

## A Universal Approach to Solvation Modeling

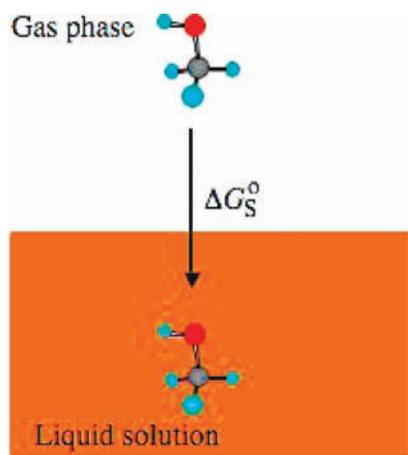
CHRISTOPHER J. CRAMER\* AND DONALD G. TRUHLAR\*

*Department of Chemistry and Supercomputer Institute, University of Minnesota,  
207 Pleasant Street SE, Minneapolis, Minnesota 55455*

RECEIVED ON JANUARY 20, 2008

### CON SPECTUS

Continuum mean-field models that have been carefully designed to address the various electrostatic and nonelectrostatic interactions that develop between a molecule and a surrounding medium are particularly efficient tools for studying the effects of condensed phases on molecular structure, energetics, properties, spectra, interaction potentials, and dynamics. The SM8 model may be combined with density functional theory or Hartree–Fock theory to describe a solute’s electronic structure and its self-consistent-field polarization by a solvent. A key feature is the use of class IV charge models to obtain accurate charge distributions (either in the vapor phase or in solution), even when using small basis sets that are affordable for large systems. A second key feature is that nonelectrostatic effects due to cavity formation, dispersion interactions, and changes in solvent structure are included in terms of empirical atomic surface tensions that depend on geometry but do not require atom-type assignments by the user. Use of an analytic surface area algorithm provides very stable energy gradients that allow geometry optimization in solution. The SM8 continuum model, the culmination of a series of SM $x$  models ( $x = 1–8$ ), permits the modeling of such diverse media as aqueous and organic solvents, soils, lipid bilayers, and air–water interfaces. In addition to predicting accurate transfer free energies between gaseous and condensed phases or between two different condensed phases, SM $x$  models have been useful for predicting the significant influence of condensed phases on processes associated with a change in molecular charge, including acid/base equilibria and oxidation/reduction processes. In this Account, we provide an overview of the algorithms associated with the computation of free energies of solvation in the SM8 model. We also compare the accuracies of the SM8 model with those of other continuum solvation models. Finally, we highlight applications of the SM8 models to compute ionic solvation free energies, oxidation and reduction potentials, and  $pK_a$  values.



### 1. Introduction

It is now possible to include continuum solvation effects in the computational modeling of molecular properties and reactions in solution, at both the classical and the quantum mechanical levels, by using efficient models based on treating the solvent as a structureless continuum.<sup>1</sup> Because of the enormous activity in this field, many reviews have appeared in the last several years.<sup>2–10</sup> In this Account, we discuss a new solvation model, called SM8,<sup>11</sup> that represents the culmination of many years of development of earlier SM $x$  models with  $x =$

1–7.<sup>12–23</sup> SM8 is a “universal” solvation model (meaning that it can be employed to model any liquid condensed phase for which certain bulk properties either are available or can be estimated, *vide infra*), and it is the most accurate continuum solvation model presently available for predicting free energies of solvation of molecular solutes. Among other things, this high accuracy has allowed quantum-mechanical modeling to assign absolute solvation free energies to the proton and other ions in various solvents<sup>11,24,25</sup> and to calculate absolute and relative oxidation and reduction potentials.

In quantum mechanical (QM) continuum models, the solute's charge distribution causes electric polarization of the surrounding solvent, which is modeled as a homogeneous medium characterized by a bulk permittivity. The polarized solvent exerts a field called the reaction field on the solute. Electronic structure calculations typically minimize the quantity

$$G_s^\circ = \left\langle \Psi \left| H^{(0)} + \frac{1}{2} V(\Psi) \right| \Psi \right\rangle \quad (1)$$

where  $\Psi$  is the solute wave function,  $H^{(0)}$  is the usual gas-phase Hamiltonian, and  $V$  is the potential energy operator associated with the reaction field. The factor of  $1/2$  in eq 1 derives from assuming a linear response of the surrounding medium to the solute's charge distribution because half of the induced favorable solute–solvent interaction is canceled by the cost of reorganizing the solvent.<sup>26</sup> Because  $V$  depends on  $\Psi$ , the corresponding Schrödinger equation is nonlinear. Iterative solutions are referred to as self-consistent reaction-field (SCRF) calculations. The absence of explicit solvent molecules means that the size of the electronic structure problem is essentially the same as in the gas phase, so the SCRF process is computationally very efficient. Importantly, if a particular situation demands the inclusion of one or more explicit solvent molecules, these can be included as part of a supermolecular QM solute; for example, it can be interesting to evaluate the degree to which charge transfer from a solute to a first-shell solvent molecule may take place.<sup>27</sup> However, the inclusion of more than a very small number of solvent molecules quickly reduces the efficiency of the SCRF procedure because of the need to sample the configurational phase space of the solute–solvent cluster in a statistically relevant fashion.

The component of the solvation free energy associated with electrostatics is

$$\Delta G_{\text{ENP}} = \left\langle \Psi^{(1)} \left| H + \frac{1}{2} V \right| \Psi^{(1)} \right\rangle - \left\langle \Psi^{(0)} \left| H \right| \Psi^{(0)} \right\rangle \quad (2)$$

where the subscript ENP on the left-hand side (lhs) of eq 2 denotes the electronic, nuclear, and polarization components of the free energy, and the superscripts (0) and (1) on the wave functions refer to their optimization in the gas phase and solution, respectively.

Continuum solvent models differ from one another in the manner in which  $V$  is constructed. They also differ in their treatment of nonelectrostatic components. The free energy of solvation, that is the free energy of transfer from the gas phase to the condensed phase, may then be written<sup>28</sup>

$$\Delta G_s^\circ = \Delta G_{\text{ENP}} + G_{\text{CDS}} + \Delta G_{\text{conc}}^\circ \quad (3)$$

where the CDS subscript on the right-hand side (rhs) emphasizes that this term includes the free energy changes associated with solvent cavitation, changes in dispersion energy, and possible changes in local solvent structure. The final term on the rhs,  $\Delta G_{\text{conc}}^\circ$ , accounts for the difference in the molar densities of the two phases in their standard states; the value of this term is 0 in this Account because we use the same 1 M density in all phases. The ENP and CDS terms are not separate thermodynamic observables; only their sum is meaningful. Nevertheless, one or the other of these terms may be expected to dominate under certain conditions, for example,  $\Delta G_{\text{ENP}}$  would be expected to be much larger than  $G_{\text{CDS}}$  for a charged solute, while the reverse would be expected to be true for a large, uncharged solute that lacks any polar functionality.

## 2. Generalized Born Theory

The reaction field generated by a charge distribution contained within a cavity that is immersed in a surrounding homogeneous medium characterized by a bulk permittivity is related to the QM charge distribution by the nonhomogeneous Poisson equation<sup>29</sup> (NPE). A number of extant QM continuum models determine the reaction field by numerical solution of the NPE.<sup>2–5,7,9</sup> Approximations must be made when using this approach because the tails of the QM charge density are not fully contained within the molecular cavity, which is typically defined as a union of atom-centered spheres having appropriate or scaled van der Waals radii or as a similar enclosed space. Furthermore, the formulation of the NPE requires one to assume an atomic-scale model of the solvent permittivity, which itself is rigorously defined only on macroscopic scales, and usually an unrealistic step function is assumed across the solute/solvent boundary.

An alternative approach is to use the generalized Born (GB) approximation.<sup>30–33</sup> In this case, the QM charge distribution is represented by atom-centered monopoles (i.e., partial atomic charges,  $q_k$ , on the atoms  $k$ ), and the reaction field potential generated at atom  $k$  by these charges is defined as

$$V_k = \left( 1 - \frac{1}{\epsilon(T)} \right) \sum_k q_k \gamma_{kk'} \quad (4)$$

where  $\epsilon$  is the permittivity of the medium, dependent on temperature  $T$ , and  $\gamma_{kk'}$  is an effective Coulomb integral. A good approximation to the latter is<sup>34</sup>

$$\gamma_{kk'} = \left( r_{kk'}^2 + b_k b_{k'} e^{-r_{kk'}^2 / (d_{kk'} b_k b_{k'})} \right)^{-1/2} \quad (5)$$

where  $r_{kk'}$  is an interatomic distance,  $b_k$  is an effective atomic Born radius, and  $d_{kk'}$  is a parameter.

In ref 34 and SM8, the effective atomic Born radius,  $b_{k_i}$  is calculated by a three-dimensional integration in which an intrinsic Coulomb radius is augmented to account for an effective distance of a given atom from the solvent (because the atom is buried under other solute atoms). The resulting effective radii are smallest when the atom is intrinsically small and is also most exposed to solvent. The determination of the intrinsic Coulomb radii relies primarily on experimental data for the solvation free energies of ionic species because ions have very large polarization free energies and are thus more sensitive to the Coulomb radii. Based on the cluster-pair approximation applied to conventional ionic solvation free energies in water,<sup>24</sup> methanol,<sup>25</sup> acetonitrile,<sup>25</sup> and dimethylsulfoxide,<sup>25</sup> we assembled an extensive compilation of absolute ionic solvation free energies that was used for parametrization.<sup>11</sup> A key difference of SM8 with earlier solvation models is that none of the prior models included ionic solvation free energies in nonaqueous solutions in the parametrization process, and the Coulomb radii were thus taken to be independent of solvent. In SM8, by contrast, some Coulomb radii vary as functions of the solvent's hydrogen-bond acidity parameter,<sup>35</sup>  $\alpha$ .

Another critical aspect associated with  $\Delta G_{\text{ENP}}$  is the assignment of partial atomic charges, which are used to calculate  $V_k$  and the solute's interaction with it. The SM8 model determines these charges using Charge Model 4 (CM4),<sup>11,16</sup> which defines a partial atomic charge according to

$$q_k^{\text{CM4}} = q_k^{(0)} + \sum_{k' \neq k} B_{kk'} (D_{kk'} + C_{kk'} B_{kk'}) \quad (6)$$

where the reference charge  $q_k^{(0)}$  is determined from Löwdin<sup>36</sup> or redistributed Löwdin<sup>37</sup> population analysis,  $B_{kk'}$  is the Mayer bond order between atoms  $k$  and  $k'$ ,<sup>38,39</sup> and  $C_{kk'}$  and  $D_{kk'}$  are empirically optimized parameters. Over a database of 397 polar molecules, CM4 models predict dipole moments with root-mean-square errors of 0.2–0.3 D,<sup>11,16</sup> even for inexpensive levels of QM theory and small basis sets. As the CM4 parameters are defined for many different combinations of electronic structure theory and basis set, the charge model effectively preprocesses the wave function in an SM8 calculation, and the SM8 parameters are independent of theoretical level. Indeed, SM8 may be regarded as being independent of CM4, too, in the sense that it ought to work with *any* good charges, but our testing has involved use of CM4 for SCRF calculations.

### 3. Nonelectrostatic Terms and SM8 Parameterization Strategy

Both the NPE and GB models of the electrostatics involve serious and essentially unremovable approximations (uncertainties in the treatment of solute charge outside the cavity and the assumption of bulk solvent permittivity on the atomic scale); they also neglect solute–solvent charge transfer. The  $G_{\text{CDS}}$  term in eq 3 must make up for these approximations, as well as account for the cavitation, dispersion, and solvent-structural effects that are missing in the bulk-electrostatics model.

The quantity  $G_{\text{CDS}}$  is then the difference between the experimental solvation free energy and  $\Delta G_{\text{ENP}}$  (see eq 3). Thus, given data for experimental solvation free energies and a protocol for computation of  $\Delta G_{\text{ENP}}$ , a set of target  $G_{\text{CDS}}$  values may be assembled. In the case of SM8, such a training set was assembled for 318 neutral solutes in 90 nonaqueous solvents and water (a total of 2346 individual solvation free energies), and 143 transfer free energies for 93 neutral solutes between water and 15 organic solvents were also used (transfer free energies provide differences in  $G_{\text{CDS}}$  values).<sup>11</sup> The functional form of  $G_{\text{CDS}}$  is

$$G_{\text{CDS}} = \sum_k (\sigma_k + \sigma^{\text{M}}) A_k \quad (7)$$

where  $\sigma_k$  and  $\sigma^{\text{M}}$  are atomic and molecular surface tensions, respectively, and  $A_k$  is the solvent-accessible surface area (SASA) of atom  $k$ , which depends on the molecular geometry and the set of all atomic van der Waals radii (we use the values of Bondi<sup>40</sup>). The physical basis for eq 7 is that deviations from bulk electrostatics, as well as cavitation, dispersion, and solvent-structural contributions, are all concentrated in the first solvation shell, and  $A_k$  is basically a continuous measure of the amount of solvent in the first solvation shell of atom  $k$ .<sup>41,42</sup>

The atomic surface tensions are sensitive to local environment and therefore they are computed according to

$$\sigma_k = \bar{\sigma}_{Z_k} + \sum_{k' \neq k} \bar{\sigma}_{Z_k Z_{k'}} T_{kk'} \quad (8)$$

where  $\bar{\sigma}_Z$  and  $\bar{\sigma}_{ZZ'}$  are monatomic and diatomic parameter functions, respectively, that depend on atomic numbers  $Z$ , and  $T_{kk'}$  is a geometry-dependent switching function called a cut-off tanh or COT.<sup>20</sup> The universal character of SM8 derives from the functional dependence of the atomic and molecular surface tensions on macroscopic solvent properties. In particular,

$$\bar{\sigma}_i = \bar{\sigma}_i^{[n]}n + \bar{\sigma}_i^{[\alpha]}\alpha + \bar{\sigma}_i^{[\beta]}\beta \quad (9)$$

where  $\bar{\sigma}_i$  is either  $\bar{\sigma}_Z$  or  $\bar{\sigma}_{ZZ'}$ ,  $n$  is the refractive index of the solvent at 293 K,  $\alpha$  and  $\beta$  are hydrogen bond acidity and basicity parameters<sup>35</sup> of the solvent, and  $\bar{\sigma}_i^{[n]}$ ,  $\bar{\sigma}_i^{[\alpha]}$ , and  $\bar{\sigma}_i^{[\beta]}$  are empirical, numerical parameters that are optimized to minimize errors in computed solvation free energies.

The molecular surface tension is also a function of solvent descriptors according to

$$\bar{\sigma}^{[M]} = \bar{\sigma}^{[\gamma]}\gamma + \bar{\sigma}^{[\varphi^2]}\varphi^2 + \bar{\sigma}^{[\psi^2]}\psi^2 + \bar{\sigma}^{[\beta^2]}\beta^2 \quad (10)$$

where  $\gamma$  is the 298 K macroscopic surface tension of the solvent at the air/solvent interface,  $\varphi^2$  is the square of the fraction of nonhydrogenic solvent atoms that are aromatic carbon atoms (carbon aromaticity),  $\psi^2$  is the square of the fraction of solvent atoms that are F, Cl, or Br (electronegative halogenicity), and  $\bar{\sigma}^{[\gamma]}$ ,  $\bar{\sigma}^{[\varphi^2]}$ ,  $\bar{\sigma}^{[\psi^2]}$ , and  $\bar{\sigma}^{[\beta^2]}$  are empirical parameters that are independent of solute or atomic number.

The SM8 model may be applied to any solvent (or indeed any medium) for which the relevant macroscopic descriptors are either known or may be estimated. We note however that water is taken to be a special case with its own parameters (one could use the general parameters for water, but water is important enough to warrant the extra accuracy achievable by having specific parameters). Both  $\Delta G_{\text{ENP}}$  and  $G_{\text{CDS}}$  are calculated with a stable analytic algorithm developed for this purpose.<sup>43</sup> This leads to very stable analytic gradients,<sup>43</sup> which facilitates geometry optimization even for difficult structures in solution.<sup>44</sup>

The SM8 solvation model additionally extends the aqueous parametrization to temperatures ( $T$ ) other than 298 K.<sup>45,46</sup> Temperature dependence is included in  $\Delta G_{\text{ENP}}$  by accounting for the  $T$  dependence of the permittivity (see eq 4). The remaining change in the total solvation free energy is associated with changes in  $G_{\text{CDS}}$ , which is modeled as

$$\Delta G_{\text{CDS}} = B(T - T_0) + C[(T - T_0) - T \ln(T/T_0)] \quad (11)$$

where  $T_0$  is 298.15 K, and  $B$  and  $C$  are computed in a fashion analogous to  $G_{\text{CDS}}(T_0)$ .

The temperature-dependent SM8T was developed by optimizing the relevant surface tension coefficients against 4403 aqueous solvation free energy measurements on 348 compounds at various temperatures ranging from 273 to 373 K.<sup>45,46</sup>

## 4. Model Performance

Table 1 compares predictions from SM8 at 298 K to those from four other readily available QM continuum solvation models for solvation free energies of neutrals and ions in

**TABLE 1.** Mean Unsigned Errors (kcal/mol) in Solvation Free Energies Calculated Using SMx and non-SMx Implicit Solvent Models<sup>a</sup>

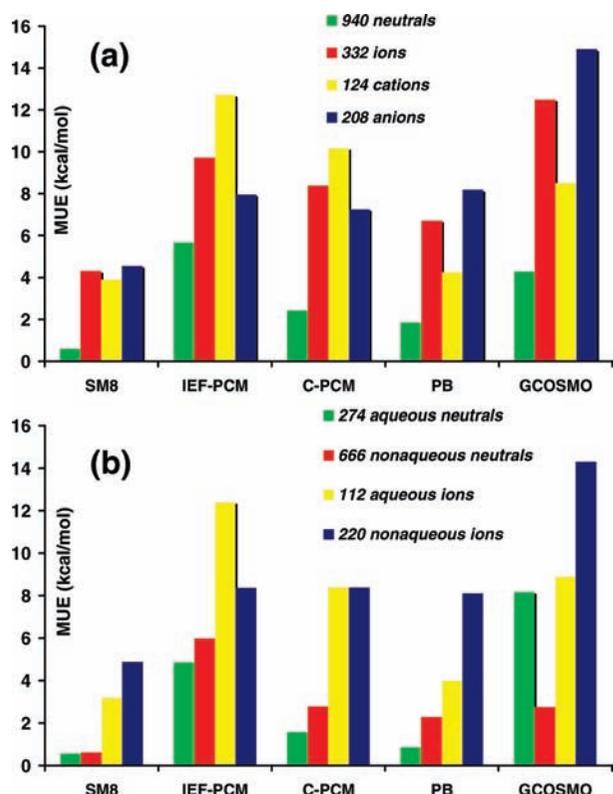
solute class	N	IEF-PCM				C-PCM*		PB*		GCOSMO*	
		SM8	G03/UA0	GAMESS	Jaguar	NWChem	NWChem	NWChem	NWChem	NWChem	
all neutrals	940	0.59	5.66	2.43	1.86	4.29 <sup>b</sup>					
all ions	332	4.31	9.73	8.39	6.72 <sup>c,d</sup>	12.49 <sup>d,e</sup>					
all cations	124	3.90	12.71	10.18	4.25	8.51					
all anions	208	4.55	7.97	7.26	8.19 <sup>c,d</sup>	14.90 <sup>d,e</sup>					
aqueous neutrals	274	0.55	4.87	1.57	0.86	8.17 <sup>b</sup>					
nonaq. neutrals	666	0.61	5.99	2.78	2.28	2.76					
aqueous ions	112	3.20	12.40	8.40	4.00	8.90 <sup>d,e</sup>					
nonaqueous ions	220	4.88	8.37	8.38	8.11 <sup>c,d</sup>	14.31					

<sup>a</sup> N is the number of data in a given solute class. The calculations were performed at the *m*PW1PW/6-31G(d) level of theory, except the calculations marked by asterisks, which used B3LYP in place of *m*PW1PW. <sup>b</sup> Excluding 11 phosphorus-containing compounds and tetramethylsilane. The count is reduced to  $N - 12$ . <sup>c</sup> Excluding 3-hydroxybenzoate. <sup>d</sup> The count is reduced to  $N - 1$ . <sup>e</sup> Excluding hydroperoxyl radical.

water and 17 organic solvents.<sup>11</sup> The other methods are the IEF-PCM model<sup>47,48</sup> as coded in *Gaussian 03*, the C-PCM model<sup>48,49</sup> as coded in *GAMESS*,<sup>48</sup> the PBSA model<sup>50,51</sup> as coded in *Jaguar*, and the GCOSMO model<sup>52,53</sup> as coded in *NWChem*. In each case, default options were used for the various models and codes; full details may be found in the original work<sup>11</sup> and in Supporting Information for this Account. The SM8 predictions show considerably smaller errors than those for the other models for aqueous and nonaqueous free energies of solvation for neutrals, cations, and anions, with particularly good performance for nonaqueous data compared with the other models. Figure 1 provides a graphical depiction of the relative model accuracies.

We note that if one makes the assumption that all ions have the same solvation free energy ( $-65.0$  kcal/mol), that all nonaqueous neutrals have the same solvation free energy ( $-5.38$  kcal/mol), and that all aqueous neutrals have the same solvation free energy ( $-2.99$  kcal/mol), these particular values being the average of all relevant experimental measurements, then the MUEs for this “three-parameter model” are 8.6, 1.5, and 2.7 kcal/mol, respectively. In only 6 of 15 cases do the non-SMx models show smaller MUEs than the three-parameter model.<sup>11</sup> The reason that the SMx models outperform other models is *not* the extent of parametrization or the choice of parameters, but rather the theoretical formulation of the problem so that the  $G_{\text{CDS}}$  term compensates for the inescapable incompleteness of *any* bulk electrostatic model for  $\Delta G_{\text{ENP}}$ .

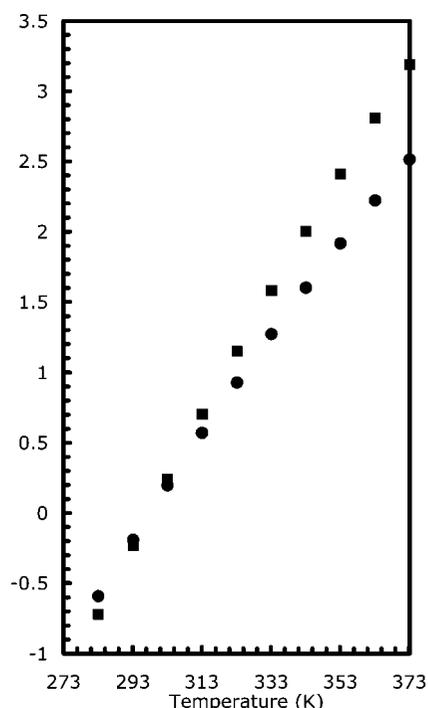
With respect to temperature dependence, the experimental aqueous solvation free energy of piperazine varies by about 4 kcal/mol over a 90 K temperature range. Figure 2 illustrates that over the complete set of data, SM8T<sup>46</sup> predicts



**FIGURE 1.** Mean unsigned errors (MUEs) in solvation free energies of neutral and ionic solutes calculated using SM8 and other continuum models including IEF-PCM/G03/UA0, C-PCM/GAMESS, PB/Jaguar, and GCOSMO/NWChem. The calculation was done for 18 solvents including acetonitrile (ions and neutral solutes), DMSO (ions and neutral solutes), methanol (only ions), water (ions and neutral solutes), and an additional 14 organic solvents from the SM8 neutral training set: (a) MUEs are given for ions and neutrals in all of the 18 solvents; (b) MUEs for solutes in aqueous solutions are compared with MUEs for solutes in nonaqueous solutions.

this temperature dependence with a mean unsigned error of 0.3 kcal/mol.

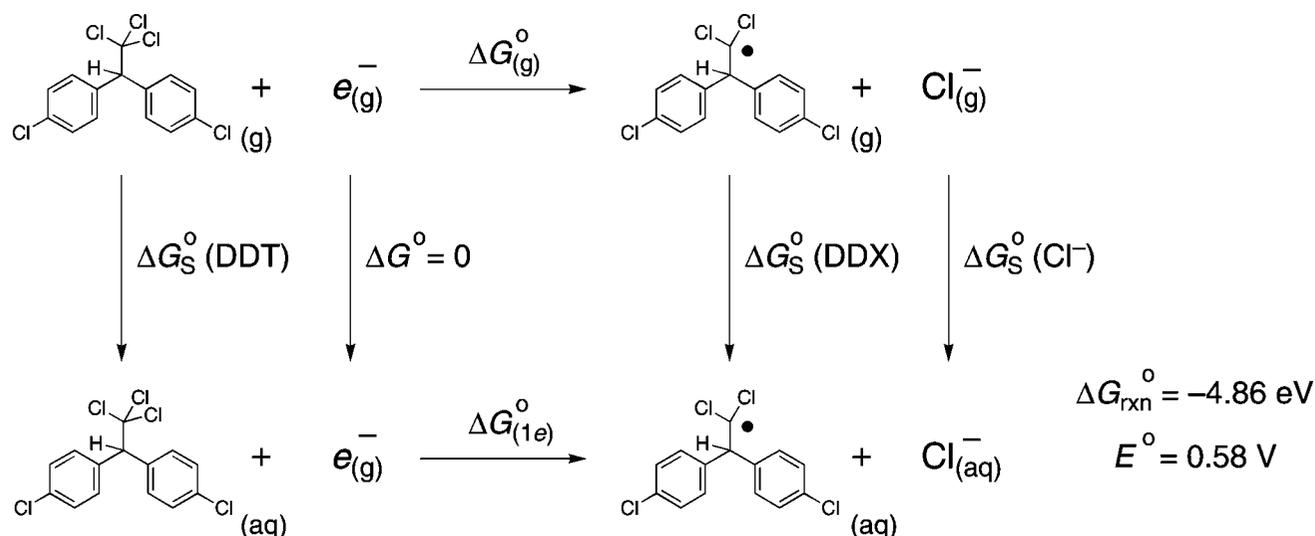
As noted above, a feature of the SM8 model that makes it particularly attractive is its universality; it can be applied to any phase that may be reasonably characterized by the bulk descriptors  $\epsilon$ ,  $n$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\varphi$ , and  $\psi$ . When values for these quantities are not available, they may be estimated, or more pragmatically, the unknown descriptors may themselves be taken to be parameters that are fit to available experimental data. Thus, one may make an estimate of the permittivity (the only nonlinear solvent parameter in the computation of  $\Delta G_S^\circ$ ), compute  $\Delta G_{\text{ENP}}$  for solutes for which solvation free energies or partition coefficients (which are differences of solvation free energies between two phases) are known, and optimize the remaining descriptors contributing to  $G_{\text{CDs}}$  through multilinear regression. A final model may be determined by optimization of  $\epsilon$ .



**FIGURE 2.** The temperature dependence of the experimental (■) free energy of solvation for piperazine compared with various possible parametrizations of SM8T.

We have put this approach into practice with an earlier SMx model for a phosphatidyl choline bilayer.<sup>54</sup> Available experimental data for partition coefficients between the bilayer and surrounding aqueous solvent led to assignment of a bilayer permittivity of 5.0 (similar to octanol), an  $\alpha$  value of 0.0 (because phosphatidyl choline has negligible hydrogen bond donating character), and optimized  $n$ ,  $\gamma$ , and  $\beta$  solvent parameters of 1.40, 27, and 1.15, respectively, these quantities being in the range of physically realistic values considering analogous molecules as solvents. This optimized model correlates the logarithms of 19 experimental partition coefficients ranging over 5 log units with a Pearson correlation coefficient of 0.9. For partition coefficients ( $P$ ) between conventional (aqueous and organic) solvents, the accuracy in  $\log P$  is typically 0.3–0.4.<sup>22</sup>

Another example of this process, applied to a problem having more available data, was the development of an SMx model for wet soil,<sup>55</sup> thus permitting the prediction of environmentally important soil/water partition coefficients,  $K_{\text{OC}}$  (these partition coefficients are normalized for organic carbon content when standard values are measured experimentally). In this case, optimization against a data set of 387 molecules led to optimized  $\epsilon$ ,  $n$ ,  $\gamma$ ,  $\alpha$ , and  $\beta$  solvent parameters of 15, 1.311, 45.3, 0.56, and 0.60, such values again



**FIGURE 3.** One-electron reduction potential vs NHE for DDT using standard state concentrations of 1 mol per 24.5 L for all species in the gas phase, 1 M for DDT and DDX in aqueous solution, and  $10^{-3}$  M for aqueous  $\text{Cl}^{-}$ .

being in a physically realistic range. The resulting soil partitioning model had a MUE over the training set of 0.98 kcal/mol.

The most novel extension of the SMx models to a different condensed phase has been the development of an SMx model to describe the surface layer of an aqueous solution that is in contact with air, that is, the air–water interface.<sup>56</sup> This formally two-dimensional phase is predicted to be characterized by  $\alpha$ ,  $\beta$ ,  $n$ , and  $\gamma$  values of 1.11, 0.59, 1.342, and  $-144.6$ , respectively. These values represent hydrogen bonding acidity and basicity values larger than those associated with bulk water<sup>57</sup>, in keeping with various studies of the aqueous surface that have rationalized this observation based on dangling OH bonds at the interface,<sup>58–61</sup> while the index of refraction is equal to that of bulk water, and the “surface tension” is negative, representing enhanced sticking as opposed to unfavorable cavitation costs. This SMx model predicts the air/surface partitioning of 85 complex organic molecules, including pesticides, with a MUE of 0.47 log units. This model can be applied to problems in atmospheric environmental chemistry, for example, contaminant transport by fog droplets.

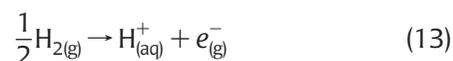
## 5. Applications to Computational Electrochemistry and $\text{pK}_a$ Prediction

Because of changes in total solute charge, solvation can have a large effect on the energetics of electron-transfer reactions. An electrochemical half-reaction for the one-electron reduction of the pesticide DDT is shown in Figure 3.<sup>62</sup> The free energy change for the aqueous portion of the thermodynamic cycle can be computed by combining gas-phase B3LYP/6-

311+G(d)//B3LYP/6-31G(d) free energies with the experimental solvation free energy for the chloride anion,<sup>63</sup> and SM5.42/BPW91/6-31G(d)//B3LYP/6-31G(d) solvation free energies for the organic molecules; this yields a free energy change in aqueous solution of  $-4.86$  eV. This  $\Delta G_{\text{rxn}}^{\circ}$  may be converted to a potential relative to the normal hydrogen electrode (NHE) by

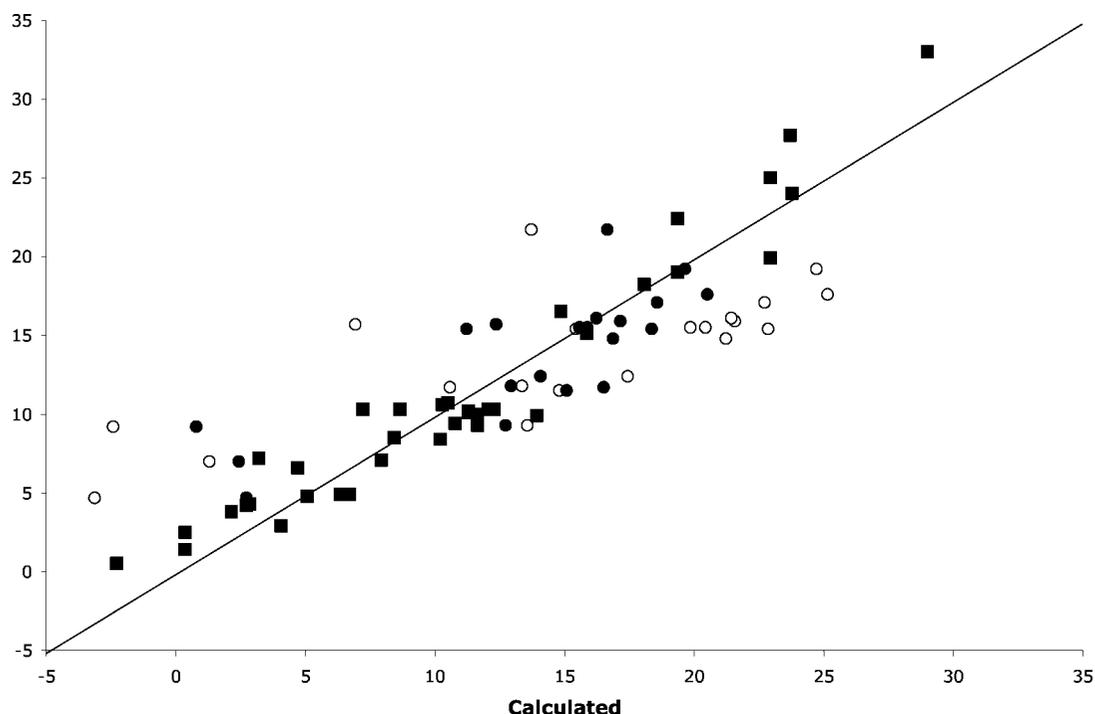
$$E^{\circ} = \frac{\Delta G_{\text{rxn}}^{\circ} - \Delta G_{\text{NHE}}^{\circ}}{nF} \quad (12)$$

where  $n$  is the number of electrons transferred,  $F$  is the Faraday constant, and  $\Delta G_{\text{NHE}}^{\circ}$  is the free energy of reaction for the reference half-reaction



which is 4.28 eV.<sup>24,62</sup> SMx models have been used to study the oxidation or reduction potentials of substituted anilines,<sup>64,65</sup> quinones,<sup>65</sup> various halogenated hydrocarbons,<sup>62,66,67</sup> flavin cofactors,<sup>68</sup> and the ruthenium ion.<sup>69</sup> When ions included in the relevant half-reactions are characterized by delocalized charge, continuum models tend to provide excellent correlation with experimental results. However, in cases of very concentrated, localized charge (e.g., the  $\text{Ru}^{2+}/\text{Ru}^{3+}$  couple<sup>69</sup>) good results can only be obtained by embedding the solute in the continuum with one or more shells of explicit solvent molecules included, that is, as a clustered supersolute.

The tendency of locally concentrated charge to lead to poor performance of the full continuum approximation and the ability of a manageable, well-defined cluster model to



**FIGURE 4.** Calculated vs experimental  $pK_a$  values for 57 acids.<sup>70</sup> Filled squares indicate unclustered conjugate bases, filled circles indicate clustered conjugate bases, and open circles indicate results for the same molecules as the solid circles but computed for the bare ion instead of the clustered ion. In almost every instance, the clustered-ion result is closer to the ideal line (shown) than is the bare-ion result (because the experimental result is a constant, connected filled and open circles may be identified by their horizontal relationship to one another).

improve this is also manifest in the calculation of  $pK_a$  values. As shown in Figure 4, when SM6 aqueous solvation free energies are added to gas-phase deprotonation free energies to compute  $pK_a$  values by

$$pK_a = \frac{\Delta G_{\text{rxn}}^\circ}{RT \ln(10)} \quad (14)$$

for the general acid deprotonation reaction



a significantly better correlation between theory and experiment is obtained for a set of 57 various acids when those conjugate bases having highly concentrated charge are considered as clusters with a single water molecule instead of as bare anions.<sup>70</sup> Importantly, in cases where the pure continuum approximation already works well, the inclusion of a clustering water molecule does not degrade the predicted  $pK_a$  values. Such behavior is the hallmark of a robust solvation model.

## 6. Conclusions

The development of the SMx models has reached a point where SM8, a universal solvation model, is particularly effective for the computation of solvation free energies and other

interesting thermodynamic properties in solution, like partition coefficients,  $pK_a$  values, and oxidation and reduction potentials. SM8 and/or earlier SMx models are available in various software packages for use by the broader modeling community, including *AMSOL*, *GAMESSPLUS*, *HONDOPLUS*, *JAGUAR*, *OMNISOL*, *SMXGAUSS*, and *SPARTAN*.

*We are grateful for the contributions of our many co-workers whose names appear as coauthors on citations in this Account; in addition to recognizing the contributions of our earlier co-workers, we especially thank our coauthors on SM8: Alek Marenich, Casey Kelly, Ryan Olson, and Adam Chamberlin. We also acknowledge financial support over the years from the Army Research Office, the Office of Naval Research, and the National Science Foundation.*

**Supporting Information Available.** Additional references for various theoretical models and full references for all software packages. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## BIOGRAPHICAL INFORMATION

**Christopher J. Cramer** was born in Freeport, Illinois, in 1961. He earned an A.B. *summa cum laude* in mathematics and chemistry from Washington University in St. Louis, where he worked with C. David Gutsche, and a Ph.D. in chemistry from the University of Illinois, where he worked with Scott E. Denmark. He joined the

faculty of the University of Minnesota in 1992 after postdoctoral work as an active-duty officer in the United States Army and presently holds the title Distinguished McKnight and University Teaching Professor of Chemistry, Chemical Physics, and Scientific Computation. His research interests are in the broad area of theoretical chemical modeling. He is recipient of the Arthur S. Fleming award and Alfred P. Sloan and John Simon Guggenheim fellowships.

**Donald G. Truhlar** was born in Chicago, Illinois, in 1944. He earned a B.A. *summa cum laude* in chemistry from St. Mary's College of Minnesota, where he worked with H. Philip Hogan and Ernest D. Kaufman, and a Ph.D. in chemistry from Caltech, where he worked with Aron Kuppermann. He joined the faculty of the University of Minnesota in 1969, and he is currently Regents Professor of Chemistry, Chemical Physics, Nanoparticle Science and Engineering, and Scientific Computation. His research interests are in the areas of theoretical and computational chemical dynamics and potential energy surfaces. He is recipient of an Alfred P. Sloan fellowship, the ACS Award for Computers in Chemical and Pharmaceutical Research, the ACS Peter Debye Award for Physical Chemistry, the NAS Award for Scientific Reviewing, and the WATOC Schrödinger Medal.

#### FOOTNOTES

\*E-mail addresses: cramer@umn.edu; truhlar@umn.edu.

#### REFERENCES

- Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*; 2nd ed.; John Wiley & Sons: Chichester, U.K., 2004, pp 385–427.
- Rivail, J.-L.; Rinaldi, D. Liquid State Quantum Chemistry: Computational Applications of the Polarizable Continuum Models. In *Computational Chemistry, Review of Current Trends*; Leszczynski, J., Ed.; World Scientific: New York, 1996, pp 139–174.
- Cramer, C. J.; Truhlar, D. G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999**, *99*, 2161–2200.
- Orozco, M.; Luque, F. J. Theoretical Methods for the Description of the Solvent Effect on Biomolecular Systems. *Chem. Rev.* **2000**, *100*, 4187–4225.
- Bashford, D.; Case, D. A. Generalized Born Models of Macromolecular Solvation Effects. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129–152.
- Cramer, C. J.; Truhlar, D. G. Thermodynamics of Solvation and the Treatment of Equilibrium and Nonequilibrium Solvation Effects by Models Based on Collective Solvent Coordinates. In *Free Energy Calculations in Rational Drug Design*; Reddy, M. R., Erion, M. D., Eds.; Kluwer Academic/Plenum: New York, 2001, pp 63–95.
- Tomasi, J. Thirty Years of Continuum Solvation Chemistry: A Review and Prospects for the near Future. *Theor. Chem. Acc.* **2004**, *112*, 184–203.
- Feig, M.; Brooks, C. L. Recent Advances in the Development and Application of Implicit Solvent Models in Biomolecule Simulations. *Curr. Opin. Struct. Biol.* **2004**, *14*, 217–224.
- Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3093.
- Cramer, C. J.; Truhlar, D. G. SMx Continuum Models for Condensed Phases In *Trends and Perspectives in Modern Computational Science*; Maroulis, G., Simos, T. E., Eds.; Brill Academic: Amsterdam, 2006; Vol. 6, pp 112–140.
- 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. Theor. Comput.* **2007**, *3*, 2011–2033.
- Cramer, C. J.; Truhlar, D. G. Molecular Orbital Theory Calculations of Aqueous Solvation Effects on Chemical Equilibria. *J. Am. Chem. Soc.* **1991**, *113*, 8552–8553.
- Cramer, C. J.; Truhlar, D. G. An SCF Solvation Model for the Hydrophobic Effect and Absolute Free Energies of Aqueous Solvation Including Specific Water Interactions. *Science* **1992**, *256*, 213–217.
- Cramer, C. J.; Truhlar, D. G. PM3-SM3: A General Parameterization for Including Aqueous Solvation Effects in the PM3 Molecular Orbital Model. *J. Comput. Chem.* **1992**, *13*, 1089–1097.
- Storer, J. W.; Giesen, D. J.; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; Truhlar, D. G.; Liotard, D. A. Solvation Modeling in Aqueous and Nonaqueous Solvents: New Techniques and a Re-examination of the Claisen Rearrangement. In *Structure and Reactivity in Aqueous Solution*; Cramer, C. J., Truhlar, D. G., Eds.; American Chemical Society: Washington, DC, 1994; Vol. 568, pp 24–49.
- 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.
- Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Parameterized Model for Aqueous Free Energies of Solvation using Geometry-dependent Atomic Surface Tensions with Implicit Electrostatics. *J. Phys. Chem. B* **1997**, *101*, 7147–7157.
- Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. OMNISOL: Fast Prediction of Free Energies of Solvation and Partition Coefficients. *J. Org. Chem.* **1998**, *63*, 4305–4313.
- Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Parameterized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J. Phys. Chem.* **1996**, *100*, 19824–19839.
- Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Universal Quantum Mechanical Model for Solvation Free Energies Based on Gas-Phase Geometries. *J. Phys. Chem. B* **1998**, *102*, 3257–3271.
- Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. Extension of the Platform of Applicability of the SM5.42R Universal Solvation Model. *Theor. Chem. Acc.* **1999**, *103*, 9–63.
- Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. New Universal Solvation Model and Comparison of the Accuracy of the SM5.42R, SM5.43R, C-PCM, D-PCM, and IEF-PCM Continuum Solvation Models for Aqueous and Organic Solvation Free Energies and for Vapor Pressures. *J. Phys. Chem. A* **2004**, *108*, 6532–6542.
- Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. Density-Functional Theory and Hybrid Density-Functional Theory Continuum Solvation Models for Aqueous and Organic Solvents: Universal SM5.43 and SM5.43R Solvation Models for Any Fraction of Hartree-Fock Exchange. *Theor. Chem. Acc.* **2005**, *113*, 107–131.
- Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Aqueous Solvation Free Energies of Ions and Ion-Water Clusters Based on An Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton. *J. Phys. Chem. B* **2006**, *110*, 16066–16081.
- Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Single-Ion Solvation Free Energies and the Normal Hydrogen Electrode Potential in Methanol, Acetonitrile, and Dimethyl Sulfoxide. *J. Phys. Chem. B* **2007**, *111*, 408–422.
- Cramer, C. J.; Truhlar, D. G. Continuum Solvation Models. In *Solvent Effects and Chemical Reactivity*; Tapia, O., Bertrán, J., Eds.; Kluwer: Dordrecht, The Netherlands, 1996; pp 1–81.
- Marenich, A. V.; Olson, R. M.; Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. Polarization Effects in Aqueous and Nonaqueous Solutions. *J. Chem. Theory Comput.* **2007**, *3*, 2055–2067.
- Ben-Naim, A. *Statistical Thermodynamics for Chemists and Biochemists*; Plenum: New York, 1992; pp 421–430.
- Wangsness, R. K. *Electromagnetic Fields*; John Wiley & Sons: New York, 1979; p 179.
- Hojtink, G. J.; de Boer, E.; Van der Meij, P. H.; Weijland, W. P. Potentials of Various Aromatic Hydrocarbons. *Recl. Trav. Chim. Pays-Bas* **1956**, *75*, 487–503.
- Paradejordi, F. On the Pariser and Parr Semiempirical Method for Computing Molecular Wave Functions. The Basic Strength of N-heteroatomic Compounds and their Monoamines. *Cahiers Phys.* **1963**, *17*, 393–447.
- Tapia, O. Local Field Representation of Surrounding Medium Effects. From Liquid Solvent to Protein Core Effects. In *Quantum Theory of Chemical Reactions*; Daudel, R., Pullman, A., Salem, L., Viellard, A., Eds.; Reidel: Dordrecht, The Netherlands, 1980; Vol. 2, pp 25–72.
- Cramer, C. J.; Truhlar, D. G. General Parameterized SCF Model for Free Energies of Solvation in Aqueous Solution. *J. Am. Chem. Soc.* **1991**, *113*, 8305–8311.
- Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
- Abraham, M. H. Scales of Solute Hydrogen-bonding: Their Construction and Application to Physicochemical and Biochemical Processes. *Chem. Soc. Rev.* **1993**, *73*–83.
- Baker, J. Classical Chemical Concepts from ab Initio SCF Calculations. *Theor. Chim. Acta* **1985**, *68*, 221–229.
- Thompson, J. D.; Xidos, J. D.; Sonbuchner, T. M.; Cramer, C. J.; Truhlar, D. G. More Reliable Partial Atomic Charges When Using Diffuse Basis Sets. *PhysChemComm* **2002**, *5*, 117–134.
- Mayer, I. Charge, Bond Order and Valence in the Ab Initio SCF Theory. *Chem. Phys. Lett.* **1983**, *97*, 270–277.

- 39 Mayer, I. Comments on the Quantum Theory of Valence and bonding: Choosing Between Alternative Definitions. *Chem. Phys. Lett.* **1984**, *110*, 440.
- 40 Bondi, A. van der Waals Volumes and Radii. *J. Phys. Chem.* **1964**, *68*, 441–451.
- 41 Nemethy, G.; Scheraga, H. A. Structure of Water and Hydrophobic Binding in Proteins. 2. Model for Thermodynamic Properties of Aqueous Solutions of Hydrocarbons. *J. Chem. Phys.* **1962**, *36*, 3401.
- 42 Hermann, R. B. Theory of Hydrophobic Bonding. 2. Correlation of Hydrocarbon Solubility in Water with Solvent Cavity Surface Area. *J. Phys. Chem.* **1972**, *76*, 2754–2759.
- 43 Liotard, D. A.; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; Truhlar, D. G. Improved Methods for Semiempirical Solvation Models. *J. Comput. Chem.* **1995**, *16*, 422–440.
- 44 Houseknecht, J. B.; McCarren, P. R.; Lowary, T. L.; Hadad, C. M. Conformational Studies of Methyl 3-*O*-Methyl- $\alpha$ -D-arabinofuranoside: An Approach for Studying the Conformation of Furanose Rings. *J. Am. Chem. Soc.* **2001**, *123*, 8811–8824.
- 45 Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. Predicting Aqueous Free Energies of Solvation As Functions of Temperature. *J. Phys. Chem. B* **2006**, *110*, 5665–5675.
- 46 Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. Extension of a Temperature-Dependent Aqueous Solvation Model to Compounds Containing Nitrogen, Fluorine, Chlorine, Bromine, and Sulfur. *J. Phys. Chem. B* **2008**, *112*, 3024–3039.
- 47 Cancès, E.; Mennucci, B.; Tomasi, J. A New Integral Equation Formalism for the Polarizable Continuum Model: Theoretical Background and Applications to Isotropic and Anisotropic Dielectrics. *J. Chem. Phys.* **1997**, *107*, 3032–3041.
- 48 Additional references for various theoretical models and full references for all software packages are provided in the Supporting Information.
- 49 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.
- 50 Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, W. A., III; Honig, B. Accurate First Principles Calculation of Molecular Charge Distributions and Solvation Energies from Ab Initio Quantum Mechanics and Continuum Dielectric Theory. *J. Am. Chem. Soc.* **1994**, *116*, 11875–11882.
- 51 Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. New Model for Calculation of Solvation Free Energies: Correction of Self-Consistent Reaction Field Continuum Dielectric Theory for Short-Range Hydrogen-Bonding Effects. *J. Phys. Chem.* **1996**, *100*, 11775–11788.
- 52 Truong, T. N.; Stefanovich, E. V. A New Method for Incorporating Solvent Effect into the Classical, ab Initio Molecular Orbital Theory and Density Functional Theory Frameworks for Arbitrary Shape Cavity. *Chem. Phys. Lett.* **1995**, *240*, 253–260.
- 53 Truong, T. N.; Stefanovich, E. V. Analytical First and Second Energy Derivatives of the Generalized Conductor-Like Screening Model for Free Energy of Solvation. *J. Chem. Phys.* **1995**, *103*, 3709–3717.
- 54 Chambers, C. C.; Giesen, D. J.; Hawkins, G. D.; Vaes, W. H. J.; Cramer, C. J.; Truhlar, D. G. Modeling the Effect of Solvation on Structure, Reactivity, and Partitioning of Organic Solutes: Utility in Drug Design In *Rational Drug Design*; Truhlar, D. G., Howe, W. J., Hopfinger, A. J., Blaney, J. M., Dammkoehler, R. A., Eds.; Springer: New York, 1999; pp 51–72.
- 55 Winget, P.; Cramer, C. J.; Truhlar, D. G. Prediction of Soil Sorption Coefficients Using a Universal Solvation Model. *Environ. Sci. Technol.* **2000**, *34*, 4733–4740.
- 56 Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Predicting Adsorption Coefficients at Air-Water Interfaces Using Universal Solvation and Surface Area Models. *J. Phys. Chem. B* **2004**, *108*, 12882–12897.
- 57 Abraham, M. H.; Andonian-Haftvan, J.; Whiting, G. S.; Leo, A.; Taft, R. S. Hydrogen Bonding. Part 34. The Factors that Influence the Solubility of Gases and Vapours in Water at 298 K, and a New Method for its Determination. *J. Chem. Soc., Perkin Trans. 2* **1994**, 1777–1791.
- 58 Du, Q.; Superfine, R.; Freysz, E.; Shen, Y. R. Vibrational Spectroscopy of Water at the Vapor/Water Interface. *Phys. Rev. Lett.* **1993**, *70*, 2313.
- 59 Du, Q.; Freysz, E.; Shen, Y. R. Surface Vibrational Spectroscopic Studies of Hydrogen-Bonding and Hydrophobicity. *Science* **1994**, *264*, 826–828.
- 60 Gragson, D. E.; Richmond, G. L. Investigations of the Structure and Hydrogen Bonding of Water Molecules at Liquid Surfaces by Vibrational Sum Frequency Spectroscopy. *J. Phys. Chem. B* **1998**, *102*, 3847–3861.
- 61 Kuo, I.-F. W.; Mundy, C. J. An ab Initio Molecular Dynamics Study of the Aqueous Liquid-Vapor Interface. *Science* **2004**, *303*, 658–660.
- 62 Lewis, A.; Bumpus, J. A.; Truhlar, D. G.; Cramer, C. J. Molecular Modeling of environmentally important processes: Reduction potentials. *J. Chem. Educ.* **2004**, *81*, 596–604; erratum **2007**, *84*, 934.
- 63 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.
- 64 Winget, P.; Weber, E. J.; Cramer, C. J.; Truhlar, D. G. Computational Electrochemistry: Aqueous One-Electron Oxidation Potentials for Substituted Anilines. *Phys. Chem. Chem. Phys.* **2000**, *2*, 1231–1239.
- 65 Winget, P.; Cramer, C. J.; Truhlar, D. G. Computation of Equilibrium Oxidation and Reduction Potentials for Reversible and Dissociative Electron-Transfer Reactions in Solution. *Theor. Chem. Acc.* **2004**, *112*, 217–227.
- 66 Patterson, E. V.; Cramer, C. J.; Truhlar, D. G. Reductive Dechlorination of Hexachloroethane in the Environment. Mechanistic Studies via Computational Electrochemistry. *J. Am. Chem. Soc.* **2001**, *123*, 2025–2031.
- 67 Arnold, W.; Winget, P.; Cramer, C. J. Reductive Dechlorination of 1,1,2,2-Tetrachloroethane. *Environ. Sci. Technol.* **2002**, *36*, 3536–3541; erratum **2002**, *36*, 4706.
- 68 Bhattacharyya, S.; Stankovich, M. T.; Truhlar, D. G.; Gao, J. L. Combined Quantum Mechanical and Molecular Mechanical Simulations of One- and Two-Electron Reduction Potentials of Flavin Cofactor in Water, Medium-Chain Acyl-CoA Dehydrogenase, and Cholesterol Oxidase. *J. Phys. Chem. A* **2007**, *111*, 5729–5742.
- 69 Jaque, P.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Computational Electrochemistry: The Aqueous Ru<sup>3+</sup>/Ru<sup>2+</sup> Reduction Potential. *J. Phys. Chem. C* **2007**, *111*, 5783–5799.
- 70 Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Adding Explicit Solvent Molecules to Continuum Solvent Calculations for the Calculation of Aqueous Acid Dissociation Constants. *J. Phys. Chem. A* **2006**, *110*, 2493–2499.