

## INTERPRETATION OF CHROMATOGRAPHIC CHARACTERISTICS OF MOBILE PHASE IN LIQUID CHROMATOGRAPHY BY MEANS OF SOLVENT PARAMETERS

Oldřich PYTELA<sup>a</sup>, Jaroslava HÁLOVÁ<sup>b</sup> and Miroslav LUDWIG<sup>a</sup>

<sup>a</sup> Department of Organic Chemistry,  
Institute of Chemical Technology, 532 10 Pardubice and

<sup>b</sup> Institute of Inorganic Chemistry, Czechoslovak Academy of Sciences, 250 68 Řež

Received March 14, 1990

Accepted May 8, 1990

The paper applies the methods of correlation analysis and analysis of latent variables (PCA, PLS, canonical correlation) to study the interrelations of general chromatographic characteristics in the adsorption and partition chromatography ( $\epsilon_0$ ,  $P'$ ,  $\chi_e$ ,  $\chi_d$ ,  $\chi_n$ ,  $I_{ap}$ ) and interpretation possibilities of these characteristics by means of general solvent parameters. The analyzed set of chromatographical parameters has been found to be relatively heterogeneous, its dominant properties being expressible by two principal components including 75.6% of source variability. The chemometrical solvent scale with three parameters PAC, PBC, PPC proved suitable for interpretation of the chromatographic parameters. This scale in the PLS method with two necessary latent variables explains 64.9% of the variability of matrix of chromatographic characteristics and enables their description. An elution solvent series has been constructed on the basis of this chemometrical solvent scale by the method of multicriterial decision ELECTRA III.

In spite of the great development of liquid chromatography on reversed phases, the chromatographical techniques on normal phases continue to retain their importance. Therefore it is necessary also to study the significant phenomenon of this method, viz. the properties of the mobile and/or stationary phases. In the liquid chromatography the solvent(s) used affect the selectivity, resolving power and velocity of passage of substances through the column. The manifestations of a solvent in chromatography can somewhat differ depending on the separation principle obeyed, and from this standpoint the properties of solvents play the greatest role in the adsorption and partition chromatography. Although an exact mechanism of interactions between stationary phase and solvent in adsorption chromatography still is a matter of discussions<sup>1</sup>, it is obvious that dipolarity and polarizability of solvents make themselves felt most significantly in this process<sup>2</sup>. The classification of solvents as a mobile phase in relation with successful separation was dealt with first of all by Snyder<sup>3</sup> but the same problems were also studied by other authors, e.g. Pusey<sup>4</sup>. The attempts at quantitative evaluation of behaviour of solvents in liquid-solid chromatography lead to the parameter  $\epsilon^0$  defined as the adsorption energy of solvent per unit

surface area of adsorbent ( $\text{Al}_2\text{O}_3$ ) with unit activity<sup>3,5-9</sup>. There also have been attempts to define  $\varepsilon^0$  for solvent mixtures<sup>5,10</sup>.

The properties of solvents make themselves felt similarly also in the partition chromatography. Along with solvent miscibility in liquid chromatography (mixotropic series, e.g. ref.<sup>2</sup>) also studied are the corresponding processes at molecular level, especially the specific interaction by hydrogen bonds and electrostatic interactions. In order to characterize solvents in the system of stationary/mobile phase, Snyder<sup>11</sup> proposed the so-called selectivity indexes  $\chi$  for some characteristic solvents (ethanol, dioxane, nitromethane and some others). Further chromatographic characteristics used include the polarity index  $P'$  and the so-called aromatic/paraffine index  $I_{\text{ap}}$  (ref.<sup>6</sup>).

Although the character of problem of the mobile phase in liquid chromatography is suitable for interpretation in terms of the analysis of latent variables, these techniques have not been sufficiently utilized so far. Older applications of the method of principal components are given in the book by Malinowski and Howery<sup>12</sup>. Extensive analyses of anchored phases in gas chromatography based on the literature data<sup>13</sup> were carried out by several groups of authors<sup>14-16</sup>. Another paper<sup>17</sup> deals with the analysis of interaction between stationary phase and solutes with regard to their chemical structure. Most often, however, the methods of analysis of latent variables are used in chromatography for calibration and resolution (e.g. refs<sup>18,19</sup>). For practical applications to the liquid chromatography it is also useful to know the similarities in behaviour of individual solvents, i.e. their classification (a survey is given e.g. in refs<sup>9,20</sup>).

The aim of the present communication is to adopt the methods of analysis of latent variables and other suitable methods to analyze the relation between chromatographic characteristics of mobile phase and general solvent characteristics inclusive of a possibility to suggest an elution solvent series (eluotropic sequence).

## COMPUTATIONAL

The empirical chemometric parameters PAC, PBC, and PPC (obtained by a treatment<sup>21</sup> of a large set of experimental data) have been selected for the general characterization of solvents. These parameters were complemented by experimental data characterizing the volume of molecule (molar volume  $v_m$ , ref.<sup>22</sup>) and by theoretical characteristics of solvents reflecting molecular shape and size (molecular connectivity  $^n\chi$ , refs<sup>23,24</sup>, molecular mass  $M$ ). Out of chromatographic characteristics we selected the elution power  $\varepsilon^0$  (refs<sup>2,6</sup>), the selectivity indexes  $\chi_c$ ,  $\chi_d$ ,  $\chi_n$  (ref.<sup>6</sup>) and further characteristics  $P'$  and  $I_{\text{ap}}$  (ref.<sup>6</sup>). Because of inaccessibility of some characteristics mentioned we reduced the original data set of 51 solvents<sup>21</sup> by excluding 6 solvents (heptane, *m*-xylene, mesitylene, methyl acetate, acetanhydride, and sulfo-

lane). The calculations by the method of principal components<sup>12</sup> (PCA) and projection of latent structures (PLS) were carried out on a PC-AT computer using an algorithm by Wold<sup>25</sup>, the calculation of canonical correlation was carried out according to our own program<sup>26</sup> using an EC 1033 computer, the calculation by the technique of multicriterial decision was carried out according to the program ELECTRA III (ref.<sup>27</sup>) using an EC 1040.

## RESULTS AND DISCUSSION

*Analysis of relation of characteristics of adsorption and partition chromatography.* The correlation analysis shows an only weak bond of the individual characteristics. The highest correlation coefficient ( $r = -0.758$ ) was found between  $\chi_c$  and  $\chi_d$ , the reciprocal dependence being noteworthy. The second greatest bond ( $r = 0.685$ ) is exhibited by two characteristics of different separation principles, viz.  $P'$  and  $\varepsilon^0$ . The application of PCA method led to two principal components according to the cross-validation criterion<sup>15</sup>. The first principal component involved 48.6% and the second one further 27.0% of variability of the source matrix. The main share in the formation of the first principal component possesses the  $\chi_c$  parameter and partially  $\varepsilon^0$ . After accepting the second principal component, the shares of the individual chromatographic characteristics in the interpretation of latent variables became more equal, however, the parameter  $\varepsilon^0$  remained dominant and the  $I_{ap}$  parameter showed a minor contribution. If the loadings corresponding to the first two principal components are plotted against each other, then the parameters  $\chi_c$ ,  $\chi_n$ , and  $\chi_d$  in this system form — almost geometrically precisely — three corners of a square. The last corner is roughly formed by the parameters  $P'$  and  $\varepsilon^0$  which are located relatively close to each other at the opposite side with regard to the  $\chi_n$  parameter. The  $I_{ap}$  parameter roughly halves the distance between  $\chi_d$  and  $\chi_n$ . From the analysis given it follows that the analyzed chromatographic characteristics have relatively few properties in common, the latent variables obtained being not interpretable from physical standpoint. The individual decomposition of selectivity indexes is surprising, on the other hand we must stress the similarity from the standpoint of the source of the different characteristics  $P'$  and  $\varepsilon^0$ .

*Analysis of relation between chromatographic and general characteristics of solvents.* This analysis was carried out by two methods — calculation of canonical correlation and the PLS method. In both the methods there was a group of explaining variables (the  $\mathbf{X}$  matrix in PLS; it included PAC, PBC, PPC,  $v_m$ ,  $^n\chi$ , and  $M$ ) and explained variables (the  $\mathbf{Y}$  matrix in PLS; it included 6 discussed chromatographic characteristics). In the canonical correlation we found a relatively close bond between both groups which was expressed by the group coefficient of canonical correlation  $R = 0.998$ . Gradual elimination of (always one) explaining quantity with the least contribution into the group canonical correlation coefficient led gra-

dually to elimination of molar volume, molecular mass, and molecular connectivity. The group canonical correlation coefficient decreased but slightly – to the value of  $R = 0.996$ . Out of the remaining parameters (PAC, PBC, PPC) the PPC parameter showed the least contribution to the interpretation of variability of the matrix of chromatographic characteristics, however, its elimination leads to a significant decrease in the group canonical correlation coefficient to the value of  $R = 0.986$ . The given general solvent parameters thus reflect the properties of the explained matrix  $Y$  relatively well. A more detailed analysis shows that the first canonical correlation (the canonical correlation coefficient  $r_1 = 0.971$ ) predominantly represents a bond between  $P'$  and the PAC, PBC parameters. The second canonical correlation ( $r_2 = 0.849$ ) reflects a bond between  $P'$  (and – with negative sign –  $\chi_e$ ) and PPC, partially also (with negative sign) a bond to PAC and PBC. The third canonical correlation ( $r_3 = 0.714$ ) already is relatively insignificant and can be interpreted only with difficulties. From the analysis it follows that there is a significant relation between the  $P'$  characteristic and the solvent parameters, on the other hand, there is a little distinct dependence for the selectivity indexes except for  $\chi_e$ . An also weak bond exists between the elution power and general solvent parameters, which can be due to the similarity of the parameters  $\varepsilon^0$  and  $P'$ .

The calculation by the PLS method was started similarly as in the canonical correlation with all the selected general solvent characteristics. On the basis of the cross-validation criterion<sup>15</sup> it was found that two latent variables are sufficient for a significant description of bond between general and chromatographic characteristics of solvents. The first latent variable describes 60.5% of the variability of the explaining matrix  $X$  and 26.1% of the variability of the explained matrix  $Y$ . The second latent variable describes further 11.2% of variability of the  $X$  matrix, but, at the same time, also further 30.0% of the variability of the  $Y$  matrix. In the next phase we excluded – on the basis of the modelling power – from the calculation the molecular mass, molecular connectivity, and at last also the molar volume. After this operation, the first latent variable explained 71.9% and the second latent variable further 11.9% of the variability of  $X$  matrix, i.e. total 83.8%. Similarly for the  $Y$  matrix the first and the second latent variables explained 29.4 and 34.5%, respectively, of the variability, i.e. total 64.9% of variability. Further elimination of parameters in the explaining matrix  $Y$  led (similarly as with the canonical correlations) to a significant decrease of the explained variability of  $X$  matrix. In the PLS model the remaining parameters PAC, PBC, and PPC make roughly the same contributions to the formation of the first score ( $t_1$ ) of  $X$  matrix and contribute especially to the description of the variability of  $P'$  and  $\varepsilon^0$  from the  $Y$  matrix. The main contribution to the formation of the second score ( $t_2$ ) of  $X$  matrix was predominantly that of the PPC parameter which contributed to interpretation of  $\chi_e$  and less to  $\chi_d$  and  $\chi_n$ . Hence it can be stated that the characteristics  $P'$ ,  $\varepsilon^0$ , and partially  $\chi_e$  are best interpreted by the general solvent parameters used. Except for  $\varepsilon^0$ , a similar

conclusion results also from the canonical correlation, however, the results of PLS method are more unequivocal. From the analysis it follows that the properties expressed by general solvent parameters are reflected first of all in the chromatographic characteristics  $P'$  and  $\varepsilon^0$ . At the same time we consider it important that each quantity is determined from a different chromatographical principle. The other analyzed chromatographic characteristics reflect either only an exceptional property of solvent or properties not connected with solvents as such. The selected chromatographic characteristics are also individually well interpreted by the PAC, PBC, and PPC parameters as it can be seen from the following regressions:

$$P' = - (0.16 \pm 0.26) + (1.18 \pm 0.48) \text{ PAC} + (4.94 \pm 0.48) \text{ PBC} + \\ + (2.82 \pm 0.58) \text{ PPC} , \\ n = 45 , \quad s = 0.366 , \quad R = 0.955 .$$

$$\varepsilon^0 = (8.69 \pm 3.06) \cdot 10^{-2} + (6.66 \pm 0.79) \cdot 10^{-1} \text{ PAC} + \\ + (6.07 \pm 0.72) \cdot 10^{-2} \text{ PBC} \\ n = 45 , \quad s = 5.76 \cdot 10^{-3} , \quad R = 0.966 .$$

As the PAC, PBC, and PPC parameters are standardized, the regression coefficients reflect the measure of contribution of the corresponding property of solvent. The  $P'$  characteristics is the most sensitive to the solvent polarity (since in the sum there appear the PBC and PPC parameters, see ref.<sup>21</sup>), being less sensitive to specific interactions, i.e. solvent acidity and basicity. On the other hand, with the  $\varepsilon_0$  characteristic these specific interactions predominate due probably to the interactions with the adsorbent surface by hydrogen bonds.

*Construction of eluotropic sequence.* The classification of solvents by the method of shortest nonclosed pathway in our previous paper<sup>20</sup> can be considered the first attempt at the construction of eluotropic sequence based on the PAC, PBC, and PPC parameters<sup>21</sup>. With regard to the distribution of solvents in the parameters space this is relatively difficult. A more suitable approach is represented by the method of multicriterial decision (using ELECTRA III, ref.<sup>27</sup>) which enables – on the basis of a chosen strategy – an estimation of sequence, in our case a sequence of solvents. Applying the ELECTRA III we used two strategies – with unit weights of contributions of the PAC, PBC, and PPC parameters and with the weights given by average relative sensitivity to the individual effects<sup>21</sup>. The calculations were carried out for 51 solvents given in the paper cited. The results in both the strategies are not much different, which documents the value of the Spearman coefficient of sequence correlation  $r = 0.998$ . However, an indifferent strategy is characteristic by its  $1.5 \times$  greater number of indifferent classes, hence this strategy has a lower

information content. Since also from physico-chemical aspect the inclusion of the weights reflecting the relative sensitivity to parameters is suitable, we will give only the sequence obtained by the nonindifferent strategy: (1) water, (2–3) 1,2-ethanediol, dimethyl sulfoxide, (4) formamide, (5) benzyl alcohol, (6) sulfolane, (7) acetic acid, (8–9) dimethylformamide, methanol, (10) 2-methoxyethanol, (11–12) acetonitrile, nitromethane, (13) dimethylacetamide, (14) acetanhydride, (15) ethanol, (16) 2-propanol, (17) benzonitrile, (18–19) 1-butanol, pyridine, (20) hexamethylphosphoramide, (21) 1,2-dichloroethane, (22–23) acetone, nitrobenzene, (24) dichloromethane, (25–26) chloroform, butanone, (27) 2-methyl-2-propanol, (28) cyclohexanone, (29) anisole, (30) methyl acetate, (31) dioxane, (32–33) chlorobenzene, phenetole, (34) tetrahydrofurane, (35) bromobenzene, (36) fluorobenzene, (37) ethyl acetate, (38) benzene, (39–40) diethyl ether, diisopropyl ether, (41) toluene, (42) dibutyl ether, (43–44) *p*-xylene, tetrachloromethane, (45) carbon disulphide, (46–47) *m*-xylene, triethylamine, (48) cyclohexane, (49) hexane, (50) mesitylene, (51) heptane. An altogether good prediction ability of the given sequence for selected chromatographic parameters is documented by the Spearman coefficient of sequence correlation which has the value of  $r = 0.913$  for  $P'$  (indifferent strategy,  $r = 0.919$ ) and  $r = 0.886$  for  $\varepsilon^0$  (indifferent strategy,  $r = 0.877$ ). If the anomalously behaving triethylamine is excluded from the correlation, the correlation coefficient for  $\varepsilon^0$  will increase to the value of  $r = 0.924$  ( $r = 0.922$ ). The given values of correlation coefficients represent sequence differences smaller than one position for each pair compared, hence the compared sequences are in good accordance.

#### REFERENCES

1. Snyder L. R., Poppe H.: *J. Chromatogr.* 184, 363 (1980).
2. Reichardt C.: *Solvents and Solvent Effects in Organic Chemistry*, 2<sup>nd</sup> ed., Appendix A. 6; A.7. Verlag Chemie, Weinheim 1988.
3. Snyder L. R., Kirkland J. J.: *Introduction to Modern Liquid Chromatography*, Chap. 9, p. 365. Wiley, New York 1969.
4. Pusey D. F. G.: *Chem. Brit.* 5, 408 (1969).
5. Snyder L. R.: *Principles of Adsorption Chromatography*. Dekker, New York 1968.
6. Snyder L. R., Kirkland J. J.: *Introduction to Modern Liquid Chromatography*, 2<sup>nd</sup> ed. Wiley, New York 1974.
7. Snyder L. R.: in: *Techniques of Chemistry* (E. S. Perry and A. Weissberger, Eds), 3<sup>rd</sup> ed., XII, p. 25 ff. Wiley, New York 1978.
8. Snyder L. R. in: *High-Performance Liquid Chromatography. Advances and Perspectives* (Cs. Horváth, Ed.), Vol. 3, p. 157. Academic Press, New York 1983.
9. Reichardt C.: *Solvents and Solvent Effects in Organic Chemistry*, 2<sup>nd</sup> ed. Verlag Chemie, Weinheim 1988.
10. Rice P. D., Bobbitt D. R.: *J. Chromatogr.* 437, 3 (198).
11. Snyder L. R.: *J. Chromatogr. Sci* 16, 223 (1978).
12. Malinowski E. R., Howery D. G.: *Factor Analysis in Chemistry*. Wiley, New York 1980.
13. McReynolds W. O.: *J. Chromatogr. Sci.* 8, 685 (1970).

14. Wold S., Andersson K.: *J. Chromatogr.* **80**, 43 (1973).
15. Wold S.: *Technometrics* **20**, 397 (1978).
16. Pytela O.: *Collect. Czech. Chem. Commun.* **55**, 42 (1990).
17. Howery D. G., Soroka J. M.: *J. Chemometrics* **1**, 91 (1987).
18. Vandeginste B. G. M., Leyter F., Gerritsen M., Noor J. W., Kateman G., Frank J.: *J. Chemometrics* **1**, 57 (1987).
19. Crilly P. B.: *J. Chemometrics* **1**, 79 (1987).
20. Pytela O.: *Collect. Czech. Chem. Commun.* **55**, 644 (1990).
21. Pytela O.: *Collect. Czech. Chem. Commun.* **55**, 634 (1990).
22. Rohrschneider L.: *Anal. Chem.* **45**, 1241 (1973).
23. Kier L. B., Hall L. H.: *Eur. J. Med. Chem.* **12**, 307 (1977).
24. Svoboda P., Pytela O., Večeřa M.: *Collect. Czech. Chem. Commun.* **48**, 3287 (1983).
25. Wold S., Geladi P., Esbensen K., Öhman J.: *J. Chemometrics* **1**, 41 (1987).
26. Pytela O., Ludwig M., Svoboda P.: *Sb. Ved. Pr. Vys. Sk. Chemickotechnol., Pardubice* **48**, 55 (1985).
27. Roy B.: *ELECTRE III. Un algorithme du Classements fonde sur une représentation floue des préférences en présence de critères multiples*. Cahiers du Centre d'Etudes de Recherche Opérationnelle No. 20, Brusel 1978.

Translated by J. Panchartek.