

# Assessment of Density Functional Theory in Predicting Structures and Free Energies of Reaction of Atmospheric Prenucleation Clusters

Jonas Elm,\* Merete Bilde, and Kurt V. Mikkelsen

Department of Chemistry, H. C. Ørsted Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark

**S** Supporting Information

**ABSTRACT:** This work assesses different computational strategies for predicting structures and Gibbs free energies of reaction of atmospheric prenucleation clusters. The performance of 22 Density Functional Theory functionals in predicting equilibrium structures of molecules and water prenucleation clusters of atmospheric relevance is evaluated against experimental data using a test set of eight molecules and prenucleation clusters:  $\text{SO}_2$ ,  $\text{H}_2\text{SO}_4$ ,  $\text{CO}_2\cdot\text{H}_2\text{O}$ ,  $\text{CS}_2\cdot\text{H}_2\text{O}$ ,  $\text{OCS}\cdot\text{H}_2\text{O}$ ,  $\text{SO}_2\cdot\text{H}_2\text{O}$ ,  $\text{SO}_3\cdot\text{H}_2\text{O}$ , and  $\text{H}_2\text{SO}_4\cdot\text{H}_2\text{O}$ . Furthermore, the functionals are tested and compared for their ability to predict the free energy of reaction for the formation of five benchmark atmospheric prenucleation clusters:  $\text{H}_2\text{SO}_4\cdot\text{H}_2\text{O}$ ,  $\text{H}_2\text{SO}_4\cdot(\text{H}_2\text{O})_2$ ,  $\text{H}_2\text{SO}_4\cdot\text{NH}_3$ ,  $\text{HSO}_4^-\cdot\text{H}_2\text{O}$ , and  $\text{HSO}_4^-\cdot(\text{H}_2\text{O})_2$ . The performance is evaluated against experimental data, coupled cluster, and complete basis set extrapolation procedure methods. Our investigation shows that the utilization of the M06-2X functional with the 6-311++G(3df,3pd) basis set represents an improved approach compared to the conventionally used PW91 functional, yielding mean absolute errors of 0.48 kcal/mol and maximum errors of 0.67 kcal/mol compared to experimental results.

## 1. INTRODUCTION

Airborne nanoparticles in the atmosphere play a central role in the global climate: They directly affect the global radiation balance by acting as scatterers or absorbers of radiation from the sun.<sup>1</sup> They also have indirect effects on the global climate by acting as cloud condensation nuclei affecting cloud properties and lifetime.<sup>2</sup> Particles can be emitted directly into the atmosphere, for example, from combustion processes, sea spray, or volcanic activity.<sup>3</sup> Particles can also be formed in situ from gas phase molecules in the atmosphere. The formation and growth of large numbers of new particles within a short time scale, so-called nucleation bursts, have been observed frequently and in a large variety of different types of air masses (e.g., urban, coastal, and remote rural) around the globe at tropic, subtropic, and mid and polar latitudes.<sup>4</sup> New particle formation in the atmosphere may significantly increase the concentration of cloud condensation nuclei, more than a factor of 2 over the course of one day.<sup>4</sup> Several nucleation mechanisms have been suggested and discussed in the literature: binary homogeneous nucleation of sulfuric acid and water,<sup>5–7</sup> ternary homogeneous nucleation of sulfuric acid and water with ammonia as a stabilizer,<sup>8–10</sup> ion mediated nucleation,<sup>11,12</sup> or organic enhanced nucleation.<sup>13–16</sup> Overall, there are strong indications that sulfuric acid plays an important role in atmospheric nucleation, but uncertainties remain with respect to the exact mechanism and to the role and nature of other participating vapors.<sup>17–19</sup> Sulfuric acid at atmospheric concentrations has been shown in laboratory experiments to be central in new particle formation,<sup>20,21</sup> and correlations between concentrations of newly formed particles in the atmosphere have been shown to correlate with sulfuric acid vapor concentrations in different atmospheric environments.<sup>22–27</sup> Recent results suggest that ammonia in the atmosphere may increase the nucleation rate of sulfuric acid particles by up to 3 orders of magnitude.<sup>28</sup>

Structures and properties of clusters composed of atmospheric precursors have been addressed in numerous investigations utilizing quantum mechanical approaches to understand the formation of prenucleation clusters on the molecular level and predict the associated activation barriers.<sup>7,12,15,16,29–33</sup>

To assess the health and climate impact of atmospheric particles and to investigate future directions of emission control and climate change mitigation, it is essential to understand the mechanisms behind nucleation and the growth of particles in the atmosphere. Here, we focus on the most likely important mechanisms for in situ particle formation in the atmosphere, namely, nucleation involving sulphuric acid. The rate of nucleation is very sensitive to the first few steps in the cluster formation, which are difficult to measure experimentally. Theoretical calculations using quantum mechanics are an excellent tool to determine the initial steps in cluster formation, but reliable high level electron correlated methods such as coupled cluster are limited to very small systems. Even the less CPU demanding second-order Møller–Plesset theory is unfeasible when a conformational screening of medium-sized clusters are needed using a large basis set. As a piece to the large puzzle of understanding atmospheric nucleation, we here review different computational approaches to calculate the structure of prenucleation clusters as well as thermochemistry of the formation of selected sulfuric acid–water and ammonia clusters.

Within recent decades, the utilization of density functional theory (DFT)<sup>34,35</sup> has proven to be extremely useful in predicting molecular equilibrium structures, thermodynamics, and properties. It is swift, applicable on larger systems due to fast basis set convergence, and thereby ideal for calculations involving relatively large clusters. The performance of DFT is

Received: March 7, 2012

Published: April 23, 2012

very dependent on the chosen form of the exchange-correlation functional, and previous investigations regarding atmospherically relevant reaction barriers, structures, and transition states have used several different functionals (B3LYP<sup>36–39</sup> BH&HLYP,<sup>40</sup> CAM-B3LYP,<sup>41</sup> BMK,<sup>42,43</sup> and PW91<sup>44–48</sup>). The functional B3LYP has been used extensively to calculate equilibrium structures, but this functional has been shown to be insufficient when describing noncovalent interactions such as the binding energy of sulfuric acid–water clusters and sulfuric acid–ammonia clusters.<sup>44,49</sup>

The BH&HLYP functional with 50% exact Hartree–Fock exchange has been used to estimate transition state properties, but at the expense of equilibrium state geometry.<sup>50</sup> The Boese and Martin functional for kinetics (BMK), which includes 42% HF exchange, has been fitted to increase the accuracy of calculating the free energy of reaction without compromising the equilibrium geometry and has been shown to give reliable results within 2 kcal/mol for the benchmark test set.<sup>51</sup> The most commonly used functional in calculating atmospheric properties is PW91, which has been shown to yield good equilibrium structures and interaction energies.<sup>52</sup>

Here, we wish to investigate which DFT functional performs the best overall in reproducing experimental observations of atmospheric prenucleation clusters in order to evaluate an optimal procedure for investigating prenucleation clusters.

## 2. COMPUTATIONAL METHODOLOGY

All DFT, MP2, and complete basis set method calculations have been performed using the Gaussian 09 program package.<sup>53</sup> The explicit coupled cluster calculations have been performed in Molpro.<sup>54</sup> Since the functionals BH&HLYP and BMK have been shown to predict accurate reaction kinetics, it could imply that the increased amount of exact HF exchange has an influence on the free energy of reaction. Thereby, it is important to test different functionals with increased HF exchange such as the long-range corrected methods:<sup>55,56</sup> LC-BLYP, LC-PBE, LC-PW91, and CAM-B3LYP. The modified Becke functionals B97-1 and B97-2 have been shown to yield good thermochemistry, and the dispersion corrected B97-D has been shown to yield good estimates of noncovalent interactions and is thereby included in the test set. The Minnesota suite of functionals (M05, M06, M06-L, M06-2X, and M06-HF) has previously shown promising applications for thermochemistry, equilibrium structures, and noncovalent interactions and is thereby included in the analysis. Several different GGAs, meta-GGAs, and hybrid functionals will be employed to evaluate the performance of different levels of theory. In order to make a comprehensive assessment of DFT functionals, the 22 functionals seen in Table 1 are tested.

Besides the explicit form of the exchange correlation functional, it is important to investigate when the desired basis set limit is reached. The most commonly used basis sets in calculations involving atmospheric species are the augmented correlation consistent basis sets with double- and triple- $\zeta$  quality (aug-cc-pVXZ, X = D, T)<sup>76</sup> and large triple- $\zeta$  Pople basis sets with different amounts of polarizations (6-311++G(X), X = (d,p), (2d,2p), (3df,3pd)).<sup>77</sup> It has been shown using a high level of coupled cluster calculations that applying additional tight d functions to the correlation consistent basis set (aug-cc-pV(X+d), X = D,T,Q<sup>78</sup>) highly improves the geometry of small sulfur-containing compounds. We wish to utilize the largest possible basis set such that the identified functional performs the best in a high basis set limit to make

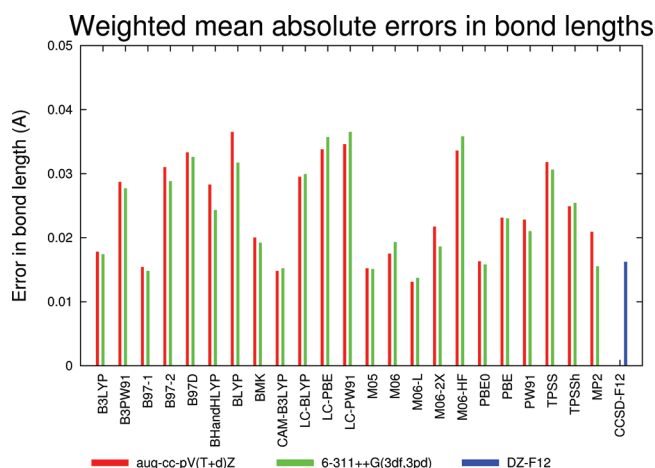
Table 1. The DFT Functionals Tested in This Investigation

functional	HF exchange	ref
B3LYP	20%	57, 58
B3PW91	20%	57
B97-1	21%	59
B97-2	21%	60
B97-D	0%	61
BH&HLYP	50%	62
BLYP	0%	63, 64
BMK	42%	51
CAM-B3LYP	19–65%	65
LC-BLYP	0–100%	56, 63, 64
LC-PBE	0–100%	56, 66
LC-PW91	0–100%	56, 67
M05	28%	68
M06	27%	69
M06-L	0%	70
M06-2X	54%	69
M06-HF	100%	71
PBE0	25%	72, 73
PBE	0%	66
PW91	0%	67
TPSS	0%	74
TPSSH	25%	75

sure that it yields the best results due to correct treatment of the system instead of yielding a “right” results due the semiempirical nature of the parametrized functional with a smaller basis set. It is unfeasible to use the extremely large correlation consistent basis sets of quadruple or higher  $\zeta$ , since the found procedure should be applicable on relatively large clusters. Thereby, the basis sets chosen in this analysis are aug-cc-pVDZ, aug-cc-pVTZ, aug-cc-pV(D+d)Z, aug-cc-pV(T+d)Z, and 6-311++G(3df,3pd),<sup>77,79,80</sup> which represent the upper limit for large prenucleation cluster calculations.

## 3. RESULTS AND DISCUSSION

**3.1. Structure of Atmospherically Relevant Compounds and Hydration Complexes.** To accurately describe cluster formation, it is important that the chosen DFT functional is able to provide accurate structures for atmospheric sulfur-containing compounds and their hydrates. A test set of the following compounds is used as a set of benchmark structures since experimental gas phase data are available and nucleation phenomena are believed to involve sulfuric acid: SO<sub>2</sub>, H<sub>2</sub>SO<sub>4</sub>, CO<sub>2</sub>·H<sub>2</sub>O, CS<sub>2</sub>·H<sub>2</sub>O, OCS·H<sub>2</sub>O, SO<sub>2</sub>·H<sub>2</sub>O, SO<sub>3</sub>·H<sub>2</sub>O, and H<sub>2</sub>SO<sub>4</sub>·H<sub>2</sub>O). The compounds SO<sub>2</sub>, CO<sub>2</sub>·H<sub>2</sub>O, CS<sub>2</sub>·H<sub>2</sub>O, OCS·H<sub>2</sub>O, SO<sub>2</sub>·H<sub>2</sub>O, and SO<sub>3</sub>·H<sub>2</sub>O have earlier been investigated with high level coupled cluster methods.<sup>78,81</sup> This test set presents huge challenges for DFT due to the noncovalent interactions involved in the formation of hydrates. By utilizing the functionals summarized in Table 1 when calculating the geometries of the above chosen test set, it is possible to determine which functional that overall performs the best in predicting the equilibrium geometry of atmospherically relevant hydrate complexes as well as finding the most suitable basis set. In Figure 1, the weighed mean absolute errors (WMAE) in all of the bond lengths of the test set can be seen depending on the functional and 6-311++G(3df,3pd)/aug-cc-pV(T+d)Z basis set in comparison with the electron correlated MP2 and singles doubles explicit coupled cluster (CCSD-F12<sup>82</sup>) with the basis set VDZ-F12. A comparison of all basis



**Figure 1.** The weighted mean absolute errors (WMAE) in all bond lengths (Å) of the test set.

sets can be seen in the Supporting Information. The explicit couple cluster method has been shown to yield converged energies 2 orders of magnitude faster than conventional coupled cluster and is thereby applicable on medium-sized systems.<sup>82</sup>

It was observed that the small correlation-consistent basis set aug-cc-pVDZ displays relatively large errors for all of the functionals tested up to 0.052 Å in the case of B97-D (Figure 1 of the Supporting Information). Thereby, it can be concluded to be too small for accurately describing the equilibrium structures of these compounds. In all cases, a significant improvement is found when increasing the basis set to aug-cc-pVTZ. A similar improvement is observed when adding an additional tight d function to sulfur. Interestingly, it is observed that the Pople style basis set 6-311++G(3df,3pd) in most cases exhibits lower errors than the larger correlation consistent aug-cc-pV(T+d)Z, and thereby it is advised that this smaller basis set be used when calculating bond lengths of prenucleation clusters. The functionals perform in the following order when looking at the errors with the 6-311++G(3df,3pd) basis set:

- group 1 (errors ∈ [0.014 Å, 0.020 Å]): M06-L, B97-1, M05, CAM-B3LYP, MP2, PBE0, CCSD-F12, B3LYP, M06-2X, BMK, M06
- group 2 (errors ∈ [0.021 Å, 0.030 Å]): PW91, PBE, BH&HLYP, TPSSH, B3PW91, B97-2, LC-BLYP
- group 3 (errors ∈ [0.031 Å, 0.037 Å]): TPSS, BLYP, B97-D, LC-PBE, M06-HF, LC-PW91

This indicates that several functionals (group 1) yield convincing structures with mean absolute errors as low as 0.014 Å in the case of the M06-L functional. It is observed that most of the conventional hybrids lie in this group together with a single long-range corrected functional (CAM-B3LYP). Several of the functionals perform on par or even better than MP2 or the CCSD-F12/VDZ-F12 calculations. The functionals in group 2 yield relatively good structures, and the group consists of a collection of both GGAs and hybrids. Group 3, which exhibits errors greater than 0.031 Å, consists mainly of the functionals with a very high amount of exact HF exchange. The maximum errors in bond lengths follow the same trend as the weighted mean absolute errors. All of the functionals designated to group 1 also yield maximum errors lower than 0.040 Å. The maximum errors in bond lengths for the basis sets

aug-cc-pV(T+d)Z and 6-311++G(3df,3pd) can be seen in the Supporting Information.

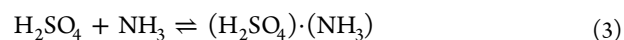
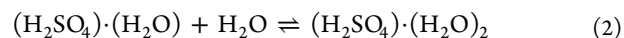
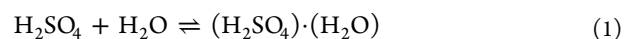
In addition to concern about errors in bond lengths, it is also important to obtain relatively good bond angles. Contrary to the bond lengths, the weighted mean absolute errors and maximum errors in bond angles were very similar for most of the functionals (see Supporting Information). They are generally found to follow the same trend as the errors in bond length regarding basis set convergence. The Pople basis set 6-311++G(3df,3pd) is sufficient to produce converged structures, and the aug-cc-pV(T+d)Z basis set is generally not an improvement. All of the functionals exhibit mean errors in the range 0.70–0.89° and maximum errors in the range 0.89–1.60°, with the exception of the functionals that include a high amount of HF exchange (M06-HF, LC-BLYP, LC-PW91, LC-PBE, BH&HLYP, BMK) where the mean errors ranged from 1.03° to 1.49° and maximum errors ranged from 1.71° to 2.53°. The electron correlated methods perform significantly better in reproducing the bond angles than DFT. MP2 yields mean errors as low as 0.59° and a max error of 0.88°, while CCSD-F12 yields a mean error of 0.32° and a max error of 0.96°.

Generally, the most challenging structural feature to describe, both experimentally and theoretically, is the noncovalent interaction between the compound and a water molecule. By exclusively investigating this intermolecular distance, we obtain a better understanding of which functional performs the best for these intermolecular interactions. It is found that the functionals assigned to group 1 above generally exhibit low errors in describing the noncovalent interaction (see the Supporting Information). The lowest error is again achieved by M06-L with 0.03 Å. Generally, the functionals assigned to group 1 are found to have little variation in errors, only ranging from 0.03 to 0.08 Å. Interestingly, the functionals PW91, PBE, and BH&HLYP, which were assigned to group 2, also exhibit low errors of 0.050, 0.057, and 0.058 Å, respectively.

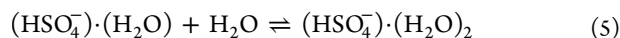
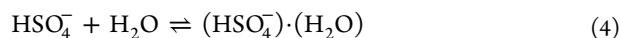
In some cases, it was difficult to achieve converged structures. This was most often experienced with the BMK, M05, M06-L, and M06-HF functionals. In the case of the BMK functional, a loose convergence criterion was needed to obtain a converged structure of the SO<sub>2</sub>·H<sub>2</sub>O compound with aug-cc-pVDZ and aug-cc-pVTZ basis sets.

From these investigations, it is apparent that all of the functionals designated in group 1 are suitable to yield equilibrium structures with both bond lengths and angles in relatively good agreement with experimental results. The LC functionals, M06-HF, BH&HLYP, and BMK, yielded large errors both regarding bond lengths and angles and should generally not be used for these types of compounds.

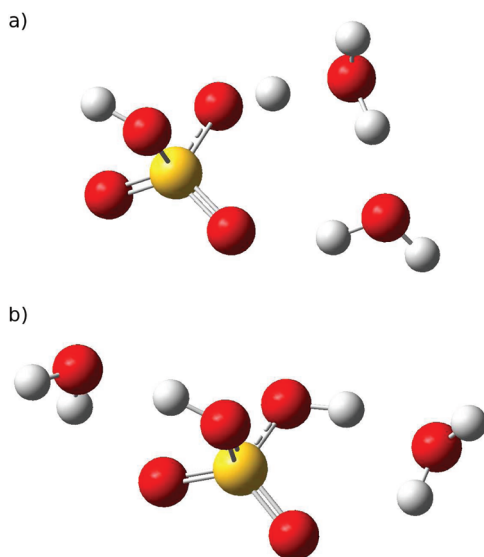
**3.2. Thermochemistry of Prenucleation Cluster Reactions.** Besides being able to predict the experimental structures of atmospherically relevant compounds, it is also utterly important that the functional is able to produce reliable free energies of reaction. In order to evaluate the overall performance of the functionals, several different nucleation reactions of sulfuric acid and bisulfate with water and ammonia are investigated. The following reactions are considered since experimental data are available.







When more than one water molecule form a complex with either  $\text{H}_2\text{SO}_4$  or  $\text{HSO}_4^-$ , it is apparent that there exist several different conformations that might influence the reaction. In reaction 2, the mechanism could be either that both water molecules hydrate the same S–OH group at one side in order to create a larger hydrogen-bonded network as seen in Figure 2a or each of the S–OH groups could attach a single water molecule each as seen in Figure 2b.



**Figure 2.** The figure shows the two different conformations of  $\text{H}_2\text{SO}_4$  with two water molecules considered for reaction 2. (a) The hydration of two water molecules at the same S–OH bond site. (b) The hydration of two water molecules at opposing S–OH sites.

Since both of these conformations lie relatively close in energy (around 0–2 kcal/mol depending on the functional), it is clear that both conformations need to be taken into consideration when calculating the overall free energy of reaction. The probability that a system is in configuration 1 is given by

$$P(1) = \frac{\exp(-\Delta G_1/RT)}{Z}$$

$$Z = \sum_s \exp(-\Delta G_s/RT)$$

Here,  $Z$  denotes the partition function and  $s$  is the label for the different configurations. By weighting each conformation by their relative populations, the total free energy of reaction ( $\Delta G_r$ ) can be calculated. In the case of reaction 5, three different conformations were taken into account to calculate the total free energy of reaction. In Table 2, the calculated free energies of reaction can be seen for all of the functionals compared to the experimental values for reactions 1–5 utilizing the basis set 6-311++G(3df,3pd). The numbers in parentheses show the free energy of reaction with counterpoise correction in order to eliminate basis set superposition errors.<sup>83</sup> The complete basis set methods CBS-QB3,<sup>84</sup> G3,<sup>85</sup> and W1<sup>86</sup> are energy calculations involving several predefined steps and extrapolation procedures to the complete basis set limit. The

CBS-QB3 and W1 methods use B3LYP geometries and frequencies and HF/higher order electron correlated methods for energies and complete basis set extrapolation. The G3 method uses MP2 for the geometry and HF/higher order electron correlated methods for energies and complete basis set extrapolation.

It is observed that the functionals have very different performances in calculating these simple prenucleation reactions. A few of the functionals predict the wrong sign of the sulfuric acid hydrations (BLYP and B97-2), and several of the functionals (BLYP, BMK, PBE, PW91) falsely predict the second hydration (eq 2) of sulfuric acid to be more favorable than the first (eq 1). Several of the functionals exhibit extremely large mean absolute errors up to 3 kcal/mol. It is observed that the long-range corrected functionals (LC-BLYP, LC-PBE, and LC-PW91) and M06-HF are the only functionals to overestimate the reaction energies. They yield a low mean error  $\sim 1$  kcal/mol and maximum error  $\sim 2$  kcal/mol, but at the same time the regression coefficient is as low as 0.73, indicating extreme fluctuation in the performance. Most of the hybrid functionals are seen to slightly underestimate the reactions. The commonly utilized B3LYP functional yields mean errors around 3 kcal/mol, a max error of 4.54 kcal/mol, and  $R^2 = 0.80$ , and it is thereby not the best choice to use in these types of reactions, as also indicated in earlier studies.<sup>44,49</sup>

The commonly used PW91 functional is seen to yield an extremely low mean error of 0.95 kcal/mol and a maximum error of 1.73, but  $R^2 = 0.87$  indicates that the correlation between the calculated data and experimental data is not in the best agreement. More severe is the incorrect assignment of the second hydration, being more favorable than the first for the hydration of sulfuric acid. A few of the functionals which were categorized in group 1 are found to also yield excellent free energies of reaction. The conventional PBE0 functional is observed to perform extremely well with mean errors as low as 1.45 kcal/mol, a max error of 2.13 kcal/mol, and  $R^2 = 0.95$ . This could be a result of the nonempirical nature of the functional in contrast to B3LYP. Interestingly, the long-range corrected CAM-B3LYP functional yields decent reaction energies with mean errors of 1.27 kcal/mol, a max error of 2.06 kcal/mol, and  $R^2 = 0.94$ . The functional that performs the best is the M06-2X functional, with mean errors of 0.48 kcal/mol, a max error of 0.67 kcal/mol, and  $R^2 = 0.97$ . It is both able to give quantitatively good energies and maintain a high correlation with the experimental data, and it yields a free energy of reaction which, in most cases, is within the uncertainty of the experiments. It is in good agreement with the calculated high level complete basis set method G3. The possible reason for the discrepancies between CBS-QB3/W1 results and G3 could be due to CBS-QB3 and W1 using B3LYP geometries and frequencies which both here and otherwise have been shown to be inaccurate regarding interaction energies and thereby limits the usefulness of these procedures.

When applying the counterpoise correction to the data, it is observed that in all cases except M06-HF the MAE and maximum error increase severely and the  $R^2$  value decreases. By close investigation of the reactions, it is observed that for each hydration step the counterpoise correction is approximately constant and for most functionals only varies from 0.5 to 0.7 kcal/mol. This indicates that it can be neglected without inducing severe errors. In the case of the M06-2X functional, the following counterpoise corrections to the energies are obtained: (1) 0.7 kcal/mol, (2) 0.67 kcal/mol, (3) 0.56 kcal/mol.

**Table 2.** The Performance of DFT in Calculating the Free Energies of Reaction of the Different Atmospheric Prenucleation Reactions 1–5<sup>a</sup>

	$\Delta G_r(1)$	$\Delta G_r(2)$	$\Delta G_r(3)$	$\Delta G_r(4)$	$\Delta G_r(5)$	MAE	ME	$R^2$
B3LYP	−0.72 (−0.12)	−0.70 (−0.24)	−5.23 (−4.77)	−1.50 (−0.90)	−2.22 (−1.81)	2.94 (3.45)	4.54 (5.14)	0.80 (0.77)
B3PW91	−0.36 (0.34)	−0.08 (0.62)	−5.46 (−4.92)	−2.23 (−1.53)	−1.94 (−1.50)	3.00 (3.23)	3.81 (4.51)	0.94 (0.84)
B971	−1.40 (−0.78)	−1.15 (−0.55)	−6.00 (−5.51)	−3.54 (−2.91)	−3.28 (−2.90)	1.94 (2.49)	2.50 (3.13)	0.95 (0.93)
B972	−0.05 (0.65)	0.41 (1.11)	−4.88 (−4.32)	−2.45 (−1.78)	−2.86 (−2.39)	2.89 (2.97)	3.62 (4.26)	0.85 (0.78)
B97D	−1.22 (−0.51)	−0.70 (0.05)	−6.25 (−5.69)	−4.43 (−3.76)	−2.70 (−1.22)	1.96 (2.77)	2.38 (3.09)	0.98 (0.96)
BH&HLYP	−1.72 (−1.12)	−1.23 (−0.64)	−5.51 (−4.31)	−3.47 (−2.85)	−2.79 (−2.39)	2.07 (2.75)	2.99 (4.19)	0.99 (0.97)
BLYP	0.33 (0.91)	0.00 (0.58)	−4.42 (−3.99)	−2.15 (−1.57)	−1.86 (−1.55)	3.26 (3.30)	4.08 (4.51)	0.96 (0.89)
BMK	−0.85 (−0.46)	−0.92 (−0.54)	−5.43 (−5.12)	−2.79 (−2.40)	−2.44 (−2.20)	2.53 (2.87)	3.25 (3.64)	0.93 (0.92)
CAM-B3LYP	−2.36 (−1.77)	−2.20 (−1.62)	−6.67 (−6.22)	−3.98 (−3.36)	−3.52 (−2.94)	1.27 (1.83)	2.06 (2.94)	0.94 (0.94)
LC-BLYP	−4.89 (−4.35)	−4.61 (−4.09)	−9.07 (−8.64)	−6.25 (−5.64)	−6.04 (−5.74)	1.16 (0.84)	2.31 (5.74)	0.93 (0.91)
LC-PBE	−4.71 (−4.04)	−4.25 (−3.59)	−10.05 (−9.51)	−4.96 (−4.23)	−5.51 (−4.75)	1.31 (0.93)	1.95 (4.75)	0.77 (0.76)
LC-PW91	−4.78 (−4.12)	−4.34 (−3.54)	−10.03 (−9.50)	−4.75 (−4.03)	−5.61 (−4.86)	1.40 (1.00)	2.04 (4.86)	0.73 (0.73)
M05	−1.73 (−1.04)	−1.24 (−0.55)	−5.35 (−4.78)	−4.11 (−3.45)	−3.18 (−2.59)	1.89 (2.53)	3.15 (3.72)	0.97 (0.97)
M06	−2.12 (−1.51)	−1.73 (−1.12)	−5.96 (−5.47)	−4.08 (−3.54)	−3.84 (−3.27)	1.47 (2.03)	2.54 (3.27)	0.95 (0.95)
M06-2X	−2.93 (−2.23)	−2.78 (−2.11)	−7.84 (−7.28)	−5.82 (−5.12)	−4.27 (−3.75)	0.48 (0.92)	0.67 (3.75)	0.97 (0.97)
M06-HF	−5.37 (−3.63)	−4.33 (−2.61)	−9.59 (−8.22)	−7.31 (−5.56)	−5.36 (−3.26)	1.38 (0.50)	2.03 (3.26)	0.96 (0.93)
M06-L	−1.81 (−1.31)	−0.78 (−0.29)	−5.79 (−5.37)	−4.59 (−4.14)	−3.33 (−2.91)	1.76 (2.21)	2.71 (3.13)	0.96 (0.96)
PBE0	−1.98 (−1.27)	−1.68 (−0.98)	−7.02 (−6.48)	−3.91 (−3.17)	−3.23 (−2.64)	1.45 (2.11)	2.13 (2.87)	0.95 (0.95)
PBE	−1.75 (−1.03)	−2.09 (−1.39)	−7.19 (−6.63)	−3.84 (−3.09)	−3.46 (−2.87)	1.35 (2.01)	2.20 (2.95)	0.90 (0.89)
PW91	−2.33 (−1.51)	−2.83 (−2.04)	−7.81 (−7.20)	−4.31 (−3.47)	−4.12 (−3.41)	0.95 (1.49)	1.73 (3.41)	0.87 (0.86)
TPSS	−0.97 (−0.11)	−0.96 (−0.12)	−6.10 (−5.48)	−2.93 (−2.04)	−2.72 (−2.12)	2.28 (3.04)	3.11 (4.00)	0.93 (0.91)
TPSSH	−1.10 (−0.26)	−0.90 (−0.07)	−6.05 (−5.43)	−3.03 (−2.17)	−2.68 (−2.09)	2.26 (3.01)	2.50 (3.87)	0.95 (0.93)
MP2	−2.33 (−0.60)	−1.94 (−0.18)	−7.16 (−5.28)	−4.83 (−3.04)	−3.81 (−2.42)	1.00 (2.71)	1.34 (3.22)	0.98 (0.97)
CBS-QB3	−1.90	−1.37	−6.91	−4.56	−4.14	1.24	1.70	0.96
G3	−2.76	−2.04	−7.16	−5.16	−4.27	0.74	1.34	0.99
W1	−1.92		−6.35	−4.21		1.89	2.15	1.00
M06-2X//M06-L	−3.07	−2.43	−6.61	−5.79	−4.01	0.63	1.89	0.95
exptl	$-3.6 \pm 1.0^{87}$	$-2.3 \pm 0.3^{87}$	$-8.5^{88}$	$-6.04 \pm 0.7^{89}$	$-4.64 \pm 0.7^{89}$			

<sup>a</sup>All of the reactions have been calculated at 1 atm of pressure and 298.15 K with the basis set 6-311++G(3df,3pd) and are shown in kcal/mol. MAE is the mean absolute error in comparison to the experimental data. ME is the maximum error, and  $R^2$  is the linear regression coefficient. The numbers in parentheses denote that a counterpoise energy correction is employed.

mol, (4) 0.70 kcal/mol, and (5) 0.52 kcal/mol. This indicates that by completely neglecting basis set superposition errors in these test reactions, the error induced is no larger than 0.18 kcal/mol from the constant offset. Thereby, it is possible to utilize M06-2X/6-311++G(3df,3pd) without using the cumbersome counterpoise correction and still achieve reliable free energy of reaction.

The M06-L functional was found to be the ultimate functional regarding both the errors in overall structure and the errors in noncovalent interactions alone, and the M06-2X functional was found to be superior in calculating free energies of reaction. The combination of the two functionals was also tested: M06-2X/6-311++G(3df,3pd)//M06-L/6-311++G(3df,3pd). By utilizing this combination, a MAE of 0.63 kcal/mol, with a maximum error of 1.89 kcal/mol, was obtained, with  $R^2 = 0.95$ . The increase in error is almost exclusively due to reaction 3 with ammonia, where the free energy of reaction is severely underestimated. This could be an indication of the M06-L functional's performance in calculating the geometry of the ammonia complex, and therefore both M06-2X/6-311++G(3df,3pd) geometries and energies might be a better overall estimate of atmospheric reactions.

#### 4. CONCLUSION

Here, we have investigated the performance of DFT in reproducing molecular equilibrium geometries and the thermochemistry of atmospherically relevant prenucleation

clusters. Looking at the overall performance, the M06-2X functional utilizing the 6-311++G(3df,3pd) basis set is an excellent density functional for estimating cluster formation through binary/ternary nucleation of sulfuric acid and water/ammonia both in neutral and ionized forms. Only a single calculation involving ammonia is incorporated in the benchmark set, and thereby caution should be taken when utilizing these findings in ternary nucleation. Previously, it has been discussed that the heavily parametrized Minnesota functionals should be used with care since the parametrization could damage transferability to other systems.<sup>90</sup> Here, we find that the M06-2X functional performs well in the case of sulfur-hydrate complexes, and this result (as well as all DFT results) should of course be handled with care when calculating nucleation involving other species. When dealing with small clusters, it is often a better idea to utilize MP2 instead of DFT; even though the methods exhibit larger mean absolute errors, they are often more systematic. Nevertheless, the M06-2X functional is worth including in future assessments of large prenucleation clusters, as it shows great promise and is an interesting alternative to the conventional PW91 functional.

#### ■ ASSOCIATED CONTENT

##### Supporting Information

The weighted mean absolute errors in bond lengths and bond angles as well as the error in noncovalent distances are available for all tested basis sets. Maximum errors in bond lengths and

bond angles for the basis sets aug-cc-pV(T+d)Z and 6-311++G(3df,3pd) can be seen also. This information is available free of charge via the Internet at <http://pubs.acs.org/>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [elm@chem.ku.dk](mailto:elm@chem.ku.dk).

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors thank the Danish Center for Scientific Computing for providing computer resources, the Danish Natural Science Research Council/The Danish Councils for Independent Research, and the Villum Kann Rasmussen Foundation for financial support

## REFERENCES

- Haywood, J.; Boucher, O. *Rev. Geophys.* **2000**, *38*, 513–543.
- Lohmann, U.; Feichter, J. *Atmos. Phys. Chem.* **2005**, *5*, 715–737.
- Poschl, U. *Angew. Chem., Int. Ed.* **2005**, *44*, 7520–7540.
- Kulmala, M.; Vehkamäki, H.; Petaja, T.; Maso, M. D.; Lauri, A.; Kerminen, V. M.; Birmili, W.; McMurry, P. H. *J. Aerosol. Sci.* **2004**, *35*, 143–176.
- Kulmala, M.; Laaksonen, A. *J. Chem. Phys.* **1990**, *93*, 696–701.
- Yu, F. *J. Geophys. Res.* **2006**, *111*, D04201:1–17.
- Natsheh, A.; Nadykto, A. B.; Mikkelsen, K. V.; Yu, F.; Ruuskanen, J. *J. Phys. Chem. A* **2006**, *110*, 7982–7083.
- Kulmala, M.; Pirjola, L.; Mäkelä, J. M. *Nature* **2000**, *404*, 66.
- Napari, I.; Noppel, M.; Vehkamäki, H.; Kulmala, M. *J. Geophys. Res.* **2002**, *107*, 1–6.
- Benson, D. R.; Yu, J. H.; Markovich, A.; Lee, S.-H. *Atmos. Chem. Phys.* **2011**, *11*, 4755–4766.
- Yu, F.; Turco, P. *Geophys. Res. Lett.* **2000**, *27*, 883–886.
- Nadykto, A. B.; Natsheh, A.; Yu, F.; Mikkelsen, K. V.; Ruuskanen, J. *Phys. Rev. Lett.* **2006**, *96*, 125701.
- O'Dowd, C. D.; Aalto, P.; Hmeri, K.; Kulmala, M.; Hoffmann, T. *Nature* **2002**, *416*, 497–498.
- Zhang, R.; Suh, I.; Zhao, J.; Zhang, D.; Fortner, E. C.; Tie, X.; Molina, L. T.; Molina, M. J. *Science* **2004**, *304*, 1487–1490.
- Sloth, M.; Bilde, M.; Mikkelsen, K. V. *Mol. Phys.* **2005**, *102*, 2361–2368.
- Falsig, H.; Gross, A.; Kongsted, J.; Osted, A.; Sloth, M.; Mikkelsen, K. V.; Christiansen, O. *J. Phys. Chem. A* **2006**, *110*, 660–670.
- Laakso, L.; Gagné, S.; Petäjä, T.; Hirsikko, A.; Aalto, P. P.; Kulmala, M.; Kerminen, V. M. *Atmos. Chem. Phys.* **2007**, *7*, 1333–1345.
- Kurtén, T.; Noppel, M.; Vehkamäki, H.; Salonen, M.; Kulmala, M. *Boreal Env. Res.* **2007**, *12*, 431–453.
- Kulmala, M.; Asmi, A.; Lappalainen, H. K.; Carslaw, K. S.; Pöschl, U.; Baltensperger, U.; Hov, Ø.; Brenquier, J.-L.; Pandis, S. N.; Facchini, M. C.; Hansson, H.-C.; Wiedensohler, A.; O'Dowd, C. D. *Atmos. Chem. Phys.* **2009**, *9*, 2825–2841.
- Sipila, M.; Berndt, T.; Petaja, T.; Brus, D.; Vanhanen, J.; Stratmann, F.; Patokoski, J.; Mauldin, R. L.; Hyvarinen, A. P.; Lihavainen, H.; Kulmala, M. *Science* **2010**, *327*, 1243–1246.
- Berndt, T.; Boge, O.; Stratmann, F.; Heintzenberg, J.; Kulmala, M. *Science* **2005**, *307*, 698–700.
- Weber, R. J.; Marti, J. J.; McMurry, P. H.; Eisele, F. L.; Tanner, D. J.; Jefferson, A. *Chem. Eng. Commun.* **1996**, *151*, 53–64.
- Weber, R. J.; Marti, J. J.; McMurry, P. H.; Eisele, F. L.; Tanner, D. J.; Jefferson, A. *J. Geophys. Res.* **1997**, *102*, 4375–4385.
- Kulmala, M.; Lehtinen, K. E. J.; Laaksonen, A. *Atmos. Chem. Phys.* **2006**, *6*, 787–793.
- Sihto, S.-L.; Kulmala, M.; Kerminen, V.-M.; Dal-Maso, M.; Petäjä, T.; Riipinen, I.; Korhonen, H.; Arnold, F.; Janson, R.; Boy, M.; Laaksonen, A.; Lehtinen, K. E. *J. Atmos. Chem. Phys.* **2006**, *6*, 4079–4091.
- Riipinen, I.; Sihto, S.-L.; Kulmala, M.; Arnold, F.; Dal-Maso, M.; Birmili, W.; Saarnio, K.; Teinilä, K.; Kerminen, V.-M.; Laaksonen, A.; Lehtinen, K. E. *J. Atmos. Chem. Phys.* **2007**, *7*, 1899–1914.
- Kuang, C.; McMurry, P. H.; McCormick, A. V.; Eisele, F. L. *J. Geophys. Res.—Atmos.* **2008**, *113*, D10209.
- Kirkby, J.; et al. *Nature* **2011**, *476*, 429–435.
- Gross, A.; Mikkelsen, K. V. *Adv. Quantum Chem.* **2005**, *50*, 125–142.
- Natsheh, A.; Nadykto, A. B.; Mikkelsen, K. V.; Yu, F.; Ruuskanen, J. *Chem. Phys. Lett.* **2006**, *426*, 20–25.
- Sloth, M.; Gross, A.; Falsig, H.; Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. *Chem. Phys.* **2008**, *348*, 21–30.
- Gross, A.; Nielsen, O. J.; Mikkelsen, K. V. *Adv. Quantum Chem.* **2008**, *55*, 355–385.
- Madsen, M. S.; Gross, A.; Mikkelsen, K. V. *Int. J. Quant. Chem.* **2011**, *111*, 1740–1747.
- Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, 864–871.
- Kohn, W.; Sham, L. J. *Phys. Rev. A* **1965**, *140*, 1133–1138.
- Natsheh, A. A.; Nadykto, A. B.; Mikkelsen, K. V.; Yu, F.; Ruuskanen, J. *Chem. Phys. Lett.* **2006**, *426*, 20–25.
- Long, B.; Zhang, W.-J.; Tan, X.-F.; Long, Z.-W.; Wang, Y.-B.; Ren, D.-S. *J. Phys. Chem. A* **2011**, *115*, 1350–1357.
- Jørgensen, S.; Gross, A. *J. Phys. Chem. A* **2009**, *113*, 10284–10290.
- Jørgensen, S.; Kjaergaard, H. G. *J. Phys. Chem. A* **2010**, *114*, 4857–4863.
- Thomsen, D. L.; Jørgensen, S. *Chem. Phys. Lett.* **2009**, *481*, 29–33.
- Bork, N.; Kurtén, T.; Enghoff, M. B.; Pedersen, J. O. P.; Mikkelsen, K. V.; Svendsmark, H. *Atmos. Phys. Chem. Discuss.* **2011**, *11*, 13947–13973.
- Vandeputte, A. G.; Sabbe, M. K.; Reyniers, M.-F.; Speybroeck, V. V.; Waroquier, M.; Marin, G. B. *J. Phys. Chem. A* **2007**, *111*, 11771–11786.
- Hemelseoet, K.; Speybroeck, V. V.; Waroquier, M. *Chem. Phys. Chem.* **2008**, *9*, 2349–2358.
- Nadykto, A. B.; Yu, F. *Chem. Phys. Lett.* **2007**, *435*, 14–18.
- Nadykto, A. B.; Yu, F.; herb, J. *Chem. Phys.* **2009**, *360*, 67–73.
- Xu, Y.; Nadykto, A. B.; Yu, F.; Jiang, L.; Wang, W. *THEOCHEM* **2010**, *951*, 28–33.
- Nadykto, A. B.; Yu, F.; Jakovleva, M. V.; Herb, J.; Xu, Y. *Entropy* **2011**, *13*, 554–569.
- Kurtén, T.; Torpo, L.; Ding, C.-G.; Vehkamäki, H.; Sundberg, M. R.; Laaksonen, K.; Kulmala, M. *J. Geophys. Res.* **2007**, *112*, 1–7.
- Kurtén, T.; Sundberg, M. R.; Vehkamäki, H.; Noppel, M.; Blomquist, J.; Kulmala, M. *J. Phys. Chem.* **2006**, *110*, 7178–7188.
- Durant, J. L. *Chem. Phys. Lett.* **1996**, *256*, 1996.
- Boese, A. D.; Martin, J. M. L. *J. Chem. Phys.* **2004**, *121*, 3405–3416.
- Tsuzuki, S.; Lüthi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949–3957.
- Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision B.01; Gaussian, Inc.: Wallingford, CT, 2010.

- (54) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; et al. *MOLPRO*, version 2010.1; Cardiff University: Cardiff, Wales, U.K.
- (55) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2004**, *120*, 8427–8433.
- (56) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2001**, *115*, 3540–3544.
- (57) Becke, A. J. *Chem. Phys.* **1993**, *98*, 5648–5652.
- (58) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (59) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. *J. Phys. Chem.* **1998**, *109*, 6264–6271.
- (60) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. *J. Phys. Chem.* **2001**, *115*, 9233–9242.
- (61) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1799.
- (62) Becke, A. J. *Chem. Phys.* **1993**, *98*, 1372–1377.
- (63) Becke, A. D. *Phys. Rev. A* **1988**, *23*, 3098–3100.
- (64) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (65) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (66) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (67) Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. *Phys. Rev. B* **1992**, *46*, 6671–6687.
- (68) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123*, 161103:1–4.
- (69) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (70) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 94101:1–18.
- (71) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 13126–13130.
- (72) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029–5036.
- (73) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (74) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401–4.
- (75) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *J. Chem. Phys.* **2003**, *119*, 12129–12137.
- (76) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (77) Franchi, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654–3665.
- (78) Wilson, A. K.; Dunning, T. H. *J. Chem. Phys.* **2003**, *119*, 11712–11714.
- (79) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.
- (80) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265–3269.
- (81) Garden, A. L.; Lane, J. R.; Kjaergaard, H. G. *J. Chem. Phys.* **2006**, *125*, 144317:1–7.
- (82) Adler, T. B.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2007**, *127*, 221106:1–4.
- (83) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.
- (84) Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, *110*, 2811–2827.
- (85) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764–7776.
- (86) Martin, J. M. L.; de Oliveira, G. *J. Chem. Phys.* **1999**, *111*, 1843–1856.
- (87) Hanson, D. R.; Eisele, F. *J. Phys. Chem. A* **2000**, *104*, 1715–1719.
- (88) Hanson, D. R.; Eisele, F. *J. Phys. Chem. A* **2000**, *104*, 830–836.
- (89) Froyd, K. D.; Lovejoy, E. R. *J. Phys. Chem. A* **2003**, *107*, 9812–9824.
- (90) Kurtén, T.; Vehkamäki, H. *Adv. Quant. Chem.* **2008**, *55*, 407–427.