Ab Initio Aqueous Thermochemistry: Application to the Oxidation of Hydroxylamine in Nitric Acid Solution

Robert W. Ashcraft,[†] Sumathy Raman,[‡] and William H. Green*,[†]

Massachusetts Institute of Technology, Department of Chemical Engineering, Cambridge, Massachusetts 02139, and Oakwood College, Department of Chemistry, Huntsville, Alabama 35896

Received: May 9, 2007; In Final Form: July 16, 2007

Ab initio molecular orbital calculations were performed and thermochemical parameters estimated for 46 species involved in the oxidation of hydroxylamine in aqueous nitric acid solution. Solution-phase properties were estimated using the several levels of theory in Gaussian03 and using COSMO*therm*. The use of computational chemistry calculations for the estimation of physical properties and constants in solution is addressed. The connection between the pseudochemical potential of Ben-Naim and the traditional standard state-based thermochemistry is shown, and the connection of these ideas to computational chemistry results is established. This theoretical framework provides a basis for the practical use of the solution-phase computational chemistry estimates for real systems, without the implicit assumptions that often hide the nuances of solution-phase thermochemistry. The effect of nonidealities and a method to account for them is also discussed. A method is presented for estimating the solvation enthalpy and entropy for dilute aqueous solutions based on the solvation free energy from the ab initio calculations. The accuracy of the estimated thermochemical parameters was determined through comparison with (i) enthalpies of formation in the gas phase and in solution, (ii) Henry's law data for aqueous solutions, and (iii) various reaction equilibria in aqueous solution. Typical mean absolute deviations (MAD) for the solvation free energy in room-temperature water appear to be ∼1.5 kcal/mol for most methods investigated. The MAD for computed enthalpies of formation in solution was 1.5−3 kcal/mol, depending on the methodology employed and the type of species (ion, radical, closed-shell) being computed. This work provides a relatively simple and unambiguous approach that can be used to estimate the thermochemical parameters needed to build detailed ab initio kinetic models of systems in aqueous solution. Technical challenges that limit the accuracy of the estimates are highlighted.

Introduction

Hydroxylamine is an important reagent used in the separation of plutonium and uranium in the PUREX process. It is used to reduce the plutonium from Pu(IV) to Pu(III), causing a change in solubility and the transfer of the plutonium from an organic phase to an aqueous phase. The aqueous solution used in these separation systems is a concentrated nitric acid solution containing hydroxylamine, trace amounts of nitrous acid, and hydrazine as a stabilizing agent. The nitrous acid is an impurity present in nitric acid solutions and acts as a catalyst in the oxidation of hydroxylamine by nitric acid. This work follows from a previous publication by Raman et al.,[1] where a more detailed introduction and background can be found; a brief introduction is presented here.

It was found that hydroxylamine can react autocatalytically with nitric and nitrous acids (R-2) or act as a nitrous acid scavenger (R-1).[2−6]

$$NH_3OH^+ + HONO \rightarrow N_2O + H_2O + H_3O^+ \quad \text{(R-1)}$$

Under certain conditions, the combination of (R-1) and (R-2) can lead to rapid evolution of N_2O(g); this is thought to have

$$NH_3OH^+ + 2HONO_2 + 2HONO \rightarrow 5HONO + H_3O^+ \quad \text{(R-2)}$$

contributed to several vessel ruptures at nuclear facilities.[7] The experimental investigation of these overall reactions has led to a proposed reaction mechanism for each. The scavenging reaction (R-1) was proposed to occur through the isomerization of an O-nitrosated intermediate, reactions R-3−R-7.[3] It has also been suggested that the key isomerization step, (R-5), may occur through the unprotonated form, NH_2ONO, which has a free lone pair that can act as the NO acceptor.[5]

$$H^+ + HONO \rightleftharpoons H_2O + NO^+ \quad \text{(R-3)}$$

$$NH_3OH^+ + NO^+ \rightleftharpoons NH_3ONO^+ + H^+ \quad \text{(R-4)}$$

$$NH_3ONO^+ \rightarrow ONNH_2OH^+ \quad \text{(R-5)}$$

$$ONNH_2OH^+ \rightarrow HONNOH + H^+ \quad \text{(R-6)}$$

$$HONNOH \rightarrow N_2O + H_2O \quad \text{(R-7)}$$

A reaction scheme has also been proposed for the autocatalytic reaction (R-2) by Pembridge and Stedman, (R-8)−(R-12).[2]

$$H^+ + HONO + NO_3^- \overset{\text{fast}}{\rightleftharpoons} N_2O_4 + H_2O \quad \text{(R-8)}$$

* Corresponding author. E-mail: whgreen@mit.edu. Phone: +1.617.253.4580. Fax: +1.617.324.0066.

[†] Massachusetts Institute of Technology, Dept. of Chemical Engineering, Cambridge, MA 02139

[‡] Oakwood College, Department of Chemistry, Kuntsville, AL 35896

10.1021/jp073539t CCC: $37.00
Published on Web 09/22/2007

$$NH_2OH + N_2O_4 \xrightarrow{slow} HNO + N_2O_3 + H_2O \quad (R\text{-}9)$$

$$HNO + N_2O_4 \rightarrow N_2O_3 + HONO \quad (R\text{-}10)$$

$$N_2O_3 + H_2O \overset{fast}{\rightleftharpoons} 2HONO \quad (R\text{-}11)$$

$$NH_3OH^+ \rightleftharpoons NH_2OH + H^+ \quad (R\text{-}12)$$

Pembridge and Stedman note that there are likely a number of other elementary steps that occur, but the available experimental data did not allow them to gain any meaningful insight into what they may be. Many of the proposed intermediates have not been observed experimentally and almost no elementary step rate constants are available, so although this mechanism seems reasonable, the exact pathway that converts NH_2OH to HONO is still very much uncertain. Prior researchers were able to find a set of rate constants which gave reasonable fits to experimental data measured with 273 K < T < 298 K. However, the PUREX process is operated between 313 and 333 K, and extrapolation of the empirically fit, simplistic model would be unwise in such an unstable and potentially hazardous system.

There are a number of other researchers working on related topics utilizing ab initio techniques that may be useful in understanding this system. Tomasi et al. have developed some of the continuum solvation methods used in this work.[8,9] Klamt and Eckert have developed the COSMO-RS theory as an extension to some of the continuum models discussed here.[10,11] The Truhlar and Cramer groups are developing an extension to the continuum models able to more accurately estimate the solvation energies as a function of temperature in aqueous solutions.[12,13] Many researchers have examined the ability of continuum models to predict aqueous acidity and basicity constants, including Chipman[14] and Klamt et al.[15] Leung et al. have examined the entropy in solution and proposed a breakdown of the solvation entropy into cavitation and polarization contributions, which is an idea drawn upon in the current work.[16] The cavitation free energy will be shown to be an important part of this work and several methods of estimating it have been investigated.[17−27]

The present work aims to use and build upon the above research to further understand the thermochemistry of elementary processes in solution and to lay the groundwork for further efforts in modeling solution-phase dynamics. This relatively complicated system also provides a test of the methodology for computing solvation enthalpies and solvation entropies. A theoretical outline of solution-phase thermochemistry is given, including the connection between pseudochemical potentials, standard state chemical potentials, and computational chemistry results. Thermochemical parameters for potentially important species are derived using the continuum solvation model available in Gaussian03 at several levels of theory.

Theoretical Basis: Solvation Thermochemistry

The treatment of solvent effects on the thermochemistry and kinetics of systems in condensed phases has been studied theoretically and empirically in a number of fields. However, often ambiguities present in the underlying assumptions and confusion with regard to standard states and reference states can cause difficulty when attempting to use the results correctly. An example of where confusion arises could be in the typical concentration-based equilibrium constant (K_C), which is a physical property of the system. However, the free energy change of reaction that it is related to in phenomenological thermodynamics is the standard state change (ΔG^{o}_{Rxn}), because the actual change in free energy at equilibrium must be zero. This means that there are an infinite number of ΔG^{o}_{Rxn} values, because the standard state can be chosen arbitrarily. The same problem is seen for gas-phase systems, but it is less problematic because there is nearly universal agreement on the standard state and the zeros of enthalpy and entropy. In solution thermochemistry there are many standard state and zero-of-energy conventions used. Often the solutions are far from ideal at the chosen standard state, and it is unclear what standard state or zero-of-energy was used in either the computations or in the analysis of the experimental data. An attempt is made here to explicitly define relationships in condensed phases between certain experimentally accessible properties and the thermodynamic variables that define them, with a particular emphasis on the relation to computational chemistry estimates. Equilibrium relationships for general vapor−liquid equilibria (VLE) and reactions in condensed phases are outlined in terms of the typical standard state thermodynamic quantities, as well as the pseudochemical potential (μ_i^* *or* G_i^*) of Ben-Naim.

Prior to delving into the equilibrium relationships, it is important to understand a few important aspects of the free energy and solvation. The first major point is that the free energy is dependent on the concentration of the species of interest. This dependence comes from the translational entropy and is directly related to the work required to compress a gas or selectively change the concentration of a noninteracting solute. For an isothermal process, as is the case in most equilibrium situations, the simple relationship between the Gibbs free energy (G_i) and the concentration (C_i) is shown in eq 13. This implicitly assumes a situation where the nonideality of the mixture does not change when the concentration is changed from C_i^+ to C_i; the more detailed discussion below will address this point. This idea is important when comparing computational results with experimental data, which often assumes a standard state of 1 mol/L for solution-phase properties. The computational results are often at a concentration equivalent to an ideal gas at a pressure of 1 atm, so one must correct the free energy and entropy estimates to a concentration of 1 M before a comparison can be made.

$$G_i(C_i) = G_i(C_i^+) + RT \ln\left(\frac{C_i}{C_i^+}\right) + (\text{nonideality}) \quad (13)$$

The pseudochemical potential (PCP) as defined by Ben-Naim is the chemical potential (CP) of a solute confined to a fixed position, and is denoted by the "*".[28,29] The concentration dependence due to the translation is no longer present in a PCP; however, any concentration dependence of the chemical potential due to nonideal interactions would be. As will be seen below, one benefit of the PCP is that it allows for a direct correspondence with experimental measurements, without the need for standard states. There are other subtleties to the PCP definition and use, and the reader is encouraged to seek the aforementioned references for more information. Equation 14 relates the traditional CP to the PCP as defined by Ben-Naim. Equation 15 is a useful relationship that will be used to convert between the standard state CP and the PCP, where ρ is the number density (or molar concentration) and Λ^3 is the ideal gas momentum partition function. Equation 15 is derived by equating the traditional and PCP-based expressions for the chemical potential.

$$\mu_i^{\#} = \mu_i^{*\#} + kT \ln(\rho_i^{\#} \Lambda_i^{3}) \quad (14)$$

$$\text{solution phase}\quad \mu_i^{\circ} = \mu_i^{*\#} + kT\ln(\rho_i^{\circ}\Lambda_i^{3}) - kT\ln\left(\frac{\gamma_i^{\#,+}}{\gamma_i^{\circ,+}}\right) \quad \text{or}$$
$$\text{gas phase}\quad \mu_i^{\circ} = \mu_i^{*\#} + kT\ln(\rho_i^{\circ}\Lambda_i^{3}) - kT\ln\left(\frac{\phi_i^{\#,+}}{\phi_i^{\circ,+}}\right) \tag{15}$$

The "state" and "behavior" terminologies will be used here regarding references, and it is important to understand the difference. Reference *behavior* describes the variations of fugacity over the entire mole fraction range, $\hat{f}_i^{+}(x_i)$, which is often linear as with Henry's Law or ideal solution behavior. A reference *state* is a single point evaluated at a specific mole fraction along the reference behavior line, e.g., $\hat{f}_i^{+}$ ($x_i = 1$). In all equations, an "o" superscript will be used to signify the standard state, a "+" will be used to signify an arbitrary reference state, a "#" will be used for the actual state of the mixture, an "∞" will be used for the dilute-limit behavior reference, and "pure" will be used for the pure *i* reference state. For example, $\mu_i^{*\#}$ is the PCP at the actual state of the mixture, and μ_i° is the CP at the standard state conditions. The symbol x_i will be used for condensed-phase mole fractions, and y_i will used for gas-phase mole fractions (or p_i for the partial pressure). Unless otherwise noted, the condensed-phase fugacity reference behavior will be defined by a linear extrapolation of dilute-limit or ideal solution behavior. The analysis is not requisite on this assumption, but having a linear reference behavior simplifies some equations. The terms $\phi_i^{\beta,\alpha}$ or $\gamma_i^{\beta,\alpha}$ represent the correction that must be made to the behavior at state "α" to achieve the behavior at state "*β*" and are known as fugacity (gas phase) and activity coefficients (condensed phase). All activity coefficients discussed here are based on the mole fraction concentration scale, as is typical in chemical engineering thermodynamics using the fugacity formalism.

An example derivation has been provided in the Supporting Information for a simple, solution-phase reaction equilibrium constant. If the reader is at all unfamiliar with the procedures and typical assumptions made in such a derivation, they are strongly encouraged to look over it to better understand where the following equations originate.

Vapor−Liquid Equilibrium. VLE is an important phenomenon in life and in many areas of research. The result is a large amount of high-quality data that can be used to determine the quality of predictions from a priori quantum chemical calculations. In general, one may think of VLE having three major regimes with regard to solute concentration: the dilute limit as described by Henry's law, the pure solute limit that is characterized by the vapor pressure, and the intermediate regime where concentration-dependent nonidealities are important. A single equation is sufficient to describe all VLE regimes, and certain assumptions can be made to simplify it on the basis of the regime of interest. Equation 16 shows the relationship in terms of the standard state chemical potential difference, and eq 17 gives it in terms of the pseudochemical potential difference. *It was assumed that the same reference state was used for the standard state and actual state in solution. The gas-phase reference state was assumed to be an ideal gas at the total pressure of the system.* The only additional assumption that had to be made was that *the total concentration of the solution (C_{soln}) is the same for the standard state and the actual state*. It is noteworthy that the PCP difference can be directly related to the actual state of the system, in this case the partial pressure of the gas and the concentration of the solute in the condensed phase, which allows one to see the allure of the PCP definition. However, the drawback to the PCP approach is that all of the details are hidden in the ΔG^{*}_{solv} term, meaning that it cannot be standardized by choosing a standard state and making assumptions about nonideal behavior. The disadvantage of the standard state approach is that one must be able to characterize the nonidealities of the solution in the actual state and the standard state to derive strictly correct solvation energies.

$$\exp\left(\frac{-\Delta G^{\circ}_{solv}}{RT}\right) = \exp\left(\frac{-(\mu^{\circ}_{i,liq} - \mu^{\circ}_{i,gas})}{RT}\right) = \left(\frac{x_i^{\#}\gamma_i^{\#,+}}{x_i^{\circ}\gamma_i^{\circ,+}}\right)\cdot\left(\frac{p_i^{\circ}\phi_i^{\circ,+}}{p_i^{\#}\phi_i^{\#,+}}\right) \tag{16}$$

$$\exp\left(\frac{-\Delta G^{*}_{solv}}{RT}\right) = \exp\left(\frac{-(\mu^{*}_{i,liq} - \mu^{*}_{i,gas})}{RT}\right) = \left(\frac{x_i^{\#}}{p_i^{\#}}\right)(RT\cdot C_{soln}) \tag{17}$$

It is relatively straightforward to define the Henry's Law and vapor pressure relationships by applying additional assumptions to the general equations. A relationship between the Henry's Law constant and the solvation energies can be obtained by making the following assumptions: *the reference state is based on Henry's Law behavior, the actual concentration is very dilute in solution* ($\gamma_i^{\#,\infty} = 1$), *and the gas phase behaves ideally*. This allows one to arrive at eq 18 when the experimental observable **H** is defined as p_i divided by x_i.

$$\mathbf{H} = \left(\frac{p_i^{\circ}}{x_i^{\circ}\gamma_i^{\circ,\infty}}\right)\cdot\exp\left(\frac{\Delta G^{\circ}_{solv}}{RT}\right) \quad \text{or} \quad \mathbf{H} = (RT\cdot C_{soln})\cdot\exp\left(\frac{\Delta G^{*}_{solv}}{RT}\right) \tag{18}$$

The vapor pressure relationship for a pure species can be obtained by applying the following assumptions: *the reference is based on Lewis−Randall (ideal) solution behavior* ($\gamma_i^{\#,pure} = 1$) *and the vapor phase behaves ideally*. The result is eq 19 describing the vapor pressure as a function of the solvation energy.

$$p_i^{vap} = \left(\frac{p_i^{\circ}}{x_i^{\circ}\gamma_i^{\circ,pure}}\right)\cdot\exp\left(\frac{\Delta G^{\circ}_{solv}}{RT}\right) \quad \text{or} \quad p_i^{vap} = (RT\cdot C_{soln})\cdot\exp\left(\frac{\Delta G^{*}_{solv}}{RT}\right) \tag{19}$$

The activity coefficient needed to relate the standard state and reference state behavior continues to be a problem because it is typically difficult to estimate accurately. To exacerbate the problem, the commonly chosen standard state of 1 M will often not obey the dilute or ideal solution behavior taken as the reference, making estimation of the activity coefficient necessary to derive a true standard state free energy change. Unless there are many experimental data available that were actually measured at 1 M, it would generally be better to choose a more dilute standard state. As shown below, the ΔG°_{solv}'s reported by Gaussian03 are of this type, using a standard state of ~0.04 M (1 atm gas equivalent) and 298 K. Note that although ΔG°_{solv} in eqs 18 and 19 have the same numerical value

assuming the standard states were chosen to be the same, ΔG^*_{solv} usually varies significantly and nonlinearly as the solute concentration is varied because the nonidealities are implicitly included. Yet another option is to choose a standard state behavior that is either a linear extrapolation from the dilute limit (for Henry's law) or is on a line connecting the zero-fugacity point with the pure solute fugacity at a mole fraction of one (for vapor pressure), and report ΔG^+_{solv} instead of ΔG^o_{solv}. This is essentially an idealized version of the standard state, allowing the activity coefficient, $\gamma_i^{o,\infty}$ or $\gamma_i^{o,pure}$, to be ignored. However, the solvation energy used to predict **H** or p_i^{vap} must be measured/calculated at conditions where the idealized behavior is valid; otherwise systematic errors will be introduced. Similarly, a ΔG^+_{solv} derived from **H** or p_i^{vap} when the activity coefficient is ignored will only be valid when the mixture obeys the dilute- and pure-limit behaviors, respectively. Because the complex nonidealities are accounted for externally with the activity coefficient, ΔG^+_{solv} depends linearly on the log of the concentration.

Chemical Reaction Equilibria. The correct computation of equilibrium constants is essential to properly constructing a detailed kinetic model, and understanding exactly what assumptions are being made is a major part of understanding potential errors. The expression for the equilibrium constant for a reaction is derived in much the same way as VLE. The expressions for the concentration-based equilibrium constant for a generic reaction are shown in eqs 20 and 21 in terms of a standard state and the pseudochemical potential differences, respectively. As before, the total concentrations at the actual state and standard states of the system were assumed to be equal.

$$K_C = \exp\left(\frac{-\Delta G^o_{Rxn}}{RT}\right)\cdot\left(\frac{\prod_{products} C_i^o}{\prod_{reactants} C_i^o}\right)\cdot\left(\frac{\prod_{reactants} \frac{\gamma_i^{\#,+}}{\gamma_i^{o,+}}}{\prod_{products} \frac{\gamma_i^{\#,+}}{\gamma_i^{o,+}}}\right) \quad (20)$$

$$K_C = \exp\left(\frac{-\Delta G^*_{Rxn}}{kT}\right)\cdot\left(\frac{\prod_{reactants} \Lambda_i^3}{\prod_{products} \Lambda_i^3}\right) \quad (21)$$

The similarity between the reaction equilibria and VLE should be obvious because a reaction is a generalization of VLE. As seen previously, if one uses the standard state free energy of reaction to estimate the equilibrium constant, then one must also use the accompanying standard state concentrations for each species and the activity coefficients for the actual and standard states relative to some arbitrary reference state. When defined in terms of the PCP difference, one has no need for activity coefficients because they are implicit in the ΔG^*_{Rxn} value. Because the nonidealities in both of the above equations must be taken into account, estimating the equilibrium constant from computational chemistry has equal difficulty regardless of which definition is used. Typical continuum-dielectric methods give only very rough estimates of the activity coefficients; however, the methodology introduced by COSMO-RS theory provides a more sophisticated way to estimate these higher-order effects for real solution conditions.

ΔG^o and ΔG^* from Computational Chemistry. Estimating the change in free energy for a reaction in solution or a solvation process was a major goal of this work. Having outlined the relationships above, the remaining task involves estimating the free energy change from computational chemistry calculations. This will be primarily concerned with data from continuum calculations, with nonideal corrections accounted for by activity coefficients from experimental data, empirical methods, or COSMO-RS theory. In most cases, continuum solution-phase calculations are performed in the infinite dilution limit, essentially an isolated molecule in a continuum that is meant to mimic the solvent. What this implies is that the only condition under which one could hope to estimate an accurate free energy change with a continuum model alone would be for solute concentrations in the dilute limit immersed in a pure solvent. In this unique situation, it is relatively straightforward to estimate the standard state free energy change or PCP difference, although the accuracy of such estimates is still subject to the general errors introduced by the continuum approximations. The standard state free energy change can be computed by directly taking the difference of the free energy values computed in the PCM, with the standard state concentration defined by the pressure (usually 1 atm) under which the partition functions were calculated. If the reference is taken to be dilute-limit behavior and the actual and standard state concentrations are small enough to also be in the dilute limit, then $\gamma_i^{\#,\infty} = \gamma_i^{o,\infty} = 1$. Taking a dilute-limit reference state is not necessary, so long as the standard state and actual state exhibit the same degree of nonideality, in which case $\gamma_i^{\#,+}/\gamma_i^{o,+} = 1$. With these criteria met, the free energy change is easily computed using the free energy values and standard state concentrations. The PCP difference under these conditions is also relatively easy to calculate from the computational chemistry output. In this case, ΔG^* is obtained by taking the free energy difference but deleting the translational partition function contribution ($RT \ln(C_i\Lambda_i^3)$). Because the computational chemistry results were assumed valid for the dilute limit in a pure solvent, these are the conditions where the PCP difference would be valid.

Additional information about the nonideality of the system is needed to estimate properties for significant solute concentrations and/or when the solvent is not pure. In general, continuum models cannot give insight into these effects and additional resources are required. For a real solution, it may be useful to think about the nonideality correction as a two-step process. Step one serves to bring the solute from the pure solvent of the quantum calculation to the actual solvent condition, with the solute being in the dilute limit in both cases. Step two takes the solute from the dilute limit in the real solvent to the real solute concentration in the solution of interest. Breaking it up this way highlights an important point about the dilute-limit activity coefficient, specifically, that $\gamma_i^{\#,\infty} = 1$ does not hold for dilute solute concentrations in any solvent. For example, if the reference state is taken to be the dilute limit in pure water, then the activity coefficient for a dilute-limit solute concentration in 3 M nitric acid solution will not be one because the chemical potential of the solute will be different in both cases. In general, the aqueous PCM calculations that are used to derive free energy changes are performed in a pure water solvent, not the real solvent mixture, making it normal to choose the reference state and standard state such that $\gamma_i^{o,+} \rightarrow 1$, which implies a dilute solute in pure water for both states. If the real solution condition is significantly different from pure water, then it should be clear that one cannot choose a reference state such that $\gamma_i^{o,+} \rightarrow 1$ and $\gamma_i^{\#,+} \rightarrow 1$. What this means is that the estimation of a reaction equilibrium constant in an impure solvent will require some estimate of the nonideality of the solution, be it for the standard

state condition or the actual state or both. To get a better idea of the assumptions that must be made, it is useful to parse the equilibrium constant expression further. Equation 22 shows another representation of the solution-phase equilibrium constant that is defined in terms of the gas-phase standard state free energy change, PCP solvation energies at infinite dilution in pure water, and the two-step activity coefficient corrections. The notation assumes that the quantum calculations were performed in pure water, and the actual solution is an aqueous nitric acid mixture.

$$K_C = \left(\frac{\prod_{\text{products}} \frac{p_i^o}{RT}}{\prod_{\text{reactants}} \frac{p_i^o}{RT}} \cdot \exp\left(\frac{-\Delta G^o_{\text{Rxn,gas}}}{RT} \right) \times \exp\left(\frac{-\Delta (\Delta G^{*,\infty H_2O}_{\text{solv}})_{\text{Rxn}}}{RT} \right) \right) \left(\frac{\prod_{\text{reactants}} \gamma_i^{\infty HNO_3,\infty H_2O} \cdot \gamma_i^{\#,\infty HNO_3}}{\prod_{\text{products}} \gamma_i^{\infty HNO_3,\infty H_2O} \cdot \gamma_i^{\#,\infty HNO_3}} \right) \quad (22)$$

This shows the corrections needed to convert the gas-phase equilibrium constant (the first two terms of the RHS) into a solution-phase K_C. The standard state activity coefficient term has been ignored in eq 22 because the gas-phase standard state is taken to be in the dilute limit, such that the species are essentially at the condensed-phase reference state once the constant-concentration solvation process has been completed. In theory, the first three terms can be obtained easily from computational chemistry, with the first two being rather trivial and well established. The $\Delta G^{*,\infty H_2O}_{\text{solv}}$ term is the PCP-based solvation energy as calculated at the dilute limit in water, which are the natural conditions in a PCM calculation. Because the translational partition function is assumed to be same in the gas and condensed phases (and is calculated as such in Gaussian03) and will cancel out, in practice this value can be obtained directly by taking the difference of the free energies from the PCM and vacuum calculations, assuming that the statistical mechanical formulas were applied at the same pressure/concentration. The final term embodies the nonideal contributions and is difficult to estimate, where $\gamma_i^{\infty HNO_3,\infty H_2O}$ represents the correction to take the solute from pure water to the real solution in the dilute limit, and $\gamma_i^{\#,\infty HNO_3}$ represents the correction to go from the dilute limit to the actual concentration in the real solvent. One method to obtain these activity coefficients would be using COSMO-RS theory as implemented in COSMO*therm*, which can provide an estimate of the PCP value at any system composition. The activity coefficient to correct from state α to state β can then be calculated using eq 23.

$$\gamma_i^{\beta,\alpha} = \exp\left(\frac{\mu_i^{*,\beta} - \mu_i^{*,\alpha}}{RT} \right) \quad (23)$$

Activity coefficients may also be derived from experimental data when available or from empirical methods assuming that parameter values can be found. The above discussion outlines the procedure for estimating equilibrium properties for condensed-phase systems from computational chemistry calculations using continuum models of the solvent and activity coefficients to account for nonidealities of the solution. These ideas can be used in the construction and implementation of detailed, solution-phase kinetics mechanisms that are thermodynamically consistent.

Computational Methodology

Gas-Phase Thermochemistry. The Gaussian03[30] suite of programs was used to perform the ab initio calculations. Gas-phase geometry optimization and frequency calculations were performed at the B3LYP/6-311G(2d,d,p) (B3LYP/CBSB7) level of theory as implemented in the CBS-QB3 method. The gas-phase heats of formation were based on the energies derived using the CBS-QB3 method, which is a compound method that seeks to approximate the complete basis set limit.[31] They were estimated through the commonly used atomization method, which includes spin−orbit and bond additivity corrections (BACs) from Petersson et al.[32] Only two of these BACs, O−H and N−H, were applicable to the species studied here. As discussed below, additional bond additivity corrections specific to this system were derived from the limited amount of experimental data available by minimizing the sum of squares of errors for the heats of formation.

Solution-Phase Thermochemistry. Solution-phase calculations were completed using several levels of theory available in Gaussian03. The IEFPCM solvation method with the UAHF radii set was used with the following levels of theory: B3LYP/6-311G(2d,d,p), B3LYP/6-311++G(3df,3pd), and CBS-QB3. The conductor-like screening model (COSMO) with the Klamt radii set was used with B3LYP/6-311G(2d,d,p) and BP86/TZVP/DGA1. The COSMO-RS model from Gaussian03 was also used at the BP86/TZVP/DGA1 level of theory. The COSMO-RS keyword, which means COSMO for real solvents, is used in Gaussian03 to generate the data that COSMO*therm* needs to perform "real solvent" calculations and results in slightly different energies than the basic COSMO keyword. However, the COSMO-RS keyword does not result in true COSMO-RS theory solvation energies; COSMO*therm* was used for this purpose. COSMO*therm*[33,34] was also used to estimate solvation energies at the BP86/TZVP/DGA1 level of theory. COSMO*therm* is a software package that takes surface charge and other molecular data from a quantum chemical calculation and uses it in the context of the statistical mechanics of interacting surfaces to estimate interaction energies between a solute and solvent mixture. The resulting energies are presumably more accurate than a typical PCM calculation because the real character of the solvent is taken into account, and not a simple, homogeneous continuum. Because it estimates solution-phase interactions, it can be used to estimate many properties, such as Henry's law constants, vapor pressures, activity coefficients, and solubility limits. It is also important understand that COSMO*therm* is parametrized to a certain degree for a given level of theory. The theory is described in detail elsewhere, and work has been done to validate its applicability to a range of systems.[10,11,15,33−35] In addition to solvation energy estimates, it could be a useful tool in estimating activity coefficients to correct dilute-limit solvation energies from quantum chemical calculations to solvation energies for more concentrated systems.

The way in which various computational chemistry packages estimate thermochemistry in solution is often unclear. There are three main approaches for estimating solvation energies: (1) empirical group or descriptor-based models such as UNIFAC or the methods developed by Abraham,[36] (2) explicit-solvent models such as Monte Carlo methods, usually employing empirical force fields, and (3) polarizable or conducting continuum quantum chemistry models. Often these approaches

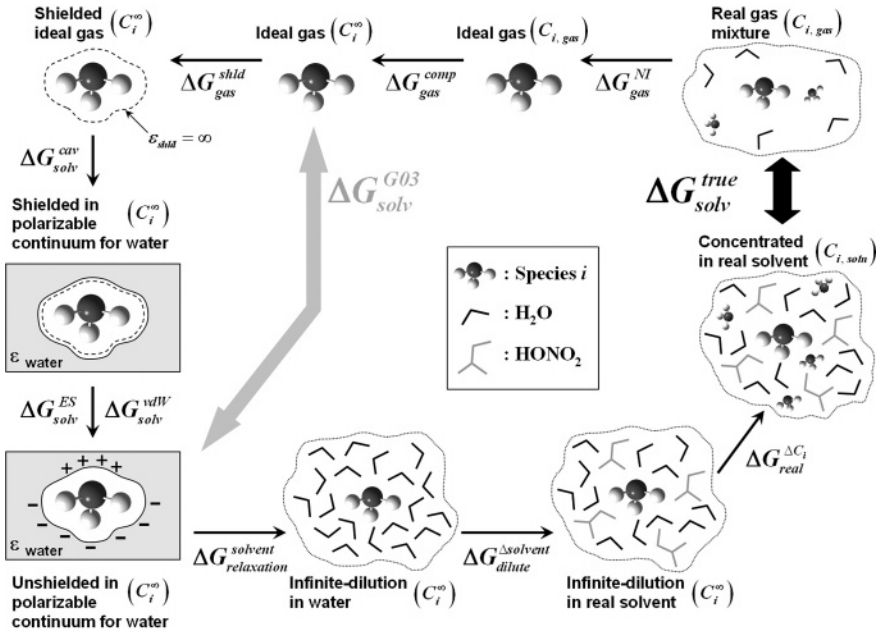

Figure 1. Thermochemical cycle for the solvation of species *i* from a real gas to a real solution.

are combined, e.g., QM/MM or adding group contributions to quantum mechanics. The different software packages often report different quantities, in part because of differences in the assumed standard states, and because most packages cannot compute all the contributions to solvation energies or have different ways of including empirical corrections. For concreteness, the discussion here will address how energies are estimated in Gaussian03 and how that compares to a more realistic solvation energy; similar issues arise with most other approaches.

The general solvation process can be broken down into eight hypothetical steps to transfer a molecule from a real gas mixture into a real solution. A schematic of this process is shown in Figure 1, for the solvation of species *i* in an aqueous nitric acid solution, for example. The processed is assumed to be isothermal, as would be the case at equilibrium.

(i) ΔG_{gas}^{NI}: The gas-phase mixture must first be taken from the real, nonideal state with concentration C_i to an ideal gas mixture at a concentration C_i. The work associated with this process is ΔG_{gas}^{NI} and removes the nonidealities of the gas mixture.

(ii) ΔG_{gas}^{comp}: The ideal gas mixture must then be expanded, reducing the concentration of the species from $C_{i,gas} \rightarrow C_i^\infty$. The dilute concentration, C_i^∞, is arbitrarily low such that the molecule can be treated as isolated.

(iii) ΔG_{gas}^{shld}: The molecule is wrapped in a hypothetical shell that blocks all interactions between molecules. This serves to screen the molecule and its charge from the surrounding environment. Because the molecule is already in the noninteracting ideal gas phase, $\Delta G_{gas}^{shld} = 0$.

(iv) ΔG_{solv}^{cav}: The shielded molecule is inserted into a polarizable continuum model of the solvent, water in this case. The solvent has a much lower free volume than the gas phase, resulting in a loss of entropy. The amount of entropy loss depends on the size of the solute molecule.

(v) ΔG_{solv}^{ES} and ΔG_{solv}^{vdW}: The shield is removed from the molecule, allowing for electronic interactions between the continuum and solute. Two free energy terms are associated with this process, the electrostatic solvation energy, ΔG_{solv}^{ES}, and the solvation energy arising from van der Waals interactions, ΔG_{solv}^{vdW}. The van der Waals interactions are typically written as repulsion and dispersion terms ($\Delta G_{solv}^{vdW} = \Delta G_{solv}^{rep} + \Delta G_{solv}^{disp}$).[9] Typically, the user allows the solute geometry to relax in response to these interactions, and the geometry of the cavity in the dielectric may also be adjusted.

(vi) $\Delta G_{relaxation}^{solvent}$: The process of converting the continuum model of the pure solvent into a discrete version of the pure solvent involves a solvent relaxation energy, $\Delta G_{relaxation}^{solvent}$. When the real solvent molecules are introduced, the solute and solvent orientations will change because the continuum model cannot accurately capture the discreet interactions. This reorientation will likely alter the electrostatic and van der Waals interactions, as well as introduce entropic and enthalpic changes within the local solvent molecules. These effects can be probed using Monte Carlo or molecular dynamics approaches.[37,38]

(vii) $\Delta G_{dilute}^{\Delta Solvent}$: The process of changing from the dilute limit in pure water to the dilute limit in a more realistic solvent will have an energy change dependent on the real solvent composition. The $\Delta G_{dilute}^{\Delta Solvent}$ term embodies the change in chemical potential of the solute when the solvent composition is changed and corresponds to the $\gamma_i^{\infty HNO_3,\infty H_2O}$ activity coefficient discussed earlier. Note that this represents the effect of the mixed solvent on the solute, not the large energy change associated with preparation of this mixed solvent in the absence of the solute.

(viii) $\Delta G_{real}^{\Delta C_i}$: The final step involves increasing the concentration of the solute from $C_i^\infty \rightarrow C_{i,soln}$ to create the true solution-phase mixture; for an ideal solution: $\Delta G_{real}^{\Delta C_i} \approx -RT \ln(C_i^\infty / C_{i,soln})$. This process is similar to step vii, but the excess energy change due to nonidealities will be mainly attributable to solute−solute interactions. This will also take the form of an activity coefficient, referred to as $\gamma_i^{\#,\infty HNO_3}$, where "#" signifies the true solution-phase mixture composition.

The free energy in solution at an arbitrary temperature can be represented by eq 24. Equation 25 expands the solvation energy in the same manner as in the thermochemical cycle, while breaking up the van der Waals terms into repulsion and dispersion. Figure 1 indicates the solvation free energy that Gaussian03 attempts to estimate in a continuum calculation,

ΔG_{solv}^{G03}, represented by steps iii−v in the cycle. This infinite-dilution solvation free energy contains contributions from electrostatic (ES) interactions, cavitation, repulsion, and dispersion, as seen in eq 26.[8,9] The latter three are known as the nonelectrostatic (non-ES) terms and the associated solvation energy is denoted ΔG_{solv}^{non-ES}.

$$\Delta G_{f,Solution} = \Delta G_{f,gas} + \Delta G_{solv}^{true} \quad (24)$$

$$\Delta G_{solv}^{true} = \Delta G_{gas}^{NI} + \Delta G_{gas}^{comp} + \Delta G_{solv}^{cav} + \Delta G_{solv}^{ES} + \Delta G_{solv}^{rep} + \Delta G_{solv}^{disp} + ... + \Delta G_{relaxation}^{solvent} + \Delta G_{dilute}^{\Delta solvent} + \Delta G_{real}^{\Delta C_i} \quad (25)$$

$$\Delta G_{solv}^{G03} = \Delta G_{solv}^{ES} + \Delta G_{solv}^{cav} + \Delta G_{solv}^{rep} + \Delta G_{solv}^{disp} \quad (26)$$

For most molecules under most conditions, the $\Delta G_{f,gas}$ term can be evaluated using conventional gas-phase quantum chemistry and needs not be discussed further. The ΔG_{gas}^{NI} term can usually be estimated from an equation of state; often the ideal gas law is sufficient. The ΔG_{gas}^{comp} term for an isothermal ideal gas is trivial to calculate: $G_{gas}^{comp} \approx -RT \ln(C_{i,gas}/C_i^{\infty})$. The ΔG_{solv}^{cav}, ΔG_{solv}^{rep}, and ΔG_{solv}^{disp} terms are estimated within Gaussian. The ΔG_{solv}^{ES} term is also estimated within Gaussian and its magnitude is dependent on the electronic structure and the dielectric constant of the solvent. The nonideal solution corrections, $\Delta G_{dilute}^{\Delta Solvent}$ and $\Delta G_{real}^{\Delta C_i}$, can take the form of an activity coefficient and are typically dependent on temperature, pressure, composition, and an appropriately chosen reference state. These quantities are not computed by Gaussian but can be estimated using COSMO-RS theory. If the species of interest is always in the dilute limit, the $\Delta G_{real}^{\Delta C_i}$ term may be neglected. The solvent relaxation term, $\Delta G_{relaxation}^{solvent}$, is not computed by Gaussian either but can be estimated empirically or computed using Monte Carlo methods.

Each of the free energy terms shown above is composed of enthalpic and entropic parts, which must be isolated to accurately estimate the enthalpy and entropy in solution and the free energy at different temperatures. The terms estimated within the PCM methodology will be addressed here. Although the precise methodology for parsing the free energy is unclear, there are a number of logical assumptions that can be made with respect to the above terms. The first is that ΔG_{solv}^{ES} can be assumed to be mostly enthalpic because it is based on the electrostatic interaction between charges. However, there will be a portion attributable to entropy due to the experimental nature of the dielectric constant (ϵ). Because an experimental ϵ is used, an entropic penalty associated with aligning solvent molecules is built into the value. If no penalty were present, the measured dielectric constant would be larger and stronger ES interactions could be obtained. This entropy effect contributes to the decrease in ϵ as the temperature is increased, increasing the entropy cost associated with orienting the solvent molecules in a specific direction. In this work, we neglected higher-order effects such as this when trying to extract enthalpic and entropic parts of a given term, making the assumption that each term is solely entropic or enthalpic, with the exception of the analytical cavitation free energy term as discussed later. The ΔG_{solv}^{non-ES} term involves major enthalpic and entropic contributions; however, a logical partitioning can still be made. The $\Delta G_{solv}^{repulsion}$ and $\Delta G_{solv}^{dispersion}$ terms are considered van der Waals interactions[9] and are assumed to contribute exclusively to the enthalpy of solvation. The free energy associated with moving the solute from the gas phase to a cavity in the solvent, ΔG_{solv}^{cav}, is primarily due to the loss of entropy because the free volume of the solvent is much less than in the gas phase. On the basis of $S = -\partial G/\partial T$ analysis using Pierotti's equations[25,26] and certain other assumptions shown later, approximately 90−93% of the cavitation free energy is entropy-based.

It is important to examine the output from PCM calculations to understand which numbers can be trusted. In Gaussian03, ΔG_{solv}^{ES} is added directly to the PCM-optimized electronic energy, along with the "gas-like" thermal correction, to obtain the energy, enthalpy, or free energy in solution at 298 K. This is significantly inaccurate because the enthalpy is being derived on the basis of a free energy of solvation, not an enthalpy of solvation. Although nonelectrostatic terms are estimated by the version of Gaussian03 used in the current work (revision B.05), the non-ES contributions are *not* included in the final energy reported at the end of an optimization calculation or in the thermochemistry results from a frequency calculation. If one wants to include the non-ES terms in the free energy, they must be included manually. Equations 27 and 28 show how the solution-phase enthalpy and free energy are calculated within a Gaussian03 frequency calculation, where E_{elec}^{0K} is the sum of the electronic and zero-point energies and $C_{p,gas-like}$ and $S_{gas-like}^{298K}$ are the heat capacity and entropy calculated using gas-phase statistical mechanics formulas with the PCM-estimated frequencies. The energies are referenced to the completely separated and ionized atoms as is typical for Gaussian.

$$H_{Gaussian03}^{298K,soln} = (E_{elec}^{0K} + \Delta G_{solv}^{ES,298K}) + \int_{0\,K}^{298\,K} C_{p,gas-like}\,dT \quad (27)$$

$$G_{Gaussian03}^{298K,soln} = (E_{elec}^{0K} + \Delta G_{solv}^{ES,298K}) + \int_{0\,K}^{298\,K} C_{p,gas-like}\,dT - T{\cdot}S_{gas-like}^{298K} \quad (28)$$

The enthalpy equation is incorrect because the enthalpic portions of the non-ES solvation free energy are excluded. The total non-ES free energies are typically small, 0−2.5 kcal/mol, but this is due to significant cancellation between enthalpic and entropic terms. Taking H_2O solvated in water using the PCM as an example, typical non-ES energy values are (in kcal/mol): $\Delta G_{solv}^{non-ES} = 0.72$, $\Delta G_{solv}^{repulsion} = 1.45$, $\Delta G_{solv}^{dispersion} = -5.18$, and $\Delta G_{solv}^{cavitation} = 4.45$. If dispersion and repulsion are assumed to be completely enthalpic, and cavitation is assumed to be completely entropic, then neglecting the non-ES terms as in eq 27 would introduce an absolute error of 3.73 kcal/mol, even though the total non-ES energy is only 0.72 kcal/mol. For a larger molecule such as N_2O_4, this error would be closer to 9 kcal/mol. The free energy equation is more reasonable, because one does not have to worry about parsing the free energy of solvation into enthalpic and entropic terms, but it still should include the total non-ES free energy. The entropy reported in the thermochemistry section of a Gaussian03 PCM frequency calculation is essentially a gas-phase entropy and should *never* be used as the true solution-phase entropy. The same can be said for the heat capacity, which also implies that the thermal correction factors are erroneous for solution-phase calculations. Discussion of a more appropriate way to estimate the entropy in solution will follow.

In the current work, the analogous equations used to estimate the thermochemistry in the dilute limit are shown in eqs 29 and 30. The values estimated by these equations are referenced to the elements in their standard states because $\Delta H_{f,CBSQB3}^{gas,0K}$ is

calculated using an atomization method that is based on experimental values for the heat of formation of isolated atoms at 0 K, which are referenced to the elements in their standard states. This allows direct comparison with experimental enthalpy of formation data for neutral species; ions require the solvated proton reference to be modified. Equation 29 attempts to include all of the enthalpic terms, including those from the non-ES free energy. Notice that the solvation enthalpy term is estimated by removing the entropic contribution ($T\Delta S_{solv}^{298K}$) from the free energy of solvation. The free energy equation shown in eq 30 includes the ES and non-ES free energy terms and an additional empirical solvent-ordering entropy correction. Because empirical terms are used to fit dilute experimental entropy data, the solvation free energy estimated in this work attempts to capture steps iii−vi in Figure 1, which is the solvation of an ideal gas in pure water at the dilute limit. The methodology for estimating the terms not explicitly calculated by Gaussian will be covered later.

$$\Delta H_{f,current}^{soln,298K} = \Delta H_{f,CBSQB3}^{gas,0K} + \Delta H_{solv}^{298K} + \int_{0\,K}^{298\,K} C_{p,gas-like}\,dT \quad (29)$$

$$\Delta G_{current}^{298K,soln} = \Delta H_{f,current}^{soln,298K} - T\cdot(\Delta S_{gas-like}^{298K} + \Delta S_{solv}^{298K}) \quad (30)$$

where

$$\Delta H_{solv}^{298K} \equiv \Delta G_{solv}^{ES} + \Delta G_{solv}^{dispersion} + \Delta G_{solv}^{repulsion} + \Delta G_{solv}^{cavitation} + T\Delta S_{solv}^{298K}$$

$$\Delta S_{solv}^{298K} \equiv \Delta S_{solv}^{cavitation} + \Delta S_{solv}^{solvent\ ordering}$$

In this work, the solvation free energy was estimated by taking the difference between the free energy of the optimized molecule in the continuum solvation model, including the non-ES contribution, and the free energy of the gas-phase-optimized molecule in a vacuum. In other words, an optimization and frequency calculation was performed in a vacuum to get the gas-phase free energy, and an optimization and frequency calculation was performed within the continuum solvation model approximation to get the solution-phase free energy. However, out of the solvation models used here, solution-phase frequency calculations may only be performed using the IEFPCM method, and are not allowed when the COSMO or COSMO-RS keywords are used. Therefore, the zero point energy and thermal corrections to enthalpy and free energy for COSMO and COSMO-RS calculations were taken to be identical to those found using the IEFPCM solvation model and the B3LYP/6-311G(2d,d,p) level of theory. The implicit assumption is that the geometries and frequencies will not change significantly with the PCM radii set. Gas-phase ZPE and thermal corrections were computed using the gas-phase vibrational−rotational frequencies in the conventional rigid rotor-harmonic oscillator methodology.

Solution-Phase and Solvation Entropy. The solvation entropy is an important term to consider when dealing with solution-phase systems. The entropy reported by Gaussian in the frequency calculation results using the IEFPCM keyword is not the total entropy in solution, but a gas-like entropy calculated using gas-phase statistical machinery with the PCM-predicted vibrational and rotational frequencies. However, the large change in translational entropy upon solvation is accounted for in the nonelectrostatic solvation free energy computed within Gaussian. The ΔG_{solv} reported by Gaussian is reasonably accurate, but the reported enthalpy and entropy in solution have significant systematic errors.

A simplistic way to think about the solvation entropy is to break it into two pieces, a cavitation entropy and a solvent-ordering entropy.[16] The cavitation entropy solely examines the effect of confining the solute in the accessible free volume of the solution. The solvent-ordering entropy only looks at the effect the solute has on increasing or decreasing the order within the local solvent network. In general, the cavitation entropy will always be negative because the movement will always be restricted as compared with the gas phase. The solvent-ordering entropy can either be positive or negative depending on the characteristics of the solute and solvent. A theoretical framework, known as scaled-particle theory, has been developed that allows one to estimate the excess chemical potential (solvation free energy) associated with creating an empty cavity in a solvent.[17,18,21] This detailed approach is theoretically sound and has been shown to work well for spherical noninteracting molecules, but it is rather cumbersome. The solvation free energy of a cavity can also be estimated with a methodology outlined by Pierotti, as shown in eq 31.[25,26] The parameters are only a function of the solvent properties and the radii of the solute and solvent, making this method useful for cases where fast estimates of solvation energy are needed with little data available.

$$\Delta G_{cav} = K_0 + K_1\cdot r_{cav} + K_2\cdot r_{cav}^2 + K_3\cdot r_{cav}^3 \quad (31)$$

$$K_0 = RT\left[-\ln(1-y) + \frac{9}{2}\left(\frac{y}{1-y}\right)^2\right] - \frac{4\pi P}{3} r_{solvent}^3$$

$$K_1 = -\frac{RT}{2r_{solvent}}\left[6\left(\frac{y}{1-y}\right) + 18\left(\frac{y}{1-y}\right)^2\right] + 4\pi P r_{solvent}^2$$

$$K_2 = \frac{RT}{4r_{solvent}^2}\left[12\left(\frac{y}{1-y}\right) + 18\left(\frac{y}{1-y}\right)^2\right] + 4\pi P r_{solvent}$$

$$K_3 = \frac{4}{3}\pi P \qquad y = \frac{4\pi\rho r_{solvent}^3}{3}$$

The K_i terms are only a function of the solvent properties, where $r_{solvent}$ is the hard-sphere radius of a solvent molecule, ρ is the number density of the solvent (molecules/Å^3), P is the pressure (1 atm = 0.01458 cal mol^{-1} Å^{-3}), R is the gas constant, and T is the temperature. The solute radius enters from the r_{cav} term, which is the hard-sphere radius of the cavity, $r_{cav} \equiv r_{solvent} + r_{solute}$. The size of the solvent and solute molecules is the main source of uncertainty in the model, apart from the approximations that went into the derivation. In this study, the radius of a water molecule was taken to be 1.35 Å, which is approximately half the distance to the first peak in the experimental O−O radial distribution function for water.[39] The volume of a solute using the UAHF radii set was estimated with the "volume" keyword in Gaussian03. This method determines the volume of the molecule by estimating the volume within an electron isodensity contour of 0.001 electrons/bohr^3. Because this makes use of a relatively inaccurate Monte Carlo integration technique, 10 such integrations were performed for each molecule, ensuring a 95% confidence interval in the mean of less than 6%. Because the isodensity contour cutoff is somewhat arbitrary, it is sensible to believe that some correction may be needed to achieve a more realistic solute radius. The solute volume based on the Klamt radii set was determined using the COSMO*therm*. For both radii sets, the radius of the solute was taken to be the radius of a sphere with equal volume to the

TABLE 1: Use of Henry's Law Constant to Estimate Solution-Phase ΔH_f[a]

molecule	Henry's law constant (**H**) expt[b]	ΔG_{solv} (kcal/mol) from **H**	$\Delta H_{f,gas}$ (kcal/mol) expt[c]	ΔS_{solv} cal/(mol K) estimate	$\Delta H_{f,soln}$ kcal/mol estimate	$\Delta H_{f,soln}$ (kcal/mol) expt	deviation (kcal/mol)
NO_2	1.2E−02	0.7	8.1	−17.2	**3.7**		
NO	1.9E−03	1.8	21.8	−12.3	**19.9**		
N_2O_4	1.6E+00	−2.2	2.2	−15.6	**−4.7**		
N_2O_3	6.0E−01	−1.6	19.8	−15.6	**13.6**		
$HONO_2$	2.1E+05	−9.2	−32.1	−16.2	−46.1	−49.6[d]	3.5
HONO	5.0E+01	−4.2	−18.3	−16.2	−27.4	−28.5[d,e]	1.1
N_2O	2.5E−02	0.3	19.6	−15.5	15.3	13.4[d]	1.9
O_2	1.3E−03	2.0	0.0	−14.0	−2.1	−2.8[e]	0.7
HOOH	8.2E+04	−8.6	−32.5	−16.3	−45.9	−45.7[d,e]	−0.2
NH_3	2.7E+01	−3.8	−11.0	−14.8	−19.2	−19.2[d]	0.0

[a] All values are at 298.15 K; standard state concentration of 1 M in the gas and solution phases. [b] Henry's Law values are in mol L^{-1} bar^{-1}, ref 45. [c] Reference 46. [d] Reference 43. [e] Reference 42.

cavity (r_{solute}^{G03}), plus an empirical parameter (α). Typical values of free energy of cavitation range from 4 to 10 kcal/mol. The cavitation entropy can be related to the free energy through a temperature derivative, as shown in eq 32. In this work, it was assumed that the radii and solvent density are temperature independent. Typical values of entropy of cavitation computed in this way range from −10 to −25 cal/(mol K).

$$\Delta S_{cav}(r_{solute}) = -\left(\frac{\partial \Delta G_{cav}}{\partial T}\right)_P \approx K_0' + K_1' \cdot r_{cav} + K_2' \cdot r_{cav}^2 \quad (32)$$

The solvent-ordering entropy contribution was accounted for empirically because a simple theoretical framework for estimating this term was not available. The principle behind the chosen empirical correction was that the amount of ordering/disordering within the solvent upon insertion of a solute molecule depends on the bonds within a solute. One would expect that inserting a solute with many polar bonds would help to retain the hydrogen-bonding network of the solvent as compared with an empty cavity, resulting in a negative value of the ordering entropy. A nonpolar solute may disrupt the network, resulting in increased entropy of the system. In the systems studied in this work, there were three major terminal bond types: N−H, O−H, and N═O. The total entropy of the molecule in solution was estimated using eq 33. Notice that the basis for this is the gas-phase entropy. The implied assumption is that the rotational−vibrational entropy of a molecule is identical in a vacuum and in solution; in our experience, the small shifts in geometry and vibrational frequencies upon solvation do not significantly change $S_{vib-rot}$.

$$S_{solution}^{298K} = S_{gas}^{298K} + \Delta S_{solv}^{298K} = S_{gas}^{298K} + \Delta S_{cav}(r_{solute}^{G03} + \alpha) + \beta \cdot N_{O-H} + \gamma \cdot N_{N-H} + \delta \cdot N_{N=0} \quad (33)$$

The Greek letters are empirically fitted parameters and N_x is the number of bonds of type x in the solute. Admittedly, this will not capture all of the interactions with the solvent, such as those by lone pairs, but the limited experimental data necessitated a small number of fitted parameters. This technique is obviously very specific to molecules in the H−N−O class in water, and the confidence in the fitted parameters is low due to the lack of experimental entropy data in solution, as only 9 of the nearly 50 species examined in this study had experimental entropy values available. Typical values of the solvent-ordering correction range from −2 to +5 cal/(mol K).

Solution-Phase Enthalpy of Formation. The estimate of the solvation entropy allows the enthalpy of formation in solution to be obtained. The data from 18 species was available for comparison with the computed values and came from several sources, with some enthalpy values estimated from Henry's law coefficients and the entropy of solvation estimated as described above.[40−44] The methodology for converting a Henry's law constant (**H**) into an enthalpy of solvation requires several assumptions, gas-phase enthalpy and entropy data, and solvation entropy data. The relationship between **H** and the free energy of solvation (ΔG_{solv}) is given in eq 18. Given the aforementioned data, the "quasi-experimental" enthalpy of formation in solution was estimated by the following equation.

$$\Delta H_{f,soln}^{298K} = \Delta H_{f,gas}^{298K} + (\Delta G_{solv}^{from\ \boldsymbol{H}} + T\Delta S_{solv}) \quad (34)$$

We say "quasi-experimental" because it is based on an experimental gas-phase heat of formation, an experimental, dilute-limit solvation free energy, and a quantum/empirical estimate of the solvation entropy. As defined here, the solvation energy is somewhat ambiguous and depends on the concentrations of the gas and solution phases and possibly on standard state properties. The errors in the estimated enthalpy in solution include uncertainties in the solution-phase entropy, gas-phase enthalpy and entropy, and any ideality assumptions. The experimental Henry's law constants[45] used in this study to derive solution-phase enthalpy values are given in Table 1.

The estimated heats of formation in solution are given for the four species of interest, and for a test set of molecules where experimental solution-phase values have been reported, yielding insight into the error of the method. The error in the method is likely within the experimental errors in determining the Henry's law constants and the heat of formation in solution. The significant error associated with $HONO_2$ is a recurring concern that will be seen in the ab initio predictions as well. One other check that is available for the four species is the equilibrium constant for N_2O_4 and N_2O_3 dissociation reactions in solution, which are believed to be known relatively accurately in dilute aqueous solution. The reported equilibrium constants for the dissociation of N_2O_4 and N_2O_3 are 1.4×10^{-5} and 7.3×10^{-5} M, respectively.[47] Again assuming an ideal behavior and using the 0.04087 M standard state (equivalent to 1 atm in the gas phase), the ΔG_{Rxn}° values based on these experimental K_C values are 4.7 and 3.8 kcal/mol, respectively. The estimated values for ΔG°_{rxn} based on Henry's law data were 5.2 and 4.1 kcal/mol, respectively, representing errors of 0.5 and 0.3 kcal/mol.

Nonideal Effects. Activity coefficients are difficult to estimate, and experimental activity coefficient measurements are only available for a few species. The COSMO-RS theory proposed by Klamt and implemented in COSMO*therm* potentially allows one to estimate the activity coefficient of any species or transition state for which an electronic structure calculation can

TABLE 2: Gas-Phase BAC Values[a]

O−H	0.02	N−N	−1.87
N−H	−0.42	N═N	−1.58
N═O	1.11	N−O	0.35

[a] Units: kcal/mol.

be performed.[10,11,33−35] The theory utilizes the statistical thermodynamics of interacting surfaces to calculate the chemical potential of interaction of the solute and solvent mixture, and the activity coefficients are derived from the chemical potential difference. Even though most species in our system of interest are in the dilute limit, their activity coefficients relative to the pure-water solvation energies estimated in Gaussian will not necessarily be close to one due to the significant concentration of nitric acid. The activity coefficient corrections for $HONO_2$, NO_3^-, H_3O^+, and H_2O, and possibly NH_3OH^+ and HONO, will likely be the major contributors to nonideal behavior and will undoubtedly be important when attempting to model the system due to their effect on species equilibria and concentrations.

Results and Discussion

Gas-Phase Estimates. The CBS-QB3 procedure for estimating thermochemistry in the gas phase has already been established and proven to be relatively accurate for gas-phase hydrocarbons. The reason for dwelling on the gas-phase estimates in this work is because the solution-phase results are highly dependent on the gas-phase heats of formation. The solution-phase estimates will be no more accurate (and likely much less accurate) than their gas-phase counterparts, making it imperative to ensure good agreement in the gas-phase heat of formation estimates prior to embarking into the solution-phase methodology. The only new aspect that was required in this work was the fitting of new parameters to deal with the nitrogen-containing species that were at the heart of this study. The gas-phase BAC values for O−H and N−H bond types were taken from previously published work using the CBS-Q method and the "extended G2 neutral test set" and were not modified here.[32] New BAC values for N/N and N/O bond types are reported in Table 2 and show larger corrections than the other derived values for C−H−O bond types.

Regardless, the values reported here have large confidence limits due to the limited amount of data used to fit them. The gas-phase BAC value for N═N bonds is based solely on experimental data for N_2H_2, because none of the species with experimental data in the current set contained an N═N bond. The quantum chemical estimates for the gas-phase heats of formation at 298.15 K are given in the Supporting Information for 47 species related to the hydroxylamine system, with and without BACs. Many of these species were reported in a previous work, but the estimates did not include BACs for N/O and N/N bonds.[1] Of the twenty species used in the fitting procedure, the largest errors without BACs were 3.5 kcal/mol for N_2H_4, −3.2 kcal/mol for NO_2^+, −2.8 kcal/mol for N_2O_4, and −2.4 kcal/mol for $HONO_2$. The largest errors when including BACs were 1.7 kcal/mol for HNO and −1.0 kcal/mol for NO_2^+, with all other species having absolute errors less than 1 kcal/mol. The NO_2^+ cation is being judged on the basis of uncertain experimental data, because there are few data points and no reported error bars. The uncertainty in NO_2^+ is likely about 2 kcal/mol because estimating its heat of formation from $\Delta H_f(NO_2)$[48] and the ionization energy[46] of NO_2 yields a $\Delta H_f(NO_2^+)$ of 229 kcal/mol, compared to the 231.1 kcal/mol[43] used in this work. Average signed errors were −0.62 and −0.02 kcal/mol, mean absolute deviations (MAD) were 1.24 and 0.43 kcal/mol, and standard deviations in the absolute errors were 1.02 and 0.41 kcal/mol, for estimates without and with BACs, respectively. The MAD values should provide a reasonable approximation of the average error involved in the estimates. It is possible, if not probable, that other ions present in this set will have errors comparable to, or larger than, NO_2^+, though it is impossible to know for certain without reliable experimental data to compare against. Because many of these species are quite uncommon, and generally only of importance in aqueous solution, gas-phase data are limited and the quantum chemistry calculations are the best tool to quickly estimate thermochemical parameters. Gas-phase entropy values are listed in the Supporting Information.

Solution-Phase Estimates. The solution-phase thermochemical parameters are broken into two parts, entropy and enthalpy. Entropy values will be reported first because they were used to derive the enthalpy of formation values. The thermochemical parameters will be reported at a state corresponding to a temperature of 298 K and a concentration of 1 M. All nonidealities are ignored in the numbers reported here, meaning that the dilute-limit solvation energy is assumed to be valid at the 1 M concentration. In the absence of strong nonidealities, one would expect that the enthalpy of formation would have a weak dependence on solute concentration. However, entropy and free energy of a system will be significantly affected by the concentration at which the data are reported due to the entropy change when altering the solute concentration. The 1 M concentration choice was required because the experimental entropy and enthalpy data were reported at that state. The solution-phase thermochemical parameters for many of these species were reported in an earlier work;[1] however, the values reported here should be considered revised and more rigorously correct estimates.

Entropy Results. The solution-phase entropy estimates at 298 K and 1 M for the species of interest are shown in Table 3, including the empirical fitting parameters discussed in the Methodology section. The entropy was estimated using Gaussian03 at four levels of theory employing the UAHF and the Klamt radii sets. The molecular volume of the solute is reported in the Supporting Information for both radii sets before modifying the solute radius by the empirical parameter α. The volume based on the Klamt radii set was computed in a deterministic manner by COSMO*therm*. In general, all of these methods yield approximately the same entropy values, with relative standard deviations of well within 5% between the methods for most molecules. It is also stressed that the solvation entropy given here is for a constant concentration solvation process (e.g., 1 M in the gas phase → 1 M in the solution phase). A full table of results, including the explicitly stated gas-phase entropy and the cavitation and empirical solvation entropy contributions are included in the Supporting Information.

The effect of the empirical parameters on the errors in entropy is detailed in Table 4, which shows statistics for four methods in three cases: without empirical corrections, with only the solute radius correction, and with the radius and bond-ordering corrections. The best-fit values of the parameters are shown for each case, which obey a similar trend and have similar values. The radius correction parameter is the most important empirical term needed to achieve reasonable entropy results because it reduces the MAD value by ∼2.25 cal/(mol K) and corrects a large systematic error by forcing the mean signed error to near-zero. The radius adjustment can be seen as a correction of an overestimate of the solute radius or the water solvent radius fixed at 1.35 Å for this analysis. The bond-ordering correction

TABLE 3: Solution-Phase and Solvation Entropy Estimates at 298 K and 1 M[a]

		IEFPCM w/ UAHF radii[b]		deviations		
species	experiment [cal/(mol K)]	ΔS(solv)	S(soln, 1M)	UAHF − Klamt[c]	UAHF − expt	Klamt − expt
H_2O	23.97[d]	−13.0	25.7	−0.2	1.7	2.0
$HONO_2$	42.20[e]	−16.2	41.1	0.1	−1.1	−1.2
HONO	34.45[e,f]	−16.2	36.8	0.2	2.3	2.1
NH_2OH		−17.8	31.9	−0.5		
N_2H_4	33.00[f]	−18.9	33.6	−0.8	0.6	1.4
NO_3	37.50[e,f]	−18.5	37.4	−0.3	−0.1	0.2
H_3O^+		−13.2	28.8	0.2		
HAN		−22.9	55.1	0.3		
$NH_3ONO_2^+$		−18.6	47.9	1.2		
$N_2H_5^+$	33.54[e,f]	−18.0	33.1	0.6	−0.5	−1.1
$ONONO_2$		−18.2	55.4	0.9		
N_2O_4		−15.6	52.8	0.6		
NO_2		−17.2	33.7	0.7		
NH_2O		−16.5	33.7	−0.1		
HNO		−13.4	33.0	0.1		
ONONO		−18.6	45.7	0.7		
NO_2^-	33.95[e,f]	−18.6	31.7	−0.3	−2.3	−1.9
NO^+		−10.3	30.7	2.4		
NO		−12.3	30.4	0.7		
NH_3OH^+		−16.8	33.7	0.8		
N_2O		−15.5	30.7	0.4		
t-NH_3ONO^+		−19.4	44.9	1.1		
c-NH_2ONO		−19.5	41.5	0.6		
c-$ONNH_2O$		−19.5	41.6	0.8		
t −ONNHOH		−20.0	39.9	−0.2		
$ONNH^+$		−15.7	34.8	0.4		
NO_2^+		−15.8	28.9	1.7		
$NH_3OH-H_2O^+$		−23.0	41.0	0.0		
NH_2ONO_2		−19.1	45.9	0.9		
NH_2O^+		−15.2	32.1	0.9		
c-HONNOH		−22.2	36.9	0.0		
$HONN^+$		−15.2	36.3	1.2		
N_2O_3		−15.6	49.6	1.3		
O_2	26.50[e]	−14.0	28.6	0.8	2.1	1.4
HO_2		−15.4	32.9	0.2		
HOOH	34.40[e,f]	−16.3	32.0	0.1	2.3	−2.4
HOONO		−19.0	41.7	0.5		
OH		−12.5	23.7	−0.3		
$HOONO_2$		−18.3	46.2	0.9		
NH_3O^+		−16.0	34.5	1.0		
c-$ONNH_2OH^+$		−19.4	45.3	1.6		
NH_3		−14.8	27.0	−0.8		
$ONNH_2ONO^+$		−22.0	57.5	1.4		
ONNHONO		−21.4	51.8	1.1		
N_2		−15.4	24.0	0.6		
$NH_2O(H)NO^+$		−19.8	44.5	1.4		

[a] Units: cal/(mol K); all estimates include empirical solvent-ordering correction. [b] IEFPCM/B3LYP/6-311G(2d,d,p) with UAHF radii set. [c] IEFPCM/UAHF and COSMO/Klamt radii set difference at B3LYP/6-311G(2d,d,p). [d] Reference 16. [e] Reference 42. [f] Reference 43.

TABLE 4: Simple Statistics for Entropy Estimates with and without Empirical Corrections[a]

	IEFPCM/B3LYP/ 6-311G(2d,d,p)			IEFPCM/B3LYP/ 6-311++G(3df,3pd)			COSMO/B3LYP/ 6-311G(2d,d,p)			COSMO/BP86/ TZVP/DGA1		
empirical correction	none[b]	radius[c]	all[d]	none	radius	all	none	radius	all	none	radius	all
mean error	4.41	−0.15	−0.05	4.28	−0.19	−0.05	4.11	−0.13	−0.04	3.75	−0.15	−0.04
mean |error|	4.41	1.92	1.46	4.41	1.81	1.40	4.11	1.83	1.52	3.94	1.64	1.12
minimum error	0.77	0.32	0.11	0.59	0.01	0.10	0.58	0.28	0.22	0.86	0.36	0.16
maximum error	9.64	4.66	2.35	9.71	4.84	2.54	9.34	4.71	2.45	8.87	4.61	2.57
standard deviation	2.77	1.41	0.89	2.76	1.83	0.76	2.65	1.43	0.67	2.35	1.68	0.89
radius correction, α (Å)		−0.15	−0.11		−0.15	−0.10		−0.14	−0.10		−0.13	−0.07
bond corrections												
β (O−H)			−0.20			−0.54			0.03			−0.33
γ (N−H)			0.07			0.41			0.11			0.45
δ (N═O)			2.50			3.05			2.48			2.85

[a] Units: cal/(mol K). [b] No solution-phase empirical corrections. [c] Fitted parameter for the solute radius correction (α). [d] Fitted parameters for the solute radius and solvent-ordering corrections (α, β, γ, δ).

to the solvation entropy also serves to reduce a slight systematic error left after the radius correction and reduces the MAD by another ∼0.5 cal/(mol K), when the radius correction is re-fit with the bond corrections. The uncertainty in the parameter

TABLE 5: Solution-Phase Enthalpy of Formation Data at 298 K[a]

species	experimental	calculated[b]	deviation	species	experimental	calculated[b]	deviation
H_2O	−68.3[c]	−68.5	−0.2	*c* −$ONNH_2O$		16.9	
$HONO_2$	−49.6[c]	−44.3	5.3	*t* -ONNHOH		−5.1	
HONO	−28.5[c,d]	−29.2	−0.7	$ONNH^+$		162.7	
NH_2OH	−23.5[c,e]	−25.5	−2.0	NO_2^+		149.9	
N_2H_4	8.2[c]	4.7	−3.5	$NH_3OH-H_2O^+$		−1.6	
NO_3	−49.0[c,d,f]	−143.2	−1.1	NH_2ONO_2		−6.1	
H_3O^+		26.7		NH_2O^+		127.9	
HAN		−76.7		*c* −HONNOH	−15.4 [c]	−17.6	−2.2
$NH_3ONO_2^+$		107.9		$HONN^+$		158.2	
$N_2H_5^+$	−1.8[c,d]	89.7	−1.5	N_2O_3	11.7 [g]	14.2	2.5
$ONONO_2$		2.7		O_2	−2.8 [d]	−3.3	−0.5
N_2O_4	−3.7[g]	−3.5	0.2	HO_2		−10.8	
NO_2	3.7[g]	2.7	−1.0	HOOH	−45.7 [c,d]	−50.3	−4.6
NH_2O		0.3		HOONO		−13.2	
HNO		16.8		OH		0.1	
ONONO		15.2		$HOONO_2$		−27.0	
NO_2^-	−25.0[c,d,f]	−118.0	0.1	NH_3O^+		131.9	
NO^+		153.8		*c* -$ONNH_2OH^+$		109.3	
NO	19.3[g]	19.8	0.5	NH_3	−19.2 [c]	−20.9	−1.7
NH_3OH^+	−32.8[c]	60.7	0.4	$ONNH_2ONO^+$		156.2	
N_2O	13.4[c]	14.4	1.0	ONNHONO		29.4	
t −NH_3ONO^+		114.8		N_2		−4.2	
c −NH_2ONO		7.2		$NH_2O(H)NO^+$		121.8	
H^+	(fitted)	93.0					

[a] Units: kcal/mol. Note: The experimental values for ions assume $\Delta H_f(H^+) = 0$. To compare with the calculated values, the estimated absolute value for $\Delta H_f(H^+)$ must be added to (cations) or subtracted from (anions) the given experimental heats of formation. This is required to properly calculate deviations for ions. [b] IEFPCM/B3LYP/6-311G(2d,d,p) with UAHF radii set. [c] Reference 43. [d] Reference 42. [e] Reference 41. [f] Reference 44. [g] From Henry's law constants in ref 45.

values when bond corrections are included is rather high, given that nine experimental data points are being fit using four adjustable parameters and that assumptions had to be made in assigning some bond types. For example, NO_3^- was assumed to have one N═O bond and two delocalized bonds, where in reality all bonds are delocalized to some degree; however, the bond assignments were done in a logical manner and not after-the-fact to reduce errors. To accurately estimate solution-phase entropy values in the manner proposed here, a radius correction is required to achieve reasonable results; however, the inclusion of higher order corrections will improve results but may not be statistically justifiable. Solution-phase entropy estimates for the case with only a radius correction are available in the supporting material.

Enthalpy of Formation Results. The solution-phase heat of formation estimates were derived from the entropy data given above and from the free energy of solvation as computed at seven levels of theory, as described in the Methodology section. However, the full results from only the most accurate method are presented here, with the entire set of data available in the Supporting Information. The method presented here is B3LYP/6-311G(2d,d,p) using IEFPCM with the UAHF radii set. The summary of errors for all methods is also present below. It is essential that the solvation free energy and entropy are both based upon the same process, taken here to be a constant concentration solvation process. The enthalpy was assumed to be independent of concentration (dilute limit), which may not be accurate at a concentration of 1 M. The enthalpy of formation at 298 K is reported for 46 species, with the enthalpy of formation of the proton in solution set to 93 kcal/mol, which is based on the fitting of experimental data and is in reasonable agreement with the value recommended by Tissandier et al.[49] This value is needed to correct the experimental ionic heats of formation, which are conventionally referenced to the proton heat of formation in solution.

The computed solution-phase thermochemical parameters are given in Table 5; experimental values are given in the table as well, with NO, NO_2, N_2O_3, and N_2O_4 derived from Henry's law data and the estimated solvation entropy. Because the solvation entropy varies slightly with each method, the "experimental" enthalpy of formation of these four species will also vary slightly. The values given are based upon the solvation entropies from IEFPCM calculations using the B3LYP/6-311G(2d,d,p) level of theory. The heat capacity data in solution is not given because the temperature dependence of the solvation effects is not well understood. In general, the heat capacities of most species in aqueous solution will not be an important factor in the overall heat capacity of a real solution, because water will typically dominate; however, they are important for nonisothermal reacting systems because the heat capacity is used to calculate temperature dependence of the equilibrium constants. All of the work here is confined to isothermal behavior, and it is likely that the solvation properties will be most accurate near 298 K.

A minimum in the potential was not found for $ONNH^+$ using the UAHF radii set, but an optimized geometry was found using the Klamt radii set and the COSMO keyword. The values shown for $ONNH^+$ with the UAHF radii set were estimated by making the assumption that the solution-phase energy difference between $ONNH^+$ and $HONN^+$ was the same for the IEFPCM and COSMO methods.

The data in Table 6 indicate the amount of error involved in the estimates, and ultimately in the solution-phase calculations because the gas-phase error is relatively small and well distributed. It also shows the best-fit proton heat of formation for each method. The experimental values to which the estimates are compared did not include uncertainties, and it is obvious that any error in those numbers may make the estimates appear better or worse than what they are in reality. The MAD values for the seven methods ranged from 1.5 to 3 kcal/mol. Solution-phase bond additivity corrections were attempted using 7 bond types; however, the statistical uncertainty in the fitted values warrants their exclusion here. The very brief summary of their effect is given here. Generally, the BAC values from estimates

TABLE 6: Simple Statistics for Solution-Phase Enthalpy Estimates[a]

	IEFPCM[b]			COSMO[c]			
statistic	B3LYP/ CBSB7	B3LYP/ 6-311++G(3df,3pd)	CBS-QB3	B3LYP/ CBSB7	BP86/ TZVP/DGA1	COSMO−RS[d] BP86/ TZVP/DGA1	COSMO*therm*[e] BP86/ TZVP/DGA1
$\Delta H_f(H^+)$	93.04	92.87	91.27	96.32	95.58	96.35	97.49
mean error	−0.59	0.05	0.82	1.83	2.09	1.59	−0.51
mean |error|	1.55	1.96	2.07	2.03	2.27	1.95	2.89
min. error	0.06	0.12	0.01	0.05	0.18	0.17	0.10
max. error	5.32	5.76	6.70	8.53	8.80	8.41	9.30
std dev	1.51	1.70	1.87	2.06	2.15	1.91	2.83

[a] Units: kcal/mol. [b] Based on Gaussian03 solvation free energy using IEFPCM keyword with UAHF radii set. [c] Based on Gaussian03 solvation free energy using COSMO keyword. [d] Based on Gaussian03 solvation free energy using COSMO-RS keyword. [e] Based on solvation free energy derived from COSMO*therm* partial pressure prediction.

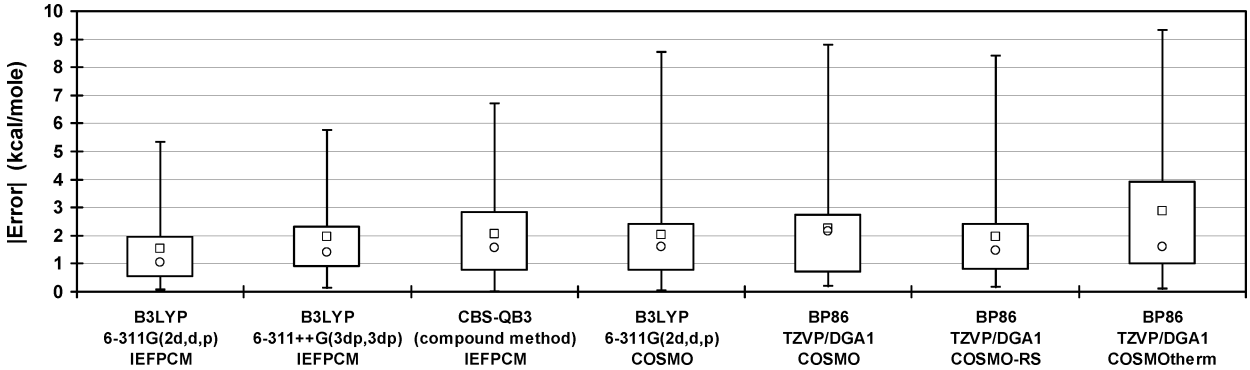

Figure 2. Absolute errors in $\Delta H_{f,soln}$ for the seven methods investigated on the basis of 18 experimental data points. Squares (□) represent the MAD, circles (○) represent the 50th percentile, the box spans from the 25th to 75th percentile, and the bars show the entire range of errors.

using the UAHF radii set were similar to each other, and the values from estimates using the Klamt radii set were similar. The inclusion of BACs could achieve two goals; it greatly reduces the systematic error in the estimates and it reduces the MAD by 25−50%. It may be possible to derive meaningful BAC values for a given method, but this would require a much larger test set and was not the goal of this work.

One significant observation is the major discrepancy in the heat of formation for H_3O^+, which is ∼27 kcal/mol when the IEFPCM/UAHF method is used and ∼45 kcal/mol when the COSMO/Klamt method is used, both at the B3LYP/CBSB7 level. The heat of formation of hydronium ion in water has been reported to be −68.3 kcal/mol based on $\Delta H_f^{298K}(H^+) = 0$.[41] Based on the fitted value for H^+ from IEFPCM/B3LYP/CBSB7, $\Delta H_f^{298K}(H^+) = 93$ kcal/mol, the experimental absolute heat of formation of H_3O^+ appears to be ∼29 kcal/mol. The experimental value was deemed to be significantly uncertain, thus not included in the fitting procedure, but nonetheless agrees well with the heat of formation predicted by the IEFPCM/UAHF method. There appears to be a large error in the solution-phase energy of H_3O^+ that is introduced if the Klamt radii set is used as implemented in this version of Gaussian03. This "error" seems to be significantly corrected once the data are processed by COSMO*therm*, which yields $\Delta H_f^{298K}(H_3O^+) = 33.5$ kcal/mol. This is in reasonable agreement with the experimental data point, given the COSMO*therm*-preferred $\Delta H_f^{298K}(H^+)$ is ∼98 kcal/mol, yielding an "experimental" $\Delta H_f^{298K}(H_3O^+)$ of ∼34 kcal/mol. It will become obvious when examining pK_A predictions that using unprocessed data based on the Klamt radii set directly from Gaussian03 will yield highly erroneous results whenever the hydronium ion is involved. This is in contrast to the reasonable results obtained for most molecules examined in this study but emphasizes that care must be taken when using numbers derived from the Klamt radii set without processing them in COSMO*therm*. A summary of the unsigned error statistics are provided in Figure 2 for the seven methods. It shows the MAD, the range or errors, and where several error percentiles lie. The MAD value is typically larger that the 50th percentile, indicating that there is a higher frequency of smaller errors.

Comparison with Experimental Aqueous Equilibrium Data. In addition to the experimental heat of formation data, there are several experimental reaction thermochemistry data available that can be used to test the quantum chemical estimates. Acid dissociation data are well-known for several of the species of interest in this study, and the relative pK_A values for these equilibria ($HA + H_2O \rightleftharpoons A^- + H_3O^+$) can be estimated with the computed thermochemical values. The experimental pK_A values were converted to their relative counterpart by taking $pK_A(H_3O^+) = -1.74$. The relative pK_A values in dilute aqueous solution at 298 K for HONO, $HONO_2$, $N_2H_5^+$, and NH_3OH^+ are 4.99, 0.24, 9.84, and 7.68 respectively,[50,51] which translate into ΔG^o_{Rxn} values of 6.80, 0.33, 13.42, and 10.47 kcal/mol for a standard state concentration of 1 M, ignoring nonidealities. As long as the standard state is identical for all reactants and products, the idealized ΔG^o_{Rxn} will be the same at all concentrations because this reaction is equimolar. The errors (estimate minus experimental) in ΔG^o_{Rxn} estimates for these acid dissociation reactions are presented in Table 7 for four of the seven methods, with surprisingly large errors for the COSMO methods in Gaussian. This highlights the point made earlier that the unprocessed COSMO results involving hydronium are highly erroneous. The results using the UAHF radii set are better, but the errors are significant and highlight why it is very difficult to model complex processes in an aqueous system, where at room temperature an error of 1.4 kcal/mol in ΔG^o_{Rxn} translates to an order of magnitude change in the equilibrium constant. The COSMO*therm* results

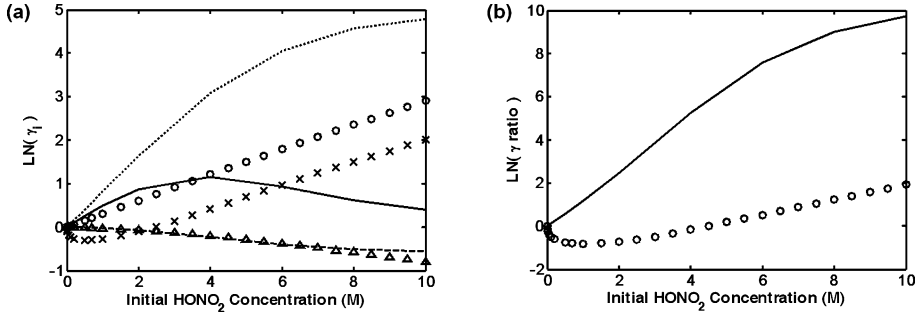

Figure 3. Activity coefficients for nitric acid equilibrium. Experimental data:[52] markers. Predictions: lines. (a) Activity coefficients for undissociated nitric acid (solid line and ○) and water (long dashes and △), as well as the mean ionic activity coefficient (short dashes and ×). (b) Log of the activity coefficient ratio for the reaction $HONO_2 + H_2O \rightleftharpoons H_3O^+ + NO_3^-$.

TABLE 7: Errors in ΔG_{Rxn} for Acid−base Equilibria for Select Methods[a]

	experimental		COSMO		
reaction	pK_A	ΔG_{Rxn}	IEFPCM B3LYP CBSB7	B3LYP CBSB7	BP86 TZVP/DGA1	COSMOtherm BP86 TZVP/DGA1
$HONO + H_2O \rightleftharpoons NO_2^- + H_3O^+$	4.99	6.80	0.23	14.96	16.81	−6.29
$HONO_2 + H_2O \rightleftharpoons NO_3^- + H_3O^+$	0.24	0.33	−3.83	11.73	13.55	−7.89
$N_2H_5^+ + H_2O \rightleftharpoons N_2H_4 + H_3O^+$	9.84	13.42	−4.79	16.64	19.37	11.65
$NH_3OH^+ + H_2O \rightleftharpoons NH_2OH + H_3O^+$	7.68	10.47	−2.30	12.74	13.37	7.72
Using IEFPCM Solvation Energy for COSMO H_3O^+ Calculations[b]						
$HONO + H_2O \rightleftharpoons NO_2^- + H_3O^+$	4.99	6.80		−3.38	−2.61	
$HONO_2 + H_2O \rightleftharpoons NO_3^- + H_3O^+$	0.24	0.33		−6.60	−5.87	
$N_2H_5^+ + H_2O \rightleftharpoons N_2H_4 + H_3O^+$	9.84	13.42		−1.70	−0.05	
$NH_3OH^+ + H_2O \rightleftharpoons NH_2OH + H_3O^+$	7.68	10.47		−5.59	−6.06	

[a] Units: kcal/mol. [b] Shows the COSMO-based estimates if the IEFPCM/B3LYP/CBSB7 ΔG_{solv} is used for H_3O^+.

are not particularly impressive, but it should be kept in mind that they used the estimated gas-phase ΔH_f. The COSMO*therm* results are improved if the dissociation free energy values are calculated without modification and are used in the empirical pK_A relationship derived by Klamt et al.[15] If the free energy in solution is calculated as the gas-phase free energy directly from Gaussian plus the solvation free energy derived from COSMO*therm*'s "partial pressure" estimate, and the ΔG°_{Rxn} value is used in the empirical relationship, the predicted pK_A value for the reactions are, as ordered in the table 5.35, −0.15, 10.29, and 7.42. These estimates are very good, with errors of less than 0.5 pK_A units. COSMO*therm* is capable of accurately estimating pK_A values semiempirically.

Activity Coefficient Estimates. Accurate estimates of system nonidealities are necessary to effectively predict the behavior of a concentrated, complex system. These effects can be important because they can easily amount to several *RT* in the effective Gibbs free energy of a molecule in a concentrated solution; however, it is somewhat less important than the errors in the solvation energies that have magnitudes of approaching 5 kcal/mol for some ionic species. For an isothermal, isobaric system, a complex dependence of the activity coefficient of each species exists and is based upon the concentration of that species, as well as the concentrations of all other species in the mixture. It is a demanding task to examine all possible dependences to get a true picture of the nonideal surface of a system. However, in the system of interest in this work, the majority of the species will be present only in very small concentrations because the main reactant, hydroxylamine, is usually only present in concentrations up to ∼0.25 M. Therefore, it is a reasonable first approximation that all of the species in the system that are not involved in the nitric acid acid/base equilibrium are present in concentrations that can be taken at the dilute limit. This means that the concentration of any compound in the infinitely dilute limit does not affect the activity coefficient of any other molecule in the system. This assumption simplifies the activity coefficients tremendously, making them only sensitive to the concentrations of H_2O, $HONO_2$, H_3O^+, and NO_3^-. This is also simplified further because these species exist in a fixed ratio if the acid−base reaction is assumed to be in rapid equilibrium. The concentrations, mole fractions, and fraction of nitric acid dissociated as a function of the initial concentration of nitric acid have been reported, taking into account the change in total solution concentration as the nitric acid mole fraction is varied.[52] Using the known concentrations of the major species, for a range of compositions, one can estimate the activity coefficient of these species, as well as those of other species at infinite dilution over a range of nitric acid concentrations.

COSMO*therm* provides a straightforward way to estimate the activity coefficient of any species in any mixture, given the appropriate quantum chemistry calculation results. This software was used to estimate the infinite dilution activity coefficient of all species in aqueous nitric acid up to a concentration of 10 mol/L. This activity coefficient corrects for the change in behavior from infinite dilution in pure water to infinite dilution in a nitric acid solution, denoted earlier by $\ln(\gamma_i^{\infty HONO_2,\infty H_2O})$. COSMO*therm* calculations were performed at ten concentrations of nitric acid solution, from initial (pre-dissociation) $HONO_2$ concentrations of 0−10 M. The nitric acid mole fraction and the fractional dissociation as a function of concentration were derived from the data given by Davis and de Bruin.[52] The activity coefficient was calculated by taking the PCP difference between the state of interest and the infinite dilution in pure water state, as in eq 23. The values given for H_2O are the deviations from the pure water condition, and not the infinite dilution condition. The table of COSMO*therm* estimates for the $\ln(\gamma_i^{\infty HONO_2,\infty H_2O})$ is given in the Supporting Information for a range of nitric acid concentrations, based on the BP86/TZVP/DGA1 level of theory.

Unfortunately, there is not a substantial amount of experimental data available for the activity coefficients of many of the species examined here. This makes it difficult to comment on the accuracy of the activity coefficient estimates, especially for ionic species that may have large chemical potential differences. However, the nitric acid equilibrium has been studied in some detail. As a result, activity coefficient data are available for undissociated nitric acid, water, and the "mean ionic" species.[52,53] The mean ionic activity coefficient is the geometric mean of the activity coefficients of the H_3O^+ and NO_3^-, which is reported in place of individual ionic activity coefficients because they are difficult to determine independently. A comparison of the experimental data and COSMO*therm* estimates are shown in Figure 3. The symbols represent the experimental data and the lines the estimates. The estimated activity coefficient of water shows quite good agreement over the entire range, and the undissociated nitric acid estimates agree reasonably well up to concentrations of 4 M. However, there is very poor agreement between the data for undissociated nitric acid above 5 M and for the mean ionic activity coefficient over the entire range. The natural log of the activity coefficient ratio for the reaction is also shown in the figure, indicating an inaccurate prediction of the nonideality of the acid/base equilibrium. Nitric acid activity coefficient calculations were also performed with H_3O^+ and NO_3^- solvated with one, two, or three explicit water molecules, but the overall accuracy of the estimates was not improved significantly. Although the amount of experimental data for inorganic activity coefficients in nitric acid solutions is limited, this result implies that estimates for neutral species at low acidities may be relatively accurate, whereas ionic species or predictions at high ionic strengths may have large errors. In the PUREX process system, the nitric acid concentration is maintained below 1.5 M in many tanks, making this methodology potentially viable for neutral species in the system.

Conclusions

In this work, a method for estimating the enthalpy of formation and entropy in aqueous solution has been proposed. It requires the decoupling of the entropic and enthalpic terms in the free energy of solvation that is returned by computational chemistry codes. The accuracy of the results with and without a variety of empirical corrections was shown. The MADs for solution-phase enthalpy of formation estimates were 1.5−3 kcal/mol. Reasonable estimates of the entropy in solution could be obtained with the use of a single empirical parameter to adjust the radius of the solute, implying that the sum of the solute and solvent radii needed to be reduced by ∼0.14 Å. Employing several additional bond correction terms could improve the estimates further, though the statistical significance of these additional terms is low. Infinite dilution activity coefficients were estimated over a range of nitric acid concentrations using COSMO-RS theory, as implemented in the COSMO*therm* software package. Although little data are available, the estimates for H_2O and $HONO_2$ showed reasonable agreement for $[HONO_2]_0 < 5$ M, whereas the estimates for the ionic species were poor over the entire range. This questions the applicability of the estimates for ionic species or for any molecule in a solution with a high ionic strength. The thermochemical estimates reported here should provide a basis for constructing a detailed kinetic model of the oxidation of hydroxylamine in aqueous nitric acid solution, which is an important but undesirable side reaction in the PUREX process.

The proper use of solution-phase computational chemistry estimates in predicting the entropy, enthalpy, and free energy has been outlined. The relationship between solvation properties and certain experimentally accessible data have been shown in a way that hopefully exhibits more clarity and detail than in the past. The implicit assumptions about standard states and ideality were kept to a minimum to allow the reader to explicitly see the assumptions that must be made to achieve certain well-known relationships that are often evoked. A clear distinction was made between the standard state chemical potential and the pseudochemical potential and how each relates to experimentally accessible properties. This approach made it easier to see how computational results can be used to estimate measurable quantities, such as Henry's law constants, vapor pressures, and reaction equilibrium constants, as well as the assumptions that are implicitly made if one directly uses the output of a code such as Gaussian03. It is sincerely hoped that the explanations provided in this work will reduce the confusion and ambiguity often present in work dealing with solvation properties and solution-phase thermodynamics in general.

Acknowledgment. R.W.A. gratefully acknowledges support from an NSF graduate research fellowship. Financial support from Duke, Cogema, Stone, and Webster is gratefully acknowledged.

Supporting Information Available: Tables of enthalpy and entropy estimates for all levels of theory discussed in the text, bond type assignments for empirical corrections, a table of COSMO*therm*-estimated activity coefficients in nitric acid solution, optimized geometries for all molecules at the IEFPCM/B3LYP/CBSB7 level of theory with the UAHF radii set, and an example derivation of a reaction equilibrium constant. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Raman, S.; Ashcraft, R. W.; Vial, M. A.; Klasky, M. L. *J. Phys. Chem. A* **2005**, *109*, 8526.
(2) Pembridge, J. R.; Stedman, G. *J. Chem. Soc., Dalton Trans.* **1979**, 1657.
(3) Bennett, M. R.; Brown, G. M.; Maya, L.; Posey, F. A. *Inorg. Chem.* **1982**, *21*, 2461.
(4) Morgan, T. D. B.; Stedman, G.; Hughes, M. N. *J. Chem. Soc. B* **1968**, 344.
(5) Bourke, G. C. M.; Stedman, G. *J. Chem. Soc., Perkin Trans. 2* **1992**, 161.
(6) Hughes, M. N.; Stedman, G. *J. Chem. Soc.* **1963**, 2824.
(7) Harlow, D. G.; Felt, R. E.; Agnew, S.; Barney, G. S.; McKibben, J. M.; Garber, R.; Lewis, M. *DOE/ EH-0555* **1998**.
(8) Tomasi, J.; Mennucci, B.; Cances, E. *J. Mol. Struct. (THEOCHEM)* **1999**, *464*, 211.
(9) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.
(10) Klamt, A.; Jonas, V.; Burger, T.; Lehrenz, J. C. W. *J. Chem. Phys. A* **1998**, *102*, 5074.
(11) Klamt, A.; Schuurmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799.
(12) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133.
(13) Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2006**, *110*, 5665.
(14) Chipman, D. M. *J. Phys. Chem. A* **2002**, *206*, 7413.
(15) Klamt, A.; Eckert, F.; Diedenhofen, M.; Beck, M. E. *J. Phys. Chem. A* **2003**, *107*, 9380.
(16) Leung, B. O.; Reid, D. L.; Armstrong, D. A.; Rauk, A. *J. Phys. Chem. A* **2004**, *108*, 2720.
(17) Chandler, D. *Nature* **2005**, *437*, 640.
(18) Garde, S.; Hummer, G.; Garcia, A. E.; Paulaitis, M. E.; Pratt, L. R. *Phys. Rev. Lett.* **1996**, *77*, 4966.
(19) Huang, D. H.; Chandler, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8324.
(20) Lum, K.; Chandler, D.; Weeks, J. D. *J. Phys. Chem. B* **1999**, *103*.
(21) Hummer, G.; Garde, S.; Garcia, A. E.; Pohorille, A.; Pratt, L. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8951.
(22) Pratt, L. R.; Chandler, D. *J. Chem. Phys.* **1977**, *67*, 3683.

(23) Pierotti, R. A. *Chem. Rev.* **1975**, *76*, 717.
(24) Pierotti, R. A. *J. Phys. Chem.* **1965**, *69*, 281.
(25) Pierotti, R. A. *J. Phys. Chem.* **1963**, *67*, 1840.
(26) Hofinger, S.; Zerbetto, F. *Eur. J. Chem.* **2003**, *9*, 566.
(27) Langlet, J.; Claverie, P.; Caillet, J.; Pullman, A. *J. Phys. Chem.* **1988**, *92*, 1617.
(28) Ben-Naim, A. *Solvation Thermodynamics*; Plenum Press: New York, 1987.
(29) Ben-Naim, A. *Molecular Theory of Solutions*; Oxford University Press Inc.: New York, 2006.
(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; 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.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(31) Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **2000**, *112*, 6532.
(32) Petersson, G. A.; Malick, D.; Wilson, W.; Ochterski, J. W.; Montgomery, J. A., Jr.; Frisch, M. J. *J. Chem. Phys.* **1998**, *109*, 10570.
(33) Eckert, F.; Klamt, A. *COSMOtherm*, Version C2.1, Release 01.05; COSMOlogic GmbH & Co. KG: Leverkusen, Germany, 2005.
(34) Eckert, F.; Klamt, A. *AIChE J.* **2002**, *48*, 369.
(35) Klamt, A. *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*; Elsevier: Amsterdam, 2005.
(36) Abraham, M. H.; Le, J. *J. Pharm. Sci.* **1999**, *88*, 868.
(37) Sese, G.; Pardro, J. A. *J. Chem. Phys.* **1998**, *108*, 6347.
(38) Gu, C.; Lustig, S.; Trout, B. *J. Phys. Chem. B* **2006**, *110*, 1476.
(39) Narten, A. H.; Levy, H. A. *J. Chem. Phys.* **1971**, *55*, 2263.
(40) *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 2005.
(41) *Thermal Constants of Substances*; Yungman, V. S., Ed.; Wiley: New York, 1999.
(42) Barner, H. E.; Scheuerman, R. V. *Handbook of Thermochemical Data for Compounds and Aqueous Species*; Wiley: New York, 1978.
(43) Wagner; Evans; Parker; Schumm; Halow; Bailey; Churney; Nuttall *J. Phys. Chem. Ref. Data* **1982**, *11*.
(44) Thermodynamic Properties of Elements and Compounds. In *Macmillan's Chemical and Physical Data*; James, A. M., Lord, M. P., Eds.; The Macmillan Press LTD: London, 1992; pp 449.
(45) Sander, R. Henry's Law Constants. In *NIST Chemistry Webbook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 2005.
(46) Afeefy, H. Y.; Liebman, J. F.; Stein, S. E. Neutral Thermochemical Data. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 2005.
(47) Schwartz, S. E.; White, W. H. Kinetics of Reactive Dissolution of Nitrogen Oxides into Aqueous Solution. In *Trace Atmospheric Constituents*; Schwartz, S. E., Ed.; John Wiley & Sons: New York, 1983; p 1.
(48) Ruscic, B.; Pinzon, R.; Morton, M.; Srinivasan, N.; Su, M.-C.; Sutherland, J.; Michael, J. *J. Phys. Chem. A* **2006**, *110*, 6592.
(49) Tissandier, M. D.; Coe, J. V.; Tuttle, Thomas R. J. *J. Phys. Chem. A* **1998**, *102*, 7787.
(50) Analytical Chemistry. In *CRC Handbook of Chemistry and Physics*, Internet Version 2007 (87th ed.); Lide, D. R., Ed.; Taylor and Francis: Boca Raton, FL, 2007.
(51) Speight, J. G. *Perry's Standard Tables and Formulas for Chemical Engineers*; McGraw-Hill Companies, Inc.: New York, 2003.
(52) Davis, W., Jr.; de Bruin, H. J. *J. Inorg. Nucl. Chem.* **1964**, *26*, 1069.
(53) Hamer, W. J.; Wu, Y.-C. *J. Phys. Chem. Ref. Data* **1972**, *1*, 1047.