

Rapid Estimation of Enthalpies of Formation from Hartree–Fock Total Energy and Partial Charges for Compounds Containing Si, S, and Cl Atoms

Adrienn Ruzsinszky* and Gábor I. Csonka*

Department of Inorganic Chemistry, Budapest University of Technology and Economics,
H-1521 Budapest, Hungary

Received: February 17, 2003; In Final Form: May 19, 2003

New atomic energy parameters for rapid estimation of basis set error, the correlation energy, zero-point energy, and thermal and relativistic corrections were obtained for HF/6-311+G(2d,p) and HF/6-31G(d) energies and atomic partial charges (natural population analysis and Mulliken). A total of 161 closed-shell neutral molecules from the G2/97 and G3/99 database composed of H, C, N, O, F, Si, S, and Cl atoms were selected to test the performance of the rapid estimation of enthalpies of formation from Hartree–Fock (REEF-HF) method. The best REEF-HF results were obtained using the HF/6-311+G(2d,p) total energies and the corresponding natural population analysis charges. The average absolute deviation from the experimental enthalpies of formation for 159 molecules is 1.55 kcal/mol. Leaving out the five most problematic molecules yielded 1.38 kcal/mol average absolute deviation for the remaining 154 molecules. This compares well to our previous best results on a smaller test set of 117 molecules composed of molecules containing only first-row elements (1.48 kcal/mol average absolute deviation). The REEF-HF method works well for molecules containing second-row elements. Our method compares favorably to G3 method in terms of computational time and to B3LYP method in terms of computational time and precision. For a set of 27 simple hydrocarbons the proposed method yielded 0.77 or 0.64 kcal/mol average absolute deviations using HF-SCF/6-31G(d) energies and NPA or Mulliken charges, respectively.

Introduction

Estimations of enthalpies of formation from Hartree–Fock self-consistent field (HF-SCF) total energies with the necessary zero-point energy and thermal corrections usually provide poor results. The reasons are known and well understood. However, the HF-SCF error is systematic enough (if the basis set quality is adequate); thus, simple methods that use HF-SCF total energy and extra parameters might provide reasonably good results. A successful method was proposed by Wiberg,¹ who used non-relativistic HF-SCF/6-31G(d) total electronic energies for a variety of hydrocarbons and derived five energy parameters for CH₃, CH₂, CH, C(saturated), and C(olefinic) groups to reproduce experimental enthalpies of formation. An estimate of the enthalpy of formation of any hydrocarbon may then be obtained as the difference of its calculated HF-SCF/6-31G(d)//HF-SCF/6-31G(d) total electronic energy and the sum of the appropriate group energy parameters. The differences between estimated and observed values were on the order of 2 kcal/mol.¹ Benzene was a notable exception with more than 10 kcal/mol deviation.¹ Later, this procedure was modified and atomic energy parameters were proposed to correct nonrelativistic HF-SCF/6-31G(d) total energies leading to average absolute (a.a.) deviations of 2 kcal/mol for enthalpies of formation calculated in this manner.² A total of 56 parameters were proposed originally for H, C, N, O, and F atoms in various molecular environments;² that number was later reduced³ to 14. The reasonable success of these methods also shows that rapid estimation of enthalpies of formation from HF-SCF energy is feasible for classical molecules. These methods treat implicitly the correlation energy, the basis set error, the zero-point energy, the thermal, and the relativistic corrections. The extension of such methods requires

new group definitions, and the definition of the group is not unique (one could define new groups to obtain better agreement with the experiment, and this procedure is almost limitless, e.g., introduction of a new CH(aromatic) energy parameter would solve the above-mentioned problem of Wiberg's method with benzene). We note that simple additivity methods exist⁴ (for about 400 groups) that rely on the structure of the molecular graphs and do not require any HF-SCF or other quantum chemical calculations.

Dewar et al.⁵ proposed an analogous but more convenient procedure called single-atom equivalent correction (the original purpose was to estimate the effective errors in ab initio and semiempirical energies). In this procedure the empirical corrections depend only on the atoms present in the molecule. The advantage of this approach is its generality and simplicity. The disadvantage is that such corrections cannot distinguish better between structural isomers than the uncorrected parent method. Dewar's calculations using HF-SCF/6-31G(d) energies and single-atom equivalent corrections showed that this approach has about 6 kcal/mol a.a. deviation for neutral molecules composed of H, C, N, and O atoms. For cations, the method has about 23 kcal/mol a.a. deviation.⁵ It was found later that the combination of Dewar's procedure with the energies of B3PW91^{6,7} or B3LYP^{8,9} DFT methods (using the 6-311G(d,p) and 6-31G(d) basis sets) reproduces the enthalpies of formation of many hydrocarbons satisfactorily.¹⁰ We note that cations were not included in this latter study.

After analyzing the nature of the correlation energy (the error of the HF-SCF energy compared to the exact energy) and its dependence on the electron content in the free atoms and ions, we developed a simple and rapid (practically instant) energy correction method for HF-SCF energies based on the depen-

dence of the correlation energy on partial atomic charges (RECEP).^{11–13} Next, we suggested an analogous method to treat basis set error (REBECEP).^{14–16} Finally, the rapid estimation of enthalpies of formation from HF-SCF total energy and partial charges (REEF-HF) method was developed for closed-shell classical molecules composed of H, C, N, O, and F atoms.¹⁷ This method unites the simplicity of Dewar's method⁵ with the precision of the method proposed by Schleyer et al.² without the necessity of non-quantum mechanical suppositions (i.e., various arbitrary definitions of chemical environment via groups). The influence of the chemical environment is captured through the atomic partial charges calculated from HF-SCF wave function.

The performance of the REBECEP method was tested¹⁵ with the experimental enthalpies of formation of the 117 closed-shell molecules¹⁸ selected from the G3/99 molecular thermochemistry database.¹⁹ First, we selected 66 smaller molecules from the G2/97^{20,21} molecular thermochemistry database to establish a relationship between the experimental enthalpies of formation and the REBECEP equation. For the test we selected 51 larger molecules of the G3/99 molecular test set¹⁹ (none of these molecules were used to obtain the energy parameters of the REBECEP equation).¹⁵ The root-mean-square (rms) deviation from the experimental enthalpies of formation for the 51 test molecules is 1.15, 3.96, and 2.92 kcal/mol for Gaussian-3, B3LYP/6-311+G(3df,2p), and REBECEP (NPA) enthalpies of formation, respectively (the corresponding a.a. deviations are 0.94, 7.09, and 2.27 kcal/mol, respectively).¹⁵ It was found that the performance of the REBECEP method is better with natural population analysis (NPA) charges²² than with Mulliken charges.¹⁵ The REBECEP method performs considerably better for these 51 test molecules with a moderate 6-31G(d) basis set than the B3LYP method with large 6-311+G(3df,2p) basis set. The performance of the fully optimized REEF-HF method¹⁷ (vide infra) is even better than that of the earlier REBECEP method.¹⁵ Using HF-SCF/6-311+G(2d,p) results combined with the NPA charges the a.a. deviation from the experimental results for the selected 115 molecules was 1.48 kcal/mol. Leaving out the five most problematic molecules (e.g., azulene, butanedinitrile, the carbonic-difluoride, etc.), the a.a. deviation decreases to 1.22 kcal/mol.¹⁷

In this paper, we add molecules containing Si, S, and Cl atoms to our molecular database,¹⁸ derive the necessary parameters for these atoms, and test the performance of the extended REEF-HF equation¹⁷ using the NPA and Mulliken atomic partial charges combined with two basis sets. We also test the performance of Dewar's method and suggest new, reoptimized parameters.

REEF-HF Equation. To obtain calculated nonrelativistic enthalpy of formation for any molecule, M, the following equation can be used

$$\Delta H_f^\circ(\text{M}) = E_T(\text{M}) + E_{ZP}(\text{M}) + \Delta E_{\text{therm}}(\text{M}) + \sum_{A \in \text{M}} [\Delta H_f^\circ(\text{A}, \text{exp}) - E_T(\text{A}) - \Delta E_{\text{therm}}(\text{A}, \text{exp})] \quad (1)$$

where $E_T(\text{M})$ is the calculated total electronic energy, $E_{ZP}(\text{M})$ is the calculated zero-point vibration energy (ZPE), and $\Delta E_{\text{therm}}(\text{M})$ is the calculated difference between the enthalpy of the molecule at $T = 298.15$ and 0 K (calculated from the molecular heat capacity). The summation is over all atom (A) of the molecule. $\Delta H_f^\circ(\text{A}, \text{exp})$'s are the experimental standard enthalpies of formation the constituent atoms of molecule M, $E_T(\text{A})$'s are the calculated total energy of these atoms, and $\Delta E_{\text{therm}}(\text{A}, \text{exp})$'s are the experimental differences between the

enthalpies at $T = 298.15$ and 0 K (calculated from the elemental heat capacities). The calculated values of $E_T(\text{M})$, $E_{ZP}(\text{M})$, $E_{\text{therm}}(\text{M})$, and $E_T(\text{A})$ are method dependent. To obtain reliable enthalpy of formation, good quality and very expensive methods must be used for the calculation of these energies and the equilibrium geometry. The HF-SCF method fails to provide good results for eq 1 for known reasons: the $E_T(\text{M})$ and $E_T(\text{A})$'s are in serious error because HF-SCF method neglects the relativistic effects, the electron correlation (Coulomb hole error), and the basis set imperfection also yields an additive error. The $E_{ZP}(\text{M})$ and $\Delta E_{\text{therm}}(\text{M})$ values are reasonably approximated by scaled HF-SCF/6-31G(d) results²³ (even the very expensive G3 method uses HF-SCF/6-31G(d) energies scaled by 0.8929).

The REEF-HF method calculates the enthalpy of formation in the following way

$$\Delta H_f^\circ(\text{M}) = E_{T_{ZT}}(\text{M}) + \sum_{A \in \text{M}} [\Delta H_f^\circ(\text{A}, \text{exp}) - E_T(\text{A}, \text{G3})] \quad (2)$$

where $E_{T_{ZT}}(\text{M})$ is the REEF-HF energy which includes the total energy (with basis set and correlation energy corrections) plus ZPE and the difference of the thermal and relativistic corrections (vide infra). $E_T(\text{A}, \text{G3})$ is the G3 atomic total energy.²⁴ The G3 atomic energies were chosen because they are sufficiently good to obtain reliable thermochemistry. By rearranging eq 2, one can obtain an energy, $E_{T_{ZT}}(\text{M}, \text{exp})$, that reproduces exactly the experimental enthalpy of formation, $\Delta H_f^\circ(\text{M}, \text{exp})$

$$E_{T_{ZT}}(\text{M}, \text{exp}) = \Delta H_f^\circ(\text{M}, \text{exp}) - \sum_{A \in \text{M}} [\Delta H_f^\circ(\text{A}, \text{exp}) - E_T(\text{A}, \text{G3})] \quad (3)$$

The calculated $E_{T_{ZT}}(\text{M})$ energy can be obtained as a sum of an $E_T(\text{HF-SCF/basis set})$ total energy and the REEF-HF energy correction

$$E_{T_{ZT}}(\text{M}) = E_T(\text{HF-SCF/basis set}) + E_{\text{corr}}(\text{REEF-HF}) \quad (4)$$

The energy correction is the sum of the atomic corrections

$$E_{\text{corr}}(\text{REEF-HF}) \equiv \sum_{A \in \text{M}} E_{\text{corr}}(Z_A, N_A) \quad (5)$$

where Z_A is the nuclear charge of atom A and N_A is the "electron content" on atom A, noninteger, calculated as (Z_A - partial charge). The $E_{\text{corr}}(Z_A, N_A)$ atomic energy terms of eq 5 are interpolated

$$E_{\text{corr}}(Z_A, N_A) = (N_A - N1)E_{\text{par}}(Z_A, N2) + (N2 - N_A)E_{\text{par}}(Z_A, N1) \quad (6)$$

where N1 and N2 are integer numbers of electrons, with $N1 \leq N_A \leq N2 = N1 + 1$. $E_{\text{par}}(Z_A, N2)$ and $E_{\text{par}}(Z_A, N1)$ in eq 6 are the so-called REEF-HF atomic energy parameters that transform the partial charge into energy correction. For hydrogen atoms we suggest using a single parameter, $E_{\text{corr}}(1, N_A) = N_A E_{\text{par}}(1, 2)/2$.

The REEF-HF atomic parameters can be obtained from the fitting procedure that finds the minimum of

$$Y = \sum_{M=1}^L [E_{T_{ZT}}(\text{M}, \text{exp}) - E_{T_{ZT}}(\text{M})]^2 \quad (7)$$

for a selected set of L molecules. The details of the fitting procedure are described elsewhere.¹² The actual values of $E_{\text{par}}(Z_A, N1)$ parameters depend on the basis set used in the

HF-SCF calculation (cf. eq 4) and the atomic charge calculation method (cf. eq 6).

Because we use experimental enthalpies of formation in eq 7, even the relativistic effects have an influence on the REEF-HF energy and parameters. Relativistic effects on energetic properties of molecules containing atoms from the first and second rows can be as large as a few kcal/mol. For relativistic effects on molecules from the G3/99 test set, see ref 25. The G3 method²⁴ includes atomic spin-orbit correction that fits well into eq 5. In first-order direct stationary perturbative approach²⁵ to the relativistic energy, the result is dependent on the HF-SCF electron density, and this density at the nucleus is a good predictor of the scalar relativistic energy for the molecules in the G3/99 test set.¹⁹ "The magnitude of the relativistic energy is greatest for the positive ion, then the neutral molecule, and smallest for the negative ion. The electron density at the nuclei is usually greatest in the positive ion, then neutral, and smallest in the anion. But the situation is reversed if the HOMO has a significant heavy-atom s orbital character."²⁵ This supports the use of eq 6 as the REEF-HF parameters are charge dependent (thus, the dominant correlation effect is partly compensated by the opposite relativistic effect); however, we do not treat explicitly the s orbital populations.

Currently we use the HF-SCF/6-31G(d) and 6-311+G(2d,p)//B3LYP/6-31G(d) total electronic energies and the corresponding NPA and Mulliken atomic charges to predict enthalpies of formation (the relativistic corrections are treated implicitly).

Results and Discussion

In Table 1 we show the fitted atomic parameters for eq 6 obtained from the HF-SCF/6-31G(d) or HF-SCF/6-311+G(2d,p) energies and Mulliken or NPA charges of the selected molecules that contain H, C, N, O, F, Si, S, and Cl atoms. The list of 161 closed-shell test molecules can be found in Tables 2 and 3. Analysis of the values of the parameters shows that due to the basis set error, the energy correction parameters are more negative on average for the smaller basis set with any partial charge. The values of the parameters decrease (they became more negative) with the increase of the number of electrons around the atom in most of the cases, because the correlation energy is roughly proportional to the number of electrons. However, some exceptions occur. The values of several energy parameters that can be related to extremely large partial charges are determined from single equations (because only a single molecule was available for that charge in our current database). This might lead to somewhat arbitrary values for those parameters because all the accumulated errors are corrected exactly by these parameters (as no other constraint exists in the database). For example, the values of $E_{\text{corr}}(17,15)$, the energy correction parameter for hypothetical Cl^{2+} , in Table 1 are too negative. The value of this parameter is determined by the ClF_3 molecule (the partial charge on Cl is larger than +1) in the database. For similar reasons, $E_{\text{corr}}(16,13)$ for S, $E_{\text{corr}}(14,11)$ for Si, and $E_{\text{corr}}(8,10)$ for O have values somewhat out of order in Table 1.

Table 2 lists the 44 new molecules in our database together with their HF-SCF energies and the necessary energy corrections (E_{corr}) to obtain $E_{\text{TZT}}(\text{M}, \text{exp})$ for eq 7 ($E_{\text{TZT}}(\text{M}, \text{exp}) = E_{\text{T}}(\text{HF-SCF}/\text{basis set}) + E_{\text{corr}}$). The performance of the REEF-HF method is shown in Tables 3 and 4. The results are in agreement with our earlier observations; the Mulliken partial charges provide larger deviations from the experimental results than the NPA charges (cf. Table 4). The statistics in Table 4 also show that the 6-311+G(2d,p) basis set yields considerably

TABLE 1: Fitted Atomic Correction Parameters, $E_{\text{corr}}(\text{N1}, \text{Z}_\text{A})$ (Hartree) for Eq 6 To Obtain $E_{\text{corr}}(\text{REEF-HF})$ from HF-SCF/6-31G(d) or 6-311+G(2d,p) Energies and NPA or Mulliken Charges

atomic number	electrons	basis: 6-31G(d)		basis: 6-311+G(2d,p)	
		NPA	Mulliken	NPA	Mulliken
1	2	-0.01363	-0.01799	-0.01830	-0.01879
6	4	-0.12929	-0.12665	-0.16275	-0.16960
6	5	-0.17204	-0.17266	-0.17765	-0.17912
6	6	-0.21503	-0.21406	-0.20604	-0.20613
6	7	-0.24527	-0.24409	-0.22549	-0.22997
6	8	-0.35058	n/a	-0.27903	-0.26292
7	6	-0.27780	-0.27990	-0.27239	-0.28312
7	7	-0.27586	-0.27692	-0.26126	-0.26121
7	8	-0.30705	-0.30514	-0.27739	-0.27694
7	9	-0.37188	n/a	-0.30575	n/a
8	7	n/a	n/a	n/a	-0.39032
8	8	-0.33583	-0.33679	-0.31402	-0.31319
8	9	-0.36115	-0.35520	-0.31932	-0.31237
8	10	-0.35751	n/a	-0.34323	n/a
9	9	-0.36206	-0.36323	-0.33615	-0.34339
9	10	-0.41035	-0.40794	-0.34946	-0.33677
14	11	-0.33671	n/a	-0.36754	-0.40545
14	12	-0.32887	-0.37058	-0.38071	-0.38145
14	13	-0.44558	-0.42449	-0.40720	-0.40145
14	14	-0.45754	-0.45757	-0.42012	-0.42119
16	13	-0.61881	n/a	-0.59482	n/a
16	14	-0.55429	-0.76592	-0.53884	-0.89346
16	15	-0.52742	-0.54667	-0.51098	-0.52745
16	16	-0.54101	-0.54008	-0.51387	-0.51244
16	17	-0.62450	-0.62259	-0.53881	-0.55046
17	15	-0.64367	-0.67072	-0.64630	n/a
17	16	-0.57803	-0.58399	-0.55890	-0.58230
17	17	-0.57711	-0.57622	-0.54856	-0.54225
17	18	-0.61331	-0.59488	-0.55477	-0.58304

better results than the 6-31G(d) basis set for NPA charges (a.a. deviation, 1.55 vs 1.97 kcal/mol, respectively), while for Mulliken charges no such improvement can be observed (Mulliken charges are known to have serious problems, divergence with large basis sets, while NPA charges remain consistent, convergent).

Largest Deviations from Experiment. In agreement with our earlier results, the azulene (117) shows the largest negative deviation from experiment (from -8.5 to -10.6 kcal/mol depending of the method). This behavior is general for the REBECEP type of method. Our analysis of the details showed that it is not possible to obtain good quality results with the same parameter set for azulene (117) and naphthalene (81). We note that for azulene and naphthalene show the same (-2.8 kcal/mol)²⁵ relativistic effect; thus, inclusion of relativistic effects do not help in this case. Naphthalene and azulene are structural isomers, so it is expected that Dewar's energy parameters combined with the HF/6-31G(d) energies will provide poor results. The original Dewar's energy parameters yielded -23 and -36 kcal/mol deviation from experiment for naphthalene and azulene, respectively. We shall return to this later.

Three other problematic molecules in Table 3 are carbonic difluoride (15), cyclopropene (22), and bicyclo[1.1.0]butane (29). We note that the G3 and G3SX²⁶ enthalpies of formation show considerable -3.5 kcal/mol deviation from experiment for carbonic difluoride (15) molecule. The experimental error is also fairly large for this molecule (cf. 1.4 kcal/mol error in Table 3 and in ref 27). The uncertainties about the enthalpy of formation suggest removing this molecule from the database; however, we kept it for comparison reasons with earlier results. Cyclopropene (22) and bicyclo[1.1.0]butane (29) have a rather strained ring structure, and our method fails to provide good quality results for such molecules (the error range is from -3.6

TABLE 2: Species, Number of Nuclei, N_a , and Electrons, N_e , the HF-SCF, $E(\text{HF-SCF})$, and Correction Energies, E_{corr} , (Hartree) To Reproduce Experimental Enthalpies of Formation for 44 Selected Molecules Calculated with 6-31G(d) and 6-311+G(2d,p) Basis Sets^a

no.	N_d^b	species	N_a	N_e	6-31G(d)		6-311+G(2d,p)	
					$E(\text{HF-SCF})$	E_{corr}	$E(\text{HF-SCF})$	E_{corr}
1	123	silane (SiH ₄)	5	18	-291.22496	-0.48723	-291.25710	-0.45509
2	125	hydrogen sulfide (H ₂ S)	3	18	-398.66667	-0.57344	-398.70473	-0.53538
3	126	hydrogen chloride (HCl)	2	18	-460.05961	-0.59540	-460.09668	-0.55833
4	127	silicon monoxide (SiO)	2	22	-363.77711	-0.77924	-363.83018	-0.72617
5	128	CS	2	22	-435.30342	-0.75431	-435.34232	-0.71541
6	129	chlorine monofluoride (FCl)	2	26	-558.81787	-0.95333	-558.88047	-0.89073
7	130	Si ₂ H ₆	8	34	-581.30482	-0.94871	-581.36283	-0.89070
8	131	methyl chloride (CH ₃ Cl)	5	26	-499.09267	-0.82352	-499.13480	-0.78139
9	132	methanethiol (H ₃ CSH)	6	26	-437.69957	-0.80666	-437.74448	-0.76175
10	133	hypochlorous acid (HOCl)	3	26	-534.83974	-0.93342	-534.90033	-0.87283
11	134	sulfur dioxide (SO ₂)	3	32	-547.16196	-1.26693	-547.25718	-1.17171
12	135	carbon tetrachloride (CCl ₄)	5	74	-1875.74324	-2.53786	-1875.86877	-2.41233
13	136	carbon oxide sulfide (COS)	3	30	-510.25808	-1.08401	-510.32157	-1.02052
14	137	carbon bisulfide (CS ₂)	3	38	-832.88332	-1.30215	-832.94571	-1.23976
15	138	silicon tetrafluoride (SiF ₄)	5	50	-686.94813	-1.91373	-687.10983	-1.75203
16	139	silicon tetrachloride (SiCl ₄)	5	82	-2127.04627	-2.74433	-2127.19186	-2.59874
17	140	nitrogen chloride oxide (ClNO)	3	32	-588.67486	-1.21056	-588.74559	-1.13983
18	142	chlorine trifluoride (ClF ₃)	4	44	-757.47717	-1.76084	-757.61595	-1.62206
19	143	ethene, tetrachloro- (C ₂ Cl ₄)	6	80	-1913.60234	-2.74566	-1913.73346	-2.61454
20	144	CH ₂ Cl ₂	5	42	-957.98436	-1.39018	-958.05420	-1.32034
21	145	CHCl ₃	5	58	-1416.86852	-1.96228	-1416.96617	-1.86463
22	146	methylsilane (CH ₃ SiH ₃)	8	26	-330.27208	-0.72950	-330.31499	-0.68659
23	147	thiirane (C ₂ H ₄ S)	7	32	-475.54584	-1.01899	-475.59463	-0.97020
24	148	dimethyl sulfoxide ((CH ₃) ₂ SO)	10	42	-551.53451	-1.41080	-551.62518	-1.32013
25	149	ethanethiol (C ₂ H ₅ SH)	9	34	-476.73478	-1.04553	-476.79025	-0.99006
26	150	dimethyl sulfide (CH ₃ SCH ₃)	9	34	-476.73450	-1.04231	-476.78642	-0.99039
27	151	ethyl chloride (C ₂ H ₅ Cl)	8	34	-538.13084	-1.06199	-538.18365	-1.00918
28	152	vinyl chloride (CH ₂ =CHCl)	6	32	-536.93292	-1.03646	-536.98541	-0.98397
29	153	CH ₃ COCl (acetyl chloride)	7	40	-611.82899	-1.37200	-611.90687	-1.29412
30	154	CH ₃ CH ₂ CH ₂ Cl (propyl chloride)	11	42	-577.16575	-1.29973	-577.22887	-1.23661
31	155	C ₄ H ₄ S (thiophene)	9	44	-551.28877	-1.46068	-551.35767	-1.39178
32	156	chlorobenzene (C ₆ H ₅ Cl)	12	58	-689.60199	-1.92279	-689.68889	-1.83589
33	157	2,5-dihydrothiophene (C ₄ H ₆ S)	11	46	-552.45438	-1.47231	-552.52399	-1.40270
34	158	tetrahydrothiophene (C ₄ H ₈ S)	13	48	-553.64302	-1.49644	-553.71200	-1.42746
35	159	<i>tert</i> -butyl chloride ((CH ₃) ₃ C-Cl)	14	50	-616.20483	-1.54494	-616.27883	-1.47094
36	160	<i>n</i> -butyl chloride (CH ₃ -CH ₂ -CH ₂ -CH ₂ -Cl)	14	50	-616.20026	-1.53915	-616.27392	-1.46549
37	161	<i>tert</i> -butanethiol ((CH ₃) ₃ C-SH)	15	50	-554.80492	-1.52978	-554.88128	-1.45342
38	162	diethyl disulfide (CH ₃ -CH ₂ -S-S-CH ₂ -CH ₃)	16	66	-952.31730	-2.06992	-952.42141	-1.96581
39	163	tetramethylsilane (Si(CH ₃) ₄)	17	50	-447.41281	-1.46186	-447.48979	-1.38488
40	164	2-methyl thiophene (C ₅ H ₆ S)	12	52	-590.32741	-1.69916	-590.40649	-1.62008
41	165	tetrahydrothiopyran (C ₅ H ₁₀ S)	16	56	-592.68029	-1.73549	-592.75945	-1.65633
42	166	SCl ₂	3	50	-1316.43343	-1.71197	-1316.53017	-1.61523
43	169	Cl ₂ O ₂ S	5	66	-1466.06944	-2.45446	-1466.23940	-2.28450
44	173	CF ₃ Cl	5	50	-795.66016	-1.88601	-795.79628	-1.74989

^a $E_{\text{TZT}}(\text{M, exp})$ of eq 3 that reproduces exactly the experimental enthalpy of formation, $\Delta H_f^\circ(\text{M, exp})$, can be obtained as $E(\text{HF-SCF}) + E_{\text{corr}}$. The experimental enthalpies of formation values can be found in Table 3. ^b Database number.

to -6.9 kcal/mol, cf. Table 3). We note that in G3 theory three hydrocarbons with strained ring systems have deviations slightly greater than 2.0 kcal/mol: cyclopropene (-2.2 kcal/mol), cyclobutene (-2.1 kcal/mol), and bicyclo[1.1.0]butane (-2.6 kcal/mol).²⁴

The calculated enthalpy of formation of the butanedinitrile (92) shows the largest positive deviation from the experiment if NPA charges are used (independent of basis set). Earlier study¹⁷ showed that this problem can be solved by using Stockholder charges. Leaving out the five most problematic molecules yielded 1.38 kcal/mol a.a. deviation for the remaining 154 molecules if the calculation is performed with NPA charges and 6-311+G(2d,p) basis set. It can be observed in Table 3 that the calculated enthalpies of formation show negative deviation for molecules that contain the *tert*-butyl group, e.g., *tert*-butyl alcohol (83), *tert*-butylamine (105), and *tert*-butyl methyl ether (111). Improving the basis set quality improves

the results in general (large deviations become less frequent); however, some of the most problematic molecules remain the same.

We note that the calculated enthalpy of formation of the 1,4-dioxane (100) shows a large, 5.66 kcal/mol (cf. Table 3), deviation from experiment if the smaller 6-31G(d) basis set is used with NPA charges. However, this deviation is basis set dependent, and the larger 6-311+G(2d,p) basis set provides reasonable agreement with experiment (cf. Table 3, NPA charges). The performance of the method using the 6-311+G(2d,p) basis set with NPA charges is reasonable. Only two more molecules show calculated values with deviation larger than 5 kcal/mol: nitrogen chloride oxide (140) and vinyl chloride (152). Interestingly, the enthalpy of formation of vinyl chloride is poorly reproduced by the G3 method (+3.6 kcal/mol deviation)²⁴ and an isodesmic scheme also gives an unusually large deviation for the latter species and may indicate a problem with the experimental value. Our results in Table 3

TABLE 3: Experimental Enthalpies of Formation and Deviations of REEF-HF Enthalpies of Formation from Experimental Result Calculated with 6-311G(d) and 6-311+G(2d,p) Basis Sets Using NPA and Mulliken Population Analysis^a

no.	species	exp value	ΔH_f° error	deviation (exp – calcd)			
				6-31G(d)		6-311+G(2d,p)	
				NPA	Mulliken	NPA	Mulliken
1	methane (CH ₄)	-17.83	0.07	-1.57	-0.79	-0.01	-0.09
2	ammonia (NH ₃)	-11.00	0.08	0.00	-2.85	0.00	-1.52
3	water (H ₂ O)	-57.80	0.01	-4.16	-6.87	-1.38	-2.96
4	hydrogen fluoride (HF)	-65.10	0.17	-3.93	-4.80	-0.27	-0.51
5	acetylene (C ₂ H ₂)	54.35	0.19	-3.17	-1.42	1.00	1.72
6	ethylene (H ₂ C=CH ₂)	12.52	0.12	-1.47	-0.14	2.42	2.36
7	ethane (H ₃ C-CH ₃)	-20.10	0.10	0.17	0.82	1.32	1.43
8	hydrogencyanide (HCN)	31.50	0.96	0.26	2.18	1.86	2.77
9	formaldehyde (H ₂ C=O)	-26.00	0.12	-1.98	1.02	0.71	2.37
10	methanol (CH ₃ -OH)	-48.00	0.14	-0.63	-1.26	1.69	0.35
11	hydrazine (H ₂ N-NH ₂)	22.75	0.12	-1.23	-0.47	0.80	-0.74
12	hydrogen peroxide (HO-OH)	-32.50	0.05	-2.26	-4.25	-0.47	-2.66
13	carbon dioxide (CO ₂)	-94.05	0.03	-1.82	-0.47	1.62	6.05
14	carbon tetrafluoride (CF ₄)	-223.00	0.30	3.18	1.80	2.64	3.68
15	carbonic difluoride (COF ₂)	-149.10	1.40	-5.64	-5.30	-3.27	-2.76
16	dinitrogen monoxide (N ₂ O)	19.60	0.10	-3.24	-1.31	-1.34	-1.54
17	nitrogen trifluoride (NF ₃)	-31.60	0.30	-0.64	2.44	0.23	5.48
18	ethene, tetrafluoro- (F ₂ C=CF ₂)	-157.40	0.70	-2.62	-3.14	0.08	-0.02
19	acetonitrile, trifluoro- (CF ₃ CN)	-118.40	0.70	0.81	-1.53	-0.19	1.49
20	propyne (C ₃ H ₄)	44.20	0.20	0.00	0.10	2.91	1.84
21	allene (C ₃ H ₄)	45.50	0.30	-1.76	-2.33	1.81	1.36
22	cyclopropene (C ₃ H ₄)	66.20	0.60	-5.69	-5.05	-3.64	-3.87
23	propylene (C ₃ H ₆)	4.80	0.20	-0.55	0.31	2.52	2.27
24	cyclopropane (C ₃ H ₆)	12.70	0.10	0.34	0.99	0.20	-1.24
25	propane (C ₃ H ₈)	-25.00	0.10	0.49	0.95	1.37	1.66
26	<i>trans</i> -1,3-butadiene (C ₄ H ₆)	26.30	0.30	-2.00	-0.58	2.64	2.17
27	dimethylacetylene (C ₄ H ₆)	34.80	0.30	2.61	2.35		
28	methylenecyclopropane (C ₄ H ₆)	47.90	0.40	0.97	0.52	1.90	1.11
29	bicyclo[1.1.0]butane (C ₄ H ₆)	51.90	0.20	-5.09	-4.83	-5.77	-6.86
30	cyclobutene (C ₄ H ₆)	37.40	0.40	-2.77	-2.49	-1.60	-2.13
31	cyclobutane (C ₄ H ₈)	6.80	0.10	1.09	1.33	0.17	-0.54
32	isobutene (C ₄ H ₈)	-4.00	0.20	-1.04	-1.34	1.48	0.75
33	<i>trans</i> -butane (C ₄ H ₁₀)	-30.00	0.20	0.46	0.82	1.11	1.70
34	isobutane (C ₄ H ₁₀)	-32.10	0.20	-1.10	-0.87	-0.46	-0.46
35	spiropentane (C ₅ H ₈)	44.30	0.20	-0.02	-0.14	-2.09	-3.98
36	benzene (C ₆ H ₆)	19.70	0.20	3.32	3.43	3.84	3.14
37	difluoromethane (CH ₂ F ₂)	-108.10	0.20	-0.82	0.84	0.86	2.74
38	trifluoromethane (CHF ₃)	-166.60	0.80	0.66	1.59	0.14	1.11
39	methylamine (H ₃ C-NH ₂)	-5.50	0.10	-0.98	-0.32	0.82	-0.18
40	acetonitrile (CH ₃ -CN)	18.00	0.10	2.94	3.43	3.36	4.14
41	nitromethane (CH ₃ -NO ₂)	-17.80	0.10	1.78	0.03	2.51	-1.73
42	methyl nitrite (CH ₃ -O-N=O)	-15.90	0.20	3.26	2.45	2.15	-4.36
43	formic acid (HCOOH)	-90.50	0.10	-0.91	-0.09	-0.50	0.10
44	methyl formate (HCOOCH ₃)	-85.00	0.20	0.65	3.23	0.34	0.87
45	acetamide (CH ₃ CONH ₂)	-57.00	0.20	-0.18	-0.54	-0.55	-1.19
46	aziridine (C ₂ H ₄ NH)	30.20	0.20	-1.62	-0.42	-1.30	-2.26
47	cyanogen (NCCN)	73.30	0.20	-2.85	-6.01	-2.51	-2.94
48	dimethylamine ((CH ₃) ₂ NH)	-4.40	0.20	-1.45	0.05	0.04	-0.11
49	<i>trans</i> -ethylamine (CH ₃ CH ₂ NH ₂)	-11.30	0.20	0.72	0.93	1.83	1.00
50	ketene (CH ₂ CO)	-11.40	0.40	-3.49	-2.05	-0.44	1.04
51	oxirane (C ₂ H ₄ O)	-12.60	0.10	-0.53	-0.04	-0.75	-2.52
52	acetaldehyde (CH ₃ CHO)	-39.70	0.10	-0.43	1.53	1.24	2.52
53	glyoxal (HCOCOH)	-50.70	0.20	1.18	4.46	2.88	4.83
54	ethanol (CH ₃ CH ₂ OH)	-56.20	0.10	0.50	-0.83	1.72	0.92
55	dimethyl ether (CH ₃ OCH ₃)	-44.00	0.10	0.49	1.57	1.86	1.14
56	vinyl fluoride (CH ₂ =CHF)	-33.2	0.40	-1.06	0.37		
57	acrylonitrile (CH ₂ =CHCN)	43.2	0.40	-0.34	0.69	1.35	2.21
58	acetone (CH ₃ COCH ₃)	-51.9	0.20	0.55	1.45	1.38	1.97
59	acetic acid (CH ₃ COOH)	-103.4	0.40	0.81	0.14	0.18	0.32
60	acetyl fluoride (CH ₃ COF)	-105.7	0.80	-0.58	0.86	-0.18	1.39
61	2-propanol ((CH ₃) ₂ CHOH)	-65.2	0.10	-0.71	-2.51	0.21	0.10
62	methyl ethyl ether (C ₂ H ₅ OCH ₃)	-51.7	0.20	1.86	2.14	2.24	1.49
63	trimethylamine ((CH ₃) ₃ N)	-5.7	0.20	-3.39	-1.21	-2.15	-0.92
64	furan (C ₄ H ₄ O)	-8.3	0.20	-2.99	-2.81	-2.96	-3.62
65	pyrrole (C ₄ H ₅ N)	25.9	0.10	-3.18	-3.63	-2.54	-3.53
66	pyridine (C ₅ H ₅ N)	33.6	0.20	3.29	3.87	2.24	2.30
67	methyl allene (C ₄ H ₆)	38.8	0.10	-0.90	-1.44	1.89	1.12
68	isoprene (C ₅ H ₈)	18	0.30	-2.71	-2.36	1.38	0.30
69	cyclopentane (C ₅ H ₁₀)	-18.3	0.20	1.35	0.96	-0.42	-0.65
70	<i>n</i> -pentane (C ₅ H ₁₂)	-35.1	0.20	0.16	0.42	0.57	1.33

TABLE 3: (Continued)

no.	species	exp value	ΔH_f° error	deviation (exp - calcd)			
				6-31G(d)		6-311+G(2d,p)	
				NPA	Mulliken	NPA	Mulliken
71	neopentane (C ₅ H ₁₂)	-40.2	0.20	-3.94	-3.91	-3.59	-4.00
72	1,3-cyclohexadiene (C ₆ H ₈)	25.4	0.20	-0.16	-0.51	0.71	0.52
73	1,4-cyclohexadiene (C ₆ H ₈)	25	0.10	0.74	-0.02	2.14	1.85
74	cyclohexane (C ₆ H ₁₂)	-29.5	0.20	0.91	0.05	-1.19	-0.54
75	<i>n</i> -hexane (C ₆ H ₁₄)	-39.9	0.20	0.31	0.49	0.50	1.74
76	3-methyl pentane (C ₆ H ₁₄)	-41.1	0.20	-2.65	-2.42	-2.60	-2.52
77	toluene (C ₆ H ₅ CH ₃)	12	0.10	1.86	1.40	1.82	0.38
78	<i>n</i> -heptane (C ₇ H ₁₆)	-44.9	0.30	0.35	0.44	0.30	1.85
79	cyclooctatetraene (C ₈ H ₈)	70.7	0.40	-2.51	-3.00	0.11	-0.55
80	<i>n</i> -octane (C ₈ H ₁₈)	-49.9	0.30	0.17	0.18	-0.12	1.75
81	naphthalene (C ₁₀ H ₈)	35.9	0.40	3.43	2.73	2.10	-0.52
82	acetic acid methyl ester (CH ₃ COOCH ₃)	-98.40	0.40	1.31	2.20	0.15	0.03
83	<i>tert</i> -butanol ((CH ₃) ₃ COH)	-74.70	0.20	-3.85	-5.70	-2.85	-3.34
84	aniline (C ₆ H ₅ NH ₂)	20.80	0.20	0.01	-1.32	0.29	-1.72
85	phenol (C ₆ H ₅ OH)	-23.00	0.20	-0.58	-3.41	-0.35	-1.90
86	divinyl ether (C ₄ H ₆ O)	-3.30	0.20	-4.30	-4.01	-0.52	-0.93
87	tetrahydrofuran (C ₄ H ₈ O)	-44.00	0.20	3.53	2.16	0.72	-1.15
88	cyclopentanone (C ₅ H ₈ O)	-45.90	0.40	3.19	3.47	1.07	1.41
89	benzoquinone (C ₆ H ₄ O ₂)	-29.40	0.80	0.99	0.12	0.83	0.19
90	urea (CH ₄ ON ₂)	-56.29	0.29	1.77	0.61	0.44	-1.20
91	pyrimidine (C ₄ H ₄ N ₂)	46.80	0.30	3.14	5.23	1.50	2.91
92	butanedinitrile (N≡C-CH ₂ -CH ₂ -C≡N)	50.10	0.20	5.66	5.97	4.79	8.18
93	pyrazine (C ₄ H ₄ N ₂)	46.90	0.30	0.94	1.48	-1.58	-1.77
94	acetyl acetylene (CH ₃ -C(=O)-C≡CH)	15.60	0.20	-3.71	-3.30	-1.57	0.00
95	crotonaldehyde (CH ₃ -CH=CH-CHO)	-24.00	0.30	1.21	3.16	3.36	3.81
96	acetic anhydride (CH ₃ -C(=O)-O-C(=O)-CH ₃)	-136.80	0.40	0.36	1.02	-2.11	0.30
97	isobutane nitrile ((CH ₃) ₂ CH-CN)	5.60	0.30	1.06	1.71	1.03	2.28
98	methyl ethyl ketone (CH ₃ -CO-CH ₂ -CH ₃)	-57.10	0.20	1.04	1.80	1.13	1.75
99	isobutanal ((CH ₃) ₂ CH-CHO)	-51.60	0.20	-1.78	0.27	-0.89	1.07
100	1,4-dioxane (C ₄ H ₈ O ₂)	-75.50	0.20	5.78	4.24	1.97	1.20
101	tetrahydropyrrole (C ₄ H ₈ NH)	-0.80	0.20	0.98	1.15	-0.21	-0.83
102	nitro- <i>s</i> -butane (CH ₃ -CH ₂ -CH(CH ₃)-NO ₂)	-39.10	0.40	2.22	0.17	0.40	-0.99
103	diethyl ether (CH ₃ -CH ₂ -O-CH ₂ -CH ₃)	-60.30	0.20	2.22	1.62	1.58	0.94
104	dimethyl acetal (CH ₃ -CH(OCH ₃) ₂)	-93.10	0.20	-0.22	0.00	-1.75	0.04
105	<i>tert</i> -butylamine ((CH ₃) ₃ C-NH ₂)	-28.90	0.20	-4.33	-4.53	-3.55	-4.65
106	<i>N</i> -methyl pyrrole (cyc-C ₄ H ₄ N-CH ₃)	24.60	0.10	-5.06	-4.42	-4.53	-1.94
107	tetrahydropyran (C ₅ H ₁₀ O)	-53.40	0.20	3.31	2.12	-0.18	0.10
108	diethyl ketone (CH ₃ -CH ₂ -CO-CH ₂ -CH ₃)	-61.60	0.20	2.08	2.70	1.48	2.21
109	isopropyl acetate (CH ₃ -C(=O)-O-CH(CH ₃) ₂)	-115.10	0.20	0.56	0.47	-2.00	1.35
110	piperidine (cyc-C ₅ H ₁₀ NH)	-11.30	0.10	0.41	0.14	-1.71	-1.50
111	<i>tert</i> -butyl methyl ether ((CH ₃) ₃ C-O-CH ₃)	-67.80	0.30	-4.31	-4.51	-4.33	-2.98
112	1,3-difluorobenzene (C ₆ H ₄ F ₂)	-73.90	0.20	-0.18	0.18	-0.17	0.95
113	1,4-difluorobenzene (C ₆ H ₄ F ₂)	-73.30	0.20	0.39	0.51	0.07	0.67
114	fluorobenzene (C ₆ H ₅ F)	-27.70	0.30	1.58	1.83	1.65	1.80
115	diisopropyl ether ((CH ₃) ₂ CH-O-CH(CH ₃) ₂)	-76.30	0.40	-1.85	-3.22	-3.12	0.57
116	ethane, hexafluoro- (C ₂ F ₆)	-321.30	0.80	2.09	1.28	-0.83	-0.26
117	azulene (C ₁₀ H ₈)	69.10	0.80	-8.54	-9.76	-8.60	-10.61
123	silane (SiH ₄)	8.20	0.50	-4.26	-4.28	-0.93	-0.92
125	hydrogen sulfide (H ₂ S)	-4.90	0.20	1.09	0.45	0.28	1.12
126	hydrogen chloride (HCl)	-22.10	0.05	-2.03	-4.91	-0.86	-0.26
127	silicon monoxide (SiO)	-24.60	n/a	-0.11	3.34	-0.29	-5.83
128	CS	66.90	n/a	1.68	0.39	3.48	3.70
129	chlorine monofluoride (FCl)	-13.20	n/a	2.24	3.07	1.97	0.33
130	Si ₂ H ₆	19.10	0.40	0.79	1.22	0.18	0.54
131	methyl chloride (CH ₃ Cl)	-19.60	0.50	3.92	3.78	3.10	1.61
132	methanethiol (H ₃ CSH)	-5.50	0.20	-1.45	0.63	1.00	1.94
133	hypochlorous acid (HOCl)	-17.80	0.50	0.56	-1.18	0.18	-6.72
134	sulfur dioxide (SO ₂)	-71.00	0.05	-0.17	0.71	-0.46	0.23
135	carbon tetrachloride (CCl ₄)	-22.90	0.50	-4.70	-3.59	-3.45	-4.03
136	carbon oxide sulfide (COS)	-33.10	0.30	0.56	-0.16	0.09	3.16
137	carbon bisulfide (CS ₂)	28.00	0.20	-0.17	-4.23	-1.24	-0.28
138	silicon tetrafluoride (SiF ₄)	-386.00	0.20	0.00	0.26	0.00	0.00
139	silicon tetrachloride (SiCl ₄)	-158.40	0.30	0.43	2.68	0.36	6.17
140	nitrogen chloride oxide (ClNO)	12.40	0.10	-4.92	-9.44	-5.59	0.00
142	chlorine trifluoride (ClF ₃)	-38.00	0.70	0.00	0.00	0.00	-13.53
143	ethene, tetrachloro- (C ₂ Cl ₄)	-3.00	0.70	0.59	1.15	-0.09	13.74
144	CH ₂ Cl ₂	-22.80	0.30	3.33	2.75	2.48	-1.37
145	CHCl ₃	-24.70	0.30	-1.24	-0.53	-0.49	-10.50
146	methylsilane (CH ₃ SiH ₃)	-7.00	1.00	2.97	-2.50	0.64	-0.50
147	thiirane (C ₂ H ₄ S)	19.60	0.30	2.19	2.92	0.75	0.49
148	dimethyl sulfoxide ((CH ₃) ₂ SO)	-36.20	0.20	0.38	-0.16	1.40	-0.22

TABLE 3: (Continued)

no.	species	exp value	ΔH_f° error	deviation (exp - calcd)			
				6-31G(d)		6-311+G(2d,p)	
				NPA	Mulliken	NPA	Mulliken
149	ethanethiol (C ₂ H ₅ SH)	-11.10	0.10	-0.89	0.86	0.72	2.36
150	dimethyl sulfide (CH ₃ SCH ₃)	-8.90	0.10	3.58	5.11	2.96	3.78
151	ethyl chloride (C ₂ H ₅ Cl)	-26.80	0.30	4.60	4.02	3.09	1.95
152	vinyl chloride (CH ₂ =CHCl) ^b	8.90	0.50	4.87	5.01	5.92	3.26
153	CH ₃ COCl (acetyl chloride)	-58.00	0.20	-0.69	-0.44	-1.28	-1.11
154	CH ₃ CH ₂ CH ₂ Cl (propyl chloride)	-31.50	0.30	5.23	4.70	3.30	1.69
155	C ₄ H ₄ S (thiophene)	27.50	0.20	-1.29	0.54	-1.29	-5.66
156	chlorobenzene (C ₆ H ₅ Cl)	12.40	0.30	4.45	3.59	3.00	0.46
157	2,5-dihydrothiophene (C ₄ H ₆ S)	20.80	0.30	4.02	4.73	2.81	3.72
158	tetrahydrothiophene (C ₄ H ₈ S)	-8.20	0.30	3.47	3.97	0.46	1.82
159	<i>tert</i> -butyl chloride ((CH ₃) ₃ C-Cl)	-43.50	0.50	1.78	0.54	-0.52	2.32
160	<i>n</i> -butyl chloride (CH ₃ -CH ₂ -CH ₂ -CH ₂ -Cl)	-36.96	0.27	4.65	4.05	2.50	0.87
161	2-propanethiol, 2-methyl- ((CH ₃) ₂ CH-SH)	-26.20	0.21	-5.10	-3.62	-3.81	-1.59
162	diethyl disulfide (CH ₃ -CH ₂ -S-S-CH ₂ -CH ₃)	-17.90	0.20	-2.13	-1.95	-1.12	-0.20
163	tetramethylsilane (Si(CH ₃) ₄)	-55.70	0.80	-0.61	-1.96	-0.13	0.00
164	2-methyl thiophene (C ₅ H ₆ S)	20.16	0.22	-1.03	0.41	-1.64	-3.89
165	tetrahydrothiopyran (C ₅ H ₁₀ S)	-15.18	0.25	3.49	3.63	-0.01	-0.45
166	SCl ₂	-4.20	0.80	-5.95	-7.73	-2.01	-9.47
169	Cl ₂ O ₂ S	-84.80	0.50	0.00	-0.33	0.00	-0.10
173	CF ₃ Cl	-169.50	0.80	-1.00	-1.55	-1.53	-0.98

^a Experimental values are taken from refs 23 and 25. ^b Please provide footnote.

TABLE 4: Statistics for the Deviation (kcal/mol) between the Experimental Enthalpies of Formation and the Calculated REEF-HF Enthalpies of Formation Depending on the Basis Sets [6-31G(d) or 6-311+G(2d,p)] and Partial Charge Calculation Methods [NPA or Mulliken]

no. of molecules	deviation (exp - calcd)			
	6-31G(d)		6-311+G(2d,p)	
	NPA	Mulliken	NPA	Mulliken
no. of molecules	161	161	159	159
root mean square deviation	2.58	2.86	2.08	3.20
average deviation	-0.13	-0.08	0.17	0.03
average absolute deviation	1.97	2.14	1.55	2.10

support this observation and agree better with G3 results than with the experiment. Only one not yet mentioned molecule shows a larger deviation than 4 kcal/mol (*N*-methyl-pyrrole (106) in Table 3). For comparison, we cite that the G3 method has a negative deviation of -3.8 kcal/mol with experiment for SO₂, and it has positive deviations of 2-3 kcal/mol with experiment for CS₂ (3.3 kcal/mol), COS (2.8 kcal/mol), and C₂Cl₄ (3.4 kcal/mol).²⁴ Inspection of the corresponding values for REEF-HF method in Table 3 yields deviations of -0.5, -1.2, 0.1, and -0.1 kcal/mol with experiment, respectively. The performance of the proposed method is rather good for the second-row elements. However, mainly for molecules containing the Cl the REEF-HF enthalpy of formation values derived from HF-SCF/6-31G(d) results are considerably less satisfactory than those derived from HF-SCF/6-311+G(2d,p) results. We also recall that the atomic enthalpy of formation of Si has large error (1.9 kcal/mol); consequently all results for molecules containing Si contain this uncertainty.

Comparison of the Performance of Single-Atom Equivalent Formula with REEF-HF. According to Dewar's proposition,⁵ the enthalpy of formation of a compound M, $\Delta H_f^\circ(M)$, can be calculated using a single-atom equivalent formula

$$\Delta H_f^\circ(M) = E_T(M, \text{HF-SCF/basis set}) - \sum_{A \in M} E_{\text{par}}(Z_A) \quad (8)$$

where $E_{\text{par}}(Z_A)$ is the single-atom equivalent that depends only on the nature of the atom. This equation relates the HF-SCF

energy error simply to the number of the electrons in the neutral atoms and ignores the finer details (the influence of the molecular environment). As the HF-SCF energy error, the correlation energy is roughly proportional to the number of electrons this approach and the REEF-HF method has common bases. The appropriate $E_{\text{par}}(Z_A)$ was found by least-squares fit to the experimental enthalpies of formation of an appropriately selected set of molecules.⁵ We do not list here the 45 molecules selected by Dewar et al.; the interested reader should refer to ref 5 (most of the molecules are listed in Table 3). Equation 8 was suggested to estimate the effective errors in ab initio and semiempirical energies. As mentioned for azulene, eq 8 cannot distinguish between structural isomers better than the HF-SCF method. Similarly, for reaction enthalpies this formula cannot give a better result than the HF-SCF method because equal correction occurs on both sides of a chemical equation. Equation 8 cannot treat at all the differences coming from thermal, zero-point, and relativistic energy contributions. "The corresponding errors will then indicate the effective errors in the energies of individual molecules, representing their contributions to the errors in the enthalpies of reaction calculated for reactions involving them."⁵ As mentioned, eq 8 was found to yield useful results for B3LYP/6-311G(d,p) energies, yielding 0.8 kcal/mol a.a. deviation from experiment for 23 simple, less problematic hydrocarbons.¹⁰ Our reoptimized HF-SCF/6-31G(d), NPA REEF-HF parametrization yields 1.17 kcal/mol a.a. deviation for 33 hydrocarbons (without azulene). Only spiroentane (35) and neopentane (71) yield larger deviations than 3 kcal/mol. If we leave these and other 'difficult' compounds out and use a reduced set of 27 simple hydrocarbons, we obtained 0.77 or 0.64 kcal/mol a.a. deviations with REEF-HF method using HF-SCF/6-31G(d) energies and NPA or Mulliken charges, respectively. This latter data set is more similar to the data set used in ref 10, but it contains several more challenging molecules such as naphthalene or spiroentane.

It can be observed that single-atom equivalent corrected HF-SCF/6-31G(d) energies (cf. eq 8) show a large and systematic error (about 23 kcal/mol) for cations⁵ if $E_{\text{par}}(Z_A)$'s obtained from neutral molecules are applied for cations. This was attributed by Dewar et al. to the poor performance of the

HF-SCF/6-31G(d) method.⁵ However, according to our analysis, the large systematic error for cations is not proof of the poor performance of the HF-SCF method; it proves the necessity of considering charges. Thus, this observation supports directly the proposed REEF-HF method. According to observations and theory, the correlation energy depends on the number of electrons. Consequently, applying the same single-atom equivalent HF-SCF/6-31G(d) correction for cations and neutral molecules must lead to large errors for the cations. To check this prediction, we reproduced the enthalpies of formation calculated by Dewar et al.⁵ for several cations (CH_3^+ , allyl⁺, NO^+ , H_3O^+ , and NH_4^+) and observed that the enthalpies are too negative due to a systematic overcorrelation effect (the correlation energy is negative). We performed the same calculations with the REEF-HF 6-31G(d) NPA energy parameters in Table 1. It can be observed that the error of our method is considerably smaller, because the correlation energy of the missing electron was not included. For example, the 30 kcal/mol deviation for CH_3^+ was decreased to 8 kcal/mol, the 12 kcal/mol deviation for H_3O^+ was decreased to 1 kcal/mol, and the 22 kcal/mol deviation for NH_4^+ was decreased to 13 kcal/mol.

Because our molecular database is considerably larger (101 molecules can be found in our database that contain H, C, N, and O atoms) than that used by Dewar, it contains controlled experimental data and our results refer to a slightly different geometry (consequently different HF-SCF/6-31G(d) energy), we reoptimized the original parameters proposed by Dewar in order to obtain the best performance. The original parametrization of Dewar et al.⁵ yielded 8.7 kcal/mol rms deviation from experiment for 101 molecules. After the reoptimization of the single-atom equivalents, we obtained a considerably better 4.7 and 3.9 kcal/mol rms and a.a. deviations, respectively. The deviations are in the range of -18.5 and $+7.2$ kcal/mol with the average deviation of -2.9 kcal/mol. The reoptimized single-atom equivalent energy parameters yielded -5 and -17 kcal/mol deviation from experiment for naphthalene and azulene, respectively. The new single-atom equivalents for H, C, N, and O atoms are -0.5711 , -37.8849 , -54.4652 , and -74.7958 Hartrees, respectively.

Conclusions

We developed four new REEF-HF parameter sets for HF/6-31G(d) or HF/6-311+G(2d,p) energies and the corresponding NPA or Mulliken charges. These new parameter sets include implicitly the ZPE, relativistic, and thermal corrections. All parameter sets were developed and tested on the same molecular database, a set of 161 closed-shell neutral molecules composed of H, C, N, O, F, Si, S, and Cl atoms. We used the B3LYP/6-31G(d) equilibrium geometries in the calculations.

The following conclusions can be drawn from our results.

The parameter set developed for the HF-SCF/6-31G(d) method and NPA charges yield slightly better results than that developed for Mulliken charges (a.a. deviation, 1.97 vs 2.14 kcal/mol) for the 161 molecules in our current database.

For a set of 27 simple hydrocarbons, the REEF-HF method (using HF-SCF/6-31G(d) energies and NPA or Mulliken charges) yielded 0.77 or 0.64 kcal/mol a.a. deviations, respectively. This compares favorably to B3LYP/6-311G(d,p) enthalpies of formation corrected by atom equivalents yielding 0.81 kcal/mol a.a. deviation from experiment for 23 simple, less problematic hydrocarbons.

The statistics show that the 6-311+G(2d,p) basis set yields better results than the 6-31G(d) basis set for NPA charges (a.a.

deviations are 1.55 and 1.97 kcal/mol, respectively). For example, for the molecules containing Cl, the calculated enthalpy of formation values derived from HF-SCF/6-311+G(2d,p) results are considerably better than those derived from HF-SCF/6-31G(d) results. For Mulliken charges, no such improvement can be observed. We conclude that the combination of HF/6-311+G(2d,p) method with Mulliken charges is the least useful for our purposes, so this REEF-HF parameter set is not recommended for general use.

The most problematic molecules independent of basis set partial charge calculation method are azulene (117), which shows by far the largest negative (about -9 kcal/mol) deviation from experiment, carbonic difluoride (15), cyclopropene (22), bicyclo[1.1.0]butane (29), and butanedinitrile (92), which shows the largest positive (about $+6$ kcal/mol) deviation from experiment. The proposed method provides systematic deviation from experiment for strained rings and crowded molecules (e.g., *tert*-butyl group).

Leaving out the five most problematic molecules yielded 1.38 kcal/mol a.a. deviation for the remaining 154 molecules (calculations with NPA charges and 6-311+G(2d,p) basis set).

We observed that applying the same single-atom equivalent HF-SCF/6-31G(d) corrections proposed by Dewar et al. for cations and neutral molecules leads to large errors for the enthalpies of formation of cations. The enthalpies are too negative due to a systematic overcorrelation effect. We performed the same calculations with the REEF-HF energy parameters for NPA charges with the 6-31G(d) basis set and observed that the error is considerably smaller in our method, because the correlation energy of the missing electron is not included. The 30 kcal/mol deviation for CH_3^+ was decreased to 8 kcal/mol, the 12 kcal/mol deviation for H_3O^+ was decreased to 1 kcal/mol, and the 22 kcal/mol deviation for NH_4^+ was decreased to 13 kcal/mol.

The original single-atom equivalents parametrization of Dewar et al. for HF-SCF/6-31G(d) method yielded 8.7 kcal/mol rms deviation from experiment for 101 molecules in our database that contain H, C, N, and O atoms. After the reoptimization of the single-atom equivalents, we obtained new single-atom equivalents for H, C, N, and O atoms (-0.5711 , -37.8849 , -54.4652 , and -74.7958 Hartrees, respectively) that provide considerably better 4.7 and 3.9 kcal/mol rms and a.a. deviations, respectively.

Acknowledgment. This paper is dedicated to Professor Jean-Louis Rivail for his birthday. This work was supported by the OTKA Grant (T 034764 Hungary).

Appendix

To facilitate the reproduction of the REEF-HF results, we summarize the values of $[\Delta H_f^\circ(\text{A, exp}) - E_T(\text{A, G3})]$ in Hartrees used in the current paper in eqs 2 and 3 for various atoms.

H	0.583279	F	99.71364
C	38.098604	Si	289.39215
N	54.74367	S	398.06575
O	75.12500	Cl	460.03652

The atomic charges are available from the authors at <http://web.inc.bme.hu/csonka>.

References and Notes

- (1) Wiberg, K. B. *J. Comput. Chem.* **1984**, *5*, 197.
- (2) Ibrahim, M. R.; Schleyer, P. v. R. *J. Comput. Chem.* **1985**, *6*, 157.
- (3) Castro, E. A. *J. Mol. Struct. THEOCHEM* **1994**, *304*, 93.

- (4) Cohen, N.; Benson, S. W. *Chem. Rev.* **1993**, *93*, 2419–2438 and references therein.
- (5) Dewar, M. J. S.; Storch, D. M. *J. Am. Chem. Soc.* **1985**, *107*, 3898.
Dewar, M. J. S. *J. Phys. Chem.* **1985**, *89*, 2145.
- (6) Perdew, J. P.; Wang, Y. *Phys. Rev.* **1992**, *B45*, 13244.
- (7) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (8) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988**, *B37*, 785.
- (9) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; 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 92/DFT*; Gaussian: Pittsburgh, PA, 1993.
- (10) Mole, S. J.; Zhou, X.; Liu, R. *J. Phys. Chem.* **1996**, *100*, 14665.
- (11) Kristyán, S.; Csonka, G. I. *Chem. Phys. Lett.* **1999**, *307*, 469.
- (12) Kristyán, S.; Csonka, G. I. *J. Comput. Chem.* **2001**, *22*, 241.
- (13) Kristyán, S.; Ruzsinszky, A.; Csonka, G. I. *J. Phys. Chem. A* **2001**, *105*, 1926.
- (14) Kristyán, S.; Ruzsinszky, A.; Csonka, G. I. *Theor. Chem. Acc.* **2001**, *106*, 319.
- (15) Kristyán, S.; Ruzsinszky, A.; Csonka, G. I. *Theor. Chem. Acc.* **2001**, *106*, 404.
- (16) Ruzsinszky, A.; Alsenoy, C.; Csonka, G. I. *J. Phys. Chem.* **2002**, *106*, 12139.
- (17) Ruzsinszky, A.; Alsenoy, C.; Csonka, G. I. *J. Phys. Chem. A* **2003**, *107*, 736.
- (18) Csonka, G. I.; Ruzsinszky, A.; Oláh, J.; Van Alsenoy, C. *J. Mol. Struct. THEOCHEM* **2002**, *589–590*, 1.
- (19) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374; <http://chemistry.anl.gov/compmat/g3-99.htm>.
- (20) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063.
- (21) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 42; <http://chemistry.anl.gov/compmat/g2-97.htm>.
- (22) Reed, A. E.; Weinstock, R. B.; Weinhold, F. J. *J. Chem. Phys.* **1985**, *83*, 735.
- (23) Ruzsinszky, A.; Kristyán, S.; Margitfalvi, L. J.; Csonka, G. I. *J. Phys. Chem. A* **2003**, *107*, 1833.
- (24) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.
- (25) Kedziora, G. S.; Pople, J. A.; Ratner, M. A.; Redfern, P. C.; Curtiss, L. A. *J. Chem. Phys.* **2001**, *115*, 718.
- (26) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **2001**, *114*, 108.
- (27) CCCBDB Release 7, Sept 2002, <http://srdata.nist.gov/cccbdb/>.