Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Talanta 64 (2004) 785–790

Talanta

www.elsevier.com/locate/talanta

Short communication

Optimization of artificial neural network for retention modeling in high-performance liquid chromatography

Tatjana Vasiljević [a,*], Antonije Onjia [b], Đuro Čokeša [b], Mila Laušević [a]

[a] *Faculty of Technology and Metallurgy, P.O. Box 494, 11001 Belgrade, Serbia and Montenegro*
[b] *Vinča Institute of Nuclear Sciences, P.O. Box 522, 11001 Belgrade, Serbia and Montenegro*

Received 6 November 2003; received in revised form 19 February 2004; accepted 22 March 2004
Available online 20 May 2004

Abstract

An artificial neural network (ANN) model for the prediction of retention times in high-performance liquid chromatography (HPLC) was developed and optimized. A three-layer feed-forward ANN has been used to model retention behavior of nine phenols as a function of mobile phase composition (methanol-acetic acid mobile phase). The number of hidden layer nodes, number of iteration steps and the number of experimental data points used for training set were optimized. By using a relatively small amount of experimental data (25 experimental data points in the training set), a very accurate prediction of the retention (percentage normalized differences between the predicted and the experimental data less than 0.6%) was obtained. It was shown that the prediction ability of ANN model linearly decreased with the reduction of number of experiments for the training data set. The results obtained demonstrate that ANN offers a straightforward way for retention modeling in isocratic HPLC separation of a complex mixture of compounds widely different in pK_a and log K_{ow} values.

Keywords: HPLC; Phenols; Experimental design; ANN; Back-propagation

1. Introduction

High-performance liquid chromatography (HPLC) was developed during the last few decades as a powerful separation technique and it has been increasingly applied to the analyses of organic pollutants in environmental samples. The method enables complex mixtures to be separated into individual compounds at ambient temperature or slightly above and therefore it is ideally suited for compounds of limited thermal stability. The separation of phenols by HPLC has already been studied [1–4]. However, the interpretation of retention behavior and the optimization of the separation performed with such a technique shows some difficulties due to different parameters (such as mobile phase composition and pH) that affect retention significantly.

Artificial neural networks (ANNs) offer attractive possibilities for non-linear modeling and optimization when underlying mechanisms are very complex. ANNs have been applied to a wide variety of chemical problems such as quantitative structure–activity relationship (QSAR) studies [5], simulation of mass spectra [6], prediction of carbon-13 NMR chemical shift [7], modeling of ion [8], ion interaction [9], gas [10] and liquid [11–17] chromatography.

ANNs are computational simulations of biological networks. Different types of neural-networks have been developed to simulate different tasks of the human brain: classification and pattern recognition [18,19]. An ANN consists of many pathways and nodes organized into a sequence of layers. The first layer is an input layer with one node for each variable or feature of the data. The last layer is an output layer consisting of one node for each variable to be investigated. In between, there is a series of one or more hidden layer(s) consisting of a number of nodes, which are responsible for learning. Nodes of one layer are connected to the nodes of the succeeding layer. Each connection is represented by a number called weight. Initially, a learning phase is defined in which each of the input parameters is applied to a processing element. The weights between these parameters are adjusted until the output is correct. The system can then be applied to unknowns. A detailed description

[*] Corresponding author. Tel.: +381-11-3303647; fax: +381-11-3370387.
E-mail address: tanjadj@eunet.yu (T. Vasiljević).

0039-9140/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2004.03.032

of the theory behind an ANN applied in chromatography has been adequately presented elsewhere [18,19]. Metting et al. [11] pointed out that the response surface for linear and non-linear changing capacity factors in HPLC can be estimated by ANNs with results better than those obtained with linear and non-linear regression models.

There are two major approaches to the ANN modeling in HPLC. One is the prediction of chromatographic behavior based on the molecular structure of compounds, i.e. quantitative structure–retention relationships (QSRR) methodology [12–14]. Loukas et al. [12] used this approach to predict retention behavior of 25 compounds on two different stationary phases and obtained a linear dependence between the predicted and obtained logarithms of capacity factors, with correlation coefficients 0.996 and 0.992. Tham et al. [13] used the QSRR method to model HPLC separation of 18 selected amino acids. They obtained a testing set Root Mean Square (RMS) error of 0.8377. The mobile phase composition was the main factor effecting the separation of amino acids with the output sensitivity of over 40 %. QSRR was used to investigate physico-chemical parameters related to the retention times of three pharmaceutical compounds [14]. The authors used 10 molecular descriptors, mobile phase composition and pH as ANN inputs. The results proved the dominant role of the concentration of the organic modifier and pH in the mobile phase to the retention properties. A sensitivity report showed that descriptor contributions to the model varied from 2 to 9%.

The other approach is the use of a preliminary experimental set as ANN parameters [11,15–17]. By this approach, Zhao et al. [15] modeled the retention behavior of 32 solutes in a methanol–tetrahydrofuran–water system and 49 solutes in a methanol–acetonitrile–water system as a function of mobile phase compositions in HPLC. The average deviation of all data points was 8.74% for the tetrahydrofuran-containing system and 7.33% for the acetonitrile-containing system. Marengo et al. [16] used the same approach to model an ion interaction HPLC method for the simultaneous separation of 20 typical antimicrobial agents. The predictions were very satisfactory with the multiple correlation coefficient higher than 0.97, except for one substance. Agatonović-Kuštrin et al. [17] used ANN for response surface modeling in HPLC optimization. They studied the combined pH and mobile phase composition effect on the reverse-phase liquid chromatographic behavior of amiloride and hydrochlorothiazide. The average error percentages obtained in this work were 0.09 and 0.13%.

In the present work an ANN was employed and optimized to model the retention behavior of nine phenols in their isocratic elution using a methanol–acetic acid–water mobile phase. Phenols are interesting not only from an ecological point of view, as priority pollutants, but also because of their HPLC behaviour due to a wide difference in their pK_a (4.09–9.6) and $\log K_{ow}$ (1.50–3.10) values, so similar model may be applied to a wide variety of compounds. To make an ANN training set, we used preliminary experiments to measure retention times according to experimental design matrices. The aproach of changing the mobile phase composition and pH as ANN parameters was chosen because of the predominant role of these two variables in retention behavior in RP-HPLC [13,14,17].

2. Experimental

2.1. Reagents

Individual stock solutions (1.0 mg ml^{-1}) of nine phenols, which belong to the U.S.EPA priority pollutant list of phenols: (1) phenol, (2) 4-nitrophenol, (3) 2-chlorophenol, (4) 2,4-dinitrophenol, (5) 2-nitrophenol, (6) 2,4-dimethylphenol, (7) 2-methyl-4,6-dinitrophenol, (8) 4-chloro-3-methylphenol, (9) 2,4-dichlorophenol, obtained from ChemService (West Chester, USA), were used to prepare a working mixture. This mixture was diluted appropriately by the mobile phase to prepare a 10 μg ml^{-1} solution of each phenol. HPLC-grade methanol, acetic acid (glac.) (Merck, Darmstadt, Germany) and Milli-Q (Millipore Co., Bedford, USA) processed water were used for these experiments.

2.2. Instrumentation

The HPLC system consisted of a Model SP8810 pump, a Spectra200 variable-wavelength detector (both from Spectra-physics, San Hose, USA), and a Rheodyne (Cotati, USA) 7125 injector fitted with a 10 μl sample loop. A Lichrochart ODS (25 cm × 4.0 mm × 10 μm) column (Merck, Darmstadt, Germany) was kept at ambient temperature. The mobile phases consisted of methanol (30–70% (v/v)) and acetic acid (0.5–1.5% (v/v)). The separation and detection were performed at ambient temperature, at a flow rate of 1.0 ml min^{-1}, and UV detection at $\lambda = 280$ nm.

2.3. Experimental design

The experimental data points (mobile phase composition) in experimental domain (30–70% (v/v) methanol and 0.5–1.5% (v/v) acetic acid in the mobile phase), used to make the ANN training set, were chosen as shown in Table 1.

2.4. Neural network

The measured retention data were used to train and test ANN. Prior to ANN training, the retention time data were normalized. Twenty-five experimental points (i.e. 25 different mobile phase compositions) were used for ANN training. A three-layer (input, hidden and output) feed-forward neural network was used to analyze nonlinear multivariate data.

To predict the retention time accurately and conveniently, the “leave–10%–out” method of cross-validation was applied. With this method, 10% of the data in the training set

Table 1
Experimental data points used to make ANN model

B (%)	A (%)				
	0.5	0.75	1.0	1.25	1.5
30	■ ● ▲ ◆	◆	▲ ◆	◆	■ ● ▲ ◆
40	◆	◆	◆	◆	◆
50	▲ ◆	◆	● ▲ ◆	◆	▲ ◆
60	◆	◆	◆	◆	◆
70	■ ● ▲ ◆	◆	▲ ◆	◆	■ ● ▲ ◆

(A) Acetic acid, (B) methanol. Designs: (■) four experimental points by the use of full factorial design; (●) five experimental points by the use of full factorial design and the central point; (▲) nine experimental points by the use of three level full factorial design; (◆) 25 experimental points evenly distributed in the experimental domain.

are not used to update the weights. Therefore, these 10% can be used as an indication of whether or not memorization has been taking place.

A new experimental point, randomly chosen and not included in the training set, was used to test the prediction power of applied ANN. The ANN systems were simulated using a QwikNet ANN simulator (Craig Jensen, Redmond, USA).

3. Results and discussion

3.1. ANN topology

A three-layer feed-forward neural network trained with an error back-propagation algorithm was used to model the retention of phenols as a function of mobile phase composition. In these networks, signals propagate from the input layer through the hidden layer to the output layer. A node thus receives signals via connections from other nodes or the outside world in the case of the input layer. The net input for a node j is given by:

$$\text{net}_j = \sum_j w_{ji} o_i \qquad (1)$$

where i represents nodes in the previous layer, w_{ji} is the weight associated with the connection from node i to node j, and o_i is the output of node i. The output of a node is determined by the transfer function and the net input of the node. The following sigmoidal transfer function in the hidden layer was used:

$$f(\text{net}_j) = \frac{1}{1 + \mathrm{e}^{-(\text{net}_j + \theta_j)}} \qquad (2)$$

where θ_j is a bias term or threshold value of node j responsible for accommodation nonzero offsets in the data.

The training algorithm used was selected by a trial-and-error process. The weights are updated after each epoch as follows

$$\Delta w_{ij} = -\eta \frac{\partial E(t)}{\partial w(t)} + \alpha\, \Delta w_{ij}(t-1) \qquad (3)$$

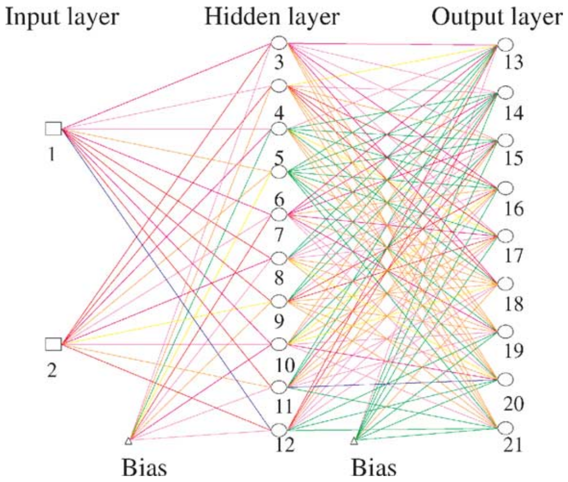

Fig. 1. Schematic representation of a three-layer feed-forward neural network used in this work. Input layer nodes—1, 2; hidden layer nodes—3, 4, 5, 6, 7, 8, 9, 10, 11, 12; output layer nodes—13, 14, 15, 16, 17, 18, 19, 20, 21.

where η is the learning rate, α is the momentum, and $\delta(t) = \partial E/\partial w$ is the actual error at time t. The learning rate, η, controls the rate at which the network learns. Here, an adaptive learning rate method, delta-bar-delta, in which each weight has its own learning rate was employed. The learning rates $\eta(t)$ are updated as follows:

$$\Delta\eta(t) = \begin{cases} \kappa, & \text{if } \bar{\delta}(t-1)\delta(t) > 0 \\ -b\eta(t), & \text{if } \bar{\delta}(t-1)\delta(t) < 0 \\ 0, & \text{else} \end{cases} \qquad (4)$$

where $\kappa = 0.06$ and $b = 0.2$ were chosen constants, and $\bar{\delta}$ is the exponential average of past values of δ:

$$\bar{\delta}(t) = (1-\theta)\delta(t) + \theta\bar{\delta}(t-1) \qquad (5)$$

The momentum, α, controlling the influence of the last weight change on the current weight update was set at zero. Pattern clipping, which specifies the degree of participation

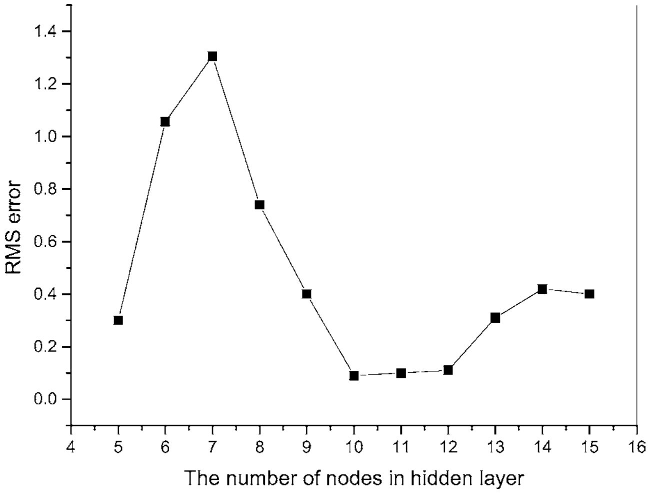

Fig. 2. Hidden layer node numbers vs. RMS error.

of each trained pattern in future learning, input noise, weight decay and error margin were set at 1, 0, 0 and 0.1, respectively.

ANN used in this work is schematically represented in Fig. 1. The input layer consists of two nodes representing eluent concentration of methanol and acetic acid in the mobile phase. The output layer consists of nine nodes representing retention times of nine phenols. In addition, there is a bias (neuron activation threshold) connected to the nodes in the hidden and output layers (but not in the input layer) via modifiable weighted connections. The weights, corresponding to the optimized ANN are presented in Table 2. The weights are arranged in rows. Each row is made up of connections from all nodes of the previous layer, to a node in the current layer.

3.2. ANN optimization

The number of nodes in the hidden layer, number of iteration steps, and the number of experimental data points used for the training set were optimized. In order to determine the optimal number of hidden layer nodes, ANNs with different numbers of hidden nodes were trained. The number of hidden nodes was varied from 5 to 15 and RMS errors were calculated:

$$\mathrm{RMS} = \sqrt{\frac{\sum_{i=1}^{n}(o_i - d_i)^2}{n}} \quad (6)$$

where d_i is a desired output (exp. values), o_i is the actual output (ANN predicted values) and n is the number of compounds in the analyzed set. According to Fig. 2, ANN with 10 hidden nodes had the lowest RMS error, and that number of nodes was chosen for further optimization.

Reduction in the number of experimental data points used for the training set is crucial for the development of the retention model without wasting time on unnecessary experiments. Also, it should be provided that the small number of experimental points in the training set do not affect the predictive ability of the model. Therefore, it is important to determine the optimal number of experimental data points used for the training set. Fig. 3 shows a significant influence of the number of experimental data points used for the training set on the ANN accuracy. The RMS error value linearly decreases with the increase in the number of experimental points (correlation coefficient $r = 0.9997$). For a good agreement between the experimental and the predicted retention times, the value of RMS error should be lower than 0.1. With 25 experimental data sets, the value of RMS error dropped below 0.1. Therefore, 25 experimental points were chosen for ANN training.

To select the best learning times, RMS error values of the training, validation and the testing set versus learning epochs were plotted (Fig. 4). The network training was stopped when the performance goal of 0.1 for RMS error was reached. It is evident from the testing curve that the number of learning epochs for RMS value below 0.1 was around 900.

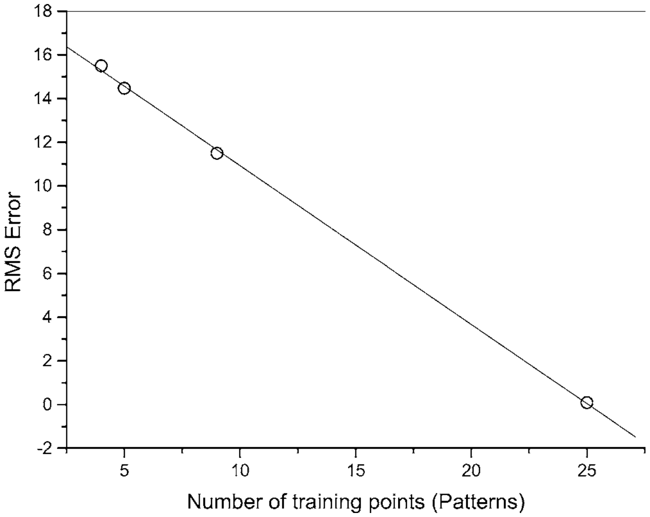

Fig. 3. Number of experimental data points used for training set vs. RMS error.

3.3. ANN validation

The optimized neural network retention model was used to predict retention times for nine phenols. A randomly selected experimental point, not previously included in the training set, was used for the method validation. From the observed and ANN predicted values of retention times of all phenols studied in this work, the percentage-normalized difference (%d) was calculated

$$\%d = \frac{t_{\mathrm{R,exp}} - t_{\mathrm{R,pred}}}{t_{\mathrm{R,exp}}} \cdot 100 \quad (7)$$

where $t_{\mathrm{R,exp}}$ is the experimentally determined retention time and $t_{\mathrm{R,pred}}$ is the ANN predicted retention time.

The results are presented in Fig. 5. In general, all %d values are in excellent agreement within ±0.003% except one

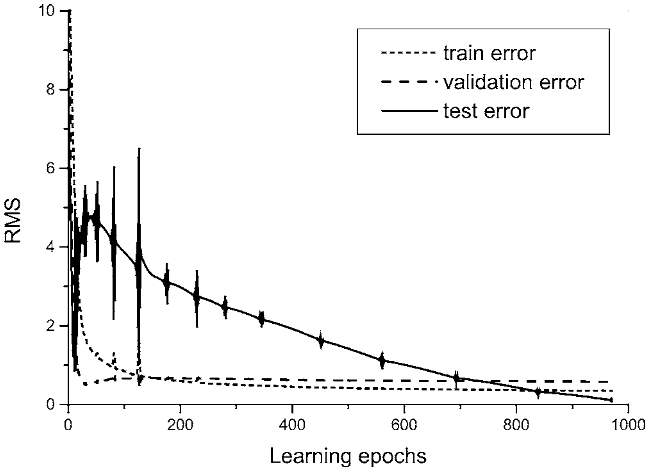

Fig. 4. Number of iteration steps vs. RMS error of training, validation and testing sets.

Table 2
Weight values in optimized neural network presented in Fig. 1 (input layer nodes–hidden layer nodes and hidden layer nodes–output layer nodes)

	Input nodes		Bias 1
	1	2	
Hidden nodes			
3	2.81	−0.77	0.92
4	−1.27	0.16	1.22
5	0.92	0.18	−0.64
6	3.39	2.51	5.75
7	11.0	0.96	9.97
8	1.67	0.14	0.91
9	1.22	0.07	−0.92
10	1.90	0.77	0.22
11	−1.81	−0.18	0.01
12	−2.03	−0.35	0.21

	Hidden nodes										
	3	4	5	6	7	8	9	10	11	12	Bias 2
Output nodes											
13	−0.64	0.62	−0.06	−0.60	−2.69	−0.56	−0.25	−0.06	3.08	0.56	0.55
14	−0.75	1.22	−0.15	−1.74	−2.39	−0.29	0.06	−0.08	2.92	0.59	0.60
15	−0.84	1.40	−0.10	−0.96	−2.77	−0.75	−0.04	−0.35	1.62	1.14	0.87
16	−0.85	2.03	0.13	−0.94	−2.66	−0.14	0.11	−0.70	1.95	0.72	0.14
17	−0.68	2.13	−0.21	−0.94	−2.69	−0.44	0.09	−0.25	1.94	0.90	−0.06
18	−0.79	1.27	−0.16	−1.20	−2.79	−1.11	−0.53	−0.26	1.65	0.66	1.70
19	−1.02	1.15	−0.02	−1.23	−2.80	−0.67	−0.75	−0.37	1.65	0.98	1.36
20	−0.62	2.15	0.01	−1.07	−2.73	−0.61	−0.33	−0.30	2.14	0.70	0.03
21	−0.81	1.59	0.26	−1.10	−2.81	−0.51	−0.46	−0.10	2.19	1.29	−0.10

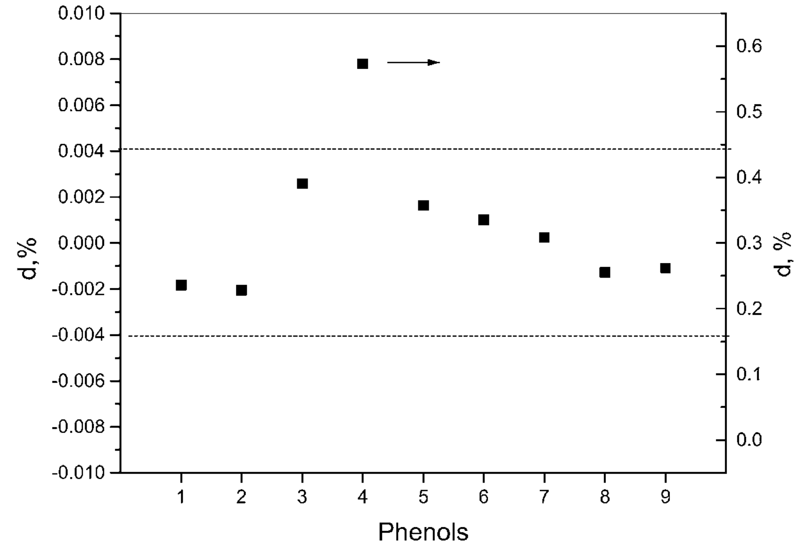

Fig. 5. Percentage-normalized difference between measured and predicted retention times for nine phenols: (1) phenol, (2) 4-nitrophenol, (3) 2-chlorophenol, (4) 2,4-dinitrophenol, (5) 2-nitrophenol, (6) 2,4-dimethylphenol, (7) 2-methyl-4,6-dinitrophenol, (8) 4-chloro-3-methylphenol, (9) 2,4-dichlorophenol.

(obtained for 2,4-dinitrophenol) having %*d* value of 0.57%. The results indicate that ANN can be used as a very promising tool for retention modeling in HPLC.

4. Conclusion

In this work, ANN was used for retention modeling of phenols in HPLC. The number of nodes in the hidden layer of ANN, number of iteration steps, and the number of experimental data points used for the training set were optimized. Through the above process, we found out that the optimum number of hidden layer nodes was 10 and that the number of experimental data points and the number of learning epochs to achieve desirable accuracy (RMS error below 0.1) were 25 and 900, respectively. The predicted and experimental retention times for eight out of nine studied phenols were in excellent agreement to within

±0.003%. However, ANN modeling of the retention of 2,4-dinitrophenol gave somewhat different but still accurate outputs (0.57%).

In general, these results show that ANN can be a very satisfactory tool in modeling of HPLC separation of compounds, such as phenols, of widely different pK_a (4.09–9.6) and log K_{ow} (1.50–3.10).

Acknowledgements

This work was supported by the Serbian Ministry of Science, Technology and Development.

References

[1] E. Pocurull, M. Calull, R.M. Marcé, F. Borrull, J. Chromatogr. A 719 (1996) 105.
[2] A. Onjia, T. Vasiljević, Đ. Čokeša, M. Laušević, J. Serb. Chem. Soc. 67 (2002) 745.
[3] S.N. Lanin, Yu.S. Nikitin, Talanta 36 (1989) 573.
[4] A. Di Corcia, A. Bellioni, M. Diab Madbouly, S. Marchese, J. Chromatogr. A 733 (1996) 383.
[5] A. Yan, G. Jiao, Z. Hu, B.T. Fan, Comput. Chem. 24 (2000) 171.
[6] M. Jalali-Heravi, M.H. Fatemi, Anal. Chim. Acta 415 (2000) 95.
[7] S.L. Anker, P.C. Jurs, Anal. Chem. 64 (1992) 1157.
[8] G. Srečnik, Z. Debeljak, S. Cerjan-Stefanović, M. Nović, T. Bolanca, J. Chromatogr. A 973 (2002) 47.
[9] E. Marengo, M.C. Gennaro, S. Angelino, J. Chromatogr. A 799 (1998) 47.
[10] Y. Gao, Y. Wang, X. Yao, X. Zhang, M. Liu, Z. Hu, B. Fan, Talanta 59 (2003) 229.
[11] H.J. Metting, P.M.J. Coenegracht, J. Chromatogr. A 728 (1996) 47.
[12] S.Y. Tham, S. Agatonović-Kuštrin, J. Pharm. Biomed. Anal. 28 (2002) 581.
[13] S. Agatonović-Kuštrin, M. Zečević, Lj. Živanović, J. Pharm. Biomed. Anal. 21 (1999) 95.
[14] Y.L. Loukas, J. Chromatogr. A 904 (2000) 119.
[15] R.H. Zhao, B.F. Yue, J.Y. Ni, H.F. Zhou, Y.K. Zhang, Chemom. Intell. Lab. Syst. 45 (1999) 163.
[16] E. Marengo, V. Gianotti, S. Angioi, M.C. Gennaro, J. Chromatogr. A 1029 (2004) 57.
[17] S. Agatonović-Kuštrin, M. Zečević, Lj. Živanović, I. Tucker, Anal. Chim. Acta 364 (1998) 265.
[18] J. Gasteiger, J. Zupan, Angew. Chem. Int. Ed. Eng. 32 (1993) 503.
[19] J. Zupan, Acta Chim. Slov. 41 (1994) 327.