Tetrahedron Vol. 38, No. 11, pp. 1615 to 1618, 1982
Printed in Great Britain.

0040-4020/82/111615-04$03.00/0

STATISTICAL STUDY OF SOLVENT EFFECTS—II

ANALYSIS OF SOME EMPIRICAL PARAMETERS OF SOLVENT POLARITY

M. CHASTRETTE*
Laboratoire de Chimie Organique Physique, ERA 689-Université Lyon I-43 boulevard du 11 Novembre 1918-69622
Villeurbanne France

and

J. CARRETTO
Ecole Normale Supérieure, 92260 Fontenay aux Roses, France

(*Received in France* 7 *December* 1981)

Abstract—A definition of the polarity of a solvent as due only to the physicochemical properties of the pure solvent is proposed. The physicochemical properties involved are the dipole moment, the dielectric constant, the molar refraction and Hildebrand's δ. A multivariate statistical analysis shows that about 90% of the variance of E_T, AN, Z (but only 70% of the variance of Π^*) are explained by these properties. The first three parameters should be considered as genuine solvent polarity parameters.

The notion of solvent polarity is often used to choose a solvent or to explain solvent effects. With the exception of some mixtures of two solvents, solvent polarity is not conveniently measured either by the dipole moment μ or by the dielectric constant ϵ. However vey useful correlations were obtained with empirical solvent polarity parameters (for reviews see Refs. 1-3). Empirical solvent polarity parameters such as Z[4], E_T[5] or Π^*[6] are derived from a model reaction involving one or several solutes. The model reaction is implicitly or explicitly supposed to involve only the solvent polarity, with the exclusion of all sorts of specific effects.

On the other hand, numerous empirical parameters were also designed to measure the capacity of the solvent to act as an acid or a base, a donor or acceptor of electron pairs (or of hydrogen bond), such as AN, DN, α, etc...

The polarity parameters should correlate very poorly or not at all with the acidity and/or basicity parameters. However a surprisingly good correlation is found between two of the most useful parameters, namely: E_T which measures the solvent polarity and AN which measures its acidity. E_T is based on the transition energy for the longest wavelength of the absorption band of a pyridinium *N*-phenoxide betaine.[5] AN is based[7] on the solvent dependent ^{31}P NMR chemical shift of the triethylphosphine oxide.

The relation (1) is observed[8] for 29 solvents:

$$AN = 1.60\,E_T - 50.5 \quad r = 0.956 \quad n = 38. \qquad (1)$$

Mayer[9] made the following comments: "*Various classical empirical solvent parameters such as the Z, E_T or Y values, which in part have been quite successfully used in the study of solvent effects, do not represent a general measure of solvent polarity or solvent ionizing power but are obviously closely related to the solvent acceptor properties*".

Other authors[10,11] give divergent interpretations of these parameters. Taft *et al.* recently proposed[12] to reconcile the different views by showing that both AN and Z depend on the parameters Π^* and α previously proposed by these authors. Π^* is a measure of solvent polarity and α measures its acidity. The following regression equations are obtained:

$$AN = 0.16 + 16.7\Pi^* + 32.9\alpha \quad r = 0.996 \quad n = 16 \qquad (2)$$

$$Z = 51.46 + 19.4\Pi^* + 20.5\alpha \quad r = 0.998 \quad n = 14. \qquad (3)$$

This approach is interesting if we accept the hypothesis that Π^* and α entirely and only measure the polarity (for Π^*) and the acidity (for α) of the solvent. It seems to us that further progress can be made only if the polarity is more rigorously defined.

The Notion of Polarity

No generally acceptable measure of the solvent polarity was obtained when using only one physicochemical parameter such as μ or ϵ. It was then a logical step to take several physicochemical quantities into account.[10] The polarity was defined[13,14] as the sum of all the molecular properties responsible for all the solute-solvent interactions. If this definition is accepted, it follows that it should be very difficult and perhaps impossible to choose a single solute to measure that general polarity. In fact, any single solute has a fixed relative sensitivity to each of the different interaction mechanisms.

We propose here a more restrictive definition of the *polarity as the sum of the non specific solute-solvent interactions*, (i.e. depending only on the solvent). In statistical terms polarity should be a linear combination of physicochemical properties of the pure solvent. It follows from this point that the measure of polarity does not require the use of a model solute.

Rather than proposing a new polarity scale we try in this paper to: (i) examine if the usual solvent polarity empirical parameters follow that definition, (ii) find the meaning of the correlation between E_T (or Z) and AN.

METHOD

Let P be a parameter accepted as a measure of the solvent polarity and $X_1, X_2 \ldots\ldots X_n$ a set of physicochemical properties of the pure solvent. To test if P follows our definition we have to determine what part of the total variance of P is accounted for by a multiple regression on $X_1, X_2 \ldots\ldots X_n$. In other terms we establish the multiple regression equation:

$$P = a_0 + a_1X_1 + a_2X_2 + \ldots\ldots\ldots a_iX_i + \ldots\ldots a_nX_n \quad (4)$$

and calculate the correponding correlation coefficient r. If eqn (4) accounts for more than 90% of the variance of P (that is if $r > 0.95$) we can say that P follows our definition, given the errors in the determination of P. If on the contrary the eqn (4) accounts for only a small part of the total variance of P, P is not a measure of the solvent polarity, according to our definition.

However a multiple regression on the X_i is not very convenient because, these properties being more or less dependent, the signification of r can be dubious. To avoid these difficulties we used the orthogonalized regression method.[15] In this method the first step is to perform a factorial analysis[18] on the n properties X_i which leads to n abstract factors F_i, linear combinations of the X_i and perfectly independent. Then we perform a multiple regression on the factors F_i and obtain eqn (5).

$$P = b_0 + b_1F_1 + \ldots\ldots\ldots b_iF_i + \ldots\ldots\ldots b_nF_n. \quad (5)$$

Conclusions upon P are then drawn from the eqn (5) and not from the eqn (4).

The problem is now to make a correct choice of the physicochemical properties X_i to be involved in the treatment. In a previous work[16] in order to classify 22 usual solvents, one of us used the dipole moment μ, the molar refraction MR and the Kirkwood function of the dielectric constant ϵ named KIR ($KIR = (\epsilon - 1)/(2\epsilon + 1)$). Acknowledging the great role played by the cohesion energy of the solvent in the solvation phenomena we also include the parameter δ, proposed by Hildebrand,[17] as a measure of the cohesion energy density. Finally our statistical analysis is based upon four more or less correlated solvent properties.

We selected 22 solvents (the same as in our previous work) for which all the variables μ, MR, KIR, δ, AN, E_T, DN are known (Table 1). We also analysed the parameters Π^* and Z although they were unknown for some solvents (the experimental values were taken from Refs. 1-3, 5-7, 16, 17).

RESULTS AND DISCUSSION

(1) *Analysis of some empirical parameters*

(a) *Factorial analysis on the physicochemical properties.* Factorial analysis[18] of the data gives four independent factors F_i which are linear combinations of the X_i properties (Table 1). It is of some interest to know in what proportion the factors F_i, which theoretically are devoid of physical significance, depend on the X_i variables. The correlation coefficient F_i/X_i are shown in Table 2.

The factors F_1 which accounts for 51.5% of the total variance of the data is strongly correlated with KIR and fairly correlated with δ, while F_2 (39.5% of the total variance of the data) is associated with the molar refraction MR. The factor F_3 accounts for 7.2% of the variance and F_4 for 1.8%. When the hydroxylated solvents (Table 1) are not included in the calculations similar results are

Table 1. Variables X_i and parameters F_i

Solvent	Variables X_i				Parameters F_i			
	μ	MR	KIR	δ	F_1	F_2	F_3	F_4
Water	1.8	3.70	0.491	23.4	0.426	0.493	0.573	-0.171
Methanol	1.7	8.23	0.477	14.5	0.164	0.269	-0.112	-0.168
Ethanol	1.7	12.9	0.470	12.7	0.086	0.180	-0.155	-0.255
Nitromethane	3.1	12.5	0.481	12.6	0.161	0.084	-0.184	0.180
Acetonitrile	3.4	11.1	0.480	11.9	0.164	0.066	-0.273	0.343
Dimethylsulfoxyde	3.9	20.1	0.485	13.0	0.172	-0.055	0.037	0.166
Dimethylformamide	3.8	19.9	0.480	11.8	0.134	-0.069	-0,075	0.178
Acetone	2.7	16.2	0.465	9.7	0.034	0.013	-0.340	0.049
Nitrobenzene	4.0	32.9	0.479	10.6	0.045	-0.268	0.121	-0.184
Benzonitrile	3.9	31.6	0.471	10,2	0.030	-0.248	0.072	-0.125
HMPA	5.5	47.7	0.475	8.9	-0.013	-0.580	0.335	-0.136
Pyridin	2.2	24.1	0.441	10.7	-0.027	-0.021	-0.026	-0.311
Tetrahydrofuran	1.7	19.9	0.405	9.3	-0.099	0.055	-0.193	-0.108
Dioxane	0.4	21.7	0.222	9.8	-0.343	0.208	0.186	0.340
Diethylether	1.2	22.5	0.340	7.8	-0.239	0.052	-0.174	-0.008
Benzene	0.0	26.2	0.232	9.2	-0.388	0.165	0.221	0.017
Hexane	0.0	29.9	0.188	7.3	-0.498	0.100	0.204	0.143
N-Methylpyrrolidone	4.1	27.0	0.478	11.3	0.094	-0.183	0.062	0.058
Dimethylacetamide	3.8	24.2	0.480	10.8	0.086	-0.143	-0.065	0.042
1,2 Dimethoxyethane	1.7	24.1	0.400	8.3	-0.151	-0.017	-0.176	-0.222
1,2 Dichloroethane	1.7	20.9	0.429	9.9	-0.064	0.041	-0.159	-0.290
Propane -1,2 diol carbonate.	5.0	21.6	0.489	13.5	0.228	-0.139	0.117	0.463

Table 2. Correlation coefficients between factors F_i and variables X_i

F_i X_i	F_1	F_2	F_3	F_4
μ	0,642	−0,777	0,019	0,181
MR	−0,456	−0,834	0,292	−0,110
KIR	0,914	−0,306	−0,208	−0,168
δ	0,777	0,486	0,400	−0,016

obtained: F_1 accounts for 61.4% of the total variance, F_2 for 31.0% and F_3 for 6.6%.

(b) *Analysis of E_T.*† The orthogonalized regression of E_T gives the equation:

$$E_T = 4.090F_1 + 1.680F_2 + 0.705F_3 - 0.717F_4 \qquad (6)$$

The correlation coefficient (r = 0.967) shows that eqn (6) accounts for 93.5% of the variance of E_T. F_1 and F_2 account for 88.8% of the total variance of E_T (76% for F_1 alone). When the hydroxylated solvents are excluded the four factors F_i account for 95.4% of the variance of E_T. Specific effects eventually associated with the model phenolbetaine amount to a maximum of 7% which is approximately the magnitude of the experimental error on E_T.[19] *Thus E_T is indeed a measure of the solvent polarity, according to our definition.*

(c) *Analysis of Z.* Z and E_T are known to be well correlated[13] and this analysis was made only as a verification. For the 17 solvents of the sample for which Z is known, one finds the regression equation:

$$Z = 3.781F_1 + 1.839F_2 + 0.916F_3 \qquad (7)$$

Equation (7) accounts for 86.6% of the variance of Z (r = 0.931). *Thus Z is an approximate measure of the solvent polarity but not so "pure" as E_T.*

(d) *Analysis of Π*.* Equation (8) shows the dependence of Π* on the factors F_i.

$$\Pi^* = 3.662F_1 - 1.367F_2 \qquad (8)$$

Thus Π* is influenced either by specific effects or by other non specific effects. These non specific effects should depend on a physical property of the pure solvent highly independant of the four properties used in this work. *Thus Π* does not seem to be a real measure of the solvent polarity according to our definition, although it is well correlated with Z and AN.*

(e) *Analysis of AN.* The regression eqn (9) accounts for 90.5% of the variance of AN (r = 0.951).

$$AN = 3.414F_1 + 2.411F_2 + 1.056F_3 - 1.145F_4 \qquad (9)$$

Thus the part of the specific effects in the variance of AN cannot exceed 10%. *AN must be considered as a measure of the solvent polarity, according to our definition.*

(f) *Analysis of DN.* The eqn (10) is obtained:

$$DN = 2.135F_1 \qquad (10)$$

Only 20.7% of the variance of DN are accounted for the equation [10]. *DN is completely different from the preceding parameters and cannot be considered as a polarity parameter.*

(2) *Empirical parameters and physicochemical properties*

Through the factors F_i, the empirical parameters depend on the physicochemical properties X_i of the solvent. The percentages of variance of each parameter explained by each variable X_i are shown in the Table 3.

On Table 3, we can immediately see that the influence of the cohesion energy density measured by δ is very preponderant: δ accounts for about 83% of the total variance of E_T, AN and Z but for only 30% of the variance of Π*. A recent work of Larsen *et al.*[20] shows how the energy required to form a cavity in the solvent is important for the explanation of Z. Excellent correlations between δ and Z (r = 0.912 for 13 solvents) or

---

†In the eqns (6)–(10) the data concerning the empirical parameters have been standardized. All the regression coefficients are significant with a critical probability lower than 5%.

Table 3. Percentage of the variance of the empirical parameters P explained by one physicochemical variable X_i

P X_i	E_T	Z^a	AN	Π^{*b}	DN
μ	7,2	0,08	0,2	52,3	14,8
MR	40,4	46,9	44,7	0,25	0,0
KIR	44,6	25,8	25,0	51,6	22,7
δ	83,6	83,6	82,7	30,5	11,4

a) for 17 solvents

b) for 22 solvents, Π* for solvent n° 22 was calculated from AN according to ref. 12.

E_T (r = 0.920 for 21 solvents) were demonstrated earlier by Herbrandson and Neufeld.[21] The similarity between the three parameters is surprising because their model molecules differ very widely in their size, dipole moment, and polarizability. However it should be noted that the dissolution of a solute necessarily involves some energy to form a cavity whatever the solute can be.

One should expect an important role of μ in the magnitude of the polarity parameters. It can be seen that μ is of importance only for Π^* and plays a negligable role in the other parameters. This is again unexpected because the model molecules used to define E_T and Z are known[3] to have high dipole moments. The triphenylphosphine oxide used to define AN must also possess a high dipole moment since there is a relatively high negative charge on its oxygen atom:

$$O \leftarrow P\,Et_3 \longleftrightarrow \bar{O}\text{—}\overset{+}{P}\,Et_3.$$

The small differences between Z, AN and E_T are related to the influence of the dielectric constant (KIR accounts for 45% of the variance of E_T and only for about 25% of the variance of Z and AN) while the molar refraction has the same influence on the three parameters. Again Π^* is very different from the others.

CONCLUSION

A definition of the polarity as only due to physicochemical properties of the pure solvent is proposed. E_T and AN (and Z to a lesser extent) are true polarity parameters according to that definition, but Π^* is not. The multidimensional statistical analysis of three empirical parameters of solvent polarity shows that E_T, Z and AN (but not Π^*) are closely related to the parameter δ of Hildebrand. This could explain the good correlation between AN and E_T or Z. It should be noted that the choice of the solvents included in the sample has a strong influence on the correlation coefficient. Further work is needed to understand how the physicochemical properties of the solute influence the polarity scales derived from it.

REFERENCES

[1] T. R. Griffiths and D. C. Pugh, *Coordn. Chem. Rev.* **29**, 129 (1979).

[2] V. Gutmann, *Ibid* **18**, 225 (1976).

[3] C. Reichardt, *Solvent Effects in Organic Chemistry*, Verlag Chemie New-York (1979).

[4] E. M. Kosower, *J. Am. Chem. Soc.* **80** 3253 (1958).

[5] K. Dimroth, C. Reichardt, T. Siepmann and F. Bohlmann, *Ann. Chem.* **661**, 1 (1963).

[6] M.J. Kamlet and R. W. Taft, *J.Am. Chem. Soc.* **98**, 377 and 2886 (1976).

[7] U. Mayer, V. Gutmann and W. Gerger, *Monatsh. Chem.* **106**, 1235 (1975).

[8] C. Reichardt, *Empirical Parameters of Solvent Polarity and Chemical Reactivity*, in *Molecular Interactions* edited by W. J. Orville-Thomas and H. Ratajczak), Vol. 3, Wiley, Baffins Lane (1982).

[9] U. Mayer, *Pure Appl. Chem.* **51**, 1967 (1979).

[10] F.W. Fowler, A. R. Katritzky and R. J. D. Rutherford, *J. Chem. Soc.* (*B*) 460 (1971).

[11] W. R. Fawcett and T. M. Krygowski, *Aust. J. Chem.* **28**, 2115 (1975).

[12a] R. W. Taft, J. N. Pienta, M. J. Kamlet and E. M. Arnett, *J. Org. Chem.* **46**, 661 (1981).

[12b] J. L. M. Abboud, M. J. Kamlet and R. W. Taft, *Progr. Phys. Org. Chem.* **13**, 485 (1981).

[13] E. sz. Kováts, *Chimia* **22**, 459 (1968).

[14] C. Reichardt, ref. 3, p. 42.

[15] P. Dagnellie, *Analyse statistique à plusieurs variables*, Presses Agronomiques de Gembloux, 185 *et sq.* (1977).

[16] M. Chastrette, Tetrahedron 1441 (1979).

[17] J. H. Hildebrand, J. M. Prausnitz and R. L. Scott, *Regular and Related Solutions*. Van Nostrand, New-York (1970).

[18] ref. 15, 191 *et sq.*

[19] T. M. Krygowski, J. P. Radomski, A. Rzeszowiak, P. K. Wrona and C. Reichardt, *Tetrahedron* **37**, 119 (1981).

[20] J. W. Larsen, A. G. Edwards and P. Dobi, *J. Am. Chem. Soc.* **102**, 6780 (1980).

[21] H. F. Herbrandson and F. R. Neufeld, *J. Org. Chem.* **31**, 1140 (1966).