Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Chromatography A, 1096 (2005) 58–68
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Artificial neural network modeling of peptide mobility and peptide mapping in capillary zone electrophoresis

Mehdi Jalali-Heravi [a,*], Yang Shen [a], Mostafa Hassanisadi [b], Morteza G Khaledi [a,**]

[a] *Department of Chemistry, North Carolina State University, NC 27695-8204, USA*
[b] *Department of Chemistry, Sharif University of Technology, P.O. Box 11365-9516, Tehran, Iran*

Available online 10 October 2005

Abstract

Recently, we have developed an artificial neural network model, which was able to predict accurately the electrophoretic mobilities of relatively small peptides. To examine the robustness of this methodology, a 3-3-1 back-propagation artificial neural network (BP-ANN) model was developed using the same inputs as the previous model, which were the Offord's charge over mass term ($Q/M^{2/3}$), corrected steric substituent constant ($E_{s,c}$) and molar refractivity (MR). The data set consisted of 102 peptides with a larger range of size than that of our earlier report – up to 42 amino acid residues as compared to 13 amino acids in the initial study – that also included highly charged and hydrophobic peptides. The entire data set was obtained from the published result by Janini and co-workers. The results of this model are compared with those obtained using multiple linear regressions (MLR) model developed in this work and the multi-variable model released by Janini et al. Better predictive ability of the BP-ANN model over the MLR indicates the non-linear characteristics of the electrophoretic mobility of peptides. The present model exhibits better robustness than the MLR models in predicting CZE mobilities of a diverse data set at different experimental conditions. To explore the utility of the ANN model in simulation of the CZE peptide maps, the profiles for the endoproteinase digests of melittin, glucagon and horse cytochrome *C* is studied in the present work.

Keywords: Peptide mapping; Capillary zone electrophoresis; Electrophoretic mobility; Artificial neural networks

1. Introduction

Peptide mapping is a widely used technique for characterization of protein structure that involves digestion of a protein through enzymatic or chemical means and subsequent separation and detection of the resultant peptide mixture. The peptide maps serve as "fingerprints" that can be applied to rapid protein identification and the detection of post-translational modifications.

Capillary liquid chromatography (combined with MS/MS) is currently one of the most commonly used techniques for peptide mapping [1,2]. While this method provides excellent resolution, it is often slow and generally consumes relatively large quantities of peptides. Capillary zone electrophoresis (CZE) has received considerable attention as peptide mapping method because of its high speed and high resolution for peptide analysis and also its small sample size requirement. Among other advantages, CZE can be multiplexed and automated so that many experiments can be performed in parallel. However, electrophoretic peptide profiles obtained are sometimes very complex owing to the complicated nature of the samples. Generally, CE separation of peptides is more successful than proteins because the smaller molecules tend to interact less compared with the capillary wall.

The key parameter for separation of peptides, especially in low ionic strength buffers, is their electrophoretic mobilities. This parameter can be converted to migration time and a CZE electropherogram can be simulated using a Gaussian function. Therefore, calculation/prediction of this parameter is very useful in peptide mapping studies.

Several empirical models, based on Stoke's law, have been developed for the prediction of electrophoretic mobilities from the charge to size ratio [3–9]. The accuracy of these predictive models are limited due to three factors: (1) the general assumption that the relationship between mobility and charge to size ratio follows a linear trend for all peptides can be inaccurate; (2) the lack of accurate pK_a values for the ionizable

* Corresponding author. On sabbatical leave from Department of Chemistry, Sharif University of Technology, P.O. Box 11365-9516, Tehran, Iran. Tel.: +98 21 6005718; fax: +98 21 6012983.
** Co-corresponding author.
E-mail address: jalali@sharif.edu (M. Jalali-Heravi).

0021-9673/$ – see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.chroma.2005.09.043

amino acids chains and the end groups in peptides result in inaccurate calculation of peptide charges and (3) molar mass is often used to describe peptide size, however, the exact effect of molar mass on electrophoretic mobility is yet to be determined. The main difference between the various reported empirical models lies in the dependence of mobility on molar mass of peptides.

Janini et al. have obtained the electrophoretic mobility of 58 peptides ranging in size from 2 to 39 amino acids and charge from 0.65 to 7.82 [10]. Based on this data set, they concluded that the Offord model is the best predictor of mobility, but it fails when applied to hydrophobic and highly charged peptides. These researchers also showed that peptide electrophoretic mobility cannot be successfully predicted with reasonable degree of accuracy for all different categories of peptides by relying on two-parameter models, namely charge (Q) and size dependence—size is usually expressed in terms of the number of amino acids residues, N, or molar mass, M [10].

In two other efforts, Janini and coworkers have developed a multi-variable computer model, based on a closest-neighbor algorithm for the prediction of the electrophoretic mobilities of peptides [11,12]. A computer algorithm was developed that matches the physical parameters of unknown peptides to their closest-neighbor in the peptide set for more accurate estimation of their mobilities. Their model is based on assumption that the electrophoretic mobilities can be calculated by a product of three functions representing peptide charge, length and width [12]. In this model, the number of amino acid residues was considered as a measure of peptide length and the width of a protein is summation of the residue widths of its amino acids. The residue width for each amino acid was assumed to be proportional to the relative mass of the residue, being 1 for glycine and 130 for tryptophane. Although, these researchers have obtained excellent results, but the main drawback of their model is that all functions appearing in the model are data dependent. In other words, their model is based on a closest-neighbor algorithm and is not robust.

We have recently developed an artificial neural network model for the following three objectives: (1) to explore the linear/non-linear characteristics of the electrophoretic mobilities of peptides; (2) to assess the predictive abilities of the various reported models; and (3) to investigate the effects of other structural factors, in addition to charge and size, on electrophoretic mobility [13]. The ANN model showed a significant improvement in the predictive ability over the empirical models and the multiple linear regression (MLR) treatment [13]. This improvement was especially pronounced for the peptides of higher charges that contain basic amino acids arginine, histidine and lysine. This model was based on a data set consisting of the electrophoretic mobilities for 125 peptides ranging in size between 2 and 14 amino acids.

The main aims of the present work, in addition to the above mentioned objectives, were: (1) evaluating the robustness of the generated model, (2) comparison of the model with the multivariable model developed by Janini et al., and (3) investigating the utility of the model in simulation of the maps of some peptides. However, long range goal of the development of such multivariate models is their applications as a tool in the construction of a data base of peptide maps. A successful model should be capable of providing accurate prediction of migration behavior in CZE for a wide range of compounds and at various experimental conditions. Therefore, in the present work, a data set that has been reported by Janini et al. was chosen [12].

2. Experimental

2.1. Data set

Development of the multiple linear regression (MLR) and artificial neural networks (ANN) in the present work relies on a data set taken from reference [12]. This data set (Table 1) consists of 102 peptides ranging in size from 2 to 42 amino acid residues, which varies in charge from 0.65 to 16. The data set includes 18 dipeptides, 32 peptides with five or less amino acid residues and 72 peptides with six or more amino acid residues. The electrophoretic mobilities of these peptides have been determined in CZE using a 50 mM phosphate buffer at pH 2.5, 22 °C and a capillary coated with 10% polyacrylamide. It should be noted that these experimental conditions are different from those we used for the development of the ANN model in our previous work [13]. According to the authors, the use of polyacrylamide coated capillary provided stability and migration time reproducibility throughout the experiments to within 1% RSD [12]. However, in order to determine the uncertainty of the experimental electrophoretic mobilities, Janini et al. have measured the mobilities of six peptides and the reference standard using two independent columns with different buffer preparations in separate days. They obtained an average RSD of 2.34%, for the measured electrophoretic mobilities that had a range of $4.73 \times 10^{-5}\,cm^2\,s^{-1}\,V^{-1}$ (for FIGITEAAANLVPMVATV) to $33.03 \times 10^{-5}\,cm^2\,s^{-1}\,V^{-1}$ (for KKKKK).

To apply the ANN modeling for the Janini et al. results, the data set was randomly divided into three groups of training, test and validation sets consisting of 70, 20 and 12 peptides, respectively. The training set was used for the model generation. The test set plays a different role in the cases of the MLR and the ANN models. For the ANN model, this set was used for early stopping to optimize learning iteration size and avoid overtraining. The validation set was used to assess the accuracy of the ANN predictions. On the other hand, in the case of the MLR model, the test set combined with the validation set was used to evaluate the model. As can be seen from Table 1, the peptides in the test and validation sets were chosen in a way that adequately represents the training set in terms of size, charge and composition.

2.2. Regression analysis

The main aim of the present work was to investigate the robustness of our previous quantitative structure–mobility relationships (QSMR) model. Therefore, as first step for developing the regression model, the three descriptors appearing in our previous QSMR models were chosen as the most suitable param-

Table 1
Experimental and calculated values of electrophoretic mobilities using MLR and ANN models together with the values of the descriptors

No.	Peptide sequence	Descriptors[a]			Exp	MLR		ANN	
		QM	$E_{s,c}$	MR	$\mu_{ef} \times 10^5$	$\mu_{ef} \times 10^5$	Percent deviation	$\mu_{ef} \times 10^5$	Percent deviation
Training set									
1	AA	0.0283	0.00	11.3	18.77	17.85	4.91	19.84	−5.68
2	DD	0.0165	−1.56	23.2	10.31	13.02	−26.33	11.92	−15.58
3	EE	0.0192	−1.24	32.5	12.52	14.22	−13.55	13.58	−8.46
4	FA	0.0218	−0.70	35.7	14.86	15.40	−3.63	15.20	−2.29
5	FF	0.0181	−1.40	60.0	12.81	13.99	−9.25	13.06	−1.95
6	FG	0.0227	−0.50	31.0	15.16	15.76	−3.98	15.57	−2.69
7	FL	0.0196	−1.94	49.6	13.33	14.30	−7.26	13.51	−1.35
8	FV	0.0202	−1.79	45.0	13.90	14.56	−4.77	13.89	0.05
9	GG	0.0321	0.40	2.1	21.70	19.36	10.79	21.70	0.00
10	HG	0.0515	−0.46	24.8	27.04	26.67	1.36	29.35	−8.53
11	LL	0.0213	−2.48	39.2	14.55	14.73	−1.21	14.61	−0.43
12	PG	0.0269	0.20	15.0	18.43	17.43	5.41	18.83	−2.15
13	RK	0.0629	−1.24	55.1	32.00	31.02	3.05	30.98	3.17
14	VV	0.0231	−2.18	29.9	15.39	15.42	−0.20	16.17	−5.08
15	WW	0.0156	−1.32	79.6	11.05	13.23	−19.77	12.80	−15.84
16	YY	0.0170	−1.40	63.7	12.10	13.59	−12.34	13.29	−9.81
17	FFF	0.0140	−2.10	90.0	10.38	12.48	−20.28	11.92	−14.80
18	SSS	0.0195	−0.84	35.5	13.22	14.48	−9.55	14.34	−8.44
19	AAAA	0.0185	0.00	22.6	13.87	14.24	−2.64	13.72	1.10
20	ANSK	0.0328	−1.68	57.0	20.91	19.46	6.93	22.53	−7.76
21	HMTE	0.0283	−2.64	75.0	18.91	17.64	6.70	19.73	−4.34
22	PARR	0.0451	−1.24	79.7	27.65	24.46	11.55	27.43	0.78
23	KKKKK	0.0770	−3.10	125.3	33.03	36.45	−10.37	31.76	3.86
24	RPPGF	0.0266	−1.12	89.0	18.36	17.55	4.42	18.72	−1.95
25	YGGFL	0.0123	−2.24	83.5	9.75	11.76	−20.57	10.58	−8.49
26	GIGAVLK	0.0243	−4.16	86.9	15.50	15.76	−1.67	16.69	−7.66
27	AAGIGILTV	0.0096	−5.68	98.9	6.50	9.83	−51.20	8.78	−35.00
28	ACHGRDRRT	0.0453	−3.63	144.0	26.54	24.39	8.08	26.90	−1.37
29	AFLPWHRLF	0.0253	−5.82	212.5	16.55	16.73	−1.09	17.53	−5.93
30	MLDLQPETT	0.0071	−6.39	146.8	6.33	9.10	−43.70	8.36	−32.11
31	RPPGFSPFR	0.0273	−2.72	174.8	19.71	18.07	8.35	18.47	6.30
32	VLQELNVTV	0.0082	−8.30	145.7	6.97	8.93	−28.10	8.14	−16.80
33	VVRRYPHHE	0.0429	−6.06	199.6	27.38	23.24	15.10	25.55	6.70
34	YLSGADLNL	0.0076	−6.06	135.1	6.23	9.28	−49.00	8.39	−34.67
35	KLVVVGAAGV	0.0195	−5.82	117.8	14.10	13.72	2.67	13.32	5.55
36	KLVVVGADGV	0.0180	−6.60	123.7	13.13	12.96	1.29	12.52	4.67
37	ACLGRDRRTEE	0.0312	−5.45	172.3	20.97	18.73	10.66	21.16	−0.88
38	CRHRRRHRRGC	0.0676	−4.84	228.9	29.68	33.25	−12.04	30.72	−3.52
39	LLGRNSFEMRV	0.0234	−7.82	210.9	17.02	15.44	9.31	17.22	−1.18
40	RPKPQQFFGLM	0.0232	−5.75	225.0	16.98	16.07	5.34	17.77	−4.64
41	YAEGDVHATSK	0.0245	−5.08	159.4	17.40	16.20	6.89	18.51	−6.37
42	ACPGKDRRTGGGN	0.0316	−3.15	146.7	19.11	19.36	−1.28	22.64	−18.49
43	ACPGTDRRTGGGN	0.0235	−3.06	133.5	15.08	16.18	−7.32	16.56	−9.84
44	ACPGRNRRTEEENL	0.0273	−6.85	219.8	19.40	17.26	11.03	20.01	−3.14
45	MGGMNWRPILTIIT	0.0134	−10.45	248.6	10.20	11.17	−9.46	12.08	−18.47
46	SPALNKMFCELAKT	0.0211	−7.46	222.0	15.71	14.73	6.25	16.32	−3.90
47	VLTTGLPALISWIK	0.0139	−10.45	233.9	10.50	11.24	−7.07	11.83	−12.63
48	HRSCRRRKRRSCRHR	0.0733	−7.46	336.7	30.27	35.57	−17.51	30.68	−1.34
49	YSPALNKMCCQLAKT	0.0201	−7.46	226.7	14.90	14.41	3.27	15.38	−3.23
50	IITLEDSSNLLGRNSF	0.0118	−12.39	263.1	11.33	10.12	10.68	10.50	7.36
51	LAPPQHHLIQVGNLRV	0.0260	−11.27	273.2	15.01	15.92	−6.07	18.07	−20.40
52	LDDRNTFRRSVVVPYE	0.0232	−11.54	307.9	18.30	15.08	17.57	16.53	9.66
53	PPPGTRVRVMAIKQSQ	0.0262	−8.33	268.2	18.20	16.84	7.47	18.26	−0.34
54	TYSPALNRMFCQLAKT	0.0188	−8.69	273.5	14.77	13.97	5.44	14.66	0.74
55	DGLAPPQHRIRVEGNLR	0.0305	−10.10	284.7	18.98	18.10	4.65	20.40	−7.47
56	KSSQYIKANSKFIGITE	0.0248	−11.28	299.5	17.05	15.70	7.94	17.48	−2.53
57	LGRNSFEVCVCACPGRD	0.0183	−7.42	215.4	13.66	13.63	0.21	14.19	−3.88
58	NHQLLSPAKTGWRIFHP	0.0304	−10.02	323.1	19.42	18.40	5.24	20.28	−4.41
59	NTFRHSVVEPYEPPEVG	0.0179	−9.20	290.2	13.55	13.61	−0.43	14.52	−7.13
60	SSCMGGMNQRPILTIIT	0.0123	−10.97	251.5	10.66	10.62	0.36	10.99	−3.09

Table 1 (*Continued*)

No.	Peptide sequence	Descriptors[a]			Exp	MLR		ANN	
		QM	$E_{s,c}$	MR	$\mu_{ef} \times 10^5$	$\mu_{ef} \times 10^5$	Percent deviation	$\mu_{ef} \times 10^5$	Percent deviation
61	YKLVVVGACGVKGSALT	0.0202	−8.99	218.9	14.33	13.92	2.88	14.95	−4.36
62	YKLVVVGANGVGKSALT	0.0201	−9.77	233.4	14.36	13.78	4.04	14.92	−3.89
63	YKLVVVGARGVGKSALT	0.0267	−9.61	249.0	17.80	16.48	7.40	18.72	−5.16
64	YKLVVVGAVGVGKSALT	0.0202	−10.08	233.9	15.06	13.74	8.77	14.97	0.59
65	YNYMCNSSGMGGMNRRP	0.0182	−7.43	277.2	14.29	14.13	1.10	15.08	−5.53
66	FIGITEAAANLVPMVATV	0.0055	−11.52	248.7	4.73	7.87	−66.33	8.62	−82.22
67	VPYEPPEVGSVYHHPLQLHV	0.0219	−12.16	354.2	15.13	14.80	2.17	16.54	−9.32
68	RTHGQSHYRRRHCSRRRLHRIHRRQ	0.0708	−15.18	565.2	29.01	34.29	−18.21	28.57	1.52
69	KSSQYIKANSKFIGITEAAANLVPMVATV	0.0182	−17.93	449.9	14.21	12.52	11.92	14.54	−2.32
70	DRVIEVVQGAYRAIRHIPRRIRQGLERRIHIGPGRAFYTTKN	0.0437	−28.09	788.1	20.83	22.13	−6.24	21.80	−4.64
Test set									
71	FD	0.0173	−1.48	41.6	13.00	13.52	−4.02	12.82	1.41
72	MM	0.0195	−1.66	46.2	13.86	14.31	−3.28	14.14	−1.99
73	AAA	0.0221	0.00	17.0	14.96	15.56	−4.04	15.83	−5.78
74	RQQ	0.0322	−1.86	68.3	24.00	19.27	19.70	22.13	7.79
75	KKK	0.0703	−1.86	75.2	33.03	33.83	−2.43	31.55	4.49
76	SSQYIK	0.0227	−4.11	119.2	16.71	15.46	7.46	16.39	1.92
77	YMDGTMSQV	0.0073	−5.46	148.4	6.62	9.45	−42.69	9.04	−36.58
78	ACSGRDRRTEE	0.0316	−4.49	164.5	21.91	19.11	12.79	21.50	1.85
79	NSFCMGGMNRR	0.0241	−5.04	179.2	18.30	16.24	11.27	17.47	4.51
80	AAANLVPMVATV	0.0076	−6.65	150.4	6.15	9.23	−50.01	8.86	−44.14
81	DAEKSDICTDEY	0.0124	−7.32	173.0	9.91	11.05	−11.53	10.81	−9.07
82	GSDCTTIHCNYM	0.0143	−6.50	160.9	12.41	11.92	3.96	11.72	5.59
83	PHRERCSDSDGL	0.0295	−5.68	181.5	19.33	18.11	6.30	20.34	−5.22
84	TTIHYNYICNSS	0.0145	−8.46	202.8	10.59	11.80	−11.38	11.90	−12.41
85	HMTEVRRYPHHER	0.0469	−8.23	289.7	26.42	24.92	5.67	26.54	−0.44
86	YAEGDVHATSKPARR	0.0338	−6.32	239.1	21.38	20.04	6.25	22.15	−3.60
87	LAKTCPVRLWVDSTPP	0.0187	−8.68	258.5	15.13	13.77	9.00	14.65	3.20
88	VVRRCPHQRCSDSDGL	0.0311	−8.48	244.3	20.75	18.44	11.12	20.92	−0.82
89	YKLVVVGAAGVGKSALT	0.0204	−8.99	224.6	14.22	14.07	1.09	15.08	−6.03
90	KQIINMWQEVGKAMYAPPISGQIRRIHIGPGRAFYTTKN	0.0288	−24.13	702.2	17.78	16.91	4.92	18.39	−3.41
Validation set									
91	KKKK	0.0737	−2.48	100.2	33.03	35.17	−6.47	31.88	3.47
92	AAAAA	0.0161	0.00	28.3	12.34	13.36	−8.27	11.40	7.64
93	YGGFM	0.0121	−1.83	87.0	9.53	11.81	−23.89	9.56	−0.35
94	VISNDVCAQV	0.0072	−7.34	127.1	5.83	8.69	−48.99	6.75	−15.81
95	CRHHRRRHRRGC	0.0710	−5.50	252.7	29.68	34.54	−16.38	30.89	−4.06
96	HMTEVRHCPHHER	0.0484	−7.57	251.6	26.41	25.35	4.02	26.36	0.21
97	LAKTCPVRLWVDS	0.0210	−8.15	218.8	10.510	9.91	5.71	16.03	−52.51
98	RTHCQSHYRRRHCSR	0.0560	−7.49	308.0	28.96	28.72	0.84	27.64	4.54
99	EPPEVGSDYHHPLQLHV	0.0238	−10.06	290.1	16.91	15.60	7.77	16.62	1.73
100	KLVVVGAGDVGKSALTI	0.0198	−10.68	218.3	13.69	13.29	2.94	13.76	−0.52
101	TPPPGTRVQQSQHMTEV	0.0184	−8.44	270.6	14.17	13.86	2.15	14.02	1.08
102	FLTPKKLQCVDLHVISNDVCAQVHPQKVTK	0.0295	−20.82	494.9	18.68	16.35	12.47	18.65	0.16

[a] Definitions of descriptors are given in the text.

eters contributing to the peptides electrophoretic mobilities as [13]:

$$\mu = p\frac{Q}{M^{2/3}} + e\sum E_{s,c} + m\sum \mathrm{MR} \quad (1)$$

where $Q/M^{2/3}$ is a hybrid parameter combining the charge over mass ratio in the form of Offord empirical model, $E_{s,c}$ is the corrected steric substituent constant and MR is molar refractivity. These descriptors were chosen from a large set of parameters such as effective net charge, molar mass, number of amino acid residue, average residue mass, molecular volume, surface area, hydrophobicity, isoelectric point value, strain parameter, Z-scale and alpha-helix content [14,15]. A detailed description of the stepwise multiple linear regression procedure used for choosing these descriptors is given in our previous paper [13]. The most common method for calculating peptide charges is to use the Henderson-Hasselbach equation [16]. According to this method, the net charge of a peptide at pH 2.5 can be calculated as the sum of all charged amino acid residues and carboxy- and amino-terminals. Each arginine (R), histidine (H), lysine (K) and N-terminal contribute a charge of +1; while each aspartic acid (D) (pK_a 3.5), glutamic acid (E) (pK_a 4.5) and the C-terminal (pK_a 3.2) contributes a charge of −0.091, −0.01 and −0.166 at pH 2.5, respectively. The parameters of $E_{s,c}$ and MR for each

Table 2
Specifications of the selected MLR model

Descriptor	Notation	Coefficient	Mean effect[a]
Charge to size ratio	QM	380.080 (±17.360)	9.845
Corrected steric substituent constant	$E_{s,c}$	0.292 (±0.170)	−1.775
Molecular refractivity	MR	0.009 (±0.006)	1.514
Constant		7.009 (±0.512)	

[a] Mean effect of a descriptor is the product of its mean and regression coefficient in the MLR model.

amino acid is taken from the references [17,18]. $\sum E_{s,c}$ and $\sum$MR for each peptide were obtained using an additive model (i.e. these parameters for a peptide is simply the sum of the corresponding parameters of amino acids included in the peptide). The calculated values of these parameters for all peptides in the training, test and validation sets are given in Table 1. The best MLR model is one that has high correlation coefficient (R) and F-values, low standard deviation and high predictive ability. The specifications for the best MLR model (Eq. (1)) are presented in Table 2.

2.3. Artificial neural networks model

A detailed description of theory behind artificial neural networks has been adequately described elsewhere [19–22]. A three-layer back-propagation network with a sigmoidal transfer function was designed in the present work. This network is written in C++ in our laboratory. The three descriptors appearing in our previous QSMR model were used as input parameters for the network generation. The signals from the output layer represent the electrophoretic mobilities of the peptides. Such an ANN may be designed as 3-y-1 net to indicate the number of nodes in input, hidden and output layers, respectively. Generally, the neural network methodology has several empirically determined parameters. These include: (1) when to stop training (i.e. the number of iterations or the convergence criterion), (2) the number of hidden nodes, and (3) learning rate and momentum terms. The values of constructed ANNs parameters were optimized with the procedure that was reported in our previous works [23–25]. The initial weights were chosen randomly. The program is written in such a way that the randomized weights depend on the number of input, hidden and output nodes. Before training, the output and inputs (except for the values of the Offord model) were normalized between 0 and 1. To evaluate the performance of the ANN, standard error of calibration (SEC) and standard error of prediction (SEP) were used [26]. The number of neurons of the hidden layer with the minimum value of SEC was selected as the optimum number. Learning rate and momentum were optimized in a similar way. We have used the validation set to examine the validity of the ANN model.

2.4. Modeling of peptide maps

The authors believe that the most important aim of the present work was investigating the application of the generated model in simulation of the peptide maps of protein digests. Therefore, to test the utility of the model, we have simulated the theoretical peptide maps of the digests of melittin, glucagon and horse cytochrome *C* polypeptide and proteins (see Section 3.4).

In order to simulate an electropherogram, first the ANN calculated electrophoretic mobilities were converted to migration times using the same values for the experimental parameters as reported by Janini et al. [10]. Values of 37 and 30 cm were used for the total length of the column and injector-to-detector length, respectively. Also, a running voltage of 8 kV was used for the purpose of this conversion. To simulate the peak of each theoretical fragment, it was assumed that the area for each peak is proportional to the number of peptide bonds. It is shown that at 200 nm, the absorbance of a peptide is largely attributed to the peptide bonds while the contribution of amino acids residues can be ignored [27,28]. For the sake of convenience, same peak width was used for each simulated peak in this work, which makes the peak height proportional to the number of peptide bonds.

3. Results and discussion

To inspect the robustness of a model, one should choose a very diverse data set consisting of peptides with wide range of structural properties. Therefore, in the present work we have applied a similar strategy (i.e. using the same structural descriptors as ANN inputs) to predict the electrophoretic mobilities of a more diverse peptide set that included significantly larger, more hydrophobic, and highly charged peptides that was obtained from the paper by Janini et al. [12].

3.1. Multiple linear regression analysis

The MLR calculated values of electrophoretic mobilities of all peptides are shown in Table 1. Also, the regression results for the selected MLR model are presented in Table 2. It can be seen that the most important descriptor is the hybrid parameter of charge-to-size ratio (Offord model). This descriptor shows a mean effect of 9.845, which is the largest among the parameters appearing in the model. This is in agreement with our previous work and confirms the conclusion reached by many researchers [4,11–13]. However, a major problem with this parameter is the accuracy of the calculated charge, which is controversial in the literature [29–31]. Although Rickard et al. have claimed that at pH 2.5 a good agreement is expected between the calculated and actual charge, this is not the case for hydrophobic or highly charged peptides [30]. At this pH, for hydrophobic peptides vari-

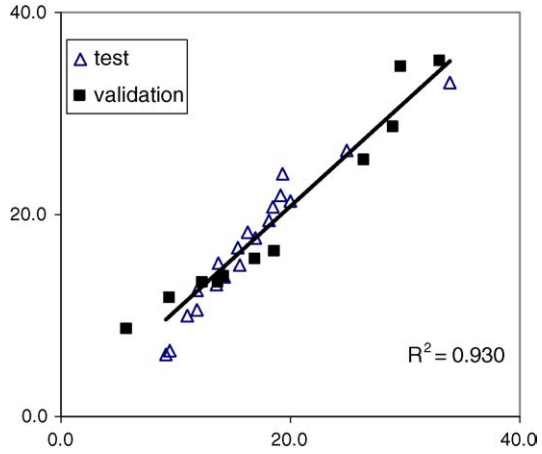

Fig. 1. Plot of the MLR calculated electrophoretic mobilities against the experimental values of peptide mobility for the test and validation sets.

ations in the ionization constants of *D* and *E* might be significant, while for highly charged peptides the calculated charges might deviate from the actual one due to mutual electrostatic interactions of charged groups in proximity of each other. Cifuentes and Poppe have presented a model for predicting pK_a values of peptides considering the mutual electrostatic interaction of charged groups, but their model is empirical in nature and time consuming [31]. Since the charge-to-mass ratio parameter plays the major role in the mechanism of electrophoretic mobilities, the MLR calculated mobilities for hydrophobic and highly charged peptides might deviate considerably from the experimental values. Inspection of Table 1 shows that this is true for most of the highly charged or hydrophobic peptides. The peptides 48, 68 and 70 with charges of 11.83, 15.83 and 12.72, respectively, are the most highly charged peptides in Table 1. The MLR calculated values for these peptides show a deviation of −17.51, −18.21 and −6.24%, respectively. On the other hand, among the peptides listed in Table 1 peptides 2, 30 and 94 show the lowest charges of 0.65, 0.73 and 0.74, respectively. The MLR calculated mobilities of these peptides also show high deviations of −26.33, −43.70 and −48.99%, respectively. Jean Luc et al. have reported a hydrophobic parameter, which can be considered as a measure for the hydrophobicity of peptides [32]. They have obtained this parameter from the partitioning of *N*-acetyl-amino acid amides in octanol/water system [32]. Based on this parameter, peptides 66, 90 and 102 of Table 1 are the most hydrophobic ones with deviations of −66.33, 4.92 and 12.47%, respectively, in their MLR calculated values. Fig. 1 shows the correlation between the MLR calculated and the experimental values of the electrophoretic mobilities of peptides included in the test and validation sets. The correlation of $R^2 = 0.930$ indicates a reasonable agreement between these values and also demonstrates some improvements over those obtained using the Offord model. These improvements can be attributed to the inclusion of the parameters of corrected steric constant and molecular refractivity into the MLR model. However, this model shows weakness in predicting the electrophoretic mobilities of highly charged, low charged and hydrophobic peptides.

Table 3
Architecture and specifications of the ANN model

Number of nodes in the input layer	3
Number of nodes in the hidden layer	3
Number of nodes in the output layer	1
Number of iterations	35000
Learning rate	0.7
Momentum	0.9

3.2. Artificial neural networks analysis

The most important advantage of the artificial neural networks over regression analyses is their ability to allow for the flexible mapping of the selected features by manipulating their functional dependence implicitly. Developing networks and comparing them with the MLR models provides us the opportunity to investigate the nonlinear characteristics of the dependence of electrophoretic mobilities of peptides on structural descriptors. In order to have a meaningful comparison, the variables for the linear and nonlinear treatments should be the same. Therefore, the three descriptors appearing in the MLR model have been considered as the inputs for generating the networks. After optimizing the parameters needed for constructing ANNs, a neural network with architecture of 3-3-1 was obtained, which its specifications are given in Table 3. We used the test set consisting of twenty peptides to optimize the learning iteration size and avoid overtraining. To evaluate the network, the electrophoretic mobilities of peptides included in the validation set were predicted and are shown in Table 1. Fig. 2 shows the plot of the ANN predicted versus the experimental values of the electrophoretic mobilities for the test and validation sets. The agreement between the predicted and observed mobilities using the ANN modeling ($R^2 = 0.970$) was considerably higher than that with the MLR ($R^2 = 0.930$) for the same set of peptides descriptors. Inspection of the results given in Table 1 shows that this improvement is mostly due to a better ability of the nonlinear model in predicting the electrophoretic mobilities of highly charged, low charged or hydrophobic peptides. A detailed discussion of this improvement is given in the next section. Fig. 3

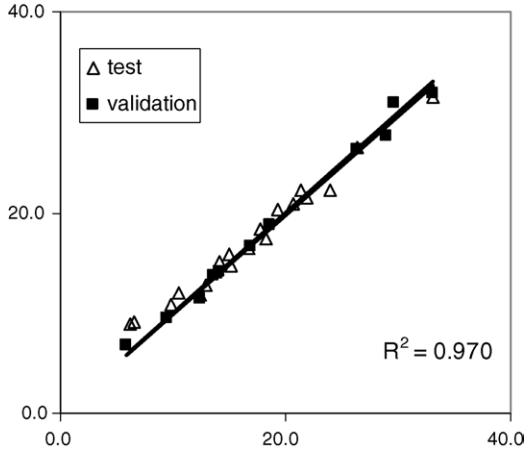

Fig. 2. Plot of the ANN calculated electrophoretic mobilities against the experimental values for the test and validation sets.

Table 4
Comparison of the statistics for the MLR and ANN models

Model	$R^2_{Training}$	R^2_{Test}	$R^2_{Validation}$	$SE_{Training}$	SE_{Test}	$SE_{Validation}$	F
MLR	0.903	0.907	0.949	1.89	1.77	0.91	872
ANN	0.959	0.960	0.993	1.04	0.96	0.65	3153

shows a plot of the residuals of ANN predicted values of the electrophoretic mobilities against the experimental values. The propagation of the residuals on both sides of zero indicates that no systematic error exists in the development of the neural network.

Table 4 lists more detailed statistics about the predictive power of the MLR and ANN models. It can be seen from this table that R^2 values for the ANN model are considerably higher than those for the MLR model for all three, training, test and validation sets. Also, the ANN predicted values show much lower standard errors as compared with those of the MLR model. The ANN model also reveals a higher value for the *F*-statistic. These results clearly reconfirm our earlier conclusion about the nonlinear behavior of the electrophoretic mobility of peptides, which is incorporated in the ANN modeling.

3.3. Prediction of highly charged and hydrophobic peptides

Despite the voluminous amount of attempt reported for calculating the electrophoretic mobilities of peptides [6–12,31], there is still no robust model to be able to predict accurately this parameter for all categories of peptides, especially highly charged and hydrophobic peptides. Notable exceptions to the previous works are a multi-variable computer model presented by Janini and coworkers [11,12] and our recent ANN model [13]. Janini and coworkers, based on a data set consisted of the electrophoretic mobilities of 58 peptides that varied in size from 2 to 39 amino acids examined the existing empirical models that correlates electrophoretic mobility with physical parameters. They reached to the conclusion that the charge-to-size parameter of Offord offers the best fit to their experimental data [10]. However, inspection of their results revealed a systematic deviation for small peptides with large positive charge, such as the lysine homologous with $n = 2–5$. To address this deficiency, Janini et al. presented a multi-variable model that takes into account other physical properties that were neglected by the Offord model [11,12]. However, the success of their model was dependent upon the accurate measurement of the electrophoretic mobilities of a large number of peptides with a wide range of charge and molar mass. However, calculations of different functions in their model depend entirely on the data set and changing the basis set requires new equations for these functions. The ANN model has overcome the problem of robustness by incorporating three general descriptors as its inputs. To demonstrate the predictive ability of the ANN model, we have highlighted the results for some of the highly charged and hydrophobic peptides in Table 5. For the sake of comparison the Offord, multi-variable [12] and MLR calculated values for the electrophoretic mobilities together with some statistics are also included in this table. The data included in Table 5 represent peptides with different sizes and were chosen based on a charge larger than 4 and a hydrophobic parameter larger than 3. The results of this table clearly demonstrate a significant improvement for the multi-variable and ANN models compared with the MLR and the Offord models. Although there is no significant differences between the R^2 values of multi-variable and ANN models, the later shows a lower standard error (SE) and relative standard deviation (RSD). It can be seen from Table 5 that the ANN model shows a RSD of 4.43% which is much lower than those of the Offord and MLR models. It should be noted that we have developed this model using the experimental mobilities which demonstrate an average RSD of 2.34% themselves. Despite the improvements, for both models several inconsistencies exist between the sequences of the real and predicted electrophoretic mobilities of highly charged and hydrophobic peptides. A summary of these inconsistencies is shown in Fig. 4. Peptides 10, 12 and 19 show an incorrect order for the ANN model, while the order of electrophoretic mobilities for peptides 9, 13, 19 and 22 is not correct for the multi-variable model. Comparison of the results in Table 5 reveals superiority for ANN model over the MLR and the Offord model. The R^2

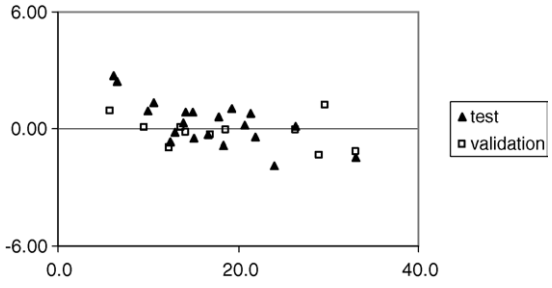

Fig. 3. Plot of the residuals vs. the experimental values of peptide mobility for the test and validation sets.

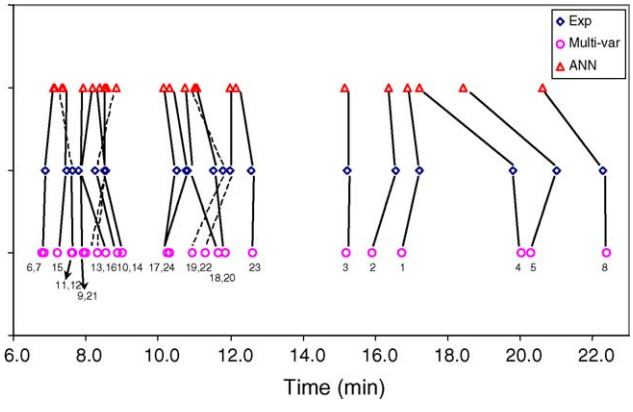

Fig. 4. Sequences of the real and predicted electrophoretic mobilities of highly charged and hydrophobic peptides using multi-variable and ANN models.

Table 5
Comparison of different models for highly charged and hydrophobic peptides

No.	Peptide	Exp	Offord	Multi-var[d]	MLR		ANN	
		$\mu_{ef} \times 10^5$	$\mu_{ef} \times 10^5$	$\mu_{ef} \times 10^5$	$\mu_{ef} \times 10^5$	Percent deviation	$\mu_{ef} \times 10^5$	Percent deviation
1	FF[a]	12.81	12.98	13.18	13.99	−9.21	13.06	−1.95
2	FL[a]	13.33	13.82	13.91	14.30	−7.28	13.51	−1.35
3	LL[a]	14.55	14.85	14.58	14.73	−1.24	14.61	−0.41
4	WW[a]	11.05	11.52	10.91	13.23	−19.73	12.80	−15.84
5	FFF[a]	10.38	10.59	10.76	12.48	−20.23	11.92	−14.84
6	KKKK[b]	33.03	45.34	33.53	35.17	−6.48	31.88	3.48
7	KKKKK[b]	33.03	47.26	33.18	36.45	−10.35	31.76	3.84
8	YGGFL[a]	9.75	9.62	9.70	11.76	−20.62	10.58	−8.51
9	ACHGRDRRT[b]	26.54	28.80	28.54	24.39	8.10	26.90	−1.36
10	VVRRYPHHE[b]	27.38	27.40	25.46	23.24	15.12	25.55	6.68
11	CRHRRRHRRGC[b]	29.68	41.78	29.75	33.25	−12.03	30.72	−3.50
12	CRHHRRRHRRGC[b]	29.68	43.74	29.61	34.54	−16.37	30.89	−4.08
13	HMTEVRRYPHHER[b]	26.42	30.32	27.11	24.92	5.68	26.54	−0.45
14	HMTEVRHCPHHER[b]	26.41	29.55	25.07	25.35	4.01	26.36	0.19
15	HRSCRRRKRRSCRHR[b]	30.27	45.09	31.41	35.57	−17.51	30.68	−1.35
16	RTHCQSHYRRRHCSR[b]	28.96	34.99	26.39	28.72	0.83	27.64	4.56
17	YAEGDVHATSKPARR[b]	21.38	22.10	21.94	20.04	6.27	22.15	−3.60
18	VVRRCPHQRCSDSDGL[b]	20.75	20.54	19.22	18.44	11.13	20.92	−0.82
19	DGLAPPQHRIRVEGNLR[b]	18.98	20.21	20.49	18.10	4.64	20.40	−7.48
20	NHQLLSPAKTGWRIFHP[c]	19.42	20.14	18.85	18.40	5.25	20.28	−4.43
21	RTHGQSHYRRRHCSRRRLHRIHRRQ[b]	29.01	43.60	28.32	34.29	−18.20	28.57	1.52
22	FLTPKKLQCVDLHVISNDVCAQVHPQKVTK[c]	18.68	19.59	19.83	16.35	12.47	18.65	0.16
23	KQIINMWQEVGKAMYAPPISGQIRRIHIGPGRAFYTTKN[c]	17.78	19.21	17.71	16.91	4.89	18.39	−3.43
24	DRVIEVVQGAYRAIRHIPRRIRQGLERRIHIGPGRAFYTTKN[c]	20.83	27.85	21.79	22.13	−6.24	21.80	−4.66
R^2			0.91	0.98	0.91		0.99	
Standard error			3.67	1.09	2.60		0.80	
RSD			33.41	4.88	12.17		4.43	

[a] Hydrophobic peptide.
[b] Highly-charged peptide.
[c] Hydrophobic and highly-charged peptide.
[d] Multi-variable model from Janini et al. work [12].

value of 0.990 for the ANN should be compared with a value of 0.910 for the MLR and Offord models. Also, SE and RSD values of 0.80 and 4.43%, respectively, for the ANN model should be compared with 2.60, 3.67 and 12.17 and 33.41% for the MLR and Offord models, respectively. Large deviations of −17.51, −18.21 and −6.24% for the MLR calculated values of the most highly charged peptides 48, 68 and 70, respectively, should be compared with the values of −1.34, 1.52 and −4.64% for their ANN calculated counterparts. Also, deviations of 4.92 and 12.4% for the MLR calculated electrophoretic mobilities of highly hydrophobic peptides of 90 and 102, respectively, should be compared with the values of −3.41 and 0.16% for their ANN counterparts. A notable exception is the hydrophobic peptide FIGITEAAANLVPMVATV (Table 1) with a large deviation of −66.33 and −82.22% for the MLR and ANN calculated values, respectively. Although, we are uncertain about the origin of these deviations, they could be due to the experimental uncertainty.

3.4. Simulation of peptide maps of protein digests

The long-range goal of developing theoretical models, which can accurately predict the CZE parameters, such as retention time or electrophoretic mobility, is the construction of a database of peptide maps. Reaching this goal means that one can easily identify unknown proteins by submitting the experimental maps to the database and searching for the closest match in terms of the migration times of the major peaks. To explore the utility of the ANN model in simulation of the CZE peptide maps, the profiles for the endoproteinase Lys-C digests of a peptide sequencing standard, melittin GIGAVLKVLTTGLPALISWIKRKRQQ, and two more complicated proteins namely glucagon and horse cytochrome *C* were studied in this work. Choosing these peptides were based on possibility of comparison of the ANN simulated maps with the experimental and multi-variable simulated ones.

Fig. 5 demonstrates the experimental and ANN simulated maps for the endoproteinase Lys-C digest of the peptide sequencing standard of melittin. The correct migration order of peptides and corresponding retention times agrees fairly with the experimental electropherogram. However, multi-variable results also show a correct order and accurate retention times [12].

Next, we have considered glucagon, a polypeptide with 29 amino acid residues. This polypeptide can be used as a control for proteolytic digestion, sequencing and amino acid analysis. Janini and coworkers digested this protein with endoproteinase Glu-C with characteristic cleavage at the C-terminal of aspar-

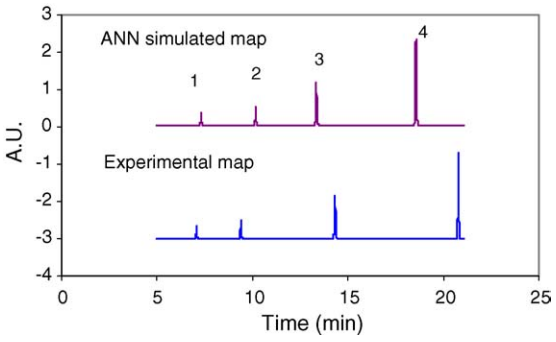

Fig. 5. Experimental and ANN simulated maps for the endoproteinase Lys-C digest of melittin.

tic acid (D) and glutamic acid (E) residues. Therefore, after a complete digestion, four fragments are expected for this protein. These fragments are listed in Table 6. Also, the values for the three descriptors together with the calculated MLR and ANN values for the electrophoretic mobilities of glucagon fragments are summarized in this table. The ANN calculated mobilities were converted to the migration times and simulated electropherogram is shown in Fig. 6. For the purpose of comparison, the experimental and simulated electropherograms reported by Janini et al., are also shown in this figure. Inspection of Fig. 6 reveals an excellent matching between the line positions and a reasonable agreement between the relative heights of the experimental and simulated peaks. It seems that both models of multi-variable and ANN overestimate the mobility for the FVQWLMNT fragment and therefore, a smaller value for the corresponding migration time. Validity of this conclusion

Table 6
Descriptor values together with MLR and ANN calculated mobilities of theoretical fragments of Lys-C digest of cytochrome *C* and Glu-C digest of Glucagon

No.	Peptide sequence	Descriptors			MLR	ANN
		QM	$E_{s,c}$	MR	$\mu_{ef} \times 10^5$	$\mu_{ef} \times 10^5$
Glucagon digest						
1	SRRAQD	0.0338	−2.92	103.6	19.88	22.05
2	YSKYLD	0.0204	−4.32	127.1	14.60	14.41
3	HSQGTFTSD	0.0177	−4.18	123.6	13.56	12.75
4	FVQWLMNT	0.0081	−6.45	172.9	9.69	7.90
Cytochrome *C* digest						
5	GK	0.0530	−0.42	26.08	27.27	29.52
6	HK	0.0657	−1.28	48.84	32.02	31.39
7	NK	0.0450	−1.40	39.5	24.03	27.27
8	GGK	0.0450	−0.22	27.11	24.27	27.29
9	ATNE	0.0144	−1.93	48.2	12.32	10.18
10	GDVEK	0.0259	−2.91	68.8	16.60	17.67
11	GITWK	0.0257	−3.22	97.3	16.66	17.61
12	IFVQK	0.0249	−4.64	108.7	16.03	17.00
13	YIPGTK	0.0238	−3.26	103.3	15.97	16.48
14	MIFAGIK	0.0217	−5.17	124.0	14.79	15.02
15	CAQCHTVEK	0.0279	−4.14	144.4	17.64	18.86
16	TEREDLIAYLK	0.0223	−8.58	207.2	14.74	15.51
17	EETLMEYLENPK	0.0137	−8.42	224.3	11.68	10.92
18	TGPNLHGLFGRK	0.0322	−5.79	191.4	19.20	20.81
19	TGQAPGFTYTDANK	0.0135	−5.39	194.8	12.22	11.11

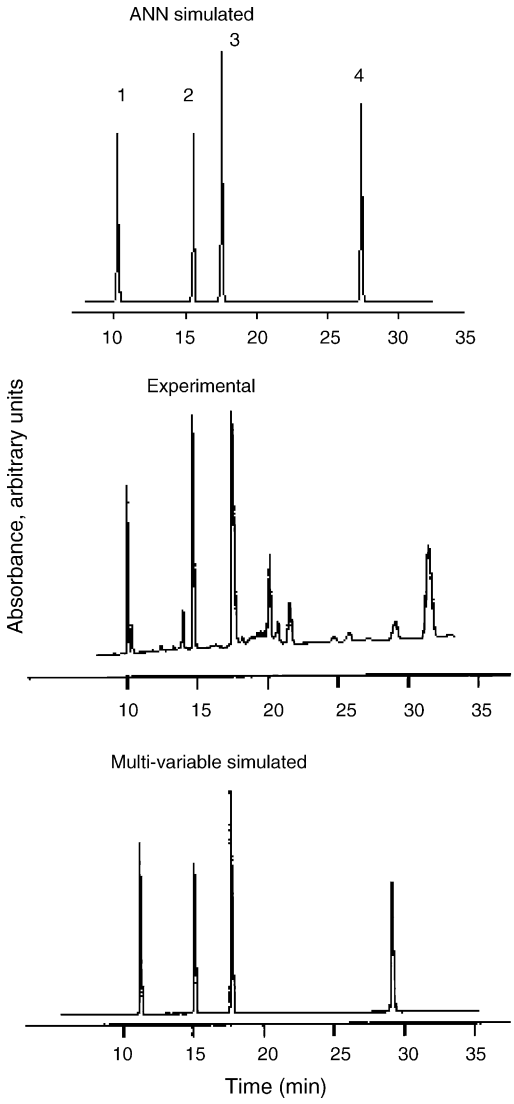

Fig. 6. Experimental, ANN and multi-variable simulated electropherogram of glucagons.

depends upon the accuracy of assignments of the peaks. However, one should consider the possibility of several shortcomings in experiment such as imperfect enzymatic digestion, impurity and autolysis of the endoproteinase.

Finally, we studied the relatively complex protein of horse cytochrome *C* with 104 amino acid residues. This protein also was digested with endoproteinase Lys-C with specificity of cleavage at the C-terminal of lysine residues [12]. The theoretical fragments of this protein together with the values of the descriptors and corresponding MLR and ANN calculated mobilities are demonstrated in Table 6. Fig. 7 shows the experimental and ANN simulated electropherogram of cytochrome *C*. For the sake of comparison, the simulated electropherogram of this protein obtained by Janini et al. is also included in this table. Inspection of Fig. 7 shows that the experimental electropherogram of cytochrome *C* have a striking similarity to the ANN simulated electropherogram. Almost each simulated peak has a counterpart in the experimental electropherogram with a good agreement between their migration times. However, Janini et al.

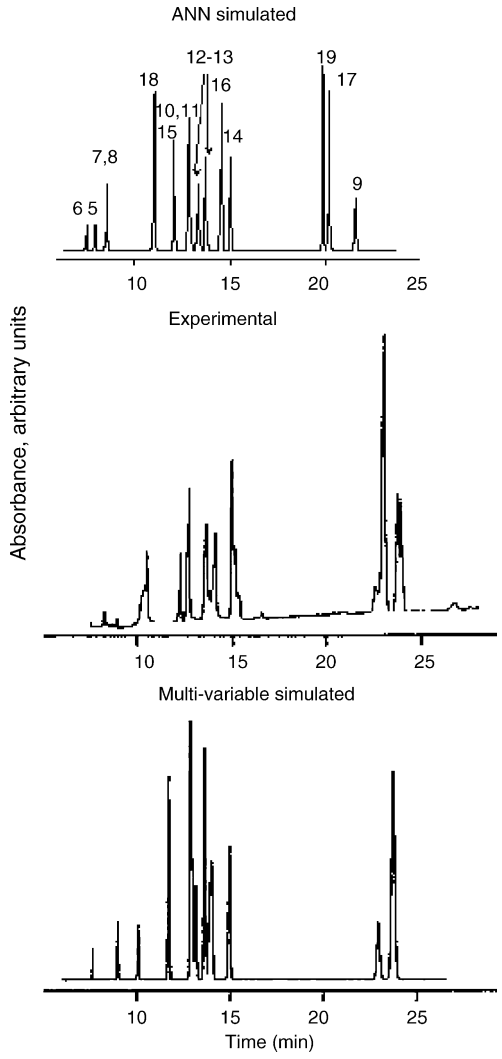

Fig. 7. Experimental, ANN and multi-variable simulated electropherogram of cytochrome *C*.

have obtained only 13 peaks for the 15 theoretical fragments. They believed that the pairs of fragments (17, 19), (5, 7) and (10, 11) co-migrated because of the closeness of their migration time. Our result confirms this conclusion except for the pairs of (5, 7), which is replaced with the pairs of (5, 8) in this work.

4. Conclusion

Recently, our research is focused on developing a multivariate model, which can be used as a tool in the construction of a database of peptide maps. As a long-range aim we have considered the specifications of simplicity, accuracy and robustness for the model in predicting the CZE parameters of peptides. Also to get this dream closer to reality, the model must be able to predict accurately the CZE parameters of peptides with identical amino acid residues but different sequences. Three steps were considered for reaching this goal: (1) Developing a model by using the descriptors as simple and general as possible and examining its ability in accurate prediction of the electrophoretic mobilities of peptides. It is shown that the three simple descriptors of charge-over-mass ratio of $Q/M^{2/3}$, corrected steric substituent constant and molar refractivity are very useful in predicting the electrophoretic mobilities of different categories of peptides. (2) To investigate the robustness of the model, a more diverse data set obtained in different experimental conditions was needed. To fulfill this, a more diverse data set consisted of 102 peptides was chosen in this work. The robustness of the neural network model was exhibited by accurate ANN calculated electrophoretic mobilities of all categories of peptides, obtained in different experimental conditions. Also, simulating the endoproteinase digests of melittin, glucagon and horse cytochrome *C* maps showed the utility of the model in simulation of the CZE peptide maps. (3) A research is under way in our group to explore the use of a new series of sequence-descriptors. The preliminary results are promising. Success in this step improves the capability of the model in simulating proteins maps consisted of isomeric peptides.

Acknowledgements

This work was supported by research grant from the U.S. National Institutes of Health (GM 38738). M. Jalali-Heravi acknowledges Sharif University of Technology, for financial support of his sabbatical leave.

References

[1] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Moris, B.M. arvik, J.R. Yates III, Nat. Biotechnol. 17 (1999) 767.
[2] M.P. Washburn, D. Wolters, J.R. Yates III, Nat. Biotechnol. 19 (2001) 242.
[3] P.D. Grossman, J.C. Colburn, H.H. Lauer, Anal. Biochem. 179 (1989) 28.
[4] R.E. Offord, Nature (London) 211 (1966) 591.
[5] B.J. Compton, J. Chromatogr. 599 (1991) 357.
[6] N.J. Adamson, E.C. Reynolds, J. Chromatogr. B 699 (1997) 133.
[7] I. Messana, D.V. Rossetti, L. Cassiano, F. Misiti, B. Giardina, M. Castagnola, J. Chromatogr. B 699 (1997) 149.
[8] V. Kasicka, Electrophoresis 20 (1999) 3084.
[9] A. Cifuentes, H. Poppe, Electrophoresis 18 (1997) 2362.
[10] G.M. Janini, C.J. Metral, H.J. Issaq, G.M. Muschik, J. Chromatogr. A 848 (1999) 417.
[11] C.J. Metral, G.M. Janini, G.M. Muschik, H.J. Issaq, J. High Resolut. Chromatogr. 22 (1999) 373.
[12] G.M. Janini, C.J. Metral, H.J. Issaq, J. Chromatogr. A 924 (2001) 291.
[13] M. Jalali-Heravi, Y. Shen, M. Hassanisadi, M.G. Khaledi, Electrophoresis 26 (2005) 1874.
[14] T. Baczek, A. Bucinski, A.R. Ivanov, R. Kaliszan, Anal. Chem. 76 (2004) 1726.
[15] K. Sak, M. Karelson, J. Jarv, Bioorg. Chem. 27 (1999) 434.
[16] B. Skoog, A. Wichman, Trends Anal. Chem. 5 (1986) 82.
[17] M. Akamatsu, T. Katayama, D. Kishimoto, Y. Kurokawa, H. Shibata, T. Ueno, T. Fujita, J. Pharm. Sci. 83 (1994) 1026.
[18] C. Hansch, A. Leo, D. Hoekman, Exploring QSAR. Hydrophobic, Electronic, and Steric Constants ACS Professional Reference Book, ACS, Washington, DC, 1995.
[19] J. Zupan, J. Gasteiger, Neural Networks in Chemistry and Drug Design, VCH, Weinheim, 1999.
[20] N.K. Bose, P. Liang, Neural Network Fundamentals, McGraw Hill, New York, 1996.
[21] D.W. Patterson, Artificial Neural Networks: Theory and Applications, Simon and Schuster, New York, 1996.

[22] P.K. Hopke, X. Song, Anal. Chim. Acta 348 (1997) 375.
[23] M. Jalali-Heravi, F. Parastar, J. Chromatogr. A 903 (2000) 145.
[24] M. Jalali-Heravi, Z. Garkani-Nejad, J. Chromatogr. A 927 (2001) 211.
[25] M. Jalali-Heravi, M. Kyani, J. Chem. Inf. Comput. Sci. 44 (2004) 1328.
[26] T.B. Blank, S.T. Brown, Anal. Chem. 65 (1993) 3084.
[27] R.K. Scopes, Anal. Biochem. 59 (1974) 277.
[28] M. Herold, G.A. Ross, R. Grimm, D.N. Heiger, in: K.D. Altria (Ed.), Capillary Electrophoresis Guidebook, Principles, Operation, and Application, Methods in Molecular Biology, vol. 52, Humana Press, Totowa, NJ, 1996, p. 285.
[29] H.J. Issaq, G.M. Janini, I.Z. Atamna, G.M. Muschik, J. Lukso, J. Liq. Chromatogr. 15 (1992) 1129.
[30] E.C. Rickard, M.M. Strohl, R.G. Nielson, Anal. Biochem. 197 (1991) 197.
[31] A. Cifuentes, H. Poppe, J. Chromatogr. A 680 (1994) 321.
[32] F. Jean Luc, P. Vladimir, Eur. J. Med. Chem. 18 (1983) 369.