THE JOURNAL OF
PHYSICAL CHEMISTRY B

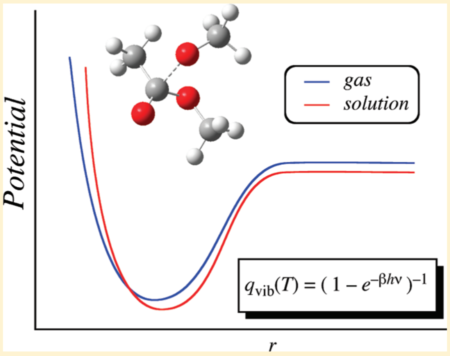
ARTICLE

pubs.acs.org/JPCB

Use of Solution-Phase Vibrational Frequencies in Continuum Models for the Free Energy of Solvation

Raphael F. Ribeiro, Aleksandr V. Marenich, Christopher J. Cramer,* and Donald G. Truhlar*

Department of Chemistry and Supercomputing Institute, University of Minnesota, 207 Pleasant Street S.E., Minneapolis, Minnesota 55455-0431, United States



ABSTRACT: We find that vibrational contributions to a solute's free energy are in general insensitive to whether the solute vibrational frequencies are computed in the gas phase or in solution. In most cases, the difference is smaller than the intrinsic error in solvation free energies associated with the continuum approximation to solvation modeling, although care must be taken to avoid spurious results associated with limitations in the quantum-mechanical harmonic-oscillator approximation for very low-frequency molecular vibrations. We compute solute vibrational partition functions in aqueous and carbon tetrachloride solution and compare them to gas-phase molecular partition functions computed with the same level of theory and the same quasiharmonic approximation for the diverse and extensive set of molecules and ions included in the training set of the SMD continuum solvation model, and we find mean unsigned differences in vibrational contributions to the solute free energy of only about 0.2 kcal/mol. On the basis of these results and a review of the theory, we conclude, in contrast to previous work (Ho, J.; Klamt, H.; Coote, M. L. J. Phys. Chem. A 2010, 114, 13442), that using partition functions computed for molecules optimized in solution is a correct and useful approach for averaging over solute degrees of freedom when computing free energies of solutes in solution, and it is moreover recommended for cases where liquid and gas-phase solute structures differ appreciably or when stationary points present in liquid solution do not exist in the gas phase, for which we provide some examples. When gas-phase and solution-phase geometries and frequencies are similar, the use of gas-phase geometries and frequencies is a useful approximation.

1. INTRODUCTION

A recent Comment[1] in *The Journal of Physical Chemistry A* about the use of implicit solvation models discussed what were perceived as "shortcomings", "confusion", and "incorrect" labeling of output in three widely available electronic structure programs,[2−4] as well as perceived misuse of that labeled output in publications. We here assert that there are no errors in the output and that what were perceived as shortcomings and misuse are instead well-defined and useful approximations; the Comment[1] needlessly narrowed the definition of free energy in solution to a single meaning that is inconsistent with the broader range of correct meanings in common use. In light of the conflicting statements in the literature, we present the theory here in a fashion that we hope will be useful for those engaged in practical work, and we follow this with computations to confirm our main theoretical points.

In order to make the situation clear, we must begin with the concept of a potential of mean force (PMF), also called a free energy surface (FES). This is the solution-phase analogue of a gas-phase potential energy surface.[5−13] Although this concept is not mentioned in the Comment, its appreciation is essential to understanding what we regard as a fundamental confusion in the Comment. A textbook[14] on molecular simulation correctly states that, "Implicit solvent models based on continuum electrostatic treatments...essentially approximate the potential of mean force.... The potential of mean force is an effective free energy potential...." The concept of a PMF or FES was introduced by Kirkwood[15] in 1935, it is used in the Debye−Hückel theory,[16a] and it has been widely used in the chemical physics literature, especially for the case of a single coordinate.[17−20] The PMF as a function of a single atomic coordinate is also called a free energy profile,[21] and it gets its name in that context because the mean force acting on the selected coordinate is given by minus the derivative of the free energy function with respect to that coordinate. In such functions, the privileged coordinate is a special case of an order parameter.[21] More generally, this interpretation applies to a PMF or FES that is a function of any number of coordinates.

2. THEORY

A general scheme for working with the FES in liquid-phase solution has been presented in a recent article,[13] and the reader may consult that for further discussion. Here we outline the major issues using the notation of that article. Let **R** denote the $3N - 6$ internal coordinates of the solute, and let **P** denote their conjugate momenta. Let **r** and **p** denote the remaining coordinates and conjugate momenta of the entire system (solute plus solvent). For simplicity in presenting our main points, we do not here distinguish between Gibbs free energy G and Helmholtz

Received: June 12, 2011
Revised: August 28, 2011
Published: August 29, 2011

free energy A. We assume the Born−Oppenheimer approximation with only the ground electronic state occupied. The thermodynamic free energy $G(S)$ of species S is given by[16b]

$$e^{-\beta G(S)} = C\int_S d\mathbf{R}\, d\mathbf{P}\, d\mathbf{r}\, d\mathbf{p}\; e^{-\beta H(\mathbf{R},\mathbf{P},\mathbf{r},\mathbf{p})} \tag{1}$$

where C is a geometry-independent constant determined by the choices for zero of energy and standard state, H is the Hamiltonian, and β is $1/\tilde{k}T$ where $\tilde{k}$ is Boltzmann's constant and T is temperature. Notice that the integral is taken only over the range of $\mathbf{R}$ that corresponds to species S. For example, for *s-cis*-4-methyl-1,3-pentadiene one integrates over the portion of the $\mathbf{R}$ space corresponding to the *s-cis* structure but not the *s-trans* and not the portion of the space corresponding to isomers of 2-methyl-1,3-pentadiene or other isomers. One could, if desired, carry out the integration of eq 1 in two steps:

$$e^{-\beta W(\mathbf{R})} = C'\int d\mathbf{R}'\, d\mathbf{P}\, d\mathbf{r}\, d\mathbf{p}\; \delta(\mathbf{R}-\mathbf{R}')e^{-\beta H(\mathbf{R}',\mathbf{P},\mathbf{r},\mathbf{p})} \tag{2}$$

$$e^{-\beta G(S)} = C''\int_S d\mathbf{R}\; e^{-\beta W(\mathbf{R})} \tag{3}$$

where C' and C'' are geometry-independent constants that again depend on zero of energy and standard-state conventions. The quantity W is called the FES or PMF. For a particular choice of C' and C'' we have

$$W(\mathbf{R}) = E(\mathbf{R}|\text{liq}) + \Delta G_{\text{coup}}(\mathbf{R}) \tag{4}$$

where $E(\mathbf{R}|\text{liq})$ is the liquid-phase expectation value of the gas-phase Hamiltonian, and[22a]

$$e^{-\beta\Delta G_{\text{coup}}(\mathbf{R})} = \langle e^{-\beta B(\mathbf{R})}\rangle \tag{5}$$

where B is the coupling energy of the rigid solute to the solvent at structure point $\mathbf{R}$, and the average is over all configurations of the solvent. We let $\mathbf{R}_e$ denote the equilibrium geometry of the solute in the gas phase, and let $\mathbf{R}_l$ be the geometry that minimizes the FES of eq 4. Then

$$G^{\circ}(S|\text{gas}) = E(\mathbf{R}_e|\text{gas}) + G^{\circ}_{\text{trans}}(S|\text{gas}) + G_{\text{con}-\text{rovib}}(S|\text{gas}) \tag{6}$$

and

$$G^{\circ}(S|\text{liq}) = W(\mathbf{R}_l) + G^{\circ}_{\text{lib}}(S|\text{liq}) + G_{\text{con}-\text{rovib}}(S|\text{liq}) \tag{7}$$

where G°_{trans} is the translational free energy, $G_{\text{con}-\text{rovib}}$ is the conformational−vibrational−rotational free energy, and $G^{\circ}_{\text{lib}}(S|\text{liq})$ is the liberational[22a] free energy. Notice that $E(\mathbf{R}|\text{gas})$ is just the gas-phase Born−Oppenheimer potential energy surface sometimes written as $V(\mathbf{R})$.

If we choose standard states in which the concentration of the solute is fixed in the gas phase to be the same as in solution (such a standard state choice is indicated by an asterisk), then it can be shown[22a] that

$$G^{*}_{\text{trans}}(S|\text{gas}) = G^{*}_{\text{lib}}(S|\text{lib}) \tag{8}$$

We define the fixed-concentration free energy of solvation of S as

$$\Delta G^{*}_{\text{S}}(S) = G^{*}(S|\text{liq}) - G^{*}(S|\text{gas}) \tag{9}$$

Combining eqs 4 and 6−9 yields

$$\Delta G^{*}_{\text{S}}(S) = \Delta G_{\text{coup}}(\mathbf{R}_l) + \Delta E_{\text{EN}}(\mathbf{R}_l, \mathbf{R}_g) + \Delta G_{\text{con}-\text{rovib}}(S) \tag{10}$$

where we have defined

$$\Delta E_{\text{EN}}(\mathbf{R}_l, \mathbf{R}_g) = E(\mathbf{R}_l|\text{liq}) - E(\mathbf{R}_g|\text{gas}) \tag{11}$$

and

$$\Delta G_{\text{con}-\text{rovib}}(S) = G_{\text{con}-\text{rovib}}(S|\text{liq}) - G_{\text{con}-\text{rovib}}(S|\text{gas}) \tag{12}$$

For future discussion, we may partition $G_{\text{con}-\text{rovib}}$ into G_{con} and G_{rovib}, where G_{rovib} is the rotational−vibrational free energy in the most populated conformation, and G_{con} is the difference of the free energy calculated by considering all significantly populated conformations from the free energy computed by considering only the most populated conformation. In the gas phase, one often assumes that vibration and rotation are separable, in which case G_{rovib} becomes the sum of a rotational contribution, G_{rot}, and a vibrational one, G_{vib}. The solution-phase analogue of this approximation is also useful, provided that one recognizes that gas-phase rotations are converted into liquid-phase librational motions; this is discussed further below.

Note also that we sometimes label the sum of the polarization free energy[7,23] G_{P} and ΔE_{EN} as ΔG_{ENP}, and similarly, it is convenient to define

$$\Delta G_{\text{EP}}(\mathbf{R}_g) = G_{\text{P}}(\mathbf{R}_g) + E(\mathbf{R}_g|\text{liq}) - E(\mathbf{R}_g|\text{gas}) \tag{13}$$

Next we consider how we proceed in the SM*x* solvation models,[23−27] although similar considerations can be applied to other continuum solvent models. We approximate the solute−solvent coupling by

$$\Delta G_{\text{coup}} = \Delta G_{\text{EP}}(\mathbf{R}_l) + G_{\text{CDS}}(\mathbf{R}_l) \tag{14}$$

where ΔG_{EP} is the electronic−polarization contribution to the solvation free energy based on bulk electrostatics and G_{CDS} is the contribution from cavitation, dispersion, and solvent structural effects. (A note on labeling: in "EP" and "electronic−polarization", the abbreviation E and the word "electronic" refer to the electronic degrees of freedom of the solute and the abbreviation P and the word "polarization" refer to the bulk electric polarization of the solvent, which—at the level of a molecular treatment of the solvent—involves translational, orientational, vibrational, and electronic reorganization of solvent molecules.) The solvent structural effects denoted by the abbreviation S in the CDS term include, by definition, hydrophobic effects, hydrogen bonding, exchange repulsion between solute and solvent, and deviations of first-solvation-shell electrostatics from bulk electrostatics. Note that G_{CDS} is called G_{nes} in ref 1.

The claim in ref 1 that the computer output is mislabeled essentially refers to labeling eq 4 as a free energy. But, as explained above, $W(\mathbf{R})$ does have the status of a free energy, and there is a large literature discussing such free energy quantities. To summarize the issue more informally, we note that the thermodynamic free energy *of a species* is obtained by averaging over all solvent *and solute* degrees of freedom. But in statistical mechanics, one may define free energy functions by averaging over all but one or a few degrees of freedom, or by averaging over all solvent degrees of freedom and no solute degrees of freedom, or indeed by averaging over any arbitrary subset of the total degrees of freedom; the free energy function is then a function of the unaveraged degrees of freedom. In the case of one unaveraged degree of freedom, the free energy function becomes a free energy profile, and in the case of more than one unaveraged degree of freedom, the free energy function becomes a free energy surface. In either case, the free energy function is a potential of mean force. What we calculate directly in our solvation models (and what is routinely calculated in electronic structure packages when self-consistent reaction (SCRF) field calculations are performed) is the free energy surface where the unaveraged degrees of freedom are all the internal degrees of freedom of the solute.

Recall that, in the gas phase, the electronic energy and hence the potential energy surface for nuclear motion are functions of all internal degrees of freedom, but in the lowest order of approximation, one sometimes approximates the thermodynamic functions of species by evaluating the potential energy at a single geometry, in particular the equilibrium geometry. Similar approximations are used as the lowest order approximation in solution; that is, the thermodynamic free energy is evaluated from the potential of mean force at a single geometry. This is a better approximation in some cases than in others. Thus, the claim in ref 1 that the results of implicit solvation models are misused in the literature must be examined on a case-by-case basis, but in most cases cited,[1] the issue is more that the example failed to follow the unnecessarily restrictive prescriptions of ref 1 than that an actual mistake was made. Calculations in solution can employ a hierarchy of approximations, just as one encounters in the gas phase. For example, in the gas phase, one sometimes computes a composite free energy based on vibrational partition functions calculated at a lower level ("double slash" calculations), and sometimes one even neglects vibrational effects. Similar approximations may sometimes be made in liquid-phase reactions. Naturally, one cannot interpret results without reference to the approximations employed, but that does not detract from the potential utility of any individual choice.

3. DISCUSSION

The Comment[1] also makes several interrelated points that bear discussion: (1) Using inaccurate models such as B3LYP for gas-phase energetics, even when augmented by accurate solvation free energies, often leads to unacceptably large errors. (2) The use of $G_{\mathrm{con-rovib}}(S|\mathrm{liq})$ may incur some degree of double counting of solvation effects because the calculated free energies of solvation may already include some of the difference between the liquid-phase and gas-phase values of $G_{\mathrm{con-rovib}}$ due to employing empirical parameters. (3) "Ideal gas [molecular] partition functions, particularly the translational and rotational contributions, are unlikely to be valid in solution. Furthermore the structural re-organization of the solvent is completely neglected through this approach," and any good results obtained this way are "serendipitous". (4) The temperature dependence of the solvation free energy should not be neglected. (5) When structural parameters in solution are substantially different from those in the gas phase, or the conformer distribution changes, the correct approach involves independent treatment of gas- and solution-phase ensembles. (6) Even with independent treatments, such systems may be beyond the reach of current models. (7) Parameterized models should be used in a way consistent with their parametrization. We agree with points (1), (4), (5), and (7), but we agree only partly with (2) and (3), and we think (6) is overly pessimistic, although one's attitude on (2), (3), and (6) depends on the level of accuracy one is pursuing. Concerning point (4), we direct the reader to our work[28−30] on temperature-dependent implicit solvation models for further discussion.

Next, consider point (2). During parametrization, we (and others) have mainly (but not always) used solute training sets for which one may safely neglect both (i) the difference between ΔG_{ENP} at the gas-phase optimized geometry and ΔG_{ENP} at the solution-phase optimized geometry and (ii) $\Delta G_{\mathrm{con-rovib}}$. This does not mean that they are zero, just that they are expected to be smaller than the mean error intrinsic to the continuum solvation model. This justifies neglecting them during parametrization. It does not justify neglecting them for arbitrary applications, where they may on a case-by-case basis be significant, and they should be included when they *are* significant. Nevertheless, one or both of them are small in many cases. When $\Delta G_{\mathrm{con-rovib}}$ is negligible, one may replace $G_{\mathrm{con-rovib}}$ $(S|\mathrm{liq})$ by $G_{\mathrm{con-rovib}}$ $(S|\mathrm{gas})$ in eq 7, and this leads to an equation like eq 4 of ref 1. While the authors of ref 1 take the attitude that one *must* use $G_{\mathrm{con-rovib}}$ $(S|\mathrm{gas})$ with implicit solvation models, our derivation shows that one should use the liquid-phase value when it is expected to differ significantly from the gas-phase value. We neglect this difference during parametrization not because we think that the parameters in G_{P} and G_{CDS} can systematically include the effect in general, but rather because we choose a training set specifically designed to make the effect as small as possible during parametrization. In fact, in the first continuum solvent paper in the SM*x* series,[23] we said,

> "In the present paper, in comparing to experiment for solvation free energies, we will simply assume that the solution-phase electronic-excitation and vibrational contributions are canceled by the gas-phase electronic-excitation and vibrational contributions. (Note especially that we do not claim that the vibrational zero point or thermal contributions are absorbed in the empirical parameters. It will often be useful or necessary to explicitly incorporate the vibrational contributions in later work, e.g., to calculate isotope effects or simply for higher accuracy, and we plan to do this. In fact, the present scheme is specifically designed with the intention that one should be able to use it to calculate vibrational frequencies in the presence of solvent.)"

Even if the effect *is* significant for some solutes in the training set, although some component of the effect might be included in G_{CDS} to the extent that it is approximately group additive, the parametrization is unlikely to capture the effect because the functional form is chosen to capture electrostatics, cavitation, dispersion, and solvent structural effects, not solvatochromic effects on solute vibrations. One exception concerns the conversion of rotational motions into low-frequency modes as discussed in the next paragraph.

Consider next point (3). First of all, translation does not cause a problem. As shown clearly by Ben-Naim, the translational partition function in the gas phase becomes the liberation free energy in solution.[22b] Second, the conversion of gas-phase rotations into solution-phase librations has long been recognized as an issue.[31] For example, a pedagogical chapter[32] in the *Techniques of Chemistry* series says that solution-phase thermodynamics may be based on "partition functions in an adiabatic bath. This may involve important changes from the gas phase in the nature of the normal modes. For example, free rotation is usually missing in solution and is replaced by hindered rotations." This total makeover of low-frequency solute rotational modes is intimately coupled to solvent structural perturbations that are included in the parametrization of cavity−dispersion−solvent-structure terms in SM*x* models, and so it is reasonable to assume that calculations of $\Delta G_{\text{con}-\text{rovib}}$ should include only the $3N - 6$ gas-phase vibrations. In fact that is how it was treated in a recent application.[13] We agree with the Comment[1] that the free energy associated with hindering rotation upon dissolution is partly included in the parametrization (for SM*x* models, it is in the solvent-structural terms) and that it is not included perfectly, but our point is that it is not treated inconsistently or confusingly: it is a well-defined approximation. The low-frequency torsions can also be strongly affected upon dissolution, and that can change the number of solute conformations. That couples this point to point (5), and it is the reason why $\Delta G_{\text{con}-\text{rovib}}$ is actually a change in the conformational−rovibrational partition function, not just a change in the rovibrational partition function in a given conformation. The torsions are also strongly coupled to rotations, and this further complicates the situation. The accurate treatment of the solution-phase thermochemistry of floppy solutes or solute−solvent complexes treated as supermolecular solutes is a challenge for the future.

If, instead of treating the conversion of rotations into librations, we approximate them as unhindered rotations in solution, we find that the dependence of the rotational partition functions on the molecular principal moments of inertia makes them very insensitive to the small changes in solute structure usually associated with solvation. Thus, irrespective of the formal question of whether the rotational partition function should be considered to be physically realistic in solution, its computation for a structure optimized in solution provides a free energy contribution quantitatively indistinguishable from that computed for a gas-phase structure (typically within 0.01 kcal/mol over the data set we describe further below), rendering it a potentially useful quantity when assembling a full solute free energy in solution.

Finally, we consider point (6); in particular, we here make explicit another key aspect of the parametrization of the SM*x* solvation models. In choosing data against which to parametrize, we usually avoided molecules with conformational issues such as multiple conformations that might be expected to be significantly different in the gas phase and solution, and we avoided molecules that might be expected to show large geometry changes upon dissolution. Therefore, we interpret the parametrization as providing those parameters that describe the physical effects involved in fixed-solute-geometry dissolution. If the parametrization were quantitatively accurate (<0.1 kcal/mol), it would include effects such as small geometry changes implicitly. But these effects are usually (not always) smaller than the error in the parametrized method. Thus, when one makes applications to systems where these effects are large (e.g., systems with multiple conformers or solvent-dependent transition states), it is philosophically consistent to include the large effects, and, if the parametrization does capture the fixed-$\mathbf{R}$ physics, the implicit model might be useful if employed with proper ensemble averaging. Another interesting question is whether solvation effects on vibrational frequencies should be treated by equilibrium or nonequilbrium[33] methods, but that is beyond our scope.

4. COMPUTATIONS

To illustrate in practice the points made above, we have computed[34] the vibrational contributions to thermal free energy in the gas phase and in aqueous and carbon tetrachloride solutions (following reoptimization) for molecular systems included in the training set that was used to parametrize the SMD[27] solvation model. Note that we continue to neglect the conformational free energy; furthermore, as recommended above, we include only the contribution of the $3N - 6$ vibrations because the effect of converting rotations to librations (which are low-frequency vibrations) is assumed to be included in the CDS terms through parametrization. We then compared the SMD free energies of solvation calculated with eq 10 (i.e., with account of geometrical changes and changes in vibrational frequencies in solution), and we also computed the SMD free energies of solvation with eq 4 of ref 1, which neglects this variation.

Equation 4 of ref 1 does not explicitly address the issue of the solute geometry used for the computation of ΔE_{EN} and ΔG_{coup}. One may in principle choose to use either the gas-phase geometry (in which case the solute nuclear repulsion remains constant) or the geometry optimized in solution. When the gas-phase geometry is used, $\Delta G_{\text{con}-\text{rovib}}$ will be neglected. However, one may take an intermediate approach where the effects of geometrical relaxation are included in the computation of the solvation free energy but the $\Delta G_{\text{con}-\text{rovib}}$ term is neglected.

We have examined these separate issues by comparing vibrational free energies for structures optimized in different phases and also by computing total solvation free energies for (i) rigid gas-phase structures (determined from a SCRF calculation for the gas-phase geometry and the gas-phase electronic energy, cf. eq 3 of ref 1), (ii) structures optimized in solution *with neglect of* $\Delta G_{\text{con}-\text{rovib}}$ (determined from a SCRF calculation for the solution-optimized geometry and the gas-phase electronic energy for the gas-phase geometry, but *not* requiring a frequency calculation in solution), and (iii) structures optimized in solution and *including* $\Delta G_{\text{con}-\text{rovib}}$ (determined from a SCRF calculation for the solution-optimized geometry, the gas-phase electronic energy for the gas-phase geometry, and a frequency calculation in solution).

In each phase, the optimizations and frequency calculations made use of the M06-2X density functional[35] and the 6-31+G(d,p) basis set.[36] Partition functions were computed using the same quasiharmonic approximation in both the gas phase and solution; the quasiharmonic approximation is the same as the usual harmonic oscillator approximation,[37] except that vibrational frequencies lower than 100 cm^{-1} were raised to 100 cm^{-1} as a way to correct for the well-known breakdown of the harmonic oscillator model for the free energies of low-frequency vibrational modes. Our chosen standard state includes zero-point vibrational energies in the free energies.

The aqueous training set is composed of 274 neutral molecules, 31 clustered ions (explained below), and 71 unclustered ions, and the carbon tetrachloride set contains 78 neutral molecular structures. Table 1 presents the mean differences between thermal contributions to the free energy computed with gas-phase and liquid-phase solute partition functions, whereas Table 2 provides the mean unsigned errors vs experiment for solvation free energies obtained by following the three protocols discussed above.

Table 1 shows that, for neutral molecules in an aqueous solution, the mean unsigned difference (MUD) in the vibrational free energies was computed to be 0.24 kcal/mol, whereas for neutral molecules in carbon tetrachloride, the MUD was 0.14 kcal/mol. In both cases, this average is smaller than the mean inaccuracy of the model (0.7−0.9 kcal/mol) when compared to experiment, thereby justifying the neglect of this effect during parametrization.

Table 1. Mean Differences (kcal/mol) between Gas-Phase Vibrational Free Energies Computed with M06-2X/6-31+G(d,p) and Solution-Phase Vibrational Free Energies Computed with SMD/M06-2X/6-31+G(d,p)

solvent	solutes	MUD[a]	MSD[b]	rmsd[c]
H_2O	274 neutrals	0.24	−0.16	0.32
CCl_4	78 neutrals	0.14	−0.10	0.20
H_2O	31 clustered ions	0.44	−0.11	0.58
		(0.67)[d]	(0.65)	(0.92)
H_2O	81 bare ions	0.14	0.28	0.36

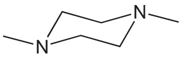
[a] Mean unsigned difference of solution results from gas-phase results. [b] Mean signed difference of solution results from gas-phase results ($G_{\text{vib}}(\text{liq}) - G_{\text{vib}}(\text{gas})$). [c] Root mean square difference of solution results from gas-phase results. [d] The values in parentheses are the results of calculating these ions unclustered, even though we recommend that that they be calculated clustered.

Table 2. Mean Unsigned Errors (MUEs, kcal/mol) as Compared to Experiment for Fixed Concentration Free Energies of Solvation Computed by SMD/M06-2X/6-31+G(d,p) for Gas-Phase and Solution Geometries without and with the Change in Vibrational Free Energies

		geometries;frequencies[a]		
solvent	solutes	gas;gas[b,c]	liq;gas[c]	liq;liq[d]
H_2O	274 neutrals	0.73	0.78	0.91
CCl_4	78 neutrals	0.58	0.59	0.66
H_2O	31 clustered ions	5.38	4.47	4.35
H_2O	81 bare ions	4.50	4.00	4.10

[a] liq denotes solution. [b] The calculations in this column are sometimes denoted as //gas. This is the method used for parametrization. [c] $\Delta G_{\text{rovib}}(S)$ assumed to be zero. [d] $\Delta G_{\text{rovib}}(S)$ computed and included.

As pointed out above, though, for cases where the effect is significantly larger than the average effect, it would be reasonable to include it. An interesting example is offered by *N,N′*-dimethylpiperazine,



which in aqueous solution presents the largest deviation, 1.6 kcal/mol, observed in the data set based on the two partition functions. In the gas phase, this molecule exhibits substantial hyperconjugative delocalization of its nitrogen lone pairs into the σ^* orbitals of vicinal C−H bonds; this well-known effect weakens these bonds, and the resulting red-shifted C−H stretch peaks in alkaloid IR spectra are known as Bohlman bands.[38] In aqueous solution, however, there is less energy penalty associated with highly localized sp^3 N lone pairs, and this manifests in a solution structure with more pyramidal nitrogen atoms, shorter C−H bonds (up to 0.003 Å shorter for those C−H bonds antiperiplanar to the N lone pairs), and an increase in essentially every C−H stretching frequency by 30−50 cm^{-1}. The net change in zero-point vibrational energy associated with these C−H stretches provides most of the 1.6 kcal/mol change in the vibrational free energy. Importantly, it *improves* the computed free energy of solvation for this solute when the change in the vibrational partition function is taken into account, leading to an error compared to experiment of 0.3 kcal/mol compared to 1.4 kcal/mol when the gas-phase vibrational partition function is used.

For calculating the free energy of aqueous solvation of ions, we have divided them into two groups. Group 1 consists of ions with three or fewer atoms, ions in which the magnitude of the partial charge on any oxygen atom is greater than that on the oxygen atom in neutral water, and all (substituted or unsubstituted) ammonium and oxonium ions. The solvation energies of group 1 ions are calculated by clustering them with a single water molecule from the solvent and using a thermodynamic cycle.[39] Other ions (group 2) are calculated unclustered (bare). Table 1 shows a mean unsigned difference between $\Delta G_{\text{vib}}(S)$ in the gas phase and in solution of 0.44 kcal/mol for the (harder-to-model) clustered ions, whereas for unclustered ions, the mean unsigned difference is reduced to 0.14 kcal/mol. Again both effects are small compared to the average inaccuracy of the model as compared to experiment, and, because mean errors in the calculated solvation free energies of ions are larger than those for neutrals (~4 kcal/mol vs 0.7−0.9 kcal/mol), the ratio of the average vibrational effect to the mean error of the model is smaller for ions than for neutrals, especially if one follows the recommendation to calculate certain ions by clustering. Importantly, *clustering* the smaller ions leads to *smaller* changes in the vibrational contributions to free energy than when members of group 1 are computed *without* clustering by a single water molecule (shown in parentheses in the table); this further justifies the use of such clusters beyond those arguments presented previously in ref 39.

The biggest outlier among ions is clustered ammonium, which has a liquid phase vibrational free energy that is less positive by 1.6 kcal/mol compared to its gas-phase counterpart. However, when eq 10 is employed to compute the solvation free energy of clustered ammonium, accounting for the change in vibrational free energy, the error given by the model is 0.5 kcal/mol, whereas neglecting the change in vibrational free energy, as suggested in ref 1, generates an error of 2.2 kcal/mol. Therefore, we conclude that, for clustered ammonium and *N,N′*-dimethylpiperazine (vide supra), adding $\Delta G_{\text{con-rovib}}$ (as approximated by ΔG_{vib}) to the free energy of solvation, as justified by eq 10, substantially improves the accuracy of their liquid-phase treatments.

Table 2 shows that the accuracy on its training set of the SMD continuum solvation model varies, for neutrals, by less than 0.2 kcal/mol when different choices are made for gas-phase vs solution optimized geometries and, for solution geometries, when different choices are made for whether or not to include changes in vibrational free energies. In the case of ions, the choice of vibrational partition functions has the same small effect as observed for neutrals, but the choice of geometry has a larger effect. Nevertheless, the predicted variation is typically only ~10−20% of the mean unsigned error and is therefore well within the mean inaccuracy of the models (and indeed within the experimental uncertainty of measurements). Thus, the results in Tables 1 and 2 show that most of the training set molecules do indeed have only small free energy effects arising from solvation effects on vibrational frequencies and from variations in ΔG_{EP} as a function of geometry choice, and this justifies the neglect of the rovibrational component of the term $\Delta G_{\text{con-rovib}}$ in eq 10 during parametrization.

As noted above, the Comment[1] singled out as inconsistent the use of solution-phase rather than gas-phase rovibrational thermal

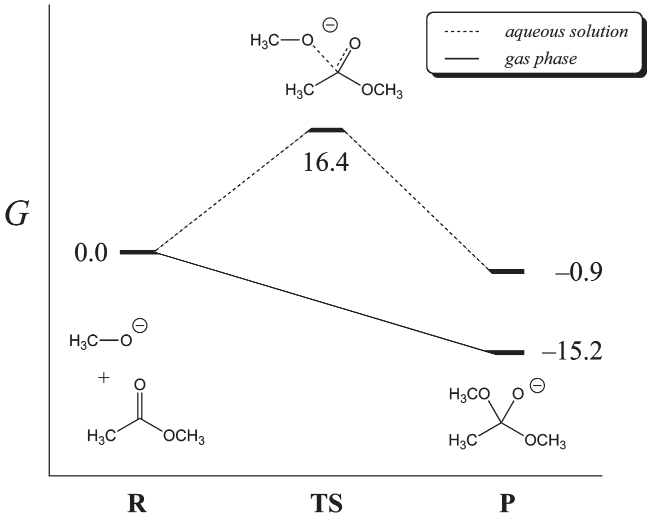


Figure 1. Stationary-point free energies (kcal/mol, 1 M standard state) in the gas phase and aqueous solution for the addition of methoxide to methyl acetate.

contributions when assembling a solute free energy designed to include such contributions (cf. eq 4 of ref 1). Although we have just demonstrated that the magnitude of the difference between the two is sufficiently small for "everyday" molecules that the choice should not be significant, it is important to note that there *are* some instances where stationary points of interest in solution will not *have* corresponding stationary points in the gas phase, making it impossible to compute relevant gas-phase rovibrational contributions for these structures.

For example, as shown in Figure 1, the addition of the methoxide anion to methyl acetate does not have a relevant transition-state structure in the gas phase; the addition to form the so-called tetrahedral intermediate product is spontaneous in the absence of solvation by a polar solvent. It is straightforward, however, to optimize the reactant, product, *and* associated transition-state structures with a solvation model and to compute ideal-gas, rigid-rotator, quasiharmonic-oscillator partition functions from which free energy contributions may be determined. At the SMD/M06-2X/6-31+G(d,p) level, including these contributions for solution-optimized structures, we compute the 298 K free energy of activation for this reaction in water to be 16.4 kcal/mol in the 1 M standard state. If the liberational and rovibrational contributions are neglected, on the other hand, the free energy of activation is predicted to be 5.9 kcal/mol. If *only* the rovibrational free energy contributions are neglected, the free energy of activation is predicted to be 18.1 kcal/mol.

Note that we do not attempt to assess the absolute accuracy of the methoxide addition reaction in Figure 1; we simply use it to highlight that the use of molecular partition functions computed for structures optimized in solution may sometimes be the best choice for a particular modeling problem.

5. SUMMARY AND CONCLUSION

When molecular free energies are defined to include free energy contributions from nuclear motion (including appropriate combinations of translational, liberational, rotational, conformational, librational, and vibrational contributions), it may sometimes be more consistent with parametrization history to determine these contributions from gas-phase partition functions, but it is not "incorrect" to employ instead partition functions for structures that are stationary points in solution. In many instances, the two approaches will give quantitatively similar results, and in instances where they do not, vibrational frequencies determined at the solution structures may well be more appropriate. We take this opportunity to note that although we disagree with some of the conclusions in ref 1, we are grateful to the authors for having raised these issues constructively because we think it is very worthwhile to have a thorough discussion of these points in the literature.

AUTHOR INFORMATION

Corresponding Author

*E-mail: cramer@umn.edu (C.J.C.), truhlar@umn.edu (D.G.T.).

ACKNOWLEDGMENT

The authors are grateful to Michelle Coote and Junming Ho for comments on the original version of the manuscript. This work was supported in part by the U.S. Army Research Laboratory under Grant No. W911NF09-1-0377 and the National Science Foundation under Grants No. CHE09-52054 and CHE09-56776.

REFERENCES

(1) Ho, J.; Klamt, H.; Coote, M. L. *J. Phys. Chem. A* **2010**, *114*, 13442.

(2) Frisch, M. J. et al. *Gaussian 03*, rev. E.01; Gaussian, Inc.: Wallingford, CT, 2004.

(3) Gordon, M. S.; Schmidt, M. W. In *Theory and Application of Quantum Chemistry: The First Forty Years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, 2005; p 1167.

(4) Higashi, M. et al. *GAMESSPLUS*, version 2008−2; University of Minnesota: Minneapolis and St. Paul, MN, 2008.

(5) Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J. *J. Phys. Chem.* **1996**, *100*, 12771.See especially page 12787.

(6) Chuang, Y.-Y.; Cramer, C. J.; Truhlar, D. G. *Int. J. Quantum Chem.* **1998**, *70*, 887.

(7) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161.See especially page 2186.

(8) Martin, F. P.-D.; Dumas, R.; Field, M. J. *J. Am. Chem. Soc.* **2000**, *122*, 7688.

(9) Schenter, G. K.; Garrett, B. C.; Truhlar, D. G. *J. Phys. Chem. B* **2001**, *105*, 9672.

(10) Poulsen, T. D.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. B* **2003**, *107*, 9567.

(11) Cummins, P. L.; Rostov, I. V.; Gready, J. E. *J. Chem. Theory Comput.* **2007**, *3*, 1203.

(12) Ruiz-Pernia, J.; Garcia-Viloca, M.; Bhattacharyay, S.; Gao, J.; Truhlar, D. G.; Tuñon, I. *J. Am. Chem. Soc.* **2009**, *131*, 2687.

(13) Kim, Y.; Mohrig, J. R.; Truhlar, D. G. *J. Am. Chem. Soc.* **2010**, *132*, 11071.

(14) Schlick, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide*; Springer: New York, 2002; p 294.

(15) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300.

(16) McQuarrie, D. A. *Statistical Mechanics*; Harper and Row: New York, 1976; (a) pp 266−267, 331−334, (b) p 58, 116.

(17) Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275.

(18) Schenter, G. K.; Garrett, B. C.; Truhlar, D. G. *J. Chem. Phys.* **2003**, *119*, 5828.

(19) Zhou, H.-X. *J. Chem. Phys.* **2008**, *128*, 114104.

(20) Trzesniak, D.; Kunz, A.-P. E; van Gunsteren, W. F. *ChemPhysChem* **2007**, *8*, 162.

(21) Darve, E. In *Free Energy Calculations*; Chipot, C., Pohorille, A., Eds.; Springer Series in Chemical Physics 86; Springer: New York, 2007; pp 119−126.

(22) Ben-Naim, A. *Statistical Thermodynamics for Chemists and Biochemists*; Plenum: New York, 1992; (a) p 423, (b) p 321.

(23) (a) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8305. (b) Cramer, C. J.; Truhlar, D. G. Erratum. *J. Am. Chem. Soc.* **1991**, *113*, 9901.

(24) Giesen, D. J.; Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. In *Computational Thermochemistry*; Irikura, K., Frurip, D. J., Eds.; ACS Symposium Series 677; American Chemical Society: Washington, D.C., 1998; pp 285−300.

(25) Cramer, C. J.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 760.

(26) Cramer, C. J.; Truhlar, D. G. *Acc. Chem. Res.* **2009**, *42*, 493.

(27) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378.

(28) Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2006**, *110*, 5675.

(29) Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2008**, *120*, 3024.

(30) Chamberlin, A. C.; Levitt, D. G.; Cramer, C. J.; Truhlar, D. G. *Mol. Pharmaceutics* **2008**, *5*, 1064.

(31) Stern, M. J.; Van Hook, W. A.; Wolfsberg, M. *J. Chem. Phys.* **1963**, *39*, 3179.

(32) Kreevoy, M. M.; Truhlar, D. G. In *Investigations of Rates and Mechanisms of Reactions Part I*; Bernasconi, C., Ed.; Techniques of Chemistry Series, 4th ed.; John Wiley and Sons: New York, 1986; Vol. 6, pp 13−95.

(33) Cappelli, C.; Corni, S.; Cammi, R.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2000**, *113*, 11270.

(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; 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 D.01; Gaussian, Inc.: Wallingford, CT, 2004.

(35) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.

(36) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.

(37) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; John Wiley and Sons: Chichester, U.K., 2004; pp 357−366.

(38) Cramer, C. J. Hyperconjugation. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley and Sons: Chichester, U.K., 1998; p 1294.

(39) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133.