Journal of Chromatography A, 1326 (2014) 110–124

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Exploratory data analysis as a tool for similarity assessment and clustering of chiral polysaccharide-based systems used to separate pharmaceuticals in supercritical fluid chromatography☆

Katrijn De Klerck, Yvan Vander Heyden, Debby Mangelings*

Department of Analytical Chemistry and Pharmaceutical Technology, Center for Pharmaceutical Research (CePhaR), Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, B-1090 Brussels, Belgium

ARTICLE INFO

Article history:
Received 30 August 2013
Received in revised form
13 December 2013
Accepted 17 December 2013
Available online 25 December 2013

Keywords:
Chiral supercritical fluid chromatography
Racemic drugs
Polysaccharide-based chiral stationary phases
Chemometrics
Similarity and dissimilarity assessment

ABSTRACT

In the search for appropriate chromatographic conditions to separate enantiomers, screening strategies are often applied because achieving chiral separations is tedious. These screenings aim to find relatively fast suitable separation conditions. However, the definition of these screenings mostly relies on years of expertise or on the labour- and time-intensive investigation of a broad range of chiral stationary- and mobile phases. A large amount of data is generated using either approach. In this study, the obtained data are investigated in a systematic manner and (dis)similar systems are searched for. For this case study, 48 chromatographic systems were characterized by the enantioresolutions of 29 racemates. Exploratory data analysis was performed by means of projection pursuit, revealing the different enantioselective patterns of the chromatographic systems. To quantify the (dis)similarity, correlation coefficients and Euclidean distances were calculated. These results were visualized in colour maps to allow investigating the degree of (dis)similarity between the systems. These maps proved to be a helpful tool in the selection of dissimilar/orthogonal chromatographic conditions. Hierarchical-cluster-analysis dendrograms were constructed next to evaluate the clustering of similar systems, i.e. with an equivalent enantioselectivity. Screening sequences were extracted and compared with the initial, defined by direct data interpretation. In a final section, selection of dissimilar systems was done by means of the Kennard and Stone algorithm. The systems selected by the applied techniques did not necessarily perform better than the selection by direct data interpretation. Nevertheless, high cumulative success rates are achieved for the selected combinations, due to the broad enantioselectivity, the high individual success rates and the complementarity of the chiral selectors.

1. Introduction

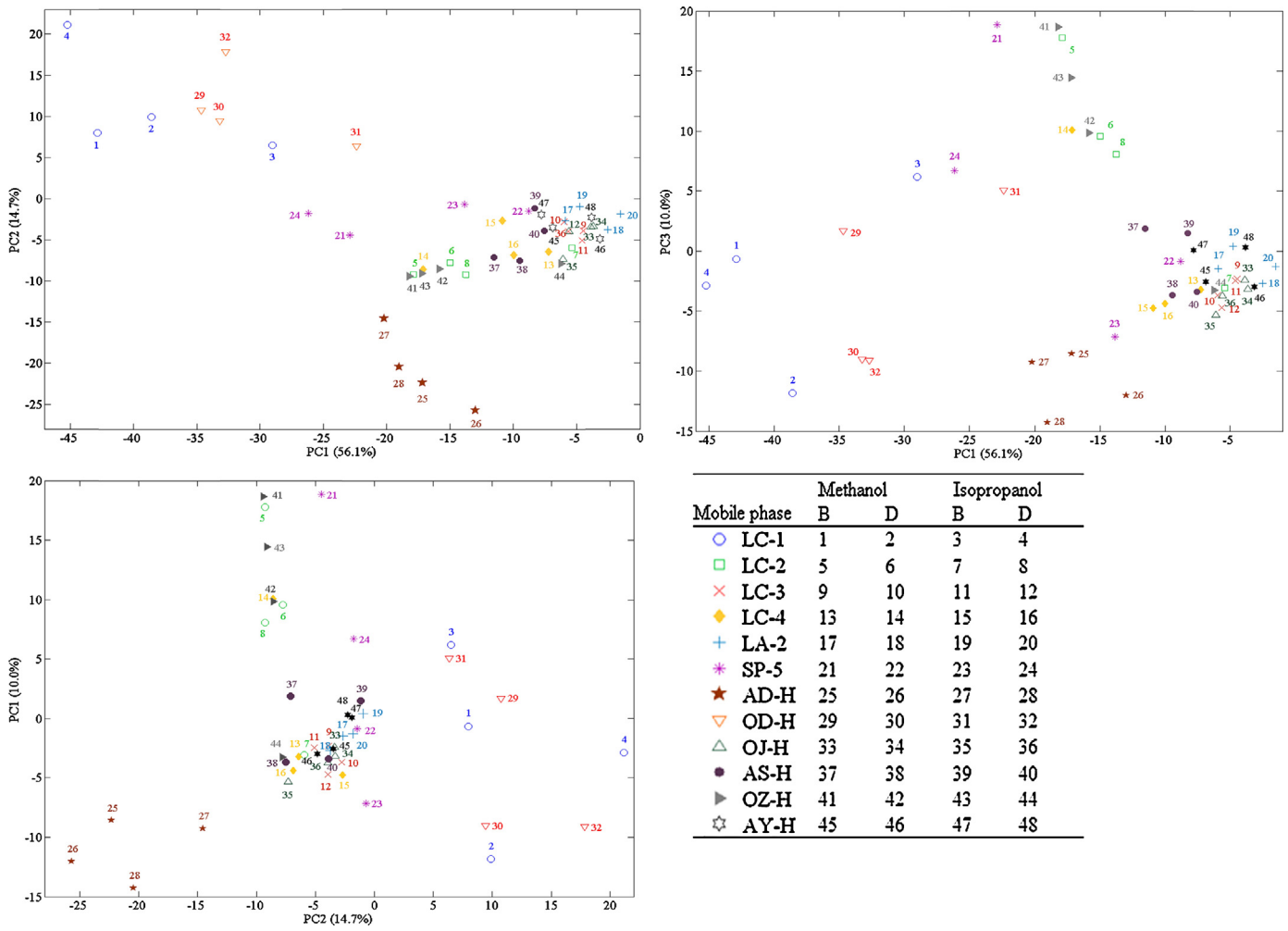
A green technique getting special attention is supercritical fluid chromatography (SFC). It has repeatedly been established as a fast and high-performance chiral-separation technique for a wide range of compounds [1–5]. However, optimal chromatographic chiral separation conditions are very tedious to achieve since enantioselectivity still remains unpredictable. The selection of a proper chromatographic system, i.e. a chiral stationary phase (CSP) and a mobile phase (MP) combination, is a problem already addressed by several groups [4,6–16]. In this context, chiral screening strategies, which propose to screen certain chromatographic systems in a given sequence, are defined. Screening strategies enable users to explore the enantioselectivity of a readily available set of complementary CSPs, thereby trying to find the most appropriate separation system systematically and efficiently. Chiral screening strategies for SFC have been defined, mainly using polysaccharide-based CSPs [1–4,17,18].

In screening strategies, the enantioselective complementarity/dissimilarity of the chromatographic systems is an important property, leading to a high rate of successful separations. In this context, it is vital to evaluate the (enantio)similarity/dissimilarity of chromatographic systems [19]. Typically, a test set of chiral racemates is analyzed. Based on the separation of these compounds, the chiral systems can be characterized. Their complementarity is assessed and a screening sequence is composed by selecting the most successful and complementary systems [11–14,20,21].

☆ Presented at the 39th International Symposium on High-Performance Liquid-Phase Separations and Related Techniques, Amsterdam, Netherlands, 16–20 June 2013.

* Corresponding author at: Vrije Universiteit Brussel (VUB), Center for Pharmaceutical Research (CePhaR), Department of Analytical Chemistry and Pharmaceutical Technology (FABI), Laarbeeklaan 103, B-1090 Brussels, Belgium.
Tel.: +32 2 477 43 29; fax: +32 2 477 47 35.
E-mail address: debby.mangelings@vub.ac.be (D. Mangelings).

0021-9673/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.chroma.2013.12.052



Mobile phase	Methanol B	Methanol D	Isopropanol B	Isopropanol D
LC-1	1	2	3	4
LC-2	5	6	7	8
LC-3	9	10	11	12
LC-4	13	14	15	16
LA-2	17	18	19	20
SP-5	21	22	23	24
AD-H	25	26	27	28
OD-H	29	30	31	32
OJ-H	33	34	35	36
AS-H	37	38	39	40
OZ-H	41	42	43	44
AY-H	45	46	47	48

Fig. 1. Score plots of the principal component analysis on 48 chromatographic systems. The legend summarizes the analyzed systems. Reproduced with permission from [15].

Chemometric techniques have the potential to improve this process and to make it more straightforward. To explore the data set systematically, exploratory data analysis methods can be used. These techniques, e.g. principal component analysis (PCA) or projection pursuit (PP), allow distinguishing chromatographic systems with similar or dissimilar properties, in this case enantioresolution patterns. This allows selecting a dissimilar system, when a first did not result in the desired separation. On the other hand, it also enables determining which systems yield similar enantioselectivities and can be used as alternatives. Finally, this chemometric approach may also simplify evaluating screening data to determine the complementarity of chromatographic systems in the context of defining screening strategies.

Stationary phases can also be characterized using a modelling approach. The research group of West [22–27] used a modified version of the linear solvation energy relationship (LSER) to assess the contribution of different interaction mechanisms to the retention on given stationary phases. Both achiral and chiral stationary phases were characterized. This allowed evaluating the (dis)similarity of the investigated systems [22,24–28].

To allow observing groups of systems with a similar enantioresolution pattern, PCA can be applied. In PCA the original variables that describe the systems are replaced by new latent variables based on the variability within the data set [27,29–31]. Euerby and Petersson [32] used this technique to characterize 135 stationary phases in terms of surface coverage, hydrophobic selectivity, shape selectivity, hydrogen bonding capacity and ion-exchange capacity. PCA proved to be a very helpful tool to allow rapid determination of the difference between phases. Later, the same research group also analyzed other stationary phases in a similar way [33,34]. Lammerhofer et al. [35] used PCA to confirm the similarity between in-house developed columns with mixed-mode ion-exchange properties and various commercially available phases.

Earlier, we applied this technique on the data set in [21] (Fig. 1) and confirmed the complementarity of the systems selected before for the proposed screening step. PCA showed that the CSP generally has a higher impact on the enantioselectivity than the MP, as expected. However, PCA did not show a clear distinction of separate groups of similar chiral systems. Most systems are grouped in a central cloud. Hence, rather than showing a totally different enantioselective pattern, many systems express a similar enantioselectivity towards the largest part of the test set. This is not surprising since polysaccharide-based CSPs are known to have a broad enantioselective recognition ability. In other words, all CSPs will separate many compounds of the test set. Therefore a considerable overlap of separated compounds between the different chromatographic systems exists.

In a screening strategy it is aimed to cover the broadest enantioselective range with the least chromatographic systems by selecting dissimilar and successful systems. In a PCA plot, the dissimilar

systems can be found on the outer edges of the data cloud. However, the plot provides no knowledge about the enantioselectivity of the chromatographic systems, so it is not evident to select successful complementary systems solely based on the PCA plots.

The aim of this paper is to further assess the potential of different chemometric techniques to visualize and quantify the degree of similarity and/or dissimilarity between chiral systems and to discover clusters of chromatographic systems with the same enantioselective behaviour. Additionally, the applicability of these techniques for the definition of screening strategies will be evaluated.

2. Material and methods

2.1. Chiral test set and chemicals

The chiral test set is composed of 29 compounds (Table 1). Solutions of all racemates were prepared in methanol at a concentration of 0.5 mg/ml. Solutions were stored at 4 °C.

CO_2 2.7 (purity ≥ 99.7%) was obtained from Linde Gas (Grimbergen, Belgium), methanol (MeOH) and isopropanol (2PrOH) (HPLC grade) from Fisher Chemical (Loughborough, Leicestershire, UK). Isopropylamine (IPA) (purity 99.5%) and trifluoroacetic acid (TFA) (purity ≥ 98%) were from Aldrich (Steinheim, Germany).

2.2. SFC instrumentation and experimental conditions

An analytical system from Waters® (Milford, Massachusetts, USA) was used, consisting of a Thar® SFC fluid delivery module (a liquid CO_2 pump and a modifier pump with a six solvent switching valve), a cooling bath of Thermo Scientific® type Neslab RTE7 controlled by a Digital One thermoregulator to cool pumpheads and CO_2-delivery tubings, a Thar® autosampler with a 48-vial plate, a Thar® SFC analytical-2-prep oven with a 10-column selection valve, a Thar® SFC automated backpressure regulator SuperPure Discovery Series and a Waters® 2998 photodiode array detector. The autosampler was equipped with a 5 μl injection loop. The instrument was controlled by Superchrom® (Thar, 2003–2009, Pittsburgh, Pennsylvania, USA) or Chromscope® Instrument Edition V1.10 software (Water Corporation, 2011, Milford, Conneticut, USA) and data were processed using the Chromscope® (TharSFC®, 2009) or Chromscope® Instrument Edition V1.10 software.

All experiments were performed with a total flow rate of 3.0 ml/min, temperature of 30 °C, backpressure of 150 bar, injection volume of 5 μl and a fixed run time of 30 min. For each enantioseparation, Rs was calculated using peak widths at half height, and α as the ratio of retention factors of the last and first eluting enantiomer of a pair.

2.3. Chromatographic systems

Twelve CSPs and four MPs, in total 48 systems, were screened. Lux® Cellulose-1 (LC-1), Lux® Cellulose-2 (LC-2), Lux® Cellulose-3 (LC-3), Lux® Cellulose-4 (LC-4), Lux® Amylose-2 (LA-2) and Sepapak®-5 (SP-5), were purchased from Phenomenex (Utrecht, The Netherlands). Chiralcel® OD-H (OD-H), Chiralcel® OJ-H (OJ-H), Chiralcel® OZ-H (OZ-H), Chiralpak® AD-H (AD-H), Chiralpak® AS-H (AS-H) and Chiralpak® AY-H (AY-H) were from Chiral Technologies Europe (Illkirch-Cedex, France). All columns had 250 × 4.6 mm i.d. dimensions with 5 μm particle size. The chiral systems are summarized in Table 2. All mobile phase compositions in this table are expressed in volume ratios (v/v).

2.4. Data analysis

All calculations were performed using m-files written in Matlab® 7.1 (The Mathworks, Natick, MA, USA).

3. Results and discussion

The enantioselectivity of 96 chiral systems was evaluated earlier [20,21]. In these studies, complementary (dissimilar) systems were selected by analysing the raw data (resolution (Rs) and selectivity (α)). High cumulative success rates were achieved by selecting the system generating the highest number of separations, combined with the results from systems generating the most complementary separations (i.e. not obtained with the previous system(s)). The resulting system sequence allowed achieving the highest success rate by screening a limited number of systems. However, this is a labour-intensive approach. The potential of chemometric techniques will therefore be assessed in an attempt to enable a faster selection of dissimilar/complementary systems. The applicability of these techniques to create generic chiral screening strategies will also be assessed.

3.1. Methodology

Earlier, 57 pharmaceutical racemates were selected to characterize the chromatographic systems. These compounds were aimed to cover a broad range of structurally, chemically and pharmacologically different pharmaceuticals. The test set consisted about 80% out of basic compounds. This was done intentionally to represent the market situation of pharmaceuticals nowadays. The same compounds as in Ref. [36] were used, with the exception of leucovorin and naproxen. Not all 57 racemates could be used to characterize the chromatographic systems with chemometric techniques. At certain separation conditions, given compounds did not elute or the results were not interpretable. Consequently, using all 57 compounds for data analysis would give an incomplete data matrix. For this reason, the data set had to be reduced to a subset of 29 racemates that gave separation results on all systems (Table 1).

Twelve polysaccharide-based CSPs in combination with four mobile phases (MPs), generating 48 chromatographic systems, were evaluated (Table 2). Concerning the use of additives, two major approaches were applied. In a first approach trifluoroacetic acid was used for acidic compounds and isopropylamine for all other compounds, while in the second both additives were used for all compounds. Very high success rates were obtained with this latter approach, which eliminates the need for different additives depending on the chemical structure of compounds. However, when combining a basic and an acidic additive one needs to be aware of the potential risk of salt formation between both additives. When keeping the additive concentrations at 0.1% in the modifier, these salts do not show solubility problems and the benefits of combined additives can be exploited without instrumental problems [36]. This concentration of 0.1(v)% corresponds to 0.012 M for IPA and 0.013 M for TFA. From the 48 systems evaluated, half use a combination of both additives in the mobile phase, and half separate additives.

Essentially, the systems were characterized by their enantioselective behaviour towards the 29 compounds, i.e. by their Rs and/or α. The resolution offers the advantage that both the separation between two peaks and their peak widths are taken into account when assessing a separation. As a result a limit value (Rs = 1.5) between partially and baseline-separated compounds can be applied. To illustrate the difference between both parameters, the chromatograms of the enantioseparations of chlorphenamine and acebutolol are presented in Fig. 2. These separations yield the

Table 1
The chiral compounds used.

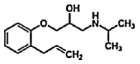
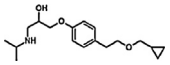
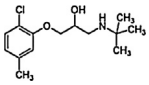
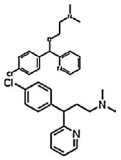
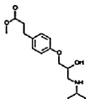
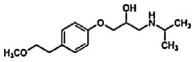
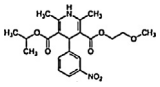
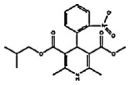
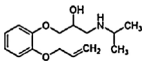
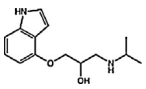
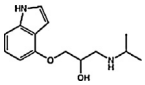
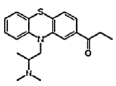
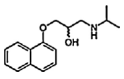
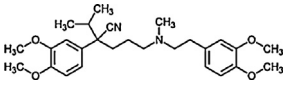
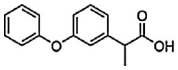
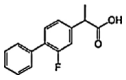
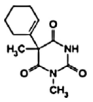
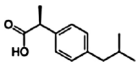
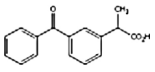
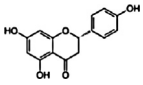
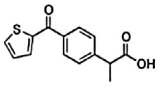
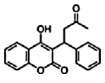
#	Compound	Structure	Supplier
1	Alprenolol		Sigma Aldrich, Steinheim, Germany
2	Ambucetamide		Janssen Pharmaceutica, Beerse, Belgium
3	Betaxolol		Sigma Aldrich, Steinheim, Germany
4	Bisoprolol		Manufacturer Unknown
5	Bupranolol		Schwarz Pharma, Monheim, Germany
6	Carbinoxamine		Manufacturer Unkown
7	Chlorphenamine		Sigma Aldrich, Steinheim, Germany
8	Esmolol		Du Pont De Nemours, Saconnex, Switzerland
9	Methadone		Federa, Brussels, Belgium
10	Metoprolol		Sigma Aldrich, Steinheim, Germany
11	Mianserine		Diosynth and Organon, Brussels, Belgium
12	Nimodipine		Bayer, Leverkussen, Germany
13	Nisoldipine		Bayer, Leverkussen, Germany
14	Nitrendipine		Bayer, Leverkussen, Germany
15	Oxazepam		Sigma Aldrich, Steinheim, Germany
16	Oxprenolol		Cynamid Benelux, Brussels, Belgium

Table 1 (*Continued*)

#	Compound	Structure	Supplier
17	Pindolol		Sigma Aldrich, Steinheim, Germany
18	Promethazine		Sigma Aldrich, Steinheim, Germany
19	Propiomazine		Manufacturer Unknown
20	Propranolol		Fluka, Neu-Ulm, Switzerland
21	Verapamil		Fluka, Neu-Ulm, Switzerland
22	Fenoprofen[a]		Sigma Aldrich, Steinheim, Germany
23	Flurbiprofen[a]		ICN Biomedicals, OH, USA
24	Hexobarbital[a]		Manufacturer Unknown
25	Ibuprofen[a]		Sigma Aldrich, Steinheim, Germany
26	Ketoprofen[a]		Sigma Aldrich, Steinheim, Germany
27	Naringenin[a]		Sigma Aldrich, Steinheim, Germany
28	Suprofen[a]		Sigma Aldrich, Steinheim, Germany
29	Warfarin[a]		Sigma Aldrich, Steinheim, Germany

[a] Acidic compounds.

same separation factor α of 1.12, although it is clear that acebutolol is only partially separated, while a baseline separation is obtained for chlorphenamine. Rs reflects these differences better; chlorphenamine yields a Rs > 1.5 (2.26), while acebutolol yields only one of 1.08. For this reason, our preference goes to Rs as parameter to assess the quality of an enantioseparation. Of course, the selectivity might be a better choice to assess other properties of separations.

When using the Rs (or α) one should be aware that a reversal of elution order of the enantiomers can occur. In this case, two systems with opposite elution orders could also be considered complementary. This can be important in certain situations, e.g. when an enantiomer is to be determined as impurity in the presence of the other. However, we are not equipped with an enantioselective detector to allow identification of each enantiomer. The pure enantiomers are also not available for all test compounds. So in fact, from a practical point of view we cannot comment on this reversal of elution orders. Moreover, the focus in this research is put on techniques to evaluate chromatographic data, more than on the generation of the data.

3.2. Identifying (dis)similar systems

3.2.1. Projection pursuit

Generally, visualizing groups or clusters in a data set can be done by exploratory techniques, such as projection pursuit (PP) methods. These aim reducing the number of informative variables while revealing clustering tendencies and/or outlying data objects. In fact, PCA is a PP method in which the data variance is used

Table 2
The chiral chromatographic systems (chiral stationary + mobile phase).

Total flow rate: 3 ml/min
Mobile phase composition: CO_2/modifier, 80/20 (v/v)
Temperature: 30 °C
Back pressure: 150 bar
Organic modifiers
A: MeOH + 0.5%TFA* or IPA**
B: MeOH + 0.1%TFA + 0.1%IPA
C: 2PrOH + 0.5%TFA* or IPA**
D: 2PrOH + 0.1%TFA + 0.1%IPA
* used for acidic compounds
** used for all other compounds

System	Chiral selector	Trade name	Modifier
1	Cellulose tris(3,5-dimethylphenylcarbamate)	Lux® cellulose-1	A
2			B
3			C
4			D
5	Cellulose tris(3-chloro-4-methylphenylcarbamate)	Lux® cellulose-2	A
6			B
7			C
8			D
9	Cellulose tris(4-methylbenzoate)	Lux® cellulose-3	A
10			B
11			C
12			D
13	Cellulose tris(4-chloro-3-methylphenylcarbamate)	Lux® cellulose-4	A
14			B
15			C
16			D
17	Amylose tris(5-chloro-2-methylphenylcarbamate)	Lux® Amylose-2	A
18			B
19			C
20			D
21	Cellulose tris(3,5-dichlorophenylcarbamate)	Sepapak®-5	A
22			B
23			C
24			D
25	Amylose tris(3,5-dimethylphenylcarbamate)	Chiralpak® AD-H	A
26			B
27			C
28			D
29	Cellulose tris(3,5-dimethylphenylcarbamate)	Chiralcel® OD-H	A
30			B
31			C
32			D
33	Cellulose tris(4-methylbenzoate)	Chiralcel® OJ-H	A
34			B
35			C
36			D
37	Amylose tris((S)-α-methylbenzylcarbamate)	Chiralpak® AS-H	A
38			B
39			C
40			D
41	Cellulose tris(3-chloro-4-methylphenylcarbamate)	Chiralcel® OZ-H	A
42			B
43			C
44			D
45	Amylose tris(5-chloro-2-methylphenylcarbamate)	Chiralpak® AY-H	A
46			B
47			C
48			D

as a projection index. As discussed higher, PCA was performed on the data set in an earlier study [21]. Another possible projection index is kurtosis, a parameter that quantifies the deviation from the normal data distribution, and mainly reveals outliers in a data set [37,38]. When characterizing enantioselective behaviour of chromatographic systems, outlying systems could potentially be interesting since they would generate different enantioselectivities. The resolution results of the 29-racemate test set on the 48 chromatographic systems are presented in Table 3. The projection pursuit method with kurtosis as a projection index was applied on this data matrix.

To overcome the influence of different Rs ranges, all results were transformed by autoscaling:

$$Rs_A = \frac{(Rs - \overline{Rs})}{s(Rs)} \quad (1)$$

Table 3
Screening results on the 48 chromatographic systems expressed as resolution (Rs) of the 29 test compounds.

Compound number (Table 1) system	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1	4.47	17.6	15.1	11.1	1.35	0.00	0.00	12.5	0.65	14.1	3.96	0.25	0.00	0.00	12.7	10.1	22.2	0.65	0.00	9.18	1.27	0.00	0.00	1.53	0.00	0.00	2.24	1.41	10.8
2	4.31	5.20	15.8	11.0	1.37	0.00	0.52	12.8	0.57	14.3	2.73	0.00	0.00	0.00	11.3	9.24	23.6	0.47	0.00	10.7	3.12	0.00	0.00	1.68	0.00	0.00	1.97	0.66	14.6
3	5.99	19.1	9.84	7.88	1.73	1.31	0.00	10.4	1.32	9.48	2.98	0.00	0.52	0.00	8.00	11.7	0.00	0.00	0.00	5.80	0.00	0.68	0.65	0.00	0.00	0.00	3.20	1.41	3.99
4	11.3	17.3	20.9	15.9	2.68	0.00	0.00	18.8	1.68	24.2	1.67	0.00	0.51	0.00	3.68	19.1	0.00	0.00	0.00	14.0	4.14	0.51	0.59	0.00	0.00	0.00	3.05	0.55	11.8
5	2.07	24.5	0.00	1.26	1.52	4.22	1.49	0.00	0.00	0.55	0.00	0.62	2.28	0.00	2.45	0.00	3.91	0.00	1.33	2.65	1.57	0.00	0.00	7.41	0.00	0.00	3.70	1.33	3.49
6	1.35	15.5	1.29	0.67	1.42	2.24	0.64	1.29	1.56	1.35	2.02	0.67	1.84	0.00	2.50	0.53	6.36	1.36	0.66	0.00	1.52	0.00	0.00	7.22	0.00	0.00	3.48	1.36	3.65
7	0.36	0.00	0.61	1.03	0.68	3.67	2.02	0.60	2.08	0.46	1.46	0.00	1.40	0.00	1.76	0.00	0.00	0.48	1.81	1.56	1.38	0.00	0.00	5.97	0.00	0.00	3.91	1.40	4.78
8	0.00	12.9	0.40	1.30	0.00	5.07	1.37	0.00	1.77	0.59	3.18	0.00	1.61	0.00	1.78	0.00	8.09	1.37	11.8	0.00	0.98	0.00	0.00	5.75	0.53	0.00	3.87	0.00	4.62
9	0.00	1.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.36	0.00	0.00	0.00	2.16	0.00	1.72	0.00	0.65	1.25	0.00	0.00	0.57	0.00	1.28	0.63	0.00	2.29	4.48
10	0.00	0.55	1.19	0.62	0.00	0.00	0.00	1.18	0.00	1.21	5.67	0.49	0.56	0.00	2.86	0.37	2.24	0.00	0.59	1.21	0.59	0.00	0.63	0.00	1.29	0.56	0.00	2.07	5.39
11	0.00	1.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.55	0.88	0.84	0.00	2.27	0.00	0.00	2.21	3.69	0.00	1.23	1.92	0.00	0.00	0.00	2.44	1.30	1.76	6.08
12	0.00	0.00	0.65	0.49	0.00	0.00	0.00	0.60	0.00	0.64	4.61	0.81	0.87	0.00	1.49	0.00	1.49	0.00	0.92	0.60	0.00	1.74	0.00	0.00	0.00	1.99	1.25	1.36	10.1
13	1.95	0.00	0.00	1.32	1.31	5.32	2.87	0.00	0.00	0.00	1.63	0.00	1.52	0.00	4.02	1.33	3.60	1.67	0.00	2.10	3.82	0.35	0.00	6.80	0.00	0.00	2.79	0.53	3.35
14	1.29	17.2	1.81	1.33	2.04	3.08	1.26	1.35	1.28	1.78	5.50	0.00	1.68	0.00	3.90	0.26	6.64	0.00	0.00	0.00	3.40	0.32	0.00	7.06	0.00	0.00	2.74	0.44	3.59
15	0.55	0.00	4.41	3.74	1.64	4.03	3.96	3.88	2.03	4.05	1.36	0.00	1.35	0.52	1.50	2.85	1.51	1.89	1.25	2.52	2.90	1.47	0.00	7.17	0.00	0.00	7.08	1.29	3.33
16	0.00	0.00	2.42	2.36	1.45	6.86	2.28	1.82	1.43	2.96	7.09	0.00	1.34	0.53	1.58	0.00	5.81	0.47	1.40	0.00	2.89	0.66	0.00	6.95	0.00	0.00	7.43	0.61	3.75
17	0.63	2.02	0.46	0.51	1.33	1.39	0.00	1.28	0.63	1.33	0.00	0.00	0.00	0.00	1.55	1.78	1.49	0.00	1.69	1.53	0.00	1.30	3.74	2.30	0.00	1.73	1.28	0.00	4.19
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.96	0.00	0.00	0.00	0.60	0.00	0.00	1.40	3.57	2.24	0.00	1.58	1.31	0.00	5.15
19	0.59	2.68	0.66	0.65	0.00	0.69	0.61	0.66	0.88	1.37	0.56	1.45	1.45	0.00	0.00	1.39	1.19	0.65	0.58	1.39	0.54	1.12	0.45	0.63	0.00	0.00	0.45	1.65	2.99
20	0.00	0.00	0.00	0.00	0.00	0.55	0.00	0.00	0.00	0.00	0.00	1.45	1.37	0.00	0.00	0.00	0.00	1.21	0.00	0.00	0.00	0.66	0.00	0.59	0.00	0.00	0.00	1.45	4.02
21	1.29	28.1	1.61	1.25	2.07	0.68	0.00	2.56	1.38	1.91	4.61	3.08	0.00	0.00	7.32	5.81	5.99	0.00	1.47	0.64	1.48	0.00	0.00	0.00	0.00	1.31	0.00	0.00	3.29
22	0.00	3.69	1.60	1.80	0.00	0.53	0.00	1.96	0.55	1.75	1.54	4.11	0.00	0.00	6.27	1.36	2.72	0.00	0.00	1.72	1.52	0.00	0.00	0.00	0.00	1.25	0.00	2.75	3.39
23	1.66	0.00	3.05	3.54	2.25	0.00	0.00	2.53	0.00	3.48	1.73	5.12	1.72	0.00	8.25	7.79	6.15	1.47	1.46	0.62	0.46	0.00	0.00	0.59	0.00	1.65	1.49	3.78	8.35
24	0.00	18.5	4.90	10.1	0.00	1.21	2.26	5.44	2.67	6.16	3.65	8.47	2.30	0.00	8.52	4.22	4.78	2.33	1.27	3.26	2.58	0.00	0.00	0.42	0.00	1.50	1.18	3.45	9.49
25	1.56	3.47	1.76	2.18	0.00	0.00	0.55	2.87	0.00	1.44	13.1	0.00	0.00	0.00	3.38	3.99	2.62	2.19	4.04	5.95	0.00	2.05	12.8	28.5	1.95	0.00	0.00	10.1	2.34
26	0.00	0.00	0.00	0.00	0.00	1.27	0.47	0.00	0.00	0.00	5.87	0.00	0.00	0.00	3.41	0.45	1.27	0.67	0.00	1.36	0.00	2.19	12.6	28.4	1.90	0.68	0.00	9.54	14.4
27	3.03	3.39	3.00	3.02	0.00	5.40	1.97	2.41	1.36	2.42	17.4	0.00	0.00	0.00	23.9	3.36	1.75	3.37	1.30	2.36	2.40	2.93	4.42	6.60	0.63	2.45	9.36	2.08	0.00
28	0.00	1.62	0.00	0.00	2.91	1.82	1.45	0.00	0.00	0.00	9.62	0.00	0.00	0.00	24.5	0.79	0.00	2.97	0.00	0.00	1.38	3.12	4.55	6.70	0.65	2.54	9.41	2.11	20.8
29	4.48	14.4	13.2	9.90	1.30	0.00	0.00	11.4	0.64	12.6	2.34	0.00	0.00	0.00	9.56	9.83	20.4	0.50	0.00	8.78	0.66	0.00	0.00	1.28	0.00	0.00	1.87	0.66	0.00
30	3.95	5.14	13.5	9.28	1.23	0.00	0.00	11.1	0.59	12.1	1.41	0.00	0.00	0.00	9.45	8.54	21.0	0.00	0.00	10.2	2.40	0.00	0.00	1.35	0.00	0.00	1.82	0.63	10.9
31	4.94	14.5	8.35	6.28	1.32	1.19	0.00	8.51	0.61	8.16	2.84	0.00	0.00	0.00	5.29	9.71	0.00	0.00	0.00	5.11	0.00	0.00	0.48	0.46	0.00	0.00	3.41	1.33	0.00
32	9.84	5.64	17.2	13.7	2.08	0.00	0.00	15.4	1.31	16.2	0.44	0.00	0.00	0.00	5.95	15.2	0.00	0.00	0.00	13.5	2.59	0.00	0.00	0.46	0.00	0.00	3.16	0.89	6.91
33	0.00	0.42	0.00	0.63	0.00	0.00	0.00	0.00	0.00	0.00	7.15	0.00	0.00	0.00	2.02	0.00	2.29	0.00	1.38	1.18	0.60	0.00	0.61	0.00	1.08	0.55	0.56	2.12	1.96
34	0.00	0.00	0.00	0.00	0.00	0.47	0.49	0.00	0.00	0.00	3.06	0.00	0.00	0.00	2.12	0.00	1.41	0.00	0.42	0.47	0.00	0.00	0.55	0.00	0.82	0.57	0.52	2.10	6.04
35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.93	0.00	6.77	1.24	1.29	0.00	4.43	0.00	0.00	2.88	3.75	0.00	1.29	1.67	0.00	1.35	0.68	2.40	0.00	1.68	9.78
36	0.43	0.52	0.66	0.59	0.00	0.00	0.00	0.59	0.00	0.59	3.65	1.22	1.25	0.00	4.45	0.00	1.19	0.00	0.61	0.46	0.51	1.73	0.00	0.67	0.95	2.67	0.59	1.83	6.05
37	0.00	7.36	1.57	0.59	0.64	0.00	0.00	1.40	0.00	1.20	0.43	0.00	0.00	0.00	12.3	0.00	2.88	23.2	1.67	1.30	0.00	0.00	0.67	2.68	0.00	0.00	0.23	0.69	0.00
38	0.00	3.28	0.89	0.00	0.00	0.00	0.00	0.00	0.00	0.53	0.94	0.00	0.00	0.00	12.4	0.00	0.64	0.00	2.97	0.00	0.00	0.00	0.70	2.70	0.00	0.00	0.42	0.65	9.55
39	0.00	5.26	3.72	1.65	0.52	0.00	0.00	1.52	0.00	1.45	0.00	0.00	0.00	0.00	6.38	0.56	2.60	0.00	1.38	0.00	0.00	0.43	0.25	1.08	0.00	0.00	2.28	1.54	0.00
40	0.00	1.84	1.65	0.54	0.00	0.80	0.59	1.25	0.00	1.37	0.00	0.00	0.00	0.00	6.52	0.00	0.48	0.00	1.51	0.00	0.00	0.00	0.00	0.98	0.00	0.00	2.37	1.51	8.85
41	1.72	25.4	0.00	0.61	0.66	4.11	1.68	0.00	0.00	0.00	0.00	0.66	1.69	0.00	2.74	1.30	4.91	1.25	0.00	2.46	1.67	0.00	0.00	8.09	0.00	0.00	2.21	1.32	2.51
42	0.92	16.1	1.11	1.37	1.35	2.51	1.20	0.98	1.71	1.66	2.37	1.45	1.76	0.00	2.85	0.44	6.99	1.27	1.28	0.00	1.42	0.00	0.00	8.06	0.00	0.00	2.27	1.32	4.06
43	0.92	21.3	1.28	1.25	1.32	5.14	2.33	1.26	2.36	0.64	1.30	0.00	1.41	0.00	2.18	0.00	3.22	1.26	2.14	1.94	1.34	0.00	0.00	6.91	0.00	0.59	3.50	1.43	4.77
44	0.00	0.00	0.00	0.00	0.00	5.10	1.39	0.00	0.00	0.00	1.65	0.00	1.47	0.00	2.92	0.00	3.80	2.69	1.33	0.57	1.21	0.00	0.00	6.61	0.00	0.00	3.51	1.40	5.58
45	1.27	1.53	1.39	0.00	1.54	1.30	0.00	0.65	0.79	1.94	1.24	0.00	0.00	0.00	1.83	2.03	2.69	0.58	0.00	2.81	0.00	1.77	5.49	3.74	0.00	3.25	2.40	1.24	2.14
46	0.00	0.00	0.00	0.00	0.00	2.33	0.00	0.00	0.00	0.00	0.54	0.00	0.00	0.00	1.96	0.00	0.00	0.00	0.84	0.00	0.00	1.50	3.93	2.58	0.00	2.15	1.44	0.61	4.92
47	0.95	3.02	1.66	1.37	1.29	2.01	3.76	1.37	1.29	1.67	1.39	1.40	1.63	0.00	0.34	1.71	4.57	1.43	1.34	2.22	1.33	1.47	0.00	1.93	0.00	1.63	1.38	4.98	1.65
48	0.00	2.55	0.00	0.00	1.44	0.57	0.00	0.00	0.00	0.00	1.37	0.65	1.53	0.00	1.21	0.60	0.65	1.68	0.00	0.94	0.45	1.23	0.00	0.00	0.00	1.50	1.50	0.00	3.61

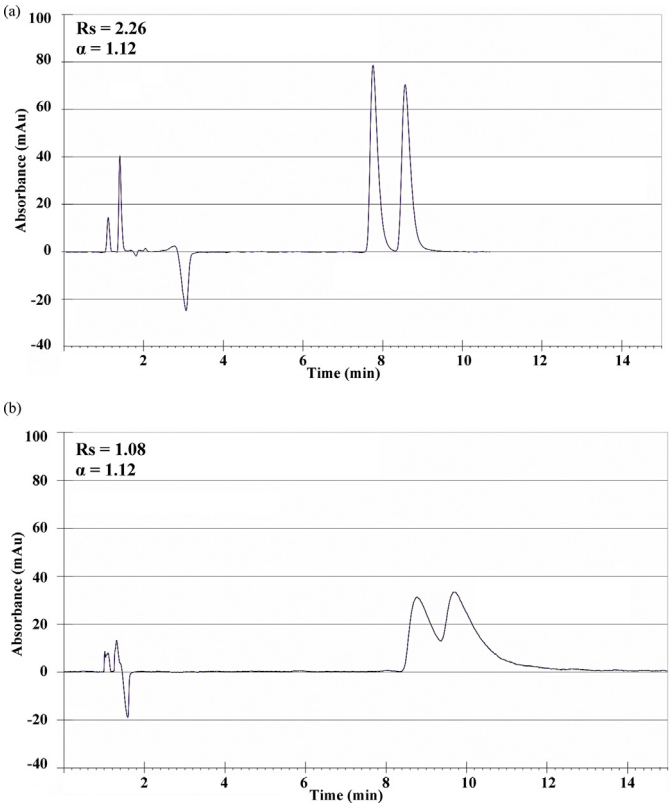


Fig. 2. Enantioseparation of (a) chlorphenamine and (b) acebutolol on Sepapak®-5 and Lux® Cellulose-1, respectively, using 80/20 (v/v), CO_2/(isopropanol + 0.1% isopropylamine + 0.1% trifluoroacetic acid) as mobile phase. Other conditions: UV-detection at 220 nm, flow rate 3.0 ml/min, 30 °C and 150 bar backpressure.

with Rs_A the autoscaled resolution, $\overline{Rs}$ the average resolution for one compound and s(Rs) the standard deviation of the obtained resolutions for one compound.

The first three projection pursuit features, i.e. the new latent variables, were considered (Fig. 3). It is clear that most systems form one central cluster. This does not necessarily imply that these systems have a similar separation pattern. The systems representation is possibly compressed, due to the strong deviating patterns of systems 25–28 and 37.

Chromatographic systems using the same CSP are generally located close to each other on the projection plot. The chiral selector is assumed to be the strongest determinant of the system's enantioselectivity. Systems 25–28 and 37 can be considered as 'outliers'. The loading plots were investigated to indicate which compounds exert the largest influence on the projection pursuit features.

For PP1, promethazine (compound 18) has the largest influence. In this direction, system 37 is located outside the data cloud. The Rs of promethazine on system 37 is 23.2 (Table 3, Fig. 4), which is significantly higher than the second highest Rs for this compound (3.37 on system 27) and the average Rs of 1.30 for system 37, which explains its location outside the data cluster. System 37 consists of Chiralpak® AS-H (amylose tris((S)-α-methylbenzylcarbamate) and 80/20 (v/v), CO_2/(MeOH + 0.5%IPA) as MP. No other tested CSP contains the same selector, thus no comparison can be made with equivalent systems.

To verify the dissimilarity of system 37 towards 38, 39 and 40 (all using AS-H as CSP), the resolutions of the 29 racemates on these systems are plotted in Fig. 4. The largest different in Rs between systems 37 and 38–40 is found for promethazine. This compound remains unseparated on systems 38–40, while a Rs of 23.2 is obtained on system 37. Warfarin is only separated on systems 38 and 40. Ambucetamide is separated on all systems, but the Rs is much higher on system 37. The same applies for oxazepam which is resolved with much higher Rs on systems 37 and 38 (the MeOH-based mobile phases). For the other compounds, the separation pattern is more similar on the four systems. Hence in this case, the dissimilarity of system 37 towards 38–40 is mainly based on the enantioseparation of promethazine. This property might be relevant in cases were a specific enantioselectivity is sought.

Systems 25–28, all with Chiralpak® AD-H are outside the central cluster. Different compounds exert a large influence on the position of the systems. The outlying positions of systems 25–28 can be explained by very high or low values for these specific compounds.

Using projection pursuit, specific aberrant enantioselective behaviours are highlighted and indicated as outliers. The position of a given system in the PP plot is a reflection of its entire enantioseparation pattern but is affected by outlying separations. When searching for dissimilar but complementary patterns, for example in the context of a screening, PP is thus less useful. The reason is that the applied patterns indicate dissimilar systems but do not take complementarity into account.

3.2.2. Correlation coefficient

Chromatographic systems that differ significantly in their enantioselective behaviour have a low correlation between their respective resolution patterns. Thus, the correlation coefficient (r) can be used as a parameter to identify systems with a similar or dissimilar enantioselectivity; $1 - |r|$ is then a measure for dissimilarity.

$$r = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2 \sum_i^n (y_i - \bar{y})^2}} \quad (2)$$

with n the number of resolutions from the test substances characterizing the chromatographic systems, i a given test substance, x_i a racemate resolution on a first system, and y_i the corresponding resolution on a second system, while $\bar{x}$ and $\bar{y}$ are the average resolutions on both systems [39].

For the resolution data of the 48 systems, the correlation matrix was determined and visualized as a colour map, where the absolute value of the correlation coefficient ($|r|$) is transferred into a colour (Fig. 5). The higher r between two systems, the stronger the linear relationship between the enantioselective pattern of the systems. This does not necessarily imply that the resolution pattern is identical, but anyway very similar and thus it would have little benefit to screen both systems. The correlation between a successful chromatographic system resulting in high resolutions and one failing to resolve most compounds, will be low [40].

The most dissimilar systems with all others are 21 till 24 (Sepapak®-5) and 37 till 40 (Chiralpak® AS-H) (Fig. 5). This can be derived from the high density of (dark) blue colour, representing a low correlation, of these systems towards the others. The selectors in these CSPs are structurally different from the others. Sepapak®-5 contains cellulose *tris*(3,5-dichlorophenylcarbamate) and is the only stationary phase with a selector containing two chlorine moieties. The chlorine substituents acts as electron withdrawing agents and significantly influence the electron density of the carbonyl group. The acidity of the —NH group increases and the possibility of forming hydrogen-bonds with analytes increases [17,18]. Chiralpak® AS-H contains amylose *tris*((S)-α-methylbenzylcarbamate) and is the only evaluated phase with a mono-substituted benzene moiety. The methyl group acts as an electron-donor and increases the electrondensity at the carbonyl oxygen atom. Consequently, analytes with electron donating substituents interact strongly with this functional group. Possibly these unique structural features yield unique enantioselective

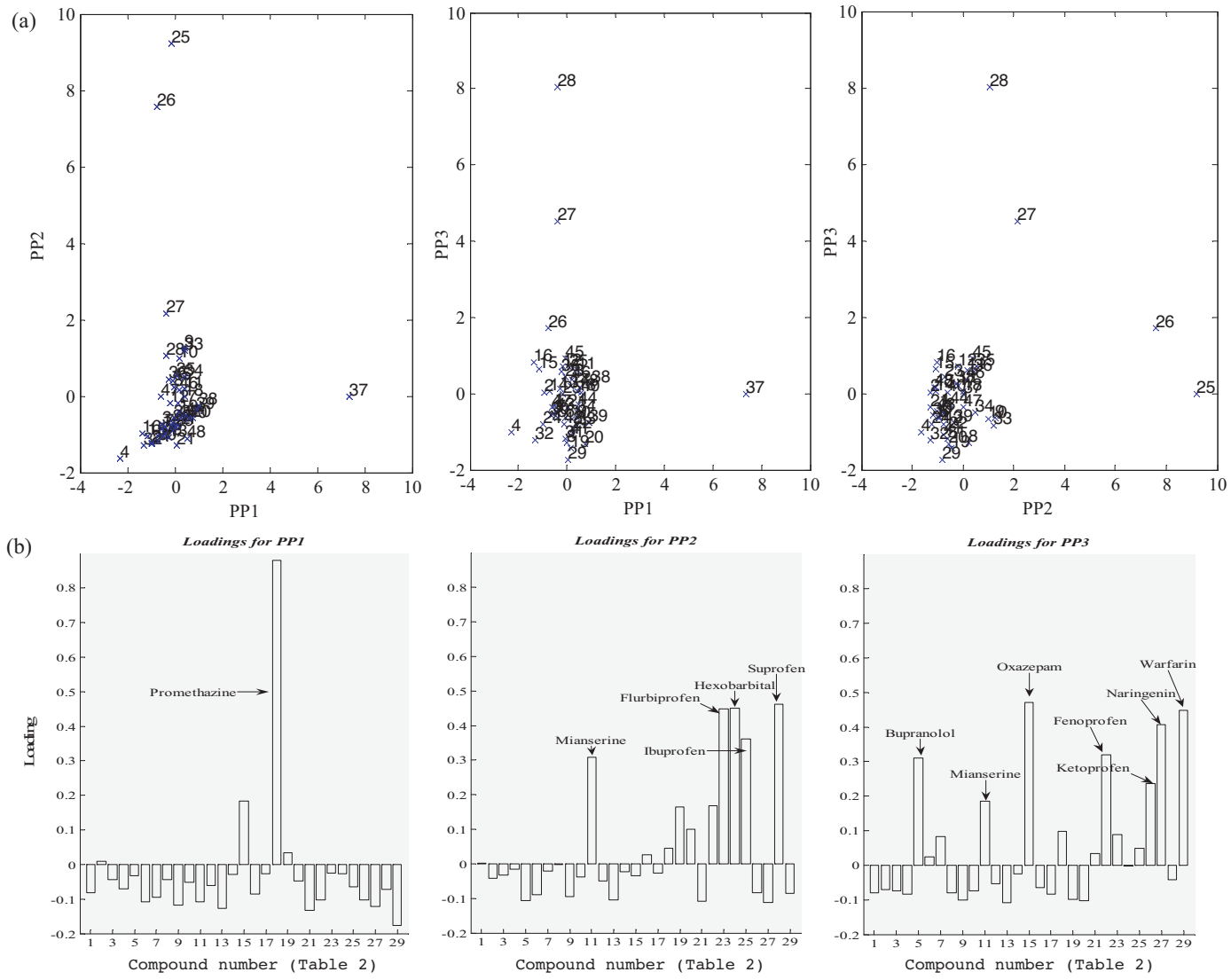


Fig. 3. (a) Score plots of the kurtosis projection pursuit of the 48 systems (29 Rs characterize each system), and (b) the loadings for each projection pursuit feature.

behaviours of these chiral selectors with high dissimilarity to the other selectors.

In general, blue shades dominate the entire colour map, implying that the correlation between the different systems is rather low. From this map it is clear that systems using different chiral selectors are strongly dissimilar. This was rather expected, since the chiral selector is assumed to be the primary determinant of enantioselective behaviour. Other tendencies, explained hereafter, can also be derived from the colour map.

The systems included in the boxes I–VIII exhibit higher correlations, i.e. orange shades are more present. The systems in these boxes use the same chiral selector. Hence, systems using the same chiral selector, but from a different manufacturer, and the same modifier are more correlated. This was also rather expected, since the chiral selector is assumed to be the primary determinant of enantioselective behaviour. Changing the modifier has less effect on dissimilarity.

Boxes VI, VII and VIII all include systems with amylose-based selectors. In these boxes, the correlation between the systems can be rather low. A certain degree of correlation exists between systems using the same selector but different MeOH-containing MPs, e.g. systems 17/18/45/46, and 25/26. The correlation between the systems with 2PrOH-containing MPs is much lower, e.g. 19/20/47/48, 27/28, and 39/40. Even less correlated are the systems using the same selector but different modifiers (MeOH or 2PrOH) e.g. 17/18/45/46 vs. 19/20/47/48; 25/26 vs. 27/28; and 37/38 vs. 39/40. Hence, the enantioselectivity of these amylose-based CSPs seems to be more influenced by the modifier type. Adding, a polar solvent to the mobile phase, induces a local conformation change in the amylose-chain. Hydrogen bonding interactions and π–π interactions become more dominant, while the carbonyl bond relaxes allowing a better inclusion of analytes [41,42]. Since, the conformational structure of cellulose-chains is less dense, their enantioselective behaviour in terms of stereoinclusion is potentially less affected. The chiral selectors of boxes VII (Chiralpak® AD-H) and VIII (Chiralpak® AS-H) seem most sensitive to changes in the mobile phase composition. Hardly any correlation is found when using these selectors with different MP compositions.

In general a higher correlation is found between systems using the same cellulose-based chiral selector and MPs with MeOH, e.g. 1/2/29/30, 5/6/41/42, 9/10/33/34, and 13/14. For the MPs containing 2PrOH, this trend is less pronounced and much lower correlations are found; e.g. 3/4/31/32, 7/8/43/44, 11/12/35/36, 15/16. The enantioresolution patterns generated by an identical CSP with different modifiers (MeOH or 2PrOH) are generally low correlated.

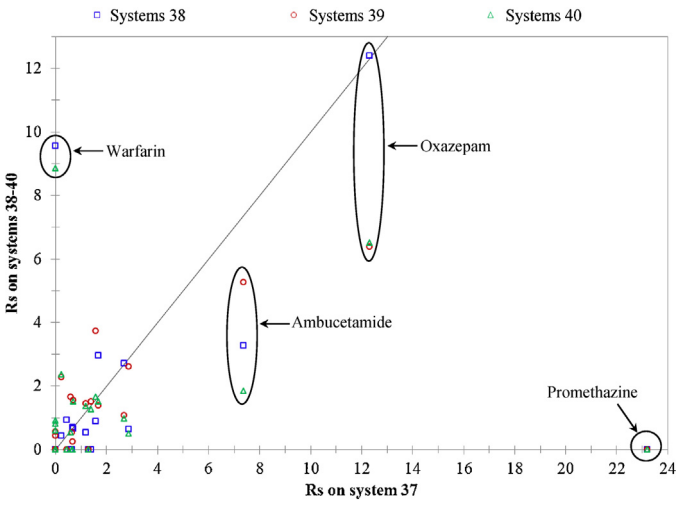


Fig. 4. Resolutions of the 29 racemates obtained on systems 38, 39, 40 plotted against the resolutions obtained for these compounds on system 37.

The chiral selector thus has the largest impact on the enantioresolution pattern of a chromatographic system. The mobile-phase composition plays a more important role in the enantioselectivity of amylose-based than of cellulose-based CSPs, although switching to another modifier type always generates a chromatographic system with some degree of dissimilarity to the previous.

To illustrate the applicability of this colour map in selecting dissimilar (and similar systems), the resolution data from highly correlated (1/2) and low correlated (23/34) systems are shown (Fig. 6). As some points coincide, e.g. compounds not separated on both systems, the map does display all 29 enantioseparations. The resolution patterns of systems 1 and 2 are indeed similar, while the patterns of systems 23 and 24 are very different.

3.2.3. Euclidean distance

Another approach to quantify the (dis)similarity between two systems (x and y) is the Euclidean distance (D). D calculates the true geometrical distance between two individual systems (x and y) according to:

$$D = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (3)$$

with n the number of resolutions [39]. The larger the Euclidean distance between two points, the more dissimilar they tend to be [19,43,44].

The higher D between two systems, the more dissimilar their enantioresolution patterns. The resulting colour map thus also illustrates the dissimilarity of the systems. In Fig. 7, a colour map of the Euclidean distances between the 48 chromatographic systems is presented. Systems from I, and VII show large D to all systems, illustrated by the relatively high density of orange/yellow in the map for these systems. Blue shades dominate within the boxes I–VIII, implying small Euclidean distances between the systems included a box. Systems in one box use the same chiral selector with different MPs. Again it is thus confirmed that the primary determinant of the enantioselective behaviour is the chiral selector. Box VII is somewhat exception to this trend seen the fact that the blue colour is less dominant here.

The two colour maps in Figs. 5 and 7 show a different degree of dissimilarity. For some systems, a low correlation is found (dark blue colour) and contradictory also a small Euclidean distance (green/blue shades). A possible explanation could be that dissimilar systems have a r value close to the lower limit zero (and a scale till one), while with D, dissimilarity results in a high value on a scale that is unlimited. Therefore high D values show more variability on a colour plot than the low r values. Compared to the r colour map, the dissimilarity between systems using the same selector and different MPs seems rather low. To verify which colour map best reflects the actual situation an example is given (systems 19 and

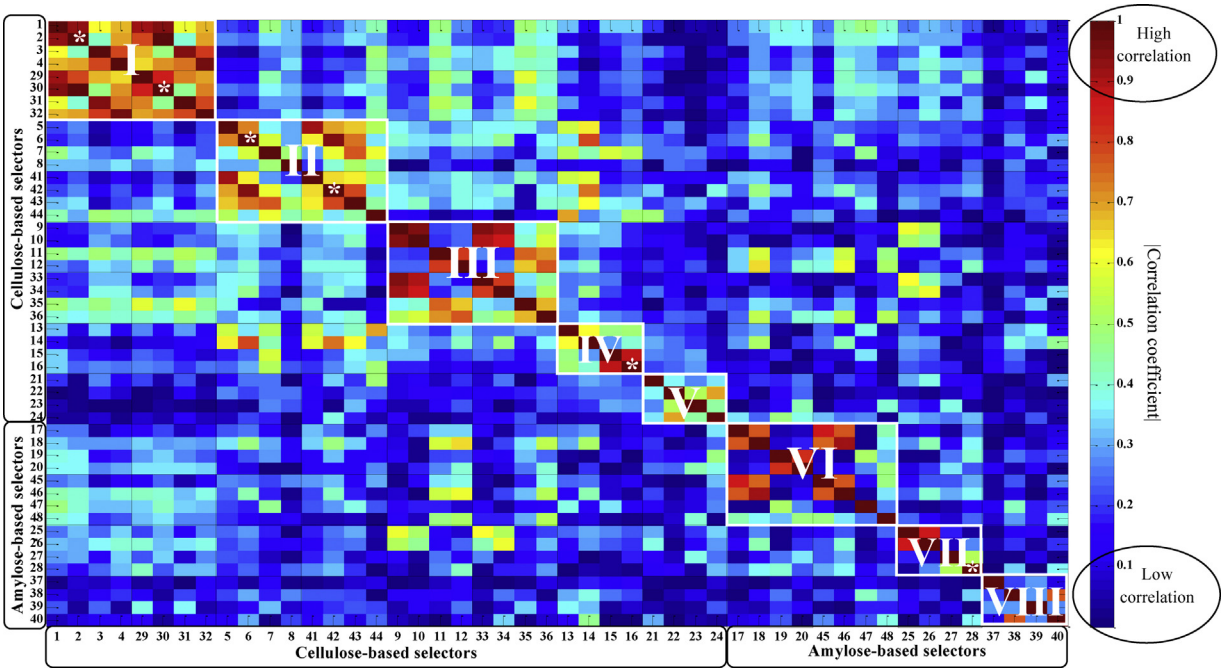


Fig. 5. Colour map of correlation coefficients between the 48 chromatographic systems (characterized by 29 enantioresolutions). The systems (numbered as in Table 2) are ranked according to the chiral selector of the stationary phase (equivalent selectors are placed sequentially and all cellulose-based selectors proceed the amylose-based). The systems with the same selector are highlighted with a box: I = LC-1 and OD-H, II = LC-2 and OZ-H, III = LC-3 and OJ-H, IV = LC-4, V = SP-5, VI = LA-2 and AY-H, VII = AD-H and VIII = AS-H. Systems marked with an asterisk are included in an earlier defined screening step [21].

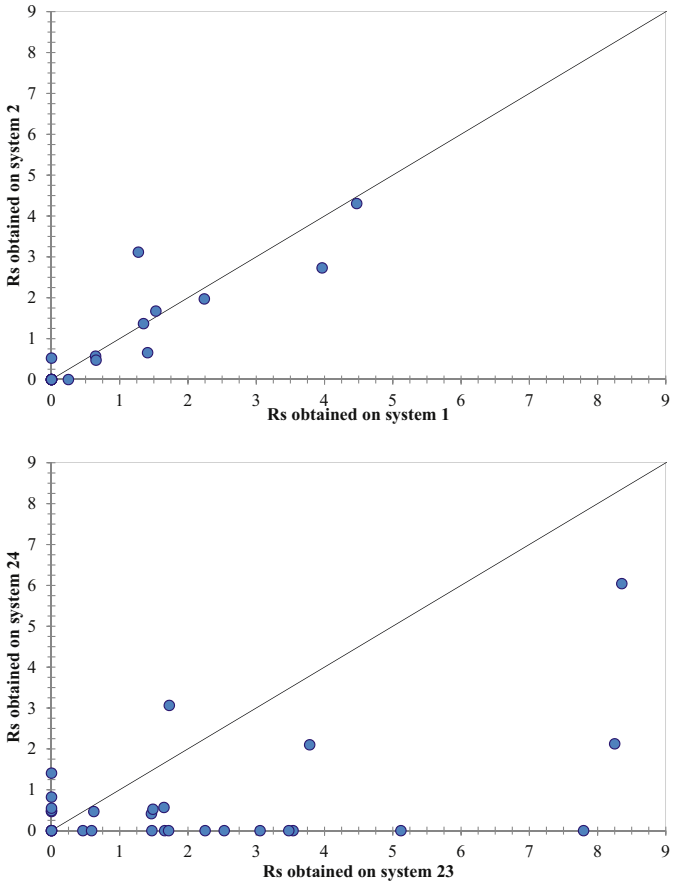


Fig. 6. Comparison of resolutions of the 29 racemates obtained on systems (a) 1 and 2 (highly correlated); (b) 23 and 34 (low correlation). System 1 = LC-1 with 80/20, CO_2/(MeOH + 0.5%TFA or IPA); system 2 = LC-1 with 80/20, CO_2/(MeOH + 0.1%TFA + 0.1%IPA); system 23 = SP-5 with 80/20, CO_2/(2PrOH + 0.5%TFA or IPA); and system 34 = OJ-H with 80/20, CO_2/(MeOH + 0.1%TFA + 0.1%IPA).

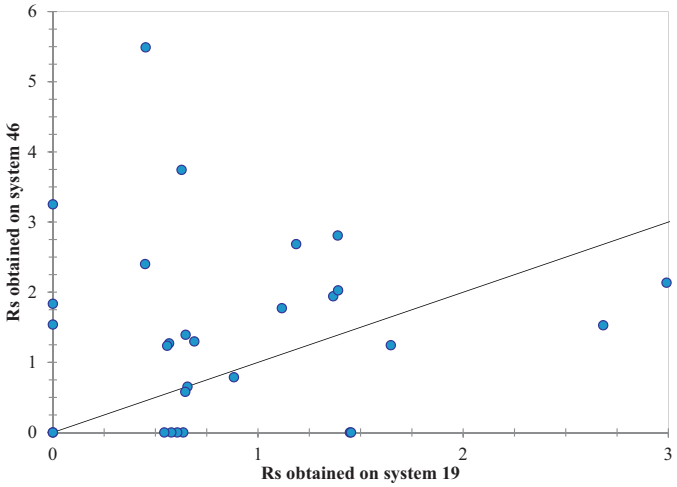


Fig. 8. Comparison of resolutions of the 29 racemates obtained on systems 19 and 46. System 19 = LA-2 with 80/20, CO_2/(2PrOH + 0.5%TFA or IPA), and system 46 = AY-H with 80/20, CO_2/(MeOH + 0.1%TFA + 0.1%IPA).

46) with obviously different degrees of dissimilarity in both colour maps. These systems seem highly dissimilar in the correlation map, while *D* is rather small (indicating similarity). The resolution patterns indicate that the first situation actually applies best (Fig. 8). The systems are rather dissimilar. Hence, the correlation-based colour map seems more relevant to select dissimilar systems.

3.3. Selection of dissimilar systems to compose a generic chiral screening procedure

3.3.1. Projection pursuit

The systems located at the edges of Fig. 3 would have to be selected, because they are dissimilar. However, several systems are located at these edges, making the choice rather empirical.

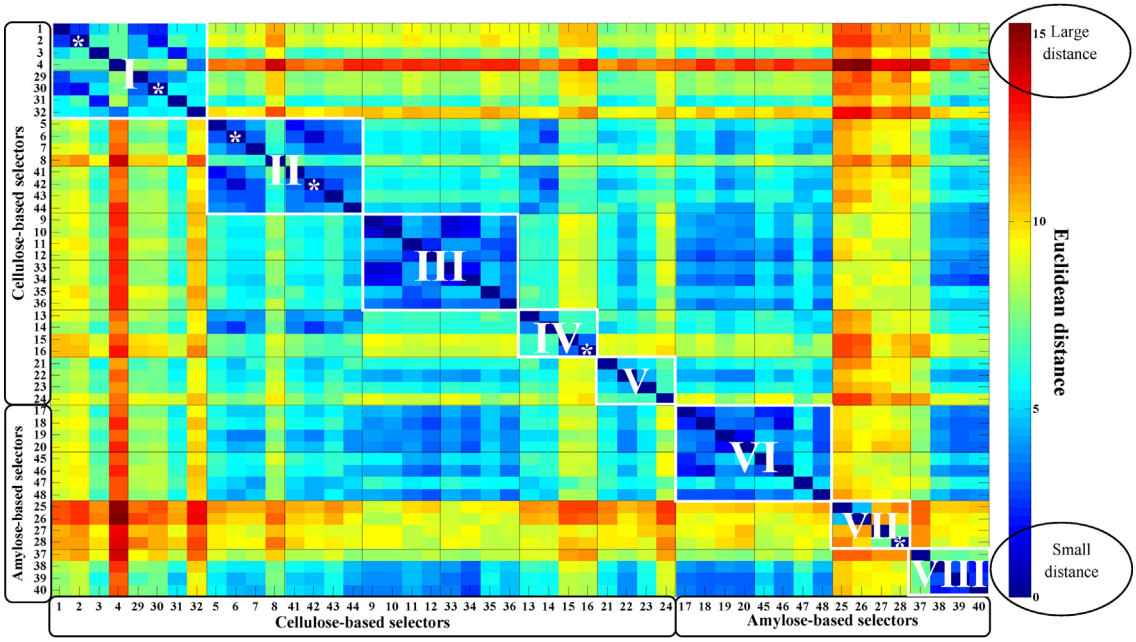


Fig. 7. Colour map of Euclidean distances between the 48 chromatographic systems (characterized by 29 enantioresolutions). The systems (numbered as in Table 2) are ranked according to the chiral selector of the stationary phase (equivalent selectors are placed sequentially; and all cellulose-based selectors proceed the amylose-based). The systems with the same selector are highlighted with a box: I = LC-1 and OD-H, II = LC-2 and OZ-H, III = LC-3 and OJ-H, IV = LC-4, V = SP-5, VI = LA-2 and AY-H, VII = AD-H and VIII = AS-H. Systems marked with an asterisk are included in an earlier defined screening step [21].

Table 4
Screening sequences derived from the PP plot (Fig. 3) with the number of partial (0 < Rs < 1.5), baseline (Rs > 1.5) and total separations (Rs > 0). A slash between two systems indicates no preference (alternative systems). Test set 1 contains 29 racemates, test set 2 contains all 57 racemates from [30].

	Screening sequence (system numbers in Table 2)	Partial separations		Baseline separations		Total separations	
		Test set 1	Test set 2	Test set 1	Test set 2	Test set 1	Test set 2
(a)	25 → 27 → 37 → 29	2	5	25	48	27	53
(b)	25 → 27 → 37 → 4	1	3	27	50	28	53
(c)	25 → 28 → 37 → 29	3	7	24	46	27	53
(d)	25 → 28 → 37 → 4	2	4	26	48	28	52
(e)	26 → 27 → 37 → 29	1	6	26	47	27	53
(f)	26 → 27 → 37 → 4	1	4	27	49	28	53
(g)	26 → 28 → 37 → 29	3	8	24	45	27	53
(h)	26 → 28 → 37 → 4	2	5	26	47	28	52
(i)[a]	6/42 → 28 → 2/30 → 16	4	9	25	48	29	57

[a] Sequence from Ref. [15].

For instance: 25 or 26, 27 or 28, 37 and 29 or 4 could be selected when screening is limited to four systems (Table 4). The choice between these options leads to eight possible screening sequences. Nevertheless, the separation rates of these sequences are similar. When system 4 is included in the screening approach instead of 29, one extra separation is obtained for the 29-racemate test set. There is however no way of deriving information about the cumulative success rate or the systems complementarity from the PP plot. In earlier research a screening step (6/42 → 28 → 2/30 → 16) was derived, based on evaluation of the systems' success rates and complementary enantioseparations [21]. This sequence allows separating the entire 29-compounds test set.

When applying the screening sequences of Table 4 on the entire 57-compound test set from [21], similar success rates are obtained for all sequences. The earlier composed screening is slightly more successful in terms of the number of total separations.

Table 4 shows, as also seen in [45], that when combining three different polysaccharide-based selectors in a screening procedures, in general, high cumulative success rates are found.

Hence, PP can be applied as a tool to investigate and compare the enantioresolution pattern of chromatographic systems, with the aim of identifying systems with the most dissimilar behaviour. These might be selected to include in a screening. However, since this approach fails to take the complementarity of systems into account, there is no a priori guarantee of the efficiency of the defined screening sequence.

3.3.2. Correlation coefficients

Dissimilar systems can also be selected using the correlation coefficient. Overall, systems 21–24 from box V (SP-5) and 37–40 from box VIII (AS-H) show the highest degree of dissimilarity to all other systems (Fig. 5). These CSPs can be used in a screening procedure. Next, the system(s) showing the largest dissimilarity to the first two systems could be selected. All systems included in a screening should have a low correlation and high cumulative success rate. This last parameter (as a consequence of complementarity) is however not verifiable from the colour map.

A possible selection (of many) derived from the colour map is: 22/34/37/45. No best screening sequence can be derived, since no information about success rate or complementarity is given in the colour map. With this screening step, 22 baseline and 5 partial separations would be achieved. When this screening step is applied on the 57-compounds test set used in [15], 10 partial and 42 baseline separations are obtained. This screening step could thus be considered relatively successful. However, other screening sequences could also be selected. In fact, the selection based on colour maps is rather arbitrary and it is not straightforward to decide which systems are best.

As mentioned earlier the following screening was derived based on the systems' success rate and complementarity: 42/6 → 28 → 30/2 → 16 [21]. These systems are marked with an asterisk in the colour map and indeed show a low correlation. With this screening step the complete test set of 29 racemates can be separated (25 baseline resolved). The screening sequence derived from the colour map is thus somewhat, but not dramatically, less successful than the one from Ref. [15]. Composing a generic screening step, solely based on these correlations, would thus not necessarily yield the best results, since complementarity is not considered. Anyway, the high success rate from the arbitrary selection from the colour map again confirms observations from [45]: when combining three (or four) different polysaccharide-based selectors in a screening procedure, high success rates are seen.

The colour map could potentially be used in other approaches where the (dis)similarity between individual systems is investigated. For example when screening a given system did not result in a separation, one could easily choose the most dissimilar system to it from the colour map, increasing the chance of a successful separation compared to a random system selection. The colour maps could also be employed to search for the most similar chromatographic systems, when one is interested in an alternative system (with the same enantioselectivity).

3.3.3. Euclidean distance

Systems from box I using LC-1 (1–4) or OD-H (29–32) have the largest Euclidean distances to the other systems. The systems using AD-H (25–28) (box VII) also show a high degree of dissimilarity to the others. For the rest of the map, rather low distances can be noticed. A screening step would preferably thus include AD-H and LC-1 or OD-H, plus some systems they are dissimilar to. A possible screening step derived from this colour map would be: 4/8/16/26. Using these systems on the present test set would yield 28 separations of which 25 baseline. For the 57-compounds test set, 7 partial and 45 baseline separation would be obtained.

The systems (16, 28, 30/2, 42/6) from [15] are again marked with an asterisk in Fig. 7. This combination of systems leads to a separation of the entire 29-compounds test set with 25 baseline separations. However, the Euclidean distances between these systems are not always the largest (mainly green/yellow shades representing an average distance). Again, since the colour map does not consider complementarity of systems, there is no a priori guarantee for a high cumulative success rate. But anyway, again the observation from [45] is confirmed from the arbitrary selection mode.

Overall, these colour maps based on r and D allow evaluating the pair-wise and partially group-wise (dis)similarity of systems, but it is not evident to select the appropriate for a screening sequence. More specifically, selection of the first system, which should have

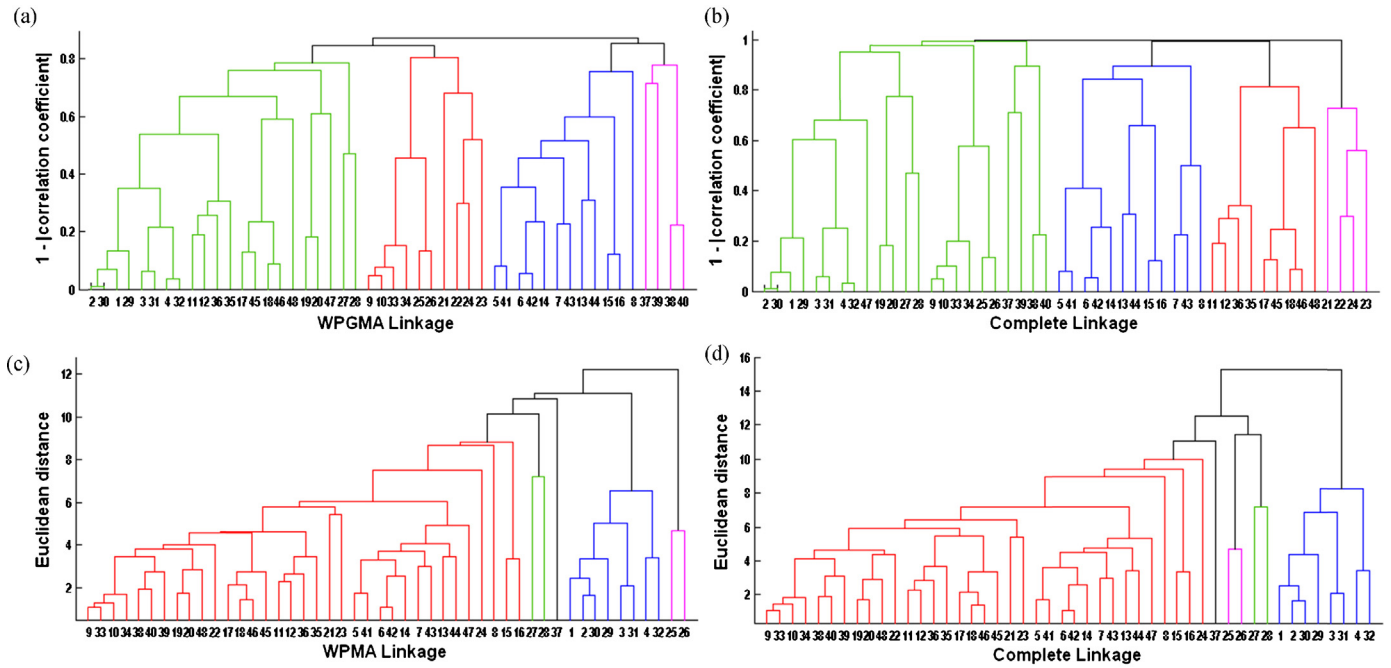


Fig. 9. Dendrograms of 48 chromatographic systems, (a, b) linked based on the dissimilarity criterion $1 - |r|$, using (a) WPGMA linkage clustering and (b) complete linkage clustering; and (c, d) on the Euclidean distance, using (c) WPGMA linkage clustering and (d) complete linkage clustering. On the *x*-axis, the system numbers are presented and divided into groups.

the broadest enantioselectivity, followed by the most complementary to it, is unfeasible from the colour maps.

3.3.4. Hierarchical cluster analysis (HCA)

Hierarchical clustering techniques group similar and dissimilar systems in dendrograms and enable visualization. West and Lesellier [23] used this technique to group chromatographic systems with similar elution orders for given test compounds. This allowed drawing conclusions about the selectivity. We evaluated this technique to group systems with a similar or dissimilar separation pattern as aim the selection for screening.

In agglomerative hierarchical cluster analysis, the most similar chromatographic systems are sequentially merged. The correlation coefficient or the Euclidean distance can be used as parameter to identify (dis)similar systems. Several methods can be applied to merge the similar systems. With complete linkage clustering, the distance between two clusters is considered to be represented by the largest distance (smallest correlation coefficient or largest Euclidean distance) between two individual systems in each cluster. With the weighted pair-group method arithmetic mean (WPGMA), the distance between two clusters is calculated as the average distance between all systems of one cluster and all systems of the other cluster [43]. Other linking methods exist, but these two appeared to be most appropriate for our data set since they allowed the best distinction of groups of systems with similar enantioselectivity.

In Fig. 9, four dendrograms are presented based on two dissimilarity criteria, i.e. $1 - |r|$ and D, respectively. Two linkage methods were applied, i.e. WPGMA and complete linkage. In each dendrogram, the most distinct clusters have a different colour. The dendrograms with the same dissimilarity criterion roughly group the same systems, while those constructed with different dissimilarity parameters (r or D) are completely different. This confirms the different information seen in the r and D colour maps.

The systems are clustered in a limited number of rather distinct groups in the dendrograms based on '$1 - |r|$' (Fig. 9a and b). No specific outliers can be indicated from these plots. The amylose-based selectors are not differentiated from the cellulose-based selectors. Systems with the same selector (amylose- or cellulose-based) are not always clustered in the same group, e.g. systems 9–12 and 33–36 (cellulose *tris*(4-methylbenzoate)), 17–20 and 45–48 (amylose *tris*(5-chloro-2-methylphenylcarbamate), and 25–28 (amylose *tris*(3,5-dimethylphenylcarbamate). Hence, by changing the MP in combination with a certain CSP, the parameter might be altered in such a way that it becomes more resembling to a system with a different chiral selector. This is also seen in the projection plot in Fig. 3 where systems with the same chiral selector (same colour) are not always located closest to each other.

Using the Euclidean distance, systems are rather sequentially grouped in one large cluster. Most distinct from the red cluster are systems 1–4 and 29–32 (blue cluster); 25 and 26 (purple cluster); and 27–28 (green cluster). Systems 1–4 and 29–32 have the same selector, cellulose *tris*(3,5-dimethylphenylcarbamate). Systems 25–28 all use Chiralpak® AD-H but are not clustered. System 37 is not included in any cluster. These latter outlying systems are thus expected to exhibit different enantioselective patterns to the other systems, which is confirmed by the colour (Fig. 7). However, it is hard to distinguish the dissimilarity between the systems included in the major red cluster.

To compose a successful screening set, dissimilar systems should be selected, e.g. system from each group. This system should have preferably the largest dissimilarity to the other groups. In other words, it should be clustered at the lowest level within the group. Because the link is always between two systems, two systems qualify. In this context, the D-based dendrograms seem less convenient, since most systems are grouped in a single cluster. The selected screening sequences from the dendrograms are presented in Table 5. In each step of the sequence, the user has the choice between two systems. With exception of the systems 22/24 and 25/26 it makes no difference which of both systems is chosen (see Table 5). Using systems 24 and 25 instead of 22 and 26, respectively, yields slightly higher success rates in terms of the number of baseline separations. The pairs 2/30 and 6/42 are selected in each sequence. They also belong to the screening

Table 5
Screening sequences derived from the dendrograms presented in Fig. 9. A slash between two systems indicates no preference (alternative systems). The resulting number of partial (0 < Rs < 1.5) and baseline separations (Rs > 1.5) after executing the sequences are also presented. A number of separations between brackets implies that the system between brackets in the screening was used. Test set 1 contains the 29 compounds, test set 2 all 57 compounds of Ref. [30].

	Dissimilarity criterion	Linkage method	Screening sequence (system numbers in Table 2)	Partial separations		Baseline separations		Total separations	
				Test set 1	Test set 2	Test set 1	Test set 2	Test set 1	Test set 2
(a)	1 – |Correlation coefficient|	WPGMA	2/30 → 6/42 → 9/10 → 38/40	7	11	20	44	27	55
(b)	1 – |Correlation coefficient|	Complete	2/30 → 6/42 → 18/46 → (22)/24	2 (4)	4	25 (23)	50	27	54
(c)	Euclidean distance	WPGMA	2/30 → 6/42 → 27/28 → 25/(26)	1 (2)	4	26 (25)	52	27	56
(d)	Euclidean distance	Complete	2/30 → 6/42 → 27/28 → 25/(26)	1 (2)	4	26 (25)	52	27	56
	Ref. [15] Initial selection	–	6/42 → 28 → 2/30 → 16	4	9	25	48	29	57

from [21]. The dendrograms as the other approaches only take into account the dissimilarity of the systems and not their complementarity. Table 5 again confirms Ref. [45]. The combination of three (or four) polysaccharide-based selectors results in high cumulative success rates.

In conclusion we further can state that dendrograms formed using $1 - |r|$ have potential to select dissimilar systems. WPGMA and complete linkage merge roughly the same systems. For the selection of systems to define a screening approach dendrograms fail to take into account the systems' complementarity. The derived screening procedure from a dendrogram is not necessarily the most efficient, although good results were obtained in our case.

3.3.5. Kennard and Stone algorithm

A final approach considered to select screening systems is the Kennard and Stone (K&S) algorithm. This algorithm adds sequentially the most distant systems, i.e. with the largest (Euclidean) distance, to the previously selected system(s). The selection can start selecting the two most distant systems or from the system closest to the mean of all systems. Both approaches were tried and four subsequent systems were selected by the algorithm. The selection starting with the system closest to the mean is: 22 → 4 → 25 → 2 and starting with the two most distant systems: 4 → 26 → 21 → 27. In this approach, a screening sequence is provided by the algorithm. For both sequences, the cumulative success rates for the 29-compound and entire 57-compounds test set were determined (Fig. 10). This cumulative rate is determined by summing the additional separations generated by a newly selected system and the total number of separations achieved on the previous system(s).

The complementarity of the four selected systems in each sequence is clear from Fig. 10. For test set 1 generate both screenings an equal total number of separated compounds, i.e. 28, but the screening starting with system 22 yields two extra baseline separations (24 vs. 26). The screenings selected with the approach of Kennard and Stone yield a similar separation rate compared to the earlier defined screening. The tendencies are the same when applying the screening sequences to the entire test set.

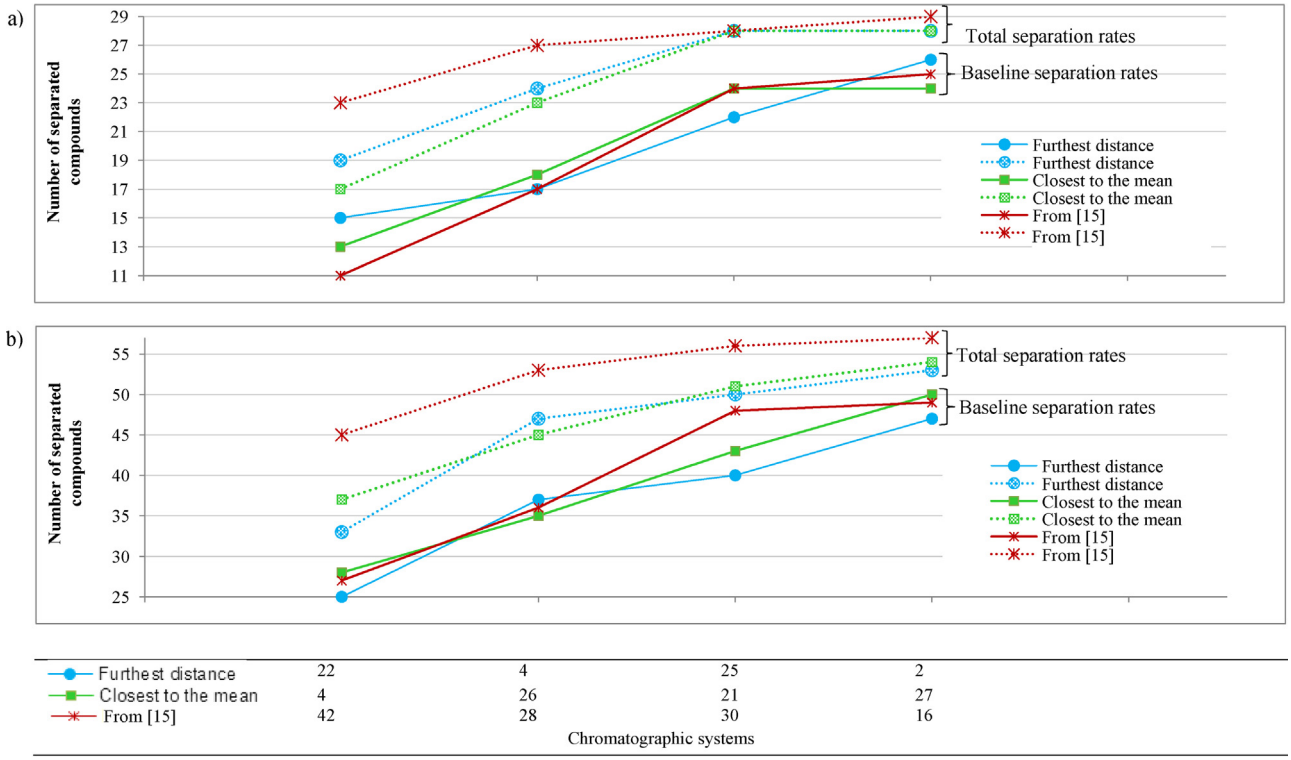


Fig. 10. Cumulative separation rates of the screening approaches defined using the Kennard and Stone algorithm. The blue curves represent the screening composed by starting with the two most distant (dissimilar) systems, and the green curves when starting with the system closest to the mean. The full lines represent the baseline separations and the dashed lines, the total separation rate. In (a) the results for the 29-compound test set are presented, in (b) for the entire 57-compound test set. The selected systems are presented below the figures (system number in Table 2). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

The results for this algorithm, but also the previous selection approaches considered, showed two things. First, the applied selections do not consider whether the first selected system has a broad (the broadest) enantioselectivity or whether the sequentially selected systems show maximal complementarity (in terms of the number of separated compounds) to the previously selected. Secondly, all selections made resulted in a high cumulative success rate for the reduced test set of 29 compounds. This seems rather the consequence of the good complementarity of the different CSPs than of the intelligence of these selection procedures which all neglected the requirements specified in the first conclusion. This second topic confirms what was also observed in [45].

4. Conclusion

A systematic evaluation of the chiral separation on 48 chromatographic systems was performed by means of chemometric techniques to identify (dis)similar chromatographic systems. A projection pursuit of the resolutions revealed several systems with an aberrant enantioselective pattern in relation to the majority of systems. This was caused by extreme Rs-values. In a next stage, the pair-wise dissimilarity of systems was investigated by means of correlation coefficient- and the Euclidean distance-based colour maps. This approach is useful in determining the (dis)similarity of two individual systems. It also confirmed that the main determinant for enantioselective behaviour is the chiral selector, rather than the mobile phase. However, this trend was less pronounced for the amylose-based stationary phases, on which the mobile phase seemed to have a larger impact. Colour maps were used to select dissimilar system for a screening procedure, but this selection is rather arbitrary.

A hierarchical cluster analysis identified the systems that express a similar enantioselectivity. Two clustering approaches were applied, i.e. WPGMA and complete linkage, with two dissimilarity parameters, i.e. $1-|r|$ and the Euclidean distance. Dendrograms were constructed, of which screening sequences with dissimilar systems were derived. The proposed screening systems showed their complementarity and high cumulative success rate. Dendrograms have potential to indicate dissimilar systems, but do not take into account the complementarity of systems. For this reason, the high cumulative success rate is possibly rather a coincidence, due to the general broad enantioselectivity of any combination of three or four polysaccharide selectors.

The Kennard and Stone algorithm was also applied for the selection of the most dissimilar systems. This technique allows the straightforward selection of dissimilar systems. However, since the success rate of the first selected system is not considered, there is no a priori guarantee that the systems selected by this approach will provide an efficient screening procedure.

Chemometric techniques indeed can be successfully applied to identify systems with a dissimilar enantioselective pattern. However, since the success rates of the first and subsequent systems are not considered, nor the complementarity, there is no a priori guarantee about success.

The selections made with the different techniques resulted in a high cumulative success rate for the reduced test set of 29 compounds as well as for the complete test set of 57. This seems rather the consequence of the good complementarity of the different polysaccharide-based CSPs than of the intelligence of these selection procedures which all neglected the requirements specified earlier.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgement

This work was financially supported by the Research Foundation Flanders (projects 1.5.114.10N/1.5.093.09N.00).

References

[1] M. Maftouh, C. Granier-Loyaux, E. Chavana, J. Marini, A. Pradines, Y. Vander Heyden, C. Picard, J. Chromatogr. A 1088 (2005) 67.
[2] Z. Pirzada, M. Personick, M. Biba, X. Gong, L. Zhou, W. Schafer, C. Roussel, C.J. Welch, J. Chromatogr. A 1217 (2010) 1134.
[3] C. White, J. Chromatogr. A 1074 (2005) 163.
[4] T. Zhang, D. Nguyen, P. Franco, J. Chromatogr. A 1191 (2008) 214.
[5] Z. Wang, in: S. Ahuja (Ed.), Chiral Separation Methods for Pharmaceutical and Biotechnological Products, Wiley, Calabash, USA, 2011, p. 299.
[6] A. Akin, F.J. Antosz, J.L. Ausec, K.F. Greve, R.L. Johnson, L.-E. Magnusson, T. Ramstad, S.L. Secreast, D.S. Seibert, G.K. Webster, Curr. Pharm. Anal. 3 (2007) 53.
[7] T.A. Berger, W.H. Wilson, J. Biochem. Biophys. Methods 43 (2000) 77.
[8] P. Franco, T. Zhang, J. Chromatogr. B: Analyt. Technol. Biomed. Life Sci. 875 (2008) 48.
[9] M.S. Villeneuve, R.J. Anderegg, J. Chromatogr. A 826 (1998) 217.
[10] T. Zhang, D. Nguyen, P. Franco, J. Chromatogr. A 1217 (2010) 1048.
[11] H. Ates, A.A. Younes, D. Mangelings, Y. Vander Heyden, J. Pharm. Biomed. Anal. 74 (2013) 1.
[12] H. Ates, D. Mangelings, Y. Vander Heyden, J. Chromatogr. B: Analyt. Technol. Biomed. Life Sci. 875 (2008) 57.
[13] A.A. Younes, H. Ates, D. Mangelings, Y. Vander Heyden, J. Pharm. Biomed. Anal. 75 (2013) 74.
[14] A. Hendrickx, D. Mangelings, B. Chankvetadze, Y. Vander Heyden, Electrophoresis 32 (2011) 2718.
[15] M.L. de la Puente, P. Lopez Soto-Yarritu, J. Burnett, J. Chromatogr. A 1218 (2011) 8551.
[16] H. Nelander, S. Andersson, K. Ohlen, J. Chromatogr. A 1218 (2011) 9397.
[17] B. Chankvetadze, J. Chromatogr. A 1269 (2012) 26.
[18] B. Chankvetadze, Methods Mol. Biol. 970 (2013) 81.
[19] E. Van Gyseghem, B. Dejaegher, R. Put, P. Forlay-Frick, A. Elkihel, M. Daszykowski, K. Heberger, D.L. Massart, Y. Vander Heyden, J. Pharm. Biomed. Anal. 41 (2006) 141.
[20] K. De Klerck, G. Parewyck, D. Mangelings, Y. Vander Heyden, J. Chromatogr. A 1269 (2012) 336.
[21] K. De Klerck, C. Tistaert, D. Mangelings, Y. Vander Heyden, J. Supercrit. Fluids 80 (2013) 50.
[22] C.G. da Silva, C.H. Collins, E. Lesellier, C. West, J. Chromatogr. A 1315 (2013) 176.
[23] C. West, E. Lesellier, J. Chromatogr. A 1302 (2013) 152.
[24] C. West, E. Lesellier, J. Chromatogr. A 1115 (2006) 233.
[25] C. West, E. Lesellier, J. Chromatogr. A 1110 (2006) 200.
[26] C. West, E. Lesellier, J. Chromatogr. A 1110 (2006) 181.
[27] E. Lesellier, C. West, J. Chromatogr. A 1158 (2007) 329.
[28] S. Khater, Y. Zhang, C. West, J. Chromatogr. A 1303 (2013) 83.
[29] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, in: B.G.M. Vandeginste, L. Kaufman (Eds.), Chemometrics: A Textbook, Elsevier Science Publishers, Amsterdam, The Netherlands, 1988, p. 339.
[30] M. Daszykowski, I. Stanimirova, B. Walczak, D. Coomans, Chemom. Intell. Lab. Syst. 78 (2005) 19.
[31] T. Cserhati, Biomed. Chromatogr. 24 (2010) 20.
[32] M.R. Euerby, P. Petersson, J. Chromatogr. A 994 (2003) 13.
[33] M.R. Euerby, P. Petersson, J. Chromatogr. A 1088 (2005) 1.
[34] M.R. Euerby, P. Petersson, W. Campbell, W. Roe, J. Chromatogr. A 1154 (2007) 138.
[35] M. Lammerhofer, M. Richter, J. Wu, R. Nogueira, W. Bicker, W. Lindner, J. Sep. Sci. 31 (2008) 2572.
[36] K. De Klerck, D. Mangelings, D. Clicq, F. De Boever, Y. Vander Heyden, J. Chromatogr. A 1234 (2012) 72.
[37] M. Daszykowski, B. Walczak, D.L. Massart, Chemom. Intell. Lab. Syst. 65 (2003) 97.
[38] S. Hou, P.D. Wentzell, Anal. Chim. Acta 704 (2011) 1.
[39] B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S. de Jong, P.J. Lewi, J. Smeyers-Verbeke, in: B.G.M. Vandeginste, S.C. Rutan (Eds.), Handbook of Chemometrics and Qualitmetrics: Part B, Elsevier, Amsterdam, 1998.
[40] E. Van Gyseghem, I. Crosiers, S. Gourvenec, D.L. Massart, Y. Vander Heyden, J. Chromatogr. A 1026 (2004) 117.
[41] B. Chankvetadze, in: G.K.E. Scriba (Ed.), Chiral Separations, Methods and Protocols, Humana Press, Jena, Germany, 2013, p. 81.
[42] S. Ma, S. Shen, H. Lee, M. Eriksson, X. Zeng, J. Xu, K. Fandrick, N. Yee, C. Senanayake, N. Grinberg, J. Chromatogr. A 1216 (2009) 3784.
[43] D.L. Massart, L. Kaufman, in: P.J. Elving, J.D. Winefordner, I.M. Kolthoff (Eds.), The Interpretation of Analytical Chemical Data by the use of Cluster Analysis, John Wiley & Sons, Brussels, Belgium, 1983, p. 1.
[44] M. Dumarey, R. Put, E. Van Gyseghem, Y. Vander Heyden, Anal. Chim. Acta 609 (2008) 223.
[45] K. De Klerck, Y. Vander Heyden, D. Mangelings, Pharmaceutical-enantiomers resolution using immobilized polysaccharide-based chiral stationary phases in supercritical fluid chromatography, J. Chromatogr. A (2014).