Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1030 (2004) 135–147
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Testing of "special base" columns in reversed-phase liquid chromatography A rational approach considering solvent effects

K. Le Mapihan, J. Vial*, A. Jardy

Laboratoire "Environnement et Chimie Analytique", UMR 7121, Ecole Supérieure de Physique et de Chimie Industrielles de la Ville de Paris, 10 Rue Vauquelin, 75005 Paris, France

Abstract

A methodology for building a chromatographic test aiming at characterizing special base stationary phases was described. Instead of choosing its conditions a priori, a "full" comprehensive test based on extended running conditions was performed on a 12 column set. The conditions were carefully chosen from their ability to take into account the solvent and the pH effects. Principal component analysis (PCA) has been combined to hierarchical cluster analysis both to provide interpreted classifications and to reduce drastically the test itself by eliminating redundant information. The final reduced test can be considered optimal because the minimized set of test conditions allows to provide as much information as in the initial full test.

Keywords: Column testing; Stationary phases, LC; Solvent effects; Principal component analysis; Cluster analysis; Drugs

1. Introduction

Reversed-phase liquid chromatography has become an indispensable method for the separation of pharmaceutical products, which frequently comprised basic properties. In order to improve the separation of such difficult compounds, a wide range of stationary phases have been developed especially for this kind of analysis. The appearance of such phases, also known as "special base", took part into the growing diversity of chromatographic supports [1], which puzzled analysts for choosing the right column within the context of a new chromatographic method development. To face this problem, a broad consensus was reached concerning the need for a test probing chromatographic properties of stationary phases. Bibliography provides a considerable number of empirical chromatographic tests that are able to discriminate between columns according to some types of properties [1–28]. Their results seemed to be partially correlated [26,27] and some of the proposed tests were prone to later refinements [2,29], making their own results challenged. These empirical tests are also commonly based on solutes that are expected to well probe the stationary phases. However, often the chosen conditions (including the solutes) were not really representative of the actual using conditions. So, it appeared necessary to build a test suitable for the new columns thanks to a new "reverse" approach: the chosen methodology consisted in the selection of chromatographic conditions a posteriori rather than postulating their probing power. A later performed selection implied the setting up of a first comprehensive test, also called "full" test that offered a wide range of variables in order to select only the most informative ones. At this end, the characterization capacity of the "full" test was extended by maximizing the diversity of test conditions. A wide range of probe solutes was injected. The analytes differed in their physico-chemical properties, the presence of various heteroatoms, their molecular mass and their three-dimensional structure. As the solvent nature had a dramatic effect on column characterization [30], the "full" test was performed systematically in two organic solvents, i.e. methanol and acetonitrile. Many of the empirical tests described include generally two pH levels, often 3 and 7. An intermediate pH level was also introduced. In order to obtain reliable chromatographic parameters, solutes were injected among four solvent fractions. Twelve chromatographic columns selected for their representativeness were then assessed thanks to the "full" test. To obtain reliable classifications, data processing

* Corresponding author. Fax: +33-140794776.
E-mail address: jerome.vial@espci.fr (J. Vial).

0021-9673/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2003.10.099

was performed by the combination of principal component analysis (PCA) and hierarchical cluster analysis. This particular data handling allowed us to construct both a rational and an objective reduction of the full test.

2. Experimental

2.1. Chemicals and reagents

Acetonitrile (MeCN; HPLC ultra gradient grade) and methanol (MeOH; HPLC gradient grade) were from Mallinckrodt Baker (Deventer, The Netherlands). Water was produced by a Milli-Q Plus ultrapure water purification system (Millipore, Molsheim, France). Orthophosphoric acid and acetic acid were obtained from Prolabo, potassium dihydrogenphosphate and sodium acetate by Aldrich whereas Tris base [tris(hydroxymethyl)aminomethane] and Tris–hydrochloride were supplied by Fluka.

To obtain a comprehensive test, the set of selected solutes had to meet the following conditions.

(i) To cover a wide distribution of physico-chemical properties, in terms of hydrophobicity, polarity and acid–base equilibria and to provide a diversity of heteroatoms, molecular masses and three-dimensional structures.
(ii) To ensure a perfect accessibility to the testing procedure, meaning cheap and not forensic products with a sufficient stability.

In order to fulfill the required conditions, most of the 30 selected probe solutes were well-known pharmaceuticals, patent of which had expired. Some "classical" compounds (like solutes of the test of Tanaka [23]) were also included. So, the test was constituted of amiodarone hydrochloride (Sigma), ampicillin sodium salt (Fluka), atropine sulfate salt (Sigma), benzylamine hydrochloride (Sigma), buspirone hydrochloride (Sigma), *n*-butylbenzene (Aldrich), caffeine (Fluka), chlorpropamide (Sigma), clofazimine (Sigma), cyanocobalamine (Sigma), dextromethorphan hydrobromide (Sigma), digitoxin (Fluka), diltiazem hydrochloride (Sigma), *p*-ethylaniline (Aldrich), ethylbenzene (Aldrich), furosemide (Sigma), imipramine hydrochloride (Sigma), loperamide hydrochloride (Sigma), nortriptyline hydrochloride (Sigma), *n*-pentylbenzene (Aldrich), phenol (Merck), primidone (Sigma), quinine sulfate dihydrate (Fluka), reserpine (Fluka), strychnine hemisulfate salt (Sigma), *o*-terphenyl (Fluka), toluene (Merck), triphenylene (Fluka), D-tubocurarine chloride (Sigma) and vancomycin hydrochloride (Sigma). Finally, the set of solutes had log *P*-values evenly distributed from −0.07 to 7.66, with molecular masses comprised between 92 and 1450 g mol^{-1} and acid–base equilibria ranging from 1.91 to 9.99.

2.2. Instrumentation

Three different LC systems were used. The first was composed of a Varian Prostar 230 ternary pump (Varian, Les Ulis, France), a Waters 715 UltraWisp autosampler, a Waters 2487 UV detector (Waters, Saint-Quentin en Yvelines, France) set at 230 nm plus a Varian 2050 UV detector (Varian, Les Ulis, France) set at 254 nm. The data acquisition was performed thanks to Class-VP 4.2 (Shimadzu Scientific Instruments, Columbia, MD, USA). The second LC system consisted of a HP 1050 quaternary pump, a HP 1050 autosampler and a HP 1100 variable wavelength detector operated at 230 or 254 nm. The third LC system was a complete HP 1090 system with a diode array detector. For the two last systems, Chemstation 6.03 (Agilent Technologies, Waldbronn, Germany) was the common data system. The acquisition frequency was at least 25 Hz. Concerning temperature regulation, all tested columns were placed in an Alltech water jacket connected to a water bath set at 40 °C (±0.1 °C with the water bath Bioblock 18205 for the first LC system, ±0.03 °C with the water bath Neslab RTE-101 for the other LC systems). All columns were operated using a flow rate of 1 ml min^{-1}. It had been verified previously that both data acquisition systems were able to produce equivalent measurements from a common chromatogram on the three kinds of recorded parameters: retention times, asymmetries and efficiencies.

2.3. Running conditions

2.3.1. Buffer preparation

In order to obtain a reliable test, it had been established that mobile phases had to be buffered [29]. At pH 3.00 and 5.00, buffers were prepared by dissolving the accurate quantity of salt (4.770 g of potassium phosphate monobasic, 3.281 g of sodium acetate for 2 l of solution) in pure water and adjusting the pH to the appropriate value at 25 °C with concentrated conjugated acid. At pH 7.00, buffer solution was obtained by adding 5.840 g of Tris–hydrochloride to 356.5 mg of Tris base. After making up to 2 l in a volumetric flask with pure water, the pH value was checked taking into account the drift due to temperature. All buffers were filtered through 0.45 μm HA type filters (Millipore, Molsheim, France), before addition of the organic modifier.

2.3.2. Mobile phase preparation

The solvent mass fraction had been determined previously to obtain isoeluotropic strength on alkylbenzenes with a Kromasil C_{18} column (250 mm × 4.6 mm, 5 μm, Akzo Nobel) (see Table 1).

Mobile phases were freshly prepared by weight for each column within the ratios indicated in Table 1.

Table 1
Solvent fraction levels for methanol and for acetonitrile

Eluent	Methanol (%, w/w)	Acetonitrile (%, w/w)
A	70	59
B	45	33
C	30	20
D	15	9

2.3.3. Injection conditions

Table 2 indicates in which mobile phases the solutes were injected; the letter represents the solvent fraction level and the figure is for the pH value.

Neutral compounds (except digitoxin) were injected at the same solvent fraction level but at different pH levels in order to verify the eluent preparation accuracy. All compounds were injected at the following concentrations: 50 ppm for the majority of solutes except for alkylbenzenes (150 ppm), *o*-terphenyl (20 ppm), triphenylene (3 ppm), benzylamine (600 ppm), strychnine (100 ppm) and D-tubocurarine (100 ppm). At least 1 h equilibration was performed before the 10 μl injection of mixtures in triplicates. In addition, each solute was injected individually in order to confirm its identity and to verify that no particular interaction biased the retention times. The Tanaka's test compounds were detected at 254 nm whereas the others were detected at 230 nm. The column void volume was determined by the injection of thiourea (Aldrich) in the acetonitrile mobile phase. All samples were stored at 4 °C or less.

2.4. Tested columns

The full test has been applied to a sample group of different columns representative of those commonly used in pharmaceutical industry. The chosen columns differed from each others in the length of alkyl moiety (C_8, C_{16} or C_{18} graft) and in the kinds of protection against residual silanol groups. For the sake of simplicity, these protections can be split into two categories.

(1) Masking silanols or decreasing the magnitude of their effect by the use of ultrapure silica, bonded with particular grafts leading to steric or shielding hindrance.
(2) Reducing partly the number of silanols by synthesizing hybrid silica, increasing bonding density, encapsulating silica, performing one or several end-capping.

These protection strategies and their combinations constitute the main origin of the wide diversity of special bases stationary phases. The available characteristics of the tested columns are reported in Table 3. The dimensions of the columns were 150 mm × 4.6 mm i.d. except for Symmetry C_{18} (250 mm × 4.6 mm i.d.) and for J'Sphere ODS H80 (150 mm × 4.0 mm i.d.).

We have not included polar end-capped columns because they commercially appeared too recently to assess their robustness and they risked to become degraded during the test itself, test that was time consuming.

Table 2
Injection conditions

Solute	Eluent
Digitoxin	A5
Clofazimine	A3; A5; A7
Amiodarone	A3; A5; A7
Butylbenzene	A5; A7
Pentylbenzene	A5; A7
o-Terphenyl	A5; A7
Triphenylene	A5; A7
Reserpine	A7; B3; B5
Toluene	B5; B7
Ethylbenzene	B5; B7
Loperamide	B3; B5; B7
Imipramine	C3; B5; B7
Nortriptyline	C3; B5; B7
Diltiazem	C3; B5; B7
Buspirone	C3; B5; B7
Dextromethorphan	C3; C5; C7
Phenol	C3; C5; C7
Furosemide	C3; C5; C7
Chlorpropamide	C3; C5; C7
p-Ethylaniline	C3; C5; C7
Quinine	D3; C5; C7
Primidone	D3; C5; C7
Strychnine	D3; D5; C7
D-Tubocurarine	D3; D5; C7
Ampicilline	D3; D5; C7
Atropine	D3; D5; D7
Cyanocobalamin	D3; D5; D7
Vancomycin	D3; D5; D7
Caffeine	D3; D5; D7
Benzylamine	D3; D5; D7

2.5. Figures of merit

Three kinds of peak parameters were recorded: retention times, asymmetries and efficiencies. Retention factors were deduced from retention time by the following equation: $k = (t_r/t_0) - 1$ where t_r and t_0 represented the retention time of the compound and the void volume retention time respectively. Peak asymmetry was calculated at 5% of the peak height from the ratio of the widths of the rear and front sides of the peak (USP definition of tailing factor). In order to get a more balanced definition, asymmetry values were transformed in such a way that front and rear tailings of the same magnitude gave opposite but comparable deviation from 1 [30]. Concerning the efficiency measurement, the reduced height h was used instead of the number of plates N (evaluated with the half height method) as the tested columns had different particle sizes (noted d_p). It was calculated as follows: $h = L/(Nd_p)$ where L was the column length.

2.6. Softwares

The Unscrambler 7.5 (Camo Asa, Oslo, Norway) was used to perform principal component analyses while cluster analysis was carried out with JMP 4.0.5 (S.A.S. Institute, Carry, NC, USA).

3. Results and discussion

3.1. Data handling

Our data were processed according to a generic three-step procedure: raw data were first pre-treated before running

Table 3
Characteristics of tested columns; the abbreviation were constructed by concatenating the nature of the graft, its length and the testing order of the stationary phases

Name	Pore diameter (nm)	Particle size (μm)	% C	Surface area (m^2/g)	Manufacturer	Comment	Abbreviation
CapcellPak C_8 UG 120	12	5	10	300	Shiseido	C_8 encapsulated silica	E8_1
Xterra RP 8	10	3.5	13.4	174	Waters	Carbamate embedded C_8 on hybrid silica	P8_1
Discovery RP Amide C_{16}	18	5	12.0	198	Supelco	Amide embedded C_{16}	P16_1
Eclipse XDB C_8	10	5	7.6	180	Zorbax	High density monomeric C_8 bonding	C8_1
Symmetry C_8	10	5	12.2	344	Waters	High density monomeric C_8 bonding	C8_2
Kromasil C_8	10	5	12	340	Akzo Nobel	High density monomeric C_8 bonding	C8_3
XTerra RP 18	10	3.5	14.4	172	Waters	Carbamate embedded C_{18} on hybrid silica	P18_1
SymmetryShield RP 18	10	3.5	17.0	339	Waters	Carbamate embedded C_{18}	P18_2
J'Sphere ODS H80	8	4	22	Not given	YMC	High density polymeric C_{18} bonding	C18_1
Stable Bond C_{18}	8	3.5	10	300	Zorbax	Diisopropyloctadecyl bonding	C18_2
Nucleosil C_{18} HD	10	5	20	350	Macherey-Nagel	High density monomeric C_{18} bonding	C18_3
Symmetry C_{18}	10	5	19.5	341	Waters	High density monomeric C_{18} bonding	C18_4

The nature of the graft is symbolized by a letter: E means encapsulated, P stands for polar embedded graft and C for pure hydrocarbonaceaous moiety.

principal component analysis; then, hierarchical cluster analysis was performed in order to obtain our final classifications. Each step will be discussed in the later parts.

3.1.1. Data pretreatment

Twelve columns were tested thanks to 30 solutes by varying three factors: solvent nature (methanol or acetonitrile), solvent fraction (four for each solvent) and pH (3, 5 and 7), leading to a 5040-data table (three peak parameters × 1680 rows, result of the combination of factor levels and solutes).

As PCA is only able to compute numerical data, before all, it was necessary to organize raw data in a compatible way. The difficulty lay in how taking into account factors like the pH or the nature of the testing solvent. Concerning the solvent nature factor, previous results had shown that it was more relevant to consider the column–solvent couple than the column by itself. The number of individuals then defined by such couples has risen to 12 columns × 2 solvents = 24. In practice and for better legible representations, the couples were identified as follow: the testing solvent was figured by a defined symbol (a triangle for acetonitrile and a circle for methanol) and the column identity by the abbreviation as given in Table 3. Regarding the pH level the best solution was to include it in the definition of the variable, i.e. variables of the PCA. So, the variables have been defined by the following construction: variable = chromatographic parameter||solute||pH. For example, *k* ampicillin five stands for the retention of ampicillin at pH 5. It led to 3 × 30 × 3 = 270 variables. As the injected solutes were distributed among the four solvent fractions in order to obtain meaningful retention factors, it was not necessary to take into account the solvent fraction level: it was already included in the data of the solute and pH. Concerning neutral compounds, they were injected at different pH levels but only one level was selected for data handling, to avoid adding willingly redundant information. After the removal of the variables in excess, the final data table was composed of 210 variables × 24 individuals instead of 3 variables × 1680 individuals.

3.1.2. Principal component analysis

PCA [31–33] is a powerful tool for the interpretation of large data tables. PCA is a projection method that is able to extract the main information from the original data set by affecting it to a dimensionally reduced space. This space is defined by linear combination of variables, called principal components (PC). The last ones are computed in such a way that the successive PCs convey less and less information while being orthogonal and maximizing their information load. The plots of the individuals, i.e. the column–solvent couples, in the new defined set of coordinate axes are called score plots whereas the representations of the initial variables constitute the loading plots. In the present study, data were centered and standardized in order to give all variables the same importance and a cross validation was performed to assess the reliability of the obtained classifications.

3.1.3. Hierarchical cluster analysis

Cluster analysis [33,34] is a tool used for pattern recognition that detects similarities between objects according to the distance between them. The hierarchical method (HCA) is based on an agglomerative process. At the beginning, each individual constitutes a cluster. At each step the distance between each point is calculated and the two closest points are gathered to form a new cluster until all the original points are together into one group. The results are represented in tree diagrams also called dendograms. In this work, the method of clustering was based on the Euclidean distance (centroid criterion) and performed on autoscaled PC-scores. Such an approach is equivalent to perform a cluster analysis based on Mahalanobis distance from the original coordinates [32,35,36]. It confers the advantage of a definitely better fit of elongated clusters.

3.2. Full test

Our first intention consisted in performing a PCA on the whole set of recorded chromatographic parameters. Due to the obvious lack of interpretability of the classifications, it was then decided to treat each kind of chromatographic parameter separately.

3.2.1. Retention factors

As shown in Fig. 1, 67% of the total information was carried on the first component, 13% on the second one.

Eighty percent of total information was described thanks to the plane constituted of the two first PCs. A distinctive feature of the PC1–PC2 score plot was that it was V-shaped, making the use of cluster analysis particularly attractive for its interpretation [35]. Moreover, by drawing the bisector, it seemed at first sight that tested columns were classified in two categories corresponding to the testing solvent. Fig. 2 shows the resulting dendogram performed on PC1–PC2 scores.

The cluster number was determined by “cutting” vertically the dendogram as indicated by the dotted line. Nine groups were suggested and led to the following interpretation. Firstly, all the considered column–solvent couples were gathered according to their testing solvent. It underlined

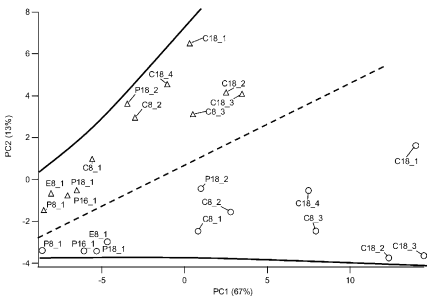

Fig. 1. PC1–PC2 score plot of the column–solvent couples constructed on all retention factors; the V-shape and its relative bisecting line are schematically represented in continuous and dotted lines, respectively.

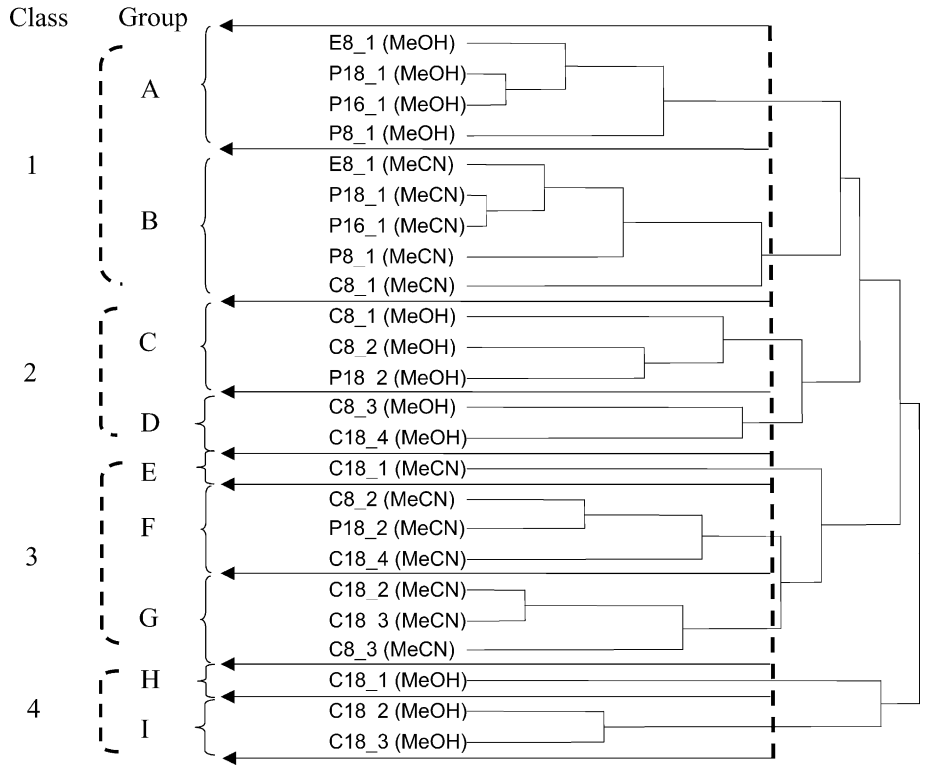

Fig. 2. Dendogram of relationships between column–solvent couples obtained by HCA of PC1–PC2 scores computed with retention factors.

the necessity of considering the couple solvent–column rather than the column by itself, confirming our previous results. If it was suggested by Fig. 1, it clearly appeared on Fig. 3. On the upper part of the PC-scores, all the columns tested in acetonitrile were found (groups B, E, F and G).

It has been noticed that the column–solvent couples were also classified according to their hydrophobicity: all poor

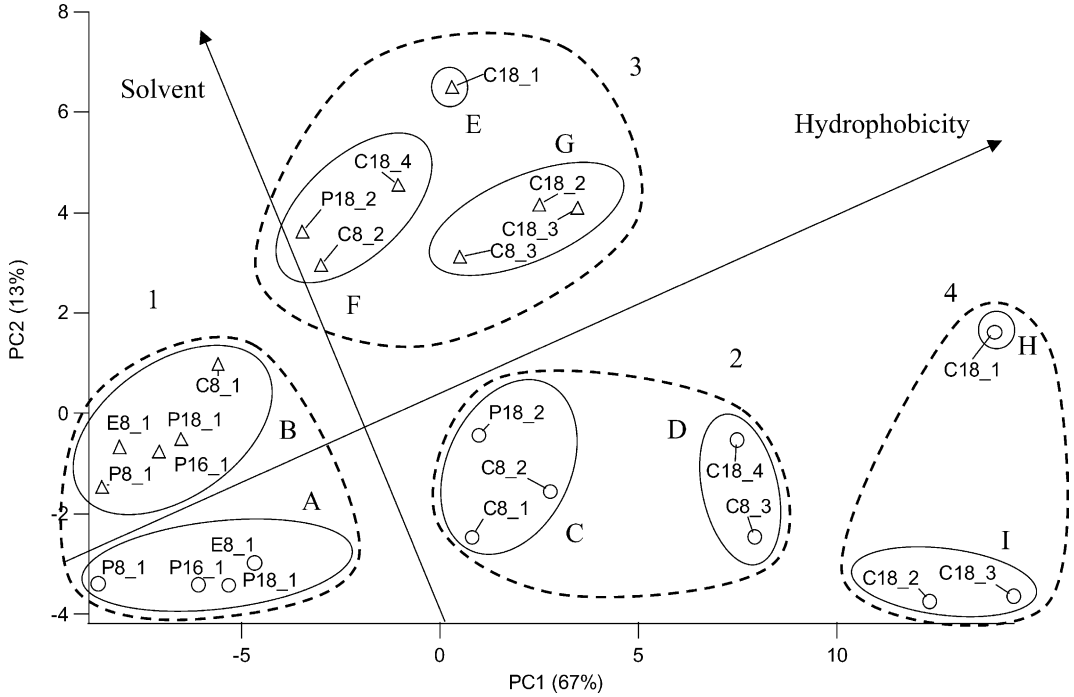

Fig. 3. Interpration of PC1–PC2 score plot of the column–solvent couples obtained with all retention factors.

retentive columns were gathered together (class 1), like for the intermediate (classes 2 and 3) or the far retentive ones (class 4). So it meant that on PC-scores, the individuals were arranged according to two main directions: hydrophobicity and solvent. It was also observed that the less hydrophobic stationary phases were together whatever the testing solvent. Such a behavior had been chromatographically previously noticed. The class 1, as a result of the merge of clusters A and B, constituted also the starting point of the V-shape of the classification. This particular shape could be attributed to a non linear effect taking into account the solvation process of alkyl chains (resulting of solvent nature).

The loading plot in relation to Figs. 1 and 3 (not shown) revealed a crescent-shaped distribution for the variables. None of the variables was in the neighborhood of the origin and then easily removable. Moreover, most of the variables were close to each other, meaning a high degree of correlation. Reduction appeared possible and necessary to extract non redundant information. However, the interpretation of the PC-scores thanks to PC-loadings was limited by the cloudy shape of the variables, confirming the relevance of combining PCA and HCA.

3.2.2. Peak asymmetries

Fig. 4 summarizes the interpreted PCA (i.e. the result of a HCA performed on a PCA) obtained with peak asymmetries as variables.

Eleven groups resulted from the HCA achieved in the space defined by the two first PC that carried 52% of total information. If the cut was slightly shifted forwards higher dissimilarities, i.e. toward the central knot, it yielded 8 groups instead of 11, as shown by the dotted lines. The solvent effect was really less marked than for the classification obtained with the retention factors. All columns having a carbamate group embedded in the alkyl graft were gathered and their relative cluster differed obviously from the alkylamide and the purely alkyl graft ones. It could be noticed that the diisopropyloctadecyl graft had been moved away from the other purely alkyl grafts. These lasts were either bonded with high density or encapsulated, strategies that were known to provide a better protection. Finally, it appeared that solvent–column couples were ordered by kinds and levels of protection of their residual silanol groups.

Concerning the efficiency variables, the same procedures had been applied but the results are not shown because of the lack of interpretability of the obtained classifications.

3.3. Reduction of the full test

Due to a totally unacceptable duration (around 1 week per column), the full test was neither applicable nor transferable in the present state. However, this step was necessary to select only the optimal testing conditions by an objective methodology.

3.3.1. Methodology

With PCA, the amount of original variables was reduced to a few principal components, leading to a drastic dimensional reduction (for example, from a space of 90

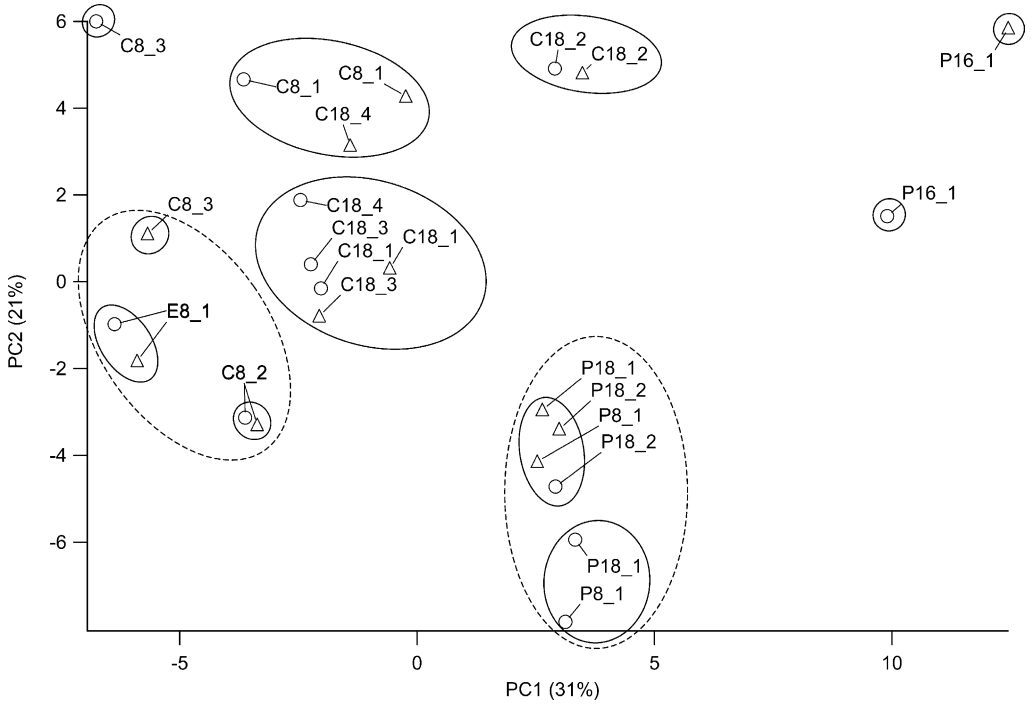

Fig. 4. Interpreted PC1–PC2 score plot of the column–solvent couples constructed on all peak asymmetries; HCA achieved in the space defined by the two first principal components computed with peak asymmetries.

variables, i.e. at 90 dimensions, to two or three calculated variables, i.e. two or three dimension space). However, because coordinates along principal components are linear combinations of the original variables, the number of variables remained the same. The aim of the following methodology was to manage a reduction of the original variable set with a minimal disturbance of the final classifications. It could be performed by different approaches. By definition, the close-to-center variables (i.e. that carried very few information) were well prone to elimination but in our case, very few variables of this kind were available. We could also work by an iterative process of elimination of neighboring variables (i.e. highly correlated) [27,28] until the PCA become unstable. Nevertheless, if the variables were gathered in a cloud of points as it was in our case, it would be difficult to choose objectively the best variables eligible for the final selected variable set. Instead, a rational reduction methodology was set up in order to select the most informative variables while preserving classifications by combining recognition pattern and physico-chemical constraints. Minimizing the duration of the test was looked for through the elimination of chromatographic conditions, i.e. pH levels, solvent fractions and solutes. Effectively, it was not conceivable in practice to obtain a "reduced" test with 24 different eluents, meaning that at least 1 day was wasted for equilibration. The reduction principle was based on finding the minimum number of levels necessary for preserving the original classifications, i.e. founding the same clusters than those revealed in the full test. Moreover, another constraint was also applied, that could serve as a second compromise: the final classifications had to be still interpretable otherwise the test would have no reason to be. The pH level reduction was applied before the solvent fraction one.

3.3.2. pH level reduction

3.3.2.1. Retention factors. Fig. 5 represents the interpreted PCA obtained for each pH level including the clusters found by HCA.

The study of the clusters, showed that the original classification was well preserved at pH 5 and 7 while it was modified at pH 3. At this last level, the split due to solvent became blurred, especially for columns with low retentive power. Moreover, the classifications at the two other pHs were a bit refined, leading to a better discrimination according to hydrophobicity: indeed, the column Symmetry C_{18} was no more gathered with intermediate retentive power columns. The fact that clustering was better in the upper pH levels could be the result of the discarding of a pH level that made the test lose its discriminating power, confirming furthermore the necessity of reducing the variable set. From a physico-chemical point of view, it was not surprising to select the upper levels: actually, at such pH values, residual silanol groups are prone to partial ionization that can be well probed by basic compounds also partially ionized. In other words, these levels were likely more balanced than the first one, for which residual silanols are nearly undissociated and basic compounds are definitely ionized, leading to interactions that are both less informative and reliable. At this step, pH 5 and 7 were both eligible.

3.3.2.2. Peak asymmetries. In order to select if possible only one pH level, it was proceeded likewise on peak asymmetries. As pH 3 level has been rejected previously, its relative results will not be shown. Fig. 6 represents the results obtained at pH 5 and 7.

The split due to solvent seemed to be more pronounced in comparison to the full test classification. This could be attributed to the dispersive effect of the solvents that is more noticeable at the considered pHs: for basic compounds, the peak asymmetries are inclined to higher values because of the appearance of a mixed nature for retention mechanism. However, the cluster number needed to perform a reliable classification remained almost the same: 11 or 12 clusters were obtained at the pH level reduction step. In addition, it turned out that there was a loss of discrimination at pH 7 between polar embedded alkyl grafts according to their nature. As a consequence, pH 7 was eliminated.

pH 5 was revealed as the only level that preserved the same clusters as those obtained with the full test for retention, asymmetries and efficiencies (results not shown). Its selection allowed us to reduce the full test by a factor of 3 without significant loss of information.

3.3.3. Solvent fraction level reduction

As said previously, the optimal constraint consisted in finding the minimal number of conditions without major disturbance in the classifications. As the choice of only one solvent fraction appeared unsuccessful, the combination of two solvent fractions was then studied.

3.3.3.1. Retention factors. Fig. 7 shows the score plot obtained with the only two solvent fractions able to preserve obviously the original classification on retention factors: A and D.

Twelve clusters were found, still allowing a classification according to the solvent nature and the hydrophobicity of the tested stationary phases. As already noticed in Fig. 3, the less retentive phases were gathered if the dendogram cut level was shift towards greater dissimilarities.

3.3.3.2. Peak asymmetries. Fig. 8 summarizes the classification obtained within the conditions formerly described.

The classification was preserved with 13 clusters instead of 11 (see Fig. 4). In such conditions, the test was still able to discriminate between polar embedded stationary phases according to the shielding group nature. The representation was also subjected to a slight counterclockwise rotation in comparison to the original one that could be

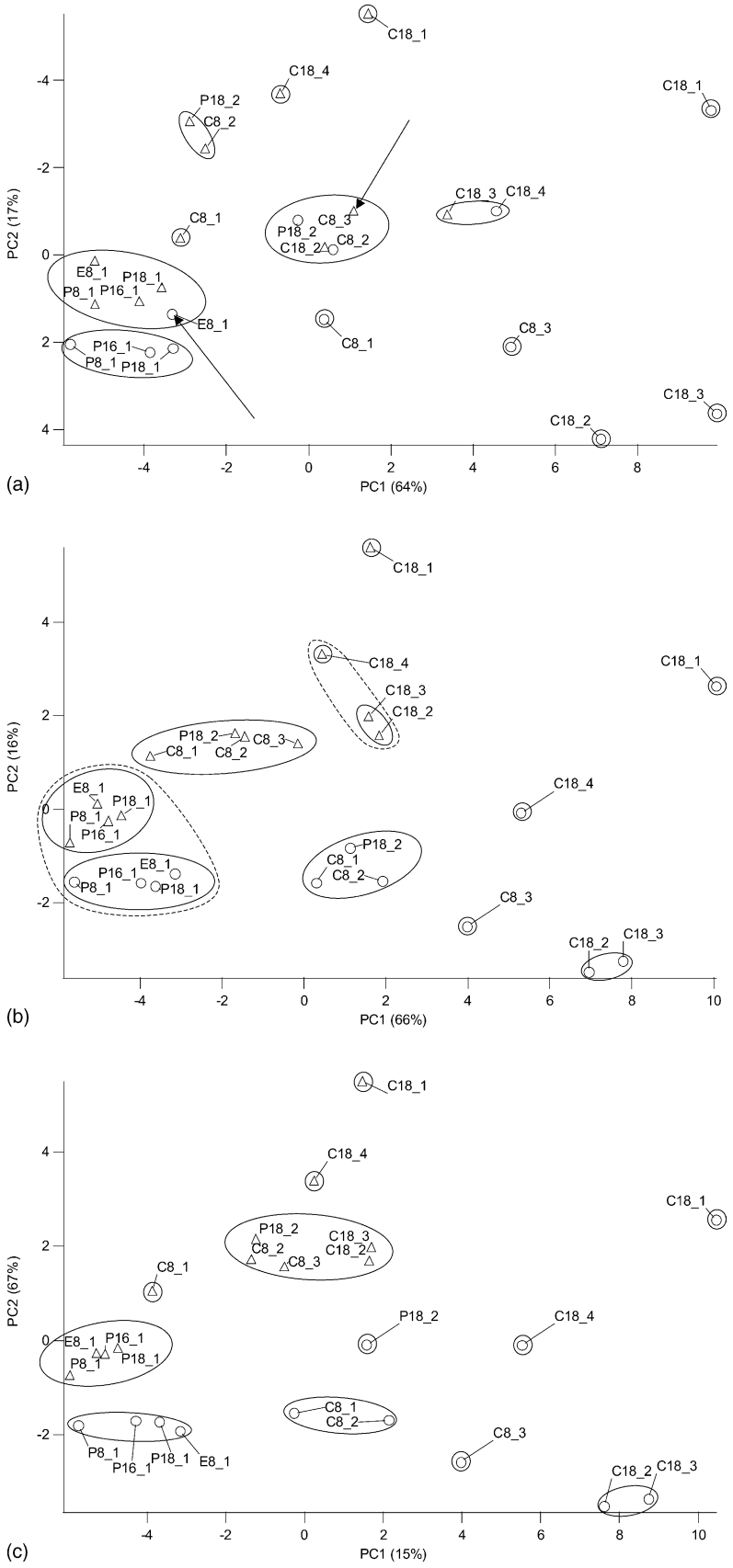

Fig. 5. Interpreted PC1–PC2 score plot of the column–solvent couples constructed on retention factors at each pH level: (a) pH 3, the arrows indicate the mismatches between the solvent–column couple and the cluster it belongs to; (b) pH 5; (c) pH 7.

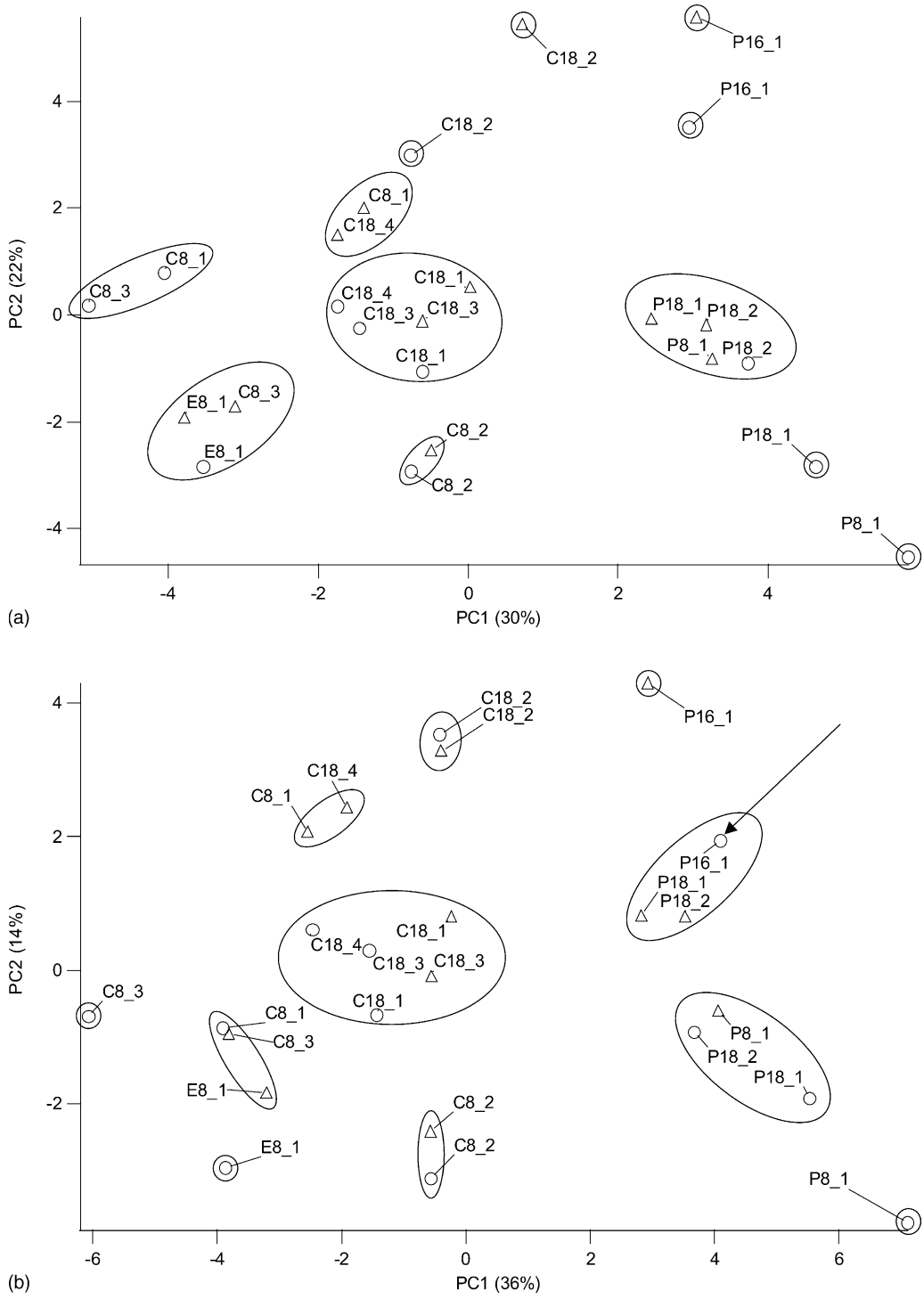

Fig. 6. Interpreted PC1–PC2 score plot of the column–solvent couples obtained with peak asymmetries at each pH level: (a) pH 5; (b) pH 7, the arrow shows the mismatch of the alkylamide grafted column clustered with alkylcarbamate bonded columns.

attributed to a switch between the two first components, considering the magnitude of carried information by each component.

Such conditions revealed themselves also sufficient for keeping the original classification on efficiencies (results not shown).

3.3.4. Final reduced test

The conditions of the final reduced test are reported in Table 4.

We managed to keep a test with a wide range of probe solutes and performed in two solvents with only two solvent fractions at an intermediate pH. The test duration had been

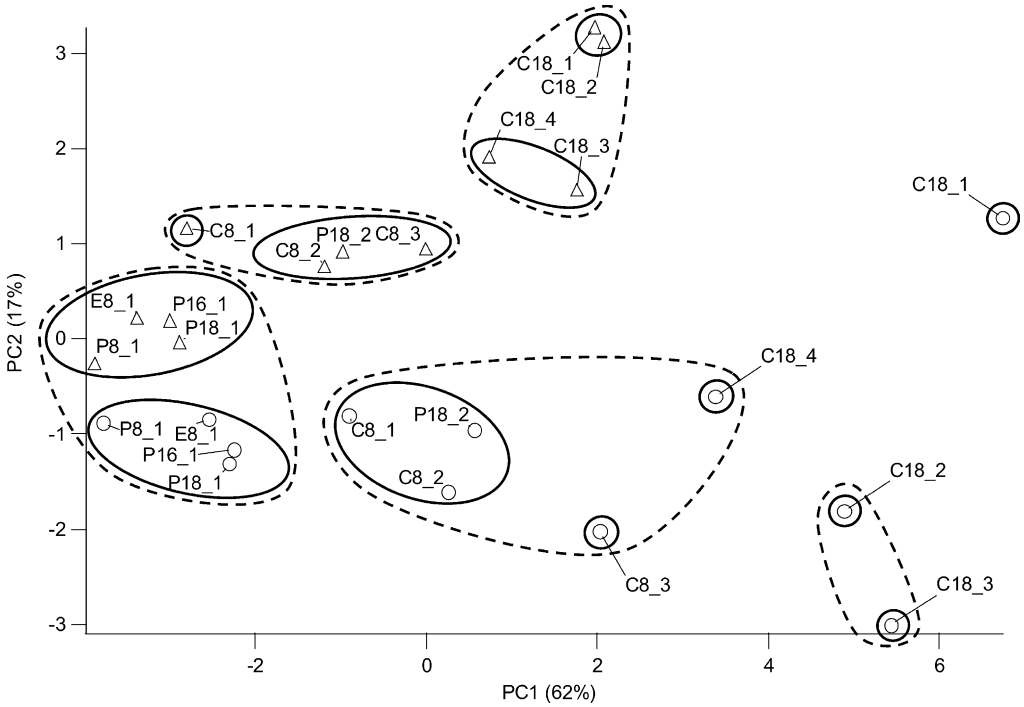

Fig. 7. Interpreted PC1–PC2 score plot of the column–solvent couples obtained with retention factors at the A–D solvent fraction levels.

divided by a factor of 6, and henceforth, the final reduced test can be carried out within 1 day or less for a whole characterization in two solvents. A parallel could be drawn between our results and the design of experiments approach: the test was based on one center point, i.e. pH, and its discrimination power was derived from wide variations operated with solvent fractions and good probing properties of our solutes. In addition, as the structure of the clusters remained almost the same before and after the reduction, and sometimes with some improvements of the classifications,

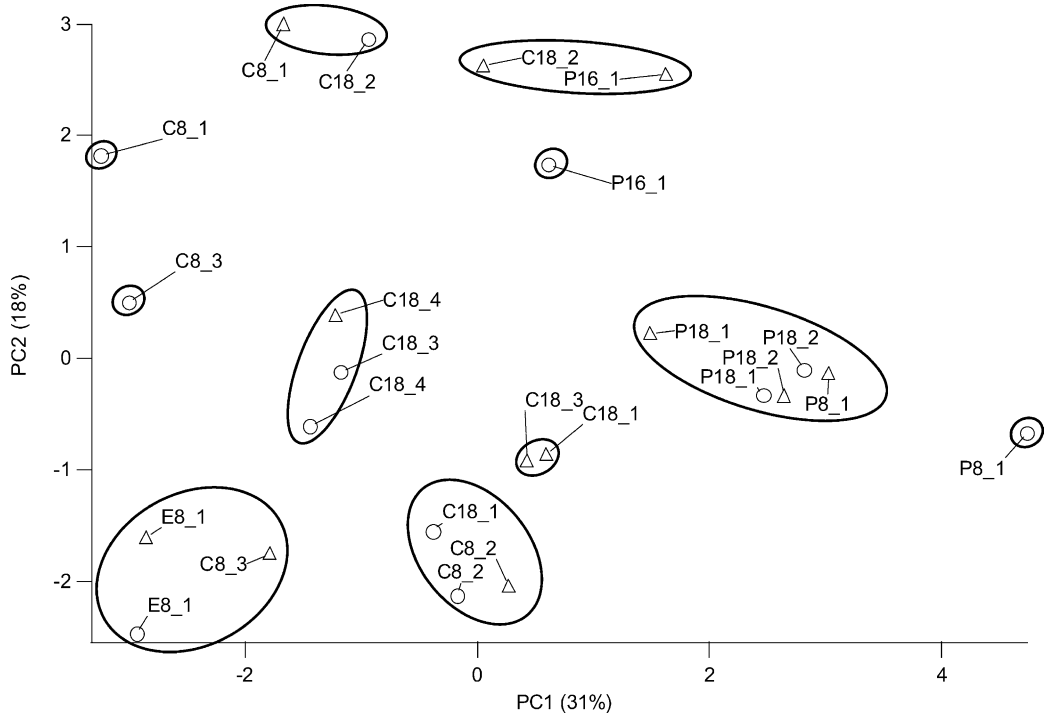

Fig. 8. Interpreted PC1–PC2 score plot of the column–solvent couples constructed on peak asymmetries at the couple A–D solvent fraction level.

Table 4
Chromatographic conditions of the final reduced test

Common condition	Solvent	Solvent fraction (%)	Solute
Acetate buffer at pH 5.00	MeOH	70	Digitoxin, clofazimine, amiodarone, butylbenzene*, pentylbenzene*, *o*-terphenyl*, triphenylene*
$T = 40\,^{\circ}C$	MeCN	59	
Flow rate $= 1\,ml\,min^{-1}$	MeOH	15	Strychnine*, benzylamine*, caffeine*, D-tubocurarine, atropine, ampicillin, vancomycin, cyanocobalamin
$\lambda = 254\,nm$ for solutes marked with '*', $\lambda = 230\,nm$ otherwise	ACN	9	

it could be asserted that the proposed methodology was performed without significant loss of information. Concerning the solutes of the final reduced test, diversity of the structures and properties was maintained, confirming the very need for different kind of solutes in case of a comprehensive testing procedure.

4. Conclusion

This study has shown that the proposed methodology for building a chromatographic test by an optimal approach was achievable. The reduced test was composed of conditions that had proved to be the more relevant ones by comparison with those of the "full" test. The combination of PCA and HCA proved to be an invaluable asset both for understanding classifications and selecting objectively the best conditions. Cluster analysis has also revealed as a useful tool for improving interpretation quality of the PC-score plots. The solvent effect was also confirmed, particularly with retention factors. Finally, the test derived from the "full" one was able to discriminate between columns of this first set. Henceforth it has to be performed on an extended column database to confirm its characterization power, to refine interpretations thanks to other kinds of special base columns and to be really of use for method developers in the pharmaceutical industry.

Acknowledgements

We would like to thank Aventis Pharma (Centre de recherche de Vitry-Alfortville) for supporting this work, and Macherey-Nagel for lending kindly its column Nucleosil C_{18} HD.

References

[1] C. Stella, P. Seuret, S. Rudaz, P.-A. Carrupt, J.-Y. Gauvrit, P. Lanteri, J.-L. Veuthey, J. Sep. Sci. 25 (2002) 1351.
[2] H.A. Claessens, M.A. van Straten, C.A. Cramers, M. Jezierska, B. Buszewski, J. Chromatogr. A 826 (1998) 135.
[3] H.A. Claessens, Trends Anal. Chem. 20 (2001) 563.
[4] U.D. Neue, E. Serowik, P. Iraneta, B.A. Alden, T.H. Walter, J. Chromatogr. A 849 (1999) 87.
[5] U.D. Neue, B.A. Alden, T.H. Walter, J. Chromatogr. A 849 (1999) 101.
[6] U.D. Neue, K. VanTran, P.C. Iraneta, B.A. Alden, J. Sep. Sci. 26 (2003) 174.
[7] L.C. Sander, S.A. Wise, J. Sep. Sci. 26 (2003) 283.
[8] L.C. Sander, S.A. Wise, in: Certificate of Analysis, NIST, Gaithersburg, MD, 1998.
[9] R.J.M. Vervoort, E. Ruyter, A.J.J. Debets, H.A. Claessens, C.A. Cramers, G.J. de Jong, J. Chromatogr. A 931 (2001) 67.
[10] R.J.M. Vervoort, E. Ruyter, A.J.J. Debets, H.A. Claessens, C.A. Cramers, G.J. de Jong, J. Chromatogr. A 964 (2002) 67.
[11] P. Jandera, M. Halama, K. Novotna, S. Buncekova, Chromatographia 57 (2003) S153.
[12] M.R. Euerby, P. Petersson, J. Chromatogr. A 994 (2003) 13.
[13] M.R. Euerby, A.P. McKeown, P. Petersson, J. Sep. Sci. 26 (2003) 295.
[14] M.R. Euerby, P. Petersson, LCGC Europe 13 (2000) 665.
[15] C. Stella, S. Rudaz, J.-L. Veuthey, A. Tchapla, Chromatographia 53 (2001) S132.
[16] A. Sandi, M. Nagy, L. Szepesy, J. Chromatogr. A 893 (2000) 215.
[17] S.D. Rogers, J.G. Dorsey, J. Chromatogr. A 892 (2000) 57.
[18] M. Kele, G. Guiochon, J. Chromatogr. A 830 (1999) 41.
[19] D.V. McCalley, J. Chromatogr. A 844 (1999) 23.
[20] D.V. McCalley, J. Chromatogr. A 769 (1997) 169.
[21] D.V. McCalley, J. Sep. Sci. 26 (2003) 187.
[22] H. Engelhardt, M. Arangio, T. Lobert, LCGC International (1997) 803.
[23] K. Kimata, K. Iwaguchi, S. Onishi, K. Jinno, R. Eksteen, K. Hosoya, M. Araki, N. Tanaka, J. Chromatogr. Sci. 27 (1989) 721.
[24] M.J. Walters, J. Assoc. Off. Anal. Chem. 70 (1987) 465.
[25] W.E. May, N.M. Trahey, in: Certificate of Analysis, NIST, Gaitherburgs, MD, 2000.
[26] D. Visky, Y. Vander Heyden, T. Ivanyi, P. Baten, J. De Beer, Z. Kovacs, B. Noszal, E. Roets, D.L. Massart, J. Hoogmartens, J. Chromatogr. A 977 (2002) 39.
[27] T. Ivanyi, Y. Vander Heyden, D. Visky, P. Baten, J. De Beer, I. Lazar, D.L. Massart, E. Roets, J. Hoogmartens, J. Chromatogr. A 954 (2002) 99.
[28] D. Visky, Y. Vander Heyden, T. Ivanyi, P. Baten, J. De Beer, Z. Kovacs, B. Noszal, P. Dehouck, E. Roets, D.L. Massart, J. Hoogmartens, J. Chromatogr. A 1012 (2003) 11.
[29] D.V. McCalley, J. Chromatogr. A 738 (1996) 169.
[30] K. Le Mapihan, J. Vial, A. Jardy, Chromatographia 57 (2003) S163.
[31] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S.D. Jong, P.J. Lewi, J. Smeyers-Verbeke, in: D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S.D. Jong, P.J. Lewi, J. Smeyers-Verbeke (Eds.), Handbook of Chemometrics and Qualimetrics, Part A, Elsevier, Amsterdam, 1997, p. 519.
[32] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S.D. Jong, P.J. Lewi, J. Smeyers-Verbeke, in: D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S.D. Jong, P.J. Lewi, J. Smeyers-Verbeke (Eds.),

Handbook of Chemometrics and Qualimetrics, Part B, Elsevier, Amsterdam, 1997, p. 87.
[33] J.N. Miller, J.C. Miller, in: J.N. Miller, J.C. Miller (Eds.), Statistics and Chemometrics for Analytical Chemistry, Prentice-Hall, Harlow, 2000, p. 214.
[34] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S.D. Jong, P.J. Lewi, J. Smeyers-Verbeke, in: D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S.D. Jong, P.J. Lewi, J. Smeyers-Verbeke (Eds.), Handbook of Chemometrics and Qualimetrics, Part B, Elsevier, Amsterdam, 1997, p. 57.
[35] B. Massart, Q. Guo, F. Questier, D.L. Massart, C. Boucon, S. de Jong, B.G.M. Vandeginste, Trends Anal. Chem. 20 (2001) 35.
[36] R. De Maesschalck, D. Jouan-Rimbaud, D.L. Massart, Chemom. Intell. Lab. Syst. 50 (2000) 1.