# Challenges in predicting $\Delta_{rxn}G$ in solution: The chelate effect

Cite as: J. Chem. Phys. **157**, 034109 (2022); https://doi.org/10.1063/5.0097291
Submitted: 27 April 2022 • Accepted: 27 June 2022 • Published Online: 20 July 2022

**A. A. Mukadam and A. L. L. East**

The Journal of Chemical Physics

ARTICLE

scitation.org/journal/jcp

# Challenges in predicting $\Delta_{rxn}G$ in solution: The chelate effect

Cite as: J. Chem. Phys. 157, 034109 (2022); doi: 10.1063/5.0097291
Submitted: 27 April 2022 • Accepted: 27 June 2022 •
Published Online: 20 July 2022

A. A. Mukadam and A. L. L. East[a)]

## AFFILIATIONS

Department of Chemistry and Biochemistry, University of Regina, Regina, Saskatchewan S4S 0A2, Canada

[a)] Author to whom correspondence should be addressed: allan.east@uregina.ca

## ABSTRACT

Gibbs energies for reactions involving aqueous ions are challenging to predict due to the large solvation energies of such ions. A stringent test would be the *ab initio* reproduction of the aqueous-phase chelate effect, an entropic effect in reactions of very small enthalpy changes. This paper examines what is required to achieve such a reproduction for the paradigmatic reaction $M(NH_3)_4^{2+} + 2\ en \rightarrow M(en)_2^{2+} + 4\ NH_3$ ($en$ = 1,2-ethylenediamine), for which $\Delta_{rxn}G^*$ and $\Delta_{rxn}H^*$ are −2.3 and +1.6 kcal mol$^{-1}$, respectively, if M = Zn. Explicit solvation via simulation was avoided in order to allow sufficiently accurate electronic structure models; this required the use of continuum solvation models (CSMs), and a great deal of effort was made in attempting to lower the relative errors of $\Delta_{solv}G^*[M(NH_3)_4^{2+}]$ vs $\Delta_{solv}G^*[M(en)_2^{2+}]$ from the CSMs available in Gaussian software. CSMs in ADF and JDFTx software were also tested. A uniform 2.2 kcal mol$^{-1}$ accuracy in $\Delta_{rxn}G^*$ for all three metal-atom choices M = {Zn, Cd, Hg} was eventually achieved, but not from any of the known CSMs tested, nor from cavity size reoptimization, nor from semicontinuum modeling: *post facto* solvation energy corrections [one per solute type, $NH_3$, $en$, $M(NH_3)_4^{2+}$, $M(en)_2^{2+}$] were needed. It is hoped that this study will aid (and encourage) further CSM development for coordination-complex ions.

## I. INTRODUCTION

Prediction of accurate solution-phase reaction energies via quantum chemistry is still a challenge.[1–4] Continuum solvation models (CSMs)[5–8] are commonly used for the solvation effect. One known source of error in these CSM-generated free energies is with aqueous ions.[3,4,9,10] Generally, CSM errors of 5 kcal mol$^{-1}$ for ion solvation are not unexpected[11–13] but can be considerably higher.[3,4,9,10] While one might hope that these errors can cancel somewhat in a reaction Gibbs energy, thus allowing for more accuracy in $\Delta_{rxn}G$ predictions than in $\Delta_{solv}G$ predictions, this should be tested more than it has been.

Recently, one of us classified the observed CSM errors on hydrated ion clusters into two types: a "q/r error" caused by high charge density that decreases with the cluster size and an "ics error" caused by imperfect cavity size that increases with the cluster size.[4] Errors of >10 kcal mol$^{-1}$, seen with $H_3O^+$,[4,9] $OH^-$,[4] and atomic ions,[9,10] have a q/r error that can be effectively eliminated with the use of explicit waters within the CSM cavity (cluster-continuum[10] or semicontinuum[14] modeling). Regrettably, for large clusters, the ics error was seen to grow to be just as large, and while error cancellation worked beautifully well for the large clusters used to solve the water autoionization problem [$2\ H_2O_{(l)} \rightarrow H_3O^+{}_{(aq)} + OH^-{}_{(aq)}$],[3] residual ics error problems limited the accuracy for alkene activations [$C_jH_{2j(aq)} + H_3O^+{}_{(aq)} \rightarrow$ transition state], with an error of roughly 6 kcal mol$^{-1}$ after semicontinuum extrapolation (large-cluster convergence).[4] Error management strategies were proposed at the time, applicable to aqueous acid-catalyzed activations.

In this paper, we wish to see if current CSMs are accurate enough, or feature enough error cancellation, to properly reproduce the entropy-driven *chelate effect*, the extra stability of metal complexes having chelating ligands (bidentate or polydentate ligands), vs those that do not. Chelate complexes are well known: they exist naturally in the blood (hemoglobin) and green plants (chlorophyll) as metal–porphyrin complexes[15] and have excellent applications in biochemistry and microbiology,[16] mineral extraction,[17] and food chemistry.[18] Although Werner[19] had first reported the synthesis of some chelate complexes in 1893, the extra stability of chelate complexes (chelate effect) was explained only 50 years later, qualitatively by Calvin[20,21] and semiquantitatively by Schwarzenbach:[22] an entropy effect due to the increased translational motion caused by the liberation of monodentate ligands. Later studies examined the quantitative variations among systems.[23–25]

The reason why the chelate effect may constitute a challenge for CSMs is the high accuracy required. The classic entropy-only chelation reactions have $\Delta_{\mathrm{rxn}}G < 0$ but $\Delta_{\mathrm{rxn}}H > 0$ at standard state (1M) conditions. The $\Delta_{\mathrm{rxn}}G$ values are, thus, sufficiently close to zero that whatever error the CSMs make would either need to be small or at least cancel well. To our knowledge, such a chelate effect (with $\Delta_{\mathrm{rxn}}G < 0$ but $\Delta_{\mathrm{rxn}}H > 0$) has never been demonstrated with CSM calculations. We here focus on the paradigm chelation reaction,

$$\mathrm{Zn(NH_3)_4}^{2+}{}_{(\mathrm{aq})} + 2\,en_{(\mathrm{aq})} \rightarrow \mathrm{Zn}(en)_2{}^{2+}{}_{(\mathrm{aq})} + 4\,\mathrm{NH}_{3(\mathrm{aq})} \tag{1}$$

($en$ = 1,2-ethylenediamine, $\mathrm{NH_2CH_2CH_2NH_2}$) for which $\Delta_{\mathrm{rxn}}G^*$ and $\Delta_{\mathrm{rxn}}H^*$ are −2.3 and +1.6 kcal $\mathrm{mol}^{-1}$, respectively ($^*$ indicates standard-state conditions of T = 298 K and 1M concentrations). These values were derived from the aqueous dissociation data of Spike and Parry:[26] $\{\Delta_{\mathrm{diss}}G, \Delta_{\mathrm{diss}}H\} = \{13.4, 14.1\}$ for $\mathrm{Zn(NH_3)_4}^{2+}{}_{(\mathrm{aq})} \rightarrow \mathrm{Zn}^{2+}{}_{(\mathrm{aq})} + 4\ \mathrm{NH}_{3(\mathrm{aq})}$ and $\{\Delta_{\mathrm{diss}}G, \Delta_{\mathrm{diss}}H\} = \{15.7, 12.5\}$ for $\mathrm{Zn}(en)_2{}^{2+}{}_{(\mathrm{aq})} \rightarrow \mathrm{Zn}^{2+}{}_{(\mathrm{aq})} + 2\ en_{(\mathrm{aq})}$. Spike and Parry obtained their data at an ionic strength of 2.1; at an ionic strength of zero, the $\Delta_{\mathrm{diss}}G$ values for $\mathrm{Zn(NH_3)_4}^{2+}$ and $\mathrm{Zn}(en)_2{}^{2+}$ are known to be reduced by a full kcal $\mathrm{mol}^{-1}$ (12.1 and 14.5 kcal $\mathrm{mol}^{-1}$, respectively).[27] The reason for the higher $\Delta_{\mathrm{diss}}G$ at I = 2.1 could be due to the dielectric decrement phenomenon[28,29] that may reduce the stability of loose amines. Fortunately, the $\Delta_{\mathrm{rxn}}G$ (and presumably $\Delta_{\mathrm{rxn}}H$) for the chelation reaction (the difference of these two dissociations) appears to be independent of the ionic strength: obtaining $\Delta_{\mathrm{rxn}}G$ from these $\Delta_{\mathrm{diss}}G$ values gives a value of −2.4 kcal $\mathrm{mol}^{-1}$ at I = 0, similar to the value of −2.3 kcal $\mathrm{mol}^{-1}$ at I = 2.1.[26] This simplifies the modeling, allowing computations of $\Delta_{\mathrm{rxn}}G$ and $\Delta_{\mathrm{rxn}}H$ here to be performed at I = 0.

As usual, the CSM-specific Gibbs energy computation for 1M solution is

$$G^* = G_{\mathrm{gas}}{}^{\circ} + \Delta G^{\circ\rightarrow *} + \Delta_{\mathrm{solv}}G^* \tag{2a}$$

$$= E_{\mathrm{el,\,gas}} + G_{\mathrm{freq}}{}^{\circ} + \Delta G^{\circ\rightarrow *} + \Delta_{\mathrm{solv}}G^* \tag{2b}$$

$$= G_{el,\varepsilon}{}^{*} + G_{\mathrm{freq}}{}^{\circ} + \Delta G^{\circ\rightarrow *}, \tag{2c}$$

where ° indicates gas-phase standard-state conditions of $T^{\circ}$ = 298 K and $P^{\circ}$ = 1 atm. $\Delta G^{\circ\rightarrow *}$ is the cratic (concentration-change) term $RT\ln(c^*/c^{\circ}) = RT\ln([1\mathrm{M}]/[P^{\circ}/RT^{\circ}]) = +1.89$ kcal $\mathrm{mol}^{-1}$. $G_{\mathrm{freq}}{}^{\circ}$ is the set of thermal motion (nuclear-motion degree-of-freedom) corrections, including zero-point vibrational energy, to convert electronic energy $E_{\mathrm{el,gas}}$ to Gibbs energy $G_{\mathrm{gas}}$, typically done with single-conformer vibrational frequency calculations that employ gas-phase rigid-rotor harmonic-oscillator equations and assume by default 1 atm gas concentrations. $\Delta_{\mathrm{solv}}G^*$ is the constant-concentration solvation Gibbs energy determined by the CSM calculation. The CSM calculation determines directly the sum $G_{\mathrm{el},\varepsilon}{}^{*} = E_{\mathrm{el,gas}} + \Delta_{\mathrm{solv}}G^*$, and hence, determinations of the $\Delta_{\mathrm{solv}}G$ term separately (if desired) would require a separate calculation of $E_{\mathrm{el,gas}}$.

To properly show that a CSM is qualitatively reproducing the chelate effect as a purely entropic effect, a means of predicting $\Delta_{\mathrm{rxn}}H$ separate from $\Delta_{\mathrm{rxn}}G$ was desired, and to do this, the recent Trouton-based approximation of Patel and East[4] for solvation entropy was employed,

$$S^* = S_{\mathrm{gas}} + \Delta S^{\circ\rightarrow *} + \Delta_{\mathrm{solv}}S^* \tag{3a}$$

$$\approx S_{\mathrm{freq}}{}^{\circ} + \Delta S_{\mathrm{Trouton}}{}^{\circ\rightarrow *}, \tag{3b}$$

where $\Delta S_{\mathrm{Trouton}}{}^{\circ\rightarrow *}$ is −21 + 8 = −13 cal $\mathrm{mol}^{-1}$ $\mathrm{K}^{-1}$, the normal Trouton constant of −21 from solvent condensation but corrected for the 1M solute concentration.[4] Enthalpy is then simply derived with $\Delta H = \Delta G + T\Delta S$.

## II. METHODS

Gaussian09 software was primarily used.[30] References for the electronic structure methods (ESMs) and continuum solvation models (CSMs) appear in the supplementary material and Sec. IV D, respectively. Our initial ESM/CSM combination (Sec. III) was B3LYP/{SDD, 6-31+G(d)}/IEFPCM, where the mixed basis set notation {SDD, 6-31+G(d)} refers to SDD[30] for Zn, and 6-31+G(d) for all other atoms, and IEFPCM (Integral-Equation-Formalism Polarizable Continuum Model) is the default CSM employed by the SCRF (self-consistent reaction field) flag (with no "non-electrostatic" terms). Since the mixed basis set runs for the Zn complexes require the "gen" keyword, which enforces the "5d" (spherical-harmonic) versions of 6-31+G(d), all runs for $en$ and $\mathrm{NH_3}$ employed the "gen" keyword as well to ensure "5d" consistency. When higher accuracy was explored (Sec. IV), several more ESMs and CSMs were tested, including CSMs not found in Gaussian09 (for which we used the ADF[31] and JDFTx[32] codes).

We alert the reader to issues regarding geometry optimization and conformer choice. (i) Our geometry optimizations were initially performed with SDD throughout, instead of the mixed {SDD, 6-31+G(d)}, but these were later abandoned because they caused ring strain errors in the complex $\mathrm{Zn}(en)_2{}^{2+}$. (ii) Unexpected asymmetric distortion of $\mathrm{Zn}(en)_2{}^{2+}$ and $\mathrm{Zn(NH_3)_4}^{2+}$ occurred with certain ESM/CSM combinations. (iii) Two of the four species, $en$ and the chelate complex $\mathrm{Zn}(en)_2{}^{2+}$, have multiple conformers. We now detail what was done regarding issues (ii) and (iii).

There were geometry distortion issues. The point group symmetries (and rotational symmetry numbers) were expected to be: $\mathrm{Zn}(en)_2{}^{2+}$, $D_2$, 4; $\mathrm{NH_3}$, $C_{3v}$, 3; $\mathrm{Zn(NH_3)_4}^{2+}$, $T_d$, 12; and $en$, $C_{2h}$, 2. However, some ESM/CSM combinations gave $C_1$-distorted minimum-energy structures of both the $\mathrm{Zn(NH_3)_4}^{2+}$ and $\mathrm{Zn}(en)_2{}^{2+}$, with negligible energy benefits. Our initial B3LYP/{SDD, 6-31+G(d)}/IEFPCM choice did this (Table I), with electronic energy benefits of only 0.11 and 0.03 kcal $\mathrm{mol}^{-1}$, respectively. It is not known if these distortions are real (e.g., dipole creation in a high dielectric?) or some artifact of the cavity discretization error. Even if the distortions are real, the energetic benefit of distortion is so small that a high-symmetry "vibrationally averaged" structure should be formally considered, with all positive vibrational frequencies but paired with high-symmetry rotational symmetry numbers of 12 and 4, respectively. Therefore, for the initial results (Sec. III), all terms in Eqs. (2) and (3) were computed using the high-symmetry saddle point except the vibrational contributions to $G_{\mathrm{freq}}$ and $S_{\mathrm{freq}}$, which were taken from the $C_1$-symmetry minimum. In Sec. IV, we more often employed B3LYP/{SDD, 6-31+G(d)}/gas optimized geometries, where no $C_1$ distortions occurred.

There were also conformer choice issues. For $en$ ($\mathrm{NH_2CH_2CH_2NH_2}$), there are estimated seven low-energy

**TABLE I.** Bond angles of $C_1$-Symmetry B3LYP/{SDD, 6-31+G(d)}/IEFPCM optimized structures of $Zn(NH_3)_4{}^{2+}{}_{(aq)}$ and $Zn(en)_2{}^{2+}{}_{(aq)}$.

| $Zn(NH_3)_4{}^{2+}{}_{(aq)}$ angle | $\theta_{NZnN}$ | $Zn(en)_2{}^{2+}{}_{(aq)}$ angle | $\theta_{NZnN}$ |
|---|---|---|---|
| 2N–Zn–5N | 104.1[a] | 14N–Zn–17N | 84.4 |
| 2N–Zn–8N | 114.3[a] | 14N–Zn–20N | 123.8[b] |
| 2N–Zn–11N | 104.4[a] | 14N–Zn–23N | 130.3[b] |
| 5N–Zn–8N | 105.3[a] | 17N–Zn–20N | 128.8[b] |
| 5N–Zn–11N | 120.0[a] | 17N–Zn–23N | 111.0[b] |
| 8N–Zn–11N | 109.0[a] | 20N–Zn–23N | 83.8 |

[a] Distortion of up to ±11° seen from ideal $T_d$-sym. angle of 109.45°.
[b] Distortion of up to ±12° seen from ideal $D_2$-sym. angle of 123°.

conformers (within 1 kcal $mol^{-1}$ of each other),[33] which may interconvert via internal rotation. Any of the high symmetry ones should be used for the rotational entropy to be correct (we chose TTT). For $Zn(en)_2{}^{2+}$, the chelate rings can have R or S chirality, and hence, four low-energy conformers should exist ($D_2$-symmetry RR and SS enantiomers and $C_2$-symmetry RS and SR enantiomers) and possibly interconvert via NCCN internal rotations. Again, a high-symmetry one should be used (we chose RR), but with B3LYP/{SDD, 6-31+G(d)}/IEFPCM, we ran into the $C_1$-symmetry distorted-minimum problem (Table I). In addition, we also found a truly $D_2$-symmetry higher energy minimum for RR, higher by only 0.06 kcal $mol^{-1}$ than the distorted minimum we employed. This unused $D_2$-symmetry minimum featured four equal Zn–N bond lengths of 2.106 Å; the $C_1$-distorted minimum we used featured Zn–N bond lengths of 2.088–2.119 Å. We suspect that, at 298 K, both of these RR variants produce one "vibrationally averaged" $D_2$-symmetry structure. We used the lower energy ($C_1$-symmetry) minimum to obtain all-positive vibrational frequencies but the $D_2$ saddle point (lying between the two minima in energy; four equal Zn–N bond lengths of 2.095 Å) for the energies and rotational entropies. The existence of multiple conformers [seven for *en* and four for $Zn(en)_2{}^{2+}$] would contribute a small amount to entropy and free energy, which we did consider (Sec. IV B).

## III. INITIAL RESULTS

The initial results for the paradigmatic chelation reaction in Eq. (1), from B3LYP/{SDD, 6-31+G(d)}/IEFPCM geometry optimization and frequency evaluation, are

$$\begin{aligned}\Delta_{rxn}G^* &= \Delta_{rxn}E_{el,\,gas} + \Delta_{rxn}G_{freq}{}^{\circ} + \Delta_{rxn}\Delta G^{\circ\to *} + \Delta_{rxn}\Delta_{solv}G^* \\ &= -3.2 - 16.7 + 3.8 + 7.0 \\ &= -9.1\ \text{kcal mol}^{-1} \quad (\text{expt} = -2.3\ \text{kcal mol}^{-1}),\end{aligned} \tag{4}$$

$$\begin{aligned}\Delta_{rxn}S^* &= \Delta_{rxn}S_{freq}{}^{\circ} + \Delta_{rxn}S_{Trouton}{}^{\circ\to *} \\ &= +44.8 - 26 \\ &= +19\ \text{cal mol}^{-1}\text{K}^{-1} \quad (\text{expt} = +13 \pm 2\ \text{cal mol}^{-1}\ \text{K}^{-1}),\end{aligned} \tag{5}$$

$$\begin{aligned}\Delta_{rxn}H^* &= \Delta_{rxn}G^* + T\Delta_{rxn}S^* \\ &= -9.1 + (298.15)(0.019) \\ &= -3.4\ \text{kcal mol}^{-1} \quad (\text{expt} = +1.6\ \text{kcal mol}^{-1}).\end{aligned} \tag{6}$$

The predicted $\Delta_{rxn}S^*$ of 19 cal $mol^{-1}$ $K^{-1}$ is reasonable, slightly higher than the experimental value of 13 ± 2 cal $mol^{-1}$ $K^{-1}$ (and much better than the 78 cal $mol^{-1}$ $K^{-1}$ value obtained from a temperature-dependent Gibbs energy method, see the supplementary material). In energy ($T\Delta_{rxn}S^*$) terms, the predicted $\Delta_{rxn}S^*$ has an error of only +2 kcal $mol^{-1}$.

However, the predicted reaction Gibbs energy is in error by −7 kcal $mol^{-1}$, an error so large that the predicted $\Delta_{rxn}H^*$ value is negative instead of positive. This means that the calculation is attributing a significant portion of the chelate effect to an enthalpic benefit, which is not true. This basic ESM/CSM result does not reproduce the chelate effect as a purely entropic effect.

The remainder of this paper focuses only on ways to cure the −7 kcal $mol^{-1}$ error in $\Delta_{rxn}G^*$. We determine the contributions to this error arising from the three approximated terms in Eq. (4), finding these errors to be {+5, −2, −10} in {$\Delta_{rxn}E_{el,gas}$, $\Delta_{rxn}G_{freq}{}^{\circ}$, $\Delta_{rxn}\Delta_{solv}G^*$}, respectively. We locate methods accurate enough to reduce these errors sufficiently for the correct prediction of the chelate effect in Eq. (1), i.e., methods that predict $\Delta_{rxn}G^* < 0$ but $\Delta_{rxn}H^* > 0$. We then test some of these methods on their ability to reproduce $\Delta_{rxn}G^*$ for Eq. (1) but with $Cd^{2+}$ or $Hg^{2+}$ in place of $Zn^{2+}$.

## IV. IMPROVING THE RESULTS FOR $\Delta_{rxn}G^*$

### A. Improving $\Delta_{rxn}E_{el,gas}$

To improve upon B3LYP/{SDD, 6-31+G(d)} for this term, over 100 alternative electronic structure models (ESMs) were examined for not only the chelate effect reaction but also the two underlying complexation reactions,

$$Zn^{2+} + \text{ligands} \rightarrow \text{complex},$$

for $Zn(NH_3)_4{}^{2+}{}_{(gas)}$ and $Zn(en)_2{}^{2+}{}_{(gas)}$. An exemplary set of results appear in Fig. 1. The best results would be the [CCSD(T)]/aV∞Z values (rightmost circles), CCSD(T) = coupled cluster with single, double, and perturbative triple excitations, which were obtained in the following focal-point-style[34,35] manner. First, MP2/aV∞Z electronic energies for each species ($Zn^{2+}$, $NH_3$, *en*, and the two complexes) were obtained by extrapolating the aVXZ series (X = D, T, Q, i.e., 2, 3, 4), by plotting vs $1/X^2$ and taking the y-intercept (since $X \to \infty$ requires $1/X^2 \to 0$) of the fitted line). Then, the resulting MP2/aV∞Z reaction energies were shifted by correlation corrections $\delta$ = CCSD(T) − MP2, taken to be +0.3 [Fig. 1(a), chelate–effect reaction, value taken from aVDZ data], +1.5 [Fig. 1(b), $Zn(NH_3)_4{}^{2+}$ complexation, value taken from aVTZ data], and +1.8 [Fig. 1(c), $Zn(en)_2{}^{2+}$ complexation, value taken as the sum of the other two due to reaction cycle closure].

For the chelate–effect reaction [Fig. 1(a)], we see that our original $\Delta_{rxn}E_{el,gas}$ = −3.2 value is in error by +5 kcal $mol^{-1}$. The original −3.2 value, from B3LYP/{SDD, 6-31+G(d)} calculation, is the leftmost plus (+) point in Fig. 1(a); the best value of −8.2 is the rightmost ○ point in Fig. 1(a).

Sufficient ESMs for sub 2 kcal $mol^{-1}$ accuracy in this $\Delta_{rxn}E_{el,gas}$ term would be MP2 or CCSD(T) paired with "converged" basis sets, which from our full set of data (some not shown) appears to be (i) any Zn-atom basis set beyond SDD, (ii) any H-atom basis set containing polarization functions, and (iii) any "first-row-atom" (second-period-atom) basis set having diffuse functions and at least two sets of *d* functions.

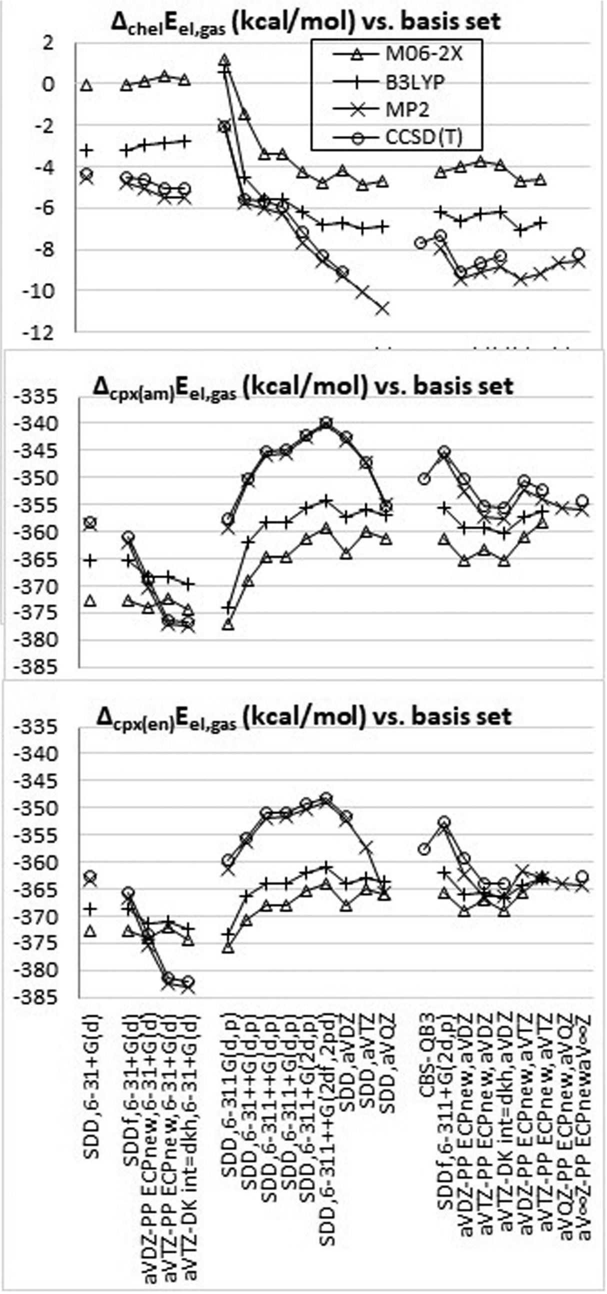

**FIG. 1.** Computed values of the electronic term $\Delta_{rxn}E_{el,gas}$ (kcal mol$^{-1}$) for the chelation reaction [Eq. (1)], and two underlying complexation reactions, plotted vs {Zn, CNH} orbital basis set. From single-point calculations using B3LYP/{SDD, 6-31+G(d)}/gas geometries. The best ("converged") result is the rightmost circle. See the supplementary material for method and basis-set references.

The individual complexation reactions [Figs. 1(b) and 1(c)] have more stringent basis set requirements for sub 2 kcal mol$^{-1}$ accuracy. They are more sensitive to metal/ligand basis set balance. The simplest ESM here that achieves this accuracy is MP2/{aVDZ-PP ECPnew, aVDZ}, and therefore, this particular ESM will be the one chosen for procedure testing (Sec. V).

### B. Improving $\Delta_{rxn}G_{freq}{}^{\circ}$

In this, the second term in Eq. (4), the most significant sources of error would be in the treatment of contributions from solute vibrational modes. We first focus on the internal rotation modes and then the non-internal-rotation modes.

The chelation reaction in Eq. (1) involves the loss of ten internal rotation modes [four of $Zn(NH_3)_4{}^{2+}$ and three of each of the two *en*]. None appear in the products. Hence, one possible inaccuracy is in the assumption of harmonic oscillators for the internal rotations of both *en* and $Zn(NH_3)_4{}^{2+}$, leaving these too "rigid" and too high in Gibbs energy, contributing to the overly negative $\Delta_{rxn}G$ prediction. To apply improved entropy modeling for internal rotations, we considered E1 theory,[36] where harmonic-oscillator entropies are replaced by free-rotor entropies [Eq. (7)] if the internal-rotation barrier height is <1.4 RT (roughly 0.8 kcal mol$^{-1}$),

$$S\left(Q_f\right) = \frac{1}{2}R \ln\left[\frac{e\pi}{\sigma^2\theta_{rot}}\right] + \frac{1}{2}\mathrm{R} \ln T, \tag{7}$$

where $S\left(Q_f\right)$ is the free rotor approximation to entropy S, $R$ = 1.987 cal mol$^{-1}$ K$^{-1}$, $\sigma$ is the symmetry number for an ammonia rotor (=3), and $\theta_{rot}$ is the rotational temperature for the $NH_3$ rotor in Kelvin, here taken to be equal to the computed overall rotational temperature of $NH_4{}^+$. To determine if this is needed, we first optimized internal-rotation transition-state geometries to determine which internal-rotation barrier heights lie below 0.8 kcal mol$^{-1}$ (Table II).

For $Zn(NH_3)_4{}^{2+}$, the internal-rotation barrier height is quite small, so improvement is possible, and under the E1 entropy improvement,[36] we would replace the four internal-rotation entropies (cal mol$^{-1}$ K$^{-1}$) from the harmonic oscillator default (3.341, 3.295, 3.272, and 2.880) with free-rotor entropies (4 × 3.47). The net effect upon $\Delta_{rxn}S$ and $\Delta_{rxn}G$, however, was rather negligible ($\Delta\Delta_{rxn}S$ = −1.1 cal mol$^{-1}$ K$^{-1}$ and $\Delta\Delta_{rxn}G$ = +0.3 kcal mol$^{-1}$).

**TABLE II.** Internal rotation barrier computation ($E_a$), B3LYP/{SDD, 6-31+G(d)}/IEFPCM.

| Internal rotation | $E_{minimum}$ (a.u.) | $E_{ts}$ (a.u.) | $E_a$ (kcal/mol) |
|---|---|---|---|
| $Zn(NH_3)_4{}^{2+}$ | −453.226 99 | −453.226 54 | 0.3 |
| en (C-C) | −190.532 06 (TTT) | −190.523 90 (TCT)[a] | 5.1 |
| en (C-C) | −190.532 06 (TTT) | −190.526 17 (TET)[a] | 3.7 |
| en (C-N) | −190.532 06 (TTT) | −190.527 74 (TTG′)[b] | 2.7 |
| en (C-N) | −190.531 84 (TGG′) | −190.526 33 (GG′G)[b] | 3.5 |

[a] TCT and TET refer to the cis (ΦNCCN = 0°) and eclipsed (ΦNCCN ≈ 120°) stationary points of *en*.

[b] TTG′ and GG′G have $E_{rel}$'s determined relative to the nearby TTT and TGG′ minima,[33] respectively.

For ethylenediamine, the internal-rotation barrier heights are sufficiently high that harmonic oscillators should work well (for a given conformer), as was found for ethylamine.[36] However, there is then the issue of its seven predicted conformers [and the four predicted conformers of $Zn(en)_2^{2+}$] predicted to be present (see Sec. II). In these cases where internal-rotation interconversion is not facile, the addition of a term for the entropy and Gibbs energy of mixing of conformers would be appropriate

$$\begin{aligned} S_{mix} &= -R\,\Sigma\, p_i \ln p_i \\ &\approx +R \ln \omega, \end{aligned} \tag{8}$$

$$G_{mix} = -T\,S_{mix}, \tag{9}$$

where $p_i = e^{-Ei/kT}$ is the probability (Boltzmann weight) that a given molecule is of conformer i, and the simple approximation $R \ln \omega$ arises by assuming $\omega$ equal-energy conformers. Use of the simple approximation would result in the following "mix" corrections to be added to "freq" terms:

$$\begin{aligned} S_{mix}(en) &= R \ln 7 = 3.87\ \text{cal mol}^{-1}\ \text{K}^{-1}, \\ S_{mix}(Zn(en)_2^{2+}) &= R \ln 4 = 2.76\ \text{cal mol}^{-1}\ \text{K}^{-1}, \\ \Delta_{rxn}S_{mix} &= 2.76 - 2\,(3.87) = -5.0\ \text{cal mol}^{-1}\ \text{K}^{-1}, \end{aligned} \tag{10}$$

$$\begin{aligned} G_{mix}(en) &= (-0.298\,15)\ S_{mix}(en) = -1.15\ \text{kcal mol}^{-1}, \\ G_{mix}(Zn(en)_2^{2+}) &= (-0.298\,15)\ S_{mix}(Zn(en)_2^{2+}) = -0.82\ \text{kcal mol}^{-1}, \\ \Delta_{rxn}G_{mix} &= (-0.298\,15)\ \Delta_{rxn}S_{mix} = +1.5\ \text{kcal mol}^{-1}. \end{aligned} \tag{11}$$

These magnitudes of $\Delta_{rxn}S_{mix}$ and $\Delta_{rxn}G_{mix}$ [Eqs. (10) and (11)] are significant enough that we believe these to be useful (i.e., non-negligible) corrections to improve the initial predictions of $\Delta_{rxn}S_{freq}$ and $\Delta_{rxn}G_{freq}$ in Eqs. (3) and (4).

As for the contributions to $\Delta_{rxn}G_{freq}^{\circ}$ from *non*-internal-rotation vibrational modes, there is certainly uncertainty in the prediction due to the harmonic approximation and the approximate nature of the B3LYP/{SDD, 6-31+G(d)} force constants, but it is not known how to improve this *ab initio*. A variety of estimates (−16.2 ± 1.1 kcal mol$^{-1}$, Table III) are obtained merely by using other ESM/CSM combinations. The MP2 predictions differ from B3LYP ones (roughly 1.5 kcal mol$^{-1}$ less negative), even after recommended[37] ESM-specific scaling factors are applied for the harmonic frequencies.

Hence, we consider the initial value of −16.7 for $\Delta_{rxn}G_{freq}^{\circ}$ [in Eq. (4)] to be in error by −2 ± 2 kcal mol$^{-1}$ due to neglect of the conformer uncertainty effect of $\Delta_{rxn}G_{mix} = +1.5$ and a shift of +0.5 ± 2 for consideration of the additional predictions in Table III.

At this point, if we use the improved values for $\Delta_{rxn}E_{el,gas}$ and $\Delta_{rxn}G_{freq}^{\circ}$ in Eq. (4), we obtain $\Delta_{rxn}G^{*} = -8.2 - 14.7 + 3.8 + 7.0 = -12.1$ kcal mol$^{-1}$ (expt = −2.3 kcal mol$^{-1}$). Since the third term has no error in it, the fourth term, $\Delta_{rxn}\Delta_{solv}G^{*}$, should have been +17 instead of the predicted +7 and hence was in error by −10 kcal$^{-1}$ mol$^{-1}$. This solvation error is due to the continuum solvation model used (Gaussian09's default, IEFPCM with UFFx1.1 radii). Subsections IV C–IV F all explore ways to reduce this solvation error.

**TABLE III.** Results for $\Delta_{rxn}G_{freq}^{\circ}$ from various ESMs.

| ESM | $\Delta_{rxn}G_{freq}^{\circ}$ (kcal mol$^{-1}$) |
|---|---|
| B3LYP/{SDD, 6-31+G(d)}/IEFPCM[a,b] | −16.7 |
| B3LYP/{SDD, 6-31+G(d)}/gas | −17.2 |
| B3LYP/{SDD, 6-311+G(d,p)}/gas | −16.8 |
| B3LYP/{SDD, aVDZ}/gas | −16.9 |
| B3LYP/{aVDZ-PP ECPnew, aVDZ}/gas | −16.8 |
| ScaledB3LYP/{aVDZ-PP ECPnew, aVDZ}/gas[c] | −16.6 |
| MP2/{SDD, 6-31+G(d)}/gas[a] | −15.7 |
| MP2/{SDD,6-311+G(d,p)}/gas[a] | −15.7 |
| MP2/{SDD, aVDZ}/gas | −15.5 |
| MP2/{aVDZ-PP ECPnew, aVDZ}/gas | −15.5 |
| ScaledMP2/{aVDZ-PP ECPnew, aVDZ}/gas[c] | −15.1 |

[a] $Zn(NH_3)_4^{2+}$ $T_d$-symmetry optimization produced imaginary frequencies; required $C_1$-symmetry optimization as well (for vibrational contribution).

[b] $Zn(en)_2^{2+}$ $D_2$-symmetry optimization produced imaginary frequencies; required $C_1$-symmetry optimization as well (for vibrational contribution).

[c] Employed vibrational frequency scale factors[37] of 0.97 (B3LYP) or 0.96 (MP2) for all four species in the reaction.

### C. Improving $\Delta_{rxn}\Delta_{solv}G^{*}$ via semicontinuum modeling?

In Sec. I, we noted the two types of CSM error as described by Patel and East:[4] q/r error and ics error. If this 10 kcal mol$^{-1}$ error is largely a q/r error, it could be reduced with semicontinuum modeling—the addition of explicit water molecules inside the continuum. We tried this (Table IV and Fig. 2) following our usual procedure[2–4,38–40] of ensuring that the added number of hydrogen bonds is identical for reactants vs products: the {8, 12, 20, 28} W models introduce {10, 14, 24, 40} H-bonds, respectively. The results stay within 1.1 kcal mol$^{-1}$ of the values from the regular continuum modeling, and thus, explicit waters hardly affect results here. Thus, the q/r error of the default CSM is not an issue: the two ions in the reaction, $Zn(NH_3)_4^{2+}$ and $Zn(en)_2^{2+}$, must be large enough that this particular error is either negligible or happened to cancel well before explicit waters were added. (This is in contrast to acid-catalyzed activation energies, where the large q/r error made on $H_3O^+$ is not canceled appreciably by the smaller error made on the cationic transition state complex.[4])

Therefore, the 10 kcal mol$^{-1}$ solvation-energy error made by B3LYP/{SDD, 6-31+G(d)}/IEFPCM, which must be fairly constant at each level of hydration in Table IV (given the constancy of $\Delta_{rxn}G_{el,\varepsilon}$ results), is due to a differential imperfect-cavity-size error made by the default CSM, and though such errors individually grow as explicit hydration is increased, the differential error is constant, meaning it exists for this reaction even in the absence of explicit hydration. It is the CSM itself that requires improvement, and we explore this in Subsections IV D–IV F of this paper.

### D. Improving $\Delta_{rxn}\Delta_{solv}G$ via alternative CSMs?

We examined the performance of other continuum solvation models (CSMs), upon not only $\Delta_{rxn}\Delta_{solv}G$ but also each individual solute $\Delta_{solv}G$ value. Recently, Patel and East gained valuable insight into CSM errors by examining solvation energies of

**TABLE IV.** Semicontinuum B3LYP/{SDD, 6-31+G(d)}/IEFPCM energies $G_{el,\varepsilon}$ (a.u.) and ${}_{rxn}G_{el,\varepsilon}$ (kcal mol$^{-1}$) for the chelation reaction [Eq. (1)].

| Model | $G_{el,\varepsilon}$ [$Zn(NH_3)_4{}^{2+}$] | $G_{el,\varepsilon}$ [*en*] | $G_{el,\varepsilon}$ [$Zn(en)_2{}^{2+}$] | $G_{el,\varepsilon}$ [$NH_3$] | $\Delta_{rxn}G_{el,\varepsilon}$ |
|---|---|---|---|---|---|
| 0W, *en* = TTT | −453.226 82 | −190.532 06 | −608.034 31 | −56.562 66 | 3.8 |
| 0W, *en* = TGT | −453.226 82 | −190.531 83 | −608.034 31 | −56.562 66 | 3.5 |
| 0W, *en* = GGG′ | −453.226 82 | −190.531 53 | −608.034 31 | −56.562 66 | 3.1 |
| 0W, *en* = TGG′ | −453.226 82 | −190.531 84 | −608.034 31 | −56.562 66 | 3.5 |
| 8W, *en* = TTT | −758.984 11 | −343.414 96 | −913.791 24 | −133.003 93 | 4.4 |
| 8W, *en* = TGT | −758.984 11 | −343.414 72 | −913.791 24 | −133.003 93 | 4.1 |
| 12W, *en* = TTT | −758.984 11 | −496.280 86 | −913.791 24 | −209.436 85 | 4.5 |
| 12W, *en* = TGT | −758.984 11 | −496.280 51 | −913.791 24 | −209.436 85 | 4.1 |
| 20W, *en* = TTT | −1064.738 00 | −649.145 69 | −1219.544 05 | −285.869 40 | 4.9 |
| 28W, *en* = TGT | −1676.233 86 | −649.158 04 | −1219.544 05 | −438.750 18 | 3.2 |

some individual species.[4] This led us to want to break down the 10 kcal mol$^{-1}$ error made on $\Delta_{rxn}\Delta_{solv}G$ into the errors made on the $\Delta_{solv}G$ solvation energies of the four individual compounds involved. Since experimental solvation Gibbs energies for the two Zn complexes were unknown, we first set about to derive some best estimates ($\Delta_{solv}G^*\{Zn(NH_3)_4{}^{2+}\}$ = −196 ± 3 kcal mol$^{-1}$ and $\Delta_{solv}G^*\{Zn(en)_2{}^{2+}\}$ = −177 ± 3 kcal mol$^{-1}$; see the supplementary material).

We then computed the solvation energies of the four species of Eq. (1) with many different CSM choices (Table V). Due to our original focus on Gaussian-code methods, several IEFPCM[41] lines are given, exploring the effects of the ESM (lines 5–7; quite minor) and the cavity radii set (lines 7, 9, 13; substantial). CPCM[42] (Conductor PCM) is the slightly faster conductor version of IEFPCM, revealing essentially identical answers, and two other variants (IEFPCM+CRD[43,44] and SMD[12]) add non-electrostatic terms to the IEFPCM purely electrostatics result (CRD = cavitation-repulsion-dispersion, SMD = Solvation Model for Densities). Several CSMs not in the Gaussian code were also tested: SM12,[45,46] COSMO (Conductor-like Screening Model),[47,48] COSMO-RS (Realistic Solvation),[49,50] SCCS (Self-Consistent Continuum Solvation model),[51] and CANDLE (Charge-Asymmetric Nonlocally Determined Local-Electric model).[52]

The difficulty of CSMs to accurately predict the solvation energies of the coordination complexes (−196 ± 3 and −177 ± 3 kcal mol$^{-1}$) is immediately seen in the second and fourth columns, with many examples of very large (>15 kcal mol$^{-1}$) errors for the cationic coordination complexes. In fact, the default CSM we had employed, IEFPCM(UFFx1.1), offered the best and second best predictions for $\Delta_{solv}G^*$ for $Zn(en)_2{}^{2+}$ and $Zn(NH_3)_4{}^{2+}$, respectively.

It could still be that the alternatives might have improved error cancellation and predicted $\Delta_{rxn}\Delta_{solv}G$ well for the chelate–effect reaction. Sadly, most of these CSMs, including the default IEFPCM(UFFx1.1), predict values for $\Delta_{rxn}\Delta_{solv}G$ in a narrow range of 5–11 kcal mol$^{-1}$ (rightmost column), a considerable underestimate of the true value of 17. This common problem is due to underestimating "gap1," the 19 kcal mol$^{-1}$ difference in solvation energy of the two Zn complexes, $\Delta_{solv}G[Zn(en)_2{}^{2+}] - \Delta_{solv}G[Zn(NH_3)_4{}^{2+}]$. The CSM methods that produced abnormal $\Delta_{rxn}\Delta_{solv}G$ values were SCCS (−17, far too low), COSMO-RS (+21, too large), and SM12 (+16, the only excellent value). SM12 happens to obtain the right value, despite large individual errors (−15 and −18 kcal mol$^{-1}$) on the solvation energies of the large ions. We included SM12 as a possibility for further testing (Sec. V).

We are aware of two previous CSM-based computations of the chelate effect. In 2003, Vallet *et al.*[53] used an early CPCM version (in Gaussian98) on the complex formation reactions for $Zn(NH_3)_2{}^{2+}$(aq) and $Zn(en)^{2+}$(aq). They obtained better error cancellation from 2 → 1 stoichiometry ($MW_6{}^{2+} + L \rightarrow \{MLW_5\}W^{2+}$, their model A), rather than balanced 2 → 2 stoichiometry ($MW_6{}^{2+} + L \rightarrow MLW_5{}^{2+} + W$, their model B), with complex formation errors of ∼0 and ∼−4 kcal mol$^{-1}$ for $Zn(NH_3)_2{}^{2+}$(aq) and $Zn(en)^{2+}$(aq), respectively. Thus, had they computed ΔG for the difference of the two formations [i.e., for $M(NH_3)_2W_n{}^{2+} + en \rightarrow M(en)W_n{}^{2+} + 2\ NH_3$, for the actual chelate effect], they would have seen an error of −4 kcal mol$^{-1}$ per *en* added, similar to our CPCM result (Table V, −10 kcal mol$^{-1}$ error for 2 *en* added). In 2018, Sengupta *et al.*[54] published (in their supplementary material) modern DLPNO-CCSD(T)/SMD results (DLPNO = Domain-based Local Pair Natural Orbital) for similar formation reactions (Zn and Cd, with $MeNH_2$ and *en*) in a paper that focused more on adjusting modern molecular mechanics force fields in order to successfully reproduce the quantitative stabilities of $M(en)_n{}^{2+}$ and $M(NH_3)_n{}^{2+}$ aqueous complexes. Their DLPNO-CCSD(T)/SMD formation reactions also did better with 2 → 1 stoichiometry than 2 → 2 stoichiometry (though they confusingly called it Method B). However, in the reaction difference, for $Cd(NH_2Me)_2W_4{}^{2+} + en \rightarrow Cd(en)W_4{}^{2+} + 2\ MeNH_2$ (the chelate effect), their error was −7 kcal mol$^{-1}$ per *en* added, worse than our SMD (and Vallet's CPCM) results for $Zn(NH_3)_{2n}{}^{2+}$ + n *en*.

### E. Improving $\Delta_{rxn}\Delta_{solv}G$ via cavity radii re-optimization?

Table V clearly gives a large variance (among CSMs) in solvation energy predictions for the two transition-metal complexes, and SM12 did not perform well in this regard. There is still, we feel, a need for improved CSMs for large polyatomic ions in water.

One possibility for improvement is in determining more optimal cavity radii for transition metal complexes. Such complexes were not used in the original training sets for CSMs; for instance, the 112 aqueous ions used in the SMD training set had no metal atoms and only +1 or −1 charges.[12] Also, the large solvation energies of dications ($\Delta_{solv}G^*$ ∼ $-50q^2$ kcal mol$^{-1}$, where q = charge

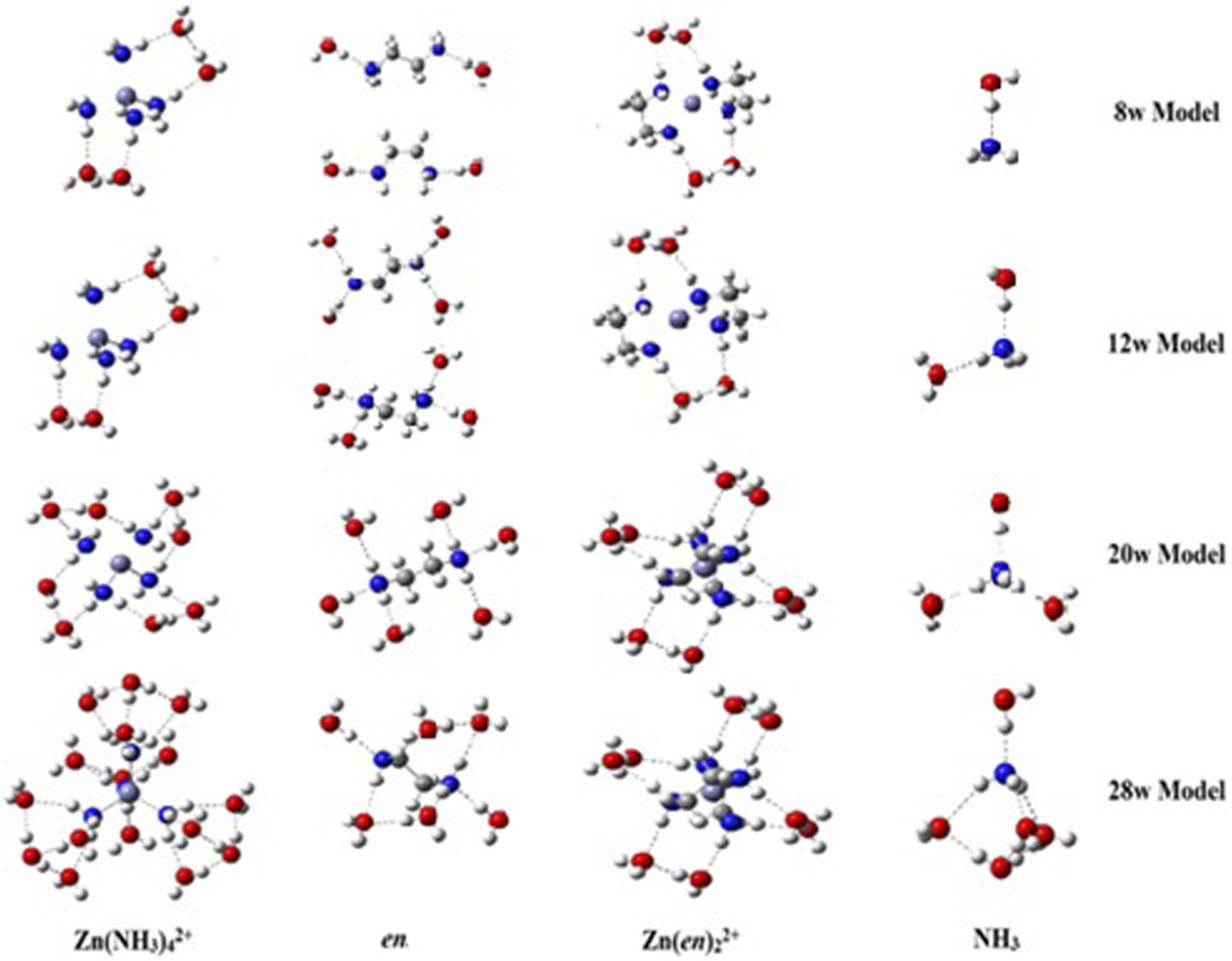

**FIG. 2.** Structures of 8W, 12W, 20W, and 28W semicontinuum models examined for the chelation reaction [Eq. (1)].

of solute) may make their solvation energies more sensitive to cavity size. Hence, we next explored re-optimization of the four cavity radii (Zn, C, N, and H) in hopes of reducing, if not eliminating, CSM solvation errors for the four solutes in the chelation reaction of interest.

We pursued re-optimization of the cavity radii for three Gaussian-code CSMs: IEFPCM,[41] IEFPCM+CRD,[43,44] and SMD,[12] paired with B3LYP/{SDD, 6-31+G(d)}. We set the electrostatic-term radii-scaling factor to be α = 1.0. [Note that the non-electrostatic CRD terms use unscaled radii always, and the non-electrostatic terms in SMD use a special set of radii (Bondi + 0.4 Å)[12] that cannot be user-adjusted.] The radii optimization was done manually via many single-point energy runs (no geometry reoptimization) with the same ESM and geometries as in Table V.

After some exploration, we settled on a species-by-species optimization procedure, working first to eliminate solvation errors for $NH_3$, en, and "gap2" (see Table V) and then the errors for $Zn(NH_3)_4^{2+}$, $Zn(en)_2^{2+}$, and "gap1." The explorations showed that generally solvation energies became more negative as cavity volumes decreased, and thus, to cure the gap1 underestimation by all CSMs with standard radii, one would need to shift some of the cavity volume away from $Zn(NH_3)_4^{2+}$ to $Zn(en)_2^{2+}$, by contracting cavities in NH-bond regions and/or expanding cavities in CH-bond regions.

Starting with $NH_3$, we found sets of $r_H$ and $r_N$ cavity radii that reproduced its known solvation energy of −4.3 kcal $mol^{-1}$ (Table VI). Then, moving to ethylenediamine (*en*), we kept these $\{r_H, r_N\}$ sets optimal for $NH_3$ and varied only $r_C$ until the experimental value of $\Delta_{solv}G(en) = -7.6$ kcal $mol^{-1}$ (Table V) was achieved. For IEFPCM (only electrostatic effects), all $r_C$ choices resulted in too high a prediction for $\Delta_{solv}G(en)$ [Fig. 3(a)]. Non-electrostatic terms (cavitation, dispersion, and solvent reorganization) must be needed. For IEFPCM+CRD, the error on *en* could be eliminated with reasonable $r_C$ values but only with small $r_H$ values [Fig. 3(b)]: the set $\{r_H = 1.00, r_N = 2.04, r_C = 2.01\}$ with α = 1.00 worked. For SMD that adds a different set of non-electrostatic terms to IEFPCM, the errors

**TABLE V.** Computed solvation Gibbs energies $\Delta_{solv}G$ (kcal mol$^{-1}$) testing various CSMs.

| CSM[a] | $Zn(NH_3)_4{}^{2+}$ | *en* | $Zn(en)_2{}^{2+}$ | $NH_3$ | gap1[b] | gap2[b] | $\Delta_{rxn}\Delta_{solv}G^*$ |
|---|---|---|---|---|---|---|---|
| CANDLE[c] | −137 | −8 | −118 | −6 | 19 | −9 | 11 |
| SCCS[c] | −116 | −6 | −128 | −4 | −12 | −5 | −17 |
| IEFPCM+CRD(UFFx1.1)[d] | −176 | 1 | −163 | −1 | 12 | −6 | 6 |
| CPCM(UFFx1.1)[d] | −190 | −6 | −179 | −4 | 11 | −4 | 7 |
| IEFPCM(UFFx1.1)[d] | −190 | −6 | −179 | −4 | 11 | −4 | 7 |
| IEFPCM(UFFx1.1)[e] | −191 | −6 | −179 | −3 | 12 | −2 | 10 |
| IEFPCM(UFFx1.1)[f] | −191 | −5 | −180 | −3 | 11 | −2 | 9 |
| COSMO(Klamt2)[g] | −199 | −8 | −184 | −5 | 14 | −5 | 10 |
| IEFPCM(Klamt)[d] | −201 | −10 | −187 | −6 | 14 | −6 | 8 |
| COSMO-RS[g] | −215 | −9 | −191 | −5 | 24 | −3 | 21 |
| SM12[g] | −211 | −8 | −195 | −4 | 16 | 0 | 16 |
| SMD[d] | −203 | −10 | −196 | −5 | 7 | 3 | 10 |
| IEFPCM(SMDx1.0)[d] | −212 | −11 | −202 | −7 | 10 | −6 | 5 |
| Expt.[h] | −196 ± 3 | −7.6 | −177 ± 3 | −4.3 | 19 | −2.0 | 17 |

[a] From single-point calculations, using B3LYP/{SDD, 6-31+G(d)}/gas optimized geometries. Terms in parentheses refer to cavity radii set used: UFF,[55] SMD,[12] "Klamt" (radii = Klamt) from Klamt and Eckert[56] plus r(Zn) = 1.9 × 1.17 = 2.223 Å, and "Klamt2" (COSMO radii in the ADF code)[31] from Klamt and Eckert plus r(Zn) = 1.908 Å. For radii values, see the supplementary material.

[b] gap1 = $\Delta_{solv}G[Zn(en)_2{}^{2+}] - \Delta_{solv}G[Zn(NH_3)_4{}^{2+}]$; gap2 = $4\ \Delta_{solv}G[NH_3] - 2\ \Delta_{solv}G[en]$.

[c] ESM = PBE+D2/{GBRV2014PP, planewaves}, JDFTx code.

[d] ESM = B3LYP/{SDD, 6-31+G(d)}.

[e] ESM = MP2/{aVDZ-PP ECPnew, aVDZ}.

[f] ESM = CCSD(T)/{aVDZ-PP ECPnew, aVDZ}.

[g] ESM = BP86/{TZP, relativistic = scalar, frozencore = small}, ADF code.

[h] $NH_3$ value from Marenich *et al.*,[12] *en* value from Cabani *et al.*,[57] cation values derived in the supplementary material.

on $NH_3$ and *en* could be eliminated for a variety of $r_H$ values but only when paired with very large $r_C$ values [$r_C$ = 2.8 Å, Fig. 3(c)].

Moving to $Zn(NH_3)_4{}^{2+}$, again we kept the {$r_H$, $r_N$} sets optimal for $NH_3$ and now varied the possible $r_{Zn}$ value when using IEFPCM+CRD or SMD (we abandoned IEFPCM on its own for failing for *en*). For IEFPCM+CRD, only a large $r_{Zn}$ value (2.99 Å) worked with the successful set so far ($r_H$ = 1.00). For SMD, reasonable values for $r_{Zn}$ (1.65–2.40 Å) worked with all the successful sets so far ($r_H$ = 1.1–1.4 Å).

At this point, we had optimal sets of {Zn, C, N, H} radii for each of IEFPCM+CRD and SMD, which produced zero error for the first three solutes. We tested them on the fourth solute, $Zn(en)_2{}^{2+}$ (Table VII). None of them worked well for this last $\Delta_{solv}G^*$. The IEFPCM+CRD model, with its only successful radii set so far, badly overpredicted the correct magnitude (−186 vs −177 kcal mol$^{-1}$). Thus, no radii set of traditional restriction (one value per elemental atom regardless of the environment) can make IEFPCM-CRD work for all four solutes simultaneously. As for SMD, two radii sets ($r_H$ = 1.3 or 1.4 Å) get within 3.5 kcal mol$^{-1}$ of the target value (which itself has an uncertainty of perhaps 3 kcal mol$^{-1}$), so these sets might seem promising. However, the problem here is the very large $r_C$ value (2.8 Å): when we tested the $r_H$ = 1.3 Å set on other solutes (methylamine and ethylamine, $\Delta_{solv}G^*{}_{expt}$ = −4.53 ± 0.03 kcal mol$^{-1}$), the resulting solvation energies were too high (−2.7 ± 0.2). Thus, although the use of such a large $r_C$ would greatly reduce the error in SMD for our paradigm chelation reaction, it will not generally work well.

We thought of introducing a fifth degree of freedom, beyond the four cavity radii re-optimized. With SMD, to alleviate the large $r_C$ problem, we tried separate $r_{H,polar}$ and $r_{H,nonpolar}$ (for CH vs NH bonds),[58] and though this did allow for smaller and more reasonable $r_C$ values, it did so only by using large $r_{H,nonpolar}$ values ~2.6 Å, still resulting in overly large $CH_2$ cavities that would cause SMD to fail for ordinary amines. With IEFPCM+CRD, to alleviate the still-remaining $Zn(en)_2{}^{2+}$ problem, we tried varying the radii scale factor α (which affects the radii for the electrostatic term only) and re-optimizing all radii for each α, but the large +9 kcal mol$^{-1}$ error made on $Zn(en)_2{}^{2+}$ could only be reduced to +8 (using α = 0.90). Introducing a separate scale factor for ions vs neutrals would not

**TABLE VI.** Sets of cavity radii ($r_H$ and $r_N$, in Å) that reproduce the experimental $\Delta_{solv}G$ of ammonia for three continuum solvation models.[a]

| H | $N^{opt}_{IEFPCM}$ | $N^{opt}_{IEFPCM+CRD}$ | $N^{opt}_{SMD}$ |
|---|---|---|---|
| 1.00 | 2.22[b] | 2.04[b] | 2.03[b] |
| 1.10 | 2.22[b] | 2.02 | 2.01 |
| 1.20 | 2.215 | 1.95 | 1.93 |
| 1.30 | 2.20 | 1.81 | 1.79 |
| 1.40 | 2.13 | 1.51 | 1.62 |
| 1.50 | 2.05 | … | … |
| 1.60 | 1.96 | … | … |

[a] Using B3LYP/{SDD, 6-31+G(d)}/CSM//B3LYP/{SDD, 6-31+G(d)}/gas.

[b] In these four cases, $r_N > r_H + R_{NH}$, since the bond length $R_{NH}$ = 1.02 Å; this means that the solute sits inside a single sphere of radius $r_N$.

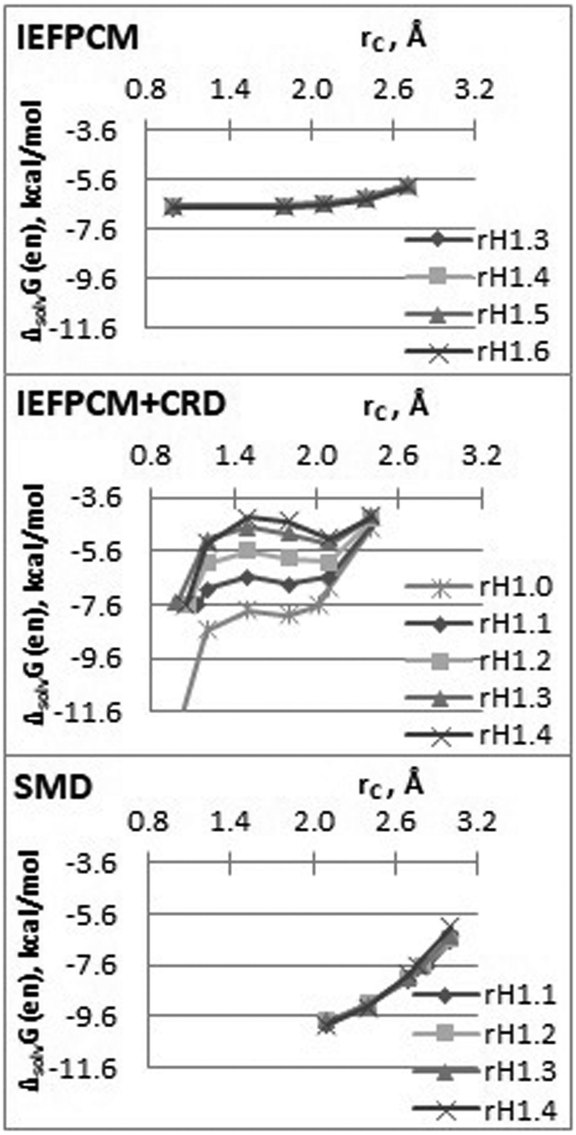

**FIG. 3.** Plots of CSM predictions (B3LYP/{SDD, 6-31+G(d)}/CSM//B3LYP/{SDD, 6-31+G(d)}/gas) of $\Delta_{solv}G(en)$, used to determine optimal carbon cavity radius values $r_C$ ones giving $\Delta_{solv}G(en) = -7.6$ kcal mol$^{-1}$: IEFPCM (top), IEFPCM+CRD (middle), and SMD (bottom). The $r_N$ values were fixed as per Table VI.

solve the gap1 problem (obtaining equal accuracy in $\Delta_{solv}G$ for both $Zn^{2+}$ complexes).

Hence, the accuracy problems of these CSMs for this chelate–effect reaction do not seem to be easily solvable by cavity radii alteration.

### F. Improving $\Delta_{rxn}\Delta_{solv}G^*$ via *post facto* corrections?

Finally, we pondered a *post facto* correction method for chelate–effect reactions, as was proposed in Options B1 and B2 of Patel and East for acid-catalysis modeling.[4] Using the data in Table V, one can determine errors in $\Delta_{solv}G^*$ made by each ESM/CSM combination on each solute, and invert their signs to consider them as *post facto* corrections to ESM/CSM data for chemically similar solutes; in our case, we will propose that such corrections for $M(en)_2{}^{2+}$ and $M(NH_3)_4{}^{2+}$ might be independent of metal-atom identity. Examples for four such ESM/CSM combinations appear in Table VIII. We test the second and third sets in Sec. V.

**TABLE VII.** Results (kcal mol$^{-1}$) for $\Delta_{solv}G^*[Zn(en)_2{}^{2+}]$ using cavity radii sets (Å) that produced zero error for the other three solutes.[a]

| CSM | $r_H$ | $r_N$ | $r_{Zn}$ | $r_C$ | $\Delta_{solv}G[Zn(en)_2{}^{2+}]$ |
|---|---|---|---|---|---|
| IEFPCM+CRD | 1.00 | 2.04 | 2.990 | 2.01 | −185.8 |
| SMD | 1.10 | 2.01 | 2.400 | 2.84 | −171.5 |
| SMD | 1.20 | 1.93 | 1.989 | 2.80 | −173.1 |
| SMD | 1.30 | 1.79 | 1.732 | 2.79 | −173.7 |
| SMD | 1.40 | 1.62 | 1.650 | 2.76 | −173.6 |
| Expt.[b] | | | | | −177 |

[a] Using B3LYP/{SDD, 6-31+G(d)}/CSM//B3LYP/{SDD, 6-31+G(d)}/gas.
[b] See Table V.

**TABLE VIII.** *Post facto* corrections for improving CSM $\Delta_{solv}G^*$ predictions.

| Solute | SMD (SMDx1.0)[a] | IEFPCM (UFFx1.1)[a] | IEFPCM (UFFx1.1)[b] | IEFPCM (UFFx1.1)[c] |
|---|---|---|---|---|
| $NH_3$ | 0.2 | −0.1 | −1.1 | −1.3 |
| *en* | 2.8 | −1.2 | −2.0 | −2.4 |
| $M(NH_3)_4{}^{2+}$ | 7.3 | −6.1 | −5.0 | −5.1 |
| $M(en)_2{}^{2+}$ | 19.5 | 2.1 | 2.6 | 2.8 |

[a] ESM = B3LYP/{SDD, 6-31+G(d)}.
[b] ESM = MP2/{aVDZ-PP ECPnew, aVDZ}.
[c] ESM = CCSD(T)/{aVDZ-PP ECPnew, aVDZ}.

## V. TESTING FOUR IMPROVED PROCEDURES

Here, we test four procedures now expected to perform ideally, based on what was learned in Sec. IV. We apply each of them to three cases of the chelate effect reaction in Eq. (1), varying the metal atom ($M = Zn^{2+}$, $Cd^{2+}$, $Hg^{2+}$). Recall that Eq. (4) calculation for $\Delta_{rxn}G^*$ involves choices for computing the three terms $\Delta_{rxn}E_{el,gas}$, $\Delta_{rxn}G_{freq}{}^{\circ}$, and $\Delta_{rxn}\Delta_{solv}G^*$. Our first three procedures for testing were the following:

Procedure A (three runs per solute):

- B3LYP/{SDD, 6-31+G(d)}/gas opt + freq [for $\Delta_{rxn}G_{freq}{}^{\circ}$],
- B3LYP/{SDD, 6-31+G(d)}/IEFPCM single points [for $\Delta_{rxn}\Delta_{solv}G^*$],
- MP2/{aVDZ ECPnew, aVDZ}/gas single points [for $\Delta_{rxn}E_{el,gas}$],

where $\Delta_{rxn}G_{freq}{}^{\circ}$ includes $\Delta_{rxn}\Delta_{mix}G = +1.5$ [Eq. (11)] and $\Delta_{rxn}\Delta_{solv}G^*$ includes the appropriate *post facto* corrections (Table VIII, second of four data columns).

Procedure B (two runs per solute):

- MP2/{aVDZ ECPnew, aVDZ}/gas opt + freq [for $\Delta_{rxn}G_{freq}{}^{\circ}$ and for $\Delta_{rxn}E_{el,gas}$],

**TABLE IX.** Chelate-effect $\Delta_{rxn}G^*$ and Eq. (4) components (kcal mol$^{-1}$) from improved procedures.

| Reaction | Procedure[a] | $\Delta_{rxn}E_{elec,gas}$ | $\Delta_{rxn}G_{freq}{}^{\circ}$ | $\Delta_{rxn}G^{\circ\to *}$ | $\Delta_{rxn}\Delta_{solv}G^*$ | $\Delta_{rxn}G_{total}{}^*$ |
|---|---|---|---|---|---|---|
| $Zn(NH_3)_4{}^{2+}$ + 2 *en* | A | −9.4 | −15.7 | 3.8 | 17.3 | −4.1 |
| | B | −9.9 | −14.0 | 3.8 | 17.6 | −2.4 |
| | C | −9.9 | −15.7 | 3.8 | 16.3 | −5.5 |
| | B2[b] | −9.9 | −14.0 | 3.8 | 17.6 | −2.4 |
| | Expt.[c] | | | | | −2.3 |
| $Cd(NH_3)_4{}^{2+}$ + 2 *en* | A | −7.7 | −15.1 | 3.8 | 11.8 | −7.1 |
| | B | −8.1 | −12.9 | 3.8 | 12.1 | −5.1 |
| | C | −8.1 | −15.1 | 3.8 | 14.0 | −5.4 |
| | B2[b] | −8.1 | −12.9 | 3.8 | 14.6 | −2.5 |
| | Expt.[c] | | | | | −4.4 |
| $Hg(NH_3)_4{}^{2+}$ + 2 *en* | A | −15.1 | −15.0 | 3.8 | 10.1 | −16.2 |
| | B | −16.0 | −12.3 | 3.8 | 11.8 | −12.8 |
| | C | −16.0 | −15.0 | 3.8 | 14.0 | −13.2 |
| | B2[b] | −16.0 | −12.3 | 3.8 | 16.5 | −8.1 |
| | Expt.[c] | | | | | −5.9 |

[a] See the text.
[b] Using Bondi's later[60] cavity radii for Cd; 1.62 Å, and Hg; 1.70 Å (with ×1.1 additional scaling).
[c] Expt. data for Zn from Spike and Parry,[26] for Cd and Hg from Martell and Smith.[27]

- MP2/{aVDZ ECPnew, aVDZ}/IEFPCM single points [for $\Delta_{rxn}\Delta_{solv}G^*$],

where $\Delta_{rxn}G_{freq}{}^{\circ}$ again includes $\Delta_{rxn}\Delta_{mix}G = +1.5$ kcal mol$^{-1}$ and $\Delta_{rxn}G_{el,\varepsilon}{}^*$ includes the appropriate *post facto* corrections (Table VIII, third of four data columns).

Procedure C (three runs per solute):

- B3LYP/{SDD, 6-31+G(d)}/gas opt + freq [for $\Delta_{rxn}G_{freq}{}^{\circ}$],
- BP86/TZP, relativistic = scalar, frozencore = small/SM12 single points [for $\Delta_{rxn}\Delta_{solv}G^*$],
- MP2/{aVDZ ECPnew, aVDZ}/gas opt [for $\Delta_{rxn}E_{el,gas}$],

where $\Delta_{rxn}G_{freq}{}^{\circ}$ again includes $\Delta_{rxn}\Delta_{mix}G = +1.5$ kcal mol$^{-1}$ but $\Delta_{rxn}G_{el,\varepsilon}{}^*$ includes no *post facto* correction.

These three procedures were applied to the chelate-effect reactions [Eq. (1)] of $Zn^{2+}$, $Cd^{2+}$, and $Hg^{2+}$ (Table IX), hoping for 2 kcal mol$^{-1}$ agreement with experimental $\Delta_{rxn}G^*$ values (so that predicted $\Delta_{rxn}H^*$ values for the Zn reaction would be positive, see Sec. III). Such agreement was achieved for the Zn reaction by procedures A and B, and for the Cd reaction by procedures B and C. However, all three procedures disagreed badly with the experiment for the Hg case, predicting greater chelation benefit (−13 to −16 kcal mol$^{-1}$) than was found experimentally (−6 kcal mol$^{-1}$).

The Hg problem was resolved: it was due to poor cavity radius values for $Hg^{2+}$ from IEFPCM (UFFx1.1 = 1.4883 Å) and the ADF version of SM12 (Bondi1 radius = 1.55 Å; see the supplementary material for a summary table of radii with references). The very small UFF (Universal Force Field)[55] choice for Hg is due to (i) it being for two-coordinate Hg (not four-coordinate) and (ii) the metal atom radii being derived from electronegativities (Hg is particularly electronegative). The somewhat small choice for Hg in ADF's SM12 is due to its choice of Zn/Cd/Hg radii values from the original Bondi paper (Bondi1[59]). Bondi, however, had a follow-up paper (Bondi2[60]), with improved radii choices for Cd and Hg. We repeated Procedure B, but for the solvation run, we replaced the UFFx1.1 radii with Bondi2x1.1 radii, and this cured the problem: with Procedure B2, all three reactions have their $\Delta_{rxn}G$ predicted to within 2.2 kcal mol$^{-1}$ of the experiment.

The apparent success of Procedure B2 for these three very similar reactions does not obviate the need for more research into improving CSMs for transition metal complexes. The procedure relies on *post facto* corrections for its success, corrections that were tailor-made for $M(NH_3)_4{}^{2+}$ and $M(en)_2{}^{2+}$ complexes; more research would be required to find systematic ways of applying *post facto* corrections for other transition-metal complexes. SM12's error cancellation success is encouraging, and an updated $Hg^{2+}$ cavity radius for SM12 would likely help, but SM12 should be tested on more ligand–exchange reaction types to see if such success will be commonplace.

## VI. CONCLUSIONS

Typical quantum-chemistry continuum solvation models (CSMs), conventionally computed, generally (with the exception of ADF's[31] SM12,[45,46] see below) fail to qualitatively reproduce the chelate effect for the paradigmatic reaction $Zn(NH_3)_4{}^{2+}{}_{(aq)} + 2\ en_{(aq)} \rightarrow Zn(en)_2{}^{2+}{}_{(aq)} + 4\ NH_{3(aq)}$. The B3LYP/{SDD, 6-31+G(d)}/IEFPCM(UFFx1.1) initial model predicts $\Delta_{rxn}G = -9$ kcal mol$^{-1}$ (Sec. III), which is 7 kcal mol$^{-1}$ too low, and the resulting $\Delta H$ of −3.5 kcal mol$^{-1}$ has the wrong sign, which would lead one to incorrectly claim that the reaction is driven by enthalpy as well as entropy.

Sources of the error from this representative model were pursued and found to be +5 kcal mol$^{-1}$ in the gas-phase electronic energies, −2 kcal mol$^{-1}$ in the thermal corrections, and −10 kcal mol$^{-1}$ in the solvation energies. The first two problems can be cured by (i) switching the electronic structure model (ESM) to MP2/{aVDZ-PP ECP = new, aVDZ} and (ii) employing $-T\Delta_{mix}S$ corrections of $-RT \ln 7$ for *en* and $-RT \ln 4$ for $M(en)_2^{2+}$, for entropy of mixing of conformers.

The third and largest problem was in the solvation Gibbs energy prediction, $\Delta_{rxn}\Delta_{solv}G^*$. Four attempted cures were explored: semicontinuum modeling (Sec. IV C), switching to other CSMs (Sec. IV D), cavity radii reoptimization (Sec. IV E), and *post facto* correction (Sec. IV F). The problems were narrowed to inaccurate and inconsistent errors in predicting the large $\Delta_{solv}G^*$ values for the two large divalent cations, $Zn(NH_3)_4{}^{2+}{}_{(aq)}$ and $Zn(en)_2{}^{2+}{}_{(aq)}$. The CSM predictions for these are extremely sensitive (in an absolute sense, in kcal mol$^{-1}$) to cavity radii choice. The CSM predictions appear to not suffer from the large q/r error seen with $H_3O^+$ and atomic ions, as such errors would have been cured with semicontinuum modeling (Sec. IV C). Of the CSMs tested (Sec. IV D), only SM12, with its large errors {−18 ± 3 and −15 ± 3 kcal mol$^{-1}$ on $\Delta_{solv}G[Zn(en)_2{}^{2+}]$ and $\Delta_{solv}G[Zn(NH_3)_4{}^{2+}]$, respectively}, happened to provide good error cancellation for the target Zn reaction. We have concluded that further work, beyond cavity radius reoptimization, is highly desirable to achieve <5 kcal mol$^{-1}$ accuracy for solvation Gibbs energies of large ions such as transition-metal complexes.

Until such improvements for large-ion solvation energies are available, we pursued ideas in Sec. V for what users could try to improve CSM accuracy for aqueous transition metal complexes in the meantime. We built improved procedures for all terms in Eq. (4) and applied them to three different versions of Eq. (1) chelate–effect reaction $M(NH_3)_4{}^{2+}$ + 2 *en*, by varying M = Zn, Cd, and Hg. The procedures involved either *post facto* corrected CSM results (Procedures A and B) or ADF's SM12 with its apparent large error cancellation (Procedure C). These worked well for the reactions when M is Zn or Cd. None of these immediately worked for the case M = Hg, but this was found to be due merely to poor cavity radii choices by IEFPCM(UFFx1.1) and ADF's SM12, and Bondi's second $Hg^{2+}$ radius[60] is now recommended for future studies involving aqueous $Hg^{2+}$. These procedures are limited only to Eq. (1) at present; much work would need to be done to determine systematic *post facto* corrections (Procedures A and B) or evaluate the general accuracy of SM12 for other ligand–exchange reactions (Procedure C).

## SUPPLEMENTARY MATERIAL

ESM references, ADF details including COSMO-RS-derived entropies, JDFTx details, derivation of "experimental" solvation Gibbs energies for $Zn(NH_3)_4{}^{2+}$ and $Zn(en)_2{}^{2+}$, a table of CSM radii sets and values, and Cartesian coordinates of optimized structures, are provided in PDF format as the supplementary material.

## ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council (Discovery Grant No. RGPIN-2017-0624) and the Canada Foundation for Innovation (Leading Edge Fund 2009, Grant No. 21625). Compute Canada (now the Digital Research Alliance of Canada) is acknowledged for access to computational resources for the ADF and JDFTx runs; the University of Regina is acknowledged for upkeep of the CFI supercomputers for the Gaussian09 runs. The Software for Chemistry and Materials (www.scm.com) support team is acknowledged for assistance in ADF computation. S. Dhillon is thanked for exploratory calculations in 2017, particularly several of the semicontinuum cluster structures in Sec. IV C. A.L.L.E. thanks R. Hempelmann (University of Saarbrücken) for hospitality and office facilities during final manuscript preparation.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**A. A. Mukadam:** Conceptualization (supporting); Investigation (equal); Visualization (supporting); Writing – original draft (supporting). **A. L. L. East:** Conceptualization (lead); Data curation (lead); Formal analysis (lead); Funding acquisition (lead); Investigation (equal); Methodology (lead); Project administration (lead); Resources (lead); Software (lead); Supervision (lead); Validation (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (lead).

## DATA AVAILABILITY

Most data are either provided in the article and the supplementary material or can be reproduced from the given data. Other data (e.g., atom coordinates to reproduce semicontinuum modeling energies in Table IV) are available from the corresponding author upon reasonable request.

## REFERENCES

[1] D. J. S. Sandbeck, C. M. Kuntz, C. Luu, R. A. Mondor, J. G. Ottaviano, A. V. Rayer, K. Z. Sumon, and A. L. L. East, J. Phys. Chem. A **118**, 11768 (2014).

[2] K. Z. Sumon, C. H. Bains, D. J. Markewich, A. Henni, and A. L. L. East, J. Phys. Chem. B **119**, 12256 (2015).

[3] S. Dhillon and A. L. L. East, Int. J. Quantum Chem. **118**, e25703 (2018).

[4] D. H. Patel and A. L. L. East, J. Phys. Chem. A **124**, 9088 (2020).

[5] B. Marten, K. Kim, C. Cortis, R. A. Friesner, R. B. Murphy, M. N. Ringnalda, D. Sitkoff, and B. Honig, J. Phys. Chem. **100**, 11775 (1996).

[6] C. J. Cramer and D. G. Truhlar, Chem. Rev. **99**, 2161 (1999).

[7] J. Tomasi, B. Mennucci, and R. Cammi, Chem. Rev. **105**, 2999 (2005).

[8] A. Klamt, B. Mennucci, J. Tomasi, V. Barone, C. Curutchet, M. Orozco, and F. J. Luque, Acc. Chem. Res. **42**, 489 (2009).

[9] V. S. Bryantsev, M. S. Diallo, and W. A. Goddard III, J. Phys. Chem. B **112**, 9709 (2008).

[10] J. R. Pliego, Jr. and J. M. Riveros, J. Phys. Chem. A **105**, 7241 (2001).

[11] C. P. Kelly, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. B **110**, 16066 (2006).

[12] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. B **113**, 6378 (2009).

[13] A. M. Maldonado, S. Hagiwara, T. H. Choi, F. Eckert, K. Schwarz, R. Sundararaman, M. Otani, and J. A. Keith, J. Phys. Chem. A **125**, 154 (2021).

[14] S. W. Rick and B. J. Berne, J. Am. Chem. Soc. **116**, 3949 (1994).

[15] H. Diehl, Chem. Rev. **21**, 39 (1937).

[16] O. Anderson, Chem. Rev. **99**, 2683 (1999).

[17] A. M. Marabini, M. Ciriachi, P. Plescia, and M. Barbaro, Miner. Eng. **20**, 1014 (2007).

[18] E. Graf, J. Agric. Food Chem. **42**, 1616 (1994).

[19] A. Werner, Z. Anorg. Chem. **3**, 267 (1893).

[20] M. Calvin and R. H. Bailes, J. Am. Chem. Soc. **68**, 949 (1946).

[21] A. E. Martell and M. Calvin, *Chemistry of the Metal Chelate Compounds* (Prentice-Hall, New York, 1952).

[22] G. Schwarzenbach, Helv. Chim. Acta **35**, 2344 (1952).

[23] C.-S. Chung, Inorg. Chem. **18**, 1321 (1979).

[24] R. T. Myers, Inorg. Chem. **17**, 952 (1978).

[25] R. D. Hancock and F. Marsicano, J. Chem. Soc., Dalton Trans. **12**, 1096 (1976).

[26] C. G. Spike and R. W. Parry, J. Am. Chem. Soc. **75**, 3770 (1953).

[27] A. E. Martell and R. M. Smith, *Critical Stability Constants: Volume 2: Amines* (Plenum Press, New York, 1975); *Critical Stability Constants: Volume 4: Inorganic Complexes* (Plenum Press, New York, 1976).

[28] A. Levy, D. Andelman, and H. Orland, Phys. Rev. Lett. **108**, 227801 (2012).

[29] Y. Z. Wei, P. Chiang, and S. Sridhar, J. Chem. Phys. **96**, 4569 (1992).

[30] M. J. Frisch *et al.*, Gaussian 09, Revision C.01, Gaussian, Inc., Wallingford, CT, 2009.

[31] E. J. Baerends, T. Ziegler *et al.*, ADF, version 2021.104, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, 2021, www.scm.com.

[32] JDFTx, version 1.7.0, R. Sundararaman, K. Letchworth-Weaver, K. A. Schwarz, D. Gunceler, Y. Ozhabes, and T. A. Arias, SoftwareX **6**, 278 (2017).

[33] A. A. Mukadam, N. P. Aravindakshan, and A. L. L. East, Phys. Chem. Chem. Phys. **22**, 7119 (2020).

[34] A. L. L. East and W. D. Allen, J. Chem. Phys. **99**, 4638 (1993).

[35] M. S. Schuurman, S. R. Muir, W. D. Allen, and H. F. Schaefer III, J. Chem. Phys. **120**, 11586 (2004).

[36] A. L. L. East and L. Radom, J. Chem. Phys. **106**, 6655 (1997).

[37] J. P. Merrick, D. Moran, and L. Radom, J. Phys. Chem. A **111**, 11683 (2007).

[38] F. Khalili, A. Henni, and A. L. L. East, J. Mol. Struct.: THEOCHEM **916**, 1 (2009).

[39] K. Z. Sumon, A. Henni, and A. L. L. East, Ind. Eng. Chem. Res. **51**, 11924 (2012).

[40] M. F. Pill, A. L. L. East, D. Marx, M. K. Beyer, and H. Clausen-Schaumann, Angew. Chem., Int. Ed. **58**, 9787 (2019).

[41] G. Scalmani and M. J. Frisch, J. Chem. Phys. **132**, 114110 (2010).

[42] M. Cossi, N. Rega, G. Scalmani, and V. Barone, J. Comput. Chem. **24**, 669 (2003).

[43] R. A. Pierotti, Chem. Rev. **76**, 717 (1976).

[44] F. M. Floris, J. Tomasi, and J. L. P. Ahuir, J. Comput. Chem. **12**, 784 (1991).

[45] C. A. Peeples and G. Schreckenbach, J. Chem. Theory Comput. **12**, 4033 (2016).

[46] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, J. Chem. Theory Comput. **9**, 609 (2013).

[47] C. C. Pye and T. Ziegler, Theor. Chem. Acc. **101**, 396 (1999).

[48] A. Klamt and G. Schüürmann, J. Chem. Soc., Perkin Trans. 2 **1993**, 799.

[49] C. C. Pye, T. Ziegler, E. van Lenthe, and J. N. Louwen, Can. J. Chem. **87**, 790 (2009).

[50] A. Klamt, J. Phys. Chem. **99**, 2224 (1995).

[51] C. Dupont, O. Andreussi, and N. Marzari, J. Chem. Phys. **139**, 214110 (2013).

[52] R. Sundararaman and W. A. Goddard III, J. Chem. Phys. **142**, 064107 (2015).

[53] V. Vallet, U. Wahlgren, and I. Grenthe, J. Am. Chem. Soc. **125**, 14941 (2003).

[54] A. Sengupta, A. Seitz, and K. M. Merz, Jr., J. Am. Chem. Soc. **140**, 15166 (2018).

[55] A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff, J. Am. Chem. Soc. **114**, 10024 (1992).

[56] A. Klamt and F. Eckert, Fluid Phase Equilib. **172**, 43 (2000).

[57] S. Cabani, P. Gianni, V. Mollica, and L. Lepori, J. Solution Chem. **10**, 563 (1981).

[58] A. A. Rashin, L. Young, and I. A. Topol, Biophys. Chem. **51**, 359 (1994).

[59] A. Bondi, J. Phys. Chem. **68**, 441 (1964).

[60] A. Bondi, J. Phys. Chem. **70**, 3006 (1966).