

# Accuracy of Effective Core Potentials and Basis Sets for Density Functional Calculations, Including Relativistic Effects, As Illustrated by Calculations on Arsenic Compounds

Xuefei Xu and Donald G. Truhlar\*

Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431, United States

**S** Supporting Information

**ABSTRACT:** For molecules containing the fourth-period element arsenic, we test (i, ii) the accuracy of all-electron (AE) basis sets from the def2- $x$ ZVP and ma- $x$ ZVP series (where  $x$ Z is S, TZ, or QZ), (iii) the accuracy of the 6-311G series of AE basis sets with additional polarization and diffuse functions, and (iv) the performance of effective core potentials (ECPs). The first set of tests involves basis-set convergence studies with eleven density functionals for five cases: equilibrium dissociation energy ( $D_e$ ) of  $\text{As}_2$ , vertical ionization potential (VIP) of  $\text{As}$ , IP of  $\text{As}$ , acid dissociation of  $\text{H}_3\text{AsO}_4$ , and  $D_e$  of  $\text{FeAs}$ . A second set of tests involves the same kinds of basis-set convergence studies for the VIP and  $D_e$  values of  $\text{As}_3$  and  $\text{As}_4$  clusters. Both relativistic and nonrelativistic calculations are considered, including in each case both AE calculations and calculations with ECPs. Convergence and accuracy are assessed by comparing to relativistic AE calculations with the cc-pV5Z-DK or ma-cc-pV5Z-DK basis and to nonrelativistic AE calculations with the cc-pV5Z or ma-cc-pV5Z basis. The primary objective of this study is to evaluate the abilities of ECPs with both their recommended basis sets and other basis sets to reproduce the results of all-electron relativistic calculations. The performance of the def2 and ma series basis sets is consistent with their sizes, and quadruple- $\zeta$  basis sets are the best. The def2-TZVP basis set performs better than most of the 6-311G series basis sets, which are the most commonly used basis sets in the previous studies of arsenic compounds. However, relativistic def2-TZVP calculations are not recommended. The large-core ECPs, which are the only available ECPs for arsenic in the popular Gaussian program, have average errors of 9–12 kcal/mol for the arsenic systems studied; therefore, these ECPs are not recommended. The triple- $\zeta$  small-core relativistic ECP (RECP) basis set cc-pVTZ-PP is found to have performance better than that of the def2-TZVP basis set, and it is highly recommended for arsenic-containing systems. The double- $\zeta$  RECP basis set ma-sc-SVP is recommended for large arsenic systems for which the def2-TZVP and cc-pVTZ-PP basis sets are unaffordable, if a basis-set error of  $\sim 2$  kcal/mol can be tolerated.

## 1. INTRODUCTION

An effective core potential (ECP) is a potential energy function added to an electronic structure calculation to replace the explicit treatment of core electrons. ECPs have two main advantages in quantum chemistry: they reduce the cost of calculations by decreasing the size of the basis set required for a given accuracy (core basis functions are not needed), and—by using a relativistic ECP (RECP)—they allow the inclusion of relativistic effects on the size and shape of core orbitals without using a relativistic treatment of the wave function. A third benefit is that using ECPs reduces basis set superposition errors. The use of ECPs is now well-established for calculating wave functions of systems containing many-electron atoms, and ECPs developed with Hartree–Fock (HF) wave functions have been found to be “very reliable and highly transferable”<sup>1</sup> for such calculations. The valence orbitals in a calculation employing an ECP differ from the valence orbitals in an all-electron (AE) calculation and are called pseudo-orbitals.

However, when one proceeds from wave function theory (WFT) to Kohn–Sham (KS) density functional theory (DFT), new issues arise because of the nonlinear dependence of the exchange–correlation (xc) functional on density.<sup>2,3</sup> This is widely recognized in solid-state physics, and so-called nonlinear core corrections (NLCCs) are widely employed in plane wave

codes for density functional calculations of extended systems. (Despite its name, an NLCC is not an add-on correction but rather constitutes a more elaborate method for defining and using an ECP by changing the treatment of the core density.) In calculations with plane wave basis sets, the ECP is usually called a pseudopotential (PP), and it is more approximate than the ECPs usually used in quantum chemistry for calculations with Gaussian basis functions because one attempts to find a compromise between high accuracy in representing the effect of core electrons and the need to include high-momentum plane waves to represent the pseudo-orbitals. Perhaps for this reason it has not been recognized that NLCCs may also be needed for calculations with Gaussian basis sets, and in fact, NLCCs are not employed in popular Gaussian-based electronic structure packages.

The need for NLCCs in Gaussian-based DFT calculations has been studied by van Wüllen,<sup>4</sup> who wrote, “Since the xc potential is nonlinear,... one cannot expect that an ECP adjusted at the HF level performs necessarily well when doing a density functional calculation on the pseudosystem with a lesser electron number and an electron density resulting from the valence orbitals alone. On the other hand, much work has been invested in the

Received: April 6, 2011

Published: June 30, 2011

development of ECPs at the HF level.... Gaussian basis sets suitable for the expansion of valence orbitals have been developed together with the core potentials. It would be very desirable to profit from this work in KS calculations.” Indeed ECPs developed at the HF level and often validated only at the HF level are present in most popular quantum chemistry software and are widely used for KS calculations. However, there is only a handful of systematic work validating this.<sup>4–7</sup> Van Wüllen also pointed out that “To assess the validity of using ECPs, an ‘internal’ consistency check is more important than is a comparison with experiment.” He used large basis sets for his tests “to avoid any bias introduced by different quality basis sets in the all-electron and ECP calculations.” He found, in a systematic set of nonrelativistic calculations with large basis sets and small-core nonrelativistic ECPs (NRECPs), that ECP calculations agree with AE calculations on average to 0.012 Å for bond lengths and 0.9 kcal/mol for bond energies, whereas these values change to 0.007 Å and 2.5 kcal/mol for KS calculations, which he judged satisfactory. Han and Hira<sup>6</sup> extended this kind of systematic test to one-component relativistic ECPs (RECPs) and found “little loss of accuracy” in using ECPs with DFT if one uses flexible basis sets and appropriate contraction coefficients; in particular, they found accuracies of 0.001 Å and 1 kcal/mol for bond lengths and bond energies.

But questions remain: (1) Are the conclusions valid for a wider variety of systems? (2) Are the conclusions valid if one uses the basis sets recommended by the ECP developers as widely used with standard quantum chemistry programs rather than the large, flexible basis sets in the kind of test summarized above? (3) What new considerations arise when one compares calculations with RECPs to calculations with NRECPs for the same system? In this paper we explore these questions, in particular for systems containing arsenic.

Arsenic and its compounds are well-known for their toxicity for most life forms and their potent physiological properties, and they have been used medicinally for thousands of years.<sup>8</sup> Now, arsenic is under the spotlight due to the report<sup>9</sup> of a proteobacterium from Mono Lake that reportedly can substitute arsenic for phosphorus to sustain its growth. This finding is highly controversial,<sup>10–12</sup> and it will stimulate further experimental and theoretical investigations on arsenic’s role in living systems. Before one starts a computational study of the biological role or other properties of arsenic compounds, it is important to choose a suitable basis set and—to efficiently include relativistic effects on core orbitals and to keep the overall cost of the calculation for a given accuracy as low as possible—to choose an accurate RECP. Exploring these choices for arsenic is of interest not only for calculations on arsenic-containing systems but also for the study’s implications for other elements, especially those in the fourth through sixth periods of the periodic table.

Motivated by the discussion above, we focus in this paper on the choice of ECP for As, along with the choice of corresponding basis set, for density-functional electronic-structure calculations on arsenic-containing compounds. To sort out the special needs for DFT as compared to those of WFT, we carry out HF calculations as well as KS calculations. To be sure that our results for DFT are not biased by the choice of a particular xc functional, all conclusions about DFT calculations are averaged over 11 different xc functionals.

In the literature,<sup>13</sup> the most commonly used all-electron (AE) basis sets for the theoretical study of small arsenic compounds are the 6-311G triple- $\zeta$  basis set<sup>14</sup> with polarization functions and sometimes augmented with diffuse functions. Although less widely used for As, the popular correlation consistent basis sets

(cc-pVxZ)<sup>15</sup> by Dunning and the most recent collection of basis sets (def2-xVZP)<sup>16</sup> from the Ahlrichs group can be used to improve the accuracy of AE calculations by increasing the size of the basis set. However, the 6-311-type, cc-pVxZ, and def2 basis sets are AE basis sets and, as such, are inefficient for the simulation of large systems. In addition, arsenic belongs to the fourth period of the elements, for which the scalar relativistic effects have been found to be very important for some properties. Calculations including relativistic effects that use AE basis sets involve higher costs for accurate results, e.g., the CPU time can be much larger than that of nonrelativistic calculations at a comparable level for large systems. Therefore, using effective core potentials fitted to relativistic atomic calculations to replace the inner shell electrons of arsenic may be a good choice for large systems, because it allows one to incorporate the most important scalar relativistic effect (the changes in size of s and p core orbitals) in formally nonrelativistic (NR) calculations while simultaneously reducing the need for basis functions to expand the core orbitals. This raises another issue, namely, the number of core electrons to be replaced by an ECP. Large-core ECPs that replace 28 electrons have been employed for As,<sup>13e,17</sup> but medium-core (18 electrons) and small-core (10 electrons) ECPs are potentially better choices. The accuracy of all these choices needs to be investigated more systematically than has been done in previous work, especially in the context of DFT.

In the present work, the performance of popular basis sets and ECPs (large-core, medium-core, and small-core) for HF and KS calculations of equilibrium bond dissociation energies ( $D_e$ ) and vertical ionization potentials (VIP) of prototype As-containing compounds will be investigated. In particular, we consider the ionization potential (IP) of As; the  $D_e$  and VIP of As<sub>2</sub>, As<sub>3</sub>, and As<sub>4</sub>; the energy of the first acid dissociation of H<sub>3</sub>AsO<sub>4</sub> (i.e., the energy to remove a proton); and the  $D_e$  of FeAs.

The choice of FeAs is motivated in part by the recent finding that solid iron arsenide is a novel superconductor material.<sup>18</sup> Furthermore, as a 3d metal in the fourth period, the choice of basis set and RECP for Fe raises some of the same questions we have raised for As; therefore, by calculating the bond dissociation energy of the diatomic molecule FeAs, we test whether our choices of basis set and ECP for As are general, by examining whether we can use the same basis set choice for Fe as for As.

## 2. COMPUTATIONAL DETAILS

Although most xc functionals have similar basis set requirements, we need to recognize that the requirements do vary. Therefore, one strategy that we employ for many of our tests of basis sets and ECPs is to consider 11 xc functionals for a given property, for example, an equilibrium bond energy ( $D_e$ ) or vertical ionization potential (VIP), and for each functional we use a large-basis relativistic AE calculation with a basis set specifically optimized for relativistic calculations to obtain a reference value for that xc functional and property. For brevity, this reference value is labeled NCBS-DK (NCBS denotes “nearly complete basis set” and “DK” is explained below). Then, for a given smaller basis, employed in an AE or ECP calculation with the rest of the Hamiltonian either relativistic or nonrelativistic, we calculate a deviation from the relativistic NCBS-DK value; we will call this the complete error. The complete mean unsigned error (C-MUE) is the mean unsigned deviation of the 11 calculated values of the property, each from its NCBS-DK reference value. Similarly, we use a large-basis nonrelativistic AE calculation

with a basis set specifically optimized for nonrelativistic calculations to obtain a nearly complete basis set nonrelativistic reference value, labeled NCBS-NR, for each functional and that property. The mean unsigned deviation from the NCBS-NR reference value is calculated for the 11 calculated values of the property with a given smaller basis set, and this error measure is called the nonrelativistic mean unsigned error (NR-MUE). The NR-MUE is a measure of how well a nonrelativistic basis set approaches the nonrelativistic limit, whereas the C-MUE is a measure of how well the whole treatment approaches the complete basis set limit including relativistic effects.

Following the above protocol, for our first set of tests, the  $D_e$  of the diatomic molecules  $\text{As}_2$  ( $^1\Sigma_g^+$ ) and  $\text{FeAs}$  ( $^2\Delta$ ), the VIP of  $\text{As}_2$  ( $^1\Sigma_g^+$ ), the IP of As atom, and the energy of the first acid dissociation of  $\text{H}_3\text{AsO}_4 \rightarrow \text{H}_2\text{AsO}_4^- + \text{H}^+$  have been calculated using 11 popular and high-performance xc functionals (M05,<sup>19</sup> M06-L,<sup>20</sup> M06,<sup>21</sup> BLYP,<sup>22</sup>  $\omega$ B97X-D,<sup>23</sup>  $\tau$ HCTHhyb,<sup>24</sup> G96LYP,<sup>22b,25</sup> mPWLYP,<sup>22b,26</sup> B3LYP,<sup>22,27</sup> X3LYP,<sup>28</sup> and BPBE<sup>22a,29</sup>), with different basis sets and ECPs. For each given xc functional, we will compare our relativistic and nonrelativistic AE calculations and our nonrelativistic RECP and NRECP calculations to Douglas–Kroll–Hess second-order scalar relativistic calculation<sup>30</sup> (sometimes called DKH and sometimes called DK) results with the nearly complete basis set NCBS-DK, and we will also compare our nonrelativistic AE calculations to nonrelativistic results with the NCBS-NR basis set for a given xc functional. By looking at such comparisons, one can largely decouple the errors that are intrinsic to a given xc functional and treatment of relativistic effects from those resulting from the choice of basis set and ECP.

For the bond dissociation energy calculations of  $\text{As}_2$  and  $\text{FeAs}$  and the VIP calculations on  $\text{As}_2$  and IP calculations on the As atom, the cc-pV5Z-DK and cc-pV5Z basis sets<sup>15c,d</sup> have been chosen as NCBS-DK and NCBS-NR basis sets, respectively. For  $D_e$  of  $\text{As}_2$ , relativistic and nonrelativistic calculations with the uncontracted cc-pV5Z basis set were also performed to confirm that the cc-pV5Z-DK and cc-pV5Z basis sets are good enough to serve as the reference basis set in our investigations. There are no diffuse functions in the NCBS basis sets for these calculations, because they were previously found to be unimportant in most density functional calculations of bond dissociation energy and ionization energy of neutral molecules.<sup>31</sup> In the calculations of the energy of reaction of the first acid dissociation of  $\text{H}_3\text{AsO}_4$ , the minimally augmented cc-pV5Z-DK basis set (ma-cc-pV5Z-DK) is used as the NCBS-DK basis set, and the minimally augmented cc-pV5Z (ma-cc-pV5Z) is used as the NCBS-NR basis set. In these minimally augmented basis sets, a set of diffuse s and p functions has been added to the cc-pV5Z-DK or cc-pV5Z basis set for non-hydrogen elements, with the exponents of the most diffuse s or p functions of cc-pV5Z-DK or cc-pV5Z basis set divided by a factor of 3, as recommended previously.<sup>31</sup>

In addition, the relativistic effects of these five cases have been respectively calculated as the difference between the results of the relativistic DKH calculations and nonrelativistic calculations with a given AE basis set and the 11 xc functionals. The value of the relativistic effect on each property is calculated as the difference between the NCBS-DK and NCBS-NR values of that property with a given xc functional, and this is taken as the NCBS relativistic effect for that property and that functional. The NCBS relativistic effect is averaged over the eleven functionals to obtain a nominal reference value for relativistic effect of that property. For each property obtained with a given smaller AE basis set and a given xc functional, we calculate a deviation from the NCBS relativistic

effect of that property with that functional; the mean unsigned deviation from the NCBS relativistic effect, with the mean obtained by averaging over the absolute values of the deviations for the 11 xc functionals, is called the mean unsigned error (MUE) of the relativistic effect for that property and that AE basis set.

In this first set of tests, the tested AE basis sets are the def2 basis sets, with size increasing from double- $\zeta$  to quadruple- $\zeta$ ; the minimally augmented diffuse def2 basis sets<sup>31</sup> (which are called ma- $x$ ZVP basis sets); and the 6-311G basis set series with various polarization functions and in some case with diffuse functions. Note that “double- $\zeta$  valence” is labeled as “split valence” (SV) rather than DZV in the def2 series and ma series. The “def2” basis sets were developed as a second generation of default basis sets for the popular TURBOMOLE program, and they constitute balanced economical basis sets of graded quality from partially polarized (P) double- $\zeta$  to heavily polarized (PP) quadruple- $\zeta$  for all elements up to radon ( $Z = 86$ ). The P-type basis sets of the “def2” series, def2-SVP, def2-TZVP, and def2-QZVP, were recommended for DFT calculations.<sup>16</sup> In general, basis sets containing diffuse functions are called augmented. Recently, we<sup>31</sup> enhanced the def2 basis sets by adding a minimal set of diffuse functions to a subset of the elements for certain kind of calculations, and the new basis sets were labeled by “ma” (which stands for minimally augmented). The new ma-TZVP basis set has been suggested for general-purpose applications of DFT.<sup>31</sup> However, the tested elements in the previous study do not include any elements of the fourth period, for which the omission of scalar relativistic effects could be a significant omission. Therefore, the performance of the def2- $x$ ZVP and ma- $x$ ZVP basis sets for both nonrelativistic and relativistic DKH calculations is tested in the present study, and all results are compared with those obtained from relativistic DKH calculations or nonrelativistic calculations with NCBS basis sets. We also provide limited testing of the cc-pVTZ and cc-pVTZ-DK basis sets.<sup>15c</sup>

The tested ECP basis sets (i.e., basis sets to be used with ECPs) and ECPs are

- (i) cc-pV $x$ Z-PP ( $x = \text{D, T, Q, 5}$ ),<sup>32</sup> which are correlation-consistent-type basis sets used with the small-core (10 electrons) multiconfiguration-Dirac–Hartree–Fock (MDF) adjusted fully relativistic Stuttgart ECP, denoted MDF10;<sup>33</sup>
  - (ii) the MDF10 RECP with its own basis set;<sup>33</sup>
  - (iii) the CRENBL medium-core (18e) RECP<sup>34</sup> with its own basis set; and
  - (iv–vii) the large-core (28e) ECPs and corresponding ECP basis sets implemented in the Gaussian program.
- The large-core ECP basis sets and ECPs are
- (iv) the Wood–Boring quasirelativistic Stuttgart large-core ECP, labeled MWB28 (also sometimes called SDD);<sup>35</sup>
  - (v) the compact effective potential, labeled CEP,<sup>36</sup> which is an RECP [and which is also sometimes called the Stevens–Basch–Krauss–Jasien (SBKJ) potential];
  - (vi) the LANL2DZ<sup>37</sup> potential, which is a large-core NRECP for As and a small-core NRECP for Fe; and
  - (vii) the relatively new large-core (28e) MDF-adjusted fully relativistic Stuttgart ECP, denoted MDF28,<sup>1</sup> with its own basis set.

In addition, we also tested some combinations of modified and unmodified def2 and ma basis sets with the MDF10 RECP. The main motivation for choosing the def2 and ma basis sets for these



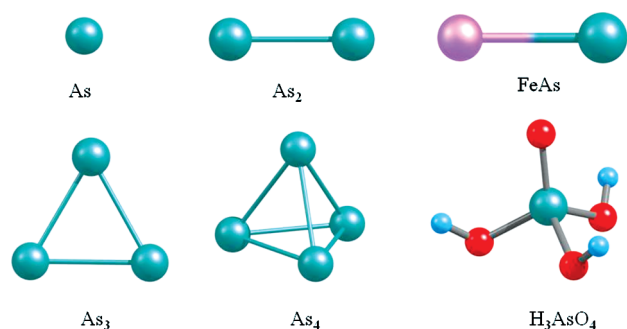


Figure 1. Structures.

combinations is that, as we already mentioned, the def2 basis sets are available for all elements up to radon.

For comparison to the ECP calculations of arsenic compounds with DFT, we also performed calculations with these ECPs by the HF method for four of the cases in the first test set:  $D_e$  and VIP of  $As_2$ , IP of As, and the first acid dissociation energy of  $H_3AsO_4$ . (The  $D_e$  of FeAs was excluded in the HF test in order to avoid complicating the comparison by having to consider transferability of ECPs for iron.) For each calculated HF property obtained with each combination of ECP and corresponding basis set, we computed the deviation from the NCBS-DK reference value of that property (obtained by relativistic HF calculations), and the mean unsigned deviation over these properties for each ECP and associated basis set was labeled MUE.

For the second set of tests, the performance of several selected basis sets and RECPs has been further tested by calculating VIPs and  $D_e$  values of arsenic clusters  $As_n$  ( $n = 3$  and 4). The NCBS-DK and NCBS-NR basis sets are again respectively taken as cc-pV5Z-DK and cc-pV5Z.

For the third set of tests, we carried out nonrelativistic cc-pVTZ<sup>15c</sup> calculations and relativistic cc-pVTZ-DK<sup>15c</sup> calculations for  $D_e$  and VIP of  $As_2$  and for the IP of As. These all-electron basis sets are de-emphasized because they are not of primary relevance to our goal of testing ECPs. Furthermore, we note that the cc-pVxZ and cc-pVxZ-DK basis sets employ general contraction schemes in which a primitive Gaussian is used in more than one contracted function. Although general contractions have some theoretical advantages over the segmented contraction schemes that the def2- $x$ ZVP, ma- $x$ ZVP, and 6-311G series basis sets use, efficient integral calculations with general contractions are not implemented in most popular ab initio packages, which effectively increases the number of primitive functions.

When we perform ECP calculations in this paper, the electrons treated explicitly are always treated nonrelativistically, although the ECP may be an RECP or an NRECP. We do not test explicitly relativistic calculations that also include effective core potentials.<sup>38</sup>

Taking  $As_2$  as an example, it has been confirmed that all 11 of the xc functionals considered here can predict reasonable geometrical structures, and the calculated properties, such as  $D_e$  values for  $As_2$ , are not very sensitive to reasonable choices of molecular geometries. (See Tables S1 and S2 of Supporting Information for this confirmation.) Therefore, in all calculations, for each molecule, the same reasonable geometry was used, which has the advantage that we are directly comparing electronic energies without the complication of using different geometries, and this strategy simplifies the interpretation of the results.

For  $As_2$ , the experimental bond length<sup>39</sup> of 2.1026 Å was used. The  $D_{3h}$  structure was used for arsenic trimer, as obtained from Igel-Mann et al.<sup>40</sup> with each bond length equal to 2.3284 Å. The experimental structure<sup>41</sup> was used for arsenic tetramer; it is a tetrahedron with each bond length equal to 2.4353 Å. The structures of FeAs,  $H_3AsO_4$ , and  $H_2AsO_4^-$  were optimized by nonrelativistic calculations with the M05 xc functional, with the def2-TZVP basis set for FeAs, and with the ma-TZVP basis set for  $H_3AsO_4$  and  $H_2AsO_4^-$ . The structures of the six molecules involved in the present article are shown in Figure 1.

All calculations are carried out using the Gaussian 09 electronic structure package.<sup>42</sup>

### 3. RESULTS AND DISCUSSION

**3.1. Performance of All-Electron Basis Sets.** Table 1 shows the mean unsigned errors (C-MUE and NR-MUE) for five cases ( $D_e$  and VIP of  $As_2$ , IP of As, and  $D_e$  of  $H_3AsO_4$  and FeAs) averaged over 11 xc functionals with the def2, ma, and 6-311G-type basis sets. The corresponding average of the C-MUEs or NR-MUEs (A-C-MUE or A-NR-MUE) over three or five cases for each basis set is also listed in Table 1. In addition, the mean unsigned deviations of  $D_e$  for  $As_2$  from the relativistic or nonrelativistic calculations with uncontracted cc-pV5Z basis set are also shown in Table 1. Comparing these results to the MUEs calculated with relativistic cc-pV5Z-DK or nonrelativistic cc-pV5Z results, we confirm that cc-pV5Z-DK and cc-pV5Z are reasonable reference basis set for As.

**3.1.1. def2 and ma Basis Sets.** **3.1.1.1. VIP of  $As_2$ ,  $D_e$  of  $As_2$ , and IP of As.** For the calculations of  $D_e$  of  $H_3AsO_4$  and FeAs, the basis set choice for other elements (H, O, or Fe) could have an impact on the conclusions. Therefore, we first analyze the calculated VIP of  $As_2$ ,  $D_e$  of  $As_2$ , and IP of As, based on their average C-MUE [A-C-MUE(3)] values, which are based on comparison to the NCBS-DK reference values.

As shown in rows 16 and 17 of Table 1, for nonrelativistic calculations, the def2-QZVP basis set has a similar, slightly lower A-C-MUE(3) value than the larger cc-pV5Z basis set (which is the NCBS-NR reference basis). The errors in the def2-TZVP nonrelativistic calculations (row 18) are comparable with those of def2-QZVP. The A-C-MUE(3) for the smaller def2-SVP basis set (row 20) is 0.4 kcal/mol larger than that of def2-TZVP. The def2-SVP basis set is worse for  $D_e$  values than for VIP and IP values. This indicates that the double split valence basis is not sufficiently flexible for a good description of the electron distribution of bonds; a triple- $\zeta$  or more accurate basis set is needed for calculations of bond dissociation energy.

When relativistic calculations are performed (rows 1–15), the average performance of both def2-QZVP and def2-SVP for the three cases is improved, while the results for the def2-TZVP basis set get worse. Examining Table 1 in more detail, we see that the complete mean unsigned errors, C-MUEs, of ionization potentials are decreased in relativistic calculations with all three def2 basis sets as compared to those of the nonrelativistic calculations with the same basis sets, while for  $D_e$  calculations on  $As_2$ , only the def2-QZVP results are improved. That is because the relativistic effects will contract the inner core and cause a more efficient screening of the nuclear charge for the outer shells, thus leading (in this case) to an expansion of the valence orbitals.<sup>43</sup> Although the shape changes of the valence orbitals are weak, the important role of valence orbitals in molecular bonding apparently requires that, to get more accurate  $D_e$  values in relativistic calculations,

**Table 1.** Mean Unsigned Errors (C-MUE and NR-MUE, kcal/mol) in All-Electron Calculations of  $D_e$  and VIP of  $\text{As}_2$  ( $^1\Sigma_g^+$ ), IP of As,  $D_e$  of  $\text{H}_3\text{AsO}_4$  ( $\text{H}_3\text{AsO}_4 \rightarrow \text{H}_2\text{AsO}_4^- + \text{H}^+$ ), and  $D_e$  of  $\text{FeAs}$  ( $^2\Delta$ ) Averaged over 11 xc Functionals, Based on Comparisons to the NCBS-DK Reference Value and NCBS-NR Reference Value of the Given Functional

basis set	type	C-MUE						
		$D_e$ , $\text{As}_2$	VIP, $\text{As}_2$	IP, As	$D_e$ , $\text{H}_3\text{AsO}_4^a$	$D_e$ , FeAs	A-C-MUE(3) <sup>b</sup>	A-C-MUE(5) <sup>a,c</sup>
NCBS-DK <sup>d</sup>	DK	0/0.17 <sup>e</sup>	0	0	0	0	0	0
def2-QZVP	DK	0.55/0.69 <sup>e</sup>	0.15	0.16	1.04	0.21	0.29	0.42
def2-TZVP	DK	2.63/2.75 <sup>e</sup>	0.79	0.58	2.01	0.23	1.33	1.25
def2-TZVPxf	DK	0.35/0.52 <sup>e</sup>				0.40 <sup>f</sup>		
def2-SVP	DK	2.11/2.11 <sup>e</sup>	0.84	0.50	10.53	3.80	1.15	3.56
def2-SVP+f	DK	1.80/1.84 <sup>e</sup>						
ma-QZVP	DK	0.45/0.58 <sup>e</sup>	0.22	0.21	0.07	0.13	0.29	0.22
ma-TZVP	DK	2.14/2.27 <sup>e</sup>	1.27	0.76	1.58	0.35	1.39	1.22
ma-SVP	DK	2.75/2.62 <sup>e</sup>	1.64	1.13	5.33	2.81	1.84	2.73
6-311++G(3df,3pd)	DK	1.17/1.33 <sup>e</sup>	0.60	0.44	0.78	0.60	0.74	0.72
6-311+G(2d,p)	DK	1.23/1.14 <sup>e</sup>	0.92	0.47	2.87	0.26	0.87	1.15
6-311+G(d,p)	DK	1.75/1.65 <sup>e</sup>	0.99	0.43	4.39	0.30	1.06	1.57
6-311G(3df,3pd)	DK	1.35/1.51 <sup>e</sup>	0.46	0.37	5.67	2.91	0.73	2.15
6-311G(2d,p)	DK	1.21/1.11 <sup>e</sup>	0.65	0.40	5.32	5.03	0.75	2.52
6-311G(d,p)	DK	1.86/1.76 <sup>e</sup>	0.76	0.35	5.79	8.02	0.99	3.36
NCBS-NR	NR	0.69/0.78 <sup>e</sup>	1.42	0.82	0.34	2.28	0.98	1.11
def2-QZVP	NR	0.70/0.80 <sup>e</sup>	1.37	0.81	1.32[0.99]	2.14	0.96	1.27 [1.20]
def2-TZVP	NR	0.78/0.94 <sup>e</sup>	1.56	0.85	2.27[1.25]	2.00	1.06	1.49 [1.29]
def2-TZVPxf	NR	1.58/1.42 <sup>e</sup>				1.65 <sup>g</sup>		
def2-SVP	NR	1.67/1.62 <sup>e</sup>	1.97	0.73	10.72[1.53]	1.87	1.46	3.39 [1.55]
def2-SVP+f	NR	1.50/1.61 <sup>e</sup>						
ma-QZVP	NR	0.79/0.88 <sup>e</sup>	1.46	0.88	0.24	2.10	1.04	1.09
ma-TZVP	NR	0.88/1.04 <sup>e</sup>	2.15	1.07	1.27	2.13	1.37	1.50
ma-SVP	NR	2.34/2.22 <sup>e</sup>	3.12	1.50	5.03	1.14	2.32	2.63
6-311++G(3df,3pd)	NR	1.24/1.41 <sup>e</sup>	1.81	0.75	0.53	2.66	1.27	1.40
6-311+G(2d,p)	NR	1.43/1.33 <sup>e</sup>	2.17	0.63	2.57	2.19	1.41	1.80
6-311+G(d,p)	NR	1.85/1.75 <sup>e</sup>	2.21	0.80	4.11	1.82	1.62	2.16
6-311G(3df,3pd)	NR	1.35/1.52 <sup>e</sup>	1.62	0.61	5.94	2.14	1.19	2.33
6-311G(2d,p)	NR	1.44/1.34 <sup>e</sup>	1.81	0.48	5.55	2.86	1.24	2.43
6-311G(d,p)	NR	2.05/1.92 <sup>e</sup>	1.93	0.66	6.05	5.55	1.55	3.25

basis set	type	NR-MUE						
		$D_e$ , $\text{As}_2$	VIP, $\text{As}_2$	IP, As	$D_e$ , $\text{H}_3\text{AsO}_4^a$	$D_e$ , FeAs	A-NR-MUE(3) <sup>b</sup>	A-NR-MUE(5) <sup>a,c</sup>
NCBS-NR <sup>h</sup>	NR	0/0.17 <sup>i</sup>	0	0	0	0	0	0
def2-QZVP	NR	0.15/0.19 <sup>j</sup>	0.07	0.07	0.98[1.33]	0.18	0.10	0.29 [0.36]
def2-TZVP	NR	0.51/0.65 <sup>i</sup>	0.54	0.36	1.99[1.59]	0.42	0.47	0.76 [0.68]
def2-SVP	NR	1.57/1.44 <sup>i</sup>	0.94	0.65	10.38[1.87]	4.14	1.05	3.54 [1.83]
ma-QZVP	NR	0.14/0.24 <sup>i</sup>	0.13	0.10	0.11	0.20	0.12	0.14
ma-TZVP	NR	0.47/0.60 <sup>i</sup>	0.73	0.33	1.61	0.28	0.51	0.68
ma-SVP	NR	2.27/2.13 <sup>i</sup>	1.70	0.77	5.38	3.18	1.58	2.66
6-311++G(3df,3pd)	NR	0.87/1.03 <sup>i</sup>	0.44	0.37	0.84	0.40	0.56	0.58
6-311+G(2d,p)	NR	1.46/1.36 <sup>i</sup>	0.77	0.40	2.91	0.18	0.88	1.14
6-311+G(d,p)	NR	1.89/1.76 <sup>i</sup>	0.84	0.36	4.45	0.46	1.03	1.60
6-311G(3df,3pd)	NR	1.00/1.17 <sup>i</sup>	0.34	0.33	5.60	3.76	0.56	2.21
6-311G(2d,p)	NR	1.47/1.37 <sup>i</sup>	0.52	0.40	5.71	5.00	0.80	2.62
6-311G(d,p)	NR	2.01/1.91 <sup>i</sup>	0.63	0.31	5.21	7.83	0.98	3.20

<sup>a</sup> The values in square brackets are obtained using the ma-TZVP basis set for O and the def2-TZVP basis set for H. <sup>b</sup> A-C-MUE(3) [A-NR-MUE(3)] is the average of C-MUE (NR-MUE) for three cases (VIP and  $D_e$  of  $\text{As}_2$ , IP of As). <sup>c</sup> A-C-MUE(5) [A-NR-MUE(5)] is the average of C-MUE (NR-MUE) for all five cases. <sup>d</sup> The NCBS-DK reference value is the relativistic result with the cc-pVSZ-DK basis set for  $D_e$  and VIP of  $\text{As}_2$ , IP of As, and  $D_e$  of FeAs, and it is the relativistic result with ma-cc-pVSZ-DK for  $D_e$  of  $\text{H}_3\text{AsO}_4$ . <sup>e</sup> For comparison, the mean unsigned deviation of  $D_e$  of  $\text{As}_2$  from the relativistic result with uncontracted cc-pVSZ-DK basis set is shown after the slash. <sup>f</sup> The def2-TZVPxf on As and def2-TZVP on Fe yields 0.38. <sup>g</sup> The def2-TZVPxf on As and def2-TZVP on Fe yields 1.71. <sup>h</sup> The NCBS-NR reference value is the nonrelativistic result with the cc-pVSZ basis set for  $D_e$  and VIP of  $\text{As}_2$ , IP of As, and  $D_e$  of FeAs, and it is the nonrelativistic result with ma-cc-pVSZ for  $D_e$  of  $\text{H}_3\text{AsO}_4$ . <sup>i</sup> For comparison, the mean unsigned deviation of  $D_e$  of  $\text{As}_2$  from the nonrelativistic result with uncontracted cc-pVSZ basis set is shown after the slash.

Table 2. Mean Unsigned Error (MUE, kcal/mol) in the Value of the Relativistic Effect<sup>a</sup> Averaged over 11 xc Functionals

basis set	MUE					
	$D_e$ , As <sub>2</sub> <sup>b</sup>	VIP, As <sub>2</sub>	IP, As	$D_e$ , H <sub>3</sub> AsO <sub>4</sub>	$D_e$ , FeAs	AMUE(S)
def2-QZVP	0.58	0.15	0.14	0.06	0.35	0.26
def2-TZVP/def2-TZVPxf	2.11/1.45	0.60	0.55	0.03	0.53	0.76
def2-SVP/def2-SVP+f	0.82/0.82	0.32	0.36	0.15	0.96	0.52
ma-QZVP	0.50	0.14	0.14	0.05	0.29	0.22
ma-TZVP	1.67	0.53	0.51	0.03	0.50	0.65
ma-SVP	0.78	0.10	0.46	0.05	0.90	0.46
6-311++G(3df,3pd)	0.40	0.18	0.22	0.06	0.20	0.21
6-311+G(2d,p)	0.17	0.41	0.21	0.04	0.23	0.21
6-311G(d,p)	0.37	0.16	0.22	0.06	0.19	0.20
6-311G(3df,3pd)	0.43	0.16	0.20	0.08	1.27	0.43
6-311G(2d,p)	0.43	0.15	0.19	0.10	1.27	0.43
6-311G(d,p)	0.41	0.14	0.21	0.08	1.26	0.42

<sup>a</sup> The relativistic effect is calculated as the difference between the value obtained in an all-electron relativistic DKH calculation and the value obtained in an all-electron nonrelativistic calculation with the same basis set. The error is the deviation from the NCBS relativistic effect, which is obtained with two different NCBS basis sets, one for the relativistic calculation and one for the nonrelativistic calculation; in particular, the NCBS relativistic effect for a given property and a given xc functional is the difference between the relativistic DKH result using the cc-pVSZ-DK basis set (ma-cc-pVSZ-DK for H<sub>3</sub>AsO<sub>4</sub>) and the nonrelativistic result with the cc-pVSZ basis set (ma-cc-pVSZ for H<sub>3</sub>AsO<sub>4</sub>). <sup>b</sup> The values after the slash are obtained with def2-TZVPxf or def2-SVP+f basis sets.

one must use valence basis functions specifically optimized for relativistic calculations or use relatively complete valence basis sets that can accommodate the change of shape. The parameters of the def2 basis sets were obtained from nonrelativistic calculations; therefore, the incomplete def2-TZVP and def2-SVP basis sets cannot accurately represent the valence orbitals in relativistic calculations. At the same time, we notice that the relativistic def2-SVP results are relatively better than those of def2-TZVP, and they have a 0.5 kcal/mol lower C-MUE for  $D_e$  of As<sub>2</sub>. This is probably due to cancellation of errors in def2-SVP relativistic calculations.

The relativistic def2-TZVP calculations overestimate the  $D_e$  values; this could be an indication that the def2-TZVP basis set is overpolarized for relativistic valence orbitals. Therefore, we deleted the f polarization subshell of def2-TZVP to get what we call def2-TZVPxf basis set, where “xf” denotes excluding f functions. The nonrelativistic def2-TZVPxf calculations give a larger C-MUE for  $D_e$  of As<sub>2</sub> than does def2-TZVP, but the relativistic def2-TZVPxf calculations have improved performance compared to def2-TZVP, even better than def2-QZVP. This confirms that the def2-TZVP basis set is overpolarized for relativistic valence orbitals. On the contrary, the f subshell added to def2-SVP basis set slightly improves the relativistic results of def2-SVP.

Previous investigations showed that the diffuse functions are usually not important for  $D_e$  and IP calculations by DFT.<sup>31</sup> Due to the relative completeness of the def2-QZVP basis set, the ma-QZVP basis has similar A-C-MUE(3) values to those of the def2-QZVP basis set for both relativistic and nonrelativistic calculations. The ma-TZVP and ma-SVP basis sets perform worse than the correspondingly sized def2 basis sets, especially for ma-SVP. This implies that augmented small basis sets might be unbalanced for  $D_e$  and IP calculations.

**3.1.1.2.  $D_e$  of H<sub>3</sub>AsO<sub>4</sub>.** Table 1 shows that, when the same basis sets are used for all the atoms of H<sub>3</sub>AsO<sub>4</sub>, the C-MUEs of the def2-TZVP and def2-SVP basis sets are respectively 2 or 10 times larger than that of def2-QZVP, for both relativistic and

nonrelativistic calculations of the energy of acid dissociation of H<sub>3</sub>AsO<sub>4</sub> (which is denoted as  $D_e$  even though we are considering dissociation to ions, i.e., proton detachment). Due to the small relativistic effect on  $D_e$  of H<sub>3</sub>AsO<sub>4</sub> (the relativistic effect will be discussed in section 3.2), the relativistic and nonrelativistic calculations with def2 basis sets have similar C-MUE values.

The products of the first acid dissociation reaction of H<sub>3</sub>AsO<sub>4</sub> are a proton and H<sub>2</sub>AsO<sub>4</sub><sup>−</sup>. Because a negative ion is involved, diffuse functions play an important role in getting an accurate  $D_e$  value for H<sub>3</sub>AsO<sub>4</sub>. As shown in Table 1, the diffuse ma basis sets improve the calculated results for  $D_e$  of H<sub>3</sub>AsO<sub>4</sub> remarkably. The C-MUE of the relativistic ma-QZVP results is only 0.07 kcal/mol, showing that ma-QZVP is very close to the complete basis set. The improvement due to adding diffuse functions is most dramatic for the double- $\zeta$  basis set; the C-MUE of ma-SVP is 5 kcal/mol lower than that of the def2-SVP basis set for both relativistic and nonrelativistic calculations.

In the calculations of  $D_e$  of H<sub>3</sub>AsO<sub>4</sub>, the choice of basis sets for oxygen and hydrogen could also affect the accuracy of the results, especially for oxygen, because the negative charge is mainly distributed on oxygen atoms. However, here we are concerned more with the choice of basis set for As. Therefore, we tested the performance of different sizes of def2 basis sets for As in nonrelativistic calculations in which the ma-TZVP basis set is used for O and H. In contrast to the large improvement discussed in the previous paragraph, when the same basis set is used for all elements, only ~0.3 kcal/mol of improvement is observed when the size of basis set increases from double- $\zeta$  to triple- $\zeta$  or from triple- $\zeta$  to quadruple- $\zeta$ . Hence, the basis set of As is less important than that of oxygen in the acid dissociation reaction of H<sub>3</sub>AsO<sub>4</sub>, and diffuse functions are not needed for As but are very important for O; using def2-TZVP or using ma-TZVP for As gives almost the same MUE (C-MUE or NR-MUE) values if the ma-TZVP basis set is used for oxygen in both cases.

**3.1.1.3.  $D_e$  of FeAs.** The relativistic def2-QZVP and def2-TZVP calculations have similar C-MUE values, which are very small (~0.2 kcal/mol), for  $D_e$  of FeAs. The nonrelativistic

Table 3. Details of the Basis Sets and ECPs for As

basis set	contracted functions	numbers of basis functions	numbers of primitive Gaussians	ECP	core <sup>a</sup>	<i>l</i> <sup>b</sup>
cc-pVSZ-DK(cc-pVSZ)	8s,7p,5d,3f,2g,1h	104	309 (306)			
def2-QZVP	11s,7p,4d,2f,1g	75	179			
def2-TZVP	6s,5p,4d,1f	48	114			
def2-SVP	5s,4p,3d	32	80			
ma-QZVP	12s,8p,4d,2f,1g	79	183			
ma-TZVP	7s,6p,4d,1f	52	118			
ma-SVP	6s,5p,3d	36	84			
6-311++G(3df,3pd)	9s,8p,5d,1f	65	116			
6-311+G(2d,p)	9s,8p,4d	53	100			
6-311+G(d,p)	9s,8p,3d	48	94			
6-311G(3df,3pd)	8s,7p,5d,1f	61	112			
6-311G(2d,p)	8s,7p,4d	49	96			
6-311G(d,p)	8s,7p,3d	44	90			
cc-pVSZ-PP	7s,6p,5d,3f,2g,1h	100	236	MDF10	10	0–3
cc-pVQZ-PP	6s,5p,4d,2f,1g	64	178	MDF10	10	0–3
cc-pVTZ-PP	5s,4p,3d,1f	39	133	MDF10	10	0–3
cc-pVDZ-PP	4s,3p,2d	23	87	MDF10	10	0–3
MDF10	6s,6p,4d	44	126	MDF10	10	0–3
CRENBL	3s,3p,4d	32	36	CRENBL	18	0–∞
SDD	2s,2p	8	16	MWB28	28	0–3
MDF28	4s,4p	16	24	MDF28	28	0–3
CEP	2s,2p	8	20	CEP	28	0–∞
LANL2DZ	2s,2p	8	12	LANL2DZ	28	0–∞
ma-sc-SVP	5s,5p,3d	35	78	MDF10	10	0–3

<sup>a</sup> This column shows the number of core electrons, that is, the number of electrons replaced by the ECP. <sup>b</sup> This column shows the range of *l* in the ECPs.

calculations with the two basis sets give similar C-MUE, which are  $\sim 1.8$ – $1.9$  kcal/mol higher than those of the relativistic calculations. It seems that the relativistic def2-TZVP calculation is good enough for calculating the  $D_e$  of FeAs. However, based on our analysis in the previous section, the def2-TZVP basis set is overpolarized for relativistic valence orbitals of As. Hence, the good results of the relativistic def2-TZVP could be fortuitous. There is only one  $\sigma$  bond in the  $^2\Delta$  state of FeAs, and it is formed by a 4p orbital of As and the 4s orbital of Fe. Table 1 shows that f functions have only a small effect on  $D_e$  for FeAs, which indicates that the  $\sigma$  bond is not strongly affected by the polarization afforded by f functions.

The nonrelativistic calculations of def2-SVP have a lower C-MUE value than those of def2-QZVP and def2-TZVP, as a result of cancellation of errors. The relative incompleteness of def2-SVP can be confirmed from the NR-MUE of nonrelativistic calculations with def2-SVP and from C-MUE for relativistic def2-SVP calculations, which are shown in Table 1.

On the basis of the C-MUE and NR-MUE values of all calculations with ma basis sets for  $D_e$  of FeAs, the diffuse functions are not needed for the larger def2 basis set, but they can improve the results of def2-SVP.

**3.1.1.4. The Overall Performance of def2 and ma Basis Sets for the Five Cases.** Consider the overall performance for the five cases, which can be evaluated through the A-C-MUE(S) in Table 1. The relativistic ma-QZVP calculations give the best results, and they have a 0.2 kcal/mol of improvement compared to the relativistic def2-QZVP results. The relativistic calculations do not systematically increase the accuracy of def2-TZVP and def2-SVP basis sets for As.

**3.1.2. The 6-311G Series of Basis Sets.** Although they are the most commonly used basis sets in studies of arsenic compounds, when the relativistic effect is ignored, all 6-311G-type basis sets have worse performance than def2-TZVP for VIP and  $D_e$  of As<sub>2</sub>; however, they have a slightly better performance for the IP of As. Except for the 6-311++G(3df,3pd) calculation of  $D_e$  for H<sub>3</sub>AsO<sub>4</sub> and the 6-311+G(d,p) calculation of  $D_e$  for FeAs, nonrelativistic calculations with basis sets of the 6-311G series also perform worse than def2-TZVP for  $D_e$  of H<sub>3</sub>AsO<sub>4</sub> and FeAs. Relativistic calculations improve the accuracy of 6-311G-type basis sets for the VIP and  $D_e$  of As<sub>2</sub> and IP of As, but they have less effect on  $D_e$  of H<sub>3</sub>AsO<sub>4</sub>. For  $D_e$  of FeAs, only augmented 6-311G-type basis sets, such as 6-311++G(3df,3pd) [which is the same as 6-311+G(3df) for FeAs], 6-311+G(2d,p) [which is the same as 6-311+G(2d) for FeAs], and 6-311+G(d,p) [which is the same as 6-311+G(d) for FeAs], can give results comparable to def2-TZVP, and the others perform very badly. The p functions of the 6-311G basis set for iron are too tight (the smallest p exponent is 0.592684) compared with the p functions of the def2 series (the smallest p exponent is 0.134915 for def2-SVP and def2-TZVP and 0.028000 for def2-QZVP). Therefore, adding diffuse functions is helpful for  $D_e$  calculations on FeAs with 6-311G-type basis sets. This is in agreement with the recommendation of Raghavachari and Trucks<sup>14e</sup> that diffuse functions are needed when using 6-311G-type (Wachters–Hay) basis sets<sup>14a,b</sup> for first row transition metals. As shown in Table 1, based on the overall results for the five tested cases, only the very large 6-311++G(3df,3pd) basis set and 6-311+G(2d,p) basis set (with relativistic treatment) perform better than the smaller def2-TZVP basis set.



**Table 4.** Complete Mean Unsigned Errors (C-MUE, kcal/mol) for  $D_e$  and VIP of  $\text{As}_2$  ( $^1\Sigma_g^+$ ), IP of As,  $D_e$  of  $\text{H}_3\text{AsO}_4$  ( $\text{H}_3\text{AsO}_4 \rightarrow \text{H}_2\text{AsO}_4^- + \text{H}^+$ ), and  $D_e$  of FeAs ( $^2\Delta$ ) Averaged over 11 xc Functionals, with Standard ECPs with Individual Optimized Valence Basis Sets, Compared to the NCBS-DK Reference Value<sup>a</sup> for the Given Functional

basis set	ECP for As	C-MUE							
		$D_e$ , $\text{As}_2$	VIP, $\text{As}_2$	IP, As	$D_e$ , $\text{H}_3\text{AsO}_4^b$	$D_e$ , FeAs	A-C-MUE(3) <sup>c</sup>	A-C-MUE(4) <sup>b,d</sup>	A-C-MUE(5) <sup>b,e</sup>
cc-pV5Z-PP	MDF10	0.57/0.44	0.60	0.61	0.75	2.35	0.59	0.63	0.98
cc-pVQZ-PP	MDF10	0.70/0.57	0.69	0.68	2.32	2.51	0.69	1.10	1.38
cc-pVTZ-PP	MDF10	0.74/0.62	0.36	0.71	5.49 (0.87) [1.44]	2.80	0.60	1.83 (0.67) [0.81]	2.02 (1.10) [1.21]
cc-pwCVTZ-PP	MDF10	0.66/0.53	0.42	0.77			0.62		
cc-pVDZ-PP	MDF10	3.64/5.51	1.03	1.13	10.82	3.52	1.93	4.16	4.03
MDF10	MDF10	2.69/2.56	0.21	0.58	0.65 <sup>f</sup> [1.91]	2.78	1.16	1.03 <sup>f</sup> [1.35]	1.38 <sup>f</sup> [1.63]
CRENBL	CRENBL	20.84/21.35	6.95	4.19	9.37	8.30	10.66	10.34	9.93
SDD	MWB28	21.48/20.71	8.65	4.32	7.91 [1.74]	6.43	11.48	10.59 [9.05]	9.76 [8.52]
MDF28	MDF28	23.92/23.79	4.62	3.90	[1.57]	9.37	10.81	[8.50]	[8.68]
CEP	CEP	25.12/24.99	9.67	2.99	13.83	7.73	12.59	12.90	11.87
LANL2DZ	LANL2DZ	30.10/29.97	11.54	4.35	6.23 [1.52]	2.87	15.33	13.06 [11.88]	11.02

<sup>a</sup> The NCBS-DK reference value is the relativistic result with the cc-pV5Z-DK basis set for  $D_e$  and VIP of  $\text{As}_2$ , IP of As, and  $D_e$  of FeAs, and it is the relativistic result with ma-cc-pV5Z-DK for  $D_e$  of  $\text{H}_3\text{AsO}_4$ . For comparison, the mean unsigned deviation of  $D_e$  of  $\text{As}_2$  from the relativistic result with uncontracted cc-pV5Z-DK basis set is shown after the slash. <sup>b</sup> The values in parentheses are obtained using the ma-cc-pVTZ-PP basis set for As and the ma-cc-pVTZ basis set for O and H; the values in square brackets are obtained using the ma-TZVP basis set for O and H. <sup>c</sup> The A-C-MUE(3) is the average of C-MUE for VIP and  $D_e$  of  $\text{As}_2$  and IP of As. <sup>d</sup> The A-C-MUE(4) is the average of C-MUE for VIP and  $D_e$  of  $\text{As}_2$ , IP of As, and  $D_e$  of  $\text{H}_3\text{AsO}_4$ . <sup>e</sup> The A-C-MUE(5) is the average of C-MUE for VIP and  $D_e$  of  $\text{As}_2$ , IP of As,  $D_e$  of  $\text{H}_3\text{AsO}_4$ , and  $D_e$  of FeAs. <sup>f</sup> The MDF10 basis set is used for As; the aug-cc-pVTZ basis set is used for O and H.

**3.2. Relativistic Effects.** So far we have discussed both relativistic and nonrelativistic calculations but we have not discussed the relativistic effect itself, that is, the difference between the two kinds of calculation. To estimate the magnitude of the relativistic effect, we calculated this difference with each of the eleven xc functionals and averaged the eleven results. These calculations show that the average absolute value of relativistic effect, calculated with the two NCBS basis sets, is respectively 1.42, 0.69, 0.82, 0.34, and 2.28 kcal/mol for VIP of  $\text{As}_2$ ,  $D_e$  of  $\text{As}_2$ , IP of As,  $D_e$  of  $\text{H}_3\text{AsO}_4$ , and  $D_e$  of FeAs. For each small AE basis set, the mean unsigned deviation (MUE) of the relativistic effects obtained with eleven xc functionals from the NCBS relativistic effects and their average MUE over the five cases (AMUE(5)) are summarized in Table 2.

As shown in Table 2, the def2-TZVP has the worst performance of all tested basis sets for calculating relativistic effects, especially for  $D_e$  of  $\text{As}_2$ , as a result of overpolarized valence orbitals in relativistic calculations. The 6-311G series of basis sets is more appropriate for describing relativistic effects than def2-TZVP and def2-SVP in most of the cases, although the 6-311G series without diffuse functions fails in the case of  $D_e$  for FeAs. Table 2 shows that diffuse functions are usually beneficial for calculations of relativistic effects. Therefore, ma basis sets perform better than the corresponding def2 basis sets. Augmented 6-311G basis sets and the ma-QZVP basis set have the smallest average mean unsigned error for calculating the relativistic effect itself.

The present investigations show that the relativistic effect is usually not very large for arsenic compounds, but it is more than 2 kcal/mol in one case. Unless there is cancellation of errors, all nonrelativistic calculations have an error equal at least to the relativistic effect; therefore, the error of nonrelativistic calculations with the best AE basis sets could be more than 2 kcal/mol. Although def2-TZVP (or ma-TZVP when diffuse functions are needed) is a good basis set for nonrelativistic studies, it fails in relativistic calculations, as we have seen. Only the relativistic calculations with the larger basis sets, def2-QZVP, ma-QZVP,

6-311+G(2d,p), or 6-311++G(3df,3pd), can give accurate results for describing the properties of arsenic compounds. However, for large systems, such basis sets could be unaffordable; even nonrelativistic def2-TZVP calculations could be hard. A good ECP (especially an RECP) and corresponding valence basis set is potentially a better choice for large systems, and we will study this next.

**3.2.1. DFT Calculations with Standard ECP Basis Sets.** Next we consider the use of ECPs in DFT calculations. Table 3 compares the size of the AE basis sets considered in this work to basis sets used in combination with ECPs. Whereas AE basis sets for As involve 32–104 contracted basis functions (80–309 primitive Gaussians), with ECPs we can reduce this to as few as eight basis functions (12 primitive Gaussians) for a large-sized core or 32 basis functions (36 primitive Gaussians) for a medium-sized core; ECP basis sets for a small core have 23–100 contracted basis functions (78–236 primitive Gaussians). The computational effort of DFT calculations usually scales as  $N^3$  or  $N^4$  as the number of atoms  $N$  in the system is increased, while HF calculations scale as  $N^4$ , and these scaling laws may be combined with the numbers of basis functions and primitive Gaussians in Table 3 for rough estimates of relative computational savings when using ECPs. (Specific timing examples will be given in section 3.6.) We should keep in mind though that the purpose of an ECP is not just to decrease the number of basis functions. As mentioned in the Introduction, another purpose that is important to us is to introduce scalar relativistic effects by using relativistic ECPs without requiring a fully relativistic calculation.

In general, the ECP on an atom is written as<sup>44</sup>

$$U = U_{L+1}(r) + \sum_{l=0}^L \sum_{m=-l}^l [U_l(r) - U_{L+1}(r)] |lm\rangle \langle lm| \quad (1)$$

where  $r$  is the distance from the nucleus, and  $|lm\rangle$  is a spherical harmonic. When  $U_{L+1}(r) \neq 0$ , the ECP affects all symmetries ( $l = 0 - \infty$ ), whereas when  $U_{L+1}(r) = 0$ , it affects only  $l = 0 - L$ .



**Table 5. Complete Mean Unsigned Errors (C-MUE, kcal/mol), Averaged over 11 xc Functionals, Using the Relativistic Small-Core RECP (MDF10) Combined with Different Valence Basis Sets, Compared to the NCBS-DK Reference Value<sup>a</sup> for the Given Functionals**

basis set	RECP	C-MUE							
		$D_e$ , As <sub>2</sub> <sup>b</sup>	VIP, As <sub>2</sub>	IP, As	$D_e$ , H <sub>3</sub> AsO <sub>4</sub> <sup>c</sup>	$D_e$ , FeAs	A-C-MUE(3) <sup>d</sup>	A-C-MUE(4) <sup>e,f</sup>	A-C-MUE(5) <sup>e,f</sup>
ma-sc-SVP	MDF10	0.77 (0.81)	1.59	2.07	4.58 [0.80]	1.42	1.48	2.25 [1.31]	2.09 [1.33]
ma-SVP	MDF10	1.11 (1.06)	0.66	3.04	4.70	1.67	1.60	2.38	2.24
def2-SVP	MDF10	0.90 (1.07)	1.07	3.72	11.09 [0.93]	2.99	1.90	4.20 [1.66]	3.95 [1.92]
def2-TZVP/def2-TZVPxf	MDF10	3.30/1.99 (3.43/1.86)	4.05	4.60	2.67 [1.01]	9.11	3.98/3.55	3.66 [3.24]	4.75 [4.41]
def2-QZVP	MDF10	0.95 (1.09)	4.12	4.43	2.01 [0.95]	2.29	3.17	2.88 [2.61]	2.76 [2.55]

<sup>a</sup> The NCBS-DK reference value is the relativistic result with the cc-pVSZ-DK basis set for  $D_e$  and VIP of As<sub>2</sub>, IP of As, and  $D_e$  of FeAs, and it is the relativistic result with ma-cc-pVSZ-DK for  $D_e$  of H<sub>3</sub>AsO<sub>4</sub>. For comparison, the mean unsigned deviation of  $D_e$  of As<sub>2</sub> from the relativistic result with uncontracted cc-pVSZ-DK basis set is shown in parentheses. <sup>b</sup> The values after the slash are obtained with def2-TZVPxf as valence basis set. <sup>c</sup> The corresponding ma-SVP or def2-*x*ZVP basis sets are used for O and H. The values in square brackets are obtained using the ma-TZVP basis set for O and H. <sup>d</sup> The A-C-MUE(3) is the average of C-MUE for VIP and  $D_e$  of As<sub>2</sub> and IP of As. <sup>e</sup> The A-C-MUE(4) is the average of C-MUE for VIP and  $D_e$  of As<sub>2</sub>, IP of As, and  $D_e$  of H<sub>3</sub>AsO<sub>4</sub>. <sup>f</sup> The A-C-MUE(5) is the average of C-MUE for all five cases.

The last column of Table 3 shows the range of  $l$  in all of the ECPs considered in this work.

The complete mean unsigned errors (C-MUE) of calculations with the standard combinations of ECP basis sets and ECPs listed in Table 3 are given in Table 4, where, as in earlier sections, the mean is over 11 xc functionals. Since these errors are with respect to the relativistic NCBS-DK (cc-pVSZ-DK or ma-cc-pVSZ-DK) results, they include both basis set incompleteness and any errors due to the ECP being an inexact substitute for an explicit relativistic core.

As shown in Table 4, the results obtained with cc-pV $x$ Z-PP ( $x \geq T$ ) basis sets are in good agreement with those of the relativistic NCBS-DK (cc-pVSZ-DK) basis set, with an average C-MUE [A-C-MUE(3)] of less than 0.7 kcal/mol for VIP and  $D_e$  of As<sub>2</sub> and IP of As, better than either the relativistic or nonrelativistic calculations with any of the AE triple- $\zeta$  basis sets and even better than nonrelativistic calculations with quadruple- and quintuple- $\zeta$  sets. The MDF10 RECP with its own basis set is also reasonably accurate with an A-C-MUE(3) of 1.2 kcal/mol. The large-core and medium-core ECPs usually are not good choices, because they could lead to misestimating of valence electron correlation and polarization energies compared to AE calculations because of a nodeless valence pseudo-orbital in the core region and core penetration effects, as well as inadequate Pauli repulsion between valence electrons on one center with the pseudopotential core on another.<sup>43</sup> Therefore, as expected, they have a very large A-C-MUE(3) (>10 kcal/mol) for the three homonuclear cases. LANL2DZ is the only tested NRECP generated from nonrelativistic AE numerical HF atomic wave functions for As; that could be one of reasons that it has the largest C-MUEs for the three cases.

In the case of H<sub>3</sub>AsO<sub>4</sub>, if the cc-pV $x$ Z AE basis set is used for oxygen and hydrogen and the cc-pV $x$ Z-PP basis set is used for As, all the tests, except for that using the quintuple- $\zeta$  basis set, give worse results than those with the def2 basis set at the same level. We took cc-pVTZ-PP as an example and tested its performance when the diffuse ma-cc-pVTZ or ma-TZVP basis set is used for O and H. The C-MUE for  $D_e$  of H<sub>3</sub>AsO<sub>4</sub> is respectively decreased from 5.5 to 0.9 or to 1.4 kcal/mol. Even when a large-core ECP basis set is used for As, but with ma-TZVP for O and H, the C-MUE could be around 1.5 kcal/mol. This confirms that the choice of basis set for oxygen is more important than the As basis for the H<sub>3</sub>AsO<sub>4</sub> acid dissociation reaction.

In the calculations of  $D_e$  for FeAs, unpublished cc-pV $x$ Z-PP basis sets for iron were kindly provided by Peterson.<sup>45</sup> For the tests of large-core ECP basis sets for As, the corresponding small-core (10 electrons) ECP is used for iron. Table 4 shows that only cc-pV $x$ Z-PP ( $x \geq T$ ), MDF10, and LANL2DZ (which is small-core NRECP for Fe) give reasonably accurate results for  $D_e$  of FeAs.

Considering all five cases, the cc-pVTZ-PP ECP basis set and the MDF10 RECP with its own basis set are recommended for arsenic, and their A-C-MUE(5) are smaller than or similar to that of the AE def2-TZVP basis. In addition, the latest new cc-pwCVTZ-PP basis set<sup>46</sup> that considers core–valence correlation effect is also tested for the VIP and  $D_e$  of As<sub>2</sub> and the IP of As. It yields a similar result to that for cc-pVTZ-PP. This indicates that the effect of core–valence correlation effect is slight for the arsenic compounds tested.

### 3.2.2. DFT Calculations with Nonstandard RECP Basis Sets.

Sufficiently accurate optimized valence basis sets are as important as the ECPs themselves because the radial shape of the pseudo-valence orbitals in the core region is different from the shape of AE valence orbitals as a result of the underlying pseudovalence orbital transformation.<sup>43</sup> The shape and size extent of pseudo-valence orbitals depend on the choice of ECP, so that the exponents of basis functions from AE nonrelativistic basis sets are not necessarily applicable for ECP calculations, nor are the contraction coefficients. For each ECP, a particular valence basis set is usually recommended by the creators of the ECP. Nevertheless, in this section we test some new combinations of def2 basis sets with the reasonably successful MDF10 small-core RECP. Selected results are presented in Table 5.

From the investigations above, we know that the def2-SVP or ma-SVP basis sets are relatively better at calculations of the relativistic effect than def2-TZVP; thus, it is not very surprising that Table 5 shows that the combinations of def2-SVP or some modified basis sets based on def2-SVP with MDF10 perform relatively well, in despite of having less basis functions than def2-TZVP. One modification that we considered is deleting the first contracted s basis function from def2-SVP to make it more compatible with the MDF10 RECP. We label the modified def2-SVP basis in which this 1s contracted function is excluded the sc-SVP basis, where “sc” denotes its intended use with a small-core RECP. The best combination that we found employs the MDF10 RECP combined with a minimally augmented sc-SVP basis set,

**Table 6.** Deviation (kcal/mol) from Relativistic NCBS-DK Reference Value<sup>a</sup> of HF Calculations and Mean Unsigned Deviation (MUEs in kcal/mol)

basis set	RECP	$D_e$ , As <sub>2</sub>	VIP, As <sub>2</sub>	IP, As	$D_e$ , H <sub>3</sub> AsO <sub>4</sub> <sup>b</sup>	MUE(3) <sup>c</sup>	MUE(4) <sup>b,d</sup>
cc-pV5Z-PP	MDF10	0.41	−0.07	0.01	0.28	0.16	0.19
cc-pVQZ-PP	MDF10	0.32	−0.02	0.06	1.15	0.13	0.39
cc-pVTZ-PP	MDF10	−0.23	0.53	0.40	3.30	0.38	1.11
cc-pVDZ-PP	MDF10	−3.33	1.78	0.56	4.65	1.89	2.58
MDF10	MDF10	−1.91	1.15	0.14	−0.37 <sup>e</sup>	1.07	0.89 <sup>e</sup>
CRENBL	CRENBL	−22.86	12.72	1.77	−13.85	12.45	12.80
SDD	MWB28	−28.15	9.41	5.70	−16.02	14.42	14.82
MDF28	MDF28	−29.16	5.90	−1.44	[−1.56]	12.17	[9.51]
CEP	CEP	−29.74	9.63	4.50	−21.08	14.62	16.23
LANL2DZ	LANL2DZ	−35.06	8.02	6.07	−14.47	16.38	15.90
ma-sc-SVP	MDF10	−0.60	1.23	−1.41	−5.42	1.08	2.16
ma-SVP	MDF10	−1.73	−0.01	−2.34	−5.54	1.36	2.41
def2-SVP	MDF10	−1.05	0.03	−2.51	4.66	1.20	2.06
def2-TZVP	MDF10	4.86	−4.67	−3.96	0.93	4.50	3.61
def2-QZVP	MDF10	1.17	−4.18	−3.64	1.34	2.99	2.58

<sup>a</sup> The NCBS-DK reference value is the HF relativistic result with the cc-pV5Z-DK basis set for  $D_e$  and VIP of As<sub>2</sub> and IP of As, and it is the relativistic result with ma-cc-pV5Z-DK for  $D_e$  of H<sub>3</sub>AsO<sub>4</sub>. <sup>b</sup> The values in square brackets are obtained using the ma-TZVP basis set for O and H. <sup>c</sup> The MUE(3) is the mean unsigned deviation from relativistic NCBS-DK reference value for VIP and  $D_e$  of As<sub>2</sub> and IP of As. <sup>d</sup> The MUE(4) is the mean unsigned deviation from relativistic NCBS-DK reference value for VIP and  $D_e$  of As<sub>2</sub>, IP of As, and  $D_e$  of H<sub>3</sub>AsO<sub>4</sub>. <sup>e</sup> The MDF10 basis set is used for As; the aug-cc-pVTZ basis set is used for O and H.

where the minimal augmentation is performed in the usual<sup>31</sup> way. We label this combination of basis set and RECP as ma-sc-SVP. The ma-sc-SVP basis set has a little worse performance than MDF10 when only three cases (VIP and  $D_e$  of As<sub>2</sub> and IP of As) are considered; however, when all five tested properties are considered, it has comparable accuracy to MDF10 and is only 0.1 kcal/mol worse than cc-pVTZ-PP (in the case of H<sub>3</sub>AsO<sub>4</sub>, the ma-TZVP basis set is used for O and H). As shown in Table 3, ma-sc-SVP has less basis functions than cc-pVTZ-PP and MDF10. In addition, in contrast to cc-pVTZ-PP (which is unpublished for 3d metal elements from Sc to Ni), ma-sc-SVP is published for all 3d metals. It could be a good compromise to use ma-sc-SVP in calculations for large arsenic systems containing 3d metals. The ma-sc-SVP basis needs to be tested further and validated, if possible, for 3d metals.

As shown in Table 5, the combination of def2-TZVP with the MDF10 RECP, denoted as def2-TZVP-MDF10, is the worst combination due to overpolarized pseudovalence orbitals. The def2-TZVPx-MDF10, which is def2-TZVP-MDF10 without *f* functions, has a substantially improved performance for  $D_e$  of As<sub>2</sub> as compared to def2-TZVP-MDF10. The larger def2-QZVP-MDF10 (def2-QZVP in conjunction with the MDF10 RECP) is worse than the combination of def2-SVP and MDF10 (def2-SVP-MDF10) or than ma-sc-SVP. We hypothesize that the better performance of double- $\zeta$  basis sets with MDF10 could be a result of cancellation of two kinds of errors: one is the intrinsic error of small basis sets; the other is the error resulting from the radial shape change of pseudovalence orbitals in the core region.

**3.2.3. HF Calculations with ECPs.** Table 6 lists the deviation and mean unsigned deviation (MUE) of HF calculations with the ECPs tested above from relativistic NCBS-DK HF reference values for  $D_e$  of As<sub>2</sub>, VIP of As<sub>2</sub>, IP of As, and  $D_e$  of H<sub>3</sub>AsO<sub>4</sub>. Comparing Table 6 with Tables 4 and 5 shows that, in most of the cases, the small-core MDF10 RECP with the basis sets optimized for it (the original MDF10 basis set and the cc-pVxZ-PP basis)

have better performance for HF calculations than for DFT, especially with the cc-pVxZ-PP ( $x \geq T$ ) basis set. The non-standard combinations of MDF10 with def2-SVP, ma-SVP, or ma-sc-SVP basis sets always perform better for IP of As in HF calculations, and when considering all four cases, these nonstandard combinations also have a smaller MUE(4) for HF calculations than for DFT calculations. The relatively worse performance in DFT confirms that the transferability of HF-based ECPs is reduced in DFT calculations because of the nonlinear nature of the xc functional's dependence on the density.

Although DFT calculations with the MDF10 RECP for arsenic have larger deviations from relativistic NCBS-DK reference values than do HF calculations, the deviations still are within an acceptable error range. Therefore, NLCCs will be helpful but are not very important for As.

All of the large- and medium-core ECPs have even larger errors for HF calculations than DFT calculations, as shown in Tables 4 and 6. This probably occurs because the errors introduced by using a large core dominate the total error of the large-core ECP calculations, so that the nonlinear error is drowned out by the large-core error.

**3.3. Further Tests for def2-TZVP, cc-pVTZ-PP, ma-sc-SVP, and MDF10 for Arsenic Clusters.** On the basis of the above investigations, the preferred basis sets and ECPs for arsenic in DFT calculations of relatively large arsenic-containing systems are RECP calculations with the cc-pVTZ-PP, ma-sc-SVP, and MDF10 basis sets and nonrelativistic def2-TZVP all-electron calculations. In this section, we test these four choices further by comparing their performance for 10  $D_e$  and three VIP calculations of arsenic clusters As<sub>*n*</sub> ( $n = 2-4$ ) plus the IP of As atom to NCBS-DK reference values. The calculated C-MUEs and the average C-MUE of the 14 properties (A-C-MUE(14)) over 11 xc functionals are listed in Table 7, along with (for comparison) the relativistic results obtained with the def2-TZVP and 6-311+G(2d,p) basis sets. We note that previous studies<sup>47</sup> of arsenic

**Table 7. Complete Mean Unsigned Errors (C-MUE, kcal/mol) for  $D_e$  Values of 10 Dissociation Reactions and IP or VIPs of Four Ionization Reactions over 11 xc Functionals, with cc-pV5Z, def2-TZVP, cc-pVTZ-PP, ma-sc-SVP, and MDF10, Compared to NCBS-DK Reference Values for the Given Functional**

	NCBS-DK <sup>a</sup>	NCBS-NR <sup>b</sup>	6-311+G(2d,p)	def2-TZVP	def2-TZVP	cc-pVTZ-PP	ma-sc-SVP	MDF10
type	DK	NR	DK	DK	NR	RECP	RECP	RECP
For $D_e$ s of 10 Dissociation Reactions								
$\text{As}_2 \rightarrow \text{As} + \text{As}$	0	0.65	1.23	2.63	0.78	0.74	0.77	2.69
$\text{As}_3 \rightarrow \text{As}_2 + \text{As}$	0	0.27	2.30	1.70	0.84	0.35	0.76	2.58
$\text{As}_4 \rightarrow \text{As}_3 + \text{As}$	0	0.35	2.07	3.23	1.36	0.26	3.77	2.89
$\text{As}_4 \rightarrow 2\text{As}_2$	0	0.62	3.54	2.30	1.76	0.54	2.87	2.79
$\text{As}_2^+ \rightarrow \text{As} + \text{As}^+$	0	0.66	1.75	2.46	0.50	1.19	3.44	3.16
$\text{As}_3^+ \rightarrow \text{As}_2 + \text{As}^+$	0	0.33	2.46	1.51	0.39	0.58	2.37	2.37
$\text{As}_3^+ \rightarrow \text{As}_2^+ + \text{As}$	0	0.40	1.54	1.67	0.93	0.20	1.30	1.90
$\text{As}_4^+ \rightarrow \text{As}_3 + \text{As}^+$	0	0.99	3.17	2.38	0.62	0.67	1.41	3.26
$\text{As}_4^+ \rightarrow \text{As}_2 + \text{As}_2^+$	0	0.48	3.73	1.62	0.57	0.45	1.39	2.68
$\text{As}_4^+ \rightarrow \text{As}_3^+ + \text{As}$	0	0.41	3.02	2.57	0.23	0.46	0.44	3.47
For IP or VIPs of Four Ionization Reactions								
$\text{As} \rightarrow \text{As}^+$	0	0.82	0.47	0.58	0.85	0.71	2.07	0.58
$\text{As}_2 \rightarrow \text{As}_2^+$	0	1.42	0.92	0.79	1.56	0.36	1.59	0.21
$\text{As}_3 \rightarrow \text{As}_3^+$	0	1.40	0.25	0.86	1.48	0.55	0.98	0.88
$\text{As}_4 \rightarrow \text{As}_4^+$	0	2.11	1.10	1.43	2.80	0.39	3.07	0.41
A-C-MUE(14) <sup>c</sup>	0	0.78	1.98	1.84	1.05	0.53	1.87	2.13
A-C-MUE(16) <sup>d</sup>	0	0.85	1.92	1.75	1.19	0.98	2.01	2.08

<sup>a</sup> NCBS-DK reference value is the relativistic result with cc-pV5Z-DK basis set. <sup>b</sup> NCBS-NR reference value is the nonrelativistic result with cc-pV5Z basis set. <sup>c</sup> A-C-MUE(14) is the average C-MUE of the 14 properties, each of which is calculated with 11 xc functionals, for a total of 154 comparisons to NCBS-DK results for each entry in the second last row. <sup>d</sup> A-C-MUE(16) is the average C-MUE for the five quantities in Tables 1 (or Table 3 or 4) and the 14 quantities in this table, counting the three quantities that appear in both tables only once. Therefore, each entry in the last row is an average over  $16 \times 11 = 176$  comparisons to NCBS-DK results.

**Table 8. Mean Unsigned Errors (C-MUE and NR-MUE, kcal/mol) of cc-pVTZ and cc-pVTZ-DK Basis Sets for  $D_e$  and VIP of  $\text{As}_2$  and IP of As Averaged over 11 xc Functionals, Based on Comparisons to the NCBS-DK Reference Value and NCBS-NR Reference Value of the Given Functional<sup>a</sup>**

		C-MUE			
basis set	type	$D_e, \text{As}_2$	VIP, $\text{As}_2$	IP, As	A-C-MUE(3) <sup>b</sup>
cc-pVTZ-DK	DK	0.61	0.13	0.32	0.35
cc-pVTZ	DK	7.99			
cc-pVTZ	NR	0.28	1.35	0.53	0.72
		NR-MUE			
basis set	type	$D_e, \text{As}_2$	VIP, $\text{As}_2$	IP, As	A-NR-MUE(3) <sup>b</sup>
cc-pVTZ	NR	0.63	0.14	0.31	0.36

<sup>a</sup> The NCBS-DK reference value is the relativistic result with cc-pV5Z-DK basis set; the NCBS-NR reference value is the nonrelativistic result with the cc-pV5Z basis set. All results in this table are for all-electron calculations. <sup>b</sup> A-C-MUE(3) [A-NR-MUE(3)] is the average of C-MUE (NR-MUE) for the three cases.

clusters showed that the def2-TZVPP basis set gives reliable results for both nonrelativistic density functional and wave function calculations. Here we restrict attention to density functional calculations, and we study the def2-TZVP basis set and compare the results to a more diverse set of other approaches, including relativistic effects.

**Table 9. Complete Error (kcal/mol) for IP of As**

basis set	def2-TZVP	6-311++G(3df,3pd)
Type	NR	NR
M05	0.87	0.99
M06-L	2.37	0.00
M06	1.26	0.43
BLYP	0.36	0.87
$\omega$ B97X-D	1.20	1.12
rHCTHhyb	0.46	0.57
G96LYP	0.92	0.85
mPWLYP	0.06	0.81
B3LYP	0.63	0.87
X3LYP	0.55	0.86
BPBE	0.70	0.83

Table 7 shows the clear superiority of cc-pVTZ-PP when errors are averaged over the 14 properties, and it has a smaller A-C-MUE(14) than nonrelativistic calculations with the cc-pV5Z (NCBS-NR) basis set, although A-C-MUE(16), which is our most comprehensive figure of merit, is worse. The nonrelativistic def2-TZVP calculations also show reasonably good performance. As we expected, the relativistic def2-TZVP calculations are worse with an A-C-MUE(14) of 1.8 kcal/mol and a similar value for A-C-MUE(16). This confirms again that the def2-TZVP basis set should only be used in nonrelativistic calculations. Although the relativistic calculations with the 6-311+G(2d,p) basis set perform



**Table 10.** Relative Computational Cost Required for Single-Point Energy Calculations with Various Basis Sets for an As<sub>20</sub> Cluster, Using the HF and BLYP Methods

	type	number of basis functions	number of primitive Gaussians	HF <sup>a</sup>	BLYP <sup>a</sup>
cc-pVTZ-DK	DK	43	197	5.95	4.52
cc-pVTZ	NR	43	196	5.68	4.15
cc-pVTZ-PP	RECP	39	133	4.81	3.28
def2-TZVP	NR	48	114	2.76	1.32
ma-sc-SVP	RECP	35	78	1.00	0.63

<sup>a</sup> The relative computational cost is estimated by dividing the CPU time required for each calculation by the CPU time for HF/ma-sc-SVP calculations.

well for the first test set, they have slightly worse A-C-MUE(14) value than the relativistic def2-TZVP results in this larger test set. The performances of ma-sc-SVP and of MDF10 with its own basis set are found to be inconsistent. For some cases they perform well, but in other cases they perform poorly. The ma-sc-SVP has a general performance worse than that of relativistic def2-TZVP calculations, and MDF10 is even worse. Therefore, although ma-sc-SVP is a relatively good RECP basis set, considering its small size and its availability for 3d metals, one must be careful about using it, and MDF10 cannot be recommended.

**3.4. cc-pVTZ and cc-pVTZ-DK.** Table 8 shows a few results for all-electron calculations with the cc-pVTZ and cc-pVTZ-DK basis sets. First we see, as already seen earlier, that the use of contracted AE basis sets in relativistic calculations where contraction coefficients are optimized for nonrelativistic calculations can lead to very large errors; in particular, the relativistic calculations with the cc-pVTZ basis set (optimized for nonrelativistic calculations) has a C-MUE of 7.99 kcal/mol for  $D_e$  of As<sub>2</sub>. For the same size and type of basis set, the results are much better when a nonrelativistically optimized basis is used nonrelativistically or when a relativistically optimized basis is used relativistically.

**3.5. Method Dependence.** We emphasize that all results presented so far are averaged over 11 choices of xc functional. To remind the reader that the results depend on the functional, Table 9 presents some results for individual functionals. Table 9 takes the IP of As as an example. The table shows the calculated complete error in the IP of As for each of the 11 xc functionals with def2-TZVP and 6-311++(3df,3pd) [it is the same as 6-311+G(3df) for As], in each case comparing to the relativistic cc-pV5Z-DK result for that functional. We see that M06-L has the smallest basis set error with 6-311++G(3df,3pd) and the largest basis set error for the def2-TZVP basis set. The mPWLYP functional is best with def2-TZVP (error = 0.06 kcal/mol) but has an error of 0.81 kcal/mol with the other basis for which results are shown.

**3.6. Timings.** The CPU time for calculating integrals depends on both the number of Gaussian primitives and the number of contracted Gaussians, while all other steps depend on only the number of contracted basis functions. Some AE basis sets and ECP basis sets have been chosen to illustrate the computational cost of HF and BLYP methods for a large arsenic cluster As<sub>20</sub>, and the corresponding CPU times are shown in Table 10. The time saving of using ECPs is confirmed. Furthermore, we note that the all-electron cc-pVTZ-DK BLYP calculation takes 7.2 times longer than the ma-sc-SVP one employing an ECP.

## 4. SUMMARY

Before summarizing the conclusions, we remind the reader that the present study is concerned with effective core potentials (ECPs) and basis-set incompleteness, not with the accuracy of density functionals. Therefore, all errors are relative to the nearly complete basis set all-electron limit for a given property and given density functional, not with respect to experiment.

We know from previous work that basis set requirements for electronic structure calculations are different for DFT and for wave function theory. We also know that scalar relativistic effects become significant for chemical accuracy for fourth-period elements. Here we test strategies for reproducing the basis-set limit of relativistic calculations on the fourth-period element As. We test four kinds of strategies for reproducing nearly complete basis set relativistic calculations: all-electron (AE) relativistic calculations; AE nonrelativistic calculations; formally nonrelativistic calculations in which a small, medium, or large core is replaced by a relativistic effective core potential (RECP); and formally nonrelativistic calculations with a nonrelativistic ECP (NRECP). All relativistic calculations are carried out by employing the Douglas–Kroll–Hess second-order scalar relativistic Hamiltonian.

The performance of AE basis sets from the def2-*x*ZVP and ma-*x*ZVP series and from the 6-311G series with additional polarization and diffuse functions was first tested by relativistic and nonrelativistic density functional calculations for five cases: VIP of As<sub>2</sub>,  $D_e$  of As<sub>2</sub>, IP of As, acid dissociation of H<sub>3</sub>AsO<sub>4</sub>, and  $D_e$  of FeAs. Overall, relativistic ma-QZVP and def2-QZVP calculations, with average mean unsigned errors of 0.2 and 0.4 kcal/mol, perform better than relativistic 6-311++G(3df,3pd) calculations, which have an average mean unsigned error of 0.7 kcal/mol. The def2-TZVP and ma-TZVP basis sets are good for nonrelativistic calculations, with average mean unsigned errors relative to nearly converged relativistic calculations of 1.3–1.5 kcal/mol, but using them in relativistic calculations does not decrease their errors. Therefore, it is not recommended to do relativistic calculations with the def2-TZVP or ma-TZVP basis set. The relativistic calculations with the 6-311+G(2d,p) basis set perform significantly better than nonrelativistic calculations with this basis set, and they perform slightly better than the def2-TZVP basis set, with an averaged mean unsigned error of ~1.2 kcal/mol. The other tested 6-311G-type basis sets [6-311G(3df,3pd), 6-311G(2d,p), 6-311G(d,p), and 6-311+G(d,p)], which are commonly used in previous studies of arsenic compounds, perform worse than def2-TZVP for both relativistic and nonrelativistic calculations, with average mean unsigned errors in the range 1.6–3.4 kcal/mol. As expected, the smallest basis set, def2-SVP, performs worst of all tested AE basis sets, except 6-311G(d,p). However, when a good basis set, ma-TZVP, is used for oxygen, the use of def2-SVP for As only has an average mean unsigned error over the five cases that is only 0.3 kcal/mol larger than using def2-TZVP. Due to the ionic product H<sub>2</sub>AsO<sub>4</sub><sup>−</sup>, the use of diffuse functions can significantly improve the results for acid dissociation of H<sub>3</sub>AsO<sub>4</sub>. For the case of FeAs, diffuse functions have less effect on the def2-TZVP and def2-QZVP basis sets, but they are helpful for the smaller def2-SVP basis set and the 6-311G-type basis set with tighter p functions for iron. The inclusion of diffuse functions does not improve the results for the other cases and may even lead to worse results.

Several ECPs available for arsenic were also investigated in the present work. First we considered the same five cases as in the previous paragraph. The cc-pV<sub>*x*</sub>Z-PP (*x* ≥ T) basis in which valence basis functions are specifically optimized for the

small-core MDF10 RECP was found to give very good results; this combination, denoted simply as cc-pVTZ-PP, has an average mean unsigned error of 1.1 kcal/mol (when the ma-cc-pVTZ augmented basis set is used for O in the case of  $\text{H}_3\text{AsO}_4$ ). The small-core MDF10 RECP with its own basis set also performs well, with an average mean unsigned error of 1.4 kcal/mol. All of the medium-core and large-core ECPs perform poorly, with average mean unsigned errors of 8.5–11.9 kcal/mol. In the popular Gaussian program, only large-core ECPs are implemented for arsenic, so one must expect large deviations from the relativistic nearly complete basis set limit.

We also introduced a very efficient nonstandard combination of basis set and RECP that is called ma-sc-SVP; this denotes the ma-SVP basis set with the first contracted s function removed and used in conjunction with the small-core MDF10 RECP. This combination has a comparable performance with MDF10, but with many less basis functions. The application of ma-sc-SVP to the rest of the fourth period needs to be validated.

The performances of small-core ECPs and associated basis sets for DFT calculations are a little worse than but still comparable to those for HF calculations for four cases:  $D_e$  of  $\text{As}_2$ , VIP of  $\text{As}_2$ , IP of As, and  $D_e$  of  $\text{H}_3\text{AsO}_4$ . Therefore, NLCCs in DFT calculations for arsenic will be helpful, but are not particularly necessary.

We selected seven of the methods for further DFT tests on a larger test set with 11 additional cases ( $D_e$  and VIP for various  $\text{As}_n$  with  $n = 3$  and 4), and we computed an average mean unsigned error for all 16 cases. Since each case is already averaged over 11 xc functionals, this final error estimate is an average over 176 basis set convergence tests. The final error for nearly complete basis set nonrelativistic calculations is 0.9 kcal/mol, and the final error for def2-TZVP nonrelativistic calculations is 1.2 kcal/mol. The latter number increases to 1.8 kcal/mol if the calculation is instead carried out relativistically. The relativistic calculations with the 6-311+G(2d,p) basis set give a slightly larger error than def2-TZVP relativistic calculations and therefore are not recommended. A much more efficient and more accurate way to include relativistic effects is with an RECP, and such a calculation with cc-pVTZ-PP has a final basis-set-incompleteness error of 1.0 kcal/mol. A very efficient but slightly less accurate option is ma-sc-SVP, with a final error of 2.0 kcal/mol.

On the basis of the present investigations, for small arsenic-containing systems, we recommend relativistic def2-QZVP (or ma-QZVP when diffuse functions are needed) calculations for getting accurate results. If the relativistic effect is insignificant for the property under consideration, nonrelativistic def2-TZVP (or ma-TZVP when diffuse functions are needed) calculations are good enough. The triple- $\zeta$  RECP basis set cc-pVTZ-PP is recommended for larger arsenic systems. However, if 3d metals are involved in the arsenic-containing system, the double- $\zeta$  RECP basis set ma-sc-SVP may be tested. The use of ma-sc-SVP for As could lead to basis-set errors up to  $\sim 2$  kcal/mol.

## ■ ASSOCIATED CONTENT

**S Supporting Information.** Bond length tests and NCBS-DK reference values. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [truhlar@umn.edu](mailto:truhlar@umn.edu).

## ■ ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation by grant no. CHE09-56776 and by the Air Force Office of Scientific Research.

## ■ REFERENCES

- (1) Stoll, H.; Metz, B.; Dolg, M. *J. Comput. Chem.* **2002**, *23*, 767.
- (2) Louie, S. G.; Froyen, S.; Cohen, M. L. *Phys. Rev. B* **1982**, *26*, 1738.
- (3) Porezag, D.; Pederson, M. R.; Liu, A. Y. *Phys. Rev. B* **1999**, *60*, 14132.
- (4) van Wüllen, C. *Int. J. Quantum Chem.* **1996**, *58*, 147.
- (5) Russo, T. V.; Martin, R. L.; Hay, P. J. *J. Phys. Chem.* **1995**, *99*, 17085.
- (6) Han, Y.-K.; Hirao, K. *Chem. Phys. Lett.* **2000**, *324*, 453.
- (7) Lovallo, C. C.; Klobukowski, M. *J. Comput. Chem.* **2004**, *25*, 1206.
- (8) Dhubhghaill, O. M. Ni; Sadler, P. J. *Struct. Bonding (Berlin)* **1991**, *78*, 129 and references cited therein.
- (9) Wolfe-Simon, F.; Blum, J. S.; Kulp, T. R.; Gordon, G. W.; Hoefft, S. E.; Pett-Ridge, J.; Stolz, J. F.; Webb, S. M.; Weber, P. K.; Davies, P. C. W.; Anbar, A. D.; Oremland, R. S. [dx.doi.org/10.1126/science.1197258](https://doi.org/10.1126/science.1197258) (accessed March 28, 2011)
- (10) Katsnelson, A. [dx.doi.org/10.1038/news.2010.645](https://doi.org/10.1038/news.2010.645) (accessed March 28, 2011)
- (11) Bortman, H. <http://www.space.com/9631-arsenic-eating-bacteria-opens-possibilities-alien-life.html> (accessed March 28, 2011)
- (12) Silver, S.; Phung, L. T. *FEMS Microbiol. Lett.* **2011**, *315*, 79.
- (13) (a) Miqueu, K.; Sotiropoulos, J. M.; Pfister-Guilouzo, G.; Ranaivonjatovo, H.; Escudé, J. J. *Mol. Struct.* **2001**, *545*, 139. (b) Tossell, J. A. *Geochim. Trans.* **2003**, *4*, 28. (c) Xu, W. G.; Jin, B. J. *Mol. Struct.: THEOCHEM* **2005**, *731*, 61. (d) Xu, W. G.; Jin, B. J. *Mol. Struct.: THEOCHEM* **2006**, *759*, 101. (e) Urban, D. R.; Wilcox, J. J. *J. Phys. Chem. A* **2006**, *110*, 5847. (f) Tossell, J. A.; Zimmermann, M. D. *Geochim. Cosmochim. Acta* **2008**, *72*, 5232. (g) Ahmed, A. B.; Feki, H.; Abid, Y.; Minot, C. *Spectrochim. Acta, Part A* **2010**, *75*, 1315. (h) Ramírez-Galicia, G.; Peñ-Méndez, E. M.; Pangavhane, S. D.; Alberti, M.; Havel, J. *Polyhedron* **2010**, *29*, 1567.
- (14) (a) Wachters, A. J. H. *J. Chem. Phys.* **1970**, *52*, 1033. (b) Hay, P. J. *J. Chem. Phys.* **1977**, *66*, 4377. (c) Raghavachari, K.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650. (d) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639. (e) Raghavachari, K.; Trucks, G. W. *J. Chem. Phys.* **1989**, *91*, 1062. (f) Binning, R. C., Jr.; Curtiss, L. A. *J. Comput. Chem.* **1990**, *11*, 1206. (g) McGrath, M. P.; Radom, L. *J. Chem. Phys.* **1991**, *94*, 511. (h) Curtiss, L. A.; McGrath, M. P.; Blaudau, J.-P.; Davis, N. E.; Binning, R. C., Jr.; Radom, L. *J. Chem. Phys.* **1995**, *103*, 6104.
- (15) (a) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007. (b) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358. (c) Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H., Jr. *J. Chem. Phys.* **1999**, *110*, 7667. (d) Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 064107.
- (16) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- (17) (a) Billes, F.; Mitsa, V.; Fejes, I.; Mateleshko, N.; Fejsa, I. *J. Mol. Struct.* **1999**, *513*, 109. (b) Ozeki, M.; Iwasa, Y.; Shimizu, Y. *Appl. Surf. Sci.* **2007**, *253*, 5914. (c) Lu, H.-L.; Chen, W.; Ding, S.-J.; Zhang, D. W.; Wang, L.-K. *Chem. Phys. Lett.* **2007**, *445*, 188. (d) Zimmermann, M. D.; Tossell, J. A. *J. Phys. Chem. A* **2009**, *113*, 5105. (e) Yang, S.; Adjaye, J.; McCaffrey, W. C.; Nelson, A. E. *J. Mol. Catal. A* **2010**, *321*, 83.
- (18) Chu, J.-H.; Analytis, J. G.; Greve, D. K.; McMahon, P. L.; Islam, Z.; Yamamoto, Y.; Fisher, I. R. *Science* **2010**, *329*, 824.
- (19) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123*, 161103.
- (20) (a) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101. (b) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.

- (21) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.
- (22) (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098. (b) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (23) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.
- (24) Boese, A. D.; Handy, N. C. *J. Chem. Phys.* **2002**, *116*, 9559.
- (25) (a) Gill, P. M. W. *Mol. Phys.* **1996**, *89*, 433. (b) Adamo, C.; Barone, V. *J. Comput. Chem.* **1998**, *19*, 418.
- (26) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664.
- (27) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
- (28) Xu, X.; Goddard, W. A., III *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 2673.
- (29) (a) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865. (b) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396.
- (30) (a) Douglas, M.; Kroll, N. M. *Ann. Phys. (NY)* **1974**, *82*, 89. (b) Hess, B. A. *Phys. Rev. A* **1985**, *32*, 756. (c) Hess, B. A. *Phys. Rev. A* **1986**, *33*, 3742. (d) Jansen, G.; Hess, B. A. *Phys. Rev. A* **1989**, *39*, 6016. (e) Barysz, M.; Sadlej, A. J. *J. Mol. Struct.: THEOCHEM* **2001**, *573*, 181. (f) deJong, W. A.; Harrison, R. J.; Dixon, D. A. *J. Chem. Phys.* **2001**, *114*, 48.
- (31) Zheng, J.; Xu, X.; Truhlar, D. G. *Theor. Chem. Acc.* **2011**, *128*, 295.
- (32) Peterson, K. A. *J. Chem. Phys.* **2003**, *119*, 11099.
- (33) (a) Metz, B.; Stoll, H.; Dolg, M. *J. Chem. Phys.* **2000**, *113*, 2563. (b) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1987**, *86*, 866.
- (34) Hurley, M. M.; Pacios, L. F.; Christiansen, P. A.; Ross, R. B.; Ermiler, W. C. *J. Chem. Phys.* **1986**, *84*, 6840.
- (35) Bergner, A.; Dolg, M.; Kuechle, W.; Stoll, H.; Preuss, H. *Mol. Phys.* **1993**, *80*, 1431.
- (36) Stevens, W. J.; Krauss, M.; Basch, H.; Jasien, P. G. *Can. J. Chem.* **1992**, *70*, 612.
- (37) (a) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270. (b) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 284. (c) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299.
- (38) Dyall, K. G. *Theor. Chem. Acc.* **2006**, *115*, 441.
- (39) Huber, K. P.; Herzberger, G. *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules*; Van Nostrand: Princeton, 1979.
- (40) Igel-Mann, G.; Stoll, H.; Preuss, H. *Mol. Phys.* **1993**, *80*, 325.
- (41) Morino, Y.; Ukaji, T.; Ito, T. *Bull. Chem. Soc. Jpn.* **1966**, *39*, 64.
- (42) 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, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02, Gaussian, Inc., Wallingford CT, 2009.
- (43) Dolg, M. *Relativistic Electronic Structure Theory, Part 1: Fundamentals Theoretical and Computational Chemistry*; Schwerdtfeger, P., Ed.; Theoretical and Computational Chemistry Series 11; Elsevier: Amsterdam, 2002; 793.
- (44) Kahn, L. R.; Baybutt, P.; Truhlar, D. G. *J. Chem. Phys.* **1976**, *65*, 3826.
- (45) The cc-pVxZ-PP basis sets for iron element were obtained from Kirk Peterson. The cc-pVxZ-PP basis sets are matched with the new RECP optimized by Dolg in 2005. See: Dolg, M. *Theor. Chem. Acc.* **2005**, *114*, 297.
- (46) Peterson, K. A.; Yousaf, K. E. *J. Chem. Phys.* **2010**, *133*, 174116.
- (47) Nava, P.; Ahlrichs, R. *Chem.—Eur. J.* **2008**, *14*, 4039.