

Available online at www.sciencedirect.com

JOURNAL OF CHROMATOGRAPHY B

Journal of Chromatography B, 845 (2007) 253–260

www.elsevier.com/locate/chromb

HPLC columns partition by chemometric methods based on peptides retention \dot{x}

Bogusław Buszewski^{a,∗}, Sylwia Kowalska^a, Tomasz Kowalkowski^a, Katarzyna Rozpędowska^a, Monika Michel^b, Tobias Jonsson^c

^a *Department of Environmental Chemistry and Ecoanalytics, Faculty of Chemistry, Nicolaus Copernicus University,* Gagarina 7 St., 87-100 Toruń, Poland

^b Department of Pesticide Residue, Plant Protection Institute, Miczurina 20 St., 60-318 Poznań, Poland ^c *SeQuant, Box 7956, SE-90719, Umea, Sweden*

> Received 13 June 2006; accepted 28 October 2006 Available online 20 November 2006

Abstract

In recent years, multivariate techniques have been utilized to evaluate reversed-phase high-performance liquid chromatographic data. In the present study, 11 high-performance liquid chromatography (HPLC) columns were divided into several groups according to the retention factors of 12 peptides. Principal component analysis (PCA) and cluster analysis (CA) were used in column and peptides' comparison and grouping. CA results indicated that all stationary phases may be generally grouped into several clusters, due to stationary phase structure and properties. On the other hand, interesting results were obtained with the use of PC. There is almost linear relationship between classified HPLC columns in the space of new PCs, which is connected with meaning of the PC's reflected in their loading values. The first component describes non-polar properties of peptides, whereas the second component is loaded with polar peptides having much lower log *P* values. PCA and CA were also used in peptides comparison however, complete explanation of peptides grouping still remains unclear.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Peptides properties; HPLC; Stationary phase type; Column grouping; Cluster analysis; Principal component analysis

1. Introduction

The significant importance of peptides for living organisms is well known. The abilities to detect, quantify, and model biologically significant molecules are important for studying basic living function. These abilities are important especially in the field of bioanalytics. The bioanalytic study focuses on specific aspects of the biomolecular detection and the analysis of problems as in the case of peptides. Their analysis has numerous applications (e.g. separation and purity control) [\[1\].](#page-6-0) The most popular method for peptide analysis is high-performance liquid chromatography (HPLC). Due to the specific structures and properties of peptides, their chromatographic analysis may provide interesting information for column grouping or can be

helpful to select proper column for peptide investigations. The correct selection of stationary and mobile phases are crucial points for the application of the HPLC technique for peptides analysis. The majority of columns used in the studies are butyl, octyl and octadecyl stationary phases [\[2\].](#page-6-0) Nowadays, there are many different packing material types and their utilization in peptide analysis seems to be a promising direction in case of choosing the best column for peptides' determination, e.g. in biological sample.

On the other hand, there are often significant differences between similarly prepared and commercially available columns, mainly due to the characteristics of the silica material used as a support and the technique applied to synthesize packing material. Not only individual adsorbents differ in their dependence on the manufacturer, but the batch-to-batch repeata-bility of the same producer is also rather poor in many cases [\[3\].](#page-6-0) Therefore, methods for column comparison and grouping are essential. One of the most popular methods for column quality evaluation are simple chromatographic tests characterizing

 $*$ This paper was presented at the 10th International Symposium on Biochromatography, Lille, France, 26–28 April 2006.

[∗] Corresponding author. Tel.: +48 56 6114308; fax: +48 56 6114837. *E-mail address:* bbusz@chem.uni.torun.pl (B. Buszewski).

^{1570-0232/\$ –} see front matter © 2006 Elsevier B.V. All rights reserved. doi[:10.1016/j.jchromb.2006.10.048](dx.doi.org/10.1016/j.jchromb.2006.10.048)

the properties of stationary phases such as column efficiency, hydrophobicity, silanol activity, and ion exchange capacity [\[4–9\].](#page-6-0)

Due to specific structures and properties of peptides, their chromatographic analysis may provide an interesting information on classification of columns or allows for the selection of proper column for peptide investigations. Chromatographic behavior of peptides is determined by the character of their side chains and substituent groups, which define their basic or acidic character (presence of ionizable groups), or the degree of hydrophobicity or hydrophilicity. Moreover, secondary structures of peptides with more than 15 amino acids in the sequence begin to play an important role during the chromatographic elution. Furthermore, all stationary phases exhibit more than one type of interaction with the given solutes and the mobile phase. Similar situation is observed in case of different peptides—there are many possible interactions, which can take part in retention.

Grouping of data (objects or variables) is possible by means of unsupervised methods [\[10,11\].](#page-6-0) Unsupervised methods identify natural clustering pattern and group objects (or variables) on the basis of similarities between the samples. In these cases, no supervisor is needed in the sense of known membership of objects to classes and the classes are not defined. There are two most common methods of partitioning, namely cluster analysis (CA) and principal component analysis (PCA) with factor analysis (FA). These methods are widely recognized as very powerful tools for getting better information about the relations within dataset [\[12\].](#page-6-0)

In recent years multivariate techniques (principal component analysis, factor analysis, cluster analysis, etc.) have also been utilized to evaluate reversed-phase high-performance liq-

uid chromatographic data [\[13–23\]. T](#page-6-0)he main goal of these investigations was to classify and select proper stationary phases, recommend preferred solvents to a given separation, and identify selectivity measures, etc. Principal component analysis and cluster analysis have frequently been used in chromatography to extract maximum information from retention data matrices of many dimensions[\[24–26\]. F](#page-7-0)elinger et al. [\[27\]](#page-7-0) used the principal component analysis to determine the ultimate factors that influence the column-to-column and batch-to-batch reproducibility of retention times and retention factors measured on five commercially available columns. He concluded that when columns from different batches are subjected to PCA, the same-batch columns form a subset or a cluster on the score plot, while the points corresponding to columns of different batches are scattered, demonstrating that there are significant differences between the batch-to-batch and the column-to-column reproducibility of retention times.

Both methods have different approaches to analyze the correlation of a dataset, but it's often observed that the results of CA and PCA are similar [\[24,28\]. T](#page-7-0)he most common application of CA and PCA in chromatographic result analysis is to find the similarities between HPLC columns and as a consequence, columns assigned for the sample group are studied.

In the present study 11 HPLC columns were divided into several groups according to the retention factors of 12 peptides. The columns utilized in the investigations contain various interaction sites. On the other hand, most of the used peptides have the same amino acid skeleton, and they differ between each other with one or two amino acids. Retention data obtained from the HPLC analysis of those peptides were used for the column comparison.

^a Values calculated with the use of HyperChem program.

2. Experimental

2.1. Materials and reagents

Standards of peptides used in the study were obtained from E. Merck (Darmstadt, Germany) and also from Professor Kaliszan laboratory (Medical University of Gdańsk, Marii Skłodowskiej-Curie 3a St., Gdańsk, Poland). Amino acids' sequence of analyzed peptides as well as their molecular masses and some of QSAR descriptors are given in [Table 1.](#page-1-0) Concentrations of substances were about 0.5 mM in water with addition of 0.1% of trifluoroacetic acid (TFA) solution.

For the preparation of mobile phases, acetonitryle of "for HPLC" purity (Lab–Scan, Dublin, Ireland), deionized water from Milli-Q system (Millipore, El Passo, TX, USA), and trifluoroacetic acid (Fluka, Busch, Switzerland) were used.

In the current studies, series of commercially available and home-made packing materials with different surface functional groups such as cholesterol (SG-CHOL), *n*acylamide (SG-CHOL, SG-AP), aminopropyl (SG-CHOL, SG-AP), cyanopropyl (SG-CN), phenyl (SG-Ph), naphthalene (SG-Ar), butyl (SG-C4), octyl (SG-C8), octadecyl (SG-C18), triacontyl (SG-C30), and residual silanols localized on the silica gel surface have been utilized. Properties of all the columns used in the study are listed in Table 2. For the chemical modification of the three columns silica surface (SG-Ph, SG-Ar, SG-CHOL), the following reagents were used: γ -aminopropyltriethoxysilane and triethylamine (Fulka, Buchs, Switzerland); cholesteryl chloroformate 98%, lauric acid chloride (Sigma–Aldrich, Gillingham, Dorset, UK); phenylpropyldimethylchlorosilane (Wacker GmbH, Munich, Germany); magnesium, 3-bromo-1-propene, allyl naphthalene, allyl benzene, bromonaphtalene (Aldrich Chemie, Steinheim, Germany). Toluene, hexane, methanol (POCh, Gliwice, Poland), tetrahydrofurane, acetonitryle, 2 propanol, petroleum ether, dichloromethane (J. T. Baker, Łódź, Poland) and morpholine (Reachim, Moscow, Russia) were also used in the studies.

2.2. Bonded phase synthesis and column packing

The reaction mechanism and the conditions for cholesterolic, phenyl and aryl stationary phases synthesis were described earlier: SG-CHOL [\[29\],](#page-7-0) SG-Ph [\[30\],](#page-7-0) SG-Ar [\[30\]. T](#page-7-0)he packing materials were prepared on the basis of silica gels: Kromasil® (Eka Nobel, Sweden) (SG-CHOL), Sirpearl 40 (SG-Ph, SG-Ar) (Sklo Union, Votice, Czech Republic). The received stationary phases were packed into $250 \text{ mm} \times 4.6 \text{ mm}$ I.D. (SG-CHOL) and $125 \text{ mm} \times 4.6 \text{ mm}$ I.D. (SG-Ph, SG-Ar) stainless-steel tubes using homemade apparatus, equipped with Haskel packing pump (Burbank, CA, USA), under constant pressure. As a packing pressurizing solvent, methanol has been used.

Structures of the stationary phases used in the investigations are presented in [Fig. 1.](#page-3-0)

2.3. Apparatus and chromatographic conditions

Two HP 1050 (Agilent Technologies, California, USA) high-performance liquid chromatography systems equipped

Characteristics of columns used in the investigations

*P*_C is the carbon content;

*P*_N is the nitrogen content.

Fig. 1. Stationary phases schematic structures modeled with HyperChem program

with a UV–vis detector (Agilent Technologies, California, USA) were selected for chromatographic measurements. One of them was equipped with autosampler and the other one with manual injection valve with $10-\mu l$ loop. Chem-Station program was used for the data collection (Agilent Technologies).

Elution was carried out with gradient mode conditions of water with the addition of 0.12% of TFA and acetonitryle with 0.10% TFA. The gradient was established as linear from 0% (v/v) to 60% (v/v) of acetonitryle in 20 min. The 'dead time' (t_0) of each column was measured by injecting acetonitryle into the system. The flow rate was 1 ml/min (except of carbon column −0.1 ml/min). All chromatographic measurements were carried out in constant room temperature (20 \degree C). The retention time was measured in peak maximum and peak asymmetry was between 0.9 and 1.2.

The obtained results have been evaluated by Statistica for Windows v. 7.1 (StatSoft, Tulsa, USA). HyperChem v.5.1 package with the ChemPlus extension (HyperCube, Waterloo, Canada) and was used in peptides geometry modeling.

3. Results and discussion

The main object of the study was the grouping of analytical columns on the basis of the retention data of peptides. [Table 3](#page-4-0) presents the retention coefficients for peptides on the selected columns. Obtained data allow to classify stationary phases on the basis of the results of CA and PCA analyzes.

3.1. Column evaluation

[Table 3](#page-4-0) presents retention factor *k*-values data used for the column comparison, which was done on the basis of all collected values. Those values were used for statistical evaluation. The Ward's method for data agglomeration and Euclidean distances for similarity measurements have been applied in the cluster analysis. The dendrogram of all the columns is presented in Fig. 2A. The grouping of different columns can clearly be distinguished. The dendrogram shows that all stationary phases may generally be grouped into several clusters. The most similar are SG-C18 and SG-C30 columns. It is probably connected with the use of long alkyl chains for the modification of the silica gel surface, which give strong hydrophobic character of stationary phase. The next cluster has been created with SG-C4 and SG-C8

Fig. 2. Columns comparison (A) relative Euclidean distances obtained by cluster analysis; (B) PCA score plot.

Retention factors values (k) obtained during the chromatographic analysis (for conditions see in text), where: t_0 , 'dead time' of the column, t_R , retention time

*Retention time longer than analysis time.

Table 3

columns. In this case, short alkyl chains are chemically bonded to the stationary phases, which reflect with less hydrophobic character of the surface and, furthermore, easy accessible silanol groups for interactions with the analyzed substances. SG-CHOL and SG-AP columns create the next conglomeration. In this case, organic groups which possess hydrophobic–hydrophilic character are located on the surface of the modified stationary phases. Both of them also have pseudo-membrane properties. The last cluster that can be clearly seen, is formed by two hydrophilic columns: SG-CN and HILIC. SG-Ph and carbon columns are very dissimilar from the rest.

The same dataset was used to perform principal component analysis. According to the eigenvalue-one criterion only the principal components (PCs) with eigenvalues greater than 1 are considered as the important ones. This criterion is based on the fact that the average eigenvalues of the autoscaled data is just 1. The scree-plot shows that only two factors have fulfilled this criterion. The cumulative explained variance for those PCs was equal to 94.6%. [Fig. 2B](#page-3-0) presents the score plot of all columns in the space of the first two components. Principal component loadings are presented in [Fig. 2B](#page-3-0) and correspond to the correlation coefficient of the particular variable. This figure indicates one group of column. The rest of the stationary phases: SG-Ar, SG-Ph and carbon cannot be assigned into this group. Those packing materials appear to be different from the rest of the columns (similar as in case of CA results). High loading values of first component were found for the following peptides: 2p, 7p, 64p, 65p, 69p, 83p, 88p while for the second one there were: 3d, 16p, 19p, 68p, 70p. The structures of the mentioned compounds were modeled in HyperChem at vacuum conditions ([Fig. 3\).](#page-5-0)

[Fig. 2B](#page-3-0) also indicates a almost linear relationship between classified HPLC columns in the space of new PCs. Enlarged view of this group has been presented in [Fig. 4.](#page-5-0) This linear correlation concerns seven stationary phases used in the study: SG-C30, SG-C18, SG-C8, SG-C4, SG-CHOL, SG-AP and HILIC. It is connected with the meaning of the PC's reflected in their loading values (Table 4). It may indicate properties of peptides, which decide about the affiliation of analyzed substances. No relation has been found between molecular mass of peptides along both components. However, another relationship has been observed. Log *P* values calculated by HyperChem program received positive values for peptides, which characterize first component, and negative in case of the second one. It can be concluded that first component describe non-polar properties of peptides

Table 4 Principal components loads calculated by Statistica program

	Component 1	Component 2
2p	-0.902429	-0.110173
7p	-0.973796	-0.197347
16p	0.430977	0.795363
19 _p	-0.170858	0.963020
64 _p	-0.995359	0.000994
65p	-0.993466	0.009125
68 _p	-0.160541	0.979691
69p	-0.995603	-0.027829
70 _p	-0.067773	0.973092
83p	-0.980514	0.014290
88p	-0.976010	-0.010894
3d	-0.281477	0.926840

Fig. 3. Structures of peptides, which have high loading values of the first (A) and second (B) principal component.

Fig. 4. Enlarged view of PCA score plot presented in the [Fig. 2B](#page-3-0).

that are able to interact through hydrophobic interactions (e.g. van der Waals) during the chromatographic analysis, whereas second component is loaded with polar peptides with much lower log *P* values. It can explain the linear relation between the above mentioned columns. Retention of hydrophobic compounds is increasing (negative values of first component loads) with the increase in hydrophobic character of stationary phase (from SG-C30 to SG-C4), while the retention of polar peptides on hydrophobic packing materials is decreasing (positive values of second component loads). The presence of SG-CHOL, SG-AP and HILIC stationary phases in this linear dependence results from their intermediate polarity and can be used in the future to predict and conclude their about hydrophobicity.

3.2. Peptides grouping

Retention factors obtained for the SG-C4, SG-C8, SG-C18, SG-C30, SG-AP, SG-CHOL, HILIC, and SG-Ph columns were

Fig. 5. Peptides grouping: relative Euclidean distances obtained by cluster analysis.

used to group peptides according to their specific chromatographic behavior. For this purpose, authors worked with less data than for column partition. Results of chromatographic analysis of peptides on SG-Ar and carbon stationary phases were omitted because of lack of data. Some peptides were retained in those columns for more than 20 min of analysis, which was the main reason of data shortage. In case of obtaining more reliable results, two chromatographic columns, for which lack of data was significant, were not considered. The second run of cluster analysis was performed to group peptides. All parameters used in CA analysis were the same as in case of column partition. The received dendrogram is presented in Fig. 5 and it allows to discriminate peptides into two main clusters. Among those two groups, some of the peptides create smaller clusters with bigger similarity. Such effect is a consequence of resemblances in peptides retention. Some of smaller clusters are formed by peptides with similar molecular mass (64p and 65p; 18p and 4p) ([Table 1\);](#page-1-0) however, few groups are created by compounds with completely different masses (21p and 62p; 69p and 70p). Several analyzed compounds are eluted from chromatographic column at the same time, although they have different size and therefore simple peptides grouping, according to the relationship between retention and molecular mass, cannot be performed. Such effect is probably the consequence of many other factors influencing the retention, e.g. changes in the mobile phase composition during the gradient elution, which can reflect in both stationary phase and peptide conformations of variations. Peptides division in two clusters is probably connected with their specific structures, being created inside the chromatographic column. Complete explanation of peptides grouping still remains unclear. We could not find any straight correlation between compounds assigned to the first or second cluster. For that purpose, we have started a new experiment set.

4. Conclusions

The use of cluster and principal component analyses leads to grouping chromatographic columns on the basis of retention

data. Cluster analysis allowed for dividing the columns into several groups, where the most similar were SG-C18 and SG-C30, while the most different were HILIC and carbon columns. PCA provided another interesting result. There is almost a linear relationship between classified HPLC columns in the space of new PCs, which is connected with meaning of the PC's reflected in their loading values. The first component describes non-polar properties of peptides, whereas the second component is loaded with polar peptides with much lower log *P* values. During every chromatographic process peptides retention depends on specific interactions between stationary phase, mobile phase and the investigated compound. Therefore, as observed in the present study, effects of peptides properties on retention will always play an important role in HPLC process. It can be concluded that the specific structure of peptides, which determines their properties, is interesting in their application for column grouping. PCA and CA were also used in peptides' comparison, however there is no clear dependence between retention of peptides and their molecular mass, because of the changes in peptides conformations.

Acknowledgements

The authors gratefully acknowledge Prof. Roman Kaliszan and Dr. Tomasz Baczek (Medical University of Gdańsk) for peptides donation.

References

- [1] C.T. Mant, R.S. Hogdes, High Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation, CRC Press, Boca Raton, FL, 1991.
- [2] K. Štulik, V. Pacáková, J. Suchánková, H.A. Claessens, Anal. Chim. Acta 352 (1997) 1.
- [3] B. Buszewski, M. Jezierska, D. Berek, J. High Resolut. Chromatogr. 21 (1998) 267.
- [4] L.C. Sander, S.A. Wise, Anal. Chem. 56 (1984) 504.
- [5] T. Daltrup, B. Kardel, Chromatographia 18 (1984) 81.
- [6] R.L. Gilpin, J. Chromatogr. Sci. 22 (1984) 371.
- [7] K. Krupczyńska, B. Buszewski, P. Jandera, Anal. Chem. 76 (2004) 227A.
- [8] P.F. Frick, J. Fekete, K. Héberger, Anal. Chim. Acta 536 (2005) 71.
- [9] K. Krupczyńska, P. Jandera, B. Buszewski, Anal. Chim. Acta 540 (2005) 127.
- [10] J.W. Einax, H.W. Zwanziger, S. Geiss, Chemometrics in Environmental Analysis, Wiley-VCH, Weinheim, 1997.
- [11] J. Mazerski, Podstawy chemmometrii, Politechnika Gdańsk, Gdańsk (in Polish) 1997.
- [12] J.W. Einax, D. Truckenbrodt, O. Kampe, Microchem. J. 58 (1998) 315.
- [13] R.G. Brereton, D.V. McCalley, Analyst 123 (1998) 1175.
- [14] R.J.M. Vervoort, M.W.J. Derksen, A.J.J. Debets, J. Chromatogr. A 765 (1997) 157.
- [15] M. Euerby, P. Peterson, J. Chromatogr. A 994 (2003) 13.
- [16] T. Iványi, Y.V. Heyden, D. Visky, P. Baten, J. De Beer, I. Lázár, D.L. Massart, E. Roets, J. Hoogmartens, J. Chromatogr. A 954 (2002) 99.
- [17] D. Visky, Y.V. Heyden, T. Iványi, P. Baten, J. De Beer, Zs. Kovács, B. Noszal, P. Dehouck, E. Roets, D.L. Massart, J. Hoogmartens, J. Chro- ´ matogr. A 1012 (2003) 11.
- [18] T. Cserháti, Anal. Chim. Acta 296 (1994) 235.
- [19] B. Walczak, L. Morin-Allory, M. Lafosse, M. Dreux, J.R. Chrétien, J. Chromatogr. 395 (1987) 183.
- [20] K. Kimata, K. Iwaguchi, S. Onishi, K. Jinno, R. Eksteen, K. Hosoya, M. Araki, N. Tanaka, J. Chromatogr. Sci. 27 (1989) 721.
- [21] S.D. Rogers, J.G. Dorsey, J. Chromatogr. A 892 (2000) 57.
- [22] S. Ounnar, M. Righezza, J. Liq. Chromatogr. Relat. Technol. 22 (1999) 2575.
- [23] M.E. Pate, M.K. Turner, N.F. Thornhill, N.J. Titchener-Hooker, Bioprocess Eng. 21 (1999) 261.
- [24] H. Schneeweiss, H. Mathes, J. Multivariate Anal. 55 (1995) 105.
- [25] Z. Kánya, E. Forgács, T. Cserháti, Z. Illes, Chromatographia 63 (2006) 129.
- [26] K.V. Mardia, J.T. Kent, J.M. Bibby, Multivariate Analysis, Academic Press, London, New York, 1979.
- [27] A. Felinger, M. Kele, G. Guiochon, J. Chromatogr. A 913 (2001) 23.
- [28] P. Forlay-Frick, E. Van Gyseghem, K. Berger, Y. Vander Heyden, Anal. Chim. Acta 539 (2005) 1.
- [29] B. Buszewski, M. Jezierska-Switała, R. Kaliszan, A. Wojtczak, K. Albert, ´ S. Bochmann, M.T. Matyska, J.J. Pesek, Chromatographia 53 (2001) 204.
- [30] R. Gadzała-Kopciuch, M. Kluska, M. Wełniak, B. Buszewski, Mater. Chem. Phys. 89 (2005) 228.