Theoretical Chemistry Accounts (2018) 137:24
https://doi.org/10.1007/s00214-018-2206-3

REGULAR ARTICLE

Empirical corrections in the G3X and G3X(CCSD) theories combined with a compact effective pseudopotential

Cleuton de Souza Silva[1,2] · Rogério Custodio[1]

Received: 12 September 2017 / Accepted: 16 January 2018 / Published online: 25 January 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

The G3X-CEP and G3X(CCSD)-CEP composite methods were developed by combining a compact effective pseudopotential (CEP) with the G3X and G3X(CCSD) all-electron methods. The G3/05 test set was used as reference data to analyze the performance of these methods by calculating enthalpies of formation, ionization energies, electron affinities, proton affinities, and molecular atomization energies. Tests were carried out considering different numbers of higher-level correction (HLC) parameters as well as a scaling of the experimental atomization energies of the elements used to determine the enthalpies of formation. The overall behavior of G3X-CEP and G3X(CCSD)-CEP calculations considering six HLC parameters presented a mean absolute error around 1.15 kcal mol^{-1} for all calculated properties. This tendency is intermediate to the G3CEP and G4CEP methods. The use of scaled experimental atomization energies of the elements provided a mean absolute error around 1.11 kcal mol^{-1}, which is similar to the one obtained by the G4CEP theory (1.09 kcal mol^{-1}). The G3X-CEP and G3X(CCSD)-CEP enthalpies of formation were significantly improved with respect to experimental results. Molecular atomization energies showed a modest improvement, while the ionization energies and electronic affinities were slightly affected indirectly by the reoptimized HLC parameters. The proton affinities were nearly unaffected.

Keywords G3X · G3X(CCSD) · G3CEP · Pseudopotential · Composite theory · Thermochemical properties

1 Introduction

G3X and G3X(CCSD) [1, 2] are composite methods belonging to the more general Gaussian-n theory [3–7], usually referred to as Gn. Both methods are modifications of the G3 theory [5], and the most significant differences with respect to the original G3 theory lie in the geometry optimization and a modification of a specific large basis set (G3large) by including g polarization for representative elements heavier than aluminum. The mean absolute error of G3X and G3X(CCSD) with respect to experimental data from a test set named G3/99 [8] is of the order of 0.95 kcal mol^{-1}, which is slightly lower than the mean absolute error of 1.06 kcal mol^{-1} for the G3 results [1, 2].

The main uses of the G3X theories have been in the study of thermochemical properties involving compounds containing phosphorous [9], chlorine [10], fluorine [11], and bromine [12] groups. As an example, Akhmetova et al. [13] carried out calculations for the homolytic dissociation energy of the N–F bond for 31 compounds and achieved an absolute mean deviation of 1.4 kJ mol^{-1} for G3X and 3.6 kJ mol^{-1} for the more elaborated G4 theory.

A recent attempt to reduce the CPU time of calculations involving the Gn theory was initiated by Pereira et al. [14, 15], which combined a compact effective potential (CEP) [16, 17] with the G3 theory resulting in a method known as G3CEP. The procedure was tested against a series of experimental reference data, also referred to as G3/05 test set [2], providing a mean absolute error of 1.29 kcal mol^{-1} compared to 1.13 kcal mol^{-1} from the original G3 method, but with a significant reduction in the CPU time. The G3CEP method was also used in the study of internal rotation barriers of different compounds [18],

---

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00214-018-2206-3) contains supplementary material, which is available to authorized users.

✉ Rogério Custodio
roger@iqm.unicamp.br

[1] Instituto de Química, Universidade Estadual de Campinas, P. O. Box 6154, Barão Geraldo, Campinas, São Paulo 13083-970, Brazil

[2] Instituto de Ciências Exatas e Tecnologia, Universidade Federal do Amazonas, Campus de Itacoatiara, Itacoatiara, Amazonas 69100-021, Brazil

demonstrating an excellent agreement between the calculated results with experimental information.

After this first successful generalization by combining pseudopotential with the G3 theory, other attempts were developed, such as: G3(MP2)//B3LYP-CEP [19, 20], G3(MP2)-CEP [21], G4-CEP [22], and W1CEP [23]. Applications of these methods demonstrated a significant economy in the CPU time preserving an excellent accuracy when compared with the original all-electron versions. The mean absolute errors with respect to the same set of experimental data of all calculated properties by calculations including and not including pseudopotential presented a systematic difference of the order of 0.20 kcal mol^{-1}. The original calculations were closer to the experimental data than the pseudopotential results. As a recent example, the mean absolute error of the G4 calculations considering the G3/05 test set is 0.83 kcal mol^{-1}, while the pseudopotential version, G4CEP, produced a deviation of 1.09 kcal mol^{-1} [22], a difference between the mean absolute errors of 0.26 kcal mol^{-1}. The sophistication of the G4 theory with respect to G4CEP suggests that the pseudopotential version may not achieve a level of accuracy as high as the all-electron calculations. The difficulty in reaching the all-electron accuracy possibly indicates that there must be a limit of accuracy for the pseudopotential versions due to the lack of correlation effects between core and valence electrons.

One feasible way to compensate for the absence of electronic correlation effects in the pseudopotential calculations is through empirical corrections. The higher-level correction (HLC) from the Gn theory, which depends on the number of valence electrons, improves the accuracy of the thermochemical properties by reducing the error between the calculated and experimental data. The optimum HLC parameters are different for each Gn theory.

Another important aspect is that enthalpies of formation are calculated considering experimental values of atomization energies of the elements of the periodic table. These atomization energies are derived from rigorous experimental measurements, but also present small uncertainties and may be suitably adjusted.

In this work, two different objectives will be explored. The first corresponds to the implementation of CEP pseudopotential in the G3X and G3X(CCSD) theories. The second objective will be to evaluate different empirical adjustments in the pseudopotential versions of the G3X theory. HLC and the scaling of experimental atomization energies are the main empirical adjustments tested in this work.

2 Theoretical methods

The combination of G3X and G3X(CCSD) along with the CEP pseudopotential yielded two methods labeled as: G3X-CEP and G3X(CCSD)-CEP. As mentioned previously, the main differences between G3X and G3 are the level of optimization of geometries, and an additional correction in the large basis set, ΔE_{HF}. The G3X reference energy is calculated at the QCISD(T)/6-31G(d) level, while the G3X(CCSD) considers the CCSD(T)/6-31G(d) level of theory. The general expression for G3X contains corrections for diffuse (ΔE_{+}) and polarization ($\Delta E_{\text{2df,p}}$) functions, electron correlation effects ($\Delta E_{\text{QCISD(T)}}$), effects from the size of the basis functions ($\Delta E_{\text{G3 large}}$), and improvement in the G3large basis set (ΔE_{HF}), spin–orbit correction (E_{SO}) from the literature, zero point energy and thermal effects (E_{ZPE}), and an empirical higher-level correction (E_{HLC}). The combination of all these effects is summarized in the expression:

$$\begin{aligned} E_{\text{G3X}} = {} & E[\text{MP4/6−31G(d)}] + \Delta E_{+} + \Delta E_{\text{2df,p}} + \Delta E_{\text{QCISD(T)}} \\ & + \Delta E_{\text{G3large}} + \Delta E_{\text{HF}} + E_{\text{SO}} + E_{\text{ZPE}} + E_{\text{HLC}} \end{aligned} \tag{1}$$

where the mathematical definition of each correction is:

$$\Delta E_{+} = E[\text{MP4/6−31} + \text{G(d)}] - E[\text{MP4/6−31G(d)}];$$

$$\Delta E_{\text{2df,p}} = E[\text{MP4/6−31G(2df,p)}] - E[\text{MP4/6−31G(d)}];$$

$$\Delta E_{\text{QCISD(T)}} = E[\text{QCISD(T)/6−31G(d)}] - E[\text{MP4/6−31G(d)}];$$

$$\begin{aligned} \Delta E_{\text{G3largeXP}} = {} & E[\text{MP2(full)/G3largeXP}] - E[\text{MP2/6−31}G\text{(2df,p)}] \\ & - E[\text{MP2/6} - 31 + \text{G(d)}] + E[\text{MP2/6} - \text{31G(d)}] \quad \text{and} \\ \Delta E_{\text{HF}} = {} & E[\text{HF/G3Xlarge}] - E[\text{HF/G3large}]. \end{aligned}$$

The determinations of all these energies are performed with the molecular equilibrium geometries obtained from B3LYP/6-31G(2df,p). The zero point energy (ZPE) is also obtained from the harmonic approximation at the B3LYP/6-31G(2df,p) level, and the frequencies are scaled by a factor of 0.9854. When necessary, thermal energies are determined from well-established expressions from the statistical thermodynamics at the temperature of 298.15 K.

Most of the adaptation of the CEP pseudopotential to the G3X and G3(CCSD) theories followed the procedure described for G3CEP [14, 15], G3(MP2)-CEP [21], and G3(MP2)//B3-CEP [19]. Only valence Gaussian functions from 6-31G(d), 6-31 + G(d), and 6-31G(2df,p) are preserved and the adapted basis sets are referred to as: P31G(d), P31G + G(d), and P31G(2df,p), where P is an indication that the original Pople's valence basis functions are being used. For the representative elements of the third period, the basis functions were adapted from Rassolov et al. [24, 25] and were truncated following the same procedure already described in references [14, 15].

Pereira et al. [14, 15] considered that for the G3CEP theory the G3large basis set for the C, N, O, F, P, Cl, As, Se, and Br atoms required some adjustment of the s and p exponents to achieve more accurate results, mainly for the enthalpy of formation. This procedure was also applied in the G3(MP2)-CEP [21] and G3(MP2)//B3-CEP [19] adaptations.

The s and p adapted exponents for G3X-CEP and G3X(CCSD)-CEP were scaled by the following parameters for: C = 0.9889, N = 0.9639, O = 0.9349, F = 0.9222, P = 0.8146, Cl = 1.0154, S = 1,100,158, As = 0.8282, Se = 0.8133, and Br = 0.8330. The same set of arbitrary molecules selected for the optimization of the basis set for the G3CEP theory were used in the present optimization [14]. All truncated large basis sets are available as supplemental material in Tables S.1 and S.2.

Finally, after adjusting all basis sets, the HLC parameters were optimized to minimize the mean absolute error taking into account all properties from the G3/05 test set. The G3, G3X, and G3X(CCSD) theories use four empirical parameters (*A*, *B*, *C,* and *D*) optimized to improve the accuracy of the properties with respect to the α and β valence electrons. *A* and *B* parameters are associated with molecules and *C* and *D* to atoms by the expressions: $E_{\text{HLC}} = -An_{\beta} - B\left(n_{\alpha} - n_{\beta}\right)$ and $E_{\text{HLC}} = -Cn_{\beta} - D\left(n_{\alpha} - n_{\beta}\right)$, respectively. The G4 theory includes two additional parameters. An additional A' parameter is used for closed-shell molecules: $E_{\text{HLC}} = -A'n_{\beta}$ and an *E* parameter for molecules and atoms containing valence orbital consisting of two s electrons, not considering 1s electrons. Some tests regarding the use of four or more HLC parameters will be discussed with respect to G3X-CEP and G3X(CCSD)-CEP.

A final adjustment is the scaling of the experimental atomization energies for the calculation of the enthalpies of formation. The experimental atomization energy of each element of the periodic table is multiplied by an adjustable parameter that minimizes the mean absolute error between the theoretical enthalpies of formation and the experimental one. Each parameter is set initially equal to 1.0 and they are optimized using the modified Simplex of Nelder and Mead [24]. This final optimization is also complemented by the reoptimization of the HLC parameters.

3 The HLC parameters

Following the original tests applied to the Gn methods, the G3X-CEP and G3X(CCSD)-CEP approximations were initially applied in the calculation of standard enthalpies of formation according to the procedure described in Curtiss et al. [25]. Proton affinities were obtained from the internal energy at 0 K as described for the G3 theory [5]. Adiabatic ionization energies and electron affinities were calculated as the difference of internal energies at 0 K between products and reactants in their respective equilibrium geometries [5]. Sets of 6d and 7f Gaussian functions were used as polarization functions at the B3LYP, MP4, CCSD(T), and QCISD(T) levels of calculation, whereas at the MP2 level sets of 5d and 7f primitives were used in this article for all calculations.

Two tests were carried out in this work regarding the HLC correction for G3X-CEP and G3X(CCSD)-CEP: four (*A*, *B*, *C,* and *D*) and six (*A*, *A′*, *B*, *C*, *D,* and *E*) parameters were optimized considering all properties available from the G3/05 test set. A summary of the optimized parameters is shown in Table 1. Columns identified as EnAt1 and EnAt2 will be discussed later. The difference between the parameters to be used along with all-electron and pseudopotential systems is evident. The pseudopotential methods present relatively similar parameters. The parameters for G3X-CEP and G3X(CCSD)-CEP are larger than the respective all-electron ones. The reason for this general tendency, as suggested in previous works [19, 21, 22], is the greater necessity of correction for the pseudopotential calculations. The most sensitive parameters are those associated with the valence closed-shell electrons (*A′* and *C*). The substantial number of closed-shell systems and electrons can be pointed out as responsible for this tendency. This argument is reinforced when six HLC parameters are used. Table 1 shows that *A′* (for closed-shell molecules) is larger than *A* (for open shell molecules) considering either G3X-CEP or G3X(CCSD)-CEP results.

Table 1 Original and optimized HLC parameters for the original G3X, G3X(CCSD), G3X-CEP, G3(CCSD)-CEP, EnAt1, and EnAt2. Data in mHartree

Parameters	4 HLC parameters				6 HLC parameters			
	G3X[a]	G3X (CCSD)[a]	G3X CEP	G3X (CCSD) CEP	G3X CEP	G3X (CCSD) CEP	EnAt1	EnAt2
A	6.783	6.635	8.08369	8.07464	8.08693	8.08606	6.76936	6.80215
B	3.083	3.085	3.60356	3.61270	2.68906	2.68548	2.70313	2.56229
C	6.877	6.645	8.38347	8.28178	8.38669	8.29224	6.91300	6.94169
D	1.152	1.076	1.75783	1.75877	1.75948	1.76604	1.09503	1.14548
E					− 0.00034	− 0.00516	0.00337	0.00338
A′					8.28948	8.28919	6.80984	6.87973

[a] Data from Refs. [1, 2]

The overall performance of the G3X and G3X(CCSD) theories for all calculated properties presents a mean absolute error of 1.01 kcal mol^{-1}, while for G3X-CEP and G3X(CCSD)-CEP with four HLC parameters the errors are 1.20 and 1.17 kcal mol^{-1}, respectively. The comparison among the all-electron and pseudopotential errors reproduces the tendency observed for other adaptations of CEP in the composite methods [14, 19, 21]. The use of six HLC parameters slightly improves the overall performance of G3X-CEP and G3X(CCSD)-CEP, which achieved mean absolute errors of 1.18 and 1.15 kcal mol^{-1}, respectively. The ionization and molecular atomization energies are the most sensitive properties to the enlargement of the number of HLC parameters. While the difference between the general performance of CEP calculations with four and six HLC parameters is 0.02 kcal mol^{-1}, both properties present a difference between the mean absolute error larger than 0.05 kcal mol^{-1}. The other properties, such as electron affinity, proton affinity, and enthalpy of formation, remain almost unchanged with the enlargement of the number of HLC parameters.

4 The performance for each property

All experimental and calculated properties considering six HLC parameters are shown as supplemental material. Table S.3 to S.7, using G3X, G3X-CEP, and G3X(CCSD)-CEP, contain the enthalpies of formation, ionization energies, electron and proton affinities, and molecular atomization energies, respectively. The performance of each property considering the three methods is summarized in Table 2 as well as some statistical properties.

A total of 57% of the enthalpies of formation calculated with G3X-CEP showed a deviation between ± 1 kcal mol^{-1}. Almost the same percentage (59%) was observed for G3X(CCSD)-CEP calculations. A greater percentage of 67% was obtained for the results calculated with G3X. All-electron calculations using G3X also concentrated only 13% of the deviations above ± 2 kcal mol^{-1}, while 17% of the results were observed for G3X-CEP and G3X(CCSD)-CEP.

The lowest deviations from the experimental data were observed for hydrocarbons (38 compounds) with mean

Table 2 Distribution considering different ranges of absolute error (in kcal mol^{-1}) of enthalpy of formation (ΔH_f^0), ionization energy (IE_0), electron affinity (EA_0), molecular atomization energy (D_0), and proton affinity (PA_0) calculated at the G3X-CEP, G3X(CCSD)-CEP, and G3X levels of theory. Mean absolute error (MAE) and standard deviation (Std) are also shown in kcal mol^{-1}

Energy range	ΔH_f^0			IE_0			EA_0		
	G3X-CEP	G3X(CCSD)-CEP	G3X	G3X-CEP	G3X(CCSD)-CEP	G3X	G3X-CEP	G3X(CCSD)-CEP	G3X
$\Delta E < -2$	18	16	10	12	12	12	4	4	4
$-2 \leq \Delta E < -1$	30	31	17	9	9	13	9	9	5
$-1 \geq \Delta E \leq 1$	141	146	167	52	52	64	39	39	37
$1 < \Delta E \leq 2$	35	30	31	16	16	10	7	8	11
$\Delta E > 2$	24	25	23	12	12	2	2	1	4
Total	248	248	248	101	101	101	61	61	61
% ($-1 \geq \Delta E \leq 1$)	57%	59%	67%	52%	52%	63%	64%	64%	61%
% ($-2 < \Delta E > 2$)	17%	17%	13%	24%	24%	14%	10%	8%	13%
MAE	1.16	1.12	1.01	1.32	1.32	1.04	0.92	0.92	0.99
Std	1.62	1.59	1.45	1.74	1.74	1.58	1.24	1.24	1.40

Energy range	D_0			PA_0		
	G3X-CEP	G3X(CCSD)-CEP	G3X	G3X-CEP	G3X(CCSD)-CEP	G3X
$\Delta E < -2$	2	2	4	0	0	0
$-2 \leq \Delta E < -1$	3	3	6	0	0	0
$-1 \geq \Delta E \leq 1$	10	10	10	3	3	3
$1 < \Delta E \leq 2$	5	5	1	4	4	4
$\Delta E > 2$	2	2	1	1	1	1
Total	22	22	22	8	8	8
% ($-1 \geq \Delta E \leq 1$)	46%	46%	46	38%	38%	38%
% ($-2 < \Delta E > 2$)	18%	18%	22%	13%	13%	13%
MAE	1.32	1.25	1.17	1.48	1.48	1.21
Std	1.75	1.67	1.50	1.44	1.44	0.83

absolute errors of 0.60 and 0.55 kcal mol^{-1} for G3X-CEP and G3X(CCSD)-CEP, while the mean absolute error for both G3X and G3X(CCSD) was 0.56 kcal mol^{-1}. The non-hydrogenated compounds were responsible for the largest deviations for all-electron and CEP data. The 79 non-hydrogenated compounds presented a mean absolute error of 1.76 and 1.78 kcal mol^{-1} for G3X-CEP and G3X(CCSD)-CEP, respectively, whereas for the all-electron original versions the mean absolute error was 1.63 kcal mol^{-1}.

G3X-CEP and G3X(CCSD)-CEP are intermediate to G3CEP and G4CEP in terms of the deviations of the enthalpies of formation with respect to experimental data. While the G3X-CEP and G3X(CCSD)-CEP had mean absolute errors of 1.16 and 1.12 kcal mol^{-1}, respectively, the equivalent original methods yielded deviations of 1.29 and 1.00 kcal mol^{-1}. The advantage of the G3X-CEP and G3X(CCSD)-CEP calculations over G4CEP is the much lower CPU time. The Hartree–Fock extrapolation technique used with G4CEP accounts for up to 70% of computational time [6, 22], which is not required in the G3 family.

Table 2 also shows the error distribution for the ionization energies and electron affinities. The mean absolute error for 101 ionization energies was 1.32 kcal mol^{-1} for both G3X-CEP and G3X(CSSD)-CEP calculations and 1.04 kcal mol^{-1} for G3X. The deviations from the experimental data showed that 52% of the results presented an accuracy between ± 1 kcal mol^{-1} for the calculations using pseudopotential. The G3X theory once again showed a better accuracy for ionization energies (63%). Similar trends were observed in other adaptations involving composite methods and pseudopotential [14, 19, 21] as well as ionization energies yielding the largest deviations when compared with the experimental data. It is worth mentioning that the mean absolute errors of the ionization energies obtained by G3X-CEP and G3X(CCSD)-CEP (1.32 kcal mol^{-1}) were identical to that presented by the G4CEP theory (1.31 kcal mol^{-1}) for the G3/05 test set.

The electronic affinities are considerably well-behaved compared to the ionization energies. Most of the deviations are concentrated below ± 2 kcal mol^{-1}. The mean absolute error of this property was 0.92 kcal mol^{-1} for G3X-CEP and G3X(CSSD)-CEP calculations. The all-electron version, G3X, is also well-behaved with a slightly larger deviation of 0.99 kcal mol^{-1}. Only 10% and 8% for the G3X-CEP and G3X(CCSD)-CEP calculations, respectively, show deviations above ± 2 kcal mol^{-1}, whereas in G3X theory, 13% of the results are above this limit. Comparing G3X-CEP and G3X(CCSD)-CEP electron affinity calculations with G4CEP indicates that the former presented slightly smaller deviations (0.92 kcal mol^{-1}) than the latter (1.02 kcal mol^{-1}). This tendency toward better results from electron affinities and ionization energies suggests a better cancelation of error from simpler methods.

The reduced number of proton affinities and molecular atomization energies containing elements of the third period of the periodic table affords a statistical analysis of little relevance. However, Table 2 shows that the 8 proton affinities present mean absolute errors of 1.27 and 1.25 kcal mol^{-1} for G3X-CEP and G3X(CSSD)-CEP, respectively, which are relatively close to the mean absolute error of 1.21 kcal mol^{-1} from the all-electron calculations.

Halogenated compounds tend to produce significant deviations for Gn calculations. Almost all molecular atomization energies have been calculated for halogenates, hydrides, or oxides. In this case, the mean absolute error for 22 energies was 1.31 and 1.25 kcal mol^{-1} for G3X-CEP and G3X(CSSD)-CEP, respectively. Calculations with G3X provided significantly lower deviations of 1.17 kcal mol^{-1}. The largest error for both G3X-CEP and G3X(CSSD)-CEP was due to KrF_2 with a deviation of 4.68 and 4.00 kcal mol^{-1}, respectively (see Table S.7). G3X performance is excellent for this example with an error of 1.3 kcal mol^{-1} with respect to the experimental results. The performance of G3X-CEP and G3X(CCSD)-CEP is comparable to G4CEP theory (with mean absolute error of 1.33 kcal mol^{-1}) for the calculation of molecular atomization energies.

The mean absolute errors from G3X-CEP and G3X(CCSD)-CEP for all properties are usually larger when compared to the original G3X theory. Even considering the optimization of HLC parameters, the effects of core-valence electron correlation cannot be completely offset by this empirical adjustment. The worst performances were for ionization energies, molecular atomization energies, and proton affinities. Ionization energies also presented the worst performance in other combinations between composite methods and pseudopotential, such as: G4CEP [22], G3CEP [14, 15], G3(MP2)-CEP [21], and G3(MP2)//B3-CEP [19].

Table 2 shows that the deviations obtained from calculations with both adaptations, G3X-CEP and G3X(CCSD)-CEP, are essentially the same. The definition of the reference energy for both methods from QCISD(T)/CEP-P31G(d) or CCSD(T)/CEP-P31G(d) calculations does not produce statistically significant advantages.

5 An additional empirical adjustment

One of the steps for the calculation of enthalpies of formation comes from experimental atomization energies of the elements at 0 K and then thermal effects are added. These data are obtained from rigorous experimental reference measurements [25, 26] and are accepted without any adjustment since the uncertainties are small and considered acceptable to yield accurate theoretical results. This alternative solves the problem of the determination of atomization energies of elements that under standard conditions are in the liquid or solid states.

The possibility of scaling the experimental atomization energies of the elements was considered in this work. The energy of each atom was multiplied by an adjustable parameter with an initial value of 1. The optimization minimized the mean absolute error considering the experimental enthalpies of formation from the G3/05 set along with the HLC parameters. The experimental atomization energies from the literature and the scaled ones are shown in Table 3. For these calculations the scaling for each atom and the HLC parameters were optimized simultaneously. Only six HLC parameters were optimized for these tests and are shown in Table 1 as EnAt1 and EnAt2. In order to avoid possible misunderstanding, EnAt1 and EnAt2 will be used in the sense that the experimental atomization energies were scaled as well as the six HLC parameters were optimized and they are associated with the G3X-CEP and G3X(CCSD)-CEP versions, respectively. Another important differentiation must be made between the atomization energy of the elements and the molecular atomization energies. The first one represents an experimental property used to calculate the enthalpy of formation. The second one is also used to calculate the enthalpy of formation, but it is an information to analyze the performance of the Gn composite methods for compounds containing elements from the third period of the periodic table.

Table 3 Experimental and scaled values of atomization energies of the elements at 0 K

Atoms	Experimental[a]	EnAt1[b]	EnAt2[c]
H	51.63 ± 0.001	51.57	51.61
Li	37.69 ± 0.2	37.56	38.30
Be	76.48 ± 1.2	77.32	77.16
B	136.2 ± 0.2	134.43	135.08
C	169.98 ± 0.1	170.18	170.01
N	112.53 ± 0.02	112.61	112.37
O	58.99 ± 0.02	59.35	59.17
F	18.47 ± 0.07	18.73	18.60
Na	25.69 ± 0.17	26.73	26.92
Mg	34.87 ± 0.2	37.60	37.39
Al	78.23 ± 1.0	83.12	82.76
Si	106.60 ± 1.9	107.57	107.28
P	75.42 ± 0.2	75.93	75.81
S	65.66 ± 0.06	66.52	66.12
Cl	28.59 ± 0.001	28.41	28.47
Br	28.18 ± 0.02	28.08	28.02

Data in kcal mol^{-1}

[a] Experimental enthalpy values taken from references [25, 26]

[b] Data optimized with respect to the G3X-CEP theory

[c] Data optimized with respect to the G3X(CCSD)-CEP theory

Table 1 shows that the optimized HLC parameters simultaneously with the atomization energies of the elements achieve a great similarity with the all-electron ones. The differences between HLC parameters for G3X-CEP and G3X(CCSD)-CEP calculations are small and do not cause significant changes in the statistical evaluation of properties.

Table 3 shows that the most significant differences in the modified atomization energies, EnAt1 and EnAt2, occur for Na, Mg, and Al. For EnAt1 the deviations are of the order of 4.0% for Na, 7.8% for Mg, and 6.3% for Al. Deviations for EnAt2 are of the same order: 4.8, 7.2, and 5.8% for Na, Mg, and Al, respectively. The deviations for other elements are significantly much smaller. Deviations above 1% using EnAt1 occur for Be (1.1%), B (1.3%), F (1.4%), and S (1.3%). In the case of EnAt2, only Li presents a deviation of 1.6%. All other elements require corrections lower than 1%.

The calculated properties with the scaled atomization energies are available as supplemental material (Tables S.3-S.7). The mean absolute errors for the G3/05 test set considering the EnAt1 (1.10 kcal mol^{-1}) and EnAt2 (1.12 kcal mol^{-1}) calculations are better than the G3X-CEP (1.18 kcal mol^{-1}) and G3X(CCSD)-CEP (1.15 kcal mol^{-1}) results. The modification of the experimental atomization energies directly affects the quality of the enthalpies of formation. The other properties are much less affected with deviations provided by the simultaneous optimization of the HLC parameters and the scaling of the atomization energies of the elements.

A better perspective of the effects of atomization energy adjustments and HLC parameters may be obtained by analyzing the errors for the different properties presented in Table 4. The enthalpies of formation present total mean absolute errors of 1.01 and 0.97 kcal mol^{-1} using EnAt1 and EnAt2, respectively. As shown previously, the mean absolute errors for G3X-CEP and G3X(CCSD)-CEP were larger, with values of 1.16 kcal mol^{-1} and 1.12 kcal mol^{-1}, respectively. The new error is identical to that obtained with the G4CEP calculations (1.00 kcal mol^{-1}). The improvement provided by EnAt1 and EnAt2 displaced several molecules to the uncertainty range of ± 1 kcal mol^{-1}, accumulating a total of 62 and 66%, respectively, of molecules in this interval and decreasing in the region above ± 2 kcal mol^{-1} to approximately 15%.

The mean absolute error improvement is mainly due to the better performance of 79 non-hydrogenated compounds. For these compounds the calculations with EnAt1 and EnAt2 presented errors of 1.50 and 1.48 kcal mol^{-1}, respectively, whereas for G3X-CEP and G3(CCSD)-CEP the same compounds presented errors of 1.76 and 1.78 kcal mol^{-1}, respectively. As examples, the molecule AlF_3 represents one of the most drastic changes with a deviation from the experimental value of 4.9 kcal mol^{-1} and 5.1 kcal mol^{-1} for G3X-CEP and G3(CCSD)-CEP, respectively, while applying

Table 4 Distribution considering different ranges of absolute error (in kcal mol^{-1}) of enthalpy of formation (ΔH_f^0), ionization energy (IE$_0$), electron affinity (EA$_0$), molecular atomization energy (D_0) and proton affinity (PA$_0$) calculated at the EnAt1 and EnAt2 levels of theory. Mean absolute error (MAE) and standard deviation (Std) are also shown in kcal mol^{-1}

Energy range	ΔH_f^0		IE$_0$		EA$_0$	
	EnAt1	EnAt2	EnAt1	EnAt2	EnAt1	EnAt2
$\Delta E < -2$	16	15	11	11	6	6
$-2 \leq \Delta E < -1$	33	31	11	11	13	12
$-1 \geq \Delta E \leq 1$	165	163	46	46	36	37
$1 < \Delta E \leq 2$	23	20	16	16	5	5
$\Delta E > 2$	21	19	17	17	1	1
Total	248		101		61	
% ($-1 \geq \Delta E \leq 1$)	62%	66%	46%	46%	59%	61%
% ($-2 < \Delta E$ & $\Delta E > 2$)	15%	15%	28%	28%	12%	12%
MAE	1.01	0.97	1.38	1.37	1.04	1.04
Std	1.46	1.41	1.76	1.76	1.23	1.23

Energy range	D_0		PA$_0$	
	EnAt1	EnAt2	EnAt1	EnAt2
$\Delta E < -2$	2	2	0	0
$-2 \leq \Delta E < -1$	4	4	0	0
$-1 \geq \Delta E \leq 1$	9	9	3	3
$1 < \Delta E \leq 2$	6	6	4	4
$\Delta E > 2$	1	1	1	1
Total	22		8	
% ($-1 \geq \Delta E \leq 1$)	41%	41%	38%	38%
% ($-2 < \Delta E$ & $\Delta E > 2$)	14%	14%	13%	13%
MAE	1.27	1.26	1.48	1.48
Std	1.67	1.65	1.44	1.44

the EnAt1 and EnAt2 energies these errors were reduced to 0.2 and 0.5 kcal mol^{-1}, respectively. NaF represented moderate changes with a deviation from the experimental value of 1.7 kcal mol^{-1} for G3X-CEP and G3(CCSD)-CEP, respectively, while applying EnAt1 and EnAt2 these errors were reduced to 0.7 and 0.5 kcal mol^{-1}, respectively. Compounds containing fluorine, chlorine, bromine, and aluminum generally show a good improvement with respect to the experimental data.

As expected, ionization energies, molecular atomization energies, and electron and proton affinities are less sensitive to the scaling of the experimental atomization energies of the elements. The mean absolute error for 101 ionization energies was 1.32 kcal mol^{-1} for both G3X-CEP and G3X(CSSD)-CEP and 1.38 and 1.37 kcal mol^{-1} for EnAt1 and EnAt2, respectively, which is essentially due to the new HLC parameters.

The mean absolute error for 61 electron affinities was 0.92 kcal mol^{-1} for G3X and G3X(CSSD)-CEP, while the error of 1.04 kcal mol^{-1} was achieved for EnAt1 and EnAt2. Similar to the ionization energies, this is only due to the HLC readjustment.

The molecular atomization energies presented a deviation of 1.32 kcal mol^{-1} for G3X-CEP and G3X(CSSD)-CEP, while for calculations using EnAt1 and EnAt2 the errors were reduced to 1.27 and 1.26 kcal mol^{-1}, respectively.

The eight proton affinities were not changed by the adjustment of EnAt1 and EnAt2 and HLC parameters.

6 Conclusion

The G3X and G3X(CCSD) composite theories were modified to consider the CEP pseudopotential. The modified versions were referred to as G3X-CEP and G3X(CCSD)-CEP. The G3/05 test set was used as reference to analyze the performance of these methods by calculating enthalpies of formation, ionization energies, electron affinities, proton affinities, and molecular atomization energies.

The optimization of four and six HLC parameters provided minor differences between both sets of calculations, but favored the six parameters, as expected. The set of six parameters improved the mean absolute error by approximately 0.02 kcal mol^{-1} compared with the mean error obtained by applying four parameters. The overall statistical behavior of G3X-CEP and G3X(CCSD)-CEP data considering six HLC parameters presented a mean absolute error around 1.15 kcal mol^{-1} for all calculated properties. These deviations are intermediate to G3CEP (1.29 kcal mol^{-1}) and G4CEP (1.09 kcal mol^{-1}) and also comparable with the tendency observed for the all-electron versions.

The scaling of the experimental atomization energies, referred to by EnAt1 and EnAt2, along with the optimization of the HLC parameters, produced mean absolute errors of 1.10 and 1.12 kcal mol^{-1}, respectively. These values are similar to the deviation of 1.09 kcal mol^{-1} obtained from the G4CEP theory. The enthalpies of formation were significantly improved, again presenting a similarity with results from the G4CEP theory. For 248 compounds from the G3/05 test set the absolute mean error of the calculations with EnAt1 (1.02 kcal mol^{-1}) and EnAt2 (0.98 kcal mol^{-1}) are significantly better than the G3X-CEP results (1.16 kcal mol^{-1}) and G3X(CCSD)-CEP (1.12 kcal mol^{-1}) with the original experimental atomization energies.

The molecular atomization energies also showed improvement in the deviations with respect to the experimental data, while ionization energies and electronic affinities were slightly affected indirectly by the new HLC parameters. The proton affinities were virtually unaffected by this correction.

The best composite method is the one where empirical adjustments may be neglected. In such a case, either an optimum cancelation of errors or a rigorous theoretical description of the electronic structure is achieved. The modification of the experimental atomization energies suggested in this work is not a desirable technique from the purely theoretical point of view. However, it is an additional source of information that either theoreticians or experimental researchers should consider in the search of accurate description of thermochemical properties.

Acknowledgements The authors would like to acknowledge financial support from FAPEAM (Fundação de Amparo à Pesquisa do Estado do Amazonas), FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo-Center for Computational Engineering and Sciences, Grants 2013/08293-7 and 2017/11485-6), and FAEPEX-UNICAMP (Fundo de Apoio ao Ensino, à Pesquisa e à Extensão da UNICAMP). The National Center of High Performance Computing in São Paulo (CENAPAD-SP) is acknowledged for access to their computational facilities.

References

1. Curtiss LA, Redfern PC, Raghavachari K, Pople JA (2002) Gaussian-3X (G3X) theory using coupled cluster and Brueckner energies. Chem Phys Lett 359:390–396. https://doi.org/10.1016/S0009-2614(02)00742-X
2. Curtiss LA, Redfern PC, Raghavachari K (2005) Assessment of Gaussian-3 and density-functional theories on the G3/05 test set of experimental energies. J Chem Phys 123:124107–124118. https://doi.org/10.1063/1.2039080
3. Pople JA, Head-Gordon M, Fox DJ et al (1989) Gaussian-1 theory: a general procedure for prediction of molecular energies. J Chem Phys 90:5622–5629. https://doi.org/10.1063/1.456415
4. Curtiss LA, Raghavachari K, Pople JA (1992) Gaussian-2 theory using reduced Moller–Plesset orders. J Chem Phys 98:1293–1298. https://doi.org/10.1063/1.464297
5. Curtiss LA, Raghavachari K, Redfern PC et al (1998) Gaussian-3 (G3) theory for molecules containing first and second-row atoms. J Chem Phys 109:7764–7776. https://doi.org/10.1063/1.477422
6. Curtiss LA, Redfern PC, Raghavachari K (2007) Gaussian-4 theory. J Chem Phys 126:084108–084119. https://doi.org/10.1063/1.2436888
7. Curtiss LA, Redfern PC, Raghavachari K (2011) G*n* theory. WIREs Interdiscip Rev Comput Mol Sci 1:810–825. https://doi.org/10.1002/wcms.59
8. Curtiss LA, Raghavachari K, Redfern PC, Pople JA (2000) Assessment of Gaussian-3 and density functional theories for a larger experimental test set. J Chem Phys 112:7374. https://doi.org/10.1063/1.481336
9. Dorofeeva OV, Moiseeva NF (2006) Computational study of the thermochemistry of organophosphorus(III) compounds. J Phys Chem A 110:8925–8932. https://doi.org/10.1021/jp060982f
10. Wang L, Lv G (2010) Prediction of gas-phase thermodynamic properties for polychlorinated naphthalenes using G3X model chemistry and density functional theory. Chemosphere 78:77–85. https://doi.org/10.1016/j.chemosphere.2009.10.041
11. He Y-L, Wang L (2009) Potential energy surfaces for protonation of hydrochlorofluoromethanes. J Mol Struct THEOCHEM 913:240–246. https://doi.org/10.1016/j.theochem.2009.08.007
12. Wang L, Liu Y (2010) The enthalpies of formation of brominated benzenes and phenols: A theoretical prediction. J Mol Struct THEOCHEM 957:72–76. https://doi.org/10.1016/j.theochem.2010.07.009
13. Akhmetova N, Kaliyeva L, O′Reilly RJ (2016) Assessment of quantum chemical methods for the calculation of homolytic N–F bond dissociation energies. Chem Data Collect 5–6:28–35. https://doi.org/10.1016/j.cdc.2016.10.003
14. Pereira DH, Ramos AF, Morgon NH, Custodio R (2011) Implementation of pseudopotential in the G3 theory for molecules containing first-, second-, and non-transition third-row atoms. J Chem Phys 135:034106–034120. https://doi.org/10.1063/1.3609241
15. Pereira DH, Ramos AF, Morgon NH, Custodio RR (2011) Erratum: "Implementation of pseudopotential in the G3 theory for molecules containing first-, second-, and non-transition third-row atoms" [J. Chem. Phys. 135, 034106 (2011)]. J Chem Phys 135:219901. https://doi.org/10.1063/1.3666235
16. Stevens WJ, Basch H, Krauss M (1984) Compact effective potentials and efficient shared-exponent basis sets for the first- and second-row atoms. J Chem Phys 81:6026–6033. https://doi.org/10.1063/1.447604
17. Stevens WJ, Krauss M, Basch H, Jasien PG (1992) Relativistic compact effective potentials and efficient, shared-exponent basis sets for the third-, fourth-, and fifth-row atoms. Can J Chem 70:612–630. https://doi.org/10.1139/v92-085
18. Pereira DH, Ducati LC, Rittner R, Custodio R (2014) A study of the rotational barriers for some organic compounds using the G3 and G3CEP theories. J Mol Model 20:2199–2212. https://doi.org/10.1007/s00894-014-2199-3
19. Rocha CMR, Pereira DH, Morgon NH, Custodio R (2013) Assessment of G3(MP2)//B3 theory including a pseudopotential for molecules containing first-, second-, and third-row representative elements. J Chem Phys 139:184108–184119. https://doi.org/10.1063/1.4826519
20. Rocha CMR, Rodrigues JAR, Moran PJS, Custodio R (2014) An interpretation of the phenol nitration mechanism in the gas phase

using G3(MP2)//B3-CEP theory. J Mol Model 20:2524–2531. https://doi.org/10.1007/s00894-014-2524-x

21. Pereira DH, Rocha CMR, Morgon NH, Custodio R (2015) G3(MP2)-CEP theory and applications for compounds containing atoms from representative first, second and third row elements of the periodic table. J Mol Model 21:204–210. https://doi.org/10.1007/s00894-015-2757-3
22. de Silva CS, Pereira DH, Custodio R (2016) G4CEP: A G4 theory modification by including pseudopotential for molecules containing first-, second- and third-row representative elements. J Chem Phys 144:204118–204126. https://doi.org/10.1063/1.4952427
23. Heerdt G, Pereira DH, Custodio R, Morgon NH (2015) W1CEP theory for computational thermochemistry. Comput Theor Chem 1067:84–92. https://doi.org/10.1016/j.comptc.2015.05.011
24. Nelder JA, Mead R (1965) A simplex method for function minimization. Comput J 7:308–313. https://doi.org/10.1093/comjnl/7.4.308
25. Curtiss LA, Raghavachari K, Redfern PC, Pople JA (1997) Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation. J Chem Phys 106:1063–1079
26. Absar I, van Wazer JR (1971) The barrier to internal rotation in monomethylphosphine: an ab initio LCAO-MO-SCF study. J Chem Soc D Chem Commun. https://doi.org/10.1039/c29710000611