

Regular article

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

Jason D. Thompson, Christopher J. Cramer, Donald G. Truhlar

Department of Chemistry and Supercomputer Institute, University of Minnesota, Minneapolis, MN 55455-0431, USA

Received: 6 August 2004 / Accepted: 9 September 2004 / Published online: 27 January 2005
© Springer-Verlag 2005

Abstract. Hybrid density functional theory, which is a combined Hartree–Fock and density functional method, provides a simple but effective way to incorporate nonlocal exchange effects and static and dynamical correlation energy into an orbital-based theory with affordable computational cost for many important problems of gas-phase chemistry. The inclusion of a reaction field representing an implicit solvent in a self-consistent hybrid density functional calculation provides an effective and efficient way to extend this approach to problems of liquid-phase chemistry. In previous work, we have parameterized several models based on this approach, and in the present article, we present several new parameterizations based on implicit solvation models SM5.43 and SM5.43R. In particular, we extend the applicability of these solvation models to several combinations of the MPW X hybrid-density functional with various one-electron basis sets, where MPW X denotes a combination of Barone and Adamo’s modified version of Perdew and Wang’s exchange functional, Perdew and Wang’s correlation functional, and a percentage X of exact Hartree–Fock exchange. SM5.43R parameter optimizations are presented for the MPW X /MIDI!, MPW X /MIDI!6D, and MPW X /6-31+G(d,p) combinations with $X=0$ (i.e., pure density functional theory), 25, 42.8, and 60.6, and for MPW X /6-31G(d) and MPW X /6-31+G(d), with $X=0$, 42.8, and 60.6; this constitutes a total of 18 new parameter sets. [Note that parameter optimizations using MPW25/6-31G(d) and MPW25/6-31+G(d) were carried out in a previous SM5.43R parameterization.] For each of the five basis sets, we found no significant loss in the accuracy of the model when parameters averaged over the four values of X are used instead of the parameters

optimized for a specific value of X . Therefore for each of the five basis sets used here, the SM5.43R and SM5.43 models are defined to have a single parameter set that can be used for any value of X between 0 and 60.6. The new models yield accurate free energies of solvation for a broad range of solutes in both water and organic solvents. On the average, the mean-unsigned errors, as compared with those from experiment, of the free energies of solvation of neutral solutes range from 0.50 to 0.55 kcal/mol and those for ions range from 4.5 to 4.9 kcal/mol. Since the SM5.43R model computes the free energy of solvation as a sum of bulk-electrostatic and non-bulk-electrostatic contributions, it may be used for detailed analysis of the physical effects underlying a calculation of the free energy of solvation. Several calculations illustrating the partitioning of these contributions for a variety of solutes in *n*-hexadecane, 1-octanol, and water are presented.

1. Introduction

Chemists have developed many theoretical models that can be used to make accurate predictions of chemical properties and processes in the gas phase. For a practical application to a given system of interest, the question one often asks is “What methods provide accurate results with minimal computational cost?” The success of Kohn–Sham density functional theory (DFT) [1] in its many and varied implementations and refinements [2–46], including hybrid DFT [10, 13–16, 18, 23, 25, 26, 29, 31–46], makes it the clear choice for many problems in chemistry, biochemistry, and materials science.

Because most chemical and biochemical phenomena occur in liquid solution, there is also great demand for

Correspondence to: C.J. Cramer and D.G. Truhlar
e-mail: cramer@pollux.chem.umn.edu; truhlar@umn.edu

accurate and inexpensive theoretical models that account for the effects of solvation. Two alternative methods that incorporate these effects are explicit-solvent models and continuum-solvent models. In the former approach, interactions between the solute and specific solvent molecules are considered. In the continuum-solvent approach, the solvent molecules are implicitly represented only by a continuous and homogeneous dielectric medium, as well as (in the more accurate models) an implicit treatment of first solvation shell effects, where “implicit” denotes that the coordinates of individual molecules are not specified, instead the thermally and spatially averaged effects of collective and equilibrated solvent molecules are represented by continuous functions of solute coordinates. Thus one considers interactions between the solute and the surrounding liquid medium as a whole. Each approach, explicit and continuum (i.e., implicit), has its strengths and weaknesses, and various DFT and hybrid DFT methods have been successfully applied within the context of both of them [47–87]. For example, Car and Parrinello [47] have implemented DFT into quantum-mechanical molecular dynamics simulations that permit the use of explicit solvent molecules and the quantum-mechanical treatment of the solute–solvent system that could provide detailed insight into the nature of the system being studied. Molecular dynamics calculations can in principle account for non-equilibrium solvent effects, and this has been illustrated in a number of investigations, such as the study of a water molecule in liquid water [64] and the study of the first step in the electrophilic bromination of ethene in liquid water [61]. DFT methods have also been used with various formulations of the polarizable continuum model (PCM) [49, 53–56, 58, 59, 66, 80, 81, 88], with continuum solvation models based on the generalized Born (GB) method [68, 72, 86, 89–95], and with continuum solvation models involving numerical solution of the Poisson or Poisson–Boltzman equations [48, 82]. The simplification of the solvent allows these continuum models to efficiently calculate free energies of transfer of solutes from one phase to another [96–101], which can be related to various interesting thermodynamic quantities, such as the free energy of solvation [48, 49, 53, 54, 56, 58, 59, 66, 68, 72, 80, 82, 86, 88, 92–95, 98–100], partition coefficient [99], vapor pressure [102, 103], solubility [103, 104], and non-equilibrium solvation effects [100, 105–114].

We have recently developed a new semiempirical continuum Solvation Model, Called Solvation Model 5.43R (SM5.43R) [86]. The SM5.43R model is based on class IV charges [115] obtained by charge model 3 (CM3) [116, 117] with redistributed Löwdin population analysis [118], the GB model [68, 72, 86, 89, 90–95], and atomic surface tensions [95, 119–123], that are multiplied by the solvent-accessible surface areas (SASAs) of atom-centered spheres. It has been parameterized for several different types of hybrid DFT methods [86]. The same parameters are also used with

an extended model, called SM5.43 (further explained later). The SM5.43R and SM5.43 solvation models are part of a more general suite of continuum solvation models, which are denoted as SM x . These models can be used to predict the standard-state free energy of transfer of a given solute between the gas phase, water, or any organic solvent, provided, in the latter case, that a certain set of solvent descriptors for that solvent are known or can be estimated.

It is interesting to compare this approach with another approach for predicting solvation free energies in a variety of solvents, in particular, with the popular PCMs. Like the SM x models, the PCMs represent the solute as a set of overlapping spheres; however, in other respects, the two kinds of models are quite different. In the SM x models, the spheres all correspond to atoms, whereas in the solute model used by the PCMs recommended in the *Gaussian03* manual [124], the spheres correspond to united-atom groups; this is called the united-atom-for-Hartree–Fock (UAHF) model [125], which was optimized [125] using HF/6-31G(d) specifically for water solvent. Furthermore, the SM x models treat cavitation, dispersion, and hydrogen bonding in a single calculation, without separating the effects, whereas the PCM treats them separately and adds the contributions. In particular, the PCM models nonelectrostatic contributions to the free energy of solvation as a sum of three terms corresponding to the free energies of cavitation, dispersion, and repulsion. Two different cavities (one for calculating the cavitation energy and one for calculating the dispersion–repulsion energy) are built from the UAHF solute model.

The free energy of cavitation in the PCM is based on an expression derived by Pierotti [126], who represented the solute and solvent as hard spheres; the model was generalized to arbitrarily shaped solutes for the PCM [59]. The cavity model used for this contribution to the free energy of solvation is a SASA defined as the surface through the center of a spherical solvent rolling on the surface of the solute model. The solvent radius for each solvent available in *Gaussian03* is one half the molecular (solvent) Lennard-Jones parameter, σ , defined by Pierotti [126].

The free energies of dispersion and repulsion in the PCM are based on atom–atom potentials proposed by Kitaigorodskii [127] that have been further improved by Huron and Claverie [128, 129, 130], and they are calculated by a procedure described by Floris and Tomasi [131, 132]. These two terms involve several cavities, one for each kind of atom (atomic number) in the solvent, with the same solute model, augmented by a different atomic radius to create an atomic-number-specific solvent-excluded surface area (SESA) for each kind of atom present in the solvent [133]. Each SESA corresponding to each atomic number in the solvent depends on a value for the radius of the solvent atom, which is taken to be the atomic van der Waals radius of Bondi [134].

In a recent comparison [86] of these two approaches, we found that the SM x approach is very accurate. In particular, as compared with the C-PCM [56, 135, 136], D-PCM [59, 80, 88, 137], and IEF-PCM [80, 138–140] models in *Gaussian03* [124], the mean unsigned errors (MUEs) for aqueous free energies of solvation of ions are a factor of 1.3–2.2 smaller (depending on the PCM method), and those for neutrals are a factor of 1.5–2.4 smaller. In addition to water solvent, *Gaussian03* supports 16 of the 90 organic solvents in the SM5.43R training set, and the PCM methods have MUEs of 3.8–3.9 kcal/mol in these, whereas SM5.43R has a MUE of 0.56 kcal/mol (a factor of 6.7–7.0 improvement). Furthermore, the SM5.43R model yields a MUE of 0.54 kcal/mol in the other 74 organic solvents. Therefore we have concluded that the SM x approach is worthy of further development.

The SM5.43R model uses any reasonable gas-phase geometry of the solute, hence the R, for “Rigid”, in its notation. In addition, geometry optimizations in the liquid phase using analytical free-energy gradients [141] and potential of mean force [96, 97, 101, 113, 142, 143] calculations can be efficiently carried out. Analytical free-energy gradients can also be used for numerical liquid-phase Hessian calculations [113, 142–145]. For liquid-phase geometry optimizations, liquid-phase Hessian calculations, and potential of mean force calculations, the R is omitted from the notation; that is, SM5.43R becomes SM5.43.

In this paper, we extend SM5.43R and SM5.43 to 18 additional combinations of electronic structure method and basis set. In particular, we carry out parameter optimizations with the MPW X hybrid-DFT method and the MIDI! [146, 147], MIDI!6D [146, 147], 6-31G(d) [148], 6-31+G(d) [148], and 6-31+G(d,p) [148] basis sets. The MPW X electronic structure method uses Barone and Adamo’s [25] modified version of Perdew and Wang’s [9] exchange functional, Perdew and Wang’s PW91 correlation functional [9], and a percentage X of exact HF exchange; we consider four values of X , namely 0 (which defines the pure DFT functional [25] mPWPW91), 25 [which gives a standard method called mPW1PW91 or (sometimes) MPW25], 42.8 (which gives a standard method [33] called MPW1K), and 60.6 (which has been called [116] MPW1KK). Note that there are 18 additional parameterizations (not 20) because we previously reported parameters for MPW25/6-31G(d) and MPW25/6-31+G(d) [86]. Note that MPW25 also goes by the names mPW0 and mPW1PW91.

The motivation for parameterizing the MPW X approach is that the optimum value of X depends on the problem. For a given property of a given molecule, or a given reaction, or a given range of molecules and/or properties, it may be useful to optimize X in the gas phase or in solution [33, 39, 41, 42, 44, 46], then use that optimum value of X for further related studies, within its range of appropriateness. (For example, it has been suggested that $X=42.8$ is optimized for kinetics [33] and

that $X=6$ is optimized for conformations of sugars [41], whereas $X=25$ is recommended for calculating heats of formation [25].) Then, having chosen X in this way, it is useful to have a solvation model already parameterized for that value of X . As long as X is in the range 0–60.6, this paper provides that parameterization.

2. Background: SM5.43R and SM5.43

The details of the SM x continuum solvation models were presented previously [68, 72, 86, 95, 149–163], so we will only give a brief overview here.

We begin by approximating the solvent as a continuous and homogeneous dielectric medium that is characterized by the solvent’s bulk dielectric constant, ϵ . In this approximation, contributions to the free energy of solvation are due to interactions between the solute and the dielectric medium rather than to interactions with specific solvent molecules, and they are estimated from a self-consistent molecular orbital calculation [98, 164, 165] employing the GB method [92–95]. This treatment of the solute–solvent system is able to account for long-range, bulk-electrostatic interactions, but it fails to account for important first-shell solvation effects due to interactions between the solute and solvent in the near vicinity of the solute; such interactions are commonly denoted as short-range interactions and they include the nonbulk character of the electrostatic interactions between the solute and solvent molecules in the first solvation shell. In order to account for these contributions to the free energy of solvation, we assume that they are proportional to the SASAs [119, 166] of the atoms in the solute. In addition, we further characterize the solvent with several solvent descriptors, which are included in the constants of proportionality that multiply the SASAs. The solvent descriptors are the refractive index (at the wavelength of the Na D line), n ; Abraham’s [167–171] hydrogen bond acidity parameter, α (which Abraham denotes as $\sum\alpha_2$); Abraham’s [167–171] hydrogen bond basicity parameter, β (which Abraham denotes as $\sum\beta_2$); the reduced surface tension, γ , which equals γ_m/γ° , where γ_m is the macroscopic surface tension at a liquid–air interface at 298 K, and γ° is 1 cal/mol/Å²; the square of the fraction of nonhydrogenic solvent atoms that are aromatic carbon atoms (carbon aromaticity), ϕ^2 ; and the square of the fraction of nonhydrogenic solvent atoms that are F, Cl, or Br (electronegative halogenicity), ψ^2 .

In the SM x models (including SM5.43R and SM5.43), the standard-state free energy of solvation, ΔG_S° , of a solute in a liquid solvent is

$$\Delta G_S^\circ = \Delta E + G_P + G_{\text{CDS}} + \Delta G_{\text{vib-rot}} + \Delta G_{\text{lib}}^\circ, \quad (1)$$

where ΔE is the change in the solute’s internal electronic energy (including solute nuclear repulsion) when it is placed in the dielectric medium, G_P is the elec-

tronic polarization energy due to mutual polarization of the solute and solvent, G_{CDS} is a semiempirical term that is the estimate of the deviation between the bulk-electrostatic contribution to the free energy of solvation ($\Delta E + G_{\text{P}}$) and the true solvation free energy, $\Delta G_{\text{vib-rot}}$ is the change in vibrational-rotational free energy upon solvation, and $\Delta G_{\text{lib}}^{\circ}$ accounts for any changes in concentration between the gas-phase and the liquid-phase, i.e., for the free energy of liberation associated with the volume available to a molecule. We employ the same concentration in both phases (1 mol/L), so $\Delta G_{\text{lib}}^{\circ}$ is zero [103, 172]. Note that if a geometry optimization in the liquid phase is carried out, then ΔE also accounts for the change in the solute's internal electronic energy due to the relaxation of the nuclei in the liquid phase. In the SM5.43R model, the solute nuclear repulsion contribution to ΔE is zero because the geometry is not relaxed in solution (i.e., geometry optimization in solution is not carried out). Furthermore, $\Delta G_{\text{vib-rot}}$ is zero because it is absorbed into the parameters. In SM5.43, both of these contributions to the free energy of solvation should be calculated (at least in principle; $\Delta G_{\text{vib-rot}}$ is neglected in this paper because it is small).

The GB approximation is an extension of the Born equation [173] for solvation of atomic ions to molecules. It predicts the electronic polarization energy to be

$$G_{\text{P}} = -\frac{1}{2} \left(1 - \frac{1}{\epsilon}\right) \sum_{k,k'} q_k q_{k'} \gamma_{kk'}, \quad (2)$$

where the summation goes over all atoms in the solute, q_k is the partial atomic charge on atom k , and $\gamma_{kk'}$ is a Coulomb integral involving atoms k and k' . In SM5.43R, the partial atomic charges are obtained from CM3 [116, 117], and they are computed self-consistently, i.e., they are free to polarize in the presence of the dielectric medium. For $k \neq k'$, we use a form [154] of the Coulomb integral that was originally proposed by Still et al. [94], which is

$$\gamma_{kk'} = \{R_{kk'}^2 + \alpha_k \alpha_{k'} \exp [(-R_{kk'}^2/d_{kk'} \alpha_k \alpha_{k'})]\}^{-1/2}, \quad (3)$$

where $R_{kk'}$ is the distance between atoms k and k' , $d_{kk'}$ is a parameter that can be optimized, and α_k is the effective Born radius of atom k . The effective Born radius of a given atom in the solute accounts for the displacement of the dielectric medium in the region of space in which the electric field lines connected to each atom of the solute reside; the other atoms effectively descreen the atom in question from the dielectric medium. The effective Born radius for atom k is [94, 95, 174]

$$\alpha_k = \left(\frac{1}{R'} + \int_{\rho_k}^{R'} \frac{A_k(R, r, \{\rho_{k'}\})}{4\pi r^4} dr \right)^{-1}, \quad (4)$$

where R' is the value of the radius of the sphere centered on atom k that completely engulfs all other spheres centered on the other atoms in the solute and $A_k(\mathbf{R}, r, \{\rho_{k'}\})$ is the exposed area of a sphere of radius r that is centered on atom k . The exposed area is determined by the ASA algorithm [174]. This algorithm takes as input the geometry of the solute, \mathbf{R} , the radius r centered on atom k , and the radii of the spheres centered on all the other atoms in the solute, which are given by the set of empirical intrinsic Coulomb radii, $\{\rho_{k'}\}$.

While the GB method makes reasonably accurate predictions of the bulk-electrostatic contribution to the free energy of solvation for small molecules [94, 175–178], at least as compared with numerical solutions of the Poisson equation, it has been shown to yield less reliable predictions for larger macromolecules [175, 177]. In particular, the GB method used in conjunction with Eq. (4) overestimated the contributions to the bulk-electrostatic solvation free energy from interior atoms in several protein molecules. Onufriev et al. [179] submitted that this was a result of the failure of Eq. (4) to account for the nonzero size of the solvent. They have therefore proposed a new equation for the calculation of the effective Born radii, which is [179]

$$\alpha_k = \left(\frac{(1-\lambda)}{\rho_k} + \frac{\lambda}{R'} + \lambda \int_{\rho_k}^{R'} \frac{A_k(\mathbf{R}, r, \{\rho_{k'}\})}{4\pi r^4} dr \right)^{-1}, \quad (5)$$

where λ has been empirically determined by Onufriev et al. to be 1.4. Note that Eq. (5) is identical to Eq. (4) when $\lambda = 1$. Onufriev et al. [179] have found that the new GB method accurately predicts $\text{p}K_{\text{a}}$ shifts in several proteins where such shifts are associated with large conformational changes. Furthermore, they have shown that the new GB method still provides reliable and accurate results for small molecules. In an effort to make our solvation models more generally applicable to larger biomolecules, we will consider using Eq. (5) instead of Eq. (4) for the calculation of effective Born radii.

The G_{CDS} term can be written as a sum of two terms:

$$G_{\text{CDS}} = G_{\text{CDS}}^{[1]} + G_{\text{CDS}}^{[2]} \quad (6)$$

where

$$G_{\text{CDS}}^{[1]} = \sum_{\delta=1}^3 S_{\delta} \sum_k A_k(\mathbf{R}, \{r_{k'} + r_{\text{S}}\}) \sum_j \sigma_{Z_k j \delta}^{[1]} f_{Z_k j}(\{Z_{k'}, R_{k'k''}\}) \quad (7)$$

and where the index δ goes over the solvent descriptors n , α , and β , S_{δ} is the value of each of these three descriptors, the index k goes over all atoms in the solute, A_k is the SASA of atom k , which depends on the geometry, \mathbf{R} , of the solute and on a set of van der Waals solute radii $r_{k'}$ augmented by a solvent radius r_{S} , Z_k is the atomic number of atom k , the index j goes over the

geometry-dependent functions $f_{Z_{kj}}$, which are used to distinguish between different types of functional groups in the solute that may be near atom k (their functional forms are described in detail elsewhere [154, 156–159, 162]), $R_{k'k''}$ is the interatomic distance between atoms k' and k'' , and $\sigma_{Z_{kj}\delta}^{[1]}$ is an atomic surface tension coefficient (a parameter to optimize). The second term in Eq. (6) is

$$G_{\text{CDS}}^{[2]} = \sum_{\delta=4}^7 S_{\delta} \sigma_{\delta}^{[2]} \sum_k A_k(\mathbf{R}, \{r_{k'} + r_S\}), \quad (8)$$

where here the index δ goes over the solvent descriptors γ , β^2 , ϕ^2 , and ψ^2 , and $\sigma_{\delta}^{[2]}$ is a molecular surface tension coefficient (this optimized parameter does not depend on atomic number). Note that in previous work [68, 72, 86, 154, 161] and in the present work, water has its own set of atomic surface tension coefficients, so neither Eq. (8) nor the solvent descriptors in Eq. (7) are needed for water, and the set of $\{\sigma_{Z_{j\delta}}^{[1]}\}$ are denoted as $\{\sigma_{Z_{j,\text{water}}}^{[1]}\}$. Thus, for water

$$G_{\text{CDS}} = \sum_k A_k(\mathbf{R}, \{r_{k'} + r_S\}) \sum_j \sigma_{Z_{kj,\text{water}}}^{[1]} f_{Z_{kj}}(\{Z_{k'}, R_{k'k''}\}). \quad (9)$$

We note that the solvent dependence is entirely contained in the S_{δ} values; r_S is independent of solvent. Furthermore, SM5.43R can make solvation free energy predictions of solutes in any organic solvent if the solvent descriptors for the solvent of interest are known experimentally or can be estimated. A database of solvent descriptors for about 175 organic solvents is available on the internet [180]. In addition, some or all of these data have been made available in previous work [159, 181]. Because the models can be applied to any solvent for which the descriptors are known or can be estimated, they are called universal solvation models.

3. Parameter optimization for SM5.43R

3.1. Parameters to optimize

There are five types of parameters to consider optimizing to obtain the new SM5.43R: (1) the values of the empirical Coulomb radii used in the calculation of the effective Born radii, (2) the values of the van der Waals radii used in Eqs. (7), (8), and (9), (3) the solvent radius, (4) the atomic and molecular surface tension parameters, and (5) the $d_{kk'}$ parameters in Eq. (3). In this parameterization, as in all of our previous parameterizations, we will use the van der Waals radii of Bondi [134] for the parameters of type 2 from the initial [86] SM5.43R parameterization. However, because we will consider using Eq. (5) instead of Eq. (4) for the calculation of the effective Born radii, we will

consider reoptimizing these radii for use with Eq. (5), and we will consider optimizing λ . We will only accept newly optimized values, however, if the accuracy of the resulting predictions of the aqueous free energies of solvation of both neutral and ionic solutes is not compromised. We will use a solvent radius of 0.4 Å, which was previously optimized for hydrophobic interactions in water [86, 182]. In terms of predicting accurate free energies of solvation, we have found that the overall optimal value of the solvent radius is zero when a single value is optimized for all solvents, and that errors in predicted free energies of solvation increase as the value of the solvent radius increases. Our motivation for using a nonzero value for the solvent radius is that it leads to a more physically meaningful definition of the SASAs of the atoms in the solute. While a different nonzero value would probably be obtained if the optimization were carried out for different potentials of mean force (PMF) in different solvents, we want to use the same value in all solvents so that our model can easily be applied in a universal way. Among the PMFs that one could consider using, the methane–methane PMF in water is certainly a good choice because of the importance of hydrophobic interactions. Thus, the value of 0.4 Å is a good compromise between obtaining a meaningful definition of the SASAs of the atoms in the solute and predicting accurate free energies of solvation. Finally, we will also optimize the parameters of type 4, and take the parameters of type 5 to be the ones determined previously [68]. (In particular, $d_{\text{HC}} = d_{\text{CH}} = 4.2$ and $d_{kk'} = 3.9$.)

3.2. Training set

We will optimize the parameters of types 1 and 4 using the training set of data used in the original [86] SM5.43R parameterization. This training set consists of 257 experimental aqueous free energies of solvation of neutral solutes, 1,980 experimental free energies of solvation of neutral solutes in 90 organic solvents, and 79 experimental free energies of transfer of neutral solutes between water and 12 organic solvents (which are a subset of the 90 organic solvents). In addition, this training set consists of 47 experimental free energies of solvation of ionic solutes in water. All solutes contain at most H, C, N, O, F, P, S, Cl, and Br.

3.3. Optimization protocol

All parameters are optimized for a temperature of 298 K, and the geometries of all solutes are optimized at the mPW1PW91/MIDI! level. We will optimize the empirical Coulomb radii and the atomic and molecular surface tension coefficients using the same procedure described in our previous SM5.43R parameter optimization [86].

3.4. Basis sets

We will carry out SM5.43R parameter optimizations for MPW X using the MIDI! [146, 147], MIDI!6D [146, 147], 6-31G(d) [148], 6-31+G(d) [148], and 6-31+G(d,p) [148] basis sets. The MIDI! and MIDI!6D basis sets were developed for H, Li, C, N, O, F, Si, P, S, Cl, Br, and I. They are small, split-valence basis sets that contain polarization functions on all heavy atoms except carbon. The MIDI! basis set uses five spherical harmonic d functions, and all other basis sets used here use six Cartesian d functions.

4. Results and discussion

4.1. The empirical Coulomb radii

Preliminary results of the parameter optimization of the empirical Coulomb radii using Eq. (5) to calculate the effective Born radii yielded larger values of the intrinsic Coulomb radii. The predicted free energies of solvation of neutral solutes were insensitive to the resulting larger effective Coulomb radii. However, this was not the case for ionic solutes, where the errors in predicted free energies of solvation were on the order of a factor of 2 larger than when Eq. (4) is used to calculate the effective Born radii. On the basis of these results, we decided to continue to use Eq. (4) and the empirical Coulomb radii determined in the first SM5.43R parameterization.

4.2. Using averaged surface tension parameters versus using specifically optimized surface tension parameters

We have carried out parameter optimizations for the SM5.43R continuum solvation model with MPW X /MIDI! [146, 147], MPW X /MIDI!6D [146, 147], MPW X /6-31G(d) [148], MPW X /6-31+G(d) [148], and MPW X /6-31+G(d,p) [148], for $X=0, 25, 42.8, \text{ and } 60.6$, i.e., 20 specific parameter sets in all. (Note that the parameter optimizations for MPW25/6-31G(d) and MPW25/6-31+G(d) were carried out in previous work [86].) For a given basis set, we found that the optimized atomic and molecular surface tension parameters for each of the four values of X in MPW X were so similar to one another that we determined the final atomic and molecular surface tension parameters to be the average of the four sets of optimized surface tension parameters, i.e., five averaged parameter sets in all, where each averaged parameter set corresponds to one of the five basis sets used in this paper.

We assess the validity of using SM5.43R models that use averaged surface tension parameters by making comparisons between the accuracy of these models and SM5.43R models using specifically optimized surface tension parameters. Table 1 shows the MUEs (in kilocalories per mole), as compared with those obtained

from experiment, of the free energies of solvation for various solute classes in the SM5.43R training set calculated by SM5.43R/MPW X /MIDI!, for $X=0, 25, 42.8, \text{ and } 60.6$, using averaged surface tension parameters. Table 1 also gives the corresponding MUEs calculated by SM5.43R models using surface tension parameters specifically optimized for each value of X in MPW X /MIDI! (e.g., column 6 of Table 1 corresponds to MUEs calculated by SM5.43R/MPW25/MIDI! using atomic and molecular surface tension parameters optimized specifically for the MPW25/MIDI! method). Table 2 shows the same data as Table 1, but the MUEs are calculated by various SM5.43R models using MPW X /6-31+G(d,p). The 6-31+G(d,p) basis set is the largest basis set considered in this paper, and the MIDI! basis set is the smallest. Tables 1 and 2 show that there is no substantial loss in accuracy of predicted free energies of solvation when averaged surface tension parameters are used instead of ones specifically optimized for a particular combination of MPW X and basis set. The largest difference between a MUE calculated with averaged parameters and a MUE calculated with specific parameters is 0.17 and 0.40 kcal/mol for MIDI! and 6-31+G(d,p), respectively. These MUEs both correspond to ionic solutes containing second-row atoms and/or bromine in water. These differences in MUEs are not too alarming, however, because the experimental uncertainty of aqueous solvation energies of ions is on the order of 5 kcal/mol [183, 184]. Note that we observe similar trends for all of the other basis sets used in this work. All results presented in the rest of this paper are therefore based on using averaged surface tension parameters.

4.3. Parameters

Tables 3, 4, 5, 6, 7, and 8 show the averaged atomic and molecular surface tension parameter sets corresponding to the MIDI!, MIDI!6D, 6-31G(d), 6-31+G(d), and 6-31+G(d,p) basis sets, respectively. The values of the parameters from the five basis sets are generally similar in magnitude. When the five sets of parameters are compared, the standard deviations of corresponding parameters from all the basis sets are small; approximately 76% of the nonzero parameters have standard deviations smaller than 10 cal/mol \AA^2 and approximately 92% of the nonzero parameters have standard deviations smaller than 20 cal/mol \AA^2 . In addition, the signs of the majority of the parameters from each basis set are the same. These are encouraging results, as they indicate that CM3 delivers consistent charge distributions. Thus, consistent predictions of ΔG_{EP} and consistent corrections to it from G_{CDS} are obtained, regardless of which basis set is used. These results also illustrate the success of the redistributed Löwdin population analysis method to predict more reliable charge distributions than the Löwdin population analysis method when diffuse functions are included in the basis set [118]. It is very

Table 1. Mean-unsigned errors (MUEs, kcal/mol) of the free energies of solvation for solutes in various solute classes of the SM5.43R training set calculated by various SM5.43R/MPW X /MIDI! solvation models

Solute class	No. of data ^a	$X=0$		$X=25.0$		$X=42.8$		$X=60.6$	
		Avg. ^b	Opt. ^c	Avg. ^b	Opt. ^c	Avg. ^b	Opt. ^c	Avg. ^b	Opt. ^c
In water solvent									
Neutral H, C, N, O, F	171	0.55	0.55	0.55	0.55	0.55	0.54	0.55	0.54
Cl, Br, S, and P neutrals	86	0.49	0.47	0.49	0.49	0.49	0.49	0.50	0.50
All neutrals	257	0.53	0.52	0.53	0.53	0.53	0.53	0.53	0.53
All H, C, N, O, F ions	32	5.03	5.03	4.98	4.98	4.94	4.94	4.90	4.90
All Cl, Br, P, S ions	15	4.59	4.41	4.45	4.28	4.37	4.21	4.31	4.15
All ions	47	4.89	4.83	4.81	4.75	4.76	4.71	4.71	4.66
In organic solvents									
All H, C, N, O absolute free energies	1,733	0.51	0.50	0.50	0.50	0.50	0.50	0.51	0.50
All H, C, N, O transfer free energies	37	0.45	0.45	0.45	0.45	0.44	0.44	0.44	0.44
All H, C, N, O data	1,770	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
F, Cl, Br, S absolute free energies	210	0.62	0.62	0.62	0.62	0.61	0.61	0.61	0.62
F, Cl, Br, S transfer free energies	33	0.63	0.62	0.62	0.62	0.62	0.62	0.62	0.62
H, C, N, O, F, S, Cl data	243	0.62	0.62	0.62	0.62	0.61	0.62	0.61	0.62
All P absolute free energies	37	1.39	1.39	1.40	1.40	1.40	1.40	1.41	1.40
All P transfer free energies	9	0.57	0.57	0.59	0.59	0.59	0.60	0.61	0.61
All P data	46	1.23	1.23	1.24	1.24	1.24	1.24	1.25	1.25
All absolute free energies	1,980	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.53
All transfer free energies	79	0.54	0.54	0.54	0.54	0.53	0.54	0.53	0.54
All data	2,059	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.53

^aNumber of experimental data in a given solute class.^bSM5.43R/MPW X /MIDI! using averaged surface tension parameters from MPW X /MIDI! ($X=0, 25, 42.8, \text{ and } 60.6$).^cSM5.43R/MPW X /MIDI! using specifically optimized surface tension parameters for the given value of X in MPW X /MIDI!.**Table 2.** MUEs (kcal/mol) of the free energies of solvation for solutes in various solute classes of the SM5.43R training set calculated by various SM5.43R/MPW X /6-31 + G(d,p) solvation models

Solute class	No. of data ^a	$X=0$		$X=25.0$		$X=42.8$		$X=60.6$	
		Avg. ^b	Opt. ^c	Avg. ^b	Opt. ^c	Avg. ^b	Opt. ^c	Avg. ^b	Opt. ^c
In water solvent									
Neutral H, C, N, O, F	171	0.56	0.55	0.55	0.55	0.55	0.55	0.55	0.55
Cl, Br, S, and P neutrals	86	0.54	0.50	0.53	0.52	0.52	0.53	0.53	0.54
All neutrals	257	0.55	0.54	0.54	0.54	0.54	0.54	0.54	0.55
All H, C, N, O, F ions	32	5.22	5.22	5.17	5.17	5.15	5.15	5.12	5.12
All Cl, Br, P, S ions	15	4.22	3.83	4.16	3.72	4.11	3.68	4.06	3.64
All ions	47	4.90	4.78	4.85	4.71	4.82	4.68	4.78	4.65
Inorganic solvents									
All H, C, N, O absolute free energies	1,733	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49
All H, C, N, O transfer free energies	37	0.46	0.46	0.45	0.46	0.45	0.45	0.46	0.45
All H, C, N, O data	1,770	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49
F, Cl, Br, S absolute free energies	210	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
F, Cl, Br, S transfer free energies	33	0.64	0.63	0.63	0.62	0.62	0.62	0.61	0.62
H, C, N, O, F, S, Cl data	243	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68
All P absolute free energies	37	1.64	1.59	1.64	1.63	1.66	1.64	1.71	1.66
All P transfer free energies	9	0.62	0.67	0.69	0.68	0.69	0.67	0.68	0.67
All P data	46	1.44	1.41	1.45	1.45	1.47	1.45	1.51	1.47
All absolute free energies	1,980	0.53	0.53	0.53	0.53	0.54	0.54	0.54	0.54
All transfer free energies	79	0.55	0.56	0.55	0.55	0.55	0.55	0.55	0.55
All data	2,059	0.53	0.53	0.53	0.53	0.54	0.54	0.54	0.54

^aNumber of experimental data in a given solute class.^bSM5.43R/MPW X /6-31 + G(d,p) using averaged surface tension parameters from MPW X /6-31 + G(d,p) ($X=0, 25, 42.8, \text{ and } 60.6$).^cSM5.43R/MPW X /6-31 + G(d,p) using specifically optimized surface tension parameters for the given value of X in MPW X /6-31 + G(d,p).

encouraging that our preliminary work in developing CM3 and the redistributed Löwdin population analysis method is confirmed in this way to have been worthwhile.

4.4. Performance of SM5.43R for solutes in water solvent

Tables 9, 10, 11, 12, and 13 give the MUEs, as compared with experimental values, of the aqueous free energies of

Table 3. Averaged atomic surface tension parameters (cal/molÅ²) for MPWX/MIDI!. An entry of 0.00 indicates that this parameter was not used. The recommended range of X is 0.0–60.6

k	$\sigma_{k,\text{water}}^{[1]}$	$\sigma_{k,n}^{[1]}$	$\sigma_{k,x}^{[1]}$	$\sigma_{k,\beta}^{[1]}$
H	59.94	20.59	0.00	0.00
C	101.89	79.16	30.76	5.85
N	102.11	36.87	-51.97	77.32
O	-125.50	-15.20	-35.96	0.00
F	33.39	-2.20	191.70	-143.64
Cl	-3.22	-27.22	0.00	0.00
Br	-13.33	-35.62	0.00	0.00
S	-64.84	-54.92	0.00	0.00
P	-44.29	-22.06	0.00	0.00
H,C	-85.79	-70.05	0.00	0.00
C,C	-67.98	-75.80	0.00	0.00
C,C (2)	-28.45	-19.32	0.00	0.00
H,O	-189.38	-19.98	-196.41	-220.23
H,O (2)	283.49	110.86	0.00	0.00
O,C	146.14	32.14	0.00	12.58
O,O	37.14	-34.34	107.90	3.11
H,N	-146.55	-38.87	0.00	-66.49
H,N (2)	-203.57	-126.81	0.00	0.00
C,N	-36.48	-73.71	40.47	0.00
N,C	-54.22	-7.40	-33.66	-10.89
N,C (2)	-18.49	0.00	0.00	0.00
N,C (3)	-62.83	-16.65	0.00	0.00
O,N	198.85	44.72	24.59	16.81
H,S	31.84	39.30	0.00	0.00
S,S	34.51	-1.70	0.00	0.00
O,P	91.94	38.85	0.00	0.00
S,P	95.34	13.84	0.00	0.00

Table 5. Averaged atomic surface tension parameters (cal/molÅ²) for MPWX/6-31G(d). An entry of 0.00 indicates that this parameter was not used. The recommended range of X is 0.0–60.6

k	$\sigma_{k,\text{water}}^{[1]}$	$\sigma_{k,n}^{[1]}$	$\sigma_{k,x}^{[1]}$	$\sigma_{k,\beta}^{[1]}$
H	59.75	20.20	0.00	0.00
C	97.06	76.19	30.85	4.98
N	89.54	31.27	-56.08	83.43
O	-128.40	-16.80	-36.54	0.00
F	30.40	-3.42	194.02	-146.97
Cl	-2.46	-27.10	0.00	0.00
Br	-13.39	-35.59	0.00	0.00
S	-58.58	-53.84	0.00	0.00
P	-27.14	-9.84	0.00	0.00
H,C	-85.42	-69.34	0.00	0.00
C,C	-64.70	-74.04	0.00	0.00
C,C (2)	-26.03	-17.65	0.00	0.00
H,O	-153.44	-5.96	-186.34	-198.99
H,O (2)	308.11	120.99	0.00	0.00
O,C	153.28	33.04	0.00	17.01
O,O	38.64	-31.45	107.77	-6.61
H,N	-127.69	-30.01	0.00	-60.66
H,N (2)	-209.30	-131.67	0.00	0.00
C,N	-15.06	-62.41	47.76	0.00
N,C	-58.17	-9.57	-34.70	-12.01
N,C (2)	11.07	0.00	0.00	0.00
N,C (3)	-38.41	-7.28	0.00	0.00
O,N	204.03	44.50	29.55	17.13
H,S	22.54	39.23	0.00	0.00
S,S	26.39	-2.36	0.00	0.00
O,P	84.57	32.09	0.00	0.00
S,P	98.05	22.36	0.00	0.00

Table 4. Averaged atomic surface tension parameters (cal/molÅ²) for MPWX/MIDI!6D. An entry of 0.00 indicates that this parameter was not used. The recommended range of X is 0.0–60.6

k	$\sigma_{k,\text{water}}^{[1]}$	$\sigma_{k,n}^{[1]}$	$\sigma_{k,x}^{[1]}$	$\sigma_{k,\beta}^{[1]}$
H	59.92	20.49	0.00	0.00
C	102.53	79.44	30.39	6.07
N	119.57	45.02	-47.47	78.43
O	-120.28	-12.59	-33.84	0.00
F	33.77	-2.02	193.90	-145.62
Cl	-3.27	-27.28	0.00	0.00
Br	-13.13	-35.68	0.00	0.00
S	-64.60	-54.71	0.00	0.00
P	-45.01	-22.63	0.00	0.00
H,C	-85.55	-69.92	0.00	0.00
C,C	-68.95	-76.36	0.00	0.00
C,C (2)	-28.28	-19.19	0.00	0.00
H,O	-204.40	-27.46	-203.34	-221.28
H,O (2)	292.32	115.34	0.00	0.00
O,C	144.24	29.79	0.00	15.87
O,O	35.03	-35.32	108.58	3.36
H,N	-151.36	-41.46	0.00	-66.54
H,N (2)	-206.57	-127.88	0.00	0.00
C,N	-46.72	-78.03	31.85	0.00
N,C	-53.86	-7.39	-32.79	-10.19
N,C (2)	-26.69	0.00	0.00	0.00
N,C (3)	-84.25	-27.80	0.00	0.00
O,N	196.31	42.36	21.48	22.18
H,S	30.63	38.31	0.00	0.00
S,S	35.60	-1.08	0.00	0.00
O,P	92.89	38.29	0.00	0.00
S,P	94.88	13.21	0.00	0.00

Table 6. Averaged atomic surface tension parameters (cal/molÅ²) for MPWX/6-31+G(d). An entry of 0.00 indicates that this parameter was not used. The recommended range of X is 0.0–60.6

k	$\sigma_{k,\text{water}}^{[1]}$	$\sigma_{k,n}^{[1]}$	$\sigma_{k,x}^{[1]}$	$\sigma_{k,\beta}^{[1]}$
H	59.89	20.68	0.00	0.00
C	100.73	79.99	31.68	3.20
N	72.32	28.83	-58.66	71.08
O	-138.12	-26.03	-40.93	0.00
F	26.47	-4.75	190.25	-148.62
Cl	-1.73	-26.33	0.00	0.00
Br	-13.41	-36.02	0.00	0.00
S	-70.27	-54.53	0.00	0.00
P	-5.09	7.33	0.00	0.00
H,C	-85.68	-68.98	0.00	0.00
C,C	-66.25	-77.13	0.00	0.00
C,C (2)	-29.97	-20.05	0.00	0.00
H,O	-152.21	0.30	-186.24	-198.01
H,O (2)	356.50	145.07	0.00	0.00
O,C	154.85	41.00	0.00	8.29
O,O	38.33	-29.15	107.48	-19.57
H,N	-107.82	-26.00	0.00	-47.28
H,N (2)	-219.20	-135.57	0.00	0.00
C,N	-15.59	-69.50	74.75	0.00
N,C	-62.78	-11.36	-38.38	-11.95
N,C (2)	0.28	0.00	0.00	0.00
N,C (3)	-37.95	-8.06	0.00	0.00
O,N	203.57	50.96	33.99	10.04
H,S	52.07	39.69	0.00	0.00
S,S	51.08	4.40	0.00	0.00
O,P	138.77	59.81	0.00	0.00
S,P	104.04	28.20	0.00	0.00

Table 7. Averaged atomic surface tension parameters (cal/molÅ²) for MPW X /6-31+G(d,p). An entry of 0.00 indicates that this parameter was not used. The recommended range of X is 0.0–60.6

k	$\sigma_{k,\text{water}}^{[1]}$	$\sigma_{k,n}^{[1]}$	$\sigma_{k,z}^{[1]}$	$\sigma_{k,\beta}^{[1]}$
H	59.90	20.59	0.00	0.00
C	96.50	78.88	31.02	1.88
N	75.99	30.21	-55.46	72.03
O	-144.50	-28.19	-45.18	0.00
F	26.29	-4.61	186.36	-147.32
Cl	-1.55	-26.18	0.00	0.00
Br	-13.12	-36.08	0.00	0.00
S	-69.90	-53.27	0.00	0.00
P	-9.72	4.70	0.00	0.00
H,C	-84.86	-68.61	0.00	0.00
C,C	-64.94	-76.77	0.00	0.00
C,C (2)	-29.65	-20.46	0.00	0.00
H,O	-143.85	4.49	-182.51	-198.96
H,O (2)	360.53	146.64	0.00	0.00
O,C	167.20	46.30	0.00	16.44
O,O	36.31	-31.10	114.60	-23.69
H,N	-106.44	-25.55	0.00	-48.05
H,N (2)	-225.07	-137.39	0.00	0.00
C,N	-24.05	-73.81	64.84	0.00
N,C	-63.53	-11.49	-38.54	-12.31
N,C (2)	-31.16	0.00	0.00	0.00
N,C (3)	-31.02	-4.09	0.00	0.00
O,N	213.04	55.55	33.80	11.87
H,S	53.81	37.92	0.00	0.00
S,S	54.54	4.73	0.00	0.00
O,P	151.26	65.89	0.00	0.00
S,P	110.76	30.57	0.00	0.00

Table 8. Averaged molecular surface tension parameters (cal/molÅ²) for MPW X and the indicated basis set. The recommended range of X is 0.0–60.6

Method	$\sigma_{\gamma}^{[2]}$	$\sigma_{\phi^2}^{[2]}$	$\sigma_{\psi^2}^{[2]}$	$\sigma_{\beta^2}^{[2]}$
MIDI!	0.26	-3.05	-6.54	2.82
MIDI!6D	0.26	-3.08	-6.51	2.82
6-31G(d)	0.27	-3.12	-6.38	2.45
6-31+G(d)	0.26	-3.02	-6.28	2.70
6-31+G(d,p)	0.25	-3.03	-6.41	2.91

solvation for various solute classes calculated by SM5.43R using MPW X , for $X=0, 25, 42.8,$ and $60.6,$ and each of the five basis sets used in this work. On average, all SM5.43R methods yield similar errors. The MUEs range from 0.50 to 0.55 kcal/mol for all 257 neutral solutes. For ionic solutes, the MUEs are larger, ranging from 4.3 to 5.2 kcal/mol.

In general, the MUEs for neutral solute classes are between 0.2 and 0.6 kcal/mol.; However, there are several solute classes that have MUEs outside of this range, in particular, amides and ureas (four data), aromatic nitrogen heterocycles (ten data), halogenated compounds containing N and/or O (nine data) or S (two data), and phosphorus compounds containing S (three data). We know from previous parameterizations that

these types of compounds are difficult to model. They do, however, add diversity to the training set, and leaving them in the training set should provide more robust and reliable solvation model parameters.

As we found with the previous SM5.43R parameterizations [86], the addition of diffuse functions to the 6-31G(d) basis set does not significantly increase the error, as one might have found if the difficulty of carrying out the population analysis with diffuse basis sets were to have caused inaccurate partial charges and hence inaccurate electrostatic contributions to solvation energies. In fact, the use of diffuse functions improves the prediction of aqueous free energies of solvation for neutral solutes containing H, C, N, and O, such as nitrohydrocarbons, amides, and ureas. This is encouraging because diffuse functions are known [41] to lead to improved relative energies of gas-phase species for DFT and hybrid DFT, and the ultimate quality of the predictions of a liquid-phase property depends both on the quality of the prediction of the gas-phase value and the prediction of the change upon solvation.

4.5. Performance of SM5.43R for solutes in organic solvents in terms of solute functional class

Tables 14, 15, 16, 17, and 18 show the MUEs, as compared with experimental values, of the free energies of solvation for various solute classes in the 90 organic solvents of our training set calculated by various SM5.43R methods. Note that the set of solvation energy data used in these tables is obtained by removing all experimental aqueous free energy of solvation data from the SM5.43R training set. On the average, the MUEs calculated by all methods are between 0.51 and 0.54 kcal/mol. For solutes containing at most H, C, N, and O, the MUEs are approximately 0.5 kcal/mol, while the MUEs increase to 0.62–0.69 kcal/mol for solutes containing second-row atoms and/or bromine. As in the aqueous training set, the MUEs for solutes containing halogens and N and/or O and solutes containing P and S are large relative to the MUEs for other solute classes. The SM5.43R methods using either the 6-31+G(d) or 6-31+G(d,p) basis set are more accurate than the other SM5.43R methods for predicting free energies of solvation of amides, ureas, and nitrohydrocarbons. However, they are not as accurate as the other methods for predicting solvation energies of solutes containing P.

There is little to no variation in the MUE with respect to the value of X in MPW X , and this trend further illustrates that CM3 provides consistent charge distributions for use in the calculation of ΔG_{EP} . For most solute classes in the organic training set and for each of the five basis sets used, the variations are typically less than 0.05 kcal/mol. The largest variations generally occur for inorganic compounds, fluoroalkanes, and fluoroarenes.

Table 9. MUEs (kcal/mol) of the aqueous free energies of solvation for various solute classes calculated by SM5.43R/MPW X /MIDI! ($X=0, 25, 42.8,$ and 60.6). Averaged values of the surface tension parameters from MPW X /MIDI! are used

Solute class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Inorganic H, C, N	4	0.31	0.27	0.30	0.37
Unbranched alkanes	8	0.64	0.64	0.64	0.64
Branched alkanes	5	0.47	0.48	0.48	0.48
Cycloalkanes	5	0.61	0.60	0.60	0.60
Alkenes	9	0.37	0.37	0.37	0.37
Alkynes	5	0.24	0.24	0.23	0.23
Arenes	8	0.35	0.36	0.36	0.38
Alcohols	12	0.55	0.53	0.51	0.50
Phenols	4	0.31	0.35	0.38	0.44
Ethers	12	0.47	0.46	0.47	0.48
Aldehydes	6	0.40	0.43	0.39	0.37
Ketones	12	0.30	0.30	0.29	0.29
Carboxylic acids	5	0.40	0.32	0.30	0.29
Esters	13	0.29	0.30	0.32	0.34
Bifunctional H, C, O compounds	5	0.39	0.42	0.43	0.44
All H, C, O compounds	109	0.41	0.41	0.41	0.42
Aliphatic amines	15	0.91	0.90	0.89	0.88
Anilines	1	0.24	0.27	0.28	0.26
Ar N heterocycles (1 N in ring)	10	1.08	1.07	1.08	1.06
Ar N heterocycles (2 Ns in ring)	2	0.47	0.47	0.53	0.51
Nitriles	4	0.61	0.61	0.63	0.62
Bifunctional H, C, N compounds	1	0.36	0.45	0.49	0.59
H, C, N compounds	33	0.86	0.86	0.86	0.85
Amides	3	1.22	1.17	1.18	1.17
Ureas	1	3.09	3.12	3.17	3.18
Nitrohydrocarbons	7	0.62	0.59	0.57	0.57
Bifunctional H, C, N, O compounds	4	0.42	0.45	0.43	0.44
H, C, N, O compounds	15	0.85	0.84	0.83	0.83
Fluorinated compounds	6	0.43	0.39	0.34	0.32
Bifunctional fluorinated compounds	4	1.09	1.09	1.10	1.12
Neutral H, C, N, O, F	171	0.55	0.55	0.55	0.55
Phosphine	1	0.13	0.12	0.12	0.11
Chlorinated compounds	27	0.38	0.39	0.40	0.41
Brominated compounds	14	0.17	0.17	0.17	0.17
Multihalogen-containing compounds	12	0.31	0.31	0.30	0.30
Bifunctional halogenated H, C, N, O compounds	5	1.34	1.30	1.26	1.24
Thiols	4	0.17	0.19	0.21	0.23
Sulfides	5	0.49	0.51	0.53	0.55
Disulfides	2	0.29	0.28	0.27	0.27
Sulfur heterocycles	1	0.54	0.56	0.55	0.57
H, C, N, O, S, halogens	2	1.62	1.76	1.84	1.91
H, C, N, O, P	3	0.19	0.13	0.11	0.12
H, C, N, O, P, S	3	1.89	1.87	1.89	1.85
H, C, N, O, P, halogens	2	0.34	0.24	0.22	0.22
H, C, N, O, P, S, halogens	5	0.83	0.88	0.87	0.88
Cl, Br, S, and P neutrals	86	0.49	0.49	0.49	0.50
All neutrals	257	0.53	0.53	0.53	0.53
H, C, N, O ions	30	5.28	5.23	5.20	5.17
F ions	2	1.28	1.12	0.99	0.89
All H, C, N, O, F ions	32	5.03	4.98	4.94	4.90
H, C, S ions	6	6.90	6.71	6.61	6.54
Cl and Br ions	3	2.75	2.55	2.43	2.33
H, C, P ions	3	2.38	2.28	2.19	2.13
Inorganic S and P ions	3	4.02	4.02	4.01	4.01
All Cl, Br, P, S ions	15	4.59	4.45	4.37	4.31
All ions	47	4.89	4.81	4.76	4.71

4.6. Performance of SM5.43R for Solutes in Organic Solvents in Terms of Solvent Functional class

Next we report the MUEs of free energies of solvation for solutes in various solvent classes calculated by SM5.43R. These MUEs are shown in Tables 19, 20, 21, 22, and 23. The SM5.43R model makes the most accurate predictions of free energies of solvation for solutes in hydrocarbon

and aromatic solvents; it is not known whether this occurs because data in these two solvents make up a large fraction, about 40%, of the data in the organic training set or because these solvents are “easier” to treat. The MUEs in other solvents, such as ketones, ethers, amines, pyridines, amides, and halogenated solvents, are between 0.5 and 0.6 kcal/mol, while the MUEs in alcohols, esters, and nitrohydrocarbons are approximately 0.7 kcal/mol.

Table 10. MUEs (kcal/mol) of the aqueous free energies of solvation for various solute classes calculated by SM5.43R/MPW X /MIDI!6D ($X=0, 25, 42.8, \text{ and } 60.6$). Averaged values of the surface tension parameters from MPW X /MIDI!6D are used

Solute class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Inorganic H, C, N	4	0.34	0.29	0.33	0.41
Unbranched alkanes	8	0.62	0.62	0.62	0.62
Branched alkanes	5	0.46	0.47	0.47	0.47
Cycloalkanes	5	0.61	0.60	0.60	0.60
Alkenes	9	0.37	0.37	0.37	0.37
Alkynes	5	0.24	0.24	0.23	0.23
Arenes	8	0.36	0.37	0.37	0.40
Alcohols	12	0.55	0.53	0.52	0.51
Phenols	4	0.28	0.33	0.35	0.39
Ethers	12	0.47	0.47	0.47	0.48
Aldehydes	6	0.41	0.41	0.42	0.38
Ketones	12	0.31	0.31	0.32	0.31
Carboxylic acids	5	0.42	0.33	0.31	0.30
Esters	13	0.29	0.30	0.33	0.35
Bifunctional H, C, O compounds	5	0.40	0.42	0.43	0.44
All H, C, O compounds	109	0.42	0.41	0.42	0.42
Aliphatic amines	15	0.93	0.92	0.91	0.91
Anilines	1	0.24	0.28	0.29	0.26
Ar N heterocycles (1 N in ring)	10	1.07	1.06	1.06	1.05
Ar N heterocycles (2 Ns in ring)	2	0.47	0.48	0.47	0.48
Nitriles	4	0.60	0.63	0.62	0.63
Bifunctional H, C, N compounds	1	0.29	0.39	0.49	0.58
H, C, N compounds	33	0.87	0.87	0.86	0.86
Amides	3	1.26	1.22	1.19	1.19
Ureas	1	3.14	3.17	3.21	3.21
Nitrohydrocarbons	7	0.68	0.65	0.63	0.63
Bifunctional H, C, N, O compounds	4	0.39	0.41	0.43	0.44
H, C, N, O compounds	15	0.89	0.87	0.86	0.86
Fluorinated compounds	6	0.43	0.40	0.36	0.33
Bifunctional fluorinated compounds	4	1.10	1.11	1.12	1.14
Neutral H, C, N, O, F	171	0.56	0.55	0.55	0.56
Phosphine	1	0.14	0.11	0.11	0.11
Chlorinated compounds	27	0.38	0.40	0.41	0.42
Brominated compounds	14	0.18	0.19	0.18	0.18
Multihalogen-containing compounds	12	0.31	0.31	0.31	0.31
Bifunctional halogenated H, C, N, O compounds	5	1.39	1.35	1.29	1.27
Thiols	4	0.17	0.18	0.20	0.22
Sulfides	5	0.47	0.50	0.51	0.53
Disulfides	2	0.30	0.29	0.28	0.27
Sulfur heterocycles	1	0.58	0.59	0.59	0.61
H, C, N, O, S, halogens	2	1.53	1.68	1.77	1.85
H, C, N, O, P	3	0.18	0.11	0.14	0.17
H, C, N, O, P, S	3	1.95	1.92	1.92	1.89
H, C, N, O, P, halogens	2	0.39	0.32	0.22	0.17
H, C, N, O, P, S, halogens	5	0.80	0.81	0.83	0.82
Cl, Br, S, and P neutrals	86	0.49	0.49	0.50	0.50
All neutrals	257	0.54	0.53	0.53	0.54
H, C, N, O ions	30	5.25	5.19	5.17	5.13
F ions	2	1.24	1.06	0.95	0.84
All H, C, N, O, F ions	32	5.00	4.94	4.91	4.87
H, C, S ions	6	6.93	6.75	6.64	6.56
Cl and Br ions	3	2.75	2.53	2.41	2.31
H, C, P ions	3	2.43	2.35	2.28	2.23
Inorganic S and P ions	3	4.08	4.10	4.12	4.12
All Cl, Br, P, S ions	15	4.63	4.50	4.42	4.36
All ions	47	4.88	4.79	4.75	4.70

4.7. Predicted free energies of solvation for selected solutes in *n*-hexadecane, *n*-octanol, and water

Finally we discuss the prediction of free energies of solvation of several solutes in three different solvents, namely *n*-hexadecane, 1-octanol, and water. These solvents have a broad range of characteristics. Of these three solvents, *n*-hexadecane has the smallest dielectric constant and does not have the capability to donate or

accept hydrogen bonds (it has zero values for the α and β solvent descriptors). The solvent 1-octanol has a smaller dielectric constant than water and has nonzero values for α and β . Both *n*-hexadecane and 1-octanol have long aliphatic tails, which should lead to favorable dispersion interactions with aliphatic and aromatic groups in solutes, and both of these solvents are sometimes used [185–195] as surrogates for biological membranes in estimating the bioavailability of drugs.

Table 11. MUEs (in kcal/mol) of the aqueous free energies of solvation for various solute classes calculated by SM5.43R/MPWX/6-31G(d) ($X=0, 25, 42.8, \text{ and } 60.6$). Averaged values of the surface tension parameters from MPWX/6-31G(d) are used

Solute class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Inorganic H, C, N	4	0.29	0.25	0.26	0.30
Unbranched alkanes	8	0.69	0.69	0.69	0.69
Branched alkanes	5	0.48	0.49	0.49	0.49
Cycloalkanes	5	0.59	0.59	0.58	0.57
Alkenes	9	0.36	0.35	0.34	0.33
Alkynes	5	0.21	0.20	0.19	0.18
Arenes	8	0.27	0.28	0.29	0.30
Alcohols	12	0.50	0.48	0.47	0.46
Phenols	4	0.26	0.28	0.31	0.38
Ethers	12	0.46	0.47	0.48	0.50
Aldehydes	6	0.36	0.33	0.27	0.24
Ketones	12	0.30	0.30	0.28	0.27
Carboxylic acids	5	0.47	0.37	0.34	0.32
Esters	13	0.30	0.32	0.33	0.34
Bifunctional H, C, O compounds	5	0.35	0.36	0.37	0.38
All H, C, O compounds	109	0.40	0.40	0.39	0.39
Aliphatic amines	15	0.87	0.85	0.85	0.83
Anilines	1	0.28	0.26	0.29	0.28
Ar N heterocycles (1 N in ring)	10	1.13	1.10	1.09	1.07
Ar N heterocycles (2 Ns in ring)	2	0.58	0.57	0.57	0.58
Nitriles	4	0.46	0.51	0.52	0.52
Bifunctional H, C, N compounds	1	0.36	0.48	0.55	0.65
H, C, N compounds	33	0.85	0.84	0.84	0.83
Amides	3	1.00	0.98	1.02	1.05
Ureas	1	2.90	2.89	2.93	2.96
Nitrohydrocarbons	7	0.42	0.40	0.37	0.36
Bifunctional H, C, N, O compounds	4	0.40	0.43	0.42	0.42
H, C, N, O compounds	15	0.70	0.69	0.69	0.69
Fluorinated compounds	6	0.39	0.37	0.35	0.34
Bifunctional fluorinated compounds	4	1.11	1.13	1.13	1.14
Neutral H, C, N, O, F	171	0.53	0.52	0.52	0.52
Phosphine	1	0.08	0.13	0.17	0.13
Chlorinated compounds	27	0.31	0.31	0.32	0.33
Brominated compounds	14	0.13	0.14	0.15	0.15
Multihalogen-containing compounds	12	0.32	0.32	0.31	0.31
Bifunctional halogenated H, C, N, O compounds	5	1.19	1.15	1.13	1.11
Thiols	4	0.20	0.21	0.22	0.23
Sulfides	5	0.51	0.54	0.56	0.59
Disulfides	2	0.23	0.23	0.22	0.21
Sulfur heterocycles	1	0.31	0.35	0.37	0.38
H, C, N, O, S, halogens	2	1.72	1.95	2.06	2.17
H, C, N, O, P	3	0.27	0.24	0.23	0.16
H, C, N, O, P, S	3	1.70	1.76	1.74	1.75
H, C, N, O, P, halogens	2	0.33	0.27	0.25	0.20
H, C, N, O, P, S, halogens	5	1.14	1.13	1.12	1.13
Cl, Br, S, and P neutrals	86	0.46	0.47	0.47	0.48
All neutrals	257	0.51	0.50	0.50	0.50
H, C, N, O ions	30	5.34	5.30	5.27	5.24
F ions	2	1.19	1.02	0.90	0.81
All H, C, N, O, F ions	32	5.08	5.03	4.99	4.96
H, C, S ions	6	5.62	5.48	5.42	5.38
Cl and Br ions	3	2.77	2.55	2.43	2.33
H, C, P ions	3	1.31	1.07	0.89	0.85
Inorganic S and P ions	3	3.73	3.74	3.74	3.83
All Cl, Br, P, S ions	15	3.81	3.66	3.58	3.55
All ions	47	4.68	4.60	4.54	4.51

Table 24 reports the values of ΔG_{EP} , G_{CDS} , and ΔG_S^0 calculated by SM5.43R using mPW1PW91 and each of the five basis sets employed in this work for a number of solutes in *n*-hexadecane, 1-octanol, and water solvents. Table 24 also gives the corresponding experimental free energies of solvation. The predicted values of ΔG_{EP} in *n*-hexadecane are more positive than the values of ΔG_{EP} in the other two solvents for all solutes.

Further insight can be obtained by analyzing the individual components; however, we should keep in mind that the partition into ΔG_{EP} and G_{CDS} depends on the Coulomb radii more strongly than their sum does [178], and G_{CDS} includes the non-bulk part of the electrostatics. For the larger solutes, such as toluene, phenol, tetrahydrofuran, anisole, benzaldehyde, pyridine, nitrobenzene, thiophene, and chlorobenzene, the

Table 12. MUEs (in kcal/mol) of the aqueous free energies of solvation for various solute classes calculated by SM5.43R/MPWX/6-31+G(d) ($X=0, 25, 42.8, \text{ and } 60.6$). Averaged values of the surface tension parameters from MPWX/6-31+G(d) are used

Solute class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Inorganic H, C, N	4	0.31	0.28	0.27	0.27
Unbranched alkanes	8	0.75	0.73	0.72	0.71
Branched alkanes	5	0.57	0.56	0.55	0.54
Cycloalkanes	5	0.50	0.49	0.49	0.49
Alkenes	9	0.28	0.27	0.26	0.26
Alkynes	5	0.24	0.24	0.23	0.22
Arenes	8	0.29	0.26	0.22	0.21
Alcohols	12	0.44	0.43	0.42	0.41
Phenols	4	0.37	0.35	0.32	0.29
Ethers	12	0.48	0.47	0.46	0.46
Aldehydes	6	0.48	0.39	0.32	0.30
Ketones	12	0.54	0.51	0.49	0.48
Carboxylic acids	5	0.53	0.53	0.54	0.54
Esters	13	0.47	0.48	0.50	0.52
Bifunctional H, C, O compounds	5	0.45	0.38	0.35	0.33
All H, C, O compounds	109	0.46	0.44	0.43	0.42
Aliphatic amines	15	0.78	0.78	0.78	0.78
Anilines	1	0.67	0.82	0.89	0.92
Ar N heterocycles (1 N in ring)	10	1.13	1.14	1.13	1.14
Ar N heterocycles (2 Ns in ring)	2	0.50	0.49	0.48	0.50
Nitriles	4	0.84	0.84	0.82	0.82
Bifunctional H, C, N compounds	1	0.89	0.70	0.62	0.57
H, C, N compounds	33	0.88	0.88	0.87	0.87
Amides	3	1.27	1.29	1.34	1.36
Ureas	1	1.70	1.93	2.08	2.20
Nitrohydrocarbons	7	0.21	0.19	0.18	0.20
Bifunctional H, C, N, O compounds	4	0.27	0.25	0.23	0.23
H, C, N,O compounds	15	0.54	0.54	0.55	0.57
Fluorinated compounds	6	0.29	0.31	0.32	0.33
Bifunctional fluorinated compounds	4	1.15	1.16	1.17	1.18
Neutral H, C, N, O, F	171	0.55	0.54	0.54	0.53
Phosphine	1	0.02	0.08	0.14	0.18
Chlorinated compounds	27	0.31	0.31	0.33	0.34
Brominated compounds	14	0.24	0.24	0.24	0.25
Multihalogen-containing compounds	12	0.37	0.36	0.36	0.36
Bifunctional halogenated H, C, N, O compounds	5	1.39	1.43	1.42	1.41
Thiols	4	0.16	0.18	0.21	0.24
Sulfides	5	0.49	0.48	0.47	0.47
Disulfides	2	0.45	0.46	0.46	0.47
Sulfur heterocycles	1	0.81	0.80	0.81	0.81
H, C, N, O, S, Halogens	2	0.37	0.42	0.46	0.48
H, C, N, O, P	3	0.20	0.26	0.41	0.63
H, C, N, O, P, S	3	1.02	0.87	0.62	0.31
H, C, N, O, P, halogens	2	1.40	1.14	0.95	0.80
H, C, N, O, P, S, halogens	5	2.29	2.23	2.19	2.35
Cl, Br, S, and P neutrals	86	0.54	0.53	0.53	0.54
All neutrals	257	0.55	0.54	0.53	0.54
H, C, N, O ions	30	5.40	5.36	5.33	5.30
F ions	2	1.22	1.05	0.95	0.87
All H, C, N, O, F ions	32	5.14	5.09	5.05	5.02
H, C, S ions	6	6.14	6.10	6.08	6.07
Cl and Br ions	3	1.84	1.67	1.57	1.50
H, C, P ions	3	2.20	1.98	1.82	1.70
Inorganic S and P ions	3	4.58	4.55	4.52	4.50
All Cl, Br, P, S ions	15	4.18	4.08	4.01	3.97
All ions	47	4.83	4.77	4.72	4.69

predicted value of G_{CDS} is large relative to that of the smaller solutes in Table 24, and this may be due in part to favorable dispersion interactions between the aliphatic groups in these solutes and *n*-hexadecane. These solutes may also benefit from favorable dispersion interactions in 1-octanol, but perhaps to a lesser extent than in *n*-hexadecane. For toluene, anisole, benzaldehyde, nitrobenzene, and chlorobenzene, this appears to be the case, as G_{CDS} is predicted to be more positive for

these solutes in 1-octanol than in *n*-hexadecane. The solutes ethanol, phenol, propanoic acid, propylamine, and pyridine are predicted to have more negative values of G_{CDS} , and this derives from favorable hydrogen-bonding interactions.

In water, all solutes are predicted to have more negative values of ΔG_{EP} than in *n*-hexadecane and 1-octanol. Moreover, the contribution to the free energy of solvation from G_{CDS} in water is more positive than in

Table 13. MUEs (in kcal/mol) of the aqueous free energies of solvation for various solute classes calculated by SM5.43R/MPWX/6-31+G(d,p) ($X=0, 25, 42.8,$ and 60.6). Averaged values of the surface tension parameters from MPWX/6-31+G(d,p) are used

Solute class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Inorganic H, C, N	4	0.29	0.26	0.26	0.26
Unbranched alkanes	8	0.73	0.72	0.71	0.71
Branched alkanes	5	0.56	0.55	0.55	0.55
Cycloalkanes	5	0.51	0.50	0.50	0.50
Alkenes	9	0.29	0.28	0.27	0.27
Alkynes	5	0.26	0.26	0.25	0.24
Arenes	8	0.26	0.25	0.24	0.24
Alcohols	12	0.44	0.44	0.44	0.44
Phenols	4	0.38	0.36	0.34	0.32
Ethers	12	0.46	0.46	0.46	0.47
Aldehydes	6	0.56	0.49	0.43	0.33
Ketones	12	0.54	0.53	0.52	0.50
Carboxylic acids	5	0.52	0.51	0.51	0.52
Esters	13	0.47	0.48	0.50	0.52
Bifunctional H, C, O compounds	5	0.38	0.32	0.32	0.32
All H, C, O compounds	109	0.46	0.45	0.44	0.44
Aliphatic amines	15	0.80	0.79	0.79	0.80
Anilines	1	0.66	0.79	0.86	0.93
Ar N heterocycles (1 N in ring)	10	1.14	1.15	1.15	1.15
Ar N heterocycles (2 Ns in ring)	2	0.50	0.54	0.54	0.58
Nitriles	4	0.81	0.82	0.81	0.80
Bifunctional H, C, N compounds	1	1.02	0.79	0.67	0.56
H, C, N compounds	33	0.89	0.89	0.89	0.89
Amides	3	1.22	1.23	1.26	1.34
Ureas	1	2.03	2.19	2.30	2.46
Nitrohydrocarbons	7	0.21	0.19	0.17	0.16
Bifunctional H, C, N, O compounds	4	0.30	0.28	0.27	0.28
H, C, N, O compounds	15	0.56	0.55	0.56	0.58
Fluorinated compounds	6	0.31	0.32	0.33	0.34
Bifunctional fluorinated compounds	4	1.15	1.17	1.17	1.19
Neutral H, C, N, O, F	171	0.56	0.55	0.55	0.55
Phosphine	1	0.12	0.02	0.13	0.18
Chlorinated compounds	27	0.33	0.34	0.34	0.35
Brominated compounds	14	0.23	0.23	0.24	0.23
Multihalogen-containing compounds	12	0.36	0.36	0.35	0.35
Bifunctional halogenated H, C, N, O compounds	5	1.36	1.39	1.37	1.36
Thiols	4	0.15	0.18	0.20	0.22
Sulfides	5	0.48	0.46	0.46	0.46
Disulfides	2	0.42	0.43	0.44	0.44
Sulfur heterocycles	1	0.82	0.81	0.80	0.81
H, C, N, O, S, halogens	2	0.40	0.41	0.44	0.45
H, C, N, O, P	3	0.58	0.44	0.53	0.77
H, C, N, O, P, S	3	0.89	0.94	0.67	0.37
H, C, N, O, P, halogens	2	1.60	0.91	0.73	0.56
H, C, N, O, P, S, halogens	5	1.98	2.09	2.10	2.24
Cl, Br, S, and P neutrals	86	0.54	0.53	0.52	0.53
All neutrals	257	0.55	0.54	0.54	0.54
H, C, N, O ions	30	5.49	5.45	5.43	5.41
F ions	2	1.18	1.00	0.89	0.78
All H, C, N, O, F ions	32	5.22	5.17	5.15	5.12
H, C, S ions	6	5.99	5.95	5.94	5.92
Cl and Br ions	3	1.89	1.71	1.62	1.53
H, C, P ions	3	2.47	2.30	2.16	2.02
Inorganic S and P ions	3	4.76	4.90	4.91	4.90
All Cl, Br, P, S ions	15	4.22	4.16	4.11	4.06
All ions	47	4.90	4.85	4.82	4.78

the other solvents for all solutes except ethanol and propanoic acid.

5. Conclusions

We have optimized parameters for the SM5.43R continuum solvation model with MPWX/MIDI!, MPWX/MIDI!6D, MPWX/6-31G(d), MPWX/6-31+G(d), and

MPWX/6-31+G(d,p), for $X=0, 25, 42.8,$ and 60.6 in this work and in a previous SM5.43R parameterization. For a given basis set, we found that it is reasonable to use surface tension parameters that are averaged over the four values of X in MPWX, i.e., for a given basis set, one set of SM5.43R parameters may be used with MPWX for any value of X between 0 and 60.6. Furthermore the parameters do not depend strongly on basis set. The SM5.43R model should be useful for

Table 14. MUEs (in kcal/mol) of the free energies of solvation for various solute classes of solutes in organic solvents calculated by SM5.43R/MPW X /MIDI! ($X=0, 25, 42.8,$ and 60.6). Averaged values of surface tension parameters from MPW X /MIDI! are used

Solute class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Inorganic HCN	13	0.99	1.04	1.08	1.11
Inorganic compounds	20	0.73	0.70	0.68	0.67
All inorganic compounds	33	0.83	0.84	0.84	0.84
Unbranched alkanes	85	0.41	0.41	0.40	0.40
Branched alkanes	7	0.50	0.50	0.49	0.49
Cycloalkanes	13	0.55	0.55	0.55	0.55
Alkenes	18	0.34	0.34	0.34	0.34
Alkynes	9	0.35	0.34	0.34	0.34
Arenes	134	0.65	0.67	0.67	0.70
Alcohols	273	0.40	0.41	0.41	0.42
Phenols	109	0.51	0.49	0.48	0.46
Ethers	87	0.40	0.39	0.38	0.38
Aldehydes	32	0.52	0.51	0.50	0.50
Ketones	195	0.51	0.51	0.50	0.50
Carboxylic acids	120	0.48	0.46	0.47	0.48
Esters	238	0.41	0.41	0.41	0.41
Nonhalogenated bifunctional compounds	28	1.79	1.79	1.79	1.79
All H, C, O, compounds	1,348	0.49	0.49	0.49	0.49
Aliphatic amines	154	0.46	0.45	0.44	0.43
Anilines	26	0.37	0.37	0.36	0.37
Ar N heterocycles (1 N in ring)	58	0.43	0.42	0.42	0.42
Ar N heterocycles (2 Ns in ring)	8	0.79	0.79	0.82	0.81
Nitriles	20	0.55	0.53	0.52	0.51
Bifunctional HCN and HCNO	4	0.88	0.91	0.93	0.97
H, C, N compounds	271	0.47	0.46	0.45	0.45
Amides	16	0.54	0.52	0.53	0.52
Ureas	7	1.43	1.43	1.45	1.45
Nitrohydrocarbons	88	0.60	0.63	0.62	0.64
Bifunctional HCN and HCNO	7	0.48	0.49	0.50	0.51
Compounds with H, C, N, O	118	0.64	0.65	0.65	0.66
All H, C, N, O absolute free energies	1,733	0.51	0.50	0.50	0.51
All H, C, N, O transfer free energies	37	0.45	0.45	0.44	0.44
All H, C, N, O data	1,770	0.50	0.50	0.50	0.50
Fluorinated hydrocarbons	5	0.85	0.83	0.77	0.75
Fluoroarenes	11	0.59	0.66	0.68	0.72
Chloroalkanes	26	0.55	0.55	0.55	0.55
Chloroalkenes	15	0.84	0.83	0.83	0.82
Chloroarenes	31	0.37	0.38	0.38	0.39
Brominated hydrocarbons	21	0.62	0.56	0.56	0.56
Bromoalkenes	2	0.13	0.11	0.11	0.11
Bromoarenes	16	0.53	0.55	0.56	0.58
Multi halogenated compounds	14	0.39	0.38	0.36	0.35
Log P : amides & ureas, halogen-containing	5	0.55	0.55	0.55	0.56
Log P : amides & ureas, halogen-containing	10	1.24	1.24	1.23	1.23
Log P : Other halogenated compounds	1	0.15	0.15	0.16	0.15
Multifunctional halogenated H, C, O compounds	33	0.94	0.92	0.90	0.87
Multifunctional halogenated H, C, N, O compounds	4	1.63	1.60	1.59	1.56
Thiols	10	0.45	0.44	0.44	0.43
Sulfides	13	0.65	0.65	0.65	0.66
Disulfides	4	0.42	0.43	0.44	0.44
Sulfur heterocycles	21	0.36	0.35	0.35	0.35
F, Cl, Br, S absolute free energies	210	0.62	0.62	0.61	0.61
F, Cl, Br, S transfer free energies	33	0.63	0.62	0.62	0.62
H, C, N, O, F, S, Cl data	243	0.62	0.62	0.61	0.61
H, C, O, P	25	1.16	1.16	1.17	1.18
H, C, N, O, P, S	6	1.59	1.61	1.62	1.64
H, C, N, O, P, halogen	10	0.96	0.96	0.96	0.97
H, C, N, O, P, S, halogen	5	1.68	1.72	1.71	1.72
All P absolute free energies	37	1.39	1.40	1.40	1.41
All P transfer free energies	9	0.57	0.59	0.59	0.61
All P data	46	1.23	1.24	1.24	1.25
All absolute free energies	1,980	0.53	0.53	0.53	0.53
All transfer free energies	79	0.54	0.54	0.53	0.53
All data	2,059	0.53	0.53	0.53	0.53

Table 15. MUEs (in kcal/mol) of the free energies of solvation for various solute classes of solutes in organic solvents calculated by SM5.43R/MPWX/MIDI!6D ($X=0, 25, 42.8, \text{ and } 60.6$). Averaged values of surface tension parameters from MPWX/MIDI!6D are used

Solute class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Inorganic HCN	13	1.02	1.08	1.13	1.17
Inorganic compounds	20	0.74	0.70	0.69	0.67
All inorganic compounds	33	0.85	0.85	0.86	0.87
Unbranched alkanes	85	0.41	0.41	0.41	0.41
Branched alkanes	7	0.50	0.50	0.50	0.50
Cycloalkanes	13	0.56	0.55	0.55	0.55
Alkenes	18	0.34	0.34	0.34	0.34
Alkynes	9	0.35	0.34	0.34	0.34
Arenes	134	0.67	0.69	0.70	0.72
Alcohols	273	0.41	0.41	0.42	0.42
Phenols	109	0.52	0.50	0.48	0.47
Ethers	87	0.41	0.40	0.39	0.39
Aldehydes	32	0.52	0.51	0.51	0.51
Aetones	195	0.51	0.51	0.52	0.51
Carboxylic acids	120	0.48	0.47	0.48	0.49
Esters	238	0.43	0.42	0.42	0.42
Nonhalogenated bifunctional compounds	28	1.82	1.82	1.82	1.81
All H, C, O, compounds	1,348	0.50	0.50	0.50	0.50
Aliphatic amines	154	0.48	0.46	0.45	0.44
Anilines	26	0.37	0.37	0.37	0.38
Ar N heterocycles (1 N in ring)	58	0.43	0.43	0.42	0.42
Ar N heterocycles (2 Ns in ring)	8	0.81	0.81	0.81	0.81
Nitriles	20	0.55	0.52	0.52	0.50
Bifunctional HCN and HCNO	4	0.87	0.91	0.95	0.99
H, C, N compounds	271	0.48	0.47	0.46	0.45
Amides	16	0.52	0.51	0.50	0.50
Ureas	7	1.48	1.48	1.49	1.49
Nitrohydrocarbons	88	0.64	0.66	0.66	0.67
Bifunctional HCN and HCNO	7	0.49	0.50	0.51	0.51
Compounds with H, C, N, O	118	0.66	0.68	0.68	0.69
All H, C, N, O absolute free energies	1,733	0.52	0.51	0.51	0.52
All H, C, N, O transfer free energies	37	0.46	0.45	0.45	0.44
All H, C, N, O data	1,770	0.52	0.51	0.51	0.51
Fluorinated hydrocarbons	5	0.85	0.82	0.79	0.76
Fluoroarenes	11	0.60	0.67	0.70	0.74
Chloroalkanes	26	0.55	0.54	0.53	0.53
Chloroalkenes	15	0.83	0.82	0.82	0.81
Chloroarenes	31	0.38	0.39	0.39	0.40
Brominated hydrocarbons	21	0.62	0.56	0.56	0.56
Bromoalkenes	2	0.12	0.11	0.11	0.11
Bromoarenes	16	0.55	0.56	0.57	0.59
Multihalogenated compounds	14	0.38	0.37	0.35	0.35
Log P : amides & ureas, halogen-containing	5	0.56	0.56	0.57	0.57
Log P : amides & ureas, halogen-containing	10	1.27	1.26	1.26	1.26
Log P : Other halogenated compounds	1	0.17	0.17	0.17	0.17
Multifunctional halogenated H, C, O compounds	33	0.97	0.95	0.93	0.90
Multifunctional halogenated H, C, N, O compounds	4	1.65	1.63	1.60	1.59
Thiols	10	0.45	0.44	0.44	0.43
Sulfides	13	0.65	0.65	0.66	0.66
Disulfides	4	0.43	0.44	0.44	0.45
Sulfur heterocycles	21	0.35	0.35	0.35	0.34
F, Cl, Br, S absolute free energies	210	0.63	0.63	0.62	0.62
F, Cl, Br, S transfer free energies	33	0.64	0.63	0.63	0.63
H, C, N, O, F, S, Cl data	243	0.63	0.63	0.62	0.62
H, C, O, P	25	1.14	1.15	1.16	1.17
H, C, N, O, P, S	6	1.60	1.62	1.63	1.65
H, C, N, O, P, halogen	10	0.99	0.99	0.99	1.00
H, C, N, O, P, S, halogen	5	1.64	1.66	1.68	1.67
All P absolute free energies	37	1.38	1.39	1.40	1.40
All P transfer free energies	9	0.58	0.59	0.62	0.63
All P data	46	1.22	1.23	1.24	1.25
All absolute free energies	1,980	0.55	0.54	0.54	0.54
All transfer free energies	79	0.55	0.54	0.54	0.54
All data	2,059	0.55	0.54	0.54	0.54

Table 16. MUEs (in kcal/mol) of the free energies of solvation for various solute classes of solutes in organic solvents calculated by SM5.43R/MPWX/6-31G(d) ($X=0, 25, 42.8, \text{ and } 60.6$). Averaged values of surface tension parameters from MPWX/6-31G(d) are used

Solute class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Inorganic HCN	13	0.97	1.00	1.00	1.01
Inorganic compounds	20	0.76	0.74	0.73	0.72
All inorganic compounds	33	0.85	0.84	0.84	0.83
Unbranched alkanes	85	0.39	0.39	0.38	0.38
Branched alkanes	7	0.49	0.48	0.48	0.48
Cycloalkanes	13	0.54	0.54	0.54	0.53
Alkenes	18	0.32	0.32	0.32	0.32
Alkynes	9	0.31	0.31	0.31	0.30
Arenes	134	0.57	0.60	0.62	0.63
Alcohols	273	0.40	0.40	0.40	0.41
Phenols	109	0.54	0.50	0.48	0.46
Ethers	87	0.38	0.38	0.38	0.38
Aldehydes	32	0.50	0.50	0.50	0.51
Acetones	195	0.50	0.50	0.49	0.48
Carboxylic acids	120	0.48	0.46	0.46	0.47
Esters	238	0.39	0.39	0.38	0.38
Nonhalogenated bifunctional compounds	28	1.68	1.68	1.69	1.67
All H, C, O, compounds	1,348	0.48	0.47	0.47	0.47
Aliphatic amines	154	0.43	0.42	0.42	0.41
Anilines	26	0.37	0.37	0.37	0.37
Ar N heterocycles (1 N in ring)	58	0.44	0.44	0.44	0.43
Ar N heterocycles (2 Ns in ring)	8	0.85	0.85	0.85	0.85
Nitriles	20	0.39	0.37	0.36	0.36
Bifunctional HCN and HCNO	4	0.83	0.87	0.89	0.93
H, C, N compounds	271	0.44	0.43	0.43	0.43
Amides	16	0.51	0.52	0.53	0.53
Ureas	7	1.28	1.27	1.27	1.27
Nitrohydrocarbons	88	0.51	0.50	0.51	0.53
Bifunctional HCN and HCNO	7	0.46	0.48	0.49	0.49
Compounds with H, C, N, O	118	0.55	0.55	0.56	0.57
All H, C, N, O absolute free energies	1,733	0.48	0.48	0.48	0.48
All H, C, N, O transfer free energies	37	0.45	0.43	0.43	0.43
All H, C, N, O data	1,770	0.48	0.48	0.48	0.48
Fluorinated hydrocarbons	5	0.87	0.82	0.80	0.78
Fluoroarenes	11	0.71	0.77	0.81	0.84
Chloroalkanes	26	0.59	0.58	0.58	0.57
Chloroalkenes	15	0.86	0.85	0.84	0.84
Chloroarenes	31	0.36	0.36	0.37	0.38
Brominated hydrocarbons	21	0.55	0.55	0.55	0.55
Bromoalkenes	2	0.13	0.15	0.15	0.15
Bromoarenes	16	0.52	0.56	0.57	0.59
Multihalogenated compounds	14	0.35	0.34	0.34	0.33
Log P : amides & ureas, halogen-containing	5	0.53	0.54	0.53	0.53
Log P : amides & ureas, halogen-containing	10	1.13	1.13	1.13	1.13
Log P : Other halogenated compounds	1	0.10	0.10	0.10	0.10
Multifunctional halogenated H, C, O compounds	33	0.94	0.90	0.87	0.84
Multifunctional halogenated H, C, N, O compounds	4	1.49	1.46	1.44	1.40
Thiols	10	0.47	0.45	0.44	0.44
Sulfides	13	0.63	0.64	0.65	0.65
Disulfides	4	0.36	0.37	0.37	0.38
Sulfur heterocycles	21	0.37	0.36	0.36	0.35
F, Cl, Br, S absolute free energies	210	0.63	0.63	0.62	0.62
F, Cl, Br, S transfer free energies	33	0.59	0.59	0.58	0.58
H, C, N, O, F, S, Cl data	243	0.62	0.62	0.62	0.62
H, C, O, P	25	1.19	1.19	1.20	1.21
H, C, N, O, P, S	6	1.46	1.50	1.51	1.49
H, C, N, O, P, halogen	10	0.95	0.95	0.96	0.96
H, C, N, O, P, S, halogen	5	1.95	1.94	1.93	1.95
All P absolute free energies	37	1.42	1.42	1.42	1.43
All P transfer free energies	9	0.59	0.61	0.62	0.65
All P data	46	1.26	1.26	1.27	1.27
All absolute free energies	1,980	0.52	0.51	0.51	0.51
All transfer free energies	79	0.52	0.52	0.52	0.52
All data	2,059	0.52	0.51	0.51	0.51

Table 17. MUEs (in kcal/mol) of the free energies of solvation for various solute classes of solutes in organic solvents calculated by SM5.43R/MPWX/6-31+G(d) ($X=0, 25, 42.8, \text{ and } 60.6$). Averaged values of surface tension parameters from MPWX/6-31+G(d) are used

Solute class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Inorganic HCN	13	1.11	1.09	1.06	1.05
Inorganic compounds	20	0.79	0.78	0.78	0.77
All inorganic compounds	33	0.92	0.90	0.89	0.88
Unbranched alkanes	85	0.36	0.36	0.36	0.37
Branched alkanes	7	0.50	0.50	0.50	0.50
Cycloalkanes	13	0.48	0.47	0.47	0.47
Alkenes	18	0.33	0.34	0.35	0.35
Alkynes	9	0.32	0.30	0.30	0.29
Arenes	134	0.58	0.57	0.58	0.59
Alcohols	273	0.41	0.41	0.41	0.41
Phenols	109	0.55	0.54	0.52	0.50
Ethers	87	0.37	0.37	0.38	0.38
Aldehydes	32	0.52	0.50	0.51	0.52
Acetones	195	0.51	0.49	0.48	0.48
Carboxylic acids	120	0.45	0.45	0.45	0.45
Esters	238	0.48	0.47	0.47	0.47
Nonhalogenated bifunctional compounds	28	1.54	1.59	1.61	1.62
All H, C, O, compounds	1,348	0.49	0.48	0.48	0.48
Aliphatic amines	154	0.38	0.38	0.39	0.39
Anilines	26	0.36	0.38	0.39	0.39
Ar N heterocycles (1 N in ring)	58	0.46	0.47	0.47	0.47
Ar N heterocycles (2 Ns in ring)	8	0.86	0.86	0.85	0.86
Nitriles	20	0.54	0.55	0.55	0.55
Bifunctional HCN and HCNO	4	0.96	0.87	0.83	0.81
H, C, N compounds	271	0.43	0.43	0.44	0.44
Amides	16	0.43	0.43	0.45	0.47
Ureas	7	0.90	0.93	0.97	1.00
Nitrohydrocarbons	88	0.38	0.39	0.40	0.41
Bifunctional HCN and HCNO	7	0.57	0.57	0.57	0.58
Compounds with H, C, N, O	118	0.43	0.44	0.45	0.46
All H, C, N, O absolute free energies	1,733	0.48	0.48	0.48	0.48
All H, C, N, O transfer free energies	37	0.46	0.45	0.44	0.44
All H, C, N, O data	1,770	0.48	0.48	0.48	0.48
Fluorinated hydrocarbons	5	0.62	0.62	0.62	0.62
Fluoroarenes	11	0.91	0.96	1.02	1.07
Chloroalkanes	26	0.59	0.58	0.56	0.55
Chloroalkenes	15	0.85	0.85	0.85	0.86
Chloroarenes	31	0.37	0.35	0.36	0.36
Brominated hydrocarbons	21	0.69	0.68	0.69	0.69
Bromoalkenes	2	0.16	0.15	0.16	0.16
Bromoarenes	16	0.50	0.51	0.53	0.54
Multihalogenated compounds	14	0.44	0.44	0.44	0.44
Log P : amides & ureas, halogen-containing	5	0.65	0.64	0.63	0.61
Log P : amides & ureas, halogen-containing	10	1.08	1.06	1.05	1.04
Log P : Other halogenated compounds	1	0.12	0.12	0.12	0.11
Multifunctional halogenated H, C, O compounds	33	1.08	1.08	1.06	1.04
Multifunctional halogenated H, C, N, O compounds	4	1.59	1.60	1.58	1.57
Thiols	10	0.44	0.43	0.41	0.39
Sulfides	13	0.65	0.65	0.66	0.66
Disulfides	4	0.50	0.51	0.51	0.51
Sulfur heterocycles	21	0.38	0.37	0.36	0.36
F, Cl, Br, S absolute free energies	210	0.67	0.67	0.67	0.67
F, Cl, Br, S transfer free energies	33	0.61	0.60	0.59	0.58
H, C, N, O, F, S, Cl data	243	0.67	0.66	0.66	0.66
H, C, O, P	25	1.25	1.25	1.30	1.36
H, C, N, O, P, S	6	1.46	1.42	1.42	1.40
H, C, N, O, P, halogen	10	1.13	1.08	1.06	1.03
H, C, N, O, P, S, halogen	5	3.01	2.95	2.90	2.84
All P absolute free energies	37	1.63	1.60	1.62	1.65
All P transfer free energies	9	0.68	0.67	0.67	0.68
All P data	46	1.44	1.42	1.44	1.46
All absolute free energies	1,980	0.53	0.52	0.52	0.53
All transfer free energies	79	0.55	0.54	0.53	0.52
All data	2,059	0.53	0.52	0.52	0.53

Table 18. MUEs (in kcal/mol) of the free energies of solvation for various solute classes of solutes in organic solvents calculated by SM5.43R/MPWX/6-31+G(d,p) ($X=0$, 25, 42.8, and 60.6). Averaged values of surface tension parameters from MPWX/6-31+G(d,p) are used

Solute class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Inorganic HCN	13	1.08	1.07	1.06	1.03
Inorganic compounds	20	0.80	0.79	0.78	0.77
All inorganic compounds	33	0.91	0.90	0.89	0.88
Unbranched alkanes	85	0.37	0.37	0.37	0.37
Branched alkanes	7	0.50	0.50	0.50	0.50
Cycloalkanes	13	0.49	0.49	0.49	0.49
Alkenes	18	0.31	0.32	0.32	0.32
Alkynes	9	0.32	0.31	0.31	0.31
Arenes	134	0.61	0.60	0.61	0.61
Alcohols	273	0.41	0.42	0.43	0.43
Phenols	109	0.58	0.56	0.55	0.53
Ethers	87	0.37	0.38	0.38	0.38
Aldehydes	32	0.52	0.51	0.51	0.52
Acetones	195	0.54	0.53	0.52	0.50
Carboxylic acids	120	0.45	0.45	0.45	0.46
Esters	238	0.47	0.47	0.47	0.48
Nonhalogenated bifunctional compounds	28	1.59	1.65	1.67	1.69
All H, C, O, compounds	1,348	0.50	0.50	0.50	0.50
Aliphatic amines	154	0.38	0.38	0.38	0.39
Anilines	26	0.36	0.37	0.38	0.39
Ar N heterocycles (1 N in ring)	58	0.47	0.47	0.47	0.47
Ar N heterocycles (2 Ns in ring)	8	0.85	0.87	0.87	0.90
Nitriles	20	0.53	0.52	0.52	0.50
Bifunctional HCN and HCNO	4	1.05	0.95	0.90	0.85
H, C, N compounds	271	0.43	0.43	0.43	0.44
Amides	16	0.35	0.38	0.40	0.43
Ureas	7	1.08	1.10	1.12	1.17
Nitrohydrocarbons	88	0.37	0.38	0.38	0.39
Bifunctional HCN and HCNO	7	0.59	0.60	0.61	0.63
Compounds with H, C, N, O	118	0.42	0.43	0.44	0.46
All H, C, N, O absolute free energies	1,733	0.49	0.49	0.49	0.49
All H, C, N, O transfer free energies	37	0.46	0.45	0.45	0.46
All H, C, N, O data	1,770	0.49	0.49	0.49	0.49
Fluorinated hydrocarbons	5	0.64	0.64	0.64	0.64
Fluoroarenes	11	0.95	1.02	1.07	1.12
Chloroalkanes	26	0.61	0.59	0.58	0.57
Chloroalkenes	15	0.87	0.88	0.88	0.88
Chloroarenes	31	0.36	0.36	0.36	0.37
Brominated hydrocarbons	21	0.70	0.70	0.70	0.70
Bromoalkenes	2	0.19	0.20	0.20	0.20
Bromoarenes	16	0.54	0.55	0.56	0.57
Multihalogenated compounds	14	0.43	0.43	0.43	0.43
Log P : amides & ureas, halogen-containing	5	0.62	0.62	0.61	0.60
Log P : amides & ureas, halogen-containing	10	1.17	1.16	1.14	1.13
Log P : Other halogenated compounds	1	0.13	0.13	0.13	0.13
Multifunctional halogenated H, C, O compounds	33	1.10	1.09	1.08	1.06
Multifunctional halogenated H, C, N, O compounds	4	1.54	1.55	1.52	1.50
Thiols	10	0.44	0.43	0.42	0.41
Sulfides	13	0.68	0.68	0.68	0.69
Disulfides	4	0.48	0.48	0.49	0.49
Sulfur heterocycles	21	0.39	0.38	0.38	0.37
F, Cl, Br, S absolute free energies	210	0.69	0.69	0.69	0.69
F, Cl, Br, S transfer free energies	33	0.64	0.63	0.62	0.61
H, C, N, O, F, S, Cl data	243	0.68	0.68	0.68	0.68
H, C, O, P	25	1.31	1.34	1.41	1.49
H, C, N, O, P, S	6	1.40	1.41	1.38	1.36
H, C, N, O, P, halogen	10	1.12	1.07	1.04	1.01
H, C, N, O, P, S, halogen	5	2.74	2.83	2.76	2.76
All P absolute free energies	37	1.64	1.64	1.66	1.71
All P transfer free energies	9	0.62	0.69	0.69	0.68
All P data	46	1.44	1.45	1.47	1.51
All absolute free energies	1,980	0.53	0.53	0.54	0.54
All transfer free energies	79	0.55	0.55	0.55	0.55
All data	2,059	0.53	0.53	0.54	0.54

Table 19. MUEs (in kcal/mol) of the free energies of solvation of solutes in each organic solvent calculated by MPWX/MIDI! ($X=0, 25, 42.8,$ and 60.6). Averaged values of the surface tension parameters from MPWX/MIDI! are used

Solvent class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Alkanes	597	0.42	0.42	0.42	0.43
Arenes	274	0.48	0.47	0.46	0.46
Alcohols	338	0.70	0.69	0.68	0.68
Ketones	43	0.61	0.62	0.61	0.64
Esters	39	0.67	0.64	0.62	0.60
Ethers	125	0.64	0.63	0.62	0.62
Amines	16	0.52	0.49	0.49	0.49
Pyridines	19	0.52	0.52	0.50	0.50
Nitriles	14	0.56	0.59	0.58	0.59
Nitrohydrocarbons	34	0.66	0.72	0.73	0.76
Amides	21	0.52	0.48	0.47	0.45
Haloaliphatics	287	0.53	0.54	0.55	0.56
Haloaromatics	116	0.50	0.53	0.54	0.56
Other C, H, O solvents	13	0.60	0.59	0.59	0.59
Other solvents	44	0.57	0.57	0.58	0.58

Table 20. MUEs (in kcal/mol) of the free energies of solvation of solutes in each organic solvent calculated by MPWX/MIDI!6D ($X=0, 25, 42.8,$ and 60.6). Averaged values of the surface tension parameters from MPWX/MIDI!6D are used

Solvent class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Alkanes	597	0.43	0.43	0.43	0.43
Arenes	274	0.49	0.48	0.47	0.46
Alcohols	338	0.72	0.70	0.69	0.69
Ketones	43	0.64	0.64	0.65	0.67
Esters	39	0.69	0.66	0.63	0.62
Ethers	125	0.66	0.65	0.64	0.63
Amines	16	0.54	0.51	0.51	0.51
Pyridines	19	0.53	0.53	0.52	0.52
Nitriles	14	0.59	0.61	0.61	0.63
Nitrohydrocarbons	34	0.70	0.75	0.77	0.79
Amides	21	0.53	0.48	0.47	0.45
Haloaliphatics	287	0.54	0.55	0.56	0.57
Haloaromatics	116	0.51	0.54	0.56	0.58
Other C, H, O solvents	13	0.61	0.60	0.61	0.60
Other solvents	44	0.59	0.59	0.59	0.59

Table 21. MUEs (in kcal/mol) of the free energies of solvation of solutes in each organic solvent calculated by MPWX/6-31G(d) ($X=0, 25, 42.8,$ and 60.6). Averaged values of the surface tension parameters from MPWX/6-31G(d) are used

Solvent class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Alkanes	597	0.41	0.41	0.41	0.41
Arenes	274	0.47	0.45	0.45	0.44
Alcohols	338	0.68	0.67	0.66	0.66
Ketones	43	0.59	0.59	0.58	0.59
Esters	39	0.65	0.62	0.60	0.58
Ethers	125	0.61	0.60	0.59	0.59
Amines	16	0.50	0.49	0.48	0.47
Pyridines	19	0.46	0.46	0.45	0.44
Nitriles	14	0.48	0.48	0.49	0.51
Nitrohydrocarbons	34	0.62	0.66	0.67	0.69
Amides	21	0.45	0.44	0.42	0.41
Haloaliphatics	287	0.51	0.52	0.52	0.52
Haloaromatics	116	0.50	0.53	0.54	0.55
Other C, H, O solvents	13	0.59	0.60	0.58	0.58
Other solvents	44	0.56	0.58	0.58	0.59

Table 22. MUEs (in kcal/mol) of the free energies of solvation of solutes in each organic solvent calculated by MPWX/6-31+G(d) ($X=0, 25, 42.8,$ and 60.6). Averaged values of the surface tension parameters from MPWX/6-31+G(d) are used

Solvent class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Alkanes	597	0.42	0.42	0.42	0.42
Arenes	274	0.48	0.48	0.48	0.48
Alcohols	338	0.69	0.69	0.69	0.69
Ketones	43	0.59	0.58	0.58	0.58
Esters	39	0.68	0.67	0.66	0.65
Ethers	125	0.59	0.58	0.58	0.58
Amines	16	0.51	0.51	0.50	0.50
Pyridines	19	0.37	0.36	0.37	0.37
Nitriles	14	0.50	0.49	0.49	0.49
Nitrohydrocarbons	34	0.62	0.61	0.60	0.60
Amides	21	0.42	0.43	0.44	0.45
Haloaliphatics	287	0.52	0.52	0.52	0.53
Haloaromatics	116	0.54	0.54	0.54	0.54
Other C, H, O solvents	13	0.59	0.59	0.59	0.59
Other solvents	44	0.56	0.58	0.59	0.60

Table 23. MUEs (in kcal/mol) of the free energies of solvation of solutes in each organic solvent calculated by MPWX/6-31+G(d,p) ($X=0, 25, 42.8,$ and 60.6). Averaged values of the surface tension parameters from MPWX/6-31+G(d,p) are used

Solvent class	No. of data	$X=0$	$X=25$	$X=42.8$	$X=60.6$
Alkanes	597	0.43	0.43	0.43	0.43
Arenes	274	0.48	0.48	0.48	0.49
Alcohols	338	0.71	0.71	0.71	0.71
Ketones	43	0.60	0.59	0.59	0.59
Esters	39	0.69	0.68	0.67	0.66
Ethers	125	0.61	0.60	0.60	0.59
Amines	16	0.50	0.50	0.50	0.50
Pyridines	19	0.40	0.38	0.38	0.37
Nitriles	14	0.51	0.51	0.51	0.50
Nitrohydrocarbons	34	0.64	0.64	0.64	0.63
Amides	21	0.42	0.42	0.43	0.44
Haloaliphatics	287	0.53	0.54	0.54	0.55
Haloaromatics	116	0.54	0.54	0.55	0.55
Other C, H, O solvents	13	0.61	0.62	0.61	0.61
Other solvents	44	0.56	0.58	0.59	0.60

predicting solvation free energies, partition coefficients, solubilities, vapor pressures, liquid-phase geometries, and potentials of mean force for a broad range of chemical compounds.

6. Availability of SM5.43R

The SM5.43R methods presented here are available in HONDOPUS [196], which incorporates our solvation model into the HONDO program [197]. For more information on HONDOPUS, see <http://comp.chem.umn.edu/hondoplus>. The new methods are also available for *Gaussian* users in a program called SMXGAUSS [198]. This program can read a *Gaussian* output file corresponding to a gas-phase calculation of a given solute and carry out a single-point calculation with SM5.43R or a geometry optimization or a Hessian calculation with SM5.43. In addition, for users who

Table 24. Solvation free energies (kcal/mol) of selected neutral solutes in three solvents calculated by SM5.43R using mPW1PW91 and the indicated basis set. Averaged values of the surface tension parameters are used for all SM5.43R methods

Basis set	<i>n</i> -Hexadecane			1-Octanol			Water		
	G_{EP}	G_{CDS}	ΔG_S^0	G_{EP}	G_{CDS}	ΔG_S^0	G_{EP}	G_{CDS}	ΔG_S^0
Toluene									
MIDI!	-1.22	-3.92	-5.14	-2.24	-2.82	-5.06	-2.48	1.17	-1.31
MIDI!6D	-1.22	-3.95	-5.17	-2.24	-2.85	-5.09	-2.48	1.15	-1.34
6-31G(d)	-1.14	-3.94	-5.08	-2.09	-2.88	-4.98	-2.32	1.14	-1.18
6-31 + G(d)	-1.12	-3.92	-5.05	-2.06	-2.89	-4.96	-2.29	1.25	-1.04
6-31 + G(d,p)	-1.10	-3.97	-5.07	-2.00	-3.00	-5.00	-2.22	1.13	-1.09
Experiment			-4.54			-4.55			-0.89
Phenol									
MIDI!	-2.25	-2.93	-5.18	-4.13	-3.68	-7.82	-4.59	-1.95	-6.53
MIDI!6D	-2.19	-2.99	-5.18	-4.03	-3.75	-7.78	-4.48	-2.02	-6.50
6-31G(d)	-2.37	-2.82	-5.19	-4.36	-3.49	-7.85	-4.85	-1.70	-6.55
6-31 + G(d)	-2.19	-2.97	-5.17	-4.01	-3.71	-7.71	-4.44	-1.77	-6.21
6-31 + G(d,p)	-2.16	-3.01	-5.17	-3.93	-3.80	-7.72	-4.34	-1.89	-6.24
Experiment				-5.14		-8.69			-6.62
Ethanol									
MIDI!	-1.62	-0.89	-2.50	-2.93	-2.58	-5.51	-3.24	-2.28	-5.53
MIDI!6D	-1.61	-0.91	-2.52	-2.93	-2.62	-5.55	-3.24	-2.32	-5.56
6-31G(d)	-1.77	-0.75	-2.52	-3.23	-2.32	-5.55	-3.57	-2.03	-5.60
6-31 + G(d)	-1.71	-0.83	-2.54	-3.09	-2.44	-5.53	-3.42	-2.15	-5.57
6-31 + G(d,p)	-1.73	-0.85	-2.57	-3.11	-2.48	-5.58	-3.43	-2.20	-5.63
Experiment			-2.03			-4.36			-5.01
Tetrahydrofuran									
MIDI!	-1.34	-2.39	-3.73	-2.46	-2.17	-4.63	-2.73	-0.74	-3.46
MIDI!6D	-1.39	-2.33	-3.71	-2.55	-2.09	-4.64	-2.83	-0.65	-3.48
6-31G(d)	-1.29	-2.43	-3.72	-2.37	-2.23	-4.60	-2.62	-0.82	-3.44
6-31 + G(d)	-1.14	-2.55	-3.69	-2.11	-2.38	-4.48	-2.34	-0.92	-3.26
6-31 + G(d,p)	-1.11	-2.61	-3.72	-2.05	-2.48	-4.53	-2.28	-1.03	-3.31
Experiment			-3.60			-3.93			-3.47
Anisole									
MIDI!	-1.82	-2.83	-4.65	-3.35	-1.88	-5.23	-3.72	0.57	-3.15
MIDI!6D	-1.84	-2.81	-4.65	-3.39	-1.85	-5.24	-3.77	0.62	-3.16
6-31G(d)	-1.78	-2.90	-4.69	-3.29	-1.99	-5.28	-3.66	0.49	-3.17
6-31 + G(d)	-1.66	-2.99	-4.65	-3.07	-2.14	-5.20	-3.41	0.49	-2.92
6-31 + G(d,p)	-1.58	-3.07	-4.66	-2.91	-2.28	-5.19	-3.23	0.31	-2.93
Experiment			-5.35			-5.47			-2.45
Benzaldehyde									
MIDI!	-2.40	-2.63	-5.03	-4.51	-1.71	-6.21	-5.02	0.47	-4.55
MIDI!6D	-2.41	-2.61	-5.02	-4.53	-1.65	-6.19	-5.06	0.54	-4.52
6-31G(d)	-2.42	-2.69	-5.12	-4.63	-1.79	-6.42	-5.20	0.47	-4.73
6-31 + G(d)	-2.52	-2.84	-5.37	-4.83	-2.07	-6.90	-5.42	0.38	-5.04
6-31 + G(d,p)	-2.50	-2.87	-5.37	-4.76	-2.13	-6.89	-5.33	0.26	-5.07
Experiment			-5.44			-6.13			-4.02
Butanone									
MIDI!	-2.31	-1.40	-3.71	-4.39	-1.09	-5.49	-4.92	0.76	-4.16
MIDI!6D	-2.34	-1.36	-3.69	-4.48	-1.01	-5.49	-5.02	0.86	-4.16
6-31G(d)	-2.22	-1.45	-3.67	-4.29	-1.14	-5.43	-4.82	0.74	-4.08
6-31 + G(d)	-2.05	-1.50	-3.54	-3.94	-1.29	-5.23	-4.43	0.61	-3.82
6-31 + G(d,p)	-2.12	-1.48	-3.60	-4.07	-1.28	-5.35	-4.57	0.60	-3.97
Experiment			-3.12			-3.78			-3.64
Cyclopentanone									
MIDI!	-2.03	-2.58	-4.60	-3.80	-2.33	-6.13	-4.24	-0.10	-4.34
MIDI!6D	-2.07	-2.54	-4.60	-3.89	-2.24	-6.13	-4.35	0.00	-4.35
6-31G(d)	-1.99	-2.60	-4.60	-3.80	-2.35	-6.15	-4.26	-0.08	-4.34
6-31 + G(d)	-1.91	-2.72	-4.63	-3.68	-2.58	-6.25	-4.12	-0.26	-4.38
6-31 + G(d,p)	-1.92	-2.69	-4.60	-3.68	-2.54	-6.22	-4.13	-0.24	-4.36
Experiment			-4.39			-5.01			-4.68

Table 24. (Contd.)

Basis set	<i>n</i> -Hexadecane			1-Octanol			Water		
	G_{EP}	G_{CDS}	ΔG_S^0	G_{EP}	G_{CDS}	ΔG_S^0	G_{EP}	G_{CDS}	ΔG_S^0
Propanoic acid									
MIDI!	-2.54	-1.32	-3.86	-4.66	-2.02	-6.68	-5.17	-1.35	-6.52
MIDI!6D	-2.51	-1.35	-3.86	-4.63	-2.01	-6.64	-5.15	-1.35	-6.49
6-31G(d)	-2.63	-1.14	-3.76	-4.87	-1.81	-6.68	-5.41	-1.03	-6.44
6-31 + G(d)	-2.32	-1.25	-3.57	-4.26	-2.21	-6.48	-4.73	-1.34	-6.07
6-31 + G(d,p)	-2.34	-1.27	-3.61	-4.30	-2.22	-6.52	-4.77	-1.38	-6.16
Experiment			-3.12			-6.86			-6.47
Ethyl acetate									
MIDI!	-2.08	-1.55	-3.63	-3.85	-0.56	-4.41	-4.28	1.01	-3.26
MIDI!6D	-2.13	-1.49	-3.62	-3.95	-0.45	-4.40	-4.39	1.13	-3.26
6-31G(d)	-2.06	-1.56	-3.62	-3.85	-0.66	-4.51	-4.28	0.98	-3.30
6-31 + G(d)	-2.14	-1.63	-3.77	-4.03	-0.98	-5.01	-4.50	0.76	-3.75
6-31 + G(d,p)	-2.07	-1.69	-3.76	-3.90	-1.03	-4.94	-4.36	0.66	-3.70
Experiment			-3.25			-4.06			-3.10
Propylamine									
MIDI!	-1.29	-1.41	-2.70	-2.31	-2.09	-4.40	-2.55	-0.93	-3.48
MIDI!6D	-1.35	-1.33	-2.67	-2.42	-1.96	-4.38	-2.66	-0.79	-3.45
6-31G(d)	-1.33	-1.40	-2.73	-2.39	-2.06	-4.45	-2.64	-0.92	-3.56
6-31 + G(d)	-1.36	-1.37	-2.73	-2.46	-2.05	-4.51	-2.72	-0.93	-3.65
6-31 + G(d,p)	-1.37	-1.37	-2.73	-2.46	-2.06	-4.52	-2.72	-0.91	-3.62
Experiment			-2.92			-4.77			-4.39
Pyridine									
MIDI!	-2.24	-2.12	-4.35	-4.20	-1.61	-5.81	-4.69	0.36	-4.33
MIDI!6D	-2.31	-2.03	-4.33	-4.34	-1.49	-5.83	-4.84	0.50	-4.34
6-31G(d)	-2.17	-2.25	-4.43	-4.12	-1.76	-5.88	-4.61	0.16	-4.45
6-31 + G(d)	-1.97	-2.48	-4.45	-3.73	-2.07	-5.80	-4.17	-0.20	-4.37
6-31 + G(d,p)	-1.91	-2.59	-4.50	-3.60	-2.25	-5.85	-4.02	-0.41	-4.43
Experiment			-4.10			-5.34			-4.70
Nitrobenzene									
MIDI!	-2.17	-3.13	-5.30	-4.09	-1.12	-5.20	-4.56	1.54	-3.02
MIDI!6D	-2.17	-3.07	-5.25	-4.10	-0.97	-5.07	-4.58	1.68	-2.90
6-31G(d)	-2.31	-3.15	-5.46	-4.42	-1.25	-5.67	-4.97	1.56	-3.41
6-31 + G(d)	-2.25	-3.53	-5.77	-4.39	-1.96	-6.35	-4.96	1.05	-3.90
6-31 + G(d,p)	-2.23	-3.63	-5.85	-4.34	-2.15	-6.48	-4.89	0.84	-4.06
Experiment			-6.22			-6.63			-4.12
Thiophene									
MIDI!	-0.84	-2.87	-3.71	-1.54	-2.08	-3.62	-1.71	-0.27	-1.98
MIDI!6D	-0.86	-2.86	-3.72	-1.57	-2.08	-3.65	-1.74	-0.27	-2.01
6-31G(d)	-0.80	-2.87	-3.67	-1.47	-2.11	-3.58	-1.63	-0.15	-1.77
6-31 + G(d)	-0.89	-2.85	-3.74	-1.63	-2.11	-3.74	-1.81	-0.41	-2.22
6-31 + G(d,p)	-0.85	-2.84	-3.70	-1.55	-2.14	-3.70	-1.72	-0.51	-2.23
Experiment			-4.01			-3.89			-1.42
Chlorobenzene									
MIDI!	-1.04	-4.22	-5.27	-1.93	-3.25	-5.18	-2.14	0.65	-1.50
MIDI!6D	-1.04	-4.25	-5.29	-1.92	-3.28	-5.20	-2.14	0.62	-1.52
6-31G(d)	-1.01	-4.23	-5.24	-1.87	-3.29	-5.16	-2.08	0.66	-1.43
6-31 + G(d)	-0.97	-4.21	-5.18	-1.78	-3.30	-5.08	-1.98	0.78	-1.20
6-31 + G(d,p)	-0.96	-4.25	-5.21	-1.76	-3.38	-5.14	-1.95	0.67	-1.28
Experiment			-4.99			-5.00			-1.12

have a *Gaussian03* executable, SMXGAUSS can be used in conjunction with the “External” option introduced in *Gaussian03*. This allows for liquid-phase geometry optimizations with the powerful optimizers available in *Gaussian*. For more information on SMXGAUSS, see <http://comp.chem.umn.edu/smxgauss>.

Acknowledgements. This work was supported by a Department of Defense Multidisciplinary University Research Initiative (MURI)

grant managed by the Army Research Office and by the National Science Foundation.

References

1. Kohn W, Sham LJ (1965) Phys Rev 140: A1133
2. Vosko SJ, Wilk L, Nusair M (1980) Can J Phys 1980: 1200
3. Levine ZH, Louie SG (1982) Phys Rev B 25: 6310
4. Perdew JP, Wang Y (1986) Phys Rev B 33: 8800
5. Perdew JP (1986) Phys Rev B 33: 8822

6. Becke AD (1986) *J Chem Phys* 84: 4524
7. Becke AD (1988) *Phys Rev A* 38: 3098
8. Lee C, Yang W, Parr RG (1988) *Phys Rev B* 37: 785
9. Perdew JP (1991) In: Ziesche P, Eschrig H (eds) *Electronic structure of solids '91*. Akademie, Berlin, p 11
10. Becke AD (1993) *J Chem Phys* 98: 1372
11. Lacks DJ, Gordon RG (1993) *Phys Rev A* 47:4681
12. Laming GJ, Termath V, Handy NC (1993) *J Chem Phys* 99:8765
13. Becke AD (1993) *J Chem Phys* 98:5648
14. Stephens PJ, Delvin JF, Chabalowski CF, Frisch MJ (1994) *J Phys Chem* 98:11623
15. Holthausen MC, Mohr M, Koch W (1995) *Chem Phys Lett* 240:245
16. Becke AD (1996) *J Chem Phys* 104:1040
17. Gill PMW (1996) *Mol Phys* 89:433
18. Kohn W, Becke AD, Parr RG (1996) *J Phys Chem* 100:12974
19. Perdew JP, Burke K, Enzerhof M (1996) *Phys Rev Lett* 77:3865
20. Perdew JP, Burke K, Enzerhof M (1997) *Phys Rev Lett* 78:1396
21. Filatov M, Thiel W (1997) *Mol Phys* 91:847
22. Burke K, Enzerhof M, Perdew JP (1997) *Chem Phys Lett* 265:115
23. Adamo C, Barone V (1997) *Chem Phys Lett* 274:242
24. Becke AJ (1997) *J Chem Phys* 107:8554
25. Adamo C, Barone V (1998) *J Chem Phys* 108:664
26. Schmider HL, Becke AD (1998) *J Chem Phys* 108:9624
27. van Voorhis T, Scuseria GE (1998) *J Chem Phys* 109:400
28. Hamprecht FA, Cohen AJ, Tozer DJ, Handy NC (1998) *J Chem Phys* 109:6264
29. Kafafi SA (1998) *J Phys Chem A* 102:10404
30. Adamson RD, Gill PMW, Pople JA (1998) *Chem Phys Lett* 284:6
31. Adamo C, DiMatteo A, Barone V (2000) *Adv Quantum Chem* 36:45
32. Becke AD (2000) *J Chem Phys* 112:4020
33. Lynch BJ, Fast PL, Harris M, Truhlar DG (2000) *J Phys Chem A* 104:4811
34. Proynov E, Chermette H, Salahub DR (2000) *J Chem Phys* 113:10013
35. Hoe W-M, Cohen AJ, Handy NC (2001) *Chem Phys Lett* 341:319
36. Wilson PJ, Bradley TJ, Tozer DJ (2001) *J Chem Phys* 115:9233
37. Handy NC, Cohen AJ (2001) *Mol Phys* 99:403
38. Poater J, Duran M, Solà M (2001) *J Comput Chem* 22:1666
39. Pu J, Truhlar DG (2002) *J Chem Phys* 116:1468
40. Sancho-García JC, Moscardó F (2003) *J Chem Phys* 119:1054
41. Lynch BJ, Zhao Y, Truhlar DG (2003) *J Phys Chem A* 107:1384
42. Kormos BL, Cramer CJ (2003) *J Org Chem* 68:6375
43. Staroverov VN, Scuseria GE, Tao J, Perdew JP (2003) *J Chem Phys* 119:12129
44. Zhao Y, Lynch BJ, Truhlar DG (2004) *J Phys Chem A* 108:2715
45. Xin X, Goddard WA III (2004) *Proc Natl Acad Sci USA* 101:2673
46. Zhao Y, Truhlar DG (2004) *J Phys Chem A* (submitted)
47. Car R, Parrinello M (1985) *Phys Rev Lett* 55:2471
48. Chen JL, Noodleman DA, Case DA, Bashford D (1994) *J Phys Chem* 98:11059
49. Adamo C, Leji F (1994) *Chem Phys Lett* 223:54
50. Rashin AA, Bukatin MA, Andelm J, Hagler AT (1994) *Biophys Chem* 51:375
51. Davidson MM, Hiller IH, Hall RJ, Burton NA (1994) *Mol Phys* 83:327
52. Fortunelli A, Tomasi J (1994) *Chem Phys Lett* 231:34
53. Truong TN, Stefanovich EV (1995) *Chem Phys Lett* 240:253
54. Truong TN, Stefanovich EV (1995) *J Chem Phys* 103:3709
55. Stefanovich EV, Truong TN (1995) *Chem Phys Lett* 244:65
56. Andzelm J, Kölmel C, Klamt A (1995) *J Chem Phys* 103:9312
57. Contreras R, Perez P, Aizman A (1995) *Int J Quantum Chem* 56:433
58. Tawa GJ, Martin RL, Pratt LR, Russo TV (1996) *J Phys Chem* 100:1515
59. Cossi M, Barone R, Cammi R, Tomasi J (1996) *Chem Phys Lett* 255:327
60. Mineva T, Russo N (1997) *Int J Quantum Chem* 61:665
61. Strnad M, Martins-Costa MTC, Millot C, Tuñón I, Ruiz-López MF, Rivail JL (1997) *J Chem Phys* 106:3643
62. Kallies B, Mitzner R (1997) *J Phys Chem B* 101:2959
63. Barone V, Bencini A, di Matteo A (1997) *J Am Chem Soc* 119:10831
64. Tuñón I, Martins-Costa MTC, Millot C, Ruiz-López MF, Rivail JL (1996) *J Comput Chem* 17:19
65. Rivail JL, Rinaldi D, Dillet V (1996) *Mol Phys* 89:1521
66. Barone V, Cossi M (1998) *J Phys Chem A* 102:1995
67. Cancès E, Mennucci B, Tomasi J (1998) *J Chem Phys* 109:260
68. Zhu T, Li J, Hawkins GD, Cramer CJ, Truhlar DG (1998) *J Chem Phys* 109:9117
69. Langgard M, Spanget-Larsen J (1998) *J Mol Struct (Theor Chem)* 431:173
70. Sorescu DC, Bennett CM, Thompson DL (1998) *J Phys Chem A* 102:10348
71. Schüürmann G (1998) *J Chem Phys* 109:9523
72. Li J, Zhu T, Hawkins GD, Winget P, Liotard DA, Cramer CJ, Truhlar DG (1999) *Theor Chem Acc* 103:9
73. Vreven T, Mennucci B, da Silva CO, Morokuma K, Tomasi J (2001) *J Chem Phys* 115:62
74. Kovalenko A, Hirata F (2001) *J Mol Liq* 90:215
75. Topol IA, Tawa GJ, Caldwell RA, Eissenstat MA, Burt SK (2000) *J Phys Chem A* 104:9619
76. Castejon H, Wiberg KB (1999) *J Am Chem Soc* 121:2139
77. Li J, Cramer CJ, Truhlar DG (1999) *Biophys Chem* 78:147
78. Kovalenko A, Hirata F (1999) *J Chem Phys* 110:10095
79. Safi B, Choho K, De Proft F, Geerlings P (1999) *Chem Phys Lett* 300:85
80. Cossi M, Scalmani G, Rega N, Barone V (2002) *J Chem Phys* 117:43
81. Eckert F, Klamt A (2002) *AIChE J* 48:369
82. Klicic JJ, Friesner RA, Liu S-Y, Guida WC (2002) *J Phys Chem A* 106:1327
83. Fattebert J-L, Gygi F (2003) *Int J Quantum Chem* 93:139
84. Ye Y, McCoy JD, Curro JG (2003) *J Chem Phys* 119:555
85. Da Silva CO, Mennucci B, Vreven T (2003) *J Phys Chem A* 107:6630
86. Thompson JD, Cramer CJ, Truhlar DG (2004) *J Phys Chem A* 108:6532
87. Rinaldi D, Bouchy A, Rivail JL (2004) *J Chem Phys* 120:2343
88. Miertus S, Scrocco E, Tomasi J (1981) *Chem Phys* 55:117
89. Hoijtink GJ, de Boer E, van der Meij PH, Weijland WP (1956) *Recl Trav Chim Pays-Bas* 75:487
90. Paradejordi F (1963) *Cahiers Phys* 17:393
91. Jano I (1965) *Compt Rend Acad Sci (Paris)* 261:103
92. Daudel R (1973) *Quantum theory of chemical reactivity*. Reidel, Dordrecht
93. Tucker SC, Truhlar DG (1989) *Chem Phys Lett* 157:164
94. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) *J Am Chem Soc* 112:6127
95. Cramer CJ, Truhlar DG (1991) *J Am Chem Soc* 113:8305
96. Kirkwood JG (1935) *J Chem Phys* 3:300
97. Tobias DJ, Brooks CLI (1987) *Chem Phys Lett* 142:472
98. Tomasi J, Persico M (1994) *Chem Rev* 94:2027
99. Cramer CJ, Truhlar DG (1996) In: Tapia O, Bertrán J (eds) *Solvent effects and chemical reactivity*. Kluwer, Dordrecht, p 1
100. Cramer CJ, Truhlar DG (1999) *Chem Rev* 99:2161
101. Roux B, Simonson T (1999) *Biophys Chem* 78:1
102. Winget P, Hawkins GD, Cramer CJ, Truhlar DG (2000) *J Phys Chem B* 104:4726
103. Cramer CJ, Truhlar DG (2001) In: Reddy MR, Erion MD (eds) *Free energy calculations in rational drug design*. Kluwer Academic/Plenum, New York, p 63
104. Thompson JD, Cramer CJ, Truhlar DG (2003) *J Chem Phys* 119:1661

105. Bonaccorsi R, Cimiraaglia R, Tomasi J (1983) *J Comput Chem* 4:567
106. Bianco R, Miertus S, Persico M, Tomasi J (1992) *Chem Phys* 168:281
107. Aguilar M, Bianco R, Miertus S, Persico M, Tomasi J (1993) *Chem Phys* 174:397
108. Basilevsky MV, Chudinov GE, Napolov DV (1993) *J Phys Chem* 97:3270
109. Truhlar DG, Schenter GK, Garrett BC (1993) *J Chem Phys* 98:5756
110. Aguilar MA, Olivares del Valle FJ, Tomasi J (1993) *J Chem Phys* 98:7375
111. Ruiz-López MF, Rinaldi D, Bertrán J (1995) *J Chem Phys* 103:9249
112. Sanchez ML, Aguilar MA, Olivares del Valle FJ (1995) *J Phys Chem* 99:15758
113. Chuang Y-Y, Truhlar DG (1999) *J Am Chem Soc* 121:10157
114. Li J, Cramer CJ, Truhlar DG (2000) *Int J Quantum Chem* 77:264
115. Storer JW, Giesen DJ, Cramer CJ, Truhlar DG (1995) *J Comput-Aid Mol Des* 9:87
116. Winget P, Thompson JD, Xidos JD, Cramer CJ, Truhlar DG (2002) *J Phys Chem A* 106:10707
117. Thompson JD, Cramer CJ, Truhlar DD (2003) *J Comput Chem* 24:1291
118. Thompson JD, Xidos JD, Sonbuchner TM, Cramer CJ, Truhlar DG (2002) *Phys Chem Commun* 5:117
119. Hermann RB (1972) *J Phys Chem* 76:2754
120. Amidon GL, Yalkowsky SH, Anik ST, Valvani SC (1975) *J Phys Chem* 79:2239
121. Rose GD, Geselowitz AR, Lesser GJ, Lee RH, Zehfus MH (1985) *Science* 229:834
122. Eisenberg D, McLachlan AD (1986) *Nature* 319:199
123. Ooi T, Oobatake M, Nemethy G, Scheraga HA (1987) *Proc Natl Acad Sci USA* 84:3086
124. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2003) *Gaussian03*, Gaussian, Pittsburgh, PA
125. Barone V, Cossi M, Tomasi J (1997) *J Chem Phys* 107:3210
126. Pierotti RA (1976) *Chem Rev* 76:717
127. Kitaigorodskii AI (1961) *Tetrahedron* 14:230
128. Huron MJ, Claverie P (1969) *Chem Phys Lett* 4:429
129. Huron MJ, Claverie P (1972) *J Phys Chem* 76:2123
130. Huron MJ, Claverie P (1974) *J Phys Chem* 78:1862
131. Floris F, Tomasi J (1989) *J Comput Chem* 10:616
132. Floris FM, Tomasi J, Pascual-Ahuir JL (1991) *J Comput Chem* 12:784
133. Cossi M, Mennucci B, Cammi R (1996) *J Comput Chem* 17:57
134. Bondi A (1964) *J Phys Chem* 68:441
135. Klamt A, Schüürmann G (1993) *J Chem Soc Perkin Trans* 2 799
136. Cossi M, Rega N, Scalmani G, Barone V (2003) *J Comput Chem* 24:669
137. Miertus S, Tomasi J (1982) *Chem Phys* 65:239
138. Cancès MT, Mennucci B, Tomasi J (1997) *J Chem Phys* 107:3032
139. Cossi M, Barone V, Mennucci B, Tomasi J (1998) *Chem Phys Lett* 286:253
140. Mennucci B, Tomasi J (1997) *J Chem Phys* 106:5151
141. Zhu T, Li J, Liotard DA, Cramer CJ, Truhlar DG (1999) *J Chem Phys* 110:5503
142. Chuang Y-Y, Cramer CJ, Truhlar DG (1998) *Int J Quantum Chem* 70:887
143. Chuang Y-Y, Radhakrishnan ML, Fast PL, Cramer CJ, Truhlar DG (1999) *J Phys Chem A* 103:4893
144. Assfeld X, Rinaldi D (1995) *AIP Con Proc* 330:59
145. Rivail D-L, Rinaldi D, Dillet V (1998) *J Chim Phys Phys-Chim Biol* 95:1818
146. Easton RE, Giesen DJ, Welch A, Cramer CJ, Truhlar DG (1996) *Theor Chim Acta* 93:281
147. Li J, Cramer CJ, Truhlar DG (1998) *Theor Chem Acc* 99:192
148. Hehre WJ, Radom L, Schleyer PvR, Pople JA (1986) *Ab initio molecular orbital theory*. Wiley, New York
149. Cramer CJ, Truhlar DG (1992) *Science* 256:213
150. Cramer CJ, Truhlar DG (1992) *J Comput Chem* 13:1089
151. Cramer CJ, Truhlar DG (1992) *J Comput-Aided Mol Des* 6:629
152. Giesen DJ, Cramer CJ, Truhlar DG (1995) *J Phys Chem* 99:7137
153. Giesen DJ, Storer JW, Cramer CJ, Truhlar DG (1995) *J Am Chem Soc* 117:1057
154. Chambers CC, Hawkins GD, Cramer CJ, Truhlar DG (1996) *J Phys Chem* 100:16385
155. Hawkins GD, Cramer CJ, Truhlar DG (1996) *J Phys Chem* 100:19824
156. Giesen DJ, Gu MZ, Cramer CJ, Truhlar DG (1996) *J Org Chem* 61:8720
157. Giesen DJ, Chambers CC, Cramer CJ, Truhlar DG (1997) *J Phys Chem B* 101:2061
158. Hawkins GD, Cramer CJ, Truhlar DG (1997) *J Phys Chem B* 101:7147
159. Giesen DJ, Hawkins GD, Liotard DA, Cramer CJ, Truhlar DG (1997) *Theor Chem Acc* 98:85
160. Hawkins GD, Cramer CJ, Truhlar DG (1998) *J Phys Chem B* 102:3257
161. Li J, Hawkins GD, Cramer CJ, Truhlar DG (1998) *Chem Phys Lett* 288:293
162. Hawkins GD, Liotard DA, Cramer CJ, Truhlar DG (1998) *J Org Chem* 63:4305
163. Winget P, Thompson JD, Cramer CJ, Truhlar DG (2002) *J Phys Chem A* 106:5160
164. Cramer CJ, Truhlar DG (1995) In: Boyd DB, Lipkowitz KB (eds) *Reviews in computational chemistry*, vol 6. VCH, New York, p 1
165. Rivail J-L, Rinaldi D (1996) In: Leszczynski J (ed.) *Computational chemistry, review of current trends*, vol 1. World Scientific, New York, p 139
166. Lee B, Richards FM (1971) *J Mol Biol* 55:379
167. Abraham MH, Grellier PL, Prior DV, Duce PP, Morris JJ, Taylor PJ (1989) *J Chem Soc Perkin Trans* 2: 699
168. Abraham MH (1993) *Chem Soc Rev* 22:73
169. Abraham MH (1993) *J Phys Org Chem* 6:660
170. Abraham MH, Weatherby PK (1994) *J Pharm Sci* 83:1087
171. Abraham MH (1994) In: Politzer P, Murray JS (eds) *Theoretical and computational chemistry series*, vol 1. Elsevier, Amsterdam, p 83
172. Ben-Naim A (1987) *Solvation thermodynamics*. Plenum, New York
173. Born M (1920) *Z Phys* 1:45
174. Liotard DA, Hawkins GD, Lynch GC, Cramer CJ, Truhlar DG (1995) *J Comput Chem* 16:422
175. Qui D, Shenkin P, Hollinger F, Still WC (1997) *J Phys Chem A* 101:3005
176. Jayaram B, Liu Y, Beveridge DJ (1998) *J Chem Phys* 109:1465
177. Srinivasan J, Trevathan MW, Beroza P, Case DA (1999) *Theor Chem Acc* 101:426
178. Curutchet C, Cramer CJ, Truhlar DG, Ruiz-López MF, Rinaldi D, Orozco M, Luque FJ (2003) *J Comput Chem* 24:284

179. Onufriev A, Bashford D, Case DA (2000) *J Phys Chem B* 104:3712
180. Winget P, Dolney DM, Giesen DJ, Cramer CJ, Truhlar DG (1999) Minnesota solvent descriptor database, version date 9 July 1999
181. Dolney DM, Hawkins GD, Winget P, Liotard DA, Cramer CJ, Truhlar DG (2000) *J Comput Chem* 21:340
182. Shinto H, Morisada S, Miyahara M, Higashitani K (2003) *J Chem Eng Japan* 36:57
183. Pearson RG (1986) *J Am Chem Soc* 108:6109
184. Florián J, Warshel A (1997) *J Phys Chem B* 101:5583
185. Hansch C, Maloney PP, Fujita T, Muir RM (1962) *Nature* 194:178
186. Fujita T, Iwasa J, Hansch C (1964) *J Am Chem Soc* 86:5175
187. Yalkowsky SH, Valvani SC (1980) *J Pharm Sci* 69:912
188. Venable RM, Zhang Y, Barry J, Pastor RW (1993) *Science* 262:223
189. Cramb DT, Wallace SC (1997) *J Phys Chem B* 101:2741
190. Ramos EU, Vaes WHJ, Verhaar HJM, Hermens JLM (1998) *J Chem Inf Comput Sci* 38:845
191. Faller B, Spanka C, Sergejew T, Tschinke V (2000) *J Med Chem* 43:1467
192. Cabrera P, Miguel A, Gonzalez DH, Fernandex TC, Pladelfina JM, Bermejo SM(2002) *Eur J Pharm Biopharm* 53:317
193. Aliste MP, MacCallum JL, Tieleman DP (2003) *Biochemistry* 42:8976
194. Carl DR, Feller SE (2003) *Langmuir* 19:8560
195. Warner KS, Li SK, He N, Suhonen TM, Chantasart D, Bolikal D, Higuchi WI (2003) *J Pharm Sci* 92:1305
196. Nakamura H, Xidos JD, Thompson JD, Li J, Hawkins GD, Zhu T, Lynch BJ, Volobuev Y, Rinaldi D, Liotard DA, Cramer CJ, Truhlar DG (2004) HONDOPLUS version 4.6. MN
197. Dupuis M, Marquez A, Davidson ER (1999) HONDO 99.6, based on HONDO 95.3, Dupuis M, Marquez A, Davidson ER Quantum Chemistry Program Exchange, Indiana University, IN
198. Thompson JD, Lynch BJ, Xidos JD, Li J, Hawkins GD, Zhu T, Volobuev Y, Rinaldi D, Liotard DA, Cramer CJ, Truhlar DG (2004) SMXGAUSS version 2.0. University of Minnesota, Minneapolis, MN