92 Å to 1.05 Å, with an H-O-H angle in the range from 93.7–116.7 degrees.

### 3. Tests of Previously Developed Methods

In order to assess the accuracy of existing DFT methods for describing small water clusters, twenty-five DFT methods were chosen to test against the data set described in the previous section. The methods chosen include 22 standard functionals that were developed for broad applicability plus three unconventional combinations of standard exchange and correlation functions (*m*PWLYP, PBELYP and TPSSLYP). The methods tested were chosen based on recent studies evaluating the performance of DFT methods for non-bonded interactions,<sup>37,38</sup> including hydrogen bonding, and they also include those methods commonly used in the

literature for the simulation of bulk water and ice. All DFT calculations were carried out using the *Gaussian03*<sup>45</sup> software package, utilizing the MG3S<sup>46</sup> basis set. For water the MG3S basis set is identical to 6-311+G(2df,2p).<sup>47,48</sup>

Table 1 gives, for each DFT method tested, the mean signed error (MSE), mean unsigned error (MUE), and root-mean-square error (RMSE) averaged over the dimers, over the trimers, and over the total data set. We note that most of the hybrid and hybrid meta methods give substantially better performance than even the best GGA method. The most accurate hybrid method is B98, with a MUE of 0.15 kcal/mol, and the best hybrid meta method is PW6B95, with a MUE of 0.13 kcal/mol. Of the non-hybrid methods, *m*PWLYP and TPSS give the best results, with MUEs of 0.28 and 0.30 kcal/mol. These errors may be compared to the mean binding energy of the 36 structures, which is 4.80 kcal/mol.

#### 4. Parameterization of New Method

The exchange-correlation energy can be written as:

$$E_{xc} = (1 - \frac{X}{100})(E_x^S + \Delta E_x^{GCE}) + \frac{X}{100} E_x^{HF} + E_c^{LSDA} + \frac{Y}{100} \Delta E_c^{GCC}$$

where  $E_x^S$  is the local spin density approximation to the exchange energy,  $E_x^{GCE}$  is the gradient correction to the exchange energy,  $E_x^{HF}$  is the Hartree-Fock exchange energy,  $E_c^{LSDA}$  is the local spin density approximation to the correlation energy,  $\Delta E_c^{GCC}$  is the gradient correction to the correlation energy, and  $X$  and  $Y$  are parameters determining, respectively, the percentage of Hartree-Fock exchange and the percentage of gradient correction to the correlation energy that are included in a given functional. For a non-hybrid method (e.g., *m*PWLYP or TPSS),  $X$  is set equal to zero.

In general  $Y$  is as an optimizable parameter. Although  $Y$  has been adjusted in some DFT functionals, including the popular B3LYP<sup>63</sup> where it has a value of 81, it has often been found

that density functionals are more sensitive to  $X$  than to  $Y$ , and much recent work has focused on one-parameter methods in which  $X$  is varied.<sup>37,53,64-66,70</sup> However, the sensitivity to  $Y$  is sometimes greater than is generally appreciated, and when one is interested in finding the best functional with  $X = 0$  for other reasons (see Introduction), it is worthwhile to re-examine the dependence of the results on  $Y$ . Since the current prescriptions for gradient corrections to correlation functionals are inexact, it seems unlikely that  $Y = 100$  is optimum. With these considerations in mind we set out to determine if the accuracy of GGAs and meta functionals could be improved by varying this single parameter to make them competitive with the hybrid and hybrid meta methods.

For simplicity we chose to investigate the effect of varying the  $Y$  values of the two most popular correlation functionals used in simulations, PBE and LYP. The *mPW*, PBE and TPSS exchange functionals were chosen as they were found to be components of the best GGA and meta functionals in the tests in Table 1. The optimal value of  $Y$  for each of four methods was found by iteratively adjusting the  $Y$  values to the nearest integer, such that the mean unsigned error (MUE) of the 36 water clusters in the database is minimized. The optimized methods found with this approach are referred to as *mPWLYP1W* ( $Y = 88$ ), *PBE1W* ( $Y = 74$ ), *PBELYP1W* ( $Y = 54$ ), and *TPSSLYP1W* ( $Y = 74$ ), where 1W denotes a one-parameter method optimized for water.

The keywords required to carry out an *mPWLYP1W* calculation in *Gaussian03* are :

```
#MPWV5LYP
IOp(3/78=0880010000)
```

The keywords required to carry out a *PBE1W* calculation in *Gaussian03* are :

```
#PBEPBE
IOp(3/78=0740010000)
```

The keywords required to carry out a PBELYP1W calculation in *Gaussian03* are:

```
#PBEV5LYP  
IOp(3/78=0540010000)
```

The keywords required to carry out a TPSSLYP1W calculation in *Gaussian03* are :

```
#TPSSV5LYP  
IOp(3/78=0740010000)
```

Since PBE is already used in many molecular dynamics and Monte Carlo codes, and since PBE1W involves only scaling the gradient correction to the correlation energy, implementation of PBE1W should be especially straightforward, and it involves changing only a single number or line of code in many programs.

## 5. Results and Discussion

Table 2 shows the mean errors for the four methods parameterized in this work. The results show a substantial improvement for all of the optimized methods over their unoptimized counterparts. Additionally, the best optimized method, PBE1W has mean unsigned errors that are comparable to those of PW6B95 and better than those of B98. This new method gives a mean unsigned error for the dimers, trimers, and total data set of 0.10, 0.20, and 0.12 kcal/mol respectively. Given that the average binding energies for these three data sets are 2.47, 12.93, and 4.80 kcal/mol, these errors represent only 4.0%, 1.5%, and 2.5% of the average binding energies.

One way to put the results in the present letter into perspective is to compare the methods to X3LYP as a standard. The X3LYP functional was specifically designed to yield improved descriptions of hydrogen bonded and van der Waals systems,<sup>62</sup> and it was found, along with PBE1PBE and *mPW1PW91*, to be particularly accurate for water dimer.<sup>71</sup> Indeed, for the present water data set, the mean unsigned errors for X3LYP, PBE1PBE, and *mPW1PW91* are

0.22, 0.45, and 0.37 kcal/mol respectively. In another study<sup>11</sup> devoted entirely to water, the authors tested X3LYP against the most accurate data available for  $(\text{H}_2\text{O})_n$  with  $n = 2-6, 8$  and concluded that X3LYP is “the most practical *ab initio* method today for calculating accurate water cluster structures and energies”. It is not clear what the authors mean by *ab initio* since X3LYP has parameters, but the message is clear that X3LYP is quite competitive for water. Therefore it is very significant to compare other functionals to X3LYP.

The present tests compare X3LYP to even more accurate data (although we are limited to smaller clusters) and — as stated above —we confirm its accuracy. It is therefore interesting to note that seven of the density functionals studied here have mean unsigned errors in kcal/mol smaller than the 0.22 of X3LYP, in particular: PBE1W (0.12), PW6B95 (0.13), B98 (0.15), MPWB1K (0.15), MPW1B95 (0.16), *m*PWLYP1W (0.17), and TPSSLYP1W (0.19). The mean unsigned error in X3LYP is 69% larger than that for PW6B95 (a general-purpose functional parameterized before the water test set was created) and 83% larger than that for PBE1W (a functional parameterized specifically for noncovalent interactions between water molecules but having the distinct advantage of no Hartree-Fock exchange and no dependence on kinetic energy density).

One question that arises is whether changing the value of  $Y$  will greatly affect the performance of the new methods when tested against other properties such as atomization energies or barrier heights. The AE6 database<sup>72</sup> is a representative set of data for organic atomization energies, and the BH6 database<sup>72</sup> is a representative set of data for chemical reaction barrier heights. Table 3 contains the results of testing these new methods against the AE6 and BH6 databases, and it compares the results to those for the other methods that are found in this study to do well for the water and to those for the popular B3LYP and PBE functionals. We see that for all of the newly parameterized methods the mean errors per bond for AE6 are a factor of



2–3 higher than for the hybrid and hybrid meta methods, while the errors in BH6 are higher by a factor of 1.3–1.8. However, a comparison of PBE1W and MPWLYP1W to their unoptimized functionals show that the new methods have smaller errors for both AE6 and BH6. Because the *m*PWLYP functional already performs quite well with  $Y = 100$ , its optimum value is close to 100, and therefore its performance is not significantly degraded by optimizing one parameter for water.

Up to this point, all reported results are for the MG3S basis set. Another issue of concern is the performance of the newly parameterized functionals with basis sets other than MG3S or when used in conjunction with pseudopotentials, which are often used in simulations. In order to address these issues, a representative set of water dimers and trimers was taken from the complete set of thirty-six structures, following the procedure of Lynch *et. al.*<sup>72</sup> This representative set, called W7, contains two literature dimers, one liquid and one vapor-phase dimer from simulations, one high-pressure dimer from simulation, and one trimer taken from the literature and simulation (additional information about W7 can be found in supporting information). Table 4 shows the results of using this representative set in testing the three new functionals against their unoptimized counterparts and the three best methods in Table 1 (PW6B95, B98, and MPWB1K) using the DIDZ<sup>73</sup> (also known as 6-31\*G(d,p)) basis set, and the aug-cc-pVDZ<sup>74</sup> and aug-cc-pVTZ<sup>74</sup> basis sets. For comparison the MG3S values for W7 are also reported.

From the data in Table 4 we calculate the mean unsigned error, averaged over four basis sets, of the four new functionals (last four rows) to be 0.47 kcal/mol, whereas the mean unsigned error of the four functionals from the literature that do not contain Hartree-Fock exchange (first four rows) is 0.72 kcal/mol. If we limit the averages to the three larger basis sets, the MUE for the four new functionals is 0.32 kcal/mol, and that for the four older functionals without

Hartree-Fock exchange is 0.49 kcal/mol. Thus the mean unsigned errors are reduced by 30-36%. The details in the numbers that contribute to these averages are also of interest. For example we see that the performance of various functionals with the DIDZ basis set follows the same trend as the MG3S basis set: all four of the newly optimized functionals do better than the unoptimized functionals, and PBE1W has errors similar to B98, MPWB1K, and PW6B95. For the Dunning basis sets, however, the trends are quite different. With the aug-cc-pVDZ basis set PBE1W still has errors similar to B98, MPWB1K, and PW6B95; however, we find that *m*PWLYP becomes the best method and that MPWLYP1W and TPSSLYP1W do not do better than *m*PWLYP and TPSSLYP. When using the aug-cc-pVTZ basis *m*PWLYP, PBE, and PBELYP outperform all of the other functionals (including the hybrid and hybrid meta functionals) by ~0.2 kcal/mol. Of the newly optimized functional only TPSSLYP1W shows improved results. We do, however, find that the optimized functionals have errors similar to B98, MPWB1K, and PW6B95.

One possible explanation for these differences is basis set superposition error. Table 5 shows the results of computing errors after applying counterpoise corrections,<sup>75</sup> again for the W7 data set for the same four basis sets. We see that the general trends and magnitude of the errors for the aug-cc-pVDZ and MG3S basis sets are similar to the aug-cc-pVTZ basis set, which changed very little after inclusion of counterpoise correction. Additionally, we see that the trends for the DIDZ basis set are quite different from the other three, and they still show that the newly optimized methods have errors similar to B98, PW6B95, and MPWB1K and are better than the unoptimized functionals. Both of the Dunning basis sets and the MG3S basis sets have reduced errors after counterpoise correction and now predict that *m*PWLYP, PBE, and TPSSLYP functionals outperform all of the other functionals by approximately 0.3 kcal/mol. Additionally, the PBE functional is shown to be the best method for all three basis sets.

Given the different results obtained when counterpoise correction is included one may

wonder if inclusion of counterpoise correction during the optimizations of the new functionals would be advantageous. However, the inclusion of counterpoise corrections is problematic. It has been argued<sup>76</sup> that counterpoise corrections are an overestimate because they involve all the orbitals of the interactive partner, whereas only the unoccupied space is actually available. Nevertheless for minimal basis sets, counterpoise corrections are usually appreciable and in the correct direction. For moderate and large basis sets, other sources of error may be comparable to or larger than the basis set superposition error, and counterpoise corrections do not always improve the results. (This is consistent with previous findings, 77.) But there are other, even more serious issues. For example, counterpoise corrections on charge transfer processes, barrier heights, and bond energies are not negligible but are invariably neglected. Thus workers who are interested in complex potential energy surfaces almost always omit counterpoise corrections, while workers entirely focused on dimer interaction energies often include them. When one considers trimers, though, counterpoise corrections become ambiguous,<sup>78</sup> and for larger oligomers or condensed-phase systems, they are essentially impossible to carry out, except by removing clusters from the liquid ensemble (generated without such corrections) for isolated analysis.<sup>79</sup> For these reasons, we chose to optimize our new functionals without counterpoise correction. One might then worry about canceling a systematic basis-set superposition error with a systematic error in the density functionals. This is a possibility, but we consider it to be just one of the many sources of error that affect practical calculations. We do point out though that counterpoise corrections on noncovalent interactions tend to be smaller with DFT than with wave function theory (WFT), so those workers who have more experience with WFT than DFT sometimes overestimate the magnitude of the problem.

We have also considered the effect of adding an effective core potential<sup>80</sup> to represent the core of the oxygen atoms. Table 6 shows the results of using the CEP effective core potential,<sup>81</sup>

as implemented in *Gaussian03*, with the MG3S all electron basis set, with and without counterpoise corrections. The results of adding the effective core potential without using counterpoise corrections show that the mean errors increase by  $\sim 0.3$  kcal/mol for almost all of the methods. The two notable exceptions to this are the PW6B95 and MPWLYP1W functionals, which have increases in their mean unsigned errors by 0.7 kcal/mol and 0.05 kcal/mol respectively. The effect of adding counterpoise corrections lowers the errors for all seven functionals taken from the literature, except for B98, and raises the errors for all for the newly parameterized methods except for PBE1W, which improves by  $\sim 0.1$  kcal/mol. Practitioners who employ effective core potentials must exercise due diligence to be sure that this does not increase the errors, especially if they also reduce the basis set.

Finally one may want to consider whether results obtained using an atom-centered Gaussian basis set are relevant to those interested in using plane wave basis sets. A recent study<sup>25</sup> compared the atomization energies for the G2-1 test set computed with the PBE and PBE1PBE (PBE0) density functionals using an atom-centered basis set in *Gaussian03* and with a plane wave basis set in the VASP code. The authors found that the results obtained using the aug-cc-pV5Z basis set versus a plane wave basis set with an energy cutoff of 1000 eV (73.5 Ry) showed good agreement. Furthermore, Ireta *et. al*<sup>82</sup> used the PBE functional and plane wave basis set, with a 70 Ry energy cut off, to determine the binding energies for a set of representative hydrogen bonded dimers. The binding energies obtained were in excellent agreement with a separate study<sup>37</sup> that calculated the binding energies for the same dimers using the PBE functional and the MG3S basis set. With these results in mind we feel confident that results obtained from this study are relevant to those using plane wave basis sets.

## 6. Concluding Remarks

In this letter we have evaluated the performance of twenty-five previously developed and broadly parameterized density functionals against a database containing twenty-eight water dimers and eight water trimers. Among these existing methods we have found the best GGA, meta GGA, hybrid GGA and hybrid meta GGA to be *mPWLYP*, TPSS, B98 and PW6B95 respectively. As far as we know, *mPWLYP*, B98, and PW6B95 have never been used for a condensed-phase simulation of any system, although TPSS has been used for water. Furthermore, we propose a new GGA functional involving a single parameter, PBE1W, that gives mean errors smaller than the best hybrid method, B98, and comparable to the best hybrid meta method, PW6B95.

We have gone on to examine the performance of PBE1W when computing atomization energies and barrier heights, as well as when used with several different basis sets and with a pseudopotential, and have found that in the absence of counterpoise corrections it consistently outperforms PBE, which is one of the most commonly used density functionals in plane wave calculations. As there are differences in the accuracy of all the density functionals when used with different basis sets, careful consideration should be given as to which basis set to use before calculations on water clusters are undertaken with any of the density functionals used in this study. Since the new PBE1W method should be easy to implement in most programs (it requires changing one parameter in the widely available PBE functional), its use for condensed-phase simulations of water would seem to be a promising avenue of research.

**Acknowledgments** The authors are grateful to Ilja Siepmann, Matthew McGrath, Renata Wentzcovich, and Koichiro Umemoto for providing structures, to Yan Zhao for help with the hybrid meta and X3LYP calculations, and to all these people and Nate Schultz for helpful suggestions. This work was supported in part by the National Science Foundation under grants

nos. CHE03-49122 (PI: Donald G. Truhlar) and ITR-0428774 (PI: Renata Wentzcovich) and by the ONR under grant no. N00014-05-1-0538 (PI: Donald G. Truhlar).

**Supporting Information Available** The accurate binding energies and geometries for all dimers and trimers in the water database and the list dimers and trimers in the W7 database are given in supporting information.

## References

- (1) Kim, K.; Jordan, K. D. *J. Phys. Chem.* **1994**, *98*, 10089.
- (2) Lee, C.; Chen, H.; Fitzgerald, G. *J. Chem. Phys.* **1994**, *101*, 4472.
- (3) Estrin, D. A.; Paglieri, L.; Corongiu, G.; Clementi, E. *J. Phys. Chem.* **1996**, *100*, 8701.
- (4) Gregory, J. K.; Clary, D. C. *J. Phys. Chem.* **1997**, *101*, 6813.
- (5) Akmatskaya, E. V.; Cooper, M. D.; Burton, N. A.; Masters, A. J.; Hillier, I. H. *Int. J. Quantum Chem.* **1999**, *74*, 709.
- (6) Kryachko, E. S. *Chem. Phys. Lett.* **1999**, *314*, 353.
- (7) Lee, H. M.; Suh, S. B.; Lee, J. Y.; Tarakeshwar, P.; Kim, K. S. *J. Chem. Phys.* **2000**, *112*, 9759.
- (8) Lee, H. M.; Suh, S. B.; Kim, K. S. *J. Chem. Phys.* **2001**, *114*, 10749.
- (9) Xantheas, S. S.; Burnham, C. J.; Harrison, R. J. *J. Chem. Phys.* **2002**, *116*, 1493.
- (10) Xin, X.; Goddard, W. A. III *J. Phys. Chem. A* **2004**, *108*, 2305.
- (11) Su, J. T.; Xu, X.; Goddard, W. A. III *J. Phys. Chem. A* **2004**, *108*, 10518.
- (12) Xantheas, S. S.; Aprà, E. *J. Chem. Phys.* **2004**, *120*, 823.
- (13) Dunn, M. E.; Pokon, E. K.; Shields, G. D. *J. Amer. Chem. Soc.* **2004**, *126*, 2647.
- (14) Belair, S. D.; Francisco, J. D.; Singer, S. J. *Phys. Rev. A* **2005**, *71*, 13204.
- (15) Laasonen, K.; Sprik, M.; Parrinello, M.; Car, R. *J. Chem. Phys.* **1993**, *99*, 9080.

- (16) Sprik, M.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1996**, *105*, 1142.
- (17) Vassiliev, P.; Hartnig, C.; Koper, M. T. M.; Frechard, F.; vanSanten, R. A. *J. Chem. Phys.* **2001**, *115*, 9815.
- (18) Izvekov, S.; Voth, G. A. *J. Chem. Phys.* **2002**, *116*, 10372.
- (19) Asthagiri, D.; Pratt, L. R.; Kress, J. D. *Phys. Rev. E* **2003**, *68*, 41505.
- (20) Grossman, J. C.; Schwegler, E.; Draeger, E. W.; Gygi, F.; Galli, G. *J. Chem. Phys.* **2004**, *120*, 300.
- (21) VandeVondele, J.; Mohamed, F.; Krack, M.; Hutter, J.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **2005**, *122*, 14515.
- (22) McGrath, M. J.; Siepmann, J. I.; Kuo, I.-F. W.; Mundy, C. J.; VandeVondele, J.; Hutter, J.; Mohamed, F.; Krack, M. *Proc. Natl. Acad. Sci.* submitted for publication.
- (23) Kuo, I-F. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I.; VandeVondele, J.; Sprik, M.; Hutter, J.; Chen, B.; Klein, M. L.; Mohamed, F.; Krack, M.; Parrinello, M. *J. Phys. Chem. B.* **2004**, *108*, 12990-12998.
- (24) Kuo, I-F. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I.; VandeVondele, J.; Sprik, M.; Hutter, J.; Chen, B.; Klein, M. L.; Mohamed, F.; Krack, M.; Parrinello, M. *J. Phys. Chem. B.* **2004**, *108*, 12990-12998.
- (25) Paier, J.; Hirschi, R.; Marsman, M.; Kresse, G. *J. Chem. Phys.*, **2005**, *122*, 234102-1.
- (26) Kresse, G.; Hafner, J. *Phys. Rev. B.*, **1993**, *47*, 558.
- (27) Schwegler, E.; Challacombe, M. *J. Chem. Phys.* **1996**, *105*, 2726.



- (28) Ochsenfeld, C.; White, C. A.; Head-Gordon, M. *J. Chem. Phys.* **1998**, *109*, 1663.
- (29) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207.
- (30) Izvekov, S.; Mazzolo, A.; VanOpdorp, K.; Voth, G. A. *J. Chem. Phys.* **2001**, *114*, 3248.
- (31) Izvekov, S.; Voth, G. A. *J. Chem. Phys.* **2001**, *115*, 7206.
- (33) Yanagisawa, S.; Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2000**, *112*, 545.
- (33) Barden, C. J.; Rienstra-Kiracofe, J. C.; Schaefer, H. F. III *J. Chem. Phys.* **2000**, *113*, 690.
- (34) Wu, Z. *J. Chem. Phys. Lett.* **2004**, *383*, 251.
- (35) Schultz, N. E.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* in press.
- (36) Hall, R. J.; Hillier, I. H.; Vincent, M. A. *Chem. Phys. Lett.* **2000**, *320*, 139.
- (37) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* in press.
- (38) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A*, in press.
- (39) Tschumper, G. S.; Leininger, M. L.; Hoffman, B. C.; Valeev, E. F.; Schaefer, H. F. III;  
Quack, M. *J. Chem. Phys.*, **2002**, *116*, 690.
- (40) Anderson, J. A.; Crager, J.; Federoff, L.; Tschumper, G. S. *J. Chem. Phys.* **2004**, *121*,  
11023.
- (41) Oliveira, G. d.; Martin, J. M. L. *J. Chem. Phys.* **1999**, *111*, 1843.
- (42) Parthiban, S.; Martin, J. M. L. *J. Chem. Phys.* **2001**, *114*, 6014.

(43) Werner, H.-L.; Knowles, P. J.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Celani, P.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; Korona, T.; Lindh, R.; Lloyd, A. W.; McNicholas, S. J.; Manby, F. R.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Rauhut, G.; Schütz, M.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. MOLPRO; 2002.6 ed.; University of Birmingham: Birmingham, 2002.

(44) Umemoto, K.; Wentzcovich, R., unpublished.

(45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian 03, Revision C.01 ed.; Gaussian, Inc., Wallingford CT, 2004.

(46) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384.

(47) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem Phys.* **1980**, *72*, 650.

- (48) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comp. Chem.* **1980**, *4*, 294.
- (49) Slater, J. C. *Quantum Theory of Molecular and Solids. Vol. 4: The Self-Consistent Field for Molecular and Solids* (McGraw-Hill, New York, 1974).
- (50) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.
- (51) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664.
- (52) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (53) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (54) Becke, A. D. *Phys. Rev. A*, **1988**, *38*, 3098.
- (55) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. *J. Chem. Phys.* **1998**, *109*, 6264.
- (56) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403.
- (57) Hoe, W.-M.; Cohen, A. J.; Handy, N. C. *Chem. Phys. Lett.* **2001**, *341*, 319.
- (58) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401-1.
- (59) Perdew, J. P.; Tao, J.; Staroverov, V. N.; Scuseria, G. E. *J. Chem. Phys.* **2004**, *120*, 6898.
- (60) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.
- (61) Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **1998**, *108*, 9624.
- (62) Xu, X.; Goddard, W. A. III *Proc. Natl. Acad. Sci.* **2004**, *101*, 2673.

- (63) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
- (64) Perdew, J. P. In *Electronic Structure of Solids '91*, Ziesche, P., Eschig, H., Eds; Akademie Verlag: Berlin, 1991; p 11.
- (65) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 4811.
- (66) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem.* **2004**, *108*, 6908.
- (67) Rey, J.; Savin, A. *Int. J. Quantum Chem.* **1998**, *69*, 581.
- (68) Kreiger, J. B.; Chen, J.; Iafrate, G. J.; Savin, A. In *Electronic Correlations and Materials Properties*; Gonis, A., Kioussis, N., Eds; Plenum; New York, 1999; p. 463.
- (69) Toulouse, J.; Savin, A.; Adamo, C. *J. Chem. Phys.* **2002**, *117*, 1441.
- (70) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 2715.
- (71) Xu, X.; Zhang, Q.; Muller, R. P.; Goddard, W. A. III *J. Chem. Phys.* **2005**, *122*, 14105.
- (72) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 8996.
- (73) Hehre, W. J.; Radon, L.; Schleyer, P. v. R.; Pople, J. A. in *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1986)
- (74) Dunning, T. H. *J. Chem. Phys.*, **1989**, *90*, 1007.
- (75) Boys, S. F.; Bernardi, D. *Mol. Phys.*, **1970**, *19*, 553.
- (76) Morokuma, K.; Kitaura, K. in *Chemical Applications of Atomic and Molecular Electrostatic*

*Potentials*, edited by P. Politer and D. G. Truhlar (Plenum, New York, 1981), pp 215-242.

(77) Schwenke, D. W.; Truhlar, D. G. *J. Chem. Phys.*, **1985**, *82*, 2481, **1987**, *86*, 3760(E)

(78) Mierzwicki, K.; Latajka, Z. *Chem. Phys. Lett.*, **2003**, *380*, 654.

(79) Malaspina, T.; Coutinho, K.; Canuto, S. *J. Chem. Phys.*, **2002**, *117*, 1692.

(80) Kahn, L. R.; Baybutt, H. P.; Truhlar, D. G. *J. Chem. Phys.*, **1976**, *65*, 3826.

(81) Stevens, W.; Basch, H.; Krauss, J. *J. Chem. Phys.*, **1984**, *81*, 6026.

(82) Ireta, J.; Neugebauer, J.; Scheffler, M. *J. Phys. Chem. A*, **2004**, *108*, 5692.

**Table 1: Mean Errors (kcal/mol) for Previously Developed Functionals<sup>a</sup>**

	references	X	Y	MSE			MUE			RMSE		
				Dimers	Trimers	All	Dimers	Trimers	All	Dimers	Trimers	All
LSDA												
SVWN5	49, 50	0	0	3.98	9.67	5.24	3.98	9.67	5.24	4.45	9.86	6.08
GGA												
mPWLYP	51, 52	0	100	0.16	0.60	0.26	0.18	0.60	0.28	0.22	0.68	0.38
PBE	53	0	100	0.41	0.81	0.50	0.43	0.81	0.52	0.53	0.91	0.63
BLYP	52, 54	0	100	-0.82	-1.77	-1.03	0.82	1.77	1.03	0.87	1.79	1.14
PBELYP	52, 53	0	100	0.80	2.35	1.14	0.80	2.35	1.14	0.82	2.39	1.34
HCTH	55	0	100	-0.81	-3.06	-1.31	0.84	3.06	1.33	0.91	3.21	1.71
OLYP	52, 56, 57	0	100	-2.41	-6.98	-3.42	2.41	6.98	3.42	2.52	7.10	4.02
meta												
TPSS	58, 59	0	100	-0.15	-0.64	-0.26	0.20	0.64	0.30	0.30	0.76	0.44
TPSSLYP	52, 58, 59	0	100	0.43	1.31	0.62	0.43	1.31	0.62	0.47	1.38	0.77
BB95	54, 60	0	100	-1.31	-3.05	-1.69	1.31	3.05	1.69	1.36	3.07	1.88
hybrid GGA												
B98	61	21.98	100	0.03	-0.33	-0.05	0.10	0.33	0.15	0.14	0.39	0.22
B97-1	55	21	100	0.19	0.19	0.19	0.20	0.26	0.22	0.22	0.27	0.23
X3LYP	51, 52, 54, 62	21.8	87.1	0.18	0.31	0.20	0.19	0.33	0.22	0.22	0.37	0.26
B3LYP	52, 54, 63	20	81	-0.25	-0.75	-0.36	0.25	0.75	0.36	0.28	0.77	0.44
MPW1K	51, 64, 65	42.8	100	0.19	-0.49	0.04	0.32	0.50	0.36	0.47	0.55	0.49
mPW1PW91	51, 64	25	100	-0.01	-0.73	-0.17	0.27	0.73	0.37	0.35	0.78	0.48
PBE1PBE	53	25	100	0.43	0.46	0.43	0.44	0.47	0.45	0.58	0.52	0.57
MPW3LYP	51, 52, 66	20	81	0.46	0.98	0.57	0.46	0.98	0.57	0.49	1.01	0.64
B97-2	55	21	100	-0.67	-2.48	-1.07	0.67	2.48	1.07	0.71	2.54	1.35
hybrid meta												
PW6B95	38, 51, 60	28	100	0.02	-0.17	-0.02	0.11	0.19	0.13	0.14	0.24	0.17
MPWB1K	51, 60, 66	44	100	0.11	-0.21	0.04	0.13	0.23	0.15	0.18	0.27	0.20
MPW1B95	51, 60, 66	31	100	-0.05	-0.43	-0.13	0.09	0.43	0.16	0.11	0.47	0.24
PBE1KCIS	37, 53, 67, 68, 69	22	100	0.27	-0.11	0.19	0.30	0.15	0.27	0.74	0.21	0.66
PWB6K	38, 51, 60	46	100	0.54	0.82	0.60	0.54	0.82	0.60	0.56	0.83	0.63
B1B95	54, 60	25	100	-0.86	-2.32	-1.18	0.86	2.32	1.18	0.89	2.34	1.35

<sup>a</sup>All results in this table were obtained using the MG3S basis set.

**Table 2: Mean Errors (kcal/mol) for DFT Functionals Newly Optimized in this Letter<sup>a</sup>**

	<i>X</i>	<i>Y</i>	MSE			MUE			RMSE		
			Dimers	Trimers	All	Dimers	Trimers	All	Dimers	Trimers	All
PBE1W	0	74	-0.01	-0.12	-0.03	0.10	0.20	0.12	0.15	0.35	0.21
MPWLYP1W	0	88	-0.08	-0.01	-0.06	0.15	0.20	0.17	0.21	0.27	0.22
TPSSLYP1W	0	74	-0.09	-0.02	-0.07	0.19	0.20	0.19	0.28	0.28	0.28
PBELYP1W	0	54	-0.11	0.00	-0.09	0.26	0.15	0.24	0.37	0.20	0.34

<sup>a</sup>All results in this letter were obtained using the MG3S basis set.

**Table 3 : Comparison of Mean Errors (kcal/mol) for the AE6 and BH6 Databases<sup>a</sup>**

	<i>X</i>	<i>Y</i>	AE6 <sup>b</sup>			BH6		
			MSE	MUE	RMSE	MSE	MUE	RMSE
<i>m</i> PWLYP	0	100	0.56	1.33	1.55	-8.85	8.85	9.32
PBE	0	100	2.39	3.04	3.47	-9.33	9.33	9.87
B98	21.98	100	-0.41	0.63	0.74	-4.00	4.00	4.49
B97-1	1	100	-0.34	0.91	0.96	-4.14	4.14	4.86
X3LYP	21.8	87.1	-0.51	0.56	0.72	-4.91	4.91	5.20
B3LYP	20	81	-0.60	0.66	0.85	-4.73	4.73	5.07
PW6B95	28	100	0.08	0.39	0.55	-3.46	3.46	3.71
MPWB1K	44	100	-0.69	0.98	1.31	-1.32	1.32	1.73
MPW1B95	31	100	0.35	0.78	0.86	-3.38	3.38	3.61
PBE1W	0	74	1.95	2.30	2.69	-8.50	8.50	8.93
MPWLYP1W	0	88	0.65	1.23	1.47	-8.47	8.47	8.96
TPSSLYP1W	0	74	-1.36	2.07	2.70	-5.95	6.35	7.06
PBELYP1W	0	54	1.92	1.92	2.28	-8.13	8.13	8.69

<sup>a</sup>The functionals are listed in the same order as in Tables 1 and 2. All results in this table were obtained using the MG3S basis set.

<sup>b</sup>For AE6 the mean errors are expressed on a per bond basis.



**Table 4 : Mean Errors (in kcal/mol) for different basis sets**

	DIDZ			aug-cc-pVDZ			aug-cc-pVTZ			MG3S			All <sup>a</sup>
	MSE	MUE	RMSE	MSE	MUE	RMSE	MSE	MUE	RMSE	MSE	MUE	RMSE	MUE
mPWLYP	1.27	1.27	1.63	0.03	0.16	0.22	-0.21	0.23	0.32	0.24	0.29	0.41	0.46
PBE	1.54	1.54	1.90	0.25	0.34	0.42	0.01	0.27	0.36	0.46	0.54	0.65	0.61
PBELYP	2.19	2.19	2.68	0.96	0.96	1.21	0.68	0.68	0.89	1.10	1.10	1.37	1.12
TPSSLYP	1.67	1.67	2.08	0.42	0.43	0.60	0.15	0.25	0.35	0.59	0.59	0.79	0.67
B98	0.98	0.98	1.19	-0.26	0.26	0.33	-0.47	0.47	0.61	-0.08	0.20	0.24	0.46
PW6B95	1.00	1.00	1.19	-0.19	0.21	0.27	-0.48	0.48	0.59	-0.03	0.10	0.12	0.44
MPWB1K	1.06	1.06	1.23	-0.15	0.16	0.26	-0.43	0.43	0.60	0.02	0.15	0.19	0.44
PBE1W	1.02	1.02	1.30	-0.24	0.24	0.29	-0.48	0.48	0.58	-0.05	0.17	0.24	0.46
MPWLYP1W	0.96	0.96	1.28	-0.28	0.28	0.35	-0.51	0.51	0.61	-0.07	0.24	0.27	0.47
TPSSLYP1W	0.99	0.99	1.31	-0.24	0.26	0.35	-0.50	0.50	0.61	-0.08	0.26	0.31	0.47
PBELYP1W	0.98	0.98	1.31	-0.22	0.25	0.38	-0.47	0.47	0.59	-0.08	0.24	0.34	0.46

<sup>a</sup>Average over four basis sets

**Table 5 : Mean Errors (in kcal/mol) for different basis sets, including counterpoise corrections**

	DIDZ			aug-cc-pVDZ			aug-cc-pVTZ			MG3S		
	MSE	MUE	RMSE	MSE	MUE	RMSE	MSE	MUE	RMSE	MSE	MUE	RMSE
<i>m</i> PWLYP	0.26	0.52	0.69	-0.42	0.42	0.52	-0.29	0.29	0.39	-0.34	0.34	0.46
PBE	0.53	0.53	0.81	-0.07	0.24	0.35	-0.07	0.24	0.35	-0.12	0.15	0.25
PBELYP	1.18	1.18	1.58	0.49	0.53	0.73	0.62	0.62	0.82	0.55	0.59	0.78
TPSSLYP	0.67	0.73	1.05	-0.03	0.26	0.31	0.05	0.24	0.30	0.01	0.24	0.32
B98	0.04	0.34	0.40	-0.67	0.67	0.81	-0.55	0.55	0.70	-0.62	0.62	0.76
PW6B95	0.07	0.39	0.45	-0.60	0.60	0.73	-0.57	0.57	0.70	-0.56	0.56	0.70
MPWB1K	0.16	0.36	0.40	-0.54	0.54	0.69	-0.52	0.52	0.70	-0.51	0.51	0.67
PBE1W	0.02	0.37	0.47	-0.67	0.67	0.78	-0.55	0.55	0.66	-0.61	0.61	0.72
MPWLYP1W	-0.05	0.44	0.59	-0.73	0.73	0.85	-0.60	0.60	0.70	-0.64	0.64	0.78
TPSSLYP1W	0.00	0.47	0.61	-0.70	0.70	0.83	-0.60	0.60	0.72	-0.65	0.65	0.80
PBELYP1W	0.00	0.49	0.64	-0.68	0.68	0.83	-0.54	0.54	0.67	-0.62	0.62	0.80

**Table 6 : Mean errors (in kcal/mol) with and without counterpoise correction for W7 database using the MG3S basis set with the CEP effective core potential on oxygen**

	MSE		MUE		RMSE	
	no-cp <sup>a</sup>	cp	no-cp	cp	no-cp	cp
<i>m</i> PWLYP	0.56	-0.24	0.56	0.29	0.74	0.43
PBE	0.96	0.12	0.96	0.25	1.21	0.29
PBELYP	1.46	0.69	1.46	0.71	1.78	0.94
TPSSLYP	1.02	0.18	1.02	0.33	1.29	0.47
B98	0.15	-0.68	0.51	0.68	0.71	0.84
PW6B95	0.64	-0.21	0.80	0.26	0.93	0.35
MPWB1K	0.26	-0.13	0.29	0.20	0.40	0.27
PBE1W	0.42	-0.39	0.48	0.40	0.58	0.49
MPWLYP1W	0.25	-0.54	0.29	0.55	0.43	0.70
TPSSLYP1W	0.36	-0.49	0.40	0.49	0.56	0.67
PBELYP1W	0.28	-0.48	0.37	0.48	0.48	0.69

<sup>a</sup> no-cp denotes no counterpoise correction, cp denotes counterpoise correction.