Journal of Chromatography, 550 (1991) 397–410
Elsevier Science Publishers B.V., Amsterdam

CHROMSYMP. 2347

Multivariate characterization of solvent strength and solvent selectivity in reversed-phase high-performance liquid chromatography

P. M. J. COENEGRACHT*, A. K. SMILDE, H. BENAK, C. H. P. BRUINS, H. J. METTING, H. DE VRIES and D. A. DOORNBOS

Chemometrics Research Group, University Centre for Pharmacy, A. Deusinglaan 2, NL-9713 AW Groningen (Netherlands)

ABSTRACT

Principal component analysis was used to determine the dimensionality and structure of three data sets consisting of the capacity factors of eleven to twenty different solutes measured in nine different mobile phase compositions consisting of water and methanol and/or acetonitrile on three reversed-phase columns. Principal component analysis showed that two principal components could account for the total variance in the data and that the percentage variance explained by the first principal component (about 80–95%) was much greater than the percentage explained by the second principal component, but that the percentage depended strongly on the choice of solutes for the sample. The first principal component could be associated with solvent strength and solvent strength selectivity and the second principal component with modifier selectivity. Solutes that showed strong modifier selectivity could be distinguished from solutes that have almost zero modifier selectivity, which could be useful for the definition of an empirical solvent strength scale.

INTRODUCTION

Solvent strength and selectivity

In reversed-phase (RP) chromatography strong interactions occur between the polar mobile phase and the molecules of the sample. The mobile phase consists of water and one to three organic solvents or modifiers. Solvent properties have been classified by Snyder [1] in terms of solvent polarity and selective interactions.

The eluting power of the mobile phase in RP chromatography depends on the strength of the pure organic solvent, which is related to the coefficient S by the following equation:

$$\log k' = \log k_w - S\varphi \quad (1)$$

Here k_w refers to the isocratic capacity factor (k') of the solute for pure water as the mobile phase and φ is the volume fraction of organic modifier in the binary mobile phase [2]. S is a constant which is not only characteristic of a given modifier, but

0021-9673/91/$03.50

depends also on the molecular size and structure of the solute [3]. Eqn. 1 is applicable for a limited range of k' values ($1 < k' < 10$), and S values also depend on the RP packing material [4]. For typical samples consisting of benzene derivatives, the average variation of S was considered small enough to assign an unique solvent strength value to a given solvent [4]. This approach makes it possible to calculate the composition of iso-eluotropic mobile phases consisting of water and different organic modifiers using transfer rules [5,6]. These rules allow the substitution of one organic solvent for another, whereas the sample k' values remain roughly constant.

This substitution changes the selectivity of the mobile phase and exploits the ability of a solvent to exhibit different selective interactions with different solutes to separate solutes of a similar polarity [7].

"The strength of a mobile phase is a major factor in controlling the retention and is a function of its quantitative composition, *i.e.* the water-to-modifier ratio, but the selectivity depends on its qualitative composition, *i.e.* the type of modifier" [8]. Descriptions of this kind and statements that it is possible to change the ratio of the carrier (water) to organic modifier with the net result that the strength of the mobile phase changes while the selectivity remains constant [5], treat the solvent strength and solvent selectivity as two independent properties of the mobile phase. Changing the solvent strength would alter the capacity factors of all the solutes of a sample but not effect the ratios of the capacity factors of the solutes so that the relative peak positions in the chromatogram would remain the same. This would mean that strength is a solvent property that has an equally strong proportional effect on all solutes and is therefore almost independent of the individual solute properties, while selectivity depends on solvent–solute interactions that also depend on the properties of the solute.

It has been shown, however, that if the water fraction of a binary, ternary or quaternary mobile phase changes, not only the strength, but also the selectivity changes [9]. The simultaneous variation of selectivity and strength is the basis of the solvent strength selectivity optimization of binary systems according to Snyder *et al.* [10] and was realized in a ternary solvent system by the combination of a statistical mixture design technique and multicriteria decision making [11].

The solvent strength concept stresses that the fraction of water in an eluent determines the range of the capacity factors of all the solutes of a sample. The solvent selectivity alters to a lesser degree the capacity factors of the individual solutes and does not significantly change the range of the capacity factors.

The empirical transfer rules for calculating the iso-eluotropic mobile phase compositions are based on retention data collected for a large number of solutes at different mobile phase compositions. The transfer rules are found by regression analysis [12] and represent an averaged solute retention behaviour. The selectivity refers to the deviation of a given solute from the average retention behaviour. This means that a good transfer rule should depend on a representative sample, *i.e.* a sample consisting of solutes that contribute equally to different selective interactions.

This paper describes the use of principal component analysis (PCA) for the determination of the number of uncorrelated factors (dimensions) that account for the total variation in the retention data of solutes in different mobile phase compositions. In addition, the contribution of each factor to the total variation is estimated and this is related to the solute composition of the sample and the selective interac-

tions of each solute. This allows an approximately quantitative estimate of the effect of solvent strength and modifier selectivity and indicated a method for the choice of solutes which are a good measure of "pure" solvent strength.

Principal component analysis

PCA constructs linear combinations of the original variables that account for as much of the total variation in the data as possible. The successive linear combinations are not correlated with each other and account for successively smaller amounts of variation. The principal components are the basic dimensions of the data necessary to define their total variance [13].

Mobile phases composed of water and methanol (MeOH) and/or acetonitrile (ACN) with different ratios of water and modifier(s) are characterized chromatographically by differences in solvent strength and selectivity. These differences can be measured by the capacity factors of the sample solutes, *i.e.*, the mobile phases are the objects and the solutes are the variables of the data matrix.

If solvent strength and solvent selectivity are different properties of the objects, then PCA should reveal at least two dimensions in the data. That is, PC1 is that linear combination of the observed variables $X_j, j = 1, 2, ..., p$

$$PC1 = wl_1X_1 + wl_2X_2 + ...wl_pX_p \quad (2)$$

where the loadings wl_1, wl_2, $...wl_p$ have been chosen to maximize the ratio of the variance of PC1 to the total variation. Component loading wl_i is a measure of the contribution of the ith variable to PC1; it measures the contribution of the ith solute to the variance explained by PC1.

The second principal component, PC2, is that weighted linear combination of the variables which is not correlated to the first PC1 and which accounts for the maximum amount of the remaining total variation. The goal of PCA is to account for most of the total variation with as few principal components as possible [13].

The principal components are statistical descriptors and do not represent physical properties, but they can be used to test hypotheses about qualitative distinctions in the data. PCA was used for the classification and selectivity characterization of different RP high-performance liquid chromatography (HPLC) packings [14–17], but not for the characterization of the strength and selectivity of RP mobile phases.

EXPERIMENTAL

General set-up

Three data matrices were investigated. The choice of the mobile phase compositions was made according to a constrained mixture design. The design aims at a regular spread of the design points in a constrained area of design space [18]. The first data matrix consisted of nine mobile phases. The solvent strength was varied at three levels and at each level two different binary and one ternary mobile phase were composed of water, MeOH and /or ACN (Fig. 1). Fifteen solutes were used to characterize the nine objects (Table I).

The second data matrix consisted of eight mobile phases. The solvent strength

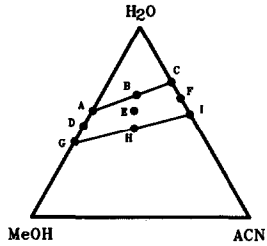

Fig. 1. Design points corresponding to the mobile phase compositions A–I of Table I.

was varied at two levels (Fig. 2). At each level the solvent strength was kept constant by experimental adjustment of the eluent composition so that the variation of the mean capacity factor of the sample solutes was smaller than 5%. At each level two different binary and two different ternary mobile phases were composed of water, MeOH and ACN (Table II). The eleven solutes used were essentially a subset of the previous sample, but an octadecylsilane (ODS) column of a different brand was used. The third data matrix was published by Weyland [19] and consisted of four binary eluents of water and MeOH, three binary eluents of water and ACN and two ternary

TABLE I

MOBILE PHASE COMPOSITIONS AT THE DESIGN POINTS A–I AND MEASURED CAPACITY FACTORS OF THE SOLUTES

Mobile phase component	A	B	C	D	E	F	G	H	I
Water	0.550	0.625	0.700	0.467	0.541	0.617	0.383	0.459	0.533
Methanol	0.450	0.225	0.000	0.533	0.267	0.000	0.617	0.308	0.000
Acetonitrile	0.000	0.150	0.300	0.000	0.192	0.383	0.000	0.233	0.467
Solute	Capacity factor								
Acetophenone (ACP)	2.747	3.960	4.030	1.534	2.237	2.437	0.910	1.262	1.692
Acetanilide (ACT)	1.145	1.450	1.111	0.683	0.871	0.729	0.435	0.525	0.571
Anisole (ANS)	5.912	8.147	9.606	3.316	4.722	5.406	1.970	2.515	3.373
p-Cresol (CRE)	2.815	3.950	3.424	1.574	2.178	2.000	0.920	1.202	1.373
Ethylaminobenzoate (EAB)	2.640	4.245	4.121	1.356	2.158	2.229	0.702	1.101	1.439
Nitrobenzene (NBZ)	3.650	5.999	7.444	2.118	3.425	4.218	1.267	1.868	2.681
Toluene (TOL)	14.213	19.735	21.909	7.683	10.237	11.218	4.198	5.090	6.395
2-Phenylethanol (PE)	2.368	2.754	1.848	1.346	1.534	1.156	0.782	0.888	0.868
Propylhydroxybenzoate (PHB)	11.038	15.147	10.090	4.722	6.445	4.343	2.217	2.818	2.395
Ethylhydroxybenzoate (EHB)	4.689	6.254	4.454	2.227	2.990	2.250	1.148	1.464	1.406
Methylhydroxybenzoate (MHB)	2.126	2.764	2.151	1.108	1.465	1.239	0.623	0.797	0.868
Dimethylphthalate (DMP)	3.310	5.745	5.414	1.603	2.811	2.947	0.851	1.454	1.890
Phenobarbital (PBL)	1.980	2.960	1.747	1.009	1.495	1.072	0.544	0.757	0.747
Prednisone (PRE)	5.592	7.058	2.242	2.158	2.653	0.947	0.960	1.181	0.593
Prednisolone (PRS)	8.213	8.343	2.060	3.227	3.217	0.812	1.465	1.414	0.505
Mean	4.830	6.568	5.444	2.378	3.230	2.867	1.267	1.623	1.787

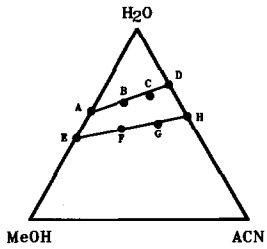

Fig. 2. Design points corresponding to the mobile phase compositions A–H of Table II.

eluents with the same modifiers. The retention times of the sample solutes, twenty sulphonamides (*p*-aminobenzoic acid analogues, Table III), were measured on a 15.0 cm × 4.6 mm stainless-steel column packed with Nucleosil C_8, particle size 5 μm.

Measurements were collected as k' values and transformed to ln k' values to obtain a constant variance of the data. The ln k' data matrix was transformed to a covariance matrix prior to PCA. Standardization was not performed because the variables are measured in the same ln k' units.

Instrumentation and chemicals

The experiments for the first data set were performed on an HPLC apparatus consisting of an automatic sampler (KA 9209, sample loop 20 μl), a solvent delivery system (KA 9208), a fixed-wavelength UV detector (KA 9202, 254 nm) and a Model B40 Kipp recorder. Stainless-steel columns, 15 cm × 4.6 mm I.D., slurry-packed with ODS Hypersil, 5 μm particle diameter, number of plates (N) = 8000 (fluoranthene in

TABLE II

MOBILE PHASE COMPOSITIONS AT THE DESIGN POINTS A–H AND MEASURED CAPACITY FACTORS OF THE SOLUTES

Mobile phase component	A	B	C	D	E	F	G	H
Water	0.560	0.590	0.640	0.700	0.420	0.460	0.490	0.540
Methanol	0.440	0.310	0.150	0.000	0.580	0.380	0.200	0.000
Acetonitrile	0.000	0.100	0.210	0.300	0.000	0.160	0.310	0.460
Solute	Capacity factor							
Acetophenone (ACP)	2.431	2.730	2.984	3.262	0.983	1.031	1.186	1.316
Acetanilide (ACT)	0.891	1.065	0.969	0.958	0.454	0.442	0.447	0.443
Aniline (AN)	1.431	1.425	1.328	1.847	0.593	0.578	0.707	0.860
p-Cresol (CRE)	2.492	2.835	2.738	2.836	1.000	1.094	1.103	1.053
Ethylaminobenzoate (EAB)	1.977	2.708	3.008	3.258	0.731	0.884	1.009	1.070
Nitrobenzene (NBZ)	3.023	3.957	4.953	6.067	1.403	1.558	1.825	2.035
Toluene (TOL)	13.538	14.357	15.308	18.607	5.085	4.800	5.190	5.193
2-Phenylethanol (PE)	2.094	2.108	1.828	1.574	0.850	0.813	0.772	0.667
Propylhydroxybenzoate (PHB)	9.453	10.258	9.297	7.656	2.417	2.620	2.368	1.789
Phenobarbital (PBL)	1.662	1.966	1.708	1.459	0.567	0.625	0.627	0.526
Prednisolone (PRS)	7.369	6.942	3.531	1.557	1.600	1.516	0.847	0.386
Mean	4.353	4.683	4.400	4.462	1.457	1.470	1.505	1.481

TABLE III

MOBILE PHASE COMPOSITIONS AT THE DESIGN POINTS A–I AND MEASURED CAPACITY FACTORS OF THE SOLUTES

Mobile phase component	A	B	C	D	E	F	G	H	I
Water	0.80	0.75	0.70	0.60	0.90	0.80	0.70	0.70	0.80
Methanol	0.20	0.25	0.30	0.40	0.00	0.00	0.00	0.15	0.10
Acetonitrile	0.00	0.00	0.00	0.00	0.10	0.20	0.30	0.15	0.10
Solute	Capacity factor								
Sulphanilamide	0.41	0.32	0.24	0.15	0.52	0.38	0.32	0.36	0.51
Sulphacetamide	1.54	1.12	0.81	0.43	1.97	0.95	0.59	0.80	1.44
Sulphapyridine	3.19	2.18	1.40	0.69	3.41	1.31	0.71	1.11	2.27
Sulphadiazine	2.76	1.92	1.26	0.62	3.00	1.25	0.69	1.02	2.06
Sulphamerazine	4.31	3.01	1.87	0.92	4.41	1.71	0.93	1.40	2.92
Sulphadimidine	5.94	4.14	2.52	1.19	5.98	2.02	1.07	1.72	3.76
Sulphamethoxydiazine	6.85	4.68	2.78	1.23	8.40	2.83	1.31	1.94	4.71
Sulphisomidine	1.75	1.24	0.73	0.33	1.87	0.51	0.21	0.50	1.19
Sulphadimethoxine	37.78	21.98	12.51	4.37	45.64	10.35	3.57	6.51	20.77
Sulphametopyrazine	8.16	5.71	3.44	1.52	10.02	3.29	1.57	2.36	5.65
Sulphamethoxypyridazine	6.54	4.56	2.78	1.20	7.27	2.32	1.07	1.79	4.26
Sulphathiazole	3.13	2.39	1.35	0.60	3.81	1.28	0.60	0.99	2.32
Succinylsulphathiazole	8.47	5.29	2.73	1.03	8.96	1.63	0.52	1.41	4.32
Phthalylsulphathiazole	27.83	15.05	8.15	2.53	34.36	5.28	1.58	3.84	13.72
Sulphafurazole	11.41	7.44	4.50	1.75	19.28	6.33	2.66	3.70	9.55
Sulphaguanidine	0.33	0.26	0.18	0.09	0.34	0.17	0.06	0.19	0.34
Sulphamethylthiodiazole	6.41	4.24	2.76	1.12	7.65	2.39	1.09	1.79	4.18
5-Methylsulphadiazine	5.71	3.87	2.50	1.08	6.44	2.35	1.21	1.76	3.84
Sulphaphenazole	26.16	13.99	8.86	3.04	44.84	12.02	4.34	6.34	19.17
Sulphamoxole	5.18	3.58	2.15	0.90	5.66	1.71	0.80	1.39	3.28

water–ACN = 3:1) were used. The mobile phase flow-rate was 1 ml/min. The dead time was measured by the injection of uracil. The capacity factors were calculated from at least duplicate injections. The reproducibility was estimated by three repetitions of all measurements at one mobile phase composition regularly spaced in the mobile phase series. The mean relative standard deviation of all retention times was 2.74%. Data sampling was performed by a Digital Minc 11 minicomputer and in-house developed software (SIP). The test solutes were used as purchased (Merck, Darmstadt, Germany, "zur Synthese"). MeOH was of analytical-reagent grade and ACN was of chromatographic quality.

The experiments for the second data matrix were performed with a Waters 6000A HPLC pump, a Shimadzu SPD-6a UV detector and an injection loop of 10 μl. Glass columns, 10 cm × 3.0 mm I.D., filled with Chromspher C_{18}, 5 μm particle diameter, were used. The mobile phase flow-rate was 0.5 ml/min. The dead time was measured by the injection of uracil. The capacity factors are usually the result of one measurement of the retention time. To estimate the repeatability, the retention times of three solutes were measured twice at every mobile phase composition. The relative standard deviation of the capacity factor was 2.4%. The reproducibility was estimated by three repetitions of all measurements at one mobile phase composition regu-

larly spaced in the mobile phase series. The mean relative standard deviation of all the capacity factors was 4.75%. Data sampling was performed by an Olivetti M24 personal computer and in-house developed software (CODA). The test solutes and chemicals were of the same quality as for the first data set.

Software

The calculations were performed on an IBM-compatible AT personal computer with a mathematical co-processor using the Unscrambler program (Camo, Norway) for PCA and the in-house developed POEM (predicting optimal eluent mixtures) package for response surface modelling by multiple linear regression and statistical model validation.

RESULTS AND DISCUSSION

First data set

The fifteen solutes of he first sample consisted of twelve benzene derivatives of different functionality often used in studies of solvent strength and selectivity [2,3,6,12], of which three formed a homologous series of increasing hydrophobicity, *i.e.* MHB, EHB and PHB (for abbreviations see Table I). The remaining three solutes were phenobarbital, PBL, and two larger molecules, the steroids PRE and PRS.

The mobile phases of the design were selected so that the capacity factors of the solutes would lie in the range 0.5–20. The solvent strength of the mobile phases A, B and C is about 1.0, of the mobile phases D, E and F 1.3, and of mobile phases G, H and I 1.5 (according to the transfer rules of Glajch and Kirkland [5]) with relative standard deviations of 11.4, 7.4 and 4.9% at each solvent strength level (Fig. 1). The capacity factors of the fifteen solutes measured at the nine mobile phase compositions are given in Table I.

The results of the PCA of these data are shown in Fig. 3a and b. The first PC accounts for 82% of the variance of the data and the second PC for 17%. This means that, given a repeatability of about 1.4% and a reproducibility of 2.74%, two principal components are sufficient to account for the variation in the data and the data have two intrinsic dimensions that are not correlated.

Parallel to PC1 in the scores plot (Fig. 3a) are three groups of mobile phase compositions for which the solvent strength decreases from left to right: G, D, A; H, E, B; and I, F, C. The water-to-modifier ratio increases in each group in the aforesaid sequence (Fig. 1) and in each group the mean capacity factor of all solutes in a given mobile phase composition increases in the same order (Table I). The mobile phase composition G has the lowest mean k' value of 1.27 and mobile phase A has the highest mean k' value of 4.83 in the first group. Therefore PC1 can be associated with a varying water fraction of the mobile phase and with a change of the mean k', *i.e.* a change in solvent strength. The scale of PC1 units does not, however, correspond quantitatively with the values of the mean k': the distance between mobile phases G and A is five PC1 units (Fig. 3a) and corresponds with a difference of 3.56 mean k' units, while for mobile phases H and B a distance of about 5.2 in PC1 units corresponds with 4.95 units of the mean k'.

Parallel to PC2 the modifier type of the mobile phases changes; mobile phases G, D and A are composed of MeOH and water and mobile phases I, F and C of ACN

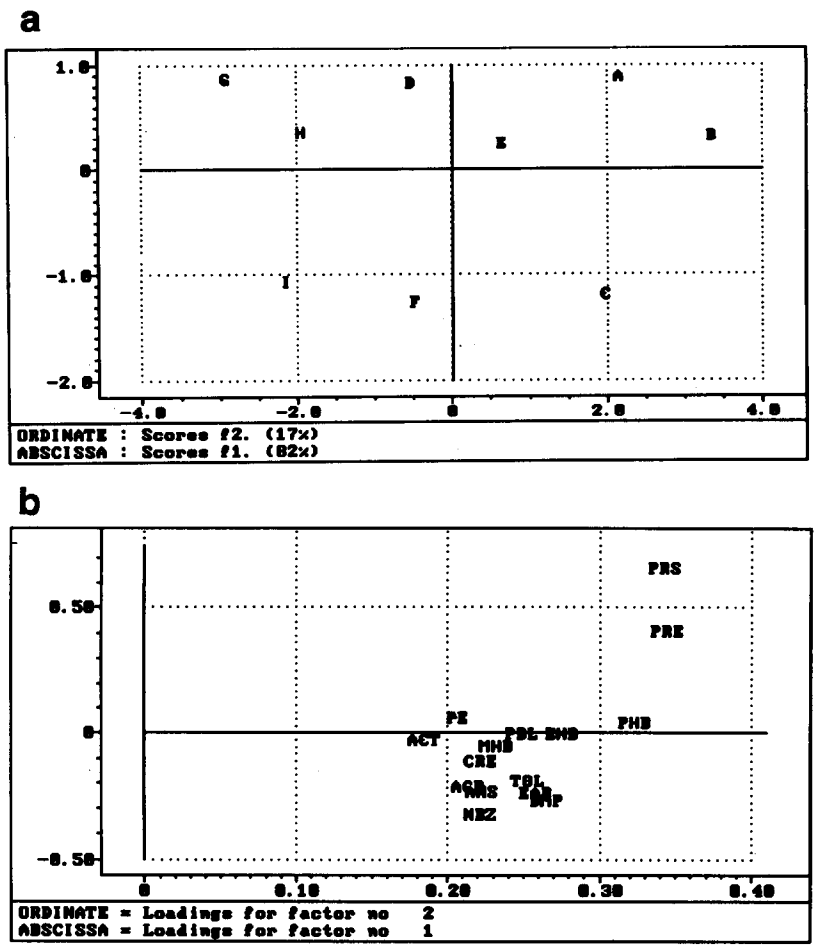

Fig. 3. (a) Scores and (b) loadings on the first two principal component axes of the data of Table I.

and water. The ternary mobile phases H, E and B contain about 40% more MeOH than ACN and are found above the PC1 axis, which does not lie in the middle between the group of MeOH-containing mobile phases (G, D, A) and the ACN-containing groups (I, F, C). This means that PC2 is associated with the substitution of ACN by MeOH, but that the magnitude of this substitution in percentages by volume is not linearly proportional to units of the PC2 axis. In addition, the units of the PC2 axis cannot be related to a change of mean k' values. The respective mean k' values of mobile phases G, D and A are 1.27, 2.38 and 4.83, and these different values all project onto the same point, about 0.8, on the PC2 axis. The same argument applies to mobile phases of the groups H, E, B and I, F, C that project different mean k' values for each group on to the points 0.3 and -1.2 of the PC2 axis, respectively. This means that along the PC2 axis a source of variation in the capacity factors is described that is not related to an increase or decrease of the mean k' value of all the

solutes in a given mobile phase, but to a variation of the capacity factor of the individual solutes. These observations indicate that PC2 can be related to modifier selectivity.

The plot of loadings (Fig. 3b) shows that ACT contributes the least to PC1 and that the loadings of PHB, PRE and PRS are highest on PC1. This means that the k' of ACT changes less along the PC1 axis, while the k' of PHB changes the most. ACT is compared with PHB and not with PRE and PRS, because these latter compounds also have high loadings on PC2. If the change in capacity factor of ACT is compared with the change in the capacity factor of PHB when the solvent strength is decreased, then the k' of ACT increases by 263% and the k' of PHB increases by 498% on going from mobile phase G to A. On going from mobile phase I to C the corresponding increase of k' is 195% for ACT and 421% for PHB. An increase of the water fraction (Table I, Fig. 1) decreases the solvent strength of the binary mobile phase and should cause a proportionally equal increase of k' for all solutes. The difference in the degree of change of k' between ACT and PHB for the same change in solvent strength can be called a solvent strength selectivity effect. This difference is not caused by modifier selectivity, because the modifier type remains the same. Both compounds also have low loadings on PC2, which represents the modifier selectivity.

The compounds MHB, EHB and PHB form a homologous series of increasing hydrophobicity. Their loadings on PC1 increase and this was reflected by a corresponding increase of their log k_w values, that were calculated from eqn. 1 using their capacity factors in water–MeOH (A, D, G) and in water–ACN (C, F, I) mobile phases, respectively. Their positions in Fig. 3b lie on a line that forms a small angle with PC1, which means that the PC1 axis problaby represents not pure hydrophobicity, but has a strong correlation with hydrophobicity. NBZ and PRS have the lowest and highest loading on PC2, respectively. This will be discussed in the next section.

Second and third data sets

In the second data set (Table II) anisole (ANS) was replaced by aniline (AN), MHB, EHB and PRE were omitted and a different brand of ODS packing material was used.

The solvent strength was adjusted at two levels (Fig. 2). The mean capacity factor of the solutes was used as a measure to adjust the solvent strength. Iso-eluotropic mobile phases have equal mean k' values according to this definition. At the first level the relative standard deviation of the mean capacity factors of all solutes measured in mobile phases A, B, C and D is 3.27%. At the second level the relative standard deviation is 1.40% among the mobile phases for the mean k' of all solutes measured in mobile phases E, F, G and H. The purposes of this experiment were firstly to see if the results obtained with the first data set could be confirmed on another packing material. The second purpose was to investigate whether the experimental adjustment of the composition of the mobile phases, to give an approximately constant mean k' value at each level of solvent strength, would lead to a better correspondence of PC1 with differences in the mean k'. Stated differently: do the scores on the PC1 axis quantitatively represent solvent strength values defined by the mean k' of all solutes measured in a given mobile phase composition?

From the scores plot (Fig. 4a) it can be concluded that this is not the case. The mean score of the mobile phases with the lower solvent strength (A, B, C, D) is 1.75

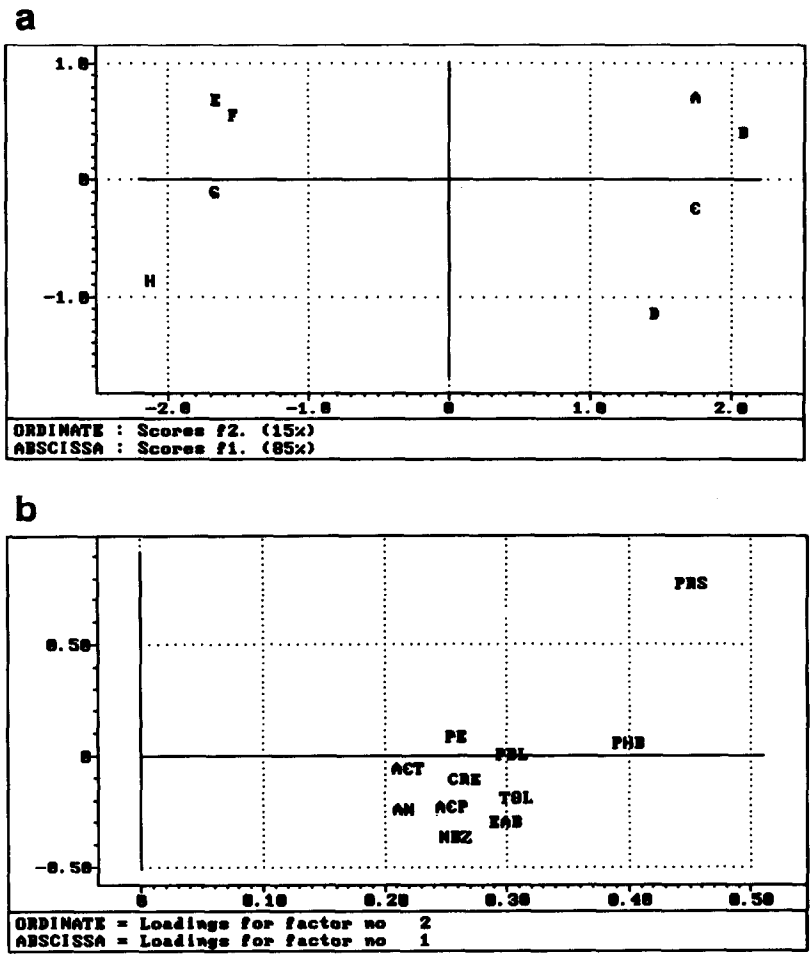

Fig. 4. (a) Scores (b) loadings on the first two principal component axes of the data of Table II.

on PC1 and the overall mean k' is 4.47; the mean score of the mobile phases E, F, G and H is -1.75 and their mean k' is 1.48. So a distance of 3.5 PC1 units corresponds to a difference of 2.99 mean k' units. This confirms that the PC1 axis can be associated with increasing mean k' values of the mobile phases. The difference in mean k' of mobile phases H and G is only 0.01 k units, but the corresponding distance in PC1 units is 0.46. This means that the PC1 scale does not correspond quantitatively with a mean k' scale and does not correspond exactly to the solvent strength defined by the mean k' of a sample. The reason is that variations in k' of the individual solutes due to selectivity effects also contribute to the mean k' value, but the variance explained by the PC1 axis is not confounded by modifier effects because they are associated with the PC2 axis.

This hypothesis is corroborated by the fact that compared with the first data set PC1 now accounts for 85% of the variance of the data and PC2 for 15%. The main

difference between the first and second data set is the presence of PRE in the first data set (see Figs. 3b and 4b), because the other solutes that have been changed have smaller loadings. By removing PRE from the data set, which has a high loading on PC2 (Fig. 3b) and is very sensitive to a change of modifier type, the variance accounted for by PC2 decreases from 17 to 15%. An even stronger confirmation can be obtained by performing a new PCA on the second data set after the removal of PRS, the solute of this data set that has the highest loading on PC2 (Fig. 4b). The results of the PCA of the data after the removal of PRS are shown in Fig. 5. The scores plot (Fig. 5a) again shows two groups of mobile phases with different solvent strength. The loadings plot (Fig. 5b) shows that the relative positions of the remaining solutes are similar to those of Fig. 4b. (The sign reversal of PC2 is irrelevant for this analysis). The variance accounted for by PC1 has been increased to 94%, whereas the variance explained by PC2 has been decreased to 5%.

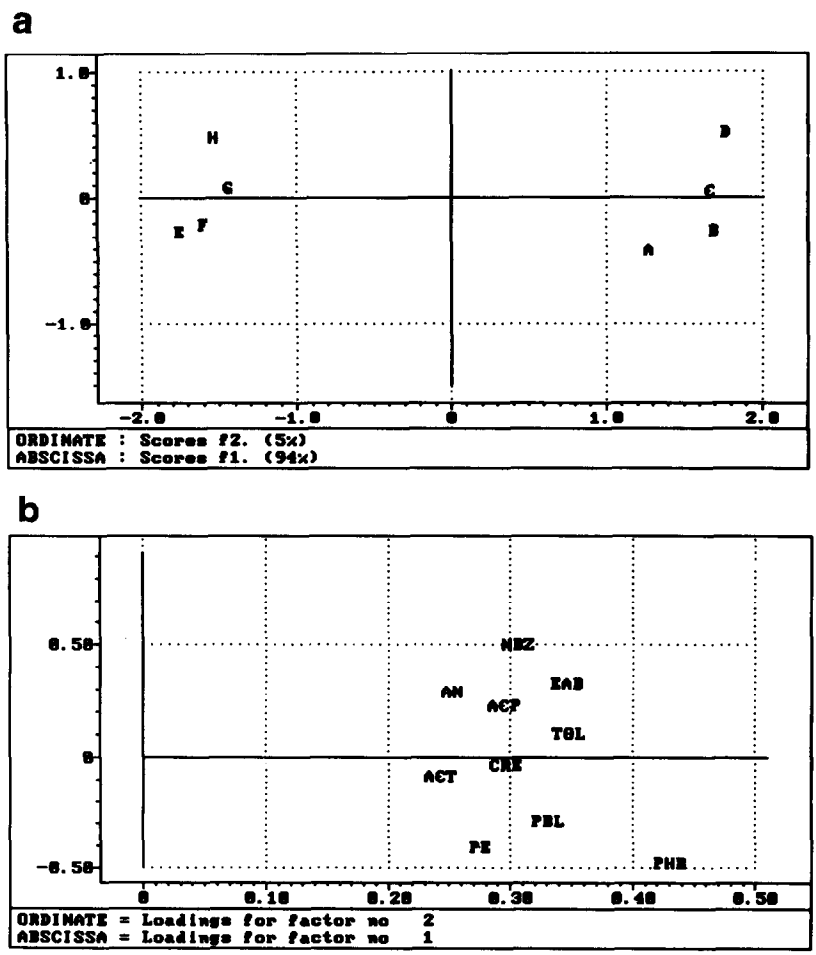

Fig. 5. (a) Scores and (b) loadings on the first two principal component axes of the data of Table II after removal of prednisolone.

This also shows that the amount of variation that is accounted for by each PC strongly depends on the selection of the variables or solutes. To obtain another impression of these amounts a third data set consisting of twenty sulphonamide derivatives was analysed and it was found that the first principal component accounted for 96% and the second for 2% of the total variance of the data. These figures suggest that for samples consisting of structurally related compounds PC1 accounts for about 95% and PC2 for about 4% of the total variance.

The PCA suggests the possibility of an unequivocal definition of solvent strength by the use of compounds that are not sensitive to a change of modifier type, *i.e.* markers of which the capacity factor remains constant when MeOH in the binary water–MeOH mobile phase is replaced by a different modifier without a change in eluotropic strength. Positive selectivity indicates a specific acceleration of the solute by the modifier [20]. Therefore if a solute has zero modifier selectivity, then changes in the capacity factor by varying the water fraction of the binary mobile phase are due to solvent strength and solvent strength selectivity.

The ideal solvent strength marker should have a zero modifier selectivity and an average solvent strength selectivity. Zero modifier selectivity means that the capacity factor of the marker is not affected by a change of modifier type provided that the solvent strength of the binary mobile phase does not change. Average solvent strength selectivity ensures that the capacity factor of the marker decreases in an average manner if the water content of a binary mobile phase is increased.

PRS has a high loading on PC2 (Fig. 4b) and is very sensitive to a change of modifier type, as is shown in Fig. 6, the contour plot of the capacity factor of PRS in the design space. The contour plot was obtained by fitting a quadratic model to the capacity factors of PRS measured at the mobile phase compostions A to H. The

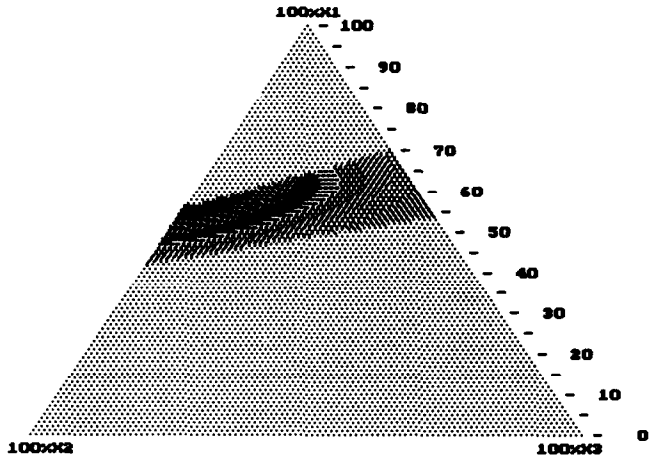

Fig. 6. Contour plot of the capacity factor of prednisolone based on the data of Table II. The capacity factor varies from 0.4 to 7.6. The different symbols correspond to ten different ranges of values of k'. The highest range from 6.9 to 7.6 is indicated by black squares (upper left corner of the design space). The lowest range of 0.4–1.1 is indicated by back-slashes (lower right corner of the design space). X_1 is water, X_2 is methanol and X_3 is acetonitrile.

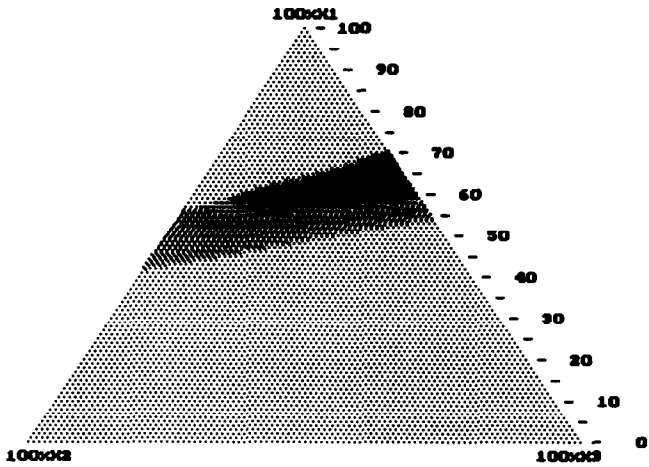

Fig. 7. Contour plot of the capacity factor of nitrobenzene based on the data of Table II. The capacity factor varies from 1.4 to 6.1. Different symbols correspond to ten different ranges of values of k'. The highest range from 5.6 to 6.1 is indicated by black squares (upper left corner of the design space). The lowest range of 1.4–1.8 is indicated by back-slashes (lower right corner of the design space). X_1 is water, X_2 is methanol and X_3 is acetonitrile.

contour lines are strongly curved and indicate that the capacity factor of PRS for a water–MeOH binary mobile phase is considerably greater than for the corresponding iso-eluotropic water–ACN binary mobile phase, for example, mobile phases A and D of Fig. 2.

The contour plots illustrate very well the sensitivity of a compound to a change of modifier type and the meaning of loading on the PC2 axis. PRS has the highest loading on PC2 and NBZ the lowest (Fig. 4b). The contour lines of the plot of NBZ (Fig. 7) indicate that the capacity factor of NBZ in a water–MeOH binary mobile phase is smaller than for the corresponding iso-eluotropic water–ACN binary mobile phase. So if a 50% water–MeOH binary mobile phase is replaced by a water–ACN binary mobile phase in such a way that the capacity factors of PRS and of NBZ remain constant, then the water–ACN binary mobile phase for PRS will contain less ACN than the water–ACN binary mobile phase for NBZ. The fraction of ACN in a water–ACN binary mobile phase that gives the same capacity factor as a 50% water–MeOH phase has been calculated for a number of solutes of the second data set, and the results are given in Table IV. The results clearly show that the fraction of ACN necessary to keep the capacity factor constant in both binary mobile phases strongly

TABLE IV

PERCENTAGE OF ACETONITRILE EQUIVALENT TO 50% METHANOL INDICATED BY A CONSTANT CAPACITY FACTOR OF A SOLUTE

Solute	PRS	PE	PHB	PBL	ACT	CRE	TOL	ACP	NBZ
Percentage ACN	20[a]	31	35	37	40	39	39	42	45

[a] Extrapolated.

correlates with the loading on PC2 (Fig. 4b). Solvent strength markers should have low absolute loadings on PC2 and as such ACT, CRE and TOL are good possible choices for a sample without steroids (Fig. 5b). The ACN fraction of these solutes corresponds well with the percentages calculated by the transfer rules of refs. 6 and 12, which are 36.5 and 40.0% of ACN, respectively.

The definition of an empirical solvent strength scale of reversed-phase mobile phases seems to be possible once the markers have been selected from the correct set of test compounds. Further research on this subject is in progress.

CONCLUSIONS

PCA of a data set of mobile phases consisting of water and different percentages of MeOH and/or ACN characterized by capacity factors of different solutes shows that there are two intrinsic dimensions in the data set. The variance in the data is caused by two independent effects that are described by two principal components.

The first principal component accounts for about 82–96% of the total variance and the second for about 17–2%. The amount of variance explained by each principal component depends strongly on the selection of the solutes that constitute the test sample. The first principal component can be associated with the water fraction of the mobile phase and is interpreted in terms of solvent strength and solvent strength selectivity. The second principal component can be associated with the modifier type of the mobile phase and is interpreted in terms of modifier selectivity.

PCA enables the selection of solutes from a sample that are the most critical test compounds for the characterization of solvent strength and modifier selectivity, and could be used for the definition of an empirical solvent strength scale of reversed-phase mobile phases.

REFERENCES

1 L. R. Snyder, *J. Chromatogr. Sci.*, 16 (1978) 223.
2 L. R. Snyder, J. W. Dolan and J. R. Gant, *J. Chromatogr.*, 165 (1979) 3.
3 P. Jandera, *J. Chromatogr.*, 314 (1984) 13.
4 J. W. Dolan, J. R. Gant and L. R. Snyder, *J. Chromatogr.*, 165 (1979) 31.
5 J. L. Glajch and J. J. Kirkland, *Anal. Chem.*, 54 (1982) 2593.
6 P. J. Schoenmakers, H. A. H. Billiet and L. de Galan, *J. Chromatogr.*, 205 (1981) 13.
7 S. R. Bakalyar, R. McIllwrick and E. Roggendorf, *J. Chromatogr.*, 142 (1977) 353.
8 H. B. Patel and T. M. Jefferies, *J. Chromatogr.*, 389 (1987) 21.
9 P. M. J. Coenegracht, H. J. Metting, A. K. Smilde and P. J. M. Coenegracht-Lamers, *Chromatographia*, 27 (1989) 135.
10 L. R. Snyder, M. A. Quarry and J. L. Glajch, *Chromatographia*, 24 (1987) 33.
11 A. K. Smilde, A. Knevelman and P. M. J. Coenegracht, *J. Chromatogr.*, 369 (1986) 1.
12 D. P. Herman, H. A. H. Billiet and L. de Galan, *J. Chromatogr.*, 463 (1989) 1.
13 W. R. Dillon and M. Goldstein, *Multivariate Analysis*, Wiley, New York, 1984, p. 23.
14 B. Walczak, M. Dreux, J. R. Chretien, K. Szymoniak, M. Lafosse, L. Morin-Allory and J. P. Doucet, *J. Chromatogr.*, 353 (1986) 121.
15 J. R. Chretien, B. Walczak, M. Dreux and M. Lafosse, *J. Chromatogr.*, 371 (1986) 253.
16 B. Walczak, L. Morin-Allory, M. Lafosse, M. Dreux and J. R. Chretien, *J. Chromatogr.*, 395 (1987) 183.
17 M. F. Delaney, A. N. Papas and M. J. Walters, *J. Chromatogr.*, 410 (1987) 31.
18 J. A. Cornell, *Experiments with Mixtures*, Wiley, New York, 1981.
19 J. W. Weyland, *Ph.D. Thesis*, State University of Groningen, Groningen, 1986, p. 56.
20 P. J. Schoenmakers, H. A. H. Billiet and L. de Galan, *J. Chromatogr.*, 218 (1981) 262.