Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 988 (2003) 77–93
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Determining orthogonal chromatographic systems prior to the development of methods to characterise impurities in drug substances

E. Van Gyseghem, S. Van Hemelryck, M. Daszykowski, F. Questier, D.L. Massart, Y. Vander Heyden*

Department of Pharmaceutical and Biomedical Analysis, Pharmaceutical Institute, Vrije Universiteit Brussel, Laarbeeklaan 103, B-1090 Brussels, Belgium

Received 19 July 2002; received in revised form 25 October 2002; accepted 3 December 2002

Abstract

To define starting conditions for the development of methods to separate impurities from the active substance and from each other in drugs with an unknown impurity profile, the parallel application of generic orthogonal chromatographic systems could be useful. The possibilities to define orthogonal chromatographic systems were examined by calculation of the correlation coefficients between retention factors *k* for a set of 68 drugs on 11 systems, by visual evaluation of the selectivity differences, by using principal component analysis, by drawing color maps and evaluating dendrograms. A zirconia-based stationary phase coated with a polybutadiene (PBD) polymer and three silica-based phases (base-deactivated, polar-embedded and monolithic) were used. Besides the stationary phase, the influence of pH and of organic modifier, on the selectivity of a system were evaluated. The dendrograms of hierarchical clusters were found good aids to assess orthogonality of chromatographic systems. The PBD–zirconia phase/methanol/pH 2.5 system is found most orthogonal towards several silica-based systems, e.g. a base-deactivated C_{16}-amide silica/methanol/pH 2.5 system. The orthogonality was validated using cross-validation, and two other validation sets, i.e. a set of non-ionizable solutes and a mixture of a drug and its impurities.
© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Orthogonal chromatographic systems; OPTICS color map; Average linkage dendrogram; Dendrogram color map; Drug substance impurities; Stationary phases, LC

1. Introduction

The characterisation of impurities in drugs is important to protect the patient from negative side effects. The ICH (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use) requires for safety reasons that impurities which are present in a quantity exceeding a certain limit should be characterised, i.e. identified and/or quantified (e.g. chromatographically) [1]. Analytically, it is a big challenge to determine all impurities potentially occur-

*Corresponding author. Tel.: +32-2-477-4723; fax: +32-2-477-4735.
E-mail address: yvanvdh@fabi.vub.ac.be (Y. Vander Heyden).

0021-9673/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0021-9673(02)02012-5

ring in a drug. Indeed, one is dealing with relatively small quantities, and moreover by- or degradation products often exhibit chemical structures which are very similar to that of the active substance. In addition, one initially does not know the structure, nor the number of impurities, which makes method development far from being evident. To determine initial starting conditions for method development and to obtain an idea about the composition of the unknown mixture, different chromatographic systems might be tested in parallel. Therefore, attempts are made to select systems which are as orthogonal as possible. Orthogonal chromatographic systems have a strongly different selectivity, because the retention in these systems is caused by different mechanisms. As a result, application of a new drug and its impurities on parallel orthogonal systems can maximize the probability that all substances will be revealed. The aim of this study is to evaluate possibilities to define a set of orthogonal reversed-phase chromatographic systems, which could be used as starting conditions for further method development to separate mixtures. It is not our intention, either to combine orthogonal systems to obtain two-dimensional chromatographic separations [2], nor to select orthogonal techniques, e.g. HPLC versus CE, as some authors described earlier [3]. In the literature, both classical stationary phases and more recently developed ones have been examined to select orthogonal systems [4–8]. Among the latter, fluorinated surfaces [4] or more pH inert materials, like base-deactivated [4] and polar-embedded silica [4,7,8] were considered. Also non silica-based materials such as graphitized carbon-clad-, polybutadiene- or polystyrene-coated zirconia [4,5], titania- [8] or alumina-based [6] stationary phases are potentially interesting. One can expect separations which differ from those obtained with conventional RP-stationary phases [4,5], because additional interactions with the solutes are possible [9]. In the literature [5,6], examples are shown of stationary phases with a different basis, having major selectivity differences.

As a chromatographic system also consists of a mobile phase, the influence of the buffer pH [5], the organic modifier type [10,11], or the temperature [12] on the selectivity can be researched. It has, for instance, been shown that the buffer pH can have a large influence on the selectivity [6]. From this point of view, some new stationary phases, e.g. zirconia-based ones, have an additional advantage because they exhibit stability over a wide pH range [4].

The aim of this work is to evaluate which data treatment is appropriate to select orthogonal chromatographic systems. Therefore, chromatographic systems with different stationary phases, buffer pH and organic modifiers were examined. Three silica-based and one zirconia-based stationary phases were evaluated. The silica-based ones were (i) a non-endcapped and intrinsically base-deactivated phase with an embedded polar hexadecylamide group, (ii) an octadecyl phase with polar endcapping, and (iii) an endcapped monolithic rod of polymerised silica [13–15]. The first two phases are modified with a polar group at or close to their silica surface, and are called high aqueous, polar-embedded [4,7,8], or intrinsically base-deactivated [4]. The zirconia-based phase was coated with polybutadiene, which is responsible for its reversed-phase properties. When using an aqueous mobile phase, ion and/or ligand exchange phenomena can occur on this column besides the reversed-phase ones [9]. Zirconia namely has an amphoteric nature and can act both as cation and anion-exchanger. Interactions with the zirconia surface due to incomplete coverage by the polymer can be a problem with these phases, which can be inhibited using phosphate- or borate-containing mobile phases [9]. On silica phases, ligand-exchange does not occur.

A silica- and an alumina-based stationary phase already proved to give different information [6]. Zirconia-based stationary phases were found to have a selectivity similar to the alumina analogues [5]. The zirconia-based phase was selected, because of its better pH and thermal stability [9,16]. On these stationary phases, mobile phases with different organic modifiers (methanol and acetonitrile) and buffers at both low and high pH were used. A total of 11 chromatographic systems were examined.

To evaluate the orthogonality of the chromatographic systems, a set of 68 drugs was injected. The substances chosen differ in structure (functional groups, ring structures), molecular mass, pK_a and pharmacological class. Eventually, the selectivity

differences of the most orthogonal chromatographic systems were validated with sets of test substances.

2. Experimental

2.1. Drugs and reagents

The 68 substances and their used concentrations are summarized in Table 1. Validation sets were created consisting of subsets of the 68 compounds. Other substances used to evaluate the orthogonality of systems were: amylbenzene (Sigma–Aldrich, Steinheim, Germany), toluene (Merck), butylbenzene, ethylbenzene, *o*-terphenyl, triphenylene and uracil (Fluka Chemie, Buchs, Switzerland), tetracycline, 4-epitetracycline, 4-epianhydrotetracycline and anhydrotetracycline (all gifts from Professor J. Hoogmartens, Katholieke Universiteit Leuven, Belgium).

The mobile phases were prepared using acetonitrile and methanol, Hypersolv for HPLC (BDH, Poole, UK), phosphoric acid solution minimum 85% (Carlo Erba, Milan, Italy), acetic acid, anhydrous disodium tetraborate, boric acid, di-sodium hydrogen phosphate dihydrate, ethylenediamine tetraacetic acid disodium salt dihydrate, sodium dihydrogen phosphate monohydrate, sodium hydroxide pellets, all pro analysi (GR) quality, sodium hydroxide solution 1 N, hydrochloric acid 5 N (all from Merck).

2.2. Chromatographic equipment and conditions

The HPLC-instrument consisted of an L-7100 pump, an L-7612 solvent degasser, an L-7250 programmable autosampler with a 100 μl loop, an L-7350 column oven, an L-7400 UV detector, and a D-7000 interface (Hitachi, Tokyo, Japan). The chromatographic data were gathered and treated with the D-7000 HPLC System Manager software (Hitachi). Detection was carried out at 254 nm and the temperature of the column oven was kept at 40 °C.

The stationary phases tested were: (a) a base-deactivated hexadecylsilica, Suplex pKb-100, (150×4.6 mm I.D., 5 μm) (Supelco, Bellefonte, PA, USA), (b) a zirconia-based phase coated with polybutadiene-polymer, ZirChrom-PBD, (100×4.6 mm I.D., 3 μm) (ZirChrom Separations, Anoka, MN, USA), (c) a monolithic phase, Chromolith Performance, RP-18e (100×4.6 mm I.D.) (Merck), and (d) a polar-embedded octadecylsilica phase, Aqua, (100×4.6 mm I.D., 3 μm) (Phenomenex, Torrance, CA, USA).

Methanol as organic modifier was used on Suplex pKb-100 and ZirChrom-PBD, acetonitrile on Suplex pKb-100, Chromolith Performance and Aqua. Initially, on the Suplex pKb-100 phase, a Britton–Robinson buffer, as in Ref. [17], was used at pH 2.5 and 7.5. The buffer contains 0.04 *M* of phosphoric, acetic and boric acid, and is brought to pH with 0.2 *M* sodium hydroxide [17]. This buffer was also used on the zirconia-based phase, at pH values of 2.5 and 7.5, while at pH 10.0, a 0.016 *M* borate buffer was used. On the other phases, and also on some systems containing the Suplex pKb-100 column, phosphate buffers with different molarities were used. These buffers resulted from the search to improve peak shapes on the systems. The different systems are described in Table 2.

The pH of the buffers was adjusted using an Orion 520A (Orion Research, Boston, MA, USA) pH-meter, which was calibrated daily. Buffers were filtered through a 0.2 μm membrane filter (Schleicher & Schuell, Dassel, Germany). In buffers, stock solutions and samples, Milli-Q water is used, prepared with the Millipore purification system (Millipore, Molsheim, France).

Each substance was dissolved in non-buffered solvent equivalent in water/modifier composition to the mobile phase at the beginning of the gradient runs (Table 2) and was individually injected. Its concentration depended on the absorbance at 254 nm.

3. Results and discussion

The orthogonality of the chromatographic systems was studied using a set of 68 pharmaceutical substances with different properties (Table 1), which were injected in the 11 systems (CS1–CS11) described in Table 2. Most substances were basic (55), because most pharmaceuticals are. Some neutral (4) and acidic substances (9) also were included.

Table 1
The 68 substances used, their concentrations and distributors

Substance (concentration, mg/l)	Distributor
(±)-Camphor (5000)	Sigma–Aldrich (Steinheim, Germany)
1,1-Dimethylbiguanide hydrochloride (1000)	Sigma–Aldrich (Steinheim, Germany)
4-Benzylphenol (1000)	Aldrich (Milwaukee, WI, USA)
5-Hydroxytryptamine hydrochloride (500)	Sigma–Aldrich (Steinheim, Germany)
5-Sulfosalicylic acid dihydrate (2000)	Merck (Darmstadt, Germany)
Acebutolol hydrochloride (1000)	Sigma (St. Louis, MO, USA)
Amiodarone hydrochloride (5000)	Clin-Midy groupe Sanofi (Montpellier, France)
Antazoline hydrochloride (1000)	Sigma–Aldrich (Steinheim, Germany)
Betaxolol hydrochloride (1000)	Synthelabo (Paris, France) (gift)
Bupranolol hydrochloride (1000)	Schwarz Pharma (Monheim, Germany)
Caffeine (1000)	Fluka (Neu-Ulm, Switzerland)
Carbamazepine (1000)	Sigma–Aldrich (Steinheim, Germany)
Celiprolol (1000)	Rhône-Poulenc-Rorer (Madrid, Spain) (gift)
Chloropyramine hydrochloride (1000)	Sigma–Aldrich (Steinheim, Germany)
Cimetidine (10000)	Penn Chemicals (Pittsburg, PA, USA) (gift)
Cirazoline hydrochloride (400)	Research Biochemicals International (Natick, MA, USA)
Cocaine hydrochloride (1000)	Bios Coutelier (Brussels, Belgium)
Codeine base (1000)	Bios Coutelier (Brussels, Belgium)
Desipramine hydrochloride (000)	Sigma–Aldrich (Steinheim, Germany)
Diclofenac sodium (5000)	Sigma–Aldrich (Steinheim, Germany)
Digitoxigenine (500)	Fluka (Neu-Ulm, Switzerland)
Digitoxine (1000)	Mann Research Laboratories (New York, NY, USA)
Dimetindene maleate (1000)	Novartis (Basel, Switzerland) (gift)
Diphenhydramine hydrochloride (5000)	Sigma–Aldrich (Steinheim, Germany)
Dopamine hydrochloride (2000)	Sigma–Aldrich (Steinheim, Germany)
Efedrine hydrochloride (2000)	Vel (Leuven, Belgium)
Famotidine (2000)	Sigma–Aldrich (Steinheim, Germany)
Fenfluramine hydrochloride (1000)	Technologie Servier (Orleans, France)
Fluphenazine dihydrochloride (USP grade) (2000)	Sigma–Aldrich (Steinheim, Germany)
Flurazepam (1000)	Dolorgiet Arzneimittel (Bonn, Germany)
Histamine dihydrochloride (1000)	Sigma–Aldrich (Steinheim, Germany)
Ibuprofen (5000)	Sigma–Aldrich (Steinheim, Germany)
Isothipendyl hydrochloride (1000)	Novartis Pharma (Wehr, Austria) (gift)
Ketotifen fumarate (1000)	Sigma–Aldrich (Steinheim, Germany)
L-(+)-ascorbic acid (1000)	Merck (Darmstadt, Germany)
Lidocaine hydrochloride (1000)	Bios Coutelier (Brussels, Belgium)
Lorazepam (1000)	MSD (Haarlem, The Netherlands)
Miconazol nitrate (1000)	Certa (Braine-l'Alleud, Belgium)
Morphine hydrochloride (2000)	Bios Coutelier (Brussels, Belgium)
Nadolol (1000)	Sigma–Aldrich (Steinheim, Germany)
Naphazoline hydrochloride (2000)	Sigma–Aldrich (Steinheim, Germany)
Nicardipine hydrochloride (1000)	UCB (Leuven, Belgium)
Nizatidine (2000)	Norgine (Marburg, Germany) (gift)
Oxeladin citrate (2000)	Sigma–Aldrich (Steinheim, Germany)
Oxprenolol hydrochloride (500)	Sigma–Aldrich (Steinheim, Germany)
Pentoxifylline (1000)	Sigma–Aldrich (Steinheim, Germany)
Phenol (1000)	Merck (Darmstadt, Germany)
Pindolol (1000)	Sigma–Aldrich (Steinheim, Germany)
Pizotifen (5000)	Novartis Pharma (Wehr, Austria) (gift)
Prazosin hydrochloride (1000)	Sigma–Aldrich (Steinheim, Germany)
Prenalterol hydrochloride (1000)	Ciba-Geigy (Basel, Switzerland)
Procaine hydrochloride (1000)	Merck (Darmstadt, Germany)

Table 1. Continued

Substance (concentration, mg/l)	Distributor
Promethazine hydrochloride (1000)	Sigma–Aldrich (Steinheim, Germany)
Propiomazine maleate (1000)	Sanofi (Paris, France) (gift)
Pyrilamine maleate (1000)	Sigma–Aldrich (Steinheim, Germany)
Ranitidine hydrochloride (2000)	Sigma (St. Louis, MO, USA)
Resorcine (1000)	Merck (Darmstadt, Germany)
Sotalol (1000)	Merck (Darmstadt, Germany)
Strychnine base (1000)	Bios Coutelier (Brussels, Belgium)
Sulfapyridine (1000)	Bios Coutelier (Brussels, Belgium)
Terazosin hydrochloride (1000)	Sigma–Aldrich (Steinheim, Germany)
Terbutaline sulphate (1000)	Astra Draco (Lund, Sweden)
Tetrahydrozolin hydrochloride (4000)	U.S.P.C. (Rockville, MD, USA)
Thiothixene (USP grade) (2000)	Sigma–Aldrich (Steinheim, Germany)
Timolol maleate (1000)	Sigma–Aldrich (Steinheim, Germany)
Tolazoline hydrochloride (5000)	Sigma–Aldrich (Steinheim, Germany)
α-Lobeline hydrochloride (1500)	Carl Roth (Karlsrhue, Germany)
β-Estradiol (500)	Sigma–Aldrich (Steinheim, Germany)

To limit the retention time of late eluting components, gradient elution was performed. For each mobile phase, it was evaluated which range of organic modifier/buffer (%, v/v) ratios can be applied. The lowest ratio (highest water fraction) was chosen so that it does not give a problem concerning the solubility of the substances, nor that stationary phase collapse occurs (monolithic or base-deactivated phases). The highest ratio for which the buffer salts do not precipitate was determined, and a margin of 5% below it was applied. For a given system (column, organic modifier and pH), higher ratios can

Table 2
Description of the chromatographic systems (CS)

CS	Stationary phase	Mobile phase
1	Suplex pKb-100	Methanol/Britton–Robinson buffer pH 2.5 from 30:70 to 75:25% (v/v) in 20 min; flow-rate 1.0 ml/min
2	Suplex pKb-100	Methanol/Britton–Robinson buffer pH 7.5 from 30:70 to 70:30% (v/v) in 10 min; flow-rate 2.0 ml/min
3	ZirChrom PBD	Methanol/Britton–Robinson buffer pH 2.5 from 30:70 to 75:25% (v/v) in 20 min; flow-rate 1.0 ml/min
4	ZirChrom PBD	Methanol/Britton–Robinson buffer pH 7.5 from 30:70 to 70:30% (v/v) in 20 min; flow-rate 1.0 ml/min
5	ZirChrom PBD	Methanol/0.016 *M* borate buffer pH 10.0 from 30:70 to 75:25% (v/v) in 8 min; flow-rate 1.5 ml/min
6	Chromolith Performance	Acetonitrile/0.08 *M* sodium phosphate buffer pH 3.0 from 10:90 to 60:40% (v/v) in 6 min; flow-rate 2.0 ml/min
7	Chromolith Performance	Acetonitrile/0.08 *M* sodium phosphate buffer pH 7.5 from 10:90 to 60:40% (v/v) in 6 min; flow-rate 2.0 ml/min
8	Aqua	Acetonitrile/0.04 *M* sodium phosphate buffer pH 3.0 from 10:90 to 70:30% (v/v) in 8 min; flow-rate 1.0 ml/min
9	Aqua	Acetonitrile/0.04 *M* sodium phosphate buffer pH 6.8 from 10:90 to 75:25% (v/v) in 4 min; flow-rate 2.0 ml/min
10	Suplex pKb-100	Acetonitrile/0.04 *M* sodium phosphate buffer pH 3.0 from 10:90 to 70:30% (v/v) in 8 min; flow-rate 1.0 ml/min
11	Suplex pKb-100	Acetonitrile/0.04 *M* sodium phosphate buffer pH 6.8 from 10:90 to 70:30% (v/v) in 8 min; flow-rate 1.0 ml/min

be used with a lower buffer concentration. For instance, 0.04 *M* phosphate buffer pH 3.0 could be used up to 70:30% (v/v) acetonitrile/buffer (CS8 and CS10), whereas 0.08 *M* phosphate buffer pH 3.0 was limited to 60:40% (CS6).

Initially, in analogy with Ref. [17], the Britton–Robinson buffer was used (CS1–CS4), for which buffering occurs in the whole pH range. Application of this buffer at pH 10.0 on CS5 gave broad asymmetrical peaks, while a 0.016 *M* borate buffer gave more symmetric and narrow peaks. Therefore, to decrease the total salt concentration on the other systems, buffers with salts only buffering around the studied pH were used. Good peak shapes were obtained with phosphate buffer concentrations of either 0.08 *M* (CS6–CS7) or 0.04 *M* (CS8–CS11).

3.1. Defining the most orthogonal systems

The goal of this study was to determine a suitable data treatment to select orthogonal systems. For each substance, its retention factor *k* was determined on each system. The orthogonality of the eleven chromatographic systems was in a first instance evaluated by calculation of the Pearson-correlation coefficients *r* between the *k*-values for each pair of systems (Table 3). Selectivity differences between two systems were visualized by plotting the retention results versus each other. Initially, this was done in two different ways, once on parallel axes [18] and once on cartesian axes, of which the latter seem to be the best (Fig. 1).

Generally, the lower a correlation coefficient, the larger the selectivity differences, and the more the systems are orthogonal (see Fig. 1a,b). A high correlation coefficient indicates that a rather linear relationship is observed between the considered responses. The solutes are eluting in a similar order and in elution times which are proportional on both systems (Fig. 1b). Such systems deliver analogous information about the solutes, and are not orthogonal. In Fig. 1a, no proportional relationship was found between the *k*-values of the solutes on CS1 and CS3, which led to a low correlation coefficient.

However, a low correlation coefficient does not always indicate generic selectivity differences. For instance, when a majority of the solutes elutes in a similar elution order on both systems, but only few substances show large selectivity differences, a relatively small correlation coefficient can be obtained as well. In Fig. 1c, a simulated example is given, for which the *k* on CS*x* are similar to those on CS*y*, except for a relatively small number of solutes (14). When comparing Fig. 1a and c, it can be seen that Fig. 1a shows general selectivity differences, implying that CS1 and CS3 are rather orthogonal, while systems CS*x*–CS*y* are not orthogonal, although the correlation coefficient is rather small. Therefore, besides calculating the *r*-values, a visual interpretation on cartesian axes of the general changes in elution order can be recommended to decide on the orthogonality of systems.

From Table 3, it can be seen that the lowest correlation coefficients are obtained when comparing

Table 3
Matrix of the correlation coefficients between the retention factors *k*, measured on the different systems

	CS1	CS2	CS3	CS4	CS5	CS6	CS7	CS8	CS9	CS10	CS11
CS1	1										
CS2	0.363	1									
CS3	0.240	0.345	1								
CS4	0.424	0.753	0.592	1							
CS5	0.439	0.684	0.590	0.951	1						
CS6	0.698	0.496	0.419	0.654	0.644	1					
CS7	0.365	0.687	0.458	0.722	0.664	0.646	1				
CS8	0.694	0.521	0.365	0.684	0.680	0.737	0.577	1			
CS9	0.489	0.783	0.383	0.804	0.748	0.728	0.851	0.673	1		
CS10	0.750	0.444	0.327	0.566	0.555	0.916	0.557	0.737	0.682	1	
CS11	0.538	0.608	0.532	0.845	0.869	0.804	0.703	0.730	0.836	0.759	1

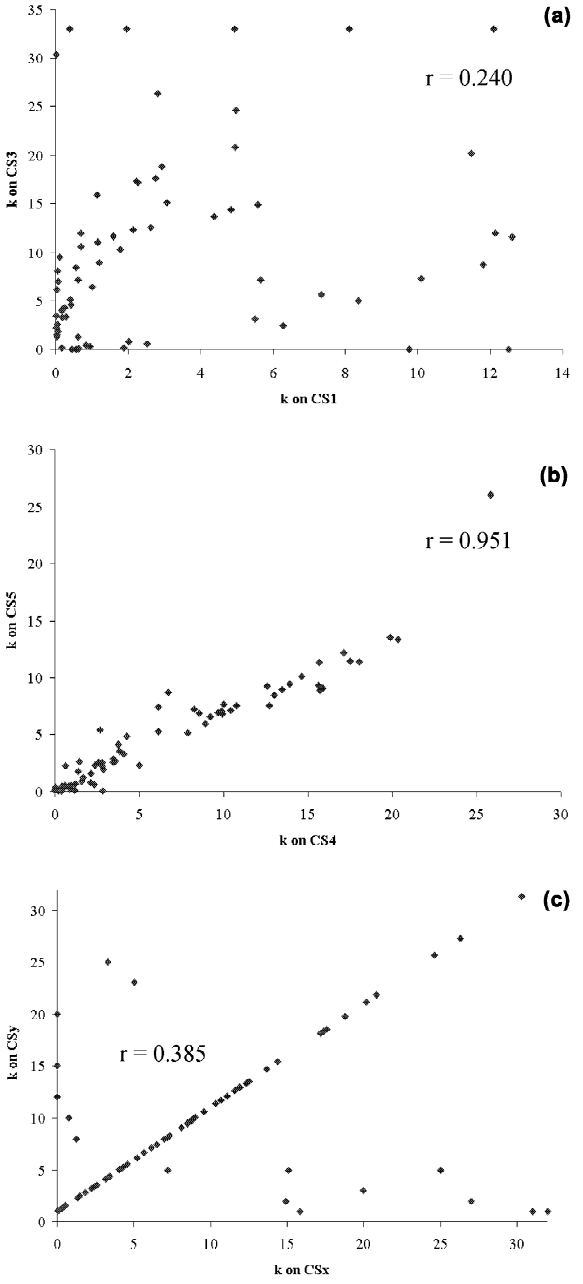


Fig. 1. Retention factors on two systems: (a) CS1–CS3, (b) CS4–CS5, and (c) hypothetical systems, CS*x*–C*sy*. *r* = correlation coefficient.

CS1–CS3 (r=0.240), CS3–CS10 (r=0.327), CS2–CS3 (r=0.345), CS3–CS8 (r=0.365), CS3–CS9 (r=0.383), CS1–CS2 (r=0.363) and CS1–CS7 (r=0.365). Figs. 1a and 2 visualize the differences in elution order of the 68 substances on the former five systems. To five substances on CS3, which were eluting late (k>33), an arbitrary large value was given (k=33). These high retention factors can be considered the consequence of both ion and ligand exchange present at a zirconia column at low pH. The zirconia-based stationary phase (CS3) is clearly behaving differently from the silica-based ones. Many changes in elution order occur between the zirconia and silica systems. Fig. 3 schematically represents the mutual relation between these systems, except for CS7, as this system only gives major selectivity differences when compared with CS1. Calculation of the correlation coefficients without considering the retention factors of the five late eluting compounds on CS3 even lowers the coefficient, for instance, r for CS1–CS3 becomes 0.170 instead of 0.240.

Large selectivity differences are obtained when comparing a PBD–zirconia phase, pH 2.5, methanol as system (CS3) with one of the systems of Group A (Fig. 3), i.e. either (i) Suplex pKb-100, pH 2.5, methanol (CS1), (ii) Suplex pKb-100, pH 3.0, acetonitrile (CS10), or (iii) Aqua, pH 3.0, acetonitrile (CS8), or with one of the systems of group B (Fig. 3), i.e. either (i) Suplex pKb-100, pH 7.5, methanol (CS2) or (ii) Aqua, pH 6.8, acetonitrile (CS9). The systems of group A all are highly correlated: CS1–CS8 (r=0.694), CS1–CS10 (r=0.750) and CS8–CS10 (r=0.737), indicating similar selectivities. This is also the case for the systems of group B: CS9 and CS2 (r=0.783). The correlation coefficients between the systems of groups A and B vary between 0.363 and 0.682 (Fig. 3). The selectivity differences for systems CS1–CS2 (r=0.363) still allow to consider them as orthogonal. The other system comparisons had intermediate correlation coefficients, and selectivity differences were more limited.

In summary, it can be said that a set of three orthogonal systems can be selected, i.e. CS3, one system of group A and one from group B. The systems from groups A and B are to be chosen in such a way that their correlation is low and that large and general selectivity differences will be observed

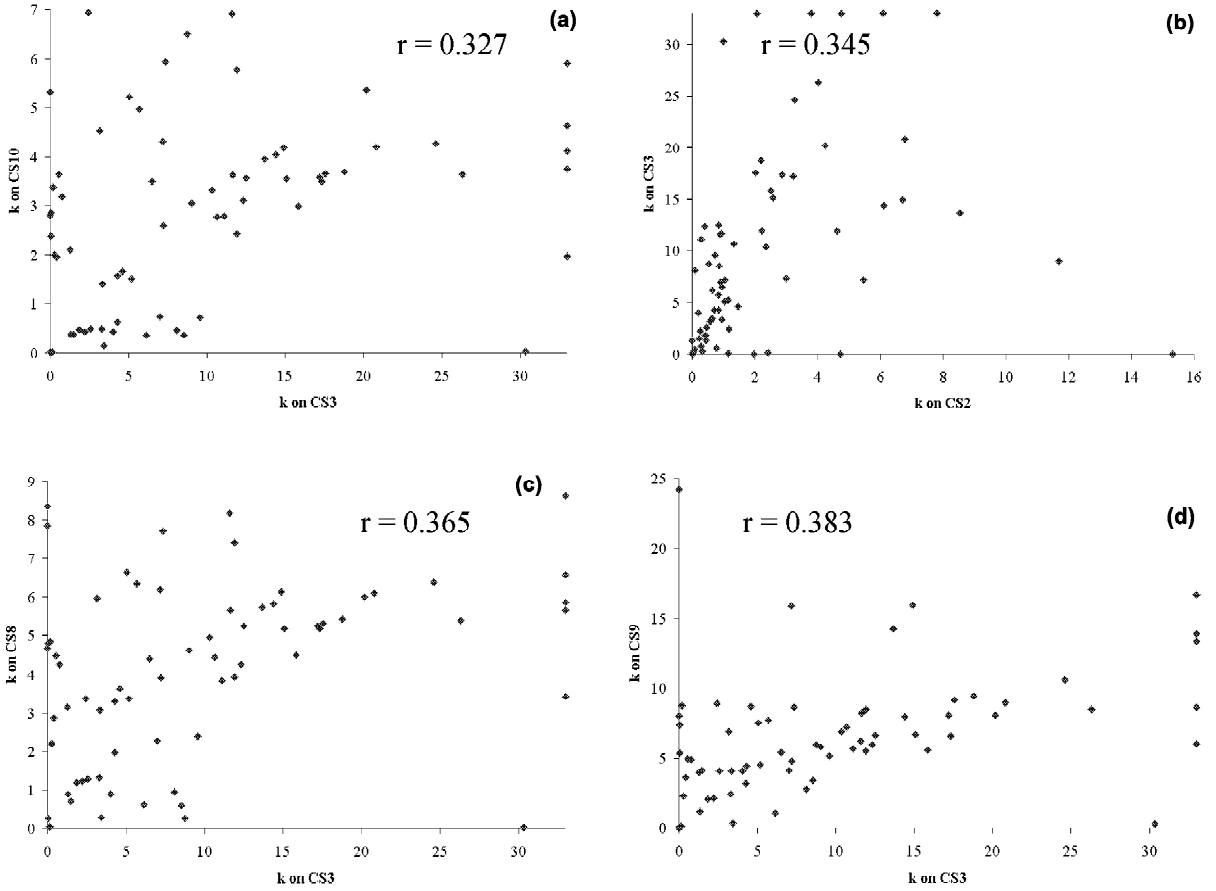


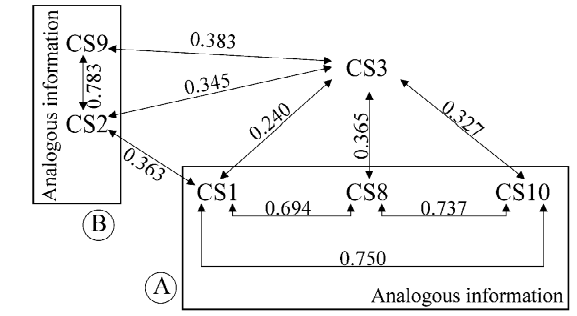
Fig. 2. Retention factors of the 68 substances on (a) CS3 versus CS10, (b) CS2 versus CS3, (c) CS3 versus CS8, and (d) CS3 versus CS9.

between the selected systems.

The above shows that the stationary phase is most responsible for the orthogonality of systems. This can be explained by the fact that the ''orthogonal'' stationary phases have different interaction possibilities with the solutes. At low pH, the zirconia stationary phase can show beside the reversed-phase properties, both anion and ligand exchange [9], and the lack of these interactions at a silica-based phase can be an explanation for the orthogonal results. At high pH, the former phase shows both ligand and cation-exchange [9], but only cation-exchange is possible on a silica-based stationary phase. Moreover, both the base-deactivated and the polar-embedded silicas are modified, so that polar silanophilic interactions are minimal.

Secondly, the pH also can have an effect on the selectivity. Indeed, the low correlation coefficient between CS1 (pH 2.5) and CS2 (pH 7.5) (r=0.363) expresses rather large changes in elution order. The influence of the pH could be explained both from changes in the ionisation of the solutes and in the stationary phase properties. Smaller differences in pH between two systems caused less selectivity differences.

Changing the organic modifier from methanol to acetonitrile did not lead to much selectivity differences, but it can enhance the selectivity differences induced by pH and/or stationary phase changes, e.g. CS10–CS11 (r=0.759), associated with a change in pH, versus CS2–CS10 (r=0.444), associated with a change in both pH and organic modifier.

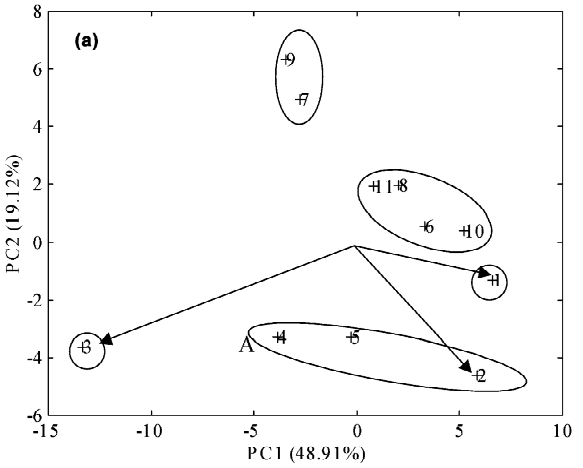


Pairs of systems	r
CS1-CS2	0.363
CS1-CS9	0.489
CS8-CS2	0.521
CS8-CS9	0.673
CS10-CS2	0.444
CS10-CS9	0.682

Fig. 3. Schematic representation of the mutual relation between the most orthogonal chromatographic systems. The values on the figure represent *r* values.

3.2. Relationships between all chromatographic systems

From the matrix of correlation coefficients (Table 3), it is not always evident, as was tried in Fig. 3, to evaluate the relationships between the researched systems, especially when their number becomes large. Therefore some visualisation methods were applied, namely principal component analysis (PCA) plots [19], an OPTICS color map [20] and dendrograms based on average linkage [19,21–24], in order to evaluate their usefulness. Fig. 4 represents the score plots of the autoscaled *k* data (11×68 matrix). The systems defined earlier as most orthogonal are indicated by means of arrows starting from the plot's origin. It can be observed that these systems are not always situated most extremely in these plots. Thus, only based on a visual inspection of the PCA plots it seems impossible to identify the most orthogonal systems. On the PCA plots also a number of clusterings/groups of columns can be distinguished (see Fig. 4). However, these clusters do not always situate the columns well relative to each other, at least not in the context of selecting orthogonal systems. For instance, cluster A (Fig. 4a)



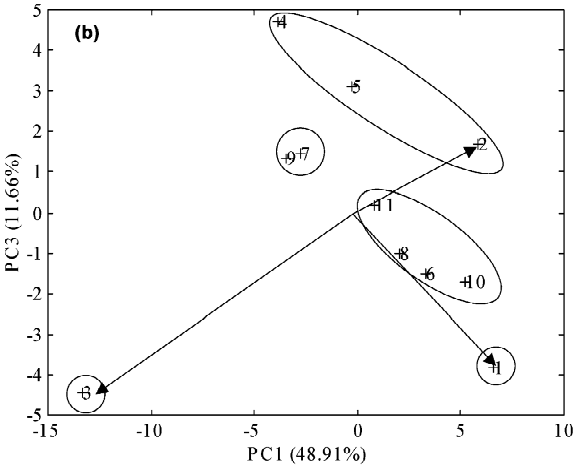


Fig. 4. Score plots of the autoscaled *k* data for 11 chromatographic systems and 68 substances (11×68 matrix) on (a) PC1–PC2 and (b) PC1–PC3.

which is situated between CS1 and CS3 is far more related to CS9. The PC1–PC3 plot (Fig. 4b) seems to represent the relation between the different columns best and it can be considered as an expansion of the correlations indicated in Fig. 3. However, the interpretation of the PCA plots to determine the relationships between the systems was not always found to be straightforward. Principal component analysis, performed on the retention data, might be useful, when applied in combination with the table of correlation coefficients. However, further evaluation

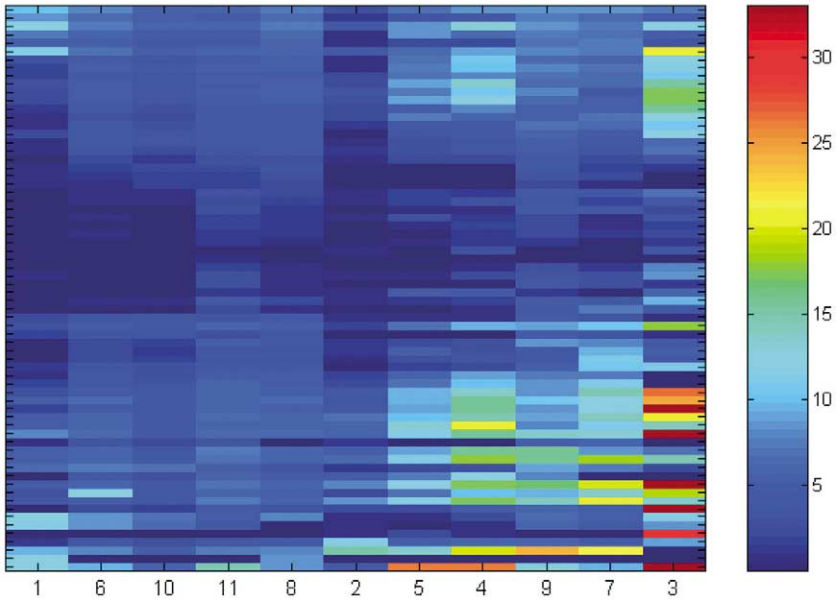


Fig. 5. OPTICS color map of the PCA data from autoscaled *k* results, for 11 chromatographic systems (abscissa) and 68 substances (ordinate). Blue colors indicate low retention and yellow to red colors high retention.

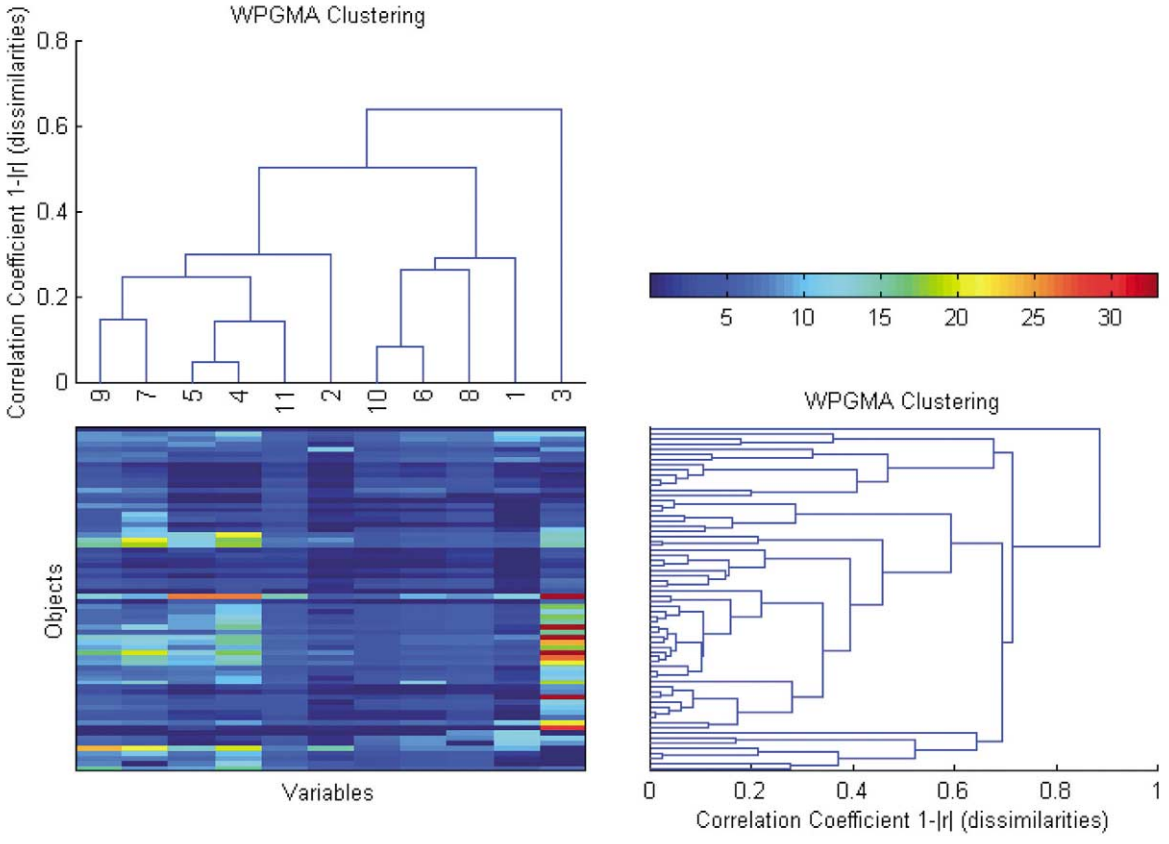


Fig. 6. Color map based on dendrograms resulting from the hierarchical average linkage technique on retention factors *k*, for 11 chromatographic systems (abscissa) and 68 substances (ordinate).

of larger data sets with more systems seems to be required to confirm this statement.

A second visualisation technique used were the so-called color maps [20,21]. In Fig. 5, an OPTICS color map representing the 11 chromatographic systems and the 68 substances, based on k data, is shown. In order to obtain the color map, PCA data from autoscaled retention factors were applied to OPTICS (Ordering Points to Investigate The Clustering Structure), a density-based method which is able to establish a unique order of the data [20]. OPTICS starts with a randomly selected object, and consecutively the most similar objects are linked according to reachability distances [20], leading to an ordering of the objects. In our situation objects can be either systems or substances. The closer the objects are to each other, the more similar they are. The color bar next to the color map links the retention factor k to a color. Similar colors indicate similar retention factor values on the different systems. For instance, in Fig. 5, CS3 has several different retention factors compared to for instance CS10 or CS8. It also can be observed that the dominating color in Fig. 5 is blue, i.e. representing low k values.

A classification of the systems can be obtained by interpreting the similarity or difference of colors between columns in the color map. Adjoining columns showing similar colors are grouped into one class, whereas differences in colors between two columns will lead to a subdivision of the systems into classes. The order of the selected systems based on their reachability distances is represented in the abscissa of Fig. 5. Once this order is obtained, it is possible to classify the systems by comparing the color patterns between the different color map columns. In Fig. 5, CS3 might be distinguished in a first group; CS7, CS9, CS4 and CS5 in a second one; CS1, CS6, CS10, CS11 and CS8 in a third, and CS2 in a fourth group. CS2 is situated between the second and the third group. The classification obtained, with the exception of CS2, is the same as when the correlation coefficients (Table 3) are interpreted. Differences between systems will be defined based on the colors which are not dominating. This means that it will be mainly based on differences in substances eluting late on one of the systems (Fig. 5).

It can be concluded that the use of an OPTICS color map is a technique that is already better than the PCA plots in order to obtain a classification of the chromatographic systems. The better applicability of the OPTICS color map can be explained by the fact that OPTICS uses as much dimensions as needed to explain 99.0–99.5% of the variance in the data, whereas looking at PCA score plots, one cannot visualise more than two or three dimensions simultaneously, which often represents less variability (see Fig. 4).

However, to define classes of orthogonal and of similar systems, the OPTICS color map was not always found to be straightforward. Therefore, a color map based on dendrograms constructed using weighted average linkage, a hierarchical two-way clustering technique [19,21–24], also was evaluated (Fig. 6). Average linkage is an agglomerative method [22] which starts with clusters each containing one object. The distance between two clusters is defined as the average distance between all pairs of objects in the two different clusters. The method successively merges the two clusters for which the dissimilarity value is the smallest into one cluster, until only one cluster is left. In the weighted mode, objects of smaller clusters carry a larger weight than objects from larger clusters. The (dis)similarity between clusters is visualised in the dendrogram by the height at which they are connected. The higher two clusters are connected, the more dissimilar they are [23].

Again, the data can be interpreted in two ways: systems on the abscissa and substances on the ordinate. From the dendrogram, it can easily be seen which systems are branched together and are interchangeable. The color map and the dendrograms from which it originates, are shown in Fig. 6. Again, the color bar relates retention to a color, scaling from blue (low retention) to red (high retention). Once the dendrograms and color map are created, the most orthogonal systems can be distinguished. Also systems delivering the most analogous information can be seen: as the ordinate refers to dissimilarities, represented by $1 - |r|$, systems which are branched at low dissimilarities are highly correlated and will have similar selectivity. Looking at Fig. 6, CS3 would be considered most orthogonal to CS1 and CS2. This figure also reveals that CS1 is considered interchangeable with CS8, CS6 and CS10. Moreover this implies that when applying arbitrarily a dissimilarity limit of 0.35, CS2 can be replaced by CS9,

CS7, CS5, CS4 or CS11. The dissimilarity limit was thus chosen because within each group the correlation coefficients are still reasonable high, i.e. generally above 0.65. Between groups, the correlation coefficients are generally low, as the next merges in the dendrogram are situated much higher, at dissimilarities of approximately 0.50 and 0.65. The conclusions drawn from the dendrogram are similar but more complete than the interpretation of Table 3, which led to Fig. 3. They also seem to be easier to interpret and more valuable than the different color maps, even the dendrogram based ones. In general, the construction and the evaluation of the dendrogram seems to be easier, especially when the number of systems becomes large.

The above discussed approaches also were applied using $\log k$ instead of k as retention parameter. Finally, we tend to prefer k above $\log k$ for a number of reasons. In the color maps $\log k$ tends to make a differentiation mainly based on substances with a low retention on (one of) the composed systems while higher retentions are reduced in importance. Therefore k represents better the real retention (differences) on the considered systems. Moreover on HPLC systems we do not expect large differences in order of magnitude between the retention of the test substances, that would recommend the use of a logarithmic transformation.

However, in practice when determining orthogonal systems, special attention is to be paid to the definition of the experimental conditions. The initial mobile phase should cause some retention for the most polar compounds and the gradient should be such that an optimal differentiation in the k of the different substances can be obtained within a reasonable analysis time. In these circumstances, we presume that the k values are suitable to evaluate orthogonality. Further evaluation is anyway required on other sets of chromatographic systems to confirm our preferences.

3.3. Validation of the most orthogonal systems

Example chromatograms of four substances on the most orthogonal systems are given in Fig. 7. It can be seen that the largest selectivity differences appear between the zirconia-based system CS3 (Fig. 7c) and the silica-based ones (CS1, CS2, CS8, CS9 and CS10). The chromatograms of CS1, CS8 and CS10 (group A in Fig. 3) are very similar. The same is valid for CS2 and CS9 (group B in Fig. 3). When comparing the elutions on the systems of groups A and B, it can be seen that the resolution of the peaks is quite different.

A first validation of the systems defined above as orthogonal was performed by means of a cross-validation procedure. Therefore the above set of substances was divided in two, a training set to ''redefine'' the orthogonal systems and a validation set to evaluate the orthogonality of the selected systems. To do so, the retention data of the 68 substances measured on CS3 were sorted, and every fourth substance was selected, leading to four representative validation sets of 17 substances, and four new training sets with the remaining 51 substances. The training sets permitted to determine whether still a similar selection of orthogonal systems was found. On the other hand, the validation sets, which can be considered as an artificial separation mixture, should evaluate whether the selected orthogonal systems indeed show general selectivity differences. This evaluation was done by plotting the data of the validation sets for the orthogonal systems (similar as in Fig. 2), and by calculating their correlation coefficients.

It was found that, based on the correlation coefficients and dendrograms, all four training sets gave a similar clustering concerning orthogonal and interchangeable systems. This can be observed in Table 4, where the table of correlation coefficients between some systems is given for the four sets.

Plotting the k data for all pairs of orthogonal systems and evaluating the correlation coefficients (Table 5) also led to large and general selectivity differences for all validation sets.

The above demonstrated that the four training sets (51 substances) give similar information as the original test set of 68 substances, and the results from the four validation sets (17 substances) indicate selectivity differences from the systems defined orthogonal with a training set not containing the substances of the validation set.

As a second evaluation, six test substances, differ-

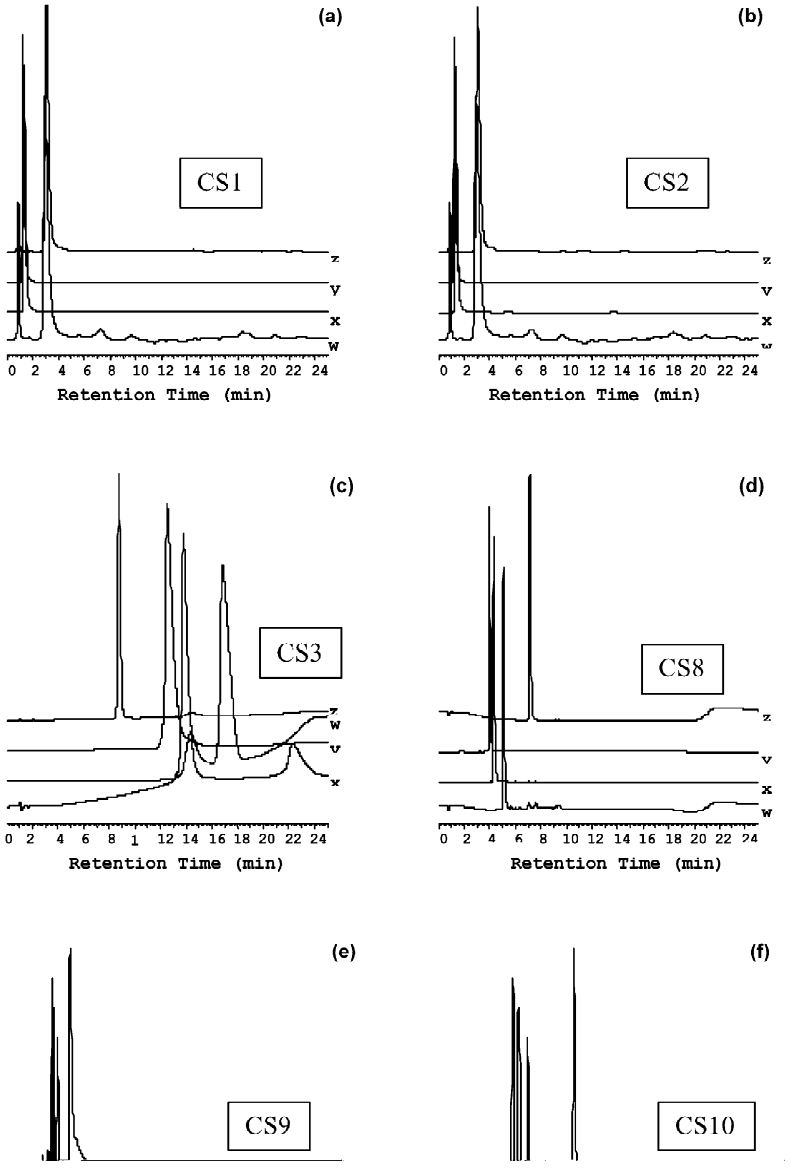


Fig. 7. Multidisplay chromatograms of bupranolol hydrochloride (w), prazosin hydrochloride (*x*), terazosin (*y*) and 4-benzylphenol (*z*) on (a) CS1, (b) CS2, (c) CS3, (d) CS8, (e) CS9, (f) CS10.

[22] D.L. Massart, L. Kaufman, The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis, Wiley, New York, 1983.
[23] W. Vogt, D. Nagel, H. Sator, Cluster Analysis in Clinical Chemistry: A Model, Wiley, Chichester, 1987.
[24] L. Kaufman, P.J. Rousseeuw, Finding Groups in Data—An Introduction to Cluster Analysis, Wiley, New York, 1990.
[25] N.H. Khan, P. Wera, E. Roets, J. Hoogmartens, J. Liq. Chromatogr. 13 (1990) 1351.