Energy Correctors for Accurate Prediction of Molecular Energies

Jorge M. Seminario,* Martha G. Maffei, Luis A. Agapito, and Pablo F. Salazar

Department of Chemical Engineering and Department of Electrical and Computer Engineering, Texas A&M University, College Station, Texas 77843-3122

Received: September 26, 2005; In Final Form: November 21, 2005

Energy correctors are introduced for the calculation of molecular energies of compounds containing first row atoms (Li−F) to modify ab initio molecular orbital calculations of energies to better reproduce experimental results. Four additive correctors are introduced to compensate for the differences in the treatment of molecules with different spin multiplicities and multiplicative correctors are also calculated for the electronic and zero-point vibrational energies. These correctors, individually and collectively yield striking improvements in the atomization energies for several ab initio methods. We use as training set the first row subset of molecules from the G1 basis of molecules; when the correctors are applied to other molecules not included in the training set, selected from the G3 basis, similar improvements in the atomization energies are obtained. The special case of the B3PW91/cc-pVTZ yields an average error of 1.2 kcal/mol, which is already within a chemical accuracy and comparable to the Gaussian-*n* theories accuracy. The very inexpensive B3PW91/6-31G** yields an average error of 2.1 kcal/mol using the correctors. Methods considered unsuitable for energetics such as HF and LSDA yield corrected energies comparable to those obtained with the best highly correlated methods.

I. Introduction

The Gaussian-*n* ($n = 1-3$) compound methods yield excellent accuracy for molecular energies, mostly within the range considered as chemical accuracy, that is, 1−2 kcal/mol. This accuracy is needed as practical applications for the design of new materials and processes require of extremely good energetics. As quantum chemistry methods extend their application to the analysis, design and simulation of nanosized systems (nanotechnology), a size region that is extremely difficult to be approached experimentally, the need for precise calculations is of paramount importance for the development of such new field avoiding trial-and-error experimentation. Gaussian-*n* and other related methods root their success on precise methods requiring computational resources that can only be practically applied to very small systems. In practice, using these methods for molecules larger than benzene becomes prohibited for an installation composed of a few modern workstations.

The Gaussian-1 (G1) method introduced by Pople et al.[1,2] gave an atomization energy accuracy better than 2 kcal/mol for a set of molecules containing only first-row elements (G1 set) and an accuracy better than 3 kcal/mol for the second row molecules.[2] The G2 method developed by the same team, Curtiss et al.,[3] yielded an accuracy of 1.2 kcal/mol for an extended set of 99 cases named the G2 set; this set includes the molecules in the G1 set plus 24 molecules that contain second-row elements. The G3 method improves on the G2 by including new corrections such as spin−orbit correction for atoms and correction for core correlation.[3] This method also improved the enthalpy of formation error from 1.56 to 0.94 kcal/mol, for the same G2/97 set of molecules.[4]

The Gaussian-*n* theories were developed to take advantage of the fact that relatively precise ab initio calculations contain systematic errors with some additive features. The nature of errors with one level of theory is possibly different from the errors with other levels of theory and might be separately estimated. Also, errors due to the finite nature of the basis sets can be decomposed in contributions with respect to the angular moment of the basis functions. Low angular momentum contributions can be found at lower levels of theory (such as MP2) using large basis sets and high angular momentum contributions can be calculated using higher levels of theory using smaller basis sets. In most of the cases, the contributions are additive.[1,2]

Additional corrections in the Gaussian-*n* methods, among others, include high-level corrections of paired and unpaired electrons using fitted parameters that reproduce the experimental energies. For example, the G1 method uses the Hartree−Fock (HF) energy, which is further corrected with the MP2, MP3, MP4SD(T)Q, and QCISD(T) energies. The zero-point energy for a molecule in this method is obtained from a HF optimization using the 6-31G(d) basis and scaled by the standard 0.8929 for such a level of theory but the geometry to be used for energy calculations is from an MP2 optimization. This MP2 geometry is used further for the MP4 and QCI methods, as no other geometry optimizations are performed for higher levels of theory.

The high cost of G1 methods is because the QCI and MP4 methods scale as N^7. This scaling means, for instance, if a molecule takes 1 day of CPU time, to calculate a double sized molecule takes 2^7, i.e., 128 days! Even with this strong restriction, these two methods are still far from chemical accuracy if an extremely good basis set is not used. For example the MP4/6-311G(d,p) and QCI/6-311G(d,p) levels of theory yield errors of 15.4 and 16.7 kcal/mol, respectively, still far from chemical accuracy and the MP4/6-311G(2df,p) yield an average error of 8.3 kcal/mol. As expected, levels of theory such as MP4/6-311G(2df,p) or QCI/6-311G(d,p) cannot be used for precise energetics due to the strong errors in energies they yield, needless to say, for lower levels of theory. For instance, the well-used HF/6-31G(d) yields an average error of 87.3 kcal/mol; although this method formally scales as N^4, modern computational algorithms have reduce this scaling to $\sim N^2$. Since

10.1021/jp055460z CCC: $33.50

the early 90's, several successful nonlocal density functional theory approaches and functionals have been introduced that have reduced the energy errors close to chemical accuracy for relatively large molecules (having more than 50−100 atoms). One of them is the so-called hybrid functional B3PW91, which using a triple-ζ basis yields an average error of 2.7 kcal/mol. The DFT methods evolved from the use of simple LSD functionals, which were a strong improvement over HF methods and comparable to MP2, however, were still far from chemical accuracy. Modern DFT methods include gradient corrections in the functionals, yielding better results than MP4 and CC when the same bases are used. For instance, hybrid functionals such as the B3PW91 performs very close to chemical accuracy and it is only surpassed by the compound methods such as the Gaussian-*n*. Back in 1992, using nonlocal DFT, Becke obtained an accuracy of 3.7 kcal/mol for this G2/99 set using gradient-corrected functionals[5] and, in 1993, he obtained, using hybrid functionals, an accuracy of 2.4 kcal/mol.[6] Recent work related to our approach using linear regressions and statistical approaches[7,8] has been reported; the reader is forwarded to these references for further information.

II. Methodology

Several ab initio methods have been used in this work, aiming to evaluate the energy correctors through a broad span of levels of accuracy, from the inexpensive HF, going through the post-HF (CCSD,QCI,MP4), the DFT (LSDA, PW91PW91, B3PW91), and to the highly accurate G1 method. In the Hartree−Fock method, which is the first approximation and least expensive of the ab initio methods, the repulsion between electrons is taken into account as an average effect but not the specific repulsion interaction between electrons, which is known as the correlation effect. However, this approximation is still useful for first-level predictions of many systems. The Møller−Plesset MP*n* ($n = 2-4$) theories add this effect of correlation between electrons. These perturbations add a second-order ($n = 2$), third-order ($n = 3$), etc. energy correction to the HF electronic energy.[9] The configuration interaction (CI)[10] method treats the wave function as a linear combination of HF wave functions for the ground state and excited states and the quadratic configuration interaction (QCI)[11] includes the size consistent term missing in the CISD, allowing a better comparison of energies. The accuracy of an MP4 calculation is approximately equivalent to the accuracy of a CISD calculation. The B3PW91 uses the Becke-3 term hybrid exchange functional and the generalized-gradient approximation (GGA) Perdew−Wang 91 correlation functional.[6,12,13]

In this work, the atomization energy of each molecule is calculated as

$$D_o = (\sum_{i=\text{atoms}} E_i) - E_{\text{molecule}} - E_{\text{ZPE}} \tag{1}$$

where E_i and E_{molecule} are the individual energies of all atoms in the molecule and the total energy of the molecule, respectively, and E_{ZPE} is the zero-point vibrational energy correction. All calculations are performed using the program Gaussian 03[14] unless specified otherwise. Average errors (average error of the mean absolute errors) of the atomization energy are with respect to experimental values reported in the Gaussian-1 paper,[1] which are reproduced in Table 1. Table 2 lists all absolute energies used in this work. The first five columns correspond to the energies reported in ref 1, G1, MP4/6-311G(d,p), MP4/6-311+G(d,p), MP4/6-311G(2df,p), QCI/6-311G(d,p); the remaining 10 columns, B3PW91/6-31G(d,f), B3PW91/cc-pVTZ, CCSD(T)/cc-pCQZ, CCSD/cc-pVTZ, HF/6-31G(d), HF/3-21G, PW91PW91/6-31G(d,f), PW91PW91/cc-pVTZ, LSDA/3-21G, and LSDA/6-31G(d,p), are computed in this work. A summary and further references about these methods and basis sets can be found in the Gaussian book.[10]

TABLE 1: Experimental Values of Atomization Energies (*D*o) for the G1 Set of Molecules of First Row

molecules	D_o,[a] kcal/mol
H_2	103.54
LiH	56.0
BeH	46.9
CH	79.9
$CH_2(^3B_1)$	179.6
$CH_2(^1A_1)$	170.6
CH_3	289.2
CH_4	392.5
NH	79.0
NH_2	170.0
NH_3	276.7
OH	101.3
OH_2	219.3
FH	135.2
Li_2	24.0
LiF	137.6
HCCH	388.9
H_2CCH_2	531.9
H_3CCH_3	666.3
CN	176.6
HCN	301.8
CO	256.2
HCO	270.3
H_2CO	357.2
H_3COH	480.8
N_2	225.1
H_2NNH_2	405.4
NO	150.1
O_2	118.0
HOOH	252.3
F_2	36.9
CO_2	381.9

[a] From ref 1.

III. Effect of the Correctors

To improve the atomization energies, we introduce corrective factors such that they minimize the average errors for a specific level of theory (method/basis set). These correctors are multiplicative factors, ϵ_E and ϵ_{ZPE}, that compensate for the deficiencies to predict the total electronic energy and the zero-point energy, respectively, as well as additive factors $\Delta E_{\text{multiplicity}}$ (ΔE_1, ΔE_2, ΔE_3, and ΔE_4 for the singlet, doublet, triplet, and quartet, respectively) to compensate for the errors in the multiplicity or total spin of the atoms and molecules. Thus, the corrected total energy ($E^{\text{corrected}}$) for each atom or molecule is given by

$$E^{\text{corrected}} = \epsilon_E E + \epsilon_{\text{ZPE}} E_{\text{ZPE}} + \Delta E_{\text{multiplicity}} \tag{2}$$

Thus an explicit expression for the theoretical dissociation energy, which includes our correctors, is constructed as

$$D_o = (\sum_{i=\text{atoms}} [\epsilon_E E + \Delta E_{\text{multiplicity}}]^i) - (\epsilon_E E + \epsilon_{\text{ZPE}} E_{\text{ZPE}} + \Delta E_{\text{multiplicity}})^{\text{molecule}} \tag{3}$$

Each set of correction factors depends on the method and basis set used to calculate the molecular energies. By minimizing the atomization energy error of the G1 set of molecules, the corrector values reported in Table 3 are found to provide the

TABLE 2: Total Electronic Energies in Hartrees for the G1 Set of Molecules[a]

molecules	M1[b]	M2[b]	M3[b]	M4[b]	M5[b]	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15
H	−0.50000	−0.49981	−0.49981	−0.49981	−0.49981	−0.50218	−0.50398	−0.49995	−0.49981	−0.49823	−0.49620	−0.49891	−0.50118	−0.49086	−0.49394
H_2	−1.17446	−1.16772	−1.16772	−1.16772	−1.16832	−1.17752	−1.17894	−1.17380	−1.17234	−1.12683	−1.12296	−1.16854	−1.17008	−1.16413	−1.17125
Li	−7.43222	−7.43203	−7.43203	−7.43203	−7.43203	−7.48363	−7.48462	−7.43272	−7.43270	−7.43137	−7.38151	−7.47206	−7.47301	−7.34453	−7.39552
Be	−14.62337	−14.61307	−14.61308	−14.61307	−14.61722	−14.65436	−14.65769	−14.61895	−14.61843	−14.56694	−14.48682	−14.64332	−14.64675	−14.43052	−14.51556
B	−24.60279	−24.58812	−24.58852	−24.59239	−24.59179	−24.63649	−24.64474	−92.58991	−24.59679	−24.52204	−24.38976	−24.62558	−24.63478	−24.30172	−24.43723
C	−37.78464	−37.76430	−37.76520	−37.77483	−37.76669	−37.82569	−37.83683	−37.78654	−37.77873	−37.68086	−37.48107	−37.81567	−37.82817	−37.36339	−37.56615
N	−54.51776	−54.49089	−54.49251	−54.50889	−54.49143	−54.56393	−54.58043	−54.52482	−54.51243	−54.38544	−54.10539	−54.55492	−54.57258	−53.95623	−54.24400
O	−74.98204	−74.93333	−74.93724	−74.96478	−74.93402	−75.03133	−75.06125	−74.99357	−74.97105	−74.78393	−74.39366	−75.02652	−75.05861	−74.24571	−74.64334
F	−99.63275	−99.56534	−99.57112	−99.60790	−99.56580	−99.68057	−99.72550	−99.65026	−99.60536	−99.36496	−98.84501	−99.67855	−99.72749	−98.69995	−99.22831
LiH	−8.02516	−8.01548	−8.01557	−8.01810	−8.01631	−8.07065	−8.07548	−8.02392	−8.02232	−7.98087	−7.92984	−8.05612	−8.06075	−7.93304	−7.98522
BeH	−15.19965	−15.18931	−15.18938	−15.19182	−15.19074	−15.24819	−15.25237	−15.19816	−15.19520	−15.14731	−15.05931	−15.23250	−15.23717	−15.01752	−15.10701
CH	−38.41842	−38.38606	−38.38775	−38.40073	−38.38959	−38.45947	−38.47372	−38.41880	−38.40710	−38.26493	−38.05191	−38.44891	−38.46420	−37.99490	−38.20796
$CH_2(^3B_1)$	−39.08345	−39.05311	−39.05425	−39.06810	−39.05466	−39.13037	−39.15400	−39.08733	−39.07446	−38.92150	−38.70907	−39.12618	−39.14137	−38.68080	−38.89612
$CH_2(^1A_1)$	−39.07314	−39.03053	−39.03253	−39.04858	−39.03467	−39.11058	−39.12739	−39.07191	−39.05657	−38.87237	−38.65185	−39.09787	−39.11581	−38.64958	−38.87181
CH_3	−39.77019	−39.73077	−39.73205	−39.74883	−39.73224	−39.82752	−39.84165	−39.77245	−39.75626	−39.55899	−39.34261	−39.81157	−39.82717	−39.37768	−39.59542
CH_4	−40.45038	−40.40503	−40.40533	−40.42466	−40.40589	−40.50869	−40.52128	−40.45089	−40.43182	−40.19517	−39.97688	−40.49085	−40.50495	−40.07271	−40.29186
NH	−55.14795	−55.10777	−55.11113	−55.13048	−55.10922	−55.20039	−55.22008	−55.15458	−55.13691	−54.95942	−54.65924	−55.19295	−55.21370	−54.58525	−54.89015
NH_2	−55.80452	−55.75310	−55.75796	−55.77971	−55.75444	−55.85708	−55.88010	−55.81081	−55.78799	−55.55770	−55.24538	−55.84953	−55.87388	−55.24300	−55.56287
NH_3	−56.48781	−56.42804	−56.43434	−56.45656	−56.42843	−56.53751	−56.56283	−56.49305	−56.46554	−56.18436	−55.87220	−56.52830	−56.55557	−55.93685	−56.26321
OH	−75.65027	−75.58825	−75.59534	−75.62361	−75.58921	−75.69942	−75.73347	−75.66163	−75.63263	−75.38228	−74.97023	−75.69674	−75.73313	−74.91007	−75.33273
OH_2	−76.34885	−76.27607	−76.28690	−76.31346	−76.27607	−76.39255	−76.43068	−76.35979	−76.32456	−76.01075	−75.58596	−76.39028	−76.43109	−75.60643	−76.05047
FH	−100.35599	−100.27374	−100.28583	−100.31962	−100.27346	−100.39428	−100.44733	−100.37318	−100.33199	−100.00291	−99.46022	−100.39476	−100.45233	−99.41106	−99.97250
Li_2	−14.90773	−14.89777	−14.89783	−14.89924	−14.90006	−14.99551	−14.99874	−14.90369	−14.90318	−14.86693	−14.76925	−14.97708	−14.97982	−14.72742	−14.82782
LiF	−107.29223	−107.19858	−107.21054	−107.25763	−107.19666	−107.36945	−107.42104	−107.30016	−107.25597	−106.93423	−106.35419	−107.36901	−107.42398	−106.28673	−106.87043
HCCH	−77.21231	−77.13994	−77.14186	−77.17987	−77.13976	−77.29452	−77.32507	−77.20930	−77.17118	−76.81783	−76.39596	−77.28735	−77.32048	−76.43942	−76.85708
H_2CCH_2	−78.46291	−78.38239	−78.38425	−78.42246	−78.38414	−78.56139	−78.58780	−78.46159	−78.42380	−78.03172	−77.60099	−78.54567	−78.57426	−77.70345	−78.13047
H_3CCH_3	−79.69777	−79.61452	−79.61518	−79.65286	−79.61579	−79.80958	−79.83264	−79.69871	−79.66098	−79.22875	−78.79395	−79.78439	−79.80981	−78.95371	−79.38555
CN	−92.58739	−92.49482	−92.49778	−92.54044	−92.51406	−92.66988	−92.70166	−92.58991	−92.54726	−92.20483	−91.68475	−92.68096	−92.71435	−91.65257	−92.13139
HCN	−93.30153	−93.22384	−93.22685	−93.27083	−93.22083	−93.38281	−93.41685	−93.30127	−93.25725	−92.87520	−92.35408	−93.38585	−93.42173	−92.37256	−92.88427
CO	−113.18218	−113.09862	−113.10234	−113.15282	−113.09356	−113.26028	−113.30509	−113.18789	−113.13855	−112.73788	−112.09330	−113.26856	−113.31562	−112.05840	−112.68957
HCO	−113.71030	−113.61895	−113.62470	−113.67603	−113.61658	−113.80451	−113.84911	−113.71783	−113.66709	−113.24766	−112.60380	−113.81212	−113.85880	−112.60764	−113.24026
H_2CO	−114.36336	−114.26263	−114.26926	−114.32149	−114.26103	−114.45657	−114.50040	−114.36897	−114.31730	−113.86633	−113.22182	−114.45966	−114.50546	−113.26521	−113.89966
H_3COH	−115.58000	−115.46847	−115.47733	−115.52787	−115.46876	−115.68061	−115.72638	−115.58889	−115.53603	−115.03542	−114.39802	−115.67285	−115.72139	−114.48674	−115.13202
N_2	−109.39933	−109.31616	−109.32091	−109.36983	−109.31021	−109.47710	−109.51944	−109.40436	−109.35536	−108.94395	−108.30095	−109.48815	−109.53156	−108.29025	−108.91514
H_2NNH_2	−111.72705	−111.61706	−111.62663	−111.67406	−111.61750	−111.82834	−111.87306	−111.73589	−111.68227	−111.16937	−110.55001	−111.82230	−111.86989	−110.65578	−111.24445
NO	−129.74364	−129.64280	−129.64887	−129.70860	−129.64088	−129.83582	−129.88633	−129.75439	−129.69841	−129.24788	−128.50064	−129.85273	−129.90471	−128.47577	−129.20946
O_2	−150.15196	−150.04161	−150.04860	−150.11869	−150.03681	−150.26254	−150.32181	−150.17382	−150.11126	−149.61791	−148.76909	−150.28544	−150.34643	−148.74276	−149.57431
HOOH	−151.38907	−151.25522	−151.26716	−151.33467	−151.25470	−151.48602	−151.55190	−151.40814	−151.33844	−150.76479	−149.94582	−151.49968	−151.56927	−149.97123	−150.81282
F_2	−199.32774	−199.17505	−199.18575	−199.27479	−199.17432	−199.42651	−199.51097	−199.35872	−199.27831	−198.67776	−197.64424	−199.45173	−199.54202	−197.54026	−198.59411
CO_2	−188.37245	−188.22507	−188.23275	−188.32539	−188.21533	−188.50695	−188.58282	−188.38452	−188.29867	−187.63418	−186.56126	−188.53160	−188.61010	−186.57270	−187.61678

[a] M1 = G1, M2 = MP4/6-311G(d,p), M3 = MP4/6-311+G(d,p), M4 = MP4/6-311G(2df,p), M5 = QCI/6-311G(d,p), M6 = B3PW91/6-31G(d,p), M7 = B3PW91/cc-pVTZ, M8 = CCSD(t)/cc-pVQZ, M9 = CCSD/cc-pVTZ, M10 = HF/6-31G(d), M11 = HF/3-21G, M12 = PW91PW91/6-31G(d,p), M13 = PW91PW91/cc-pVTZ, M14 = LSDA/3-21G, M15 = LSDA/6-31G(d,p). [b] From ref 1.

TABLE 3: Values of ΔE_4, ΔE_3, ΔE_2, ΔE_1, ϵ_{ZPE} and ϵ_E That Minimize the Average Error of the Atomization Energies for the 32 Molecules of the G1 Set Plus H_2

method/basis set	ΔE_4 (kcal/mol)	ΔE_3 (kcal/mol)	ΔE_2 (kcal/mol)	ΔE_1 (kcal/mol)	ϵ_{ZPE}	ϵ_E
G1	1.1	−0.1	0.8	2.1	0.99	1.00
MP4/ 6-311G(d,p)	15.9	11.9	6.7	4.9	0.99	0.96
MP4/6-311+G(d,p)	13.9	10.2	5.3	3.7	0.99	0.97
MP4/6-311G(2df,p)	10.3	7.2	6.1	4.9	0.99	0.96
QCI/6-311G(d,p)	15.0	10.6	4.6	3.6	0.99	0.98
B3PW91/6-31G(d,p)	4.8	3.3	0.1	−5.9	0.97	0.97
B3PW91/cc-pVTZ	0.0	0.5	−1.4	−6.8	0.72	0.98
CCSD(T)/cc-pVQZ	2.3	1.1	−0.1	−0.8	0.96	1.01
CCSD/cc-pVTZ	10.6	8.0	0.6	0.4	0.96	1.00
HF/6-31G(d)	71.0	61.9	22.6	3.8	0.99	0.84
HF/3-21G	89.0	75.1	22.6	1.0	0.99	0.80
PW91PW91/6-31G(d,p)	−4.3	−3.5	1.6	−2.9	0.97	0.97
PW91PW91/cc-pVTZ	−9.5	−7.0	0.9	−1.8	0.97	1.00
LSDA/3-21G	−11.6	−18.8	−7.7	−5.5	0.99	0.98
LSDA/6-31G(d,p)	2.6	2.1	5.3	0.9	0.99	0.84

TABLE 4: Average Error of the Atomization Energies, before and after Corrections, for the First Row Molecules (G1 Set Plus H_2) with Respect to Experimental Values

method/basis set	error before (kcal/mol)	errors after (kcal/mol)
G1	1.5[a]	0.8
MP4/ 6-311G(d,p)	15.4[a]	2.4
MP4/6-311+G(d,p)	14.6[a]	1.8
MP4/6-311G(2df,p)	8.3[a]	1.9
QCI/6-311G(d,p)	16.7[a]	2.1
B3PW91/6-31G(d,p)	4.1	2.1
B3PW91/cc-pVTZ	2.7	1.2
CCSD/cc-pVTZ	12.5	2.1
CCSD(t)/cc-pVQZ	3.6	0.8
HF/6-31G(d)	87.3	9.2
HF/3-21G	105.5	12.9
PW91PW91/6-31G(d,p)	9.4	3.9
PW91PW91/cc-pVTZ	9.0	3.6
LSDA/3-21G	29.1	9.2
LSDA/6-31G(d,p)	40.0	7.0

[a] From ref 1.

best fittings for each specific level of theory, i.e., method/basis set; applying these corrective factors allows us to drastically improving the accuracy in the energetic predictions. The magnitude of these corrective factors already provides an idea of the quality of the uncorrected methods. The smallest additive factors correspond to the G1 method and the largest to the HF levels of theory. It is important to notice how the geometries and ZPE are calculated. For the G1 method the geometries are calculated at the MP2/6-31G(d,p) level of theory and the ZPE is calculated at the HF/6-31G(d,p) level. However, for all DFT and HF methods the geometry and ZPE are calculated at the same level of theory as the energies. For the CC methods, the geometries and ZPEs are calculated at the CCSD/cc-pVTZ, as shown in Table 3.

We also find the values of the correctors when the ϵ_{ZPE} is constrained to exactly 1.00. No major changes are noted in the values of the correctors; however, when we constrain the multiplicative correctors, ϵ_{ZPE} and ϵ_E, at exactly 1.00, relatively large changes can be observed in the values of the additive correctors, especially for the less precise methods.

Table 4 shows the actual improvement using the correctors reported in Table 3. The average error of the G1 method can be brought down to 0.8 kcal/mol. Interestingly, all the standard ab initio methods can also be brought to chemical accuracy. The CCSD(T)/cc-pVQZ also reaches an average error with respect to experiment of 0.8 kcal/mol using a large basis set, which splits the valence basis into four levels and uses a total of 55 functions per each first row atom. Interestingly, all nonlocal DFT methods yield excellent energetics at a relatively low cost. Methods such as B3PW91/6-31G(d,p) yield errors of only 2.1 kcal/mol, already at the chemical accuracy range, and with the triple-ζ basis, cc-pVTZ, the error goes down to only 1.2 kcal/mol, in excellent agreement with experiment. However, it is worthwhile to mention that the B3PW91 already yield excellent energetics without the corrections. The errors of 4.1 and 2.7 kcal/mol with the 6-31G(d,p) and cc-pVTZ basis sets,

TABLE 5: Atomization Energy Errors before and after Using the Energy Correctors to 15 Molecules Belonging to the G3 Set

	B3PW91/ cc-pVTZ		B3PW91/ 6-31G(d,p)		PW91PW91/ 6-31G(d,p)		PW91PW91/ cc-pVTZ		HF/ 3-21G		HF/ 6-31G(d)		LSDA/ 3-21G		LSDA/ 6-31G(d,p)	
molecule	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A
C_6H_6	7.0	0.1	12.0	1.9	53.9	4.3	44.6	7.8	364	28	397	31.8	212	31.8	220	1.5
CF_3CN	1.1	3.4	5.2	5.5	55.3	31.2	46.60	26.0	240	48	301	66.1	146	66.1	175	63.5
C_4H_{10}	3.0	10.0	8.0	5.2	31.3	1.5	16.56	1.8	310	0.2	338	10.3	181	10.4	190	7.5
CH_3NH_2	2.0	1.4	0.5	0.7	12.7	0.8	9.80	0.8	176	0.2	200	3.3	71	3.3	88	9.2
CH_3NO_2	2.7	4.1	2.8	8.1	45.6	20.0	42.31	15.4	270	3.6	357	10.1	108	10.1	157	57.4
CH_3CONH_2	0.2	2.6	1.3	0.1	33.2	3.5	29.27	3.9	274	5.2	324	13.9	132	13.9	159	25.6
$CH_3CH_2NH_2$	1.2	3.6	2.9	1.3	22.9	0.3	16.36	0.2	247	0.8	278	3.4	117	3.4	137	14.4
$HCOOCH_3$	0.4	1.3	2.9	2.0	36.7	8.4	30.12	6.4	245	1.1	299	11.0	127	11.0	155	30.5
CH_3COCH_3	0.0	3.6	5.1	0.6	33.8	3.4	25.65	3.6	264	0.6	310	7.7	14	7.7	163	16.4
C_2H_4NH	1.8	1.3	4.2	3.6	25.8	4.2	21.23	3.2	222	1.1	266	1.2	97	1.2	126	20.3
$C_2H_5OCH_3$	2.3	6.6	5.7	2.4	33.3	3.0	21.58	1.5	290	5.0	328	7.5	160	7.5	180	21.8
CH_3CH_2OH	3.6	5.0	0.3	3.8	20.3	1.2	13.99	0.5	224	1.2	256	5.3	107	5.3	129	12.8
C_2F_4	9.2	3.9	14.3	11.3	63.4	46.6	52.48	42.4	213	62.4	263	85.9	162	85.9	186	84.3
CH_3OCH_3	2.3	3.6	3.0	0.9	24.2	2.9	15.93	1.3	219	5.1	250	2.2	115	2.2	132	17.8
CH_3CH_2O	2.8	3.4	8.3	0.5	27.7	3.0	20.69	2.5	184	9.0	212	4.7	105	4.7	119	13.6
Average	2.7	3.6	5.1	3.2	34.7	9.0	27.14	7.8	249	11.4	292	17.7	132	17.6	154	26.4

respectively, are much better than any of the standard highly correlated ab initio methods.

Methods with strong and unacceptable errors such as HF and the LSDA can be practically corrected to provide acceptable energetics, as indicated in Table 4. The strong improvement of the LSDA was already a matter of study when this method was compared with the G1 set in the early 90's,[15] and then extended to the use of pure nonlocal methods.[16] As shown in Table 3, the HF/3-21G, which only requires 9 basis functions per first row atom, makes an improvement from an average error of 105.5 kcal/mol to 12.9 kcal/mol., which is much better than several of the uncorrected highly correlated methods.

To validate the use of the reported energy correctors, we used them in molecules outside the G1 training group of molecules. We used the largest molecules from the G3 set. The results for each specific case are shown in Table 5. Notice that the corrections for the already good methods are not as strong as the corrections for the more imprecise methods. In addition, the errors for this set of molecules increase tremendously. The HF/6-31G(d,p) average error is now 292 kcal/mol, and the correctors lower it to 17.7 kcal/mol. Interestingly, for these larger molecules, the simpl and economic HF/3-21G method yields an error of 250 kcal/mol, which is reduced to 17.7 kcal/mol after the corrections. These systems are relatively large to test the highly correlated methods.

IV. Conclusions

We use as training set the first row subset of molecules from the G1 basis of molecules; when the correctors are applied to other molecules not included in the training set, selected from the G3 basis, similar improvements in the atomization energies are obtained. The special case of the B3PW91/cc-pVTZ yields an average error of 1.2 kcal/mol, which is already within a chemical accuracy and comparable to the Gaussian-*n* theories accuracy. The very inexpensive B3PW91/6-31G** yields an average error of 2.1 kcal/mol using the correctors. Methods considered unsuitable for energetics such as HF and LSDA yield corrected energies comparable to those obtained with the best highly correlated methods. We demonstrate that simple corrections taking into account the multiplicity of the molecules yield extraordinary improvements in the total energies using a variety of precise and imprecise methods. It is concluded that simple systematic corrections can yield acceptable energetics without the need to perform further brute force calculations. It is also concluded that different multiplicities are not treated equally by the different methods no matter their origin. This fact needs to be considered to develop spin dependent functionals.

Acknowledgment. We highly appreciate the support of ARO (DURINT and DURIP grants), DOE, and DARPA/ONR.

References and Notes

(1) Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622.

(2) Curtiss, L. A.; Jones, C.; Trucks, G. W.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1990**, *93*, 2537.

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

(4) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.

(5) Becke, A. D. *J. Chem. Phys.* **1992**, *97*, 9173.

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

(7) Duan, X.-M.; Song, G.-L.; Li, Z.-H.; Wang, X.-J.; Chen, G.-H.; Fan, K.-N. *J. Phys. Chem.* **2004**, *121*, 7086.

(8) Wang, X.; Wong, L.; Hu, L.; Chan, C.; Su, Z.; Chen, G. *J. Phys. Chem. A* **2004**, *108*, 8514.

(9) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(10) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*; Gaussian Inc.: Pittsburgh, PA, 1996.

(11) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968.

(12) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.

(13) Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. *Phys. Rev. B* **1992**, *46*, 6671.

(14) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T., Jr.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian-2003*, revision C.2; Gaussian, Inc.: Pittsburgh, PA, 2003.

(15) Seminario, J. M.; Concha, M. C.; Politzer, P. *Int. J. Quantum Chem.* **1991**, *25*, 249.

(16) Seminario, J. M. *Chem. Phys. Lett.* **1993**, *206*, 547.