# An extended hybrid density functional (X3LYP) with improved descriptions of nonbond interactions and thermodynamic properties of molecular systems

Xin Xu[a)]
*State Key Laboratory for Physical Chemistry of Solid Surfaces; Center for Theoretical Chemistry, Department of Chemistry, Xiamen University, Xiamen 361005, China and Materials and Process Simulation Center, Beckman Institute (139-74), Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125*

Qingsong Zhang, Richard P. Muller, and William A. Goddard III[b)]
*Materials and Process Simulation Center, Beckman Institute (139-74), Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125*

(Received 15 June 2004; accepted 10 September 2004; published online 13 December 2004)

We derive here the form for the exact exchange energy density for a density that decays with Gaussian-type behavior at long range. This functional is intermediate between the B88 and the PW91 exchange functionals. Using this modified functional to match the form expected for Gaussian densities, we propose the X3LYP extended functional. We find that X3LYP significantly outperforms Becke three parameter Lee–Yang–Parr (B3LYP) for describing van der Waals and hydrogen bond interactions, while performing slightly better than B3LYP for predicting heats of formation, ionization potentials, electron affinities, proton affinities, and total atomic energies as validated with the extended G2 set of atoms and molecules. Thus X3LYP greatly enlarges the field of applications for density functional theory. In particular the success of X3LYP in describing the water dimer (with $R_e$ and $D_e$ within the error bars of the most accurate determinations) makes it an excellent candidate for predicting accurate ligand–protein and ligand–DNA interactions. © *2005 American Institute of Physics.* [DOI: 10.1063/1.1812257]

## I. INTRODUCTION

Density-functional theory[1] (DFT) has become the method of choice for first principles quantum chemical calculations of the electronic structure and properties of many molecular and solid systems. With the exact exchange-correlation functional, DFT could take into full account of all complex many-body effects at a computational cost characteristic of mean-field approximations.[1] However, the exact exchange-correlation functional is unknown, making it essential to pursue more and more accurate and reliable approximate functionals.

Various approximations to the exchange-correlation energy have been developed and tested in recent decades.[2–41] A foundation of most approaches is the local density approximation (LDA) based on solutions of the uniform electron gas (UEG).[2–4] It is well documented that LDA yields results of good or moderate accuracy for such properties as lattice constants, bulk moduli, equilibrium geometries, and vibrational frequencies.[5] However, LDA leads to bond energies and cohesive energies far too large, making it ''not useful for thermochemistry.''[6]

The generalized gradient approximation (GGA) includes the first-order gradient of the density.[7] The most popular GGA functionals[7–17] include

(i) the B88 exchange functional (Becke[8]) which is often combined with the LYP correlation functional, due to Lee–Yang–Parr;[11,12] and
(ii) the ''nonempirical'' exchange-correlation functionals, PW91[9] and PBE,[10] due to Perdew and co-workers.

These GGAs significantly reduce the overbinding tendency of LDA, but generally remain inadequate for thermochemistry of molecules.[6,21–23]

A big step toward greater accuracy was the introduction of hybrid methods[24] that include some amount of ''exact exchange'' on the basis of the adiabatic connection formula.[24–26] The most effective hybrid method is B3LYP,[27] which is formulated as:

$$E_{\mathrm{xc}}^{\mathrm{B3LYP}} = a_0 E_x^{\mathrm{exact}} + (1-a_0) E_{\mathrm{x}}^{\mathrm{Slater}} + a_x \Delta E_x^{\mathrm{B88}} + a_c E_c^{\mathrm{VWN}} + (1-a_c) E_c^{\mathrm{LYP}}, \quad (1)$$

where $a_0 = 0.20$, $a_x = 0.72$, and $a_c = 0.19$. These parameters were derived by Becke from a linear least-squares fit to 56 atomization energies, 42 ionization potentials and 8 proton affinities.[24] These hybrid methods fail to account for van der Waals interactions.[6,29,30]

Numerous efforts have been made to extend GGAs to include these long-range interactions.[31–35] Within the framework of GGA, Adamo and Barone optimized the exponent of the $s^d$ term and the $\beta$ constant of the PW91 functional, fitting at the same time the exact exchange energies of isolated atoms and the differential exchange energy of noble gas

[a)]Author to whom correspondence should be addressed. Electronic mail xinxu@xmu.edu.cn
[b)]Electronic mail: wag@wag.caltech.edu

 

dimers ($He_2$,$Ne_2$) near their van der Waals minima.[16] The resulting mPWPW model is superior to the original PWPW functional for these interactions.[16] Perdew, Burke, and Ernzerhof (PBE[10]) presented a simplified GGA claimed to improve six shortcomings over PW91. Although PBE is not specifically designed for van der Waals systems, it does show the best performance among conventional GGAs.[36–38] However, PBE does not pass the test of 93 chemical systems designed by Handy and co-workers, who concluded that PBE "cannot be recommended for chemistry."[39]

We present here the X3LYP extended functional, which predicts accurate electronic and thermodynamic properties of molecular systems with improved descriptions of the equilibrium properties of hydrogen bonded and van der Waals systems, thus greatly extending the applicability of density-functional theory.

## II. FORMULATION OF THE EXTENDED FUNCTIONAL

### A. Background

We will assume that the exchange-correlation functional is separable

$$E_{xc}=E_x+E_c. \tag{2}$$

Since the magnitude of the correlation energy is generally less than 10% of the exchange energy, we consider that it is most important that the exchange functional be accurate.[41] Thus in the present work we choose the LYP[11,12] correlation functional and focus on developing an improved exchange functional.

The exchange energy, $E_x$, is expressed as

$$E_x=\int \epsilon_x[\rho(\mathbf{r}),|\nabla\rho(\mathbf{r})|,\ldots]d\mathbf{r}, \tag{3}$$

where $\epsilon_x$ is the exchange energy distribution per unit volume.[1]

In LDA[2]

$$\epsilon_x^{\mathrm{LDA}}(\rho)=A_x\rho(\mathbf{r})^{4/3}, \tag{4}$$

where $A_x=-\frac{3}{4}(3/\pi)^{1/3}$. Thus $\epsilon_x^{\mathrm{LDA}}$ depends on the density only at the point where it is evaluated.

In GGA[7]

$$\epsilon_x^{\mathrm{GGA}}(\rho,|\nabla\rho|)=\epsilon_x^{\mathrm{LDA}}\cdot F(s), \tag{5}$$

where $F(s)$ is an *enhancement factor* and $s$ is the *dimensionless gradient* defined as[7]

$$s=\frac{|\nabla\rho|}{(24\pi^2)^{1/3}\rho^{4/3}}. \tag{6}$$

The well-established B88 exchange functional takes the form[8]

$$F^{\mathrm{B88}}(s)=\frac{1+s\cdot a_1\cdot\sinh^{-1}(s\cdot a_2)+a_3\cdot s^2}{1+s\cdot a_1\cdot\sinh^{-1}(s\cdot a_2)}. \tag{7}$$

Another popular exchange functional is PW91[9]

$$F^{\mathrm{PW91}}(s)=\frac{1+s\cdot a_1\cdot\sinh^{-1}(s\cdot a_2)+(a_3+a_4\cdot e^{-100s^2})s^2}{1+s\cdot a_1\cdot\sinh^{-1}(s\cdot a_2)+a_5\cdot s^d}. \tag{8}$$

Here $a_2=(48\pi^2)^{1/3}$, $a_1=6\beta\cdot a_2$, $a_3=-a_2^2/(2^{1/3}A_x)\cdot\beta$, $a_4=10/81-a_3$, $a_5=-a_2^4\times10^{-6}/(2^{1/3}A_x)$, and $d=4$. Becke obtained $\beta=0.0042$ from fitting to Hartree–Fock (HF) exchange energies for the noble gas atoms.[8] Note that if $a_4$ and $a_5$ are set to zero, PW91 exchange has the same form as B88.

The PBE exchange functional takes the form[10]

$$F^{\mathrm{PBE}}(s)=\frac{1+0.492\,53s^2}{1+0.273\,02s^2}. \tag{9}$$

The $F^{\mathrm{B88}}(s)$, $F^{\mathrm{PW91}}$, and $F^{\mathrm{PBE}}(s)$ functions are plotted in Fig. 1. These three functions are similar for small $s$, but differ significantly for large $s$, which is the region believed to be important for describing van der Waals systems.[41]

It has been shown that as $r$ approaches infinity, $\rho(r)$ approaches $\exp(-\alpha r)$ so that

$$\lim_{r\to\infty}\epsilon_x=-\frac{\rho(r)}{2r}\ \ (\text{Condition 1}). \tag{10}$$

$F^{\mathrm{B88}}(s)$ assures this correct asymptotic behavior of the exchange energy density.[8]

Levy and Perdew showed that some scaling properties can be satisfied if the asymptotic form of the functional for large $s$ is $s^{-\alpha}$, where $\alpha\geqslant1/2$ (Condition 2).[42] Another condition is the Lieb–Oxford bound (Condition 3),[43] which in its global version states that[43]

$$E_x\geqslant E_{xc}\geqslant-1.679\int\rho(r)^{4/3}dr. \tag{11}$$

The local Lieb–Oxford bound suggests that $F(s)$ should be bound from above at large $s$.

B88 violates Conditions 2 and 3. It is the $a_5\cdot s^4$ term in $F^{\mathrm{PW91}}(s)$ that leads the PW91 exchange functional to obey the Levy scaling inequalities and the Lieb–Oxford bound. However, PW91 violates Condition 1.

Condition 1 and Conditions 2 and 3 cannot be simultaneously satisfied by functionals of the GGA form. Thus the large $s$ behavior cannot be uniquely fixed with these mathematical conditions. It has been argued that the Lieb–Oxford bound is more important than other conditions for a weakly bound system.[44] In fact, PBE sacrifices Condition 2 to avoid the $F(s)$ turnover of PW91, which is suspected to cause spurious wiggles in the potential for large $s$.[10]

### B. Gaussian-type density decay

According to the Fermi–Amaldi model,[18,19] the exchange energy $E_x(\rho)$ may be approximated by the classical Coulomb repulsion $J(\rho)$ via:

$$E_x(r)=-\frac{J(\rho)}{N}, \tag{12}$$

$$J(\rho)=\frac{1}{2}\int\int\frac{\rho(\mathrm{r})\rho(\mathrm{r}')}{|r-r'|}d\mathbf{r}d\mathbf{r}', \tag{13}$$

where $N$ is the total number of electrons of the system concerned.

The exchange potential $V_x(\mathbf{r})$ associated with Eq. (12) is

$$V_x(\mathbf{r})=\frac{\delta E_x(\rho)}{\delta\rho(r)}=-\frac{1}{\mathrm{N}}\int\frac{\rho(r')}{|r-r'|}d\mathbf{r}'=-\frac{V_c(\mathbf{r})}{N}, \quad (14)$$

where $V_c(\mathbf{r})$ is the Coulomb potential and $N$ has been kept fixed in deriving (14).

To get $E_x(\rho)$ from $V_x(r)$, one may use

$$E_x(\rho)=\frac{1}{2}\int\rho(r)V_x(r)d\mathbf{r}. \quad (15)$$

As most DFT calculations on finite molecules use Gaussian basis functions, we therefore, consider that the long-range behavior of the electron density may have the form of a Gaussian function[20]

$$\lim_{r\to\infty}\rho(r)=2\left(\frac{2Z}{\pi}\right)^{3/2}e^{-2Zr^2}. \quad (16)$$

Inserting (16) into (14), one finds

$$V_x^{\mathrm{Gauss}}(r)=-\frac{\mathrm{erf}(\sqrt{2Z}r)}{r}. \quad (17)$$

Since $\lim_{r\to\infty}(\mathrm{erf}(\sqrt{2Z}r))=1$, it is clear from Eqs. (3), (15), and (17) that Condition 1 is fulfilled for a Gaussian-type density.

Combining Eqs. (3)–(5) with Eq. (15) gives

$$F(r)=\frac{V_x(r)}{2A_x\rho^{1/3}(r)}. \quad (18)$$

Thus we arrive at

$$F^{\mathrm{Gauss}}(r)=\frac{(2\pi^5)^{1/6}\mathrm{erf}(\sqrt{2Z}r)}{3^{4/3}e^{-2Zr^2/3}\sqrt{Z}r} \quad (19)$$

by inserting Eqs. (16) and (17) into Eq. (18).

Using Eq. (16), we can rewrite Eq. (6) as

$$s(r)=\frac{\left(\frac{2}{9\pi}\right)^{1/6}\sqrt{Z}r}{e^{-2Zr^2/3}}. \quad (20)$$

Equations (19) and (20) determine the $F^{\mathrm{Gauss}}(s)$ for a Gaussian-type asymptotic density, which for $s\geqslant 1.5$ (Fig. 1) lies between $F^{\mathrm{B88}}(s)$ and $F^{\mathrm{PW91}}(s)$, but closer to $F^{\mathrm{B88}}(s)$. Note that as $s\to 0$, $F^{\mathrm{Gauss}}(s)\to(2^5\pi/3^4)^{1/3}=1.074\,66$, instead of 1.0 as required to obey the limit within the local density approximation (LDA). This may not be necessary for a finite system. Thus Handy *et al.* recently developed a local exchange functional, OPTX, by fitting to the unrestricted HF energies of the first- and the second-row atoms and found the LDA term to be 1.051 51 rather than 1.0.[17,40]

## III. X3LYP FUNCTIONAL

Based on the $F^{\mathrm{Gauss}}(s)$ behavior for $s\geqslant 1.5$, as shown in Fig. 1, we propose the extended exchange functional

$$F^X(s)=1+a_{x1}(F^{\mathrm{B88}}(s)-1)+a_{x2}(F^{\mathrm{PW91}}(s)-1). \quad (21)$$

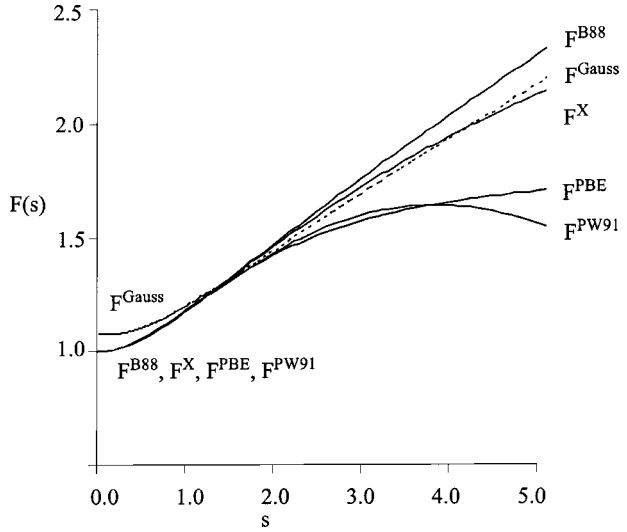


FIG. 1. Enhancement factors, $F^{\mathrm{B88}}(s)$, $F^{\mathrm{PW91}}(s)$, $F^{\mathrm{PBE}}(s)$, $F^{\mathrm{Gauss}}(s)$, and $F^{\mathrm{X}}(s)$, for a set of GGA functionals. $F^{\mathrm{Gauss}}$ is shown with a dashed line.

Here we choose to obey the LDA limit as $s\to 0$, as usually done in the framework of GGA. It is not necessary to restrict the $F^X$ to be a linear combination of B88 and PW91; however, we found that this form allows sufficient flexibility, and we considered that using these well known functions would make it easier to incorporate X3LYP into existing DFT codes.

Following the form of the B3LYP functional, we formulate X3LYP, as

$$E_{xc}^{\mathrm{X3LYP}}=a_{x0}E_X^{\mathrm{exact}}+(1-a_{x0})E_x^{\mathrm{Slater}}+a_x\Delta E_x^{\mathrm{extended}}+a_cE_c^{\mathrm{VWN}}+(1-a_c)E_c^{\mathrm{LYP}}. \quad (22)$$

The parameters $\{a_{x0},a_x,a_c\}$ in Eq. (22) and $\{a_{x1},a_{x2}\}$ in Eq. (21) are determined through least-square fitting of the chemical properties for a small set of atoms and molecules listed in Table I:

(i) Group (1) in Table I contains total energies of the first 10 atoms (H to Ne), including 8 cations, and 5 anions. This includes the first ionization potentials (IPs) and electron affinities (EAs). The exact total energies of these atoms and ions are taken from Refs. 45 and 46;

TABLE I. Sets of atoms and molecules used to determine the parameters in X3LYP.

| |
|---|
| (1) Total energies: |
| (a) Atoms: H, He, Li, Be, B, C, N, O F, Ne |
| (b) Cations: $Li^+$, $Be^+$, $B^+$, $C^+$, $N^+$, $O^+$, $F^+$, $Ne^+$ |
| (c) Anions: $H^-$, $B^-$, $C^-$, $O^-$, $F^-$ |
| (2) Ionization potentials |
| Na, Mg, Al, Si, P, S, Cl, Ar |
| (3) Electron affinities |
| Al, Si, P, S, Cl |
| (4) Atomization energies |
| $H_2$, $He_2$, $Li_2$, $Be_2$, $C_2$, $N_2$, $O_2$, $F_2$, $Ne_2$, $Na_2$, $Mg_2$, $Si_2$, $P_2$, $S_2$, $Cl_2$, CN, CO, CS, NO, SO, ClO, SiO, ClF, PF, AlF, SiF, CCl, SiCl, NaCl, CH, NH, OH, HF, $CO_2$, $O_3$, $SO_3$, OCS, $CS_2$ |

(ii) group (2) contains 8 IPs of the second-row atoms[23] and
(iii) group (3) contains 5 EAs of the second-row atoms;[23]
(iv) group (4) contains 33 diatomic and 5 triatomic molecules from the first and second rows, selected to include a variety of bonding situations including open- and closed-shell molecules; molecules with single, double, and triple bonds; ionic systems, and systems which require multiple configurations. In particular, we include $He_2$ and $Ne_2$ to represent van der Waals systems.

The atomization energies are computed at the experimental geometries.[47] The experimental atomization energies are taken from Refs. 22 and 47–50.

The parameters are optimized by minimizing

$$\nabla = \sum_{i=1}^{n} \left( \frac{E_i}{E_i^{\mathrm{ref}}} - 1 \right)^2 w_i , \tag{23}$$

self-consistently by solving the Kohn–Sham orbital equations. Here $E_i$ is the calculated energy and $E_i^{\mathrm{ref}}$ is the corresponding reference energy[23,45–50] in subsets (1)–(4). All energies are in a.u. The relative weights $w_i$ are adjusted to give a reasonable balance of different contributions. For atomic energies, we use unit weight, except that of H, for which we use a value of $1\times10^3$. For IPs and EAs, we set $w_i=10$. For covalent bindings, the weights are around 50. For weak bindings of $Be_2$, $Mg_2$, $He_2$, and $Ne_2$, large weights of $7\times10^2$, $1\times10^4$, $2\times10^7$, and $3\times10^7$, respectively, are used. All fitting calculations were performed using the aug-cc-pVTZ basis sets.[51,52] The final results for the parameters of X3LYP are

$$\{a_{x0}, a_x, a_c\} = \{0.218,\ 0.709,\ 0.129\},$$
$$\{a_{x1}, a_{x2}\} = \{0.764\,457,\ 0.235\,543\}. \tag{24}$$

Since codes for DFT calculations on solids often use plane wave basis sets that make it expensive to include exact exchange, we also optimized the parameters for the XLYP functional in which the exact exchange term, as well as the VWN term, is deleted. Thus

$$E_{xc}^{\mathrm{XLYP}} = E_x^{\mathrm{Slater}} + a_{x1}\Delta E_x^{\mathrm{B88}} + a_{x2}\Delta E_x^{\mathrm{PW91}} + E_c^{\mathrm{LYP}}, \tag{25}$$

where $\{a_{x1}, a_{x2}\} = \{0.722,\ 0.347\}$.

The $F^X(s)$ from Eq. (24) is plotted in Fig. 1, where we see that it is quite close to $F^{\mathrm{Gauss}}(s)$ for larger $s$.

To validate the accuracy of X3LYP for thermochemistry, we use the extended G2 set of molecules for which there are quite accurate experimental data available.[22,23,46] This set contains:

(i) the heats of formation of 148 molecules;
(ii) 42 ionization potentials;
(iii) 25 electron affinities;
(iv) 8 proton affinities.

These 148 molecules include inorganic compounds and organic compounds; radicals, saturated hydrocarbons, and unsaturated (aromatic) rings. Thus the heats of formation of these molecules provide a good test of the functionals for the thermochemistry of the covalent systems.

We also include 10 total atomic energies of the first row atoms.[45,46]

In addition, we included $He_2$, $Ne_2$, and $Ar_2$ to assess the accuracy of the van der Waals systems. Here we also demanded that exclusion of the LYP correlation function would lead to a repulsive potential curve, similar to HF. This is to eliminate cancellation of errors between the correlation functional and the exchange functional.

We did not include $(H_2O)_2$ in the training set for X3LYP but we use it to validate the accuracy of X3LYP for hydrogen bonded interactions.

For the validation against the G2 set, we use the same second-order Moller–Plesset (MP2) molecular geometries as in G2 theory,[22,23,48] and we use the same scaled HF vibrational frequencies for zero-point energies and finite-temperature corrections. Here we employ the 6-311 $+G(3df,2p)$ basis sets.[22,23,48] This choice of geometries and basis sets allows our results to be compared directly with previously published data obtained with other functionals.[22,53]

For $He_2$ and $Ne_2$, we used the aug-cc-pVTZ basis sets. For $(H_2O)_2$ we used the aug-cc-pVTZ(-f) basis sets. These bonding energies are BSSE-corrected.

All calculations were performed with JAGUAR 4.0,[54] but we did *not* use the pseudospectral method, making it easier to compare with literature data. The ultrafine DFT grids of Jaguar were used in all calculations.

## IV. RESULTS AND DISCUSSION

### A. Heats of formation

Table II lists the experimental heats of formation (298 K) for the extended G2 set of 148 molecules.[22,48] The mean-absolute-deviations (MADs) from experiment (theory-exptl.) for B3LYP,[22] PBE1PBE (PBE0)[53] and X3LYP are presented. The performance of other functionals like PWPW, PWLYP, PW3PW, PW1PW (PW0), and PW3LYP, which have not previously been fully tested over the G2 set, is given in the supplemental material (Table S1). A detailed assessment for the newly developed OLYP[40] and O3LYP[17] functionals will be published elsewhere.

As done elsewhere,[10] we use the notation PBE to signify the PBE exchange functional combined with the PBE correlation functional, and we use PBE1PBE to denote the one-parameter hybrid scheme.[38,53] As the hybrid coefficient was deduced from the perturbation theory arguments, PBE1PBE is considered a parameter-free hybrid model, and often labeled as PBE0.[38] Similarly, we use the notations PW1PW (PW0)[16] or mPW1PW (mPW0).[16]

The best result is for X3LYP with MAD=2.8 kcal/mol. Next best is B3LYP with MAD=3.1 kcal/mol, while PBE0 (PBE1PBE) leads to MAD=4.8 kcal/mol. In contrast, LDA overbinds strongly, leading to MAD=90.9 kcal/mol!

In PBE, hydrogen atoms have a self-correlation energy (3.6 kcal/mol per H), which leads to a spurious lowering of the energy of H.[53] Thus the heat of formation of $H_2$ in PBE

TABLE II. Experimental heats of formation (kcal/mol at 298 K) for the G2 test set (148 molecules) and the deviations (theory-exptl.) obtained from B3LYP, PBE0, and X3LYP.

| No. | Molecule | Exp.[a] | B3LYP | PBE0 | X3LYP |
|---|---|---|---|---|---|
| 1 | $H_2$ | .00 | −1.0 | 6 | −0.431 |
| 2 | LiH | 33.30 | −0.4 | 6 | −0.178 |
| 3 | BeH | 81.70 | −8.2 | −8 | −7.936 |
| 4 | CH | 142.50 | −1.7 | 1 | −1.497 |
| 5 | $CH_2$ ($^3B_1$) | 93.70 | −2.1 | −5 | −1.954 |
| 6 | $CH_2$ ($^1A_1$) | 102.75 | −0.2 | −2 | −0.005 |
| 7 | $CH_3$ | 35.00 | −3.3 | −2 | −3.222 |
| 8 | $CH_4$ | −17.90 | −1.6 | 2 | −1.515 |
| 9 | NH | 85.20 | −4.6 | −3 | −4.380 |
| 10 | $NH_2$ | 45.10 | −6.5 | −2 | −6.188 |
| 11 | $NH_3$ | −10.97 | −3.5 | 2 | −3.117 |
| 12 | OH | 9.40 | −1.8 | 1 | −1.683 |
| 13 | $H_2O$ | −57.80 | 1.3 | 6 | 1.648 |
| 14 | HF | −65.14 | 1.6 | 5 | 1.837 |
| 15 | $SiH_2$ ($^1A_1$) | 65.20 | −2.1 | 7 | −1.976 |
| 16 | $SiH_2$ ($^3B_1$) | 86.20 | −2.3 | −1 | −2.097 |
| 17 | $SiH_3$ | 47.90 | −3.2 | 3 | −2.949 |
| 18 | $SiH_4$ | 8.20 | −1.9 | 8 | −1.720 |
| 19 | $PH_2$ | 33.10 | −6.0 | 0 | −5.812 |
| 20 | $PH_3$ | 1.30 | −3.3 | 3 | −3.055 |
| 21 | $H_2S$ | −4.90 | 0.3 | 2 | 0.486 |
| 22 | HCl | −22.06 | 1.0 | 2 | 1.005 |
| 23 | $Li_2$ | 51.60 | 3.5 | 7 | 3.007 |
| 24 | LiF | −80.10 | 0.5 | 7 | 0.472 |
| 25 | $C_2H_2$ | 54.19 | 2.5 | 0 | 2.524 |
| 26 | $H_2C{=}CH_2$ | 12.54 | −0.6 | −2 | −0.828 |
| 27 | $H_3C{-}CH_3$ | −20.08 | −0.6 | −1 | −1.069 |
| 28 | CN | 104.90 | 2.2 | 0 | 2.972 |
| 29 | HCN | 31.50 | 0.0 | 2 | 0.622 |
| 30 | CO | −26.42 | 3.9 | 6 | 4.168 |
| 31 | HCO | 10.00 | −2.2 | −2 | −1.944 |
| 32 | $H_2C{=}O$ | −25.96 | −0.4 | 3 | −0.177 |
| 33 | $CH_3{-}OH$ | −48.00 | −0.1 | 3 | −0.279 |
| 34 | $N_2$ | .00 | 1.4 | 4 | 2.502 |
| 35 | $H_2N{-}NH_2$ | 22.79 | −6.3 | −1 | −5.850 |
| 36 | NO | 21.58 | −3.0 | −1 | −2.013 |
| 37 | $O_2$ | .00 | −2.0 | −4 | −2.753 |
| 38 | HO−OH | −32.53 | 1.8 | 6 | 2.257 |
| 39 | $F_2$ | .00 | 2.6 | 5 | 3.138 |
| 40 | $CO_2$ | −94.05 | 0.2 | −1 | 0.633 |
| 41 | $Na_2$ | 33.96 | −0.2 | 3 | −0.857 |
| 42 | $Si_2$ | 139.87 | 5.4 | 3 | 5.384 |
| 43 | $P_2$ | 34.31 | 1.4 | 5 | 1.622 |
| 44 | $S_2$ | 30.74 | −1.2 | −8 | −1.302 |
| 45 | $Cl_2$ | .00 | 2.9 | −2 | 2.760 |
| 46 | NaCl | −43.56 | 4.6 | 5 | 4.301 |
| 47 | SiO | −24.64 | 5.5 | 9 | 5.772 |
| 48 | CS | 66.90 | 4.9 | 4 | 5.175 |
| 49 | SO | 1.20 | −0.7 | −3 | −0.304 |
| 50 | ClO | 24.19 | −1.6 | −4 | −1.383 |
| 51 | ClF | −13.24 | 1.1 | 2 | 1.099 |
| 52 | $H_3Si{-}SiH_3$ | 19.10 | −0.2 | 10 | −0.345 |
| 53 | $CH_3Cl$ | −19.56 | 0.8 | −1 | 0.478 |
| 54 | $H_3C{-}SH$ | −5.50 | 1.2 | 0 | 0.956 |
| 55 | HOCl | −17.80 | 1.5 | 2 | 1.619 |
| 56 | $SO_2$ | −70.95 | 10.0 | 4 | 10.515 |
| 57 | $BF_3$ | −271.41 | 3.9 | 3 | 2.824 |
| 58 | $BCl_3$ | −96.30 | 6.3 | −7 | 5.120 |
| 59 | $AlF_3$ | −289.03 | 11.9 | 14 | 11.253 |
| 60 | $AlCl_3$ | −139.72 | 10.2 | 2 | 9.137 |
| 61 | $CF_4$ | −223.04 | 4.5 | 0 | 2.845 |
| 62 | $CCl_4$ | −22.94 | 14.0 | −6 | 12.291 |
| 63 | $O{=}C{=}S$ | −33.08 | −0.5 | −3 | −0.181 |
| 64 | $CS_2$ | 27.95 | 0.2 | −5 | 0.512 |
| 65 | $COF_2$ | −152.70 | 9.1 | −2 | 8.520 |
| 66 | $SiF_4$ | −385.98 | 20.1 | 15 | 18.536 |

TABLE II. (*Continued.*)

| No. | Molecule | Exp.[a] | B3LYP | PBE0 | X3LYP |
|---|---|---|---|---|---|
| 67 | $SiCl_4$ | −158.40 | 18.8 | 3 | 17.034 |
| 68 | $N_2O$ | 19.61 | −2.9 | −3 | −1.162 |
| 69 | ClNO | 12.36 | −2.0 | 0 | −0.737 |
| 70 | $NF_3$ | −31.57 | −4.0 | −2 | −4.049 |
| 71 | $PF_3$ | −229.07 | 7.1 | 5 | 6.175 |
| 72 | $O_3$ | 34.10 | 8.6 | 7 | 7.350 |
| 73 | $F_2O$ | 5.86 | 0.4 | 4 | 1.074 |
| 74 | $ClF_3$ | −37.97 | −1.9 | −3 | −1.907 |
| 75 | $C_2F_4$ | −157.40 | −3.2 | −8 | −4.567 |
| 76 | $C_2Cl_4$ | −2.97 | 11.3 | −12 | 9.618 |
| 77 | $CF_3CN$ | −118.40 | 3.7 | −2 | 2.735 |
| 78 | $C_3H_4$ (propyne) | 44.20 | 1.9 | −5 | 1.448 |
| 79 | $C_3H_4$ (allene) | 45.50 | −1.9 | −8 | −2.408 |
| 80 | $C_3H_4$ (cyclopropene) | 66.20 | 3.2 | −10 | 2.793 |
| 81 | $C_3H_6$ (propylene) | 4.78 | 0.6 | −5 | −0.160 |
| 82 | $C_3H_6$ (cyclopropane) | 12.70 | 2.2 | −9 | 1.340 |
| 83 | $C_3H_8$ (propane) | −25.00 | 1.5 | −3 | 0.335 |
| 84 | $C_4H_6$ (butadiene) | 26.30 | 1.5 | −9 | 0.454 |
| 85 | $C_4H_6$ (2-butyne) | 34.80 | 2.4 | −9 | 1.527 |
| 86 | $C_4H_6$ (methylene cyclopropane) | 47.90 | 0.0 | −14 | −1.105 |
| 87 | $C_4H_6$ (bicyclobutane) | 51.90 | 7.1 | −16 | 5.849 |
| 88 | $C_4H_6$ (cyclobutene) | 37.40 | 6.1 | −12 | 4.711 |
| 89 | $C_4H_8$ (cyclobutane) | 6.80 | 5.2 | −10 | 3.559 |
| 90 | $C_4H_8$ (isobutene) | −4.00 | 3.1 | −7 | 1.593 |
| 91 | $C_4H_{10}$ (trans butane) | −30.00 | 3.7 | −5 | 1.895 |
| 92 | $C_4H_{10}$ (isobutane) | −32.07 | 4.8 | −4 | 2.984 |
| 93 | $C_5H_8$ (spiropentane) | 44.30 | 5.4 | −19 | 3.748 |
| 94 | $C_6H_6$ (benzene) | 19.74 | 4.5 | −24 | 2.175 |
| 95 | $H_2CF_2$ | −107.71 | 0.0 | 1 | −0.622 |
| 96 | $HCF_3$ | −166.60 | 2.2 | 1 | 1.117 |
| 97 | $H_2CCl_2$ | −22.83 | 4.6 | −4 | 3.868 |
| 98 | $HCCl_3$ | −24.66 | 9.0 | −6 | 7.786 |
| 99 | $H_3C–NH_2$ (methylamine) | −5.50 | −3.2 | 0 | −3.260 |
| 100 | $CH_3–CN$ (methyl cyanide) | 18.00 | −0.6 | −2 | −0.426 |
| 101 | $CH_3–NO_2$ (nitromethane) | −17.80 | −2.4 | −2 | −1.800 |
| 102 | $CH_3–O–N=O$ (methyl nitrite) | −15.90 | −1.3 | 0 | −0.805 |
| 103 | $CH_3–SiH_3$ (methyl silane) | −7.00 | 1.0 | 6 | 0.637 |
| 104 | HCOOH (formic acid) | −90.50 | 0.9 | 1 | 0.741 |
| 105 | $HCOOCH_3$ (methyl formate) | −85.00 | 0.2 | 0 | −0.441 |
| 106 | $CH_3CONH_2$ (acetamide) | −57.00 | −1.6 | −6 | −2.320 |
| 107 | $CH_2–NH–CH_2$ (aziridine) | 30.20 | −1.0 | −8 | −1.395 |
| 108 | NCCN (cyanogen) | 73.30 | 0.4 | −3 | 1.344 |
| 109 | $(CH_3)_2NH$ (dimethylamine) | −4.40 | −2.0 | −2 | −2.699 |
| 110 | $CH_3–CH_2–NH_2$ (trans ethylamine) | −11.30 | −2.2 | −3 | −2.947 |
| 111 | $H_2C=C=O$ (ketene) | −11.35 | −2.4 | −6 | −2.456 |
| 112 | $CH_2–O–CH_2$ (oxirane) | −12.57 | 1.4 | −4 | 0.969 |
| 113 | $CH_3CHO$ (acetaldehyde) | −39.70 | 0.3 | −1 | −0.136 |
| 114 | $O=CH–CH=O$ (glyoxal) | −50.70 | 1.6 | 0 | 1.332 |
| 115 | $CH_3CH_2OH$ (ethanol) | −56.21 | 1.9 | 0 | 1.068 |
| 116 | $CH_3–O–CH_3$ (dimethylether) | −44.00 | 0.0 | 1 | −0.739 |
| 117 | $CH_2–S–CH_2$ (thiooxirane) | 19.60 | 3.1 | −7 | 2.574 |
| 118 | $CH_3CH_3SO$ (dimethyl sulfoxide) | −36.20 | 6.5 | −2 | 5.540 |
| 119 | $CH_3–CH_2–SH$ (ethanethiol) | −11.10 | 3.6 | −2 | 2.793 |
| 120 | $CH_3–S–CH_3$ (dimethyl sulphide) | −8.90 | 2.8 | −2 | 1.868 |
| 121 | $H_2C=CHF$ | −33.20 | −1.5 | −3 | −2.048 |
| 122 | $CH_3–CH_2–Cl$ (ethyl chloride) | −26.80 | 2.7 | −3 | 1.859 |
| 123 | $H_2C=CHCl$ (vinyl chloride) | 8.90 | −1.6 | −5 | −2.224 |
| 124 | $H_2C=CHCN$ (acrylonitrile) | 43.20 | 2.0 | −6 | 1.883 |
| 125 | $CH_3–CO–CH_3$ (acetone) | −51.93 | 2.0 | −4 | 0.947 |
| 126 | $CH_3COOH$ (acetic acid) | −103.40 | 2.6 | −2 | 1.790 |
| 127 | $CH_3COF$ (acetyl fluoride) | −105.70 | 1.5 | −3 | 0.822 |
| 128 | $CH_3COCl$ (acetyl chloride) | −58.00 | 2.5 | −5 | 1.805 |
| 129 | $CH_3CH_2CH_2Cl$ (propyl chloride) | −31.52 | 4.6 | −5 | 3.136 |
| 130 | $(CH_3)_2CH–OH$ (isopropanol) | −65.20 | 4.5 | −1 | 2.990 |
| 131 | $C_2H_5–O–CH_3$ (methyl ethyl ether) | −51.70 | 1.5 | −1 | 0.153 |
| 132 | $(CH_3)_3N$ (trimethylamine) | −5.70 | −0.2 | −3 | −1.675 |
| 133 | $C_4H_4O$ (furan) | −8.30 | 4.2 | −15 | 2.653 |

TABLE II. (*Continued.*)

| No. | Molecule | Exp.[a] | B3LYP | PBE0 | X3LYP |
|---|---|---|---|---|---|
| 134 | $C_4H_4S$ (thiophene) | 27.50 | 7.9 | −17 | 6.162 |
| 135 | $C_4H_4NH$ (pyrrole) | 25.90 | 0.8 | −19 | −0.737 |
| 136 | $C_5H_5N$ (pyridine) | 33.60 | 0.2 | −22 | −1.580 |
| 137 | SH | 34.18 | −1.4 | 0 | −1.253 |
| 138 | CCH | 135.10 | 3.4 | −4 | 3.510 |
| 139 | $C_2H_3$ ($^2A'$) | 71.60 | −3.3 | −7 | −3.400 |
| 140 | $CH_3CO$ ($^2A'$) | −2.40 | −2.1 | −6 | −2.376 |
| 141 | $H_2COH$ ($^2A$) | −4.08 | −2.4 | −3 | −2.495 |
| 142 | $CH_3O$ ($^2A'$) | 4.10 | −3.7 | −3 | −3.811 |
| 143 | $CH_3CH_2O$ ($^2A''$) | −3.70 | −1.4 | −5 | 0.486 |
| 144 | $CH_3S$ ($^2A'$) | 29.80 | −1.7 | −3 | −1.989 |
| 145 | $C_2H_5$ ($^2A'$) | 28.90 | −2.8 | −6 | −3.239 |
| 146 | $(CH_3)_2CH$ ($^2A'$) | 21.50 | −1.8 | −9 | −2.773 |
| 147 | $(CH_3)_3C$ | 12.30 | 1.1 | −12 | −0.575 |
| 148 | $NO_2$ | 7.91 | −5.2 | −5 | −3.492 |
| MAD[b] | | ··· | 3.1[c] | 4.8[d] | **2.804**[e] |

[a]Experimental data taken from Refs. 22 and 48.
[b]Mean-absolute-deviation.
[c]Data taken from Ref. 22 obtained with Gaussian. Jaguar leads to MAD=3.14 kcal/mol.
[d]Data taken from Ref. 53 obtained with Gaussian. Jaguar leads to MAD=4.93 kcal/mol.
[e]Data obtained with Jaguar.

is under estimated by 6 kcal/mol, while $H_2$ is overbound in B3LYP and X3LYP by 1.0 and 0.4 kcal/mol, respectively. Concerning the importance of hydrogen in chemistry, it is unfortunate that PBE0 has a deviation of 6 kcal/mol for the heat of formation of $H_2O$, while B3LYP and X3LYP lead to errors of 1.3 and 1.6 kcal/mol, respectively.

For the subset of inorganic hydrides ($X_nH_m$, X=H, Li, N, O, F, Si, P, S, Cl; $n=1, 2$; $m=1-6$), MADs are 1.84 (B3LYP), 4.42 (PBE0), and 1.78 kcal/mol (X3LYP). The maximum errors occur at $N_2H_4$ for B3LYP (6.3)[22] and X3LYP (5.9); while the maximum error is 10 kcal/mol ($Si_2H_6$) for PBE0.[53]

The performance of PBE0 for larger hydrocarbons (Nos. 78–94 in Table II) is also less satisfactory. The MAD of this subset is 9.9 kcal/mol, with the maximum error of 24 kcal/mol for benzene. B3LYP performs much better. The MAD of this subset is 3.2 kcal/mol. The maximum error (7.1 kcal/mol) occurs at bicyclobutane. For benzene, B3LYP deviates from experiment by 4.5 kcal/mol. X3LYP is the best for this subset. The MAD of this subset is 2.2 kcal/mol, with maximum error of 5.8 kcal/mol at bicyclobutane. For benzene, X3LYP leads to deviation of 2.2 kcal/mol from the experimental result.

For a subset of substituted hydrocarbons (e.g., Nos. 95–136 in Table II), the performance of B3LYP, PBE0, and X3LYP are comparable except for the molecules from No. 133 to 136. PBE0 is particularly poor for furan, thiophene, pyrrole, and pyridine. The MADs of the substituted hydrocarbon subset are 2.11 (B3LYP[22]), 4.11 (PBE0[53]), and 1.85 (X3LYP). The maximum deviations are 9.2 and 7.9 at methylamine for B3LYP[22] and X3LYP, respectively, and 22 at pyridine for PBE0.[53]

For a subset of radicals (e.g., Nos. 138–148 in Table II), results from X3LYP and B3LYP are close, leading to MAD =3.00 and 2.89 with the maximum error being −8.2 and −7.9 kcal/mol at BeH for B3LYP and X3LYP, respectively. MAD (3.93) for PBE0 is larger with the maximum error of 12 kcal/mol at $(CH_3)_3C$.

Although X3LYP and B3LYP are generally more accurate than PBE0, there are cases where PBE0 is better. For example, X3LYP and B3LYP are poor for $SO_2$ (~10 kcal/mol error for both), $AlCl_3$ [10.2 (B3LYP); 9.1 (X3LYP)], and $SiCl_4$ [18.8 (B3LYP); 17.0 (X3LYP)]. Errors for PBE0 of these systems are significantly smaller, being 4 ($SO_2$), 2 ($AlCl_3$), and 3 kcal/mol ($SiCl_4$).

Table III presents a statistical evaluation of 18 different flavors of GGAs for the calculations of the heats of formation of the extended G2 set. From Table III it is clear that the MAD=90.9 kcal/mol for LDA (SVWN) is too high to be useful for thermochemistry.

GGAs greatly reduce the errors. OLYP leads to the smallest MAD (4.66 kcal/mol), being the best GGA up-to-date. The BLYP and BPW91 functionals give MAD=7.09 and 7.85, respectively. PWPW (MAD=17.8) and PWLYP (12.9) are less satisfactory, showing a larger tendency of overbinding. Thus for thermochemistry the PW91 exchange functional is poorer than B88 exchange functional and the PW91 correlation functional is poorer than the LYP correlation functional. The performance of PBE (MAD=17.1)[53] is very similar to PW91, unacceptable for thermochemistry. XLYP leads to MAD=7.56, similar to that of BLYP.

Table III shows that for thermochemistry the hybrid methods give an overall improvement compared to pure GGAs. Thus the performance of PBE is significantly improved going from pure PBE (MAD=17.1) to one-parameter hybrid PBE0 (4.8).[53] Keeping in mind that PBE and PBE0 are parameter-free, their overall performances are impressive. It is interesting to notice that the three-parameter hybrid PW3PW (MAD=10.3) is actually much worse than the one-parameter hybrid PW1PW (PW0) (MAD=5.2), lending support to the Perdew theoretical hybrid scheme.[38,53]

KMLYP was specially designed for activation barriers

TABLE III. Mean absolute deviations (MAD, theory-exptl.) for heats of formation at 298 K, in kcal/mol, obtained from various flavors of DFT methods against the extended G2 set (148 molecules).[a]

| | Max +[b] | Max −[c] | MAD |
|---|---|---|---|
| HF | +344.3 ($C_5H_5N$, pyridine) | −0.6 (BeH) | 149.2 |
| MP2 (full)[d] | +96.0 ($C_5H_5N$, pyridine) | −8.3 ($BF_3$) | 35.4 |
| LDA (SVWN)[e] | +0.4 ($Li_2$) | −228.7 ($C_6H_6$, benzene) | 90.9 |
| BLYP[e] | +24.8 ($SiCl_4$) | −28.4 ($NO_2$) | 7.09 |
| BPW91[e] | +15.7 ($SiF_4$) | −32.2 ($NO_2$) | 7.85 |
| BP86[e] | +6.3 ($SiF_4$) | −49.7 ($C_5H_5N$, pyridine) | 20.2 |
| PBE (PBEPBE)[f] | +11.0 ($Si_2H_6$) | −52.0 ($C_5H_5N$, pyridine) | 17.1 |
| PWPW (GGA II) | +6.4 ($Si_2H_6$) | −52.8 ($C_4F_4$) | 17.8 |
| mPWPW | +8.5 ($Si_2H_6$) | −47.3 ($C_2F_4$) | 15.1 |
| PWLYP | +12.7 ($SiCl_4$) | −39.0 ($NF_3$) | 12.9 |
| OLYP | +26.1 ($SiF_4$) | −22.0 ($NO_2$) | 4.66 |
| XLYP[g] | +28.1 ($SiCl_4$) | −25.5 ($NO_2$) | 7.56 |
| B3LYP[e] | +20.1 ($SiF_4$) | −8.2 (BeH) | 3.1 |
| B3PW91[e] | +21.8 ($SiF_4$) | −12.0 ($C_2F_4$) | 3.51 |
| B3P86[e] | +7.8 ($SiF_4$) | −49.2 ($C_5H_8$, spiropentane) | 18.0 |
| PBE0 (PBE1PBE)[f] | +15.0 ($SiF_4$) | −24.0 ($C_6H_6$, benzene) | 4.8 |
| PW3PW[h] | +10.1 ($SiF_4$) | −33.9 ($C_5H_5N$, pyridine) | 10.3 |
| PW1PW (PW0)[i] | +17.9 ($SiF_4$) | −20.5 ($C_6H_6$, benzene) | 5.24 |
| mPW1PW (mPW0)[j] | +20.7 ($SiF_4$) | −14.8 ($C_6H_6$, benzene) | 3.88 |
| PW3LYP[k] | +10.4 ($SiCl_4$) | −21.8 ($C_5H_5N$, pyridine) | 7.86 |
| KMLYP[l] | +43.6 ($O_3$, ozone) | −64.1 ($C_5H_8$, spiropentane) | 20.4 |
| O3LYP[m] | +25.9 ($SiF_4$) | −9.4 ($NO_2$) | 4.13 |
| X3LYP[n] | +18.5 ($SiF_4$) | −7.9 (BeH) | **2.80** |

[a]The basis sets used in all calculations are 6-311+G($3df$,$2p$). All geometries are optimized at MP2(Full)/6-31G* (Refs. 22, 23, 48, 57). Scaled HF/6-31G($d$) frequencies are used for zero-point energies and thermo-corrections (Refs. 22, 23, 48, and 57). The present calculations are performed with Jaguar (Ref. 54). Other data are taken from the corresponding literature.

[b]Maximum positive deviations.

[c]Maximum negative deviations.

[d]Data taken from Refs. 56 and 57. Basis set used is 6-31G*.

[e]Data taken from Ref. 22.

[f]Data taken from Ref. 53. PBE0 (also called PBE1PBE) is according to the formula: $0.25\,E_x(\mathrm{HF}) + 0.75\,E_x(\mathrm{Slater}) + 0.75\,\Delta E_x(\mathrm{PBE}) + 1.0\,E_c(\mathrm{PW91,local}) + 1.0\,\Delta E_c(\mathrm{PBE,nonlocal})$.

[g]$1.0\,E_x(\mathrm{Slater}) + 0.722\,\Delta E_x(\mathrm{B88}) + 0.347\,\Delta E_x(\mathrm{PW91}) + 1.0\,E_c(\mathrm{LYP})$.

[h]$0.20\,E_x(\mathrm{HF}) + 0.80\,E_x(\mathrm{Slater}) + 0.72\,\Delta E_x(\mathrm{PW91}) + 1.0\,E_c(\mathrm{PW91,local}) + 0.81\,\Delta E_c(\mathrm{PW91,nonlocal})$.

[i]$0.25\,E_x(\mathrm{HF}) + 0.75\,E_x(\mathrm{Slater}) + 0.75\,\Delta E_x(\mathrm{PW91}) + 1.0\,E_c(\mathrm{PW91,local}) + 1.0\,\Delta E_c(\mathrm{PW91,nonlocal})$.

[j]$0.25\,E_x(\mathrm{HF}) + 0.75\,E_x(\mathrm{Slater}) + 0.75\,\Delta E_x(\mathrm{mPW}) + 1.0\,E_c(\mathrm{PW91,local}) + 1.0\,\Delta E_c(\mathrm{PW91,nonlocal})$.

[k]$0.20\,E_x(\mathrm{HF}) + 0.80\,E_x(\mathrm{Slater}) + 0.72\,\Delta E_x(\mathrm{PW91}) + 0.19\,E_c(\mathrm{VWN}) + 0.81\,E_c(\mathrm{LYP})$.

[l]$0.557\,E_x(\mathrm{HF}) + 0.443\,E_x(\mathrm{Slater}) + 0.552\,E_c(\mathrm{VWN}) + 0.448\,E_c(\mathrm{LYP})$.

[m]$0.1161\,E_x(\mathrm{HF}) + 0.9262\,E_x(\mathrm{Slater}) + 0.8133\,\Delta E_x(\mathrm{OPTX}) + 0.19\,E_c(\mathrm{VWN5}) + 0.81\,E_c(\mathrm{LYP})$.

[n]$0.218\,E_x(\mathrm{HF}) + 0.782\,E_x(\mathrm{Slater}) + 0.542\,\Delta E_x(\mathrm{B88}) + 0.167\,\Delta E_x(\mathrm{PW91}) + 0.129\,E_c(\mathrm{VWN}) + 0.871\,E_c(\mathrm{LYP})$.

(kinetics) by omitting the GGA contribution to the exchange energy and emphasizing the role of exact exchange [the mixing coefficient for $E_x$(HF) is 0.557 in KMLYP vs. 0.20 in B3LYP].[55] KMLYP is reported to achieve activation barriers that are more accurate than B3LYP. We find that the MAD (20.4) of KMLYP is quite high for thermochemistry, however, these errors can be greatly reduced using ''high-level corrections'' in which energy corrections are included for each bond pair.[55]

O3LYP uses much less exact exchange (0.1161) than most hybrid functions, leading to similar MADs for O3LYP (4.1) and OLYP (4.7).

Overall, X3LYP (MAD=2.8), B3LYP (3.1[22]), B3PW91 (3.5[22]), mPW1PW (3.9), O3LYP (4.1), OLYP (4.7), and PBE0 (4.8[53]) show the best performance for thermochemistry.

G2 theory leads to a MAD of only 1.58 kcal/mol.[22,23,48,56,57] G2 theory is a composite, based on the 6-311G** basis set but with several basis set extensions. Electron correlation is treated by Moller–Plesset (MP) perturbation theory and by quadratic configuration interaction [QCISD(T)]. However we must emphasize that G2 theory is not *ab initio*. It includes an empirical ''high-level corrections'' for each covalent bond, assuming additivity. Removing these empirical corrections leads to much poorer thermochemistry. Thus based on the data in Refs. 56 and 57, we deduce that for the heats of formation of the first 56 molecules in Table II the MP4/6-311G** calculations lead to MAD=21.8 kcal/mol while QCISD(T)/6-311G** leads to MAD=16.8 kcal/mol. This can be compared to the results from DFT on the same systems with the same basis set (6-311G**): MAD=4.2 kcal/mol for B3LYP, 5.5 PBE0 and 4.9 X3LYP. Thus the current generation of DFT functionals lead to results significantly better than the standard *ab initio* methods, if empirical corrections are excluded from the *ab initio*. Since the empirical corrections in G2 theory are pointwise, there is no information on the forces corresponding to

TABLE IV. Mean absolute deviations (MAD, theory-exptl.) for ionization potentials at 0 K, in eV, obtained from various flavors of DFT methods against the G2 test set (42 systems).

| | Max +[a] | Max −[b] | MAD |
|---|---|---|---|
| SVWN[c] | +1.31 ($Ne \rightarrow Ne^+$) | −0.10 ($H \rightarrow H^+$) | 0.666 |
| BLYP[c] | +0.51 ($O \rightarrow O^+$) | −0.44($Cl_2 \rightarrow Cl_2^+$) | 0.183 |
| BPW91[c] | +0.44 ($O_2 \rightarrow O_2^+$) | −0.38($Be \rightarrow Be^+$) | 0.163 |
| BP86[c] | +1.07 ($O \rightarrow O^+$) | NA | 0.593 |
| PBE[d] | +0.46 ($O \rightarrow O^+$) | −0.34($Cl_2 \rightarrow Cl_2^+$) | 0.160 |
| PWPW | +0.48 ($O_2 \rightarrow O_2^+$) | −0.29($Cl_2 \rightarrow Cl_2^+$) | 0.164 |
| mPWPW | +0.48 ($O_2 \rightarrow O_2^+$) | −0.30($Cl_2 \rightarrow Cl_2^+$) | 0.163 |
| PWLYP | +0.57 ($O \rightarrow O^+$) | −0.39($Cl_2 \rightarrow Cl_2^+$) | 0.170 |
| OLYP | +0.91 ($C_2H_4 \rightarrow C_2H_4^+$) | −0.44($Cl_2 \rightarrow Cl_2^+$) | 0.185 |
| XLYP | +0.52 ($O \rightarrow O^+$) | −0.43($Cl_2 \rightarrow Cl_2^+$) | 0.179 |
| B3LYP[c] | +0.80 ($O_2 \rightarrow O_2^+$) | −0.20($Be \rightarrow Be^+$) | 0.163 |
| B3PW91[c] | +0.74 ($O_2 \rightarrow O_2^+$) | −0.32($Be \rightarrow Be^+$) | 0.163 |
| B3P86[c] | +1.29 ($O_2 \rightarrow O_2^+$) | NA | 0.638 |
| PBE0[d] | +0.69 ($O_2 \rightarrow O_2^+$) | −0.34($Be \rightarrow Be^+$) | 0.162 |
| PW3PW | +0.77 ($O_2 \rightarrow O_2^+$) | −0.23($Be \rightarrow Be^+$) | 0.166 |
| PW1PW | +0.70 ($O_2 \rightarrow O_2^+$) | −0.30($Be \rightarrow Be^+$) | 0.162 |
| mPW1PW | +0.77 ($O_2 \rightarrow O_2^+$) | −0.32($Be \rightarrow Be^+$) | 0.163 |
| PW3LYP | +0.84 ($O_2 \rightarrow O_2^+$) | −0.12($Be \rightarrow Be^+$) | 0.180 |
| KMLYP | +1.47 ($O_2 \rightarrow O_2^+$) | −0.04($Be \rightarrow Be^+$) | 0.376 |
| O3LYP | +0.58 ($SiH_4 \rightarrow SiH_4^+$) | −0.30 ($O_2 \rightarrow O_2^+$) | **0.139** |
| X3LYP | +0.78 ($O_2 \rightarrow O_2^+$) | −0.25 ($P_2 \rightarrow P_2^+$) | 0.154 |

[a]Maximum positive deviations.
[b]Maximum negative deviations.
[c]Data taken from Ref. 23.
[d]Data taken from Ref. 53.

this correction and hence one cannot include the corrections in the potential surface (barrier heights, etc).

### B. Ionization potentials (IPs)

Table IV and Table S2 list experimental IPs and theoretical deviations from experiment for the 18 atoms up to Ar and the 24 molecules in the G2 data set.[23,48] The MADs for the total 42 systems are 0.163 eV (B3LYP), 0.162 (PBE0), and 0.154 (X3LYP). The only case better than X3LYP is O3LYP with error of 0.139 eV.

Very accurate experimental IPs for atoms are known to provide a good test of the functionals for describing positively charged systems. For atomic systems, MADs for B3LYP,[23] PBE0[53] and X3LYP are 0.204, 0.151, and 0.178 eV, respectively. For molecular systems, MADs for B3LYP, PBE0, and X3LYP are 0.132, 0.172, 0.136 eV, respectively.

The IP of $O_2$ is a problem for all three functionals (errors of 0.80, 0.69, and 0.78 eV, respectively), possibly because the MP2 geometry for $O_2$ is very bad ($R_{OO}=1.246$ Å rather than 1.207 Å).

Generally, cations are more inhomogeneous than the neutral system. Thus it is not surprising that GGAs (except BP86) dramatically improve the predictions of IPs over LDA (SVWN) (MAD=0.67 eV, Table IV). However inclusion of exact exchange has little benefit.

### C. Electron affinities (EAs)

There has been some debate in the literature concerning whether DFT methods are suitable for calculating EAs.[53,58–60] The ''self-interaction error'' artificially shifts the Kohn–Sham orbital energies upwards, often leading to an unstable (positive) highest occupied orbital energy of an anion. On the other hand, use of finite basis sets with functions localized at the anion provide an artificial stabilization. In any case the numerical results demonstrate that DFT calculations predict EAs with an accuracy comparable to conventional *ab initio* calculations.[23,53,58]

Table V and Table S3 summarize the experimental EAs and the theoretical deviations from experiment for 7 atoms and 18 molecules.[23,48] Over these 25 systems the best performance is for X3LYP (MAD=0.087 eV), but B3LYP (0.11 eV) and PBE0 (0.13 eV) are comparable.

As expected, LDA (SVWN) overbinds (by MAD=0.75 eV) the extra electron (relative to the neutral system) and most GGAs (except BP86) remove most of this error, leading to MAD from 0.11 to 0.14 eV. Although HF exchange is self-interaction error free, inclusion of exact exchange leads to errors of 0.08–0.14 eV, indicating no improvement in the performance over the corresponding pure DFT methods [e.g., 0.107 (BLYP) versus 0.108 eV (B3LYP); 0.111 (PBE) versus 0.126 eV (PBE0)].

For atomic systems, the MADs for B3LYP, PBE0, and X3LYP are 0.106, 0.090, and 0.080 eV, respectively, but PBE0 performs significantly better for the second low atoms. For the molecular systems, B3LYP, PBE0, and X3LYP lead to MADs of 0.111, 0.146, 0.096 eV, respectively. The EA of $Cl_2$ is problematic for both B3LYP and X3LYP.

### D. Proton affinities (PAs)

Protonation makes the molecules more inhomogeneous. Thus it is anticipated that PAs may be systematically underestimated by LDA. Table VI shows that the MAD for the prediction of PA by LDA (SVWN) is ~6.3 kcal/mol with a maximum negative deviation of −10.1 kcal/mol. GGAs reduce the LDA errors effectively, although PA are still underestimated in PWLYP and PW3LYP as shown by the lack of positive deviations with these methods. B3PW91 and B3P86 show the best performance, with MADs being 0.73 and 0.71 kcal/mol, respectively.

Over these 8 systems the MADs are 1.6 (B3LYP),[23] 1.7 (X3LYP), and 2.4 kcal/mol (PBE0).[53] These error statistics are impressive, but the sample space (8 data) may be too small to draw a definitive conclusion.

### E. Total energies

Total energies for the first 10 atoms are summarized in Table VII. Comparing to the experimental values,[45,46] we see that LDA (SVWN) makes huge errors (MAD, 0.245 a.u. =6.67 eV=153.7 kcal/mol). GGAs remove a large part of this error. For the pure DFT methods, BLYP and BPW91 perform best (MAD 0.007 and 0.006 a.u., respectively); while BP86 and PBE behave worst (MAD 0.112 and 0.046 a.u., respectively). Inclusion of some exact exchange does not make hybrid DFT methods superior to the corresponding pure DFT methods. The MADs are 0.004 (X3LYP), 0.013 (B3LYP), 0.010 (B3PW91), 0.040 (PBE0), and 0.002 a.u. (O3LYP).

TABLE V. Mean absolute deviations (MAD, theory-exptl.) for electron affinities at 0 K, in eV, obtained from various flavors of DFT methods against the G2 test set (25 systems).

| | Max +[a] | Max −[b] | MAD |
|---|---|---|---|
| SVWN[c] | +1.20 ($F \leftarrow F^-$) | NA | 0.754 |
| BLYP[c] | +0.37 ($Cl_2 \leftarrow Cl_2^-$) | −0.16 ($Si \leftarrow Si^-$, $S_2 \leftarrow S_2^-$) | 0.107 |
| BPW91[c] | +0.31 ($C \leftarrow C^-$) | −0.11 ($S_2 \leftarrow S_2^-$) | 0.102 |
| BP86[c] | +0.84 ($C \leftarrow C^-$, $Cl_2 \leftarrow Cl_2^-$) | NA | 0.659 |
| PBE[d] | +0.29 ($C \leftarrow C^-$) | −0.12 ($S_2 \leftarrow S_2^-$) | 0.111 |
| PWPW | +0.37 ($C \leftarrow C^-$) | −0.05 ($S_2 \leftarrow S_2^-$) | 0.141 |
| mPWPW | +0.36 ($C \leftarrow C^-$) | −0.07 ($S_2 \leftarrow S_2^-$) | 0.130 |
| PWLYP | +0.39 ($Cl_2 \leftarrow Cl_2^-$) | −0.13 ($Si \leftarrow Si^-$) | 0.134 |
| OLYP | +0.12 ($Cl_2 \leftarrow Cl_2^-$) | −0.33 ($O_2 \leftarrow O_2^-$) | 0.133 |
| XLYP | +0.37 ($Cl_2 \leftarrow Cl_2^-$) | −0.13 ($Si \leftarrow Si^-$) | 0.112 |
| B3LYP[c] | +0.45 ($Cl_2 \leftarrow Cl_2^-$) | −0.06 ($OH \leftarrow OH^-$) | 0.108 |
| B3PW91[c] | +0.29 ($Cl_2 \leftarrow Cl_2^-$) | −0.17 ($OH \leftarrow OH^-$) | 0.101 |
| B3P86[c] | +0.84 ($Cl_2 \leftarrow Cl_2^-$) | NA | 0.601 |
| PBE0[d] | +0.22 ($PO \leftarrow PO^-$) | −0.28 ($OH \leftarrow OH^-$) | 0.126 |
| PW3PW | +0.31 ($Cl_2 \leftarrow Cl_2^-$) | −0.10 ($OH \leftarrow OH^-$) | 0.108 |
| PW1PW | +0.25 ($Cl_2 \leftarrow Cl_2^-$) | −0.24 ($OH \leftarrow OH^-$) | 0.117 |
| mPW1PW | +0.25 ($Cl_2 \leftarrow Cl_2^-$) | −0.26 ($OH \leftarrow OH^-$) | 0.120 |
| PW3LYP | +0.47 ($Cl_2 \leftarrow Cl_2^-$) | NA | 0.137 |
| KMLYP | +0.52 ($CN \leftarrow CN^-$) | −0.04 ($NH_2 \leftarrow NH_2^-$) | 0.221 |
| O3LYP | +0.20 ($Cl_2 \leftarrow Cl_2^-$) | −0.25 ($O_2 \leftarrow O_2^-$) | 0.107 |
| X3LYP | +0.40 ($Cl_2 \leftarrow Cl_2^-$) | −0.12 ($OH \leftarrow OH^-$) | **0.087** |

[a]Maximum positive deviations.
[b]Maximum negative deviations.
[c]Data taken from Ref. 23.
[d]Data taken from Ref. 53.

### F. Bonding properties of noble-gas dimers

For a neutral atom the effective potential seen by an electron far from atom should have the form $-1/r$, but none of the conventional density functionals have this form.[29–35] [The GGGA method[61] does lead to $-1/r$ but this has not been tested thoroughly.] With the wrong long-range potential, we cannot expect to have the correct long range density and hence we would expect problems getting the correct

TABLE VI. Mean absolute deviations (MAD, theory-exptl.) for proton affinities at 0 K, in kcal/mol, obtained from various flavors of DFT methods against the G2 test set (8 systems).

| | Max +[a] | Max −[b] | MAD |
|---|---|---|---|
| SVWN[c] | NA | −10.1 ($PH_3 \leftarrow PH_4^+$) | 6.32 |
| BLYP[c] | 0.55($C_2H_2 \leftarrow C_2H_3^+$) | −3.9 ($H_2O \leftarrow H_3O^+$) | 1.90 |
| BPW91[c] | 1.89($C_2H_2 \leftarrow C_2H_3^+$) | −1.59 ($PH_3 \leftarrow PH_4^+$) | 1.03 |
| BP86[c] | 0.95($C_2H_2 \leftarrow C_2H_3^+$) | −2.28 ($PH_3 \leftarrow PH_4^+$) | 0.84 |
| PBE | 3 ($C_2H_2 \leftarrow C_2H_3^+$) | −5 ($PH_3 \leftarrow PH_4^+$) | 2.7 |
| PWPW | 0.02 ($HCl \leftarrow H_2Cl^+$) | −3.77 ($PH_3 \leftarrow PH_4^+$) | 1.43 |
| mPWPW | 0.46 ($HCl \leftarrow H_2Cl^+$) | −3.20 ($PH_3 \leftarrow PH_4^+$) | 1.20 |
| PWLYP | NA | −5.31 ($H_2O \leftarrow H_3O^+$) | 3.49 |
| OLYP | 3.42($C_2H_2 \leftarrow C_2H_3^+$) | −0.74 ($PH_3 \leftarrow PH_4^+$) | 1.38 |
| XLYP | 0.12($C_2H_2 \leftarrow C_2H_3^+$) | −4.19 ($H_2O \leftarrow H_3O^+$) | 2.17 |
| B3LYP[c] | 1.28($C_2H_2 \leftarrow C_2H_3^+$) | −2.80 ($H_2 \leftarrow H_3^+$) | 1.63 |
| B3PW91[c] | 2.12($C_2H_2 \leftarrow C_2H_3^+$) | −0.45 ($H_2 \leftarrow H_3^+$) | 0.73 |
| B3P86[c] | 1.48($C_2H_2 \leftarrow C_2H_3^+$) | −0.74 ($H_2 \leftarrow H_3^+$) | **0.71** |
| PBE0[d] | 5 ($C_2H_2 \leftarrow C_2H_3^+$) | −3 ($PH_3 \leftarrow PH_4^+$) | 2.4 |
| PW3PW | 0.75($C_2H_2 \leftarrow C_2H_3^+$) | −1.82 ($PH_3 \leftarrow PH_4^+$) | 1.08 |
| PW1PW | 1.12 ($NH_3 \leftarrow NH_4^+$) | −1.37 ($SiH_4 \leftarrow SiH_5^+$) | 0.98 |
| mPWPW | 1.70($C_2H_2 \leftarrow C_2H_3^+$) | −1.02 ($SiH_4 \leftarrow SiH_5^+$) | 0.86 |
| PW3LYP | NA | −3.80 ($H_2 \leftarrow H_3^+$) | 2.29 |
| KMLYP | 0.88 ($NH_3 \leftarrow NH_4^+$) | −3.05 ($H_2 \leftarrow H_3^+$) | 1.93 |
| O3LYP | 3.25($C_2H_2 \leftarrow C_2H_3^+$) | −0.59 ($H_2 \leftarrow H_3^+$) | 1.13 |
| X3LYP | 0.81($C_2H_2 \leftarrow C_2H_3^+$) | −3.20 ($H_2 \leftarrow H_3^+$) | 1.71 |

[a]Maximum positive deviations.
[b]Maximum negative deviations.
[c]Data taken from Ref. 23.
[d]Data taken from Ref. 53.

TABLE VII. Total energies (Hartree) and deviations (theory-exptl.) for 10 atoms.

| | System | Exp.[a] | LDA(SVWN) | BLYP | BPW91 | BP86 | PBE | PWPW |
|---|---|---|---|---|---|---|---|---|
| 1 | H | −0.500 | 0.004 | 0.002 | −0.004 | −0.018 | 0.000 | −0.001 |
| 2 | He | −2.904 | 0.034 | −0.001 | −0.002 | −0.038 | 0.013 | 0.006 |
| 3 | Li | −7.478 | 0.081 | −0.002 | −0.006 | −0.057 | 0.018 | 0.006 |
| 4 | Be | −14.667 | 0.149 | 0.008 | 0.009 | −0.064 | 0.040 | 0.022 |
| 5 | B | −24.654 | 0.209 | 0.005 | 0.007 | −0.089 | 0.047 | 0.022 |
| 6 | C | −37.845 | 0.269 | 0.001 | 0.002 | −0.115 | 0.052 | 0.019 |
| 7 | N | −54.589 | 0.328 | 0.003 | −0.001 | −0.140 | 0.060 | 0.018 |
| 8 | O | −75.067 | 0.394 | −0.011 | −0.008 | −0.173 | 0.065 | 0.012 |
| 9 | F | −99.734 | 0.459 | −0.017 | −0.012 | −0.202 | 0.075 | 0.010 |
| 10 | Ne | −128.938 | 0.525 | −0.014 | −0.009 | −0.224 | 0.092 | 0.013 |
| | MAD | 0.000 | 0.245 | 0.007 | 0.006 | 0.112 | 0.046 | 0.013 |
| | System | X3LYP | KMLYP | B3LYP | B3PW91 | B3P86 | PBE0 | O3LYP |
| 1 | H | 0.000 | −0.002 | −0.002 | −0.001 | −0.019 | −0.001 | −0.000 |
| 2 | He | −0.002 | 0.004 | −0.009 | 0.006 | −0.040 | 0.011 | 0.006 |
| 3 | Li | −0.004 | 0.016 | −0.013 | 0.006 | −0.058 | 0.012 | 0.009 |
| 4 | Be | 0.008 | 0.045 | −0.004 | 0.022 | −0.063 | 0.032 | 0.000 |
| 5 | B | 0.006 | 0.066 | −0.007 | 0.022 | −0.083 | 0.039 | 0.001 |
| 6 | C | 0.005 | 0.086 | −0.011 | 0.019 | −0.106 | 0.043 | 0.000 |
| 7 | N | 0.009 | 0.106 | −0.012 | 0.018 | −0.127 | 0.049 | 0.001 |
| 8 | O | 0.000 | 0.132 | −0.021 | 0.012 | −0.153 | 0.057 | 0.002 |
| 9 | F | 0.000 | 0.160 | −0.026 | 0.010 | −0.176 | 0.069 | 0.000 |
| 10 | Ne | 0.008 | 0.189 | −0.023 | 0.013 | −0.194 | 0.086 | 0.003 |
| | MAD | 0.004 | 0.081 | 0.013 | 0.010 | 0.102 | 0.040 | **0.002** |

[a]Experimental data taken from Refs. 45 and 46.

long-range dispersion interactions (particularly near the minimum for the noble-gas dimers). We find below that X3LYP and mPWPW lead to fairly good descriptions for $He_2$ and $Ne_2$; however, this does not mean that the underlying problem of DFT has been solved. Indeed these methods do not do well at describing larger noble-gas dimers such as $Xe_2$, and even $Ar_2$. On the other hand, for the biological and other organic materials in which we are interested, the nonbonded contacts are dominated by H, C, N, and O, which have dispersion interactions similar to He and Ne. Thus we focused on these two cases to ensure a good description of biological systems.

Noble-gas dimers are the least ambiguous test molecules for determining how well the van der Waals attraction (London dispersion) is described. Table VIII summarizes the bonding properties of $He_2$, $Ne_2$, and $Ar_2$ calculated by different flavors of DFT functionals.

Although the B88 exchange functional has been very successful in describing the thermochemistry of covalent systems, it fails completely to describe van der Waals interactions. As shown in Table VIII, every DFT methods using B88 as exchange functional, pure or hybrid, gives unbounded noble-gas dimers.

On the other hand, Table VIII shows that the PW91 exchange functional severely overbinds noble-gas dimers. Adamo and Barone modified PW91[16] by fitting the differential exchange energies of noble-gas dimers to HF values, removing most of the overbinding tendency of PW91. This mPWPW model yields $r_e(\text{He–He}) = 3.14\,\text{Å}$ and $D_e(\text{He–He}) = 0.069\,\text{kcal/mol}$,[16] as compared the PWPW values of $r_e(\text{He–He}) = 2.645\,\text{Å}$ and $D_e(\text{He–He}) = 0.231$ kcal/mol[16] and the experimental values of $r_e(\text{He–He}) = 2.970\,\text{Å}$, $D_e(\text{He–He}) = 0.022\,\text{kcal/mol}$.[50] The PBE functional gives a good description of noble-gas dimers. For $He_2$, PBE0 yields $r_e = 2.818\,\text{Å}$ and $D_e = 0.042\,\text{kcal/mol}$, although, critically,[44] PBE0 still overestimates $D_e$ by 91%.

The van der Waals attraction between noble-gas atoms is entirely due to electron correlation, originating from the interactions between instantaneous fluctuating dipoles as shown by London. Thus when the correlation functional is eliminated to obtain the exchange-only potential, the noble gas dimers should lead to totally repulsive interactions for all interatomic distances, similar to the HF potential. However, the PW91, mPW, PBE, and corresponding hybrid models, without correlation all lead to a bound state, indicating that some electron correlation is implicitly included in the exchange-only potentials.

In this context, we conclude that X3LYP outperforms all the other functionals listed in Table VIII. For $He_2$, X3LYP yields $r_e = 2.726\,\text{Å}$ and $D_e = 0.021\,\text{kcal/mol}$; while the exchange-only (X3) potential is repulsive.

Further improvement on the correlation functional is needed to describe correctly the van der Waals attraction.

### G. Bonding properties of water dimer

Hydrogen bonding plays a critical role in a wide range of chemical and biological phenomena. Consequently water dimer, a prototypical hydrogen bonded system, has received much experimental and theoretical attention.[62–69] The equilibrium geometry and dissociation energy of $(H_2O)_2$ are now known quite accurately: $r_e(\text{O}\cdots\text{O}) = 2.912 \pm 0.005\,\text{Å}$ and $D_e = 5.02 \pm 0.10\,\text{kcal/mol}$.[65] These results come from a high-level *ab initio* theory [coupled cluster including single and double excitations plus triples (CCSD(T)) (full)] using basis sets that are extrapolated to infinity.[65] Accurate experimental

TABLE VIII. Bonding properties of $He_2$, $Ne_2$, and $Ar_2$.[a] Bond lengths are in Å and bond energies are in kcal/mol.

| | $He_2$ | | $Ne_2$ | | $Ar_2$ | |
|---|---|---|---|---|---|---|
| | $R_e$ | $\Delta E$ | $R_e$ | $\Delta E$ | $R_e$ | $\Delta E$ |
| HF | unbounded | | unbounded | | unbounded | |
| SVWN | 2.377 | 0.251 | 2.595 | 0.533 | 3.379 | 0.787 |
| BLYP | unbounded | | unbounded | | unbounded | |
| BPW91 | unbounded | | unbounded | | unbounded | |
| mPWPW[b] | 3.14 | 0.069 | 3.25 | **0.092** | 4.45 | 0.115 |
| PBE | 2.752 | 0.073 | 3.097 | 0.111 | 4.000 | 0.126 |
| PWPW | 2.645 | 0.231 | 3.016 | 0.316 | 3.954 | 0.295 |
| PWLYP | 2.400 | 0.510 | 2.753 | 0.751 | 3.728 | 0.560 |
| OLYP | 2.887 | 0.079 | 3.283 | 0.123 | 4.836 | 0.050 |
| X (Ex-only) | unbounded | | unbounded | | unbounded | |
| XLYP | 2.805 | **0.023** | 3.126 | 0.069 | 4.384 | 0.020 |
| B3LYP | unbounded | | unbounded | | unbounded | |
| B3PW91 | unbounded | | unbounded | | unbounded | |
| mPW0 (Ex-only) | 3.105 | 0.039 | 3.467 | 0.051 | 4.724 | 0.029 |
| mPW0(mPW1PW) | **3.052** | 0.045 | 3.254 | 0.053 | 4.435 | 0.036 |
| mPW0[b] | 3.11 | 0.046 | 3.23 | 0.069 | 4.42 | 0.069 |
| PBE0 (Ex-only) | 3.016 | 0.032 | 3.161 | 0.030 | 4.338 | 0.035 |
| PBE0(PBE1PBE) | 2.818 | 0.042 | 3.118 | 0.061 | 4.040 | 0.081 |
| PW3PW | 2.660 | 0.164 | **3.003** | 0.221 | 3.943 | **0.225** |
| PW3LYP | 2.420 | 0.379 | 2.750 | 0.566 | **3.722** | 0.449 |
| KMLYP | 2.448 | 0.140 | 2.671 | 0.303 | 3.584 | 0.332 |
| O3LYP | 2.860 | 0.072 | 3.225 | 0.109 | 4.473 | 0.032 |
| X3 (Ex-only) | unbounded | | unbounded | | unbounded | |
| X3LYP | 2.726 | **0.021** | 2.904 | 0.063 | 4.234 | 0.007 |
| exptl.[c] | 2.970 | 0.022 | 3.091 | 0.084 | 3.757 | 0.285 |

[a]All calculations are performed with aug-cc-pVTZ. Bond energies are corrected for basis set superposition error (BSSE) effects.
[b]Reference 38. Basis sets are modified cc-pV5X.
[c]Reference 50.

determination of $r_e$ and $D_e$ have proven to be an elusive goal. Microwave measurements lead to a vibrationally averaged O··O distance $R_0 = 2.976$ Å, from which it was estimated that $R_e = 2.946$ Å.[68] The widely accepted experimental $D_e = 5.4 \pm 0.7$ kcal/mol[69] was based on measurements of the thermal conductivity of the water vapor and involved complex interpretations. We conclude that the *ab initio* values are the most reliable.

Table IX lists the calculated bonding properties of $(H_2O)_2$ using a variety of DFT methods. The most accurate overall description is from X3LYP, which leads to a bond distance just 0.004 Å from the exact value (and within the error bars) and a bond energy within 0.05 kcal/mol of the exact value (and within the error bars), predicting bonding properties of $(H_2O)_2$ with better quality than the other functionals listed in Table IX.

Other DFT methods with either $R_e$ or $D_e$ within experimental error of the exact result are PBE0 for $D_e$ and mPWPW for $R_e$.

BLYP gives $r_e = 2.952$ Å, which is 0.04 Å too long, and $D_e$ too weak by 0.84 kcal/mol (16%), indicating too weak a hydrogen bond. B3LYP leads to some improvement but still underestimates hydrogen bonds, leading to $r_e$ too long by 0.014 Å and $D_e$ too weak by 0.45 kcal/mol (9%).

In contrast PWPW overestimates hydrogen bonds, leading to $r_e$ too short by 0.026 Å and $D_e$ too strong by 0.41 kcal/mol. The modified PW91 functional (mPWPW) improves $R_e$ to within 0.001 Å, but over-corrects the overbinding, leading to a bond too weak by 0.54 kcal/mol.

Although OLYP is very promising for thermochemistry, it is not good for hydrogen bonding, leading to $R_e$(O··O) too long by 0.263 Å and $D_e$ too weak by 2.26 kcal/mol, indicating that hydrogen bonds are significantly underestimated by this functional. O3LYP improves slightly from OLYP, but $R_e$(O··O) too long by 0.183 Å, with $D_e$ too weak by 1.82 kcal/mol.

## V. CONCLUDING REMARKS

Development of improved approximations to the exchange-correlation functional has been critical to the success of Kohn–Sham density functional theory, with several exchange-correlation functionals that do quite well for particular properties. This success has been achieved either by construction of functional forms to satisfy physical constraints or by fitting a few scale parameters to experimental data. We have combined these two approaches to obtain the X3LYP exchange-correlation functional whose form matches well the behavior of a Gaussian-type decaying density. To aid those that would like to test X3LYP, we express the $F^{X}$ GGA as a linear combination of $F^{B88}$ and $F^{PW91}$. The four mixing coefficients in X3LYP were determined by fitting to the atomization energies of a set of 33 diatomic and 5 triatomic molecules involving single, double, and triple bonds.

TABLE IX. Bonding properties of water dimer.[a] Bond lengths are in Å and bond energies are in kcal/mol.

| | $R_e(O\cdots O)$ | $D_e$ |
|---|---|---|
| BLYP | 2.952 | 4.18 |
| PWPW | 2.886 | 5.43 |
| mPWPW | **2.911** | 4.48 |
| PBE | 2.899 | **5.11** |
| OLYP | 3.175 | 2.76 |
| XLYP | 2.953 | 4.43 |
| B3LYP | 2.926 | 4.57 |
| PW0 (PW1PW) | 2.884 | 5.23 |
| mPW0 (mPW1PW) | 2.898 | 4.60 |
| PBE0 (PBE1PBE) | 2.896 | **4.98** |
| O3LYP | 3.095 | 3.20 |
| X3LYP | **2.908** | **4.97** |
| Best *ab initio*[b] | **2.912±0.005** | **5.02±0.10** |
| exptl. | 2.948[c] | 5.44±0.7[d] |

[a]Our calculated $D_e$ are BSSE-corrected. In bold face are the results within the uncertainty of the most accurate values [*ab initio* CCSD(T) full calculations with the basis set extrapolated to infinity].
[b]Reference 65. CCSD(T)(FULL)/IO275→∞ (IO275: interaction optimized basis set with 275 basis functions for $H_2O$ dimer. O: $7s5p5d3f2g1h$; $H_d$: $2s4p1d$, H: $2s3p$, BF: $3s3p2d1f$).
[c]Geometric parameters are taken from Ref. 67.
[d]Experimental $D_e$ was estimated by adding the zero-point energy calculated at HF/4-21G level in Ref. 69.

In addition we included $He_2$, $Ne_2$, and $Ar_2$ to test accuracy of van der Waals interactions, and here we demanded that exclusion of the LYP correlation function would lead to a repulsive potential curve, similar to HF. Total energies of the first 10 atoms as well as ionization potentials and electron affinities of the first and second row atoms were also included in the fitting set.

The accuracy of X3LYP was validated by testing against experimental data for the extended G2 set, which contains 148 standard heats of formation, 42 ionization potentials, 25 electron affinities, 8 proton affinities, and 10 total atomic energies for H through Ne. In addition we tested X3LYP for noble gas and water dimers. Among all the DFT functionals tested here, the mean absolute deviations achieved by X3LYP are:

(i) Heats of formation: 2.804 kcal/mol (best for all DFT);
(ii) ionization potential: 0.154 eV (2nd best of all DFT, best is O3LYP with 0.139 error);
(iii) electron affinities: 0.087 eV (best for all DFT);
(iv) proton affinities: 1.714 kcal/mol (best is B3P86 with 0.71 error);
(v) total atomic energies: 0.004 a.u. (2nd best of all DFT, best is O3LYP with 0.002 error);
(vi) $He_2$ bond energy: error of 0.001 kcal/mol or 5% (tied with XLYP for the best of all DFT);
(vii) $He_2$ bond distance: error of 0.244 Å or 8% (best is mPW0 with an error of 0.082 Å but a bond energy too large by 0.023 or 110%);
(viii) $Ne_2$ bond energy: error of 0.021 kcal/mol or 25% (best is mPWPW with an error of 0.008);
(ix) $Ne_2$ bond distance: error of 0.187 Å or 5% (the best is PW3PW with an error of 0.088 Å);
(x) $H_2O$ dimer bond energy: error of 0.005 kcal/mol or 0.1% (2nd best of all DFT, best is PBE0 with 0.004 error);
(xi) $H_2O$ dimer bond distance: error of 0.004 Å or 0.1% (2nd best of all DFT, best is mPWPW with 0.001 Å error).

Thus X3LYP is the most accurate DFT for most properties and is competitive with the best DFT for most other properties, making it the most consistent overall. In particular the accuracy for van der Waals and hydrogen bond interactions should make X3LYP useful for applications over a wide range of important chemical and biological systems.

Supplementary material available (Ref. 70):

Table S1 Deviations (theory-exptl.) from experiment for the heats of formation (kcal/mol at 298 K) for the extended G2 set (148 molecules), calculated by PWPW, PWLYP, PW3PW, PW1PW, and PW3LYP in Jaguar.

Table S2 Ionization potentials (in eV) at 0 K of 42 systems of G2 set and the deviations (theory-exptl.) obtained from B3LYP, PBE0, and X3LYP.

Table S3 Electron affinities (in eV) at 0 K of 25 systems of G2 set and the deviations (theory-exptl.) obtained from B3LYP, PBE0, and X3LYP.

Table S4 Proton affinities (in kcal/mol) at 0 K of 8 systems of G2 set and the deviations (theory-exptl.) obtained from B3LYP, PBE0, and X3LYP.

## ACKNOWLEDGMENTS

We thank Dr. Y. X. Cao, Dr. Dale Braden, and Dr. Jason Perry at Schrödinger Inc. for technical support of using and programming with Jaguar. This research was funded by DOE (ASCI), DARPA-CMDF, DARPA-PROM, National Institutes of Health (HD 36385-02), National Natural Science Foundation of China (20021002), National Natural Science Foundation of Fujian (2002F010), the Ministry of Science and Technology of China (2001CB610506) and TRAPOYT from the Ministry of Education of China. The facilities of the Materials and Process Simulation Center (MSC) used in these studies were funded by DURIP-ARO, DURIP-ONR, IBM (SUR), NSF (MRI), and the Beckman Institute. In addition, the MSC is funded by grants from DOE-FETL, ARO-MURI, ONR-MURI, NIH, ChevronTexaco, Aventis Pharma, General Motors, Seiko-Epson, Berlex Biopharma, and Asahi Kasei.

[1] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
[2] P. A. M. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).
[3] J. C. Slater, *Quantum Theory of Molecules and Solids*, Vol. 4 (McGraw-Hill, New York, 1974).
[4] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
[5] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
[6] W. Kohn, A. D. Becke, and R. G. Parr, J. Phys. Chem. **100**, 12974 (1996).
[7] J. P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).
[8] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[9] J. P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991), p. 11.
[10] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[11] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[12] B. Miehlich, A. Savin, H. Stoll, and H. Preuss, Chem. Phys. Lett. **157**, 200 (1989).
[13] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
[14] M. Filatov and W. Thiel, Mol. Phys. **91**, 847 (1997).
[15] P. M. W. Gill, Mol. Phys. **89**, 433 (1996).
[16] C. Adamo and V. Barone, J. Chem. Phys. **108**, 664 (1998).
[17] W.-M. Hoe, A. J. Cohen, and N. C. Handy, Chem. Phys. Lett. **341**, 319 (2001).
[18] E. Fermi and G. Amaldi, Men. R. Acad. Italia. **6**, 117 (1934).
[19] S. B. Liu and R. G. Parr, J. Comput. Chem. **20**, 2 (1999).
[20] G. J. Laming, V. Termath, and N. C. Handy, J. Chem. Phys. **99**, 8765 (1993).
[21] B. G. Johnson, P. M. W. Gill, and J. A. Pople, J. Chem. Phys. **98**, 5612 (1993).
[22] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **106**, 1063 (1997).
[23] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **109**, 42 (1998).
[24] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[25] A. D. Becke, J. Chem. Phys. **107**, 8554 (1997).
[26] H. L. Schmider and A. D. Becke, J. Chem. Phys. **108**, 9624 (1998).
[27] P. J. Stephens, F. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).
[28] C. Tuma, A. D. Boese, and N. C. Handy, Phys. Chem. Chem. Phys. **1**, 3939 (1999).
[29] J. M. Pérez-Jordá and A. D. Becke, Chem. Phys. Lett. **233**, 134 (1995).
[30] S. Kristyan and P. Pulay, Chem. Phys. Lett. **229**, 175 (1994).
[31] T. A. Wesolowski, O. Parisel, Y. Ellinger, and J. Weber, J. Phys. Chem. A **101**, 7818 (1997).
[32] Y. Andersson, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **76**, 102 (1996).
[33] J. F. Dobson and B. D. Dinte, Chem. Phys. Lett. **76**, 1780 (1996).
[34] W. Kohn, Y. Meir, and D. E. Makarov, Phys. Rev. Lett. **80**, 4153 (1998).
[35] Q. Wu and W. Yang, J. Chem. Phys. **116**, 515 (2002).
[36] Y. Zhang, W. Pan, and W. Yang, J. Chem. Phys. **107**, 7921 (1997).
[37] D. C. Patton and M. R. Pederson, Phys. Rev. A **56**, R2495 (1997).
[38] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).
[39] A. J. Cohen and N. C. Handy, Chem. Phys. Lett. **316**, 160 (2000).
[40] N. C. Handy and A. J. Cohen, Mol. Phys. **99**, 403 (2001).
[41] D. J. Lacks and R. G. Gordon, Phys. Rev. A **47**, 4681 (1993).
[42] M. Levy and J. P. Perdew, Phys. Rev. B **48**, 11638 (1993).
[43] E. H. Lieb and S. Oxford, Int. J. Quantum Chem. **19**, 427 (1981).
[44] T. A. Wesolowski, J. Chem. Phys. **113**, 1666 (2000).
[45] E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, and C. F. Fischer, Phys. Rev. A **44**, 7071 (1991).
[46] S. J. Chakravorty, S. R. Gwaltney, and E. R. Davidson, Phys. Rev. A **47**, 3649 (1993).
[47] D. Feller and K. A. Peterson, J. Chem. Phys. **108**, 154 (1997).
[48] L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991).
[49] V. E. Bondybey and J. H. English, J. Chem. Phys. **80**, 568 (1984).
[50] J. F. Ogilvie and F. Y. H. Wang, J. Mol. Struct. **273**, 277 (1992).
[51] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[52] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[53] M. Ernzerhof and G. E. Scuseria, J. Chem. Phys. **110**, 5029 (1999).
[54] JAGUAR 4.0 Schrödinger, Inc., Portland, Oregon, 2000.
[55] J. K. Kang and C. B. Musgrave, J. Chem. Phys. **115**, 11040 (2001).
[56] J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, and L. A. Curtiss, J. Chem. Phys. **90**, 5622 (1991).
[57] L. A. Curtiss, C. Jones, G. W. Trucks, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **93**, 2537 (1990).
[58] J. M. Galbraith and H. F. Schaefer, J. Chem. Phys. **105**, 862 (1996).
[59] N. Rösch and S. B. Trickey, J. Chem. Phys. **106**, 8940 (1997).
[60] T. V. Mourik and R. J. Gdanitz, J. Chem. Phys. **116**, 9620 (2002).
[61] X. Hua, X. Chen, and W. A. Goddard III, Phys. Rev. **55**, 16103 (1997).
[62] K. Kim and K. D. Jordan, J. Phys. Chem. **98**, 10089 (1994).
[63] D. A. Estrin, L. Daglieri, G. Corongiu, and E. Clementi, J. Phys. Chem. **100**, 8701 (1996).
[64] M. W. Feyereisen, D. Feller, and D. A. Dixon, J. Phys. Chem. **100**, 2993 (1996).
[65] W. Klopper, J. G. C. M. van Duijneveldt-vande Rijdt, and F. B. van Duijneveldt, Phys. Chem. Chem. Phys. **2**, 227 (2000).
[66] T. R. Dyke, K. M. Mack, and J. S. Muenter, J. Chem. Phys. **66**, 498 (1977).
[67] J. A. Odutola and T. R. Dyke, J. Chem. Phys. **72**, 5062 (1980).
[68] K. Kuchitsu and Y. Morino, Bull. Chem. Soc. Jpn. **38**, 805 (1965).
[69] L. A. Curtiss, D. J. Frurip, and M. Blander, J. Chem. Phys. **71**, 2703 (1979).
[70] See EPAPS Document No. E-JCPSA6-121-313445 for Tables S1¡S4. A direct link to this document may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory /epaps/. See the EPAPS homepage for more information.