Assessment of Gaussian-3 and density-functional theories on the G3/05 test set of experimental energies

Larry A. Curtiss[a)] and Paul C. Redfern
Materials Science and Chemistry Divisions, Argonne National Laboratory, Argonne, Illinois 60439

Krishnan Raghavachari
Department of Chemistry, Indiana University, Bloomington, Indiana 47401

The G3/99 test set [L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **112**, 7374 (2000)] of thermochemical data for validation of quantum chemical methods is expanded to include 78 additional energies including 14 enthalpies of formation of the first- and second-row nonhydrogen molecules, 58 energies of molecules containing the third-row elements K, Ca, and Ga–Kr, and 6 hydrogen-bonded complexes. The criterion used for selecting the additional systems is the same as before, i.e., experimental uncertainties less than ±1 kcal/mol. This new set, referred to as the G3/05 test set, has a total of 454 energies. The G3 and G3X theories are found to have mean absolute deviations of 1.13 and 1.01 kcal/mol, respectively, when applied to the G3/05 test set. Both methods have larger errors for the nonhydrogen subset of 79 species for which they have mean absolute deviations of 2.10 and 1.64 kcal/mol, respectively. On all of the other types of energies the G3 and G3X methods are very reliable. The G3/05 test set is also used to assess density-functional methods including a series of new functionals. The most accurate functional for the G3/05 test set is B98 with a mean absolute deviation of 3.33 kcal/mol, compared to 4.14 kcal/mol for B3LYP. The latter functional has especially large errors for larger molecules with a mean absolute deviation of 9 kcal/mol for molecules having 28 or more valence electrons. For smaller molecules B3LYP does as well or better than B98 and the other functionals. It is found that many of the density-functional methods have significant errors for the larger molecules in the test set. © *2005 American Institute of Physics.* [DOI: 10.1063/1.2039080]

I. INTRODUCTION

There is much interest in the development of reliable quantum chemical methods for the prediction of thermochemical data.[1–3] Accurate test sets of experimental data are important for the critical evaluation of these new methods as well as for the development of new methods. The Gn test sets have been widely used for these purposes.[4–7] The first in this series, the "G2 test set"[4] contained 125 energies (55 atomization energies, 38 ionization potentials, 25 electron affinities, and 7 proton affinities) and was limited to 1 or 2 nonhydrogen atoms, with the exception of SO_2 and CO_2. In the second in this series, G2/97,[5,6] the G2 test set was expanded from 125 to 301 energies (147 atomization energies, 88 ionization potentials, 58 electron affinities, and 8 proton affinities). The aim of the test set was to include larger molecules such as benzene, butane, and SiF_4. In the latest in the series, G3/99,[7] the number of energies was expanded from 301 to 376 with 75 additional enthalpies of formation. The G3/99 test set included larger test species such as octane, naphthalene, and SF_6. These test sets of accurate experimental data have been used in the development of the Gaussian-n theories.[4,8–12] They have also been used in the development of a number of density-functional methods. The criterion used in compiling the Gn test sets includes (1) use of a wide variety of species, (2) each entry should have an experimental accuracy of ±1 kcal/mol, and (3) none of the values should be biased by theoretical input.

A number of other test sets have also been proposed and have been used in testing of quantum chemical methods. Boese and Handy[13] have developed the HCTH407 test set, which contains 407 test energies as well as some subsets of a smaller number of energies. The criterion used in selecting these is not as strict in some cases. This test set and modifications of it have been used by Boese *et al.* to assess a variety of density-functional methods.[14] Cioslowski *et al.*[15] have developed a set of 600 molecules, but again the uncertainties of some of the entries are much greater than 1 kcal/mol. Lynch *et al.*[16] have compiled test sets of different quantities including 109 atomization energies, 42 barriers, 13 ionization energies, and 13 electron affinities. Most of the species are the same as in the G2/97 test set. The test set of Lynch *et al.* is based on electronic energies obtained by removing temperature corrections and zero-point energies from experimental enthalpies in many cases. The use of electronic energies facilitates comparison with theory, but introduces theoretical input, and possibly additional error.

In this paper we present a test set referred to as G3/05 that is an expansion of the G3/99 test set.[7] Seventy-eight energies are added including (1) 14 enthalpies of formation of the first- and second-row nonhydrogen molecules, (2) 58 energies of molecules containing the third-row nontransition

---

[a)]Electronic mail: curtiss@anl.gov

elements K, Ca, and Ga–Kr, and (3) 6 hydrogen-bonded complexes. The addition of the 14 nonhydrogen molecules was done because of an apparent weakness in G3 methods for these species that we found in previous assessments on the G3/99 test set. The G3 mean absolute deviation (MAD) for the 47 nonhydrogens was 2.11 kcal/mol, significantly larger than for other subsets of molecules.[7] The implementation of the G3X methodology,[9] which included a larger basis set at the Hartree-Fock (HF) level and better geometries, reduced the mean absolute deviation for this set to 1.49 kcal/mol. There were still some molecules with larger deviations, as large as 4.3 kcal/mol, indicating that the G3X method did not completely correct the problem. Thus, there is a clear need for further assessments on these nonhydrogen systems. The 58 third-row nontransition and 6 hydrogen-bonded complexes are added to expand the variety of energy types in the test set. The details of the G3/05 test set are given in Sec. II. The results of the assessment of the G3 and G3X methods on this expanded test set are given in Sec. III.

In this paper we also report the assessment of a number of density functionals on the G3/05 test set. Previously, we reported the assessment of seven density functionals (SVWN, BLYP, BP86, BPW91, B3LYP, B3P86, and B3PW91).[5,6] The B3LYP functional was subsequently tested on the G3/99 test set.[7] In another paper we examined the use of B3LYP geometries and zero-point energies.[17] This did not change the results significantly. Since then there have been numerous new density functionals that have been proposed by different groups. In this paper we examine 23 functionals with the G2/97 test set and 11 of these are further assessed on the G3/05 test set. The results of these assessments are given in Sec. IV. In this section we also investigate the dependence of the accuracy of density-functional methods on the size of the molecules in the test set. In two previous studies[7,18] we have found that B3LYP tends to have larger errors as the molecules increase in size. Therefore, it is of interest to carry out a systematic investigation of the size dependence of other functionals.

II. THE G3/05 TEST SET

The G3/05 test set is composed of the G3/99 test set and 78 new energies as described in the Introduction. The 14 new enthalpies of the formation for the first- and second-row non-hydrogen molecules are taken largely from the JANAF-NIST thermochemical compilation.[19] The 58 energies of molecules containing the third-row nontransition elements K, Ca, and Ga–Kr include 47 energies from our previous paper[20] that extended G3 theory to the third-row nonhydrogen molecules. This includes 22 atomization energies, 17 ionization potentials, 4 electron affinities, and 2 proton affinities. The other 11 third-row energies contain bromine and are enthalpies taken largely from the thermochemical tables of Pedley *et al.*[21] The six hydrogen-bonded complexes were chosen to cover a range of experimental hydrogen bond energies from 1 to 15 kcal/mol. These are the water dimer, methanol dimer, acetone dimer, acetic acid dimer, hydrogen fluoride dimer, and hydrogen chloride dimer. Again the quoted uncertainties in all of these energies are less than 1 kcal/mol.

The total number of entries in the G3/05 test set is 454 energies. The energies are divided into subgroups as before (neutral enthalpies of formation, ionization energies, electron affinities, and proton affinities) along with a new subset containing hydrogen-bonded complexes. In some cases atomization energies of D_0 values are used (third-row and hydrogen-bonded complexes) as conversion of these values to enthalpies of formation is not possible without introducing theoretical input into the "experimental" energy. The enthalpies of the neutrals are further divided into types of molecules (hydrocarbons, substituted hydrocarbons, nonhydrogens, inorganic hydrides, and radicals).

III. ASSESSMENT OF GAUSSIAN-3 THEORIES

A. Theoretical methods

Gaussian-3 (G3) and Gaussian-3X (G3X) theories have been previously described. The G3 theory is a composite electronic structure method that is based on a sequence of *ab initio* molecular-orbital calculations. It approximates a QCISD(T)/G3Large level of calculation at the MP2/6-31G* geometry and uses scaled HF/6-31G* zero-point energies. Details of the G3 theory are given elsewhere.[8] The G3X theory is an extension of the G3 theory that uses B3LYP/6-31G(2df,p) geometries and zero-point energies (scaled by 0.9854).[9] It also includes a HF calculation at a larger basis set than G3Large that includes g functions for second-row elements. The G3 theory has been developed for the first-, second-, and third-row nontransition-metal elements. It has been previously tested on 47 of the 58 third-row molecules in the G3/05 test set.[20] The G3X theory has been developed and tested thus far only on the first- and second-row systems.

The extension of the G3X theory for the third-row nontransition-metal elements is done in a similar manner as for the second-row elements with the same higher level correction.[9] We use geometries at the B3LYP/6-31G(2df) level and zero-point energies at that level scaled by 0.9854.[22] The G3XLarge basis set used in the final Hartree-Fock level calculation has g functions added to the G3Large basis set. The g function exponents are taken from Dunning's cc-pVQZ basis sets for Ga–Kr (Ga: 0.403, Ge: 0.468, As: 0.546, Se: 0.573, Br: 0.649, and Kr: 0.739).[23]

The enthalpies of formation in this paper were calculated in the same manner as in Ref. 5. The temperature correction to 298 K for molecules was obtained using the rigid-rotor/harmonic-oscillator approximation. The vibrational part of the correction was obtained using scaled HF/6-31G* frequencies for the G3 theory and scaled B3LYP/6-31G(2df,p) frequencies for the G3X theory.

B. Results and discussion

The G3 and G3X results for the new entries in the G3/05 test set are given in Tables I–III. Table I contains the results for the 14 new nonhydrogens based on the first- and second-row elements. Table II contains the results for the 58 species containing the third-row elements, and Table III contains the

TABLE I. Deviations of different theoretical methods from the experiment of the 14 additional nonhydrogen enthalpies of formation in the G3/05 test set. [Density-functional results based on single-point 6-311+G($3df,2p$) energies at MP2/6-31G* geometries with scaled (0.89) HF/6-31G* zero-point energies.]

Species	Expt.[a] ΔH_f^0(298 K)	Deviation (Expt.—Theor.) (kcal/mol)								
		G3	G3X	B3LYP	B98	B97-2	B1B95	O3LYP	VSXC	TPSS
AlF	−63.5±0.8	2.4	2.9	−2.1	−1.6	−1.8	−1.9	−4.7	3.9	−2.8
Al$_2$Cl$_6$	−309.7±0.8	9.9	7.8	−29.1	−3.3	−2.7	2.7	−31.9	30.0	−2.8
ClNO$_2$	2.9±0.4	0.4	1.0	4.5	7.4	11.4	5.4	10.9	8.3	18.6
NaF	−69.4±0.5	1.2	0.8	−2.6	0.3	1.3	−3.7	−3.1	4.9	−0.1
ClFO$_3$	−5.1±0.7	−9.2	−5.0	−18.3	−11.6	−3.9	−8.3	−3.6	−8.2	4.9
N$_2$O$_3$	19.8±0.2	−2.7	0.7	2.5	2.9	7.7	−2.1	7.7	7.8	20.1
N$_2$O$_4$	2.2±0.4	−1.3	1.6	6.4	8.4	15.8	3.6	16.4	15.6	28.5
C$_6$F$_6$	−228.4±0.4[b]	4.1	3.1	−2.9	16.1	30.5	24.4	6.4	22.2	20.3
C$_6$F$_5$Cl	−194.1±0.7[b]	3.2	2.0	−7.9	12.3	27.9	22.6	1.7	19.8	16.5
COCl$_2$	−52.4±0.1[b]	0.9	0.7	−3.6	4.0	8.4	6.7	2.2	2.9	5.7
LiNa	43.4±0.3[c]	3.3	3.8	−1.4	2.1	2.9	−4.1	0.7	3.9	1.0
BeF$_2$	−190.3±1.0	−2.3	−3.5	2.6	1.6	0.1	−0.8	−2.4	2.2	0.6
ClCN	32.9±1.0[b]	1.0	1.6	−0.9	0.1	4.9	2.3	−0.3	−0.4	3.8
MgCl$_2$	−93.8±0.5	0.5	−0.3	−6.1	0.7	1.8	0.6	−4.5	5.8	2.7
MAD[d]		3.04	2.49	6.50	5.18	8.65	6.37	6.90	9.70	9.18

[a]Except where noted the experimental data is from Ref. 19.
[b]Reference 21.
[c]References 52 and 15.
[d]Mean absolute deviation.

results for the six hydrogen-bonded complexes. A summary of the mean absolute deviations for the G3 and G3X theories is given in Table IV.

Overall the mean absolute deviations of G3 and G3X for the G3/05 test set are 1.13 and 1.01 kcal/mol, respectively. This is slightly larger than for the G3/99 test set where the deviations were 1.07 and 0.95 kcal/mol, respectively.[7] The reason for the increase is largely due to the 32 additional nonhydrogen species (14 containing only the first- and second-row elements and 18 containing the third-row elements also) that are included in the G3/05 test set. These 32 new molecules have mean absolute deviations of 2.08 and 1.87 kcal/mol, respectively, at the G3 and G3X levels. Contributing to this large mean average absolute deviation are some of the 14 first- and second-row nonhydrogen species in Table I such as Al$_2$Cl$_6$ (deviation from the experiment of 9.9 for the G3 theory and 7.8 for the G3X theory), ClFO$_3$ (−9.2 kcal/mol at G3 and −5.0 at G3X), and C$_6$F$_6$ (4.1 at G3 and 3.1 at G3X). G3X has a mean absolute deviation of 2.49 kcal/mol for this set of 14 species compared to 3.04 for the G3 theory. Thus, while the G3X method is an improvement over G3 for nonhydrogen species, it is still well above the target accuracy goal of 1 kcal/mol. This may be due to remaining basis deficiencies in the G3X theory.

Overall, the mean absolute deviations of G3 and G3X on the 58 third-row species in G3/05 are 1.14 and 1.11 kcal/mol, respectively. Thus, the two methods perform similarly on the third-row species.

The mean absolute deviation of the G3 and G3X theories for hydrogen-bonded complexes are 0.60 and 1.06 kcal/mol, respectively. The larger deviation for G3X is due to the use of B3LYP/6-31G($2df,p$) for geometry optimization, which fails for the water and hydrogen fluoride dimers where it gives incorrect cyclic structures (double hydrogen bonding).

Also, the MP2/6-31G* geometry in the G3 theory is the reason for the large error in the G3 binding energy for the hydrogen fluoride dimer. The inclusion of diffuse functions in the basis sets corrects the problems in the density-functional geometries of the water dimer and the hydrogen fluoride dimer. The G3 and G3X binding-energy results based on B3LYP/6-31+G($2df,p$) optimized geometries are given in parentheses in Table III. The use of this geometry gives very good agreement with the experiment for the G3X theory and improves the mean absolute deviation to 0.49 kcal/mol. It also improves agreement of the G3 theory with the experiment for the HF dimer (see Table III). Thus, the G3 methods do well on the hydrogen bonding energies, if an accurate enough method for geometry optimization is used.

We note that, while we have used great care in selecting accurate (±1 kcal/mol) experimental energies, it is possible that some experimental values may be in error by more than this amount. Since the number of cases is likely to be small, we expect that this will have little effect on the conclusions of the assessments in this paper. Recently, Ruscic *et al.*[24] have developed a methodology for critical analysis of thermochemical data using a network approach. They plan to carry out a comprehensive reevaluation of thermochemical data in the major compilations using this approach. In the future, we plan to update our experimental test set based on the improved thermochemical data from this work.

IV. ASSESSMENT OF DENSITY-FUNCTIONAL THEORIES

A. Theoretical methods

There have been numerous new density-functional methods that have been proposed since our papers on the G2/97

TABLE II. Deviations of different theoretical methods from the experiment of the 58 third-row molecules added to G3/05 test set. [Density-functional results based on single-point 6-311+G($3df,2p$) energies at MP2/6-31G* geometries with scaled (0.89) HF/6-31G* zero-point energies.]

		Expt.[a] (kcal/mol)	Deviation (Expt.-Theor.) (kcal/mol)								
	Species		G3	G3X	B3LYP	B98	B97-2	B1B95	O3LYP	VSXC	TPSS
D_0	K_2	12.6	−3.1	−2.2	1.0	−6.1	−7.5	2.2	−1.8	−8.6	−2.3
	KBr	90.5	−1.0	0.0	1.3	−4.1	−5.7	−0.5	2.4	−5.4	−3.8
	KCl	101.0	−0.6	0.3	3.4	−2.4	−3.9	1.0	3.5	−4.7	−1.3
	KF	117.6	1.7	2.6	1.7	−1.6	−3.0	2.4	2.2	−6.1	−2.6
	GeH_4	270.5	−2.5	−2.8	−5.7	−4.0	−1.9	−3.8	−5.6	−5.5	−15.7
	AsH	64.6	−0.1	−0.4	−2.9	−1.0	−1.0	−0.8	−1.6	−0.2	−3.7
	AsH_2	131.1	−0.8	−1.2	−6.3	−3.4	−3.3	−3.2	−4.0	−3.1	−8.7
	AsH_3	206.0	1.4	0.8	−3.9	−0.6	−0.3	−0.9	−1.3	−1.0	−8.8
	SeH	74.3	−1.1	−1.1	−1.7	−0.6	−0.6	−1.1	−1.3	−3.7	−3.6
	SeH_2	153.2	0.9	0.6	−2.0	−0.5	−0.5	−1.9	−2.2	−4.3	−6.6
	Hbr	86.5	−0.2	0.0	−1.9	−1.5	−1.6	−2.6	−3.1	−2.5	−4.6
	GaCl	109.9	−1.5	−1.3	0.7	−3.1	−3.6	−3.4	1.5	−4.6	−3.2
	GeO	155.2	−1.6	−3.0	0.0	−0.9	−1.7	0.7	2.3	−6.5	−3.7
	As_2	91.3	−0.4	−1.6	−1.0	0.4	−4.7	−0.7	7.8	0.1	−2.7
	BrCl	51.5	0.3	0.5	−0.5	−3.5	−4.1	−4.9	−1.4	−0.1	−6.7
	BrF	58.9	0.3	−0.2	−1.9	−2.6	−3.0	−3.2	−2.2	−2.9	−8.3
	BrO	55.3	0.1	−0.1	−2.2	−4.7	−4.9	−4.2	−4.4	−4.8	−10.6
	Br_2	45.4	−0.1	0.1	−3.2	−5.5	−6.3	−6.9	−2.8	−2.1	−9.5
	BBr	103.5	0.7	1.1	2.5	−1.0	−2.0	−1.3	1.4	−3.0	−0.5
	NaBr	86.2	−2.1	−1.9	1.8	−3.2	−4.5	0.4	2.6	−4.5	−2.4
	CH_3Br	358.2	−0.3	−0.1	−1.1	−1.4	−1.4	−2.2	−3.8	−0.2	−8.2
	GeS_2	191.7	−1.9	−2.5	3.0	−5.1	−9.4	−9.2	1.0	−8.9	−9.7
	KrF_2	21.9	−0.6	−1.3	−3.4	−3.2	−2.6	−0.9	−5.6	−8.7	−18.4
ΔH_f^0(298 K)	CF_3Br	−155.0±0.7	2.3	1.3	−1.9	6.8	9.6	8.6	2.6	4.2	9.6
	CCl_3Br[b]	−10.0±0.4	2.9	2.2	−11.4	1.1	5.5	6.5	−7.2	−0.1	5.0
	C_2H_3Br[b]	18.9±0.4	2.0	1.9	1.1	2.3	4.6	4.3	4.1	3.5	9.2
	C_2H_5Br	−14.8±0.4	1.2	0.9	−0.2	0.6	0.9	1.9	4.0	−0.1	9.8
	C_3H_7Br[c]	−23.8±0.8	0.6	0.1	−4.4	−3.0	−2.3	−0.9	1.0	−0.4	8.6
	C_6H_5Br[d]	25.2±1.0	1.5	1.4	−6.0	1.2	11.6	8.8	0.1	6.5	9.3
	$C_6H_{13}Br$[e]	−35.4±0.4	1.2	0.7	−9.7	−7.0	−5.1	−3.3	0.4	−5.7	11.6
	$C_3H_6Br_2$[f]	−17.1±0.3	2.7	2.2	−5.3	−0.2	1.4	2.8	−1.3	2.6	11.2
	CHF_2Br	−101.6±0.2	1.2	0.4	−0.5	5.1	6.8	6.3	3.2	2.6	10.2
	$COBr_2$	−27.1±0.1	2.7	2.6	1.5	8.1	11.4	9.3	4.3	6.5	12.1
	$C_5H_8Br_2$[g]	−13.1±0.4	3.2	3.1	−14.4	−4.5	0.2	2.9	−7.3	4.0	8.3
IP	K	100.1	0.8	0.8	−3.6	2.8	6.4	2.1	0.4	4.8	2.3
	Ca	140.9	−1.7	−2.1	−1.0	−2.5	−2.0	4.7	2.8	0.3	3.0
	Ga	138.3	−0.2	−0.2	−0.7	2.1	3.7	3.6	2.5	4.1	−0.2
	Ge	182.2	−0.1	−0.1	−0.1	1.5	3.4	3.0	2.3	1.8	−0.8
	As	225.7	−0.4	−0.3	−1.8	−1.6	0.3	−0.1	−0.7	−2.5	−3.5
	Se	224.9	1.0	0.5	−2.4	0.8	1.6		4.4	6.2	1.6
	Br	272.4	0.5	0.0	−3.3	−1.4	−0.7	0.4	1.6	0.1	−0.8
	Kr	322.8	1.3	0.9	−4.0	−3.2	−2.7	−1.6	−1.5	−2.1	−2.5
	AsH	222.3	−1.0	−1.0	−1.4	−0.5	1.3	1.1	0.2	−0.6	−2.2
	AsH_2	217.8	−0.8	−0.8	−1.6	−0.1	1.8	1.7	0.8	0.2	−1.5
	SeH	227.0	0.1	−0.1	−0.3	1.7	2.5	3.5	5.3	3.4	2.6
	SeH_2	228.0	−0.3	−0.4	0.3	1.4	2.3	2.8	4.3	2.9	2.1
	HBr	268.9	0.8	0.8	−3.0	−1.9	−1.3	−0.5	0.3	−0.5	−1.4
	Br_2	242.6	−0.2	−0.2	−0.8	0.2	1.6	2.0	3.8	5.5	3.7
	HOBr	245.3	−0.4	−0.6	2.1	3.5	5.0	5.6	6.9	7.8	6.7
	BrF	271.7	0.7	0.8	−1.9	−0.4	0.8	1.0	3.0	3.8	2.8
	NaBr	191.6	−4.9	−5.1	−8.8	−6.9	−5.3	−4.8	−4.7	−5.2	−7.0
EA	K	11.5	−2.6	−3.0	−0.6	−2.7	−2.8	3.2	0.1	−2.4	−0.9
	Ge	28.4	−0.5	−0.4	−2.1	−0.3	2.1	0.1	0.4	−1.0	−3.4
	Br	77.6	−0.5	−0.9	−4.4	−2.9	−1.8	−0.8	−0.6	−1.5	−2.7
	BrO	54.4	−1.3	−1.2	−1.4	2.0	3.8	4.5	5.0	5.9	4.7

TABLE II.   (*Continued.*)

	Species	Expt.[a] (kcal/mol)	Deviation (Expt.-Theor.) (kcal/mol)								
			G3	G3X	B3LYP	B98	B97-2	B1B95	O3LYP	VSXC	TPSS
	SeH	51.0	−0.4	−0.5	−3.2	−1.8	−0.7	0.3	1.0	−0.6	−1.6
PA	Br⁻	322.6	−0.3	0.3	1.1	−0.4	−2.4	−0.7	−2.4	−2.5	−1.8
	CH₃Br	157.3	0.4	0.4	−1.0	−2.0	−3.4	−1.3	−3.5	−4.5	−3.7
MAD[h]			1.14	1.11	2.76	2.50	3.40	2.87	2.75	3.42	5.41

[a]For experimental references and uncertainties see Refs. 20, 53, and 54, except for ΔH_f values which are from Ref. 21.
[b]Vinyl bromide.
[c]2-Bromopropane.
[d]Bromobenzene.
[e]1-Bromohexane.
[f]1,2-Dibromopropane-*trans*.
[g]1,2-Dibromocyclopentane axial.
[h]Mean absolute deviation.

test set appeared.[5,6] In that study we examined seven functionals. In this investigation we have included 16 additional functionals for a total of 23. The functionals fall into three different general categories. The first category contains nonempirical functionals that are constructed to satisfy several exact constraints on the spin density. In addition to the local spin-density approximation (LSDA), this includes nonempirical generalized gradient approximation (GGA) functionals that depend on the local density and its gradient as well as meta-GGA functionals that also depend on the kinetic-energy density for the occupied Kohn-Sham orbitals. The three functionals that we have investigated in this category are PW91, PBE, and TPSS. PW91 is the Perdew and Wang's 1991 GGA functional.[25] PBE is Perdew, Burke, and Ernzerhof's 1996 GGA exchange and correlation functional.[26] TPSS is the nonempirical meta-GGA functional developed by Staroerov *et al.*[27] The second category includes functionals based on semiempirical GGA or meta-GGA. These functionals include one or more parameters that are obtained by fitting to the experimental data on atoms and, in many cases, on molecules/solids. The five functionals we considered in this category are BLYP,[28,29] BPW91,[29,25] BP86,[29,30] OLYP,[31] and VSXC.[32] The first three use Becke's 1988 exchange functional with three different correlation functionals and

were included in our first study. OLYP is a new functional developed by Handy and Cohen[31] that employs a new optimization of the exchange functional. VSXC is a meta-GGA with 20 parameters.[32] The third category includes hybrid semiempirical functionals. These have additional parametrization based on including exact (HF) exchange into the functional. In this category we have considered the three functionals from our previous study (B3LYP, B3P86, and B3PW91, all using Becke's three-parameter hybrid exchange functional)[33,34] and 11 additional functionals. These include (1) Becke's B98 functional[35] that is a ten-parameter fit to the G2/97 test set, closely related to his previous B97 functional,[36] (2) B97-1, a modification to Becke's B97 functional proposed by Hamprecht *et al.*,[37] (3) B97-2, a modification to Becke's B97 functional proposed by Wilson *et al.*,[38] (4) O3LYP, a three-parameter hybrid of Hoe *et al.*[39] based on OLYP, (5) X3LYP, a three-parameter hybrid of Xu and Goddard[40] designed especially for van der Waals interactions, (6) PBE1PBE (also referred to as PBE0), a one-parameter hybrid based on the PBE functional,[41] (7) TPSSh, a one-parameter hybrid of the TPSS functional,[27] (8) B1B95, a one-parameter hybrid functional using Becke's B88 exchange and his 1995 kinetic-energy-dependent correlation functional,[42] (9) mPW1PW91, a one-parameter hybrid func-

TABLE III. Deviations of different theoretical methods from the experiment of the six hydrogen-bonded complexes in the G3/05 test set. [Density-functional results based on single-point 6-311+G(3*df*,2*p*) energies at MP2/6-31G* geometries with scaled (0.89) HF/6-31G* zero-point energies. Values in parentheses are at B3LYP/6-31+G(2*df*,*p*) geometries and scaled 0.9854 B3LYP/6-31+G(2*df*,*p*) zero-point energies.]

Species	Expt.	Deviation (Expt.-Theor.) (kcal/mol)								
		G3	G3X	B3LYP	B98	O3LYP	VSXC	TPSS	OLYP	X3LYP
(H₂O)₂	−3.6±0.3[a]	−0.3(−0.1)	−1.9(−0.1)	−0.7(−0.6)	−0.4(−0.4)	−2.2(−2.0)	−0.8(−0.7)	−0.6(−0.5)	−2.7(−2.6)	−0.3(−0.2)
(CH₃OH)₂	−3.5±0.2[a]	0.5(0.8)	0.4(0.8)	−1.2(−0.3)	−0.8(−0.1)	−3.2(−1.9)	1.4(−0.8)	−0.9(−0.2)	−3.9(−2.4)	−0.6(0.1)
(CH₃COCH₃)₂	−3.2±0.4[a]	−1.0(−1.1)	−1.1(−1.1)	−2.7(−2.6)	−2.1(−2.1)	−4.3(−3.9)	1.1(0.8)	−2.7(−2.7)	−5.0(−4.5)	−2.2(−2.2)
(HCl)₂	−1.23±0.1[b]	0.1(−0.1)	−0.1(−0.1)	−0.4(−0.6)	0.0(−0.2)	−1.1(−1.3)	0.6(0.3)	−0.2(−0.4)	−1.4(−1.6)	−0.2(−0.3)
(CH₃COOH)₂	−14.9[c]	0.3(0.7)	0.7(0.7)	−1.0(−0.5)	−0.6(−0.2)	−5.2(−5.0)	−0.3(−0.7)	−0.7(0.1)	−6.8(−6.7)	−0.2(0.2)
(HF)₂	−2.97±0.1[b]	−1.4(0.2)	−2.1(0.2)	−1.3(0.1)	−1.2(0.1)	−3.1(−1.3)	0.7(−0.2)	−1.2(0.0)	−3.8(−1.8)	−0.8(0.4)
MAD		0.60(0.49)	1.06(0.49)	1.19(0.79)	0.85(0.49)	3.17(2.59)	0.80(0.59)	1.05(0.66)	3.93(3.26)	0.70(0.58)

[a]Reference 55. (H₂O)₂: ΔH_f^0 measured over the range of 360–385 K. (CH₃OH)₂: ΔH_r^0 measured over the range of 305–335 K. (CH₃COCH₃)₂: ΔH_f^0 measured over the range of 340–380 K.
[b]Reference 56. D_0 value.
[c]Reference 57. ΔH_f^0 at 373 K.

TABLE IV. Mean absolute deviations (kcal/mol) of different methods for different types of energies in the G3/05 test set. [Density-functional results based on single-point 6-311+G($3df,2p$) energies at MP2/6-31G* geometries with scaled (0.89) HF/6-31G* zero-point energies.]

	Mean absolute deviation (kcal/mol)								
	G3	G3X	B3LYP	B98	B97-2	B1B95	O3LYP	VSXC	TPSS
Enthalpies of formation (270)	1.19	1.01	4.63	3.46	4.52	3.66	3.94	3.65	6.70
Nonhydrogens (79)[a]	2.10	1.63	5.77	4.20	5.90	5.12	5.75	5.13	6.57
Hydrocarbons (38)	0.69	0.56	5.40	4.87	4.73	3.63	3.43	3.29	7.12
Subst. hydrocarbons (100)	0.82	0.79	4.52	3.16	4.48	3.37	2.87	3.03	6.75
Inorganic hydrides (19)	0.95	0.86	2.16	1.23	1.92	1.79	1.99	2.78	5.61
Radicals (34)	0.83	0.75	2.87	2.28	2.68	2.20	4.60	2.91	7.03
Ionization energies (105)	1.10	1.04	3.83	3.59	3.86	3.93	4.45	4.62	4.85
Electron affinities (63)	0.98	0.99	2.99	2.73	3.29	3.56	3.09	3.06	3.18
Proton affinities (10)	1.14	1.04	1.39	1.07	2.25	1.32	1.76	1.96	1.94
Hydrogen-bonded complexes (6)	0.60	1.06	1.19	0.85	1.79	1.77	3.17	0.80	1.05
All (454)	1.13	1.01	4.11	3.30	4.11	3.63	3.88	3.71	5.61

[a]The 32 new nonhydrogens added to the G3/05 test set include the 14 molecules in Table I and 18 molecules in Table II (K_2, KBr, KCl, KF, GaCl, GeO, As_2, BrCl, BrF, BrO, Br_2, BBr, NaBr, GeS_2, KrF_2, CF_3Br, CCl_3Br, and $COBr_2$).

tional developed by Adamo and Barone[43] using their modified Perdew-Wang exchange functional and the PW91 correlation functional, (10) MPW1B95, a functional proposed by Zhao and Truhlar[44] that replaces the PW91 correlational functional in mPW1PW91 with B95 and reoptimizes the parameter, and (11) the KMLYP functional of Kang and Musgrave,[45] a two-parameter hybrid optimized for the energy of a hydrogen atom and the electron affinity of an oxygen atom. We also have investigated a modified KMLYP version[45] that includes G3-like higher-level correction parameters based on number of unpaired and lone pair/pi electrons in the atoms or molecules, optimized for atomization energies. Among these, PBE1PBE can also be classified as a nonempirical functional since the fraction of exact exchange (0.25) can be justified based on fundamental considerations. For completeness, we have included the nonempirical LSDA functional, SVWN.[46]

B. Results and discussion

1. Comparison to experiment

We have used the G2/97 test set of 301 energies for an initial assessment on the 23 density functionals described above and we then chose eleven of the functionals for further assessment on the full G3/05 test set. The mean absolute deviations for all of the test sets are given in Table V. The detailed results for seven representative functionals for the new energies in the G3/05 test set are given in Tables I–III and a breakdown of the performance of these functionals on the whole set is given in Table IV.

For consistency, we have used the same basis set [6-311+G($3df,2p$)], geometries (MP2/6-31G*), zero-point energies (scaled HF/6-31G*), and thermal corrections (scaled HF/6-31G*) that we used in previous assessments and have not included any spin-orbit corrections or relativistic effects, except where noted. We have previously investigated the dependence on the geometry and zero-point energy for the B3LYP functional.[7,17] We found that the choice of the geometry (MP2 versus B3LYP) had little effect on the mean absolute deviation for the G2/97 test set. We found that there

was a more significant dependence on choice of zero-point energy (0.4 kcal/mol increase in the mean absolute deviation) if B3LYP/6-31G* with a scale factor of 0.98 is used instead of HF/6-31G* with a scale factor of 0.89. Although the use of B3LYP/6-31G* with the 0.98 scale factor is more appropriate for zero-point energies,[47] the slight increase in the B3LYP mean absolute deviation is probably because the B3LYP parametrization was implicitly consistent with the scaled HF zero-point energies.[12]

Because of the dependence on choice of zero-point energy found in Ref. 17, we have examined the effect of the use of the B3LYP zero-point energies on the results of the some of the density-functional methods. We also examined the effect of inclusion of atomic spin-orbit effects. The mean absolute deviations including the B3LYP zero-point energies and spin-orbit effects are given in parentheses in Table V. For the hybrid methods, with the exception of MPW1B95, all of the mean absolute deviations increase when spin-orbit and the B3LYP zero-point energies are included. The largest increases are found for B3LYP and X3LYP, which is probably due to the fact that they were implicitly optimized for the HF zero-point energies. The largest increase (~0.8 kcal/mol) for these two hybrids occur for the larger G3/99 and G3/05 test sets. In contrast, the MPW1B95 functional shows significant improvement (0.89 kcal/mol) when the spin-orbit and scaled B3LYP zero-point energies are included because Zhao and Truhlar[44] optimized it on D_e values, i.e., with zero-point energies subtracted and with spin-orbit corrections included. The TPSS functional does not involve empirical parameters and also has improved agreement with the experiment (by about 1 kcal/mol) with the use of the scaled B3LYP zero-point energies and spin-orbit effects. In the following discussions, we refer to mean absolute deviations calculated using the scaled HF zero-point energies, except for MPW1B95 and TPSS.

The results for the 23 functionals tested on the G2/97 test set (Table V) indicate that the B98 functional has the smallest mean absolute deviation (2.90 kcal/mol). This is the only functional that has a significantly smaller mean ab-

TABLE V. Mean absolute deviations of density-functional methods. [Density-functional results based on single-point 6-311+G($3d2f,2p$) energies at MP2/6-31G* geometries with scaled (0.89) HF/6-31G* zero-point energies. Values in parentheses are based on B3LYP/6-31G($2df$) zero-point energies (scaled by 0.9854) and spin-orbit corrections (Ref. 8) for the atoms.]

		MAD (kcal/mol)		
		G2/97	G3/99	G3/05
LSDA	SVWN	52.20	84.68	
Nonempirical GGA	PW91	10.87		
	PBE	10.44		
	TPSS	5.30 (4.86)	5.58 (4.76)	5.67 (4.71)
Semiempirical GGAs	BLYP	5.78	7.32	
	BPW91	5.90		
	BP86	12.24		
	OLYP	4.78	5.41	5.55
	VSXC	3.36 (3.35)	3.58 (3.70)	3.75 (3.80)
Hybrid semiempirical GGAs	B3LYP	3.31 (3.60)	4.27 (5.01)	4.14 (4.97)
	B3PW91	3.65 (3.44)		
	B3P86	15.39 (14.84)		
	B98	2.90 (3.12)	3.39 (3.95)	3.33 (3.81)
	B97-1	3.48 (3.59)	4.08 (4.37)	4.10 (4.32)
	B97-2	3.50 (3.52)	4.09 (4.25)	4.14 (4.23)
	O3LYP	4.08 (3.81)	3.95 (3.96)	3.89 (4.03)
	X3LYP	3.26 (3.54)	3.96 (4.69)	3.80 (4.64)
	PBE1	4.66 (4.35)		
	TPSSh	4.54 (4.34)		
	B1B95	3.35 (3.36)	3.67 (3.85)	3.66 (3.77)
	mPW1PW91	3.86 (4.06)		
	MPW1B95	3.98 (3.55)	4.88 (4.08)	4.85 (3.96)
	KMYLP	20.55 [3.66][a]		

[a]KMLYP results in square brackets include a higher-level correction based on three parameters (number of unpaired electrons for atoms and molecules, lone electron pairs for atoms, and lone electron pairs or pi electrons for molecules. We have reoptimized the three parameters for the enthalpies in the G2/97 test set (the values are 2.33, 14.89, and 15.50 mhartree, respectively).

solute deviation than B3LYP (3.31 kcal/mol). Several other functionals have mean absolute deviations of a similar magnitude as B3LYP. Among these are VSCX (3.36 kcal/mol), MPW1B95 (3.55 kcal/mol), B1B95 (3.35 kcal/mol), B97-1 (3.48 kcal/mol), B97-2 (3.50 kcal/mol), and X3LYP (3.26 kcal/mol). The X3LYP functional is very close in accuracy to B3LYP because it is a combination of B3LYP and B3PW91. The KMLYP functional, which has been shown to be very good for energy barriers, does very poorly with a mean absolute deviation of 20.55 kcal/mol. The poor performance is probably due to the method used for obtaining the exchange and correlation mixing parameters, i.e., from the H-atom energy and the O-atom electron affinity. We also examined the effect of including higher-level correction parameters as in the modified version of KMLYP.[45] We optimized them for the G2/97 set of enthalpies and found that this greatly improves the accuracy (3.66 kcal/mol), although it is still not as good as B3LYP. The most accurate nonempirical functional is TPSS with a mean absolute deviation of 4.86 kcal/mol, compared to 10.44 and 10.87 for the other two functionals PBE and PW91, respectively.

Eleven of the functionals were further assessed on the full G3/05 test set and the results are given in Table V. The B98 functional has the smallest mean absolute deviation of the 11 functionals (3.33 kcal/mol). The B3LYP functional

does poorly for larger molecules and, thus, its mean absolute deviation increases to 4.14 kcal/mol. The other hybrid functionals that we investigated have similar or smaller mean absolute deviations than B3LYP on the full test set with values ranging from 3.66 kcal/mol (B1B95) to 4.14 kcal/mol (B97-2). The nonempirical functional TPSS has a mean absolute deviation of 4.71 kcal/mol while the VSXC functional has a deviation of 3.75 kcal/mol. As in the case of the Gn methods the density-functional methods have larger errors for the nonhydrogen species although the increase is not nearly as dramatic (see Table IV). It is notable that the density-functional methods perform well for the third-row systems with mean absolute deviations of 2.5–3.5 kcal/mol on this subset, even though most were parametrized on the first- and second-row molecules.

Finally, the density-functional theory (DFT) methods perform about as well as the G3 methods for hydrogen-bonded complexes. The inclusion of diffuse functions in the basis sets corrects the problems in the density-functional geometries of the water dimer and the hydrogen fluoride dimer. The binding-energy results based on B3LYP/6-31+G($2df,p$) optimized geometries are given in parentheses in Table III. The use of this geometry gives very good agreement with the experiment for these two dimers for all of the density-functional methods, except for O3LYP and OLYP,

TABLE VI. Size dependence of the G3 and DFT methods for enthalpies of formation in G3/05 test set. [Density-functional results based on single-point 6-311+G($3d2f,2p$) energies at MP2/6-31G* geometries with scaled (0.89) HF/6-31G* zero-point energies.]

No. of pairs of valence electrons in molecule	No. of molecules in subset	MAD (kcal/mol)								
		G3	G3X	B3LYP	B98	B97-2	B1B95	O3LYP	VSXC	TPSS
≤ 8	101	0.99	1.00	2.25	2.05	2.69	2.32	3.27	2.99	5.18
$>8 \leq 14$	89	1.00	0.83	3.40	3.10	3.96	3.46	4.39	3.11	7.80
>14	80	1.64	1.22	9.01	5.64	7.46	5.58	4.29	5.08	7.42

and improves the mean absolute deviation to <1 kcal/mol. The O3LYP and OLYP functionals have substantial errors for hydrogen-bonded complexes.

2. Size dependence

Errors in quantum-mechanical methods should, in general, grow with the size of the system. Since the enthalpies of formation are computed from the calculated atomization energies, a roughly linear dependence on the size of the system can be expected, particularly in a sequence of molecules containing similar kinds of bonds (e.g., alkane chains). In this analysis, we consider the number of valence electrons in the system as a rough measure of the size of the molecule. The dependence of the G3 and DFT methods on the number of electrons in the molecule is shown in Table VI where mean absolute deviations are given for enthalpies of formation in the G3/05 test set for three sets of molecules: those containing <8 pairs of valence electrons, >8≤14 pairs valence electrons, and >14 pairs of valence electrons. The separation into the different groups is somewhat arbitrary, but nevertheless it shows interesting trends. Some density-functional methods show significantly increasing errors with molecule size. The B3LYP functional shows a particularly strong dependence on the number of electrons as its error increases dramatically from 2.25 to 3.40 to 9.01 kcal/mol for the three subsets. The error as a function of number of electrons is plotted in Fig. 1 and indicates an approximate linear dependence with number of valence electrons. With the exception of O3LYP the other functionals also show an increase in error with number of electrons, although not as large as B3LYP. The O3LYP functional is an exception in that it shows little dependence on size. The reason for this is not clear. Both G3X and G3 theories do not depend as much on
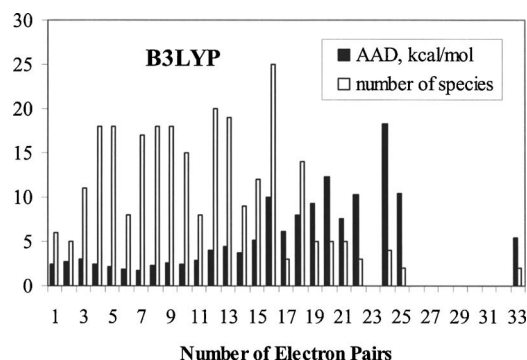

FIG. 1. Mean absolute deviation from the experiment as a function of number of pairs of valence electrons for B3LYP on the enthalpies of formation in the G3/99 test set.

the number of electrons as the DFT methods with G3X having the smallest dependence. The mean absolute deviation of G3X stays about level at 1.00, 0.83, and 1.22 for the three subsets whereas G3 shows a larger increase for the last subset to 1.64 kcal/mol. In the case of the G3 methods the errors tend to be more systematic than in the DFT methods so that the higher-level correction (HLC) used in the G3 methods is able to correct for the size dependence. We have applied the HLC to the DFT methods and, while, this provides for some improvement it does not do nearly as well as for the G3 methods, apparently because the errors are not systematic for the DFT methods.

The dependence on molecular size was further investigated for seven representative DFT methods on a set of larger molecules with accurate experimental energies (quoted uncertainty of less than 1 kcal/mol). The molecules in this set have 56 or more valence electrons and thus are larger than the molecules in the G3/05 test set and most are not feasible for the G3 methods. The 15 molecules in the set include 2 linear alkanes (dodecane and hexadecane), 2 fluorinated hydrocarbons (C_6F_6 and C_6F_5Cl), 3 two-dimensional aromatics (fluoranthane, anthracene, and pyrene), 2 three-dimensional sp^3 hydrocarbons (adamantane and diadamantane), and 5 linked aromatic rings. The deviations from the experiment for seven representative functionals are given in Table VII for these 15 large molecules along with mean absolute deviations. The results are consistent with the trends found on the G3/05 test set in that they indicate significantly poorer performance on molecules with large numbers of electrons. The B3LYP method has one of the largest mean absolute deviation (21.2 kcal/mol), which is more than double its error (9.01 kcal/mol) for the third subset of molecules (>28 valence electrons) in the G3/05 test set. The B3LYP error is especially large for large sp^3 hydrocarbons and contributes to its poor performance in our larger test sets. Of special note is the error in the three-dimensional cyclic sp^3 hydrocarbons diadamantane and adamantane, which have errors of −51.0 and −32.9 kcal/mol, respectively. This is significantly greater than the error in the linear alkanes of similar size, dodecane (−20.9 kcal/mol) and hexadecane (−30.2 kcal/mol). The difference in performance for the linear and cyclic alkanes is due to the higher proportion of C–C bonds (relative to C–H bonds) in the latter, and the relatively poor description of the C–C bonds in B3LYP. Similar differences can be seen for some of the other methods such as TPSS and O3LYP. The B3LYP functional also does poorly for larger aromatics such as pyrene (−24.8 kcal/mol) and fluoranthene (−23.6). Three of the

TABLE VII. Assessment on very large molecules (>56 valence electrons).

Molecule	Expt.[a] $\Delta H_f(298)$	Deviation (Expt.-Theor.) (kcal/mol)						
		B3LYP	B98	B97-2	B1B95	O3LYP	VSXC	TPSS
C_6F_6	−228.35	−2.90	16.11	30.54	24.39	6.42	22.17	20.29
C_6F_5Cl	−194.14	−7.85	12.34	27.87	22.57	1.71	19.80	16.52
$C_{12}H_{26}$ dodecane	−69.24	−20.89	−18.93	−15.72	−12.35	−0.84	−11.58	14.06
$C_{16}H_{34}$ hexadecane	−89.58	−30.25	−26.37	−21.62	−17.46	−4.30	−16.66	15.92
$C_{10}H_{16}$ adamantane	−32.17	−32.93	−21.18	−11.71	−6.93	−17.27	−4.99	−4.39
$C_{14}H_{20}$ diadamantane	−34.87	−51.01	−30.18	−14.30	−6.58	−30.12	−1.72	−11.58
$C_{16}H_{10}$ pyrene	53.94	−24.77	−4.63	25.34	15.79	−11.31	14.40	4.75
$C_{16}H_{10}$ fluoranthene	69.07	−23.63	−3.16	27.34	18.64	−7.84	15.06	7.25
$C_{14}H_{10}$ anthracene	55.19	−18.74	−3.26	21.35	13.22	−6.44	13.74	7.68
$C_{12}H_8N_2$ phenazine	82.17	−7.75	6.00	28.11	17.45	−2.21	10.08	15.76
$C_{12}H_{10}N_2$ trans-azobenzene	96.58	−9.69	0.44	20.47	10.85	−4.49	9.15	15.51
$C_{13}H_{10}O$ benzophenone	13.12	−17.87	−3.13	17.90	10.84	−7.58	18.04	6.78
$C_{12}H_8S$ dibenzothiophene	49.02	−22.87	−5.33	19.17	13.72	−9.16	8.96	4.90
$C_{16}H_{12}S_2$ 3,6-diphenyl-1,2-dithiin	100.96	−25.49	−1.20	28.13	22.43	−10.95	26.51	14.88
MAD (kcal/mol)		21.19	10.88	22.11	15.23	8.62	13.78	11.45

[a]From Ref. 20. All experimental uncertainties are less than 1 kcal/mol. Density-functional results based on single-point 6-311+G($3d2f,2p$) energies at MP2/6-31G* geometries with scaled (0.89) HF/6-31G* zero-point energies.

density functionals have smaller mean absolute deviations: B98 (10.88 kcal/mol), O3LYP (8.62 kcal/mol), and TPSS (11.45). However, in each case substantial errors can be seen for specific cases. This assessment on very large molecules indicates that all of the density functionals have significant deficiencies with certain types of molecules.

3. Basis set dependence

We have investigated the basis set dependence of 6 selected functionals in this work of the 147 enthalpies of formation in the G2/97 test set. We have not investigated ion-ization potentials and electron affinities since in previous work[6] we found little dependence on basis sets for these quantities. The results are given in Table VIII for the B3LYP, B98, O3LYP, MPW1B95, BLYP, and TPSS functionals, which represent a range of functional types. The results show a quite surprising dependence on the polarization functions used in the basis sets. In some cases expanding the polarization set significantly improves agreement with the experiment (B3LYP and B98) and in some cases it makes agreement significantly worse (MPW1B95, O3LYP, and TPSS). This can be seen in the series of calculations done with the

TABLE VIII. Basis set dependence of density-functional methods for the 147 enthalpies of formation in the G2/97 test set. [B3LYP, B98, O3LYP, and BLYP results are based on 0.8929 scaled HF ZPEs with no atomic spin-orbit correction. All methods are based on MP2(full)/6-31G* geometries. The numbers in parentheses for MPW1B95 and TPSS results are based on 0.9854 scaled B3LYP/6-31G* ZPEs and atomic spin-orbit correction.]

Basis set[a]	Polarization set[b]	Mean absolute deviations, kcal/mol					
		B3LYP	B98	O3LYP	MPW1B95	TPSS	BLYP
6-31G(d)	d,d	5.31	5.26	6.27	6.04(5.71)	7.31(6.80)	7.80
6-31+G(d)	d,d	8.66	7.93	5.57	4.49(4.99)	5.41(5.41)	9.44
6-31+G($2df,p$)	$2df,2df,p$	3.97	2.69	5.89	5.34(4.49)	7.38(6.66)	6.63
6311+G(d,p)	d,d,p	8.22	7.49	5.49	4.61(5.29)	4.72(4.70)	8.89
6311+G($2df,p$)	$2df,2df,p$	3.86	2.95	3.53	3.44(2.95)	5.40(4.67)	7.35
6311+G($3df,2p$)	$3df,3df,2p$	3.04	2.50	4.20	4.34(3.48)	6.37(5.49)	7.12
6311+G($3d2f,2pd$)	$3d2f,3d2f,2pd$	2.88	2.25	4.38	4.53(3.59)	6.55(5.64)	6.92
G3MP2Large[c]	$3df2f,2df,p$	3.41	2.51	3.54	3.66(2.91)	5.69(4.96)	7.03
G3Large	$3d2f,2df,p$	3.22	2.37	3.80	3.91(3.08)	6.03(5.25)	6.99
G3XLarge	$3df2fg,2df,p$	3.14	2.32	3.77	⋯	⋯	⋯
aug-cc-pVDZ	d,d,p	11.17	10.63	7.11	⋯	⋯	⋯
aug-cc-pVTZ	$2df,2df,2pd$	4.07	3.32	3.64	⋯	⋯	⋯
aug-cc-pVQZ	$3d2fg,3d2fg,3p2df$	3.26	2.40	4.09	⋯	⋯	⋯

[a]The correlation-consistent basis sets are from Ref. 48. The G3 basis sets are from Refs. 8 and 10.
[b]Order is second row, first row, and H polarization functions.
[c]Does not have core-polarization functions that are included in the G3Large and G3XLarge basis sets.
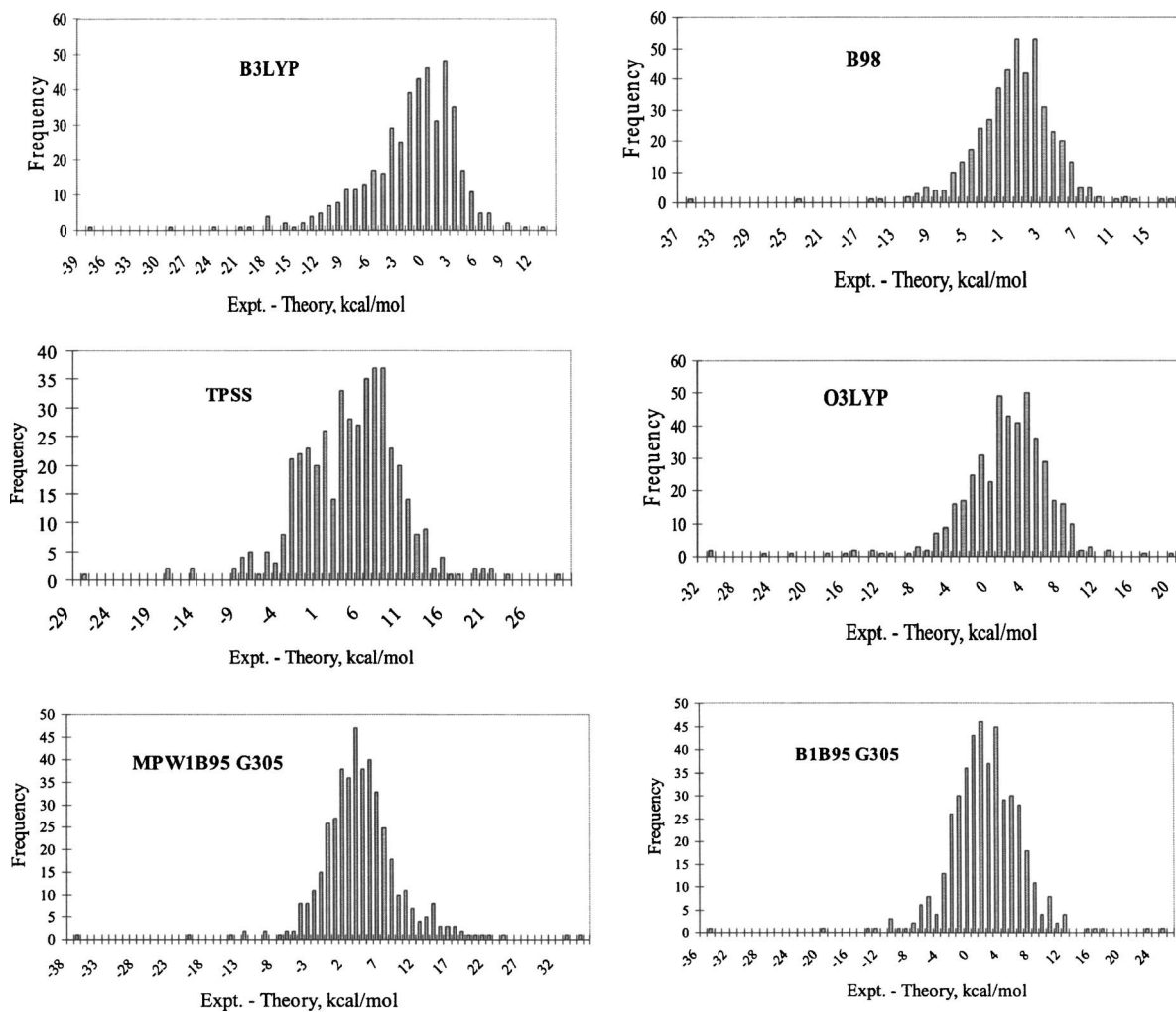
FIG. 2. Distribution of errors in kcal/mol for B3LYP, B98, O3LYP, TPSS, B1B95, and MPW1B95 density-functional methods for the G3/05 test set. The mean and standard deviations in kcal/mol are B3LYP (−2.47 and 5.51), B98 (−0.74 and 4.7), TPSS (3.59 and 6.13), O3LYP (1.09 and 5.29), MPW1B95 (3.23 and 6.01), and B1B95 (1.09 and 4.98).

6-311G basis set with different sets of polarization functions. For example, increasing the polarization set from $2df,p$ to $3d2f,2pd$ decreases the mean absolute deviation by ~1 kcal/mol for B3LYP, while it increases it by about the same amount for MPW1B95. The latter increase occurs because the MPW1B95 functional was optimized for a modified G3MP2Large basis set that has a $2df$ polarization set on the first-row elements. The large effect of the polarization functions on the mean absolute deviations is also noted in the correlation-consistent basis sets[48] where the aug-cc-pVTZ has an error of 4.07 for B3LYP compared to 3.31 for 6-311+G($3df,2p$). This is apparently due to use of a $2df$ set on the pVTZ basis set. The quadruple zeta basis set, aug-cc-pVQZ with a $3d2fg$ polarization set, has a mean absolute deviation (3.26 kcal/mol) comparable to triple zeta 6-311+G($3df,2p$) for B3LYP. The results in Table VIII also indicate that the use of core-polarization functions in the G3Large basis set has only a small effect on the deviations when compared to G3MP2Large, which does not have core-polarization functions. Also the addition of diffuse functions leads to an increase in error relative to 6-31G* basis set results for B3LYP (from 5.31 to 8.66 kcal/mol) and B98

(from 5.26 to 7.93 kcal/mol), although for other functionals it leads to improved agreement with the experiment.

The effect of increasing the polarization set can be explained in terms of the distribution of errors. From looking at the distribution of errors, it is evident that most of the DFT methods are skewed either in the direction of negative or positive deviations. This is illustrated for several of the functionals in Fig. 2 including B3LYP, B98, O3LYP, MPW1B95, B1B95, and TPSS. For example, B3LYP tends to underbind (more negative deviations from the experiment than positive) and thus increasing the basis set will improve agreement with the experiment. Alternatively TPSS, O3LYP, and MPW1B95 tend to overbind (more positive deviations from the experiment than negative) and thus decreasing the polarization set makes for better agreement with the experiment. The B98 functional has only slightly more negative deviations than positive and is the most equally distributed of the functionals that we investigated. It is noted that the distribution does not shift uniformly as the effect of polarization functions is dependent on the type of molecule. For comparison with the density-functional results, the distribution of errors of the G3 and G3X theories is given in Fig. 3. The
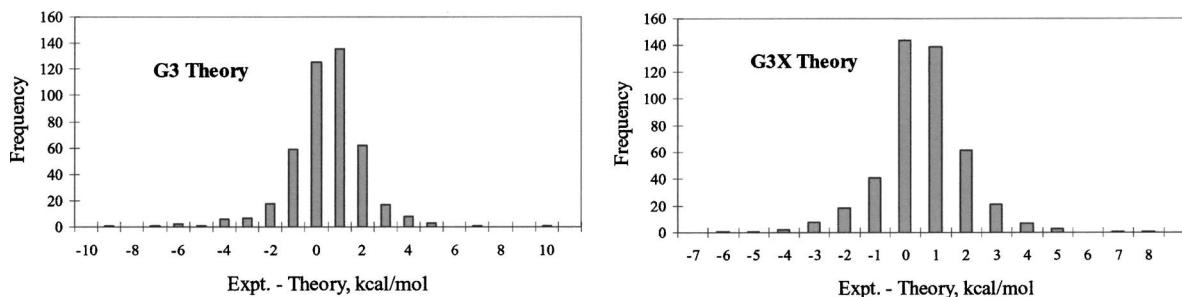
FIG. 3. Distribution of errors in kcal/mol for the G3 and G3X theories for the G3/05 test set.

figure shows that the deviations from experiment for the *ab initio* based methods are quite equally distributed.

There have been a number of studies on the dependence of density-functional methods on basis sets. Based on a study of the HCTH functional, Boese *et al.*[14] concluded that the triple zeta basis set level [e.g., the 6-311G$(3df, 2pd)$ and TZ2P basis sets] is likely to be sufficient for the use in density-functional calculations. Wang and Wilson[49] examined the performance of the correlation-consistent basis sets for a small set of test molecules and found that while increasing the basis set improved performance for the hybrid functionals tested (B3LYP, B3PW91, and B3P86), the performance actually became worse for the nonhybrids tested (BLYP, BP86, and BPW91). Our results for the TPSS functional are consistent with this finding, but show that hybrids do not all show the same dependence. Jensen[50] has developed polarization-consistent (pc) basis sets for the first row for use in extrapolating density-functional calculations. Wang and Wilson[51] concluded that these basis sets give only a small improvement over correlation-consistent basis sets for six different functionals that they examined.

A number of general observations can be made from our basis set study of density functionals. First, the distribution of errors of the density functionals can lead to some surprisingly large basis set dependences. Second, improvement in the basis sets does not always lead to improvement in results as is the case with most *ab initio* methods. This is true not only for the heavily parametrized methods, but also for those with few parameters. Finally, the use of a smaller basis set size can improve agreement with the experiment if the reduction of binding improves the distribution of errors.

V. CONCLUSIONS

In this paper we presented an expansion of the G3/99 test set of thermochemical data to include 78 additional energies. This new set, referred to as the G3/05 test set, has 454 total energies including 14 new enthalpies of formation of the first- and second-row nonhydrogen molecules, 58 energies of molecules containing the third-row elements K, Ca, and Ga-Kr, and 6 hydrogen-bonded complexes. The following conclusions can be drawn from this assessment of Gn and DFT methods.

(1)   The G3 and G3X theories have an accuracy of about 1 kcal/mol for the third-row molecules and hydrogen-bonded complexes that are added to the test set. As has been found previously, these methods do not perform as

well for the nonhydrogen species. For the subset of 79 nonhydrogens in the G3/05 test, G3 and G3X have mean absolute deviations of 2.05 and 1.60 kcal/mol, respectively.

(2)   On the full G3/05 test set, the G3X theory is an improvement over G3 theory with a mean absolute deviation of 1.01 kcal/mol compared to 1.13 kcal/mol, respectively.

(3)   Twenty-three density functionals were tested. The most accurate functional was found to be the B98 functional with a mean absolute deviation of 3.33 kcal/mol, compared to 4.14 kcal/mol for B3LYP. The B98 functional does better for larger molecules where B3LYP tends to have much larger errors. For smaller molecules B3LYP does as well or better than B98 and the other functionals.

(4)   It is found that for most of the density-functional methods the errors increase significantly with increasing number of electrons. The B3LYP method is notable in this respect with an error of 9 kcal/mol for molecules having 28 or more valence electrons.

(5)   As in the case of the Gn methods the density-functional methods have the larger errors for the nonhydrogen species although the increase is not nearly as dramatic. The density-functional methods, including B3LYP, perform well for the third-row systems even though most were parametrized on the first- and second-row molecules.

(6)   The performance of density-functional methods are found to be quite sensitive to the number of polarization functions used on a triple zeta basis set and in many cases functionals are found to decrease in accuracy with increasing basis set size.

[1] In *Quantum-Mechanical Prediction of Thermochemical Data*, Understanding Chemical Reactivity Vol. 22, edited by J. Cioslowski (Kluwer Academic, Dordrecht, 2001).

[2] In *Computational Thermochemistry*, ACS Symposium Series 677, edited by K. K. Irikura and D. J. Frurip (American Chemical Society, Washington DC, 1998).

[3] In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, edited by S. R. Langhoff (Kluwer Academic, Dordrecht, 1995).

[4] L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991).

[5] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **106**, 1063 (1997).

[6] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **109**, 42 (1998).

[7] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **112**, 7374 (2000).

[8] L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, and J. A. Pople, J. Chem. Phys. **109**, 7764 (1998).

[9] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **114**, 108 (2001).

[10] L. A. Curtiss, P. C. Redfern, K. Raghavachari, V. Rassolov, and J. A. Pople, J. Chem. Phys. **110**, 4703 (1999).

[11] A. G. Baboul, L. A. Curtiss, P. C. Redfern, and K. Raghavachari, J. Chem. Phys. **110**, 7650 (1999).

[12] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, Chem. Phys. Lett. **359**, 390 (2002).

[13] A. D. Boese and N. C. Handy, J. Chem. Phys. **114**, 5497 (2001).

[14] A. D. Boese, J. M. L. Martin, and N. C. Handy, J. Chem. Phys. **119**, 3005 (2003).

[15] J. Ciolowski, M. Schmieczek, G. Liu, and V. Stoyanov, J. Chem. Phys. **113**, 9377 (2000).

[16] B. J. Lynch, Y. Zhao, and D. G. Truhlar, J. Phys. Chem. A **107**, 1384 (2003); B. J. Lynch and D. G. Truhlar, *ibid.* **107**, 3898 (2003); **107**, 8996 (2003).

[17] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, Chem. Phys. Lett. **270**, 419 (1997).

[18] P. C. Redfern, P. Zapol, L. A. Curtiss, and K. Raghavchari, J. Phys. Chem. A **104**, 5850 (2000).

[19] M. 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 Suppl. **14**, 1 (1985).

[20] L. A. Curtiss, P. C. Redfern, V. Rassolov, G. Kedziora, and J. A. Pople, J. Chem. Phys. **114**, 9287 (2001).

[21] J. B. Pedley, R. D. Naylor, and S. P. Kirby, *Thermochemical Data of Organic Compounds*, 2nd ed. (Chapman and Hall, New York, 1986).

[22] The geometries and zero-point energies are available at the website chemistry.anl.gov/compmat/comptherm.htm

[23] A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., J. Chem. Phys. **110**, 7667 (1999).

[24] B. Ruscic, R. E. Pinzon, M. L. Morton *et al.*, J. Phys. Chem. A **108**, 9979 (2004).

[25] J. P. Perdew, in *Electronic Structure of Solids*, edited by P. Zieche and H. Eschrig (Akademie, Berlin, 1991); J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992); J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[26] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); **78**, 1396 (1997).

[27] V. N. Staroerov, G. E. Scuseria, J. Tao, and J. P. Perdew, J. Chem. Phys. **119**, 12129 (2003).

[28] C. Lee, C. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[29] A. D. Becke, Phys. Rev. A **38**, 3098 (1988); C. Lee, C. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[30] J. P. Perdew, Phys. Rev. B **33**, 8822 (1986); **34**, 7406 (1986).

[31] N. C. Handy and A. J. Cohen, Mol. Phys. **99**, 403 (2001).

[32] T. Van Voorhis and G. E. Scuseria, J. Chem. Phys. **109**, 400 (2001).

[33] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[34] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).

[35] H. L. Schmider and A. D. Becke, J. Chem. Phys. **108**, 9624 (1998).

[36] A. D. Becke, J. Chem. Phys. **107**, 8554 (1997).

[37] F. A. Hamprecht, A. J. Cohen, D. J. Tozer, and N. C. Handy, J. Chem. Phys. **109**, 6264 (1998).

[38] P. J. Wilson, T. J. Bradley, and D. J. Tozer, J. Chem. Phys. **115**, 9233 (2001).

[39] W.-M. Hoe, A. J. Cohen, and N. C. Handy, Chem. Phys. Lett. **341**, 319 (2001).

[40] X. Xu and W. A. Goddard III, Proc. Natl. Acad. Sci. U.S.A. **101**, 2673 (2004).

[41] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996); M. Ernerhof and G. E. Scuseria, *ibid.* **110**, 5029 (1999); C. Adamo and V. Barone, *ibid.* **110**, 6158 (1999).

[42] A. Becke, J. Chem. Phys. **104**, 1040 (1996).

[43] C. Adamo and V. Barone, J. Chem. Phys. **108**, 664 (1998).

[44] Y. Zhao and D. G. Truhlar, J. Phys. Chem. A **108**, 6908 (2004).

[45] J. K. Kang and C. B. Musgrave, J. Chem. Phys. **115**, 11040 (2001).

[46] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980); J. C. Slater, *The Self-Consistent Field for Molecules and Solids*, Quantum Theory of Molecules and Solids, Vol. 4 (McGraw-Hill, New York, 1974).

[47] A. P. Scott and L. Radom, J. Phys. Chem. **100**, 16502 (1996).

[48] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989); D. E. Woon and T. H. Dunning, Jr., *ibid.* **98**, 1358 (1993); R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *ibid.* **96**, 6796 (1992); K. A. Peterson, D. E. Woon, and T. H. Dunning, Jr., *ibid.* **100**, 7410 (1994); A. Wilson, T. van Mourik, and T. H. Dunning, Jr., J. Mol. Struct.: THEOCHEM **388**, 339 (1997).

[49] N. X. Wang and A. K. Wilson, J. Chem. Phys. **121**, 7632 (2004).

[50] F. Jensen, J. Chem. Phys. **115**, 9113 (2001); **116**, 7372 (2002); Chem. Phys. Lett. **402**, 510 (2005).

[51] N. K. Wang and A. K. Wilson, Mol. Phys. **103**, 345 (2005).

[52] C. E. Fellow, J. Chem. Phys. **94**, 5855 (1991).

[53] L. A. Curtiss, M. P. McGrath, J.-P. Blaudeau, N. E. Davis, R. Binning, and L. Radom, J. Chem. Phys. **103**, 6104 (1995).

[54] J.-P. Blaudeau, M. P. McGrath, L. A. Curtiss, and L. Radom, J. Chem. Phys. **107**, 5016 (1997).

[55] L. A. Curtiss and M. Blander, Chem. Rev. (Washington, D.C.) **88**, 827 (1988).

[56] A. S. Pine and B. J. Howard, J. Chem. Phys. **84**, 590 (1986).

[57] J. Chao and Z. Zwolinski, J. Phys. Chem. Ref. Data **7**, 363 (1978).