# Acidities of Arsenic (III) and Arsenic (V) Thio- and Oxyacids in Aqueous Solution using the CBS-QB3/CPCM Method

**Merle D. Zimmermann**[†] **and John A. Tossell***

*Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742*

*Received: October 15, 2008; Revised Manuscript Received: January 6, 2009*

A number of calculations of acidic As(III) and As(V) species formed with sulfur and oxygen ($H_3AsS_3$, $H_2AsS_3^-$, $HAsS_3^{2-}$, $AsS_3^{3-}$, $H_3AsS_4$, $H_2AsS_4^-$, $HAsS_4^{2-}$, $AsS_4^{3-}$, $H_3AsO_3$, $H_2AsO_3^-$, $HAsO_3^{2-}$, $AsO_3^{3-}$, $H_3AsO_4$, $H_2AsO_4^-$, $HAsO_4^{2-}$, and $AsO_4^{3-}$) are presented. p$K_a$s for successive deprotonations in both the gas phase and aqueous solution (using both explicit water molecules and a self-consistent reactive field conductive polarizable continuum model (SCRF CPCM) for solvation) are fitted to known experimental values for the $H_3PO_4$ and $H_3AsO_4$ series' of deprotonations with a linear extrapolation showing $r^2$ values of 0.97 for the CBS-QB3 method with a single explicit water molecule in the CPCM. Though the unfitted p$K_a$s of $H_3AsO_4$ (p$K_{a1}$ = 4.6, p$K_{a2}$ = 17.7, and p$K_{a3}$ = 28.6) compare unfavorably with experimental values of p$K_{a1}$ = 2.3, p$K_{a2}$ = 7, p$K_{a3}$ = 13, by the linear fit: 0.42907 × p$K_a$ − 0.23, the predicted p$K_a$s become p$K_{a1}$ = 1.7, p$K_{a2}$ = 7.3, and p$K_{a3}$ = 12.1, which correspond well. The $H_3PO_4$ calculation shows a similar trend; p$K_{a1}$ = 4.9, p$K_{a2}$ = 19.9, and p$K_{a3}$ = 29.8 becomes p$K_{a1}$ = 1.9, p$K_{a2}$ = 8.3, p$K_{a3}$ = 12.6. Experimentally, $H_3PO_4$ has p$K_{a1}$ = 2.2, p$K_{a2}$ = 7.1, and p$K_{a3}$ = 12.3. With the same extrapolation, we predict $H_2SO_3$ to have p$K_{a1}$ = 1.1 and p$K_{a2}$ = 8.1, which compare favorably with the experimental values of p$K_{a1}$ = 1.9 and p$K_{a2}$ = 7.2.

## 1. Introduction

Arsenic chemistry is intriguing from several perspectives. Both a naturally occurring contaminant in wells around the world and a geochemically available mineral,[1,2] a better understanding of arsenic species' behavior and properties would be to the benefit of all. This work is intended to compliment the related work on speciation models and analysis of vibrational spectra. During the last 20 years, a number of irregularities were observed with respect to some types of arsenic speciation.[3] Most published experimental information about arsenic acidities are from 20 or 30 years ago, except for work by Helz, Tossell, and co-workers.[4−7]

The continual improvement of computational capabilities allows us to take another look at the polyprotic acids of interest using more recent theoretical methods than have previously been attainable. Although the challenge of multiple deprotonations in polyprotic acids leaves quantitative results for the third deprotonation out of reach, we will show that qualitatively correct modeling for both first and second deprotonations can and should be attempted with methods readily available today. Primary deprotonations of $H_3PO_4$ and related small acids have been studied before by Dixon and co-workers[8,9] by use of higher level CCSD(T) correlation consistent complete basis set extrapolations to good result, and by adjusting the cavity sizes for the $H_2PO_4^-$ anion, results within 1 p$K_a$ unit of the experimental measurements were obtained.

In our approach, we focus on a variety of acids with structures isoelectronic to our acids of arsenic, and have found a significant linear correspondence between the results of the CBS-QB3 model chemistry approach and experimental measurement. An extrapolation is therefore provided which scales the calculated results to known experimental data, adjusting points of all but one of the modeled p$K_a$s ($H_3AsO_3$'s p$K_{a1}$) to within 1 p$K_a$ unit of available measurements without requiring changes to the basic theoretical approach.

## 2. Methods

Primary calculations used the CBS-QB3 complete basis set extrapolation[10] as implemented in Gaussian 03,[11] which was used to conduct our computations. CBS-QB3 is a general purpose model chemistry which tries to overcome the limitations of a finite basis set by using carefully selected basis sets and methods to estimate the complete basis set energy. This allows for more accurate results at a smaller total computational cost. With heavier atoms such as As, relativistic effects can affect the electronic structure of the element. We therefore conducted a series of calculations with CCSD(T) and the relativistically corrected CEP-121G basis set[12−14] for two of the systems studied, $H_3AsO_4$ and $H_3PO_4$, to determine if the relativistic contributions to the calculated p$K_a$s could significantly influence the outcome relative to the CBS-QB3 method.

CBS-QB3 uses B3LYP[15] to calculate geometries followed by a series of high level calculations (CCSD(T),[16] MP4SDQ,[17] and MP2[18]) to build its extrapolation. B3LYP[15] is a hybrid Hartree-Fock/density functional theory method which combines both approaches to calculate energetics. It is known for having both a low computational cost and unpredictable errors in the energies produced but generally good geometries. CCSD(T)[16] is a coupled cluster method with both single, double, and triple excitations. CCSD(T) scales with the seventh power of the number of basis functions, making larger basis sets such as the popular aug-cc-pV*Z correlation consistent basis sets computationally expensive, though quite possible through molecules with second or third group elements.[8,9,19] In CBS-QB3, CCSD(T) is used with a minimal 6-31+G(d′) basis set. MP4SDQ[17] and MP2[18] are two different levels of Moller−Plesset perturbation theory, taking into account single, double, and quadruple excitations. CBS-QB3 uses CBSB4 for its MP4SDQ calcula-

* To whom correspondence should be addressed. E-mail: tossell@umd.edu.
† E-mail: merle@umd.edu.

10.1021/jp809123q CCC: $40.75 
Published on Web 04/07/2009

tions, and CBSB3 for the MP2 calculation. In most cases, CBS-QB3 produces energetics similar to the G2[20] and G3[21] methods, within 1−2 kcal/mol of experimental measurements on the test set.[22]

Hydration energies were calculated using the conductive polarizable continuum method, CPCM.[23−25] In this method, molecules are placed in conductive cavities made of spherical segments to estimate the amount of stabilization caused by the solvent. Although not as effective in duplicating experimental measurements as explicit solvation, the computational cost is linearly related to the size of the system, a significant cost savings compared to adding additional molecules.

In cases where explicit solvation is used, there is also an additional source of noise in the results, as changes in the solvent structure cause variations in the calculated energies. A single hydrogen bond would make a difference of about 4.3 kcal/mol, which is significant when calculating p$K_a$s.

**2.1. CBS-QB3 Calculations.** CBS-QB3 was used to calculate energetics in our experiment. The CBS-QB3 method provided us with B3LYP/CBSB7 gas phase geometries and CBS-QB3 energies for each of the arsenic compounds we considered. When CBS-QB3 was extended to this part of the periodic table, the G2 test set gave CBS-QB3 reaction energies with errors on the order of 1.12 kcal/mol.[22] The As G2 test set compounds considered were $As_2$, AsH, $AsH_2$, and $AsH_3$ and not the oxides and sulfides we are focusing on here. The initial step of B3LYP geometry optimization occasionally runs into trouble. In these cases, several manual restarts of the optimization sequence were required to bring the structures to their final geometries. In more stubborn situations, the geometries were optimized using MP2 first and then reoptimized with B3LYP from final MP2 geometries. Though this used more CPU time, it required less manual intervention.

**2.2. Solvation Approaches.** We approach solvation in two ways. Very similar approaches have been previously used by others.[26,27]

In the first method we take the final geometry from the CBS-QB3 gas phase job, and put it into the polarizable continuum (PCM) directly without further alteration of the geometry. We then run a CPCM solvation calculation[23] to find the $\Delta G$ of solvation. The CPCM job is configured with the orbitals and electron distribution from a HF/6-31+G** single-point calculation, requesting a tight optimization of the self-consistent field, and using the SCFVAC and RADII=UAHF tags to extend the SCRF=CPCM calculation. Both SCFVAC and RADII=UAHF have been used in the past in our experiments and are chosen to ensure easy and accurate comparison with other papers from our laboratory.

Further optimization might be conducted in a future investigation by optimizing the solvation shell sizes for polarizable continuum model, as was done in the studies on small acids in the previously cited work.[8,9] Using CPCM at the gas phase geometries is limited in quality because, while it does estimate the stabilization of the compounds when placed in solution, it introduces errors because the geometries are not updated to account for the presence of the simulated solute. However, it allows for fast and easily reproducible solvation calculations with less intervention required than our second method detailed below.

The second method we use to find the hydration energy is to conduct the entire CBS-QB3 procedure in the CPCM. The stabilized compounds have geometries differing from the gas phase ones, as the optimization step of the CBS-QB3 process occurs in the CPCM reaction field. There are occasional issues with this second method because the B3LYP/CBSB7 optimization during the first stage of the CBS-QB3 calculation has even more difficulties converging to the optimized geometry on the flatter potential energy surface caused by the presence of the CPCM.

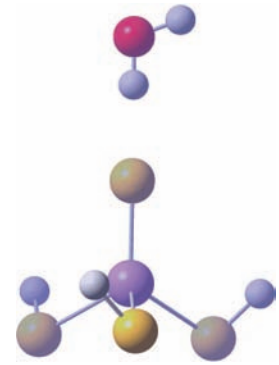

**Figure 1.** $H_3AsS_4$ with explicit $H_2O$ molecule showing initial geometry preoptimization hydrogen bonding arrangement.

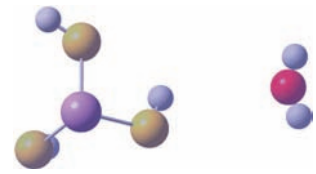

**Figure 2.** $H_3AsS_3$ with explicit $H_2O$ molecule showing initial geometry preoptimization hydrogen bonding arrangement.

In both cases, we take the energy of the solvated $H^+$ ion produced by the deprotonation in solution, corrected for concentration, at −269.0 kcal/mol (eq 8 in the cited paper):[28]

$$pK_a = [\Delta G(A_{gas}^-) - \Delta G(AH_{gas}) + \Delta G_s(A^-) - \Delta G_s(AH) - 269.0\text{kcal/mol}]/1.3644 \quad (1)$$

This figure was determined in the cited paper[28] (as eqs 4−8) by combining a gas phase contribution with the solvated free energy of the hydrogen ion (a $\Delta G_{solv}(H^+)$ of −264.61 kcal/mol) along with a correction for the differences in concentration of the gas and solvated reference states. The gas phase energy contribution was found in the reference by using the Sackur−Tetrode equation to evaluate the entropy of the gaseous $H^+$, then adding in the expected translational energy at 298 K to get a $G_{gas}(H^+)$ of −6.28 kcal/mol. The result is identical to that found by using the NIST-JANAF Thermochemical Tables,[29] which present us with an $S$ of 108.946 J $K^{-1}$ $mol^{-1}$, finding $-T\Delta S$ at 298 K, then converting to kcal/mol units (via 4.184 J/cal) to get −7.76 kcal/mol, and finally adding the thermal correction to the enthalpy of 0.002360 hartree/particle (1.481 kcal/mol). The conversion to go between the 1 atm concentration of $H^+$ in the standard state of the gas phase to the 1 M standard solution phase was $RT$ ln(24.46). This approach has previously been used in experiments from both our laboratory and in future works from the cited source. Other laboratories have used a value in the neighborhood of −264 kcal/mol for the energy of the solvated $H^+$ ion in their calculations and take care of the necessary conversions of state separately.[30−34] As any inaccuracies in this figure would cause a purely systematic error in the raw results, the accuracy of our later linear extrapolation will therefore not be influenced by our choice of method here.

**2.3. Higher Level Methods.** We also conduct CCSD(T) calculations at each of the final CBS-QB3 geometries of $H_3AsO_4$, $H_2AsO_4^-$, $HAsO_4^{-2}$, $AsO_4^{-3}$, $H_3PO_4$, $H_2PO_4^-$, $HPO_4^{-2}$,

**TABLE 1: Deprotonation Energetics**

| | gas phase $\Delta G_{\text{CBS-QB3}}$ (kcal/mol) | $\Delta G_{\text{CBS-QB3}} + \Delta G_{\text{CPCM}}$ (for CBS-QB3 and CPCM, handled separately) (kcal/mol) | p$K_a$ (aq) (for CBS-QB3 and CPCM, handled separately) | $\Delta G$ (aq) (CBS-QB3 in CPCM) | p$K_a$ (aq) (CBS-QB3 in CPCM) |
|---|---|---|---|---|---|
| $H_3AsS_4 \rightarrow H_2AsS_4^- + H^+$ | 306.1 | −5.0 | −3.6 | −5.3 | −3.9 |
| $H_2AsS_4^- \rightarrow HAsS_4^{-2} + H^+$ | 407.4 | 11.6 | 8.5 | 10.7 | 7.9 |
| $HAsS_4^{-2} \rightarrow AsS_4^{-3} + H^+$ | 499.5 | 18.3 | 13.4 | 19.2 | 14.1 |
| $H_3AsO_4 \rightarrow H_2AsO_4^- + H^+$ | 325.3 | 8.9 | 6.5 | 8.3 | 6.1 |
| $H_2AsO_4^- \rightarrow HAsO_4^{-2} + H^+$ | 452.3 | 26.8 | 19.6 | 26.3 | 19.3 |
| $HAsO_4^{-2} \rightarrow AsO_4^{-3} + H^+$ | 567.8 | 28.5 | 20.9 | 38.6 | 28.3 |
| $H_3AsS_3 \rightarrow H_2AsS_3^- + H^+$ | 319.3 | 7.4 | 5.5 | 7.4 | 5.4 |
| $H_2AsS_3^- \rightarrow HAsS_3^{-2} + H^+$ | [a] | [a] | [a] | 18.8 | 13.8 |
| $HAsS_3^{-2} \rightarrow AsS_3^{-3} + H^+$ | [a] | [a] | [a] | 26.2 | 19.2 |
| $H_2AsS_3^- \rightarrow AsS_3^{-3} + 2H^+$ | 937.7 | 45.5 | 33.3 | 45.0 | 33.0 |
| $H_3AsO_3 \rightarrow H_2AsO_3^- + H^+$ | 343.1 | 21.0 | 15.4 | 21.5 | 15.7 |
| $H_2AsO_3^- \rightarrow HAsO_3^{-2} + H^+$ | 470.5 | 37.8 | 27.7 | 38.8 | 28.4 |
| $HAsO_3^{-2} \rightarrow AsO_3^{-3} + H^+$ | 574.5 | 37.6 | 27.5 | 49.4 | 36.2 |
| $H_3PO_4 \rightarrow H_2PO_4^- + H^+$ | 327.0 | 8.3 | 6.1 | 6.6 | 4.8 |
| $H_2PO_4^- \rightarrow HPO_4^{-2} + H^+$ | 457.3 | 29.6 | 21.7 | 28.9 | 21.2 |
| $HPO_4^{-2} \rightarrow PO_4^{-3} + H^+$ | 580.9 | 34.1 | 25.0 | 40.7 | 29.8 |
| $H_3AsS_4 \cdot H_2O \rightarrow H_2AsS_4^- \cdot H_2O + H^+$ | 298.9 | −3.6 | −2.6 | −6.3 | −4.6 |
| $H_2AsS_4^- \cdot H_2O \rightarrow HAsS_4^{-2} \cdot H_2O + H^+$ | 402.2 | 10.8 | 7.9 | 11.3 | 8.3 |
| $HAsS_4^{-2} \cdot H_2O \rightarrow AsS_4^{-3} \cdot H_2O + H^+$ | 489.5 | 18.7 | 13.8 | 17.6 | 12.9 |
| $H_3AsO_4 \cdot H_2O \rightarrow H_2AsO_4^- \cdot H_2O + H^+$ | 321.6 | 7.8 | 5.7 | 6.3 | 4.6 |
| $H_2AsO_4^- \cdot H_2O \rightarrow HAsO_4^{-2} \cdot H_2O + H^+$ | 439.7 | 27.7 | 20.3 | 24.1 | 17.7 |
| $HAsO_4^{-2} \cdot H_2O \rightarrow AsO_4^{-3} \cdot H_2O + H^+$ | 551.2 | 29.7 | 21.8 | 39.1 | 28.6 |
| $H_3AsS_3 \cdot H_2O \rightarrow H_2AsS_3^- \cdot H_2O + H^+$ | 313.8 | 5.8 | 4.3 | 6.9 | 5.0 |
| $H_2AsS_3^- \cdot H_2O \rightarrow HAsS_3^{-2} \cdot H_2O + H^+$ | 412.1 | 20.3 | 14.9 | 18.9 | 13.8 |
| $HAsS_3^{-2} \cdot H_2O \rightarrow AsS_3^{-3} \cdot H_2O + H^+$ | 508.1 | 24.8 | 18.2 | 25.3 | 18.5 |
| $H_3AsO_3 \cdot H_2O \rightarrow H_2AsO_3^- \cdot H_2O + H^+$ | 333.8 | 14.1 | 10.3 | 15.8 | 11.5 |
| $H_2AsO_3^- \cdot H_2O \rightarrow HAsO_3^{-2} \cdot H_2O + H^+$ | 458.0 | 41.2 | 30.2 | 41.6 | 30.5 |
| $HAsO_3^{-2} \cdot H_2O \rightarrow AsO_3^{-3} \cdot H_2O + H^+$ | 560.8 | 38.5 | 28.2 | 47.9 | 35.1 |
| $H_3PO_4 \cdot H_2O \rightarrow H_2PO_4^- \cdot H_2O + H^+$ | 322.1 | 8.5 | 6.2 | 6.7 | 4.9 |
| $H_2PO_4^- \cdot H_2O \rightarrow HPO_4^{-2} \cdot H_2O + H^+$ | 443.0 | 27.9 | 20.5 | 27.2 | 19.9 |
| $HPO_4^{-2} \cdot H_2O \rightarrow PO_4^{-3} \cdot H_2O + H^+$ | 562.4 | 35.8 | 26.2 | 40.7 | 29.8 |
| $H_2SO_3 \rightarrow HSO_3^- + H^+$ | 323.6 | 4.1 | 3.0 | 3.3 | 2.4 |
| $HSO_3^- \rightarrow SO_3^{-2} + H^+$ | 466.0 | 29.9 | 21.9 | 31.6 | 23.1 |
| $H_2SO_3 \cdot H_2O \rightarrow HSO_3^- \cdot H_2O + H^+$ | 312.7 | 2.6 | 1.9 | 4.2 | 3.1 |
| $HSO_3^- \cdot H_2O \rightarrow SO_3^{-2} \cdot H_2O + H^+$ | 454.4 | 30.0 | 22.0 | 26.5 | 19.4 |

[a] As noted before, we were unable to find an equilibrium geometry for $HAsS_3^{-2}$ under gas phase CBS-QB3 conditions; these deprotonations could not be evaluated.

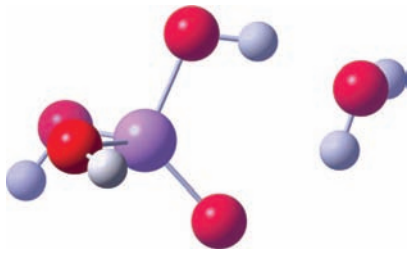

**Figure 3.** Hydrated $H_3AsO_4 \cdot H_2O$ complex at optimized CBS-QB3 geometry calculated in the PCM.

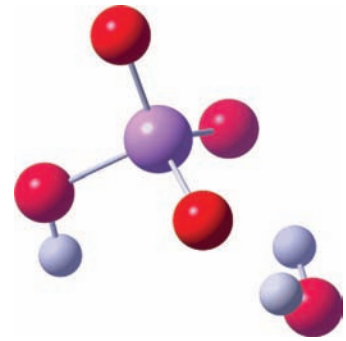

**Figure 4.** Hydrated $HAsO_4^{-2} \cdot H_2O$ complex at optimized CBS-QB3 geometry calculated in the PCM.

and $PO_4^{-3}$ (each without explicit $H_2O$ solvation) for comparison purposes using the 6−311G(2d,p), aug-cc-pVDZ, and CEP-121G basis sets. These calculations are each conducted at two different geometries: the gas phase CBS-QB3 geometry and the CPCM-optimized CBS-QB3 CPCM geometry. As these single-point calculations do not include zero-point energies, we added in the thermal correction to the Gibbs free energy found in the B3LYP/CBSB7 frequency calculation step of the CBS-QB3 process to the calculated CCSD(T) values.

These CCSD(T) calculations allow us to check whether the additional computational costs of the higher level method yields significantly better results and to measure the effects of additional d-type functions in the case of 6-311G(2d,p) basis set, compare the performance and accuracy of CBS-QB3's CBSB7 and CBSB8 basis sets against the CCSD(T)/aug-cc-pVDZ approach, and test the size of possible relativistic effects which would be accounted for by the relativistically aware core potential of As in CEP-121G.

**2.4. Explicit Water Molecules.** We have run the same series of CBS-QB3 calculations with a single explicit water molecule associated with the compounds. The water molecule was positioned so that before optimization it initially had a single hydrogen bond with the deprotonated oxygen or sulfur molecule in those compounds with one, as shown in Figure 1. For the cases of $H_3AsS_3$ (seen in Figure 2) and $H_3AsO_3$, the molecule

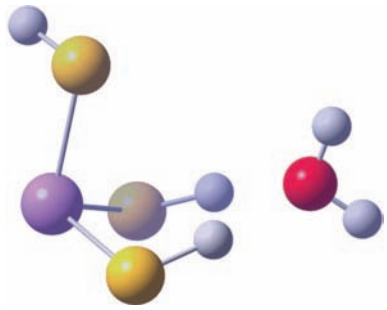

**Figure 5.** Hydrated $H_3AsS_3 \cdot H_2O$ complex at optimized CBS-QB3 geometry calculated in the PCM.

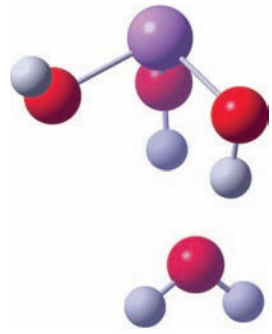

**Figure 6.** Hydrated $H_3AsO_3 \cdot H_2O$ complex at optimized CBS-QB3 geometry calculated in the PCM.

**Figure 7.** Hydrated $H_3AsS_4 \cdot H_2O$ complex at optimized CBS-QB3 geometry calculated in the PCM.

was placed so the oxygen atom of the water molecule interacted with one of the protonating hydrogen atoms of the arsenic compound.

The calculations were then conducted in the three ways detailed in the CBS-QB3 and solvation sections above (2.1 and 2.2): at the CBS-QB3 derived B3LYP/CBSB7 optimized geometries (in the gas phase), with the CPCM water (at the gas phase geometry), and in the full CBS-QB3 CPCM environment.

## 3. Results and Discussion

The CBS-QB3 gas phase reaction $\Delta G$ values were calculated by the standard thermodynamic method:

$$\text{reaction } \Delta G = (\text{sum of } \Delta G\text{s for products}) - (\text{sum of } \Delta G\text{s for reactants})$$

The calculated $\Delta G$ values for each of the compounds in our investigation are reported in the Supporting Information section. We used these to find the gas deprotonation energies in the gas phase and $pK_a$ values for deprotonation in the solvated environment, collected in Table 1. $pK_a$s are determined from the reaction $\Delta G$s by dividing by 1.3644 kcal/mol. For the gas phase CBS-QB3 + CPCM data, we add the gas phase CBS-QB3 energies to the $\Delta G$s of solvation calculated at the gas phase CBS-QB3 geometries to rapidly estimate the solution $pK_a$s, using a $\Delta G$ of −269.0 kcal/mol for the concentration corrected, solvated $H^+$ ion[28] as explained above.

Note that, because we were unable to find a stable gas phase CBS-QB3 geometry for $HAsS_3^{-2}$, we were unable to calculate the gas phase + CPCM $pK_a$s in Table 1 for the second or third deprotonations of the compound. In all cases, starting from a variety of reasonable geometries, $HAsS_3^{-2}$ dissociated into an ($AsS_2^-$/$SH^-$) pair during optimization. This dissociation did not occur in the case of $AsS_3^{-3}$. For comparison purposes, the double-deprotonation reaction was constructed from the available information.

In the cases with explicit hydration of the compounds, we often observed the water molecule finding a local minima where it had two hydrogen-bonding type interactions with the solute as seen in Figures 3−6. In a few cases, the explicit water molecule only interacted with a single molecule of the complex, as seen in Figure 7, of the explicitly hydrated $H_3AsS_4 \cdot H_2O$ complex, as optimized in the CPCM via the CBS-QB3 process. We did not observe back proton transfer in any of the cases we calculated.

While explicit hydration is taking place, it appears that the proton continues to belong to the explicit $H_2O$ molecule. The bond lengths between the oxygen and its hydrogens on the $H_2O$ differ by about 10%, 0.1 Å, between bonding to the explicitly solvated $H_3AsO_4$ and the explicitly solvated $AsO_3^{-3}$, going from an OH bond length of 0.97 to 1.07 Å, and a hydrogen bonding length of 2.08 Å in the uncharged case to 1.70 Å in the most charged case. The highly charged anions have a larger share of the proton than the neutral ones, but the proton is still part of the water molecule.

In cases with an anion, one or two of the oxygen atoms are interacting with the explicit $H_2O$ molecule with hydrogen bonding interaction. The explicitly hydrated, uncharged complexes generally optimized to geometries where there were cyclic hydrogen bonding arrangements; in $H_3AsO_4$ in the CBS-QB3 CPCM-optimized conditions, we saw the water molecule position to have one incoming hydrogen bond from a proton in the complex and one outgoing one to an adjacent oxygen atom on the complex. The most interesting case was probably the $H_3AsS_3$ cluster in the gas phase CBS-QB3 with an associated explicit $H_2O$. In that case, the water molecule moved so that it had two incoming hydrogen bonds to the oxygen atom and what appears to be a loose one going outward to the third oxygen in the cluster.

Table 1contains calculations of the reaction $\Delta G$ and $pK_a$ values calculated in the CPCM throughout the CBS-QB3 process. We suspect the lower predicted $pK_{a3}$ values in the gas phase CBS-QB3 + CPCM solvation model are a result of the large stabilizations that the highly charged, small anionic compounds have when in the CPCM environment. We note that the raw energies calculated by the CBS-QB3 method in the CPCM are significantly lower than that of the two-part approach of combining the gas phase and CPCM calculations done separately (see supplemental table of raw energies), except in the case of the small and highly charged $AsO_4^{-3}$, $AsO_3^{-3}$, and $AsO_3^{-3} \cdot H_2O$ calculations. We suspect this is because in the case of highly charged anions the CPCM

**TABLE 2: $H_3AsO_4$ and $H_3PO_4$ Comparison between Experiment and Theory**

| | exp. p$K_a$[a] | CBS-QB3 + (gas phase) CPCM p$K_a$ | CBS-QB3 CPCM p$K_a$ (all in PCM) | CBS-QB3 + (gas phase) CPCM w/$H_2O$ | CBS-QB3 CPCM p$K_a$ w/$H_2O$ (all in PCM) |
|---|---|---|---|---|---|
| $H_3AsO_4 \rightarrow H_2AsO_4^- + H^+$ | 2.3 | 6.5 | 6.1 | 5.7 | 4.6 |
| $H_2AsO_4^- \rightarrow HAsO_4^{-2} + H^+$ | 7 | 19.6 | 19.3 | 20.3 | 17.7 |
| $HAsO_4^{-2} \rightarrow AsO_4^{-3} + H^+$ | 13 | 21.0 | 28.3 | 21.8 | 28.6 |
| $H_3PO_4 \rightarrow H_2PO_4^- + H^+$ | 2.2 | 6.1 | 4.8 | 6.2 | 4.9 |
| $H_2PO_4^- \rightarrow HPO_4^{-2} + H^+$ | 7.1 | 21.7 | 21.2 | 20.5 | 19.9 |
| $HPO_4^{-2} \rightarrow PO_4^{-3} + H^+$ | 12.3 | 25.0 | 29.8 | 26.2 | 29.8 |

[a] Experimental p$K_a$s from William L. Jolly, *Modern Inorganic Chemistry*, 2nd edition.

**TABLE 3: Calculated p$K_a$s Using CCSD(T) Method**

| | | CCSD(T)/6-311G(2d,p) | | CCSD(T)/aug-cc-pVDZ | | CCSD(T)/CEP-121G | |
|---|---|---|---|---|---|---|---|
| | expt. p$K_a$[a] | gas phase geometry + CPCM solvation energy p$K_a$ | p$K_a$ with CPCM | Gas phase geometry + CPCM solvation energy p$K_a$ | p$K_a$ with CPCM | Gas phase geometry + CPCM solvation energy p$K_a$ | p$K_a$ with CPCM |
| $H_3AsO_4 \rightarrow H_2AsO_4^- + H^+$ | 2.3 | 13.6 | 12.8 | 6.6 | 6.4 | 0.2 | −0.5 |
| $H_2AsO_4\text{-} \rightarrow HAsO_4^{-2} + H^+$ | 7 | 32.6 | 11.0 | 17.6 | 19.7 | 11.8 | 14.1 |
| $HAsO_4^{-2} \rightarrow AsO_4^{-3} + H^+$ | 13 | 50.3 | 43.1 | 19.1 | 29.8 | 23.7 | 33.3 |
| $H_3PO_4 \rightarrow H_2PO_4^- + H^+$ | 2.2 | 12.6 | 27.1 | 5.2 | 4.9 | −2.4 | −1.2 |
| $H_2PO_4\text{-} \rightarrow HPO_4^{-2} + H^+$ | 7.1 | 35.9 | 16.8 | 22.2 | 22.2 | 15.9 | 14.1 |
| $HPO_4^{-2} \rightarrow PO_4^{-3} + H^+$ | 12.3 | 51.1 | 44.1 | 19.7 | 27.8 | 19.1 | 19.6 |

[a] Experimental p$K_a$s from William L. Jolly, *Modern Inorganic Chemistry*, 2nd edition.

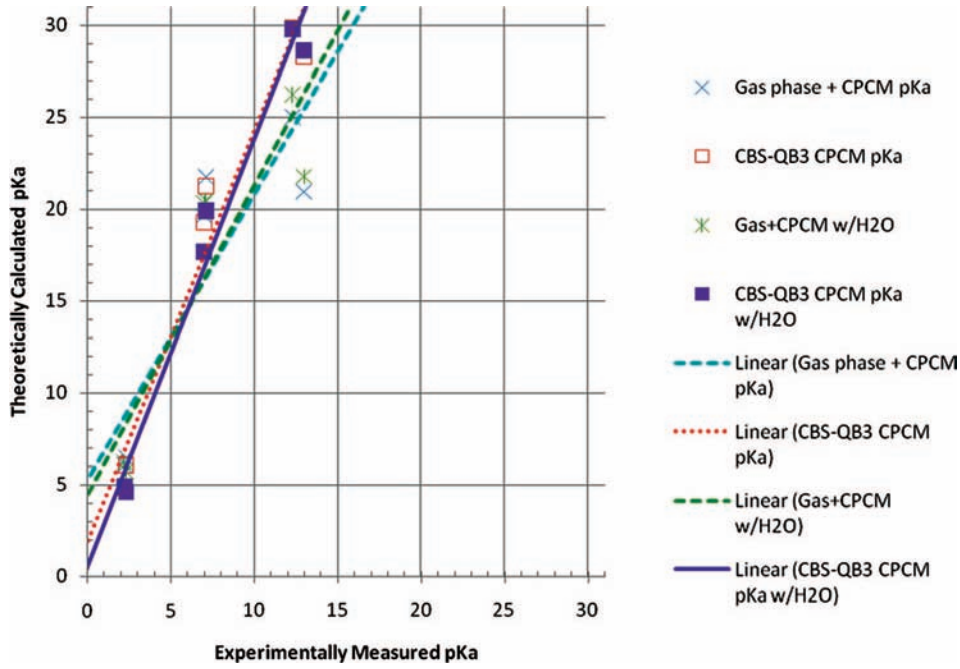


**Figure 8.** Graph of experimental vs theoretical p$K_a$s for As(V) and P(V) species.

contribution to the final energy of the complex is very large. Any slight improvements in geometry caused by doing the CBS-QB3 calculation in the CPCM in these cases are therefore overshadowed by the purely electrostatic stabilization, leading to the observed trend.

To obtain more accurate p$K_a$s we can reference our CBS-QB3 calculated p$K_a$ values for the As species to known ones for measurements of $H_3AsO_4$ and the triprotic acid $H_3PO_4$, as shown in Table 2. Similarly, p$K_a$s from the CCSD(T) calculations are summarized in Table 3 (the calculated CCSD(T) energies are collected in the supplemental Table 3a in the Supporting Information). By graphing the theoretical data against measured experimental values (as seen in the linear extrapolation Figures 8 and 9) it appears that the points for many of them have a linear relation.

By conducting linear extrapolations between the theoretical data and the experimental (see Table 4), we can graph their correspondence in Figure 8 for the CBS-QB3 calculations and Figure 9 for the more linear of the CCSD(T) calculations. The error in using the extrapolation to go from the theoretical calculation to the experimental measurement is the horizontal distance between the marked points and the corresponding colored line. For the CBS-QB3 CPCM p$K_a$ with explicit single-molecule $H_2O$ solvation method, it is evident that those errors are relatively small. Using the extrapolation determined from the $H_3PO_4$ and $H_3AsO_4$ data, we consolidate the fitted p$K_a$ values

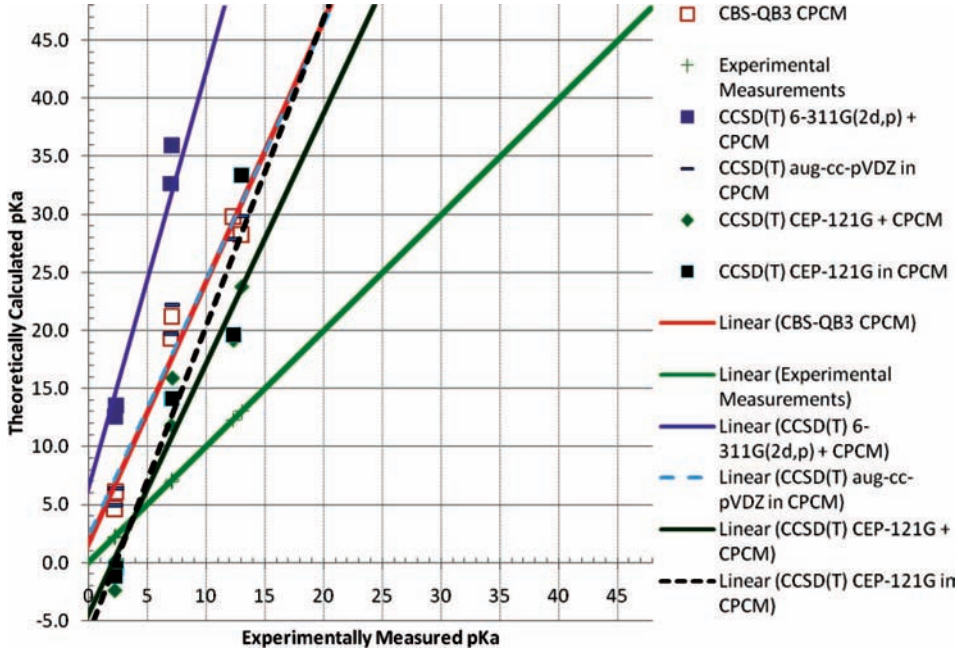


**Figure 9.** Graph of experimental vs theoretical p$K_a$s for As(V) and P(V) species with CCSD(T) method.

**TABLE 4: Linear Extrapolation Slope/Intercepts Converting Theoretical Values to Experimental Ones**

| $y = mx + b$ | CBS-QB3 method | | CBS-QB3 with explicit solvation | | CCSD(T)/6−311G(2d,p) | | CCSD(T)/aug-cc-pVDZ | | CCSD(T)/CEP-121G | |
|---|---|---|---|---|---|---|---|---|---|---|
| $y$ = expt. p$K_a$ $x$ = theo. calculated p$K_a$ | CBS-QB3 + (gas phase geometry's) CPCM stabilization p$K_a$ | CBS-QB3 CPCM p$K_a$ (all structures optimized & energy calculated in PCM) | CBS-QB3 + CPCM w/$H_2O$ | CBS-QB3 CPCM p$K_a$ w/$H_2O$ | + CPCM solvation energy | optimized with CPCM | + CPCM solvation energy | optimized with CPCM | + CPCM solvation energy | Optimized with CPCM |
| $m$ | 0.643 | 0.444 | 0.593 | 0.429 | 0.273 | 0.231 | 0.516 | 0.428 | 0.428 | 0.342 |
| $b$ | −3.398 | −0.767 | −2.635 | −0.233 | −1.593 | 1.353 | −0.462 | −0.597 | 2.448 | 2.795 |
| $r^2$ | 0.778 | 0.952 | 0.824 | 0.974 | 0.979 | 0.542 | 0.649 | 0.944 | 0.923 | 0.907 |

**TABLE 5: Fitted CBS-QB3 CPCM p$K_a$s (with Explicit $H_2O$ Molecule Solvation)**

| | | exp. p$K_a$[a] | CBS-QB3 CPCM p$K_a$ w/$H_2O$ | fitted CBS-QB3 CPCM p$K_a$ w/$H_2O$ |
|---|---|---|---|---|
| $H_3AsO_4$ | p$K_{a1}$ | 2.3 | 4.6 | 1.7 |
| | p$K_{a2}$ | 7 | 17.7 | 7.3 |
| | p$K_{a3}$ | 13 | 28.6 | 12.1 |
| $H_3PO_4$ | p$K_{a1}$ | 2.2 | 4.9 | 1.9 |
| | p$K_{a2}$ | 7.1 | 19.9 | 8.3 |
| | p$K_{a3}$ | 12.3 | 29.8 | 12.6 |
| $H_3AsO_3$ | p$K_{a1}$ | 9.2 | 11.5 | 4.7 |
| | p$K_{a2}$ | | 30.5 | 12.9 |
| | p$K_{a3}$ | | 35.1 | 14.8 |
| $H_3AsS_4$ | p$K_{a1}$ | | −4.6 | −2.2 |
| | p$K_{a2}$ | | 8.3 | 3.3 |
| | p$K_{a3}$ | | 12.9 | 5.3 |
| $H_3AsS_3$ | p$K_{a1}$ | | 5.0 | 1.9 |
| | p$K_{a2}$ | | 13.8 | 5.7 |
| | p$K_{a3}$ | | 18.5 | 7.7 |
| $H_2SO_3$ | p$K_{a1}$ | 1.8 | 3.1 | 1.1 |
| | p$K_{a2}$ | 7.2 | 19.4 | 8.1 |

[a] Experimental p$K_a$s from William L. Jolly, Modern Inorganic Chemistry, Second Edition.

for the other compounds which we studied in Table 5. This fitting method takes care of any systematic errors that might be observed from our choice of the solvated $H^+$ ion energy,[28] as well as ones caused by our choice of method.[10,23]

There was a lack of experimental data for a similar polyprotic acid in the case of the $H_3AsO_3$, $H_3AsS_4$, and $H_3AsS_3$, preventing us from doing a similar extrapolation for our other calculated data, and therefore, the caveat remains that the extrapolated points there are expected to be of lower quality than the former ones, as there are significant chemical differences expected between both oxy- and thioacids and As(V) and As(III). This expectation is confirmed by the single experimental data point we have available, for $H_3AsO_3$, which was experimentally measured to be 9.2, but our fitted model predicted to be 4.7, a difference of 4.5 p$K_a$ units, 3 more than expected. Experimental data for the As(III) oxoacid and the thioacids of arsenic would be useful to evaluate the quality of these predictions.

## 4. Conclusions

The CBS-QB3 (gas phase) + CPCM (gas phase geometry) and CBS-QB3/CPCM (optimizing and calculating in solution) approaches provided quite similar predicted p$K_a$s in all cases except for the smaller, harder anions, $AsO_3^{-3}$ and $AsO_4^{-3}$. The CBS-QB3 (gas phase) + CPCM (gas phase geometry) reports much lower p$K_{a3}$ values in both cases. However, the differences between the two methods for $AsS_3^{-3}$ and $AsS_4^{-3}$ were significantly smaller.

Examining the linear extrapolations in Figure 8 and Table 4, we see that the linear relation for the CBS-QB3 CPCM calculations with explicit $H_2O$ molecule solvation has one of the highest observed $r^2$ values: 0.97. Though the CCSD(T) calculation with the 6-311G(2d,p) basis has a slightly better $r^2$ value, it is clear from a visual comparison (see Figure 9) that the CBS-QB3-based calculation's values are much closer to the

experimentally measured ones than those of the CCSD(T) calculation. Interestingly, the aug-cc-pVDZ calculation at the CBS-QB3 CPCM geometry yielded nearly the same results (and extrapolation) as our best CBS-QB3 calculation, though the intercept was not quite as close to zero as it was for the CBS-QB3 CPCM method.

We would expect that the most accurate algorithm would provide us with the closest linear relation, as in that case the errors would be concentrated in systematic factors. Therefore, we recommend using the CBS-QB3 CPCM method with explicit hydration above the other methods explored in this paper.

As we expect a similar distortion in the results from the other calculations due to solvation effects, we present the extrapolated values in Table 5 with a word of warning; they are almost certainly of lower accuracy than the $H_3AsO_4$ and $H_3PO_4$ points, because the compounds have significant structural and chemical differences.

Of the As(III) species, the $H_3AsS_3$ is likely to be deprotonated at least once in solutions near pH 7. We have been unable to find a stable geometry for the doubly deprotonated species in the gas phase.

It would also be interesting to follow up with an examination of $H_3AsS_3$ and its deprotonation products using molecular mechanics techniques in an explicit solvent. We would predict that hydrogen bonding between the solvent and the molecule would result in deprotonated geometries closer to the original species, lowering the cost for dissociating into more anionic forms.

**Acknowledgment.** We thank NSF Grant EAR0539109 for providing funding for the project, and G. R. Helz for discussion and comments on the arsenic systems.

**Supporting Information Available:** Additional figures, tables, and a collection of geometries for each of the compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

## References and Notes

(1) Duker, A. A.; Carranza, E. J. M.; Hale, A. *Environ. Int.* **2005**, *31*, 631.
(2) Yudovich, Y. E.; Ketris, M. P. *Int. J. Coal Geol.* **2005**, *61*, 141.
(3) Korte, N. E.; Fernando, Q. *Crit. Rev. Environ. Contr.* **1991**, *21*, 1.
(4) Tossell, J. A. *Geochem. Trans.* **2003**, *4*, 28.
(5) Tossell, J. A. *Geochim. Cosmochim. Acta* **1997**, *61*, 1613.
(6) Helz, G. R.; Tossell, J. A.; Charnock, J. M.; Pattrick, R. A. D.; Vaughan, D. J.; Garner, C. D. *Geochim. Cosmochim. Acta* **1995**, *59*, 4591.
(7) Helz, G. R.; Tossell, J. A. *Geochim. Cosmochim. Acta* **2008**, *72*, 4457.
(8) Alexeev, Y.; Windus, T. L.; Dixon, D. A.; Zhan, C. G. *Intl. J. Quantum Chem.* **2005**, *102*, 775.
(9) Alexeev, Y.; Windus, T. L.; Dixon, D. A.; Zhan, C. G. *Intl. J. Quantum Chem.* **2005**, *104*, 379.
(10) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, *110*, 2822.
(11) Frisch, M. J. Trucks, G. W. Schlegel, H. B. Scuseria, G. E. Robb, M. A. Cheeseman, J. R. Montgomery, Jr., J. A. Vreven, T. Kudin, K. N. Burant, J. C. Millam, J. M. Iyengar, S. S. Tomasi, J. Barone, V. Mennucci, B. Cossi, M. Scalmani, G. Rega, N. Petersson, G. A. Nakatsuji, H. Hada, M. Ehara, M. Toyota, K. Fukuda, R. Hasegawa, J. Ishida, M. Nakajima, T. Honda, Y. Kitao, O. Nakai, H. Klene, M. Li, X. Knox, J. E. Hratchian, H. P. 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. Ayala, P. Y. Morokuma, K. Voth, G. A. Salvador, P. Dannenberg, J. J. Zakrzewski, V. G. Dapprich, S. Daniels, A. D. Strain, M. C. Farkas, O. Malick, D. K. Rabuck, A. D. Raghavachari, K. Foresman, J. B. Ortiz, J. V. Cui, Q. Baboul, A. G. Clifford, S. Cioslowski, J. Stefanov, B. B. Liu, G. Liashenko, A. Piskorz, P. Komaromi, I. Martin, R. L. Fox, D. J. Keith, T. Al-Laham, M. A. Peng, C. Y. Nanayakkara, A. Challacombe, M. Gill, P. M. W. Johnson, B. Chen, W. Wong, M. W. Gonzalez, C. and Pople, J. A. *Gaussian 03, revision D.01*; Gaussian, Inc.: Wallingford, CT, 2004.
(12) Stevens, W.; Basch, H; Krauss, J. *J. Chem. Phys.* **1984**, *81*, 6026.
(13) Stevens, W. J.; Krauss, M.; Basch, H.; Jasien, P. G. *Can. J. Chem.* **1992**, *70*, 612.
(14) Cundari, T. R.; Stevens, W. J. *J. Chem. Phys.* **1993**, *98*, 5555.
(15) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(16) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968.
(17) Krishnan, R.; Pople, J. A. *Int. J. Quantum Chem.* **1978**, *14*, 91.
(18) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503.
(19) Gutowski, K. E.; Dixon, D. A. *J. Phys. Chem. A* **2006**, *110*, 12044.
(20) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.
(21) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.
(22) Ramakrishna, V.; Duke, B. J. *J. Chem. Phys.* **2003**, *118*, 6137.
(23) Truong, T. N.; Stefanovich, E. V. *Chem. Phys. Lett.* **1995**, *240*, 253.
(24) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105* (8), 2999.
(25) Gutowski, K. E.; Dixon, D. A. *J. Phys. Chem. A* **2006**, *110*, 8840.
(26) Paine, S. W.; Kresge, A. J.; Salam, A. *J. Phys. Chem. A* **2005**, *109*, 4149.
(27) Dutton, A. S.; Fukuto, J. M.; Houk, K. N. *Inorg. Chem.* **2004**, *43*, 1039.
(28) Liptak, M. D.; Shields, G. C. *Int. J. Quantum Chem.* **2001**, *85*, 727.
(29) Chase, M. W., Jr., Ed.; *NIST-JANAF Thermochemical Tables*, 4th ed.; The American Chemical Society: Washington, DC, and the American Institute of Physics: New York, 1998.
(30) Camaioni, D. *J. Phys. Chem. A* **2005**, *109*, 10795.
(31) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. *J. Phys. Chem. A* **1998**, *102*, 7787.
(32) Zhan, C. G.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 11534.
(33) Pliego, J. R.; Riveros, J. M. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1622.
(34) Pliego, J. R.; Riveros, J. M. *Chem. Phys. Lett.* **2000**, *332*, 597.

JP809123Q