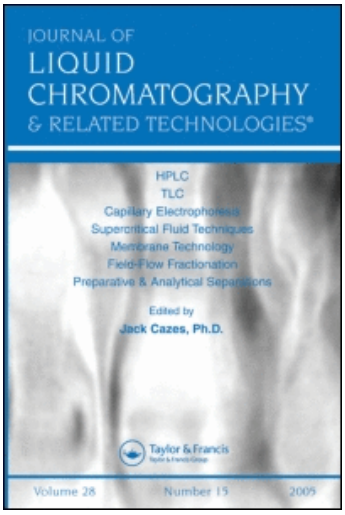


Journal of Liquid Chromatography & Related Technologies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

Artificial Neural Network Modeling of pH Dependent Structural Descriptor-Mobility Relationship for Capillary Zone Electrophoresis of Tripeptides

M. Olajos[ab]; T. Chován[a]; S. Mittermayr[bc]; T. Kenesei[a]; P. Hajos[a]; I. Molnár[d]; F. Darvas[c]; A. Guttman[b]
[a] Institute of Chemical Engineering, University of Pannonia, Veszprém, Hungary [b] Horváth Laboratory of Bioseparation Sciences, Institute of Analytical Chemistry, Innsbruck, Austria [c] Institute of Biomedical Engineering, University for Health Sciences and Technology, Hall, Austria [d] Molnár-Institute of Applied Chromatography, Berlin, Germany

To cite this Article Olajos, M. , Chován, T. , Mittermayr, S. , Kenesei, T. , Hajos, P. , Molnár, I. , Darvas, F. and Guttman, A.(2008) 'Artificial Neural Network Modeling of pH Dependent Structural Descriptor-Mobility Relationship for Capillary Zone Electrophoresis of Tripeptides', Journal of Liquid Chromatography & Related Technologies, 31: 15, 2348 — 2362
To link to this Article: DOI: 10.1080/10826070802281935
URL: http://dx.doi.org/10.1080/10826070802281935

Journal of Liquid Chromatography & Related Technologies®, 31: 2348–2362, 2008

ISSN: 1082-6076 print/1520-572X online
DOI: 10.1080/10826070802281935

Artificial Neural Network Modeling of pH Dependent Structural Descriptor-Mobility Relationship for Capillary Zone Electrophoresis of Tripeptides

M. Olajos,[1,2] T. Chován,[1] S. Mittermayr,[2,3] T. Kenesei,[1] P. Hajos,[1] I. Molnár,[4] F. Darvas,[3] and A. Guttman[2]

[1]Institute of Chemical Engineering, University of Pannonia, Veszprém, Hungary
[2]Horváth Laboratory of Bioseparation Sciences, Institute of Analytical Chemistry, Innsbruck, Austria
[3]Institute of Biomedical Engineering, University for Health Sciences and Technology, Hall, Austria
[4]Molnár-Institute of Applied Chromatography, Berlin, Germany

Abstract: The aim of this work was to study the structural descriptor-mobility relationship of representative tripeptides in capillary zone electrophoresis (CZE) with the change of such separation parameters as pH, applied voltage and separation length in respect to their influence on electrophoretic migration properties. At the present stage of the work, the ionic charge was considered as structural descriptor. A multivariable linear regression (MLR) model and a back-propagation artificial neural network (BP-ANN) were applied to predict the electrophoretic mobilities of the model tripeptides with non-polar, polar, positively charged, negatively charged and aromatic *R* group characteristics. Here we present a comprehensive analysis on electrophoretic mobilities measured at pHs 2.5, 4.5, 7.5 and 9.5 at two different capillary lengths of 10 cm and 30 cm, as well as four applied electric field strengths ranging from 100 to 400 V/cm to teach and evaluate our mobility predicting models. The anticipated mobilities predicted by MLR and BP-ANN were compared to each other and to the experimental data, respectively. The BP-ANN model resulted in considerable higher precision in predictability that of the MLR method.

Correspondence: A. Guttman, Horváth Laboratory of Bioseparation Sciences, Institute of Analytical Chemistry, Innsbruck, Austria. E-mail: Andras.Guttman@uibk.ac.at

Keywords: Back-propagation artificial neural network, Capillary zone electrophoresis, Multivariable linear regression, Peptide mobility modeling

INTRODUCTION

Peptides are among the most important biomolecules, being responsible for cellular structure and function. They play a decisive role in regulation and control of many vitally important processes in all living organisms, acting, e.g., as hormones, neurotransmitters, immunomodulators, coenzymes, enzyme substrates and inhibitors, receptor ligands, drugs, toxins, and antibiotics.[1] In recent proteomics endeavors, comprehensive analysis of a proteome and/or peptidome represents the main directions to get a better understanding of the molecular bases of biological processes. In drug target discovery the importance of peptides is ever increasing, since both the structure and function of many proteins are identified by their peptide fragments.[2]

Capillary electrophoresis (CE) is of high importance in the analysis, isolation and characterization of peptides. This is attributed to its simplicity, low running cost, high separation efficiency, small sample volume and speed of separation. Furthermore, CE has proven to be a very effective technique for the measurement of physicochemical characteristics of peptides, such as molecular mass, charge state and electromigration properties.[3,4]

For many years, introduction of models to predict and optimize electromigration properties of proteins and peptides represented a challenge. The main parameter in CE separation of peptides, especially when using low ionic strength buffers, is their electrophoretic mobility. Based on the Stoke's Law, electrophoretic mobility (μ_{ef}) in CE can be given as:

$$\mu_{ef} = \frac{q}{6\pi\eta r} \tag{1}$$

where r the effective ion radius, q is the charge and η is the solution viscosity.

The development of quantitative structural descriptor-mobility models (QSMR), which correlate mobility, mass and charge of peptides, offers a powerful tool, not only for predicting electrophoretic mobility, but also for optimizing CE separations, studying structural modifications (e.g. phosphorylation, glycosylation, deamidation, etc.) and for the investigation of surface charge characteristics and conformation.[5] In the literature the semi-empirical Offord model[6] was frequently reported to provide satisfactory results in predicting electrophoretic mobilities.[7] However it should be mentioned, that electrophoretic

mobilities of peptides predicted by QSMR models that rely only on charge and mass data have limited accuracy when applied to a range of different peptides.[8] Recently non-linear modeling approaches, such as artificial neural networks (ANN) were effectively used to address the reported shortcomings of the semi-empirical models.[9–12] ANNs are computational models consisting of elements, connected and ordered in layers, capable of processing information. Machine learning techniques like ANNs for designing QSMR models have several advantages over semi-empirical approaches including the capability of self-learning and modeling complex data without the need for a detailed understanding of the underlying phenomena.

A detailed description of the theory behind the application of artificial neural networks for mobility prediction has been adequately described elsewhere.[13] Briefly, artificial neural networks (ANN) are based on a simplified mathematical model of the information processing elements of the human brain. On the other hand, an ANN is not more than a nonlinear regression model and therefore it can be efficiently used as a general approximation function, e.g., the description of structure-CZE parameters-mobility relationship. The neuron model contains a weight factor for each of its inputs. The output of the neuron is calculated from the weighted sum of the inputs by the neuron transfer function. The most frequently used transfer functions are the sigmoid $\left(\frac{1}{1+e^{-x}}\right)$ and the $\tanh(x)$ functions. In a usual feed-forward network, each neuron transfers the information to all neurons of the following layer. The input values must be normalized to the range of 0–1, while the outputs are denormalized. Learning basically means searching the weight factors by nonlinear optimization. One of these methods is the back-propagation algorithm which minimizes sum of the squared output errors using a gradient type approach layer by layer. An ANN consisting of an input, a hidden and an output layer can provide arbitrarily accurate approximation by using an adequate number of hidden neurons. Apparently, ANNs can be universally used to solve almost any approximation problem. It is very important, that the teaching patterns should cover the full range of independent and dependent variables. When the number of the hidden neurons is adequate, the regression model is accurate. If the number of the neurons is too high, the generalization ability of the ANN gets compromised. The learning process is relatively slow; however the evaluation of the model is fast.

The main goal of this work was to investigate and model the structural descriptor-mobility relationship for capillary zone electrophoresis (CZE) of tripeptides. The structure can be adequately represented by suitably selected structural descriptors. At this stage of the work the ionic charge was selected as first structural descriptor,

since it can be readily calculated for the studied tripeptides. Separation parameter changes were carefully designed in respect to their influence on electrophoretic migration. Both the multivariable linear regression (MLR) method and a back-propagation artificial neural network (BP-ANN) were applied to electrophoretic mobility prediction of the model peptides with non-polar, polar, positively and negatively charged as well as aromatic *R* group characteristics.

EXPERIMENTAL

Chemicals

Tripeptides of alanine (AAA), lysine (KKK), and tyrosine (YYY) were purchased from Sigma-Aldrich (St. Louis, MO). Tripeptides of aspartic acid (DDD) and serine (SSS) were from Bachem (Bubendorf, Switzerland). The charges of the tripeptides were calculated from the pK values of the protonated groups by the Pallas software (Comgenex, Budapest, Hungary). A mixture of mesityl oxide (2.5%) and benzyl alcohol (2.5%), both from Sigma-Aldrich, was used as combined EOF marker. HCl and NaOH (Sigma-Aldrich) were used for conditioning solutions and to adjust the pH of the background electrolytes. Acetic acid, phosphoric acid and CHES were used as running buffers components (all from Sigma-Aldrich). All solutions were prepared with HPLC grade water (Sigma-Aldrich) and all chemicals used were of analytical grade.

Instrumentation

All experiments were carried out on a Beckman P/ACE MDQ capillary electrophoresis instrument (Beckman Coulter, Fullerton, CA), using bare fused-silica capillaries (Polymicro Technologies, Tucson, AZ; 360 μm OD, 50 μm I.D.) of a total length of 40 cm. Effective separation lengths were 30 cm and 10 cm, respectively. Detection wavelength was 214 nm and separation temperature was set at 25°C.

Separation Conditions

The electrophoretic migration properties of the tripeptide samples were investigated at four different pHs. Buffers for pH 2.5 and 7.5 were prepared using phosphoric acid. The pH 4.5 buffer was prepared from acetic acid and the pH 9.5 buffer was prepared from CHES. All running

buffers were 30 mM in concentration. The investigated tripeptide samples were chosen to represent the five main amino acid characteristics of apolar (alanine tripeptide), polar (serine tripeptide), acidic (aspartic acid tripeptide), basic (lysine tripeptide) and aromatic (tyrosine tripeptide). Four different voltages of 4, 8, 12 and 16 kV were applied during the study. Every tripeptide sample was analyzed in four different pH background electrolytes at four different electric fields and two effective capillary lengths (10 and 30 cm). Samples were injected by pressure (1 psi for 4 s) for both separation lengths. Each experiment was done in triplicates. Before every set of injections, the capillary was conditioned by rinsing with water and 1 M NaOH, 5 min each, followed by water and buffer rinse. The capillary was pressure-rinsed by the relevant background electrolyte solution between the injections. The peak of the reference EOF marker (injected for 4 s at 1 psi) was used to calculate the electrophoretic mobilities of the tripeptides.

RESULTS AND DISCUSSION

The electrophoretic migration times and electroosmotic flow (EOF) values at the four different pHs applying four different voltages with the sample test mixture of all five tripeptides is listed in Table 1. As one can follow, at pH 2.5 the elution order was lysine, alanine, serine, tyrosine and aspartic acid tripeptides. Increasing the pH to 4.5 resulted in co-migration of three of the five tripeptides: alanine, serine and tyrosine tripeptides followed by the aspartic acid tripeptide at around 20 minutes. At pH 7.5 the migration times of all tripeptides decreased with increasing resolution and the migration order changed to lysine, alanine, tyrosine, serine and aspartic acid. When the running buffer pH was further increased to 9.5, the resolution became poor again, resulting in overlapping of the alanine, serine and tyrosine tripeptides, followed by the aspartic acid at 10 min.

A representative example of the capillary zone electrophoresis separation of the 5 tripeptides attained at pH 7.5 is shown in Figure 1. The migration order was: KKK (lysine), AAA (alanine), YYY (tyrosine), SSS (serine) and DDD (aspartic acid) tripeptides. Please note that at this stage the goal of this work was to study the effects of background electrolyte pH, separation voltage and separation length on the electromigration properties, not to attain the best possible peak shapes. The average standard deviations of migration times were quite adequate for all the samples due to the carefully designed conditioning of the capillary. The overall RSD values were around 1%. For all the analytes, the average standard deviations of the migration times were: 1.54% for the alanine tripeptide, 1.02% for the aspartic acid tripeptide,

Table 1. Migration times of the lysine, alanine, serine, tyrosine and aspartic acid tripeptides and the EOF marker at pH 2.5, 4.5, 7.5 and 9.5 with 4, 8, 12 and 16 V applied voltages. Other separation conditions were the same as in Figure 1

	Voltage (kV)	AAA (sec)	DDD (sec)	KKK (sec)	SSS (sec)	YYY (sec)	EOF marker (sec)
pH 2.5	4	453.36 ± 3.25	4639.17 ± 57.36	233.57 ± 1.96	509.08 ± 5.04	764.17 ± 5.37	5243.41 ± 83.31
	8	234.01 ± 0.71	2355.92 ± 19.84	118.19 ± 0.10	268.61 ± 4.48	389.31 ± 4.18	2666.50 ± 52.97
	12	157.50 ± 0.28	1550.51 ± 8.91	79.61 ± 0.15	179.99 ± 0.08	247.66 ± 1.97	1690.51 ± 54.12
	16	121.87 ± 0.89	1227.35 ± 11.07	60.84 ± 0.82	139.66 ± 4.48	184.75 ± 2.66	1255.96 ± 77.76
pH 4.5	4	295.69 ± 2.65	3015.20 ± 41.32	148.75 ± 2.31	302.09 ± 1.28	329.86 ± 7.43	326.14 ± 0.88
	8	148.68 ± 0.17	1617.96 ± 10.64	76.81 ± 0.16	152.14 ± 0.50	166.17 ± 1.68	167.08 ± 0.33
	12	99.36 ± 0.47	1205.05 ± 14.71	51.35 ± 0.29	101.24 ± 0.48	110.54 ± 2.28	112.95 ± 0.14
	16	74.15 ± 0.18	1098.48 ± 6.65	39.85 ± 0.24	75.57 ± 0.77	83.93 ± 0.63	85.74 ± 0.56
pH 7.5	4	230.19 ± 1.16	2752.55 ± 21.44	156.77 ± 0.57	283.28 ± 3.47	255.29 ± 2.76	250.08 ± 1.43
	8	182.77 ± 1.44	3735.55 ± 49.53	105.36 ± 0.82	258.94 ± 1.05	237.42 ± 2.34	160.72 ± 2.69
	12	77.44 ± 0.22	1967.55 ± 8.98	82.01 ± 0.81	93.42 ± 0.41	85.01 ± 0.28	119.89 ± 0.23
	16	59.73 ± 1.67	1259.55 ± 9.12	53.52 ± 0.21	71.47 ± 1.72	63.02 ± 0.26	51.67 ± 0.09
pH 9.5	4	279.37 ± 4.66	598.44 ± 9.11	181.38 ± 1.04	491.22 ± 2.79	271.44 ± 10.93	420.36 ± 0.97
	8	186.56 ± 5.77	249.73 ± 0.94	75.68 ± 1.43	202.10 ± 6.63	175.54 ± 2.19	104.66 ± 5.05
	12	80.86 ± 0.49	162.87 ± 0.89	44.16 ± 0.25	85.51 ± 0.22	73.83 ± 0.18	54.31 ± 0.22
	16	67.76 ± 0.78	241.72 ± 3.04	36.99 ± 0.59	71.45 ± 0.77	57.36 ± 0.30	46.40 ± 1.06

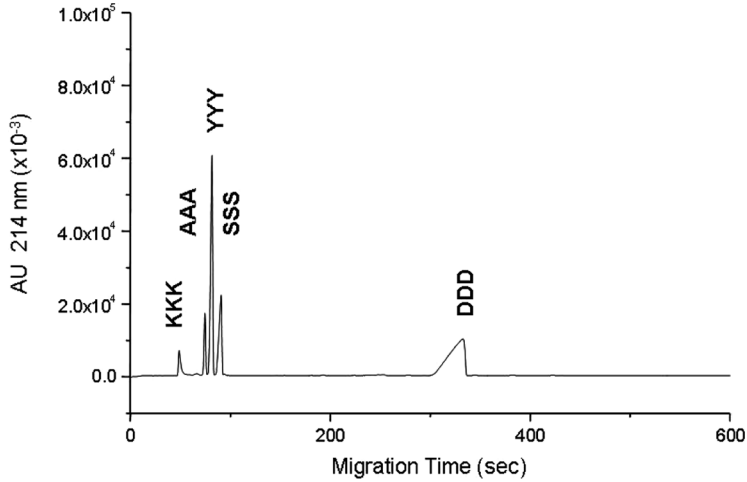


Figure 1. Capillary zone electrophoresis separation of a tripeptide test mixture in 30 mM phosphate buffer (pH 7.5). Peaks: lysine (KKK), alanine (AAA), tyrosine (YYY), serine (SSS) and aspartic acid (DDD) tripeptides. Conditions: 50 μm i.d. bare fused silica capillary with 10 cm effective length; Applied voltage: 12 kV, Temperature: 25°C. Injection: 1 psi/4 sec.

0.86% for the lysine tripeptide, 0.99% for the serine tripeptide, 1.08% for the tyrosine tripeptide and 1.41% for the EOF marker.

Mobility and Charge Calculation

The first statistical momentum (m_1) of the tripeptide peaks was considered to be the migration time, and was calculated as:

$$m_1 = \frac{\sum (t \cdot AU)}{\sum AU} \tag{2}$$

where t is the elapsed time of the separation and AU is the corresponding UV absorbance value. Peptide mobilities for all the test mixture components were calculated as:

$$\mu_{ef} = \frac{v_{\text{sample}} - v_{\text{EOF}}}{E} \tag{3}$$

where v_{sample} represents the migration velocity of the samples, v_{EOF} the migration velocity of the EOF marker, and E stands for the applied electric field strength.

Development of the structural descriptor-mobility related model required determination of peptide charges at each pH value. The molar fractions of the single ($i = 1$), double ($i = 2$), m times ($i = m$) protonated forms of the tripeptides were derived by:

$$\phi_{i(i=1,2,3,\ldots,m)} = \frac{[H^+]^i \prod_{n=1}^{i} K_n}{1 + \sum_{k=1}^{m} [H^+]^k \prod_{n=1}^{k} K_n} \tag{4}$$

where K_n is the dissociation constant for the nth protonable group.

The molar fractions of the completely deprotonated form of the tripeptides were determined as:

$$\phi_0 = \frac{1}{1 + \sum_{k=1}^{m} [H^+]^k \prod_{n=1}^{k} K_n} \tag{5}$$

The average charges of the sample tripeptides at each pH were calculated as:

$$q = \sum_{i=0}^{m} \phi_i \cdot q_i \tag{6}$$

where q_i is the charge corresponding to the i-times protonated form of the tripeptides. The calculated charge states of the tripeptides as the function of pH is shown in Figure 2.

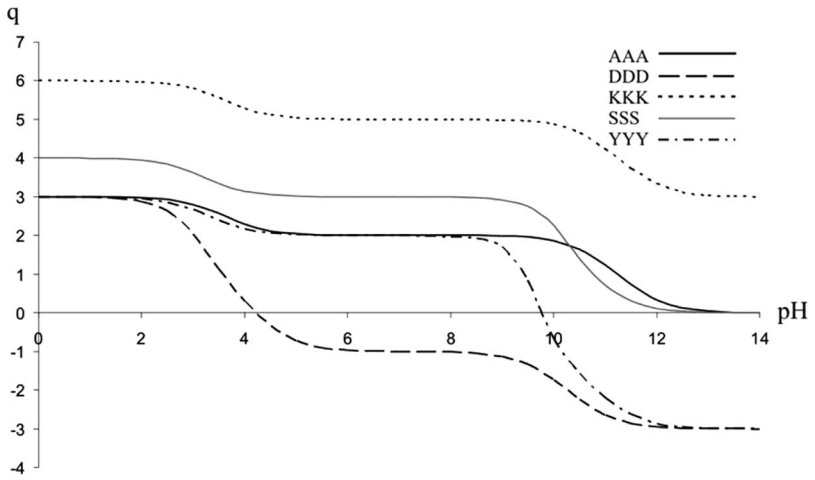


Figure 2. Calculated charges of the tripeptide test compounds as a function of the pH. AAA: alanine tripeptide, DDD: aspartic acid tripeptide, KKK: lysine tripeptide, SSS: serine tripeptide, YYY: tyrosine tripeptide.

Electrophoretic Mobility Prediction Models

In order to achieve more accurate and higher prediction performance than available with the semi-empirical models such as the Offord's equation,[6] Grossman's equation[14] and Compton's equation,[15] more complex empirical models were developed. Both a multivariable linear regression (MLR) method and a back-propagation artificial neural network (BP-ANN) model were both applied to predict the

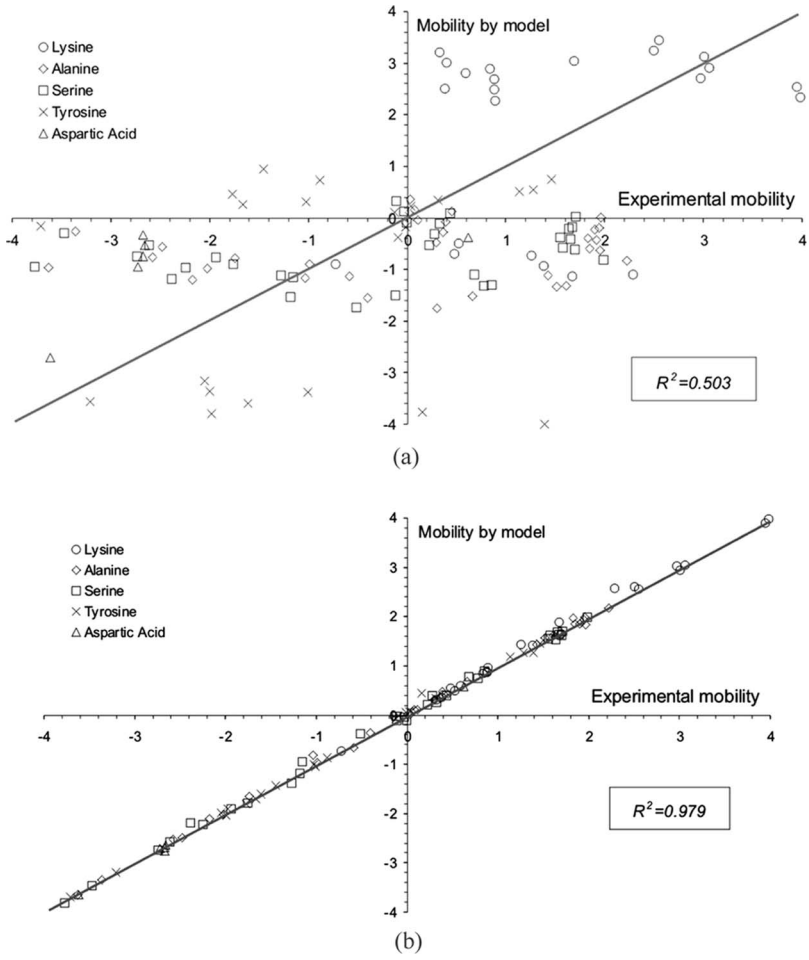
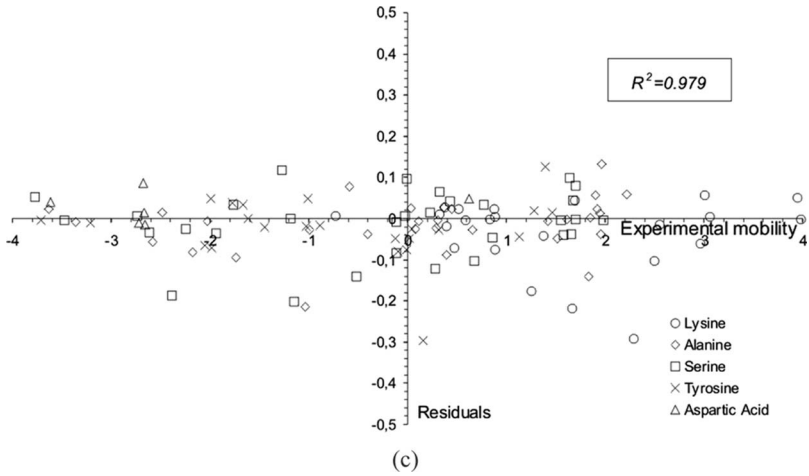


Figure 3. Correlation diagrams of the linear model (a), artificial neural network model (b) and the residuals of the artificial neural network model (c). Mobilities are given in 10^{-4} cm^2/Vs.



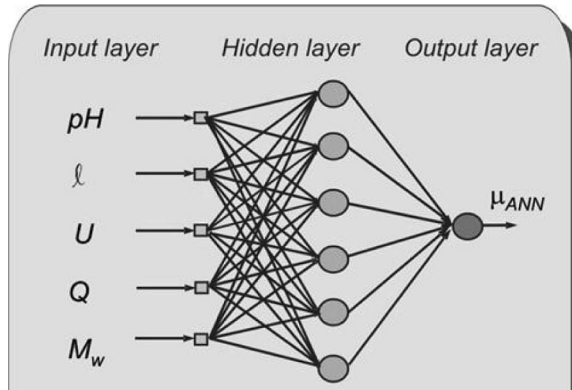
(c)

Figure 3. Continued

electrophoretic mobilities of the test compounds with non-polar, polar, aromatic, as well as positively and negatively charged *R* groups. For mobility prediction, the following input parameters were selected: pH, *l* (effective capillary length), *U* (applied voltage), *Q* (peptide charge) and *Mw* (molecular weight), both for the linear and the neural network models.

The linear model was a simple multivariable one with its parameters determined by linear regression using Matlab (Natick, MA). This model did not give accurate modeling performance because of the highly nonlinear characteristics of the studied relationship as depicted in the correlation diagram of Figure 3(a). The average deviation of the multivariable linear regression modeled mobility values from the experimentally determined values was 208.5%. Therefore, we considered that the predictive capability of this model was not satisfactory in case of complex sample mixtures consisting of peptides with similar structures. The correlation coefficient and the residuals of the artificial neural network model, on the other hand, provided reasonably good prediction in spite of its relatively simple structure, as delineated in Figure 3(b) and 3(c).

The structure of our artificial neural network model is given in the upper panel of Figure 4, and its main parameters are summarized in the lower panel of the same figure. A sigmoid transfer function was applied for the neurons in the hidden layer. The network was trained by a usual back-propagation algorithm implemented in Matlab. The measured data was divided into training and validation sets. The number of hidden layer neurons was increased until no further significant performance

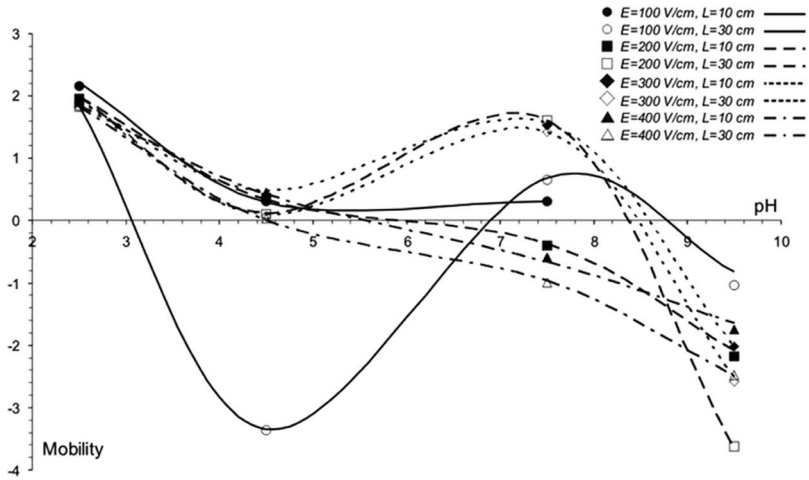


Number of iterations	1000
Learning rate	0.1
Number of nodes in the input layer	5
Number of nodes in the hidden layer	15
Number of nodes in the output layer	1

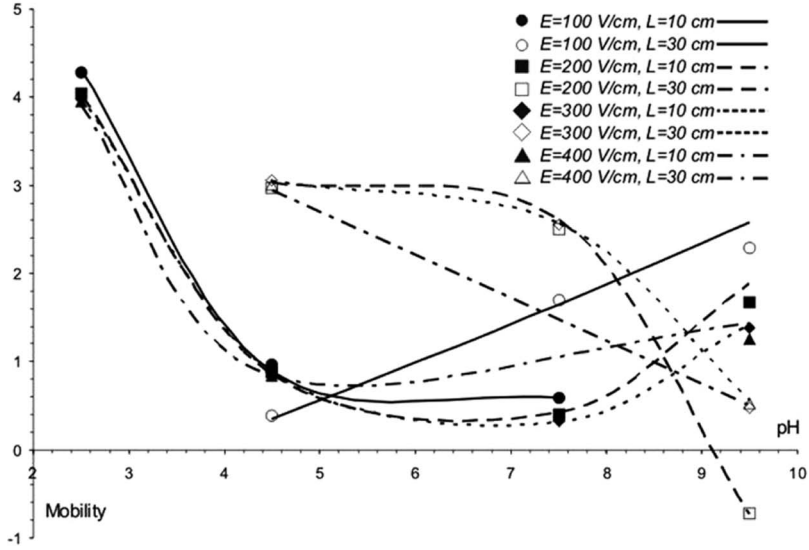
Figure 4. Structure (upper panel) and parameters (lower panel) of the ANN model. Q represents the charge of the sample, l is the effective capillary length, U is the applied voltage and Mw is the molecular weight.

improvement was obtained. The number of training periods was selected at the minimum of the validation curve to avoid ANN overtraining.

Please note that in the following diagrams the validation data set was applied. The structure-CZE parameters-mobility prediction capabilities of the artificial neural network model are demonstrated in Figures 5(a)–(c). The predictions are indicated by the lines, while the dots represent the experimental data. Figure 5(a) delineates the electrophoretic mobility versus pH relationship for the alanine tripeptide at different applied electric field strengths. The results follow the expected behavior. Figure 5(b) depicts the same relationship for the lysine tripeptide, while in the case of serine tripeptide the relationship is shown in Figure 5(c). Figures 5(a)–(c), all suggest that the studied relationship was rather complex and nonlinear. This complexity can be accounted for the stepwise changes of ionic charges, the interactions between the different tripeptides in the sample and for other non studied structural characteristic factors.

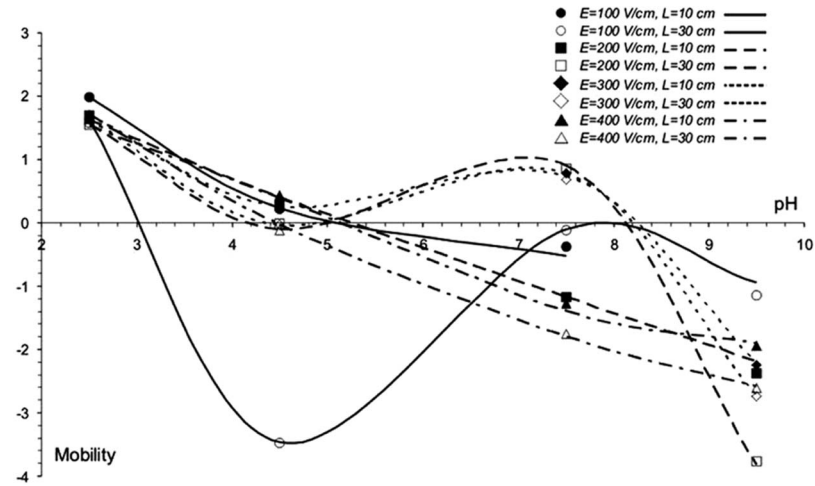


(a)



(b)

Figure 5. Relationship between the electrophoretic mobility and the background electrolyte pH of the model tripeptides as determined by the artificial neural network model (lines) and the experimentally obtained data (point markers). Alanine tripeptide (a), lysine tripeptide (b), serine tripeptide (c).



(c)

Figure 5. Continued

At the same time the neural model reflected this function with good accuracy. The average deviation of the ANN modeled mobility values from the experimentally determined mobilities was 1.95%, exhibiting apparently good correlation.

The most important advantage of artificial neural networks over regression analyses was their ability to allow flexible mapping of the selected features by manipulating their functional dependence implicitly. Developing networks and comparing them with the MLR models provided us the opportunity to investigate the nonlinear characteristics of the electrophoretic mobility dependences of the peptides on structural descriptors and pH.

CONCLUSIONS

The carefully selected tripeptide test mixture of Lysine (KKK), alanine (AAA), serine (SSS), tyrosine (YYY) and aspartic acid (DDD), enabled study and modeling of the ionic charge-CZE parameters-mobility relationship. The ionic charge is considered as one of the structural descriptors. The basis for the selection was to represent all major amino acid types (*R*-groups). The performance of the multivariable linear and the artificial neural network model was compared based on the experimental data. The back-propagation artificial neural network model

showed superior performance. As a first approximation, the inadequacy of the linear model was considered to be due to the highly nonlinear effects of pH and some possible structural parameters. Continuation of this work will include extending the experimental database and including available literature data sources. Additionally, searching for other, more adequate descriptors for the structural influence on the electric field mediated differential migration of various model peptides.

ACKNOWLEDGMENT

This work was supported by the Marie Curie Chair (006733) of the European Commission, the Tyrolean Government's HITT–38 HLBS grant and the Hungarian TéT A-10/2006 grant.

REFERENCES

1. Lodish, H.; Berk, A.; Matsudaira, P.; Kaiser, C.A.; Krieger, M.; Scott, M.P.; Zipursky, L.; Darnell, J. *Molecular Cell Biology*; 5th ed.; Freeman, W.H., New York, 2003.
2. Kasicka, V. Recent developments in capillary electrophoresis and capillary electrochromatography of peptides. Electrophoresis **2006**, *27*(1), 142–175.
3. Dolnik, V. Capillary electrophoresis of proteins 2003–2005. Electrophoresis **2006**, *27*(1), 126–141.
4. Hutterer, K.; Dolnik, V. Capillary electrophoresis of proteins 2001–2003. Electrophoresis *2003*, *24*(22–23), 3998–4012.
5. Adamson, N.J.; Reynolds, E.C. Rules relating electrophoretic mobility, charge and molecular size of peptides and proteins. Journal of Chromatography, B: Biomedical Sciences and Applications **1997**, *699*(1–2), 133–147.
6. Offord, R.E. Electrophoretic mobilities of peptides on paper and their use in the determination of amide groups. Nature **1966**, *211*(5049), 591–593.
7. Cifuentes, A.; Poppe, H. Behavior of peptides in capillary electrophoresis. Effect of peptide charge, mass, and structure. Electrophoresis. **1997**, *18*(12–13), 2362–2376.
8. Tessier, B.; Blanchard, F.; Vanderesse, R.; Harscoat, C.; Marc, I. Applicability of predictive models to the peptide mobility analysis by capillary electrophoresis-electrospray mass spectrometry. Journal of Chromatography, A **2004**, *1024*(1–2), 255–266.
9. Jalali-Heravi, M.; Shen, Y.; Hassanisadi, M.; Khaledi, M.G. Artificial neural network modeling of peptide mobility and peptide mapping in capillary zone electrophoresis. Journal of Chromatography, A **2005**, *1096*(1–2), 58–68.
10. Yu, K.; Cheng, Y. Machine learning techniques for the prediction of the peptide mobility in capillary zone electrophoresis. Talanta **2007**, *71*(2), 676–682.

11. Janini, G.M.; Metral, C.J.; Issaq, H.J. Peptide mapping by capillary zone electrophoresis: How close is theoretical simulation to experimental determination. Journal of Chromatography, A **2001**, *924*(1–2), 291–306.
12. Ma, W.; Luan, F.; Zhang, H.; Zhang, X.; Liu, M.; Hu, Z.; Fan, B. Accurate quantitative structure-property relationship model of mobilities of peptides in capillary zone electrophoresis. Analyst **2006**, *131*(11), 1254–1260.
13. Hopke, P.K.; Song, X.-H. The chemical mass balance as a multivariate calibration problem. Chemometrics and Intelligent Laboratory Systems **1997**, *37*(1), 5–14.
14. Grossman, P.D.; Colburn, J.C.; Lauer, H.H. A semiempirical model for the electrophoretic mobilities of peptides in free-solution capillary electrophoresis. Analytical Biochemistry **1989**, *179*(1), 28–33.
15. Compton, B.J. Electrophoretic mobility modeling of proteins in free zone capillary electrophoresis and its application to monoclonal antibody microheterogeneity analysis. Journal of Chromatography **1991**, *559*(1–2), 357–366.

Manuscript 6312K