

Toni Ann Riveros¹
Lyra Porcasi²
Sarah Muliadi²
Grady Hanrahan^{2*}
Frank A. Gomez¹

¹Department of Chemistry and Biochemistry, California State University, State University Drive, Los Angeles, CA, USA

²Department of Chemistry, California Lutheran University, Thousand Oaks, CA, USA

Received September 4, 2008

Revised January 23, 2009

Accepted January 26, 2009

Short Communication

Application of artificial neural networks in the prediction of product distribution in electrophoretically mediated microanalysis

The successful application of artificial neural networks toward the prediction of product distribution in electrophoretically mediated microanalysis is presented. To illustrate this concept, we examined the factors and levels required for optimization of reaction conditions for the conversion of nicotinamide adenine dinucleotide to nicotinamide adenine dinucleotide, reduced form by glucose-6-phosphate dehydrogenase in the conversion of glucose-6-phosphate to 6-phosphogluconate. A full factorial experimental design examining the factors voltage, enzyme concentration, and mixing time of reaction was utilized as input-output data sources for suitable artificial neural networks training for prediction purposes. This approach proved successful in predicting optimal values in a reduced number of experiments. Model validation addressing the extent of reaction and product ratios were subsequently determined experimentally in replicate analyses, with results shown to be in good agreement (<10% discrepancy difference) with predicted data.

Keywords:

Artificial neural networks / CE / Electrophoretically mediated microanalysis / Experimental design
DOI 10.1002/elps.200800703

Enzymes are highly specific and efficient organic catalysts that initiate, facilitate, and regulate the vast majority of chemical reactions that propel metabolic processes. Without them, most of the important processes involved in life would not occur. The activity of an enzyme is highly dependent on many factors including temperature, pH, and salt concentration. Hence, these parameters must be considered when measuring an enzyme's activity. There are a variety of techniques currently available to measure enzyme activity including spectrophotometric, fluorometric, radiometric, and chromatographic methods. One technique that has grown in use, since it was initially described in 1992 by Regnier and coworkers is electrophoretically mediated microanalysis (EMMA) [1]. In classical on-capillary EMMA, differential electrophoretic mobility is utilized to merge distinct zones of analyte and analytical reagent(s) under the influence of an electric field [2–10]. The reaction is then

allowed to proceed within the region of reagent overlap either in the presence or absence of an applied potential, and the resultant product is transported to the detector under the influence of an electric field. Pre- and post-capillary assays are also widely used.

Although EMMA and related CE methods have been widely applied to measuring enzyme activities and other parameters, little work has been devoted to optimizing the conditions for these techniques. For example, in free solution other experimental parameters (mobility differences, voltage, length of plug-plug overlap, or Joule heating) unique to CE must be considered when developing and optimizing conditions for the EMMA assay. In a very early study, we used a univariate approach in optimizing the conditions for EMMA [4]. This concept was demonstrated by examining the optimization of reaction conditions for the conversion of nicotinamide adenine dinucleotide (NAD) to nicotinamide adenine dinucleotide, reduced form (NADH) by glucose-6-phosphate dehydrogenase (G6PDH, EC 1.1.1.49) in the conversion of glucose-6-phosphate (G6P) to 6-phosphogluconate. Here, voltage (V), enzyme concentration (E), and mixing time of reaction (M) at the applied voltage were studied. Since only one parameter was varied at a time, no interactive effects could be determined. Earlier attempts at modifying substrate and enzyme concentrations were performed, but these changes resulted in either

Correspondence: Dr. Frank A. Gomez, Department of Chemistry and Biochemistry, California State University, Los Angeles, 5151 State University Drive, Los Angeles, CA 90032, USA

E-mail: fgomez2@calstatela.edu

Fax: +1-323-343-6490

Abbreviations: ANN, artificial neural networks; EMMA, electrophoretically mediated microanalysis; G6P, glucose-6-phosphate; G6PDH, glucose-6-phosphate dehydrogenase; MO, mesityl oxide; NAD, nicotinamide adenine dinucleotide; NADH, nicotinamide adenine dinucleotide, reduced form; RSM, response surface methodology; SSE, sum square error

*Additional corresponding author: Dr. Grady Hanrahan
E-mail: ghanraha@clunet.edu

nominal conversion to product, poor signal to noise ratios, or peak areas yielding high SD.

We recently described the use of response surface methodology (RSM) and chemometrics in EMMA [11]. Using the same G6PDH model system, the product distribution (product/(substrate+product)) of the reaction was predicted by examining the same parameters in a Box-Behnken response surface design. The model predicted results were in good agreement (7.1% discrepancy difference) with the experimental data. The use of chemometric RSM provided a direct relationship between electrophoretic conditions and product distribution of the microscale reactions in CE and has provided scientists with a new and versatile approach to optimizing enzymatic experimental conditions. Such knowledge was also useful in providing input data to the model described in the current study.

There have also been a variety of other studies incorporating chemometrics in CE including, for example, optimizing the separation of two or more components *via* artificial neural networks (ANN) [12,13]. The use of natural computing methods, in particular ANN, has recently proved valuable in a variety of separation techniques. The popularity of ANN is derived from its ability to perform a variety of complex modeling tasks without the need for detailed understanding of underlying phenomena or the establishment of defined mathematical models [14]. ANN has been particularly popular in the CE community, highlighting studies in CZE [12,13,15–21], micellar electrokinetic chromatography [22–25], and a selection of CE-related applications [26–29]. When combined with chemometric experimental design techniques, they have been shown to provide a reduction in the number of required experiments and analysis time, as well as enhancing the statistical data evaluation capabilities [15,16,18].

The aim of the current study was to demonstrate the use of ANN in improving prediction capabilities and enzyme conversion in EMMA. It presents the successful use of a combined experimental design/ANN approach in EMMA, ultimately providing robust and enhanced optimization capabilities. This work is significant in that a combined approach for understanding and characterizing a given enzymatic experimental system is presented and provides a basis for future analysis of other enzyme systems using ANN.

All chemicals were of analytical grade. G6P, G6PDH (G6PDH, EC 1.1.1.49) and NAD were purchased from SigmaAldrich (St. Louis, MO, USA) and used without further purification. Mesityl oxide (MO) was purchased from Calbiochem (San Diego, CA, USA). Stock solutions of G6PDH (9 mg mL⁻¹), MO (100 µL/1000 µL buffer), and NAD (3 mg mL⁻¹) were each prepared by dissolving in buffer (50 mM boric acid, G6P-200 µM; pH 9.44).

All analyses were carried out using a P/ACE MDQ CE instrument (Beckman Instruments, Fullerton, CA, USA). The capillary tubing (Polymicro Technologies, Phoenix, AZ, USA) was of coated fused silica with an internal diameter of 50 µm, length from inlet to detector 40 cm, with a length from detector to outlet of 10.2 cm, resulting in a total length of 50.2 cm. Data was collected and analyzed with Beckman 32 Karat software.

A sample solution (3.6 nL) containing NAD and MO was introduced by pressure injection (3 s; 0.5 psi) into the capillary equilibrated with buffer. Next, a solution (3.6 nL) of G6PDH was introduced by pressure injection (3 s, 0.5 psi), and electrophoresis was run at a given contact voltage (+1.0, +13, and +25 kV) for the desired mixing time (0.2, 0.8, and 1.4 min). The voltage was then increased to +25 kV to complete elution of all species.

To evaluate the influence of mixing time, voltage, and enzyme concentration on the percentage conversion of NAD to NADH by G6PDH we employed the use of a two-level, full factorial design (2³). Such designs are described as 2^k-designs where the base two stands for the number of factor levels and *k* the number of factors each with a high (+) and low (–) value [30,31]. The randomized eight runs and acquired data obtained are highlighted in Table 1. Experimental factors and levels were selected on the basis of previous univariate and RSM studies. In addition, we carefully considered our knowledge of the reactivity of the enzyme based on known experimental protocols for providing a foundation for the chosen experimental boundary values.

In this work, a multilayered, feed-forward ANN analyzed in JMP (SAS Institute) statistical software was utilized. ANN are computational modeling tools defined by structures comprised of interconnected adaptive processing elements (neurons or nodes) that perform parallel computations for data processing and interpretation [32,33]. Input and output

Table 1. Results from the 2³ factorial design

Experiment	Mixing time (min)	Voltage (kV)	Enzyme concentration (mg/mL)	Mean percentage conversion (experimental)	RSD (%) (experimental, <i>n</i> = 3)	Percentage conversion (predicted)	Percentage difference
1	0.2	1	1	8.68	8.21	7.99	+7.9
2	0.2	1	7	12.6	4.91	13.3	–5.6
3	0.2	25	7	10.5	5.15	12.1	–15.2
4	0.2	25	1	4.99	13.6	5.17	–3.6
5	1.4	25	7	11.2	1.17	10.5	+6.3
6	1.4	25	1	8.93	7.88	9.21	–3.1
7	1.4	1	1	17.8	2.99	18.5	–3.9
8	1.4	1	7	36.4	6.70	34.7	+4.7

layers are present and represent data inputs into the neural network and response (or often multiple responses) of the network to the inputs, respectively. One or more intermediate layers, termed hidden layers, may exist between the input and output layer, which are utilized to compute associations between parameters. Ostensibly, ANN is a collection of processing units, which communicate by activation of neurons over a number of weighted connections [32]:

$$\xi_j^l = \sum_{i=1}^{N_{l-1}} w_{ji}^l x_i^{l-1} \quad (1)$$

where ξ_j^l = the net effect and w_{ji}^l = a connection weight. In a given layer l , a typical neuron, j , integrates the signals, x_j , impinging upon it and producing the net effect.

The most commonly reported algorithm for training ANN is backpropagation [34]. This algorithm involves running each input pattern in a training set through a feed forward ANN, calculating the error incurred by the difference in actual and target network output, backpropagating and adjusting weights according to error. Most units in ANN transform their input by using a scalar-to-scalar activation function. A basic continuous sigmoid function typically used in backpropagation is as follows [32]:

$$\sigma(\xi) = \frac{1}{1 + e^{-\xi}} \quad (2)$$

This process is repeated iteratively until the total error across the training set is below a specified maximum. Investigators must take great care in choosing the number of nodes *per* layer and maximum training error because inappropriate selections will lead to over/underfitting of training set pattern data. Overfitting is especially detrimental to analyses as it can easily lead to predictions that are far beyond the range of the training data. As with the 2^3 screening design, ANN data were analyzed in JMP (SAS Institute) statistical software. More detailed theory of ANN can be found in a variety of informative publications [32–34].

From the experimental design matrix (Table 1), eight experiments (performed in triplicate) were analyzed in randomized order with the mean response (percentage conversion) reported. Included in the table are the (RSD) values (%) between replicate analyses. Factorial ANOVA results (Table 2) confirmed good agreement between measured and predicted values. ANOVA for a linear regres-

Table 2. Factorial ANOVA results

Source	D.F. ^{a)}	Sum of squares	Mean square	F-ratio ^{b)}
Model	3	1469.82	489.94	17.75
Error	20	551.998	27.600	Prob>F
C. total ^{c)}	23	2021.82	-	<0.001

a) D.F. = degrees of freedom.

b) F-Ratio = ratio of mean square for lack of fit to mean square for Pure Error.

c) The total (C. total) is the sum of squared distances of each response from the sample mean.

sion partitions the total variation of a sample into components. These components are then used to compute an *F*-ratio (ratio of mean square for lack of fit to mean square for pure error) that evaluates the effectiveness of the model.

Statistical analysis of the model equations revealed r^2 (0.93) and adjusted r^2 (0.91) values with good precision for all responses. An examination of Prob>F from the effect test results revealed that enzyme concentration had the greatest single effect (Prob>F = <0.001). Prob>F is the significance probability for the *F*-ratio, which states that if the null hypothesis is true, a larger *F*-statistic would only occur due to random error. Significant probabilities of 0.05 or less are often considered evidence of a significant regression factor in the model. Additionally, a significant interactive effect (Prob>F = 0.031) between mixing time and voltage was revealed. As reported in an earlier study, this can be explained due to the extent of contact between substrate and enzyme being dictated by the differential in electrophoretic mobilities, which is in turn dictated by mixing time and voltage [11].

Linear models that take into account interaction effects obtained from two level designs, however, have been reported to lead to erroneous conclusions about factor effects where second order effects (curvatures) are neglected [35]. Here, the use of three level designs (*e.g.* response surface modeling) or ANN can be extremely beneficial. ANN itself has proved superior to mathematical modeling in terms of chromatographic prediction by a variety of published sources [13, 35, 36].

In order to optimize the conversion of NAD to NADH by G6PDH, an optimal 3:4:1 ANN structure generated using information obtained from the 2^3 factorial screening design was developed. Refer to Fig. 1 for visualization of optimal hidden node determination. Here, the number of nodes was varied from three to nine and plotted against the sum square error (SSE), with the latter computed from:

$$SSE = \sum_{i=1}^n (d_i - a_i)^2 \quad (3)$$

where d_i = the desired output, a = the actual output and n = the number of experiments in the training set. As shown, four hidden nodes resulted in the lowest SSE with no further improvement upon increasing the hidden node

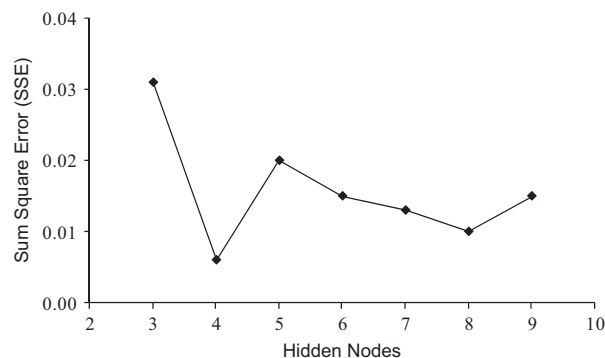


Figure 1. SSE values versus the number of hidden nodes of input data. Note that the SSE is the sum of the errors squared. See Eq. (3) in the text for a detailed definition.

number. To select the optimum number of iterations, examination of the mean square error, given by:

$$MSE = \frac{SSE}{n} \quad (4)$$

of the training set and testing set *versus* learning iterations was performed. Here, the number of iterations was stopped at 7500, a value where the error for the data set ceased to decrease. Upon adequate ANN structure determination (3:4:1) and model development, a data subset in the range selected in the experimental design was created and ANN utilized for prediction purposes, ultimately searching for optimized percentage conversion. This process helps ensure that true generalization has occurred. Generalization refers to ANN's ability to automatically classify data that were not seen during training.

As further justification for our model, we ran a general linear model (running a neural network without a hidden layer) and compared this with our hidden layer model in terms of training data (with each model run five times) to assess which model generalized the best. Comparison was made in terms of the corrected concordance index (*c*-index), where a *c*-index of 1 indicates a “perfect” model and a *c*-index of 0.5 indicates a model that cannot predict any better than a random model. The mean *c*-index for the hidden layer model was 0.8 ± 0.1 , whereas the general linear model registered 0.6 ± 0.1 . Additionally, we employed the use of the akaike information criteria (AIC) for further estimation of ANN model generalizability. More specifically, the AIC was used in comparing the two models with the same training set data. Here, we assessed the error term (the model that had the lowest AIC was considered to be the best). This proved valuable in our selection of the ANN hidden layer model.

From the data patterned by the ANN, a contour profile function was utilized to produce a response surface for the two interactive factors (mixing time and voltage). This interactive profiling facility was used for optimizing the response surface graphically with optimum predicted values of: mixing time = 1.41 min, voltage = 1.2 kV, and enzyme concentration held constant at 1.00 mg/mL. These conditions resulted in a predicted conversion of 42.5%.

To demonstrate the predictive ability of the developed ANN model, a series of three replicate experiments using the ANN model predicted optimal conditions were performed with the results highlighted in Table 3. Here, sound statistical evidence was obtained by good RSD values (%) between replicates and overall agreement (percentage difference) between model predicted and experimental conversions. Note the systematic negative relative differences between the model predicted and experimental. A likely criticism comes in the form of the “Black Box” discussion, where models are considered applicable only within a given system space. We acknowledge that our training data set is not overly large and likely resulted in predictions slightly beyond the range of the training data. We have, however, presented a representative subset (in statistical terms) through the incorporation of proper experimental design. More noteworthy, our ANN model allowed extrapolation and prediction beyond our initial range of chosen factors in the factorial

Table 3. Percent conversion results of experimental *versus* ANN model predicted under optimal conditions (mixing time = 1.41 min, voltage = +1.2 kV and enzyme concentration = 1.00 mg/mL)

Experiment	Percentage conversion (experimental)	RSD (%) (experimental, <i>n</i> = 3)	Percentage conversion (predicted)	Relative difference
1	40.4	4.18	42.5	−5.2
2	38.7	2.34	42.5	−9.8
3	41.3	3.14	42.5	−2.9

design. As a result, percentage conversion (experimental) increased substantially from the factorial design and when compared with our previous RSM study [11].

A representative electropherogram from replicate number 2 is shown in Fig. 2. While the peak for NAD is sharp in the electropherogram, the peak for NADH is broad and tails in the front end. Reason for this can be that as the front end of the zone of enzyme solution migrates into the zone of NAD (which is also moving away from the zone of enzyme but at a slower velocity), the initial conversion of NAD to NADH has yet to reach the maximal rate and is still accelerating. On continued electrophoresis, the concentration of NAD in the plug that is overlapped with the plug of enzyme has reached its maximum thereby resulting in the optimal conversion rate to product. Stacking of the product plug occurs on continued electrophoresis resulting in the classic peak shape at the end of the overlap of the two plug zones. Since the length of the plugs of both enzyme and NAD are small (0.18 mm), the peak for NADH appears as a peak and does not resemble a rectangular-shaped peak as is observed with much larger sample plugs.

These results clearly show a distinct improvement in prediction power when compared with our previous use of a Box-Behnken response surface model alone in a similar EMMA study [11]. The input-output patterns for ANN training in this work required the use of only eight experimental runs through a full factorial design. This is compared with our previous work using the Box-Behnken, which required a total of 15 experimental runs to obtain appropriate model predicted values. It is widely known that the Box-Behnken approach has regions of poor prediction in the complete factor space. The use of ANN has reduced the amount of NAD required in the optimization studies from 500 to 130 pmol as compared with our previous RSM study [11]. The amount of enzyme is approximately the same under both techniques (RSM 2.8 pmol, ANN 3.0 pmol).

In summary, a combined experimental design/ANN approach in EMMA has been described. Here, the conversion of NAD to NADH by G6PDH (G6PDH, EC 1.1.1.49) in the conversion of G6P to 6-phosphogluconate was used as the model system. Such an approach allowed powerful prediction capabilities in fewer numbers of experiments than in previous studies using univariate approaches or RSM techniques. Input data gathered from properly

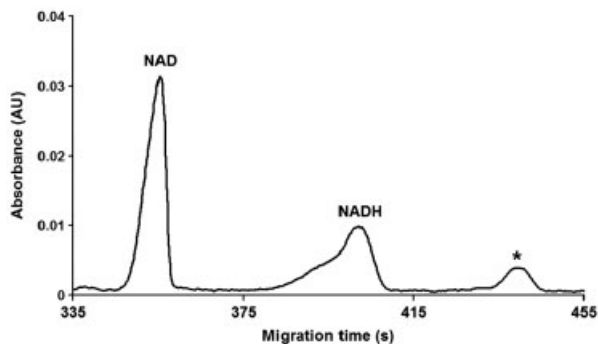


Figure 2. Representative electropherogram showing the separation of NAD and NADH after reaction with G6DPH in 50 mM borate, 200 μ M G6P buffer (pH 9.44). The total analysis time was 8.45 min at +1.0 kV (92.8 μ A) using a 40.0 cm (inlet to detector, 50.2 cm total length) coated capillary. The peak marked * is an impurity.

designed (and statistically validated) experiments helped ensure proper model development. To this end, such a study provides guidance for the development and application of ANN tools that can increase the robustness and transparency, and therefore confidence and acceptance of ANN modeling techniques as methods for optimizing conditions for an enzymatic reaction. Current work is focused on expanding the present study to include examination of additional parameters including kinetics, EOF, temperature and flow rate, as well as additional enzymatic systems.

The authors gratefully acknowledge financial support for this research by grants from the National Science Foundation (CHE-0515363, DMR-0351848 and CBET-0723271). The authors also acknowledge support from the John Stauffer Charitable Trust and the Swenson Summer Research Internship Program. Finally, we thank Joseph Rower for his initial work on neural network development which led to a greater understanding of the experimental details presented in this paper.

The authors have declared no conflict of interest.

References

- [1] Bao, J., Regnier, F. E., *J. Chromatogr. A* 1992, 608, 217–224.
- [2] Harmon, B. J., Patterson, D. H., Regnier, F. E., *Anal. Chem.* 1993, 65, 2655–2662.
- [3] Patterson, D. H., Harmon, B. J., Regnier, F. E., *J. Chromatogr. A* 1994, 662, 389–394.
- [4] Kwak, E.-S., Esquivel, S., Gomez, F. A., *Anal. Chim. Acta* 1999, 397, 183–190.
- [5] Nováková, S., Van Dyck, S., Van Schepdael, A., Hoogmartens, J., Glatz, Z., *J. Chromatogr. A* 2004, 1032, 173–184.
- [6] Urban, P. L., Goodall, D. M., Bergström, E. T., Bruce, N. C., *J. Chromatogr. A* 2007, 1162, 132–140.
- [7] Whisnant, A. R., Johnston, S. E., Gilman, S. D., *Electrophoresis* 2000, 21, 1341–1348.
- [8] Latorre, R. M., Saurina, J., Hernandez-Cassou, S., *Electrophoresis* 2001, 22, 4355–4362.
- [9] Jimidar, M., Vennekens, T., Van Ael, W., Redlich, D., De Smet, M., *Electrophoresis* 2004, 25, 2876–2884.
- [10] Zhang, J., Hoogmartens, J., Van Schepdael, A., *Electrophoresis* 2008, 29, 56–65.
- [11] Montes, R. E., Gomez, F. A., Hanrahan, G., *Electrophoresis* 2008, 29, 375–380.
- [12] Pazourek, J., Gajdošová, D., Spanilá, M., Farková, M., Novotná, K., Havel, J., *J. Chromatogr. A* 2005, 1081, 48–54.
- [13] Zhang, Y., Li, H., Hou, A., Havel, J., *Talanta* 2005, 65, 118–128.
- [14] Despagne, F., Massart, D. L., *Analyst* 1998, 123, 157–178.
- [15] Havel, J., Peña-Méndez, E. M., Rojas-Hernández, A., Doucet, J.-P., Panaye, A., *J. Chromatogr. A* 1998, 793, 317–329.
- [16] Dohnal, V., Farková, M., Havel, J., *Chirality* 1999, 848, 616–621.
- [17] Hamed, A. B., Elosta, S., Havel, J., *J. Chromatogr. A* 2005, 1084, 7–12.
- [18] Jalali-Heravi, M., Shen, Y., Hassanisadi, M., Khaledi, M. G., *J. Chromatogr. A* 2005, 1096, 58–68.
- [19] Spanilá, M., Pazourek, J., Farková, M., Havel, J., *J. Chromatogr. A* 2005, 1084, 180–185.
- [20] Bocaz-Beneventi, G., Latorre, R., Farková, M., Havel, J., *Anal. Chim. Acta* 2002, 452, 47–63.
- [21] Yu, K., Cheng, Y., *Talanta* 2007, 71, 676–682.
- [22] Dobiášová, Z., Pazourek, J., Havel, J., *Electrophoresis* 2002, 23, 268–277.
- [23] Havel, J., Madden, J. E., Haddad, P. R., *Chromatographia* 1999, 49, 481–488.
- [24] Friáz-García, S., Sánchez, M. J., Rodríguez-Delgado, M. A., *Electrophoresis* 2004, 25, 1042–1050.
- [25] Jalali-Heravi, M., Garkani-Nejad, Z., *J. Chromatogr. A* 2001, 927, 211–218.
- [26] Zhao, R., Xu, G., Yue, B., Liebich, H. M., Zhang, Y., *J. Chromatogr. A* 1998, 828, 489–496.
- [27] Latorre, R. M., Hernandez-Cassou, S., Saurina, J., *J. Sep. Sci.* 2001, 24, 427–434.
- [28] Fatemi, M. H., Goudarzi, N., *Electrophoresis* 2005, 26, 2968–2973.
- [29] Zhang, F., Li, H., *Electrophoresis* 2005, 26, 1692–1702.
- [30] Lundstedt, T., Seifert, E., Abramo, L., Thelin, B., Nyström, A., Pettersen, J., Bergman, R., *Chemometr. Intell. Lab. Syst.* 1998, 42, 3–40.
- [31] Hanrahan, G., Lu, K., *Crit. Rev. Anal. Chem.* 2006, 36, 141–151.
- [32] Basheer, I. A., Hajmeer, M., *J. Microbiol. Methods* 2000, 43, 3–31.
- [33] Schalkoff, R. J., *Artificial neural networks*. McGraw-Hill, New York, 1997.
- [34] Rumelhart, D., Hinton, G. E., Williams, R. J., *Parallel distributed processing: explorations in the microstructure of cognition*, MIT Press, Cambridge, 1986.
- [35] Siouffi, A. M., Than-Tan-Luu, R., *J. Chromatogr. A* 2000, 892, 75–106.
- [36] Sacchero, G., Bruzzoniti, M. C., Sarzanini, C., Menstati, E., Metting, H. J., *J. Chromatogr. A* 1998, 799, 35–45.