QSPR Treatment of the Unified Nonspecific Solvent Polarity Scale

Alan R. Katritzky* and Lan Mu

Florida Center of Heterocyclic Compounds, Department of Chemistry, University of Florida, Gainesville, Florida 32611-7200

Mati Karelson

Department of Chemistry, University of Tartu, 2 Jakobi Street, Tartu EE2400, Estonia

A three-parameter QSPR equation with $R^2 = 0.936$ was developed for the unified nonspecific solvent polarity scale (S') on the basis of theoretical molecular descriptors. It correlates S' for 25 structurally diverse solvents within a 5% average absolute error. The correlation equation includes the following three orthogonal theoretical molecular descriptors: (i) the average structural information content (order 0); (ii) the weighted partial negative surface area; and (iii) the hydrogen-bonding acceptor surface area. These descriptors provide insight into nonspecific solvation at the molecular level. Predictions using this three-parameter model are used to extend available S' values to a total of 67 solvents.

INTRODUCTION

Many of the chemical and physical properties of compounds in solution are profoundly influenced by the nature of the solvent. Early on, solvent effects were evaluated using simple macroscopic solvent parameters such as the static dielectric constant, permanent dipole moment, and refractive index or by functions combining them according to simple physical models.[1−5] Such approaches are not appropriate to account for the specific solute/solvent interactions, taking place at the molecular level, and indeed these macroscopic solvent parameters often fail to correlate observed solvent effects. Even in the absence of specific interactions, solvents reorganize in real systems to form cavities which accommodate the solute molecules with stabilization resulting from the interaction of the solute dipole (and induced dipole) with the internal relative permittivity of the cavity.[6] Solvent reorganization and induced dipole moments tend to create an internal permittivity, different from the bulk relative permittivity and the dimensions of the cavity surrounding each solute molecule in a solvent and the internal permittivity of the organized solvent region are difficult to quantify without making assumptions.[6] It is thus not possible to define the solvent polarity in terms of simple physical solvent characteristics, and consequently empirical solvent polarity scales evolved as an alternative approach for predicting or analyzing solvent effects. Numerous polarity scales, some of which involve a number of parameters, have been proposed for the correlation of various chemical phenomena.[7−12] A disadvantage of such scales is that they reflect, in addition to the nonspecific effects, variable specific effects, which are dependent on the type of probe (solute) used to develop each scale. As a result, a scale that works for one system may not be suitable for another.

$$\Delta\chi = PS' + E_A{}^*E_B + C_A{}^*C_B + W \qquad (1)$$

A unified solvation model has been proposed by Drago for solvation (eq 1).[6,13] In this treatment, solvent effects are separated into nonspecific (PS′) and specific ($E_A{}^*E_B + C_A{}^*C_B$) interactions. The former arises from the electrostatic forces including polarization forces that arise from dipole-induced dipole moments between the solvent and charged ions or dipolar molecules. The latter comprise electron-pair donor and acceptor interactions which include hydrogen-bonding and $\pi-\pi^*$ charge transfer. In eq 1, $\Delta\chi$ is the solvent dependent physicochemical property of a series of solutes, P is the susceptibility of the solute probe to polarity, and S' is the solvent polarity scale derived from experimental observations, which encompasses a wide variety of solutes and solvents but excludes the specific solute−solvent interactions. Thus the nonspecific interactions are unified and the disparities with this scale are attributed to specific effects. Deviations resulting from predictions using S' allow an analysis of the factors that underlie the unusual effects, instead of finding correlations by searching for another scale.[6]

As has been recently reviewed,[13] S' has been successful in many applications, such as the prediction of nonspecific solvation interactions for various solutes and solvents, the analysis of physical data in a series of solvents for nonspecific solvation contributions, and the selection of appropriate solvents for dissolving solids, recrystallizing mixtures, operating high pressure liquid chromatography (HPLC) separations, growing single crystals, increasing the rates of reactions, improving the selectivity of reactions that involve competing pathways, preparing emulsions, and synthesizing copolymers. The S' scale also provides general guidance for solubility and reactivity in the absence of specific interactions.[13]

However, similar to other empirical parameters, up until now S' values have always been derived from experimental observations, which has hindered the use of S' to evaluate solvents whose experimental properties are difficult or impossible to measure and for which S' values are not available. Furthermore, as an empirical scale, the direct relationship of S' with the molecular structure of the solvent is unclear. These limitations have encouraged our present application of theoretical QSPR to the S' scale of solvation.

QSPR Treatment of the Solvent Polarity Scale

J. Chem. Inf. Comput. Sci., Vol. 37, No. 4, 1997 **757**

Table 1. The One-Parameter Correlations for the 25 Diverse Solvents

descriptor	$R^2_{weight\ 0.8-1.0}$
average information content (order 0)	0.8208
average structural information content (order 0)	0.8054
total dipole moment	0.7538
weighted partial negative surface area	0.4377

Table 2. The Best Three-Parameter Correlation of the S' for the 25 Diverse Solvents

descriptor	x ± Δx	t-test	R^2	R_{cv}^2
intercept	0.60 ± 0.11	5.38		
ASIC°	29.99 ± 0.32	9.39	0.8054	0.7397
WPNS	−0.08 ± 0.01	−5.99	0.8954	0.8576
HASA	0.04 ± 0.01	3.62	0.9361	0.9001

The QSPR (the quantitative structure−property relationships) technique attempts to establish a relationship between a physical or chemical property and defined features of the molecular structures. The molecular structure contains encoded within it all conceivable information regarding the physical properties, chemical behavior, and biological effects of a compound and a developed QSPR model attempts to deduce, identify, and rationalize the links between chemical structure and their properties. By the same token, QSPR should, in principle, be able to link any empirical solvent scale with theoretical molecular descriptors. Since the descriptors involved in a QSPR model are calculated solely from the chemical structures of compounds, the influence

of structural features on solvation should therefore be capable of elucidation, and thus ideally the applicability of S' could be extended to any compound, known or unknown without synthesizing it or making any measurements.

Famini and co-workers[14−16] pioneered the development of theoretical solvent polarity scales. Their TLSER (theoretical linear solvation energy relation) (eq 2) was developed to be analogous to the LSER (linear solvation energy relation) of Kamlet and Taft.[8] In eq 2, P is the experimental property under discussion and may relate to either solute or solvent; V_{mc} is the steric term, π_1 is the molecular polarizability, ϵ_a is the LUMO energy relative to the HOMO energy of water, ϵ_b is HOMO energy relative to LUMO energy of water, Q_+ is the partial charge of the most positive hydrogen atom of a molecule, and Q_- is the partial charge of the most negative atom of a molecule, all parameters are calculated for either solutes or for solvents but not for both. TLSER has been used successfully to correlate a range of properties including HPLC retention indices,[16] octanol−water partition coefficients, rate constants for the hydrolysis of organophosphonothioates,[16,17] and rate constants for the decarboxylation of substituted 3-carboxybenzisoxazoles.[18] The development of TLSER demonstrated that complex solvent effects can be explored using fundamental molecular properties determined by quantum chemical techniques. However, eq 2 is constrained by certain solvation processes, from which the equation is developed. Thus, the coefficients developed from one type of solute/solvent may not be suitable for other types

Table 3. The Experimental and Predicted S' of 67 Compounds[a]

no.	solvent	formula	$S'_{(exp)}$	$S'_{(calc)}$	$\Delta S'$	wt[c]	no.	solvent	formula	$S'_{(exp)}$	$S'_{(calc)}$	$\Delta S'$	wt[c]
1	1,1,1-trichloroethane	Cl_3CCH_3	1.93	2.40	0.47	0.2	35	*N,N*-diethylformamide	$HC(O)NEt_2$	2.80	2.71	0.09	0.6
2	1,1,2-trichloroethane	ClH_2CHCl_2	2.35	2.44	0.09	0.2	36	*N,N*-dimethylacetamide	$CH_3CON(CH_3)_2$	2.70	2.60	0.10	1
3	1,1,2-trichloroethene	$Cl_2C=CHCl$	1.90	2.33	0.43	0.4	37	*N,N*-dimethylcyanamide	$(CH_3)_2NCN$	2.81	2.84	0.03	0.2
4	1,2-dichlorobenzene	$1,2-Cl_2C_6H_4$	2.10	2.03	0.07	0.6	38	*N,N*-dimethylformamide	$HCON(CH_3)_2$	2.80	2.71	0.09	1
5	1,4-dioxane	$O(CH_2CH_2)_2O$	1.93	1.93	0.00	1	39	*n*-pentane	C_5H_{12}	0.57	1.34	0.77	0.2
6	2-butanone	$CH_3C(O)C_2H_5$	2.50	2.33	0.17	0.8	40	*n*-nonane	C_9H_{20}	0.90	1.27	0.37	0.2
7	3-methylsulfolane	$C_5H_{10}SO_2$	2.55	3.67	0.12	0.4	41	*N*-methylpyrrolidinone	$CH_2CH_2CH_2CONCH_3$	2.62	2.60	0.02	0.6
8	3-pentanone	$CH_3COC_3H_7$	2.37	2.17	0.20	0.8	42	*N*-methylimidazole	$C_4H_6N_2$	2.60	2.70	0.10	0.8
9	4-butyrolactone	$C_4H_6O_2$	2.86	2.67	0.19	1	43	*n*-hexane	C_6H_{14}	0.68	1.28	0.60	0.2
10	4-methylpyridine	$4-CH_3C_5H_4N$	2.31	2.36	0.05	0.8	44	*n*-heptane	C_7H_{16}	0.79	1.29	0.50	0.2
11	acetone	$(CH_3)_2CO$	2.58	2.62	0.04	1	45	*n*-decane	$C_{10}H_{22}$	0.90	1.27	0.37	0.2
12	acetonitrile	CH_3CN	3.00	3.17	0.17	1	46	*n*-butyronitrile	$CH_3CH_2CH_2CN$	2.70	2.40	0.30	0.6
13	acetophenone	$C_6H_5C(O)CH_3$	2.52	2.62	0.10	0.8	47	nitrobenzene	$C_6H_5NO_2$	2.61	2.84	0.23	1
14	anisole	$C_6H_6OCH_3$	2.04	2.16	0.08	0.6	48	nitroethane	$C_2H_5NO_2$	2.78	2.78	0.00	0.2
15	benzene	C_6H_6	1.73	1.76	0.03	1	49	nitromethane	CH_3NO_2	3.07	3.10	0.03	0.8
16	benzonitrile	C_6H_5CN	2.63	2.61	0.02	1	50	pentahydropyran	$(CH_2)_5O$	1.98	1.64	0.34	0.4
17	bromobenzene	C_6H_5Br	2.10	1.99	0.11	0.4	51	propionitrile	C_2H_5CN	2.80	2.64	0.16	0.6
18	butyl acetate	$CH_3C(O)OBu$	1.99	2.25	0.26	0.4	52	propyl acetate	$CH_3C(O)OPr$	2.05	2.34	0.29	0.4
19	carbon disulfide	CS_2	1.51	2.72	1.21	0.4	53	propylene carbonate	$(CH_2)_3(O-)_2CO$	3.10	2.83	0.27	0.6
20	carbon tetrachloride	CCl_4	1.49	1.57	0.08	1	54	pyridine	C_5H_5N	2.44	2.27	0.17	1
21	chlorobenzene	C_6H_5Cl	1.98	2.02	0.04	0.8	55	quinoline	C_9H_7N	2.30	2.14	0.16	0.2
22	cyclohexane	C_6H_{12}	1.11	1.29	0.18	1	56	tetrahydrofuran	$(CH_2)_4O$	2.08	1.75	0.33	1
23	cyclohexanone	$(CH_2)_5CO$	2.35	2.29	0.06	0.6	57	tetrahydrothiophene	$(CH_2)_4S$	1.99	1.64	0.35	0.2
24	diisopropyl ether	$i-Pr_2O$	1.76	1.59	0.07	0.8	58	tetramethylene sulfone	$(CH_2)_4SO_2$	2.88	3.59	0.71	0.4
25	di-*n*-butyl ether	$(n-C_4H_9)_2O$	1.58	1.51	0.07	0.6	59	tetramethylurea	$[(CH_3)_2N]_2CO$	2.48	2.50	0.02	0.6
26	dichloromethane	CCl_2H_2	2.08	2.78	0.30	0	60	thiophene	$(CH)_4S$	1.83	2.32	0.49	0.2
27	diethyl ether	$(C_2H_5)_2O$	1.73	1.66	0.07	1	61	toluene	$C_6H_5CH_3$	1.66	1.89	0.33	0.8
28	dimethylaniline	$C_6H_5N(CH_3)_2$	1.96	2.07	0.11	0.6	62	tributyl phosphate	$(n-C_4H_9O)_3PO$	2.30	2.93	0.63	0.4
29	DMSO	$(CH_3)_2SO$	3.00	2.80	0.20	1	63	trichloromethane	CCl_3H	1.74	2.53	0.79	0
30	ethyl acetate	$CH_3C(O)OC_2H_5$	2.50	2.51	0.01	0.8	64	triethylamine	$(C_2H_5)_3N$	1.43	1.54	0.11	1
31	ethyl formate	$HCOOC_2H_5$	2.24	2.98	0.74	0.4	65	triethyl phosphate	$(C_2H_5O)_3PO$	2.55	2.53	0.02	0.6
32	hexamethylphosphoramide	$[(CH_3)_2N]_3PO$	2.52	2.55	0.03	0	66	trimethylbenzene	$C_6H_3(CH_3)_3$	1.54	1.85	0.31	0.4
33	hexyl acetate	$CH_3C(O)Ohex$	1.94	2.19	0.25	0.4	67	trimethyl phosphate	$(CH_3O)_3PO$	2.79	3.62	0.83	0.2
34	methyl acetate	$CH_3C(O)OCH_3$	2.35	2.66	0.31	0.6							

[a] S' are collected from four sources: Drago, R. S. *Applications of Electrostatic-Covalent Models in Chemistry*; Surfside Scientific Publishers: Gainesville, FL, 1994. Bustamante, P.; Drago, R. S. Extension of the Unified Solvation Model to Systems Requiring a Cavity Term. *J. Chem. Soc., Perkin Trans. 2* submitted for publication. George, J. E.; Drago, R. S. Extension of the Unified Solvation Model to Organometallic and Polycyclic Aromatic Probes. *Inorg. Chem.* **1996**, *35*, 239−241. Personal communication by R. S. Drago. [b] Calculated using the best three-parameter model. [c] The statistical weights are cited from reference 27.
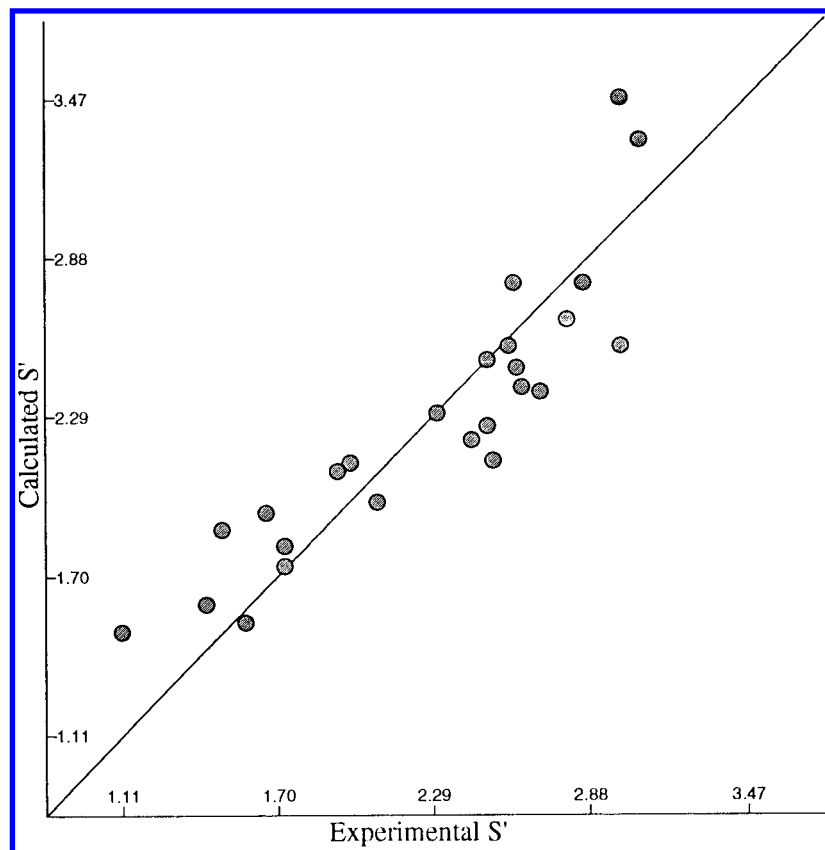
Figure 1. The calculated *vs* experimental S′ of 25 diverse compounds using the best one-parameter correlation equation.

of solute/solvent, *i.e.*, the regression coefficients vary from different species and processes.

$$\log P = mV_{mc} + p\pi_1 + a\epsilon_a + a'Q_+ +$$
$$b\epsilon_b + b'Q_- + \text{intercept} \quad (2)$$

Our groups have recently developed successful QSPR correlation equations using the CODESSA (Comprehensive Descriptors for Structural and Statistical Analysis) approach for boiling points,[19,20] melting points,[19,21] flash points,[22] gas solubilities,[23] gas chromatographic response factors and retention times,[24] polymer glass transition temperatures,[25] and critical micelle concentrations.[26] Furthermore, our previous investigations[20] demonstrated that, by deliberately screening the redundant and overlapping factors in descriptor space, quantum-chemical parameters that resemble the intrinsic dimensions of a property can be established. We therefore anticipated that the use of CODESSA to develop a QSPR equation for *S′* would be useful both for the prediction of *S′* for additional solvents and at the same time provide insight into the physical mechanism by which chemical structure determines *S′*. In this work, *S′* values of 25 solvents, whose accuracy was weighted between 0.8−1.0, were used to develop the QSPR model. The resulted correlation equation was then used to predict *S′* values for another 42 solvents and to interpret the physical meaning of *S′*.

METHODOLOGY

Experimental *S′* data for 67 solvents (Table 3) were assembled from previous publications.[13,27,28] This solvent set included saturated and unsaturated hydrocarbons, halogenated solvents, and compounds containing cyano, nitro, amide, sulfide, mercapto, sulfone, phosphate, ester, ether, and carbonyl groups and furan, pyran, dioxane, pyridine,

aniline, quinoline, imidazole, pyrrolidinone, and pyrazine rings. Structures of the solvents were drawn and preoptimized by the MMX molecular mechanics method using the PCMODEL[29] program. The final geometry optimization of these compounds was performed on an IBM RISC/6000 model 320 computer using the semiempirical quantum-chemical AM1 parameterization with the MOPAC 6.0 program which was modified by incorporation of the self-consistent reaction field (SCRF) method.[30] The MOPAC results for individual compounds were loaded into the CODESSA program[31] along with the *S′* data.

The CODESSA program implements procedures which enable the calculation of a large selection of descriptors including molecular volume and many other geometrical characteristics of the molecule, such as Jurs' shape parameters,[32] and a large variety of topological indices (Wiener indices,[33] Randić connectivity indices,[34] Kier and Hall connectivity indices,[35] and information content indices[36]). Numerous quantum-chemical descriptors are extracted and calculated from the AM1 output.[19,20,24,31,37] These include the most positive and the most negative Mulliken net atomic charges, frontier molecular orbital (FMO) energies, and the respective Fukui FMO nucleophilic, electrophilic, and one-electron reactivity indices. The total dipole moment of the molecule, dipole moment components, and molecular bond orders were also used as descriptors. Additional, more specific, descriptors included the valence state energies of atoms and total Coulombic and exchange energies between atoms in the molecule. The CPSA descriptors proposed by Jurs *et al.*[38,39] were also included in the CODESSA program based on AM1 calculated atomic partial charges. In the present work, about 300 molecular descriptors were calculated using CODESSA for all 67 compounds. Descriptors that are solely associated with a specific constituent, such
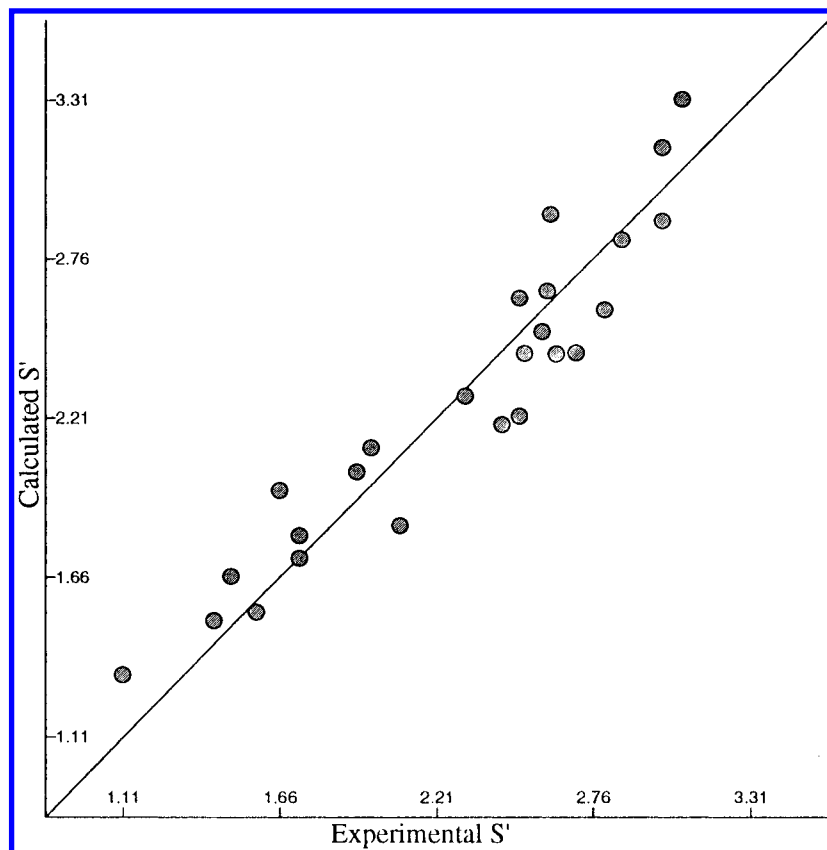
QSPR TREATMENT OF THE SOLVENT POLARITY SCALE

J. Chem. Inf. Comput. Sci., Vol. 37, No. 4, 1997 **759**

Figure 2. The calculated *vs* experimental S′ of 25 diverse compounds using the best two-parameter correlation equation.

as the number of C atoms, the minimum atomic state energy for a C atom, and the average valence of a C atom *etc*. were not included as they were considered to be of little relevance. The correlation analysis to find the best QSPR model of a given size was carried out using the procedure based on the stepwise scale addition method.[31] The following heuristic procedure was carried out using this approach.

(1) To reduce the number of descriptors in the starting set, the following criteria were applied to eliminate descriptors for which

(a) the *F*-value for the one-parameter correlation with the descriptor was below 1,

(b) the correlation coefficient for this one-parameter equation was less than R_{min}, a user-defined value for insignificant correlations ($R_{min} = 0.01$ was used in the present work),

(c) the *t*-value for descriptor in the one-parameter correlation was less than $t1 = 1.5$,

(d) the descriptor which was highly intercorrelated with another descriptor characterized by a higher single-parameter correlation coefficient value for a given property.

All the remaining descriptors were then listed in decreasing order according to the correlation coefficient of the corresponding one-parameter correlation equation.

(2) Starting with the first descriptor from this preselected list of descriptors, two-parameter correlations were calculated using the following pairs: (a) the first descriptor with each of the remaining descriptors, the second descriptor with each of the remaining descriptors, *etc*. This procedure was continued until for some *n*th descriptor no correlations with an *F*-value above one-third of the maximum *F*-value for a given set had been found. Pairs of descriptors with an

intercorrelation above R_{nc} were not considered. The *NS* (number of descriptor sets to select for the next recursion level) best pairs showing the highest *F*-values in the two-parameter correlations were selected and processed further as the *working* sets.

(3) Each of the remaining descriptors, if not correlated over R_{nc} with the descriptors already included, were added to the selected *working* set of descriptors. When the resulting correlation had an *F*-value above $F_{working}n/(n + 1)$ (where *n* is a number of descriptors in the *working* set plus one), *i.e.*, when this correlation was more significant than the *working* correlation, this extended set of descriptors was treated further.

(4) After all the descriptors have been processed one-by-one, if the maximum number of descriptors ND_{max} allowed by the user have not yet been achieved, then the NS best extended *working* sets, *i.e.*, the sets with the highest *F*-values, were submitted to the procedure of step 2. Otherwise, the procedure was completed, and the correlation with the highest *F*-value was stored as the final result. The same method was previously employed in QSPR studies by our groups of many other properties of diverse organic compounds and polymers. The reliability of the heuristic procedure was compared to that of other methods such as the best multilinear regression in ref 20.

RESULTS AND DISCUSSION

A group of 25 structurally diverse solvents, whose *S′* accuracy has been weighted on a 0.0−1.0 scale as 0.8 or 1.0,[27] was selected to develop the QSPR model. Several descriptors were found important for one-parameter correlations: the average information content (order 0), denoted as AIC^0 (eq 3);[35] the average structural information content
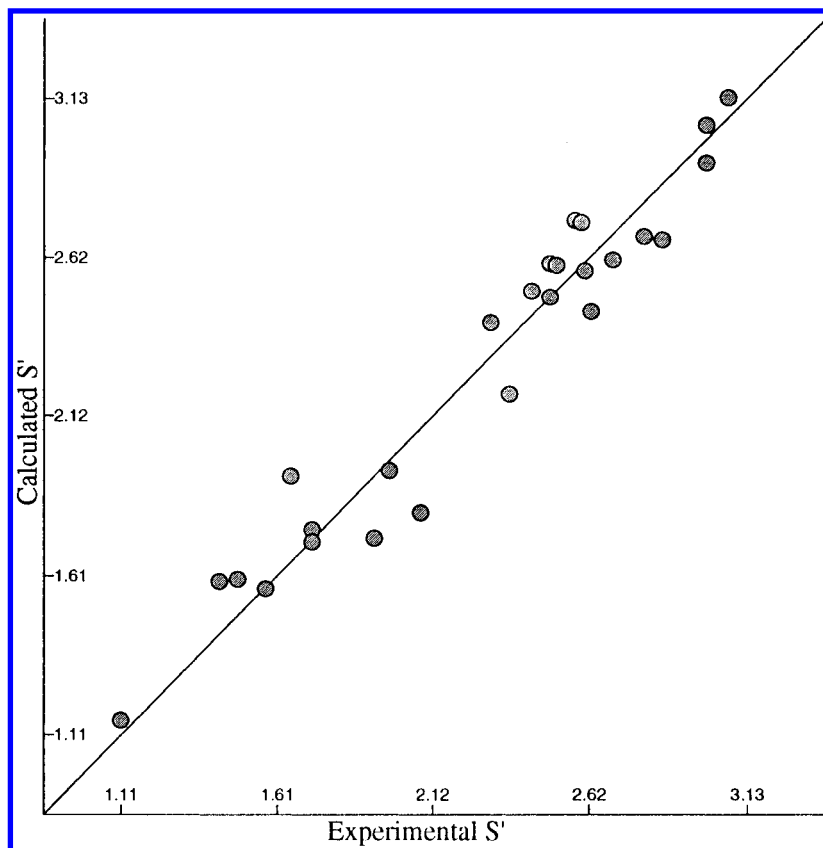
Figure 3. The calculated *vs* experimental S′ of 25 diverse compounds using the best three-parameter correlation equation.

(order 0), denoted as ASIC⁰ (eq 4);[35] the total dipole moment of the molecule; the weighted partial negative surface area (WPNS, eq 5) *etc.* In both eqs 3 and 4, n_i is the number of atoms in the *i*th class, and n is the total number of atoms in the molecule. In eq 5, S_{tot} is the total molecular surface area, calculated from the van der Waals' radii of the atoms (overlapping spheres), q_- is the partial charge on the negatively charged atom(s), and s_- is the partial surface area of the negatively charged atom(s). The squared correlation coefficients for the best one-parameter correlations are provided in Table 1.

$$\text{AIC}^0 = -n\sum_i \frac{n_i}{n}\log_2\frac{n_i}{n} \qquad (3)$$

$$\text{ASIC}^\circ = -n\sum_i \frac{n_i}{n}\log_2\frac{n_i}{n}/\log_2 n \qquad (4)$$

$$\text{WPNS} = \frac{S_{tot}}{1000}\sum q_-s_- \qquad (5)$$

The various average information content indices encode the branching ratio and constitutional diversity of a molecule. Their significance in solvation has been demonstrated in a previous study on solubility of gaseous hydrocarbons in water.[23] The importance of the total dipole moment of the molecule to S′ was already well established in previous discussion of correlations between S′ and this quantity.[6,13] The WPNS measures the accessible surface areas weighted charges of the negatively charged atoms in a molecule, which relates to the intermolecular electrostatic interactions between the solute and solvent.

The best three-parameter correlation selected by the heuristic procedure, described before, involved: (i) ASIC⁰;

(ii) the weighted partial negative surface area (WPNS); and (iii) the hydrogen-bonding acceptor surface area (HASA). The HASA is defined by eq 6, where S_A is the surface area on hydrogen-bonding acceptors. The HASA measures the solute accessible surface area of nitrogen and oxygen atoms in a molecule. Table 2 provides the successive regression coefficients for each descriptor and the respective standard errors, the R^2 and R_{cv}^2 values, and the *t*-test value for this correlation.

$$\text{HASA} = \sum S_A \qquad (6)$$

Notice that the most important descriptor, ASIC⁰, in the three-parameter model has $R^2 = 0.8054$, $F = 95$, and $s = 0.25$ for the single parameter correlation (Figure 1), which is slightly worse than the best one-parameter correlation (Table 1). Addition of WPNS improved the correlation to $R^2 = 0.8954$, $F = 95$, and $s = 0.19$ (Figure 2). By including HASA as the third descriptor, the correlation enhanced to $R^2 = 0.9361$, $F = 103$, and $s = 0.15$ with an average absolute error of 0.11 unit (Figure 3). The stability of the model is characterized by the corresponding cross-validated correlation coefficients (R_{cv}^2).[31] The predicted S′ value for all 67 solvents along with their observed values are given in Table 3. Knowing that the S′ values weighted at 0.8 could have an averaged absolute error about 0.2, the 42 predicted S′ data, for which the accuracy of the observations were weighted from 0–0.6 on a 0–1.0 scale, should be better references for nonspecific solvent polarity scale.

The three descriptors involved in the final QSPR model are almost orthogonal. The R^2 for the intercorrelations (the correlation of any one of the three descriptors against the other two) range from 0.16 to 0.22. These descriptors directly reflect the molecular structural features of the molecule, which quantitatively revealed the structural con-

QSPR Treatment of the Solvent Polarity Scale

J. Chem. Inf. Comput. Sci., Vol. 37, No. 4, 1997 **761**

tribution of a solvent to S'. Namely, the surface area weighted negative partial charges, the surface areas on H-bonding acceptors, and the topological layouts of the atoms contribute to the electrostatic forces including polarization forces that arise from the dipole-induced dipole moments. We note that the negative sign in front of the coefficient of the WPNS (Table 2) does not indicate a negative contribution of this descriptor to nonspecific solvation. On the contrary, since the WPNS has negative values for all solvents, the negative sign shows a positive contribution of the WPNS toward S'. However, the three descriptors may encode a more profound physical meaning which relates to S'. Various approaches are being applied to explore the physical meaning of these descriptors and S', and the results will be reported elsewhere.

CONCLUSION

We have established a unified nonspecific solvent polarity scale using a correlation equation with three theoretical molecular descriptors. The correlation equation allows the prediction of S' of structurally diverse solvents with an average absolute error of 0.11 units, which accounts for a 5% average absolute error. All parameters in the correlation equation are derived from the quantum-chemical calculated descriptors of the solvents and adequately reflect the solvent/solute interactions in the solvent's internal cavity. The prediction of the S' no longer requires several measurements of physical and chemical quantities. The correlation and analysis of nonspecific solvent effects can be more conveniently applied in the development of new technologically and biomedically important solvents.

ACKNOWLEDGMENT

REFERENCES AND NOTES

(1) Kirkwood, J. G. Theory of Solutions of Molecules Containing Widely Separated Charges with Special Application to Zwitterions. *J. Chem. Phys.* **1934**, *2*, 351−361.
(2) Onsager, L. Electric Moments of Molecules in Liquids. *J. Am. Chem. Soc.* **1936**, *58*, 1486−1493.
(3) Kirkwood, J. G. The Dielectric Polarization of Polar Liquids. *J. Chem. Phys.* **1939**, *7*, 911−919.
(4) Block, H.; Walker, S. M. A Modification of the Onsager Theory for a Dielectric. *Chem. Phys. Lett.* **1973**, *19*, 363−364.
(5) Brady, J. E.; Carr, P. W. An Analysis of Dielectric Models of Solvatochromism. *J. Phys. Chem.* **1985**, *89*, 5759−5766.
(6) Drago, R. S. A Unified Scale for Understanding and Predicting Nonspecific Solvent Polarity: A Dynamic Cavity Model. *J. Chem. Soc. Perkin Trans. 2* **1992**, 1827−1838.
(7) Abboud, J. L.; Kamlet, M. J.; Taft, R. W. Regarding a Generalized Scale of Solvent Polarities. *J. Am. Chem. Soc.* **1977**, *99*, 8325−8327.
(8) Kamlet, M. J.; Abbound J.-L. M.; Abraham M. H.; Taft, R. W. Linear Solvation Energy Relationships. 23. A Comprehensive Collection of the Solvatochromic Parameters, π^*, α, and β, and Some Methods for Simplifying the Generalized Solvatochromic Equation *J. Org. Chem.* **1983**, *48*, 2877−2887.
(9) Kamlet, M. J.; Abboud, J. L. M.; Taft, R. W. An Examination of Linear Solvation Energy Relationships. *Prog. Phys. Org. Chem.* **1981**, *13*, 485−623.
(10) Reichardt, C. *Solvents and Solvent Effects in Organic Chemistry*, 2nd ed.; VCH Publishers: Weinheim, 1988.
(11) Reichardt, C. Solvatochromic Dyes as Solvent Polarity Indicators. *Chem. Rev.* **1994**, *94*, 2319−2358.
(12) Abraham, M. H. Scales of Solute Hydrogen-bonding: Their Construction and Application to Physicochemical and Biochemical Processes. *Chem. Soc. Rev.* **1993**, *22*, 73−83.
(13) Drago, R. S. *Applications of Electrostatic-Covalent Models in Chemistry*; Surfside Scientific Publishers: Gainesville, FL, 1994.
(14) Lowrey, A. H.; Cramer, C. J.; Urban, J. J.; Famini, G. R. Quantum Chemical Descriptors for Linear Solvation Energy Relationships. *Computers Chem.* **1995**, *19*, 209−215.

(15) (a)Famini, G. R.; Ashman, W. P.; Mickiewicz, A. P; Wilson, L. Y. Using Theoretical Descriptors in Quantitative Structure−Activity Relationships: Opiate Receptor Activity by Fentanyl-like Compounds. *Quant. Struct.-Act. Relat.* **1992**, *11*, 162−170. (b)Donovan, W. H.; Famini, G. R. Using Theoretical Descriptors in Structure Activity Relationships: Retention Indices of Sulfur Vesicants and Related Compounds. *J Chem. Soc., Perkin Trans. 2* **1996**, 83−89.
(16) Cramer, C. J.; Famini, G. R.; Lowrey, A. H. Use of Calculated Quantum Chemical Properties as Surrogates for Solvatochromic Parameters in Structure-Activity Relationships. *Acc. Chem. Res.* **1993**, *26*, 599−605.
(17) Trinajstic, N. *Chemical Graph Theory*, 2nd ed.; CRC Press: Boca Raton, FL, 1992.
(18) Famini, G. R.; Wilson, L. Y. Using Theoretical Descriptors in Quantitative Structure-Property Relationships: 3-Carboxybenzisoxazole Decarboxylation Kinetics. *J. Chem. Soc., Perkin Trans. 2* **1994**, 1641−1650.
(19) Murugan, R.; Grendze, M. P.; Toomey, J. E., Jr.; Katritzky, A. R.; Karelson, M.; Lobanov, V.; Rachwal, P. Predicting Physical Properties from Molecular Structure. *CHEMTECH* **1994**, *24*, 17−23.
(20) Katritzky, A. R.; Mu, L.; Lobanov, V. S.; Karelson, M. Correlation of Boiling Points with Molecular Structure. I. A Training Set of 298 Diverse Organics and a Test Set of 9 Simple Inorganics. *J. Phys. Chem.* **1996**, *100*, 10400−10407.
(21) Katritzky, A. R.; Maran, U.; Karelson, M.; Lobanov, V. S. Prediction of Melting Points for the Substituted Benzenes: A QSPR Approach manuscript in preparation.
(22) Katritzky, A. R.; Lobanov, V. S.; Karelson, M.; Murugan, R; Grendze, M. J.; Toomey, J. E., Jr. Comprehensive Descriptors for Structural and Statistical Analysis. 1. Correlations Between Structure and Physical Properties of Substituted Pyridines. *Rev. Roum. Chim.* **1996**, *41*, 851−867.
(23) Katritzky, A. R.; Mu, L.; Karelson, M. A QSPR Study of the Solubility of Gases and Vapors in Water. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1162−1168.
(24) Katritzky, A. R.; Ignatchenko, E. S.; Barcock, R. A.; Lobanov, V. S.; Karelson, M. Prediction of Gas Chromatographic Retention Times and Response Factors Using a General Quantitative Structure-Property Relationship Treatment. *Anal. Chem.* **1994**, *66*, 1799−1807.
(25) Katritzky, A. R.; Rachwal, P.; Law, K. W.; Karelson M; Lobanov, V. S. Prediction of Polymer Glass Transition Temperatures Using a General Quantitative Structure-Property Relationship Treatment. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 879−884.
(26) Huibers, P. D. T.; Lobanov, V. S.; Katritzky, A. R.; Shah, D. O.; Karelson, M. Prediction of Critical Micelle Concentration Using a Quantitative Structure-Property Relationship Approach. 1. Nonionic Surfactants. *Langmuir* **1996**, *12*, 1462−1470.
(27) Bustamante, P.; Drago, R. S. Extension of the Unified Solvation Model to Systems Requiring a Cavity Term. *J. Chem. Soc., Perkin Trans. 2* Submitted for publication.
(28) George, J. E.; Drago, R. S. Extension of the Unified Solvation Model to Organometallic and Polycyclic Aromatic Probes. *Inorg. Chem.* **1996**, *35*, 239−241.
(29) *PCMODEL*, *Manual*, 5th ed.; Serena Software: Bloomington, IN, 1992.
(30) Karelson, M. M.; Tamm, T.; Katritzky, A. R.; Cato, S. J.; Zerner, M. C. M. O. Calculations Applicable to Condensed Phases: The Combination of Self-Consistent Reaction Field Theory with Semi-Empirical Quantum-Chemical Methods. *Tetrahedron Comput. Methodol.* **1989**, *2*, 295−304.
(31) Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *CODESSA, Training Manual*; Gainesville, FL, 1995.
(32) Rohrbaugh, R. H.; Jurs, P. C. Descriptions of Molecular Shape Applied in Studies of Structure/Activity and Structure/Property Relationships. *Anal. Chim. Acta* **1987**, *199*, 99−109.
(33) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17−20.
(34) Randić, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609−6615.
(35) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure-Activity Analysis*; John Wiley & Sons Inc.: New York, 1986.
(36) Stankevich, M. I.; Stankevich, I. V.; Zefirov, N. S. Topological Indices in Organic Chemistry. *Russ. Chem. Rev. (Engl. Transl.)* **1988**, *57*, 191−208.
(37) Karelson, M.; Lobanov, V. S.; Katritzky, A. R. Quantum-Chemical Descriptors in QSAR/QSPR Studies. *Chem. Rev.* **1996**, *96*, 1027−1043.
(38) Stanton, D. T.; Jurs, P. C. Development and Use of Charged Partial Surface Area Structural Descriptors in Computer-Assisted Quantitative Structure-Property Relationship Studies. *Anal. Chem.* **1990**, *62*, 2323−2329.
(39) Stanton, D. T.; Egolf, L. M.; Jurs, P. C.; Hicks, M. G. Computer-Assisted Prediction of Normal Boiling Points of Pyrans and Pyrroles. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 306−316.

CI970011K