A Unified Treatment of Solvent Properties

Alan R. Katritzky,*,§ Tarmo Tamm,§,‡ Yilin Wang,§ and Mati Karelson‡

Center for Heterocyclic Compounds, Department of Chemistry, University of Florida, P.O. Box 117200, Gainesville, Florida 32611, and Department of Chemistry, University of Tartu, 2 Jakobi Str., Tartu, 51014, Estonia

Principal component analysis (PCA) has been carried out with 40 solvent scales as variables, each having 40 data points for 40 solvents as objects. The first three components account for 74% of the total variance. For 36 of the scales, an average of 88% of the variance is described by the first three principal components. The solvents and the solvent scales are grouped according to the scores and loadings obtained from PCA treatment. This allows comparison of both solvent scales and characterization of individual solvents.

INTRODUCTION

The preceding paper[1] assembles the literature data for 45 scales that have been widely used to characterize the properties of solvents. For each scale, a QSPR equation was derived on the basis of theoretically calculated molecular descriptors. The equations allow, in principle, the prediction of the magnitude of that scale for any solvent, including those for solvents where no experimental values are available. In the present paper, a principal component analysis (PCA) is used to search for insights into (i) inter-relationships between the diverse solvent scales and the mechanisms by which chemical structure influences the solvent properties and (ii) similarity/diversity of the individual solvents.

Previous Work. It was realized early that no single solvent scale could offer a general correlation for solvent effects. Many previous attempts have been made to formulate multiparameter equations, either by a combination of two or more existing scales or by postulating various specific parameters to account for distinct types of effects. One of the earliest such approaches, made almost 30 years ago in one of our groups,[2] combined several popular solvent parameters into multiparameter models and presented correlations with experimental data from different types of solvent dependent processes. Independently of this, a highly significant and wide-ranging early contribution by Koppel and Palm[3] defined their fundamental four parameters for characterizing solvents: polarity, polarizability, basicity, and acidity. The basicity parameter B was derived from IR spectra and the acidity parameter E was calculated using the E_T, P, and Y parameters. Of them, E_T was taken from the work of Reichardt,[4] P was derived from measurements of refractive index,[3] and Y was derived from the dielectric constant. This work was continued by Koppel and Paju,[5] and more recently by Palm and Palm,[6,7] who developed a nine-parameter model and used it to correlate the data from 359 individual solvent dependent processes.

Much insight into solvent properties has been provided by Kamlet and Taft who constructed scales of solvent hydrogen bond acceptor basicity (HBA)[8] and hydrogen bond donor acidity (HBD)[9] and then used these to evaluate the contribution of basicity and acidity to several other solvent scales. The solvatochromic comparison method was also applied by the same authors with Abboud to assemble the π^* scale[10] which combines polarity and polarizability. The Kamlet–Taft expression[11] on linear solvation energy relationship (LSER) successfully describes the relationship between several solvent scales. Importantly, the LSER approach was extended by Famini et al. by the definition of corresponding theoretically derived scales (theoretical LSER or TLSER).[12–14]

Chastrette et. al.[15] applied PCA to classify 83 solvents using eight solvent parameters and concluded that explaining the dimensionality in nine classes of solvents is most reasonable.

Drago and his group proposed[16,17] a unified solution model (USM) which distinguishes between specific and nonspecific contributions of solvent scale parameters. This has been correlated by QSPR to structure descriptors of the solvents.[18,19] Marcus[20] investigated correlations between the nine solvent parameters α, β, π^*, $E_T(30)$, DN, AN, Z, acidity, and basicity (see Table 3 for definition of these symbols) and found several to be interrelated. He concluded that there were four essentially independent solvent parameters: hydrogen bond donation ability (described best by α), hydrogen bond acceptor ability (β), polarity/polarizability (π^*), and solvent stiffness (cohesive energy density, δ_H^2).

Catalan[21] used the $E_T(30)$, π^*, P_y, S', and SPP empirical scales as descriptors of nonspecific solvent effects in an effort to obtain a generalized solvent polarity scale.[22] Sergent, Luu, and Elguero[23] analyzed solvent scales statistically by an optimum design methodology, which is an equilibrated balance between classical design and uniform distribution, and tested the validity of solvent models.

More recently Cramer, Truhlar, and the Minnesota group[24–27] have developed several parametrization models describing free energies of solvation and partition coefficients and building a universal organic solvatation model.

Objectives of the Present Work. Despite the substantial progress summarized above, there is no generally accepted single treatment of solvent effects presently available.

§ University of Florida.
‡ University of Tartu.

A Unified Treatment of Solvent Properties

J. Chem. Inf. Comput. Sci., Vol. 39, No. 4, 1999 **693**

Table 2. Variance Covered by the 10 Components

component	% of variance	tot. variance
1	44.86	44.86
2	15.38	60.24
3	13.96	74.20
4	6.32	80.52
5	5.09	85.61
6	4.48	90.09
7	1.77	91.86
8	1.55	93.40
9	1.34	94.75
10	1.04	95.79

Obviously, no single scale can give an overall treatment. However, the previous attempts to compare and unify the different scales have been of limited scope; thus, no such treatment has considered more than 10 scales, and a constant difficulty has been the large number of missing experimental values.

We envisaged a general treatment in the form of principal component analysis (PCA) of a matrix comprising the most important scales and solvents. Such an approach would be similar to that which we took some years ago[28] to shed light on these many scales of aromaticity and demonstrated that

aromaticity was a multiparameter concept with magnetic and classical aromaticity being orthogonal. This conclusion has successfully withstood and rejected criticism.[29] Thus the objective behind the present study was to assemble, for a representative number of solvents, as full as possible a set of scales associated with diverse solvent characteristics (including both measured and calculated properties) and to determine the principal components of the variance of the scales. We desired to test these components for their ability to group solvent scales by their loadings and solvents by their scores. The components should further allow predictions regarding the characteristics of both solvent scales and of other solvents not used in the original survey. In principle it should be possible to estimate values of the characteristics of scales and solvents previously unavailable.

METHODOLOGY

Recently, we developed[1] QSPR models for 45 individual polarity scales using theoretically derived molecular descriptors. The CODESSA program[30] has now been used to predict for 40 of these polarity scales all missing values for 40 solvents, thus providing a complete matrix 40 × 40 suitable

Table 3. Loadings of the Three Principal Components

group	scale	description	PC1	PC2	PC3	R
i	J	expression of dielectric constant	0.852	0.399	0.040	0.941
	S′	statistical, $\Delta\chi = PS' + W$ (30 probes, 31 solvents, >300 shifts)	0.938	0.163	−0.136	0.962
	d	dielectric constant	0.885	0.361	−0.044	0.957
	Y	(polarity) expression of dielectric constant	0.893	0.205	−0.068	0.919
	N	combination of M and P	0.856	0.336	0.109	0.926
ii	E_T	electronic transition energy of pyridinium *N*-phenolate betaine dye	0.969	−0.135	−0.030	0.979
	S	derived from Kosower Z values, uses *R* for process sensitivity	0.897	−0.265	−0.063	0.937
	$E_{CT(A)}$	UV CT absorption maximums of tetra-*n*-hexylammonium iodide trinitrobenzene	0.840	−0.362	−0.150	0.927
	AN	^{31}P NMR of triethylphosphine oxide in different solvents	0.850	−0.333	−0.115	0.920
	A	free energy changes in 77 reactions (data from electronic, IR, ESR, NMR spectra	0.887	−0.305	−0.065	0.940
	E_B	N→π* transition of 2,2,6,6-tetramethylpiperidine-*N*-oxyl	0.893	−0.363	−0.091	0.968
	Z	transition energy for the CT absorption of 1-ethyl-4 methoxycarbonylpyridinium iodide	0.938	−0.250	0.024	0.971
	E	(acidity) derived from ET and P and Y	0.658	−0.476	−0.188	0.833
	ETSO	UV/vis spectra of *N,N*-(dimethyl)thiobenzamide-*S*-oxide	0.821	−0.381	0.013	0.906
	SA	vis spectra of *o-t*-butylstilbazolium betaine dye and its di-*t*-butyl homomorph	0.779	−0.545	−0.074	0.953
	α	various probes (longest wavelength absorption)	0.755	−0.476	−0.476	0.899
	δ	square root of cohesive pressure	0.891	−0.112	−0.179	0.915
	C_B	(susceptibility to covalent interaction of a base) statistical from ΔH data of different bases and acids	0.494	−0.454	−0.262	0.721
iii	BB	free energy changes in 77 reactions (data from electronic, IR, ESR, NMR spectra	0.505	0.627	−0.308	0.863
	P_s	bathochromic UV/vis spectra shifts of γ_{max} of (α-perfluoroheptyl-β,β-dicyanovinyl)aminostyrenes	0.569	0.718	−0.265	0.954
	π*	shifts of π→π* absorption band of a set of nitroaromatic compounds	0.657	0.536	−0.414	0.944
	P_y	vibronic fine structure of pyrene fluorescence spectra	0.526	0.462	−0.376	0.800
	E^*_{MLCT}	electronic spectra of metal to ligand charge transfer of W(CO)4(2,2′bipyridine)	0.606	0.530	−0.186	0.827
	E_{BB}	(susceptibility to electrostatic interaction of a base) statistical from ΔH data of different bases and acids	0.687	0.471	0.204	0.858
	SPP^N	UV/vis spectra of 2-dimethylamino-7-nitrofluorene and 2-flouro-7-nitrofluorene	0.521	0.722	−0.127	0.899
iv	B-2	acid−base hydrogen bond formation induced shifts of phenol OH group stretching frequency	0.256	0.256	0.803	0.881
	ΔH^0_{BF3}	enthalpy of complexation of solvents with BF3 in dichloromethane	0.473	0.168	0.497	0.708
	D_S	(for soft acceptors) Raman/IR stretching freq of Hg$_2$Br$_2$ in gas phase and solutions	0.069	−0.038	0.628	0.632
	DN	ΔH of reaction between SbCl$_5$ and the solvent	0.409	0.055	0.800	0.900
	D_π	second-order rate constant of the reaction of DDM and TCNE	0.325	0.061	0.558	0.649
	π^*_{aso}	bathochromic shifts of six azo merocyanine dyes	0.348	−0.056	−0.406	0.539
	$\Delta\nu_{OH}$	IR frequency shifts of phenol hydroxyl group	0.133	0.098	0.906	0.921
	B	(basicity) from stretching frequency of CH30D in different solvents	0.585	0.148	0.595	0.848
	β	various probes (longest wavelength absorption)	0.518	0.137	0.670	0.858
v	SB	UV/vis spectra of 5-nitroindoline and 1-Me-5-nitroindoline	−0.107	−0.616	−0.483	0.791
vi	P	(polarizability) expression of refractive index	−0.364	0.539	−0.437	0.784
	M	expression of refractive index	−0.312	0.521	−0.445	0.753
	μ	ΔG of the transfer of Na$^+$ and K$^+$ ions from water to solvent	−0.370	0.400	−0.079	0.551
	$\Delta\nu_{CI}$	IR stretching of iodine cyanide C−I bonds	−0.504	0.438	0.264	0.718
vii	χ_R	transition energy of merocyanine dye	−0.849	−0.260	0.216	0.914

for a general principal component analysis of solvent properties.

PCA[31,32] is a relatively straight-forward method for transforming a given set of data into principal components (PC) that are orthogonal (unrelated) to each other. In contrast to multiregression analysis (MRA), PCA requires no particular assumption about the underlying structure of the variables. Meister and Schwarz[33] studied the principal components of solvent ionicity using factor analysis and found a single principal component of ionicity, which is common to all the various operational charge definitions. Héberger and Lopata[34] performed PCA on experimental and calculated parameters of radical addition reactions to assess nucleophilicity and electrophilicity of radicals. Two principal components were extracted, accounting for electrophilic and nucleophilic properties of radicals.

The principal component model may be described by eq 1 where $\vec{x}_{ik}$ is the mean scaled value of the experimental quantities (variables) (scaling weights, w_k, transfer $\vec{x}_{ik}$ to unscaled data, $\vec{x}'_{ik} = w_k^{-1}\vec{x}_{ik}$); t_{ia} are scores; P_{ak} are loadings; e_{ik} are residuals; i is the chemical compound (object); k is the experimental measurement (variable); and a is the principal component.

$$x_{ik} = \vec{x}_{ik} + \sum_{a=1}^{A} t_{ia} P_{ak} + e_{ik} \qquad (1)$$

The number of PCs (scores) existing in a characteristic vector space is equal to, or less than, the number of variables in the data set. Each and every PC is orthogonal to all the other PCs. The first principal component is defined as that giving the largest contribution to the respective PCA of linear relationship exhibited in the data. The second component may be viewed as the second best linear combination of variables that accounts for the maximum possible of the residual variance after the effect of the first component is removed from the data. Subsequent components are defined similarly until practically all the variance in the data is exhausted.

PCA allows the examination of a set of characteristics (variables) of a class of compounds (objects) to investigate the relations between them. It enables the identification of one, two, three, or more PCs derived from the characteristics for the compounds examined. These components have defined values for each of the compounds (t_{1i}, t_{2i}, t_{3i}, the "scores") and are taken in certain proportions (p_{1k}, p_{2k}, p_{3k}, etc., the "loadings") for each type of characteristic. Graphical representations of these values, the "scores" plot for the compounds and the "loadings" plot for the characteristics, provide pictures that allow the recognition of systematic patterns that is otherwise difficult to deduce from the original data matrix.

Examples of some of the applications of PC analysis in heterocyclic chemistry include investigations of (i) aromaticity[28,29] and of (ii) the simultaneous dependence of S_N2 rates on alkyl group structure and leaving group nucleofugacity in nucleophilic displacements in which heterocycles act as leaving groups.[35,36] A multivariate statistical treatment is particularly suitable for solvent characteristics where large numbers of solvents and many scales add new dimensions to the problems generally investigated in LFERs.

Table 4. Scores of the First Three Principal Components

group	solvent	PC1	PC2	PC3
i	formamide	2.020	0.587	−0.260
ii	water	3.140	−2.685	−1.386
	methanol	1.269	−0.935	0.596
	1,2-ethanediol	1.355	−0.875	0.078
	acetic acid	0.971	−1.153	−0.151
	ethanol	0.984	−0.660	0.867
	1-propanol	0.776	−0.615	0.805
	2-propanol	0.657	−0.614	0.900
	I-butanol	0.681	−0.629	0.734
	tert-butanol	0.320	−0.445	1.008
iii	nitrobenzene	0.343	1.774	−1.451
	benzonitrile	0.111	1.249	−0.698
	nitromethane	0.623	0.965	−1.065
	acetophenone	−0.006	1.266	−0.566
	acetonitrile	0.438	0.523	−0.131
	dimethyl sulfoxide	0.684	1.801	0.424
	N,N-dimethylformamide	0.524	1.268	0.514
	N,N-dimethylacetamide	0.486	1.462	0.726
	acetone	0.101	0.483	0.376
	pyridine	0.017	1.329	0.970
	cyclohexanone	−0.087	0.796	0.140
	2-butanone	0.004	0.552	0.259
iv	THF	−0.428	0.341	0.807
	methylacetate	−0.381	−0.083	0.180
	ethyl acetate	−0.489	−0.082	0.360
	1,4-dioxane	−0.800	−0.339	0.583
	diethyl ether	−0.827	−0.429	1.055
	di-*n*-butyl ether	−1.039	−0.515	0.969
	triethylamine	−0.950	−0.340	2.924
	chloroform	−0.641	−0.610	−1.845
	1,2-dichloroethane	−0.321	0.225	−1.192
	dichloromethane	−0.320	0.037	−1.110
	bromobenzene	−0.734	0.592	−1.222
	chlorobenzene	−0.710	0.415	−0.984
	anisole	−0.611	0.469	−0.475
	benzene	−1.214	−0.240	−0.798
	toluene	−1.168	−0.299	−0.420
v	*n*-hexane	−1.723	−1.692	0.407
	cyclohexane	−1.603	−1.252	0.322
	carbon tetrachloride	−1.454	−1.643	−2.251

Our initial data set of 45 polarity scales and 65 solvents was first reduced to that set of 40 scales and 40 solvents with the most experimental values. The matrix used is given in Table 1 (Supporting Information).

$$x^n = \frac{x - \bar{x}}{\sigma^2} \cdot 10 + 50 \qquad (2)$$

The values of solvent scales were normalized and mean-centered around the value of 50 with deviations multiplied by 10 (eq 2, where x^n is the normalized value, x is the original value, and $\bar{x}$ and σ^2 stand for the mean value and standard deviation of the scale, respectively). PCA was carried out on the normalized matrix using the STATISTICA program package[38] with these 40 polarity scales as variables, each having 40 data points for the 40 solvents (objects). Table 2 lists the percentage of variance covered by the different components. The first principal component is responsible for 45% of the variance, the second for 15%, the third for 14%; these three components thus account for 74% of the total variance. The next three make up 4% to 6% each, so the total variance covered by six components is 90%. Thus, it appears that three major orthogonal components determine solvent polarity; probably some less important interactions are described by minor components (see discussion later).
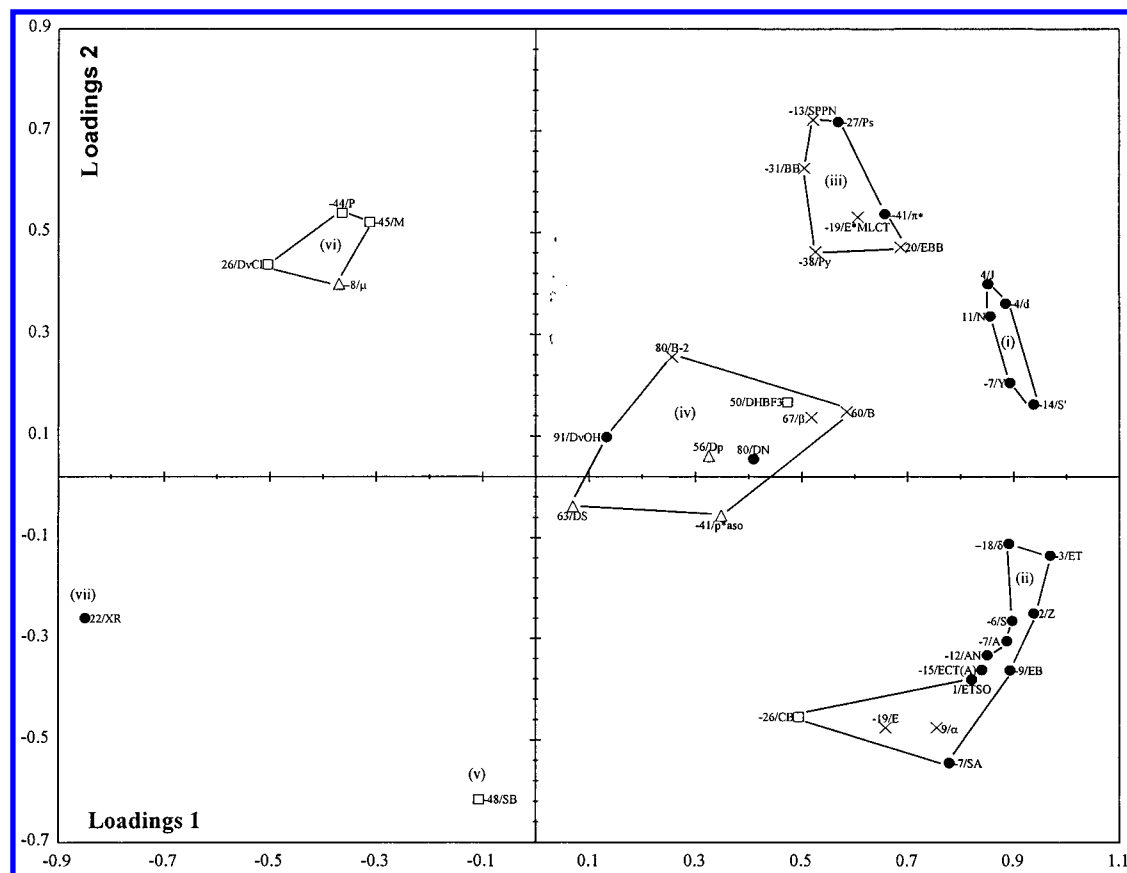
Figure 1. Loadings of the second PCA component plotted versus the loadings of the first component with the third component loading and scale classification given as labels to the data points. The ●, ×, □, and △ symbols represent $R \geq 90\%$, $80 \leq R < 90\%$, $70 \leq R < 80\%$, and $54\% \leq R < 70\%$, respectively, which reflect the goodness of fit (see text). Scales are grouped as explained in text.

Table 5. Variation of the Loadings of the Three Principal Components for Each Group of Scales[a]

group	no. of scales	type of scales	PCI		PC2		PC3	
			M	S	M	S	M	S
i	5	expression of dielectric constant	0.89	0.03	0.29	0.09	−0.02	0.09
ii	13	solvent stabilization of charge transfer in the UV/vis absorption spectral maximum of large and highly polarized conjugated systems	0.82	0.12	−0.34	0.13	−0.13	0.12
iii	7	ability of solvent to change UV absorption maxima	0.58	0.07	0.58	0.10	−0.21	0.19
iv	9	solvent basicity	0.35	0.16	0.09	0.09	0.56	0.36
v	1	UV/vis spectra of 5-nitroindoline and 1-Me-5-nitroindoline	−0.11	0	−0.62	0	−0.48	0
vi	4	solvent refractive index	−0.39	0.07	0.48	0.06	−0.17	0.29
vii	1	transition energy of merocyanine dye	−0.85	0	−0.26	0	0.22	0

[a] M and S represent the mean value and standard deviation of loadings for each group.

DISCUSSION OF SOLVENT SCALE INTER-RELATIONSHIPS

Table 3 presents the loadings of the polarity scales in the first three principal components. The total variance in each scale covered by the first three components is also given in the final column: (i) for 22 of the 40 scales, the three components describe 90% or more of their variance; (ii) another eight scales are described rather well with 80−89% of the variance; (iii) a further six scales are described with 70−79% of the variance; (iv) just four of the scales are poorly described with only 54−65% of their variance accounted for by the three main components.

Figure 1 shows the loadings[39] of the second PCA component plotted against the loadings of the first component, with the loadings of the third component as labels next to each data point. In Figure 1, the scales are distinguished as (i) well described by the first three components as solid

points; (ii) rather well described as crosses; (iii) less well described as open squares; (iv) poorly correlated by open triangles. The clustering of the scales in the space defined by the three first components suggests their classification into seven distinct groups as shown in Table 5:

(i) The five scales J, S′, d, Y, and N have large positive loadings (0.85 to 0.94) for the first component, small positive loadings (0.16 to 0.40) for the second component, and small loadings (−0.14 to 0.11) for the third component. These scales depend heavily on the solvent dielectric constant.

(ii) The 13 scales E_T, S, $E_{CT(A)}$, AN, A, E_B, Z, E, E_T^{SO}, SA, α, δ, and C_B have large positive loadings (0.49 to 0.97) for the first component, small to medium negative loadings (−0.11 to −0.55) for the second component, and relatively small loadings (−0.48 to 0.02) for the third component. Most of these scales are highly influenced by the solvent stabilization of charge transfer in the UV/vis absorption spectral
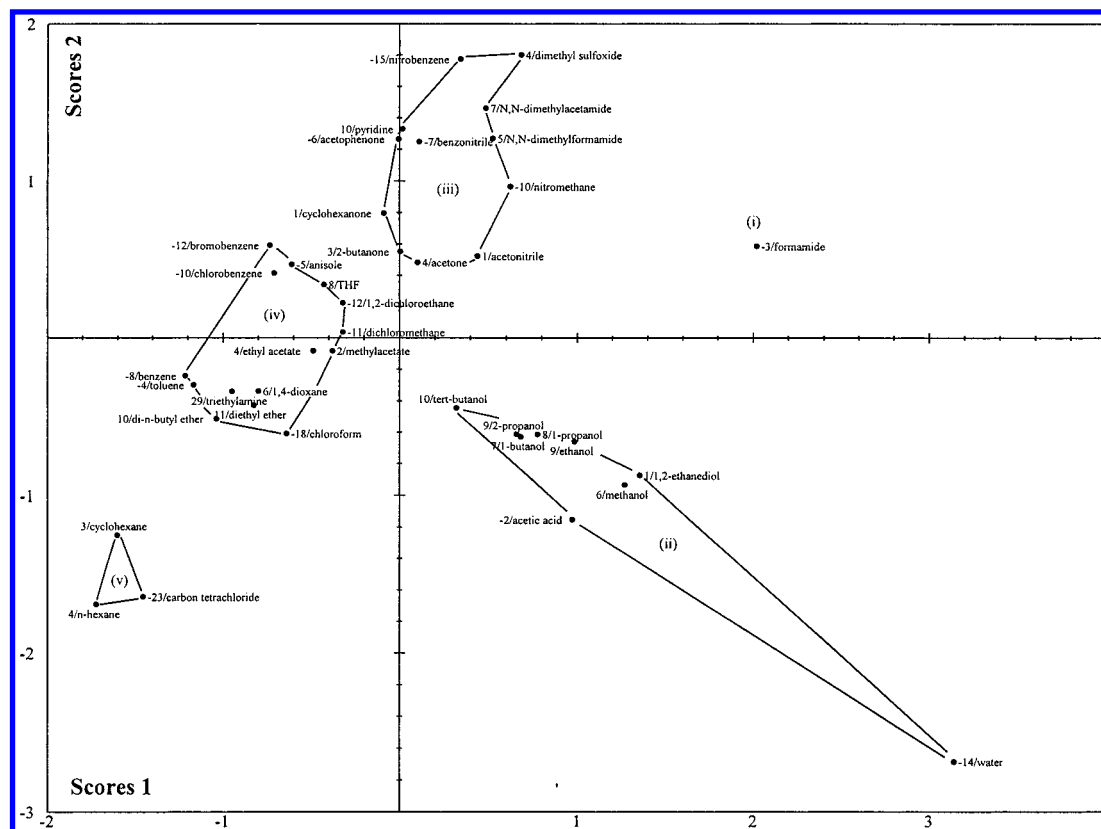
Figure 2. Plot of the scores of the second component versus the scores of the first component with the third component loading and scale classification given as labels to the data points. The point density reflects the goodness of fit (see text). Solvents are grouped as explained in text.

Table 6. Variation of the Scores of the Three Principal Components for Each Group of Solvents[a]

group	no. of solvents	type of solvents	PC1		PC2		PC3	
			M	S	M	S	M	S
i	1	formamide	2.02	0	0.59	0	−0.26	0
ii	9	hydroxylic solvents	1.13	0.77	−0.96	0.64	0.38	0.73
iii	12	dipolar aprotic	0.27	0.26	1.12	0.44	−0.04	0.72
iv	15	ethers, esters, amine, alkyl halides aromatic ring solvents	−0.71	0.28	−0.06	0.37	−0.08	1.19
v	3	*n*-hexane, cyclohexane, carbon tetrachloride	−1.59	0.11	−1.53	0.20	−0.51	1.23

a M and S represent mean value and standard deviation of scores for each group.

maximum of large and highly polarized conjugated systems. According to our findings in the previous paper, these scales account mainly for acidic and electrophilic properties of solvents.

(iii) The seven scales BB, P_S, π^*, P_y, E^*_{MLCT}, E_{BB}, and SPP^N process medium to large positive loadings for the first component (0.51 to 0.69) and the second component (0.46 to 0.72) and small loadings (−0.41 to 0.20) for the third component. Most of these scales measure the ability of the solvent to change UV absorption maxima, but unlike group ii these scales account more for nonspecific solvent dipolarity effects.

(iv) The nine scales B-2, ΔH^0_{BF3}, D_S, DN, D_π, π^*_{aso}, $\Delta\nu_{OH}$, B, and β show small to moderate positive loadings (0.07 to 0.59) for the first component, small loadings (−0.06 to 0.26) for the second component, and large positive loadings (0.50 to 0.91) for the third component (except for the loading of π^*_{aso}, which is −0.41). These scales reflect the solvent basicity, in agreement with the preceding paper.

(v) Scale SB is unique with a small negative first component and medium negative second (−0.62) and third

components (−0.48). The SB scale can be used to describe solvent basicity effects based on experimental evidence from UV/vis spectroscopy. According to our findings of the preceding paper this scale includes some specific influences.

(vi) The four scales P, M, μ, and $\Delta\nu_{Cl}$ have medium negative first components (−0.31 to −0.504), medium positive second components (0.40 to 0.54), and small to medium positive or negative third components (−0.45 to 0.26). These scales reflect the solvent refractive index (polarizability).

(vii) Scale χ_R has a large negative first component (−0.85), a small negative second component (−0.26), and a small positive third component (0.22). χ_R is based on the transition energy of merocyanine dyes.

Significantly, 34 of the 40 scales are concentrated into four major groups: i−iv. This is the same dimensionality of solvent effects as was suggested in 1972 by Koppel and Palm on the basis of their classic analysis of the experimental data.[2] Thus, our results confirm their observation. However, the six scales of groups v−vii deviate in the PCA loadings plot from all of the above-mentioned major groups i−iv. This

A Unified Treatment of Solvent Properties

J. Chem. Inf. Comput. Sci., Vol. 39, No. 4, 1999 **697**

suggests that the complex solvation phenomena may involve nonlinear inter-relations between the different mechanisms for which the previous (multi)linear treatments framework cannot account for.

DISCUSSION OF SOLVENT CLASSIFICATION

The scores of the first three principal components for the 40 solvents are presented in Table 4. Figure 2 presents the plot of the scores of the second component against the scores of the first component, with the scores of the third component as labels next to the data points. The solvents also show clear clustering into five groups (see Table 6):

(i) Formamide has a very large positive score (2.02) for the first component and medium positive score (0.59) and small negative score (−0.26) for the second and third component, respectively. It is highly polar and can form both intermolecular and intramolecular hydrogen bonds, which distinguishes it from the solvents in group iii.

(ii) The nine hydroxylic solvents have moderate to large positive scores (water 3.14, others 0.32 to 1.36) for the first component and moderate to large negative scores (water −2.69, others −0.45 to −1.15) for the second component, together with variable scores (−1.39 to 1.01) for the third component. Water is separated from the alcohols by the very high first and second component scores. The solvents of this group are all highly polar protic solvents and all form strong hydrogen bonds.

(iii) The 12 dipolar aprotic solvents, which are strong H-bond acceptors but not donors, have small negative to medium positive scores (−0.09 to 0.68) for the first component, medium to large positive scores for the second component (0.48 to 1.80), and highly variable scores (−1.45 to 0.97) for the third component.

(iv) The ethers, esters, amine, alkyl halides, and five aromatic-ring-containing solvents have small to large negative scores (−0.32 to −1.21) for the first component, moderate negative to moderate positive scores (−0.61 to 0.34) for the second component, and very variable scores (−1.85 to 2.92) for the third component. These solvents possess notably smaller polarity than the solvents in groups ii and iii.

(v) *n*-Hexane, cyclohexane, and carbon tetrachloride are distinct from other solvents with large negative scores for the first (−1.45 to −1.72) and second component (−1.25 to −1.69) and medium positive (0.32 to 0.41) or large negative scores (−2.25) for the third component. These three solvents are all apolar.

The PCA score grouping directly reflects the chemical nature of the solvents. The number of separate solvent groups is of course somewhat dependent on the total number and type of solvents included in the study. Nevertheless, solvents from our set of 40 fall into six distinct groups representing solvent classes from highly polar protic solvents (acetic acid, water) to apolar alkanes (*n*-hexane).

CONCLUSION

We believe that the present result demonstrates that PCA allows a rational classification and grouping of both solvents and solvent scales. In particular, the PCA loadings allow a logical classification of solvent scales into separate groups, and the PCA scores directly reflect the chemical nature of

the solvents. The manifold implications of these results are under active investigation.

Supporting Information Available: Table 1 gives the experimental and predicted values of the 40 scales and 40 solvents included in the PCA. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Katritzky, A. R.; Tamm, T.; Wang, Y.; Sild, S.; Karelson, M. QSPR Treatment of Solvent Scales. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 684−691.

(2) Fowler, F. W.; Katritzky, A. R.; Rutherford, R. J. D. The Correlation of Solvent Effects on Physical and Chemical Properties. *J. Chem. Soc. B,* **1971,** 460−489.

(3) Koppel, I. A.; Palm, V. A. In *Advances in Linear Free Energy Relationships;* Chapman, N. B., Shorter, J., Eds.; Plenum: London, 1972; pp 203−280.

(4) Reichardt, C. Solvatochromic Dyes as Solvent Polarity Indicators. *Chem. Rev.* **1994,** *94,* 2319−2358.

(5) Koppel, I. A.; Paju, A. I. Parameters of General Basicity of Solvents. *Org. React.* **1974,** *39,* 121−136.

(6) Palm, N.; Palm, V. Check of a Modified Approach to Solvent Effects by Processing of an Extended Set of Solvent-Dependent Data Series. 1. Test of an Earlier Defined Set of Nine Solvent Parameters. *Org. React.* **1997,** *31,* 141−158.

(7) Palm, V.; Palm, N. On the Total Number and list of Parameter Scales Significant for General Quantitative Description of Solvent Effects. *Org. React.* **1993,** *28,* 125−150.

(8) Kamlet, M. J.; Taft, R. W. The Solvatochromic Comparison Method. I. The β-Scale of Solvent Hydrogen-Bond Acceptor (HBA) Basicities. *J Am. Chem. Soc.* **1976,** *98,* 377−383.

(9) Taft, R. W.; Kamlet, M. J. The Solvatochromic Comparison Method. 2. The α-Scale of Solvent Hydrogen-Bond Donor (HBD) Acidities. *J. Am. Chem. Soc.* **1976,** *98,* 2886−2894.

(10) Kamlet, M. J.; Abboud, J. L.; Taft, R. W. The Solvatochromic Comparison Method. 6. The π* Scale of Solvent Polarities. *J. Am. Chem. Soc.* **1977,** *99,* 6027−6038.

(11) Kamlet, M. J.; Abboud, J. L.; Abraham, M. H.; Taft, R. W. J. *J. Org. Chem.* **1983,** *48,* 2877−2887.

(12) Famini, G. R.; Wilson, L. Y. Using Theoretical Descriptors in Linear Solvation Energy Relationships. *Theor. Comput. Chem.* **1994,** *1,* 213−241.

(13) Headley, A. D.; Starnes, S. D.; Wilson, L. Y.; Famini, G. R. Analysis of Solute Solvent Interactions for the Acidity of Acetic-acids by Theoretical Descriptors. *J. Org. Chem.* **1994,** *59,* 8040−8046.

(14) Cronce, D. T.; Famini, G. R.; De Soto, J. A.; Wilson, L. Y. Using Theoretical Descriptors in Quantitative Structure-Property Relationship: Some Distribution Equilibria. *J. Chem. Soc., Perkin Trans. 2* **1998,** 1293−1301.

(15) Chastrette, M.; Rajzmann, M.; Chanon, M.; Purcell, K. F. Approach to a General Classification of Solvents Using a Multivariate Statistical Treatment of Quantitative Solvent Parameters. *J. Am. Chem. Soc.* **1985,** *107,* 1−11.

(16) Drago, R. S. A Unified Scale for Understanding and Predicting Non-Specific Solvent Polarity: A Dynamic Cavity Model. *J. Chem. Soc., Perkin Trans. 2* **1992,** 1827−1838.

(17) Joerg, S.; Drago, R. S.; Adams, J. Donor-Acceptor and Polarity Parameters for Hydrogen Bonding Solvents. *J. Chem. Soc., Perkin Trans. 2* **1997,** 2431−2438.

(18) Mu, L.; Drago, R. S.; Richardson, D. E. A Model Based QSPR Analysis of the Unified Non-specific Solvent Polarity Scale. *J. Chem. Soc., Perkin Trans. 2* **1998,** 159−167.

(19) Katritzky, A. R.; Mu, L.; Karelson, M. QSPR Treatment of the Unified Non-specific Solvent Polarity Scale. *J. Chem. Inf Comput. Sci.* **1997,** *37,* 756−761.

(20) Marcus, Y. The Properties of Organic Liquids that are Relevant to their Use as Solvating Solvents. *Chem. Soc. Rev.* **1993,** 409−416.

(21) Catalán, J. On the E$_T$(30), π*, P$_y$, S′, and SPP Empirical Scales as Descriptors of Nonspecific Solvent Effects. *J. Org. Chem.* **1997,** *62,* 8231−8134.

(22) Catalán, J.; López, V.; Pérez, P.; Martín-Villamil, R.; Rodríguez, J.-G. Progress Towards a Generalized Solvent Polarity Scale: The

Solvatochromism of 2-(Dimethylamino)-7-nitrofluorene and Its Homomorph 2-Fluoro-7-nitrofluorene. *Liebigs Ann.* **1995**, 241−252.

(23) Sergent, M.; Luu, R. P. T.; Elguero, J. Statistical Analysis of Solvent Scales. Part 1. *Anales de Química Int. Ed.* **1997**, *93*, 71−75.

(24) Hawkins, D. H.; Cramer, C. J.; Truhlar, D. G. Universal Quantum Mechanical Model for Solvation Free Energies Based on Gas-Phase Geometries. *J. Phys. Chem.* **1998**, 102, 3257−3271.

(25) Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. OMNISOL: Fast prediction of Free Energies of Solvation and Partition Coefficients. *J. Org. Chem.* **1998**, *63*, 4307−4313.

(26) Giesen, D. J.; Gu, M. Z.; Cramer, C. J.; Truhlar, D. G. A Universal Organic Solvation Model. *J. Org. Chem.* **1996**, 61, 8720−8721.

(27) Giesen, D. J.; Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Modeling Free Energies of Solvation and Transfer. In *Computational Thermochemistry*; American Chemical Society: Washington, DC, 1998; pp 285−300.

(28) (a) Katritzky, A. R.; Barczynski, P.; Musumarra, G.; Pisano, D.; Szafran, M. Aromaticity as a Quantitative Concept. 1. A Statistical Demonstration of the Orthogonality of "Classical" and "Magnetic" Aromaticity in Five- and Six-Membered Heterocycles. *J. Am. Chem. Soc.* **1989**, *111*, 7−15. (b) Katritzky, A. R.; Feygelman, V.; Musumarra, G.; Barczynski, P.; Szafran, M., Aromaticity as a quantitative concept. 2. Sixteen Familiar Five- and Six-Membered Monocyclic Heterocycles. *J. Prakt. Chem.* **1990**, *332*, 853−869. (c) Katritzky, A. R.; Musumarra, G.; Barczynski, P.; Szafran, M., Aromaticity as a Quantitative Concept. 3. Benzo-fused Five- and Six-membered Heterocycles. *J. Prakt. Chem.* **1990**, 332, 870−884. (d) Katritzky, A. R.; Barczynski, P., Aromaticity as a Quantitative Concept. 4. Less Familiar Five- and Six-Membered Monocyclic Heterocycles. *J Prakt. Chem.* **1990**, *332*, 885−897.

(29) Katritzky, A. R.; Karelson, M.; Sild, S.; Krygowski, T. M.; Jug, K. Aromaticity as a Quantitative Concept. 7. Aromaticity Reaffirmed as a Multidimensional Characteristic. *J. Org. Chem.* **1998**, *63*, 5228−5231.

(30) Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *CODESSA Reference Manual Version 2.0*; Gainesville, FL, 1994.

(31) Draper, N.; Smith, M. In *Applied Regression Analysis*, 2nd ed.; Wiley-Interscience: New York, 1981.

(32) Mather, P. M. In *Computational Methods of Multivariate Analysis in Physical Geography*; Wiley: London, 1976.

(33) Meister, J. M.; Schwarz, W. H. E. Principal Component of Ionicity. *J. Phys. Chem.* **1994**, *98*, 8245−8252.

(34) Héberger, K.; Lopata, A. Assessment of Nucleophilicity and Electrophilicity of Radicals, and Polar and Enthalpy Effects on Radical Addition Reactions. *J. Org. Chem.* **1998**, *63*, 8646−8653.

(35) Sjöström, M.; Wold, W. Linear Free Energy Relationships. Local Empirical Rules-Or Fundamental Laws of Chemistry. *Acta Chem. Scand., Ser. B* **1981**, *35*, 537−554.

(36) Albano, C.; Wold, S. Multivariate Analysis of Solvolysis Kinetic Data: an Empirical Classification Paralleling Charge Delocalization in the Transition State. *J. Chem. Soc., Perkin Trans. 2* **1980**, 1447−1451.

(37) Musumarra, G.; Bruno, M.; Katritzky, A. R.; Sakizadeh, K.; Alunni, S.; Clementi, S. Kinetics and Mechanisms of Nucleophilic Displacements with Heterocycles as Leaving Groups. Part 19. Chemometric Investigation of the Simultaneous Dependence of S_N2 Rates on Alkyl Group Structure and Leaving Group Nucleofugacity. *J. Chem. Soc., Perkin Trans. 2* **1985**, 1887−1894.

(38) StatSoft, Inc. *STATISTICA* Release 5, 97 ed., 1997.

(39) The loadings are normalized: the sum of the squares of loadings for each component is the eigenvalue−percentage of variance covered by that component.