

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/231664182>

# Optimized Parameters for Scaling Correlation Energy

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY A · APRIL 1999

Impact Factor: 2.69 · DOI: 10.1021/jp9900382

CITATIONS

153

READS

19

4 AUTHORS, INCLUDING:



**Jose C Corchado**

Universidad de Extremadura

**118** PUBLICATIONS **3,296** CITATIONS

SEE PROFILE



**M. Luz Sánchez**

Universidad de Extremadura

**42** PUBLICATIONS **1,671** CITATIONS

SEE PROFILE



**Donald Truhlar**

University of Minnesota Twin Cities

**1,342** PUBLICATIONS **81,184** CITATIONS

SEE PROFILE

## LETTERS

### Optimized Parameters for Scaling Correlation Energy

Patton L. Fast, Jose Corchado, Maria Luz Sanchez, and Donald G. Truhlar\*

*Department of Chemistry and Supercomputer Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431*

*Received: January 4, 1999; In Final Form: March 8, 1999*

Twelve general parametrizations of the SAC (scaling-all-correlation) method for semiempirical extrapolation of electronic structure calculations are presented. The methods are based on Møller–Plesset perturbation theory and coupled-cluster theory with correlation-consistent basis sets, and the parametrizations are based on 49 equilibrium atomization energies. This paper also presents an optimized scale factor for estimating the total anharmonic zero-point vibrational energy of a molecule with a root-mean-square accuracy of 0.17 kcal/mol.

#### 1. Introduction

In previous papers<sup>1–5</sup> a very simple approach has been suggested for extrapolating correlated electronic structure calculations to the limit of full dynamical correlation of the valence electrons and a complete one-electron basis set for the valence electrons. In the most broadly applicable approach, Gordon and one of the authors suggested using the relation<sup>2</sup>

$$E_{\text{SAC}} = E_{\text{HF}} + \frac{E_{\text{AC}}}{F} \quad (1)$$

where  $E_{\text{AC}}$  is the calculated valence correlation energy,  $E_{\text{HF}}$  is the Hartree–Fock (HF) energy, and  $F$  is a parameter. This expression is a simplified version of an earlier method in which the energy was approximated as<sup>1</sup>

$$E_{\text{SEC}} = E_{\text{CASSCF}} + \frac{E_{\text{EC}}}{F} \quad (2)$$

where  $E_{\text{EC}}$  is the calculated valence external correlation energy,  $E_{\text{CASSCF}}$  is the energy calculated by the full-valence complete-active-space self-consistent-field method. These methods are called scaling-all-correlation<sup>2–5</sup> (SAC) and scaling-external-

correlation<sup>1</sup> (SEC), respectively. The justification for eq 2 is that the fraction of external correlation energy recovered in a practical calculation is approximately a constant (i.e., independent of geometry) for a given correlation level and one-electron basis set.<sup>1</sup> Thus the energy predicted by eq 2 will yield a potential surface that parallels the exact one. The justification for eq 1 is that in many cases  $E_{\text{HF}}$  provides a good approximation to  $E_{\text{CASSCF}}$ . The SEC<sup>1,6,7</sup> and SAC<sup>2–5,7–10</sup> methods have been used for many successful applications.

The parameter  $F$ , which depends on the level of electron correlation and the one-electron basis set, may be determined from accurate ab initio calculations or from experiment. Furthermore it may be parametrized in a general way by averaging over a set of molecules, or the parameter may be specific to a particular reaction or a limited range of similar systems. The former is called a general parametrization, and the latter is called SRP (specific reaction parameters or specific range parameters). Previous general parametrizations have been based on 4–5,<sup>2,4</sup> 4–10,<sup>3</sup> and 13<sup>5</sup> molecules, and have yielded  $F$  values ranging from 0.31 to 0.98.<sup>2–5</sup> (Note that we would not normally recommend using the method when  $F \lesssim 0.65$  because we believe that extrapolation based on a calculation that includes less than about 65% of the valence correlation

energy is too unreliable for most purposes.) Siegbahn and co-workers<sup>10</sup> have used  $F = 0.80$  as a general parameter for this approach, which they call PCI-X or PCI-80. In the present paper, we present twelve general parametrizations (for various levels of electron correlation and two one-electron basis sets), each based on 49 experimental atomization energies. The resulting  $F$  parameters should be more widely useful than the previous ones for calculating thermochemical properties such as bond energies as well as for calculating potential curves and potential surfaces, including barrier heights for chemical reactions.

The use of equilibrium atomization energies (also called dissociation energies) to find optimum values of the  $F$  parameters is motivated by the fact that we are parametrizing the methods for the calculation of potential energy surface features, such as bond energies and barrier heights. The equilibrium dissociation energy  $\mathcal{D}_e$  is the sum of the bond energies of the molecule, and a barrier height is the difference between  $\mathcal{D}_e$  for reactants and  $\mathcal{D}_e$  for the transition state. (The atomization energies  $\mathcal{D}_e$  and  $\mathcal{D}_0$  should not be confused with the individual bond energies, which are usually denoted  $D_e$  and  $D_0$ .)

Section 3 presents the experimental data, which is obtained from a subset of the experimental data<sup>11</sup> used to parametrize the G2 training set, plus recent updates,<sup>12,13</sup> along with our own calculations of zero-point and thermal vibrational energy. Section 5 presents our new parameters and predicted atomization energies. Section 6 contains discussion.

## 2. Electronic Structure Calculations

All electronic structure methods used in this paper are frozen-core nonrelativistic calculations for molecules containing no atoms heavier than Cl. All calculations were carried out with the *Gaussian94* program.<sup>14</sup>

We use Dunning's<sup>15</sup> correlation-consistent polarized valence-double- $\zeta$  and valence-triple- $\zeta$  basis sets, denoted cc-pVDZ and cc-pVTZ, respectively. In the rest of the paper these basis sets are abbreviated pDZ and pTZ, respectively.

We employ seven levels of electron correlation:

MP2	Møller–Plesset (MP) second-order perturbation theory <sup>16</sup>
MP4SDQ	MP fourth-order perturbation theory with single, double, and quadruple excitations <sup>16</sup>
MP4	full MP4, i.e., MP4 with single, double, triple, and quadruple excitations <sup>15</sup>
CCD	coupled-cluster theory with double excitations <sup>17–19</sup>
QCISD	quadratic configuration interaction with single and double excitations <sup>20</sup>
CCSD	coupled-cluster theory with single and double excitations <sup>21</sup>
CCSD(T)	CCSD with a quasiperturbative treatment of connected triple excitations, including both a fourth-order-like term resulting from triples–doubles interaction and a fifth-order-like term resulting from triples–singles interaction. <sup>22</sup>

The reference wave function for correlated methods is restricted Hartree–Fock (RHF) for closed-shell species and unrestricted Hartree–Fock (UHF) for open-shell species.

All energy calculations reported in this page are single-point energies at MP2/pDZ geometries.

To compare methods for electronic structure calculations, it is useful to have a measure of computational effort, which we

call “cost.” To gauge the cost of the various methods and basis sets, we calculated single-point energies for six molecules with *Gaussian94* on an IBM SP computer, and we define cost as the average CPU time in seconds for these six calculations. The molecules chosen are the six largest molecules in the database of the next section of this paper, namely,  $\text{C}_2\text{H}_4$ ,  $\text{C}_2\text{H}_6$ ,  $\text{H}_3\text{COH}$ ,  $\text{H}_2\text{NNH}_2$ ,  $\text{Si}_2\text{H}_6$ , and  $\text{CH}_3\text{SH}$ . Although any cost measure is somewhat arbitrary, it is useful for evaluation purposes to have some measure of cost, and this is a reasonable measure of typical cost for single-point calculations on molecules with 5–8 atoms.

## 3. Equilibrium Atomization Energy Database

In this paper we use  $E$  and “energy” to denote the sum of the Born–Oppenheimer electronic energy and the nuclear repulsion energy for a given nuclear geometry. Thus the equilibrium dissociation energy of a molecule is given by

$$\mathcal{D}_e = E(\text{atoms}) - E(\mathbf{R} = \mathbf{R}_e) \quad (3)$$

The first term on the right is the energy of the infinitely separated atoms, and the second term is the energy of the molecule when all its internal coordinates, denoted collectively by  $\mathbf{R}$ , are at their so-called equilibrium values, i.e., the values that minimize  $E$ . The quantity  $\mathcal{D}_e$  will be called the equilibrium atomization energy.

Experimentally, one cannot directly measure  $E$ , but thermodynamic measurements extrapolated to 0 K can yield the ground-state dissociation energy defined by

$$\mathcal{D}_0 = \mathcal{D}_e - \text{ZPVE} \quad (4)$$

where ZPVE is the zero-point vibrational energy. For example

$$\mathcal{D}_0 = \left[ \sum_A \Delta H_{f,0}(\text{atom } A) \right] - \Delta H_{f,0}(\text{molecule}) \quad (5)$$

where  $\Delta H_{f,0}$  is the heat of formation at 0 K, and the sum in eq 5 is over all the atoms  $A$  in the molecule. The notations  $\mathcal{D}_e$  and  $\mathcal{D}_0$  follow Herzberg<sup>23</sup> with the understanding that in this paper  $\mathcal{D}_e$  and  $\mathcal{D}_0$  refer to simultaneous or successive breaking of *all* bonds in the molecule, i.e., the sum of all the bond dissociation energies. Since electronic structure calculations directly yield  $\mathcal{D}_e$  (as opposed to say  $\mathcal{D}_0$  or  $\Delta H_{f,298}$ ), it is convenient to have a table of best estimates of this quantity. We will create such a table by combining previously tabulated data for heats of formation with a consistent set of estimates of vibrational energy contributions. In most cases we use experimental  $\Delta H_{f,0}$  data, but in a few cases we use  $\Delta H_{f,298}$  experimental data. One can estimate  $\Delta H_{f,0}$  from  $\Delta H_{f,298}$  using the harmonic oscillator–rigid rotor approximation.<sup>24</sup>

The original Gaussian-2 (G2) test set<sup>11</sup> contains 55 molecules. All six molecules containing Li, Be, or Na were removed from the test set because correlation-consistent basis sets are not available for these elements and because six metal-containing molecules is not representative enough of the wide variety of metal bond types. This leaves 49 molecules. For 36 of these we used the experimental values of  $\mathcal{D}_0$  given in the original G2 paper,<sup>11</sup> but for 13 cases, namely,  $\text{CH}_4$ ,  $\text{SiH}_2(^1A_1)$ ,  $\text{SiH}_2(^3B_1)$ ,  $\text{SiH}_3$ ,  $\text{SiH}_4$ ,  $\text{HCCH}$ ,  $\text{HCN}$ ,  $\text{H}_3\text{COH}$ ,  $\text{Si}_2$ ,  $\text{SiO}$ ,  $\text{Si}_2\text{H}_6$ ,  $\text{CH}_3\text{Cl}$ , and  $\text{SO}_2$ , we used more recently compiled values of  $\Delta H_{f,298}^0$  for molecules<sup>12</sup> and atoms,<sup>13</sup> from which we estimated  $\Delta H_{f,0}^0$  and then  $\mathcal{D}_0$  as discussed above.

To estimate ZPVE from eq 4, we proceed as follows. First we note that Martin has<sup>25</sup> carefully estimated the experimental anharmonic zero-point energies for 13 molecules. We calculated

**TABLE 1: The MP2/cc-pVDZ and Experimental Anharmonic Zero-Point Energies (kcal/mol) for the 13 Molecules Used to Obtain the MP2/cc-pVDZ Zero-Point Energy Scaling Factor**

molecule	anharmonic ZPE	MP2/cc-pVDZ ZPE	MP2/cc-pVDZ <sup>a</sup> scaled ZPE
H <sub>2</sub>	6.21	6.43	6.30
CH <sub>4</sub>	27.71	28.49	27.89
NH <sub>3</sub>	21.20	21.85	21.39
H <sub>2</sub> O	13.25	13.57	13.29
HF	5.85	5.96	5.84
CO	3.11	3.02	2.96
N <sub>2</sub>	3.36	3.11	3.04
F <sub>2</sub>	1.30	1.33	1.31
C <sub>2</sub> H <sub>2</sub>	16.46	16.55	16.20
HCN	9.95	9.93	9.72
H <sub>2</sub> CO	16.53	16.87	16.52
CO <sub>2</sub>	7.24	7.24	7.09
N <sub>2</sub> O	6.77	6.74	6.59
RMS error		0.33	0.17

<sup>a</sup> Scale factor for MP2/cc-pVDZ zero-point energies is 0.9790.

**TABLE 2: Experimental Atomization Energies (kcal/mol)**

molecule	$\mathcal{D}_e$	molecule	$\mathcal{D}_e$
CH	84.00	HCO	278.39
CH <sub>2</sub> ( <sup>3</sup> B <sub>1</sub> )	190.07	H <sub>2</sub> CO	373.73
CH <sub>2</sub> ( <sup>1</sup> A <sub>1</sub> )	181.51	H <sub>3</sub> COH	512.90 <sup>a</sup>
CH <sub>3</sub>	307.65	N <sub>2</sub>	228.46
CH <sub>4</sub>	420.11 <sup>a</sup>	H <sub>2</sub> NNH <sub>2</sub>	438.60
NH	83.67	NO	155.22
NH <sub>2</sub>	181.90	O <sub>2</sub>	119.99
NH <sub>3</sub>	297.90	HOOH	268.57
OH	106.60	F <sub>2</sub>	38.20
OH <sub>2</sub>	232.55	CO <sub>2</sub>	389.14
FH	141.05	Si <sub>2</sub>	71.99 <sup>a</sup>
SiH <sub>2</sub> ( <sup>1</sup> A <sub>1</sub> )	151.79 <sup>a</sup>	P <sub>2</sub>	117.09
SiH <sub>2</sub> ( <sup>3</sup> B <sub>1</sub> )	131.05 <sup>a</sup>	S <sub>2</sub>	101.67
SiH <sub>3</sub>	227.37 <sup>a</sup>	Cl <sub>2</sub>	57.97
SiH <sub>4</sub>	322.40 <sup>a</sup>	SiO	192.08 <sup>a</sup>
PH <sub>2</sub>	153.20	SC	171.31
PH <sub>3</sub>	242.55	SO	125.00
SH <sub>2</sub>	182.74	ClO	64.49
ClH	106.50	ClF	61.36
HCCH	405.39 <sup>a</sup>	Si <sub>2</sub> H <sub>6</sub>	530.81 <sup>a</sup>
H <sub>2</sub> CCH <sub>2</sub>	563.47	CH <sub>3</sub> Cl	394.64 <sup>a</sup>
H <sub>3</sub> CCH <sub>3</sub>	712.80	CH <sub>3</sub> SH	473.84
CN	180.58	HOCl	164.36
HCN	313.20 <sup>a</sup>	SO <sub>2</sub>	257.86 <sup>a</sup>
CO	259.31		

<sup>a</sup> Calculated from  $\Delta H_{f,298}^0$  given in refs 12 and 13.

these zero-point energies in the harmonic approximation at the MP2/cc-pVDZ level, and we found that a scale factor of 0.9790 minimizes the root-mean-square (RMS) deviation between calculated and experimental ZPVEs. The RMS deviation for these 13 molecules is 0.33 kcal/mol without scaling and 0.17 kcal/mol with scaling. The individual scaled and unscaled ZPVEs are compared to experiment in Table 1. In creating our  $\mathcal{D}_e$  database we will use Martin's estimate of ZPVE for the 11 molecules in his set that also appear in our database and the scaled MP2/cc-pVDZ value for the remaining molecules. The ZPVE is removed from  $\mathcal{D}_0$  to obtain equilibrium energies of atomization,  $\mathcal{D}_e$ .

To estimate  $\Delta H_{f,0}^0$  from  $\Delta H_{f,298}^0$  for the seven cases where we used  $\Delta H_{f,298}^0$  data, we also used MP2/cc-pVDZ frequencies scaled by 0.9790.

The resulting  $\mathcal{D}_e$  data set is given in Table 2.

**TABLE 3: Spin–Orbit Energies (kcal/mol) Used in Eq 6<sup>a</sup>**

species	$E_{SO}$	species	$E_{SO}$
C	−0.09 <sup>b</sup>	CH( <sup>2</sup> Π)	−0.04 <sup>c</sup>
O	−0.023 <sup>b</sup>	OH( <sup>2</sup> Π)	−0.20 <sup>c</sup>
F	−0.38 <sup>b</sup>	NO( <sup>2</sup> Π)	−0.18 <sup>c</sup>
Si	−0.43 <sup>b</sup>	ClO	−0.46 <sup>c</sup>
S	−0.56 <sup>b</sup>	Si <sub>2</sub> ( <sup>3</sup> Π)	−0.20 <sup>c</sup>
Cl	−0.84 <sup>b</sup>	CH <sub>2</sub> ( <sup>3</sup> B <sub>1</sub> )	0.00 <sup>d</sup>
		SiH <sub>2</sub> ( <sup>3</sup> B <sub>2</sub> )	0.00 <sup>d</sup>

<sup>a</sup> Closed-shell molecules, atoms in *S* states, linear molecules in  $\Sigma$  states, and singlet and doublet molecules in *A* or *B* states are not listed since  $E_{SO}$  is necessarily zero in the Russell–Saunders scheme used here. <sup>b</sup> Ref 26. <sup>c</sup> Ref 23. <sup>d</sup> Ref 27; in the Table 0.00 means less than 0.005.

#### 4. Spin–Orbit Energy

All electronic structure methods considered in this paper are frozen-core, nonrelativistic calculations for molecules containing no atom heavier than Cl. The most important relativistic correction to atomization energies for molecules with no atom heavier than Cl is spin–orbit coupling. For closed-shell species, spin–orbit coupling is zero. For open-shell species, it is always negative, since we use Russell–Saunders (LS or  $\Sigma\Lambda$ ) coupling and consider only dissociation of the lowest-energy electronic state with a given spin multiplicity. The magnitude of  $E_{SO}$  is the difference in energy between the ground-state energy and the degeneracy-weighted average energy of the multiplet. The spin–orbit energies<sup>23,26,27</sup> for all atoms and molecules in this paper for which  $E_{SO}$  is nonzero are tabulated in Table 3.

#### 5. Optimized SAC Parameters

The calculated value of the equilibrium dissociation energy is

$$\mathcal{D}_e = \sum_A [E_{SAC}(A) + E_{SO}(A)] - [E_{SAC}(\text{molecule}) + E_{SO}(\text{molecule})] \quad (6)$$

where  $E_{SAC}$  is given by eq 1, and the sum is over all atoms in the molecule. In eq 1,  $E_{AC}$  is the difference between the MP2, CCSD, or CCSD(T) energy and the HF energy. Notice that when eq 1 is substituted into eq 6, it is linear in  $1/F$ , and therefore  $F$  was optimized by the least-squares method to minimize the root-mean-square (RMS) deviation of the calculated  $\mathcal{D}_e$  values from the accurate values in Table 2. SAC atomization energies for a few of the methods are given as examples in Table 4, and the optimum  $F$  values for all cases are given in Table 5.

#### 6. Discussion

Table 5 gives the mean signed error (MSE), mean unsigned error (MUS), RMS error (RMSE), and cost for each of the SAC methods. Table 6 gives the same quantities for the ab initio methods, i.e., using eq 6 with  $F = 1$ . Comparison of these tables shows that the SAC procedure lowers the MUE by large factors.

Ideally, we would provide explicit recommendations about the accuracy to be expected for predictions on additional molecules. This is difficult because we found that the errors are not normally distributed. Particularly large errors were obtained for CN, P<sub>2</sub>, S<sub>2</sub>, and SO<sub>2</sub>. Clearly, further testing of the method would be useful, but in the absence of further testing, the values in Table 5 provide guidelines.

We note that the mean errors in the atomization energies are approximately 2.3 times larger than the mean error in the individual bond energies because our 49 molecules have 113

**TABLE 4: SAC Calculations of Equilibrium Atomization Energies (kcal/mol)**

molecule	atomization energy $\mathcal{D}_e$			
	MP4SDQ-SAC/pDZ	MP4SDQ-SAC/pTZ	CCSD(T)-SAC/pDZ	CCSD(T)-SAC/pTZ
CH	83.09	84.67	83.17	84.39
CH <sub>2</sub> ( <sup>3</sup> B <sub>1</sub> )	188.68	192.07	186.85	190.31
CH <sub>2</sub> ( <sup>1</sup> A <sub>1</sub> )	181.05	183.01	180.90	182.22
CH <sub>3</sub>	308.09	311.35	305.42	308.60
CH <sub>4</sub>	425.29	426.60	421.74	422.95
NH	80.19	82.96	80.04	82.60
NH <sub>2</sub>	179.57	183.43	178.70	182.31
NH <sub>3</sub>	297.05	300.63	295.20	298.63
OH	105.95	107.76	105.14	107.10
OH <sub>2</sub>	233.29	235.30	231.39	233.76
FH	140.40	142.56	138.91	141.47
SiH <sub>2</sub> ( <sup>1</sup> A <sub>1</sub> )	152.61	154.74	152.40	153.70
SiH <sub>2</sub> ( <sup>3</sup> B <sub>1</sub> )	128.86	132.46	128.56	131.79
SiH <sub>3</sub>	224.55	228.71	223.37	226.91
SiH <sub>4</sub>	322.72	326.96	320.61	323.98
PH <sub>2</sub>	149.33	153.32	149.79	152.91
PH <sub>3</sub>	238.58	242.34	238.27	241.00
SH <sub>2</sub>	182.47	183.54	181.69	182.57
ClH	106.49	106.77	105.93	106.38
HCCH	402.94	404.38	403.23	404.71
H <sub>2</sub> CCH <sub>2</sub>	566.47	567.73	564.27	565.45
H <sub>3</sub> CCH <sub>3</sub>	720.62	721.77	715.25	716.50
CN	162.57	163.10	179.15	177.16
HCN	316.09	311.84	317.49	313.16
CO	268.26	260.20	267.80	260.59
HCO	284.28	279.26	284.22	279.82
H <sub>2</sub> CO	381.14	377.03	379.92	376.14
H <sub>3</sub> COH	517.41	518.54	513.83	515.38
N <sub>2</sub>	233.82	226.10	235.78	227.94
H <sub>2</sub> NNH <sub>2</sub>	437.15	440.68	435.47	438.92
NO	158.71	151.36	160.83	153.60
O <sub>2</sub>	131.80	121.13	132.67	122.97
HOOH	275.82	272.37	275.36	272.42
F <sub>2</sub>	49.35	40.83	51.04	42.72
CO <sub>2</sub>	398.62	388.40	399.89	390.64
Si <sub>2</sub>	68.26	68.87	72.41	72.92
P <sub>2</sub>	104.31	104.38	112.40	111.84
S <sub>2</sub>	92.00	93.01	95.47	97.40
Cl <sub>2</sub>	49.60	52.53	50.97	54.59
SiO	187.72	188.00	186.70	188.48
SC	171.89	166.01	175.42	169.76
SO	116.03	117.91	117.97	120.82
ClO	52.93	55.81	57.68	60.96
ClF	57.80	58.40	58.50	59.53
Si <sub>2</sub> H <sub>6</sub>	527.34	536.75	524.55	532.62
CH <sub>3</sub> Cl	394.96	396.69	392.81	394.92
CH <sub>3</sub> SH	474.18	476.61	471.80	474.29
HOCl	163.19	163.79	163.66	164.81
SO <sub>2</sub>	232.75	241.34	236.44	246.18

bonds. For the mean signed error this approximation relation becomes exact. To emphasize this, Tables 5 and 6 also show the mean signed errors in the individual bond dissociation energies.

Although we have based the present parametrization on single-point energies calculated by MP2/pDZ geometries, these geometries should be close enough to the accurate ones that the  $F$  factors should also be valid at other geometries. In fact, although this is not exploited here, SAC calculations can also be used to optimize geometries.<sup>8</sup>

In previous work using the SAC method with smaller databases and the less systematic Pople-type basis sets, the advantages of MP4 over MP2 were not very clear, but the present study, based on a much larger data set and the more systematic correlation-consistent basis sets, shows that MP4-SAC is almost twice as accurate as MP2-SAC. Perhaps surprisingly, the MP4SDQ-SAC method emerges as a very

**TABLE 5: Scaling Factors, Mean Errors (kcal/mol), and Cost for SAC Methods**

method	$F$	$\mathcal{D}_e$			bond energies <sup>a</sup>	
		MSE	MUE	RMSE	MSE	cost
MP2-SAC/pDZ	0.7766	-2.66	9.39	11.53	-1.15	66
MP4SDQ-SAC/pDZ	0.6989	-0.83	4.82	6.96	-0.36	94
MP4-SAC/pDZ	0.7488	-1.07	5.46	6.80	-0.46	129
QCISD-SAC/pDZ	0.6955	-0.53	4.19	6.28	-0.23	166
CCD-SAC/pDZ	0.6750	-0.83	5.39	8.10	-0.36	153
CCSD-SAC/pDZ	0.6898	-0.55	4.43	6.66	-0.24	205
CCSD(T)-SAC/pDZ	0.7323	-0.53	4.27	5.78	-0.23	256
MP2-SAC/pTZ	0.9454	-1.56	5.89	7.29	-0.68	893
MP4SDQ-SAC/pTZ	0.8436	-0.47	3.73	5.33	-0.20	1287
MP4-SAC/pTZ	0.9214	-0.53	2.44	3.54	-0.23	2787
CCSD-SAC/pTZ	0.8318	-0.26	4.06	5.29	-0.11	2021
CCSD(T)-SAC/pTZ	0.8928	-0.13	1.88	2.73	-0.06	6920

<sup>a</sup> 113 bonds in data set (counting a double or triple bond as one bond).

**TABLE 6: Mean Errors (kcal/mol) and Cost for ab Initio Methods**

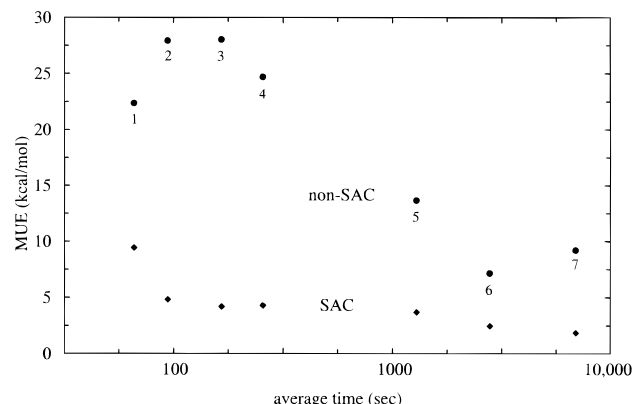
method	$\mathcal{D}_e$			bond energies <sup>a</sup>	
	MSE	MUE	RMSE	MSE	cost
HF/pDZ	-90.81	90.81	99.90	-39.38	34
MP2/pDZ	-22.35	22.35	24.99	-9.69	66
MP4SDQ/pDZ	-27.93	27.93	30.81	-12.11	94
MP4/pDZ	-23.61	23.61	25.95	-10.24	129
QCISD/pDZ	-28.02	28.02	31.00	-12.15	166
CCD/pDZ	-30.07	30.07	33.36	-13.04	153
CCSD/pDZ	-28.55	28.55	31.63	-12.38	205
CCSD(T)/pDZ	-24.70	24.70	27.32	-10.71	256
HF/pTZ	-84.82	84.82	93.37	-36.78	450
MP2/pTZ	-6.11	7.14	8.89	-2.65	893
MP4SDQ/pTZ	-13.66	13.66	15.52	-5.92	1287
MP4/pTZ	-7.15	7.15	8.14	-3.10	2787
CCSD/pTZ	-14.49	14.49	16.55	-6.28	2021
CCSD(T)/pTZ	-9.21	9.21	10.37	-3.99	6920

<sup>a</sup> 113 bonds in data set (counting a double or triple bond as one bond).

competitive scheme in terms of performance for a given cost. The computational effort of the MP4SDQ method, like CCSD, scales as  $o^2v$ ,<sup>4</sup> where  $o$  is the number of occupied orbitals and  $v$  is the number of virtual orbitals, whereas methods including triples, i.e., MP4 and CCSD(T) scale as  $o^3v^4$ .<sup>16,28</sup> (The MP2 method scales as  $ov$ .<sup>4</sup>) The inclusion of triple excitations, i.e., MP4 vs MP4SDQ and CCSD(T) vs CCSD, is very useful with the pTZ basis sets, but their effect is much less dramatic at the pDZ level.

Examination of Tables 5 and 6 shows that some of the methods are particularly attractive in terms of quality of the results at minimal cost. In particular, we most strongly recommend the following seven combinations: MP2-SAC/pDZ, MP4SDQ-SAC/pDZ, QCISD-SAC/pDZ, CCSD(T)-SAC/pDZ, MP4SDQ-SAC/pTZ, MP4-SAC/pTZ, and CCSD(T)-SAC/pTZ. In Figure 1, the mean unsigned error is plotted vs cost for these methods, and in each case we also plot the mean unsigned error that is obtained by the unscaled version of that same electron correlation method and basis set. Scaling the MP4SDQ/pDZ, QCISD/pDZ, and MP4SDQ/pTZ results is particularly powerful, reducing the mean unsigned error by factors of 5.8, 6.7, and 3.7, respectively. The MP4SDQ-SAC/pDZ calculation gives a mean unsigned error of 4.82 kcal/mol with an average cost of 94 s, whereas the best method without scaling (Table 6) gives a mean unsigned error of 7.14 kcal/mol with a cost of 893 s—the error is 1.5 times larger despite the fact that the cost is 9.5 times larger. Similarly the QCISD-SAC/pDZ method gives a





**Figure 1.** The mean unsigned error (MUE) vs the logarithm of the cost (average CPU time for the six largest molecules) for selected SAC (◆) and non-SAC (●) methods: (1) MP2/pDZ, (2) MP4SDQ/pDZ, (3) QCISD/pDZ, (4) CCSD(T)/pDZ, (5) MP4SDQ/pTZ, (6) MP4/pTZ, and (7) CCSD(T)/pTZ.

mean unsigned error 1.7 times smaller than MP2/pTZ at a cost 5.8 times smaller. As a consequence of these and the other results in this paper, it would seem advisable to scale the correlation energy whenever one wants the most accurate possible energies for a given cost.

**Acknowledgment.** This work was supported in part by the U.S. Department of Energy, Office of Basic Energy Sciences.

## References and Notes

- (1) Brown, F. B.; Truhlar, D. G. *Chem. Phys. Lett.* **1985**, *117*, 307.
- (2) Gordon, M. S.; Truhlar, D. G. *J. Am. Chem. Soc.* **1986**, *108*, 5412.
- (3) Gordon, M. S.; Truhlar, D. G. *Int. J. Quantum Chem.* **1987**, *31*, 81.
- (4) Gordon, M. S.; Nguyen, K. A.; Truhlar, D. G. *J. Phys. Chem.* **1989**, *93*, 7356.
- (5) Rossi, I.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *234*, 64.
- (6) (a) Steckler, R.; Schwenke, D. W.; Brown, F. B.; Truhlar, D. G. *Chem. Phys. Lett.* **1985**, *121*, 475. (b) Schwenke, D. W.; Steckler, R.; Brown, F. B.; Truhlar, D. G. *J. Chem. Phys.* **1986**, *84*, 5706. (c) Schwenke, D. W.; Tucker, S. C.; Steckler, R.; Brown, F. B.; Lynch, G. C.; Truhlar, D. G.; Garrett, B. C. *J. Chem. Phys.* **1989**, *90*, 3110. (d) Lynch, G. C.; Truhlar, D. G.; Brown, F. B.; Zhao, J.-g. *Hyperfine Interactions* **1994**, *87*, 885.
- (7) Corchado, J. C.; Truhlar, D. G. *ACS Symp. Ser.* **1998**, *712*, 201–219.
- (8) Truong, T. N.; Truhlar, D. G.; Baldrige, K. K.; Gordon, M. S.; Steckler, R. *J. Chem. Phys.* **1989**, *90*, 7137.
- (9) (a) Melissas, V. S.; Truhlar, D. G. *J. Chem. Phys.* **1993**, *99*, 1013. (b) Melissas, V. S.; Truhlar, D. G. *J. Phys. Chem.* **1994**, *98*, 875. (c) Hu, W.-P.; Rossi, I.; Corchado, J. C.; Truhlar, D. G. *J. Phys. Chem. A*, **1997**, *101*, 6911. (d) Truong, T. N.; Truhlar, D. G. *J. Chem. Phys.* **1990**, *93*, 1761. Erratum: **1992**, *97*, 8820. (e) Nguyen, A. A.; Gordon, M. S.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 1596. (f) Espinosa-García, J.; Corchado, J. C.; Sana, M. J. *Chem. Phys.* **1993**, *90*, 1181. (g) Espinosa-García, J.; Tolosa, S.; Corchado, J. C. *J. Phys. Chem.* **1994**, *98*, 2337. (h) Espinosa-García, J.; Corchado, J. C. *J. Chem. Phys.* **1994**, *101*, 1333. (i) Corchado, J. C.; Olivares del Valle, F. J.; Espinosa-García, J. *J. Phys. Chem.* **1993**, *97*, 9129. Erratum: **1994**, *98*, 5796. (j) Espinosa-García, J.; Corchado, J. C. *J. Chem. Phys.* **1994**, *101*, 8700. (k) Allison, T.; Lynch, G. C.; Truhlar, D. G.; Gordon, M. S. *J. Phys. Chem.* **1996**, *100*, 13575. (l) Melissas, V. S.; Truhlar, D. G. *J. Chem. Phys.* **1993**, *99*, 3542. (m) Irikura, K. K.; Frurip, D. J. *ACS Symp. Ser.* **1998**, *677*, 419.
- (10) (a) Siegbahn, P. E. M.; Blomberg, M. R. A.; Svensson, M.; Boussard, P. J. E. *J. Chem. Phys.* **1995**, *102*, 5377. (b) Siegbahn, P. E. M.; Crabtree, R. H. *J. Am. Chem. Soc.* **1996**, *118*, 4442. (c) Blomberg, M. R. A.; Siegbahn, P. E. M. *ACS Symp. Ser.* **1998**, *677*, 197.
- (11) Pople, J.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622.
- (12) Petersson, G. A.; Malick, D. K.; Wilson, W. G.; Ochterski, J. W.; Montgomery, J. A., Jr.; Frisch, M. J. *J. Chem. Phys.* **1998**, *109*, 10570.
- (13) Ochterski, J. W.; Petersson, G. A.; Wiberg, K. B. *J. Am. Chem. Soc.* **1995**, *117*, 11299.
- (14) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian94* (Revision E.2); Gaussian, Inc.: Pittsburgh, PA, 1995.
- (15) (a) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007. (b) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1994**, *100*, 2975. (c) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007. (d) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358.
- (16) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.
- (17) Cizek, J. *Adv. Chem. Phys.* **1969**, *14*, 35.
- (18) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem. Symp.* **1978**, *14*, 545.
- (19) Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359.
- (20) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968.
- (21) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910.
- (22) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.
- (23) Herzberg K. G. In *Molecular Spectra and Molecular Structure. I*, 2nd ed.; D. van Nostrand: Princeton, 1950.
- (24) Irikura, K. K. *ACS Symp. Ser.* **1998**, *677*, 402.
- (25) Martin, J. M. L. *J. Chem. Phys.* **1992**, *97*, 5012.
- (26) Moore, C. National Bureau of Standards (U.S.) Circular 467, 1952.
- (27) (a) Langhoff, S. R. *J. Chem. Phys.* **1974**, *61*, 3881. (b) Havlas, Z.; Downing, J. W.; Michl, J. *J. Phys. Chem.* **1998**, *102*, 5681.
- (28) Raghavachari, K.; Anderson, J. B. *J. Phys. Chem.* **1996**, *100*, 12960.