Fitting Atomic Correlation Parameters for RECEP (Rapid Estimation of Correlation Energy from Partial Charges) Method to Estimate Molecular Correlation Energies within Chemical Accuracy

SÁNDOR KRISTYÁN,[1] GÁBOR I. CSONKA[2]

[1]*Department of Chemistry, George Washington University, 725 21st Street, NW, Washington, DC 20052*
[2]*Department of Inorganic Chemistry, Budapest University of Technology, H-1521 Budapest, Hungary*

ABSTRACT: The accuracy of the RECEP method [Chem Phys 1997, 224, 33 and Chem Phys Lett 1999, 307, 469] has been increased considerably by the use of fitted atomic correlation parameters. This method allows an extremely rapid, practically prompt calculation of the correlation energy of molecules after an HF-SCF calculation. The G2 level correlation energy and HF-SCF charge distribution of 41 closed-shell neutral molecules (composed of H, C, N, O, and F atoms) of the G2 thermochemistry database were used to obtain the fitted RECEP atomic correlation parameters. Four different mathematical definitions of partial charges, as a multiple choice, were used to calculate the molecular correlation energies. The best results were obtained using the natural population analysis, although the other three are also recommended for use. For the 41 molecules, the G2 results were approached within a 1.8 kcal/mol standard deviation (the mean absolute difference was 1.5 kcal/mol). The RECEP atomic correlation parameters were also tested on a different, nonoverlapping set of other 24 molecules from the G2 thermochemistry database. The G2 results of these 24 molecules were approached within a 2.3 kcal/mol standard deviation (the mean absolute difference was 1.9 kcal/mol). This method is recommended to estimate total correlation energies of closed shell ground-state neutral molecules at stationary (minimums and transition states) points on the potential surface. Extension of the work for charged molecules, radicals, and molecules containing

other atoms is straightforward. Numerical example as a recipe is also provided.
© 2000 John Wiley & Sons, Inc.   J Comput Chem 22: 241–254, 2001

Keywords: correlation energy; partial charges

Introduction

The term "electron correlation energy" (E_{corr}) for the electronic ground state is usually defined[1] as the difference between the exact nonrelativistic full-CI (configurations interactions) electronic ground-state energy of the system [E_0(CI)], and the single determinant ground-state HF-SCF (Hartree–Fock self-consistent field) basis set limit energy approximation (E_0(HF-SCF)):

$$E_{corr} = E_0(CI) - E_0(HF\text{-}SCF). \tag{1}$$

The HF theory is well defined and unique for closed-shell molecules. We used the restricted Hartree–Fock (RHF) in our calculation where the spatial orbitals for α and β spins are identical. Although E_{corr} depends[2–5] mainly (quasi-linearly) on the number of electrons, N, (more precisely on $N - 1$) the considerably smaller dependence on the nuclear configuration must be considered to obtain chemical accuracy.

The efficient calculation of correlation energy plays a central role in modern theoretical chemistry.[1–7] The well-known, very expensive approximations for correlation energy are summarized elsewhere, for example, in refs. 1–9. These calculations have various problems, as large disc space demand, slow convergence, even divergence is not uncommon. The very poor convergence and the $O(n^5) - O(n^7)$ scaling[10] with the basis set increase led to the development of the so-called composed (or extrapolation) methods, for example, G2.[11, 12] Considerably faster alternatives for the estimation of the correlation energy are the density functional theory (DFT) methods,[13–24] for example, the widely used B3LYP.[11] The recent correlation functionals provide usually rather good results, and show a basis set convergence similar to the HF-SCF method. However, it should be noted that the DFT methods also introduce new problems: there is no simple way to improve the results (cf. use of semiempirical functionals), and numerical instabilities might occur as well (cf. numerical integrals).

In our previous works[4, 5] we analyzed the applicability of the RECEP method to several small ions and closed-shell molecules without any fitting procedure. The results showed that although the G2 energy was approximated qualitatively correctly, the chemical accuracy (1–2 kcal/mol) was not reached. Now we refine the RECEP method (via its parameters) using a considerably larger subset of the G2 thermochemistry database.[25] Forty-one closed-shell molecules were used to obtain the necessary "RECEP atomic correlation parameters"[4, 5] and another 24 closed-shell molecules were used to test the RECEP method with these new parameters. We have performed HF-SCF/6-311+G(2d,p) and G2 calculations, using the G2 geometries.[25] The ChelpG,[26] Merz–Kollman (MK),[27] Mulliken,[1] and natural population analysis (NPA)[28] partial charges were calculated from the HF-SCF wave function using the Gaussian 94[11] program package. We can conclude that any reasonable partial charge definition combined with a suitably fitted "RECEP atomic correlation parameters" (i.e., the fit must use the same partial charge definition as well) is adequate for estimating molecular correlation energies.

For various estimations of the molecular correlation energy we use the following notations: E_{corr}(G2), E_{corr}(RECEP), etc., where the method is in the argument, while E_{corr} itself notates the accurate complete-CI value as defined in eq. (1). For atomic correlation energies (e.g., refs. 29–31) we use the E_{corr}(CI, N, Z) or E_{corr}(B3LYP, N, Z) notations, where the method, the number of electrons (N), and the atomic charge (Z) are also indicated in the argument.

The Rapid Estimation of Correlation Energy from the Partial Charges (RECEP) Method

Empirically, the correlation energy of any N electronic system can be roughly approximated[1–5] by the following linear relationship:

$$E_{corr} \approx a\,(N - 1), \tag{2}$$

where $-0.035 \geq a[\text{hartree}] \geq -0.045$. The CI[29–31] and B3LYP results[2–5] for atoms ($1 \leq N \leq Z \leq 18$) justify eq. (2). Although the linear approximation in eq. (2) cannot reach the chemical accuracy, it can serve as a basis for the RECEP method.[4, 5] The

essence of the RECEP method is that a quasi-linear dependence of the correlation energy on the (fractional) number of electrons is conserved in molecules. The RECEP formula to estimate the molecular E_{corr} value is the following:

$$E_{corr}(RECEP) \equiv \Sigma_{(A=1,M)}E_{corr}(N_A, Z_A). \quad (3)$$

In eq. (3), N_A is the electron content around atom A, calculated from an adequately defined partial charge. The N_A values in eq. (3) are not necessarily integer numbers, while the Z_As (atomic charge) are. The summation in eq. (3) runs for all M atoms in the molecule. The two basic assumptions of eq. (3) can be summarized as follows: the correlation energy is the sum of the "RECEP atomic correlation energies or parameters" and the value of the latter can be estimated from the partial charges. It is important to note that the "RECEP atomic correlation energies or parameters," $E_{corr}(N_A, Z_A)$, in a molecule are necessarily different from the atomic correlation energies of the corresponding atoms in free space, although they have similar values, or at least, trends (see refs. 4 and 5 or Table I). Equation (3) is supposed to be used for molecules in the vicinity of equilibrium nuclear configurations and transition states, i.e., in chemical bond, but not for van der Waals or London dispersion forces. However, in a limit case, when the molecule is infinitely separated into its constituent neutral or ionic atoms (nonbond state on the potential energy surface), eq. (3) provides a CI quality correlation calculation if the $E_{corr}(CI, N, Z)$ values are chosen for $E_{corr}(N_A, Z_A)$ with $N = N_A$ and $Z = Z_A$ for all M atoms. This limit behavior may also indicate why eq. (3) is able to provide high-quality correlation calculation for molecules using proper $E_{corr}(N_A, Z_A)$ values representing atoms in equilibrium chemical bond (bond state on the potential energy surface). (We use the following notation: $E_{corr}(CI, N, Z)$ notes the CI correlation energy of atoms in free space—as above, and $E_{corr}(N_A, Z_A)$ notes the RECEP atomic correlation parameters for eq. (3) for a chosen partial charge definition and noninteger N_A.)

At this point we recall that partial charges are essentially mathematical constructions that serve to reflect the electron content around the selected atom of the molecule. These partial charges are not physically measurable quantities, as they depend on how we define the partition scheme of the electron density. However, partial charges can be defined to reproduce the measurable dipole moment of the molecule, for example. Their primary use is to help chemists to establish empirical rules, for example, for describing chemical bonds. Indeed,

the partial charges might differ substantially depending on the mathematical formula—but again, when it is defined, the definition is unique. In our previous work,[5] we considered the Mulliken charges,[1] the charges derived from the electrostatic potential (ESP),[32, 33] the natural population analysis (NPA),[28] and the charges derived from Bader's population analysis.[34] The advantages and disadvantages of these charge definitions are discussed in the literature, and will not be discussed here. Although any partial charge can be used in a fit for $E_{corr}(N_A, Z_A)$ in eq. (3), we selected four widely used ones: ChelpG,[26] Merz–Kollman (MK),[27] Mulliken,[1] and NPA[28] partial charges as it was mentioned in the introduction. The ChelpG charges are widely used in molecular mechanics programs, the Mulliken charges are very easy to calculate, and the NPA charges have very good basis set convergence properties.

The use of the HF-SCF one-electron density, $\rho_{HF-SCF}(\mathbf{r})$, to derive electron correlation effects can be justified readily. It was observed that DFT correlation energy functionals provide adequate correlation energy using moderate basis set for $\rho_{HF-SCF}(\mathbf{r})$.[2, 3, 20] Similar argument[4, 5] holds for the HF-SCF level partial charges used in eq. (3). This observation originates in the fact that the E_{corr} in DFT is an integrated quantity with respect to the $\rho(\mathbf{r})$, and the one-electron density integrates to the number of electrons in any case. Thus, more accurate electron density causes relatively small change in E_{corr}.

An important question is how to calculate the $E_{corr}(N_A, Z_A)$ terms in eq. (3). As an initial guess, these terms can be derived from the CI correlation energy of the atomic cations, neutral atoms, and anions in free space. The noninteger N_A can be treated by linear interpolation between nearby integer values ($N \leq N_A \leq N + 1$) as proposed in refs. 4 and 5. The data of atomic correlation energies in free space for integer (N, Z) pairs can be found in refs. 29–31. However, these atomic values are not the best choice for eq. (3). The reason for this was analyzed in refs. 4 and 5 in detail. The results show that the correlation energy is very sensitive to spin pairing effects (the opposite spin electrons have different correlation energy than the parallel spin electrons.[5] To illustrate this, we present the exact full CI correlation energy differences for the first row elements in Figure 1. The figure shows how the correlation energy changes as a result of adding the electrons one by one to the 2s (Li, Be) and 2p (B-Ne) shells. We present three alternative ways to add one electron to these shells.

TABLE I.

Atomic Correlation Parameters, $E_{corr}(N_A, Z_A)$, in hartree for eq. (3) to Obtain E_{corr}(RECEP) Estimation for Molecular Correlation Energies for Closed-Shell, Singlet, Ground-State Neutral Molecules in the Vicinity of Stationary Points.

Atom	N[a]	Z	Corrected			Fitted Parameters[e]			
			Exact[b]	Exact[c]	B3LYP[d]	ChelpG[f]	MK[g]	Mulliken	NPA[h]
H	2	1	−0.0395	−0.0395	−0.0432	−0.0406	−0.0408	−0.0397	−0.0376
C	4	6	−0.1264	−0.1264	n.a.	n.a.	n.a.	−0.1171	−0.1105
	5	6	−0.1388	−0.1388	−0.1400	−0.1385	−0.1381	−0.1423	−0.1387
	6	6	−0.1564	−0.1754	−0.1911	−0.1653	−0.1650	−0.1660	−0.1659
	7	6	−0.1827	−0.2087	−0.2258	−0.1868	−0.1864	−0.1866	−0.1909
N	6	7	−0.1666	−0.1856	−0.2005	−0.2259	−0.2265	−0.2515	−0.2227
	7	7	−0.1883	−0.2143	−0.2373	−0.2281	−0.2280	−0.2240	−0.2259
	8	7	−0.2617	−0.2877	−0.3035	−0.2333	−0.2333	−0.2302	−0.2351
	9	7	n.a.	n.a.	−0.3622	n.a.	n.a.	n.a.	−0.3804
O	8	8	−0.2579	−0.2839	−0.3079	−0.2692	−0.2690	−0.2712	−0.2703
	9	8	−0.3314	−0.3314	−0.3619	−0.2743	−0.2738	−0.2646	−0.2790
F	9	9	−0.3245	−0.3245	−0.3599	−0.2901	−0.2903	−0.2879	−0.2892
	10	9	−0.3995	−0.3995	−0.4430	−0.2956	−0.2940	−0.3048	−0.3061

(While values under footnotes b and d are true atomic correlation energies in free space as well, the values under footnotes c and e are RECEP atomic correlation parameters only to estimate molecular correlation energies.) Values under footnotes b, c, and d can be used with any partial charge but give only moderate results via eq. (3) (cf. Fig. 2), while for values under footnote e the type of partial charge is restricted. However, all these four under footnote e have about the same quality, and can provide the chemical accuracy via eq. (3) (cf. Fig. 3).

[a] N is the number of electrons and Z is the nuclear charge for the $E_{corr}(N_A, Z_A)$ values.

[b] CI correlation energies for "high-spin" ground-state atoms in free space (RHF) can be found in refs. 29–31. N.a. = not available (double anions are not stable[4] in free space).

[c] Corrected atomic CI correlation energies (cf. footnote b): −0.019 hartree correction for six electronic systems, −0.026 hartree correction for 7 and 8 electronic systems,[5] as a quick adjustment to eq. (3).) N.a. = same as in footnote b. Using these for $E_{corr}(N_A, Z_A)$ is called RECEP-c in the text.

[d] Atomic correlation energies for singlet or doublet (not necessarily ground state) "low-spin" state (e.g., for carbon the $1s^2 2s^2 p_x^2$ singlet low-spin state) systems in free space calculated as a difference of B3LYP/6-311+G(3df,2pd) and HF/6-311+G(3df,2pd) energies. N.a. = it was not necessary to calculate for the molecules used in ref. 5 and here. Using these for $E_{corr}(N_A, Z_A)$ is called RECEP-d in the text.

[e] Optimized parameters, calculated with a linear fit of eq. (3) to the correlation energies of the 41 molecules in Table III. N.a. = not available, because partial charge values requiring these (N_A, Z_A) value pairs did not occur in the used 41 molecules during the linear fit. (However, a proper larger set can target these and other, unlisted values.) Using these for $E_{corr}(N_A, Z_A)$ in the case of different partial charges (for a multiple choice) is called RECEP-fit atomic correlation parameters in the text. It is also important that the HF-SCF/6-311+G(2d,p) level partial charges (and HF energies) are recommended to calculate for eq. (3) (i.e., in this way the deviation from the basis set limit is incorporated in E_{corr}(RECEP) as well). Table II shows how to use these parameters. These values are the main results of this work.

[f, g, h] partial charges available, for example, in package in ref. 11, based on ref. 26, ref. 27, and ref. 28, respectively, when estimating $E_{corr}(N_A, Z_A)$ values. For the Mulliken charge see refs. 1 and 11.

First we show the correlation energy difference for cations and neutral atoms (cf. "cat. neut." in Fig. 1): $E_{corr}(\text{CI}, N' = Z, Z) - E_{corr}(\text{CI}, N = Z - 1, Z)$, where $Z = 3 \ldots 10$ [e.g., if $Z = 3$ (Li) we show the correlation energy difference between Li and Li$^+$, etc.]. Next, we show the correlation energy difference for cations (cf. "cat. cat." in Fig. 1): $E_{corr}(\text{CI}, N' = Z, Z + 1) - E_{corr}(\text{CI}, N = Z - 1, Z)$, where $Z = 3 \ldots 10$ [e.g., if $Z = 3$ (Li) we show the correlation energy difference between Be$^+$ and Li$^+$, etc.]. The third case is for neutral atoms (cf. "neut. neut." in Fig. 1): $E_{corr}(\text{CI}, N' = Z, Z) - E_{corr}(\text{CI}, N = Z - 1, Z - 1)$, where $Z = 3 \ldots 10$ [e.g., if $Z = 3$ (Li) we show the correlation energy difference between Li and He, etc.]. In summary, Figure 1 shows that adding the first electron to the 2s shell (Li) has little effect on the correlation energy. Adding the second electron to the 2s shell (Be) effects a considerably greater change in the correlation energy [cf. the three bars under Be in Figure 1, i.e., $E_{corr}(\text{CI}, 4, 4) - E_{corr}(\text{CI}, 3, 4)$, $E_{corr}(\text{CI}, 4, 5) - E_{corr}(\text{CI}, 3, 4)$, and $E_{corr}(\text{CI}, 4, 4) -$
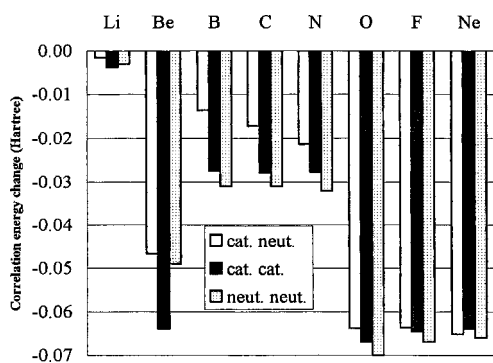
FIGURE 1. The exact ground-state correlation energy differences for the elements of the second row of the periodic table. These differences are calculated in three different ways for each element. Cat. neut. in the legend notes the correlation energy difference for the corresponding cation and neutral atom (e.g., for Li it notes the correlation energy difference between Li^+ and Li). Cat. cat. notes the correlation energy difference between the neighboring cations (e.g., for Li it notes the correlation energy difference between Li^+ and Be^+). Neut. neut. notes the correlation energy difference between the neighboring neutral elements (e.g., for Li it notes the correlation energy difference between Li and He). In this way the figure represents the correlation energy change due to adding one electron to the 2s (Li, Be) or 2p (B-Ne) shells. The figure shows the importance of the spin pairing effect in the correlation energy and its Z dependence (e.g., the spin pairing occurs in the 2s shell of Be and in the 2p shells of O, F, and Ne). The values were derived from the exact ground-state correlation energies by Davidson.[28-30]

$E_{corr}(CI, 3, 3)$]. The analysis of Figure 1 reveals that the dominant effect of the correlation energy change comes from the spin pairing effect (a considerably smaller Z dependence can also be seen in Fig. 1). This indicates that accurate high spin atomic correlation energies [$E_{corr}(CI, N, Z)$] in free space cannot provide adequate values for atomic correlation parameters [$E_{corr}(N_A, Z_A)$] in eq. (3) for molecules in the vicinity of their equilibrium nuclear configuration.

Table I shows some suggested RECEP atomic correlation parameters. The first two derived from high-spin full-CI atomic correlation energies by a quick correction[5] (column "corrected exact") as well as low spin B3LYP atomic correlation energies (column "B3LYP"). These two choices are called as RECEP-c, and RECEP-d, respectively, as in ref. 5. In more detail: instead of using the correlation energy of high spin atomic states in free space (e.g., triplet $1s^2 2s^2 2p_x 2p_y$ for Carbon atom, cf. footnote b or column "Exact" in Table I), we proposed[4, 5] to

use the correlation energy of the excited or low spin states (e.g., singlet $1s^2 2s^2 2p_x^2$ for Carbon atom) for $E_{corr}(N_A, Z_A)$ values in eq. (3) to obtain better results for near equilibrium molecular environments (cf. footnote d or column "B3LYP" in Table I). The correlation energy of these low-spin states (column "B3LYP" in Table I) can easily be approximated using calculated atomic values in free space, for example, with package in ref. 11. In our previous study,[5] we compared the high-spin and low-spin B3LYP atomic correlation energies in free space. For 6, 7, and 8 electronic systems the differences between the high-spin and low-spin correlation energies are approximately -0.019, -0.026, and -0.026 hartree, respectively, these were used for the "quick correction" to yield the "Corrected exact" (footnote c) column in Table I. The values in Table I no longer refer to the individual atomic ground or excited states in free space (except columns "Exact" and "B3LYP" which are); i.e., these are not "atomic correlation energies," these are only "atomic correlation parameters" to be used in eq. (3) for $E_{corr}(N_A, Z_A)$. These parameters are used to estimate molecular correlation energies in the vicinity of stationary points in ground state. RECEP-c parameters provide slightly better results than RECEP-d parameters (cf. Table III); however, the results are far from the chemical accuracy. These parameters RECEP-c and RECEP-d can be calculated for the whole periodic system (although for atoms with $Z > 18$ some ralativistic correction is necessary). Similar parameters can be obtained from other reliable methods, for example CCSD(T)[1, 9, 11] results as well, instead of using B3LYP like above. However, columns under footnotes b, c, and d in Table I provide only moderate results via eq. (3) in this way (b is the weakest), because it relies on purely atomic parameters when predicting molecular bonds. But, at least, they can justify the power of eq. (3).

Alternatively, a fitting procedure can be applied (*vide infra*) for $E_{corr}(N_A, Z_A)$ parameters in eq. (3). This way one can obtain the optimal values for the "RECEP atomic correlation parameters" from the known (e.g., G2) molecular correlation energies for a larger set of small molecules (e.g., the G2 molecule set). It was forecasted in refs. 4, 5 that this fitting procedure would provide the most accurate results for molecular correlation energies via eq. (3). Using the fitted parameters (noted as RECEP-fit) eq. (3) approaches chemical accuracy (cf. footnote e in Table I and RECEP-c, RECEP-d, RECEP-fit results in Table III).

In Table II we provide a numerical example of RECEP method, as a recipe, for methyl-nitrite mole-

TABLE II.

The Numerical Details Necessary for the Calculation of the Molecular Correlation Energy [cf. eq. (3)] of Methylnitrite (CH_3—O—N=O) Using Fitted RECEP Atomic Correlation Parameters obtained from G2 Correlation Energies and HF-SCF NPA Partial Charges (cf. Footnote h in Table I).[a]

Z_A	Parc.chrg.$_A$	$N1$	N_A	$N2$	$E_{corr}(N1, Z_A)$	$E_{corr}(N2, Z_A)$	$E_{interpolated,A} \equiv E_{corr}(N_A, Z_A)$
6	−0.133	6	6.133	7	−0.1659	−0.1909	−0.1692
8	−0.490	8	8.490	9	−0.2703	−0.2790	−0.2746
1	0.171	0	0.829	2	0.0	−0.0376	−0.0156
1	0.165	0	0.835	2	0.0	−0.0376	−0.0157
1	0.165	0	0.835	2	0.0	−0.0376	−0.0157
7	0.504	6	6.496	7	−0.2227	−0.2259	−0.2243
8	−0.382	8	8.382	9	−0.2703	−0.2790	−0.2736

[a] Z_A denotes the nuclear charge. The partial NPA charge[28] (Parc.chrg.$_A$) in a.u. was calculated at the HF-SCF/6-311+G(2d,p) level, and $N_A = Z_A -$ Parc.chrg.$_A$. $N1$ is the integer part of N_A, and $N2 = N1 + 1$, for hydrogen ($Z_A = 1$):[39] $N1 = 0$ and $N2 = 2$. $E_{interpolated,A}(N_A, Z_A) = (N_A - N1)E_{corr}(N2, Z_A) + (N2 - N_A)E_{corr}(N1, Z_A)$ cf. ref. 5, and text. For hydrogen atoms[39] the atomic correlation parameter $E_{corr}(1, 1) = E_{corr}(2, 1)/2$ and $E_{interpolated,A}(N_A, Z_A = 1) = N_A E_{corr}(N2 = 2, Z_A = 1)/2$. $E_{corr}(N1, Z_A)$ and $E_{corr}(N2, Z_A)$ are from column h of Table I. E_{corr}(RECEP-fit) $= \Sigma_{(A)} E_{interpolated,A} = -0.9886$ hartreee in accord with eq. (3) (cf. G2 value in Table III line 41). The use of any other type of partial charges under footnote e in Table I provides a similar result.

cule. In this example we use the best, the linearly fitted "RECEP atomic correlation parameters" for the NPA charges (column h in Table I). Notice that a HF-SCF calculation with partial charge calculation (NPA) on this molecule is the only *ab initio* calculation, and it demands much less disc space and CPU time than, for example, a G2 calculation. The calculation of E_{corr}(RECEP-fit) by eq. (3) can be done even on a pocket calculator as follows here: The atomic nuclear charges (Z_A), and the NPA partial charges[28] (Parc.chrg.$_A$) of the atoms in the molecule are listed in Table II. The NPA charges are calculated with Gaussian 94[11] at HF-SCF level using 6-311+G(2d,p) basis set (as in the fitting procedure). The equilibrium geometry of the molecule is taken from the G2 molecular database.[25] In this article we focus on the accuracy of RECEP method; thus, we do not address the problems arising from the geometry optimization. The HF/6-31G(d) geometries are not generally reliable enough, because they do not contain the correlation effects. In this respect the DFT method can provide a decent geometries with small basis set [B3LYP/6-31G(d)],[40] however, this level of DFT does not provide reliable thermochemistry.

The column N_A is the electron content around the atoms as $N_A = Z_A -$ Parc.chrg.$_A$; notice that this definition is fundamental in the calculation and in the fitting procedure. E_{corr}(RECEP-fit) $= \Sigma_{(A)} E_{corr}(N_A, Z_A) = -0.9886$ hartree in accordance with eq. (3), where $E_{corr}(N_A, Z_A) \equiv (N2 - N_A)E_{corr}(N1, Z_A) + (N_A - N1)E_{corr}(N2, Z_A)$ interpolated for noninteger N_A and integer $Z_A > 1$.[4, 5] For hydrogen atoms[5] $E_{corr}(N_A, Z_A = 1) \equiv N_A E_{corr}$

$(N2 = 2, Z_A = 1)/2$ for noninteger $0 < N_A < 2$. (These two definitions of interpolation were used in the fitting procedure as well.) As it was mentioned earlier, this linear interpolation[4, 5] provides reasonable correlation energy via eq. (3).[39] (This linear interpolation may be replaced by an interpolating function; however, this question will be investigated elsewhere.) For $Z_A > 1$, the $N1$ is the integer part of the N_A, and $N2 = N1 + 1$; thus, $N1 \leq N_A \leq N2$, and for $Z_A = 1$ (Hydrogen) $N1 = 0$ and $N2 = 2$. (Index A is omitted for convenience for $N1$ and $N2$.) The $E_{corr}(N1, Z_A)$ and $E_{corr}(N2, Z_A)$ values for the integer $N1$ and $N2$ are the fitted RECEP atomic correlation parameters $[E_{corr}(N_A, Z_A)]$ shown in Table I for NPA charges. In the linear interpolation $E_{corr}(N1 = 0, Z_A = 1) = 0.0$ represents zero correlation energy contribution in the absence of electrons. The RECEP-fit instead of RECEP in the argument of E_{corr}(RECEP-fit) emphasizes that $E_{corr}(N_A, Z_A)$ values come from a linear fit using known (here G2) correlation energies, the topic of this work. E_{corr}(G2) $= -0.9875$ hartree, and the deviation is E_{corr}(G2) $- E_{corr}$(RECEP-fit) $= 0.001131$ hartree ≈ 0.7 kcal/mol. This is an excellent agreement, as indicated in Table III (cf. row 41).

In Table III we show the result for the set of 41 neutral closed-shell molecules composed of H, C, N, O, and F atoms, selected from the G2 thermochemistry database.[25] We selected those molecules for which the most reliable experimental heat of formations were available. We excluded the complications arising from the spin pairing effects (radicals), molecular charges, relativistic effects (third-row el-

TABLE III.

The Set of 41 Molecules,[25] Their E_{corr}(G2) Values Were Used to Obtain the Fitted $E_{corr}(N_A, Z_A)$ RECEP Atomic Correlation Parameters for eq. (3) Listed in Column e in Table I.

Molecule	E_0(HF-SCF)	E_{corr}(G2)	Correlation Energy Differences,[a] E_{corr}(G2) − E_{corr}(RECEP)		
			RECEP-d	RECEP-c	RECEP-fit
1 Methane (CH_4)	−40.2102	−0.2433	27.2	13.0	1.3
2 Ammonia (NH_3)	−56.2150	−0.2767	44.5	32.0	0.0
3 Water (H_2O)	−76.0527	−0.2999	50.8	30.8	−0.9
4 Hydrogenfluoride (HF)	−100.0526	−0.3063	68.7	43.2	0.4
5 Acetylene (C_2H_2)	−76.8422	−0.3698	38.4	16.6	1.5
6 Ethylene ($H_2C=CH_2$)	−78.0584	−0.4064	44.4	20.4	3.0
7 Ethane ($H_3C—CH_3$)	−79.2541	−0.4480	48.0	21.7	1.7
8 Hydrogencyanide (HCN)	−92.8979	−0.4031	36.8	14.4	1.9
9 Formaldehyde ($H_2C=O$)	−113.9033	−0.4617	54.8	29.2	1.7
10 Methanol ($CH_3—OH$)	−115.0815	−0.5028	68.0	36.4	0.8
11 Hydrazine ($H_2N—NH_2$)	−111.2174	−0.5149	65.3	39.4	−0.3
12 Hydrogenperoxide (HO—OH)	−150.8235	−0.5684	75.7	40.6	0.2
13 Carbondioxide (CO_2)	−187.6892	−0.6835	82.3	30.1	−0.2
14 CF_4	−435.7780	−1.3054	240.1	107.9	0.6
15 COF_2	−311.7100	−0.9954	160.3	68.7	−0.3
16 N_2O	−183.7207	−0.7274	40.6	−1.9	−1.5
17 NF_3	−352.6474	−1.1020	156.1	75.4	1.1
18 C_2F_4 ($F_2C=CF_2$)	−473.5672	−1.4763	245.1	142.8	−1.8
19 CF_3CN	−428.6011	−1.4111	206.7	101.8	−1.3
20 Propyne (C_3H_4)	−115.8984	−0.5744	60.1	26.4	2.3
21 Allene (C_3H_4)	−115.8970	−0.5739	59.2	26.2	2.3
22 Cyclopropene (C_3H_4)	−115.8553	−0.5800	55.5	21.7	−1.8
23 Propylene (C_3H_6)	−117.1082	−0.6132	64.6	28.6	2.3
24 Cyclopropane (C_3H_6)	−117.0916	−0.6175	62.3	26.3	−0.2
25 Propane (C_3H_8)	−118.2994	−0.6551	67.6	29.4	0.7
26 Trans-butadiene	−154.9667	−0.7793	80.2	34.4	2.2
27 Dimethylacetylene (2-butyne)	−154.9525	−0.7795	81.3	35.5	2.6
28 Methylenecyclopropane (C_4H_6)	−154.9303	−0.7843	77.8	32.0	−0.6
29 Bicyclobutane	−154.9120	−0.7916	73.0	27.2	−5.3
30 Cyclobutane (C_4H_8)	−156.1390	−0.8257	80.5	32.5	−2.2
31 Isobutene (C_4H_8)	−156.1577	−0.8219	83.5	35.6	0.5
32 Trans-butane (C_4H_{10})	−157.3446	−0.8626	87.0	36.7	−0.6
33 Isobutane (C_4H_{10})	−157.3451	−0.8649	85.8	35.6	−1.9
34 Spiropentane (C_5H_8)	−193.9673	−0.9957	94.9	37.1	−4.3
35 Benzene (C_6H_6)	−230.7633	−1.1134	111.1	45.7	1.1
36 Difluoromethane (H_2CF_2)	−237.9779	−0.7722	134.6	79.8	1.7
37 Trifluoromethane (HCF_3)	−336.8798	−1.0393	185.5	104.3	0.8
38 Methylamine ($H_3C—NH_2$)	−95.2473	−0.4811	60.3	35.3	−1.2
39 Acetonitrile ($CH_3—CN$)	−131.9605	−0.6062	58.5	26.0	3.2
40 Nitromethane ($CH_3—NO_2$)	−243.7359	−0.9920	88.3	30.5	−0.6
41 Methylnitrite ($CH_3—O—N=O$)	−243.7366	−0.9875	89.9	31.7	0.7

The HF/6-311+G(2d,p) level total energies, E_0(HF-SCF), and G2 correlation energies, E_{corr}(G2), are in hartree, while the correlation energy differences are in kcal/mol.

[a] Simple addition of HF-SCF total and G2 correlation energy yields the total G2 energy, E_0(G2), for ground electronic state [see eq. (1)], thus the energy differences directly provide the deviation compared to the G2 energy (E_{corr}(G2) − E_{corr}(RECEP) = E_0(G2) − E_0(RECEP) and E_0(RECEP) = E_0(HF-SCF) + E_{corr}(RECEP) in accord with eq. (1). The basis set error is incorporated into the E_{corr}(RECEP-fit) correlation energy as well; see text. [The geometry optimization was on MP2/6-31G(d) level.] The HF/6-311+G(2d,p) NPA charges[11, 28] were used for all three columns in these correlation calculations, and in this way column h of Table I was used for the RECEP-fit. For RECEP-d and RECEP-c, any partial charge[5] can be used, but they do not reach chemical accuracy, while RECEP-fit does.

ements), and the experimentally under represented elements in the molecules (e.g., Li and B). We focus on most important organic molecules. The inclusion of the S-containing molecules will be investigated in the future studies. For the present study it was necessary to separate the G2 total energy, E_0(G2), into HF-SCF, E_0(HF-SCF), and correlation, E_{corr}(G2) parts. For HF-SCF energies[5] we selected the HF/6-311+G(2d,p) single-point energies calculated at the G2 [e.g., MP2=full/6-31G(d)] equilibrium geometries. Subtracting this HF-SCF energy from the G2 total energy yields the G2 correlation energy, E_{corr}(G2), cf. eq. (1). (Obviously, all the zero point energies and thermal corrections are left out, because these are not necessary for the present purpose.) If we increase the quality of the basis set used in the HF-SCF method, the HF-SCF total energy would decrease, and this would provide slightly different G2 correlation energy. (In this respect the HF-SCF basis set limit total energy would play a special role; we shall return to this question later in this article.) However, the HF-SCF limit calculations are very demanding computationally; thus, only small molecules could be treated this way. It is advisable to use the smallest but yet good-quality basis set possible. The G2 estimation for E_0 [the E_0(G2) in eq. (1)] is also not full-CI quality yet, only a relatively good (and expensive) compromise. In our future work, we plan to perform a full-CI level and HF-SCF limit calculations to obtain the true E_{corr} via eq. (1). Knowing the true correlation energy would lend a possibility to control the quality of the present fitted atomic correlation parameters and would provide the basis for development of a new fitted $E_{corr}(N_A, Z_A)$ parameter set if necessary. (On the other side, including the basis set error in the parameter set values is also worth to investigate, because in this way one does not have to reach the HF limit.)

In Table III we show the HF-SCF total and G2 correlation energies according to the definitions above for 41 molecules of the G2 database. We also show the RECEP-d, RECEP-c, and RECEP-fit correlation energy differences (in kcal/mol) between the G2 and the RECEP estimations of the correlation energies for the same molecules. The RECEP correlation energies were calculated using the NPA charges (cf. Table I). Although the differences are fairly big in case of RECEP-c and RECEP-d, the inspection of Figure 2 reveals a very good linear relationship between the G2 and RECEP correlation energies ($R^2 \approx 0.99$). The essential difference is that the G2 correlation energy is considerably smaller in absolute value than the RECEP estimation of the correla-
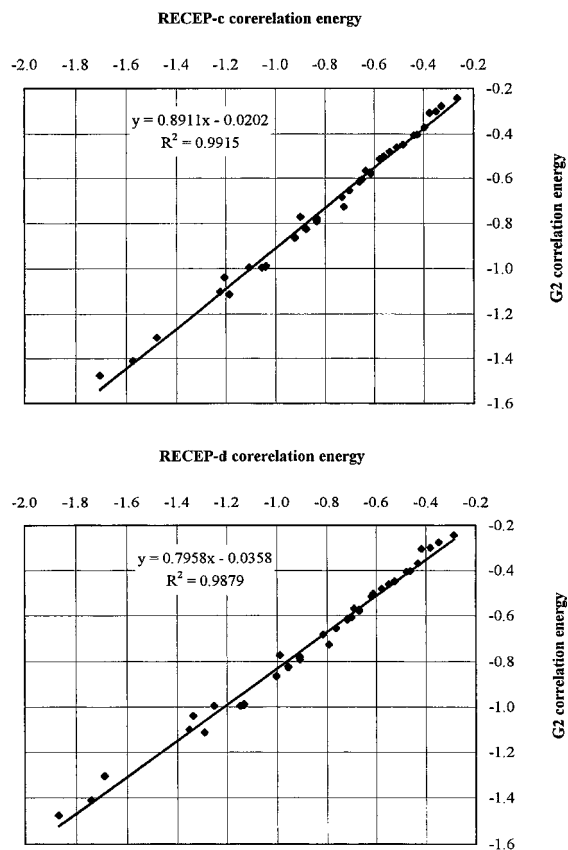

FIGURE 2. Statistical data for E_{corr}(RECEP-c) and E_{corr}(RECEP-d) correlation energies vs. G2 correlation energy (E_{corr}(G2)) in hartree listed in Table III.

tion energies (by a factor of 0.8911 and 0.7958 for RECEP-c and RECEP-d, respectively, cf. Fig. 2). Although the linear fit between the G2 and RECEP correlation energies is qualitatively good (cf. Fig. 2) it stops short of the required chemical accuracy. Using the linear equation parameters in Figure 2 the largest difference is reduced to 40 kcal/mol and the standard deviation is about 18 kcal/mol. This is a considerable improvement compared to the errors in Table III; however, this clearly shows that such an *a posteriori*, two-parameter linear fit cannot provide accurate results for a larger set of molecules.

Fitting the RECEP Atomic Correlation Parameters

In this section we shall investigate the optimization of $E_{corr}(N_A, Z_A)$ parameters to reach the nearest chemical accuracy with eq. (3). We name these atomic correlation parameters as RECEP-fit parameters (cf. column e in Table I and the RECEP-fit

TABLE IV.
The Set of 24 Molecules (Not Included in the Linear Fit) Used to Test the (Self-Consistency of) RECEP Atomic Correlation Parameters in Table I (cf. Footnote e and h).

Molecule	E_0(HF-SCF)	E_{corr}(G2)	Correlation Energy Difference, E_{corr}(G2) $-$ E_{corr}(RECEP-fit)
1 HCOOH (formic acid)	-188.8266	-0.7230	-0.4
2 HCOOCH$_3$ (methyl formate)	-227.8588	-0.9296	-1.2
3 CH$_3$CONH$_2$ (acetamide)	-208.0454	-0.9055	-1.7
4 C$_2$H$_4$NH (aziridine)	-133.0800	-0.6519	-3.0
5 NCCN (cyanogen)	-184.6243	-0.7785	-1.4
6 (CH$_3$)$_2$NH (dimethylamine)	-134.2834	-0.6882	-1.6
7 CH$_3$CH$_2$NH$_2$ (*trans*-ethylamine)	-134.2953	-0.6882	-2.0
8 CH$_2$CO (ketene)	-151.7713	-0.6284	1.9
9 C$_2$H$_4$O (oxirane)	-152.9137	-0.6748	-2.4
10 CH$_3$CHO (acetaldehyde)	-152.9638	-0.6665	2.2
11 HCOCOH (glyoxal)	-226.6589	-0.8878	2.2
12 CH$_3$CH$_2$OH (ethanol)	-154.1321	-0.7091	0.3
13 CH$_3$OCH$_3$ (dimethylether)	-154.1148	-0.7088	0.3
14 CH$_2$=CHF (vinyl fluoride)	-176.9412	-0.6734	2.0
15 CH$_2$=CHCN (acrylonitrile)	-169.8097	-0.7733	2.5
16 CH$_3$COCH$_3$ (acetone)	-192.0208	-0.8731	1.7
17 CH$_3$COOH (acetic acid)	-227.8859	-0.9276	0.4
18 CH$_3$COF (acetyl fluoride)	-251.8768	-0.9340	0.9
19 (CH$_3$)$_2$CHOH (isopropanol)	-193.1819	-0.9181	-1.8
20 C$_2$H$_5$OCH$_3$ (methyl ethyl ether)	-193.1655	-0.9154	-0.5
21 (CH$_3$)$_3$N (trimethylamine)	-173.3212	-0.8984	-4.3
22 C$_4$H$_4$O (furan)	-228.6888	-1.0114	-4.6
23 C$_4$H$_5$N (pyrrole)	-208.8676	-0.9900	-5.5
24 C$_5$H$_5$N (pyridine)	-246.7597	-1.1468	-0.3

HF/6-311+G(2d,p) total energies, E_0(HF-SCF), and G2 correlation energies, E_{corr}(G2), are in hartree. Correlation energy differences between the G2 and the RECEP-fit estimations of correlation energies are in kcal/mol (footnote a in Table III holds here as well).

values in Table III). As it was indicated in the previous section, the next logical step is to optimize the RECEP-c and -d values further for the atomic correlation parameters to obtain the best possible fit to G2 molecular correlation energy. As was noted earlier, the G2 correlation energy, E_{corr}(G2), is certainly not the best possible choice for the correlation energy, yet provides near chemical accuracy energies. We have chosen the G2 method, because it is well tested on a large set of molecules, and full-CI or CCSD(T) basis set limit energies are not readily available for a larger set of molecules. However, as soon as better results than G2 are available, *vide infra* the RECEP atomic correlation energy parameters can be fitted accordingly.

The following criteria were used to obtain the fitted atomic $E_{corr}(N_A, Z_A)$ values in eq. (3). $Y \equiv$ minimum $\Sigma_{(i=1,41)}(E_{corr}(G2)_i - E_{corr}(RECEP)_i)^2$. $E_{corr}(RECEP)_i = \Sigma_{(A)}[(N_A - N1)_i E_{corr}(N2, Z_A) + (N2 - N_A)_i E_{corr}(N1, Z_A)]$, where index A is running over the atoms of ith molecule (see explanation in

Table II). This system of linear equations require L number of $E_{corr}(N, Z)$ parameters. Let us denote the kth $E_{corr}(N, Z)$ parameter by a_k and regroup the terms to yield $E_{corr}(RECEP)_i = \Sigma_{(k=1,L)}x_{ki}a_k$, where x_{ki}'s are the factors ($N_A - N1$ and $N2 - N_A$) arising from the summation of fractional electron numbers requiring the same a_k parameter in the molecule. The $\partial Y / \partial a_k = 0$ minimization yields a system of linear equations: $a_1 \Sigma_{(i=1,41)}x_{1i}x_{ki} + a_2 \Sigma_{(i=1,41)}x_{2i}x_{ki} + \cdots + a_L \Sigma_{(i=1,41)}x_{Li}x_{ki} = \Sigma_{(i=1,41)}y_i x_{ki}$ with $k = 1, 2, \ldots, L$, and $y_i \equiv E_{corr}(G2)_i$. The solution yields the desired a_k, i.e., $E_{corr}(N_A, Z_A)$, values (listed under column e in Table I for four different partial charges). Depending on the partial charge definitions used for the fitting procedure we obtained $L = 11$–13 atomic correlation parameters, cf. Table I. The values under footnote e in Table I were obtained in less than one second on a 25 Mflop computer.

Optimizing the RECEP atomic correlation parameters compensates for the errors arising from the spin pairing effects, the partial charges, and limited

basis set used for HF-SCF energy calculation. The results in Tables III and IV (cf. column RECEP-fit) show that the fitted $E_{corr}(N_A, Z_A)$ RECEP atomic correlation parameters provide good results for closed-shell ground-state neutral molecules at stationary points. In this way, the RECEP method can reach its practical limit.

Optimally this practical limit approximates sufficiently the G2 total energy after a simple HF-SCF partial charge and energy calculation. These calculations are four to five orders of magnitude faster than the more expensive G2 calculations. (Preliminary research, not shown here, indicates that different parametrization of eq. (3) is necessary for calculating correlation energy of radicals (e.g., $NH_2^{\bullet}$, $CH_3^{\bullet}$, etc.); however, the concept is useable. We think this is in agreement with the known fact that the HF-SCF method itself needs two different kinds of "wave function treatments:" one for closed-shell molecules, and another for radicals.)

To obtain the RECEP-fit correlation energies for molecules in Table III, we used the NPA charges, and these linearly fitted RECEP atomic correlation parameters shown in Table I (cf. column h). The same fitting procedure was repeated for the ChelpG, MK, and Mulliken charges, and the resulting parameters are also shown in Table I. According to our results (not shown here for the sake of brevity) these different definitions of partial charges did not lead to significantly different results, which is why we show only the NPA results for RECEP-fit in Table III. The root-mean-square deviations from the G2 total energy for the 41 molecules listed in Table III are 2.0, 2.0, 2.1, and 1.8 kcal/mol for ChelpG, MK, Mulliken, and NPA charges, respectively (the mean absolute difference (MAD) is 1.5 kcal/mol in the latter case). The G2 correlation energy (and total energy) is approximated usually within 3 kcal/mol, with the exceptions of bicyclobutane and spiropentane (-5.3 and -4.3 kcal/mol deviations, respectively, cf. Table III). This shows that the spiro compounds probably require a different parameter set for the correlation energy.

The quality of the RECEP atomic correlation parameters was checked against the G2 total energies of other 24 molecules that were not used in the fitting procedure. The results in Table IV show that the G2 total energy is sufficiently approximated in this test set as well. This indicates the reliability of the present RECEP-fit results. And more importantly, the RECEP atomic correlation parameters in eq. (3) can be considered quasi-constants within the chemical accuracy for closed-shell singlet ground-state neutral molecules in the vicinity of their station-
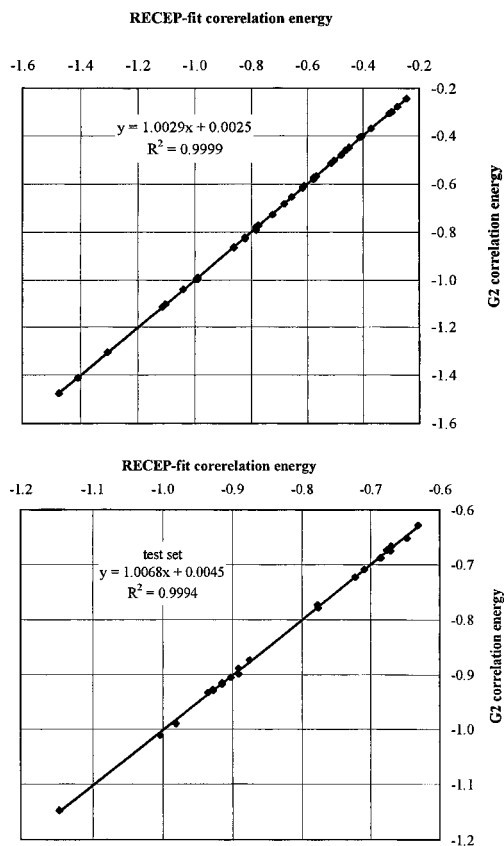

FIGURE 3. Statistical data for E_{corr}(RECEP-fit) correlation energies vs. G2 correlation energy (E_{corr}(G2)) in hartree listed in Tables III and IV (upper figure: training set of 41 molecules, lower figure: test set of 24 molecules).

ary points (minimums and transition states). The largest difference is -5.5 kcal/mol (cf. Table IV), the standard deviation is 2.3 kcal/mol, and the MAD is 1.9 kcal/mol using the NPA charges and fitted parameters (column h in Table I). (Figure 3 shows how good is the agreement between the G2 and the RECEP-fit correlation energies.) In summary, a comparison of Figure 2 to Figure 3 and the RECEP columns in Tables III–IV shows that RECEP-c and RECEP-d do not achieve the chemical accuracy, while RECEP-fit does (more precisely, the RECEP-fit results approximate with MAD < 2 kcal/mol the prestigious G2 results).

Note on Improved Approximations to the Correlation Energy

As mentioned, the G2 total energies are of high quality and easily available,[11] although, further improvements are necessary and conceivable. Bench-

TABLE V.

The Set of 17 Molecules with Known HF-SCF, G2, and CCSD(T) Total Energies,[a] as Well as G2 and CCSD(T) Limit (as "Best Estimate" in Fig. 4) Correlation Energies (hartree).

	HF/6-311+G(2d,p)	HF-SCF/limit	G2	CCSD(T) Limit + Core Correlation	Correlation Energy	
					G2	Limit
1 H_2	−1.1325	−1.1337	−1.1758	−1.1746	−0.0433	−0.0409
2 CH	−38.2754	−38.2799	−38.4188	−38.4719	−0.1434	−0.1920
3 CH_3	−39.5705	−39.5767	−39.7727	−39.8286	−0.2022	−0.2519
4 CH_4	−40.2102	−40.2172	−40.4535	−40.5080	−0.2433	−0.2908
5 NH_3	−56.2150	−56.2251	−56.4917	−56.5573	−0.2767	−0.3323
6 H_2O	−76.0527	−76.0677	−76.3526	−76.4317	−0.2998	−0.3640
7 HF	−100.0526	−100.0714	−100.3589	−100.4528	−0.3063	−0.3814
8 C_2H_2	−76.8422	−76.8554	−77.2120	−77.3220	−0.3699	−0.4666
9 C_2H_4	−78.0584	−78.0709	−78.4648	−78.5751	−0.4064	−0.5042
10 CO	−112.7699	−112.7912	−113.1825	−113.3119	−0.4125	−0.5207
11 N_2	−108.9680	−108.9928	−109.3982	−109.5280	−0.4302	−0.5352
12 H_2CO	−113.9033	−113.9235	−114.3650	−114.4951	−0.4616	−0.5717
13 O_2	−149.6337	−149.6680	−150.1523	−150.3120	−0.5186	−0.6440
14 F_2	−198.7380	−198.7744	−199.3265	−199.5150	−0.5885	−0.7406
15 CO_2	−187.6892	−187.7253	−188.3727	−188.5798	−0.6835	−0.8545
16 *trans*-butadiene[b]	−154.9667	−154.9929	−155.7460	−155.9666	−0.7793	−0.9737
17 benzene[b]	−230.7633	−230.8013	−231.8767	−232.2103	−1.1133	−1.4090

[a] For HF-SCF limit the TQ5 extrapolation, and for CCSD(T) limit the Q5 extrapolation were used (cf. text, the values were derived from the non relativistic values by J. M. L. Martin[36]).

[b] Simplified W1 extrapolation scheme was used.

mark quality *ab initio* total atomization energies (TAE) became recently available for small neutral molecules up to *trans*-butadiene and benzene;[35] this provides an opportunity to judge the quality of the HF-SCF/6-311+G(2d,p) and the G2 total energies. It should be noted that in the original article the extrapolation for the infinite basis set was done for the TAE and not for the component total energies (the explicit total energies are not given). Thus, to obtain the HF-SCF and CCSD(T) limit total energy components we used the appropriate total energies of the corresponding atoms published on the Internet.[36] In Table V we show the results for the selected 17 molecules containing H, C, N, O, and F atoms. We note that the recalculated CCSD(T) total energies do not necessarily agree exactly with the true extrapolated limit energies. In Table V we present the sum of core correlation and CCSD(T) limit valence correlation energy. The inclusion of the core correlation contributes considerably to the differences between the G2 and the CCSD(T) limit + core correlation energies.

To see the possible errors we compared the true extrapolated HF-SCF and CCSD(T) total energies to our results in Table V. We selected the HF (hydrogen-fluorid) molecule for this test. The HF-SCF extrapolated total energy after a TQ5 extrapolation is −100.0714 hartree (the $E_\infty = E[5] − (E[5] − E[Q])^2/(E[5] − 2E[Q] + E[T])$ formula was used, according to the $A + B/C^n$ type extrapolation,[37] where $n = 3, 4, 5$) that agrees well with the value in Table V for HF molecule. The extrapolated CCSD(T) valence correlation + core correlation energy is −100.4555 hartree (the $E_\infty = E[Q] + (E[Q] − E[T])/((4/3)^3 − 1)$ formula[38] was used for the valence correlation, according to the $A + B/n^3$ type extrapolation, where $n = 3, 4$, and −0.0579 hartree core correlation energy was added[36]). This shows a slight (1 milli-hartree) difference compared to the value in Table V.

Comparison of the HF-SCF energies in Table V shows that there is a simple linear relationship between HF/6-311+G(2d,p) and HF-SCF limit total energies: E_0(HF-SCF limit) $= 1.000178 \cdot E_0$(HF/6-311+G(2d,p)) within ±5 kcal/mol error bar for the above mentioned 17 molecules (r.m.s. deviation is 2 kcal/mol, MAD is 1.4 kcal/mol). The largest errors were observed for O_2 (4.8 kcal/mol) and for benzene (−4.2 kcal/mol).

Comparison of the correlation energy defined as a difference of the G2 total energy and the HF/6-311+G(2d,p) energy, and the estimate of the correlation energy after the extrapolation to the infinite
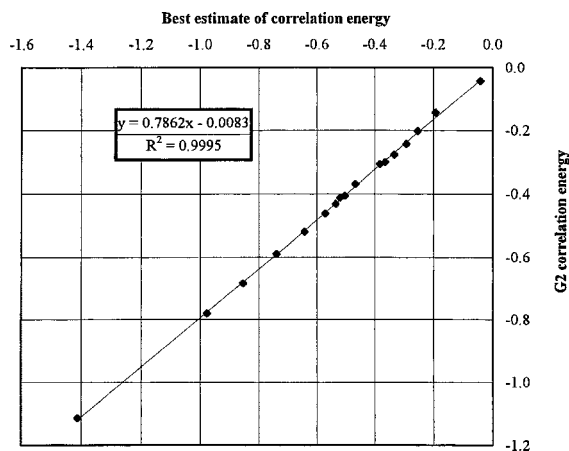
FIGURE 4. Statistical data for best estimate (cf. Table V) vs. G2 correlation energy (hartree).

basis set with the core correlation energy added, provides that the G2 correlation energy is about the 79% of the latter one. The statistical correlation between the two correlation energies is excellent (cf. $R^2 = 0.9995$ in Fig. 4). The G2 method provides a reliable thermochemistry; however, the correlation energies derived from G2 calculations should be scaled up. This is in line with our RECEP-c and -d results (cf. Fig. 2) and with the recent G3 total energies available on the Internet.[25]

Finally, we note that it is possible to derive very precise correlation energies from the experimental TAE values. After a zero point vibration energy correction and a relativistic correction, the correlated total energy can be obtained. If the molecular geometry is known, an extrapolated HF-SCF limit total energy can be calculated. The difference of the two energies could provide good-quality, well-defined correlation energy. A fitting procedure for these correlation energies could provide the ultimate RECEP atomic correlation parameters. However, the requirement of the HF-SCF limit total energy makes such calculations rather expensive for large molecules. Thus, currently it is advisable to perform only relatively rapid HF-SCF calculations [e.g., a HF-SCF/6-311+G(2d,p)] and absorb the imperfections into the RECEP atomic correlation parameters, as was done in the present paper (cf. values under footnote e in Table I).

Conclusions

The quasi-linear dependence of the correlation energy on the partial charges in closed-shell singlet ground-state neutral molecules was used and ana-

lyzed in eq. (3) to obtain RECEP atomic correlation parameters for rapid and accurate estimations of molecular correlation energies. It is literally a density functional method, based on a radically new approach, using a very simple relationship between the partial charges and atomic correlation energies (parameters) in the molecules. The total correlation energy of the molecule is approximated as the sum of the atomic correlation energies (parameters). We have proposed four alternative sets of parameters for four different type of partial charges (Table I) to use in eq. (3) as a multiple choice. These sets can provide approximately the same quality results. The atomic correlation parameters were obtained from a linear regression of eq. (3) for its $E_{corr}(N_A, Z_A)$ RE-CEP atomic correlation parameters using G2 level correlation energies of 41 closed-shell ground-state small molecules from the "G2 molecule set." We can conclude that these parameters are suited to estimate molecular correlation energies of any small or large closed-shell ground-state molecules at stationary points (minimums and transition states) until an HF-SCF/6-311+G(2d,p) calculation can be done. Using our method, the CPU time and disc space demands for a "close to CI quality" E_0 are practically limited only by the HF-SCF calculation. The four types of chosen partial atomic charges used for eq. (3) can be derived from the analysis of HF/6-311+G(2d,p) wave function, readily available in the Gaussian94 package.[11] The results show that our E_{corr}(RECEP) estimation for molecular correlation energies has a standard deviation of 1.8 Kcal/mol (in case of NPA partial charges) compared to the G2 correlation energies. These correlation calculations are always stable, and need neither CPU time nor disc space. Future consideration should be given to atoms, which also occur frequently in organic, inorganic, and biomolecules, such as sulfur (S) and chlorine (Cl). The inclusion of charged molecules into the parametrization is also feasible. (These latter two extensions can be achieved with simply using larger molecule set than the one in Table III for the fit, including molecules containing other atoms (S, Cl, etc.) and/or that are charged. It extends the (N, Z) domain of Table I.)

We think that the extension of this method is straightforward for radicals (e.g., open-shell doublet ground-state molecules), only the RECEP atomic correlation parameters (cf. Table I) should be fitted (spin corrections) for them and treated as a separate parameter set. (This is in agreement with the known fact that the HF-SCF error is different for closed and open shells.) In summary, our hypothesis is strongly supported by our experience

such that, within the chemical accuracy, RECEP atomic correlation parameters in eq. (3) constitute a quasi-constant set for closed shell (singlet) ground-state molecules in the vicinity of stationary points (cf. footnote e in Table I). However, another quasi-constant set describes the open shell (e.g., doublet) ground-state radicals, and both of these sets can be obtained from a fit described in this work. Furthermore, even the error of failing to reach the limit of the HF-SCF calculation can be incorporated in these sets of parameters by fixing the type of basis. We plan to perform RECEP calculations for spin corrections (E_{corr}) of doublet radicals, as well as to obtain more atom types than are listed in Table I.

Supplementary Material

A simple Fortran program that performs the linear fit for eq. (3) to obtain $E_{corr}(N_A, Z_A)$ RECEP atomic correlation parameters (for any number of molecules for which an adequate level correlation energy and a HF-SCF level partial charges (any) are precalculated with any fixed basis set as input) can be downloaded from http://web.inc.bme.hu/~kristyan, or can be obtained via e-mail (kristyan@gwu.edu or kristyan@web.inc.bme.hu) or surface mail. An input for this program is also available, which uses the 41 molecules in Table III with their NPA charges and G2 correlation energies (calculated by us) for the linear fit, and calculates the E_{corr}(RECEP) correlation energy for any additional input molecules based on eq. (3) and Table II using precalculated HF-SCF/6-311+G(2d,p) level NPA partial charges by the user. (It reproduces column h of Table I in every run as well.) On a 20-Mflop machine or faster it runs almost instantly without considerable disc space demand.

Acknowledgment

References

1. Szabo, A.; Ostlund, N. S. Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory; McMillan: New York, 1982.
2. Kristyan, S. Chem Phys Lett 1995, 247, 101.
3. Kristyan, S. Chem Phys Lett 1996, 256, 229.
4. Kristyan, S. Chem Phys 1997, 224, 33.
5. Kristyan, S.; Csonka, G. I. Chem Phys Lett 1999, 307,469.
6. Kristyan, S. Comput Phys 1994, 8, 556.
7. Peng, Z.; Kristyan, S.; Kuppermann, A.; Wright, J. Phys Rev A 1995, 52, 1005.
8. Moller, C.; Plesset, M. S. Phys Rev 1934, 46, 618.
9. Cizek, J. Adv Chem Phys 1969, 14, 35.
10. Raghavachari, K.; Anderson, J. B. J Phys Chem 1996, 100, 129.
11. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Wong, M. W.; Foresman, J. B.; Robb, M. A.; Head–Gordon, M.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. Gaussian 94, Revision F2; Gaussian, Inc.: Pittsburgh, PA, 1993.
12. Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. J Chem Phys 1991, 94, 7221.
13. Lee, C.; Yang, W.; Parr, R. G. Phys Rev B 1988, 37, 785.
14. Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Chem Phys Lett 1989, 157, 200.
15. Becke, A. D. J Chem Phys 1988, 88, 1053.
16. Becke, A. D. J Chem Phys 1988, 88, 2547.
17. Becke, A. D. J Chem Phys 1992, 96, 2155.
18. Becke, A. D. J Chem Phys 1992, 97, 9173.
19. Becke, A. D. J Chem Phys 1993, 98, 5648.
20. Parr, R. G.; Yang, W. Density-Functional Theory of Atoms and Molecules; Oxford University Press: New York, 1989.
21. Kristyan, S.; Pulay, P. Chem Phys Lett 1994, 229, 175.
22. Kristyan, S. J Chem Phys 1995, 102, 278.
23. Dirac, P. A. M. Proc Cambridge Philos Soc 1930, 26, 376.
24. Wigner, E. P. Trans Faraday Soc 1938, 34, 678.
25. http://chemistry.anl.gov/compmat/comptherm.htm.
26. Breneman, C. M.; Wiberg, K. B. J Comput Chem 1990, 11, 361.
27. Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. J Comput Chem 1990, 11, 431.
28. Reed, A. E.; Weinstock, R. B.; Weinhold, F. J. J Chem Phys 1985, 83, 735.
29. Davidson, E. R.; Hagstrom, S. A.; Chakravorty, S. J.; Umar, V. M.; Fischer, C. F. Phys Rev A 1991, 44, 7071.
30. Chakravorty, S. J.; Gwaltney, S. R.; Davidson, E. R.; Parpia, F. A.; Fischer, C. F. Phys Rev A 1993, 47, 3649.
31. Davidson, E. R. www home page: http://ezinfo.ucs.indiana.edu/~davidson and http://phb.indiana.edu/~davidson/atom/tub.3.
32. Biosym Technologies, 9685 Scranton Road, San Diego, CA 92121-2777, USA: Turbomole version 2.3 program, 1993 December; InsightII/Turbomole/Parameters/Properties pulldown: "ESP Charges" for van der Waals fit to the molecular wave function to calculate partial charges in the molecule.
33. Chirlian, L. E.; Francl, M. M. J Comput Chem 1987, 8, 894.
34. Bader, R. F. W. Adv Quant Chem 1981, 14, 63.
35. Martin, M. L.; Oliveira, G. J Chem Phys 1999, 111, 1843.
36. http://winston.weizmann.ac.il/~comartin/wl/.
37. Feller, D. J Chem Phys 1992, 96, 6104.
38. Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Chem Phys Lett 1998, 286, 243.
39. As it can be seen, the hydrogen atoms in molecules need special treatment[5] in eq. (3). It comes from the fact, that

$E_{corr}(CI, N_A = 0, Z_A = 1) = E_{corr}(CI, N_A = 1, Z_A = 1) = 0$ for free proton (H^+) and free hydrogen atom (H), respectively, i.e., no correlation effect because no electron pair in the system at all. However, electrons around an H atom in a stable molecular bond, even with near to zero partial charge (i.e., Parc.chrg.$_A \approx 0$ or $N_A \approx 1$ on the H atom) exhibit substantial correlation effect, because all electrons are in interaction with the other $N - 1$ electrons. It follows that for an H atom with any $0 < N_A < 1$ electron content on it, like in Table II, the interpolation, based on the zero correlation effect in free proton and free H atom would yield a poor $E_{corr}(N_A, Z_A = 1) = 0$ value.[4] A more educated choice for the interpolation is one, for example, between free proton ($E_{corr}(N_A = 0, Z_A = 1) = 0$) and a H^- anion

($E_{corr}(N_A = 2, Z_A = 1) < 0$), or equivalently between free proton and a "half" H^- anion ($E_{corr}(N_A = 2, Z_A = 1)/2$) values, where the latter is the correlation energy of "H atom in molecular bond with zero partial charge on it." The values for $E_{corr}(N_A = 2, Z_A = 1)$ are around -0.04 hartree [Table I, eq. (2)], and in Table II the case of NPA charge (see -0.0376 hartree in Table I) is taken. It must also be noted that the H atom has a very frequent occurrence (Tables II–V) in chemically important molecules. Furthermore, in the case of neutral molecules, the electron content on H atoms in a stable molecular bond is mostly between $0.7 < N_A < 1$ as in the case of Methylnitrite (Table II).

40. Baboul, A. G.; Curtiss, A. L.; Redfern, P. C. J Chem Phys 1999, 110, 7650.