International Journal of Pharmaceutics, 93 (1993) 183–188

IJP 03117

Application of the extended Hildebrand solubility parameter treatment for optimizing reversed-phase high-performance liquid chromatography determination of pharmaceuticals

A. Gustavo Gonzalez and Agustín G. Asuero

Department of Analytical Chemistry, University of Seville, 41012 Seville (Spain)

(4 September 1992)
(Accepted 16 November 1992)

Key words: Solvent classification; Hildebrand solubility parameter; Reversed-phase HPLC; Optimization; Simplex lattice design

Summary

Retention in HPLC is described in terms of Hildebrand solubility parameters. A set of solvents commonly used in reversed-phase HPLC was classified into selectivity groups by using the extended Hildebrand solubility parameters as descriptors. Optimization rules according to this system are given for improving the selectivity of separation.

Introduction

High-performance liquid chromatography (HPLC) is widely considered to be a technique mainly for biotechnological, biomedical and biochemical research as well as for the pharmaceutical industry (Brown, 1990). For a given stationary phase, the interaction process between the solute and the mobile phase results from one or more of the following phenomena, often referred to using the term 'polarity': dipole-dipole interactions (permanent, induced or instantaneous dipole-dipole interactions) and hydrogen bond donor-acceptor abilities. The most commonly used measurements of polarity are dielectric constants ϵ and Hildebrand's solubility parameters δ_H (Barton, 1974). A recent expanded version of this theory permits one to assign type solubility parameters for each kind of intermolecular interaction (Keller et al., 1971; Snyder, 1974). Accordingly, parameters δd, δo, δa and δb are defined for any solvent. Solvents with large δd values exhibit increasing dispersion interaction with solutes which also have large δd values. Solvent and solutes which possess large δo values likewise show strong orientation dipole interactions. Strong hydrogen bond donor (HBD) solvents (large δa values) undergo strong interaction with strong hydrogen bonding acceptor (HBA) solutes (large δb values) and vice versa. An induction parameter δ_{in} must also be considered, but its effects cancel in liquid-liquid chromatographic systems.

The one-component solubility parameter δ_H has proved to be useful for regular solutions. It is still used for various purposes in non-regular

---

Correspondence to: A.G. Gonzalez, Department of Analytical Chemistry, University of Seville, 41012 Seville, Spain.

solutions, but to some extent has been superseded by multicomponent solubility parameters. It is usual to evaluate the 'total' solubilty parameter as follows (Chretien et al., 1986)

$$\delta = (\delta d^2 + \delta o^2 + 2\delta a \delta b)^{1/2} \quad (1)$$

The aim of the present paper is to present an easy way for predicting capacity factors (k'_i) and improving selectivity factors (α) by using the extended Hildebrand solubility parameters (EHSP). By using these parameters as descriptors, mobile phases may be classified according to pattern recognition techniques and selectivity groups are derived therefrom. For this purpose, 13 solvents widely used as mobile phases in reversed-phase HPLC (RP-HPLC) have been selected. Table 1 includes the solvents considered, and the values of the 'total' and extended Hildebrand solubility parameters (δ, δd, δo, δa and δb).

Theoretical Background

The fundamental quantity of interest under linear elution conditions is the distribution constant

$$K_i = [i]_S/[i]_M \quad (2)$$

of a solute i between the stationary (S) and mobile (M) phases. The distribution constant can be expressed in terms of the molar volume of solute ($\bar{V}_i$) and the solubility parameters of solute (δ_i), stationary (δ_S) and mobile (δ_M) phases according to the theory of regular solutions (Karger et al., 1973):

$$\ln K_i = (\bar{V}/RT)\left[(\delta_i - \delta_M)^2 - (\delta_i - \delta_S)^2\right] \quad (3)$$

putting $x = \delta_i - \delta_M$, $y = \delta_i - \delta_S$ and $a = RT \ln K/\bar{V}_i$ we obtain the canonical equation of a hyperbola

$$x^2 - y^2 = a^2 \quad (4)$$

Phase immiscibility generally requires $\delta_S - \delta_M > 4$ while the combination of very polar phases and non-polar phases leads to values $\delta_S - \delta_M < 17$. Therefore, the optimum working range will be defined by the equation of the hyperbola and the two inequalities

$$x - y > 4$$

$$x - y < 17$$

Relatively non-polar solutes such as hydrocarbons cannot be succesfully separated in normal phase HPLC columns (S more polar than M) because for hydrocarbons $x \sim 0$ and accordingly,

TABLE 1

Total and extended Hildebrand solubility parameters for several solvents widely used as cosolvents in RP-HPLC

Solvent	δ	δd	δo	δa	δb
(1) Ethyl acetate	7.6	7.0	3.0	0.0	2.0
(2) Tetrahydrofuran	8.6	7.6	4.0	0.0	3.0
(3) Methyl acetate	8.2	6.8	4.5	0.0	2.0
(4) Acetone	8.4	6.8	5.0	0.0	2.5
(5) Dioxane	8.8	7.8	4.0	0.0	3.0
(6) Propanol	9.5	7.2	2.5	4.0	4.0
(7) Pyridine	9.8	9.0	4.0	0.0	5.0
(8) Benzonitrile	9.8	9.2	3.5	0.0	1.5
(9) Nitromethane	10.8	7.3	8.0	0.0	1.0
(10) Ethanol	10.6	6.8	4.0	5.0	5.0
(11) Acetonitrile	10.3	6.5	8.0	0.0	2.5
(12) Dimethylsulphoxide	11.3	8.4	7.5	0.0	5.0
(13) Methanol	10.7	6.2	5.0	5.0	5.0

Values taken from Karger et al. (1973).

the condition $-17 < y < -4$ leads to a value for K_i close to zero.

$$K_i = \exp\left[-\bar{V}_i y^2/RT\right] \quad (5)$$

For $\bar{V}_i = 100$ and $T = 298$ K, the values of K_i range from 0.52 to 8.5×10^{-6}. Reversed-phase HPLC provides a solution for this problem, since here M is more polar than S, and this ensures large K_i values for non-polar samples.

Capacity factors $k_i' = K_i V_M V_S$ (V_M and V_S being the volume ratio of mobile and stationary phases) may be easily controlled by changing the liquid S and M. A reliable prediction of k_i' as a function of solvent composition requires the knowledge of the different kinds of intermolecular interactions. For a given solute and a given column (stationary phase) the k_i' value may be first adjusted into approximately the right range (2–5) by varying the polarity of M (δ_M). When at this point, selectivity factors α for sample solutes are not sufficient different from unity, we should select another mobile phase with similar polarity (δ_M) but different values of δd, δo, δa and δb (isoeluotropic series) in order to improve the selectivity ($\alpha \approx 1.5$) owing to different solute-solvent interactions.

Calculation procedure

Each selected mobile phase is featured by four descriptor variables (corresponding to the EHSP) and therefore, it may be regarded as a data vector belonging to a four-dimensional pattern space. In order to reduce the dimensionality of this pattern space a factor analysis (FA) has been carried out from a principal component analysis (PCA). PCA derives a set of orthogonal basis vectors called PCs (principal components) so that each successive PC explains the maximum amount of variance (information) in the data not accounted for by the previous PCs (Aries et al., 1991). We will attempt to describe the formulation of the observed data on the basis of a minor number of unique underlying factors selecting the PCs which mainly contribute to explain the total variance. The technique of combining two or more PCs to provide physically or chemically meaningful factors is called FA. Unfortunately, the terms 'PCA' and 'FA' are used indiscriminately in the literature. The task of interpreting the significance of these factors is not always straightforward (Sharaf et al., 1982; Malinowski, 1991). Various methods have been proposed for rotating the factors to find new ones which may be easier to interpret. In order to improve this, the technique of Varimax Rotation was used. The usual aim of this method is to make the loadings 'large' or 'small' so that most variables have a high loading on a small number of factors. The scores data plot may then used for classification purposes. One common difficulty in the analysis of multivariate data occurs when data differ greatly in scale. This difficulty is easily overcome by normalization procedures (e.g., autoscaling). After normalization, the descriptor variables are called features.

Results and Discussion

Classification of mobile phases

PCA was applied to a data matrix formed by 13 rows (samples) and four columns (descriptors) leading to four PCs. The first three PCs explain the 95.2% of the variance, so we select these PCs as underlying factors. In order to improve the interpretation of these factors we rotate them by using the Varimax technique giving the new rotated factors.

In Table 2 it can be seen at a glance that the information of rotated factors (varivectors) is compressed into fewer descriptors with a much

TABLE 2

Varimax rotated factor matrix

Descriptor/ factor	1	2	3
δd	−0.00287	0.97426	−0.10930
δo	−0.08416	−0.06758	0.98791
δa	0.72152	−0.54639	−0.29597
δb	0.97287	0.08166	−0.00919

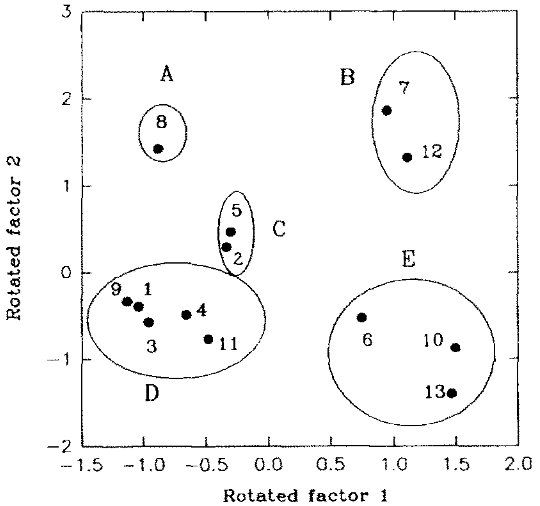

Fig. 1. Scores plot of the two first rotated factors showing the clustering of the 13 solvents.

more dramatic cutoff between large and small loadings. The first rotated factor clearly represents the δb variable, the second, δd and the third, δo. Evidently, according to the dimensionality reduction of the pattern space, the descriptor δa will be a linear combination of the other, roughly represented by the rotated factors. For classification purposes the data are projected down onto the first rotated factors as axis. This scores plot can be used to perceive obvious grouping among the solvents by analogy to cluster analysis (CA), as can be seen in Fig. 1. Accordingly, the solvents studied may be classified into five groups:

Class A: benzonitrile.
Class B: pyridine and dimethylsulfoxide.
Class C: dioxane and tetrahydrofuran.
Class D: nitromethane, methyl acetate, ethyl acetate, acetone and acetonitrile.
Class E: propanol, ethanol and methanol.

Each class shows grouping of solvents with similar solvency properties. Consequently, a small number of solvents would be sufficient for most separations. The classes A–E are well separated (according to the centroidal cluster distance) and accordingly represent the principal selectivity solvent groups in the case of RP-HPLC.

Optimization of selectivity

An important challenge in RP-HPLC is to improve the selectivity of the separation process after the optimal value of k_i' has been established. This goal may be attained by using isoeluotropic mobile phases which belong to different selectivity groups. For RP-HPLC, the polar classes C–E are available selectivity solvent groups, tetrahydrofuran (THF), acetonitrile (ACN) and methanol (MeOH) being typical representatives, respectively.

This fact agrees well with Snyder's rule, because these solvents are placed at the apices of the solvent selectivity triangle of Snyder (1974) and accordingly, great differences in chemical selectivity can be expected. It should be noted that solvent mixtures provide almost unlimited possibilities for modifying polarity, water-cosolvent mixtures being the common ones in RP-HPLC. Generally, the properties of a solvent mixture will be intermediate between those of component of the mixture. Values of the parameters δ for a solvent mixture are generally equal to the arithmetic average for each component of the mixture, weighted according to its volume fraction (Barton, 1975). For a binary mixture of water and an organic modifier we have

$$\delta_M = \delta_w + f(\delta_{OM} - \delta_M) \quad (6)$$

where δ_w, δ_{OM} and f are, respectively, the solubility parameter for water, for the organic modifier and the volume fraction of the organic modifier (cosolvent) in the mobile phase. From Eqns 1 and 6 and according to the relationship between k_i' and K, we obtain

$$\begin{aligned}\ln(k'_i) = (V_i/RT)\{\delta_w + f(\delta_{OM} - \delta_w) + \delta_s - 2\delta_i\} \\ \times\{\delta_w + f(\delta_{OM} - \delta_w) - \delta_s\} \\ + \ln(V_s/V_M) \quad (7)\end{aligned}$$

This equation may be rewritten as

$$\ln(k_i') = Af^2 + Bf + C \quad (8)$$

in which the coefficient A is expected to be positive, B large and negative (because $\delta_w > \delta_{OM}$ and $\delta_w \gg \delta_i$) and C is the natural logarithm of the capacity factor in pure water (Schoenmakers, 1988). A similar quadratic equation may be utilized for describing the retention in a ternary eluent where water is mixed with two cosolvents whose volume fractions are f_1 and f_2

$$\ln(k'_i) = A_1 f_1^2 + A_2 f_2^2 + B_1 f_1 + B_2 f_2 + C + D_{12} f_1 f_2 \qquad (9)$$

and for three cosolvents

$$\ln(k'_i) = A_1 f_1^2 + A_2 f_2^2 + A_3 f_3^2 + B_1 f_1 + B_2 f_2 + B_3 f_3 + C + D_{12} f_1 f_2 + D_{13} f_1 f_3 + D_{23} f_2 f_3 \qquad (10)$$

The first step in optimizing mobile phase composition is to adjust the polarity of the organic modifier in water for giving a binary mixture to produce a k'_i range of about 1–10 for all the solutes (Glajch et al., 1980). Once the necessary composition has been experimentally determined for say, MeOH, the composition of the isoeluotropic aqueous mixtures of ACN and THF can be predicted approximately by using the rough rules:

$$f_{OM} = (\delta_{MeoH} - \delta_w)/(\delta_{OM} - \delta_w) \qquad (11)$$

where OM refers to ACN or THF. Thus, we will have three nearly isoeluotropic aqueous mixtures, namely, A (water-MeOH), B (water-ACN) and C (water-THF). The following step is to apply Scheffe's simplex-centroid design (Akhnazarova and Kafarov, 1982) as modified by Glajch et al. (1980) called the sentinel method (Schoenmakers, 1988). In this experimental design the response surface is the chromatographic optimization function COF defined as

$$\mathrm{COF} = \sum \ln(R_i/R_d) \qquad (12)$$

It may be applied when the samples to be analyzed are not very complex. In this formula, R_i is the resolution of the adjacent pairs of peaks i and $i+1$, and R_d is the desired resolution for each pair. In the case of very complex mixtures the so-called overlapping resolution maps are powerful tools for attaining the optimal conditions (Glajch et al., 1980). For common samples, a value of COF is obtained at each chromatogram from Eqn 12 by selecting a target value for resolution, e.g., $R_s = 1.2$–1.8. The actual resolution for each pair of adjacent peaks i and $i+1$ is calculated from the chromatogram by utilizing the experimental values for the capacity factors and the average number of theoretical plates N, according to

$$R_i = N^{1/2}(k'_{i+1} - k'_i)/2(k'_i + 1 + k'_i + 2) \qquad (13)$$

So, at each mobile phase composition we obtain a chromatogram and from it, a value of COF.

When using the sentinel method for optimizing the COF, a 10-run design is applied. For the sake of D-optimality, an incomplete third degree polynomial response surface in terms of the volume fractions of initial mixtures A–C, namely, ϕ_A, ϕ_B and ϕ_C is required for fitting COF (Akhnazarova and Kafarov, 1982)

$$\mathrm{COF} = \beta_A\phi_A + \beta_B + \beta_C\phi_C + \beta_{AB}\phi_A\phi_B + \beta_{AC}\phi_A\phi_C + \beta_{BC}\phi_B\phi_C + \beta_{ABC}\phi_A\phi_B\phi_C \qquad (14)$$

Runs 1–7 are used to estimate the coefficients of Eqn 14. In experimentation following simplex-lattice designs there are no degrees of freedom to test the equation for the adequacy because the designs are saturated. Thus runs 8–10 are used to check the precision of the predicted optimum in solvent composition. The summary of mobile phase compositions to be used in the experimental design for the sentinel method is given in Table 3.

Once the coefficients have been obtained and the adequacy of the fit has been checked, the

TABLE 3

Summary of mobile phase composition for the experimental design of the sentinel method

Run	ϕ_A	ϕ_B	ϕ_C
1	100	0	0
2	0	100	0
3	0	0	100
4	50	50	0
5	50	0	50
6	0	50	50
7	33	33	33
8	67	16	16
9	16	67	16
10	16	16	67

ϕ_A, ϕ_B and ϕ_C are the volume fractions in % of the chosen eluotropic mixtures of MeOH-water, ACN-water and THF-water, respectively. Values taken from Glajch et al. (1980).

optimum coordinates of the response surface function may be easily obtained by solving the non-linear system $\delta\text{COF}/\delta\phi_i = 0$ (where $i =$ A–C) (Berenton, 1991; Morgan, 1991). At the optimum coordinates, the COF response surface should give a value close enough to the target value COF = 0.0.

Acknowledgement

Financial support from the Dirección General de Investigación Científica y Técnica de España through Project PB89-0630 is gratefully acknowledged.

References

Akhnazarova, S. and Kafarov, V., *Experiment Optimization in Chemistry and Chemical Engineering,* Mir, Moscow, 1982, pp. 245, 262.

Aries, R.E., Lidiard, D.P. and Spragg, R.A., Principal component analysis. *Chem. Br.,* 27 (1991) 821–824.

Barton, A.F., Solubility parameters. *Chem. Rev.,* 75 (1975) 731–753.

Berenton, R.G., Chemometrics as an aid to liquid chromatography optimization. *Anal. Proc.,* 28 (1991) 145–147.

Brown, P.R., High-performance liquid chromatography. Past developments, present status and future trends. *Anal. Chem.,* 62 (1990) 995A–1008A.

Chretien, J.R., Walczak, B., Morin-Allory, L., Dreux, M. and Lafosse, M., Factor analysis and experiment design in high-performance liquid chromatography. VI: Comparison of the retention mechanisms for fourteen octadecylsilica packings. *J. Chromatogr.,* 371 (1986) 253–267.

Glajch, J.L., Kirkland, J.J., Squire, K.M. and Minor, J.M., Optimization of solvent strength and selectivity for reversed phase liquid chromatography using an interactive mixture-design statistical technique. *J. Chromatogr.,* 199 (1980) 57–79.

Karger, B.L., Snyder, L.R. and Horvath, C., *An Introduction to Separation Science,* Wiley, New York, 1973, p. 269.

Keller, R.A., Karger, B.L. and Snyder, L.R., In Stock, R. and Perry, S.G. (Eds), *Gas Chromatography 1970,* Institute of Petroleum, London, 1971, p. 125.

Malinowski, E.R., *Factor Analysis in Chemistry,* 2nd Edn, Wiley, New York, 1991, p. 24.

Morgan, E., *Chemometrics: Experimental Design, Analytical Chemistry by Open Learning,* Wiley, Chichester, 1991, p. 256.

Schoenmakers, P.J., *Optimization of Chromatographic Selectivity. A Guide to Method Development,* Elsevier, Amsterdam, 1988, p. 60.

Sharaf, M.A., Illman, D.L. and Kowalski, B.R., *Chemometrics,* Wiley, New York, 1986, p. 209.

Snyder, L.R., Classification of the solvent properties of common liquids. *J. Chromatogr.,* 92 (1974) 223–230.