

# Prediction of Vapor Pressures from Self-Solvation Free Energies Calculated by the SM5 Series of Universal Solvation Models

Paul Winget, Gregory D. Hawkins, Christopher J. Cramer, and Donald G. Truhlar\*

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

Received: July 15, 1999

The SM5.4, SM5.2R, and SM5.0R solvation models are applied to calculate the vapor pressure of 156 molecules. This provides a test of the solvation models, but—even more so—it provides a useful scheme for estimating the vapor pressure at 298 K of any organic molecule for which a few standard descriptors are known or can be estimated.

## 1. Introduction

Vapor pressure is a fundamental property of any substance, and there is considerable interest in its prediction for innumerable reasons, especially in chemical engineering and in environmental fate assessment. In many cases, experimental vapor pressure measurements are not available, and a variety of semiempirical methods are available for estimating the vapor pressure in such cases.

Most methods for estimating vapor pressure require several other experimental parameters, e.g., critical properties, heats of vaporization, boiling points, second virial coefficients, or even vapor pressures at other temperatures.<sup>1–9</sup> Recently developed methods based upon group-additive contributions reduce the needed input to the molecular structure and are reasonably accurate for chemical functionalities similar to those for which they were parameterized.<sup>10,11</sup> Multilinear relationships between vapor pressure and descriptors other than functional groups have also been explored. Quantitative structure–property relationships (QSPRs) between vapor pressure and various gas-phase molecular properties computed at either classical<sup>12–14</sup> or quantum mechanical<sup>15</sup> levels of theory have been explored. Vapor pressures have also been predicted from classical mechanical molecular simulations<sup>16</sup> and quantum mechanical reaction fields.<sup>17–19</sup> The present paper presents a method with certain similarities to these latter approaches. In particular, we test how well vapor pressures can be predicted by the SM5 series of universal solvation models.

The SM5 series of solvation models have been previously parameterized using experimental free energies of solvation.<sup>20–30</sup> No new parameters are required for the prediction of vapor pressures, and the only properties required for the molecule whose vapor pressure is to be predicted are its molecular structure, its liquid density, six solvent descriptors, and also (in two of the three SM5 models considered here) its dielectric constant. (Of course the density, any of the molecular descriptors, or the dielectric constant may be estimated if they are not known, and these properties are typically more widely available or easily estimated than are the vapor pressures of liquids.) An advantage of the SM5 series of solvation models is that they are applicable to virtually any substance composed of H, C, N, O, F, P, S, Cl, Br, and/or I. A disadvantage is that they are only parameterized at 298 K, so one may calculate the vapor pressure directly only at this one temperature. (One could

imagine extending the predictions to other temperatures by using additional assumptions such as Trouton's rule and the Clausius–Clapeyron equation, but that is beyond the scope of this article.)

## 2. Theory

**2.1. Relation of Vapor Pressure to Free Energy of Solvation.** Very often in chemical thermodynamics, one uses a standard state of one atmosphere of partial pressure for gases and one mole per liter of molarity for liquid-phase solutes. However, the relationship between intermolecular interactions and free energies of transfer between phases is most direct if one uses the same standard concentration for both phases, for example, a standard state of one mole per liter for both gas-phase and liquid-phase components. When this consistent standard-state convention is adopted, and assuming ideal behavior, the standard-state molar free energy of solvation of A in B at temperature  $T$  is given by<sup>31,32</sup>

$$\Delta G_s^\circ = -RT \ln \frac{M_{\text{liq}}}{M_{\text{vapor}}} \quad (1)$$

where  $R$  is the universal gas constant,  $M_{\text{liq}}$  is the molarity of A in liquid B, and  $M_{\text{vapor}}$  is the molarity of the vapor of A in equilibrium with this solution. If we now consider the case where A is the same as B,  $M_{\text{liq}}$  becomes the molarity of the pure liquid and  $M_{\text{vapor}}$  becomes the molarity of the vapor in equilibrium with it. Let the standard-state molarity  $M^0$  be 1 mole per liter. At 298 K, the pressure of an ideal gas of this molarity is 24.45 atm (or 18583 Torr). We call this the standard state pressure  $P^0$ , and we denote the equilibrium vapor pressure of the pure liquid as  $P_{\text{vapor}}$ . Since

$$\frac{M_{\text{liq}}}{M_{\text{vapor}}} = \frac{M_{\text{liq}}/M^0}{M_{\text{vapor}}/M^0} = \frac{M_{\text{liq}}/M^0}{P_{\text{vapor}}/P^0} \quad (2)$$

and since  $M^0 = 1$ , eq 1 becomes

$$\Delta G_s^\circ = RT \ln \frac{P_{\text{vapor}}}{P^0 M_{\text{liq}}} \quad (3)$$

where  $\Delta G_s^\circ$  is the standard-state solvation energy of A in A. Rearranging yields

$$P_{\text{vapor}} = P^0 M_{\text{liq}} \exp(\Delta G_{\text{S}}^{\circ}/RT) \quad (4)$$

Using this equation we can predict the vapor pressure of any liquid from its molecular weight and density (which are used to calculate  $M_{\text{liq}}$ ) and its predicted solvation free energy in self-solution.

**2.2. SM5 Universal Solvation Models.** The SM5 series of universal solvation models is presented in detail elsewhere. Here we summarize the essential features of the three models we test here, namely SM5.4,<sup>20–22,24</sup> SM5.2R,<sup>26</sup> and SM5.0R.<sup>23,28</sup> The expressions of the standard-state free energy of solvation in these three models are

$$\Delta G_{\text{S}}^{\circ}(\text{SM5.4}) = \Delta G_{\text{ENP}} + G_{\text{CDS}} \quad (5)$$

$$\Delta G_{\text{S}}^{\circ}(\text{SM5.2R}) = \Delta G_{\text{EP}} + G_{\text{CDS}} \quad (6)$$

$$\Delta G_{\text{S}}^{\circ}(\text{SM5.0R}) = G_{\text{CDS}} \quad (7)$$

where  $\Delta G_{\text{ENP}}$  is the change in solute electronic and nuclear energy and solvent electronic polarization energy when a solute is immersed in a continuous medium of dielectric constant  $\epsilon$ ,  $\Delta G_{\text{EP}}$  is an approximation to this in which solute nuclear relaxation is not included (i.e., for computational purposes, the geometry in liquid solution or neat liquid is assumed to be the same as in the gas phase), and  $G_{\text{CDS}}$  is a cavitation–dispersion–solvent–structure term. The  $\Delta G_{\text{ENP}}$  and  $\Delta G_{\text{EP}}$  terms account for bulk electrostatic effects. The  $G_{\text{CDS}}$  term in its most general form is written as a sum over contributions from individual atoms  $k$  of the solute as follows:

$$G_{\text{CDS}} = \sum_k \sum_i \sum_j \sum_{\delta} A_k(\mathbf{R}, r_i) \sigma_{Z_k ij \delta} f_j(\{Z_{k'}, R_{k'k''}\}) S_{\delta} \quad (8)$$

In this expression, the sum over  $i$  involves one or two terms corresponding to effective solvent radii  $r_i$ ;  $A_k(\mathbf{R}, r_i)$  is the solvent-accessible surface area of atom  $k$ , which depends on the complete three-dimensional geometry  $\mathbf{R}$  of the solute and on the effective solvent radius;  $\sigma_{Z_k ij \delta}$  is an atomic surface tension coefficient that depends on the atomic number  $Z_k$  of atom  $k$  ( $Z_k = 1, 6, 7, \dots$  for H, C, N, ...) and on the term indices  $i, j$ , and  $\delta$ ;  $f_j(\{Z_{k'}, R_{k'k''}\})$  is a geometrical factor that depends on the atom ( $k$ ) and the collection of all the atomic numbers  $Z_{k'}$  and interatomic distances  $R_{k'k''}$  in the molecule; and  $S_{\delta}$  is a solvent descriptor. We use one solvent descriptor (arbitrarily set equal to unity) for water solvent and six for organic solvents (although there are seven terms in the sum over  $\delta$  since one of the descriptors,  $\beta$ , appears as  $\beta$  for  $S_3$  and as  $\beta^2$  for  $S_5$ ). The six descriptors are

- $n$  refractive index at the wavelength of the Na D line
- $\alpha$  Abraham's<sup>33–37</sup> hydrogen bond acidity parameters  $\sum \alpha_2$
- $\beta$  Abraham's<sup>33–37</sup> hydrogen bond basicity parameter  $\sum \beta_2$
- $\gamma$  macroscopic molecular surface tension in units of cal mol<sup>−1</sup> Å<sup>−2</sup>
- $\phi^2$  square of the fraction  $\phi$  of nonhydrogenic solvent atoms that are aromatic carbon atoms (aromaticity)
- $\psi^2$  square of the fraction  $\psi$  of nonhydrogenic solvent atoms that are F, Cl, or Br (electronegative halogenicity)

A critical element in the SM5 series of solvation models is that the atomic surface tension coefficients are determined semiempirically by regression on a large number (more than 2000) of solvation free energies. Thus, not only do they account for short-range cavitation, dispersion, and solvent-structure

interactions, but also they make up semiempirically for any systematic deficiencies in the modeling of long-range electrostatic interactions by  $\Delta G_{\text{ENP}}$  in SM5.4 or by  $\Delta G_{\text{EP}}$  in SM5.2R and for the total omission of explicit electrostatic modeling in SM5.0R. The SM5.2R and SM5.0R models are based on gas-phase geometries, and the semiempirical surface tension coefficients also make up for any systematic errors in the free energies of solvation due to that aspect of the model. The SM5.4 model, in contrast, includes separate geometry optimization both in the gas phase and in liquid solution.

Another difference of the SM5.4 and SM5.2R models is that the electrostatics part of the SM5.4 calculation is based on class IV charges<sup>33</sup>, whereas the SM5.2R electrostatics are based on class II charges.<sup>38</sup>

All three models treat water solvent as a special case. The SM5.4 model also treats chloroform, benzene, and toluene solvents as special cases. The nonspecial solvents are modeled with  $S_{\delta} = n, \alpha, \beta$ , and  $\gamma$  in the SM5.4 model and with  $S_{\delta} = n, \alpha, \beta, \gamma, \beta^2, \phi^2$ , and  $\psi^2$  in the SM5.2R and SM5.0R models.

The electrostatic terms  $\Delta G_{\text{ENP}}$  and  $\Delta G_{\text{EP}}$  are computed by quantum mechanical self-consistent reaction field methods. The present paper uses the semiempirical AM1<sup>39–41</sup> and PM3<sup>42</sup> molecular orbital methods for these steps. There are separate sets of surface tension coefficients for each choice of molecular orbital method.

### 3. Calculations

We selected 156 molecules as test cases. Of these, 77 are organic solvents that were used in the development of the universal solvation models, and the values of  $\epsilon, n, \alpha, \beta$ , and  $\gamma$  used for parameterization for 76 of these are given in ref 21. However, since then, we have improved our solvent descriptor database, and in the present work we always used the updated values. The updated values of all solvent descriptors used in this paper are given in the Minnesota Solvent Descriptor Database, which is available at <http://comp.chem.umn.edu/solvation>. The values of  $\phi^2$  and  $\psi^2$  are easily calculated from the molecular structures. The values of  $\epsilon, n, \alpha, \beta$ , and  $\gamma$  used in this paper, as well as the liquid densities, were taken from standard sources.<sup>34,35,37,43–46</sup>

The molecular geometries used for the calculations depend on the model. For the SM5.4/AM1 calculations, gas-phase geometries were optimized by the AM1 method,<sup>39–41</sup> and liquid-phase geometries were optimized by the SM5.4/AM1 method.<sup>20–22,24</sup> For the SM5.4/PM3 calculations, gas-phase geometries were optimized by the PM3 method<sup>42</sup> and liquid-phase geometries were optimized by the SM5.4/PM3 method.<sup>20–22,24</sup> For the SM5.2R/AM1 and SM5.2R/PM3 calculations, one set of calculations was performed with the geometries optimized by classical mechanics using the MM2<sup>47</sup> molecular mechanics force field, and another set was performed with the geometries optimized by the MM3<sup>48,49</sup> molecular mechanics force field. For the SM5.0R calculations, we also used these MM2 and MM3 geometries.

In a few cases, in the sets of calculations employing MM2 and MM3 force fields, a needed molecular mechanics parameter was unavailable. In those cases we substituted a different geometry for the unavailable geometry. The cases where this occurs are CS<sub>2</sub> and nitro compounds. For CS<sub>2</sub>, MM2 parameters are unavailable. For nitro compounds, optimized parameters are unavailable in both the MM2 and MM3 force fields. For these compounds AM1 gas-phase geometries were used for SM5.2R/AM1 calculations, and PM3 gas-phase geometries were used for SM5.2R/PM3 and for SM5.0R calculations. All cases where substituted geometries were used are footnoted in the tables.

**TABLE 1: Natural log of Vapor Pressures in Torr from the Free Energies of Self-Solvation**

compd	exptl	SM5.4		SM5.2R				SM5.0R	
		AM1	PM3	AM1/MM3	PM3/MM3	AM1/MM2	PM3/MM2	MM3	MM2
<i>n</i> -pentane	6.24	7.03	7.18	6.39	6.72	6.39	6.72	7.35	7.36
<i>n</i> -hexane	5.03	5.80	5.98	5.29	5.64	5.30	5.65	6.28	6.29
<i>n</i> -heptane	3.82	4.55	4.75	4.19	4.56	4.20	4.57	5.21	5.22
<i>n</i> -octane	2.64	3.30	3.52	3.12	3.51	3.13	3.53	4.16	4.17
<i>n</i> -nonane	1.49	2.05	2.30	2.06	2.47	2.09	2.49	3.11	3.13
<i>n</i> -decane	0.36	0.81	1.08	1.03	1.45	1.05	1.48	2.09	2.11
<i>n</i> -undecane	-0.89	-0.72	-0.41	-0.44	0.12	-0.27	0.19	0.90	0.87
<i>n</i> -dodecane	-2.00	-1.74	-1.41	-1.08	-0.62	-1.05	-0.58	0.01	0.05
<i>n</i> -pentadecane	-5.68	-5.50	-5.09	-4.14	-3.63	-4.09	-3.58	-3.03	-2.99
<i>n</i> -hexadecane	-6.55	-6.77	-6.33	-5.17	-4.64	-5.12	-4.59	-4.05	-4.00
2,4-dimethylpentane	4.59	5.14	5.27	5.23	5.56	5.21	5.55	6.31	6.32
2,2,4-trimethylpentane	3.90	4.46	4.54	4.77	5.14	4.74	5.12	6.02	6.00
cyclopentane	5.76	8.18	8.13	5.93	6.35	5.91	6.32	6.72	6.71
cyclohexane	4.59	7.54	7.70	4.88	5.37	4.90	5.38	5.81	5.83
methylcyclohexane	3.83	6.32	6.47	4.27	4.74	4.29	4.76	5.27	5.30
<i>cis</i> -1,2-dimethylcyclohexane	2.67	5.74	5.82	4.11	4.52	4.08	4.50	5.01	5.00
<i>cis</i> -decalin	-0.24	4.62	4.73	1.72	2.20	1.41	2.01	2.40	2.31
<i>trans</i> -decalin	0.20	4.31	4.43	1.42	1.94	1.43	1.95	2.27	2.29
1-pentene	6.46	7.01	6.97	7.15	7.10	7.15	7.09	7.26	7.26
<i>E</i> -2-pentene	6.23	6.65	6.73	6.64	6.67	6.63	6.66	6.90	6.90
1-hexene	5.23	5.78	5.76	6.04	6.01	6.05	6.01	6.17	6.17
1-hexyne	4.89	5.09	5.36	5.58	5.78	5.50	5.71	6.35	6.36
benzene	4.55	4.50	4.14	3.61	3.58	3.62	3.59	3.83	3.83
toluene	3.35	3.29	3.16	10.63	8.40	2.72	2.73	5.92	3.12
ethylbenzene	2.26	2.92	2.80	1.92	1.86	1.93	1.86	2.04	2.04
<i>o</i> -xylene	1.89	2.73	2.72	1.82	1.90	1.81	1.89	2.36	2.37
<i>m</i> -xylene	2.13	2.74	2.72	2.08	2.05	2.07	2.02	2.36	2.35
<i>p</i> -xylene	2.19	2.79	2.74	2.15	2.07	2.14	2.06	2.33	2.33
isopropylbenzene	1.51	2.03	1.97	1.60	1.46	1.54	1.40	1.52	1.49
<i>n</i> -butylbenzene	0.02	0.75	0.73	0.17	0.00	0.21	0.05	-0.10	-0.04
<i>tert</i> -butylbenzene	0.76	1.31	1.23	1.33	1.22	1.32	1.21	1.32	1.31
mesitylene	0.91	1.49	1.61	10.77	7.68	1.34	1.22	6.68	1.49
1,2,4-trimethylbenzene	0.82	1.61	1.71	1.24	1.20	1.28	1.22	1.51	1.54
benzyl alcohol	-2.51	-3.21	-3.40	-4.64	-4.52	-4.79	-4.64	-3.97	-4.06
tetralin	-1.00	0.79	1.01	-1.46	-1.08	-1.50	-1.12	-0.78	-0.78
methanol	4.84	5.42	5.20	4.91	5.14	4.83	5.06	5.77	5.76
ethanol	4.08	3.78	3.78	3.36	3.66	3.29	3.58	4.21	4.19
1-propanol	3.03	2.75	2.73	2.53	2.66	2.46	2.58	2.97	2.94
2-propanol	3.82	4.29	3.71	2.84	2.87	2.84	2.85	3.05	3.03
1-butanol	1.95	2.33	1.86	1.74	1.75	1.75	1.73	1.85	1.86
2-methyl-2-propanol	3.74	3.37	2.93	2.05	2.26	2.06	2.25	2.53	2.56
2-methyl-1-propanol	2.35	3.32	2.96	2.37	2.58	2.42	2.60	2.92	2.94
2-butanol	2.91	3.70	3.08	1.78	1.89	1.75	1.85	2.04	2.01
1-pentanol	0.90	1.04	0.59	0.81	0.75	0.82	0.73	0.75	0.76
1-hexanol	-0.08	-0.32	-0.94	0.07	-0.16	0.14	-0.12	-0.20	-0.12
<i>m</i> -cresol	-2.18	-2.18	-1.96	12.33	7.68	-3.02	-2.56	8.36	-2.28
1-heptanol	-1.53	-1.53	-2.15	-0.76	-1.05	-0.69	-1.01	-1.22	-1.14
1-octanol	-2.54	-2.63	-3.05	-1.81	-2.03	-1.77	-2.03	-2.35	-2.33
1-nonanol	-3.78	-4.44	-4.44	-2.71	-2.95	-2.74	-2.99	-3.30	-3.31
1-decanol	-4.77	-5.55	-5.55	-3.48	-3.78	-3.50	-3.81	-4.28	-4.27
tetrahydrofuran	5.09	5.61	5.96	4.99	5.36	5.06	5.37	6.02	6.01
1,4-dioxane	3.64	6.38	6.04	6.15	5.68	6.20	5.69	4.94	5.02
diethyl ether	6.28	5.05	5.59	5.92	6.16	5.92	6.15	6.53	6.54
anisole	1.28	1.84	2.04	1.68	2.00	1.69	1.95	1.78	1.71
diisopropyl ether	5.00	3.95	4.12	4.85	4.74	4.88	4.75	4.82	4.78
thiophene	4.36	3.65	4.20	2.73	3.54	2.63	3.50	4.36	4.34
ethyl phenyl ether	0.42	0.07	0.34	0.31	0.50	0.34	0.47	0.10	0.05
dibutyl ether	1.80	1.06	1.40	2.83	2.57	2.87	2.59	2.10	2.15
propanal	5.76	4.47	4.30	4.32	4.46	4.34	4.47	6.38	6.39
butanal	4.71	3.69	3.52	3.74	3.74	3.73	3.72	5.18	5.17
pentanal	3.54	2.94	2.62	3.17	2.85	3.19	2.84	3.87	3.87
benzaldehyde	0.17	-0.84	-1.11	-2.51	-2.20	-2.63	-2.35	-0.88	-0.96
acetone	5.44	3.82	3.82	3.82	4.03	3.77	3.97	6.03	6.04
butanone	4.57	2.82	2.86	2.92	3.21	2.89	3.17	4.95	4.96
cyclopentanone	2.43	3.74	3.64	2.12	2.17	1.99	2.01	2.88	2.87
2-pentanone	3.62	1.78	1.82	2.16	2.32	2.10	2.25	3.74	3.73
3-pentanone	3.61	1.87	1.90	2.07	2.39	2.05	2.36	3.87	3.88
2-hexanone	2.37	0.61	0.66	1.22	1.36	1.18	1.30	2.64	2.64
4-methyl-2-pentanone	2.99	1.16	1.21	2.02	2.18	1.98	2.13	3.27	3.30
cyclohexanone	1.46	2.63	2.81	1.09	1.34	1.07	1.30	2.04	2.05
2-heptanone	1.33	-0.47	-0.40	0.41	0.47	0.38	0.42	1.51	1.52
acetophenone	-0.92	-2.30	-2.44	-3.39	-3.06	-3.78	-3.47	-2.19	-2.24

TABLE 1 (Continued)

compd	exptl	SM5.4		SM5.2R				SM5.0R	
		AM1	PM3	AM1//MM3	PM3//MM3	AM1//MM2	PM3//MM2	MM3	MM2
2-octanone	-0.60	-1.69	-1.58	-0.56	-0.53	-0.59	-0.57	0.34	0.36
acetic acid	2.75	4.72	4.52	0.87	1.06	1.38	1.49	1.26	1.32
propanoic acid	1.31	5.29	5.13	1.52	1.60	1.92	1.93	0.14	0.16
butanoic acid	-0.06	4.72	4.52	1.24	1.16	1.57	1.44	-0.92	-0.90
pentanoic acid	-1.41	3.84	3.66	0.58	0.48	0.90	0.75	-1.94	-1.93
hexanoic acid	-3.14	2.75	2.59	-0.14	-0.28	0.17	-0.02	-2.92	-2.91
methyl methanoate	6.37	5.32	5.26	5.67	5.95	5.63	5.88	7.97	7.91
ethyl methanoate	5.50	3.67	3.70	4.35	4.53	4.28	4.43	6.25	6.20
methyl ethanoate	5.37	3.77	3.98	3.86	4.39	4.43	4.80	5.94	5.90
methyl propanoate	4.43	2.97	3.17	3.40	3.80	3.91	4.16	4.74	4.70
ethyl ethanoate	4.54	2.14	2.45	2.73	3.10	3.23	3.47	4.27	4.23
methyl butanoate	3.47	2.14	2.27	2.89	3.08	3.33	3.39	3.57	3.52
propyl ethanoate	3.51	1.10	1.33	1.93	2.14	2.37	2.45	3.01	2.97
butyl ethanoate	2.44	0.12	0.31	1.16	1.26	1.61	1.59	1.88	1.84
pentyl ethanoate	1.25	-1.10	-0.89	0.20	0.25	0.64	0.58	0.71	0.68
methyl benzoate	-0.96	-1.32	-1.57	-1.68	-1.63	-1.84	-1.93	-2.17	-2.27
2-propen-1-ol	3.26	4.10	3.34	3.68	3.47	3.76	3.49	3.04	3.01
formic acid	3.75	5.73	5.25	1.17	1.24	1.16	1.24	3.21	3.26
water	3.17	-2.16	-0.72	3.43	3.38	3.50	3.42	0.00	0.00
dimethyl sulfoxide	-0.50	-6.49	-9.89	-15.85	-13.82	-5.29	-17.21	7.65	7.95
carbon disulfide	5.88	-0.74	1.19	2.97 <sup>a</sup>	3.16 <sup>b</sup>	2.97	3.16	1.42	1.42
propylamine	5.73	4.98	5.24	5.06	5.06	5.08	5.09	4.93	4.98
butylamine	4.52	3.78	4.04	4.03	3.99	4.06	4.03	3.83	3.87
diethylamine	5.47	3.94	2.70	3.10	2.59	2.51	2.02	3.20	2.86
pentylamine	3.40	2.33	2.58	2.83	2.75	2.87	2.79	2.54	2.59
dipropylamine	3.00	3.90	3.23	3.72	3.30	3.61	3.21	3.37	3.42
triethylamine	4.23	5.62	5.48	5.87	5.81	5.29	5.19	6.02	6.01
<i>N,N</i> -dimethylformamide	1.38	1.31	2.27	2.43	2.50	1.86	1.82	6.07	5.96
<i>N</i> -methylformamide	-1.38	-0.46	-2.92	0.38	-0.59	0.60	-0.60	3.77	3.79
<i>N,N</i> -dimethylacetamide	0.69	-0.93	0.78	0.56	0.97	-0.08	0.15	4.84	4.71
formamide	-2.81	1.02	-3.46	-0.82	-2.32	-0.84	-2.63	2.28	2.28
pyridine	3.03	2.42	3.92	1.97	2.95	2.09	3.02	3.44	3.43
aniline	-0.71	0.82	-0.09	-1.03	-1.66	-1.02	-1.33	0.27	0.26
2-methylpyridine	2.43	0.91	2.42	0.93	1.90	1.03	1.97	2.61	2.61
3-methylpyridine	1.80	1.52	2.81	1.39	2.20	1.47	2.24	2.70	2.69
4-methylpyridine	1.78	1.26	2.76	1.30	2.17	1.39	2.21	2.72	2.71
<i>N</i> -methylaniline	-0.79	0.34	-0.53	-0.74	-1.20	-0.63	-0.94	0.21	0.25
2,4-dimethylpyridine	-1.99	-0.21	1.35	0.36	1.21	0.41	1.23	1.96	1.96
2,6-dimethylpyridine	1.69	-0.48	0.94	0.07	0.96	0.14	1.01	1.81	1.82
acetonitrile	4.51	4.72	4.17	6.58	5.75	6.57	5.74	9.12	9.12
propanonitrile	3.86	4.35	4.13	6.10	5.30	6.07	5.24	7.85	7.86
butanonitrile	2.97	3.70	3.53	5.55	4.63	5.51	4.57	6.63	6.64
benzonitrile	-0.27	1.24	1.25	0.14	-0.09	0.14	-0.09	0.77	0.77
nitroethane	3.04	-0.17	2.31	-1.66 <sup>a</sup>	0.09 <sup>b</sup>	-1.66 <sup>a</sup>	0.09 <sup>b</sup>	6.56	6.56 <sup>b</sup>
1-nitropropane	2.32	-0.81	1.16	-1.90 <sup>a</sup>	-1.06 <sup>b</sup>	-1.90 <sup>a</sup>	-1.06 <sup>b</sup>	5.26	5.26 <sup>b</sup>
2-nitropropane	2.85	-1.03	1.68	-1.64 <sup>a</sup>	0.40 <sup>b</sup>	-1.64 <sup>a</sup>	0.40 <sup>b</sup>	5.79	5.79 <sup>b</sup>
nitrobenzene	-1.39	-3.07	-2.94	-4.72 <sup>a</sup>	-4.74 <sup>b</sup>	-4.72 <sup>a</sup>	-4.74 <sup>b</sup>	-0.47	-0.47 <sup>b</sup>
nitromethane	3.58	0.70	2.88	-1.47 <sup>a</sup>	0.08 <sup>b</sup>	-1.47 <sup>a</sup>	0.08 <sup>b</sup>	8.02	8.02 <sup>b</sup>
<i>o</i> -nitrotoluene	-1.69	-3.35	-3.20	-4.62 <sup>a</sup>	-5.34 <sup>b</sup>	-4.62 <sup>a</sup>	-5.34 <sup>b</sup>	-1.06	-1.06 <sup>b</sup>
ethanethiol	6.27	7.43	7.65	8.22	8.28	8.20	8.26	8.35	8.37
thiophenol	0.41	0.68	0.93	0.70	0.73	0.78	0.84	1.29	1.22
diethyl sulfide	4.07	4.00	4.27	5.68	5.65	5.72	5.69	5.49	5.53
dimethyl disulfide	3.36	4.19	4.69	4.32	4.42	4.22	4.34	5.40	5.39
fluorobenzene	4.35	4.88	4.71	2.97	3.30	2.99	3.32	4.69	4.69
perfluorobenzene	4.44	12.16	11.25	12.24	10.51	12.26	10.52	7.37	7.37
dichloromethane	6.07	6.91	6.78	5.98	7.95	5.80	7.88	7.31	7.34
chloroform	5.28	5.48	5.33	4.62	5.73	4.74	5.75	4.78	4.74
1,1,1-trichloroethane	4.82	4.15	4.60	3.25	4.61	3.30	4.62	4.38	4.35
1,1,2-trichloroethane	3.15	4.02	3.90	2.90	4.97	2.68	4.88	3.81	3.82
1-chloropropane	6.27	6.49	6.43	6.04	7.11	5.86	7.00	7.29	7.29
2-chlorobutane	5.06	5.36	5.39	5.03	6.13	4.91	6.07	6.55	6.55
1-chloropentane	3.49	4.04	4.01	4.01	5.01	3.86	4.92	5.07	5.08
1,2-dichloroethane	4.37	5.18	4.99	4.19	6.50	4.94	6.32	6.12	6.14
carbon tetrachloride	4.74	5.18	5.27	3.89	3.58	3.95	3.64	2.89	2.81
<i>Z</i> -1,2-dichloroethene	5.31	6.23	6.17	6.28	7.87	6.27	7.88	7.99	8.01
<i>E</i> -1,2-dichloroethene	5.81	7.66	7.47	8.02	7.85	8.01	7.86	7.37	7.38
trichloroethene	4.30	5.58	5.61	5.84	5.48	5.82	5.50	5.51	5.54
chlorobenzene	2.48	2.49	2.68	0.99	1.33	1.00	1.34	2.73	2.73
<i>o</i> -dichlorobenzene	0.30	0.36	0.77	-0.98	-0.54	-0.97	-0.52	1.39	1.39
$\alpha$ -chlorotoluene	0.27	0.34	0.29	-0.68	0.33	-0.90	0.22	0.53	0.55
<i>o</i> -chlorotoluene	1.26	1.60	1.92	0.40	0.66	0.40	0.65	2.05	2.05
dibromomethane	3.81	5.70	5.60	5.22	4.89	5.21	4.80	5.29	5.30



TABLE 1 (Continued)

compd	exptl	SM5.4		SM5.2R				SM5.0R	
		AM1	PM3	AM1//MM3	PM3//MM3	AM1//MM2	PM3//MM2	MM3	MM2
bromoform	1.71	3.59	3.49	1.87	1.77	1.90	1.79	1.29	1.31
bromoethane	6.15	6.44	6.27	6.72	6.20	6.66	6.05	7.55	7.55
1-bromopropane	4.93	5.50	5.37	5.80	5.35	5.73	5.21	6.37	6.36
2-bromopropane	5.38	5.13	4.97	5.60	5.03	5.56	4.96	6.50	6.50
bromobenzene	1.45	1.20	1.11	-0.61	0.07	-0.60	0.07	1.14	1.14
diiodomethane	0.19	3.89	3.89	4.32	4.55	4.34	4.39	2.57	2.58
iodomethane	6.00	7.08	7.32	8.14	6.94	8.11	6.64	7.90	7.90
iodoethane	4.91	5.40	5.68	6.11	4.87	6.09	4.61	6.07	6.06
1-iodopropane	3.77	4.35	4.55	4.99	4.12	4.97	3.90	4.82	4.81
iodobenzene	0.06	0.00	-0.10	-2.00	-0.33	-1.93	-0.45	-0.91	-1.01
2,2,2-trifluoroethanol	4.27	-0.36	0.04	1.91	2.80	1.76	2.72	4.94	5.00
tetrachloroethene	2.92	4.46	4.56	4.93	3.49	4.94	3.49	3.72	3.73

<sup>a</sup> Geometry by AM1. <sup>b</sup> Geometry by PM3.TABLE 2: Mean Unsigned Errors in the Free Energies of Self-Solvation (in kcal/mol)<sup>a</sup>

functional group	data points	SM5.4		SM5.2R				SM5.0R	
		AM1	PM3	AM1//MM3	PM3//MM3	AM1//MM2	PM3//MM2	MM3	MM2
unbranched alkanes	10	0.28	0.42	0.40	0.66	0.42	0.68	1.04	1.05
branched alkanes	2	0.33	0.39	0.45	0.66	0.43	0.65	1.14	1.13
alkenes and alkynes	4	0.25	0.30	0.39	0.41	0.37	0.39	0.57	0.58
arenes	13	0.39	0.42	1.02	0.75	0.30	0.26	0.60	0.26
cycloalkanes	6	1.97	2.02	0.55	0.82	0.51	0.80	1.05	1.05
alcohols and water	16	0.42	0.39	0.90	0.67	0.40	0.33	0.74	0.36
aldehydes	4	0.58	0.72	0.81	0.79	0.82	0.81	0.37	0.38
aliphatic amines	10	0.76	0.55	0.67	0.55	0.68	0.49	1.53	1.53
aromatic amines	8	0.71	0.58	0.58	0.50	0.55	0.45	0.63	0.63
carboxylic acids	5	2.59	2.48	0.99	0.94	1.09	1.03	0.51	0.50
esters	10	1.01	0.94	0.68	0.54	0.49	0.42	0.38	0.38
ethers	8	0.58	0.46	0.47	0.38	0.47	0.38	0.28	0.29
ketones	11	0.93	0.92	0.67	0.57	0.72	0.62	0.29	0.29
nitriles	4	0.44	0.40	1.08	0.67	1.07	0.65	1.97	1.97
nitrohydrocarbons	6	1.62	0.67	2.44	1.90	2.44	1.90	1.52	1.52
non-halo bifunctional	2	0.84	0.47	0.89	0.81	0.91	0.81	0.22	0.22
chloroalkanes	12	0.44	0.38	0.45	0.83	0.48	0.80	0.79	0.80
chloroarenes	4	0.07	0.20	0.68	0.39	0.71	0.39	0.35	0.36
brominated hydrocarbons	6	0.51	0.48	0.52	0.33	0.51	0.32	0.61	0.61
fluorinated hydrocarbons	2	2.45	2.13	2.72	2.11	2.72	2.11	0.97	0.97
iodinated hydrocarbons	5	0.70	0.80	1.27	0.72	1.26	0.69	0.88	0.89
thiols	2	0.42	0.56	0.66	0.69	0.68	0.71	0.87	0.86
other halo compds	2	1.83	1.74	1.30	0.60	1.34	0.63	0.44	0.46
other sulfur compds	4	2.00	2.32	3.09	2.77	1.51	3.26	2.38	2.43
overall <sup>b</sup>	156	0.80	0.74	0.86	0.75	0.71	0.68	0.80	0.74
subset <sup>c</sup>	145	0.70	0.68	0.79	0.70	0.63	0.62	0.78	0.71

<sup>a</sup> See footnotes *a* and *b* in Table 1; in a few cases, as explained in section 3, AM1 or PM3 geometries were used instead of MM3 or MM2 geometries. <sup>b</sup> All molecules. <sup>c</sup> 145 molecules, excluding carboxylic acids and nitro compounds.

In most cases, calculations were based on a single dominant conformation. Unbranched alkyl chains were always modeled in their most extended form. In the case of 1,2-dichloroethane, the two separate conformations were averaged using the standard<sup>26</sup> formula.

**Software.** AM1, PM3, SM5.4, and SM5.2R calculations were carried out with AMSOL.<sup>50</sup> MM2 calculations were carried out with CHEM3D-PRO.<sup>51</sup> MM3 calculations were carried out with MACROMODEL.<sup>52</sup> SM5.0R calculations were carried out with OMNISOL.<sup>53</sup>

**Results.** Standard-state free energies of self-solvation were calculated by eqs 5–8 and converted to vapor pressures by eq 4. The results are given in Table 1. The calculated results are compared to experimental vapor pressures taken from the following standard sources in the following priority order: (1) *Handbook of Chemistry and Physics*,<sup>46</sup> (2) *Handbook of Vapor Pressure*,<sup>54</sup> and (3) *Physical and Thermodynamic Properties of Pure Chemicals: Data Compilation*.<sup>45</sup> Data from any of these sources were rejected if the measurement was marked with an error estimate greater than 4%.

#### 4. Discussion

Since the vapor pressures span several orders of magnitude, it is more convenient to look at the errors in terms of free energies of self-solvation. Experimental values are computed from the vapor pressures by eq 3 and compared directly to the calculations. The mean unsigned errors (i.e., the mean absolute values of the deviation of theory from experiment) are given in Table 2. This table shows that the mean unsigned errors in the free energies of self-solvation range from 0.7 kcal/mol to 0.8 kcal/mol for the various models. This range corresponds to mean unsigned deviations of 1.2 to 1.5 in the natural logarithm of the vapor pressure, and this in turn corresponds to typical errors of factors of 3 to 5 in the vapor pressures themselves.

A few of the functional groups have particularly large errors for some or all of the models. We especially notice the large errors for carboxylic acids, nitrohydrocarbons, and fluorinated hydrocarbons. There may be specific reasons for some of these errors. For example, carboxylic acids dimerize to a large extent in the liquid, and this violates our ideality assumption. Nitro-

**TABLE 3: Breakdown of the Free Energy of Self-Solvation (kcal/mol) for Primary Alcohols for SM5.4/AM1 Calculations**

compd	<i>n</i>	$\alpha$	$\beta$	$\gamma^a$	$\epsilon$	$\Delta G_S^\circ$ (exp)	$\Delta G_{\text{ENP}}$	$G_{\text{CDS}}$	$G_{\text{CDS}}^{[2]}$	$\Delta G_S^\circ$ (calc)
methanol	1.3288	0.43	0.47	31.8	32.61	-4.8	-4.4	0.1	13.4	-4.3
ethanol	1.3611	0.37	0.48	31.6	24.85	-5.0	-4.3	-0.8	15.6	-5.1
1-propanol	1.3850	0.37	0.48	33.6	20.52	-5.5	-4.0	-1.5	17.6	-5.5
1-butanol	1.3993	0.37	0.48	35.9	17.33	-6.0	-3.5	-2.3	19.4	-5.7
1-pentanol	1.4101	0.37	0.48	36.5	15.13	-6.6	-3.4	-3.0	21.2	-6.4
1-hexanol	1.4178	0.37	0.48	37.2	12.51	-7.1	-3.3	-3.8	23.0	-7.1
1-heptanol	1.4249	0.37	0.48	38.5	11.32	-7.8	-3.2	-4.5	24.9	-7.8
1-octanol	1.4295	0.37	0.48	39.0	9.86	-8.4	-3.1	-5.2	26.7	-8.3
1-nonanol	1.4333	0.37	0.48	40.1	8.60	-9.0	-3.5	-5.8	28.6	-9.3
1-decanol	1.4372	0.37	0.48	41.0	7.53	-9.6	-3.4	-6.6	30.4	-9.9

<sup>a</sup> Units of cal mol<sup>-1</sup> Å<sup>-1</sup>.

hydrocarbons, in contrast, do not have optimized parameters in MM2 (CHEM3D-PRO gives a warning) and they cannot be calculated with MM3. The large errors for nitro compounds may be caused in part by bad AM1 geometries, not by the solvation models. Thus these two classes of compounds do not provide valid tests of the models, and we have also computed mean unsigned errors for the 145-molecule subset in which these two classes of molecules are excluded. This lowers the mean unsigned errors of the various models by 0.02 to 0.10 kcal/mol, depending on the model (see Table 2). The resulting mean unsigned errors of 0.62–0.79 kcal/mol are probably a more fair assessment of the accuracy of the present approach.

There are also relatively large errors in several individual molecules. The modeling of carbon disulfide is difficult since this molecule has unique functionality that was not parameterized as a solute, and CS<sub>2</sub> exhibits a high error. Another molecule that is distinct is dimethyl sulfoxide. While there is a relatively high error in all of the models for this molecule, the error in SM5.2R/PM3/MM2 is exceptionally high, about 10 kcal/mol. This error may be explained by the fact that the PM3 model tends to give unphysical charges for hypervalent atoms, and this adversely affects the computed free energies in the liquid phase.<sup>55</sup>

The mean error in each of the models for fluorinated compounds is due to the unusually high error for perfluorobenzene. Fluorobenzene and other halo compounds showed errors that were more acceptable. Perfluoro compounds exhibit low refractive indices, low surface tensions, and low dielectric constants. While these solvent descriptors are within the range that the models were parameterized for, they are all at the low extreme. The parameterization did not include the solvation of a perfluorous solute in perfluorobenzene. Partitioning data for solutes in perfluorobenzene that are contained within the MedChem database<sup>56</sup> are limited to ketone and ester functional groups. This lack of partitioning data makes it difficult to determine the parameterization of the model for a solvent exhibiting these descriptors.

The SM5.0R model exhibits a high degree of accuracy, especially when we allow for the simplicity of the model. Since this model has implicit electrostatics and depends only on the atomic surface tensions, the subtle changes between the MM3 and MM2 geometries generally changed the free energy of self-solvation by less than 0.1 kcal/mol. However, in the SM5.2R models, partial atomic charges are used, and, when combined with the calculation of the surface tensions, a small difference in the geometry of a molecule can lead to a large change in the free energy of solvation. In cases where the geometry is well known, such as when high-level ab initio structures are used, the SM5.2R model should provide a better model.

**TABLE 4: Breakdown of  $G_{\text{CDS}}$  (kcal/mol) for Primary Alcohols for SM5.4/AM1 Calculations<sup>a</sup>**

compd	<i>n</i>	$\alpha$	$\beta$	$\gamma$	total
methanol	-0.1	-0.3	-0.4	1.0	0.1
ethanol	-1.2	-0.2	-0.5	1.1	-0.8
1-propanol	-2.1	-0.2	-0.5	1.3	-1.5
1-butanol	-3.1	-0.3	-0.4	1.5	-2.3
1-pentanol	-4.0	-0.3	-0.4	1.6	-3.0
1-hexanol	-5.0	-0.3	-0.3	1.8	-3.8
1-heptanol	-5.9	-0.3	-0.3	2.0	-4.5
1-octanol	-6.7	-0.3	-0.4	2.1	-5.2
1-nonanol	-7.5	-0.2	-0.5	2.4	-5.8
1-decanol	-8.4	-0.2	-0.5	2.6	-6.6

<sup>a</sup> The values tabulated are the CDS contributions in kcal/mol from the terms proportional to *n*,  $\alpha$ ,  $\beta$ , and  $\gamma$ .

Although the mean unsigned errors in the free energies of self-solvation are larger on average than the errors in the free energies of dilute-solution solvation that were considered previously,<sup>20–24,26,28</sup> they are only slightly larger. This is very encouraging for two reasons: (1) *No* vapor pressure data were used in parameterizing the models, and (2) the rigid models (SM5.2R and SM5.0R) are applied here with very inexpensive estimates of geometries that are *not* the geometries that were used for parameterization of the models.

A valuable feature of the new models is that they provide insight into the sources of the variations in the vapor pressure among a series of compounds. Consider, for example, the SM5.4 model for general organic solvents. If eq 8 is specialized for the organic SM5.4 model,  $G_{\text{CDS}}$  is a sum of terms  $G_{\text{CDS}}^{[1]}$  and  $G_{\text{CDS}}^{[2]}$  involving, respectively, the smaller effective solvent radius,  $r_1 = 1.4$  Å and the larger effective solvent radius  $r_2 = 2.8$  Å. In particular

$$G_{\text{CDS}} = G_{\text{CDS}}^{[1]} + G_{\text{CDS}}^{[2]} \quad (9)$$

where

$$G_{\text{CDS}}^{[1]} = \sum_j \sum_{\delta} \sum_k A_k(\mathbf{R}, r_1) \sigma_{Z_k}^{[1]} f_j(\{Z_k, R_{k''}\}) S_{\delta} \quad (10)$$

and

$$G_{\text{CDS}}^{[2]} = \sum_{\delta} \sigma_{\delta}^{[2]} S_{\delta} \sum_k A_k(\mathbf{R}, r_2) \quad (11)$$

The variables are the same as those defined in section 2.2. In the sums,  $\delta$  runs over *n*,  $\alpha$ , and  $\beta$  for  $G_{\text{CDS}}^{[1]}$  and over *n* and  $\gamma$  for  $G_{\text{CDS}}^{[2]}$ . The major difference between the terms  $G_{\text{CDS}}^{[1]}$  and  $G_{\text{CDS}}^{[2]}$  is that the first term is a function of the atomic numbers of the atoms and the solvent-accessible surface areas of the individual atoms, whereas the second term is a function of the

**TABLE 5: Breakdown of the Free Energy of Self-Solvation (kcal/mol) for Primary Aldehydes for SM5.4/AM1 Calculations**

compd	<i>n</i>	$\alpha$	$\beta$	$\gamma^a$	$\epsilon$	$\Delta G_s^\circ$ (exp)	$\Delta G_{\text{ENP}}$	$G_{\text{CDS}}$	$G_{\text{CDS}}^{[2]}$	$\Delta G_s^\circ$ (calc)
propanal	1.3636	0.00	0.45	32.5	18.50	-3.9	-4.2	-0.4	16.6	-4.6
butanal	1.3843	0.00	0.45	35.1	13.45	-4.4	-3.9	-1.0	18.6	-4.9
pentanal	1.3944	0.00	0.45	36.6	10.00	-5.0	-3.6	-1.7	20.5	-5.3
benzaldehyde	1.5463	0.00	0.39	54.7	18.22	-7.0	-5.7	-1.9	23.1	-7.5

<sup>a</sup> Units of cal mol<sup>-1</sup> Å<sup>-1</sup>.**TABLE 6: Breakdown of  $G_{\text{CDS}}$  (kcal/mol) for Aldehydes for SM5.4/AM1 Calculations<sup>a</sup>**

compd	<i>n</i>	$\alpha$	$\beta$	$\gamma$	total
propanal	-1.8	0.0	0.2	1.2	-0.4
butanal	-2.7	0.0	0.2	1.4	-1.0
pentanal	-3.5	0.0	0.2	1.6	-1.7
benzaldehyde	-4.7	0.0	0.5	2.3	-1.9

<sup>a</sup> The values tabulated are the CDS contributions in kcal/mol from the terms proportional to *n*,  $\alpha$ ,  $\beta$ , and  $\gamma$ .

entire accessible surface area of the molecule. In interpreting the individual terms one should keep in mind that (i) hydrogen bonding is included both in the electrostatic term and also the  $\alpha$  and  $\beta$  terms, (ii) the terms in CDS may to some extent make up for systematic deficiencies in other contributions, and (iii) the dispersion interactions are not completely separable from other effects since *n* terms in  $G_{\text{CDS}}^{[2]}$  are associated with structure breaking as well as dispersion. As a consequence of these complications, the interpretation of individual terms is not as definitive as using their sum. With these caveats, the individual contributions are still very interesting. For example, if the total of all terms proportional to *n* becomes more negative, it probably indicates that dispersion is becoming more net favorable, although the quantitative value of the sum would be expected to be smaller in magnitude than the total solute-solvent dispersion interaction. With these considerations as background, we turn to some examples of the individual contribution, which are illustrated by the data in Tables 3–8.

Consider, for example, Tables 3 and 4 for alcohols. If the series of primary alcohols from ethanol to decanol is considered, the hydrogen bonding acidity and basicity parameters remain the same, 0.37 and 0.48, respectively, so in the SM5.4/AM1 model the change in free energy of self-solvation is a result of three molecular properties, namely the dielectric constant, the macroscopic surface tension, and the index of refraction. The dielectric constant ranges from 24.85 for ethanol to 7.53 for decanol. This change is reflected in the less negative electrostatic free energy of self-solvation,  $\Delta G_{\text{ENP}}$ , which ranges from -4.3 kcal/mol to -3.1 kcal/mol. This, however, is in distinct contrast to the experimental change in the free energy of self-solvation of -5.0 kcal/mol to -9.6 kcal/mol. This change is accounted for in the cavitation-dispersion-solvent-structure term,  $G_{\text{CDS}}$ . There is an increase in the  $G_{\text{CDS}}^{[2]}$  contribution which may arise mostly from the increase in the structural reorganization energy needed to form the cavity, but which also serves as part of the estimate of the dispersion term. The  $G_{\text{CDS}}^{[1]}$  term has contribu-

tions from the acidity, basicity, and index of refraction. Since  $\alpha$  and  $\beta$  are constants, the overall trend in this term is dominated by the increase in favorable dispersion interactions, reflected in the change in the index of refraction. Table 4 shows that the contributions from terms proportional to the index of refraction decrease from -1.2 to -8.4 kcal/mol. The  $\alpha$  and  $\beta$  components are nearly uniform, with less than a 0.2 kcal/mol change over the ethanol-to-decanol series. The values change slightly since the terms depend on the solvent-accessible surface areas of the atoms. Methanol has different values for  $\alpha$  and  $\beta$  but is included for completeness.

Next consider the aldehydes in Tables 5 and 6. The calculations in this case reproduce all qualitative trends in the data, although they underestimate the extent of the variation. For the aldehydes, we must consider  $\epsilon$ , *n*,  $\beta$ , and  $\gamma$ . For propanal through pentanal, the electrostatic term (controlled by partial charges and  $\epsilon$ ) varies in the opposite direction to the total, just as it did for primary alcohols. Just as for the alcohols, the variation of the first-solvation-shell terms with *n* dominates those due to  $\gamma$ , and the dispersion trend primarily explains the result.

Finally, consider the series of nitriles, where all five of the parameters vary upon going from acetonitrile to butanonitrile. The breakdowns are shown in Tables 7 and 8. Here again, the decrease in the dielectric constant is reflected in the more positive  $\Delta G_{\text{ENP}}$ . The  $G_{\text{CDS}}$  term is slightly more complicated, with an increase (less negative contribution) due to the  $\alpha$ ,  $\beta$ , and  $\gamma$  contributions, but a marked decrease (negative value getting larger in magnitude) in the contributions from *n*, leading to an overall decrease in  $G_{\text{CDS}}$ . If we again associate *n* contributions with dispersion, we see that dispersion is again the dominant effect.

Although some of the other methods that are available offer greater accuracy, this analysis has shown that the present approach offers the ability to decipher the causes of the changes in the physical properties.

## 5. Concluding Remarks

The SM5 series of solvation models predicts free energies of self-solvation to better than 1 kcal/mol and thus can be used to predict vapor pressures at 298 K with useful accuracy. The size of the mean unsigned errors in the vapor pressure would preclude the usefulness in some applications, but the generality of this model allows its application to molecules or series of molecules where other methods fall short. For instance, group-additive methods are not applicable when parameters for a specific functional group are unavailable, or they may be

**TABLE 7: Breakdown of the Free Energy of Self-Solvation (kcal/mol) for Nitriles for SM5.4/AM1 Calculations**

compd	<i>n</i>	$\alpha$	$\beta$	$\gamma^a$	$\epsilon$	$\Delta G_s^\circ$ (exp)	$\Delta G_{\text{ENP}}$	$G_{\text{CDS}}$	$G_{\text{CDS}}^{[2]}$	$\Delta G_s^\circ$ (calc)
acetonitrile	1.3442	0.07	0.32	41.3	35.69	-4.8	-5.9	1.2	14.9	-4.7
propanonitrile	1.3655	0.02	0.36	38.5	29.32	-5.1	-5.1	0.4	16.9	-4.7
butanonitrile	1.3842	0.00	0.36	38.8	24.29	-5.5	-4.7	-0.3	18.7	-5.0
benzonitrile	1.5289	0.00	0.33	55.8	5.59	-7.3	-4.9	-1.4	23.0	-6.4

<sup>a</sup> Units of cal mol<sup>-1</sup> Å<sup>-1</sup>.



**TABLE 8: Breakdown of  $G_{\text{CDS}}$  (kcal/mol) for Nitriles for SM5.4/AM1 Calculations<sup>a</sup>**

compd	$n$	$\alpha$	$\beta$	$\gamma$	total
acetonitrile	-1.1	-0.1	1.1	1.3	1.2
propanonitrile	-2.1	0.0	1.2	1.4	0.4
butanonitrile	-3.0	0.0	1.2	1.5	-0.3
benzonitrile	-5.1	0.0	1.3	2.4	-1.4

<sup>a</sup> The values tabulated are the CDS contributions in kcal/mol from the terms proportional to  $n$ ,  $\alpha$ ,  $\beta$ , and  $\gamma$ .

inaccurate because of nonadditive interactions between functional groups. In addition, the SM5 models, being founded on physical principles, permit analysis of the factors (e.g., electrostatics, dispersion, hydrogen bonding, etc.) contributing to the differences in vapor pressures of related compounds. Finally, these methods are inexpensive (SM5.0R is very inexpensive), and they are available in well documented, easy-to-use computer codes,<sup>50,53</sup> so they may be used for quick surveys or scans of large data sets. In this regard we note that the only input data required for any of the methods are the molecular formula and geometry, index of refraction, surface tension, Abraham's acidity and basicity parameters, and, in the quantum mechanical models, the dielectric constant. These latter quantities are widely tabulated in the literature, and if a value is not readily available it can probably be estimated from the values for similar compounds.

The vapor pressure data also provide an opportunity to test the accuracy of the SM5 family of models on cases that are not in the parameterization set. The errors in the free energies of solvation for each of the models are not unreasonably greater than those seen in the parameterization set.

**Acknowledgment.** This work was supported in part by the National Science Foundation and the Alfred P. Sloan Foundation.

## References and Notes

- (1) Fredenslund, A.; Gmehling, J.; Rasmussen, P. *Vapor-Liquid Equilibria Using UNIFAC*; Elsevier: New York, 1977.
- (2) Mackay, D. M.; Bobra, A.; Chan, D. W.; Shiu, W. Y. *Environ. Sci. Technol.* **1982**, 16, 645.
- (3) McGarry, J. *Ind. Eng. Chem. Process Des. Dev.* **1983**, 22, 313.
- (4) Lyman, J. L.; Reehl, W. F.; Rosenblatt, D. H. *Handbook of Chemical Property Estimation Methods*; American Chemical Society: Washington, DC, 1990; chapter 14.
- (5) Banerjee, S.; Howard, P. H.; Lande, S. S. *Chemosphere* **1990**, 21, 1173.
- (6) Mishra, D. S.; Yalkowsky, S. H. *Ind. Eng. Chem. Rev.* **1991**, 30, 1609.
- (7) Haranurda, S. *Chem. Eng.* **1996**, 135.
- (8) Myrdal, P. B.; Yalkowsky, S. H. *Ind. Eng. Chem. Res.* **1997**, 36, 2494.
- (9) Potukuchi, S.; Wexler, A. S. *Atmos. Environ.* **1997**, 31, 741.
- (10) Tu, C. *Fluid Phase Equilibria* **1994**, 99, 105.
- (11) Simmons, K. A. *J. Agric. Food Chem.* **1999**, 47, 1711.
- (12) Hilal, S. H.; Carreira, L. A.; Karickhoff, S. W.; Melton, C. M. *J. Chromatogr. A* **1994**, 662, 269.
- (13) Basak, S. C.; Gute, B. D.; Grunwald, G. D. *J. Chem. Inf. Comput. Sci.* **1997**, 37, 651.
- (14) Katritzky, A. R.; Wang, Y.; Sild, S.; Tamm, T.; Karelson, M. *J. Chem. Inf. Comput. Sci.* **1998**, 38, 720.
- (15) Liang, C.; Gallagher, D. A. *J. Chem. Inf. Comput. Sci.* **1998**, 38, 321.
- (16) Lago, S.; Garzon, B.; Calero, S.; Vega, C. *J. Phys. Chem. B* **1997**, 101, 6763.
- (17) Langlet, J.; Claverie, P.; Caillet, J.; Pullman, A. *J. Phys. Chem.* **1988**, 92, 1617.
- (18) Rosseinsky, D. R.; Stead, K.; Mowforth, C. W. *Mol. Phys.* **1998**, 95, 409.
- (19) Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. *J. Phys. Chem. A* **1998**, 102, 5074.
- (20) Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, 100, 16385.
- (21) Giesen, D. J.; Gu, M. Z.; Cramer, C. J.; Truhlar, D. G. *J. Org. Chem.* **1996**, 61, 8720.
- (22) Giesen, D. J.; Chambers, C. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, 101, 2061.
- (23) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, 101, 7147.
- (24) Giesen, D. J.; Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1997**, 98, 85.
- (25) Giesen, D. J.; Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. In *Computational Thermochemistry*; Irikura, K. K., Frurip, D. J., Eds.; ACS Symposium Series 677; American Chemical Society: Washington, DC, 1998; pp 285–300.
- (26) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1998**, 102, 3257.
- (27) Chambers, C. C.; Giesen, D. J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G.; Vaes, W. H. J. In *Rational Drug Design*; Truhlar, D. G., Howe, W. J., Hopfinger, A. J., Blaney, J., Dammkoehler, R. A., Eds.; IMA Volumes in Mathematics and Its Applications 108; Springer: New York, 1999; pp 51–72.
- (28) Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *J. Org. Chem.* **1998**, 63, 4305.
- (29) (a) Zhu, T.; Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **1998**, 109, 9117. (b) Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1998**, 288, 293.
- (30) (a) Hawkins, G. D.; Li, J.; Zhu, T.; Chambers, C. C.; Giesen, D. J.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. In *Rational Drug Design*; Parrill, A. L., Reddy, M. R., Eds.; ACS Symposium Series 719; American Chemical Society: Washington, DC; pp 121–140. (b) Hawkins, G. D.; Zhu, T.; Li, J.; Chambers, C. C.; Giesen, D. J.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. In *Combined Quantum Mechanical and Molecular Mechanics Methods*; Gao, J., Thompson, M. A., Eds.; ACS Symposium Series 712; American Chemical Society: Washington, DC, 1998; pp 201–219.
- (31) Ben-Naim, A. *Statistical Thermodynamics for Chemists and Biochemists*; Plenum: New York, 1992.
- (32) Cramer, C. J.; Truhlar, D. G. *Rev. Comput. Chem.* **1995**, 6, 1.
- (33) Abraham, M. H.; Grellier, P. L.; Prior, D. V.; Duce, P. P.; Morris, J. J.; Taylor, P. *J. Chem. Soc., Perkin Trans. 2* **1989**, 699.
- (34) Abraham, M. H. *Chem. Soc. Rev.* **1993**, 22, 73.
- (35) Abraham, M. H. *J. Phys. Org. Chem.* **1993**, 6, 660.
- (36) Abraham, M. H. In *Quantitative Treatments of Solute/Solvent Interactions*; Politzer, P., Murray, J. S., Eds.; Theoretical and Computational Chemistry 1; Elsevier: Amsterdam, 1994; pp 83–134.
- (37) Abraham, M. H.; Weatherby, P. K. *J. Pharm. Sci.* **1994**, 83, 1087.
- (38) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1995**, 9, 87.
- (39) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, 107, 3932.
- (40) Dewar, M. J. S.; Zoebisch, E. G. *J. Mol. Struct. (Theochem)* **1988**, 180, 1.
- (41) Dewar, M. J. S.; Yate-Ching, Y. *Inorg. Chem.* **1990**, 29, 3881.
- (42) Stewart, J. J. P. *J. Comput. Chem.* **1989**, 10, 221.
- (43) *Beilstein: Handbook of Organic Chemistry*, 4th ed.; Luckenback, R., Ed. Springer-Verlag: New York 1984. (b) Lide, D. R. *Handbook of organic solvents*; CRC Press: Boca Raton, FL, 1995. (c) *Handbook of Analytical Chemistry*, 1st ed.; Meites, L., Ed., McGraw-Hill: New York, 1963. (d) *American Institute of Physics Handbook*; 3rd ed.; Gray, D. E., Ed.; McGraw-Hill: New York, 1972. (e) *Organic Solvents: Physical Properties and Methods of Purification*, 3rd ed.; Riddick, J., Bunger, W. B., Eds.; Wiley-Interscience: New York, 1970.
- (44) Jasper, J. J. *J. Phys. Chem. Ref. Data* **1972**, 1, 841.
- (45) Daubert, T. E.; Danner, R. P. *Physical and Thermodynamic Properties of Pure Chemicals: Data Compilation*; Hemisphere Publishing Corp.: New York, 1989.
- (46) Lide, D. R., Ed.; *CRC Handbook of Chemistry and Physics*, 76th ed.; CRC Press: Boca Raton, FL, 1996.
- (47) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, 99, 8127.
- (48) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, 111, 8551.
- (49) Allinger, N. L.; Rahman, M.; Lii, J. H. *J. Am. Chem. Soc.* **1990**, 112, 8293.
- (50) Hawkins, G. D.; Giesen, D. J.; Lynch, G. C.; Chambers, C. C.; Rossi, I.; Storer, J. W.; Li, J.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. AMSOL version 6.5.3; University of Minnesota: Minneapolis, MN, 1998.
- (51) CHEM3D-PRO version 3.5.1; Cambridge Scientific Computing; Cambridge, MA.
- (52) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Hendrickson, T.; Still, W. C.; MACROMODEL version 6.0; Columbia University, New York, 1990.



(53) Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. OMNISOL version 1.0; University of Minnesota: Minneapolis, MN, 1997.

(54) Yaws, C. L. *Handbook of Vapor Pressure*; Gulf Publishing Co.: Houston, TX, 1994.

(55) Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1992**, 6, 625.

(56) Leo, A. J. *Pomona Medchem Database*; Biobyte Corporation: Claremont, CA, 1994.