13706 *J. Phys. Chem. A* **2008,** *112,* 13706–13711

Enhanced Enthalpies of Formation from Density Functional Theory through Molecular Reference States

Edward N. Brothers* and Gustavo E. Scuseria

Department of Chemistry, Mail Stop 60, Rice University, Houston, Texas 77005-1892

Received: July 12, 2008; Revised Manuscript Received: September 19, 2008

Accurate prediction of enthalpy of formation is an important goal for theoretical methods. As such, the ability of a density functional to accurately predict enthalpies of formation for a wide variety of compounds is often used as a critical test of its efficacy. These enthalpies are typically calculated by modeling formation reactions from the isolated atoms that make up the molecule. However, the enthalpy of formation can be calculated from any valid reference state, e.g., as in isodesmic reactions, and using different reference states can alter the accuracy of prediction. We have had excellent results using a single molecular reference state per element, namely C_{60} for carbon and the diatomic standard reference states for hydrogen, nitrogen, oxygen, and fluorine. This molecular reference scheme can be viewed as a better measure of the upper limit of accuracy of a density functional/basis set pair, as it leads to generally more accurate predictions than are possible using atomic energies. For example, LSDA's unsigned average error drops from 158.8 to 11.6 kcal/mol, and PBE's error improves to 5.1 kcal/mol from 35.8 kcal/mol with the 6-311G(2df,2p) basis set. This scheme also makes small basis sets far more accurate, indicates that a revision of the relative thermochemical accuracy of functionals may be required, and can remove qualitative failures for some functional/basis set pairs.

The proposal of a new density functional[1,2] typically includes (as part of the standard validation) predicting the enthalpies of formation (ΔH_f) of some set of compounds, such as the Gn/XX sets of Curtiss and co-workers[3−6] or the various sets proposed by Truhlar and co-workers.[7,8] This test can be viewed in some sense as a pass/fail test, i.e., proposed functionals with ΔH_f errors above a certain level are considered unsuccessful. While ΔH_f is not the only quantity of interest, it does usually make up the majority of data points a functional's accuracy is assessed against. For example, the popular G3/99 set[5] has 223 ΔH_f, 88 ionization potentials, 58 electron affinities, and 8 proton affinities, i.e. ΔH_f prediction is 60% of the evaluation (if all energies in the set are given equal weights). Using these benchmark sets, statements can be made that functional "A" is better than functional "B" based on their respective performance, and thus it is imperative to adequately determine the limits of the accuracy of both functionals. This paper describes what we deem to be a better way to assess this (in the sense that it more accurately reflects the limits of prediction) by utilizing a nonempirical, nonatomic scheme for evaluating ΔH_f. It should be noted in passing that, in addition to being a good way to calculate ΔH_f in an assessment of a functional, this procedure is also a practical method for assessing the ΔH_f of compounds of specific interest, such as (for example) in determining whether a reaction is exothermic or endothermic by calculating ΔH_f on both sides of the reaction.

Before discussing the method, some background is required. The history of evaluating molecular thermochemistry with density functionals as a measure of functional accuracy goes back to Becke,[9] who tested the accuracy of his 1988 functional[10] in reproducing the atomization energies of the G2 test set.[3] This set of compounds was originally developed to evaluate the G2 model chemistry.[3] The G2 model chemistry was subsequently used to investigate ΔH_f of a series of silicon molecules.[11] Later Gn/XX test sets used ΔH_f in place of atomization energies, until the most recent benchmark set (G3/05)[6] which uses a mixture of atomization energies and ΔH_f. We choose to use ΔH_f for the reasons given by Staroverov and co-workers,[12] namely the coupled rationale that ΔH_f's are observable quantities and more readily available than atomization energies. However, the procedure described in this paper is readily adaptable to use with atomization energies; our selection is simply a matter of choice.

The ΔH_f of a compound is formally the enthalpy of reaction for forming a compound from its constituent elements at their standard states, where the standard states have ΔH_f defined as zero. For example, the ΔH_f of fluoromethane is from the reaction

$$\frac{3}{2}H_2^{gas} + \frac{1}{2}F_2^{gas} + C^{graphite} \rightarrow CH_3F \qquad (1)$$

However, reactions of this type are often computationally intractable, due to the difficulty of performing calculations on the standard states of many elements, such as graphite in the reaction above or other elements whose standard states are complicated solids, such as sulfur. Most commonly, ΔH_f is actually calculated by modeling the formation reaction of a compound of interest from its constituent atoms in the gas phase, where the ΔH_f of the atoms are experimentally known. (As in all reactions for calculating ΔH_f, the ΔH_f of the reactants are known from experimental sources, and the quantity that is theoretically modeled is the enthalpy of reaction.) Again using fluoromethane as an example, the reaction modeled is most commonly

$$3H^{gas} + F^{gas} + C^{gas} \rightarrow CH_3F \qquad (2)$$

While this type of reaction is computationally the most efficient due to the minimal size of the reactants, it requires that the method in question be equally accurate for atoms and molecules,

---

* To whom correspondence should be addressed. E-mail: ed.brothers@qatar.tamu.edu.

10.1021/jp8061652 CCC: $40.75

which is known to not be the case for density functional theory (DFT),[13−15] at least for functionals currently in use. Thus while this is the way the ΔH_f which are quoted in the literature are often generated, this is not a perfect measurement of the ability of a functional to predict ΔH_f, as using different reference states alters (and in many cases increases) the accuracy of prediction.

There have been a number of methods that been used previously as a way to avoid the problems associated with using atomic energies in predicting ΔH_f. One technique is the use of isodesmic reactions, where the reactants and products of the reaction have the same number and type of bonds, e.g.,

$$CH_4 + CH_2F_2 \rightarrow 2CH_3F \qquad (3)$$

Isodesmic reactions do not require atomic energies, which can enhance their accuracy for DFT, but they have three significant drawbacks. First, for large and/or complicated molecules, the reaction scheme can become very complex. Second, the reaction requires a choice of reactants, which fundamentally introduces ambiguity, as a variety of different reactant sets can be used resulting in varying experimental accuracies for the ΔH_f of the reactants. (While there is always some choice made in determining reference states, the number of reference states that would be required for an evaluation of a large test set using only isodesmic reactions would be unwieldy.) Third, isodesmic reactions preclude the investigation of novel bond structures. Thus while isodesmic reactions can enhance accuracy of prediction versus the atomic formation reaction, they are necessarily impractical for large scale testing.

Another scheme that can be used is replacing the atomic energies with parameters, which removes the need for atomic energies and (due to parameter optimization) necessarily enhances the accuracy of prediction.[13,14] In this case, the example reaction becomes

$$3H^{\text{parameter}} + F^{\text{parameter}} + C^{\text{parameter}} \rightarrow CH_3F \qquad (4)$$

This method is quite useful if what is desired is simply accurate prediction with a single functional/basis pair, and in addition the method dramatically increases the power of small basis sets for prediction.[14] However, this requires a reparameterization for each pair to be used in an investigation, with all the computational requirements (and pitfalls) of semiempirical parametrization. Thus it is of little use in evaluating a new functional, and becomes impractical (for example) comparing different methods or examining the effect of increasing basis set size.

It is obvious that when evaluating a functional or a functional/basis pairing the ultimate goal is to determine the limits of accuracy for the method in question. As stated previously, using standard states as a reference, using parameters in place of atomic energies, or modeling isodesmic reactions are impractical for this purpose. However, it is easy to see that formation reactions from atoms are fundamentally not the limit of predictive power of a method, as there are practical nonparametric schemes that can be used (such as formation reactions from molecular reference states) which result in greater accuracy.

A useful set of reference states must be computationally inexpensive (for practicality), nonatomic (to avoid known issues), and contain a minimum number of compounds (to reduce complexity of application). In keeping this, we further propose that only one reference compound per element should be used. (This differs from isodesmic reactions, which need multiple reference states for each element to adequately cover all bonding situations.) For some elements, the standard state is already a practical choice. For example, the standard states of hydrogen, oxygen, nitrogen, and fluorine are all diatomic gas phase molecules, thus these are used as reference states in the new scheme. These reference states have the added benefit of unambiguous ΔH_f, namely zero. The problem of selecting a new reference state can be seen, however, in deciding on a reference state for carbon. Graphite is computationally intensive especially if out-of-phase vibrations are to be included,[15] and C_2 is a poor choice for use with DFT due to its intrinsic multireference character.[16] One could argue for the use of any hydrocarbon whose ΔH_f is well-known, but we decided that it is preferable to use single-element compounds to avoid one reference state from directly influencing another reference state. (With a hydrocarbon reference state, any issues with H_2 would effect even non-hydrogen carbon compounds.) In addition, we specifically desire closed shell molecules with large HOMO−LUMO gaps for ease and accuracy of computation, as is already the case with the selected diatomics. Thus for carbon, we chose C_{60}. This structure is computationally tractable even with large basis sets, due to its high symmetry. However, this molecule's computational cost will make up a substantial fraction of the total cost of evaluating most test sets, and thus it should be useful for functional evaluations, but may be too expensive for functional parametrization without some approximation such as fixed geometries or frequency contributions from another level of theory. (It should be noted that both of these approximations are routinely used in evaluating density functionals.[12]) Finally, C_{60}'s ΔH_f is experimentally known to be 634.8 kcal/mol; please note that this value is not without controversy and is discussed in greater detail below.[17] Using these choices of reference states, the example formation reaction of fluoromethane becomes

$$\frac{3}{2}H_2^{\text{gas}} + \frac{1}{2}F_2^{\text{gas}} + \frac{1}{60}C_{60} \rightarrow CH_3F \qquad (5)$$

A complete list of our selected selected reference states is H_2, O_2, N_2, F_2, and C_{60}. (Note that extension of this group of reference molecules would require further choices, which is not trivial, and will be dealt with in a subsequent paper.)

It should be noted that a similar set of reference states to calculate ΔH_f (at 0 K) was proposed as part of the evaluation of the HEAT method of Tajiti et al., which is a high-accuracy wave function based method in the Gn vein.[18] However, the nonatomic reference states were used in that work to avoid potential issues with the experimental accuracy of carbon and fluorine atoms, not with calculated atomic energies, as HEAT does not have the same issues as DFT.[18] In the HEAT paper, the reference states were the four diatomic standard reference states listed above plus carbon monoxide. This carbon source is a suboptimal selection for DFT, as shall seen below, but worked very well for HEAT.

In order to evaluate the proposed scheme, we formed as a test bed a subset of 141 of the 223 ΔH_f in the G3/99 test set[5] by selecting compounds consisting only of the five desired elements (CHONF) and removing the four diatomic standard reference states from the test bed. We then conducted tests with a wide array of functional-basis set pairs, and selected results are presented in Tables 1–6. Before discussing the results, some technical details must be mentioned. One deciding influence on our computational procedure was a desire for a simple, consistent, "black-box" procedure that would allow straightforward calculation of ΔH_f once reference states were chosen, with a single level of theory being used for all intermediate quantities for a selected molecule. Toward this end, all structures, energies, and vibrational frequencies were calculated using the functional/basis pair being investigated. This is different than the procedure used in many papers, where while

TABLE 1: Root-Mean-Square Error of ΔH_f (in kcal/mol) Using Various Reference States for Carbon Calculated with the 3-21G Basis Set[a]

	carbon reference states					
method	atoms	C_{60}	CH_4	CO	HCN	CF_4
LSDA	152.32	18.70	16.92	120.40	19.16	48.36
BLYP	18.61	11.21	12.30	45.33	39.80	136.42
PBE	21.93	11.64	13.07	77.51	20.79	120.45
TPSS	18.50	11.44	12.30	59.06	27.91	138.41
VSXC	22.18	14.12	26.70	59.67	30.60	124.38
B3LYP	27.88	11.79	14.53	50.46	29.82	115.17
PBEh	21.17	15.53	17.63	79.00	15.50	102.73
TPSSh	27.54	11.48	13.52	61.04	24.35	129.81
M05	20.14	15.36	14.63	48.68	18.26	10.29
M052X	28.15	12.79	10.88	57.60	17.99	51.05
HF	336.33	25.00	27.61	23.35	35.17	67.37

[a] The "atoms" column is generated from atomic reference states for all elements and is included for comparison.

TABLE 2: Root-Mean-Square Error of ΔH_f (in kcal/mol) Using Various Molecular Reference States for Carbon Calculated with the 6-311G(2df,2p)[a]

	carbon reference states					
method	atoms	C_{60}	CH_4	CO	HCN	CF_4
LSDA	167.61	13.51	12.46	84.21	32.12	53.44
BLYP	12.08	7.83	7.83	9.06	21.38	42.57
PBE	34.69	5.02	5.51	41.61	5.30	21.23
TPSS	8.69	7.41	5.94	23.84	8.26	33.31
VSXC	5.63	6.37	20.08	25.92	6.89	7.68
B3LYP	4.89	5.03	7.31	11.96	8.89	6.29
PBEh	10.53	6.65	8.56	43.18	13.23	26.54
TPSSh	4.54	5.07	5.23	25.64	4.88	13.31
M05	5.39	10.03	8.64	11.79	7.63	105.80
M052X	5.28	7.83	6.25	17.30	8.69	81.36
HF	287.34	19.65	20.40	20.96	14.87	115.17

[a] The "atoms" column is generated from atomic reference states for all elements and is included for comparison.

testing a level of theory and basis the geometries are kept fixed from previous optimization at another level of theory and basis, and the frequency data is obtained from scaled values at yet a third level of theory and basis set.[5] The one detriment of our choice is that results are not directly comparable to previously published results. However, we believe that this procedure more accurately reflects how the ΔH_f of a new structure would be generated. The ultimate goals of this paper are both to detail a consistent procedure for determining the total efficacy of various functional/basis set pairs to predict ΔH_f, and provide a path to accurately calculating the ΔH_f of a molecule of interest with a functional/basis set pair, thus all data was generated using the pairing under consideration. It should be noted that all frequency corrections were from unscaled harmonic frequencies and done at 298 K. Finally, all calculations were done in the development version of the Gaussian suite of programs.[19]

We selected a wide array of density functionals, which will only be briefly discussed. The local spin density functional (LSDA) was used, applying the correlation functional of Vosko, Wilk, and Nusair, specifically the fifth proposed functional of their paper.[20] The empirical GGA of Becke and Lee, Yang, and Parr was tested (BLYP),[10,21] as was the nonempirical GGA of Perdew, Burke, and Ernzerhof (PBE).[22] Hybrids of each of these functionals, the empirical B3LYP[23] and the nonempirical PBEh,[24,25] respectively, were also used. The meta-GGAs of Van Voorhis and Scuseria (VSXC,[26] empirical) and that of Tao, Perdew, Staroverov, and Scuseria (TPSS,[27] nonempirical) were applied, as was the corresponding nonempirical hybrid of TPSS, TPSSh.[12] The functionals by Truhlar and co-workers, the highly accurate and heavily parametrized hybrid meta-GGAs M05[7] and M05-2X,[8] were also used. Finally, for comparison purposes, standard Hartree−Fock theory was included.[28]

It is necessary at this point to justify the choice of C_{60} as a carbon reference state given the nearly infinite number of other choices, especially since many of these choices would be smaller and hence computationally less intensive, as even given symmetry of C_{60}, CO or CH_4 will always be computationally less intensive. Tests were thus run with various carbon sources, specifically C_{60}, CH_4, CO, HCN, and CF_4. Note that all of these compounds are in the test set except C_{60}, and thus their statistics will be slightly enhanced as any carbon source will get its own ΔH_f correct by definition. As the largest basis set used in this study was 6-311G(2df,2p)[29] and one of the smallest was 3-21G,[30] these basis sets were used as test cases, and the root-mean-square errors are tabulated in Tables 1 and 2. For the smaller 3-21G basis set the choice is clear: C_{60} used as a reference state provides the best overall ΔH_f prediction for 9 of the 11 functionals tested. (Note that this is effectiveness is discussed further below.) For the larger basis set, we can immediately eliminate CO, HCN, and CF_4 as reference states as they are consistently outperformed by the standard atomic reference states as well as C_{60} and CH_4 reference states. The choice between C_{60} and CH_4 can be made by noting that C_{60} does better for 7 of 11 functionals.

After selecting C_{60}, a related question arises, namely what ΔH_f to use for this reference state. The NIST Thermochemistry webbook gives the value as 610 ± 30 kcal/mol, and lists a set of several possibilities, all based on the value of Steele et al.[17] (the value used in this study) which is at the high end of the NIST range, i.e., 634.8 kcal/mol. However, it is necessary to consider what would happen to the results in this study if a value at the low end of the range was used instead. RMS error variation with reference value is shown graphically for a few functionals with the biggest basis set used in this study in Figure 1. RMS was chosen rather than unsigned error as it better reflects overall performance. Effectively, the results are unpredictable. Three functionals are nearly unaffected (BLYP, PBE, and TPSSh), three functionals do better (PBEh, VSXC, and B3LYP) and one does worse (TPSS). The four functionals that change significantly may be explained by examining their mean error, as large negative errors mean improvement as ΔH_f increases and vice-versa. However, this does not explain all behavior. Consider the BLYP and B3LYP results. At the low end of the C_{60} range, they perform equivalently, around 7.4 kcal/mol rms, and their unsigned error at the high end has about the same magnitude, 2.8 and −2.4 kcal/mol respectively, but the change in reference value dramatically improves B3LYP while BLYP gets only slightly worse. Thus if a hugely different reference value was selected, the order of the functionals from best to worst would change somewhat and the numerical results presented herein would be affected. However, the Steele et al. number has an uncertainty of ±6 kcal/mol, and an alteration of that size, namely 1%, would have a minimal effect on the results, as can be seen in Figure 1.

We now turn to detailed examination of the results of a few example basis sets, starting again with 6-311G(2df,2p) as shown in Table 3. (Please note that this is slightly smaller than the basis set used in previous evaluations of functionals on this test set, namely 6-311+G(3df,2p).)[5,12] Using atomic reference states, the prediction of ΔH_f with a "large" basis set holds few surprises. For example, functionals that are known for accuracy

TABLE 3: 6-311G(2df,2p) ΔH_f in kcal/mol[a]

	atomic reference					molecular reference				
method	ME	MAE	RMS	max (+)	max (−)	ME	MAE	RMS	max (+)	max (−)
LSDA	−149.27	149.27	167.61	0.0 ()	−353.2 (Azulene)	−5.49	10.72	13.51	28.1 (CH radical)	−40.0 (2-Nitrobutane)
BLYP	0.03	9.21	12.08	32.0 (n-Octane)	−31.3 (NF_3)	2.83	6.23	7.83	18.4 (C_2F_6)	−13.5 (Cyanogen)
PBE	−29.89	29.98	34.69	2.3 (CH_2 (1A1))	−87.3 (Azulene)	1.38	3.93	5.02	11.3 (CO)	−14.8 (N_2O)
TPSS	−7.46	7.72	8.69	7.5 (H_2O)	−22.0 (NF_3)	4.45	5.80	7.41	17.7 (t-Butanol)	−9.8 (N_2O)
VSXC	−2.21	4.27	5.63	11.4 (H_2NNH_2)	−17.8 ($CF_2{=}CF_2$)	−3.56	5.24	6.37	11.3 (H_2O_2)	−16.5 (3-Methyl-pentane)
B3LYP	2.22	3.75	4.89	12.7 (n-Octane)	−8.3 ($CF_2{=}CF_2$)	−2.43	3.90	5.03	12.1 (Ozone)	−13.1 (Naphthalene)
PBEh	−6.88	7.82	10.53	9.2 (H_2O)	−37.4 (Naphthalene)	−3.60	5.50	6.65	13.8 (Ozone)	−15.6 (n-Octane)
TPSSh	−1.52	3.72	4.54	9.8 (1,4-Benzoquinone)	−10.0 (t-Butyl radical)	1.99	4.02	5.07	13.2 (H_2O)	−10.0 (Pyrimidine)
M05	−1.76	3.67	5.39	8.0 (H_2O_2)	−23.2 (C_2F_6)	−7.10	7.61	10.03	11.5 (Ozone)	−43.8 (C_2F_6)
M052X	−2.33	3.80	5.28	20.7 (Ozone)	−17.3 (C_2F_6)	−4.69	5.89	7.83	21.8 (Ozone)	−34.3 (C_2F_6)
HF	258.51	258.51	287.34	592.3 (n-Octane)	0.0 ()	−11.74	16.06	19.65	68.8 (Ozone)	−48.3 (C_2F_6)

[a] "ME" is mean (signed) error, "MAE" is mean absolute (unsigned) error, and "RMS" is root-mean-square error. max (+) and max (−) are the errors and names of the extreme outliers.

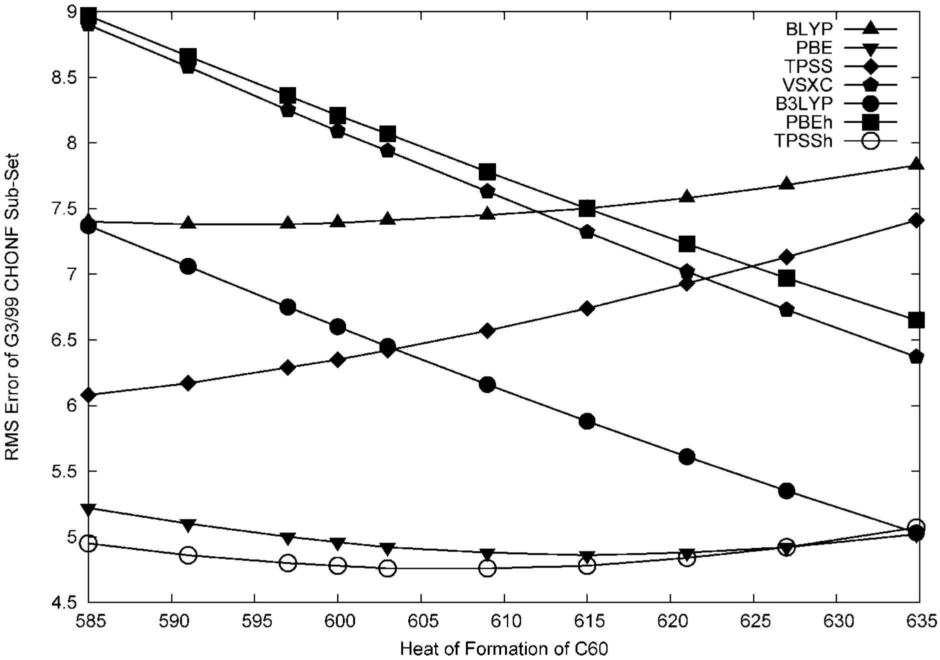

Figure 1. rms error of several density functionals when C_{60}'s ΔH_f is altered. Please note the points on the right-hand side of the graph correspond to the value used for all other calculations. All calculations were done with 6-311G(2df,2p).

TABLE 4: 6-31G(2d,2p) ΔH_f in kcal/mol[a]

	atomic reference					molecular reference				
method	ME	MAE	RMS	max (+)	max (−)	ME	MAE	RMS	max (+)	max (−)
LSDA	−155.80	155.80	175.48	0.0 ()	−364.5 (Azulene)	−4.06	11.57	14.44	31.1 (CCH)	−41.3 (2-Nitrobutane)
BLYP	−5.45	7.53	11.06	10.2 (Cyclohexane)	−39.1 (NF_3)	4.51	6.22	8.02	18.7 (t-Butanol)	−11.3 (Pyrimidine)
PBE	−35.70	35.76	40.69	2.6 (HF)	−97.2 (Azulene)	2.53	5.05	6.21	14.5 (CO)	−14.7 (N_2O)
TPSS	−13.41	13.71	15.58	6.9 (H_2O)	−36.7 (n-Octane)	5.39	6.53	7.98	17.9 (t-Butanol)	−10.0 (Pyrimidine)
VSXC	−8.01	8.42	10.18	6.7 (H_2NNH_2)	−25.1 (C_2F_6)	−2.86	5.36	6.55	12.2 (H_2O_2)	−17.8 (3-Methyl-pentane)
B3LYP	−1.80	3.23	4.03	6.9 (Cyclooctatetraene)	−12.7 ($CF_2{=}CF_2$)	−1.03	3.48	4.42	12.0 (Ozone)	−14.2 (Pyrimidine)
PBEh	−10.90	11.80	14.66	8.7 (H_2O)	−43.1 (Naphthalene)	−2.86	5.98	7.41	13.9 (Ozone)	−19.4 (n-Octane)
TPSSh	−6.77	8.54	10.41	10.6 (Cyanogen)	−31.6 (n-Octane)	2.73	4.39	5.63	14.6 (H_2O)	−11.3 (Pyrimidine)
M05	−9.80	10.55	12.54	6.4 (Ozone)	−32.3 (n-Octane)	−6.35	7.30	9.68	12.4 (Ozone)	−41.8 (C_2F_6)
M052X	−0.08	3.56	5.04	20.2 (Ozone)	−24.4 (C_2F_6)	−4.13	6.04	8.20	21.7 (Ozone)	−37.4 (C_2F_6)
HF	257.48	257.48	285.65	578.4 (n-Octane)	0.0 ()	−11.76	16.13	19.65	69.3 (Ozone)	−43.6 (Phenyl radical)

[a] "ME" is mean (signed) error, "MAE" is mean absolute (unsigned) error, and "RMS" is root-mean-square error. max (+) and max (−) are the errors and names of the extreme outliers.

of ΔH_f prediction such as M05, M05-2X, B3LYP, and TPSSh work very well for this benchmark set. Other well-known phenomena can be seen from the atomic reference state data, such as the addition of exact exchange improving accuracy (TPSSh versus TPSS, PBEh versus PBE, B3LYP versus BLYP), and HF and LSDA failing qualitatively. However, if the molecular reference states are used instead, the worst functionals are greatly improved. For example, the qualitative failures HF and LSDA, while still the least accurate functionals, become competitive with the other methods. There is also a decrease in error with PBE, which goes from fairly large overbinding to being the second best functional tested. Put another way, our scheme demonstrates that PBE can more accurately determine ΔH_f than originally thought; this should not be taken as an an endorsement of PBE for all problems, but rather as another example of the improvement possible from changing reference

TABLE 5: 3-21G ΔH_f in kcal/mol[a]

	atomic reference					molecular reference				
method	ME	MAE	RMS	max (+)	max (−)	ME	MAE	RMS	max (+)	max (−)
LSDA	−131.98	132.03	152.32	3.4 (HF)	−333.3 (n-Octane)	−7.10	14.61	18.70	29.9 (HF)	−47.9 (Piperidine)
BLYP	14.91	16.50	18.61	39.3 (Naphthalene)	−35.7 (F_2O)	2.78	7.50	11.21	53.4 (C_2F_6)	−31.4 (Cyanogen)
PBE	−15.28	17.85	21.93	26.1 (H_2O)	−56.7 (Naphthalene)	−0.26	8.81	11.64	45.8 (C_2F_6)	−22.5 (Cyanogen)
TPSS	14.33	15.30	18.50	47.0 (1,4-Benzoquinone)	−24.2 (F_2O)	2.91	7.80	11.44	55.2 (C_2F_6)	−24.6 (Cyanogen)
VSXC	19.23	19.45	22.18	49.6 (Pyrazine)	−14.9 (F_2O)	−4.26	11.27	14.12	46.1 (C_2F_6)	−32.4 (3-Methyl-Pentane)
B3LYP	24.99	25.09	27.88	59.8 (1,4-Benzoquinone)	−6.8 (F_2O)	−3.40	9.31	11.79	43.8 (C_2F_6)	−27.1 (Succinonitrile)
PBEh	17.13	17.24	21.17	55.4 (N_2O)	−3.6 (Cyclopentane)	−6.41	12.35	15.53	38.3 (C_2F_6)	−40.4 (n-Octane)
TPSSh	23.23	23.49	27.54	66.7 (1,4-Benzoquinone)	−9.1 (F_2O)	−0.04	8.51	11.48	51.3 (C_2F_6)	−22.7 (Cyanogen)
M05	13.25	15.65	20.14	60.4 (Ozone)	−16.7 (n-Octane)	−10.59	12.42	15.36	21.3 (HF)	−41.3 (n-Octane)
M052X	24.65	24.65	28.15	67.9 (N_2O)	0.0 ()	−5.57	10.16	12.79	22.2 (HF)	−31.4 (n-Octane)
HF	305.82	305.82	336.33	657.0 (Azulene)	0.0 ()	−15.62	21.50	25.00	68.7 (Ozone)	−54.6 (n-Octane)

[a] "ME" is mean (signed) error, "MAE" is mean absolute (unsigned) error, and "RMS" is root-mean-square error. max (+) and max (−) are the errors and names of the extreme outliers.

TABLE 6: STO-3G ΔH_f in kcal/mol[a]

	atomic reference					molecular reference				
method	ME	MAE	RMS	max (+)	max (−)	ME	MAE	RMS	max (+)	max (−)
LSDA	−320.86	320.86	377.21	0.0 ()	−936.1 (n-Octane)	8.98	26.97	33.75	116.9 (C_2F_6)	−64.5 (n-Octane)
BLYP	−139.17	139.91	170.26	17.9 (HF)	−466.4 (n-Octane)	17.60	22.35	31.28	139.7 (C_2F_6)	−22.7 (n-Octane)
PBE	−175.84	176.26	212.42	18.0 (HF)	−555.6 (n-Octane)	16.34	23.99	32.30	137.2 (C_2F_6)	−36.3 (n-Octane)
TPSS	−143.52	145.28	179.80	34.3 (N_2O)	−503.3 (n-Octane)	20.80	24.76	34.11	147.4 (C_2F_6)	−20.8 (n-Octane)
VSXC	−121.70	124.91	157.49	56.0 (N_2O)	−448.1 (n-Octane)	14.16	23.46	32.98	151.1 (C_2F_6)	−43.7 (n-Octane)
B3LYP	−128.06	131.65	165.54	56.7 (N_2O)	−477.9 (n-Octane)	11.75	23.28	32.30	145.3 (C_2F_6)	−50.5 (n-Octane)
PBEh	−139.34	143.14	180.45	61.3 (N_2O)	−510.8 (n-Octane)	10.54	25.96	34.92	146.4 (C_2F_6)	−64.9 (n-Octane)
TPSSh	−133.31	136.93	172.24	54.9 (N_2O)	−495.8 (n-Octane)	17.92	24.60	34.08	150.5 (C_2F_6)	−34.3 (n-Octane)
M05	−131.62	135.59	170.58	58.8 (N_2O)	−488.2 (n-Octane)	11.30	22.61	29.41	109.7 (C_2F_6)	−54.0 (n-Octane)
M052X	−108.29	117.55	150.15	90.9 (N_2O)	−447.6 (n-Octane)	7.65	27.06	36.16	146.8 (C_2F_6)	−75.9 (n-Octane)
HF	152.66	152.66	173.20	372.6 (1,4-Benzoquinone)	0.0 ()	−3.89	34.20	43.96	151.1 (C_2F_6)	−114.0 (n-Octane)

[a] "ME" is mean (signed) error, "MAE" is mean absolute (unsigned) error, and "RMS" is root-mean-square error. max (+) and max (−) are the errors and names of the extreme outliers.

states. Also noteworthy is that the best functionals with atomic reference states lose a bit of accuracy upon the change of reference. This most notable with M05, and occurs to a lesser extent with most of the former top performers. While it is tempting to argue that this is a function of parametrization of the functional, as the heavily optimized functionals M05, M05-2X, and VSXC all lose accuracy, B3LYP is barely affected and TPSSh, a nonempirical functional, also has its accuracy reduced. The critical points to be taken from this data are that using molecular reference states modifies the accuracy of the functionals, and for worst functionals, the accuracy is dramatically improved.

We next reduced the basis set to smaller double-ζ plus polarization functions basis set 6-31G(2d,2p),[29] the results of which are in Table 4. Using atomic reference states still provides good results for some functionals, such as B3LYP and M05−2X. In fact, the switch of reference states actually reduces the accuracy of these two functionals, albeit very slightly in the case of B3LYP. (In fact, with this basis set B3LYP is the best regardless of reference state.) However, in general the result of changing the reference state is more pronounced with this basis set. With 6-31G(2d,2p), only two functionals are capable of sub-10 kcal/mol accuracy (measured by rms error) using the standard atomic reference states, where as all but HF and LSDA reach this level of accuracy with molecular reference states. In addition, with this basis set there is an obvious systematic error with atomic reference state results as all density functionals overbind, which is effectively removed by our procedure. Finally, the reduction in basis set size has a lesser effect on the molecular reference state results than on those calculated with atomic reference states; this reduction in basis set sensitivity may allow for cheaper evaluations of new functional accuracy through the use of smaller basis sets, although examining individual compounds or reaction enthalpies may require larger basis sets.

The next basis set to be considered is the smallest double- ζ basis set, 3-21G. (Note that for first row compounds such as those considered here, 3-21G and 3-21G* are the same.) While the accuracy of functionals typically increases with larger basis sets, it has been shown in the past that using this basis set results in surprisingly reasonable ΔH_f,[31] and this is shown again by the atomic reference state data in Table 5. There are are two striking features from the atomic reference data for this basis set. First, for many functionals, the addition/increase of exact exchange decreases the accuracy of prediction, as seen with BLYP/B3LYP, TPSS/TPSSh, and M05/M05-2X. This may be attributed to the insufficiency of the basis set for exact exchange,[31] since hybrid functionals have increased basis set requirements. In fact, if the B3LYP functional is reoptimized in a small basis set, the optimal percentage of exact exchange is zero.[31] (After the change in reference state, it can be noted that the inclusion/increase of exact exchange is still not clearly beneficial with this basis set.) Also, the systematic error, evidenced by the signed errors, reduces greatly, removing the qualitative errors for HF and LSDA. For the other nine functionals, the unsigned errors and rms are reduced by approximately one-half. There is also significant reordering in the best-to-worst list of functionals. Finally, it should be noted that with molecular reference states this selection of this basis set becomes at least as important as functional choice, as illustrated by many functionals having the same outliers, namely an underbinding of C_2F_6 and an overbinding of cyanogen or n-octane.

We turn finally to tests with the "toy" (minimum or single-ζ) basis set STO-3G.[32] The ΔH_f calculated from atomic reference states are considered first. It can be easily seen from Table 6 that this basis set does remarkably poorly with all functionals for ΔH_f, to the point of qualitative failure and ~150 kcal/mol errors. All signed and unsigned errors for the functionals have nearly the same value, with massive overbinding for DFT and underbinding for HF. If the molecular reference states are used instead, there are a number of striking things about the data. First, the signed errors go from qualitatively incorrect values of hundreds of kcal/mol to ~25 kcal/mol for all functionals, i.e. removing atomic reference state dramatically improves the overbinding of functionals and the underbinding of HF. In fact, the signs are now reversed, with DFT underbinding and HF overbinding, and the unsigned error is ~10% of the atomic reference state values. This is not meant as an endorsement of the accuracy of minimal basis sets for ΔH_f prediction; the error bars are still unacceptably large. However, the error has been dramatically reduced without empirical parametrization, and the argument is that these values more credibly describe the limits of accuracy of ΔH_f prediction with these basis sets. Note also that the behavior is universal, with all functionals being affected in the same way, regardless of type or functional parametrization. Finally, the point noticed with 3-21G about basis selection being as important as functional choice can be seen in the outliers in Table 6.

A brief mention should be made at this point of a use for the known atomic energies and atomic ΔH_f in functional development. A functional can have good thermochemistry and poor experimental atomic energy prediction, as shown by the parametrization of atomic energies versus ΔH_f not bringing the resultant atomic energy plus parameter value closer to the experimental energies.[14] If the new reference states are used, there are two additional tests that should be conducted during functional evaluations. The first test is simply the error in atomic energy prediction, which is already commonly done. The second test would be the prediction of atomic ΔH_f, i.e. modeling reactions such as

$$\frac{1}{2}\mathrm{F}_2 \rightarrow F^{\mathrm{gas}} \qquad (6)$$

(Results for atomic ΔH_f with the 6-311G(2df,2p) basis set is included in the Supporting Information.) An excellent functional should get these quantities and the molecular ΔH_f correctly.

In conclusion, this paper proposes a new set of reference states from which molecular ΔH_f can be calculated. Using this set results in enhanced accuracy as well as a qualitative change in functional evaluation, with previously failing methods now providing reasonable results, and a reordering of the hierarchy of functional accuracy with respect to this property, especially for PBE. Future work will develop nonatomic reference states for other elements, especially sulfur and phosphorus, as well as testing this scheme on other ab initio methods. It is hoped that these reference states will be used by developers of new functionals and basis sets, as it provides a more accurate assessment of a functional's predictive power.

Acknowledgment. This work was supported by NSF CHE-0807194 and the Welch Foundation (C-0036). Special thanks are owed to Drs. Benjamin Janesko and Thomas Henderson for useful comments on this manuscript.

Supporting Information Available: Tables equivalent to Table 6 are available for the basis sets STO-2G, STO-6G, MIDI!, 3-21+G*, 6-31G, 6-31G(d), 6-31G(d,p), and 6-31+G(d,p), as well as the atomic ΔH_f calculated from the molecular reference states with the 6-311G(2df,2p) basis set. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

(2) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

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

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

(5) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374.

(6) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2005**, *123*, 124107.

(7) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123*, 161103.

(8) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comp.* **2006**, *2*, 364.

(9) Becke, A. D. *J. Chem. Phys.* **1992**, *96*, 2155.

(10) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(11) Curtiss, L. A.; Raghavachari, K.; Deutsch, P. W.; Pople, J. A. *J. Chem. Phys.* **1991b**, *95*, 2433.

(12) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *J. Chem. Phys.* **2003**, *119*, 12129.

(13) Csonka, G. I.; Ruzsinszky, A.; Tao, J.; Perdew, J. P. *Int. J. Quantum Chem.* **2005**, *101*, 506.

(14) Brothers, E. N.; Scuseria, G. E. *J. Chem. Theory Comp.* **2006**, *2*, 1045.

(15) Brothers, E. N.; Izmaylov, A. F.; Rusakov, A. A.; Scuseria, G. E. *J. Phys. Chem. B* **2007**, *111*, 13869.

(16) Watts, J. D.; Bartlett, R. J. *J. Chem. Phys.* **1992**, *96*, 6073.

(17) Steele, W. V.; Chirico, R. D.; Smith, N. K.; Billups, W. E.; Elmore, P. P. R.; Wheeler, A. E. *J. Phys. Chem.* **1992**, *96*, 4731.

(18) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599 URL http://link.aip.org/link/?JCP/121/11599/1.

(19) Gaussian Development Version, Revision F.01, Frisch, M. J. Trucks, G. W. Schlegel, H. B. Scuseria, G. E. Robb, M. A. Cheeseman, J. R. Montgomery, J. A., Jr., Vreven, T. Scalmani, G. Kudin, K. N. Iyengar, S. S. Tomasi, J. Barone, V. Mennucci, B. Cossi, M. 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. Li, X. Hratchian, H. P. Peralta, J. E. Izmaylov, A. F. Heyd, J. J. Brothers, E. Staroverov, V. Zheng, G. Kobayashi, R. Normand, J. Burant, J. C. Millam, J. M. Klene, M. Knox, J. E. Cross, J. B. Bakken, V. Adamo, C. Jaramillo, J. Gomperts, R. Stratmann, R. E. Yazyev, O. Austin, A. J. Cammi, R. Pomelli, C. Ochterski, J. W. 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. Chen, W. Wong, M. W. and Pople, J. A. Gaussian, Inc., Wallingford CT, 2006.

(20) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(21) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(22) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(23) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

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

(25) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029.

(26) Van Voorhis, T.; Scuseria, G. E. *J. Chem. Phys.* **1998**, *109*, 400.

(27) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(28) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.

(29) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.

(30) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.

(31) Brothers, E. N.; Merz, K. M. *J. Phys. Chem. A* **2004**, *108*, 2904.

(32) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657.

JP8061652