THE JOURNAL OF PHYSICAL CHEMISTRY A

pubs.acs.org/JPCA Article

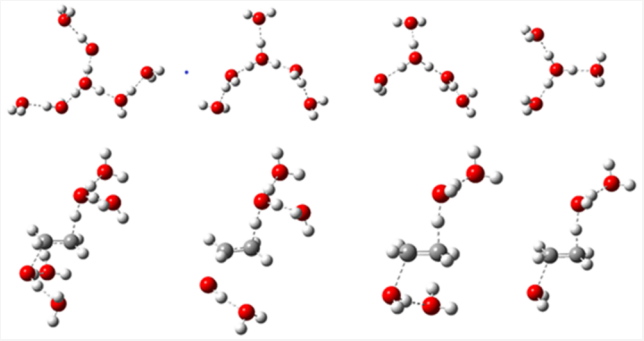
Semicontinuum (Cluster-Continuum) Modeling of Acid-Catalyzed Aqueous Reactions: Alkene Hydration

Darpan H. Patel and Allan L. L. East*

Cite This: *J. Phys. Chem. A* 2020, 124, 9088–9104 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: An effort is made to reduce the errors of continuum solvation models (CSMs) with semicontinuum modeling to achieve 3 kcal mol^{-1} agreement with experiment for acid-catalysis activation Gibbs energies. First, two underappreciated CSM issues are reviewed: errors in the CSM solvation Gibbs energies grow beyond 5 kcal mol^{-1} (i) as ions are made smaller and (ii) as water clusters grow larger. Second, the computational reproduction of the known Gibbs energies (Δ_rG and $\Delta^{\ddagger}G$) of the paradigmatic reaction ethene + H_2O + H_3O^+ → TS^+ → ethanol + H_3O^+ is attempted. It is argued that, despite the >5 kcal mol^{-1} solvation errors for ions, it is possible to employ error cancellation strategies to reduce the errors in the reaction and activation Gibbs energies to 3 kcal mol^{-1} accuracy. A new 3 kcal mol^{-1} effect due to solvent-molecule "placement" (confinement from 1 M bulk concentration) was isolated and proved useful. Third, computational reproduction of the known entropies (Δ_rS and $\Delta^{\ddagger}S$) of the paradigmatic reaction is attempted using Trouton's constant and neglect of solvent structure reorganization effects (which must cancel well for this reaction); this worked well for Δ_rS but needs empirical correction of ∼11 cal mol^{-1} K^{-1} for $\Delta^{\ddagger}S$ due to solvent disorientation when H_3O^+ is consumed. These entropy estimates allow for enthalpy (Δ_rH and $\Delta^{\ddagger}H$) estimation from the Gibbs energy values. Fourth, two recommended options, including A + H_3O^+·2W → $[AHOH_2^+$·2W$]^{\ddagger}$, are shown to also work well for the activations of propene and isobutene.



1. INTRODUCTION

In quantum chemistry, continuum solvation models (CSMs)[1−3] are commonly used to account for the effects of solvation. This paper will focus on the most popular methods available in Gaussian09 and Gaussian16: the default model (IEFPCM[4−6] or its sister CPCM[7]), which computes the main long-range dielectric effect, and the more elaborate SMD model of Marenich, Cramer, and Truhlar,[8] which contains additional energy terms (and some empirical parameters) to account for additional solvation effects such as cavitation, dispersion-attraction, and strong local interactions with the first solvation shell. An underappreciated problem is that the expected errors from these CSMs can be large (>5 kcal mol^{-1}). We are in pursuit of ways that can either reduce or better manage these errors, for predicting reaction (Δ_rG) and activation ($\Delta^{\ddagger}G$) Gibbs energies.

One example of large CSM errors is with the aqueous solvation of ions. The original SMD paper of 2009,[8] which employed an explicit water molecule with the ion inside the CSM cavity for 31 of the 112 ions studied (including H_3O^+), reported (its Table 9) a $\Delta_{solv}G$ root-mean-square error of 5.5 kcal mol^{-1} for these 112 aqueous ions. That added explicit water molecule was rather important; without it, SMD makes an error of +25 kcal mol^{-1} on the water autoionization energy $2H_2O_{(l)}$ → $H_3O^+_{(aq)}$ + $OH^-_{(aq)}$.[9] The problem with small aqueous ions is their strong local interaction with the first solvation shell, due to a high ion charge density (high q/r) and/or perhaps specifically to the resulting high partial charges on solute atoms in contact with solvent water. This CSM error is termed the *high q/r error* (ε_{hqr}) in the present paper.

An old (1994−2002) means for treating the high q/r error was empirical cavity radius scaling for ions.[10−13] Unfortunately, such corrections become very system specific,[14] limiting wide applicability. An additional objection is the unphysical nature of the small and inconsistent cavities required to match solvation energies of anions.[12,13] In 2002, Chipman,[13] who noted that these problems did not exist with less polar solvents like DMSO, referred to cavity scaling for ions in water as an "artificial distortion" of the cavity that tries "to force dielectric continuum theory by itself to reproduce other things that it is ill-suited to describe." The Gaussian programs,[15] as well as SMx models,[16] soon dropped the idea of cavity scaling for ions. Surprisingly,

Received: July 30, 2020
Revised: October 2, 2020
Published: October 19, 2020

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

cavity scaling is still occasionally invoked, particularly for pK_a computations.[17−22]

A second means of dealing with the high q/r error is to add terms to the CSM energy calculation to account for first-solvation shell effects. The ΔG_{CDS} term of the Minnesota SMx[16,23,24] and SMD[8] methods is meant to include such effects, requiring atom-based empirical parameters and a solvent-specific cavity radius correction r_s. Despite these empirical parameters, the accuracy of SMD for ions was still limited, and the 5.5 kcal mol^{-1} RMS accuracy in the original paper was achieved by also incorporating the third means of dealing with high q/r error: addition of an explicit solvent molecule inside the cavity (in 31 of the 112 cases), to reduce the q/r charge density to a more manageable level.

This third means of dealing with the high q/r error has various names: semicontinuum,[25] cluster-continuum,[26] and discrete-continuum[27] modeling. This technique reduces high q/r by spreading the charge across not only the solute but also one (or more) explicit water molecules. It treats the strong short-range effects not with system-specific cavity scaling or added terms with empirical parameters, but with the power and rigor of ab initio quantum chemistry. Not only might such modeling reduce the level of empiricism in dealing with short-range effects, but two particular studies, by Bryantsev/Diallo/Goddard[28] and Dhillon/East,[9] offered tantalizing evidence that *asymptotic convergence* (with continual addition of explicit waters) can lead to <3 kcal mol^{-1} accuracy even for notoriously difficult ions. Bryantsev et al. investigated the solvation Gibbs energies of H_3O^+ and Cu^{2+} while Dhillon et al. investigated the reaction Gibbs energy for water autoionization, and in both papers, CSM errors of >20 kcal mol^{-1} were systematically cured by addition of explicit waters W in a balanced way ($AW_n^{q\pm}$ vs W_n clusters), achieving asymptotic convergence to the correct answers.

There is another large CSM error of concern. Although not understood at the time, the need for balanced clusters in those ion studies is shown here to be due to the need to cancel out a different type of CSM error that accumulates with increasing cluster size. This CSM error is termed the *imperfect cavity size error* (ε_{ics}) in the present paper, only because we can show (section 3) that a simple scaling of cavity radii can reproduce (and hence arbitrarily eliminate) the error. This rising error will be demonstrated for a variety of CSM models.

As will be demonstrated, both the high q/r error and the imperfect cavity size error can individually rise above 5 kcal mol^{-1}, even for SMD, and hence more work is needed to establish reliable and accurate protocols for semicontinuum modeling, to reduce the impact of these errors. From our work to date[9,17,29−31] we had developed three semicontinuum modeling principles we thought reasonable for aqueous reactions: conservation (reactants vs products) of total hydrogen-bonds added, conservation of number of cavities via stoichiometry ($\Delta\nu = 0$, e.g., A + B → C + D), and the need (when using basic CSMs) to fill the first solvation shell for H_3O^+ and OH^-. The two conservation principles served to cancel out the imperfect cavity size errors, while the third principle served to minimize or eliminate the high q/r error of small ions. Now, in moving on to applying these principles to the case of activation Gibbs energy of acid catalysis, complications arise with the application of the conservation principles, since (i) *stoichiometric change* ($\Delta\nu \neq 0$) challenges the cavity conservation effort and (ii) the *proton-transfer transition state* challenges the H-bond conservation effort. While the use of three explicit waters for H_3O^+ (the third principle) is still anticipated to solve any issues with high q/r, this principle arose from an asymptotic study of water autoionization only and certainly needs to be tested for other reactions and activations involving H_3O^+.

An additional issue we sought to examine in this paper is the division of CSM-generated ΔG values into *entropy and enthalpy* components.

For an exemplary acid-catalyzed reaction to study, a paradigmatic one was chosen where the change in not only Gibbs energy but also enthalpy and entropy are all known, for both reaction and activation: acid-catalyzed hydration of ethene to ethanol. For activation at T = 298.15 K, $\{\Delta^{\ddagger}S, \Delta^{\ddagger}H, \Delta^{\ddagger}G\}$ = {−5 ± 3 cal mol^{-1} K^{-1}, 32.7 ± 1.0 kcal mol^{-1}, 34.1 ± 1.0 kcal mol^{-1}}, derivable (see Supporting Information) from the reaction rate constants of Baliga and Whalley.[32] For overall reaction at T = 298.15 K, $\{\Delta_r S, \Delta_r H, \Delta_r G\}$ = {−18 cal mol^{-1} K^{-1}, −9.2 kcal mol^{-1}, −3.8 kcal mol^{-1}}, derivable (see Supporting Information) from heats of formation, third-law entropies, and the hydration (solvation) data tabulated by Cabani et al.[33] All these values correspond to 1 M of each species (i.e., liquid H_2O as well as solutes). In the literature are three previous attempts to reproduce via modeling the activation energy for this model reaction. Two were poor $\Delta^{\ddagger}G$ estimates from molecular dynamics (MD) simulations: a paper in 2004 by van Erp and Meijer[34] reported $\Delta^{\ddagger}G$ = 23 kcal mol^{-1} (too low by 10 kcal mol^{-1}) from a density-functional-theory (DFT) MD simulation (BLYP[35,36] forces), and a paper in 2007 by Tolosa Arroyo et al.[37] reported $\Delta G^{\ddagger}$ = 36 kcal mol^{-1} but for the *uncatalyzed* hydration of ethene (hence also too low by a significant amount) from a classical MD simulation employing an empirical force field. However, a third paper, in 2017 by Yamabe, Tsuchida, and Yamazaki,[38] reported an apparently successful reproduction of the activation energy ($E_a = \Delta^{\ddagger}H + RT$ = 33.3 kcal mol^{-1}), obtaining $\Delta^{\ddagger}E_{0K}$ = 32.5 kcal mol^{-1} using a Berkeley DFT,[39] ωB97x-D/6-311+G(d,p), with a semicontinuum model consisting of 13 explicit waters and the default polarizable continuum model (SCRF = IEFPCM[4−6]) in Gaussian09;[40] however, we are unable to reproduce their result (see the Supporting Information), and we anticipated that far fewer water molecules would be sufficient.

The paper proceeds as follows: section 2 is the Methods section, section 3 is the examination of CSM solvation errors and the new solvent confinement correction (SCC), section 4 is the thermodynamic results for alkene hydrations (reaction and activation Gibbs energies, followed by entropies and enthalpies), and section 5 is the Conclusions.

2. METHODS

All calculations employed Gaussian09.[40] Several CSMs (continuum solvation models) and ESMs (electronic structure methods) were tested. Geometry optimizations primarily used M06-2X/6-31G(d)/CPCM, following Sturgeon et al.;[41] here M06-2X[42,43] refers to the ESM and CPCM[7] refers to the CSM. Other ESMs tested include ωB97x-D,[39] BLYP,[35,36] MP2,[44] MP4,[44] and CCSD(T),[45] and in the last three the post-Hartree−Fock correlation was applied with no further alteration of the Hartree−Fock solvent reaction field (i.e., PTE mode[46]). Extended basis sets were also tested. To derive CSM errors for water cluster solvation, accurate gas-phase energies were needed for water cluster dissociation, and for this, the composite ESM known as Gaussian-4 (G4) theory was employed.[47]

A conventional[2,48] approach to CSMs (continuum solvation models) is followed but a new term unique to semicontinuum modeling, $\Delta_{SCC}G$, is introduced. Let $\Psi(0)$ and $\Psi(f)$ be the wave

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

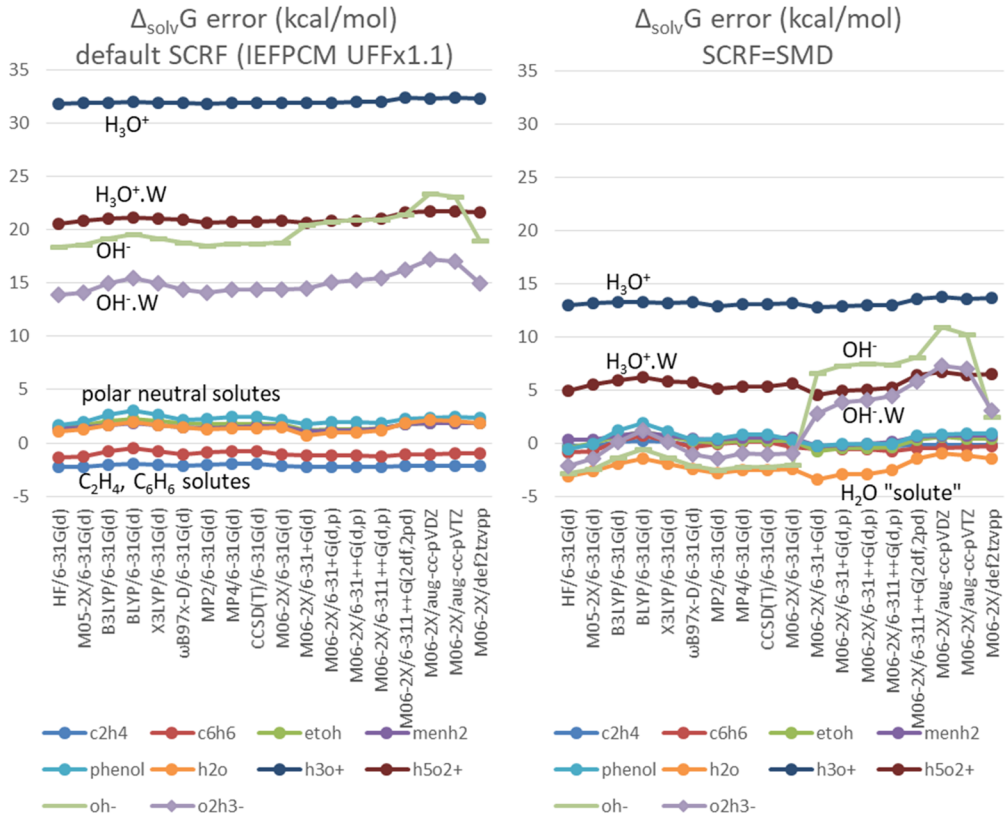


Figure 1. Errors in computed $\Delta_{solv}G$ for 1 M gas → 1 M solution, using two popular CSMs available in Gaussian09: IEFPCM (left plot) and SMD (right plot). Eighteen ESMs (electronic structure methods) were tried with each CSM, in single-point runs (gas, SCRF, SCRF = SMD) using geometries preoptimized using M06-2X/6-31G(d)/CPCM. Note for both IEFPCM and SMD the large errors for the ions, H_3O^+ (uppermost curve) and $H_5O_2^+$ (second highest curve).

function for the solute electrons in gas and solution, respectively. In solution, the free energy of activation or reaction at standard concentration (1 M) can be written as the stoichiometrically weighted difference of $G_{solution}$ values, these being the molar Gibbs energies of a 1 M solution of each solute (eq 1):

$$G(1\text{ M}) = G_{gas}(1\text{ atm}) + \Delta_{conc}G(1\text{ atm} \rightarrow 1\text{ M}) + \Delta_{solv}G(1\text{ M}) + \Delta_{SCC}G(1\text{ M}) \quad (1)$$

$$G_{gas} = E_{elec,gas} - TS_{elec} + (PV)_{gas} + E_{tm,gas} - TS_{tm,gas}(1\text{ atm}) \quad (2a)$$

$$\Delta_{conc}G = +1.9\ \frac{\text{kcal}}{\text{mol}} \quad (2b)$$

$$\Delta_{solv}G = \Delta_{solv}E_{elec} + \Delta_{solv}G_{\varepsilon} + \Delta_{solv}G_{other} \quad (2c)$$

$$\Delta_{SCC}G = -3n\ \frac{\text{kcal}}{\text{mol}} \quad (2d)$$

In eq 2a, $E_{elec,gas} = \langle\Psi(0)|\hat{H}|\Psi(0)\rangle$, $(PV)_{gas} = RT$, and the subscript tm indicates contribution from thermal motions (translation, rotation, vibration). In eq 2b, $\Delta_{conc}G = +1.9$ is the concentration effect $-T\Delta_{conc}S = -RT\ln(V_f/V_i) = -RT\ln(1/24.466)$ for compression of the perfect gas from its standard concentration ($P = 1$ atm, $T = 298.15$ K, $V = 24.466$ L mol^{-1}) to a 1 M concentration ($V = 1$ L mol^{-1}). In eq 2c, the first term is the polarization of E_{elec} due to solvation ($\langle\Psi(f)|\hat{H}|\Psi(f)\rangle - \langle\Psi(0)|\hat{H}|\Psi(0)\rangle$), while the second term $\Delta_{solv}G_{\varepsilon} = \langle\Psi(f)|\hat{V}_{\varepsilon}/2|\Psi(f)\rangle$ is the dielectric interaction integral with its traditional labeling as a free energy ($\Delta_{solv}G_{\varepsilon}$) in consideration of an isothermal solvation process.[49] Two other symbols require special mention.

The third term in eq 2c, $\Delta_{solv}G_{other}$, refers to all other solvation terms and corrections not computed by default in Gaussian09 and Gaussian16. Some examples of extra terms we will label as CAV (cavitation), REP (repulsion), DIS (dispersion), SSR (solvent structure reorganization), and DAM (solute entropy damping due to solvent). For corrections for CSM errors of unknown origin, we will use the labels HQRC (high q/r correction) and ICSC (imperfect cavity size correction) for components of $\Delta_{solv}G_{other}$. The SMx CSMs of Cramer and Truhlar[8,16,23,24] also have a $\Delta_{solv}G_{other}$ term, $\Delta_{solv}G_{CDS}$, with its subscript highlighting the CAV/DIS/SSR effects.

The new fourth term in eq 1, $\Delta_{SCC}G$, is a *solvent confinement correction*, to be applied to explicitly hydrated solutes AW_n (A = solute, W = water) to correct for the fact that each explicit hydration, under the 1 M concentration convention, results in a significant artificial entropy loss. The value $3n$ is derived in section 3.3.

Substituting eq 2a−2d into eq 1 gives 10 terms. Let us discard one ($TS_{elec} = 0$ for species in closed-shell singlet electronic states, as in ethene hydration), and group the remaining 9 down to 5 user-convenient terms (eq 3):

$$G(1\text{ M}) = G_{el,\varepsilon} + G_{freq}(1\text{ atm}) + \Delta_{conc}G(1\text{ atm} \rightarrow 1\text{ M}) + \Delta_{solv}G_{other} + \Delta_{SCC}G(1\text{ M}) \quad (3)$$

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

$$G_{\mathrm{el},\varepsilon} \equiv E_{\mathrm{elec,gas}} + \Delta_{\mathrm{solv}}E_{\mathrm{elec}} + \Delta_{\mathrm{solv}}G_{\varepsilon} = \langle \Psi(f) \left| \hat{H} + \frac{1}{2}\hat{V}_{\varepsilon} \right| \Psi(f) \rangle \quad (4a)$$

$$G_{\mathrm{freq}}(1\ \mathrm{atm}) \equiv E_{\mathrm{tm,gas}} + (PV)_{\mathrm{gas}} - TS_{\mathrm{tm,gas}}(1\ \mathrm{atm}) \quad (4b)$$

In eq 3, $G_{\mathrm{el},\varepsilon}$ is the energy reported by default from a Gaussian09 (or Gaussian16) SCRF run, and G_{freq} is the "thermal correction to the Gibbs free energy" reported by Gaussian from a traditional vibrational frequency run, which uses gas-phase rigid-rotor/harmonic-oscillator (RRHO) and $PV = RT$ formulas. For G_{freq} it matters little if the continuum is on or not during the frequency run,[50] and therefore, the terms in eq 3 can be computed with a single electronic structure method (ESM) in a single run per solute, appending tail-end corrections $\Delta_{\mathrm{conc}}G$ = 1.9 and $\Delta_{\mathrm{SCC}}G = -3n$. Others may use different ESMs for different pieces of eq 3. For instance, Pliego[51,52] often uses three different ESMs for this equation: a high-accuracy ESM for $E_{\mathrm{elec,gas}}$, a medium-accuracy one for $\Delta_{\mathrm{solv}}G$, and a low-cost one for G_{freq}. Gaussian09 and Gaussian16 users employing PCM CSMs (except SMD) are omitting $\Delta_{\mathrm{solv}}G_{\mathrm{other}}$ (e.g., cavitation) by default.

The semicontinuum models, needed to cure the high q/r error by CSMs on H_3O^+, are presented with their results in section 4. A referee requested that we comment on the concern, held by some, that sampling of solvent-molecule positions may be needed to ensure sufficient accuracy (for instance, an energy or time average of energies from 100 snapshots from a Monte-Carlo or molecular dynamics simulation). The usual hypothesis, which so far has stood up well, is that any error caused by choosing only one configuration either is small in an absolute sense or is canceling out well in a $\Delta_r G$ or $\Delta^{\ddagger}G$ calculation. Here, as in water autoionization, we have tested placement options and spotted no evidence that such sampling is required.

3. RESULTS AND DISCUSSION I: CSM SOLVATION ERRORS

3.1. Single Molecules. We start with Figure 1, which shows errors from computed CSM solvation energies $\Delta_{\mathrm{solv}}G$ (eq 2c; no frequency runs needed) for 10 exemplary solutes in water: 5 neutrals, 4 ions, and the case of H_2O into H_2O (relevant when water is a reactant or product, as in ethene hydration or semicontinuum clusters). The Zundel ion (monohydrated hydronium ion) is symmetrically bridged, $H_2O{\cdot}H{\cdot}OH_2^+$, but the monohydrated hydroxide ion is not, $HOH{\cdot}OH^-$. The basic default CSM (IEFPCM with UFF[53] x1.1 radii) and the more elaborate and popular alternative SMD were both tested, with 18 different ESMs to show the degree of insensitivity to ESMs.

The errors from the default algorithm (IEFPCM UFFx1.1, left plot) are reasonable for the neutral molecules: the errors are less than 3 kcal mol^{-1}, negative for the hydrocarbons while positive for the polar species. However, the errors for the cations are enormous, and rather imbalanced for H_3O^+ vs OH^-. SMD does much better, of course: exceptionally well for the small neutral solutes, an error of −2 to −3 for H_2O into H_2O (unless large basis sets are used, not normally recommended[8] for SMD) and +5 for the two monohydrated ions (assuming diffuse basis sets for anions).

According to Figure 1, adding one water to H_3O^+ (making $H_5O_2^+$) or OH^- (making $O_2H_3^-$) clearly reduces the *high q/r error* of both these CSMs, but in both cases a sizable error yet remains. For some reactions involving small ions (e.g., HCl + $Br^- \rightarrow HBr + Cl^-$), this error may cancel out well in the reaction Gibbs energy. Unfortunately, for acid-catalyzed reactions, in steps in which H^+ is transferred from H_3O^+ to a larger organic substrate R, the q/r of H_3O^+ vs RH^+ are significantly different, leading to imbalanced q/r errors, and the reaction or activation Gibbs energy may be severely underpredicted. The proposal of Dhillon and East,[9] to use the Eigen ion ($H_3O^+{\cdot}3H_2O$) in place of H_3O^+, eliminated the problem of the high q/r error of the default CSM algorithm by reducing q/r: making the ion radius r larger, dispersing the cationic charge over a greater volume and surface, effectively lowering q/r below a threshold that the CSM can properly handle. We are thus optimistic that semicontinuum models that add more than one water molecule to ions can reduce CSM solvation errors further than the 5 kcal/mol achievement of the original SMD paper.

Adding more than one water molecule could, however, accumulate *imperfect cavity size* error as mentioned in the Introduction. Note that the errors observed for "solute" H_2O (H_2O into H_2O) are +2 kcal mol^{-1} from IEFPCM and −2 kcal mol^{-1} from SMD. This error would accumulate as more loose waters are considered, e.g., 8 kcal mol^{-1} error for four loose waters. Would this error also accumulate for the increasingly hydrated clusters of semicontinuum modeling? The answer is yes, but often in the opposite direction, as we shall show by examining the water cluster cycles of Bryantsev et al.[28] next.

3.2. Water Clusters. In their paper, Bryantsev, Diallo, and Goddard[28] were examining semicontinuum methods to predict ion solvation energies, $\Delta_{\mathrm{solv}}G(A^{q+})$, to improve upon error-prone direct CSM calculation on bare ions. They tried two techniques, which they called the cluster and monomer methods. The cluster method (using CSMs for AW_n^{q+} and W_n) worked well, converging to the right answer as the number n of added waters W increases, but the monomer method (using CSMs for AW_n^{q+} and n W) failed miserably, never getting within 14 kcal mol^{-1} of the right answer and getting worse (diverging) as n increased beyond ∼7.

They went further and isolated the problem to the act of clustering. Their water clustering cycle (Figure 2) became a useful point of focus. The ΔG for the process of W_n(gas) → nW(liquid) should have been independent of whether the CSM is used before or after cluster dissociation (which we call the "slush" or "vapor" pathways, respectively, in Figure 2). Bryantsev et al. found that the results disagreed. Calling this discrepancy ($\Delta G_{\mathrm{via\ slush}} - \Delta G_{\mathrm{via\ vapor}}$) the "residual error of closing the cycle", they showed that the discrepancy (i) grew linearly with size of cluster n, becoming quite alarming (up to 50 kcal mol^{-1} and beyond for Jaguar and SM6) and (ii) had a growth rate that depended on the CSM used. Riccardi et al.[54] repeated the analysis using Gaussian's default IEFPCM and SMD and showed that its discrepancy was dependent on cavity radii used, with mildest (but still positive) error growth from default (UFFx1.1) cavity radii.

The underlying reasons for these growing residual errors, as we now show, are ever-growing CSM errors for both nW and W_n as n increases. The residual error grows with n because the CSM errors made on nW and W_n grow with n at different rates. Bryantsev et al. speculated that perhaps the error in the W_n calculation was a rigidity problem in G_{freq}. It is not. The error in the W_n calculation is in the CSM term $G_{\mathrm{el},\varepsilon}$ and grows with cluster size. Our group has been calling this type of CSM error the *imperfect cavity size error* of CSMs, for reasons soon explained.

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

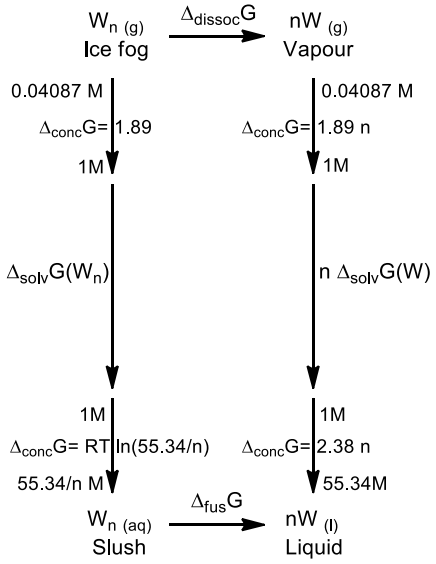


Figure 2. Water cluster cycle. The concentration terms $\Delta_{conc}G = RT \ln c_2/c_1 = -RT \ln V_2/V_1$ are explicitly shown, with values in kcal mol^{-1}. The terms "ice fog" and "slush" reflect our *nanocrystal assumption* (that the water cluster W_n is solid). The original cycle of Bryantsev et al.[28] did not use these terms, for they had assumed $\Delta G = 0$ for the bottom leg, considering the water cluster W_n to represent a liquid (the *nanodroplet assumption*).

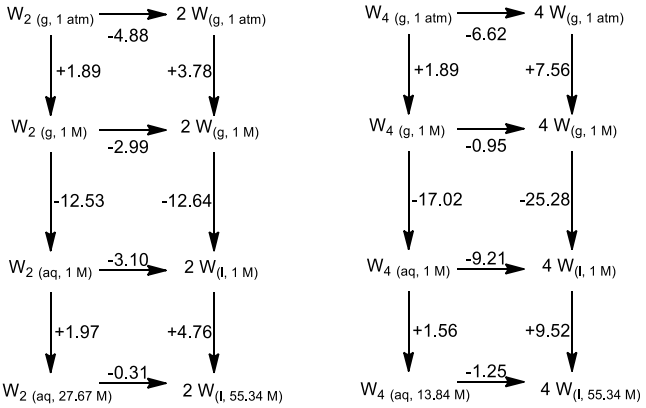


Figure 3. Completed water cluster cycles (zero residual error) for W_2 and ring-W_4 clusters under the nanocrystal assumption (see Figure 2), showing "expt" values (in kcal per mole of W_n) for Gibbs energies of each leg. Values in the outermost cycles are experimentally known ones, except the top horizontal leg (1 atm → 1 atm dissociation: G4 values) and the vertical leg for W_n solvation (1 M → 1 M: derived from requiring zero residual error). The values of roughly $-3(n-1)$ for the third horizontal leg (1 M → 1 M melting of slush) relate to $\Delta_{SCC}G$ (section 3.3).

First, we need reasonably accurate estimates of $\Delta_{solv}G$ for W_n clusters (W = one H_2O molecule), using the water clustering cycle first discussed by Bryantsev et al.[28] (Figure 2). One only needs an accurate estimation of $\Delta_{dissoc}G$, the gas-phase cluster dissociation, which we obtained with G4[47] computations. With such predictions for $\Delta_{dissoc}G$, and the use of experimentally known $\Delta_{vap}G$ and $\Delta_{fus}G$ values for water/ice (using $\Delta_{solv}G(W) = -\Delta_{vap}G[1\text{ M} \rightarrow 1\text{ M}] = 2.05$ kcal mol^{-1}), accurate ("expt") predictions for $\Delta_{solv}G(W_n)$, were obtained. This allows quantitative "completion" (zero residual error) of water cluster cycles, and Figure 3 shows two such completed cycles for the W_2 and W_4 clusters.

Second, these derived "expt" values of of $\Delta_{solv}G$ for W_n clusters are compared to the published CSM results (Table 1). The table shows two versions of "expt" values, derived using either the nanocrystal assumption (nca) used in the present paper (which employs $\Delta_{fus}G$ in the bottom leg of Figure 2) or the nanodroplet assumption (nda), made by Bryantsev et al. (which assumes no Gibbs energy difference in the bottom leg of Figure 2). All the CSMs in Table 1 (except IEFPCM UFFx0.8) underpredict the magnitude of $\Delta_{solv}G$ for the large clusters. To make this even clearer, the CSM errors (under the nanocrystal assumption for W_n) are plotted in Figure 4, clearly showing that the CSM predictions (except UFFx0.8) increasingly deviate from the derived "expt" values: the CSM error grows as the cluster W_n grows. The CSM predictions would be slightly worse (the error curves in Figure 4 would rise faster) under the nanodroplet assumption.

The magnitude and rate of rise of the errors in Figure 4 are strongly dependent on cavity radii (see Supporting Information for a summary table of radii). SM6[16] and Jaguar[55] provide tight cavities and hence more mildly rising errors, while UFFx1.1 offers the largest cavities and hence more dramatically rising errors. The UFFx1.1 and UFFx0.8 data, which bracket all others, differ only by the scaling factor used on all the UFF atomic radii (1.1 vs 0.8). This implies that one could very well reproduce the Figure 4 results of any of those CSMs simply by taking simple IEFPCM and scaling the UFF radii by a particular factor. This is why our group has been calling this particular CSM error the *imperfect cavity size* error (see Introduction). Unlike the high q/r error (section 3.1), which is *reduced* as waters are added to ions, the imperfect cavity size error is likely *increasing* as waters are added to solutes.

This imperfect cavity size error will asymptotically increase for any water-solvated solute, but it may start off being negative in the absence of explicit hydration. An "error-switch" from negative to positive would then have to occur at some level of explicit hydration. This *error-switch phenomenon* even exists in pure water clusters (as exposed by Figure 4): all CSMs there (except UFFx1.1 and UFFx0.8) produce negative errors for water monomer but positive errors for larger clusters. One way to explain this phenomenon is with a group additivity model for the solvation of water clusters:

$$\Delta G_{solv}(W_n) \approx \sum_{i=1}^{n} c_i \approx \sum_{j=1}^{\text{W types}} n_j c_j \qquad (5)$$

where the contribution c_i of water submolecule i to the total solvation energy of the cluster depends roughly on the contribution of the submolecule to the cavity size, perhaps related to the submolecule's coordination number in the cluster (the number of explicit waters j hydrogen-bonding to the particular contributing submolecule). Suppose the contributions c_i are of only five types, c_j. A least-squares fitting of this five-parameter linear function was performed for each particular set of data for $n = \{1, 2, 4, 5, 8, 10\}$ or $\{1, 2, 4, 5, 8, 10, 12, 14, 16, 18\}$ when such extended data was available. The resulting contributions c_j are plotted versus j in Figure 5 for each CSM, as well as for the "true" (expt(nca)) solvation energies (dashed line). As expected, the contribution becomes less negative as the contribution to cavity size decreases (as j increases from 0 to 4). What is interesting is that all CSMs produce too steep a rise in the trend, relative to the dashed line of true values. Contributions to $\Delta_{solv}G$ are too negative for $j = 0$ and some $j = 1$ waters (only in W_1–W_3), causing negative errors in $\Delta_{solv}G$,

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

Table 1. $\Delta_{solv}G$ (kcal mol^{-1}) for Water Clusters (W_n) in Water, from Various CSMs[a] and "Experimental" Values

cluster	UFFx1.1[c]	Bondix1.7x1.1[c]	COSMO[b]	Bondix1.1[c]	SMD[c]	SM6[b]	Jaguar[b]	UFFx0.8	expt(nca)[d]	expt(nda)[d]
W	−5.0	−6.7	−6.7	−7.9	−8.5	−8.8	−8.8	−11.0	*-6.32*	*-6.32*
W_2	−8.3	−11.0	−11.6	−12.9	−13.9	−15.4	−15.5	−18.7	*-12.5*	*-12.8*
W_4	−7.8	−11.1	−12.0	−13.6	−14.6	−17.3	−18.1	−22.8	*-17.0*	*-18.3*
W_5	−9.5	−13.5	−14.5	−16.5	−17.4	−21.2	−21.3	−27.1	*-19.3*	*-20.8*
W_8	−10.0	−14.3	−15.2	−17.8	−18.7	−23.9	−21.9	−30.3	*-28.2*	*-31.9*
W_{10}	−12.9	−18.1	−19.2	−22.1	−22.9	−30.2	−29.3	−37.5	*-33.4*	*-38.0*

[a]B3LYP/6-311++G(d,p) values. UFFx1.1, Bondix1.7x1.1, Bondix1.1, and UFFx0.8 refer to choice of radii used in IEFPCM computations. [b]Bryantsev et al.[28] data. [c]Riccardi et al.[54] data. [d]nca/nda refer to nanocrystal/nanodroplet assumptions for W_n; see text.

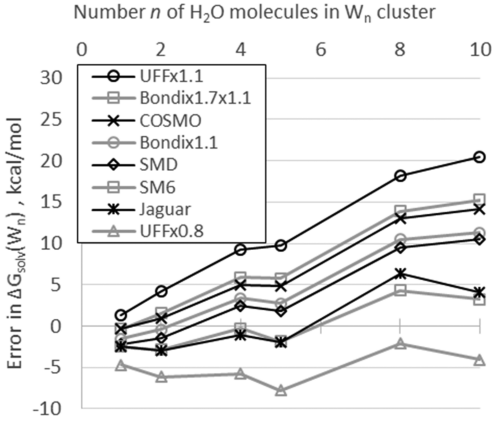


Figure 4. Errors in computed B3LYP/6-311++G(d,p) Gibbs energies of solvation of water clusters W_n (nanocrystal assumption), plotted vs cluster size n. Data from Bryantsev et al.[28] and Riccardi et al.[54] and this work (UFFx0.8). Note the rising error with cluster size.

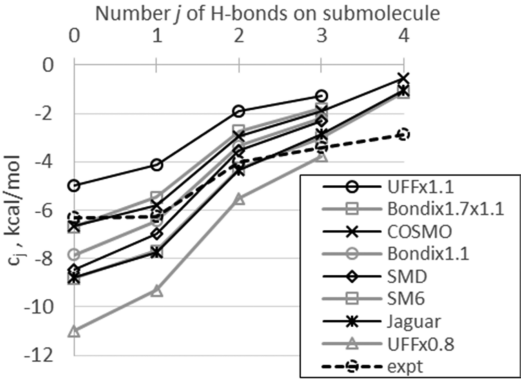


Figure 5. Submolecule contributions c_j (kcal mol^{-1}) to Gibbs energies of aqueous solvation of water clusters, plotted versus submolecule type j (0-, 1-, 2-, 3-, 4-coordinated in cluster). Most CSMs predict overly small contibutions for submolecules with $j = 3$ and $j = 4$ coordination, but overly large contributions for $j = 0$ (water monomer). Values obtained from fitting to Bryantsev et al.[28] and Riccardi et al.[54] data (which appears in Table 1 for the smaller clusters).

but too large for $j > 1$ (two- and three-coordinate H_2O, in W_4 and up), causing positive errors in $\Delta_{solv}G$. We point this out to CSM developers interested in addressing this size-dependence problem.

Returning to the issue of the growing "residual error of closing the cycle" in Figure 2, one might wonder why the residual errors are *mildest* for UFFx1.1, given that it gives the largest CSM errors for W_n solvation in Figure 4 here. This is due to the error-switch phenomenon in the previous paragraph. Most of the CSMs (except UFFx1.1) in Figure 4 overestimate the stability of water monomer ($m = 0$) in water, while underestimating the stability of clusters W_8 ($m = 12$) and larger. Thus, the growing errors in W_n vs nW accumulate in *opposite* directions, causing the particularly large residual errors seen and reported by Bryantsev et al.[28] and Riccardi et al.[54] for the cycle test in Figure 2. The UFFx1.1 results do not feature such an error-switch, and thus the growing errors in W_n vs nW accumulate in the *same* direction, producing partial error cancellation (smaller residual errors) for the cycle.

For ion solvation Gibbs energies, the cluster method of Bryantsev et al. obtains $\Delta_{solv}G(A^{q+})$ using the difference $\Delta_{solv}G(AW_n^{q+}) - \Delta_{solv}G(W_n)$, which employs two clusters with (when n is large) very similar sizes and hence similar CSM errors. This results in excellent error cancellation, and as n grows, the contributions to both AW_n and W_n should become sufficiently equal that asymptotic convergence to a steady value for the ion solvation Gibbs energy should result. In contrast, the monomer method of Bryantsev et al. uses the difference $\Delta_{solv}G(AW_n^{q+}) - n\Delta_{solv}G(W)$, with nW having larger net cavity size than AW_n^{q+}, an imbalance that causes an error imbalance that grows with n and produces divergent results, as is observed. The cluster method is the proper way to deal with these growing CSM errors. To extend the cluster method defined by Bryantsev et al. from solvation energies to reactions, one should thus employ a principle of balanced total cavity sizes. We had been doing this with our two conservation principles mentioned in the Introduction: conserving both the number of cavities (stoichiometry) and the number of H-bonds introduced. Since satisfying the two principles may not always be possible or convenient, a means for correcting the error introduced by a mismatch would be desirable. That correction is now presented.

3.3. Solvent Confinement Correction, $\Delta_{SCC}G$.

Consider the following two semicontinuum models for predicting the reaction Gibbs energy for ethene hydration:

$$C_2H_4 + H_2O{\cdot}4W \rightarrow C_2H_5OH{\cdot}3W + W \quad \text{(cavity-conserving)} \tag{6a}$$

$$C_2H_4 + H_2O{\cdot}3W \rightarrow C_2H_5OH{\cdot}3W \quad \text{(solvent-confinement-conserving)} \tag{6b}$$

The first model obeys the cavity-conserving principle (2 for reactants, 2 for products) but not the H-bond-conserving principle (4 for reactants, 3 for products). The second model does the opposite, violating cavity conservation (2 for reactants, 1 for products) to achieve H-bond conservation (3 for reactants, 3 for products). Which, then, is the more theoretically correct model?

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

The question posed is not a moot one. Note that the difference between these two models is the act of *confining* one solvent water molecule W to be attached to a cluster: $H_2O{\cdot}3W + W \rightarrow H_2O{\cdot}4W$. Under the 1 M concentration convention, ΔG for such a confinement is +3 kcal mol^{-1} (third horizontal legs of Figure 3), due to a significant entropy loss. The important consequence of this modeling act of solvent confinement upon estimates of $\Delta_r G$ for ethene hydration is that the cavity-conserving model (eq 6a) will inherently predict a reaction Gibbs energy 3 kcal mol^{-1} lower than the confinement-conserving model (eq 6b), even if the ESM and CSM are perfect. If one were using the 55.4 M concentration convention for water, the two models would be in better agreement, for then the ΔG for such a confinement is only −0.3 or −0.4 kcal mol^{-1} under the nanocrystal assumption (fourth horizontal legs of Figure 3), or 0 under the nanodroplet assumption. However, mistakes are more often made in this convention, and quantum chemistry users would find it simpler to adhere to the 1 M convention for everything they compute. Hence the 3 kcal mol^{-1} effect for a confinement of an explicit water (from hypothetical 1 M bulk) is a legitimate concern.

Here is the answer to the question posed: under the 1 M convention, eq 6b is the theoretically correct model, because it conserves the number of solvent confinements: it confines three solvating waters (denoted by W) onto both the reactants and the products, and although this artificially raises the Gibbs energy of both by 9 kcal mol^{-1}, this balance cancels away when $\Delta_r G$ is calculated. The reactant water, H_2O, is not a "confined" water; the modeling confines the three solvating W molecules to it.

To properly employ models that do not conserve the number of solvent confinements, such as the cavity conserving model 6a, one should invoke $\Delta_{SCC}G = -3n$ kcal mol^{-1}, the *solvent confinement correction* (eq 2d), for each species employing n explicit solvent waters W inside the CSM cavity. In model 6a, this applies corrections (kcal mol^{-1}) of −12 to reactants and −9 to products, raising its $\Delta_r G$ prediction by 3. This use of $\Delta_{SCC}G$ allows proper comparison of the $\Delta_r G$ of nonconserving models to experimental values generated under the 1 M convention.

To restate: the principle of *solvent confinement conservation* is the "theoretically correct" way to try to reproduce experimental $\Delta_r G$ (or $\Delta^{\ddagger}G$) values from the 1 M convention, as it provides cancellation of these $\Delta_{SCC}G$ terms to eliminate their need. When confinements are not conserved, $\Delta_{SCC}G$ should be employed.

As an aside, to support the claim $\Delta_{SCC}G = -3n$ kcal mol^{-1} in eq 2d, a short mathematical derivation is here provided, using equations shown in Figure 2. To release n confined waters under the 1 M convention, we consider the n confined waters of the water cluster W_{n+1} (call this W_k) and apply

$$\begin{aligned}\Delta_{SCC}G &= kG(W) - G(W_k)\\ &= RT\ln\left(\frac{55.34}{k}\right) + \Delta_{fus}G - kRT\ln\left(\frac{55.34}{1}\right)\\ &= -2.38(k-1) - 0.59\ln k + \Delta_{fus}G\end{aligned}$$

Inserting $\Delta_{fus}G \approx -0.3m_{\text{H-bonds}} \approx -0.3\left(\frac{2}{3}\right)k\ln k$, together with $\ln k \approx 3(k-1)/(k+3)$, gives

$$\begin{aligned}\Delta_{SCC}G &\approx -2.38(k-1) - 0.59\left(\frac{3(k-1)}{k+3}\right)\\ &\quad - 0.3\left(\frac{2}{3}\right)k\left(\frac{3(k-1)}{k+3}\right)\\ &\approx -3(k-1)\\ &\approx -3n\end{aligned}$$

4. RESULTS AND DISCUSSION II: ALKENE HYDRATION

4.1. Overall Reaction $\Delta_r G$: Ethene. The known overall reaction Gibbs energy for ethene hydration, $C_2H_4(aq) + H_2O(l) \rightarrow C_2H_5OH(aq)$, is −3.8 kcal mol^{-1} (1 M convention for all species), derived from experimentally determined components (see Supporting Information). It is straightforward to reproduce computationally to within the desired 3 kcal mol^{-1} accuracy. From eq 3 for each species, $\Delta_r G$ prediction involves:

$$\Delta_r G = \Delta_r G_{el,\varepsilon} + \Delta_r G_{freq} + \Delta_r(\Delta_{solv}G_{other}) + \Delta_r(\Delta_{conc}G) + \Delta_r(\Delta_{SCC}G) \quad (7)$$

and the $\Delta_r(\Delta_{solv}G_{other})$ term has been allowed to default (to 0 in the CPCM calculations).

Desiring an ESM (electronic structure method) less resource-consuming than CCSD(T)/aug-cc-pVTZ to compute hydrated clusters for semicontinuum calculations, we first tested several ESMs for their ability to reproduce (as closely as possible) the CCSD(T)/aug-cc-pVTZ/CPCM estimate of −13.0 kcal mol^{-1} for the $\Delta_r G_{el,\varepsilon}$ component of $\Delta_r G$. Table 2 shows that MP2 outperformed the density functional theories in this regard, and hence MP2/aug-cc-pVTZ/CPCM//M06-2X/6-31G(d)/CPCM was used for ensuing $\Delta_r G$ (eq 7) predictions.

Table 2. Results of Various ESMs for Predicting $\Delta_r G_{el,\varepsilon}$ (eq 4a, kcal mol^{-1}) for Ethene + H_2O → Ethanol

ESM	$\Delta_r G_{el,\varepsilon}$
M06-2X/6-31G(d)/CPCM[a]	−24.0
M06-2X/6-31+G(d)/CPCM[a]	−20.2
M06-2X/6-31+G(d,p)/CPCM[a]	−16.8
M06-2X/6-311++G(d,p)/CPCM[a]	−15.3
M06-2X/aug-cc-pVTZ/CPCM[a]	−15.0
BLYP/aug-cc-pVTZ/CPCM[a]	−7.4
B3LYP/aug-cc-pVTZ/CPCM[a]	−10.6
X3LYP/aug-cc-pVTZ/CPCM[a]	−11.5
ωB97x-D/aug-cc-pVTZ/CPCM[a]	−14.8
MP2/aug-cc-pVTZ/CPCM[a]	−13.8
MP4/aug-cc-pVTZ/CPCM[a]	−13.5
CCSD(T)/aug-cc-pVTZ/CPCM[a]	−13.0
CCSD(T)/aug-cc-pVTZ/CPCM[b]	−13.0

[a]Single-point energy using geometries optimized with M06-2X/6-31G(d)/CPCM. [b]Single-point energy using geometries optimized with M06-2X/6-31+G(d)/CPCM.

Next, calculations were performed for 20 different semicontinuum choices, including the traditional model with no explicit solvating waters W. The hydrated clusters employed are shown in Figure 6. As mentioned in the Methods section, our experience with testing variations in explicit water placement or orientation has shown that such variations have negligible effects, smaller than the CSM solvation-energy errors. What is important is the avoidance of structures with spurious H-bonds

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

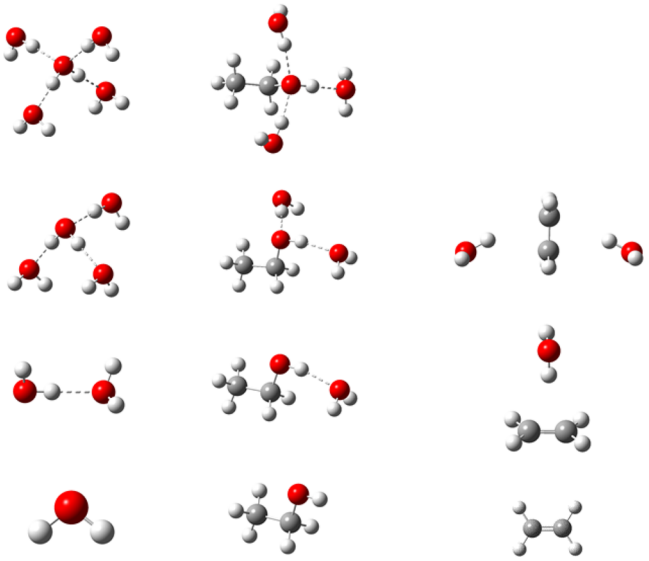


Figure 6. Optimized structures used for Table 3.

between solvent molecules, to appropriately balance water (H-bond) confinements. Table 3 shows the results.

We first comment on the first two rows in Table 3, the results without explicit solvating water molecules, differing only in CSM used (SMD vs CPCM). Both predictions satisfy the goal of 3 kcal mol^{-1} accuracy. The CPCM Δ_rG error is $-2.4 - (-3.8) = +1.4$, due principally to the sum of individual CSM errors, $+2 - (+2 - 2) = +2$, with minor errors from Δ_rG_{freq} and ESM. (These individual CSM errors were already determined for IEFPCM, Figure 1, and CPCM and IEFPCM errors are nearly identical.) The SMD Δ_rG error is $-(2.6) - (-3.8) = +1.2$, also due principally to the sum of individual CSM errors, $0 - (-2 - 0) = +2$.

The remainder of Table 3 tests semicontinuum models. Traditional results without a solvent confinement correction (penultimate column of Table 3) show values of 0 to −6 kcal mol^{-1} (and thus errors of −2 to +4 kcal mol^{-1}). Use of the new solvent-confinement correction $\Delta_r(\Delta_{SCC}G)$ dramatically tidies up the errors, grouping the results into three sets: the first and third sets make fine predictions of -2.4 ± 0.5 kcal mol^{-1}, while the second set predicts slightly worse: -0.4 ± 0.8 kcal mol^{-1}. Other than solvent confinement errors, if we suppose a −0.8 kcal mol^{-1} error from MP2 (vs CCSD(T), Table 2), and assign all remaining error to MP2 CSM solvation errors, one could derive every individual CSM error: {1.9, 5.1, 7.0, 9.2) for {EtOH, EtOH·W, EtOH·2W, EtOH·3W}, {1.9, 3.7, 6.9, 9.2} for {W, W_2, W_4, W_5}, and {−2.3, −0.9, 3.1} for {C_2H_4, C_2H_4·W, C_2H_4·2W}. Such derived errors reveal the systematic growth of CSM's imperfect-cavity-size error: the rise is roughly 1.9 kcal mol^{-1} upon water addition to any of these solutes, matching our expectation of $2n$ kcal mol^{-1} (Figure 4), but with two exceptions: larger error jumps caused by the first water on ethanol and the second water on ethene. If this error of 1.9 kcal mol^{-1} per added water were consistent right from the first water added, then this accumulating error would beautifully cancel when adding waters W to both sides of a reaction model, resulting in a constant prediction at all levels of hydration. The reason that the middle 12 models in the table produce worse results (errors of +2.7 to +4.2 kcal mol^{-1}) is due to excessive

Table 3. Δ_rG and Its Components (kcal mol^{-1}) for the Overall Reaction $C_2H_4(aq) + H_2O(liq, 1\ M) \rightarrow C_2H_5OH(aq)$, from MP2/aug-cc-pVTZ/CPCM//M06-2X/6-31G(d)/CPCM Semicontinuum Calculations

model	$\Delta_rG_{el,\varepsilon}$	Δ_rG_{freq}	$\Delta_r(\Delta_{conc}G)$	$\Delta_r(\Delta_{SCC}G)$	Δ_rG^a	Δ_rG^b
$C_2H_4 + H_2O \rightarrow C_2H_5OH$[c]	−14.0	13.2	−1.9	0	−2.6	−2.6
$C_2H_4 + H_2O \rightarrow C_2H_5OH$	−13.8	13.2	−1.9	0	−2.4	−2.4
$C_2H_4 + H_2O$·**W** $\rightarrow C_2H_5OH$ + **W**	−9.7	4.5	0	3.0	−5.3	−2.3
$C_2H_4 + H_2O$·**4W** $\rightarrow C_2H_5OH + W_4$	−8.5	2.7	0	3.0	−5.9	−2.9
C_2H_4·**W** $\rightarrow C_2H_5OH$	−11.3	6.4	0	3.0	−4.9	−1.9
$C_2H_4 + H_2O$·**W** $\rightarrow C_2H_5OH$·**W**	−13.7	14.6	−1.9	0	−1.0	−1.0
$C_2H_4 + H_2O$·**3W** $\rightarrow C_2H_5OH$·**3W**	−14.9	16.7	−1.9	0	0.0	0.0
C_2H_4·**W** $+ H_2O \rightarrow C_2H_5OH$·**W**	−15.3	16.5	−1.9	0	−0.6	−0.6
C_2H_4·**W** $+ H_2O$·**W** $\rightarrow C_2H_5OH$·**2W**	−16.5	17.9	−1.9	0	−0.6	−0.6
$C_2H_4 + H_2O$·**3W** $\rightarrow C_2H_5OH$·**W** $+ W_2$	−8.9	5.4	0	3.0	−3.5	−0.5
$C_2H_4 + H_2O$·**3W** $\rightarrow C_2H_5OH$·**2W** + **W**	−10.1	6.7	0	3.0	−3.4	−0.4
$C_2H_4 + H_2O$·**4W** $\rightarrow C_2H_5OH$·**2W** $+ W_2$	−8.9	4.9	0	3.0	−4.0	−1.0
$C_2H_4 + H_2O$·**4W** $\rightarrow C_2H_5OH$·**3W** + **W**	−9.7	6.2	0	3.0	−3.5	−0.5
C_2H_4·**W** $+ H_2O$·**3W** $\rightarrow C_2H_5OH$·**2W** $+ W_2$	−11.7	8.6	0	3.0	−3.1	−0.1
C_2H_4·**W** $+ H_2O$·**3W** $\rightarrow C_2H_5OH$·**3W** + **W**	−12.4	9.9	0	3.0	−2.6	0.4
C_2H_4·**W** $+ H_2O$·**4W** $\rightarrow C_2H_5OH$·**3W** $+ W_2$	−11.2	8.1	0	3.0	−3.2	−0.2
C_2H_4·**W** $+ H_2O$·**4W** $\rightarrow C_2H_5OH$·**W** $+ W_4$	−10.1	6.0	0	3.0	−4.1	−1.1
C_2H_4·**2W** $+ H_2O \rightarrow C_2H_5OH$·**2W**	−19.1	18.2	−1.9	0	−2.8	−2.8
C_2H_4·**2W** $+ H_2O$·**W** $\rightarrow C_2H_5OH$·**3W**	−19.9	19.5	−1.9	0	−2.3	−2.3
C_2H_4·**2W** $+ H_2O$·**W** $\rightarrow C_2H_5OH$·**2W** + **W**	−15.1	9.5	0	3.0	−5.6	−2.6
C_2H_4·**2W** $\rightarrow C_2H_5OH$·**W**	−13.9	8.2	0	3.0	−5.7	−2.7
experiment (see Introduction)						−3.8

[a]Sum of first three components, i.e., omitting $\Delta_r(\Delta_{SCC}G)$. [b]Sum including $\Delta_r(\Delta_{SCC}G)$. [c]MP2/aug-cc-pVTZ/SMD rather than MP2/aug-cc-pVTZ/CPCM.

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

error from the first water on ethanol. However, the third set of models returns to good values (errors <2 kcal mol^{-1}) because of the second water on ethene, which brings with it a large error jump to counter the large error jump of the first water on ethanol.

To summarize:

(i) There is no need to use semicontinuum models for this reaction involving neutral species. The predictions without any explicit waters added (top 2 rows of Table 3) achieve 2 kcal mol^{-1} accuracy, using CPCM or SMD.

(ii) Although we now know (section 3) that the addition of explicit waters W increases (accumulates) CSM error on both sides of the reaction, it cancels very well.

(iii) Use of the new solvent confinement correction brings *all* Set 1 and Set 3 semicontinuum models into agreement, still producing $\Delta_r G = -2.4 \pm 0.5$. Only the second set of models produces worse values, $\Delta_r G = -0.4 \pm 0.8$, due to the excessive CSM error of the first W on ethanol that is not counterbalanced until both sides of ethene are explicitly hydrated (C_2H_4·2W).

4.2. Activation $\Delta^{\ddagger}G$: Ethene. This activation involves conversion of H_3O^+ (and alkene) into a large cationic transition state, and thus semicontinuum extrapolation is expected to be needed to remove the high q/r error made by CSMs on H_3O^+ (see Introduction).

For ethene hydration, the activation Gibbs energy, $\Delta^{\ddagger}G$, obtained from the Eyring-equation $\Delta^{\ddagger}H$ and $\Delta^{\ddagger}S$ values experimentally determined by Baliga and Whalley,[32] is 34.1 kcal mol^{-1}. Their use of the Eyring equation assumes a 1 M concentration of all species, and to determine the nature of the activated complex, Baliga and Whalley determined not only activation entropies but also activation volumes. Both were sufficiently negative to draw their conclusion that the activated complex consisted of C_2H_4, H^+, and at least one water molecule. They pointed out the possibility that more than one water might be involved in the activated complex, but that contemporary methods were unable to discriminate further. We shall make the assumption that the importance of any additional water molecule (beyond 1) in the activated complex is at the level of importance of a water molecule in the first solvation shell of the hydronium reactant, so that the negative $\Delta^{\ddagger}S$ and $\Delta^{\ddagger}V$ valued determined experimentally would still correspond to a stoichiometric change of −1 in going from reactants to transition state, as they preferred to represent it.

Therefore, of the following four series of aqueous semicontinuum models,

$$C_2H_4 + H_3O^+\cdot nW + H_2O \rightarrow [C_2H_5OH_2{}^+\cdot(n+1)W]^{\ddagger} \quad \text{(Series I)}$$

$$C_2H_4 + H_3O^+\cdot nW \rightarrow [C_2H_5OH_2{}^+\cdot nW]^{\ddagger} \quad \text{(Series II)}$$

$$C_2H_4 + H_3O^+\cdot nW \rightarrow [C_2H_5OH_2{}^+\cdot(n-1)W]^{\ddagger} + H_2O \quad \text{(Series III)}$$

$$C_2H_4 + H_3O^+\cdot nW \rightarrow [C_2H_5OH_2{}^+\cdot(n-2)W]^{\ddagger} + 2H_2O \quad \text{(Series IV)}$$

it is Series II that should be considered to have the right stoichiometry and thus be the series of models that adheres to the new principle of solvent confinement conservation. Models form the other series would thus need to invoke the solvent confinement correction, $\Delta_{SCC}G$, to allow "theoretically correct" comparison to the experimental values from the Eyring-equation 1 M convention: $\Delta^{\ddagger}(\Delta_{SCC}G) = \{-3, 0, +3, +6\}$ for {I, II, III, IV} respectively.

Analogous to eq 7 for reaction Gibbs energy, for activation Gibbs energy one obtains

$$\Delta^{\ddagger}G = \Delta^{\ddagger}G_{el,\varepsilon} + \Delta^{\ddagger}G_{freq} + \Delta^{\ddagger}(\Delta_{solv}G_{other}) + \Delta^{\ddagger}(\Delta_{conc}G) + \Delta^{\ddagger}(\Delta_{SCC}G)) \quad (8)$$

and we again begin by allowing $\Delta^{\ddagger}(\Delta_{solv}G_{other})$ to default. The preliminary ESM (electronic structure method) search, for an accurate but inexpensive method for the semicontinuum modeling, tested the same Table 2 ESMs but now for their ability to reproduce the CCSD(T)/aug-cc-pVTZ/CPCM//M06-2X/6-31+G(d)/CPCM estimate of $\Delta^{\ddagger}G_{el,\varepsilon}$ = 3.3 kcal mol^{-1} for the elementary model $C_2H_4 + H_3O^+ + H_2O \rightarrow [H_2O\cdot H^+\cdot C_2H_4\cdot OH_2]^{\ddagger}$. This time M06-2X did as well as MP2 (Table 4), and hence M06-2X/aug-cc-pVTZ/CPCM//M06-2X/6-31G(d)/CPCM single points were employed for ensuing $\Delta^{\ddagger}G$ (eq 8) predictions.

Table 4. Results of Various ESMs for Predicting $\Delta^{\ddagger}G_{el,\varepsilon}$ (Eq 4a, kcal mol^{-1}) for $C_2H_4 + H_3O^+ + H_2O \rightarrow [H_2O\cdot H^+\cdot C_2H_4\cdot OH_2]^{\ddagger}$

ESM*	$\Delta^{\ddagger}G_{el,\varepsilon}$
M06-2X/6-31G(d)/CPCM[a]	−2.9
M06-2X/6-31+G(d)/CPCM[a]	0.4
M06-2X/6-31+G(d,p)/CPCM[a]	1.7
M06-2X/6-311++G(d,p)/CPCM[a]	2.0
M06-2X/aug-cc-pVTZ/CPCM[a]	3.0
BLYP/aug-cc-pVTZ/CPCM[a]	1.1
B3LYP/aug-cc-pVTZ/CPCM[a]	2.7
X3LYP/aug-cc-pVTZ/CPCM[a]	1.9
ωB97x-D/aug-cc-pVTZ/CPCM[a]	1.5
MP2/aug-cc-pVTZ/CPCM[a]	3.9
MP4/aug-cc-pVTZ/CPCM[a]	3.6
CCSD(T)/aug-cc-pVTZ/CPCM[a]	4.5
CCSD(T)/aug-cc-pVTZ/CPCM[b]	3.3

[a]Single-point energy using geometries optimized with M06-2X/6-31G(d)/CPCM. [b]Single-point energy using geometries optimized with M06-2X/6-31+G(d)/CPCM.

Next, calculations were performed for 19 different semicontinuum choices, categorized by the four series above. The hydrated ion clusters employed are shown in Figure 7, again avoiding structures with spurious H-bonds between solvent molecules to appropriately balance solvent water (H-bond) confinements. Results (including $\Delta^{\ddagger}(\Delta_{SCC}G)$ where appropriate) are plotted in Figure 8.

From the CPCM data, all four series are seen to extrapolate to an asymptote of roughly 40 kcal mol^{-1}, in error by +6 kcal mol^{-1} (outside the desired goal of 3 kcal mol^{-1} accuracy). This surprised us, since the same extrapolation for another hydronium-containing process ($\Delta_r G$ of the water autoionization reaction) resulted in only ∼1 kcal mol^{-1} of error.[9] We repeated the calculations with SMD (with the caveat that SMD was not meant to be used with $n > 1$ waters on ions), and again each series of models extrapolates to asymptotic values, but now these values depend upon the series. Series II, the theoretically most

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

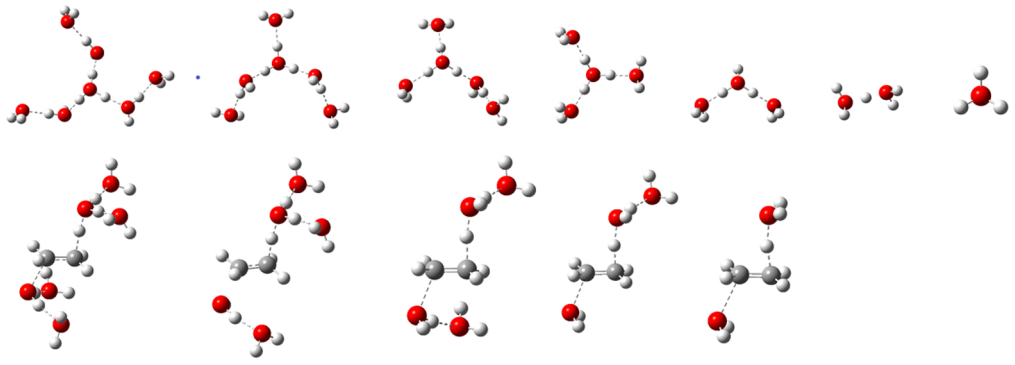


Figure 7. Optimized structures of $H_3O^+ \cdot nW$ (top) and transition states $[C_2H_5OH_2^+ \cdot nW]^‡$ (bottom), used for Figure 8.

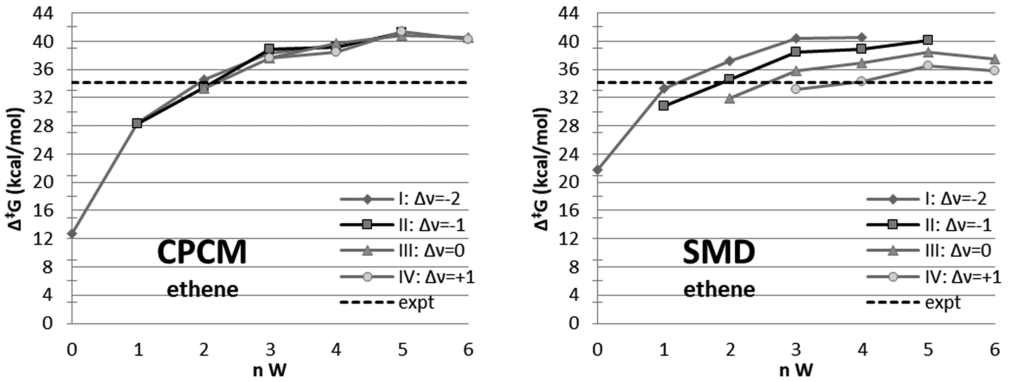


Figure 8. Computed $\Delta^‡G$ values for acid-catalyzed ethane hydration, M06-2X/aug-cc-pVTZ/{CPCM or SMD} using geometries from M06-2X/6-31G(d)/CPCM, plotted against number of explicit waters added. Results include $\Delta^‡(\Delta_{conc}G)$ and $\Delta^‡(\Delta_{SCC}G)$. Dashed line: 34.1 kcal mol^{-1}, the experimental value.[32] Dark line with squares: Series II ($\Delta\nu = -1$), the series with balanced solvent confinements (see text).

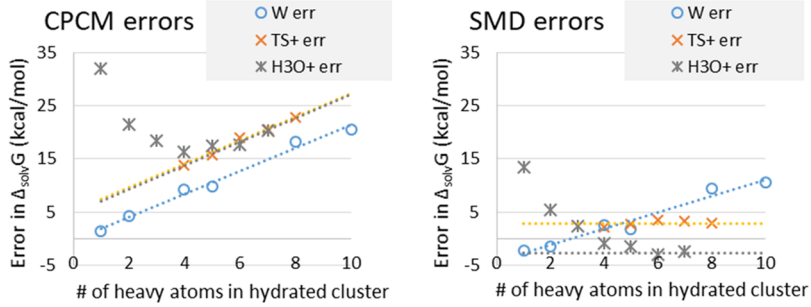


Figure 9. Errors in computed $\Delta_{solv}G$ values for explicitly hydrated clusters. Circles: water clusters W_n. X's: transition state complexes $[C_2H_5OH_2^+ \cdot nW]^‡$. Stars: $H_3O^+ \cdot nW$, whose data reveal the high q/r error (deviation from dashed line) at small cluster sizes.

savory one with its balanced solvent confinements, has nearly the same asymptote (40 kcal mol^{-1}) with either SMD or CPCM.

Hence the first two conclusions to be made from Figure 8 regarding $\Delta^‡G$ predictions here are the following: (i) As expected, neither $n = 0$ nor $n = 1$ modeling is sufficient to achieve the goal of 3 kcal mol^{-1} agreement with Eyring-equation $\Delta^‡G$ from experiment, with either of these popular CSMs. (ii) Surprisingly, the accuracy goal *also cannot be achieved* with semicontinuum extrapolation of the most savory Series II models, which should have eliminated the high q/r error of CSMs. Both CPCM and SMD leave a roughly +6 kcal mol^{-1} error upon Series II extrapolation.

4.3. Activation $\Delta^‡G$: Error Management Options. We next provide an error analysis that will lead to better error management, which we believe will achieve the accuracy goal more often, and hence be methods of first choice in acid-catalyzed activation energy prediction. The errors in the Figure 8 $\Delta^‡G$ predictions were probed by deriving the CSM errors made on each hydrated cation cluster. This can be done using the known CSM solvation errors for C_2H_4, H_3O^+, and $H_3O^+ \cdot H_2O$ (Figure 1, taking the Marenich et al.[8] estimate of $\Delta_{solv}G = -87.8$ kcal mol^{-1} for $H_3O^+ \cdot H_2O$), and assuming that all 19 semicontinuum models were attempts to reproduce the expt $\Delta^‡G$ = 34.1 kcal mol^{-1}. These derived CSM errors, as well as those for pure water clusters from section 3.2, are plotted in Figure 9. We interpret these CSM errors as sums of two contributions mentioned in the Introduction: (i) the imperfect cavity size error ε_{ics}, which was expected to rise with increasing hydration (Figure 4), indeed does in Figure 9 with CPCM, but with CSM = SMD this rise is seen only when hydrating a water molecule, not when hydrating the cations; (ii) the high q/r error ε_{hqr}, which is expected to decrease with increasing explicit hydration, indeed does in Figure 9 in the hydronium cluster data but is too small to be seen in the transition state data.

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

These error components can be fit with the phenomenological functions:

$$\varepsilon_{ics} = a + bh \tag{9}$$

$$\varepsilon_{hqr} = ce^{-dh} \tag{10}$$

where h is the number of heavy atoms in the cluster: $h = n$ for W_n, $n + 1$ for $H_3O^+{\cdot}nW$, $n + 3$ for $[C_2H_5OH_2^+{\cdot}nW]^{\ddagger}$. Together with the parameter values in Table 5, they reproduce the error data in Figure 9 very well. These functions can then be used to explain all the error points for the activation Gibbs energy of ethene hydration (Figure 8) as well. For example, the error $\varepsilon(n)$ of a Series II prediction is

Table 5. Parameter Values for the Phenomenological Error Functions $\varepsilon = a + bh + ce^{-dh}$ That Reproduce the CSM Solvation Errors of the Hydrated Ions in Figure 9

parameter	H_3O^+ CPCM	TS^+ CPCM	H_3O^+ SMD	TS^+ SMD
a	6.4	6.4	−2.7	2.8
b	2	2	0	0
c	50	10	30	6
d	0.75	0.75	0.64	0.64

$$\begin{aligned}\varepsilon(n) &= \varepsilon(TS^+.nW) - \varepsilon(H_3O^+.nW) - \varepsilon(C_2H_4)\\ &= \{a_{TS} + b(n+3) + c_{TS}e^{-d(n+3)}\}\\ &\quad - \{a_{H_3O} + b(n+1) + c_{H_3O}e^{-d(n+1)}\} - \varepsilon(C_2H_4)\end{aligned} \tag{11}$$

$$\begin{aligned}\lim_{n\to\infty} \varepsilon &= \{a_{TS} - a_{H_3O}\} + 2b - \varepsilon(C_2H_4)\\ &= \{6.4 - 6.4\} + 2(2) - (-2) = +6\\ &\quad \text{(CPCM data)}\\ &= \{2.8 - (-2.7)\} + 2(0) - (0) = +5.5\\ &\quad \text{(SMD data)}\end{aligned} \tag{12}$$

For a Series III prediction the error is

$$\begin{aligned}\varepsilon(n) &= \varepsilon(TS^+.(n-1)W) + \varepsilon(H_2O) - \varepsilon(H_3O^+.nW)\\ &\quad - \varepsilon(C_2H_4)\\ &= \{a_{TS} + b(n+2) + c_{TS}e^{-d(n+2)}\} + \varepsilon(H_2O)\\ &\quad - \{a_{H_3O} + b(n+1) + c_{H_3O}e^{-d(n+1)}\} - \varepsilon(C_2H_4)\end{aligned} \tag{13}$$

$$\begin{aligned}\lim_{n\to\infty} \varepsilon &= \{a_{TS} - a_{H_3O}\} + b + \varepsilon(H_2O) - \varepsilon(C_2H_4)\\ &= \{6.4 - 6.4\} + (2) + (2) - (-2) = +6\\ &\quad \text{(CPCM data)}\\ &= \{2.8 - (-2.7)\} + (0) + (-2) - (0) = +3.5\\ &\quad \text{(SMD data)}\end{aligned} \tag{14}$$

To probe ideas for better error management, we first address what "went wrong" with Series II semicontinuum extrapolation (eq 12) to large n. As eqs 12 and 14 show, extrapolated series results will eliminate the high q/r error ce^{-dh} (which is really the point of semicontinuum modeling), and while these equations show that there can still be a remaining imperfect-cavity-size error, eqs 12 and 14 show that some beneficial error cancellation should have been possible, which we had hoped would produce $|\varepsilon| < 3$ kcal mol^{-1}. The Series II extrapolation with CPCM (+6 error) "went wrong" because of the negative error CPCM makes on nonpolar hydrocarbons in water ($\varepsilon(C_2H_4) = -2$ kcal mol^{-1}). Optimistically, if CPCM generally gives this negative error for reactant hydrocarbon segments (see Figure 1), this knowledge might afford better error management. The Series II extrapolation with SMD (+5.5 error) "went wrong" because of the poor cancellation of the a parameters of the two cations, $a_{TS} - a_{H_3O}$. Optimistically, if SMD were to be consistent with this particular discrepancy for acid catalysis, this knowledge might also afford better error management.

We now assemble options for better error management for these acid-catalysis $\Delta^{\ddagger}G$ predictions.

Options A1−A3, the least complicated, avoid modification of $\Delta_{solv}G_{other}$ beyond default settings, by relying on semicontinuum models that happen to provide better CSM cancellation and have a fathomable chance of reliably doing so. Examination of Figure 8 offers several possibilities:

Option A1: Series II, $n = 2$ ($H_3O^+{\cdot}2W$), either CPCM or SMD
Option A2: Series I, $n = 1$ ($H_3O^+{\cdot}W$), with SMD
Option A3: Series IV, $n = 3$ ($H_3O^+{\cdot}3W$), with SMD

Options A1 and A2 keep some high-q/r error (ε_{hqr}) in the H_3O^+ calculation to help offset imperfect cavity size errors (ε_{ics}). Option A3, by having two product water molecules (Series IV), adding a counterbalancing −4 kcal mol^{-1} of SMD ICS error to the products to offset the ICS error caused by SMD's $a_{TS} - a_{H_3O}$ mismatch. Option A1, using the Series II $n = 2$ semicontinuum model:

$$A + H_3O^+{\cdot}2W \rightarrow [AHOH_2{}^+{\cdot}2W]^{\ddagger} \tag{15}$$

is the simplest of the three, due to its conservation of solvent confinements (Series II modeling) which avoids the need for $\Delta_{SCC}G$ as well as $\Delta_{solv}G_{other}$ alteration in eq 8. Figure 8 shows that this model achieves 1 kcal mol^{-1} agreement with experiment using either CPCM or SMD.

Some, however, might consider the error-cancellation strategies of Options A1−A3 as "getting the right answer for the wrong reason." Option A1, for instance, deliberately keeps quite a bit of HQR error in the H_3O^+ calculation to offset a completely different kind of CSM error (ICS). Options B1 and B2 recognize the ICS CSM errors by adding simple *post facto* empirical corrections for them, into $\Delta^{\ddagger}(\Delta_{solv}G_{other})$:

Option B1: Series II, $n = 3$ ($H_3O^+{\cdot}3W$), CPCM or SMD, and add −4 kcal mol^{-1} to $\Delta^{\ddagger}(\Delta_{solv}G_{other})$
Option B2: Series II, $n = 5$ ($H_3O^+{\cdot}5W$), CPCM or SMD, and add −6 kcal mol^{-1} to $\Delta^{\ddagger}(\Delta_{solv}G_{other})$

Both options use Series II to avoid the need for $\Delta_{SCC}G$. Option B uses easy-to-remember "Eigen ion" modeling ($n = 3$, $H_3O^+{\cdot}3W$) and a milder correction. Option B2 eliminates HQR error on H_3O^+ via semicontinuum extrapolation, leaving only the ICS errors of the CSMs, which for $\Delta^{\ddagger}G$ require a −6 kcal mol^{-1} correction.

Some, however, may also object to correcting CSM errors with *post facto* "fudge factor" corrections to $\Delta^{\ddagger}G$. An Option C would be to improve the CSM accuracy on the ions by invoking better $\Delta_{solv}G_{other}$ corrections for each ion, by writing $\Delta_{solv}G_{other}$ = default + HQRC + ICSC, where HQRC = $-\varepsilon_{hqr} = -ce^{-dh}$ and

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

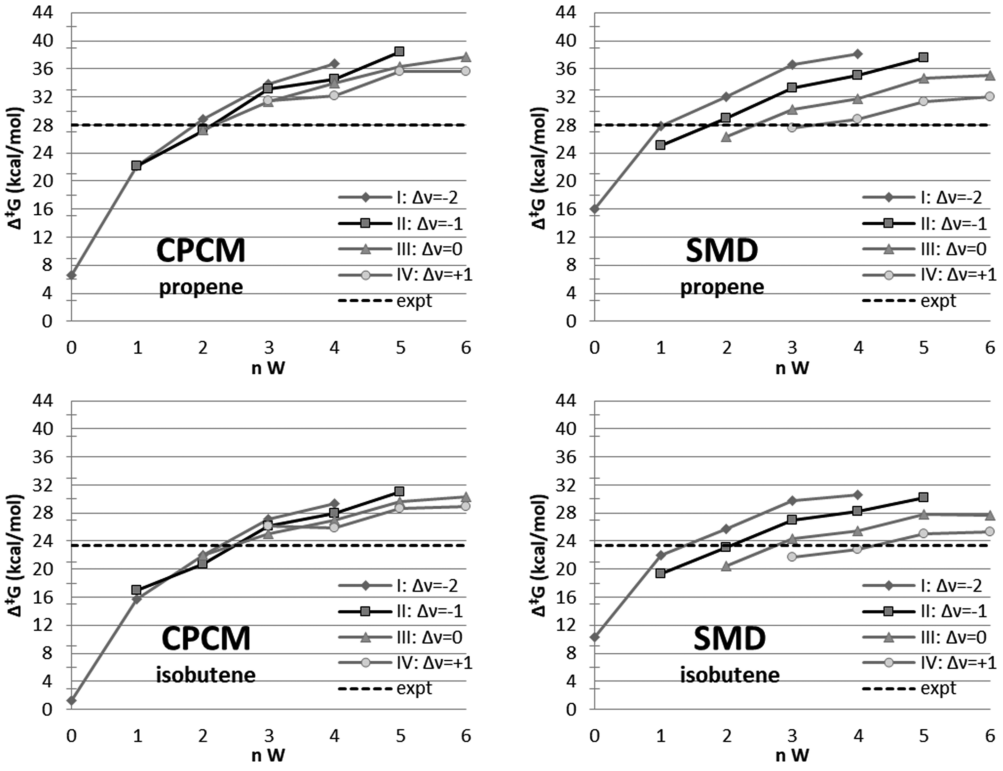


Figure 10. Computed $\Delta^{\ddagger}G$ values for acid-catalyzed hydration of higher alkenes. See Figure 8 for computational details. Dashed lines: experimental values[22] of 28.0 and 23.4 kcal mol^{-1}.

ICSC = $-\varepsilon_{ics} = -a - bh$, and using the parameter values of Table 5. This would eliminate CSM errors on the ions, leaving only CSM errors on C_2H_4 and reactant H_2O of ±2 kcal mol^{-1}. All 19 semicontinuum models would then achieve the goal of 3 kcal mol^{-1} agreement with experiment (except Series IV with SMD which involved 2 H_2O). The drawback we see with this option for general acid catalysis is that it may be very difficult to guess the *abcd* parameter values appropriate for the particular transition state of interest. However, the philosophy of Option C is, we think, important enough to stress. If a CSM method like SMD already contains $\Delta_{solv}G_{other}$ terms internally, why not ask CSM developers to consider adding to it these HQRC and ICSC terms, after further development, and thereby improve CSM predictions for ions?

Until CSMs improve for ions, we recommend Options A1 and B1. Option A1, the simplest option, happens to work for acid-catalyzed ethene hydration $\Delta^{\ddagger}G$ prediction because it happens to provide the appropriate cancellation of CSM errors in the difference $\Delta^{\ddagger}G = G(TS) - G(reactants)$. Option B1 uses improved and easy-to-remember Eigen-ion modeling for H_3O^+ and a simple *post facto* lowering of 4 kcal mol^{-1} to cure the remaining CSM error discrepancies. The question is, will these protocols generally work for all acid-catalysis activations? The fair response is probably not. However, they do work for the hydrations of propene and isobutene (next section), which we think justifies these options as solid first choices when using CPCM or SMD for acid-catalysis activation energies.

4.4. Activation $\Delta^{\ddagger}G$: Other Alkenes. We applied the same 19 semicontinuum models of section 4.2 (ethene activation $\Delta^{\ddagger}G$) to the prediction of $\Delta^{\ddagger}G$ for propene and isobutene activation as well (systems also studied experimentally by Baliga and Whalley[32]). The results appear in Figure 10. Note that the discussed Options of section 4.3 continue to work well, with Options A1 (Series II, $n = 2$) and B1 (Series II, $n = 3$, post facto correction of −4 kcal mol^{-1}) still achieving 3 kcal mol^{-1} accuracy.

4.5. Entropy: $\Delta_r S$ and $\Delta^{\ddagger}S$. There is a long history of study of solvation entropy theory, with particular discussion of solute vs solvent apportionment. Table 6 shows an exemplary (but not exhaustive) list of notable examples of entropy component considerations.

This history has several interesting features. One is the long-time concern of solvent structure reorganization ("struc" or SSR) effects when solvating ions, often (but not always) absent in the solvation entropy apportioning of neutral species. Hydrophobic effects dictate that it should be a term present for neutrals as well as ions. A second interesting feature is the ambiguity of apportioning the entropy loss to solute vs solvent: note Pierotti's scaled-particle theory attributed the entire entropy loss (other than the cratic[69] concentration adjustment) to solvent (primarily cavitation),[59] while Wertz attributed it all to solute (primarily free volume restriction).[60] Henchman points out that the ambiguity can be linked to an arbitrary choice of reference frame: he states that use of the molecule frame (MF) puts no restriction on the solute but requires cavitation of the solvent, while use of the system frame (SF) forces consideration of local hindrance of all molecules and should not require a cavitation term.[64] We also mention in passing that we,[70] like De-Cai Fang,[71−73] have in the past switched from cavitation to the damping of solute entropy (a system frame idea) to improve default free-energy predictions in organic solvents.

To predict specifically the entropy, suppose one considers the entropy S of a solution as the temperature derivative of eq 3, generating the terms

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

Table 6. Historical Partitioning of Solvation Entropy

reference	expression[a]
Eley and Evans (1938), atomic ions[56]	$S = S_t' + \Delta S_{strucI} + \Delta S_B$
Krestov (1962), ions[57]	$\Delta_{solv}S = \Delta S_t + \Delta S_r + \Delta S_{struc}$
Bockris and Saluja (1972), atomic ions[58]	$S = S_t' + \Delta S_{strucI} + \Delta S_{strucII} + \Delta S_B$
Pierotti (1976), neutrals[59]	$\Delta_{solv}S = S_{cav} + \Delta S_{conc} + \alpha RT$ (α = thermal expansion coefficient of solvent)
Wertz (1980), neutrals[60]	$\Delta_{solv}S = 0.46\ (S_{gas} + \Delta S_{conc})$ ($\Delta S_{conc} = -14.3$ eu)
Abraham (1982), atomic ions[61]	$\Delta_{solv}S = \Delta_{solv}S_{ng} + \Delta S_{B,I} + \Delta S_{B,II} + \Delta S_{B,III}$
Marcus (1986), ions[62]	$\Delta_{solv}S = \Delta S_{conc} + \Delta S_{strucI} + \Delta S_{strucII} + \Delta S_B$
Marcus (1994), ions[63]	$\Delta_{solv}S = \Delta_{solv}S_{ng} + \Delta S_{struc} + \Delta S_{B,I} + \Delta S_{B,III}$
Henchman (2010/11), atomic ions and neutrals[64,65]	$\Delta_{solv}S = \Delta S_t + \Delta S_{struc,v} + \Delta S_{struc,lib} + \Delta S_{struc,or}$
Garza (2019), neutrals[66]	$S = S_v^{gas} + S_t' + S_r' + S_{cav}$

[a]Subscripts: t = translation, r = rotation, v = vibration, cav = solvent cavitation, lib = libration, or = orientation, struc = solvent structure, ng = noble gas, B = "Born charging" term[67,68] with permittivity derivatives $d\varepsilon/dT$, conc = standard-state concentration adjustment. A prime (′) indicates a damped gas-phase term. I, II, and III indicate first and second solvent shells and bulk-solvent region, respectively.

$$S = S_\varepsilon + S_{freq} + \Delta_{solv}S_{other} + \Delta_{conc}S + \Delta_{SCC}S \quad (16)$$

Here S_{freq} is the gas-phase entropy computed from a vibrational frequency run, $\Delta_{conc}S = -6.4$ eu (cal mol^{-1} K^{-1}) from eq 2b, and $\Delta_{SCC}S = +10n$ eu from eq 2d. Unfortunately, the terms S_ε and $\Delta_{solv}S_{other}$ (particularly ΔS_{cav}, in the molecule frame convention) require subroutines for temperature derivatives not available to Gaussian users. In addition, we wondered if it was safe to neglect SSR effects, since we are dealing with a mix of ions and neutrals and an ordered solvent (acid-catalyzed aqueous reactions).

To obtain a rough gauge of how large SSR entropy might be, we used an idea used occasionally for ions: take experimental solvation entropies and remove some intrinsic amount. Abraham[61] and Marcus[63] removed the entropy of solvation of noble gases, though this may already contain some SSR entropy. Instead, we chose to subtract off 21 eu (cal mol^{-1} K^{-1}), the Trouton constant,[74] believing that this entropy of condensation represents the same intrinsic entropy loss as the act of solvation (see Supporting Information for this justification). Results of this subtraction are presented in Table 7. Indeed, SSR entropies can be significant. (Interestingly, they are negative in all the aqueous cases, suggesting that perhaps all solutes are structure-makers in water, a point already noted by Henchman for ions.[65]) Unfortunately, such terms are very difficult to predict ab initio. (However, we did note that the SSR effects cancel well for our reaction of interest, $C_2H_4 + H_2O \rightarrow C_2H_5OH$: $-18.5 + 10 + 7 = -1.5$ eu).

Given the lack of algorithms for Gaussian users to compute ΔS_{cav} and ΔS_{SSR}, and the possibility of switching from the molecule-frame idea of cavitation to the system-frame idea of solute hindrance, we chose here merely to take the Trouton constant for the entropy of solvation and apply the semicontinuum solvent-confinement correction $\Delta_{SCC}S = +10n$ eu for any confined waters:

$$S = S_{freq}(1\ \text{atm}) + \Delta_{solv}S_{Trouton}(1\ \text{atm} \rightarrow 1\ \text{M}) + \Delta_{SCC}S \quad (17)$$

The S_{freq} values were taken from our geometry-optimization level of theory (M06-2X/6-31G(d)/CPCM); no differences are expected with SMD. Although the Trouton constant is −21 eu for the condensation standard-state convention (1 atm → 55.4 M for water), it is only −13 eu for the 1 atm → 1 M solution convention (incorporating the 1 atm → 1 M concentration adjustment as well: see Supporting Information).

The results, using the 8 "preferred" semicontinuum models from section 4.1 and 7 exemplary ones from section 4.2, appear in Table 8. For Δ_rS_{total} (the reaction entropy), predictions are very good (errors < 5 eu) for models avoiding explicit waters on ethene, getting worse for $C_2H_4{\cdot}n$W cases due to artificially high entropy predictions for weakly bound waters from the harmonic-oscillator assumption.[78] As with Δ_rG_{total}, for Δ_rS_{total} there is no need for explicit waters (semicontinuum modeling).

For $\Delta^‡S_{total}$ (the activation entropy), the predictions are too negative by 5−17 eu. Among a variety of reasons for this is the approximate nature of the Eyring equation for deriving $\Delta^‡S_{total}$. A likely contributor, however, is the neglect of SSR (solvent structure reorientation), particularly that of the strongly ordering reactant H_3O^+. From Table 7, if we imagine the SSR effect of the cationic transition state to be a little more ordering than that of ethanol (e.g., −22 eu), then the omitted $\Delta^‡\Delta_{SSR}S$ effect for $C_2H_4 + H_3O^+ \rightarrow TS^+$ would be $-22 - (-10 - 27) = +15$, which if included would bring the Trouton-predicted entropies into better agreement with the experimentally determined Eyring-equation value. We propose an empirical correction of $\Delta^‡\Delta_{solv}S_{other} = \Delta^‡\Delta_{SSR}S = +11$ eu (the midpoint of

Table 7. Entropies (eu) of Solvation (from Experiment[a]) and the Derived Solvent-Structure-Reordering Portion ΔS_{SSR} (after Removing the Trouton Constant of −21 eu)

solute	solvent	$\Delta_{solv}S_{expt,298K}$	ΔS_{SSR}
benzene	benzene	−27	−6
C_2H_4	benzene	−16	+5
Ar	benzene	−13	+8
Ne	benzene	−08	+13
He	benzene	−11	+10
water	water	−28	−7
C_2H_5OH	water	−39.5	−18.5
C_2H_4	water	−31	−10
Ar	water	−31	−10
Ne	water	−26	−5
He	water	−24	−3
H_3O^+	water	−48	−27
H^+	water	−44.5	−23.5
Li^+	water	−48	−27
Na^+	water	−40	−19
K^+	water	−31	−10
Rb^+	water	−29	−8

[a]1 atm → unit mole fraction convention, matching the normal convention for condensation and Trouton's Rule. $\Delta_{solv}S_{expt}$ values from Wilhelm (1973 benzene,[75] 1977 water[76]), except for C_2H_5OH and the ions which come from Cabani et al.[33] and Tissandier et al.,[77] respectively (after standard state conversion).

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

Table 8. Predicted Reaction and Activation Entropies (eu = cal mol^{-1} K^{-1})

model	$\Delta_r S_{freq}$	$\Delta_r\Delta_{Trouton}S$	$\Delta_r\Delta_{SCC}S$	$\Delta_r S_{total}$
reaction				
$C_2H_4 + H_2O \rightarrow C_2H_5OH$	−32.95	13	0	−20
$C_2H_4 + H_2O\cdot\mathbf{W} \rightarrow C_2H_5OH + \mathbf{W}$	−9.33	0	−10	−19
$C_2H_4 + H_2O\cdot\mathbf{4W} \rightarrow C_2H_5OH + \mathbf{W_4}$	−2.81	0	−10	−13
$C_2H_4\cdot\mathbf{W} \rightarrow C_2H_5OH$	−14.52	0	−10	−25
$C_2H_4\cdot\mathbf{2W} + \mathbf{W} \rightarrow C_2H_5OH\cdot\mathbf{2W}$	−47.43	13	0	−34
$C_2H_4\cdot\mathbf{2W} + \mathbf{W_2} \rightarrow C_2H_5OH\cdot\mathbf{3W}$	−51.67	13	0	−39
$C_2H_4\cdot\mathbf{2W} + \mathbf{W_2} \rightarrow C_2H_5OH\cdot\mathbf{2W} + \mathbf{W}$	−23.81	0	−10	−34
$C_2H_4\cdot\mathbf{2W} \rightarrow C_2H_5OH\cdot\mathbf{W}$	−19.36	0	−10	−29
experiment (see Introduction)				*−18*
model	$\Delta^{\ddagger}S_{freq}$	$\Delta^{\ddagger}\Delta_{Trouton}S$	$\Delta^{\ddagger}\Delta_{SCC}S$	$\Delta^{\ddagger}S_{total}$
activation				
$C_2H_4 + H_3O^+\cdot\mathbf{W} + \mathbf{W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{2W}]^{\ddagger}$	−59.51	26	10	−23
$C_2H_4 + H_3O^+\cdot\mathbf{W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{W}]^{\ddagger}$	−30.83	13	0	−18
$C_2H_4 + H_3O^+\cdot\mathbf{2W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{2W}]^{\ddagger}$	−28.94	13	0	−16
$C_2H_4 + H_3O^+\cdot\mathbf{3W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{3W}]^{\ddagger}$	−34.61	13	0	−22
$C_2H_4 + H_3O^+\cdot\mathbf{5W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{5W}]^{\ddagger}$	−35.25	13	0	−22
$C_2H_4 + H_3O^+\cdot\mathbf{5W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{4W}]^{\ddagger} + \mathbf{W}$	−5.82	0	−10	−16
$C_2H_4 + H_3O^+\cdot\mathbf{5W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{3W}]^{\ddagger} + \mathbf{2W}$	22.02	−13	−20	−11
experiment (see Introduction)				*−5.7 ± 3*

Table 9. Summary of Best Results (kcal mol^{-1}, except eu for ΔS)

model	$\Delta_r S$[a]	$T\Delta_r S$	$\Delta_r G$	$\Delta_r H$[e]
reaction				
$C_2H_4 + H_2O \rightarrow C_2H_5OH$ (CPCM)	−20	−5.9	−2.4	−8.4
$C_2H_4 + H_2O \rightarrow C_2H_5OH$ (SMD)			−2.6	−8.6
experiment (see Introduction)	*−18*	*−5.4*	*−3.8*	*−9.2*
model	$\Delta^{\ddagger}S$[b]	$T\Delta^{\ddagger}S$	$\Delta^{\ddagger}G$	$\Delta^{\ddagger}H$[e]
activation				
$C_2H_4 + H_3O^+\cdot\mathbf{2W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{2W}]^{\ddagger}$ (CPCM A1[c])	−5	−1.5	33.4	32.0
$C_2H_4 + H_3O^+\cdot\mathbf{2W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{2W}]^{\ddagger}$ (SMD A1)			34.5	33.0
$C_2H_4 + H_3O^+\cdot\mathbf{3W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{3W}]^{\ddagger}$ (CPCM B1[d])	−11	−3.2	34.8	31.7
$C_2H_4 + H_3O^+\cdot\mathbf{3W} \rightarrow [C_2H_5OH_2^+\cdot\mathbf{3W}]^{\ddagger}$ (SMD B1)			34.4	31.2
experiment[22]	*−5.7 ± 2.5*	*−1.7*	*34.1*	*32.4 ± 1.0*
$C_3H_6 + H_3O^+\cdot\mathbf{2W} \rightarrow [C_3H_7OH_2^+\cdot\mathbf{2W}]^{\ddagger}$ (CPCM A1)	−7	−2.0	27.1	25.1
$C_3H_6 + H_3O^+\cdot\mathbf{2W} \rightarrow [C_3H_7OH_2^+\cdot\mathbf{2W}]^{\ddagger}$ (SMD A1)			29.0	27.0
$C_3H_6 + H_3O^+\cdot\mathbf{3W} \rightarrow [C_3H_7OH_2^+\cdot\mathbf{3W}]^{\ddagger}$ (CPCM B1)	−11	−3.2	29.1	25.9
$C_3H_6 + H_3O^+\cdot\mathbf{3W} \rightarrow [C_3H_7OH_2^+\cdot\mathbf{3W}]^{\ddagger}$ (SMD B1)			29.3	26.1
experiment[22]	*−5.4 ± 2.5*	*−1.6*	*28.0*	*26.4 ± 1.0*
i-$C_4H_8 + H_3O^+\cdot\mathbf{2W} \rightarrow [C_4H_9OH_2^+\cdot\mathbf{2W}]^{\ddagger}$ (CPCM A1)	−5	−1.5	20.8	19.3
i-$C_4H_8 + H_3O^+\cdot\mathbf{2W} \rightarrow [C_4H_9OH_2^+\cdot\mathbf{2W}]^{\ddagger}$ (SMD A1)			23.1	21.6
i-$C_4H_8 + H_3O^+\cdot\mathbf{3W} \rightarrow [C_4H_9OH_2^+\cdot\mathbf{3W}]^{\ddagger}$ (CPCM B1)	−9	−2.7	22.2	19.5
i-$C_4H_8 + H_3O^+\cdot\mathbf{3W} \rightarrow [C_4H_9OH_2^+\cdot\mathbf{3W}]^{\ddagger}$ (SMD B1)			23.0	20.3
experiment[22]	*−7.7 ± 0.7*	*−2.3*	*23.4*	*21.1 ± 0.2*

[a]From $\Delta^{\ddagger}S_{freq}$ + 13 (Trouton constant × $\Delta^{\ddagger}\nu$). [b]From $\Delta^{\ddagger}S_{freq}$ + 13 (Trouton constant × $\Delta^{\ddagger}\nu$) + 11 (solvent structure reordering of H_3O^+). [c]Option A1: $\Delta^{\ddagger}G$ avoids $\Delta^{\ddagger}\Delta_{SCC}G$ and any empirical corrections. [d]Option B1: $\Delta^{\ddagger}G$ avoids $\Delta^{\ddagger}\Delta_{SCC}G$ and includes *a posteriori* −4 kcal mol^{-1} empirical correction. [e]ΔH computed from $\Delta G + T\Delta S$.

the +5 to +17 corrections that would be needed for the −5 to −17 errors seen in Table 8) to be added to the Trouton-based predictions for acid catalysis activations.

4.6. Summary of Results, Including Enthalpy $\Delta_r H$ and $\Delta^{\ddagger}H$. Finally, enthalpy predictions can be derived for each semicontinuum model by combining free energy and entropy results. We present in Table 9 a summary of results from the recommended protocols.

5. CONCLUSIONS

Two types of CSM errors are categorized: (i) the high q/r error ε_{hqr} when solvating ions, an error that can be cured with semicontinuum modeling, and (ii) an error that grows with increasing explicit hydration of solute, which we term the imperfect cavity size error ε_{ics}. The growing ε_{ics} errors can cancel well in $\Delta_r G$ or $\Delta^{\ddagger}G$ computation, leading to asymptotic values.

For the overall reaction ethane + water → ethanol in aqueous media, no explicit water molecules (other than the reactant one) are needed. A 2 kcal mol^{-1} accuracy is obtained for $\Delta_r G$ from the common CPCM and SMD models of Gaussian09 and

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

Gaussian16. For $\Delta_r S$ a simple Trouton-based solvation entropy (−21 eu for 1 atm → 55 M, adjusted to −13 eu for 1 atm → 1 M) applied to gas-phase entropies achieves 2 eu accuracy, due to significant cancellation of the omitted large solvent-structure-reordering terms in the individual entropies of solvation. The resulting enthalpy $\Delta_r H$, from $\Delta_r G + T\Delta_r S$, also achieved 2 kcal mol^{-1} accuracy. Addition of explicit waters tended to make these predictions worse and is not recommended.

For the acid-catalyzed activation Gibbs energy $\Delta^{\ddagger}G$, explicit water molecules are needed to reduce the high q/r error that CSMs make for H_3O^+. The addition of explicit waters to the ions only, i.e., $H_3O^+ \cdot n$W and the transition state $[C_2H_5OH_2^+ \cdot m\text{W}]^{\ddagger}$, results in converged asymptotic values at $n \approx 5$ or 6. Use of the newly introduced *solvent-confinement correction*, $\Delta_{SCC}G$, allows offset series (e.g., $m = n \pm 1$) to be used. The asymptotic values are generally too high, however, due to poor cancellation of CSM errors for the two CSMs employed: CPCM with default settings produces an unusual negative error for hydrocarbons in water, while SMD has an unusual base error for H_3O^+ versus the transition state cation. Options (labeled A1, A2, B1, ...) are offered for "error management" to improve error cancellation when computing $\Delta^{\ddagger}G$ with CPCM or SMD, and they are shown to achieve 3 kcal mol^{-1} accuracy for three acid-catalysis activations (propene and isobutene as well as ethene). These options should constitute recommended first choices for modellers interested in predicting acid-catalysis activation Gibbs energies. The simplest, Option A1, is the semicontinuum model

$$A + H_3O^+ \cdot 2W \rightarrow [AHOH_2^+ \cdot 2W]^{\ddagger}$$

which requires no corrections (solvent-confinement or otherwise) to default CPCM or SMD computations, other than the usual + 1.9 kcal mol^{-1} $\Delta_{conc}G$ correction on each solute for 1 atm → 1 M concentration conversion.

For activation entropy, the simple Trouton-based solvation entropy idea correctly did not work terribly well, predicting $\Delta^{\ddagger}S$ to be generally 5−17 eu more negative than experimental Eyring-equation values. One significant cause of the discrepancy could be the neglect of solvent structure reorientation, as H_3O^+ may have a unique (and exploitable) large error here. For acid catalysis, a proposed *a posteriori* correction of −11 eu on H_3O^+ (so +11 eu on $\Delta^{\ddagger}S$) is proposed and works for all three activations considered (ethene, propene, isobutene).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.0c07011.

Derivation of experimental values for ethene hydration energies and entropies, justification for using Trouton's constant for solvation, table of typical CSM cavity radii, and Cartesian coordinates and computed energies and entropies of species (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Allan L. L. East − *Department of Chemistry and Biochemistry, University of Regina, Regina, SK S4S 0A2, Canada;* orcid.org/0000-0003-1898-4370; Email: allan.east@uregina.ca

Author

Darpan H. Patel − *Department of Chemistry and Biochemistry, University of Regina, Regina, SK S4S 0A2, Canada*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.0c07011

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work was supported by the Natural Sciences and Engineering Research Council (Discovery Grant RGPIN-2017-06247) and the Canada Foundation for Innovation (Leading Edge Fund 2009, grant 21625). A.L.L.E. thanks his 2017 Visiting Professorship at the Center for Solvation Science ZEMOS@RUB (Ruhr University, Bochum, Germany), a part of the Cluster of Excellence RESOLV funded by the German Research Foundation DFG (EXC 1069/2033), allowing discussions with D. Marx on semicontinuum modelling strategies. A. Mukadam is thanked for exploratory testing of many PCM CSM variations for their effects on solvation energies of some neutral solutes. Several referees are also thanked for valuable comments leading to a more definitive manuscript.

■ REFERENCES

(1) Cramer, C. J.; Truhlar, D. G. Implicit solvation models: equilibria, structure, spectra, and dynamics. *Chem. Rev.* **1999**, *99*, 2161−2200.

(2) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999−3093.

(3) Klamt, A. The COSMO and COSMO-RS solvation models. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 699−709.

(4) Tomasi, J.; Mennucci, B.; Cancès, E. The IEF version of the PCM solvation method: An overview of a new method addressed to study molecular solutes at the QM ab initio level. *J. Mol. Struct.: THEOCHEM* **1999**, *464*, 211−216.

(5) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. New developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution. *J. Chem. Phys.* **2002**, *117*, 43−54.

(6) Scalmani, G.; Frisch, M. J. Continuous surface charge polarizable continuum models of solvation. I. General formalism. *J. Chem. Phys.* **2010**, *132* (15p), 114110.

(7) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **2003**, *24*, 669−681.

(8) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **2009**, *113*, 6378−6396.

(9) Dhillon, S.; East, A. L. L. Challenges in predicting $\Delta_{rxn}G$ in solution: hydronium, hydroxide, and water autoionization. *Int. J. Quantum Chem.* **2018**, *118*, No. e25703.

(10) Orozco, M.; Luque, F. J. Optimization of the cavity size for ab initio MST-SCRF calculations of monovalent ions. *Chem. Phys.* **1994**, *182*, 237−248.

(11) Johnston, K. P.; Bennett, G. E.; Balbuena, P. B.; Rossky, P. J. Continuum electrostatics model for ion solvation and relative acidity of HCl in supercritical water. *J. Am. Chem. Soc.* **1996**, *118*, 6746−6752.

(12) Barone, V.; Cossi, M.; Tomasi, J. A new definition of cavities for the computation of solvation free energies by the polarizable continuum model. *J. Chem. Phys.* **1997**, *107*, 3210−3221.

(13) Chipman, D. M. Computation of pKa from dielectric continuum theory. *J. Phys. Chem. A* **2002**, *106*, 7413−7422.

(14) Ginovska, B.; Camaioni, D. M.; Dupuis, M.; Schwerdtfeger, C. A.; Gil, Q. Charge-dependent cavity radii for an accurate dielectric

continuum model for solvation with emphasis on ions: aqueous solute with oxo, hydroxo, amino, methyl, chloro, bromo, and fluoro functionalities. *J. Phys. Chem. A* **2008**, *112*, 10604−10613.
(15) 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*; Gaussian, Inc.: Wallingford, CT, 2003.
(16) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. SM6: a density functional theory continuum solvation model for calculating aqueous solvation free energies of neutrals, ions and solute-water clusters. *J. Chem. Theory Comput.* **2005**, *1*, 1133−1152.
(17) Khalili, F.; Henni, A.; East, A. L. L. Entropy contributions in pKa computation: application to alkanolamines and piperazines. *J. Mol. Struct.: THEOCHEM* **2009**, *916*, 1−9.
(18) Benn, M. H.; Rauk, A.; Swaddle, T. W. Measurement of the interaction of aqueous copper(II) with a model amyloid-β protein fragment-Interference from buffers. *Can. J. Chem.* **2011**, *89*, 1429−1444.
(19) Thapa, B.; Schlegel, H. B. Density functional theory calculation of pKa's of thiols in aqueous solution using explicit water molecules and the polarizable continuum model. *J. Phys. Chem. A* **2016**, *120*, 5726−5735.
(20) Lian, P.; Johnston, R. C.; Parks, J. M.; Smith, J. C. Quantum chemical calculaions of pKa's on environmentally rlevent functional groups: carboxylic acids, mines, and thiols in aqueous solution. *J. Phys. Chem. A* **2018**, *122*, 4366−4373.
(21) Xu, L.; Coote, M. L. Improving the accuracy of PCM-UAHF and PCM-UAKS calculations using optimized electrostatic scaling factors. *J. Chem. Theory Comput.* **2019**, *15*, 6958−6967.
(22) Xu, L.; Coote, M. L. Correction to improving the accuracy of PCM-UAHF and PCM-UAKS calculations using optimized electrostatic scaling factors. *J. Chem. Theory Comput.* **2020**, *16*, 816−817.
(23) Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Self-consistent reaction field model for aqueous and nonaqueous solutions based on accurate polarized partial charges. *J. Chem. Theory Comput.* **2007**, *3*, 2011−2033.
(24) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Generalized Born Solvation Model SM12. *J. Chem. Theory Comput.* **2013**, *9*, 609−620.
(25) Rick, S. W.; Berne, B. J. The aqueous solvation of water: a comparison of continuum methods with molecular dynamics. *J. Am. Chem. Soc.* **1994**, *116*, 3949−3954.
(26) Pliego, J. R., Jr.; Riveros, J. M. The cluster-continuum model for the calculation of the solvation free energy of ionic species. *J. Phys. Chem. A* **2001**, *105*, 7241−7247.
(27) Tuñón, I.; Silla, E.; Bertrán, J. Proton solvation in liquid water. An ab initio study using the continuum model. *J. Phys. Chem.* **1993**, *97*, 5547−5552.
(28) Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A., III. Calculation of solvation free energies of charged solutes using mixed cluster/continuum models. *J. Phys. Chem. B* **2008**, *112*, 9709−9719.
(29) Sumon, K. Z.; Bains, C. H.; Markewich, D. J.; Henni, A.; East, A. L. L. Semicontinuum solvation modeling improves predictions of carbamate stability in the CO_2 + aqueous amine reaction. *J. Phys. Chem. B* **2015**, *119*, 12256−12264.
(30) Sumon, K. Z.; Henni, A.; East, A. L. L. Predicting pK_a of amines for CO_2 capture: computer versus pencil-and-paper. *Ind. Eng. Chem. Res.* **2012**, *51*, 11924−11930.
(31) Pill, M. F.; East, A. L. L.; Marx, D.; Beyer, M. K.; Clausen-Schaumann, H. Mechanical activation drastically accelerates amide bond hydrolysis, matching enzyme activity. *Angew. Chem., Int. Ed.* **2019**, *58*, 9787−9790.
(32) Baliga, B. T.; Whalley, E. Effect of pressure and temperature on the rate of the acid-catalyzed hydration of ethylene. *Can. J. Chem.* **1965**, *43*, 2453−2456.
(33) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J. Solution Chem.* **1981**, *10*, 563−595.
(34) van Erp, T. S.; Meijer, E. J. Proton-assisted ethylene hydration in aqueous solution. *Angew. Chem., Int. Ed.* **2004**, *43*, 1660−1662.
(35) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098−3100.
(36) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(37) Tolosa Arroyo, S.; Corchado Martin-Romo, J. C.; Hidalgo Garcia, A.; Sansón Martin, J. A. Molecular simulation of the hydration of ethene to ethanol using ab initio potentials and free energy curves. *J. Phys. Chem. A* **2007**, *111*, 13515−13520.
(38) Yamabe, S.; Tsuchida, N.; Yamazaki, S. One-step paths of the alkene hydration revealed by a DFT study. *ChemistrySelect* **2017**, *2*, 6857−6864.
(39) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom−atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(40) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Varone, B.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, R.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian09*, Rev. C01; Gaussian Inc.: Wallingford, CT, 2010.
(41) Sturgeon, M. R.; Kim, S.; Lawrence, K.; Paton, R. S.; Chmely, S. C.; Nimlos, M.; Foust, T. D.; Beckham, G. T. A mechanistic investigation of acid-catalyzed cleavage of aryl-ether linkages: implications for lignin depolymerization in acidic environments. *ACS Sustainable Chem. Eng.* **2014**, *2*, 472−485.
(42) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06 functionals and twelve other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(43) Zhao, Y.; Truhlar, D. G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **2008**, *41*, 157−167.
(44) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618−622.
(45) Scuseria, G. E.; Schaefer, H. F., III. Is coupled cluster singles and doubles (CCSD) more computationally intensive than quadratic configuration interaction (QCISD)? *J. Chem. Phys.* **1989**, *90*, 3700−3703.
(46) Caricato, M.; Scalmani, G.; Trucks, G. W.; Frisch, M. J. Coupled cluster calculations in solution with the polarizable continuum model of solvation. *J. Phys. Chem. Lett.* **2010**, *1*, 2369−2373.

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104

(47) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 theory. *J. Chem. Phys.* **2007**, *126*, No. 084108.
(48) Ho, J.; Klamt, A.; Coote, M. L. Comment on the correct use of continuum solvent models. *J. Phys. Chem. A* **2010**, *114*, 13442−13444.
(49) Onsager, L. Theories of concentrated electrolytes. *Chem. Rev.* **1933**, *13*, 73−89.
(50) Ribeiro, R. F.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Use of solution-phase vibrational frequencies in continuum models for the free energy of solvation. *J. Phys. Chem. B* **2011**, *115*, 14556−14562.
(51) Silva, C. M.; Silva, P. L.; Pliego, J. R. Prediction of the pH-rate profile for dimethyl sulfide oxidation by hydrogen peroxide: the role of elusive $H_3O_2^+$ ion. *Int. J. Quantum Chem.* **2014**, *114*, 501−507.
(52) Silva, P. L.; Silva, C. M.; Guimaraes, L.; Pliego, J. R., Jr. Acid-catalyzed transesterification and esterification in methanol: a theoretical cluster-continuum investigation of the mechanisms and free energy barriers. *Theor. Chem. Acc.* **2015**, *134* (13p), 1591.
(53) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.
(54) Riccardi, D.; Guo, H.-B.; Parks, J. M.; Gu, B.; Liang, L.; Smith, J. C. Cluster-continuum calculations of hydration free energies of anions and group 12 divalent cations. *J. Chem. Theory Comput.* **2013**, *9*, 555−569.
(55) *Jaguar*, version 7.0; Schrodinger, LLC: New York, NY, 2007.
(56) Eley, D. D.; Evans, M. G. Heats and entropy changes accompanying the solution of ions in water. *Trans. Faraday Soc.* **1938**, *34*, 1093−1112.
(57) Krestov, G. A. Thermodynamic characteristics of structural changes in water associated with the hydration of ions. *Zh. Strukt. Khim.* **1962**, *3*, 137−142.
(58) Bockris, J. O'M.; Saluja, P. P. S. Approximate calculations of the heats and entropies of hydration according to various models. *J. Phys. Chem.* **1972**, *76*, 2298−2310.
(59) Pierotti, R. A. A scaled particle theory of aqueous and nonaqueous solutions. *Chem. Rev.* **1976**, *76*, 717−726.
(60) Wertz, D. H. Relationship between the gas-phase entropies of molecules and their entropies of solvation in water and 1-octanol. *J. Am. Chem. Soc.* **1980**, *102*, 5316−5322.
(61) Abraham, M. H.; Liszi, J.; Papp, E. Calculations on ion solvation. *J. Chem. Soc., Faraday Trans. 1* **1982**, *78*, 197−211.
(62) Marcus, Y. The hydration entropies of ions and their effects on the structure of water. *J. Chem. Soc., Faraday Trans. 1* **1986**, *82*, 233−242.
(63) Marcus, Y. Viscosity B-Coefficients, Structural entropies and heat capacities, and the effects of ions on the structure of water. *J. Solution Chem.* **1994**, *23*, 831−848.
(64) Irudayam, S. J.; Henchman, R. H. Solvation theory to provide a molecular interpretation of the hydrophobic entropy loss of noble-gas hydration. *J. Phys.: Condens. Matter* **2010**, *22*, No. 284108.
(65) Irudayam, S. J.; Henchman, R. H. Prediction and interpretation of the hydration entropies of monovalent cations and anions. *Mol. Phys.* **2011**, *109*, 37−48.
(66) Garza, A. J. Solvation entropy made simple. *J. Chem. Theory Comput.* **2019**, *15*, 3204−3214.
(67) Bjerrum, N.; Larsson, E. Studien über ionenverteilungskoeffizienten. I. *Z. Phys. Chem.* **1927**, *127*, 358−384.
(68) Latimer, W. M.; Kasper, C. The theoretical evaluation of the entropies of aqueous ions. *J. Am. Chem. Soc.* **1929**, *51*, 2293−2299.
(69) Gurney, R. W. *Ionic Processes in Solution*; Dover Publications: New York, 1953.
(70) Sandbeck, D. J. S.; Kuntz, C. M.; Luu, C.; Mondor, R. A.; Ottaviano, J. G.; Rayer, A. V.; Sumon, K. Z.; East, A. L. L. Challenges in predicting $\Delta_{rxn}G$ in solution: The mechanism of ether-catalyzed hydroboration of alkenes. *J. Phys. Chem. A* **2014**, *118*, 11768−11779.
(71) Chen, Y.-M.; Chass, G. A.; Fang, D.-C. Between a reactant rock and a solvent hard place − molecular corrals guide aromatic substitutions. *Phys. Chem. Chem. Phys.* **2014**, *16*, 1078−1083.
(72) Zhao, L.; Li, S.-J.; Fang, D.-C. A theoretical study of ene reactions in solution: a solution-phase translational entropy model. *ChemPhysChem* **2015**, *16*, 3711−3718.
(73) Li, S.-J.; Fang, D.-C. A DFT kinetic study on 1,3-dipolar cycloaddition reactions in solution. *Phys. Chem. Chem. Phys.* **2016**, *18*, 30815−30823.
(74) McLachlan, D., Jr.; Marcus, R. J. The statistical-mechanical basis of Trouton's Rule. *J. Chem. Educ.* **1957**, *34*, 460−462.
(75) Wilhelm, E.; Battino, R. Thermodynamic functions of the solubilities of gases in liquids at 25°C. *Chem. Rev.* **1973**, *73*, 1−9.
(76) Wilhelm, E.; Battino, R.; Wilcock, R. J. Low-pressure solubility of gases in liquid water. *Chem. Rev.* **1977**, *77*, 219−262.
(77) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. The proton's absolute aqueous enthalpy and Gibbs free energy of solvation from cluster-ion solvation data. *J. Phys. Chem. A* **1998**, *102*, 7787−7794.
(78) East, A. L. L.; Radom, L. Ab initio statistical thermodynamical models for the computation of third-law entropies. *J. Chem. Phys.* **1997**, *106*, 6655−6674.

https://dx.doi.org/10.1021/acs.jpca.0c07011
J. Phys. Chem. A 2020, 124, 9088−9104