

## Performance of the correlation consistent composite approach for transition states: A comparison to G3B theory

Thomas V. Grimes, Angela K. Wilson,<sup>a)</sup> Nathan J. DeYonker, and Thomas R. Cundari<sup>b)</sup>

*Department of Chemistry and Center for Advanced Scientific Computing and Modeling (CASCaM), University of North Texas, P.O. Box 305070, Denton, Texas 76203-5070, USA*

(Received 16 April 2007; accepted 4 September 2007; published online 19 October 2007)

The correlation consistent composite approach (ccCA) was applied to the prediction of reaction barrier heights (i.e., transition state energy relative to reactants and products) for a standard benchmark set of reactions comprised of both hydrogen transfer reactions and nonhydrogen transfer reactions (i.e., heavy-atom transfer,  $S_N2$ , and unimolecular reactions). The ccCA method was compared against G3B for the same set of reactions. Error metrics indicate that ccCA achieves “chemical accuracy” with a mean unsigned error (MUE) of 0.89 kcal/mol with respect to the benchmark data for barrier heights; G3B has a mean unsigned error of 1.94 kcal/mol. Further, the greater accuracy of ccCA for predicted reaction barriers is compared to other benchmarked literature methods, including density functional (B3LYP, MUE=1.16 kcal/mol) and wavefunction-based [QCISD(T), MUE=1.10 kcal/mol] methods. © 2007 American Institute of Physics. [DOI: 10.1063/1.2790011]

### I. INTRODUCTION

Using composite model chemistries to accurately calculate energies at a reduced computational cost, the theoretical modeling of chemical species has made great progress. Composite methods have even progressed to the point of predicting thermodynamic properties of ground state systems containing up to approximately a dozen heavy atoms to within “chemical accuracy,” usually considered to be  $\pm 1$  kcal/mol from experimental values. Some composite methods even claim to be within 1 kJ/mol of experiment.<sup>1,2</sup> However, the prediction of kinetic rate constants, which implicitly require accurate energetics for both ground and transition states, has not, in general, enjoyed the same degree of success. The correlation consistent composite approach (ccCA) is a composite model chemistry that aims to achieve chemical accuracy without empirical parameters and at a reasonable computational cost in comparison to large basis set coupled cluster calculations.<sup>3</sup> The ccCA method has been benchmarked against a large set of atomization energies, heats of formation, etc., for stable molecules, atoms, radicals, and even *s*- and *d*-block metal systems.<sup>4</sup> This paper extends the benchmarking of ccCA to transition states and represents a first step toward accurate theoretical predictions of reaction kinetics using composite methods.

There are two basic problems that hinder any theoretical attempt at reproducing experimental kinetic data. The first is the exponential dependence of the chemical rate constant ( $k$ ) on the activation energy ( $E_a$ ) that is indicated by the Arrhenius equation,  $k=A \exp[-E_a/RT]$ . Achieving chemical accuracy ( $\pm 1$  kcal/mol) for the activation barrier can still lead to large errors in the predicted rate constant. The second prob-

lem is the nature of the  $E_a$  term itself. As given, it is a purely phenomenological quantity: Arrhenius plots are never completely linear for all temperatures. Furthermore, the direct experimental measurement of transition state energies is not possible. Association of the observed activation barrier with the adiabatic transition state energy is thus strictly incorrect, as can be easily rationalized by the fact that the adiabatic transition state geometry and energy do not change with temperature. This association has led to substantial error in comparing predicted and observed rate constants (e.g., see Ref. 5). Contributions from tunneling, barrier recrossing, diabatic effects (i.e., coupling of different electronic states), the presence of multiple reaction channels, and potential van der Waals complexes in the reaction channels are all neglected within the simplest transition state theory. Thus, the observed rate constant depends on the entire potential energy surface (PES) rather than just the reactant and product asymptotes and a first-order adiabatic saddle point.

As is the case in thermochemistry modeling, in choosing a transition state theory (TST) to model a particular reaction there is a trade-off between accuracy and available resources. In theory, direct integration of the quantum mechanical scattering equations or the application of quantum statistical mechanics could provide quantitatively accurate kinetic data. However, the prohibitive cost of such a procedure currently limits its use to very small chemical systems.<sup>6</sup> If a computationally expensive level of theory is employed, fewer points on that PES can be feasibly computed within a reasonable amount of time and, therefore, a lower level of TST must often be used. An example of such an approach is the use of expensive levels of theory to calculate the product/reactant asymptotes and a transition state (i.e., the application of conventional transition state theory).<sup>7</sup> Conversely, a cheaper level of theory will allow more of the PES to be explored,

<sup>a)</sup>Electronic mail: akwilson@unt.edu

<sup>b)</sup>Electronic mail: tomc@unt.edu

but at the cost of uncertainty in those points. An alternative approach that is sometimes employed<sup>8</sup> is to use a high level of theory to calculate the product/reactant asymptotes and the transition state, but use a lower level of theory for a limited number of points on the PES (usually points along the minimum energy path, MEP). While often a good compromise, the latter approach suffers, not only from reduced accuracy along the MEP, but also from a potential mismatch between PESs described at two different levels of theory. It is highly desirable, therefore, to construct a method that balances speed and accuracy uniformly across the PES. Given the success of composite model chemistries in predicting thermochemistry, it is natural to extend such approaches to kinetic modeling.

Rather than attempt to benchmark ccCA against experimental observations directly, two databases of “best estimate” transition state energies were used (HTBH38/04 and NHTBH38/04).<sup>9</sup> The barrier heights contained therein are either (a) adjusted against experimental rates combined with dynamical simulations<sup>10</sup> in order to extract a better estimate of the transition state energy (the classical barrier height), or (b) from very accurate *ab initio* calculations. By benchmarking ccCA against these values, it is possible to obtain a more balanced comparison of how ccCA performs for transition states in relation to the best available data. It should be stressed that the benchmarking described in this paper does not directly yield an estimate of how well ccCA predicts rate constants, but rather how well ccCA describes the adiabatic PES for the most important stationary points (reactant and product asymptotes and the transition state) relative to levels of theory currently employed in the literature for kinetic modeling.

Among composite model chemistries, G3 and G3B were found to have an accuracy (mean unsigned error) of 0.99 and 1.01 kcal/mol, respectively, in the literature with respect to the G2/97 test set, for ground state energies.<sup>11</sup> The application of Gaussian-*n* methods to kinetics has also been reported in the literature.<sup>5,12,13</sup> For comparison, results computed using G3B are given along with the ccCA results in this paper. The G3B variant of G3 theory was used in this research to provide the most even standard of comparison, since both ccCA and G3B model chemistries use B3LYP optimizations with comparable basis sets to provide the reference geometry. Furthermore, it is typically asserted that hybrid DFT methods provide good transition state geometries.<sup>14</sup> The combination of a fast, yet accurate geometry optimization with a carefully constructed series of additive basis set and correlation energy corrections allows ccCA to provide more accurate thermochemical data than G3B at a computational expense much cheaper than *Wn* model chemistries for ground state properties.<sup>3</sup> We seek to investigate the potential of ccCA for modeling kinetics in the current paper.

## II. COMPUTATIONAL METHODS

Composite model chemistries attempt to produce very accurate predictions with reasonable computational costs by computing a series of additive basis set and electron correlation corrections to a reference energy at a geometry provided

by a lower level of theory. While G3 theory specifies the use of (all-electron) MP2/6-31G(d) geometries,<sup>15</sup> the G3B modification uses B3LYP/6-31G(d) geometries.<sup>11</sup> Since the most recent and successful formulation of ccCA uses B3LYP/6-31G(2df,p) geometries,<sup>3(b)</sup> the current study compares ccCA to G3B for prediction of barrier heights. Two major differences between G3B and ccCA are discussed below. Detailed analyses of G3B and ccCA for the prediction of ground state properties can be found in Refs. 11 and 3, respectively.

While the Gaussian-*n* methods employ Pople-style basis sets for the computation of energies, ccCA attempts to systematically reduce basis set error using the correlation consistent basis sets originally developed by Dunning.<sup>16</sup> In the recommended formulation of ccCA, the reference energy is the result of extrapolating MP2 energies to the complete basis set limit.<sup>3</sup> The two variants of ccCA, ccCA-S4 and ccCA-P, differ only in the fit function used in the reference energy extrapolation. The former fits MP2/aug-cc-pVxZ (*x*=T,Q) energies to the Schwartz fit function<sup>17</sup> given in Eq. (1), while the latter is a three-point fit for MP2/aug-cc-pVxZ (*x*=D,T,Q) energies to the Peterson equation<sup>18</sup> [Eq. (2)],

$$E = E_0 + B \left( l_{\max} + \frac{1}{2} \right)^{-4}, \quad (1)$$

$$E = E_0 + B \exp[-(x-1)] + C \exp[-(x-1)^2], \quad (2)$$

where  $l_{\max}$  is the highest angular momentum function in the basis set,  $x$  is the cardinal number,  $E$  is the fit energy at the aug-cc-pVxZ level,  $E_0$  is the extrapolated energy (i.e.,  $E$  for  $x, l_{\max} \rightarrow \infty$ ), and  $B$  and  $C$  are fit parameters.

The other significant difference between G3B and ccCA is the use of empirical corrections. G3 theory employs a higher-level correction (HLC) that originally corresponded to a physical quantity, the isogyric correction, in G1 theory.<sup>19</sup> Within subsequent *Gn* variants, the HLC was generalized to account for deficiencies of the model chemistry in an average way by being parametrized to minimize the magnitude of the mean signed deviation of the overall training set. The only “parameter” in ccCA is the harmonic vibrational scale factor, which is a nonoptimized parameter used to account for well-known deficiencies in the harmonic approximation of vibrational frequencies, and not for deficiencies in the model chemistry as a whole. No zero-point energy corrections were used in the current work, so the vibrational scale factor is irrelevant for this specific application. All calculations were performed within the GAUSSIAN 03 suite of programs.<sup>20</sup>

## III. DATABASES

To provide a standard of comparison, the HTBH38/04 and NHTBH38/04 reaction barrier databases (henceforth HTBH and NHTBH, respectively) compiled by Zhao *et al.* were used.<sup>21</sup> The HTBH database contains 19 hydrogen transfer reactions with values for the forward and reverse classical barrier heights (i.e.,  $\Delta E^\ddagger$  exclusive of zero-point corrections). All computed geometries and energies are given in the supporting information.<sup>22</sup> The NHTBH database contains 19 reactions, including heavy-atom transfer, bimolecular nucleophilic substitution ( $S_N2$ ), association, and unimo-

molecular reactions.<sup>21</sup> Most of the barrier heights in NHTBH are derived from Weizmann-1 (W1) theory,<sup>23</sup> though some are derived from its successor [W2h (Ref. 23)] or literature values from other levels of theory expected to be of at least chemical accuracy.

While it may seem undesirable to benchmark one composite method against another, W1 is designed to obtain the complete basis set limit of CCSD(T) calculations and is expected to consistently give better than 1 kcal/mol accuracy for barrier heights.<sup>21,24</sup> Although W1 contains a single fit parameter (the exponent for the valence correlation extrapolation), this parameter is derived from W2 theory rather than the experiment. Confidence in the theoretically derived benchmark values is largely based on the known performance of W1 and W2 theories for stable molecules, producing errors on the order of only  $\pm 1$  kJ/mol.<sup>21,24</sup> A larger error is observed with  $W_n$  theory for systems showing a high degree of nondynamical correlation effects, due to the inherent limits of CCSD(T).<sup>25</sup> The inclusion of higher-level excitations (connected quadruple, pentuple, etc.) in W3 and W4 theories was found to yield a better treatment of nondynamical correlation effects.<sup>1(b)</sup>

For the purposes of benchmarking, the “best estimate” values provided in the HTBH and NHTBH databases are assumed to be correct. All geometries used in this research were fully optimized and characterized by energy Hessian calculations at the level used for the geometry optimization [B3LYP/6-31G(d) for G3B and B3LYP/6-31G(2df,p) for ccCA] unless otherwise indicated.

## IV. RESULTS AND DISCUSSION

### A. Problematic structures

While DFT often provides good geometries,<sup>26,27</sup> even for transition states,<sup>14</sup> like any theory it is not infallible. In particular, there are three systems for which the default B3LYP-based geometry optimization failed.

#### 1. The $\text{H} + \text{F}_2 \rightarrow \text{HF} + \text{F}$ transition state

This reaction is apparently barrierless at the B3LYP/6-31G(2df,p) level of theory, as suggested by a scan of the PES. As a consequence, the QCISD/MG3 (Ref. 28) optimized transition state geometry provided by Lynch *et al.*<sup>29</sup> was used without modification for both the ccCA and G3B calculations. At the given geometry, both B3LYP/6-31G(2df,p) and B3LYP/6-31G(d) methods have one imaginary frequency corresponding to the correct transition mode.

#### 2. The $\text{CH}_3 + \text{FCl} \rightarrow \text{CH}_3\text{F} + \text{Cl}$ transition state

No transition state could be located using B3LYP/6-31G(2df,p) or B3LYP/6-31G(d) methods. The reaction is assumed to be barrierless at these levels of theory, so the transition state geometry provided by Lynch *et al.*<sup>29</sup> was used. At the given geometry, B3LYP/6-31G(2df,p) calculations predict one imaginary frequency corresponding to the correct transition mode. Using the B3LYP/6-31G(d) level of theory,

TABLE I. Hammond postulate analysis of problematic species. [ $V_f$  and  $V_r$  are the benchmark values of the forward and reverse barrier heights, respectively, in kcal/mol. The completion estimate is defined as  $V_f/(V_f+V_r)$ .]

Reaction	$V_f$	$V_r$	Completion at transition state
$\text{H} + \text{F}_2 \rightarrow \text{HF} + \text{F}$	2.27	106.18	2%
$\text{CH}_3 + \text{FCl} \rightarrow \text{CH}_3\text{F} + \text{Cl}$	7.43	60.17	11%
$\text{F} \cdots \text{CH}_3\text{Cl} \rightarrow \text{FCH}_3 \cdots \text{Cl}$	2.89	29.62	9%

a reasonable transition mode is found ( $459i$   $\text{cm}^{-1}$ ), plus two small imaginary modes corresponding to C–F–Cl linear bending modes ( $37i$  and  $36i$   $\text{cm}^{-1}$ ).

### 3. The $\text{F}^- \cdots \text{CH}_3\text{Cl}$ ion pair

The stationary point corresponding to this separated ion pair (a stable minimum, not a transition state) could not be located using B3LYP/6-31G(2df,p), B3LYP/6-31G(d), or even B3LYP/cc-pVTZ. Upon attempted geometry optimization using these basis sets, the product ion pair  $\text{CH}_3\text{F} \cdots \text{Cl}^-$  resulting from moving through the transition state is always obtained.  $S_N2$  reactions typically have a double-well structure<sup>13</sup> corresponding to a bound ion pair in the entrance and exit channels of the reaction. Martin and Parthiban found it impossible to obtain this bound ion pair using B3LYP/6-311G(2d,d,p),<sup>30</sup> but were able to find it once diffuse functions were added to the basis set.<sup>13</sup> To investigate basis set and correlation effects, optimizations were attempted at the CCSD/cc-pVTZ, B3LYP/cc-pVQZ, and B3LYP/MG3 levels of theory to locate this elusive ion pair. We were able to locate the bound ion pair at the CCSD/cc-pVTZ, B3LYP/MG3, and B3LYP/aug-cc-pVTZ levels, but not using the B3LYP/cc-pVQZ method. To retain consistency with the above-mentioned troublesome stationary points, the QCISD/MG3 geometry was used.<sup>29</sup> Both B3LYP/6-31G(2df,p) and B3LYP/6-31G(d) predict two small imaginary modes (nearly degenerate  $157i$  and  $145i$   $\text{cm}^{-1}$ , respectively) at the QCISD/MG3 geometry, roughly corresponding to linear F–C–Cl bending.

Besides being in the NHTBH database, all three problematic structures have one feature in common. The transition state geometry for each “problem” reaction is predicted to be close to the reactants, via the Hammond postulate.<sup>31</sup> That is, the forward barrier is small, leading to a transition state that strongly resembles the reactants both in energy and in structure. The linear interpolation proposed independently by Agmon<sup>32</sup> and Miller<sup>33</sup> for quantitative application of the Hammond postulate can be used to estimate the degree of completion of a reaction at the transition state using the known forward and reverse reaction barrier heights. While quantitative formulations of the Hammond postulate have been compared to results using Marcus theory,<sup>34</sup> the use of these parameters as an estimation of how well DFT will perform appears to be unique to the current study. The completion estimates for the three troublesome reactions above are given in Table I. All of the NHTBH reactions have completion estimates ranging from 2% to 59%, and it is noticeable that the troublesome reactions mentioned above have completion estimates less than 12%, i.e., have very “early”

transition states. There is one other reaction ( $\text{H}+\text{C}_2\text{H}_4 \rightarrow \text{CH}_3\text{CH}_2$ ) in the NHTBH set and one reaction in the HTBH set ( $\text{F}+\text{H}_2 \rightarrow \text{HF}+\text{H}$ ) with completion estimates under 12% (4% and 5%, respectively) but B3LYP geometries were found for the corresponding transition states for these reactions. It is notable that all the transition states for proton-transfer reactions were found, despite the small completion estimates for two of them. This leads one to believe that the completion estimates of Agmon<sup>32</sup> and Miller<sup>33</sup> may be a good indicator of when a transition state for nonproton-transfer reactions will be difficult (if not impossible) to locate using DFT and perhaps whether modification of the geometry optimizations will be necessary.

## B. “Black box” method performance

Analysis of the reaction set is slightly complicated by the fact that the forward and reverse reactions are not absolutely independent because both the forward and reverse barriers depend on the transition state structure and energy. The reverse and forward barriers are inextricably linked through the equilibrium constant, which is a thermodynamic quantity dependent on ground state energies that is thus expected to be accurately predicted by composite methods. For nonsymmetric reactions, the dependency is irrelevant from a statis-

TABLE II. Error metrics (in kcal/mol) for ccCA and G3B.

	ccCA-S4	ccCA-P	G3B
HTBH—Forward reaction barriers			
MSE	-0.15	-0.15	1.03
MUE	0.79	0.79	1.59
Max error	2.12	2.14	5.04
HTBH—Reverse reaction barriers			
MSE	-0.11	-0.11	1.44
MUE	0.99	1.01	1.75
Max error	2.88	2.95	4.92
NHTBH—Forward reaction barriers			
MSE	-0.29	-0.27	0.62
MUE	1.01	1.00	1.82
Max error	3.19	3.26	6.70
NHTBH—Reverse reaction barriers			
MSE	0.06	0.06	2.31
MUE	0.76	0.77	2.57
Max error	3.67	3.75	6.68
HTBH aggregate (36 independent data points)			
MSE	-0.10	-0.10	1.25
MUE	0.91	0.92	1.71
Max error	2.88	2.95	5.04
NHTBH aggregate (32 independent data points)			
MSE	-0.19	-0.18	1.41
MUE	0.98	0.99	2.28
Max error	3.67	3.75	6.70
HTBH+NHTBH aggregate			
MSE	-0.14	-0.14	1.32
MUE	0.94	0.95	1.98
Max error	3.67	3.75	6.70

TABLE III. “Double-counted” statistics for direct comparison with other benchmarks. (All errors are defined as expected minus calculated, e.g., a positive error indicates that the calculated value is lower than the true value. MSE=mean signed error; MUE=mean unsigned error.)

	ccCA-S4	ccCA-P	G3B
HTBH			
MSE	-0.13	-0.13	1.23
MUE	0.89	0.90	1.67
NHTBH			
MSE	-0.11	-0.11	1.47
MUE	0.89	0.89	2.20
HTBH+NHTBH			
MSE	-0.12	-0.12	1.35
MUE	0.89	0.89	1.94

tical standpoint because two degrees of freedom are produced (forward and reverse barriers) from two independent variables (transition state energy with respect to reactants and products). However, the products and reactants are identical in symmetric reactions and the corresponding energies are constrained to be equal. Therefore, the forward barrier is not independent of the reverse barrier (they are, of course, equal). Zhao *et al.*'s reported statistics double count symmetric reactions (two in the HTBH set and five in the NHTBH set),<sup>9,21</sup> so both unique and “double-counted” statistical approaches were investigated (Tables II and III, respectively). There are only minor differences in the reported error arising from double-counting symmetric reactions, as is expected of randomly distributed error. Detailed comparisons of the calculated values versus best estimates are given in the supporting information<sup>22</sup> and are summarized in Table II. Overall results are reported for ccCA and G3B.

### 1. ccCA versus G3B

In a global sense, the mean unsigned error for the entire data set (aggregate HTBH+NHTBH, Table II and Fig. 1) indicates that ccCA [mean signed error (MSE) of -0.14 kcal/mol and mean unsigned error (MUE) of 0.95 kcal/mol] achieves chemical accuracy for the predic-

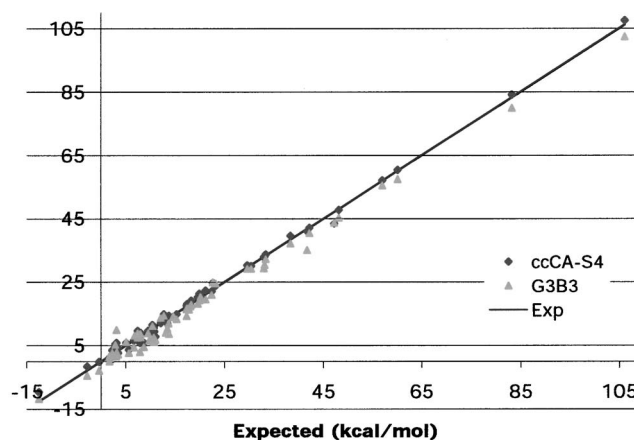


FIG. 1. Graph of computed versus best estimate reaction barriers for ccCA and G3B.

TABLE IV. Problematic reactions (reactions 1 and 2 for ccCA; reactions 3 and 4 for G3B). (All energies are in kcal/mol.)

Reaction	Lit $V^\ddagger$	ccCA-S4	ccCA-P	G3B
(1) $\text{H}_2 + \text{Cl} \rightarrow \text{H} + \text{HCl}$	8.70	5.82	5.75	4.59
(2) $\text{F}^- \cdots \text{CH}_3\text{OH} \rightarrow \text{HO}^- \cdots \text{CH}_3\text{F}$	47.20	43.53	43.45	43.67
(3) $\text{O} + \text{CH}_4 \rightarrow \text{HO} + \text{CH}_3$	13.70	14.37	14.36	8.66
(4) $\text{H} + \text{CO} \rightarrow \text{HCO}$	3.17	5.82	5.92	9.87

tion of reaction barrier heights. On the other hand, G3B performs less well with a mean unsigned error of 1.98 kcal/mol for the entire set. Furthermore, the maximum errors of the ccCA method are 3.67 and 3.75 kcal/mol for ccCA-S4 and ccCA-P, respectively, whereas the maximum error for G3B is much higher (6.70 kcal/mol). The consistently negative mean signed error for ccCA (except for the NHTBH reverse barrier subset, see Table II) means that the transition state energy is consistently overestimated, whereas the positive mean signed error for G3B indicates that the barrier is consistently underestimated.

## 2. ccCA versus density and wavefunction-based methods

Since the NHTBH and HTBH databases were used, a comparison can be made between the accuracy of ccCA and G3B versus the 29 DFT functionals and six *ab initio* wavefunction-based methods benchmarked in Ref. 21. The original benchmark calculations are equivalent to  $X/\text{MG3S}$  (Ref. 35)  $\parallel \text{QCISD}/\text{MG3}$  (where  $X$  is a DFT functional) or  $Y/\text{MG3} \parallel \text{QCISD}/\text{MG3}$  (where  $Y$  is an *ab initio* wavefunction method). Analysis of the mean unsigned errors indicates that four DFT functionals (MPW1K, BB1K, MPWB1K, and MPWKIS1K) and one *ab initio* method, QCISD(T), outperform G3B for the calculation of activation barriers. It is perhaps unsurprising that the DFT functionals specifically parametrized for kinetics outperform G3B for this data set. However, both the ccCA-P and ccCA-S4 variants outperform *all* methods benchmarked in this study, including the highest *ab initio* level reported by Zhao *et al.*:<sup>9</sup> QCISD(T). The overall MSE and MUE for QCISD(T) are 0.64 and 1.10 kcal/mol, respectively, whereas ccCA (both variants) gives  $-0.12$  and  $0.89$  kcal/mol, respectively. Although the MG3(S) basis sets are fairly large (equivalent to the 6-311+G(3d2f,2df,2p) Pople basis set), there remains the possibility that the difference can be attributed to a better treatment of basis set completeness in ccCA. Furthermore, extrapolation to the basis set limit as is done in ccCA should also ameliorate basis set superposition error (BSSE);<sup>36</sup> the BSSE was not explicitly corrected for in the benchmarked calculations.<sup>21</sup>

## C. Reliability: ccCA versus G3B

As a measure of reliability, the reactions with the largest error for ccCA and for G3B in both HTBH and NHTBH sets are given in Table IV (that is, only the reaction with the greatest error for each method and subset is selected). Reactions 1 and 2 ( $\text{H}_2 + \text{Cl} \rightarrow \text{H} + \text{HCl}$  in HTBH and  $\text{F}^- \cdots \text{CH}_3\text{OH} \rightarrow \text{HO}^- \cdots \text{CH}_3\text{F}$  in NHTBH, respectively) are

the reactions for which ccCA has the greatest error. Both ccCA variants perform similarly: the average difference between the unsigned errors for ccCA-S4 and ccCA-P for all reactions is only 0.03 kcal/mol, with a maximum deviation of 0.09 kcal/mol, therefore the worst reactions for either ccCA variant are the same and are listed only once in Table IV. Similarly, reactions 3 and 4 ( $\text{O} + \text{CH}_4 \rightarrow \text{HO} + \text{CH}_3$  in HTBH and  $\text{H} + \text{CO} \rightarrow \text{HCO}$  in NHTBH, respectively) are the reactions for which G3B shows the greatest error. For reactions 1 and 2, ccCA and G3B yield similar reaction barriers (approximately 3–4 kcal/mol lower than the benchmark value), but the ccCA predicted barriers are significantly closer to the best estimate values than G3B for reactions 3 and 4 (approximately 1 and 3 kcal/mol error for ccCA; approximately 5 and 7 kcal/mol for G3B). The fact that both methods err in the same direction and with roughly the same magnitude for reactions 1 and 2 reactions whereas ccCA provides superior barrier heights for the other two reactions potentially indicates a limitation of G3B that is not in common with ccCA, (Table IV).

The accuracy of the calculated barrier height is affected, not only by the performance of the method for the transition state, but also by the accuracy with which the reactant and product energies can be predicted. It is therefore useful to investigate the reliability of ccCA and G3B by comparing the unsigned errors (i.e., calculated versus best estimate from databases) of ccCA (Ref. 37) to the unsigned errors of G3B for each reaction rather than just the mean unsigned error across an entire set. If both methods have similar absolute errors for a particular reaction, then both are similarly reliable for that reaction (and presumably for analogous reactions). Across the HTBH and NHTBH databases, the ccCA unsigned error is more than 1 kcal/mol greater than the unsigned error of G3B for only *one* reaction ( $\text{F}^- + \text{CH}_3\text{Cl} \rightarrow \text{FCH}_3 + \text{Cl}^-$ , in the NHTBH set). Conversely, the G3B unsigned error is greater than the ccCA unsigned error by more than 1 kcal/mol for no fewer than 28 reactions in the aggregate set.

## D. Effect of modifying the methods

### 1. Effect of HLC on calculated barrier heights

Another test of reliability is to modify the methods to determine what impact a particular component has on the performance of the composite method as a whole. Rather than perform an exhaustive study of each individual correction, which is outside the scope of this research, two specific perturbations were selected as the most interesting: the impact of removing the HLC from G3B, and placing the geometries used by each method on equal footing. When the empirical HLC is removed from G3B, the mean signed and unsigned errors improve noticeably while the maximum error gets worse (Table V). In the HTBH set, none of the reactions to which the HLC contributes<sup>38</sup> are improved by inclusion of the HLC. In fact, the HLC degrades all of the reactions to which it contributes by an average of 1.13 kcal/mol for the HTBH set.

Of the 15 reactions in the NHTBH set for which the HLC contributes, seven are improved by addition of the

TABLE V. HLC-exclusive Error Metrics for G3B. (MUE=mean unsigned error; MSE=mean signed error; Max=maximum error; values listed in kcal/mol. Unique forward and reverse reactions were combined for each database, as in Table II.)

Database	MSE	MUE	Max
HTBH	0.74	1.24	4.57
NHTBH	1.08	2.21	7.93
HTBH+NHTBH	0.90	1.70	7.93

HLC (the unsigned error is reduced an average of 0.44 kcal/mol), and eight are degraded by the addition of the HLC (0.64 kcal/mol on average) resulting in an average degradation of 0.13 kcal/mol across the NHTBH set. It is apparent, therefore, that the HLC in G3B contributes to both the systematic and nonsystematic errors. Across the entire reaction set (HTBH and NHTBH combined), the HLC improves the accuracy (as determined by the unsigned error against best estimate) only 10% (7/67) of the time, while it actually degrades the accuracy 34% (23/67) of the time. Perhaps more troubling than the net loss of accuracy due to the HLC<sup>39</sup> is the unpredictability with which it will improve or degrade the computed values.

Since the HLC distinguishes between atoms and molecules, a transition state-consistent variant of *Gn* methods may reasonably be expected to have a separate set of HLC parameter values for transition states.<sup>40</sup> Indeed, Fast *et al.* encountered this very problem while formulating a variant of G3 [multicoefficient G3, MCG3 (Ref. 41)] that could be used to examine any arbitrary point on the PES rather than being confined to minima. Since MCG3 parametrizes the contribution of each explicit correction, the HLC was completely removed for consistent treatment at every point on the PES. From a physical interpretation, the fact that the atomic and molecular HLC parameters differ in the G3 method indicates that the parameters are linked to the electronic environment of the system. Since transition states almost invariably have at least one bond that is stretched beyond a normal covalent equilibrium length, the electronic environment of the stretched bond is expected to be somewhere between an equilibrium bond (molecular HLC parameters) and a broken bond (atomic HLC parameters).

## 2. Effect of geometry on calculated barrier heights

To decouple G3 and ccCA from dependence on the ability of the B3LYP functional to predict geometries, all structures were reoptimized at the CCSD/6-31G(d) level of theory. The error metrics, analogous to Table II, are reported in Table VI for the values computed at these new geometries. As with B3LYP, the structure for the  $F^- \cdots CH_3Cl$  bound ion pair could not be located, and hence the QCISD/MG3 geometry from Lynch *et al.*<sup>29</sup> was used as discussed above. Strictly speaking, the use of geometries obtained by levels of theory other than B3LYP/6-31G(d) requires a reoptimization of the HLC component of G3B. However, the standard HLC parameters in G3B were used without modification for the sake of simplicity and because the reoptimized parameters are not expected to differ greatly from the standard values.<sup>42</sup>

TABLE VI. Error metrics for ccCA and G3B, modified to use the CCSD/6-31 G(d). (The mean signed error (MSE), mean unsigned error (MUE), and maximum error (Max error) are quoted in kcal/mol.)

	ccCA-S4	ccCA-P	G3B
HTBH—Forward reaction barriers			
MSE	-0.35	-0.34	0.60
MUE	0.69	0.70	1.38
Max error	2.07	2.08	4.27
HTBH—Reverse reaction barriers			
MSE	-0.30	-0.30	0.99
MUE	0.68	0.68	1.53
Max error	2.15	2.18	3.86
HTBH—Forward reaction barriers			
MSE	-0.30	-0.29	0.47
MUE	0.88	0.89	1.58
Max error	2.71	2.80	6.66
NHTBH—Reverse reaction barriers			
MSE	0.06	0.06	2.15
MUE	1.01	1.02	2.38
Max error	3.58	3.54	4.59
HTBH aggregate (36 “independent” reactions)			
MSE	-0.30	-0.30	0.88
MUE	0.69	0.69	1.49
Max error	2.15	2.18	4.27
NHTBH aggregate (32 “independent” reactions)			
MSE	-0.11	-0.11	1.35
MUE	1.05	1.06	2.15
Max error	3.58	3.54	6.66
HTBH+NHTBH aggregate			
MSE	-0.21	-0.21	1.10
MUE	0.86	0.86	1.80
Max error	3.58	3.54	6.66

The use of CCSD/6-31G(d) geometries systematically improves the G3B values, though the mean errors still indicate that chemical accuracy is not achieved (MUE decreases from 1.98 to 1.80 kcal/mol across the entire database, see Tables III and VI). On the other hand, ccCA performs only slightly better with CCSD-based geometries than with the standard DFT-based geometries (MUE is decreased from 0.95 to 0.86 kcal/mol). The maximum error across all subsets and the mean unsigned error of the HTBH subset is decreased, Table VI, but the mean signed error becomes more negative and the mean unsigned error of the NHTBH subset is very slightly increased. The more negative mean signed error indicates that the reaction barrier is predicted consistently too high (as compared to the best estimate values), but the spread of the error and its magnitude is decreased overall. The reactions causing the maximum error for ccCA and G3B with the CCSD geometries are given in Table VII. It is interesting to note that the two reactions with the most error from ccCA (reactions 1 and 2, Table VII) are not the same as with the standard B3LYP geometries (cf. Table IV), whereas G3B errs most significantly for the same reactions as before (reactions 3 and 4, Table VII), potentially indicating a systematic error in G3B that may not be

TABLE VII. Problematic reactions (reactions 1 and 2 for ccCA; reactions 3 and 4 for G3B) using CCSD/6-31 G(d) geometries. (All energies are in kcal/mol.)

Reaction	Lit V <sup>#</sup>	ccCA-S4	ccCA-P	G3B
(1) H <sub>2</sub> O+NH <sub>2</sub> →HO+CH <sub>4</sub>	12.70	14.85	14.88	15.14
(2) FCH <sub>3</sub> +Cl <sup>-</sup> →F <sup>-</sup> +CH <sub>3</sub> Cl	20.11	16.53	16.57	16.58
(3) O+CH <sub>4</sub> →HO+CH <sub>3</sub>	13.70	14.51	14.50	9.43
(4) H+CO→HCO	3.17	5.88	5.97	9.83

present in ccCA.<sup>43</sup> Again, for reactions 1 and 2, ccCA and G3B produce similar values, but ccCA gives much better values than G3B for reactions 3 and 4, once again suggesting that ccCA is more reliable than G3B for predicting barrier heights.

## V. CONCLUSIONS

A standard benchmark suite of reactions from the literature<sup>9</sup> was employed to compare the accuracy of ccCA to G3B for the prediction of reaction barrier heights. The mean signed and unsigned errors as well as the maximum error metrics were computed to support the assertion that both variants of ccCA (ccCA-S4 and ccCA-P) achieve chemical accuracy whereas G3B does not. Since a standard database was used,<sup>9</sup> ccCA was also compared to several DFT and wavefunction-based methods and found to yield greater accuracy for reaction barriers, including the highest wavefunction-based theory tested, QCISD(T).

Unmodified, G3B was found to have a mean signed error of +1.32 kcal/mol (i.e., barriers predicted too low on average) and a mean unsigned error of 1.98 kcal/mol with respect to the test set. The largest error for G3B was 6.70 kcal/mol. A mean signed error of -0.14 kcal/mol (i.e., barriers predicted slightly too high on average) and a mean unsigned error of 0.94 kcal/mol as well as a maximum error of 3.67 kcal/mol were computed for ccCA.<sup>44</sup> An analysis of the problematic reactions for ccCA and for G3B indicates that in addition to being more accurate, ccCA is more reliable than G3B for the reaction set used.

The impact of modified reference geometries and removal of the HLC for G3B was also investigated. Even with the modifications, ccCA is still more accurate and more reliable (mean signed error of -0.22 kcal/mol, mean unsigned error of 0.86 kcal/mol) than G3B (mean signed error of 1.11 kcal/mol, mean unsigned error of 1.86 kcal/mol) for the prediction of transition state properties. Hence, the ccCA methodology represents a chemically robust composite approach for the modeling of both the thermodynamics and kinetics of chemical reactions with chemical accuracy.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the U.S. Department of Education for its support through CASCaM resources. Partial funding was obtained from the Office of Basic Energy Sciences, U.S. Department of Energy (Grant Nos. DEFG02-03ER15387) and from the NSF (Grant No. CHE-0342824). One of the authors (A.K.W.) acknowledges the NSF for support through a CAREER Award (Grant No. CHE-0239555).

- <sup>1</sup>(a) J. M. L. Martin and G. de Oliveira, *J. Chem. Phys.* **111**, 5 (1999); (b) A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay, and J. Gauss, *ibid.* **120**, 9 (2004); (c) A. Karton, E. Rabinovich, J. M. L. Martin, and B. Ruscic, *ibid.* **125**, 144108 (2006).
- <sup>2</sup>A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton, *J. Chem. Phys.* **121**, 23 (2004).
- <sup>3</sup>(a) N. J. DeYonker, T. R. Cundari, and A. K. Wilson, *J. Chem. Phys.* **124**, 114104 (2006); (b) N. J. DeYonker, T. V. Grimes, S. Yockel, A. Dinescu, B. Mintz, T. R. Cundari, and A. K. Wilson, *ibid.* **125**, 104111 (2006).
- <sup>4</sup>N. J. DeYonker, K. A. Peterson, G. Steyl, A. K. Wilson, and T. R. Cundari, "Quantitative Computational Thermochemistry of Transition Metal Complexes," *J. Phys. Chem. A* (in press); D. S. Ho, N. J. DeYonker, A. K. Wilson, and T. R. Cundari, *ibid.* **110**, 31 (2006); N. J. DeYonker, D. S. Ho, A. K. Wilson, and T. R. Cundari, "Computational *s*-block Thermochemistry with the Correlation Consistent Composite Approach and G3," *J. Phys. Chem.* (in press).
- <sup>5</sup>P. Blowers, X. Zheng, and K. Homan, *Chem. Eng. Commun.* **190**, 1233 (2003).
- <sup>6</sup>S. C. Althorpe and D. C. Clary, *Annu. Rev. Phys. Chem.* **54**, 493 (2003).
- <sup>7</sup>A. Fernández-Ramos, J. A. Miller, S. J. Klippenstein, and D. G. Truhlar, *Chem. Rev. (Washington, D.C.)* **106**, 4518 (2006).
- <sup>8</sup>T. N. Truong and T. J. Evans, *J. Phys. Chem.* **98**, 9558 (1994); J. C. Corchado, J. Espinosa-García, W.-P. Hu, I. Rossi, and D. G. Truhlar, *ibid.* **99**, 687 (1995); J. C. Corchado, J. Espinosa-García, O. Roberto-Neto, Y.-Y. Chuang, and D. G. Truhlar, *J. Phys. Chem. A* **102**, 4899 (1998); Z.-F. Xu, S.-M. Li, Y.-X. Yu, Z.-S. Li, and C.-C. Sun, *ibid.* **103**, 4910 (1999).
- <sup>9</sup>Y. Zhao, N. González-García, and D. G. Truhlar, *J. Phys. Chem. A* **109**, 2012 (2005).
- <sup>10</sup>B. J. Lynch, P. L. Fast, M. Harris, and D. G. Truhlar, *J. Phys. Chem. A* **104**, 4811 (2000).
- <sup>11</sup>A. G. Baboul, L. A. Curtiss, P. C. Redfern, and K. Raghavachari, *J. Chem. Phys.* **110**, 7650 (1999).
- <sup>12</sup>J. T. Jodkowski, M.-T. Rayez, J.-C. Rayez, T. Bérces, and S. Dóbbé, *J. Phys. Chem. A* **103**, 3750 (1999); M. R. Hand, C. F. Rodriguez, I. H. Williams, and G. C. Balint-Kurt, *ibid.* **102**, 5958 (1998); T. D. Fridgen and T. B. McMahon, *ibid.* **107**, 668 (2003); P. L. Fast, M. L. Sánchez, and D. G. Truhlar, *Chem. Phys. Lett.* **306**, 407 (1999); M. G. Bryukov, I. R. Slagle, and V. D. Knyazev, *J. Phys. Chem. A* **105**, 6900 (2001); H. Basch and S. Hoz, *ibid.* **101**, 4416 (1997); S. P. So, *Chem. Phys. Lett.* **313**, 307 (1999); M. Schwartz, P. Marshall, R. J. Berry, C. J. Ehlers, and G. A. Petersson, *J. Phys. Chem. A* **102**, 10074 (1998); J. Korchowiec, S. Kawahara, K. Matsumura, T. Uchimaru, and M. Sugie, *ibid.* **103**, 3548 (1999); B. S. Jursic, *THEOCHEM* **499**, 223 (2000); X. Yu, S.-M. Li, J.-Y. Liu, Z.-F. Xu, Z.-S. Li, and C.-C. Sun, *J. Phys. Chem. A* **103**, 6402 (1999); B. Wang, H. Hou, and Y. Gu, *ibid.* **105**, 156 (2001); Y.-X. Yu, S.-M. Li, Z.-F. Xu, Z. S. Li, and C.-C. Sun, *Chem. Phys. Lett.* **302**, 281 (1999); Z.-F. Xu, S.-M. Li, Y.-X. Yu, Z.-S. Li, and C. C. Sun, *J. Phys. Chem. A* **103**, 4910 (1999); J. Pu and D. G. Truhlar, *ibid.* **109**, 773 (2005); I. Aliagad and S. Gronert, *ibid.* **102**, 2609 (1998); A. Goumri, J.-D. R. Rocha, D. Laakso, C. E. Smith, and P. Marshall, *ibid.* **103**, 11328 (1999); *J. Chem. Phys.* **101**, 9405 (1994); A. Goumri, D. Laakso, J.-D. R. Rocha, C. E. Smith, and P. Marshall, *ibid.* **102**, 161 (1995).
- <sup>13</sup>S. Parthiban, G. Oliveira, and J. M. L. Martin, *J. Phys. Chem. A* **105**, 895 (2001).
- <sup>14</sup>B. L. Lynch and D. G. Truhlar, *J. Phys. Chem. A* **105**, 2936 (2001).
- <sup>15</sup>L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, and J. A. Pople, *J. Chem. Phys.* **109**, 7764 (1998).
- <sup>16</sup>T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989); T. H. Dunning, Jr., K. A. Peterson, and A. K. Wilson, *ibid.* **114**, 9244 (2001); A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., *ibid.* **110**, 7667 (1999); A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., *Abstr. Pap. - Am. Chem. Soc.* **213**, 60 (1997); R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992); D. E. Woon and T. H. Dunning, Jr., *ibid.* **100**, 2975 (1994); **103**, 4572 (1995); **117**, 10548 (2002).
- <sup>17</sup>J. M. L. Martin, *Chem. Phys. Lett.* **259**, 669 (1996).
- <sup>18</sup>K. A. Peterson, D. E. Woon, and T. H. Dunning, Jr., *J. Chem. Phys.* **100**, 7410 (1994).
- <sup>19</sup>J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, and L. A. Curtiss, *J. Chem. Phys.* **90**, 5622 (1989).
- <sup>20</sup>M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C.

- Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, *GAUSSIAN 03*, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.
- <sup>21</sup> Y. Zhao, N. González-García, and D. G. Truhlar, *J. Phys. Chem. A* **109**, 2012 (2005).
- <sup>22</sup> See EPAPS Document No. E-JCPSA6-127-302739 for all the relevant computed geometries and energies. This document can be reached through a direct link in the online article's HTML reference section or via the EPAPS homepage (<http://www.aip.org/pubservs/epaps.html>).
- <sup>23</sup> J. M. L. Martin and G. Oliveira, *J. Chem. Phys.* **111**, 1843 (1999).
- <sup>24</sup> W2h differs from W2 in the fact that the former employs standard cc-pVxZ basis sets for elements in groups I–IV rather than aug-cc-pVxZ basis sets for all elements. See J. M. L. Martin and S. Parthiban, in *Quantum Mechanical Prediction of Thermochemical Data*, edited by J. Cioslowski (Kluwer, Dordrecht, 2001).
- <sup>25</sup> S. Parthiban and J. M. L. Martin, *J. Chem. Phys.* **114**, 6014 (2001).
- <sup>26</sup> C. W. Bauschlicher, *Chem. Phys. Lett.* **246**, 40 (1995).
- <sup>27</sup> B. G. Johnson, P. M. W. Gill, and J. A. Pople, *J. Chem. Phys.* **98**, 5612 (1993).
- <sup>28</sup> P. L. Fast, M. L. Sanchez, and D. G. Truhlar, *Chem. Phys. Lett.* **306**, 407 (1999); L. A. Curtiss, P. C. Redfern, K. Raghavachari, V. Rassolov, and J. A. Pople, *J. Chem. Phys.* **110**, 4703 (1999); L. A. Curtiss, K. Raghavachari, P. C. Radfern, V. Rassolov, and J. A. Pople, *ibid.* **109**, 7764 (1998).
- <sup>29</sup> B. J. Lynch, Y. Zhao, and D. G. Truhlar, <http://comp.chem.umn.edu/database/>
- <sup>30</sup> J. A. Montgomery, Jr., M. J. Frisch, J. W. Ochterski, and G. A. Petersson, *J. Chem. Phys.* **110**, 2822 (1999).
- <sup>31</sup> G. S. Hammond, *J. Am. Chem. Soc.* **77**, 334 (1955).
- <sup>32</sup> N. Agmon, *J. Chem. Soc., Faraday Trans. 2* **74**, 388 (1978).
- <sup>33</sup> A. R. Miller, *J. Am. Chem. Soc.* **100**, 1984 (1978).
- <sup>34</sup> W. T. Lee and R. I. Masel, *J. Phys. Chem. A* **102**, 2332 (1998); F. A. Bulat and A. Toro-Labbé, *ibid.* **107**, 3987 (2003).
- <sup>35</sup> According to Zhao, *et al.*, “The MG3S basis is the same as MG3 except it omits the diffuse functions on hydrogens.” (cf. Ref. 21).
- <sup>36</sup> T. Van Mourik, A. K. Wilson, K. A. Peterson, D. E. Woon, and T. H. Dunning, Jr., in *Advances in Quantum Chemistry*, edited by P. Löwdin (Academic, San Diego, CA, 1998), Vol. 31, pp. 105–135.
- <sup>37</sup> Given how close the ccCA-P and ccCA-S4 values are to each other (*vide supra*), the average of the unsigned error for the ccCA variants for each reaction was taken as the ccCA unsigned error for the reaction.
- <sup>38</sup> Note that since the HLC is calculated solely using the number of paired and unpaired electrons, many reaction barriers have no contribution from the HLC.
- <sup>39</sup> Assuming that the databases used provide a good sampling of reactions on average the HLC is expected to degrade accuracy by  $(0.96 \text{ kcal/mol}) \times (23/67) - (0.44 \text{ kcal/mol}) \times (7/67) = 0.28 \text{ kcal/mol}$ . Or, looking just at cases in which the HLC contributes, an average degradation of  $(0.96 \text{ kcal/mol}) \times (23/30) - (0.44 \text{ kcal/mol}) \times (7/30) = 0.63 \text{ kcal/mol}$  results.
- <sup>40</sup> Since the HLC is calibrated for stable species, there is no *a priori* reason to believe that it is appropriate for transition states. Further, the HLC cancels unless (a) at least one of the reactants is an atomic species, or (b) the spin multiplicity of the transition state is different from the infinitely separated reactants, leading to an inconsistency in the way the HLC is handled for general points on the PES. Specifically, a large class of reactions will not have any contribution from the HLC, which, as the empirical fit parameter set for the method, is intended to account for higher-order physical effects not explicitly accounted for elsewhere.
- <sup>41</sup> P. L. Fast, M. L. Sánchez, and D. G. Truhlar, *Chem. Phys. Lett.* **306**, 407 (1999).
- <sup>42</sup> In fact, when G3 was modified to use DFT geometries and ZPE (the G3B method), the HLC parameters changed from  $[A, B, C, D] = [6.386, 2.977, 6.219, 1.185 \text{ (mhartree)}]$  to  $[6.760, 3.233, 6.786, 1.269 \text{ (mhartree)}]$  (cf. ref. 11). That is a maximum change of 0.567 mhartree for the C parameter, corresponding to the correction for electron pairs in atomic species.
- <sup>43</sup> To further disambiguate errors arising from geometry selection versus a fundamental limitation in the G3B protocol, the G3B barrier heights were recomputed using QCISD/MG3 geometries from Truhlar *et al.* (Ref. 29). MG3 is a much larger basis than 6-31G(d), *vide supra*, and should clearly illustrate any issues arising from basis set deficiencies in the computed geometries. The improvement in calculated geometries effected no clear improvement, increasing the overall MSE from 1.35 to 1.36 kcal/mol and decreasing the overall MUE from 1.94 to 1.65 kcal/mol (all “double-counted” values), reinforcing the idea of a fundamental limitation in the G3B protocol that cannot be resolved by using a more accurate geometry.
- <sup>44</sup> For simplicity, the ccCA-S4 variant results are quoted. The ccCA-P results are similar.