Localized orbital corrections applied to thermochemical errors in density functional theory: The role of basis set and application to molecular reactions

Dahlia A. Goldfeld, Arteum D. Bochevarov, and Richard A. Friesner^{a)}

Department of Chemistry, Columbia University, New York, New York 10027, USA

(Received 2 August 2008; accepted 7 October 2008; published online 2 December 2008)

This paper is a logical continuation of the 22 parameter, localized orbital correction (LOC) methodology that we developed in previous papers [R. A. Friesner *et al.*, J. Chem. Phys. **125**, 124107 (2006); E. H. Knoll and R. A. Friesner, J. Phys. Chem. B **110**, 18787 (2006).] This methodology allows one to redress systematic density functional theory (DFT) errors, rooted in DFT's inherent inability to accurately describe nondynamical correlation. Variants of the LOC scheme, in conjunction with B3LYP (denoted as B3LYP-LOC), were previously applied to enthalpies of formation, ionization potentials, and electron affinities and showed impressive reduction in the errors. In this paper, we demonstrate for the first time that the B3LYP-LOC scheme is robust across different basis sets [6-31G*, 6-311++G(3df,3pd), cc-pVTZ, and aug-cc-pVTZ] and reaction types (atomization reactions and molecular reactions). For example, for a test set of 70 molecular reactions, the LOC scheme reduces their mean unsigned error from 4.7 kcal/mol [obtained with B3LYP/6-311++G(3df,3pd)] to 0.8 kcal/mol. We also verified whether the LOC methodology would be equally successful if applied to the promising M05-2X functional. We conclude that although M05-2X produces better reaction enthalpies than B3LYP, the LOC scheme does not combine nearly as successfully with M05-2X than with B3LYP. A brief analysis of another functional, M06-2X, reveals that it is more accurate than M05-2X but its combination with LOC still cannot compete in accuracy with B3LYP-LOC. Indeed, B3LYP-LOC remains the best method of computing reaction enthalpies. © 2008 American Institute of Physics. [DOI: [10.1063/1.3008062](https://doi.org/10.1063/1.3008062)]

I. INTRODUCTION

In a number of previous papers^{1,2} we have developed a novel approach to improving the results obtained from density functional theory³ (DFT) based quantum chemical calculations. We argued in these papers that the predominate errors in modern DFT methods, and most clearly in the popular B3LYP functional,^{4,5} could be ascribed to inaccurate treatment of nondynamical electron correlation in localized electron pairs (and the erroneous assignment of excessive nondynamical correlation to singly occupied orbitals). An empirical localized orbital correction (LOC) scheme, DFT-LOC, was constructed² using physical reasoning to identify the key parameters to be optimized and fitting the parameters themselves to a database of enthalpies of formation compiled by Curtiss *et al.*⁶ With a total of 22 parameters, the average error for the 222 molecules in the G3 data set was reduced, for the B3LYP functional, from 4.8 to 0.8 kcal/mol using the B3LYP-LOC model. Further studies⁷ extended the correction scheme to ionization potential and electron affinities, and to transition metal containing species, with similarly improved results.

In the present paper, we return to the subject in Ref. 1, enthalpies of formation of the G3 database (whose molecules are composed exclusively of first and second row atoms), and carry out a more in depth examination of the DFT-LOC

model using the JAGUAR *ab initio* quantum chemistry code⁸ for all calculations. Specifically, we focus on the effect of basis set on the average errors obtained using LOCs and examine errors for a series of chemical reactions (in many cases the realistic objective in practice) as well as enthalpies. In addition to the B3LYP functional (our best performer in Ref. 1, as compared to other widely used functionals), we also investigate the performance of the M05-2X functional of Zhao *et al.*,⁹ both in its native form and incorporating LOCs. We have selected M05-2X for investigation because it has displayed interesting abilities to achieve reasonable results for properties that other DFT methods have great difficulty with (e.g., dispersion interactions),^{10–12} while at the same time preserving respectable results for the traditional atomization energy benchmarks (the M06-2X functional¹³ has similar advantages^{11,12,14–16}) and is the most recent version of this family of functionals; its limited analysis is also carried out in this work). Furthermore, M05-2X has previously been benchmarked on only a subset of the G3 molecules; evaluation of results for the complete G3 set provides a fuller picture of its performance. Consequently, we present not only numerical averages but also an analysis of the distribution of outliers for both the corrected and uncorrected functionals considered below.

The basis set dependence of DFT errors has been examined in a number of previous papers.^{17–20} However, the data sets used in most of these were either smaller or less diverse than what we investigate here. Understanding the basis set

^{a)}Electronic mail: rich@chem.columbia.edu.

dependence of the DFT-LOC methods is, on the other hand, a new topic; the relatively normal behavior that is observed (particularly for B3LYP-LOC) provides additional evidence that the physics of the method is solidly grounded.

A final aspect of the present work is that it has been carried out using an automated algorithm for assigning the correction terms, as opposed to the hand assignment that has been employed in previous publications. In fact, some errors in assignment were detected in the process of converting over, and these are noted below. The overall quality of the results is very similar to that reported in Ref. 1, and the vast majority of parameter values and predicted enthalpies of formation are very close (typically a few tenths of a kcal/mol) to the values reported therein (assuming that the same basis set is used). However, this step is of great importance in the process of implementing the methodology in a form that can be distributed and used by others. We discuss our plans along those lines briefly in Sec. V, below.

This paper is organized as follows. In Sec. II, we briefly review the DFT-LOC approach, outlining its rationale, implementation, and prior results. Then, in Sec. III, we specify the technical details of our computations. Section IV presents results for enthalpies of formation and reaction energies for several different basis sets and for uncorrected and corrected versions of the B3LYP and M05-2X DFT functionals, and discusses the implications of these results for the overall project of developing an empirically corrected DFT methodology. In Sec. VI, the conclusion, we summarize our results and briefly discuss future directions.

II. THE LOCs OVERVIEW

The LOC methodology developed in our laboratory is described in detail in Refs. 1 and 2. The former paper is the original work and demonstrates how the LOC approach improves the performance of DFT in predicting enthalpies. The latter paper is an extension of this methodology that improves DFT calculations of ionization potentials and electron affinities. The basic idea of our approach is simple and will be described primarily for enthalpies of formation as they are the focus of the present paper. The goal of LOC is to reduce the enthalpy errors via a collection of empirically derived parameters. The parameters were first optimized by multiple linear regression against a training set of experimental enthalpies of formation of 147 molecules (the so-called G2 data set⁶) and were then tested on a larger set of Curtiss *et al.*,²¹ the G3 data set of 222 enthalpies of formation. Twenty-two correction parameters were utilized, and initially were fitted to reproduce enthalpies of formation of the molecules in the training set. The final parameters are determined by fitting to the entire G3 data set.

The parameters are assigned based on the presence of single unpaired electrons or pairs of electrons occupying localized valence orbitals of specific types, as well as on the local environments of these orbitals and the hybridization state of the atoms in the molecule. The corrections applied to a given molecule are uniquely determined by its valence bond structure. When applied to the G3 data set of 222 enthalpies of Curtiss *et al.*,²¹ these parameters reduce the mean

unsigned error (MUE) of B3LYP calculations of the G3 set from 4.8 to 0.8 kcal/mol. This method also contributes no extra computing cost to the original DFT calculation, as the LOCs are precomputed and trivially applied *a posteriori* in an additive way. We note that since the publication of our first paper,¹ we found a few insignificant mistakes in our parameter assignments to molecules from the G3 data set, which have been corrected in the results discussed below. These errors change the numerical values presented in Ref. 1 only marginally and lead to negligible alteration of the MUE for the G3 enthalpies of formation; they also do not affect the physical interpretation or overall assessment of the LOC methodology contained therein.

The necessity for improving DFT prediction of enthalpies, even for the favorable case of organic molecules (as opposed to, e.g., transition metal containing systems, where errors are substantially larger), can be outlined as follows. A typical error of an enthalpy of formation of a small molecule containing first and second row atoms, obtained by using common functionals, is on the order of a few kcal/mol.^{6,21} This result is still significantly larger than chemical accuracy, which is commonly taken to be on the order of 1 kcal/mol error. Furthermore, even for B3LYP, which is considered one of the most accurate functionals for predicting enthalpies, there are striking outliers. For example, B3LYP/6-311+G(3df,2p) calculations of ozone and perfluorosilane enthalpies give errors of approximately 9 and 20 kcal/mol, respectively. Finally, many DFT studies suggest that enthalpy errors scale with molecular size,^{1,21,22} which is an indication that DFT errors tend to be systematic. The fact that there are systematic errors in DFT energies is of paramount importance to our LOC approach.

Let us discuss the possible sources of these systematic errors and the rationale this provides for the construction of the LOC scheme. In the hydrogen molecule, there is an electron pair contained within a chemical bond, and regardless of how stretched the molecule is, the exact [Hartree-Fock (HF)] exchange hole is delocalized over both hydrogens. However, the true exchange-correlation hole is mostly localized on an atomic center. Localized correlation components combined with HF exchange produce a delocalized exchange-correlation hole. This delocalization is incorrect and leads to large errors in enthalpies. To produce a qualitatively correct localized exchange-correlation hole, one needs to combine localized correlation with a localized exchange hole. One can obtain a (more or less) localized exchange hole by combining exact exchange with both Becke exchange and generalized gradient approximation correlation. This explains why the B3LYP functional works so well, particularly in molecules without a significant amount of delocalization of the electron pair: the resulting exchange-correlation hole is localized (we can think of the delocalized component of the correlation hole canceling out the delocalized component of the exchange hole).

Electron correlation can be pictorially thought of as existing in two length scales: correlation on the scale of an atom and on the scale of a chemical bond. We denote the former as “dynamical” correlation and the latter as “nondynamical” correlation. A localized exchange-correlation func-

tional does well at determining dynamical correlation but encounters problems in predicting nondynamical correlation. The question as to how DFT functionals such as B3LYP represent nondynamical correlation is the starting part for the development of the LOC methodology. We briefly review the arguments in Ref. 1, which link the representation of nondynamical correlation with the self-interaction terms of the DFT functional.

In an orbital with a single electron, such as the orbital in the hydrogen molecule ion, the wave function is fully delocalized by symmetry. The “self-interaction” of the electron must be zero, as the Coulomb and exchange-correlation terms, i.e., the electron interacting with itself, completely cancel. This means that there is only a kinetic energy and electron-nuclear interaction term in the remaining functional. However, the Coulomb and exchange-correlation terms in current DFT functionals do not cancel exactly, resulting in a self-interaction energy of the electron, which is artificial. This effect is most detrimental to DFT calculations at long distances when the self-interaction term becomes very negative and large. In this case, because of symmetry constraints, the density of the single electron must be equal on both atom centers, yielding a large error in the total energy because of increased self-interaction of the electron. This effect is dramatically observed when the distance between two atoms in the hydrogen molecule ion is increased substantially, as was shown by Zhang and Yang.²³

In Ref. 2 we argue (based on a number of results in literature, and also on the data contained therein) that the self-interaction “error” introduced by DFT functionals, in the case of doubly occupied orbitals, is actually used to represent the intrapair nondynamical correlation associated with those orbitals. Adjustment of parameters in the functional to fit experimental data (as is done for many, if not most, modern DFT functionals) can then be thought of as in great part adjusting the self-interaction term to reproduce the nondynamical correlation energy for the “average” chemical bond or lone pair. In singly occupied orbitals, there is no intrapair nondynamical correlation, and the self-interaction energy of the electron is simply erroneous; this leads to specific types of errors for systems containing unpaired electrons, as discussed in Ref. 1 in more detail. The core concept of the LOC approach is that the actual nondynamical correlation error in a given localized orbital will deviate from the average value implicitly assumed by the functional, and that this deviation will be transferable when the bond is found in different molecules. One key idea is that as bond length increases with respect to the size of the atomic component of the bonding orbital, the nondynamical correlation energy becomes increasingly underestimated, as the electrons in the bond have more room to avoid one another. This concept is validated in the LOC model as the bond energy of single bonds becomes increasingly underestimated as the ratio of the bond length to the orbital “size” (estimated by a simple metric in Ref. 1) increases. Each molecule may contain several structural features responsible for different sources of nondynamical correlation errors. These might include various types of atom hybridizations, bonds, and bonding environments. Since the sources of nondynamical correlation are similar for all mol-

ecules, our LOC methodology is applicable beyond the scope of the G3 data set. Table III lists 22 parameters that account for these sources of error. These parameters were found to be most necessary only for certain characteristics of a valence bond structure. For a detailed explanation of how to apply these parameters, we refer the reader to Ref. 1 and here we give only one explicit example. The DFT-computed enthalpy of formation of ammonia should be corrected additively by one instance of the N/P_sp^3 parameter (ammonia contains an sp^3 -hybridized nitrogen) and three instances of the NPOLH parameter (which account for the three N–H bonds). Our scheme does not include a parameter specifically for the hydrogen atom.

In our previous papers, we showed that the LOC methodology works extremely well, as illustrated by its application to the G3 data set of Curtiss *et al.*⁶ Since the obtained parameters were fitted to the DFT calculations of Ref. 1 for the molecules in the G2 and G3 data sets, rigorous application of our parameters requires the use of the exact calculation scheme employed therein (including how one optimizes geometry, calculation of zero-point vibrational energies (ZPEs), and does single point energy calculations). Their scheme worked in the following way: all molecular geometries were optimized using MP2(full)/6-31G* level of theory, and ZPEs were obtained using scaled HF/6-31G* calculations.⁶ Single point energy calculations were then obtained at the B3LYP/6-311+G(3df,2p) level. To assess the performance of our LOC methodology with a variety of basis sets, methods of geometry optimization energy, and method for computing ZPE's, we have performed our own DFT calculations on the G3 data set entirely at the DFT level. Thus, we first recalculated all molecular geometries using the popular B3LYP/6-31G* combination and used scaled B3LYP/6-31G* to find ZPEs. We then did single point energy calculations using two different functionals with four different basis sets for each. They are B3LYP/6-31G*, B3LYP/6-311++(3df,3pd), B3LYP/cc-pVTZ, B3LYP/aug-cc-pVTZ, M05-2X/6-31G*, M05-2X/6-311++(3df,3pd), M05-2X/cc-pVTZ, and M05-2X/aug-cc-pVTZ. Both cc-pVTZ and aug-cc-pVTZ were taken to test the effect of diffuse functions on the LOCs. The success of this protocol, as shown below, demonstrates that DFT-LOC calculations can be successfully performed using only DFT based algorithms (i.e., eliminating the MP2 geometry optimization step) with no loss of accuracy or robustness. This result is not surprising but validates a protocol with substantially reduced CPU times for practical applications on large systems.

III. COMPUTATIONAL DETAILS

Atomization energies are obtained by subtracting the energy of the molecule from the energy of the atoms, which constitute the molecule. The molecule and the atoms are assumed to be in the gaseous state. For the reaction

the atomization energy E_{at} is written as

TABLE I. The MUEs of enthalpies of formation at 298.15 K of Pople's extended G2 and full G3 set in kcal/mol. Pure values are the MUEs without any LOCs. LOC^a denotes the use of LOCs fit to the extended G2 data set. LOC^b denotes the use of LOCs fit to the entire G3 data set.

Basis	B3LYP						M05-2X					
	G2		G3		G3		G2		G3		G3	
	Pure	LOC ^a	Pure	LOC ^a	Pure	LOC ^b	Pure	LOC ^a	Pure	LOC ^a	Pure	LOC ^b
6-31G*	5.49	2.16	7.61	3.13	7.61	2.48	6.76	2.56	7.44	3.34	7.44	2.87
cc-pVTZ	4.01	0.94	6.62	1.46	6.62	1.13	2.80	1.84	2.99	2.25	2.99	2.06
aug-cc-pVTZ	4.29	0.88	7.24	1.37	7.24	1.10	2.64	1.84	2.99	2.29	2.99	2.15
6-311++G(3df,3pd)	3.15	0.67	5.09	0.81	5.09	0.77	3.15	1.94	4.00	2.22	4.00	1.93

$$E_{\text{at}}(A_x B_y \cdots C_z) = E(A_x B_y \cdots C_z) - xE(A) - yE(B) - \cdots - zE(C), \quad (3.2)$$

where E is the total energy. In this study E is approximated by a nonrelativistic total energy of the system in a finite basis set (the sum of the nuclear repulsion energy and the electronic energy). Then the standard enthalpy of formation of the compound $A_x B_y \cdots C_z$ at zero kelvin (K) is

$$\begin{aligned} \Delta H_f^\circ(A_x B_y \cdots C_z; 0\text{K}) &= x\Delta H_f^\circ(A; 0\text{K}) + y\Delta H_f^\circ(B; 0\text{K}) \\ &+ \cdots + z\Delta H_f^\circ(C; 0\text{K}) \\ &- E_{\text{at}}(A_x B_y \cdots C_z). \end{aligned} \quad (3.3)$$

The terms such as $x\Delta H_f^\circ(A; 0\text{K})$ effect a transition from $E_{\text{at}}(A_x B_y \cdots C_z)$ to $\Delta H_f^\circ(A_x B_y \cdots C_z; 0\text{K})$ since the former is defined through the gaseous states of the atoms and the latter is defined through the reference states of the elements. For an arbitrary temperature T the enthalpy of formation $\Delta H_f^\circ(A_x B_y \cdots C_z; T)$ can be defined as

$$\Delta H_f^\circ(A_x B_y \cdots C_z; T) = H^\circ(A_x B_y \cdots C_z; T) - \sum H^\circ(T), \quad (3.4)$$

where

$$\sum H^\circ(T) = xH^\circ(A; T) + yH^\circ(B; T) + \cdots + zH^\circ(C; T). \quad (3.5)$$

Taking the difference of relations (3.4) at 298.15 and 0 K, we arrive at the working equation

$$\begin{aligned} \Delta H_f^\circ(A_x B_y \cdots C_z; 298.15\text{K}) \\ &= \Delta H_f^\circ(A_x B_y \cdots C_z; 0\text{K}) \\ &+ [H^\circ(A_x B_y \cdots C_z; 298.15\text{K}) - H^\circ(A_x B_y \cdots C_z; 0\text{K})] \\ &- [\sum H^\circ(298.15\text{K}) - \sum H^\circ(0\text{K})]. \end{aligned} \quad (3.6)$$

This equation was used to compute all the theoretical values $\Delta H_f^\circ(A_x B_y \cdots C_z; 298.15\text{K})$ presented in this paper. Different terms of Eq. (3.6) were obtained from different sources, a procedure very similar to that of Curtiss and co-workers.^{6,24} The geometries, optimized at the B3LYP/6-31G* level, were used to obtain the atomization energies with the functional and basis set of interest. The atomic terms such as $\Delta H_f^\circ(A; 0\text{K})$ needed for formula (3.3) were taken from Ref. 6, which in its turn acquired them from the experimental measurements.²⁵⁻²⁷ The first bracket in Eq. (3.6), which in-

cluded ZPE, was estimated by B3LYP/6-31G* (with the scaling factor of 0.9806 for the vibrational frequencies). The second bracket in Eq. (3.6) has the same sources as $\Delta H_f^\circ(A; 0\text{K})$ —namely, Refs. 6 and 25–27, and is the enthalpy difference of a reference state per one mol of atoms. All the DFT computations were carried out by JAGUAR using very accurate grids and fully analytic integrals.

IV. RESULTS

A. Atomization reactions

Although atomization energies [Eq. (3.2)] are not necessarily relevant to experimental chemists, they are very common in the field of quantum chemistry for the purposes of benchmarking. A more pertinent characteristic is the enthalpy of formation at room temperature [Eq. (3.6)] which can be directly compared to experiment. The extensive data sets of such enthalpies of Curtiss and co-workers, believed to be accurate to better than 1 kcal/mol, provide an excellent set of experimental data with which to quantitatively compare the enthalpies obtained by different quantum chemistry methods. We do this in what follows, comparing experimental values to those obtained by B3LYP, B3LYP with applied LOC corrections (which we denote B3LYP-LOC), M05-2X, and M05-2X-LOC in combination with four different basis sets.

The data sets we use are denoted in the following way: the G2 data set contains 147 molecules (and their associated enthalpies of formation) and is the training set for the LOC parameters. The G2 complement is an additional data set of 75 molecules, and the full G3 data set contains all 222 molecules. The G2 data set is comprised of the enthalpies of 147 molecules made up only of first and second row elements. Table I shows the performance of our eight functional-basis set combinations on both G2 and G3 data sets. The first thing to note, as displayed in Table I, is that for both B3LYP and M05-2X, the LOC methodology reduces the enthalpy errors significantly: for B3LYP, the average error is decreased by more than 3 kcal/mol, and for M05-2X, between 1 and 3 kcal/mol. The average error of B3LYP-LOC/6-311++G(3df,3pd) is a mere 0.67 kcal/mol, while the plain and augmented versions of cc-pVTZ produce errors below 1 kcal/mol. The application of the LOC scheme to M05-2X reduces the MUE of the G2 data set substantially but not to the same degree as for B3LYP. However, the combination of M05-2X-LOC with all tested basis sets other than 6-31G*

FIG. 1. The histogram of errors of $\Delta H_f(298.15 \text{ K})$ across the G3 data set in the aug-cc-pVTZ basis: the white and black bars represent the pure and the LOC-corrected functionals, respectively.

still results in enthalpy errors of under 2 kcal/mol. Therefore we see that although the MUE resulting from B3LYP-LOC of the G2 data set is less than that of M05-2X-LOC, both still yield quantitatively good results. Lastly, it must be emphasized that for enthalpies obtained with both B3LYP-LOC and M05-2X-LOC and the 6-31G* basis set, the average enthalpy

FIG. 2. The histogram of errors of $\Delta H_f(298.15 \text{ K})$ across the G3 data set in the 6-311++G(3df,3pd) basis: the white and black bars represent the pure and the LOC-corrected functionals, respectively.

TABLE II. The MUEs of enthalpies of formation at 298.15 K of Pople's complement of G2 with respect to G3 (75 compounds in total) in kcal/mol. The error is defined as the experimental value minus the DFT value. Pure values are the MUEs without any LOCs. LOC^a denotes the use of LOCs fit to the extended G2 data set. LOC^b denotes the use of LOCs fit to the entire G3 data set.

Basis	B3LYP			M05-2X		
	Pure	LOC ^a	LOC ^b	Pure	LOC ^a	LOC ^b
6-31G*	11.77	5.05	2.42	8.77	4.86	2.93
cc-pVTZ	11.73	2.45	1.23	3.37	3.05	2.22
aug-cc-pVTZ	13.03	2.35	1.26	3.37	2.88	2.19
6-311++G(3df,3pd)	8.90	1.10	0.91	5.66	2.77	1.91

errors are 2.16 and 2.56 kcal/mol, respectively. While these MUEs might seem large at first glance, they are, in fact, smaller than any of the enthalpy average errors obtained by any non-LOC functional/basis set combination.

The full G3 data set, comprised of 222 molecules, includes molecules with up to 10 nonhydrogen, first and second row atoms. As in the G2 data set, the enthalpy MUE is most reduced when using B3LYP-LOC scheme. B3LYP-LOC/6-311++G(3df,3pd) also produces the lowest MUE of the enthalpies of 0.81 and 0.77 kcal/mol when the parameters are fit to the G2 or G3 data set, respectively. B3LYP-LOC/cc-pVTZ and B3LYP-LOC/aug-cc-pVTZ also both reduce the MUE from around 7 kcal/mol to below 1.4 kcal/mol. Thus, we see that for both the G2 and G3 data sets, B3LYP and B3LYP-LOC produce the best results when coupled with 6-311++G(3df,3pd). For the B3LYP-LOC/cc-pVTZ and B3LYP-LOC/aug-cc-pVTZ combinations, there is a small difference when the LOCs used are fit to the G2 training set or the full G3 set. This was also seen in the original paper,¹ and again the difference is small enough and the results are good enough so that we can be confident in the effectiveness of the parameters fit to the G2 data set. On the other hand, M05-2X-LOC does not reduce the MUE of enthalpies nearly as well as B3LYP-LOC. As we see from Table I, the MUE of the G3 set of enthalpies is only reduced to 2.22 kcal/mol when the parameters are fit to the G2 training set. When the parameters are fit to the entire G3 data set, the best MUE is 1.93 kcal/mol, which is more than two times higher than the best MUE obtained with B3LYP-LOC/6-311++G(3df,3pd). These observations are illustrated by the histograms in Figs. 1 and 2, which characterize the performance of the corrected and uncorrected B3LYP and M05-2X in the two largest basis sets. While the LOC scheme reduces the number of B3LYP outliers to a remarkably small number, the effect on the number of large M05-2X outliers is less dramatic (and similar to what we observed for functions other than B3LYP in Ref. 1). Interestingly, however, it shifts the distribution of errors from the positive to the negative values.

Next, we see the trends observed for the G2 set occurring for the G2 complement, as evident from Table II. B3LYP-LOC performs better than M05-2X-LOC, and B3LYP-LOC/6-311++G(3df,3pd) is the best method for calculating enthalpies. The fact that B3LYP-LOC consis-

TABLE III. The 22 parameters fit to both the extended G2 and G3 data sets for each basis set using the classic B3LYP hybrid functional.

	B3LYP							
	6-31G*		cc-pVTZ		aug-cc-pVTZ		6-311++G(3df,3pd)	
	G2	G3	G2	G3	G2	G3	G2	G3
Be _{sp}	7.91	7.83	7.10	7.08	7.24	7.21	7.27	7.29
N/P _{sp} ²	2.11	2.46	4.06	3.80	4.33	4.01	4.37	4.49
N/P _{sp} ³	-1.80	-3.07	2.40	2.48	3.16	3.10	3.71	3.71
N/P _{quart}	7.93	7.65	9.31	8.45	9.09	8.58	7.97	6.70
O _{sp} ²	-0.07	0.69	0.32	0.75	-0.06	0.42	0.93	1.33
O _{sp} ³	1.81	-1.71	1.90	0.44	2.10	0.58	2.25	1.78
OCT_EXP	-13.80	-4.93	-1.05	2.76	2.93	3.54	4.73	4.92
NPOLH	0.51	0.63	0.35	0.35	0.28	0.28	0.27	0.24
POLH	-6.99	-4.71	-2.43	-1.45	-1.81	-0.81	-1.38	-1.05
NPOLF	2.59	2.42	0.61	0.40	-0.04	-0.27	0.98	0.82
POLF	0.98	1.00	-0.97	-1.16	-1.74	-1.86	-0.86	-0.91
SSBC	-1.51	-1.19	-1.83	-1.72	-1.93	-1.81	-1.46	-1.36
MSBC	-2.20	-2.37	-2.39	-2.45	-2.44	-2.52	-1.89	-1.90
LSBC	-6.52	-7.40	-4.01	-4.32	-3.89	-4.21	-2.45	-2.57
DBC	-2.06	-1.99	-1.20	-1.25	-1.25	-1.34	-0.75	-1.00
TBNPOL	-6.36	-6.56	-1.57	-1.80	-1.59	-1.82	-0.72	-1.03
TBPOL	-2.23	-2.03	0.65	1.02	0.40	0.91	1.25	1.56
CT	-5.86	-5.50	-5.61	-5.82	-5.88	-6.44	-4.44	-4.52
ESBC	-0.66	-0.19	-0.74	-0.45	-0.76	-0.51	-0.56	-0.50
RH	-0.40	-4.40	0.24	0.27	0.31	0.34	0.53	0.54
R1A	1.78	1.52	1.69	1.65	1.74	1.71	1.53	1.60
RT	-1.56	-1.61	-2.20	-2.28	-2.29	-2.43	-2.34	-2.33

tently performs better than M05-2X-LOC points to the idea that the errors of M05-2X are less regular than those of B3LYP. The LOC method is successful because DFT errors are consistent among molecules. For example, the OCT_EXP parameter assumes a consistent error among all molecules with expanded octets, implying that only one parameter is sufficient to address this error. The octet expansion errors resulting from calculations using M05-2X are apparently not as uniform as those from calculations using B3LYP.

The B3LYP-optimized LOC parameters, displayed in Table III, also point to the idea that B3LYP produces errors very consistently. For the combination of all basis sets with B3LYP, the parameters do not change very much when parametrized with the G2 training set or the full G3 data set. This supports the hypothesis that nondynamical correlation errors arise from specific valence bond properties endemic to molecules with different structural properties, such as the presence of four-membered rings, or as mentioned before, expanded octets. In contrast, some of the M05-2X optimized LOC parameters dramatically change depending on whether or not they were fitted to the G2 or G3 data set (see Table IV). One particularly striking example of this change is the OCT_EXP parameters, which, for example, goes from a value of -8.67 to -0.96 for the aug-cc-pVTZ basis set. However, this large change in value occurs for each basis set, implying that it is not a singular effect. Most likely, the molecules with an expanded octet in the G2 complement are treated in a very different manner by M05-2X than those in the G2 training set, leading to a very different value of the OCT_EXP parameter. These parameter values are another

indication of M05-2X's inconsistency in the way it treats nondynamical correlation. Possibly, better results in correcting M05-2X (and other functionals as well) could be obtained by expanding the number of fitting parameters (e.g., to differentiate different types of octet expansion); however, we believe that with the current data set, this carries significant risk of overfitting. However, with additional data, perhaps generated via accurate high level *ab initio* calculations, the use of additional parameters could likely be profitably explored.

In our previous paper,¹ we chose 22 parameters to be in our final set, although we tested a few additional ones that did not prove themselves essential. In this work, we attempted to include the parameters originally omitted as well as a few new ones to our original set of 22. Examples of these were atomic hybridization corrections to sulfur and chlorine atoms. When our new, extended set of parameters was applied to both B3LYP and M05-2X, the MUEs of enthalpies of formation were slightly lowered with respect to the MUEs of the original B3LYP-LOC or M05-2X-LOC schemes. However, the differences were not large enough to warrant the inclusion of the new parameters to the original set of 22 parameters. It is obvious that larger numbers of parameters would result in smaller MUEs, but in an effort to keep the number of parameters, we use commensurate with that of other theories (for example, VSXC, which uses 21 parameters²⁸), we choose to only keep the parameters that have the largest effect. We would also like to mention that the extra parameters did not qualitatively change the general pattern of M05-2X-LOC producing substantially worse re-

TABLE IV. The 22 parameters fit to both the extended G2 and G3 data sets for each basis set using the classic M05-2X hybrid functional.

	M05-2X							
	6-31G*		cc-pVTZ		aug-cc-pVTZ		6-311++G(3df,3pd)	
	G2	G3	G2	G3	G2	G3	G2	G3
Be_sp	0.13	0.05	-1.08	-1.02	-0.96	-0.89	-0.72	-0.64
N/P_sp ²	-3.04	-2.39	-1.04	-0.59	-0.71	-0.29	-0.48	0.40
N/P_sp ³	-5.92	-7.12	-1.24	-0.25	-0.43	0.24	0.11	0.67
N/P_quart	-0.24	0.96	0.33	0.99	-0.07	0.76	-0.69	-0.65
O_sp ²	-4.82	-3.71	-3.59	-2.91	-3.64	-2.91	-2.47	-1.90
O_sp ³	-4.72	-7.44	-3.86	-4.24	-3.56	-4.10	-3.61	-3.12
OCT_EXP	-25.69	-10.14	-12.53	-2.43	-8.67	-0.96	-8.04	-0.08
NPOLH	-0.41	-0.30	0.00	-0.12	-0.06	-0.24	0.07	-0.03
POLH	-3.91	-2.25	1.43	1.56	1.85	2.07	2.16	1.75
NPOLF	2.83	2.88	2.12	2.21	2.07	2.16	3.06	3.13
POLF	2.31	2.26	1.64	1.43	1.54	1.35	2.12	2.04
SSBC	0.77	1.02	0.80	0.92	0.76	1.06	1.40	1.47
MSBC	-0.73	-0.73	-0.61	-0.32	-0.55	-0.12	0.09	0.37
LSBC	-4.32	-4.96	-1.56	-1.59	-1.45	-1.35	-0.12	0.01
DBC	-0.26	-0.81	1.07	0.50	1.11	0.43	1.56	0.90
TBNPOL	-5.51	-5.97	-0.61	-1.00	-0.52	-0.84	0.76	0.25
TBPOL	-6.22	-6.08	-3.21	-2.83	-3.29	-2.79	-1.44	-1.19
CT	-2.06	-1.91	-0.10	-0.84	0.22	-0.78	0.65	0.43
ESBC	0.21	0.54	0.30	0.42	0.31	0.38	0.44	0.35
RH	-1.05	-1.08	-0.40	-0.34	-0.32	-0.21	-0.19	-0.17
R1A	1.16	0.74	0.84	0.72	0.83	0.78	0.68	0.65
RT	-4.91	-4.81	-5.49	-5.42	-5.50	-5.50	-5.43	-5.25

sults than B3LYP-LOC. The molecules in the G2 data set are small, up to five nonhydrogen atoms. For these molecules, as displayed in Table I, B3LYP/6-31G* and B3LYP/6-311++G(3df,3pd), without the application of the LOCs, produce enthalpies with average errors lower than M05-2X by about 1 kcal/mol with these two basis sets. For the cc-pVTZ and aug-cc-pVTZ basis sets, however, M05-2X leads to lower average errors than B3LYP by a little over 1 kcal/mol. Thus, for the G2 data set, B3LYP produces relatively better results than M05-2X with Pople-type basis sets, while M05-2X works better than B3LYP with Dunning-type basis sets. The G2 complement, which contains only the bigger molecules, tells a different story. For these molecules, as shown in Table II, M05-2X performs better than B3LYP with all basis sets, although we still see that M05-2X performs better with cc-pVTZ and aug-cc-pVTZ than with either 6-31G* or 6-311++G(3df,3pd). For these molecules, the differences are striking. For example, with the aug-cc-pVTZ basis set, the average enthalpy error obtained with B3LYP is 9.66 kcal/mol higher than with M05-2X. For the full G3 data set, we again see that M05-2X outperforms B3LYP with all tested basis sets, although the differences between the two are less drastic. Overall, we find that the M05-2X does indeed predict enthalpies of formation better than B3LYP, without the application of the LOCs.

All of these results point to several features of our LOC methodology. We also gained new insights into a comparison between the seminal and popular B3LYP functional and the new M05-2X functional. With regard to our study of how well the LOC methodology works using a range of basis sets,

B3LYP/6-31G* geometries, and scaled B3LYP/6-31G* frequency calculations, we show that it is indeed a robust scheme. As we demonstrated in our original paper,¹ the 22 LOC parameters do not generally change that much when fitted to the G2 or G3 data sets. Just as in our former paper, the LOC methodology lowers the MUE of enthalpies of both data sets to less than 1 kcal/mol when using the 6-311++G(3df,3pd) basis set, which is very similar to the basis set of Curtiss *et al.* used in Ref. 6. Thus, the B3LYP-LOC/6311++G(3df,3pd) combination produces enthalpies of formation within chemical accuracy, which is a remarkable result. In fact, such a good performance has never been shown before with other DFT-related methods. While B3LYP-LOC does not perform quite as well for the G3 data set when using either the cc-pVTZ or aug-cc-pVTZ basis sets, overall, it still produces enthalpies of formation with the lowest MUEs of any M05-2X method tested in this paper. The M05-2X functional does not respond nearly as well to the applied LOC parameters. As explained previously, this, in conjunction with the fact that several parameters change radically when fitted to either the G2 set or G3 set using the functional, implies that errors formed when using M05-2X are less consistent in nature than those of B3LYP. However, when comparing plain, non-LOC-corrected M05-2X and B3LYP calculations of enthalpies of formation for the entire G3 set, M05-2X produces considerably lower MUEs than B3LYP does. The best basis set for doing these types of calculations using M05-2X is aug-cc-pVTZ, while the best basis set for B3LYP is 6-311++G(3df,3pd). M05-2X produces more and larger outlying

errors than B3LYP does. The best overall methodology for calculating enthalpies is B3LYP-LOC, which not only reduces the MUE but also dramatically reduces the number of large outliers and shrinks the spread of errors.

B. Molecular reactions

One can think of the enthalpies of formation as ΔH of the “atomization” reactions. Such reactions are convenient for testing purposes, but the majority of concrete chemical problems involve more general types of reactions, where various types of chemical compounds are found on both sides of the equation and diverse chemical bonds break and form. We will refer to this latter, more general type of reactions as *molecular* reactions. A particular set of parameters, which is parametrized to perform well on the atomization reactions, is not guaranteed to do equally well, in a least-squares sense, on an ensemble of molecular reactions even in the case when the latter consists entirely of the compounds from the atomization data set. At first glance, this statement might look surprising since the molecular reactions can be written as linear combinations of the atomization reactions. However, carrying out the so-called elementary operations on the set of linear equations (such as constructing their linear combinations) leaves the solution invariant only when the unique solution of the original system exists. When it does not, as in the case of an overdetermined system (i.e., there are more equations than the unknowns), the solution in the least-squares sense is generally not invariant under linear combinations. This is because the least-squares solution satisfies some of the equations in the original overdetermined system only approximately, and therefore adding the left- and right-hand sides of such equations to the other equations modifies the latter. This, in turn, modifies the whole system and leads to a different least-squares solution. That is why one might need to add typical molecular reactions to the original training set of the atomization reactions if one expects the LOC scheme to do well on test sets of arbitrary composition. Besides, augmenting the atomization reactions with the molecular reactions increases the number of equations to be solved and makes the solved-for coefficients more accurate. In essence, fitting only to atomization reactions can incorporate subtle systematic biases, leading to less robust cancellation of error when treating molecular reactions than one would intuitively expect. By explicitly incorporating molecular reactions into the least-squares fitting process, this difficulty can be overcome, and relatively uniform performance across the possible ensemble of molecular reactions that can be built from the training set can be obtained.

Following the above analysis, the G2 set of atomization reactions was complemented by the set of 90 molecular reactions built by us from the compounds of the G2 set [this set listed in Table V, will be referred to as the MRG2 training set (MRG2-TRN) in the rest of this paper]. We selected this set in such a way that it covered the G2 set more or less uniformly. A few ubiquitous compounds (for example, water, ammonia, hydrogen, and oxygen) occur quite often in MRG2-TRN but such a “bias” is perhaps compensated by the fact that these common compounds are likely to occur

more frequently in practical applications. The “experimental” enthalpies of the reactions from the MRG2-TRN were determined from the corresponding G2 experimental enthalpies of formation. These experimental enthalpies served to measure the errors of the corresponding enthalpies predicted by a DFT approach. Another collection of 70 reactions, also consisting entirely of the G2 compounds and called the MRG2 test set (MRG2-TST), was used for the testing purposes (for the contents of MRG2-TST see Supplementary Information²⁹).

Table VI gives an account of the performance of two parametrization schemes (trained on the G2 set and on G2 simultaneously with MRG2-TRN) applied to the test set of molecular reactions (MRG2-TST). Let us first discuss the results obtained with B3LYP. With a remarkable constancy, the error of both the pure and the LOC-assisted B3LYP method decreases in the row 6-31G*, cc-pVTZ, aug-cc-pVTZ, 6-311++G(3df,3pd). The last basis is smaller than aug-cc-pVTZ but stably gives the smallest errors. Another observation concerns the comparison of the MUEs that are associated with different parametrization schemes. It is clear that the parameters derived from the optimization over the G2 test set (the A set of coefficients in Table VI) are capable of good performance on both the atomization and molecular reactions. However, the set of coefficients trained on both the molecular and atomization reactions performs noticeably better on either the test molecular reactions or the mixture thereof with the atomization reactions.

Next, we turn to the M05-2X method. In all the instances of the application of the LOC to this functional, the reduction in error was unsatisfactory. The reduction ratio was about 2.5 in favorable cases and well below 2.0 in most other cases, even when the larger bases were used (compare this to a typical reduction factor of 4.5 in the case of B3LYP). It must be concluded that either the errors of M05-2X are much less systematic than those of B3LYP or their nature is entirely different from that on which the LOC scheme is erected.

A typical pattern of molecular reaction errors produced by B3LYP, M05-2X, and their LOC-corrected analogs can be observed by examining Table V, where these methods are applied to the MRG2-TRN set of reactions (using the MRG2-TRN-trained set of coefficients). The “pure” B3LYP and M05-2X have similar MUEs and a comparable number of outliers, although M05-2X has a greater number of exceptionally large outliers (over 15 kcal/mol in absolute error). The LOC scheme reduces the average error of B3LYP sharply (by a factor of 7.3), but it does not perform very convincingly on M05-2X (the reduction factor is only 2.1). It is interesting that there is little correlation in how the two functionals and their LOC-analogs handle a particular reaction: it is generally not true that when one functional has a large error the other one does too.

As a test of the robustness of our scheme, the coefficients trained on the MRG2-TRN set were applied to a set of 70 reactions composed of the G2 compounds, the MRG2-TST set. This test set was selected in such a way that its MUE computed with B3LYP/6-311++G(3df,3pd) would be similar to that of the MRG2-TRN set so that the complexities of the two sets might be considered roughly equivalent.

TABLE V. The MRG2-TRN set and the $\Delta H^\circ(298.15K)$ errors of the constituent reactions obtained with B3LYP and M05-2X in the 6-311++G(3df,3pd) basis set. The $\Delta H^\circ(298.15K)$ errors are given in kcal/mol. The error is defined as the experimental value minus the DFT value. All errors are with respect to the reaction enthalpies obtained from the experimental enthalpies of formation of the individual compounds. The coefficients used in the LOC scheme were obtained from the solution of the linear system constructed after the MRG2-TRN set.

Reactions	B3LYP	B3LYP+LOC	M05-2X	M05-2X+LOC
1. $H_2NNH_2 = N_2 + 2H_2$	-4.80	-0.69	-8.17	-3.16
2. $CH_3Cl = CH_3 + Cl$	3.67	0.06	-1.67	-1.29
3. $CH_2 = C = CH_2 = CH_3CCH$	-3.62	-0.16	-2.18	-0.15
4. $C_4H_6 = \text{Bicyclobutane}$	-7.03	-1.00	-2.21	3.38
5. $\text{Cyclobutane} = \text{Cyclobutene} + H_2$	0.09	-0.75	-4.27	-2.34
6. $\text{Isobutane} = \text{trans-butane}$	-1.02	0.28	0.13	-3.2
7. $\text{trans-butane} = 2C_2H_5$	9.35	0.30	0.69	1.38
8. $CH_2Cl_2 = CH_2(1A_1) + Cl_2$	1.51	-0.88	-4.58	-2.58
9. $CH_2Cl_2 = CH_2(3B_1) + 2Cl$	6.22	-0.73	-4.00	-2.30
10. $CH_3CN = CH_3 + CN$	-0.28	-0.33	-8.53	-1.25
11. $CH_3NO_2 = CH_3ONO$	-1.31	-0.81	-0.96	-3.60
12. $CH_3NO_2 = CH_3 + NO_2$	5.14	-0.45	-4.03	-3.37
13. $\text{Cyanogen} = 2CN$	-5.75	-0.46	-12.25	1.59
14. $\text{Dimethylamine} = \text{trans-ethylamine}$	0.20	1.16	0.66	3.08
15. $\text{trans-ethylamine} = NH_2 + C_2H_5$	7.19	-0.38	-0.55	1.49
16. $CH_3CH_2SH = CH_3SCH_3$	0.65	0.51	0.81	1.30
17. $CH_3CH_2SH = C_2H_5 + SH$	7.72	1.34	0.18	-0.13
18. $\text{Acetone} = CH_3 + CH_3CO$	7.11	-0.40	-0.52	-1.11
19. $\text{Acetone} = 2CH_3 + CO$	4.47	-1.01	-2.97	0.04
20. $C_2H_5O + HCl = C_2H_5OH + Cl$	-2.45	-0.98	0.08	1.15
21. $BeH + OH = H_2O + Be$	-11.15	-1.09	-0.43	-0.84
22. $BeH + Cl = HCl + Be$	-8.84	0.62	0.71	0.41
23. $BeH + F = HF + Be$	-8.98	0.47	0.73	0.43
24. $2CH_3 = 2C_2H_6$	-5.73	-0.32	0.21	-0.50
25. $CH_3 + NH_2 = CH_3NH_2$	-6.49	-0.74	0.02	-2.03
26. $CH_3 + CH_3CClO = \text{Acetone} + Cl$	-2.54	-0.73	-0.32	-0.66
27. $NH + H = NH_2$	2.11	-0.29	0.76	-0.51
28. $NH_2 + OH = NH + H_2O$	-5.19	-1.34	-2.03	-2.50
29. $NH_3 + \text{Oxirane} = H_2O + \text{Aziridine}$	-1.84	0.83	-0.32	-2.98
30. $NH_3 + CH_3OCH_3 = CH_3OH + CH_3NH_2$	0.00	0.38	-0.31	-0.48
31. $NH_3 + CH_3CFO = HF + CH_3CONH_2$	-1.12	0.38	0.51	2.10
32. $NH_3 + CH_3CClO = HCl + CH_3CONH_2$	-0.05	0.32	1.37	-2.38
33. $NH_3 + \text{Furan} = H_2O + \text{Pyrrole}$	-0.91	-0.15	1.28	-2.36
34. $H_2O + CH_2F_2 = 2HF + H_2CO$	-0.44	-0.58	-5.44	3.40
35. $H_2O + CH_2Cl_2 = 2HCl + H_2CO$	4.41	0.55	-2.84	-3.24
36. $H_2O + CH_3OCH_3 = 2CH_3OH$	1.49	0.54	0.59	1.76
37. $H_2O + CH_3SCH_3 = H_2S + CH_3OCH_3$	3.69	-0.62	0.79	-2.36
38. $H_2O + CH_3CFO = HF + CH_3COOH$	-0.87	-0.71	0.28	3.21
39. $P_2 + 2H_2 = 2PH_2$	11.84	-0.29	5.11	-1.36
40. $SiH_3 + HCl = SiH_4 + Cl$	-0.58	-0.11	-1.51	-0.52
41. $PH_2 + H = PH_3$	-2.72	-0.10	-2.90	-2.45
42. $H_2S + Cl_2 = 2HCl + S$	1.49	1.02	-0.47	-4.48
43. $H_2S + \text{Oxirane} = H_2O + \text{Thiirane}$	-2.83	-1.15	0.36	-2.91
44. $CN + H_2 = HCN + H$	2.30	-1.24	8.83	1.93
45. $CH_3OH + HCOOH = H_2O + HCOOCH_3$	-0.71	0.24	0.54	-0.63
46. $N_2 + O_2 = 2NO$	1.30	0.11	-1.33	-1.81
47. $O_2 + S_2 = 2OS$	-2.87	0.08	-3.67	-1.53
48. $O_2 + COS = CO_2 + OS$	-3.03	-0.08	-1.29	0.85
49. $2O + H_2 = H_2O_2$	-1.31	0.06	-1.81	1.40
50. $F_2 + CH_3CO = CH_3CFO + F$	-1.42	1.34	11.37	7.00
51. $Cl_2 + CH_3CO = CH_3CClO + Cl$	-1.57	1.98	1.11	1.29
52. $Cl_2 + C_2H_5 = C_2H_5Cl + Cl$	-2.24	1.03	1.50	0.38
53. $CH_3SH + H_2 = CH_4 + H_2S$	1.55	-0.60	-0.40	-1.15
54. $BF_3 + AlCl_3 = BCl_3 + AlF_3$	-4.32	-0.48	-9.04	3.49
55. $COS + H_2 = H_2CO + S$	-0.61	-0.64	-0.92	0.37
56. $CH_3OCH_3 + H_2 = CH_4 + CH_3OH$	0.70	-0.24	-0.53	1.47
57. $C_2H_5Cl + H_2 = HCl + C_2H_6$	1.28	-0.63	-0.49	-0.91

TABLE V. (Continued.)

Reactions	B3LYP	B3LYP+LOC	M05-2X	M05-2X+LOC
58. $\text{H}_2 + \text{SH} = \text{H}_2\text{S} + \text{H}$	-2.18	-1.09	1.05	1.24
59. $\text{H}_2 + \text{NO}_2 = \text{H}_2\text{O} + \text{NO}$	-4.82	0.12	4.20	3.00
60. $(\text{CH}_3)_3\text{C} + \text{H} = \text{Isobutane}$	-4.26	-0.32	-0.82	2.09
61. $2\text{CH}_4 + \text{O}_2 = 2\text{CH}_3\text{OH}$	-4.93	1.44	5.13	4.59
62. $2\text{H}_2\text{O} + \text{CF}_4 = 4\text{HF} + \text{CO}_2$	2.69	1.49	-7.80	6.69
63. $2\text{H}_2\text{O} + \text{CCl}_4 = 4\text{HCl} + \text{CO}_2$	12.48	0.05	-2.96	-2.25
64. $\text{NO} + \text{O}_3 = \text{NO}_2 + \text{O}_2$	12.63	-1.03	18.30	13.97
65. $2\text{H}_2\text{O} + \text{CHCl}_3 = 3\text{HCl} + \text{HCOOH}$	7.72	1.01	-0.55	-1.03
66. $\text{H}_2\text{O} + 2\text{F}_2 = 2\text{HF} + \text{F}_2\text{O}$	3.93	1.16	11.11	0.70
67. $3\text{C}_2\text{H}_2 = \text{Benzene}$	0.40	-0.67	11.04	2.22
68. $2\text{N}_2 + \text{O}_2 = 2\text{N}_2\text{O}$	6.56	0.80	-3.82	-1.50
69. $\text{N}_2 + 2\text{O}_2 = 2\text{NO}_2$	4.43	-0.26	-6.84	-3.80
70. $2\text{NO} + \text{O}_2 = 2\text{NO}_2$	3.13	-0.37	-5.51	-1.99
71. $2\text{O}_2 + \text{CS} = \text{CO}_2 + \text{SO}_2$	-10.43	1.73	-6.39	-0.19
72. $2\text{O}_2 + \text{DMSO} = \text{H}_2\text{O} + 2\text{H}_2\text{CO} + \text{SO}_2$	-7.91	-0.51	-10.62	-4.59
73. $\text{O}_2 + 2\text{ClF}_3 = 3\text{F}_2 + 2\text{ClO}$	-12.09	1.21	-28.37	0.19
74. $2\text{trans-butadiene} + 11\text{O}_2 = 6\text{H}_2\text{O} + 8\text{CO}_2$	-37.55	-0.98	-2.58	0.71
75. $\text{O}_2 + 2\text{CH}_3\text{CHO} = 2\text{CH}_3\text{COOH}$	-7.85	0.41	4.34	1.45
76. $2\text{F}_2 + \text{CH}_2 = \text{CHCl} + \text{ClF}_3 + \text{CH}_2 = \text{CHF}$	6.48	0.09	15.97	-2.40
77. $\text{F}_2 + 2\text{Li} = 2\text{LiF}$	1.09	0.12	15.12	4.67
78. $\text{CO}_2 + 2\text{H}_2 = 2\text{H}_2\text{O} + \text{C}$	-3.27	0.63	2.33	2.88
79. $\text{Cl}_2 + \text{CH}_3\text{ONO} + \text{H}_2 = \text{HCl} + \text{CH}_3\text{OH} + \text{ClNO}$	2.25	1.33	-0.49	-3.04
80. $\text{CH}_2 = \text{C} = \text{CH}_2 + 2\text{H}_2 = \text{C}_3\text{H}_8$	-5.66	-2.10	2.81	0.96
81. $\text{C}_3\text{H}_4 + 2\text{H}_2 = \text{C}_3\text{H}_8$	-0.49	0.94	3.94	5.42
82. $3\text{F}_2 + \text{Cl}_2 = 2\text{ClF}_3$	13.94	1.17	25.90	1.39
83. $\text{Cl}_2 + 2\text{CH}_2 = \text{C} = \text{O} + \text{H}_2 = 2\text{CH}_3\text{CClO}$	-7.81	-0.44	1.44	1.06
84. $\text{F}_2 + 2\text{CH}_2 = \text{C} = \text{O} + \text{H}_2 = 2\text{CH}_3\text{CFO}$	-6.74	-1.28	12.57	2.31
85. $\text{SiH}_4 + 4\text{Cl}_2 = 4\text{HCl} + \text{SiCl}_4$	-11.57	-0.55	8.30	-4.05
86. $3\text{F}_2 + 3\text{CH}_3\text{CN} = 3\text{HF} + 2\text{CF}_3\text{CN}$	-2.42	0.14	39.99	5.55
87. $4\text{Cl}_2 + \text{CH}_3\text{SiH}_3 = 3\text{HCl} + \text{CH}_3\text{Cl} + \text{SiCl}_4$	-8.38	0.95	9.58	-2.23
88. $\text{Thiophene} + 4\text{O}_2 = 2\text{H}_2\text{O} + 4\text{CO} + \text{SO}_2$	-29.31	0.28	-19.95	0.69
89. $\text{PH}_2 + \text{HF} = \text{F} + \text{PH}_3$	-1.81	-0.03	-2.79	-0.71
90. $\text{CH}_3\text{CHO} + \text{F} = \text{CH}_3\text{CFO} + \text{H}$	-1.30	0.81	2.34	-3.9
MUE	4.87	0.67	4.74	2.22

TABLE VI. The performance of different parametrizations of the LOC approach on various sets of atomization and molecular reactions with compounds from G2 set only. The $\Delta H^\circ(298.15 \text{ K})$ errors are given in kcal/mol. The coefficients of the LOC parametrization were first found by solving a system of linear equations in the least-square sense. Then these coefficients were used in the application of the LOC scheme to a set of reactions. The first row of the table indicates the equations from which the coefficients were derived, whereas the second row shows the reactions to which the coefficients were applied. A stands for the atomization reactions of the G2 set (147 reactions), TRN stands for the MRG2 training set (90 reactions), and TST stands for the MRG2 test set (60 reactions). The MUEs of the enthalpies of formation for the LOC scheme are compared to the analogous values of the pure (without LOC) DFT method.

Coefficients→ Reactions→	A		A		A+TRN		A+TRN	
	A		TST		TST		A+TST	
	Pure	LOC	Pure	LOC	Pure	LOC	Pure	LOC
B3LYP/631G*	5.49	2.16	11.84	5.63	11.84	4.15	7.54	3.10
B3LYP/cc-pVTZ	4.01	0.94	6.56	2.31	6.56	1.53	4.83	1.25
B3LYP/aug-cc-pVTZ	4.29	0.88	5.79	1.88	5.79	1.20	4.77	1.10
B3LYP/6-311++G(3df,3pd)	3.15	0.67	4.72	1.08	4.72	0.79	3.66	0.77
M05-2X/631G*	6.76	2.56	9.21	5.67	9.21	3.91	7.55	3.36
M05-2X/cc-pVTZ	2.80	1.84	4.51	3.68	4.51	2.97	3.35	2.44
M05-2X/aug-cc-pVTZ	2.64	1.84	4.11	3.74	4.11	3.00	3.11	2.44
M05-2X/6-311++G(3df,3pd)	3.15	1.94	3.68	3.30	3.68	2.99	3.32	2.53

lent. Supplementary Material²⁹ contains the reactions comprising the MRG2-TST set as well as their deviations from the experimental enthalpies. The data for several parametrization schemes are shown. Let us first focus on the coefficients parametrized on the MRG2-TRN set (the LOC-TRN columns). There is a modest increase in the MUE (from 0.67 to 0.89 kcal/mol) for B3LYP-LOC when compared to the MUE registered in the case of the MRG2-TRN set. Somewhat surprisingly, this time M05-2X shows a better result (its MUE is 3.68 kcal/mol) although it was doing only slightly better than B3LYP when applied to the training set. Nevertheless, M05-2X-LOC fails to greatly improve this MUE, and lowers it only to 3.35 kcal/mol (a factor of 1.1, compared to 5.3 in the case of B3LYP-LOC). This result further strengthens our position that B3LYP and M05-2X are similar in how they predict enthalpies and that B3LYP-LOC is a much more successful method than the analogous M05-2X-LOC. B3LYP-LOC is successful not only when compared to M05-2X-LOC but also in the absolute sense: the average unsigned error of 0.89 kcal/mol obtained on a set of 70 diverse reactions should be regarded as a fine chemical accuracy.

In Sec. IV A, we showed that B3LYP-LOC parametrized on the G2 atomization reactions retained its good performance when applied to the G3 test and especially to the complement of G2 with respect to G3 (see Table I). After the encouraging testing with the reactions based on G2, it is now logical to test whether B3LYP-LOC will be effective in predicting the enthalpies of reactions containing the molecules from the complement of G2. There are only 75 molecules in the latter set and their chemical structures are such that it is very hard to construct a few dozen more or less realistic reactions composed only of these molecules. For this reason our MRG3-TST set containing 70 reactions included compounds from G2, but in such a way that each reaction included at least one compound from these 75 entries making up the complement. The MRG3-TST should be considered “difficult” since the compounds from the G2 complement are generally larger and more challenging molecules. A detailed report for the MRG3-TST set can be found in the Supplementary Material.²⁹ As before, all methods give roughly the same quality of results. B3LYP works worse than M05-2X on average (the MUEs are 4.62 and 3.24 kcal/mol, respectively) but B3LYP-LOC convincingly outperforms M05-2X-LOC. The MUEs of B3LYP-LOC are in the range of 0.97–1.52 (depending on the used LOC parametrization), whereas the analogous range for M05-2X-LOC is 3.48–4.10. For B3LYP, the LOC-TRN parametrization, which did not include the molecules from the G2 complement increases the MUE roughly by a factor of 1.6 in comparison with the MUE of MRG2-TST set. This is consistent with the case of the G3 complement of atomization reactions (see Table II). Here the addition of either the atomization reactions or the compounds from the G2 complement during parametrization helps significantly: in both cases the MUE diminishes to approximately chemical accuracy. The various parametrizations of M05-2X demonstrate a significantly worse, and systematic errors are not eliminated to the extent seen for B3LYP-LOC. It is striking how large an error the pure and the LOC-

corrected M05-2X methods give for reaction 54 of the MRG3-TST set which involves five molecules of F₂: for the pure M05-2X it is 41.14 kcal/mol and for the LOC-corrected M05-2X it is from 27.72 to 35.86 kcal/mol (depending on the parametrization scheme). Reaction 12, which has two molecules of F₂ is also quite an unfortunate case with the error 25.08 kcal/mol for the pure M05-2X. This should come as no surprise given the large error the M05-2X method predicts for the enthalpy of formation of F₂ and other fluorine-containing compounds (see Supplementary Material²⁹). In view of this, M05-2X also does not perform well on reactions containing F₂ in the MRG2-TRN and MRG2-TST sets.

V. DISCUSSION

The performance of B3LYP-LOC that we report here is consistent with that presented in Ref. 1. When we employ a basis set that is very similar (although not identical) to that used by Curtiss *et al.* in obtaining their results in Ref. 6 (the data from which were used in the original B3LYP-LOC parametrization), the MUE obtained for both the G2 and G3 data sets is within 0.1 kcal/mol of the results of Ref. 1, and the parameter values are also quite similar (although not identical).

Generally, the quality of the B3LYP-LOC results increases as the size of the basis set increases. This is an important new result, as it supports the assertion that the B3LYP-LOC methodology has a sound underlying physical basis. A robust model chemistry should provide optimal results as the basis set is converged, as opposed to requiring a particular basis set, of intermediate size, to generate its best results; if this is the case, it implies that the functional is compensating for basis set errors. Such compensation does not make the method useless (indeed, the B3LYP-LOC results for the smaller basis sets have an element of compensation of this type) but this is an additional burden upon the method that is likely to restrict its range of validity with regard to novel chemical entities.

The slightly better performance of the 6-311++G(3df,3pd) basis set (in combination with B3LYP), as compared to the aug-cc-pVTZ basis, is a little puzzling, as the latter is somewhat larger than the former. However, these bases have a different structure: the former is of the Pople type and the latter is of the Dunning type. For now, we regard this issue as a minor discrepancy, which does not affect the general conclusions of our study, and intend to pursue it further in future work. It may, in fact, be that further increase in basis set size, in the appropriate regions of the basis set space, would lead to results that are better still; understanding the sensitivity of the results for each class of bonds, atoms, molecules, etc., to particular components of the basis set is the best way to address this issue and design the most compact, but fully converged, basis set possible. It should be noted that the size of the basis set needed to achieve the highest accuracy and greatest robustness is not a serious computational issue since the final energy calculations are required only for a single point (after geometry optimization) and, given JAGUAR's performance for large basis sets, this is unlikely to constitute a substantial fraction of the CPU time

as contrasted to what is required for geometry optimization.

The consistency of performance for molecular reactions, while not surprising, further confirms the stability and reliability of the B3LYP-LOC model. The use of reaction data enables selection of the most stable parameters, which yield consistent energy differences as a function of chemical transformation, from a parameter space that is relatively flat in some regions. Stability in analyzing molecular reaction thermodynamics is critical to the next phase of our LOC effort, which is designing models capable of handling transition states. To do this effectively, the thermodynamics of the forward and reverse reactions must be treated consistently, and the current model clearly accomplishes this objective.

The M05-2X results are consistent with previously published work. For the G2 set, the errors for M05-2X are slightly smaller (using the optimal basis sets for each) than for B3LYP. M05-2X does not yield a qualitative reduction in the number or magnitude of large outliers in the G2 set (see the histograms in Figs. 1 and 2). For the G2 complement, M05-2X outperforms B3LYP significantly (again assuming optimal basis sets are used for each, with an average error of 3.37 kcal/mol versus 8.90 kcal/mol for B3LYP). This is unsurprising given the fact that M05-2X contains many more parameters than B3LYP. The most significant cause of the diminished errors in the G2 complement appears to be the better performance for hydrocarbons (or molecules containing a large component of hydrocarbon structure) as the system size increases. This improved performance has been noted previously in a study that considered only hydrocarbons.³⁰ It is important to note, however, that the superior treatment of hydrocarbons does not necessarily carry over to other chemical groups and, hence, does not guarantee or even suggest any improvement in robustness and generality. The G2/G3 set is highly enriched with hydrocarbon and near-hydrocarbon species; if the fitting protocol is, in effect, specially optimized for these species (whether by design or inadvertently), then a rather unbalanced result can be obtained.

The M05-2X-LOC results are clearly not nearly as good as the B3LYP-LOC results despite using the same number of fitting parameters. Given the rather different concentration of errors in M05-2X (e.g., in Si containing compounds), we investigated the effects of expanding the LOC parameter list to parameters previously discarded as irrelevant (in Ref. 1), e.g., an atomic correction parameter for Si. None of these experiments yielded a significant improvement in the MUE for either M05-2X-LOC or B3LYP-LOC. We can only conclude, as was suggested above, that B3LYP has an error distribution that is particularly well suited to LOC-type corrections. It is nevertheless the case that the M05-2X-LOC model does represent a nontrivial improvement upon the M05-2X model itself. However, B3LYP-LOC remains in a class by itself with regard to the ability to achieve (on average) chemical accuracy for the tests we have performed herein.

It is interesting to examine the outliers in the best B3LYP-LOC and M05-2X-LOC calculations. For B3LYP-LOC, there are a few errors in the 2–4 kcal/mol range. Some of these may be due to experimental error (based on excel-

TABLE VII. The performance of the M06-2X functional and its LOC analog on the G2 set. The $\Delta H^\circ(298.15\text{ K})$ errors are given in kcal/mol.

Basis	M06-2X	M06-2X-LOC
6-31G*	4.65	2.36
cc-pVTZ	2.55	1.58
aug-cc-pVTZ	2.43	1.52
6-311++G(3df,3pd)	2.39	1.66

lent agreement between the B3LYP-LOC and G3 results, as noted previously in Ref. 1), while others might be improved by increasing the number and type of parameters. There is only one truly large outlier in this data set, POCl₃, the error for which increased from −5 to −7 kcal/mol with the new parametrization in the present paper (as compared to Ref. 1). This molecule contains three halogens (always difficult atoms), an octet expansion, and a charge transfer bond. Possibly there are some issues about how to combine these three factors; alternatively, there may be a problem with assigning a single parameter for octet expansion or charge transfer. In our view, more data are required before making any attempt to improve the results by adding more parameters; such data could, in principle, be obtained by carrying out chemically accurate CCSD(T) calculations for molecules with features similar to POCl₃. Tracking down the origin of this and others, outliers will become a high priority once a fully functioning LOC methodology, including transition state and transition metal capabilities, has been implemented in JAGUAR. However, it should be recognized that the number of outliers is remarkably small considering the size and diversity of the full G3 data set.

In contrast, the M05-2X and M05-2X-LOC results contain a significant number of large errors, and in many cases the LOCs make little or no impact on the major errors already present in M05-2X; examples of this type include SiH₄, F₂, and a number of other Si and F containing compounds. In other cases, the LOCs improve, but do not eliminate, large M05-2X errors, examples being ozone, SO₂, and other oxygen and sulfur-containing compounds. These error patterns, which are very different from those manifested in B3LYP-LOC, suggest that new types of errors in the physics are causing systematic problems, and perhaps a modified version of the LOC approach could provide greatly improved systematic answers (with minimal increase in the number of parameters).

A detailed investigation of the error patterns of other functionals extends beyond the reach of this work, but we nevertheless decided to make a step in this direction by assessing the performance of the promising M06-2X (and its LOC analog) on the G2 set. Even though we did not consider the G2 complement or the molecular reactions for this functional, the G2 set is a good starting point for judging the consistency of errors. Table VII shows the results of M06-2X and M06-2X-LOC applied to the G2 test set. Compared to M05-2X, pure M06-2X displays a substantial reduction in the MUEs for 6-31G* and 6-311++G(3df,3pd) and a modest improvement in the two Dunning-type bases. The LOC-corrected M06-2X also shows an improvement over M05-

2X-LOC in all the bases, with the smallest MUE of 1.52 in aug-cc-pVTZ. In spite of this improvement, the M06-2X-LOC's best result still cannot compete with B3LYP-LOC's 0.67 kcal/mol in 6-311++G (3df, 3pd) and even 0.88 kcal/mol in aug-cc-pVTZ. Besides, the ratio of the pure and LOC-corrected MUEs is roughly the same for both M05-2X and M06-2X (and is much smaller than that of B3LYP). These facts indicate that M06-2X, although being an improvement over M05-2X, may still be suffering from the same type of errors as M05-2X.

VI. CONCLUSIONS

In this work we extended our previously published LOC methodology to several basis sets [6-31G*, 6-311++G(3df,3pd), cc-pVTZ, and aug-cc-pVTZ], one more functional (M05-2X), and molecular reactions constructed from the G3 set of molecules. We wanted to confirm that our LOC methodology, which showed such remarkable progress in lowering the errors of atomization energies in DFT, is transferable to other DFT computational schemes such as different basis sets or different ways of optimizing molecular geometries. We chose to specifically examine only two functionals: B3LYP and M05-2X. The former is the functional for which this LOC scheme was designed and happens to be the method when combined with LOC among all the other functionals tested in the original paper.¹ M05-2X was a compelling choice for this paper because it had not been tested with LOCs before and because it has shown itself as an excellent alternative to B3LYP. We confirm that M05-2X actually does produce enthalpies with considerably lower MUEs than B3LYP for both atomization and molecular heats of formation. This effect is particularly striking for the G2 complement, which consists of mostly larger molecules. Furthermore, we show that Dunning-type basis sets are best used in combination with M05-2X, while Pople-type basis sets suit B3LYP better. However, the most noticeable result of this paper is that B3LYP combined with LOC works significantly better than either M05-2X or M05-2X-LOC. The M06-2X functional, which we applied only to the G2 set, shows an overall improvement over M05-2X (in both its pure or LOC forms) but still is not a rival for B3LYP-LOC. Indeed, B3LYP-LOC is again shown to be the best known DFT-related method of producing accurate enthalpies of reactions. The progress demonstrated in this paper marks an important step in the development of our LOC methodology.

We have recently completed an initial study of the effects of LOCs upon small transition metal containing species, which has yielded highly promising results. Initial investigation of transition states suggests that the errors obtained with B3LYP for activation enthalpies are highly systematic, and that corrections consistent with those derived for enthalpies of formation will yield a similar level of re-

duction in the observed errors. Finally, the current LOC approach provides parameters, which can be applied at single geometry points; the development of a continuous function, which agrees with the present LOCs at stationary points but can be applied across the entire range of geometries, is a high priority and appears to be quite feasible based on preliminary evaluations.

ACKNOWLEDGMENTS

This work was supported in part by grants from the NIH (Grant No. GM40526) and DOE (Grant No. DE-FG02-90ER-14162).

- ¹R. A. Friesner, E. H. Knoll, and Y. Cao, *J. Chem. Phys.* **125**, 124107 (2006).
- ²E. H. Knoll and R. A. Friesner, *J. Phys. Chem. B* **110**, 18787 (2006).
- ³R. C. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
- ⁴A. D. Becke, *J. Chem. Phys.* **98**, 1372 (1993).
- ⁵A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- ⁶L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **106**, 1063 (1997).
- ⁷D. Rinaldo and R. A. Friesner (unpublished).
- ⁸JAGUAR 7.0, Schrödinger, Inc., Portland, OR, 2007.
- ⁹Y. Zhao, N. E. Shultz, and D. G. Truhlar, *J. Chem. Theory Comput.* **2**, 364 (2006).
- ¹⁰Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.* **3**, 289 (2007).
- ¹¹Y. Zhao and D. G. Truhlar, *Acc. Chem. Res.* **41**, 157 (2008).
- ¹²E. E. Dahlke, R. M. Olson, H. R. Leverenz, and D. G. Truhlar, *J. Phys. Chem. A* **112**, 3976 (2008).
- ¹³Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.* **120**, 215 (2008).
- ¹⁴F. Cozzi, R. Annunziata, M. Benaglia, K. K. Baldridge, G. Aguirre, J. Estrada, Y. Sritana-Anant, and J. S. Siegel, *Phys. Chem. Chem. Phys.* **10**, 2686 (2008).
- ¹⁵H. Valdes, K. Pluháčková, M. Pitonák, J. Řezáč, and P. Hobza, *Phys. Chem. Chem. Phys.* **10**, 2747 (2008).
- ¹⁶Y. Zhao and D. G. Truhlar, *Phys. Chem. Chem. Phys.* **10**, 2813 (2008).
- ¹⁷N. X. Wang and A. K. Wilson, *J. Chem. Phys.* **121**, 7632 (2004).
- ¹⁸N. X. Wang and A. K. Wilson, *Mol. Phys.* **103**, 345 (2005).
- ¹⁹R. Vessecchi and S. E. Galembeck, *J. Phys. Chem. A* **112**, 4060 (2008).
- ²⁰J. Tirado-Rives and W. L. Jorgensen, *J. Chem. Theory Comput.* **4**, 297 (2008).
- ²¹L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **112**, 7374 (2000).
- ²²M. D. Wodrich, C. Corminboeuf, and P. v. R. Schleyer, *Org. Lett.* **8**, 3631 (2006).
- ²³Y. Zhang and W. Yang, *J. Chem. Phys.* **109**, 2604 (1998).
- ²⁴L. A. Curtiss, K. Raghavachari, P. W. Deutsch, and J. A. Pople, *J. Chem. Phys.* **95**, 2433 (1991).
- ²⁵E. Storms and B. Mueller, *J. Chem. Phys.* **81**, 318 (1977).
- ²⁶B. Ruscic, C. A. Mayhew, and J. Berkowitz, *J. Chem. Phys.* **88**, 5580 (1988).
- ²⁷M. W. Chase, Jr., C. A. Davies, J. R. Downey, Jr., D. J. Frurip, R. A. McDonald, and A. N. Syverud, *J. Phys. Chem. Ref. Data* **14**, Suppl. No. 1, 421 (1985).
- ²⁸T. van Voorhis and G. E. Scuseria, *J. Chem. Phys.* **109**, 400 (1998).
- ²⁹See EPAPS Document No. E-JCPSA6-129-007842 for supplementary tables. For more information on EPAPS, see <http://www.aip.org/pubservs/epaps.html>.
- ³⁰M. D. Wodrich, C. Corminboeuf, P. R. Schreiner, A. A. Fokin, and P. v. R. Schleyer, *Org. Lett.* **9**, 1851 (2007).