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

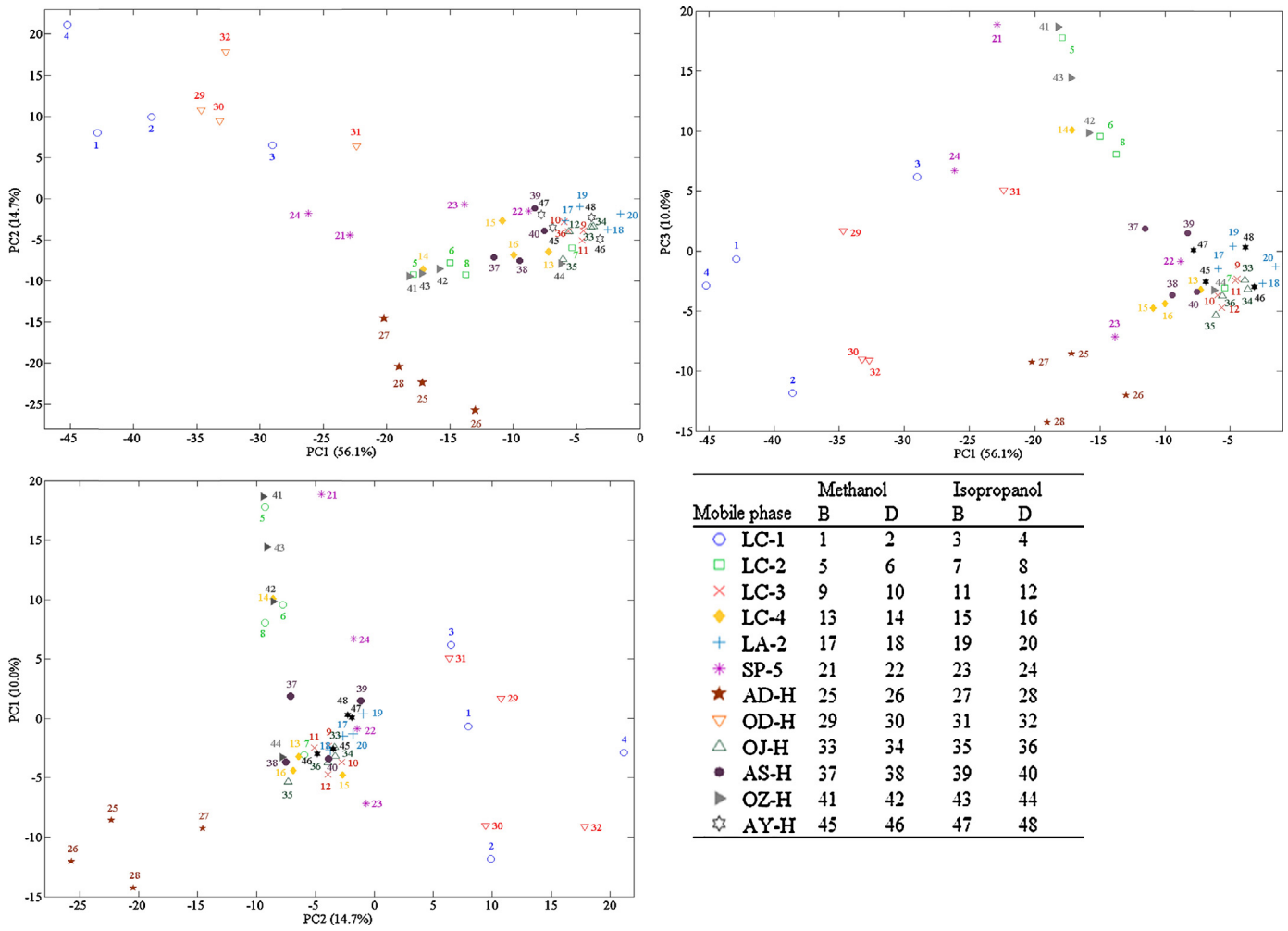
A green technique getting special attention is supercritical fluid chromatography (SFC). It has repeatedly been established as a fast and high-performance chiral-separation technique for a wide range of compounds [1–5]. However, optimal chromatographic chiral separation conditions are very tedious to 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