Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1062 (2005) 79–86
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Evaluation of novel sample identification approach based on chromatographic fingerprint set correlation homogeneity analysis

Željko Debeljak[a,1], Goran Srečnik[a], Tomislav Madić[a], Marinko Petrović[a], Natalija Knežević[a], Marica Medić-Šarić[b,*]

[a] *Analytical Development, R&D, PLIVA d.d., Prilaz baruna Filipovića 25, 10000 Zagreb, Croatia*
[b] *Department of Pharmaceutical Chemistry, Faculty of Pharmacy and Biochemistry, A. Kovačića 1, 10000 Zagreb, Croatia*

Received 2 July 2004; received in revised form 2 November 2004; accepted 3 November 2004

Abstract

Instead of usual rationale for chromatographic fingerprint based sample identification which relies upon visual inspection or principal component analysis of raw or aligned chromatograms novel nonparametric statistical measure of fingerprint set homogeneity is proposed. Randomization test is applied for significance analysis of fingerprint set homogeneity while average maximum crosscorrelation is used as a merit function. Chromatogram sets generated by random selection from standard and unknown sample chromatogram collections are compared with respect to merit function values with set of chromatograms that represents standard and/or unknown sample. In that instance fingerprint homogeneity significance is represented by the fraction of random chromatogram sets that have higher merit values than the standard and/or unknown sample sets. A set of peptide maps corresponding to different haemoglobin variants has been selected for evaluation of proposed test. This approach is compared to chromatogram alignment based on correlation optimized warping coupled with principal component or cluster analysis. Proposed method is simple i.e. straightforward sample identification procedure which reliability has been evaluated here. Impact of this approach on peptide mapping validation and system suitability analysis is discussed.

Keywords: Sample identification; Chromatographic fingerprint; Correlation homogeneity; Randomization test; Peptide map

1. Introduction

Visual inspection of chromatographic fingerprint patterns has been used for decades as a sample identification procedure [1,2]. Moreover, regulatory authorities recognized this type of analysis as a valid procedure for identification of protein samples [3,4]. But there is still a problem of visual comparison of complex patterns that is prone to subjective decision-making. Visual inspection of fingerprints encounters problems caused by time shifts, variable peak number and corresponding signal intensities in chromatographic fingerprint [1,2]. For example, an attempt to completely control chromatographic variables that could cause misidentification of fingerprints failed due to instrument-to-instrument variability in case of anti-CD4 monoclonal antibody peptide map based sample identification [5]. As a consequence authors suggested one of the alignment procedures to obtain reproducible retention times. On the other hand the "major peak" scheme on which visual inspection relies upon is not quite reliable, also. It has been shown that 26 peaks are assigned as "major peaks" of human haemoglobin A peptide map although only 23 peptide fragments could be present in sample solution [6]! Authors suggested that this difference could be caused by trypsin autodigestion or by existence of haemoglobin degradation products. This type of reasoning is strongly dependant on chromatographic integration events, thresholds in particular and sample preparation or instrumental conditions. Therefore obtained conclusions are questionable. Review of numerous causes of deviations from expected

* Corresponding author. Tel.: +385 1 4818 304; fax: +385 1 4856 201.
E-mail address: bebamms@pharma.hr (M. Medić-Šarić).
[1] Present address: Department of Medicinal Biochemistry, Clinical Hospital Osijek, J. Huttlera 4, 31000 Osijek, Croatia.

0021-9673/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2004.11.012

number of peaks in peptide map analysis is given in a paper written by Malmquist [7].

Instead of raw data comparison many authors attempted to develop reproducible alignment procedures and to compare chromatogram transformants [2,7–10]. These attempts showed that in order to obtain maximum fingerprint correspondence one needs piecewise procedure. This means that retention time shifts are not linear [2,9] and, depending on mobile phase gradients one sometime needs two or more internal standards to compensate for different retention shifts for every individual time "piece" if one wants to solve this problem by experimental approach [11]. Still, theoretical approach based on chromatographic alignment procedures proved to be useful [7–10]. One of the first successful attempts to develop quantitative measure of peptide fingerprint concordance was published by Malmquist [7]. Although the intention was to develop fast and automatic procedure for determination of fingerprint concordance proposed procedure proved to be quite complex. First of all, it requires quite large number of training samples. Besides that, it is based on complex chromatogram alignment preprocessing after which jackknife procedure should be applied on simulated chromatograms for determination of number of important principal components. Finally, principal component analysis (PCA) derivative named soft independent modeling of class analogy (SIMCA) should be applied in order to make a decision whether sample of interest belongs to a certain category or not. Application of maximum alignment principle and chromatogram simulations resulted in 10% cut-off needed for reduction of false positive identification [7]. Since usual cut-off is 5% or lower this result indicates that the methodology could have problems with differentiation of highly similar but still nonidentical sample classes i.e. false positive identification. The first problem here is possibility of chromatogram overfitting. The second problem is how to classify simulated chromatograms—to the original or to some other sample class? Peaks generated by the simulation does not necessary transform starting sample chromatogram to a new or different category since the number of peptide map peaks in the real world experimental settings could be variable [7]. Similar problems are common to new alignment approaches [9,10]. Besides alignment based approaches some even more complex alternatives were proposed [11,12] which description falls outside the scope of this article.

All these findings lead to the conclusion that the major obstacle for quantitative fingerprint based sample identification is variability due to instrumental and/or sample preparation conditions i.e. variability of retention times and/or number of peaks in chromatographic fingerprints. Unfortunately this type of dual variability is quite a complex problem to handle. Namely, when at least one of the chromatographic axes is not tightly controlled it is not possible to use some of the sample classification methods. Strategy used in alignment approaches is based on sequential approach. The first step is alignment of chromatograms which more or less solves retention time axis variability problem. But alignment procedure leads to possible chromatogram overfitting problems [7]. Remaining variability should be addressed to peak number or peak intensity variations and sample identification could be provided by some classification method.

Another approach to dual variability has been considered in this article. Dual variability problem is reflected in raw fingerprint correlations which consequently vary more or less among different pairs of samples. Correlation variability is composed of random component and systematic differences among samples if they exist. In order to extract systematic correlation differences randomization[2] test has been considered [13–15]. The rationale for such test relies upon the fact that in case of two or more non-identical samples sample variability measured among randomly chosen chromatograms could not result in higher average pairwise correlation coefficient than within-group correlation except by the chance i.e. random differences. If random retention time shifts and/or peak number variability exist it will have approximately the same effect on correlation variability of both groups. Therefore, it would not compromise the final result. Still, randomization imposes request for relatively large number of chromatograms, but it solves the dual variability problem. To construct described test one needs to calculate average pairwise correlation between chromatograms corresponding to unknown sample and to compare it to average pairwise correlations between randomly combined chromatograms from complete set of chromatograms. The complete set of chromatograms should contain all chromatograms corresponding to both, unknown sample and standard. Fraction of between-group correlations that are higher than within-group correlation determines whether all analyzed samples have the same origin or not. This fraction represents significance of the test and it also represents the statistical basis for sample identification.

Described method measures complete sample set homogeneity based on correlation. Authors considered ANOVA approach and Hotelling's test [13] as well. ANOVA was not analyzed in details due to the fact that average pairwise correlation coefficients are interdependent variables. Moreover, type of average pairwise correlation distribution is generally not known while usage of ANOVA presumes normal variable distribution. Randomization test avoids these obstacles. Hotelling's test was preliminary considered but it has been proven that its strength is too low. Almost all comparisons resulted in false positive sample identifications.

Except the average pairwise correlation Kendall's concordance coefficient [16] could be used as merit function. This approach avoids analysis of separate pairs of chromatograms but CPU requirements are considerably higher in this case. Therefore this possibility wasn't analyzed in details.

The major goal of this paper is to provide quantitative chromatographic fingerprint based sample identification proce-

[2] Besides the term "randomization test" similar terms like "bootstrap test", "resampling test", "Monte Carlo test" and "rerandomization test" could be found in referenced literature.

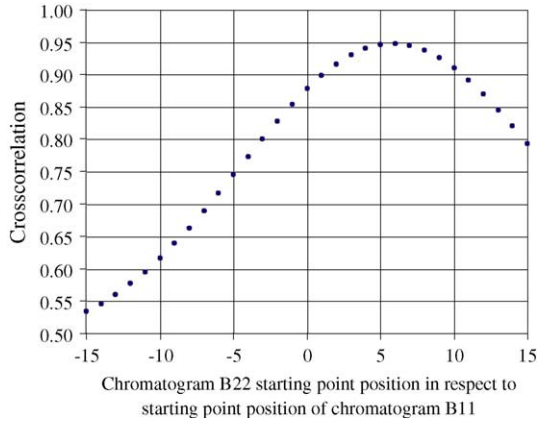

Fig. 1. Selection of maximum crosscorrelation. Starting point of B12 chromatogram is shifted from −15 to 15 points in respect to starting point of B11 chromatogram in order to find best matching pair of starting positions in terms of crosscorrelation.

dure that is suitable for routine analysis and analytical method validation. Proposed unsupervised classification method fits these requirements. Therefore it has been selected for experimental reliability evaluation.

2. Theory

In order to avoid chromatogram disconcordance caused by small differences in the starting point of different chromatograms due to less than a perfect data collection[3] instead of average pairwise Pearson correlation coefficient average pairwise maximum crosscorrelation coefficient <*r*> has been selected as a merit function for chromatographic fingerprint set comparison [17]. Rationale for selection of maximum crosscorrelation value is graphically presented in Fig. 1.

It is visible that Pearson correlation significantly changes its value for slightly different selection of starting points of analyzed chromatogram pair (<±20 data points shift or <±5 s shift). Pearson correlation coefficient calculated for chromatogram pairs that differ only in starting points selection is known as crosscorrelation coefficient [17]. Therefore, instead of a simple Pearson correlation coefficient between two chromatograms maximal crosscorrelation is chosen for merit function development. Average pairwise maximum crosscorrelation <*r*> has been selected for comparison of chromatogram collections that represent a real world or random sample. Starting point adjustment window width is the only user-defined variable used in this approach. Although such variables could introduce subjective errors this problem could be easily avoided since ten or twenty data point shifts are generally sufficient. Moreover, as the window width increases chances for missing correct <*r*> drop to zero.

If N represents the number of chromatograms per sample number of pairs across which averaging is made is $N(N-1)/2$. Since simple comparison of <*r*> does not include *r*'s variability additional *randomization test* is applied.

If M different samples with N corresponding chromatograms are to be classified $\binom{M \times N}{N}$ chromatogram combinations are possible. For example, in benchmark experiment four chromatograms per sample have been recorded. Therefore, 70 combinations for analysis whether two samples belong to the same class or not are possible. In order to avoid misinterpretation of results all combinations are used for significance calculations. In case of let's say five chromatograms per sample and correlation homogeneity comparison of two samples there are 252 possible chromatogram combinations. In this case or in case of even more replicate chromatograms per sample calculation of all possible <*r*> is not needed. Only a fraction of randomly selected chromatogram pairs (*nt*) and corresponding correlations will satisfy the request for representative sample.

For any combination all pairwise crosscorrelations have to be calculated and corresponding maximum crosscorrelations should be averaged. Finally, decision about two or more samples correspondence is based on the fraction of between-group <*r*> values ($n1$) that are higher than within-group <*r*> value. This rationale could be given in a form of hypotheses:

H0. Analyzed set is homogeneous, $n1/nt > c$,

H1. Analyzed set is not homogeneous, $n1/nt < c$, where c stands for critical fraction value.

Five percent value has been widely accepted as a significance limit for different statistical tests [3,4] and therefore this value has been set here as decision criterion (c). In case of two samples there are two within-group <*r*> values. It is expected that both of them are larger than the vast majority of between-group <*r*> values if samples belong to the different classes (H1) i.e. analyzed set is not homogeneous. If both within-group <*r*> values are lower than considerable fraction of between-group <*r*> values, either H0 hypothesis should be accepted or both samples contain uncorrelated chromatograms. The last situation indicates that sample preparation or instrumental conditions are not controlled i.e. system suitability conditions are not met. This implies mandatory preanalysis of crosscorrelation relative standard deviation (R.S.D.) and subsequent identification of sources of variation in case of low within-group <*r*> values. Finally, it is possible that one within-sample <*r*> satisfies H0 acceptance criterion and the other not. If one sample indicates that H0 is acceptable while the second sample homogeneity test favors H1 hypothesis more chromatograms should be provided. In case of very similar samples both sample chromatogram sets could have very similar within-group and between-group

[3] These small shifts caused by instrumental error as well as small variations of data acquisition interval, especially in case of two dimensional signals (diode array or MS signals) are the major technical obstacles for application of proposed method. In the second case some improvements could be made by increasing the data acquisition interval.

Table 1
HPLC gradient elution conditions used for haemoglobin fingerprint analysis

Time/min	A (%)	B (%)
0.0	97	3
40.0	70	30
85.0	50	50
115.0	30	70

Mobile phase A: 1 mL trifluoroacetic acid + 1000 mL water. Mobile phase B: 1 mL trifluoroacetic acid + 99 mL water + 900 mL acetonitrile.

<*r*> values. Final decision could be made based on increased number of chromatograms per sample or higher *nt* value.

Described algorithm has been implemented in C programming language and tested on personal computer architecture running both, Microsoft Windows™ and Linux operating systems.[4]

3. Experimental

Benchmark experiment included tryptic digestion of haemoglobin samples and HPLC analysis of corresponding digests. All separations were made on Agilent 1100 HPLC instrument equipped with diode array UV/VIS detector and autosampler. 215 nm wavelength was used for randomization test evaluation. 0.8 ml/min flow rate was applied. All separations were made on Vydac 218TP54, 250* 4.6 mm analytical column with 5 μm particles and 300 Å pore size thermostated on 35 °C. Elution gradient conditions are given in the Table 1.

Trifluoroacetic acid and gradient grade acetonitrile were obtained from Merck, Darmstadt, Germany. Doubly distilled 18 MΩ cm^{-1} water was used. Guanidine hydrochloride and dithiothreitol were obtained from Sigma-Aldrich, St. Louis, USA same as human haemoglobin A (H), human haemoglobin S (S), bovine haemoglobin (B1) and TPCK treated trypsin while the second lyophilized bovine haemoglobin (B2) was obtained from Calbiochem, San Diego, USA.[5] Haemoglobin samples were diluted with water to 1.5 mg/ml concentration. Two hundred fifty microliter of this solution was mixed with 16 μL of 1 M phosphate buffer pH 8.0, 28 μL of 1 mg/ml trypsine solution and 36 μL of water. These solutions were stored at 37 °C for 18 h. Digestion was ceased by adding 100 μL of guanidine HCl, 7 μL of dithiothreitol (1.542 g/ml) and by elevating sample temperature to 90 °C for 1 min.

In order to include daily variations and possible degradation effects due to sample freezing all samples were prepared in duplicate and injected twice per day two days in a row.

Only a fraction of freshly prepared trypsin and haemoglobin solutions was used for tryptic digestion while the rest of the solutions was frozen at −20 °C. The second day these solutions were defrosted and used for a new preparation of tryptic digest. Mobile phases were also freshly prepared every day since both of these experimental conditions are possible causes of misidentification of resulting peptide map fingerprints. This way high sample variability has been ensured making these experimental conditions quite complex and therefore suitable for identification procedure benchmarking.

4. Results and discussion

Two preparations of bovine haemoglobin from different manufacturers have been included in benchmark analysis in order to test false negative type of statistical error. To test subtle differences between very similar samples human haemoglobin A and human haemoglobin S were chosen. They differ in only 1 amino acid (6β Glu → Val).[6] These settings enable analysis of false positive errors. To make identification more complex this kind of point mutation does not change the number of tryptic fragments. Unfortunately, this point mutation also causes lower affinity of trypsin towards S substrate compared to affinity for H. Accordingly, 50 μL of digested S sample was injected instead of 25 μL. This way small peaks are not lost and chromatographic fingerprint of S is brought to the comparable intensity scale as the rest fingerprints. Corresponding chromatograms are given in Fig. 2A–D.

When H and S chromatograms are compared only one fragment peak (marked by an arrow) among analyzed set of more than 25 peptide fragment peaks has different retention time due to described point mutation.

To show the complexity of the selected chromatographic fingerprint based identification problem as the first choice for fingerprint analysis PCA and cluster analysis have been chosen. In order to extract information about variation among analyzed chromatograms covariance based PCA and cluster analysis based on application of single linkage rule and Euclidean distances calculation are used. Results are given in Fig. 3A and B.

It is clearly visible that in case of PCA separation of samples of different species is achieved. Dendrogram is less selective since B11 chromatogram is not classified to bovine chromatograms cluster. Since bovine samples are grouped together by PCA approach false negative type of error has been avoided. Still, human haemoglobin chromatograms are completely misclassified by both methods. This example shows that false positive errors are the main obstacle in chromatographic fingerprint based sample identification. These results are consistent with previous comment on significance limit selection from paper [7]. In order to simplify identification problem only H and S samples have been analyzed by PCA. Although this simplification par-

[4] Source codes and all chromatograms are available on request that should be addressed to: zdebelja@inet.hr.

[5] In addition to given symbols all haemoglobin chromatograms are enumerated by numbers 1–4.

[6] Human haemoglobin A consists of 2α and 2β chains which contain 141 and 146 aminoacid residues, respectively.

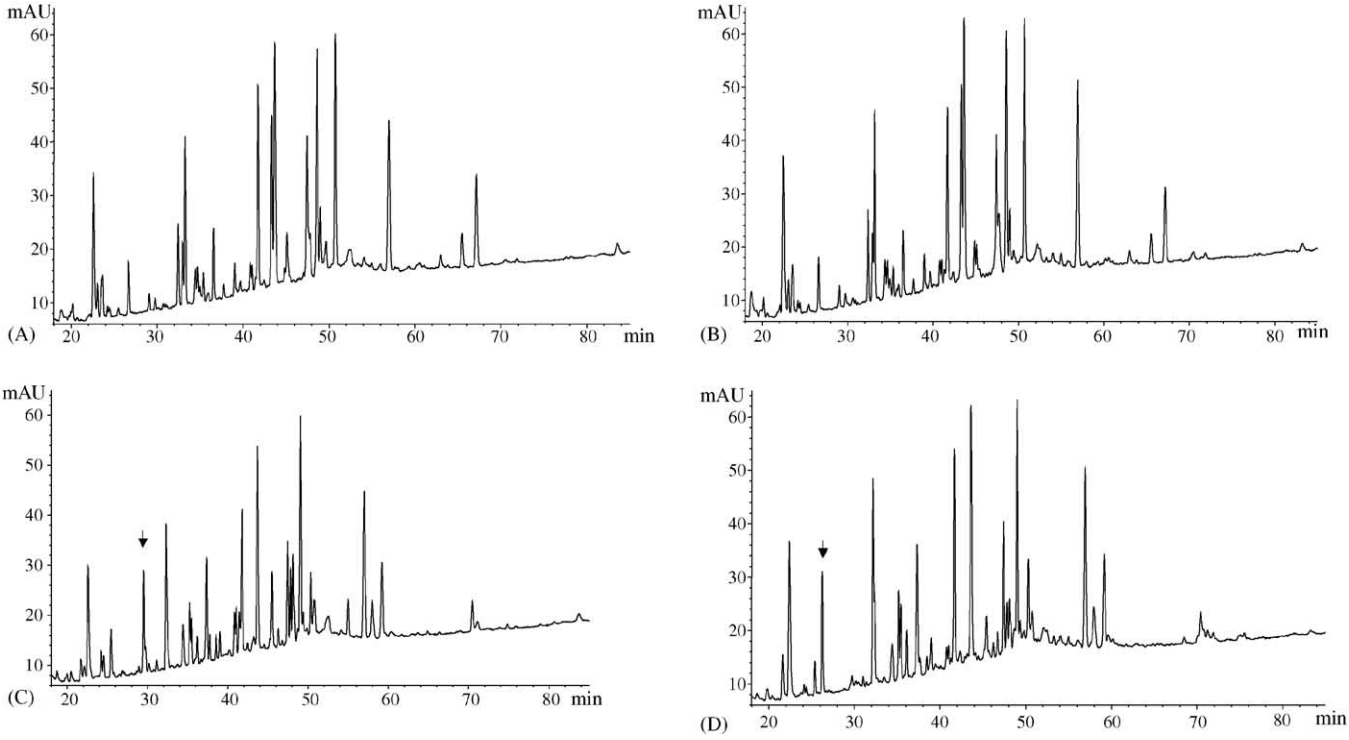

Fig. 2. (A–D) Chromatograms of tryptic digests of B1, B2, S and H samples. (A) B1 tryptic digest, (B) B2 tryptic digest, (C) S tryptic digest, (D) H tryptic digest.

tially improved separation between classes some of S chromatograms were still very closely positioned to some of H chromatograms.

The next step was application of correlation optimized warping (COW) alignment procedure [9]. This approach demands selection of at least three user-defined variables: target chromatogram, segment size and slack parameter. Chromatographic peak width in analyzed chromatograms ranges from 10 to 50 data points. Therefore slack parameter has been varied from 10 to 50 while segment size parameter was set either to 1000 or to 100 points. No significant differences between analyzed settings regarding resultant PCA has been detected in any of the following examples. Percentage of variance explained by the first principal component differs less than 2% between analyzed sets of parameter values. Therefore only PCA corresponding to COW aligned chromatograms based on the segment size and slack variable set to 100 and 10 points, respectively is presented. The same selection of parameter values is used for cluster analysis. Selection of proper limits of variable space that should be searched for the best COW alignment makes this type of analysis prone to subjective decision making while variable space search itself is quite time consuming.

As the first target for COW alignment B11 chromatogram has been selected. Cluster analysis and PCA were applied on aligned chromatograms and results are given by Fig. 4A and B.

The improvement is obvious. All bovine samples are classified correctly by both methods but the S–H classification problem remains. H3 and H4 have been set closer to S cluster instead of H1, H2 pair in dendrogram. PCA resulted in misclassification of H3 and S1 chromatograms.

To simplify the problem only S–H pair was analyzed by PCA. This attempt resulted in two convex classes (Fig. 5). S1 chromatogram is still positioned near H class. Nevertheless, it could be concluded that satisfactory classification has been achieved i.e. false positive sample identification problem has been resolved.

To test the influence of different target chromatogram selection on COW/PCA and COW/dendrogram based sample identification performance experiment has been repeated. Only H–S pair classification was evaluated while S1 has been selected for target chromatogram. These experimental settings represent typical pharmaceutical or biomedical sample identification problem [5,6]. Results are given by Fig. 6A and B.

Although human S haemoglobin was used as target chromatogram dendrogram results have not improved. Surprisingly, PCA results deteriorated and possibility of false positive identification emerged again.It seems that target chromatogram selection represents crucial problem. This example confirms the necessity of jackknife procedure and chromatogram simulations proposed by Malmquist if alignment procedure is used for numerical sample classification [7].

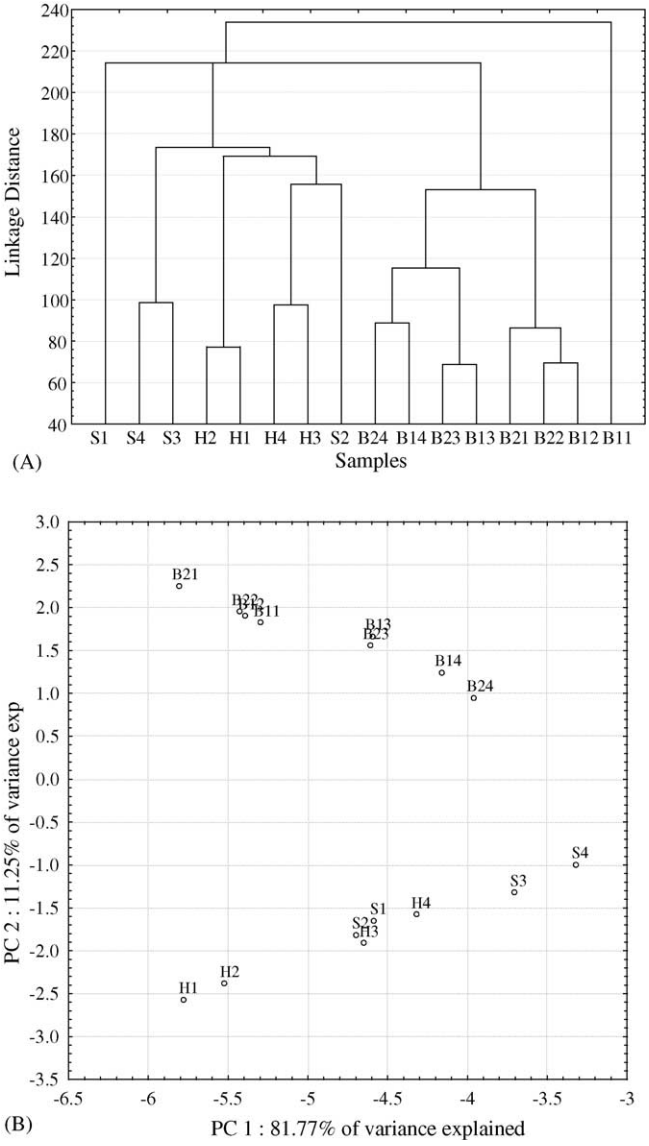

Fig. 3. (A and B) Cluster analysis and PCA of complete set of raw chromatograms. (A) Cluster analysis, (B) PCA.

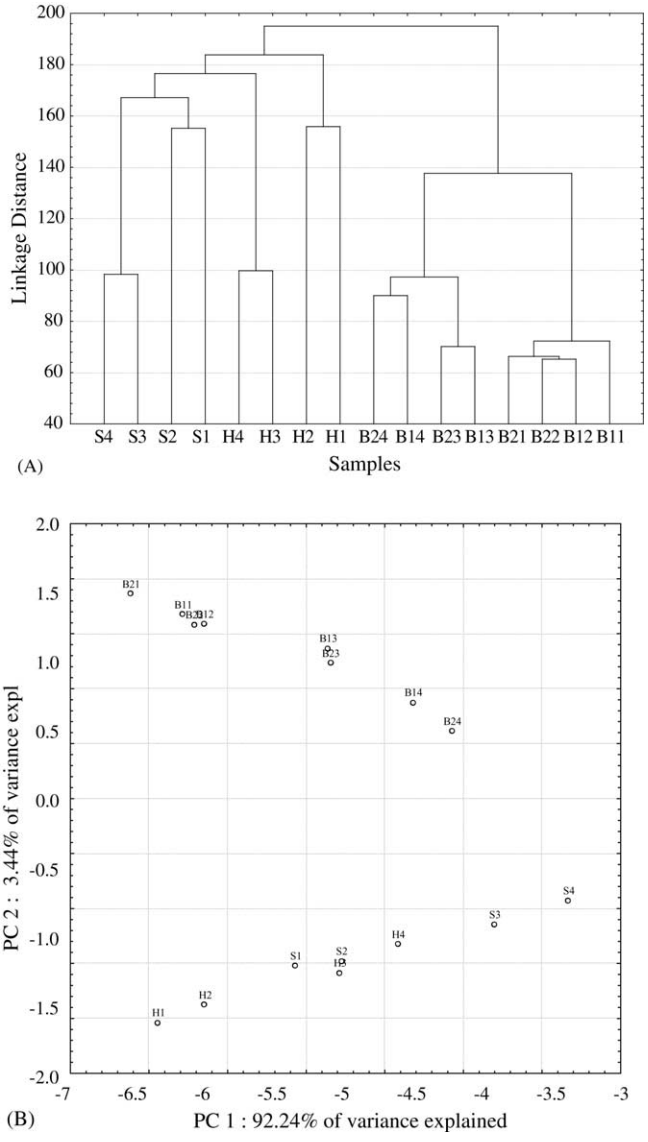

Fig. 4. (A and B) Cluster analysis and PCA of complete set of chromatograms aligned against B11 chromatogram based on COW. (A) Cluster analysis, (B) PCA.

Based on COW/PCA approach it is not clear whether H and S samples belong to the same class or not because selection of S1 as target chromatogram results in alignment overfitting. As a consequence of this false positive identification error inevitably emerges.

Finally, instead of alignment preprocessing all raw chromatograms were analyzed by described crosscorrelation homogeneity test method. Starting point window width was set to ±100. Results of these analyses are given in Table 2.

According to homogeneity significance results all chromatograms have been correctly identified. The same analysis was repeated on first set of chromatograms aligned against B11. Corresponding results are given in Table 3.

Although $<r>$ for bovine samples have increased, homogeneity results are essentially the same. This finding shows that proposed test could be used for quantitative sample identification analysis of aligned chromatograms. Moreover, it correctly classifies and/or identifies samples that have not been classified correctly by PCA and dendrogram approach.

Randomization homogeneity test was applied on partial S–H set of chromatograms aligned against S1. Results are given in Table 4.

These results confirm previous findings. As between and within-group $<r>$ increase due to overfitting probability of false positive error also increases. That is clearly visible in Table 4.

Some authors [5] stressed out the importance of peptide mapping validation. Therefore impact of previous findings on peptide mapping validation is analyzed. Described approach proved to be appropriate for fingerprint identification and

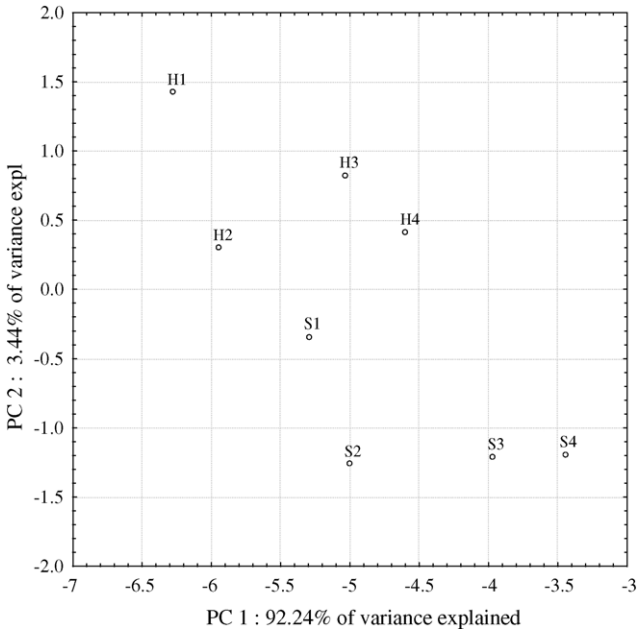

Fig. 5. PCA of S and H samples represented by chromatograms aligned against B11 chromatogram based on COW.

Table 2
Homogeneity of complete set of raw chromatograms

	<r>	R.S.D. (%)		
B1	0.955	0.05		
B2	0.936	0.29		
S	0.949	0.13		
H	0.958	0.10		

Significance (%)

	B1	B2	S	H
B1		40.00	0.00	1.43
B2	71.43		1.43	1.43
S	1.43	0.00		4.29
H	0.00	0.00	1.43	

Significance is calculated based on randomization test. Critical significance value (c) is 5%. Sample pairs that represented by specific cell position form a unique set of chromatograms which homogeneity has been analyzed. Since both samples that form a pair could be compared to randomized sets of chromatograms in respect to corresponding <r> values two significance values have been calculated for each sample pair.

Table 3
Homogeneity of complete set of chromatograms aligned against B11 chromatogram based on COW

	<r>	R.S.D. (%)		
B1	0.971	0.05		
B2	0.943	0.29		
S	0.941	0.17		
H	0.936	0.15		

Significance (%)

	B1	B2	S	H
B1		25.71	0.00	0.00
B2	67.14		0.00	0.00
S	1.43	1.43		0.00
H	1.43	1.43	2.86	

Significance is calculated based on randomization test. Critical significance value (c) is 5%.

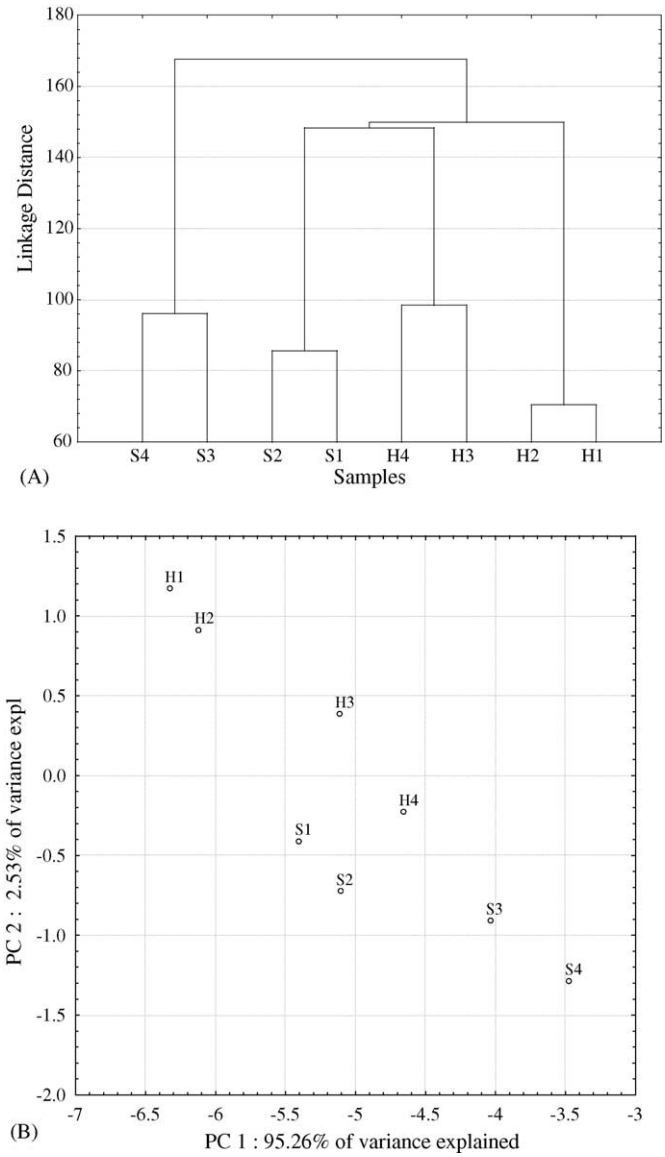

Fig. 6. (A and B) Cluster analysis and PCA of S and H represented by chromatograms aligned against S1 chromatogram based on COW. (A) Cluster analysis, (B) PCA.

Table 4
Homogeneity of a set containing S and H sample chromatograms aligned against S1 based on COW

	<r>	R.S.D. (%)
S	0.960	0.11
H	0.965	0.07

Significance (%)

	S	H
S		7.14
H	0.00	

Significance is calculated based on randomization test. Critical significance value (c) is 5%.

therefore it is suitable for chromatographic selectivity analysis that is a part of any chromatographic method validation. In case of peptide mapping of samples put under temperature and humidity stress conditions it is particularly important to accurately classify samples to the same or different category in comparison to some standard since such samples could contain degradation products which should be identified, or classified to a new class.

Proposed test is suited for sample preparation repeatability. Critical values for $<r>$ obtained this way could be used for system suitability analysis. As any other chromatographic method peptide mapping analysis should be made under controlled analytical conditions. This means that instrumentation and sample preparation should be tested. In completely analogous way to usual injection precision suitability analysis fingerprint concordance could be made. Before any analytical sequence two preparations of standard peptide map preparation are to be analyzed. In case of acceptable standard set $<r>$ and R.S.D. values sequence could continue. Unacceptable results indicate nonuniform sample preparation conditions or unconditioned instruments. This scheme is similar to resolution testing for system suitability purposes. Since chromatographic conditions could change during sequence it would be useful to inject standard samples at the end of analytical sequence to account for possible condition changes.

When sample and standard chromatograms are recorded homogeneity test could be made. In case of acceptable between-group homogeneity sample identity is confirmed. In the opposite case one should consider possible existence of degradation products or related substances in analyzed sample. Proposed procedure represents quantitative way of peptide map identification and it resembles usual system suitability routines.

5. Conclusions

Analyzed randomization test of chromatographic fingerprint crosscorrelation homogeneity proved to be a suitable tool for sample classification or identification purposes. It has been shown to be straightforward and accurate quantitative measure of chromatographic fingerprint based sample identification capable of detecting single amino acid mutation based on a peptide map analysis. Benchmark test has shown that randomization test of homogeneity is at least reliable as COW/PCA approach. It has also shown that chromatographic alignment procedure could lead to misidentification. Randomization test provides numerical measure of set homogeneity. In order to numerically evaluate set homogeneity based on existing alignment procedures introduction of a substantial number of user-defined variables is unavoidable. Because of that fact proposed randomization homogeneity test is less prone towards subjective sample identification in comparison to existing alternatives. Although proposed method doesn't require any alignment procedure it has been shown that analysis of aligned chromatograms is possible and it is as reliable chromatographic fingerprint identification or classification method as PCA. Application of proposed test to analytical method validation and its application to MS fingerprint identification, which is based on analysis of two-dimensional detector signals, are currently under research.

Acknowledgement

Authors would like to express their gratitude to Dr. Dan Bylund from Department of Analytical Chemistry, Uppsala University, Sweden for endowed Matlab® implementation of COW.

References

[1] R.L. Garnick, N.J. Solli, P.A. Papa, Anal. Chem. 60 (1988) 2546.
[2] G. Malmquist, R. Danielsson, J. Chromatogr. A 687 (1994) 71.
[3] European Pharmacopoeia, 4th ed., Council of Europe, 2001.
[4] US 26-NF 21, United States Pharmacopoeial Convention Inc., 2003.
[5] K.R. Lee, J. Bongers, D. Gulati, S. Burman, Drug Dev. Ind. Pharm. 21 (2000) 1045.
[6] G.A. Ross, J. Chromatogr. A 636 (1993) 69.
[7] G. Malmquist, J. Chromatogr. A 687 (1994) 89.
[8] V. Pravdova, B. Walczak, D.L. Massart, Anal. Chim. Acta 456 (2002) 77.
[9] D. Bylund, R. Danielsson, G. Malmquist, K.E. Markides, J. Chromatogr. A 961 (2002) 237.
[10] N.-P.V. Nielsen, J.M. Carstensen, J. Smedsgaard, J. Chromatogr. A 805 (1998) 17.
[11] I.V. Tetko, T.I. Aksenova, A.A. Patiokha, A.E.P. Villa, W.J. Welsh, W.L. Zielinski, D.J. Livingstone, Anal. Chem. 71 (1999) 2431.
[12] N.-P.V. Nielsen, J. Smedsgaard, J.C. Frisvad, Anal. Chem. 71 (1999) 727.
[13] R.P. Brown, Genetica 101 (1997) 67.
[14] E.A. Maharaj, Comput. Stat. Data Anal. 40 (2002) 131.
[15] C.E. Lunneborg, Data Analysis by Resampling: Concepts and Applications, Duxbury Press, Pacific Grove, 2000.
[16] M.G. Kendall, B. Babington Smith, Ann. Math. Stat. 10 (3) (1939) 275.
[17] G.E.P. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, 1976.