J. Phys. Chem. A **2000,** *104,* 6111−6116

MC-QCISD: Multi-Coefficient Correlation Method Based on Quadratic Configuration Interaction with Single and Double Excitations

Patton L. Fast and Donald G. Truhlar*

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

Received: February 2, 2000; In Final Form: April 28, 2000

This paper presents a multi-coefficient correlation method based on quadratic configuration interaction with single and double excitations (MC-QCISD) and basis sets using segmented contraction and having the same exponential parameters in the *s* and *p* spaces. The results are comparable to a previous multi-coefficient correlation method based on coupled cluster theory with less efficient correlation-consistent basis sets, and they are better than a previous multi-coefficient correlation method based on Møller−Plesset fourth order perturbation theory with single, double, and quadruple excitations with correlation-consistent basis functions. The mean unsigned error per bond of the MC-QCISD method is 0.72 kcal/mol. The new method should be very efficient for computing geometries of open-shell transition states.

1. Introduction

Multilevel techniques based on scaling (e.g., SEC,[1−3] SAC,[3−12] MCSAC[11,12]), extrapolation (e.g., CBS,[13−16] exponential and power law extrapolation,[17−22] IB,[25,26] and EIB[11,12]), additive energy corrections (e.g., G2,[27,28] G3,[29] and their many variants[30]), and multi-coefficient correlation methods (e.g., MCCM,[11,12] MCG2,[31] MCG3,[32] and G3S[33]) have proved to be powerful methods for calculating high-accuracy thermochemistry and potential energy surfaces (or potential energy surface features such as saddle point energies) with less expense than brute force methods. For example, a brute force method of attaining the accuracy of Møller−Plesset perturbation theory of fourth order[34] (MP4) with a polarized triple-ζ (pTZ) basis set would actually carry out an MP4/pTZ calculation, whereas a multilevel calculation might base this on separate MP2/pTZ and MP4/pDZ calculations (where 2 denotes second order and D double-ζ). In empirical methods such as SAC, G2, MCCM, etc., the calculations contain semiempirical parameters so that the result should actually be more accurate than the brute force result (MP4/pTZ in our example). In effect, one is extrapolating toward infinite-order electron correlation (full configuration interaction, for short FCI) with an infinite one-electron basis (for short, infinite basis or IB); this combination is called complete configuration interaction.

We have recently shown that multilevel methods of the linear combination type may also be used to extrapolate energy gradients and thereby to calculate more accurate molecular geometries.[35] In practice, molecular geometries of stable species can typically be calculated accurately with lower orders of electron correlation and smaller one-electron basis sets than are required for accurate energies. However, the calculation of saddle point geometries is much more difficult. Saddle point geometries tend to be sensitive to having a balanced treatment of the multiconfigurational character of the wave function and (because of Hammond's principle[36]) to having a balanced treatment of the energetics of the breaking and making bonds. Furthermore, saddle point wave functions for radical reactions are very susceptible to contamination by components representing incorrect spin states. Durant and Rohlfing[30a,37] and our group[38] have both recommended the QCISD method (quadratic configuration interaction with single and double excitations[39]) for the calculation of saddle point geometries of radical reactions. In fact, of all of the methods for analytical energy gradients that are available in the popular and versatile *Gaussian* electronic structure package,[40] QCISD would be expected to yield the most accurate saddle point geometries.

Therefore, it would be desirable to have a multi-coefficient correlation method based on QCISD and on one-electron basis functions that are handled efficiently by *Gaussian*. (We note that *Gaussian* is designed to be particularly efficient for Gaussian basis sets in which the same exponential parameters are used for *s* and *p* functions[41] and in which each primitive Gaussian function occurs in only one contracted function, which is called segmented contraction.) The goal of the present paper is to parameterize such a method. It will utilize the 6-31G(d),[34] 6-31G(2df,p),[33] and MG3 basis sets,[29,32] and the resulting method will be called MC-QCISD (multi-coefficient QCISD). We note that the MG3 basis set is also called G3MP2large[30b] basis set, and it is essentially an improved[29] version of the 6-311++G(3d2f,2df,2p) basis set.

2. Theory and Parameterization

The MC-QCISD energy is defined by

$$E(\text{MC-QCISD}) = c_0E(\text{HF/6-31G(d)}) + c_1\Delta E(\text{MP2}|\text{HF/6-31G(d)}) + c_2\Delta E(\text{MP2/MG3}|\text{6-31G(d)}) + c_3\Delta E(\text{QCISD}|\text{MP2/6-31G(d)}) \quad (1)$$

where

$$\Delta E(\text{L2}|\text{L1/B}) \equiv E(\text{L2/B}) - E(\text{L1/B}) \quad (2)$$

and

$$\Delta E(\text{L/B2}|\text{B1}) \equiv E(\text{L/B2}) - E(\text{L/B1}) \quad (3)$$

The motivation for this kind of form was explained previously:[11] c_0 and c_2 are motivated by extrapolating to an IB; c_1 and c_3

10.1021/jp000408i CCC: $19.00

TABLE 1: Coefficients for the Minimal Methods Optimized in This Work over the 82-Molecule Data Set[a]

method	version	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7
MC-QCISD	v2m	1.0038	1.0940	1.2047	1.0441				
SAC-MP2/6-31G(d)	v2m	1.0000	1.2207						
SAC-MP2/6-31+G(d)	v2m	1.0000	1.2290						
SAC-QCISD/6-31G(d)	v2m	1.0000	1.4457						
SAC-QCISD/6-31G(2df,p)	v2m	1.0000	1.1913						
MCSAC-QCISD/6-31G(d)	v2m	1.0000	1.6024	2.5063					
MCSAC-QCISD/6-31G(2df,p)	v2m	1.0000	1.1955	1.2199					
MCCM-CO-MP2;6-31G(2df,p);6-31G(d)	v2m	0.9869	1.8358	0.8175	2.0673				
MCCM-CO-QCISD;6-31G(2df,p);6-31G(d)	v2m	0.9782	1.1504	1.1190	1.7517	1.1121	0.7136		
MCCM-UT-QCISD;6-31G(2df,p);6-31G(d)	v2m	0.9804	1.0091	1.1449	1.6604	1.3032			
MCCM-CO-MP2;MG3;6-31+G(d)	v2m	0.9724	1.2936	0.8577	2.0067				
SAC-MP2/pDZ	v2m	1.0000	1.2318						
SAC-MP4SDQ/pDZ	v2m	1.0000	1.4370						
MCCM-CO-MP2	v2m	0.9918	1.0276	0.7833	2.6875				
MCCM-UT-MP4SDQ	v2m	0.9934	1.2606	1.0363	1.6307	0.8541			
MCCM-UT-CCSD	v2m	0.9969	1.2625	1.0610	1.4813	0.8312			
MCG3	v2m	1.0121	1.2047	1.0646	1.0975	1.1859	0.8139	1.4470	1.414

[a] The notation for MC-QCISD is defined by eq 1; the definitions of the coefficients for the other methods are explained in ref 12.

are motivated by extrapolating to FCI. The four parameters are determined by the linear least-squares methods to minimize the root-mean-square error in the predicted zero-point-exclusive atomization energies of an 82-molecule training set. These 82 data are tabulated (with references) in a previous paper.[12]

All MP2 and QCISD calculations in the present paper have a frozen core. We do not include core correlation, spin−orbit, or scalar relativistic contributions explicitly in the present model, and thus they are implicit in the model. This approach is called a "minimal" model in previous work.[12,31,32]

The atomization energies were calculated for all the component methods of eq 1 using geometries optimized by the MP2-(full)/6-31G(d) method.[42] The parameters, c_0 through c_3, were optimized, and the resulting values are 1.0038, 1.0940, 1.2047, and 1.0441, respectively. The parameters are all positive numbers, which is physical, and they are all less than 1.5 (in fact considerably so), which should make for stable calculations. Thus we accept the fit as physical.

To place the present method in perspective, we optimized the parameters for 16 other minimal methods, using the 82 experimental atomization energies. Six of these multilevel methods are the ones that we previously recommended (out of a total of 42 multilevel methods examined),[11,12] on the basis of performance/cost ratios, as the most promising general parameterizations for molecules containing first- and second-row elements. These methods have all been presented elsewhere,[10−12,29] although the "minimal"-type parameterization is new. The other 11 methods are various ways to combine subsets of the components available from the calculations used in the MC-QCISD method plus (in one case) the 6-31+G(d) basis set. The notation for these methods is based on our standard nomenclature[11,12] for multilevel methods and requires no further explanation. In addition, we will compare our results to the G3[29] and G3S[33] methods, for which we accept the parameters of Curtiss et al.[29,33]

In optimizing the 17 sets of new coefficients, we used MP2/cc-pVDZ[43] (for short pDZ) geometries for the five methods based on correlation consistent basis sets[43] (these methods are identified in Table 1), and MP2(full)/6-31G(d) geometries for all other methods. In this respect we note a philosophical difference between our approach and that used in the G2, G3, and G3S methods. In G2, G3, and G3S, the geometry choice is considered part of the method. In our work, although of course we must choose a geometry for obtaining parameters and calculating mean errors, we consider it perfectly acceptable to use our methods with other geometries, provided the geometry choice is stated. (A second difference is that we focus on zero-point-exclusive atomization energies, which are a difference between two points on the potential energy surface, whereas the G2, G3, and G3S methods focus more directly on heats of formation.)

For the six previously recommended[12] methods, the new parameters presented here, based on our now "standard" 82-molecule training set, are denoted version 2m (second minimal-type parameterization) coefficients to distinguish them from versions 1, 1m, 1sc, 2s, and 2sc presented previously.[10−12] The differences between version 1 and version 2 are very important since version 2 is based on a larger, better training set. The differences between m, s, and sc versions are small and simply distinguish between explicit or implicit core correlation and spin−orbit contributions; nevertheless, for consistency this paper uses only the minimal-type parameterizations.

Note that there is only one version of MC-QCISD coefficients at this time; it is called version 2m. The "2" denotes that it was parameterized with the version-2 training set (not that it is a second set of coefficients).

3. Results

The mean errors for all 18 methods over the 82-molecule data set are given in Table 2. For consistency, the errors were calculated using MP2(full)/6-31G(d) geometries for methods based on segmented, equalized-exponent basis sets (the G3 and G3S methods and methods based on 6-31G(d), 6-31+G(d), 6-31G(2df,p), and MG3 basis sets) and were calculated using MP2/cc-pVDZ geometries for the other methods, since these other methods may be used with any reasonable geometries, but they all involve MP2/cc-pVDZ as one of their components or as a stage of one of their components.

An estimation of the cost for each of the methods is also given in Table 2. This estimation is the time for one single-point gradient calculation on furan (C_4H_4O) divided by the time for an HF/6-31G(d) gradient calculation. This time is provided to given an example of the differing computational costs of the various calculations; the timings would be similar but not identical for other first-row atoms of comparable size. We are of course aware that the "costs" would be different if we choose a different molecule as a standard or if we emphasized energies or Hessians rather than gradients. Nevertheless, gradients were selected because they are a central component of geometry

TABLE 2: Mean Errors (kcal/mol) over the 82-Molecule Data Set

method	n	N	MSE	MUE	RMSE	cost[a]	HR[b]
SAC-MP2/6-31G(d)	5	1	−6.62	17.76	21.76	3	
SAC-MP2/6-31+G(d)	5	1	−6.04	16.95	20.53	3	
SAC-MP2/pDZ	5	1	−3.64	10.69	12.73	4	yes[c]
SAC-QCISD/6-31G(d)	6	1	−3.35	11.57	14.41	12	
SAC-MP4SDQ/pDZ	6	1	−1.53	7.06	10.77	13	yes[c]
MCSAC-QCISD/6-31G(d)	6	2	−1.76	8.91	12.22	14	yes[d]
MCCM-CO-MP2;6-31G(2df,p);6-31G(d)	5	4	−0.65	6.47	8.67	31	
MCCM-UT-QCISD;6-31G(2df,p);6-31G(d)	6	5	−0.08	3.64	5.11	42	yes[d]
SAC-QCISD/6-31G(2df,p)	6	1	−0.70	4.28	5.99	48	
MC-QCISD	6	4	0.15	2.51	3.81	70	yes[d]
MCSAC-QCISD/6-31G(2df,p)	6	2	−0.66	4.24	5.98	70	
MCCM-CO-MP2;MG3;6-31+G(d)	5	4	−0.48	4.64	6.36	84	yes[d]
MCCM-CO-QCISD;6-31G(2df,p);6-31G(d)	6	6	−0.08	3.44	4.92	84	
MCCM-CO-MP2	5	4	−1.41	5.41	6.74	99	yes[c]
MCCM-UT-MP4SDQ	6	5	−0.32	2.92	4.32	111	yes[c]
G3(MP2)	7	4	−1.53	2.13	2.96	728	
MCCM-UT-CCSD	6	5	−0.15	2.73	3.75	939	yes[c]
MCG3	7	8	0.15	1.24	1.63	1110	yes[c]
G3	7	4	−0.44	1.26	1.69	6374	yes[c]
G3S	7	6	−0.83	1.40	1.84	6515	yes[d]

[a] CPU time for one C_4H_4O gradient calculation on a single processor of an Origin 2000 with R12000 processors as calculated by Gaussian98 (Rev. A.7) and normalized to the time (83 s) for one HF/6-31G(d) gradient. All gradients were calculated with C_{2v} symmetry. [b] HR denotes the highly recommended methods. [c] highly recommended in ref 12. [d] highly recommended in this paper.

TABLE 3: Mean Errors (kcal/mol) over the 82-Molecule Data Set for All of the Components[a]

method	n	MSE	MUE	RMSE	cost
HF/6-31G(d)	4	−123.08	123.08	142.48	1
HF/6-31+G(d)	4	−123.93	123.93	143.81	1
HF/pDZ	4	−128.19	128.19	148.96	1
MP2/6-31G(d)	5	−27.67	27.91	33.49	2
MP2/6-31+G(d)	5	−28.00	28.00	33.54	2
MP2/pDZ	5	−27.08	27.08	30.70	2
MP3/6-31G(d)	6	−40.32	40.32	45.15	5
HF/6-31G(2df,p)	4	−115.87	115.87	135.40	6
MP4SDQ/6-31G(d)	6	−39.82	39.82	45.27	7
MP3/6-31+G(d)	6	−41.29	41.29	46.36	10
QCISD/6-31G(d)	6	−40.27	40.27	46.02	11
MP4SDQ/pDZ	6	−40.05	40.05	46.45	12
HF/pTZ	4	−119.30	119.30	139.37	21
MP2/6-31G(2df,p)	5	−3.86	10.34	12.82	22
HF/MG3	4	−117.23	117.23	137.25	23
HF/G3large	4	−116.65	116.65	136.58	29
QCISD/6-31G(2df,p)	6	−19.19	19.19	22.53	42
MP2/MG3	5	−1.85	8.48	10.51	57
MP2/pTZ	5	−3.17	6.97	8.79	74
MP3/6-31G(2df,p)	6	−17.68	17.68	20.21	97
MP2(full)/G3large	5	0.05	8.93	11.03	111
MP4SDQ/6-31G(2df,p)	6	−18.51	18.51	21.49	154
MP4/6-31G(d)	7	−33.47	33.47	38.32	329
QCISD(T)/6-31G(d)	7	−35.55	35.55	40.66	670
MP4/6-31+G(d)	7	−33.85	33.85	38.84	785
CCSD/pDZ	6	−41.53	41.53	48.46	840
MP4/6-31G(2df,p)	7	−9.88	10.34	12.26	4453

[a] Note that $N = 0$ for all rows of this table.

optimizations, classical trajectory calculations, and reaction path calculations as well as some methods for fitting potential energy surfaces. Additionally, furan was chosen as a reasonably representative medium-sized molecule. To give an idea of the scaling for larger systems, we note that all methods scale as the size of the system to the power *n*, and Table 2 gives the values of *n*.

Finally, Table 2 also gives the number *N* of empirical parameters for each method.

Table 3 has the same columns of information as Table 2, but it is for single-level methods. Any level that forms a component of any of the multilevel methods in Table 2 is included in Table 3.

TABLE 4: Exponent of Tightest *d* Functions for Second-Row Atoms

basis set	Si	P	S	Cl
6-31G(d)	0.45	0.55	0.65	0.75
6-31+G(d)	0.45	0.55	0.65	0.75
6-31G(2df,p)	0.9	1.1	1.3	1.5
MG3	1.8	2.2	2.6	3.0
G3large	8.0	10.0	11.0	13.0
pDZ[a]	0.275	0.373	0.479	0.6
pTZ[b]	0.481	0.652	0.819	1.046

[a] cc-pVDZ. [b] cc-pVTZ.

4. Discussion

In choosing the basis sets to be used here, we were particularly concerned that the basis sets for second row atoms would include tight *d* functions, since this seems to be a prerequisite for consistent accuracy on that row.[44] The MG3 basis satisfies this criterion. Consider, for example, the subrow of the periodic table from Si to Cl. In all basis sets used in this paper, the tightest *d* function increases monotonically in proceding from Si to Cl. Table 4 gives the exponent of the tightest *d* function for Si, P, S, and Cl for all of these basis sets. The tightest *d* exponent in the 6-31G(2df,p) basis for Si is a factor of 1.4 to 1.9 tighter than that in the pTZ basis. The MG3 *d* functions are larger yet, by another factor of 2. We believe that these *d* functions are key elements in a consistent treatment of the two rows.

Before discussing the numerical results, we should say a few words about the philosophy of our approach in order to avoid misconceptions. We recognize that correlated electronic structure calculations have many uses. In some cases they are used with geometries obtained by less expensive methods to predict accurate electronic energies; in other cases they are combined with vibrational−rotational energies or vibrational−rotational energies and entropies to estimate thermochemical quantities such as enthalpies and free energies, and in yet other cases they are used to optimize geometries of stable species or saddle points or to scan potential energy surfaces. We want to design methods that can be used for all of these purposes and that can be used with a variety of lower-level geometries when the application involves single-point energy calculations at lower-level or experimental geometries. Therefore, we focus on electronic energies, and we made a database of D_e values (i.e., zero-point-exclusive dissociation energies) for testing our electronic energies. When we compare our results and timings to other methods, we test those methods for electronic energies. The reader should keep in mind that slightly different comparisons would ensue if we followed other protocols. For example, Gaussian-*x* methods are designed for use with specific geometries and specific well-defined procedures for calculating vibrational−rotational energies. It is beyond our scope here to comment on the accuracy of those geometries or vibrational−rotational energies, but we believe that our conclusions would not be significantly different if we tested our methods on enthalpies of formation instead of equilibrium dissociation energies.

First compare the MC-QCISD results in Table 2 to those of all the single-level methods in Table 3. The MC-QCISD method yields a mean unsigned error (MUE) of 2.5 kcal/mol for a cost of 70 reduced cost units. None of the single-component methods in Table 4 has an error this small or even within a factor of 2 or 2.5 kcal/mol. (The best MUE in Table 3 is 6.8 kcal/mol.) The mean signed error and RMS error of the MC-QCISD method are also excellent. Although it is no longer surprising in light of our previous work, it is still worth noting the high overall accuracy of all multilevel methods in Table 2 compared to the single-level ones in Table 3. For example, Table 2 shows that one can attain an accuracy (MUE) of 10.7 kcal/mol at a cost of 4 and an accuracy of 7.1 kcal/mol at a cost of 13, whereas comparable accuracies with single-level methods require costs of 57 and 74 reduced cost units, respectively. For costs of 42 and 70 reduced cost units, multilevel methods can reduce the error to 3.6 and 2.5 kcal/mol, respectively.

It is particularly interesting to compare MC-QCISD to MP2/MG3 and QCISD/6-31G(d), which are its two most expensive components. The mean unsigned error is reduced by a factor of 3.4 with respect to the former and a factor of 16.1 with respect to the latter. (Note that, as stated above, all errors and timings in Table 2 are based solely on electronics energies.) It is also interesting to compare MC-QCISD to G3(MP2). The MC-QCISD and G3(MP2) methods are similar in that they use the same basis sets, and the larger basis set (MG3, also called G3MP2large) is used only for MP2 calculations. The main differences are the neglect of triple excitations in the QCISD calculation of MC-QCISD and the use of four empirical coefficients to scale energy components in MC-QCISD. In contrast, G3(MP2) includes a QCISD(T)/6-31G(d) calculation and has four empirical parameters in a “higher level correction,” which, unlike the present method, leads to discontinuous potential energy surfaces along dissociation coordinates leading to atoms or atomic ions. The G3(MP2) method is parameterized against a broad set of 299 data for neutral and charged species, whereas MC-QCISD is parameterized against 82 atomization energies of neutral molecules. The importance of triple excitations for obtaining accurate energies of some molecules by methods that do not scale the energy components is well known, but the present method attempts to make up for neglecting triples by the multi-coefficient scaling approach. The advantage of not explicitly including triples is that MC-QCISD scales (in computer time) as N^6 (where N is the number of atoms), whereas G3(MP2) scales as N^7. (We also draw the reader’s attention to our own MCG3[12,32] method for consideration in choosing among methods that scale as N^7.) In addition to its improved computer-time scaling properties, MC-QCISD is much more affordable for geometry optimization[35] than is G3(MP2). We note that G3(MP2) uses MP2(full)/6-31G(d) geometries as an intrinsic aspect of the method, but it would require expensive QCISD T gradients if one wished to use the G3(MP2) energy expression to optimize geometries. Our approach is to focus on electronic energies rather than thermodynamic functions of equilibrium states. Thus, although MC-QCISD is parameterized and tested here with MP2(full)6-31G(d) geometries, we explicitly anticipate the possibility that geometries could be optimized at the MC-QCISD level or that the method could be used for scans of potential energy surfaces. With all of these considerations in mind, we turn our attention to the G3(MP2) results in Table 2 (note that, as stated above), all timings and errors in Table 2 are based solely on the electronic energies). Table 2 shows that G3(MP2) has a slightly better mean unsigned error than MC-QCISD, but it is about a factor of 10 more expensive.

Next we compare the MC-QCISD method to the other multilevel methods in Table 2. As mentioned above, six of the multilevel methods have already been pre-selected from over forty multilevel combinations as the ones with the best performance/price ratios, so they represent the state-of-the-art for a range of cost categories, and indeed MC-QCISD is not particularly better in accuracy for approximately the same sizes of basis sets. The error is similar to the MCCM-UT-CCSD method, which is encouraging because the physical effects and basis set sizes of the two methods are similar, but we note that the cost of MC-QCISD is a factor of 13.4 less, and in fact it was the anticipation of such a savings that motivated this paper. (We are aware that CCSD gradients can be done more efficiently with other computer codes, but the present method should be preferred by those who do not have access to such codes, and it may even be preferred by those who do have access to efficient CCSD codes because it has one less parameter and the ratio of costs is such that it should be more efficient even with analytic CCSD gradients.)

Only the two methods whose computer time scales as the seventh power of the size of the system are unequivocally better than MC-QCISD. These methods are already 16 and 91 times more expensive than MC-QCISD for furan gradients, and with their less favorable scaling these relative timing factors will be even larger for bigger systems.

The 82-molecule training set has 3.48 bonds per molecule (on average) so that the mean unsigned error per bond in MC-QCISD is only 0.72 kcal/mol. This is quite acceptable for many applications. The error in transition state barrier heights is expected to be larger, though, because transition states are notoriously difficult.

Next consider the other new methods developed here. Many of the trends are transparently clear in the tables and require little comment. For example, the additional coefficients in the MCSAC-QCISD/6-31G(d) scheme, as compared to the SAC-QCISD/6-31G(d) scheme, lower the mean unsigned error by 23%, a factor of 1.9. Adding an HF and MP2 calculation with the 6-31G(2df,p) basis set, which yields the MCCM-UT-QCISD;6-31G(2df,p);6-31G(d) method, then lowers the mean unsigned error a further factor of 2.5 to 3.64 kcal/mol.

Considering all of the performance vs cost tradeoffs in Table 2, we single out the MCCM-UT-QCISD;6-31G(2df,p);6-31G(d) and MC-QCISD methods as particularly highly recommended. We also add MCSAC-QCISD/6-31G(d) to the “honor roll” of particularly highly recommended methods as a less expensive alternative based on QCISD; it may be valuable for large-molecule cases where QCISD is desired because of spin contamination, multireference character, or balance, but where the more expensive QCISD methods are unaffordable.

MCCM-CO-MP2;MG3;6-31+G(d) has also been placed on the honor roll (see last column of Table 2) because it uses only basis sets with diffuse functions and is relatively inexpensive; the diffuse character can be important when making and breaking bonds involving polar radicals. The methods not placed on the highly recommended list should also be useful, but one of our goals has been to identify the most highly successful or most highly recommended multilevel combinations, and the twelve methods marked "yes" in the last column of Table 2 are the recommended set selected from a total of over 50 combinations examined in previous work and here.

5. Concluding Remarks

A multi-coefficient correlation method based on QCISD has been parameterized against 82 zero-point-exclusive atomization energies. The mean unsigned error per bond is only 0.72 kcal/mol, yet the cost is quite reasonable. The MC-QCISD method has been parameterized entirely on the basis of bond energies, but we believe it will be especially useful for calculating transition state geometries. There are two reasons for not using transition state geometries and energies in the parameterization: (1) because of Hammond's postulate,[36] the location and energy of transition states is very sensitive to errors in the overall exoergicity, a source of error that does not affect reactant and product geometries; (2) accurate transition state geometries or energies are available only for a very few simple reactions. Nevertheless, based on previous favorable experience with QCISD,[30a,37,38] we anticipate that the MC-QCISD method will be especially useful for optimizing the geometries of transition states, but it should also be useful for a wide variety of thermochemical calculations when the MCG3, G3, and G3S methods are too expensive. The present paper has also identified two less expensive methods based on QCISD that are particularly highly recommended for cases where the MC-QCISD method is unaffordable, and we have parameterized a method that uses diffuse functions for all components because that may be more accurate for polar radicals.

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

Supporting Information Available: Table of calculated atomization energies (kcal/mol) for the 82-molecule data set by G3(MP2) and the twelve most highly recommended methods. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Brown, F. B.; Truhlar, D. G. *Chem. Phys. Lett*. **1985**, *117*, 307.

(2) (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.

(3) Corchado, J. C.; Truhlar, D. G. In *Combined Quantum Mechanical and Molecular Mechanical Methods*; Gao, J., Thompson, M. A., Eds.; ACS Symposium Series 712; American Chemical Society: Washington, DC, 1998; p 106.

(4) Gordon, M. S.; Truhlar, D. G. *J. Am. Chem. Soc*. **1986**, *108*, 5412.

(5) Gordon, M. S.; Truhlar D. G. *Int. J. Quantum Chem*. **1987**, *31*, 81.

(6) Gordon, M. S.; Nguyen, K. A.; Truhlar, D. G. *J. Phys. Chem*. **1989**, *93*, 7356.

(7) Rossi, I.; Truhlar, D. G. *Chem. Phys. Lett*. **1995**, *234*, 64.

(8) (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.

(9) Blomberg, M. R. A.; Siegbahn, P. E. M. In *Computational Thermochemistry*; Irikura, K. K., Frurip, D. J., Eds.; ACS Symposium Series 677; Americal Chemical Society: Washington, DC, 1998; p 197.

(10) Fast, P. L.; Corchado, J. C.; Sánchez, M. L.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 3139.

(11) Fast, P. L.; Corchado, J. C.; Sánchez, M. L.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 5129.

(12) Tratz, C. M.; Fast, P. L.; Truhlar, D. G. *PhysChemComm* **1999**, *2*, article 14.

(13) Nyden, M. R.; Petersson, G. A. *J. Chem. Phys.* **1981**, *75*, 1843.

(14) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A., Jr. *J. Chem. Phys.* **1996**, *104*, 2598.

(15) 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.

(16) Ochterski, J. W.; Petersson, G. A.; Wiberg, K. B. *J. Am. Chem. Soc.* **1995**, *117*, 11299.

(17) Dunning, T. H., Jr., *J. Chem. Phys.* **1989**, *90*, 1007.

(18) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104.

(19) Martin, J. M. L.; Taylor, P. R. *Chem. Phys. Lett.* **1996**, *248*, 336.

(20) Martin, J. M. L. *Chem. Phys. Lett.* **1996**, *259*, 669.

(21) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.

(22) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.

(23) Martin, J. M. L. In *Computational Thermochemistry*; Irikura, K. K., Frurip, D. J., Eds.; ACS Symposium Series 677; American Chemical Society: Washington, DC, 1998, p 212.

(24) Petersson, G. A. In *Computational Thermochemistry*; Irikura, K. K., Frurip, D. J., Eds.; ACS Symposium Series 677; American Chemical Society: Washington, DC, 1998, p 237.

(25) Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *294*, 45.

(26) Fast, P. L.; Sánchez, M. L.; Truhlar, D. G. *J. Chem. Phys.* **1999**, *111*, 2921.

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

(28) Curtiss, L. A.; Raghavachari, K. In *Computational Thermochemistry*; Irikura, K. K., Frurip, D. J., Eds.; ACS Symposium Series 677; American Chemical Society: Washington, DC, 1998, p 176.

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

(30) See, e.g., (a) Durant, J. L.; Rohlfing, C. M. *J. Chem. Phys.* **1993**, *98*, 8031. (b) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1999**, *110*, 4703. (c) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 42. (d) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1999**, *110*, 4703.

(31) Fast, P. L.; Sánchez, M. L.; Corchado, J. C.; Truhlar, D. G. *J. Chem. Phys.* **1999**, *110*, 11679.

(32) Fast, P. L.; Sánchez, M. L.; Truhlar, D. G. *Chem. Phys. Lett.* **1999**, *306*, 407.

(33) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 1125.

(34) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.

(35) (a) Chuang, Y.-Y.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 651. (b) Rodgers, J. M.; Fast, P. L.; Truhlar, D. G. *J. Chem. Phys.* **2000**, *112*, 3141.

(36) Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334.

(37) Durant, J. L. In *Computational Thermochemistry*; Irikura, K. K., Frurip, D. J., Eds.; ACS Symposium Series 677; American Chemical Society: Washington, DC, 1998, p 267.

(38) Chuang, Y.-Y.; Coitiño, E. L.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 446.

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

(40) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle E. S.; Pople, J. A. *Gaussian* 98 (Revision A.7), Gaussian, Inc.: Pittsburgh, PA, 1998.

(41) (a) Pople, J. A.; Hehre, W. J. *J. Comput. Phys.* **1978**, *27*, 161. (b) Schlegel, H. B. *J. Chem. Phys.* **1982**, *77*, 3676.

(42) The MP2(full)/6-31G(d) geometries are available at the website http://chemistry.anl.gov/compmat/G2-97.htm.

(43) (a) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007. (b) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358.

(44) (a) Martin, J. M. L.; de Oliveira, G. *J. Chem. Phys.* **1999**, *111*, 1843. (b) Bauschlicher, C. W.; Ricca, A. *J. Phys. Chem. A* **1998**, *102*, 8044. (c) Martin, J. M. L. *J. Chem. Phys.* **1998**, *108*, 2791. (d) Bauschlicher, C. W., Jr.; Partridge, H. *Chem. Phys. Lett.* **1995**, *240*, 533.