Quantitative MALDI-TOFMS Analysis of Amino Acids Applying Soft Modeling Methods

P. MARŠÁLEK, M. FARKOVÁ, and J. HAVEL

Department of Analytical Chemistry, Faculty of Science, Masaryk University, CZ-611 37 Brno
e-mail: havel@chemi.muni.cz

Received 3 July 2001

This study demonstrates the applicability of Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOFMS) in multicomponent analysis of amino acids with various chemometric methods like experimental design, partial least squares (PLS), and artificial neural networks (ANN). Several negative factors influencing reproducibility of intensities were systematically studied and their influence was eliminated. Binary mixtures of L-alanine and L-isoleucine were analyzed and evaluated using PLS and ANN. The results obtained show the possibility to use MALDI-TOFMS in quantitative analysis of amino acids.

Recent development of modern analytical methods extended the spectrum for the characterization of complex samples, while among the most exciting innovations in the last decade is Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOFMS) introduced independently in Germany and Japan [1, 2]. New developments have rapidly increased mass accuracy and resolution of mass spectra. There are several papers [3—11] dealing with the way how to improve reproducibility of MALDI-TOF mass spectra. The basic problem in the quantification of MALDI data is caused by poorly controlled sample preparation, strong influence of dependence of the signal on the laser power, and many other factors influencing complicated processes of ionization in MALDI. Some authors have tried to optimize the sample preparation by the use of co-matrices [12], fast evaporation [13, 14] or the use of nitrocellulose [15]. Nevertheless, sample preparation still remains something of an art and requires special attention [7]. This is a serious disadvantage of MALDI-TOFMS. Analyte concentration can be determined only with special procedures, as intensities are of very limited reproducibility caused mainly by heterogeneity of the semi-crystalline matrix, laser influence fluctuations, the formation of various gas-phase adducts, and other factors. Addition of a known quantity of a compound that is chemically similar to the analyte as an internal standard should solve this problem [6], but complete and reliable quantification in MALDI-TOFMS is still not resolved.

Important requirement for reproducibility is the homogeneous mixing of the matrix with the analyte to achieve approximately the same amount of molecular ions by each laser shot. This can be achieved by optimization of the sample preparation procedure and the data collection protocol [3].

The aim of this work was to investigate further some basic factors of reproducibility and to examine possibilities to apply chemometric methods like experimental design, partial least squares (PLS) with projection to latent variables and "soft" modeling with artificial neural networks (ANN) to reach or to improve quantification of multicomponent mixtures by MALDI-TOFMS technique.

THEORETICAL

PLS is the name for a class of methods used for relating blocks of variables measured on sets of objects. This multivariate regression method provides an overview of large data sets [16] and it is gaining increasing importance in many fields of chemistry: analytical, physical, clinical, and industrial process and product control [17—19].

Artificial neural network is capable of modeling extremely complex functions. There are several architectures, one of the most common is the feed-forward neural network of multilayer perceptrons. ANN consist of interconnected network of elementary units, called neurons, which are arranged in layers and operate in parallel. There are three types of neurons contained in a network: input neurons – which accept the input data characterizing each observation, output neurons – which provide the predicted value or pattern, and hidden neurons – which neither receive inputs directly nor provide output values directly. Neurons are connected by weights, which are modified in the course of

network operation. Each neuron has an activity level that is determined by an input signal received from the other units in the network. Training data for this network contain examples of inputs together with the corresponding outputs, and the network learns to infer a relationship between the two [20, 21].

The goal of the use of ANN is to find relationships between input and output data, which then consist of two steps – "learning" and "prediction". In the "learning" phase the optimum structure, weight coefficients, and biases are searched for and used then in the next step for the prediction.

EXPERIMENTAL

L-Serine, L-alanine, L-proline, and L-isoleucine were purchased from Merck (Darmstadt, Germany), 5-chlorosalicylic acid, 5-methoxysalicylic acid, 5-hydroxysalicylic acid, α-cyano-4-hydroxycinnamic acid, and 3,4-dihydroxycinnamic acid were from Sigma-Aldrich (Prague, Czech Republic). Bidistilled water used through the work was obtained by a quartz distillation still from Heraeus (Hanau, Germany).

Mass spectra were acquired using Kratos Kompact III (Manchester, GB) MALDI-TOF mass spectrometer with Kratos Kompact software V5.2.0. Laser operated at 337 nm, maximal pulse energy was 250 μJ with pulse duration 10 ns.

The data in this work were processed on a Pentium PC computer using the program PLS5050 [22] for MS-DOS system and Trajan 3.0 software package [23].

Experimental design is a planned interference in the natural order of events by the researcher [24—26]. There are several types of ED and selection of a specific type of design depends primarily on both the nature and the extent of the information we want to obtain.

RESULTS AND DISCUSSION

The matrices tested were: 5-chlorosalicylic acid, 5-methoxysalicylic acid, 5-hydroxysalicylic acid, α-cyano-4-hydroxycinnamic acid, and 3,4-dihydroxycinnamic acid. The most optimal 5-methoxysalicylic acid was chosen as the matrix for the later studies because of its good ability to ionize amino acids and good stability under vacuum.

Methanol, acetone, and acetonitrile were tested as matrix solvents and acetonitrile was found the most suitable.

Two methods for the sample and matrix preparation for analysis were tested:

1. Coating of the slide with a matrix and consecutively depositing the sample;

2. Coating of the slide with a mixture of matrix and sample.

The procedures were tested using L-proline as analyte and 5-methoxysalicylic acid as a matrix. Average

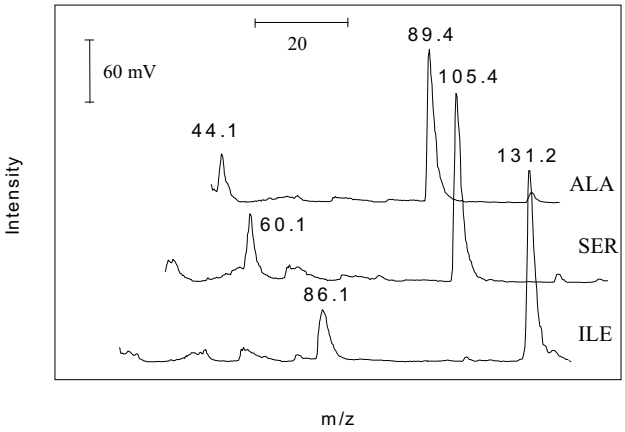

Fig. 1. Mass spectra of L-serine, L-alanine, and L-isoleucine. Experimental conditions: all amino acids of $c = 0.5$ mmol dm^{-3}; sample volume 0.5 mm^3; volume of matrix solution 1.0 mm^3; matrix: 5-methoxysalicylic acid; linear positive mode; relative laser power 95; 100 shots.

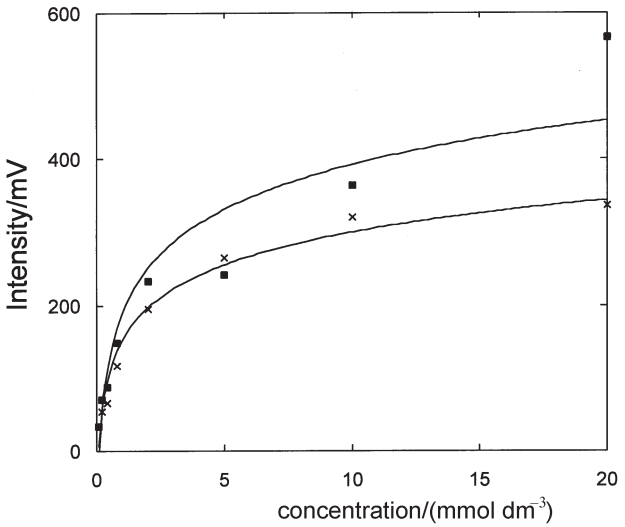

Fig. 2. Dependence of alanine (×) and isoleucine (■) molecular ions peaks intensities on concentration of amino acids. Experimental conditions: amino acids of $c =$ 0.1—20 mmol dm^{-3}; sample volume 0.5 mm^3; volume of matrix solution 1.0 mm^3; matrix: 5-methoxysalicylic acid; linear positive mode; relative laser power 95; 100 shots.

of 20 intensities for successive transfer was (160.6 ± 24.4) mV and for simultaneous transfer (53.9 ± 15.3) mV. It can be seen that higher signal intensities and better reproducibility were obtained using the second method. This method of sample preparation was also simpler and matrix consumption was lower. Experimental conditions: L-proline concentration 1 mmol dm^{-3}; sample volume 0.5 mm^3; volume of matrix solution 1.0 mm^3; volume of sample and matrix mixture solution 1.0 mm^3; matrix: 5-methoxysalicylic acid; linear positive mode; relative laser power 95; 100 shots.

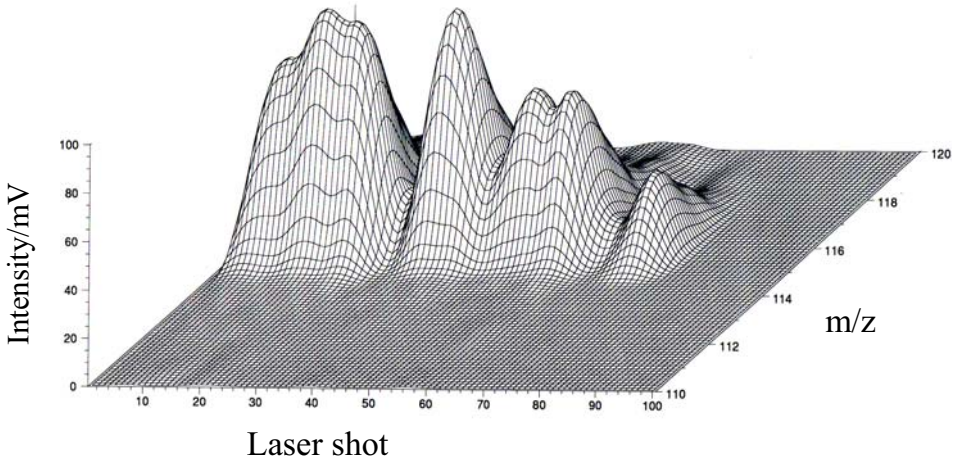

Fig. 3. "Chromatogram" – the dependence of the intensity on molecular mass and order number of the laser shot. Experimental conditions: isoleucine of $c = 1.0$ mmol dm^{-3}; sample volume 0.5 mm^3; volume of matrix solution 1.0 mm^3; matrix: 5-methoxysalicylic acid; linear positive mode; relative laser power 95; 100 shots.

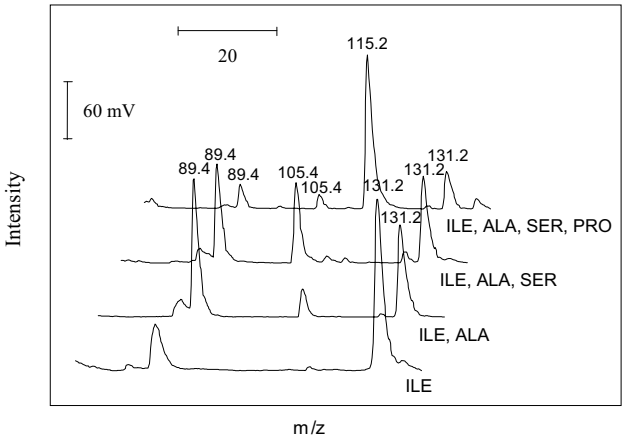

Fig. 4. The dependence of intensity of the isoleucine molecular ion peak on addition of amino acids. Experimental conditions: amino acids of $c = 1.0$ mmol dm^{-3}; sample volume 0.5 mm^3; volume of matrix solution 1.0 mm^3; matrix: 5-methoxysalicylic acid; linear positive mode; relative laser power 95; 100 shots.

Laser power determines the degree of the desorption, ionization, and fragmentation of analytes during MALDI process. The highest intensity values of L-proline ions were obtained in relative laser power signal to 95.

Fig. 1 shows MALDI-TOF mass spectra of L-serine, L-alanine, and L-isoleucine amino acids. The following peaks were observed at m/z: 89.4 (89.09), 105.4 (105.09), 131.2 (131.18). Theoretical values are given in the brackets.

Intensity of the amino acids molecular ion peaks depends on concentration as shown in Fig. 2, the plot shows a logarithmic character. Linear part is obtained in the concentration range 0.1—2 mmol dm^{-3}, which is the most suitable and was therefore used in further measurements and analysis.

Several factors influencing reproducibility of intensities were systematically studied.

The homogeneity of the layer of matrix and sample crystals on the plate is considered to be the key problem for the reproducibility and quantitative analysis on MALDI-TOFMS. Fig. 3 presents the dependence of the intensity on the molecular mass and on the order number of laser shot across the spot. It is evident from Fig. 3 that analyte crystals are on the plate distributed unevenly.

Laser energy was changed in the region from 70 to 150 relative units, but the ratio of the intensity of the fragment ($m/z = 86$) *vs.* the intensity of the molecular ion ($m/z = 131.2$) was not completely reproducible. The ratio of the intensity of isoleucine molecular ion and of its fragment varied from 0.96 to 2.75 (average 1.47 ± 0.58).

The dependence of intensity of the isoleucine molecular ion peak on sequential addition of amino acids (alanine, serine, proline) to the mixture was studied. The numerical value of the intensity of isoleucine, 584, changed after the addition of alanine or a mixture of alanine and serine, or alanine, serine, and proline to 335, 303, and 128, respectively. It can be seen from the results in Fig. 4 that the intensity of isoleucine peak is remarkably decreasing with the addition of other amino acids. L-Alanine, L-serine, and L-isoleucine were selected for the following experiments.

For some matrices the intensity of amino acids peaks depends on the time the sample was inside the spectrometer. In Fig. 5 the dependence of studied amino acids intensity on this exposition time is demonstrated. It is evident that intensities of alanine, serine, and isoleucine peaks pronouncedly decrease with time. This phenomenon was the highest for 5-chlorosalicylic acid. Most probably, the sublimation (evaporation) of matrix decreases amount of analyte deposited in matrix surface. Therefore, 5-chlorosalicylic acid was eliminated in further study.

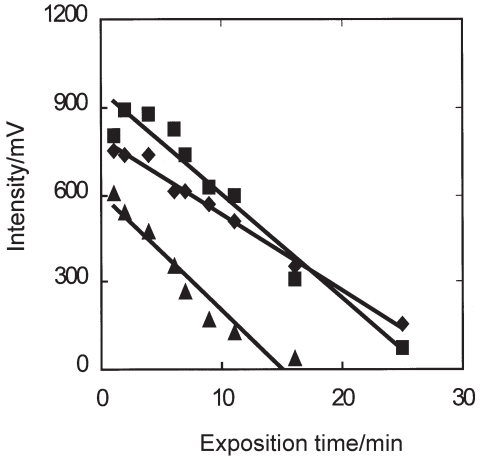

Fig. 5. The dependence of studied alanine (♦), serine (■), and isoleucine (▲) intensity I on time of samples emplacement in vacuum of mass spectrometer cell. Experimental conditions: amino acids of c = 1.0 mmol dm^{-3}; sample volume 0.5 mm^3; volume of matrix solution 1.0 mm^3; matrix: 5-chlorosalicylic acid; linear positive mode; relative laser power 95; 100 shots.

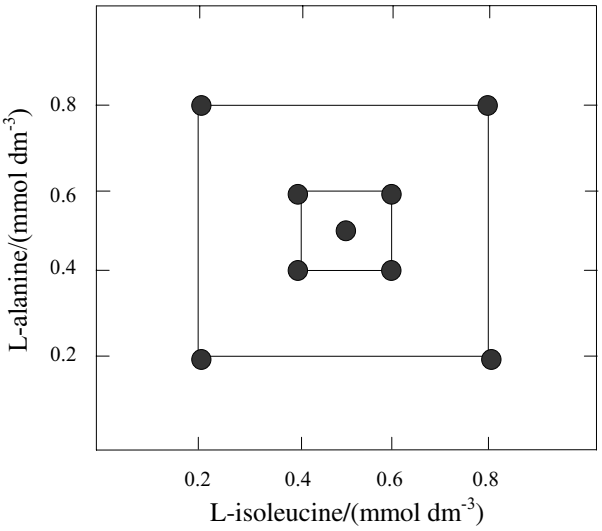

Fig. 6. Experimental design for mixtures of alanine and isoleucine.

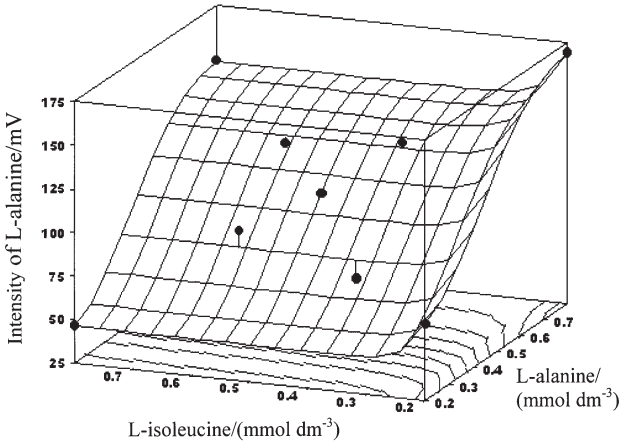

Fig. 7. The dependence of alanine molecular ion peak intensity on alanine and isoleucine concentrations response. Experimental conditions: amino acids of c = 0.2—0.8 mmol dm^{-3}; sample volume 0.5 mm^3; volume of matrix solution 1.0 mm^3; matrix: 5-methoxysalicylic acid; linear positive mode; relative laser power 95; 100 shots.

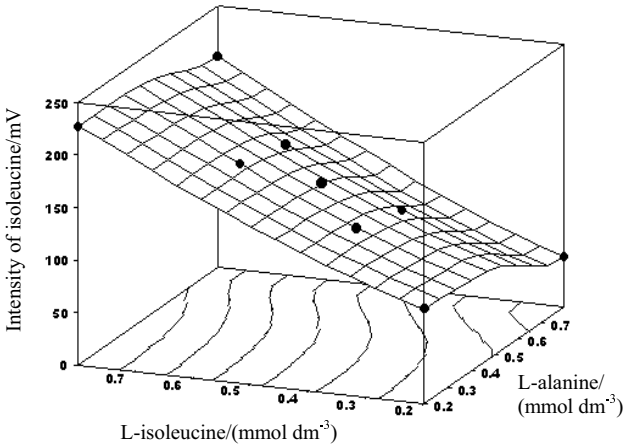

Fig. 8. The dependence of isoleucine molecular ion peak intensity on alanine and isoleucine concentrations response. Experimental conditions: amino acids of c = 0.2—0.8 mmol dm^{-3}; sample volume 0.5 mm^3; volume of matrix solution 1.0 mm^3; matrix: 5-methoxysalicylic acid; linear positive mode; relative laser power 95; 100 shots.

Binary mixtures of L-alanine and L-isoleucine were analyzed in three concentration ranges, namely 6—30 mmol dm^{-3}, 5—10 mmol dm^{-3}, and 0.2—0.8 mmol dm^{-3}. The most convenient for quantitative analysis using MALDI-TOFMS execution is the concentration range 0.2—0.8 mmol dm^{-3}. The mixtures of the components were prepared using the experimental design method. In Fig. 6 an example of the experimental design applied for preparation of mixtures of alanine and isoleucine is demonstrated.

Figs. 7 and 8 represent response surface of alanine and isoleucine. Fig. 7 shows dependence of alanine molecular ion peak intensity on alanine and isoleucine concentrations and Fig. 8 shows influence of alanine and isoleucine on the intensity of isoleucine molecular ion peak.

In Table 1 the results of analysis for mixtures of alanine and isoleucine as obtained using the PLS method are shown. The results obtained by PLS can be regarded as satisfactory.

The structure of ANN was searched for by the Trajan program itself and used for further training process. Fig. 9 demonstrates the optimal structure of neural network used. The results of computation are given in Table 2. It can be seen that results obtained using the PLS method are comparable with results obtained using the ANN method.

Table 1. Results of Analysis for Mixtures of Alanine and Isoleucine as Obtained Using the PLS Method

Alanine concentration/ (mmol dm^{-3})		Relative error/%	Isoleucine concentration/ (mmol dm^{-3})		Relative error/%
obtained	given		obtained	given	
		Calibration set			
0.228	0.200	14.00	0.209	0.200	4.50
0.201	0.200	0.50	0.777	0.800	−2.88
0.498	0.500	−0.40	0.534	0.500	6.80
0.796	0.800	−0.50	0.181	0.200	−9.50
0.793	0.800	−0.88	0.811	0.800	1.38
0.346	0.400	−13.50	0.413	0.400	3.25
0.410	0.400	2.50	0.586	0.600	−2.33
0.596	0.600	−0.67	0.391	0.400	−2.25
0.630	0.600	5.00	0.595	0.600	−0.83
Average prediction error		0.67	Average prediction error		−0.21
		Test set			
0.252	0.200	−20.63	0.159	0.200	25.79
0.825	0.800	−3.03	0.180	0.200	11.11
0.416	0.400	−3.85	0.597	0.600	0.50
Average prediction error		−9.17	Average prediction error		12.47

Table 2. Results of Analysis for Mixtures of Alanine and Isoleucine as Obtained Using the ANN Method

Alanine concentration/ (mmol dm^{-3})		Relative error/%	Isoleucine concentration/ (mmol dm^{-3})		Relative error/%
obtained	given		obtained	given	
		Calibration set			
0.215	0.200	7.60	0.227	0.200	13.38
0.219	0.200	9.66	0.780	0.800	−2.55
0.497	0.500	−0.64	0.506	0.500	1.23
0.786	0.800	−1.77	0.220	0.200	10.04
0.776	0.800	−2.99	0.788	0.800	−1.50
0.392	0.400	−1.92	0.397	0.400	−0.69
0.606	0.600	0.91	0.389	0.400	−2.84
0.598	0.600	−0.42	0.600	0.600	0.03
0.399	0.400	−0.30	0.604	0.600	0.64
Average prediction error		1.13	Average prediction error		1.97
		Test set			
0.234	0.200	17.03	0.233	0.200	16.34
0.792	0.800	−1.03	0.241	0.200	20.57
0.415	0.400	0.02	0.620	0.600	3.33
Average prediction error		5.34	Average prediction error		13.41

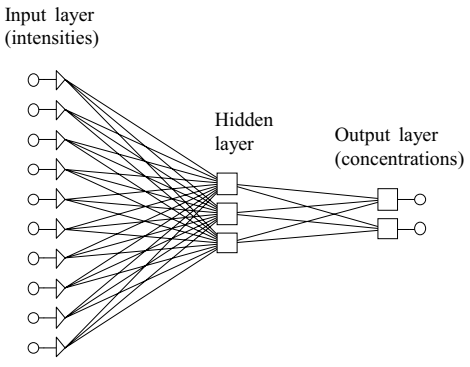

Fig. 9. Optimal structure of the neural network used.

CONCLUSION

Conditions for simultaneous determination of amino acids in their mixtures were found and several interfering factors were found and eliminated. By optimization of MALDI-TOFMS analysis parameters the enhancement of data reproducibility was achieved. Quantitative analysis of amino acids in mixtures by MALDI-TOFMS is possible. The standard deviation of the determination was in the range ± 0.7—13 rel. %.

Acknowledgements. This work is a part of the project FRVŠ 541/2001 of the Ministry of Education, Youth, and Sports, Prague, Czech Republic.

REFERENCES

1. Hillenkamp, F., Karas, M., Beavis, R. C., and Chait, B. T., *Anal. Chem. 63*, 1193A (1991).
2. Tanaka, K., Waiki, H., Ido, Y., Akita, S., Yoshida, Y., and Yoshida, T., *Rapid Commun. Mass Spectrom. 2*, 151 (1988).
3. Kazmaier, T., Roth, S., Zapp, J., Harding, M., and Kuhn, J., *Fresenius' J. Anal. Chem. 361*, 473 (1998).
4. Wang, J., Sporns, P., and Low, N. H., *Food Chem. 47*, 1594 (1999).
5. Ling, Y. Ch., Lin, L., and Chen, Y. T., *Rapid Commun. Mass Spectrom. 12*, 317 (1998).
6. Sarracino, D. and Richert, C., *Bioorg. Med. Chem. Lett. 6*, 2543 (1996).
7. Benard, S., Arnhold, J., Lehnert, M., Schiller, J., and Arnold, K., *Chem. Phys. Lipids 100*, 115 (1999).
8. Asbury, G., Al-Saad, K., Siems, W. F., Hannan, R. M., and Hill, H. H., *Am. Soc. Mass Spectrom. 10*, 983 (1999).
9. Wu, J., Chatman, K., Harris, K., and Siuzdak, G., *Anal. Chem. 69*, 3767 (1997).
10. Hensel, R. R., King, R. C., and Owens, K. G., *Rapid Commun. Mass Spectrom. 11*, 1785 (1997).
11. Chromá-Keull, H., Havliš, J., and Havel, J., *Rapid Commun. Mass Spectrom. 14*, 40 (2000).
12. Gusev, A. I., Wilkinson, W. R., Proctor, A., and Hercules, D. M., *Anal. Chem. 67*, 1034 (1995).
13. Nicola, A. J., Gusev, A. I., Proctor, A., Jackson, E. K., and Hercules, D. M., *Rapid Commun. Mass Spectrom. 9*, 1164 (1995).
14. Vorm, O., Roepstorff, P., and Mann, M., *Anal. Chem. 66*, 3281 (1994).
15. Prestn, L. M., Murray, K. K., and Russell, B. H., *Biological Mass Spectrom. 22*, 543 (1993).
16. Geladi, P., *J. Chemometrics 2*, 231 (1988).
17. Havel, J., Cuesta, F., and Jančář, L., *React. Kinet. Catal. Lett. 49*, 189 (1993).
18. Revilla, A. L., Vargas, G., and Havel, J., *Chem. Pap. 53*, 34 (1999).
19. Perůtka, J. and Havel, J., *J. Chromatogr., A 745* (1—2), 225 (1996).
20. Gasteiger, J. and Zupan, J., *Angew. Chem. 32*, 503 (1993).
21. Aleksander, I. and Morton, H., *An Introduction to Neural Computing*. Chapman and Hall, London, 1990.
22. Havel, J., unpublished results.
23. *Trajan Neural Network Simulator, Release 3.0 D*. Copyright Trajan Software Ltd., 1996—1998.
24. Deming, S. N. and Morgan, S. L., *Experimental Design: a Chemometric Approach*. Elsevier, New York, 1987.
25. Massart, D. L., Vandeginste, B. G. M., Deming, S. M., Michotte, Y., and Kaufman, L., *Chemometrics*: *A Textbook*. Elsevier, Amsterdam, 1988.
26. Morgan, E., *Chemometrics: Experimental Design*. Wiley & Sons, London, 1991.