2178 *J. Phys. Chem. A* **2000,** *104,* 2178−2182

A Universal Solvation Model Based on Class IV Charges and the Intermediate Neglect of Differential Overlap for the Spectroscopy Molecular Orbital Method

Jiabo Li, Tianhai Zhu, Christopher J. Cramer,* and Donald G. Truhlar*

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

Received: June 7, 1999; In Final Form: September 20, 1999

The SM5.42R solvation model, which is based on SM5-type atomic surface tensions, class IV point charges based on charge model 2, and rigid geometries, is parametrized for the intermediate neglect of differential overlap for spectroscopy (INDO/S) method, both in the original (INDO/S) and more recent (INDO/S2) versions. The parametrization is based on 2184 free energies of solvation of 275 neutral solutes and 49 ions in water and 90 organic solvents.

1. Introduction

Solvation effects play an important role in determining chemical structure and reactivity, and there has been much recent progress in modeling these effects, especially by quantum mechanical methods.[1] Our group has developed a sequence of solvation models, called SM*x* models, based on combining a quantum mechanical self-consistent reaction field calculation with an estimate of residual first-solvation-shell effects based on empirical atomic surface tensions.[2−6] This approach has subsequently been adopted by other groups as well.[7] Our own work has culminated in the SM5 suite of solvation models[3] and, most recently, in the SM5.42R model.[4−6] This model uses SM5 functional forms for surface tensions[3] and class IV partial atomic charges[8] based on the CM2 charge model,[9,10] and it is parametrized using rigid geometries. It has been parametrized for equilibrium free energies of solvation calculated by neglect of diatomic differential overlap (NDDO) semiempirical molecular orbital theory,[6] ab initio Hartree−Fock calculations,[4,6] density functional theory,[5,6] and hybrid Hartree−Fock density-functional theory.[6]

Solvation effects are also very important for electronic spectroscopy.[1a,c,11] In the present paper we extend the SM5.42R parametrizations to INDO/S[12] and INDO/S2[10] versions of intermediate neglect of differential overlap for spectroscopy molecular orbital theory. The INDO/S method is the most widely used quantum mechanical modeling method for the spectroscopy of large molecules, and INDO/S2 is a recent reparametrization that gives better results for carbonyl compounds. It is therefore of interest to see whether these methods can be used to predict quantitative solvent effects, and the present paper will show that the answer is affirmative. This is encouraging for the future development of quantitative models of solvatochromic shifts of large molecules calculated by INDO/S or INDO/S2.

2. Theory

The INDO/S,[12] INDO/S2,[10] CM2,[9,10] and SM5.42R[4−6] models are all fully described elsewhere. INDO/S is a refinement of the intermediate neglect of differential overlap[13] molecular orbital method in which the parameters are chosen to reproduce electronic spectra.[12,14] INDO/S2 is a reparametrization designed to provide a more accurate treatment of carbonyl groups.[10] CM2 is a class IV charge model based on the linear mapping of charges obtained by population analysis.[9,10] The SM5.42R model approximates the standard-state free energy of solvation of solute X in solvent Y at 298 K as

$$\Delta G_S^o = \Delta G_{EP} + G_{CDS} \qquad (1)$$

where ΔG_{EP} is the change in solute electronic energy and solvent electric polarization free energy calculated by a self-consistent reaction-field calculation based on CM2 partial atomic charges, and G_{CDS} accounts for residual solvation effects in the first solvation shell, especially cavitation, dispersion, and solvent structural effects (such as hydrogen bonding and hydrophobic effects). The G_{CDS} term is a sum over atomic contributions, each of which is an atomic surface tension times the surface area that the given atom exposes to the solvent. The atomic surface tensions are themselves functions of solute geometry and solvent properties. These functions contain linear parameters called atomic surface tension coefficients that are determined by regression against experimental data. For the present work, the surface tension functionals and all parameters therein are identical to those in the SM5.2R model[3e] and in all previous SM5.42R parametrizations;[4−6] these functional forms represent refinements of those used in the SM5.4 model[3a,b,d] and the SM5.0R model.[15] One regression is carried out for water and another for organic solvents. Because the model is general enough to treat water and virtually any organic solvent, it is called a universal solvation model.

Equation 1 is valid for a choice of standard states that involves the same number density in the gas phase and in the liquid solution. In particular, we take the gas-phase standard state as an ideal gas at a partial pressure corresponding to a concentration of 1 M (not a pressure of 1 atm), and we take the standard state for the liquid solution as an ideal solute at a concentration of 1 M.

We previously[4−6] presented two schemes for SCF calculations employing CM2 charges in reaction fields. In the INDO/S and INDO/S2 models, we have coded both schemes. They generally give similar results, and the calculations in this paper are based on scheme II.

---

* Corresponding authors.

10.1021/jp991858h CCC: $19.00

TABLE 1: Surface Tension Coefficients (cal mol^{-1} Å^{-2}) for SM5.42R/INDO/S2

atom (or atom pair)	$\sigma^{(n)}$	$\sigma^{(\alpha)}$	$\sigma^{(\beta)}$	$\sigma^{(\text{water})}$
H	39.65			98.67
C	36.29	24.38	−34.40	58.78
N	23.28	−59.00	107.38	−30.23
O	−39.64	−31.92		−188.50
F	3.28			44.28
P	−23.93			−66.25
S	−73.52	−65.18	61.79	−64.11
Cl	−31.32			0.92
Br	−41.77			−12.40
I	−47.76			−22.88
H, C	−92.07			−119.10
H, N	−75.48		−154.92	−203.73
H, N (2)	−212.99			−277.90
H, O	−8.53	−299.34	−346.42	−238.41
H, O (2)	176.81			384.69
H, S	69.02			30.49
C, C	−64.33			−68.29
C, C (2)	12.57			2.73
O, C	89.46		54.71	284.94
O, O	−34.68	150.42	−28.50	44.76
C, N	−68.30	38.98		19.86
N, C	−14.04	−56.95	2.02	−72.14
N, C (2)	78.72			149.20
O, N	174.03	99.46	174.49	573.38
P, O	134.01			309.82
P, S	233.42			396.12
S, S	−2.40			−3.89

TABLE 2: Surface Tension Coefficients (cal mol^{-1} Å^{-2}) for SM5.42R/INDO/S

atom (or atom pair)	$\sigma^{(n)}$	$\sigma^{(\alpha)}$	$\sigma^{(\beta)}$	$\sigma^{(\text{water})}$
H	39.93			98.71
C	37.31	24.50	−21.48	59.17
N	26.28	−66.08	108.38	−27.27
O	−43.18	−32.36		−196.03
F	4.42			45.55
P	1.74			−87.07
S	−73.68	−69.26	60.74	−65.58
Cl	−31.49			0.35
Br	−41.77			−12.44
I	−47.68			−22.91
H, C	−92.42			−119.26
H, N	−86.41		−150.19	−227.44
H, N (2)	−207.41			−265.20
H, O	−5.20	−291.72	−333.36	−231.38
H, O (2)	174.64			377.94
H, S	69.76			34.76
C, C	−64.76			−68.54
C, C (2)	11.98			2.45
O, C	90.97		54.71	289.05
O, O	−38.00	140.69	−34.40	34.09
C, N	−71.06	50.13		19.88
N, C	−13.81	−58.16	1.86	−72.09
N, C (2)	74.72			142.38
O, N	160.19	78.21	174.49	532.94
P, O	125.47			306.76
P, S	249.71			433.38
S, S	−2.21			−2.04

TABLE 3: Molecular Surface Tension Coefficients (cal mol^{-1} Å^{-2})

parametrization	$\sigma^{(\gamma)}$	$\sigma^{(\beta 2)}$	$\sigma^{(\phi 2)}$	$\sigma^{(\psi 2)}$
INDO/S2	0.3118	8.42	−3.75	−8.59
INDO/S	0.3045	8.81	−3.65	−8.77

TABLE 4: Performance of the SM5.42R/INDO/S2 Model by Solute Functional Class for Solutes Composed of H, C, N, O, F, S, Cl, Br, and I

	no. of solvent				
solute class	solutes^a	classes^b	data	MSE^c	MUE^c
unbranched alkanes	9	19	84	0.25	0.43
branched alkanes	5	3	12	0.17	0.44
cycloalkanes	5	6	18	−0.06	0.42
alkenes	9	4	27	−0.15	0.29
alkynes	5	3	14	0.07	0.11
arenes	9	19	134	−0.27	0.42
alcohols	17	19	385	0.17	0.45
ethers	12	19	93	0.29	0.46
aldehydes	7	8	38	0.05	0.54
ketones	12	18	203	−0.09	0.50
carboxylic acids	5	14	124	−0.01	0.45
esters	14	8	249	0.23	0.47
CH and CHO bifunctional solutes	5	8	28	0.47	0.99
H_2O, H_2	2	9	22	0.02	0.58
aliphatic amines	15	10	168	0.15	0.31
aromatic amines	11	12	81	0.20	0.49
nitriles	4	6	22	−1.44	1.44
nitrohydrocarbons	6	8	38	0.97	0.97
amides & ureas	4	6	11	0.73	1.35
HCN and HCNO bifunctional solutes	6	3	11	1.09	1.74
NH_3, N_2H_2	2	8	15	−0.15	0.58
thiols	4	5	14	0.30	0.37
sulfides	6	6	23	0.19	0.72
disulfides	2	3	5	−0.01	0.15
fluorinated hydrocarbons	9	5	19	−0.51	0.68
chloroalkanes	13	5	35	0.32	0.35
chloroalkenes	5	4	16	0.44	0.44
chloroarenes	8	6	37	−0.86	0.87
brominated hydrocarbons	14	6	50	−0.20	0.42
iodinated hydrocarbons	9	6	28	−0.03	0.40
other halo compounds	26	9	80	0.13	0.80
total	260	19	2084	0.08	0.51

[a] Number of solutes in this solute class. [b] Number of solvent classes for which there are data for this solute class. [c] Mean errors in kcal/mol.

TABLE 5: Performance of the SM5.42R/INDO/S2 Model for Neutral Solutes Containing Phosphorus

	no. of solvent				
solute class	solutes^a	classes^b	data	MSE^c	MUE^c
P, H, C and O compounds	5	6	28	−0.05	1.25
other P compounds	10	4	23	0.35	1.36
total	15	9	51	0.13	1.30

[a] Number of solutes in this solvent class. [b] Number of solvents classes for which data was included for this solute class. [c] Mean errors in kcal/mol.

3. Parametrization

The parametrization was carried out as for previous SM5.42R models.[4−6] All geometries were optimized by the ab initio Hartree−Fock method[16] with the MIDI![17,18] basis set. The aqueous parametrization is based on free energies of solvation for 275 neutral solutes and 49 ions in water.[3e,4−6] The organic parametrization is based on 1860 free energies of solvation for 240 solutes in 90 organic solvents.[3e,4−6] All nonlinear parameters, including atomic radii, were held constant at the same values that they have in the original SM5.42R parametrizations,[4,5] and the surface tension coefficients were regressed on the experimental free energies of solvation. The resulting surface tension coefficients are given in Tables 1−3, where the notation is the same as used previously.[3e,4−6]

4. Results

Errors are measured in terms of the mean signed error (MSE) and the mean unsigned (absolute) error (MUE) in the standard-

TABLE 6: Performance of the SM5.42R/INDO/S2 Model by Solvent Functional Class for Solutes Composed of H, C, N, O, F, S, Cl, Br, and I

	no. of solute				
solvent class	solvents[a]	classes[b]	data	MSE[c]	MUS[c]
alkanes	11	30	475	−0.02	0.35
cycloalkanes	2	24	106	−0.03	0.50
arenes	12	16	256	0.49	0.57
aliphatic alcohols	12	31	299	−0.09	0.56
aromatic alcohols	2	7	12	0.42	0.66
ketones	4	10	35	−0.13	0.46
esters	2	8	36	0.28	0.50
aliphatic ethers	4	19	99	−0.12	0.52
aromatic ethers	3	5	15	−0.21	0.38
amines	2	6	12	0.18	0.42
pyridines	3	5	15	0.00	0.36
nitriles	2	5	10	−0.41	0.44
nitrohydrocarbons	4	8	27	−0.11	0.73
tertiary amides	2	5	10	−0.03	0.24
haloaliphatics	12	27	269	0.37	0.69
haloaromatics	6	11	106	0.10	0.40
misc acidic	3	5	15	0.08	0.43
misc basic	4	12	39	−0.02	0.44
aqueous	1	31	248	−0.05	0.55
total	91	31	2084	0.08	0.51

[a] Number of solvents in this solvent class. [b] Number of solute classes for which data was included for this solvent class. [c] Mean errors in kcal/mol.

TABLE 7: Performance of the SM5.42R/INDO/S2 Model for Ions in Water

solute class	data	MSE	MUE
non-phosphorus cations	17	−0.3	4.8
non-phosphorus anions	26	1.8	3.1
all non-phosphorus ions	43	1.0	4.0
phosphorus-containing cations	4	−0.8	2.6
phosphorus-containing anions	2	−7.4	7.4
all phosphorus-containing ions	6	−3.0	4.2
all cations	21	−0.4	4.4
all anions	28	1.1	3.4
all ions	49	0.5	4.0

TABLE 8: Mean Unsigned Errors (kcal/mol)

			neutrals		ions	
model	parametrization	ref	non-P[a]	all[b]	non-P[c]	all[d]
SM5.0R	[no quantal component]	16	0.40	n.a.[e]	n.a.	n.a.
SM5.2R	AM1	3e	0.41	n.a.	3.7	n.a.
	PM3	3e	0.40	n.a.	3.6	n.a.
	MNDO	3e	0.39	n.a.	3.8	n.a.
	MNDO/d	3e	0.38	n.a.	3.6	n.a.
SM5.42R	BPW91/MIDI!6D	5	0.43	0.45	3.9	3.7
	BPW91/DZVP	5	0.43	0.44	3.6	3.6
	BPW91/6-31G*	5	0.41	0.43	3.9	3.9
SM5.42R	HF/MIDI!6D	4	0.43	0.45	3.9	3.9
SM5.42R	HF/MIDI!	6	0.42	0.44	3.8	3.8
	HF/6-31G*	6	0.41	0.43	3.7	3.7
	HF/6-31+G*	6	0.50	0.53	3.6	3.5
	HF/cc-pVDZ	6	0.43	0.45	3.8	3.9
	AM1	6	0.43	0.45	4.0	4.1
	PM3	6	0.48	0.46	3.9	4.0
	BPW91/MIDI!	6	0.43	0.45	3.9	3.8
	B3LYP/MIDI!	6	0.43	0.45	3.9	3.7
SM5.42R	INDO/S2	present	0.51	0.53	4.0	4.0
	INDO/S	present	0.56	0.58	4.0	3.9

[a] 2084 data. [b] 2135 data. [c] 43 data. [d] 49 data. [e] n.a. denotes not available.

TABLE 9: SM5.42R/INDO/S2 and Experimental Free Energies of Solvation (kcal/mol) for Ions in Aqueous Solution

	ΔG_{EN}	G_{CDS}	$\Delta G_{\mathrm{S}}^{\mathrm{o}}$	$\Delta G_{\mathrm{S}}^{\mathrm{o}}$ [a]
HC_2^-	−79.5	0.6	−78.9	−73
$CH_3OH_2^+$	−85.4	1.7	−83.6	−87
$(CH_3)_2OH^+$	−72.9	0.4	−72.6	−70
$CH_3CH_2OH_2^+$	−82.3	1.6	−80.6	−81
$CH_3C(OH)CH_3^+$	−68.1	−0.1	−68.2	−64
H_3O^+	−100.6	6.6	−94.0	−105
CH_3O^-	−88.9	0.0	−88.9	−98
$CH_3CO_2^-$	−76.1	1.1	−75.0	−77
$CH_3COCH_2^-$	−72.6	0.4	−72.1	−81
$C_6H_5O^-$	−65.8	−1.7	−67.5	−75
$C_6H_5CH_2^-$	−56.0	−0.9	−56.9	−59
OH^-	−106.3	−5.1	−111.3	−110
HO_2^-	−96.8	−6.7	−103.6	−101
O_2^-	−86.9	−6.7	−93.6	−87
$CH_3NH_3^+$	−77.3	−1.3	−78.6	−73
$HC(OH)NH_2^+$	−67.3	−2.8	−70.1	−78
CH_3CNH^+	−67.7	0.0	−67.7	−69
$CH_3C(OH)NH_2^+$	−67.1	−2.7	−69.9	−70
$(CH_3)_2NH_2^+$	−70.7	0.1	−70.7	−66
$(CH_3)_3NH^+$	−64.5	2.0	−62.5	−59
imidazoleH$^+$	−62.0	−1.9	−63.9	−64
pyridineH$^+$	−58.8	−1.6	−60.4	−58
$C_6H_5NH_3^+$	−69.2	−2.8	−72.1	−68
NH_4^+	−84.3	−1.8	−86.2	−81
CN^-	−78.0	1.3	−76.7	−75
CH_2CN^-	−67.0	0.0	−67.0	−75
NH_2^-	−87.0	−1.4	−88.4	−95
NO_2^-	−75.8	6.5	−69.3	−73
NO_3^-	−68.9	10.7	−58.2	−66
N_3^-	−66.0	−1.6	−67.6	−74
$CH_3SH_2^+$	−71.3	2.0	−69.4	−74
$(CH_3)_2SH^+$	−64.3	2.1	−62.2	−61
HS^-	−85.2	−1.2	−86.3	−76
CH_3S^-	−75.1	−0.9	−76.1	−76
$CH_3CH_2S^-$	−72.5	−1.0	−73.4	−74
n-$C_3H_7S^-$	−72.1	−0.8	−72.9	−76
$C_6H_5S^-$	−64.1	−3.1	−67.2	−65
F^-	−109.3	1.2	−108.1	−107
CHF_2COO^-	−68.7	2.2	−66.5	−70
Cl^-	−77.0	0.0	−76.9	−78
$CHCl_2COO^-$	−60.4	0.9	−59.4	−66
Br^-	−71.0	−0.5	−71.5	−72
I^-	−61.6	−1.1	−62.7	−63
PH_2^-	−72.2	−0.1	−72.3	−67
PH_4^+	−71.9	2.4	−69.5	−73
$CH_3PH_3^+$	−67.3	2.8	−64.4	−63
$(CH_3)_2PH_2^+$	−63.1	3.3	−59.8	−57
$(CH_3)_3PH^+$	−59.2	3.7	−55.6	−53
$H_2PO_4^-$	−79.7	2.3	−77.4	−68

[a] Last column: experimental values.

state free energies of solvation. Tables 4−6 summarize the performance of the model for neutral solutes when using the INDO/S2 Hamiltonian; Table 7 summarizes the results for ions. The performance of the model when using the INDO/S Hamiltonian is similar but slightly worse and is summarized in the Supporting Information.

5. Discussion

Since the results are very similar for the INDO/S2 and INDO/S Hamiltonians, we will concentrate the discussion on the results obtained with INDO/S2.

Tables 4 and 5 show good accuracy for all solute classes except those containing nitrogen and phosphorus. Table 6 shows good accuracy for all solvent classes, although there are systematic errors for aromatic solvents and for nitrile solvents. Table 7 shows reasonably good results for all ion classes except phosphorus-containing anions.

Table 8 compares the overall mean unsigned errors for all the parametrizations[3e,4−6,19] that use the present training set.

TABLE 10: SM5.42R/INDO/S2 Calculations of Selected Solutes in Different Solvents

		SM5.42R/INDO/S2			expt			SM5.42R/INDO/S2			expt
solute	solvent	ΔG_{EP}	G_{CDS}	ΔG_S^o	ΔG_S^o	solute	solvent	ΔG_{EP}	G_{CDS}	ΔG_S^o	ΔG_S^o
butanone	hexane	−2.03	−1.53	−3.56	−3.48	butanone	chloroform	−3.67	−1.18	−4.85	−5.43
butanone	heptane	−2.07	−1.47	−3.54	−3.36	butanone	carbon tetrachloride	−2.43	−1.25	−3.68	−4.09
butanone	octane	−2.11	−1.42	−3.52	−3.24	butanone	chlorohexane	−3.92	−1.21	−5.13	−4.10
butanone	nonane	−2.13	−1.37	−3.50	−3.20	butanone	bromoethane	−4.25	−1.32	−5.57	−5.13
butanone	decane	−2.17	−1.32	−3.49	−3.30	butanone	fluorobenzene	−3.84	−1.25	−5.09	−4.60
butanone	hexadecane	−2.25	−1.16	−3.42	−3.12	butanone	chlorobenzene	−3.88	−1.02	−4.90	−4.47
butanone	isooctane	−2.11	−1.54	−3.65	−3.40	butanone	bromobenzene	−3.83	−0.97	−4.80	−4.37
butanone	cyclohexane	−2.20	−1.29	−3.49	−3.48	butanone	iodobenzene	−3.63	−0.90	−4.53	−4.22
butanone	benzene	−2.47	−1.22	−3.69	−4.46	butanone	methylformamide	−4.89	−0.20	−5.09	−4.34
butanone	toluene	−2.58	−1.24	−3.81	−4.27	butanone	aceticacid	−3.97	−0.74	−4.71	−4.80
butanone	ethylbenzene	−2.62	−1.19	−3.81	−4.12	butanone	methoxyethanol	−4.57	−0.22	−4.78	−4.28
butanone	isopropylbenzene	−2.57	−1.23	−3.79	−4.02	butanone	THTPDO[a]	−4.78	0.82	−3.96	−4.09
butanone	xylene	−2.58	−1.19	−3.77	−4.23	butanone	dimethyl sulfoxide	−4.79	0.35	−4.44	−4.23
butanone	trimethylbenzene	−2.57	−1.16	−3.72	−3.97	butanone	carbon disulfide	−2.76	−1.29	−4.05	−3.85
butanone	mesitylene	−2.47	−1.24	−3.72	−3.95	butanone	water	−4.84	1.11	−3.73	−3.64
butanone	tetralin	−2.88	−1.02	−3.90	−3.12	fluorobenzene	hexane	−0.53	−4.22	−4.75	−4.15
butanone	ethanol	−4.68	−1.02	−5.69	−4.32	fluorobenzene	heptane	−0.54	−4.20	−4.74	−4.13
butanone	propanol	−4.62	−1.00	−5.63	−4.15	fluorobenzene	decane	−0.56	−4.09	−4.65	−3.48
butanone	2-propanol	−4.61	−1.04	−5.65	−4.07	fluorobenzene	hexadecane	−0.58	−3.97	−4.55	−4.03
butanone	butanol	−4.57	−0.94	−5.51	−4.12	fluorobenzene	cyclohexane	−0.57	−4.09	−4.67	−3.59
butanone	octanol	−4.31	−0.88	−5.19	−3.78	fluorobenzene	decalin	−0.62	−3.98	−4.60	−3.44
butanone	benzyl alcohol	−4.41	−0.61	−5.03	−4.57	fluorobenzene	octanol	−1.10	−3.80	−4.90	−3.87
butanone	*m*-cresol	−4.44	−0.68	−5.12	−5.98	fluorobenzene	chloroform	−0.94	−3.88	−4.82	−4.25
butanone	cyclohexanone	−4.53	−0.55	−5.08	−4.42	fluorobenzene	carbon tetrachloride	−0.63	−4.12	−4.74	−3.64
butanone	acetophenone	−4.57	−0.56	−5.13	−4.39	fluorobenzene	water	−1.23	−0.43	−1.66	−0.78
butanone	dibutyl ether	−3.05	−1.19	−4.24	−3.78	pyridine	hexane	−0.90	−4.06	−4.97	−3.81
butanone	tetrahydrofuran	−4.12	−0.98	−5.09	−4.54	pyridine	heptane	−0.92	−4.05	−4.97	−4.28
butanone	diisopropyl ether	−3.40	−1.43	−4.83	−3.96	pyridine	hexadecane	−1.00	−3.85	−4.85	−4.10
butanone	diethyl ether	−3.54	−1.43	−4.97	−4.09	pyridine	cyclohexane	−0.98	−3.96	−4.94	−4.30
butanone	anisole	−3.54	−0.85	−4.39	−4.43	pyridine	benzene	−1.10	−4.01	−5.11	−5.28
butanone	ethoxybenzene	−3.53	−0.95	−4.48	−4.28	pyridine	toluene	−1.14	−4.02	−5.17	−5.13
butanone	diphenyl ether	−3.37	−1.43	−4.79	−4.08	pyridine	xylene	−1.15	−3.98	−5.12	−5.12
butanone	triethylamine	−2.58	−0.92	−3.50	−3.86	pyridine	isobutanol	−2.01	−4.05	−6.06	−5.87
butanone	aniline	−4.05	−0.47	−4.53	−4.87	pyridine	octanol	−1.90	−3.95	−5.85	−5.34
butanone	pyridine	−4.45	−0.59	−5.05	−4.61	pyridine	4-methyl-2-pentanone	−1.96	−3.62	−5.59	−5.33
butanone	2-methylpyridine	−4.32	−0.70	−5.02	−4.52	pyridine	butyl acetate	−1.65	−3.57	−5.22	−5.31
butanone	dimethylpyridine	−4.09	−0.75	−4.84	−4.34	pyridine	dibutyl ether	−1.35	−3.72	−5.07	−4.65
butanone	acetonitrile	−4.76	−0.83	−5.58	−4.73	pyridine	diisopropyl ether	−1.50	−3.88	−5.39	−4.88
butanone	benzonitrile	−4.68	−1.07	−5.75	−4.58	pyridine	diethyl ether	−1.57	−3.84	−5.41	−4.81
butanone	nitromethane	−4.75	−0.50	−5.25	−4.72	pyridine	chloroform	−1.62	−4.08	−5.71	−6.45
butanone	nitroethane	−4.71	−0.74	−5.45	−4.73	pyridine	carbon tetrachloride	−1.08	−3.99	−5.07	−5.01
butanone	dimethylformamide	−4.76	−0.27	−5.03	−4.56	pyridine	dichloroethane	−1.91	−3.74	−5.64	−5.53
butanone	dimethylacetamide	−4.76	−0.34	−5.10	−4.52	pyridine	water	−2.13	−2.72	−4.85	−4.70

[a] Tetrahydrothiophene 1,1-dioxide.

Note that both the CM2 parameters and the surface tension parameters are individually optimized for each of the 14 SM5.42R parametrizations in Table 8, including INDO/S and INDO/S2. The table shows that we obtain almost as good performance with the INDO/S2 parametrization as with the wide variety of previous parametrizations that we have reported. This is particularly remarkable since the INDO/S and INDO/S2 Hamiltonians are much simpler than any of the other Hamiltonians in Table 8.

Although the ZINDO program has parameters for most of the periodic table, the present parametrization is limited to molecules composed of H, C, N, O, F, P, S, Cl, Br, and I. Extensions to the other elements provide a worthwhile challenge but will be difficult due to the sparsity of data. One possible procedure is the bootstrap method previously applied to parametrize MNDO/d for Zr.[20] Thus, the present parametrizations could form the starting point for extension to a wide variety of inorganic compounds, which could be very useful since INDO/S is already widely applied to such molecules.

Table 9 provides the breakdown into ΔG_{EP} and G_{CDS} for all ions in the training set. For 6 of the 49 ions, the INDO/S2 value of $|G_{CDS}|$ is bigger than 4.0 kcal/mol (the mean unsigned error), and for 19 cases (39%) it is greater than or equal to 2.0 kcal/mol. Thus, the nonelectrostatic contributions are not negligible for ions. To keep the results for ions in perspective, we note that we estimate the uncertainty in the experimental value of ΔG_S^o to be up to ∼5 kcal/mol for the typical ion. Thus, the accuracy obtained here is quite acceptable.

Table 10 provides a breakdown of ΔG_S^o into its components for three neutral solutes, butanone, fluorobenzene, and pyridine, in all the solvents for which they are represented in the training set. The table shows that, as the solvent is varied, the electrostatic component, ΔG_{EP}, varies by a factor of 2.4 for butanone and pyridine and a factor of 2.3 for fluorobenzene. The G_{CDS} term is usually negative, although butanone shows a positive G_{CDS} in tetrahydrothiophene 1,1-dioxide, dimethyl sulfoxide, and water. In all three of these cases inclusion of the solvophobic first-solvation-shell term improves the agreement with experiment. Whereas the electrostatic term dominates G_{CDS} for butanone, the opposite is true for the two aromatic solutes in Table 10. The solvent dependence of ΔG_S^o is particularly well reproduced for pyridine, even though it has a nitrogen atom. For fluorobenzene the calculated ΔG_S^o is systematically too negative by an average of 0.87 kcal/mol. Some of the experimental trends are hard to understand, e.g., decane solvent vs heptane or hexadecane, and may result from experimental error. In this regard, we note that we estimate an uncertainty in the experimental $\Delta G^{\mathrm{o}}_{\mathrm{S}}$ for a typical solute/solvent pair to be $\sim$0.2 kcal/mol.

6. Concluding Remarks

Despite its simplicity, the INDO/S Hamiltonian is shown to be capable of predicting reasonably accurate free energies of solvation when combined with the CM2 charge model and SM5.42R solvation model. We have explored its usefulness for solvatochromic shifts in a followup study.[21] These methods are now included in the ZINDO program.[22]

Acknowledgment. We are pleased to contribute this paper to the W. A. Goddard III 60th birthday issue of *The Journal of Physical Chemistry A*; we appreciatively acknowledge his pioneering and original contributions to electronic structure theory and his dedicated teaching. This work was supported in part by the National Science Foundation.

Supporting Information Available: Analogs of Tables 4−7 and 9 for the INDO/S Hamiltonian plus tabulations of the parameters used in the generalized Born part of the calculations and of the parameters that are different in the INDO/S and INDO/S2 Hamiltonians. This information is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) For reviews, see: (a) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027. (b) Cramer, C. J.; Truhlar, D. G. *Rev. Comput. Chem.* **1995**, *6*, 1. (c) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161. (d) The collection of articles by: Tomasi, J.; Lim, C.; Pratt, L. R.; Warshel, A.; Rivail, J.-L.; Bertrán, J.; Tucker, S. C.; Gao, J.; Cramer, C. J.; Truhlar, D. G.; and co-workers. In *Structure and Reactivity in Aqueous Solution*; ACS Symposium Series 568; Cramer, C. J., Truhlar, D. G., Eds.; American Chemical Society: Washington, DC, 1994.

(2) (a) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8305; **1991**, *113*, 9901(E). (b) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213. (c) Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **1992**, *13*, 1089. (d) Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Design* **1992**, *6*, 629. (e) Giesen, D. J.; Storer, J. W.; Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1995**, *117*, 1057. (f) Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1995**, *99*, 7137.

(3) (a) Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 16385. (b) Giesen, D. J.; Gu, M. Z.; Cramer, C. J.; Truhlar, D. G. *J. Org. Chem.* **1996**, *61*, 8720. (c) Giesen, D. J.; Chambers, C. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, *101*, 2061. (d) Giesen, D. J.; Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1997**, *98*, 8. (e) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1998**, *102*, 3257.

(4) Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *288*, 293.

(5) Zhu, T.; Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **1998**, *109*, 9117.

(6) Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.*, in press.

(7) (a) Bachs, M.; Luque, F. J.; Orozco, M. *J. Comput. Chem.* **1994**, *15*, 446. (b) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 11775.

(8) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Design* **1995**, *9*, 87.

(9) Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 1820.

(10) Li, J.; Williams, B.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **1999**, *110*, 724; Erratum **1999**, *111*, 5624.

(11) (a) Amos, A.; Burrows, B. L. *Adv. Quantum Chem.* **1973**, *7*, 289. (b) Nicol, M. F. *Appl. Spectrosc. Rev.* **1974**, *8*, 183. (c) Rao, C. N. R.; Singh, S.; Senthilnathan, V. P. *Chem. Soc. Rev.* **1976**, *5*, 297. (d) Suppan, P. *J. Photochem. Photobiol. A* **1990**, *50*, 293.

(12) (a) Ridley, J. E.; Zerner, M. C. *Theor. Chim. Acta* **1973**, *32*, 111. (b) Ridley, J. E.; Zerner, M. C. *Theor. Chim. Acta* **1976**, *42*, 223.

(13) Pople, J. A.; Beveridge, D. A.; Dobash, P. A. *J. Chem. Phys.* **1967**, *47*, 2026.

(14) (a) Bacon, A. D.; Zerner, M. C. *Theor. Chim. Acta* **1979**, *53*, 21. (b) Zerner, M. C.; Loew, G. H.; Kirchner, R. F.; Mueller-Westerhof, U. T. *J. Am. Chem. Soc.* **1980**, *102*, 589. (c) Anderson, W. P.; Cundari, T.; Drago, D.; Zerner, M. C. *Inorg. Chem.* **1990**, *29*, 1. (d) Anderson, W. P.; Cundari, T.; Zerner, M. C. *Int. J. Quantum Chem.* **1990**, *39*, 31.

(15) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, *101*, 7147.

(16) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.

(17) Easton, R. E.; Giesen, D. J.; Welch, A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chim. Acta* **1996**, *93*, 281.

(18) Li, J.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1998**, *99*, 192.

(19) (a) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, *101*, 7147. (b) Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *J. Org. Chem.* **1998**, *63*, 4305.

(20) Das, P. K.; Dockter, D. W.; Fahey, D. R.; Lauffer, D. E.; Hawkins, G. D.; Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G.; Dapprich, S.; Froese,R. D. J.; Holthausen, M. C.; Liu, Z.; Mogi, K.; Vyboishchikov, S.; Musaev, D. G.; Morokuma, K. In *Transition State Modeling for Catalysis*; ACS Symposium Series 712; Truhlar, D. G., Morokuma, K., Eds.; American Chemical Society: Washington, DC, 1999; p 208.

(21) Li, J.; Cramer, C. J.; Truhlar, D. G. *Int. J. Quantum Chem.*, in press.

(22) Zerner, M. C.; Ridley, J. E.; Bacon, A. D.; Edwards, W. D.; Head, J. D.; McKelvey, J.; Culberson, J. C.; Knappe, P.; Cory, M. G.; Weiner, B.; Baker, J. D.; Parkinson, W. A.; Kannis, D.; Yu, J.; Roesch, N.; Kotzian, M. Tamm, T.; Karelson, M. M.; Zheng, X.; Pearl, G.; Broo, A.; Albert, K.; Cullen, J. M.; Cramer, C. J.; Truhlar, D. G.; Li, J.; Hawkins, G. D.; Liotard, D. A. ZINDO-version 99.1, 1999.