J Mol Model (2005) 11: 55–60
DOI 10.1007/s00894-004-0221-x

ORIGINAL PAPER

Hakan Kayı · S. Ali Tuncel · Ali Elkamel
Erdoğan Alper

Prediction of lower critical solution temperature of *N*-isopropylacrylamide – acrylic acid copolymer by an artificial neural network model

Received: 8 October 2004 / Accepted: 25 October 2004 / Published online: 8 December 2004

Abstract In this paper, we have investigated the lower critical solution temperature (LCST) of *N*-isopropylacrylamide–acrylic acid (NIPAAm-AAc) copolymer as a function of chain-transfer agent/initiator mole ratio, acrylic acid content of copolymer, concentration, pH and ionic strength of aqueous copolymer solution. Aqueous solutions with the desired properties were prepared from previously purified polymers, synthesized at 65 °C by solution polymerization using ethanol. The effects of each parameter on the LCST were examined experimentally.In addition, an artificial neural network model that is able to predict the lower cretical solution temperature was develeped. The predictions from this model compare well against both training and test data sets with an average error less than 2.53%.

Keywords Lower critical solution temperature · neural networks · *N*-isopropylacrylamide–acrylic acid copolymer

H. Kayı · S. A. Tuncel · E. Alper
Chemical Engineering Department,
Hacettepe University,
Beytepe, Ankara, Turkey

A. Elkamel
Chemical Engineering Department,
University of Waterloo,
Ontario, Canada

H. Kayı (✉)
Computer-Chemie-Centrum,
Erlangen-Nürnberg University,
Erlangen, Germany
E-mail: hakan.kayi@chemie.uni-erlangen.de
Tel.: +49-9131-8526583
Fax: +49-9131-8526565

Introduction

Stimuli-responsive or smart polymers have been widely used in different biological applications, such as the delivery of therapeutics, cell culture, tissue engineering, bioseperations, sensors or actuator systems and gene transfection. These polymers respond to small external stimuli with large changes in their properties. The external stimuli may be temperature, pH, electric field, chemicals or ionic strength and the responses are large changes in the properties such as shape, surface characteristics, solubility, or homogeneous solution phase to gel phase transition.

Lower critical solution temperature (LCST) is defined succinctly as the phase transition temperature of a thermo-sensitive polymer. If the temperature is above the LCST, a reversible macromolecular phase transition from hydrophilic to hydrophobic structure occurs and a phase change resulting in precipitation of the polymer is observed. Below LCST, the interaction of polymer and water molecules leads to dissolution of polymer chains due to interactions *via* hydrogen bonding between the polymer and water. Poly(*N*-isopropylacrylamide) (polyNIPAAm) is a temperature-sensitive polymer and it has an LCST of 32 °C [1, 2]. Below this temperature, the polymer is hydrophilic and soluble, and above LCST it is hydrophobic and becomes collapsed. Some polymers respond to a combination of two or more stimuli. Temperature- and pH-responsive polymers have been prepared by copolymerization of the temperature-sensitive monomer NIPAAm with hydrophilic comonomers containing ionizable groups such as acrylic acid (AAc). Therefore, the NIPAAm–AAc copolymer has both temperature- and pH-sensitivity. Addition of AAc comonomer also results in an increase in LCST [3].

The dependency of the LCST of the polymer on parameters such as pH, comonomer content, chain-transfer agent/initiator mole ratio in feed, ionic strength and concentration of aqueous polymer solution is highly

nonlinear. Although there are now several nonlinear multivariate statistical methods that are able to approximate any nonlinear relationship, the assumption of functional dependency is a serious drawback of such procedures. An alternative approach would be the use of artificial neural networks (ANNs) for extracting information without any assumptions of the nature of the nonlinearity, from measured data in the form of predictive input–output models, or as a way of representing the input space efficiently. [4]

ANNs, also classified as black-box models, can handle multivariable problems. ANNs are based on a principle stating that a system of highly interconnected simple processing elements can learn complex interrelationships between independent and dependent variables [5].

In this study, an ANN that can predict the LCST of temperature- and pH-responsive NIPAAm–AAc copolymer depending on the AAc content of copolymer, chain transfer agent/initiator mole ratio, concentration, pH and ionic strength of aqueous copolymer solution parameters is presented. The validity of the network for NIPAAm–AAc copolymer was examined.

Materials and methods

Materials

The monomer, NIPAAm, (Aldrich Chemical Co., Milwaukee,WI USA) was recrystallized from *n*-hexane–acetone solution. The comonomer, AAc, which was supplied by Aldrich Chemical Co., was used as received. Azoisobutyronitrile (AIBN, Merck AG, Darmstad, Germany) was recrystallized from methanol (Aldrich Chemical Co.) and used as initiator in the copolymerization. Absolute ethanol (Merck AG) was selected as the solvent for free radical solution polymerization. A chain-transfer agent, 2-mercaptoethylamine (MEA, Sigma Chemical Co, USA) was used in copolymerization as received.

Methods

Preparation of NIPAAm–AAc Copolymers

A series of random copolymers of NIPAAm and AAc was prepared by free radical solution polymerization of the two monomers in ethanol using AIBN as the initiator in the presence of the chain transfer agent MEA. NIPAAm and AAc were dissolved in the organic phase containing ethanol in a sealed cylindrical glass polymerization reactor. The initiator, AIBN and the chain transfer agent, MEA were then dissolved in the resulting homogeneous mixture. The polymerization medium was purged by bubbling dry nitrogen for 10 min to remove any oxygen dissolved in the reaction mixture and then sealed. The copolymerization was performed at 65 °C for 24 h (Scheme 1) under nitrogen atmosphere with shaking at 120 cpm in a shaking water-bath equipped with a temperature-control system. After cooling to room temperature, the organic phase was removed in a rotary evaporator under vacuum at 40 °C. The resulting copolymers were dissolved in water and again precipitated in water by heating and adjusting the pH. The redissolution–reprecipitation was repeated several times for the complete removal of impurities in the precipitated copolymer. The isolated copolymer was dried under vacuum at 40 °C and stored. The details of the feed solutions used in copolymerizations, as well as the AAc content of the resulting copolymers, are shown in Table 1. AAc content was determined by titration method.

Artificial neural networks

ANNs are biologically inspired computer programs designed to simulate the way in which the human brain processes information. ANNs gather their knowledge by detecting the patterns and relationships in data and learn (or are trained) through experience, not from programming [6]. They are extensively interconnected parallel structures containing processing elements called neurons. A neuron processes an input vector **x** with components $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$,..., $\mathbf{x}_n$ to give an output **y**. The output **y** can serve as an input to other neurons. Several factors other than the input vector **x** determine the output **y**. Synaptic weights, an adder for summing the input signals and a transfer function for limiting the amplitude of the output of a neuron are the three basic

Table 1 Feed compositions used in the preparation of NIPAAm–AAc copolymers and the AAc content of the copolymers

NIPAAm/AAc (%mol/mol)	AIBN (%mol)	MEA (%mol)	AAc content of copolymer (mmol AAc/g copolymer)
96.66/0	3.34	0	0
92.95/3.84	3.21	0	0.36
89.51/7.40	3.09	0	0.49
83.35/13.77	2.88	0	1.08
69.08/28.53	2.39	0	2.56
53.74/44.40	1.86	0	3.68
43.98/54.50	1.52	0	4.25
95.08/0	3.28	1.64	0
91.48/3.78	3.16	1.58	0.34
88.16/7.28	3.04	1.52	0.45
82.17/13.57	2.84	1.42	1.05
93.54/0	3.23	3.23	0
90.06/3.72	3.11	3.11	0.34
86.83/7.17	3.0	3.0	0.44
81.02/13.38	2.80	2.80	1.03
90.62/0	3.13	6.25	0
87.35/3.61	3.01	6.03	0.31
84.31/6.96	2.91	5.82	0.41
78.82/13.02	2.72	5.44	1.01

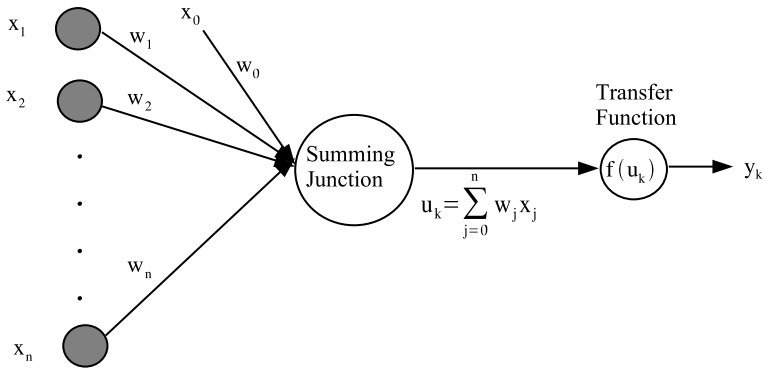

Fig. 1 A nonlinear neuron model

elements of the neuron model. Also, there is an internal threshold, $\mathbf{x}_0$. The weight vector **w** multiplies the input vector **x** in order to obtain a weighted input $\mathbf{w_j x_j}$ on which further calculations are performed. The output of the neuron **k** is obtained by applying a transfer function **f** on summed weighted inputs $\mathbf{u_k}$ (i.e., $\mathbf{y_k = f(u_k)}$) (Fig. 1).

There are several forms of the transfer function that can be used. S-shaped sigmoidal functions are widely used as nonlinear transfer functions because of their advantages, especially in supervised learning [7]. They are continuously differentiable, monotonic, symmetric, bounded between 0 and 1 for logistic sigmoid and, −1 and 1 for tangent hyperbolic sigmoid functions. To avoid overflows due to very large or very small data, input and output data are normalized to the range of the transfer function used. The data z can be normalized in an interval $[\sigma_1, \sigma_2]$ corresponding to the transfer function by using:

$$n_j = \sigma_1 + (\sigma_2 - \sigma_1)\left(\frac{z - z_{\min}}{z_{\max} - z_{\min}}\right) \quad (1)$$

where n_j is the normalized value and can be an input or an output datum. Also, $z_{\min}$ and $z_{\max}$ are the minimum and maximum values of z, respectively.

The processing elements within an ANN are arranged in a certain structure. A typical multilayer feed-forward network consists of an input layer, an output layer and at least one layer between input and output layer, termed the hidden layer (Fig. 2). By adding one or more hidden layers, the network is enabled to extract higher-order statistics. In a rather loose sense, the network acquires a global perspective despite its local connectivity due to the extra set of synaptic connections and the extra dimension of neural interactions [8].

As shown in Fig. 2, each neuron is connected to every other neuron in the next adjacent layer. The input neurons transmit their inputs to the hidden neurons without any processing because they do not posses any transfer function. Thus, they act as distribution channels. The neurons in the hidden and output layers calculate their inputs by performing a weighted summation of all the outputs they receive from the preceding layer. Their outputs are calculated by transforming their inputs using a nonlinear transfer function.

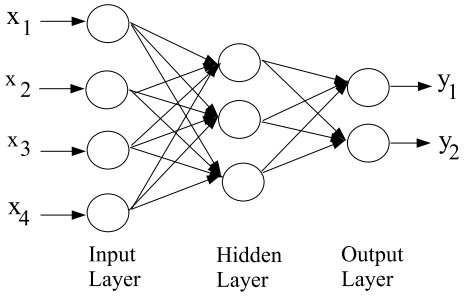

Fig. 2 Architecture of a feed-forward multilayer neural network with one hidden layer

The procedure for selecting the best neural network architecture is a trial and error one. Thus, training an ANN is an optimization process and includes modifying the network architecture, which involves adjusting the weights of the links, pruning or creating some connection links, and/or changing the firing rules of the individual neurons [9]. Different architectures are trained and the network that best mimics the real system is chosen. An ANN is said to have learned if it can (1) handle imprecise, fuzzy, noisy and probabilistic information without noticeable adverse effect on response quality, and (2) generalize from the examples it has learned to unknown ones [10]. Training is usually performed via a backpropagation-with-momentum algorithm. This algorithm minimizes the sum squared error between the values predicted by the neural network and desired values [11].

Details about the training algorithms, design and the mathematics behind the neural networks can be found in many sources. [12, 13, 14].

LCST data collection

The focus of the data-collection step is to generate data for LCST of the NIPAAm–AAc copolymer as a function of AAc content of copolymer, chain-transfer agent/initiator mole ratio in feed, concentration, pH and ionic strength of the aqueous copolymer solution.

Table 2 Details of the input parameters

Parameter	Measurement range	
	Minimum value	Maximum value
pH	2	12
AAc content of copolymer (mmol/g)	0	4.25
Ionic strength (N)	0	3
MEA/AIBN mole ratio (mol/mol)	0/1	2/1
Concentration (weight%)	0.15	0.5

LCST variations for each of the copolymers were obtained from cloud-point experiments. The cloud points of the solutions were determined by repeated heating and cooling the samples in 2- ml test tubes, and the temperature at which the solutions either became visually turbid or cleared was measured with a Model 3750-K thermometer (Digitron, England).

Table 2 summarizes the details of the input parameters. Up to 1.08 mmol AAc/g copolymer content, LCST values of the copolymer were measured in the pH range 2–12 and ionic strength range 0–1 N. For the copolymers with the contents of 2.56, 3.68 and 4.25 mmol AAc/g copolymer, LCST was measured in the pH range 2–7 although ionic strength was adjusted up to 3 N. Since the carboxylic acid groups are completely ionized, the temperature sensitivity is lost and no LCST was observed for NIPAAm–AAc copolymer as from 2.56 mmol/g AAc content above pH of 7.

It was found that the low pH and the high ionic strength values can tolerate a larger AAc content within the copolymer. When the pH was lowered, the ratio of nonionized carboxylic acid groups to ionized ones increased. Therefore, intra- and inter-molecular hydrogen bonds increased, then the hydophilicity of the copolymer decreased. In addition, increasing the ionic strength by addition of NaCl reduces the solvent quality for NIPAAm, e.g., hydrophobic hydration around polymer side chains was weakened by the solvation of salt ions, while at the same time the electrostatic repulsion has been diminished. As for low salt conditions, we have to consider a reduced dissociation of protons from AAc with temperature [15].

Results and Discussion

To determine the variation of LCST of NIPAAm–AAc copolymer with five parameters as pointed out before, 944 experiments were carried out. Then the data were randomly divided into two sets. One set for training the network and the remaining set for testing. The training set was formed from 850 patterns, approximately 90% of all data. Since the LCST is a function of five different parameters, the architecture of the artificial neural network consisted of an input layer of five neurons and an output layer of one neuron only. Only one hidden layer was used initially. First, five hidden neurons were tried and the number of neurons was increased systematically, checking each time if the prepared neural network succeeded in reaching a mean squared error given as 10^{-3}. All variables were normalized using their minimum and maximum values according to Eq. 1 and Table 2. The Levenberg–Marquardt backpropagation algorithm was used to train the network [16]. It was found that a network with one hidden layer of 14 neurons gave sufficiently accurate predictions. The use of more neurons or more hidden layers did not result in improved accuracy. The design, the training, the testing and the validation of neural networks were performed using the Neural Network Toolbox of MATLAB [17].

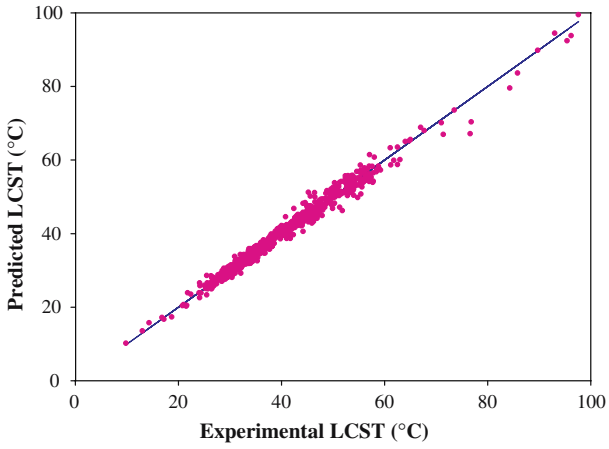

Fig. 3 Cross plot of predicted and experimental LCST values for the training data set

Figure 3 shows a cross plot of predicted and measured LCST values for the training set. The average absolute error for this data set was 2.53% and the minimum and maximum errors were 0.0008 and 13.38%, respectively.

The trained network was used to furnish predictions of LCST for data points that were not used in training the network (the testing set). This test step is important to check the generalization characteristics of the neural network prepared. In this case, the average absolute error was 2.38% and the minimum and maximum errors were 0.001 and 8.049%, respectively. Figure 4 shows a cross plot of predicted and measured LCST values for the testing set. The data fall close to the 1 : 1 parity line, indicating that the predictions are also good for this testing set.

Finally, in order to illustrate the validity of the neural network model, effects of each the five parameters on LCST were predicted. For predicting the effect of one parameter on LCST, the remaining four parameters were held constant. The average percentage error values for LCST vs pH, AAc content, ionic strength, concen-

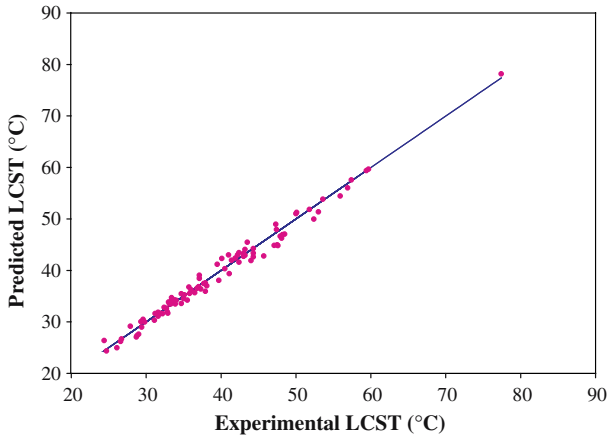

Fig. 4 Cross plot of predicted and experimental LCST values for the testing data set

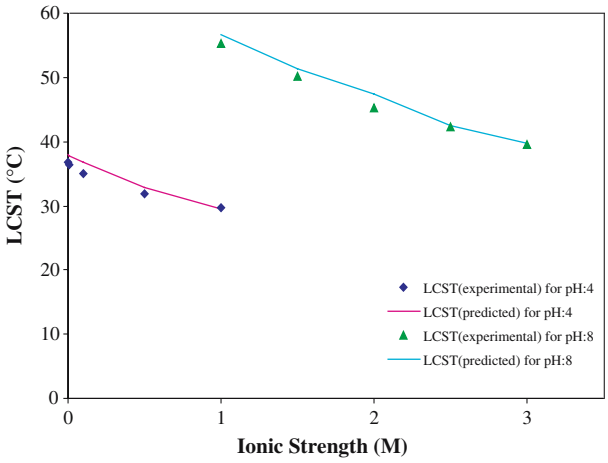

Fig. 7 Validation of network: experimental and predicted LCST vs ionic strength. AAc = 1.08 mmol/g, MEA/AIBN = 1/2 mol/mol, C = 0.4% by mass

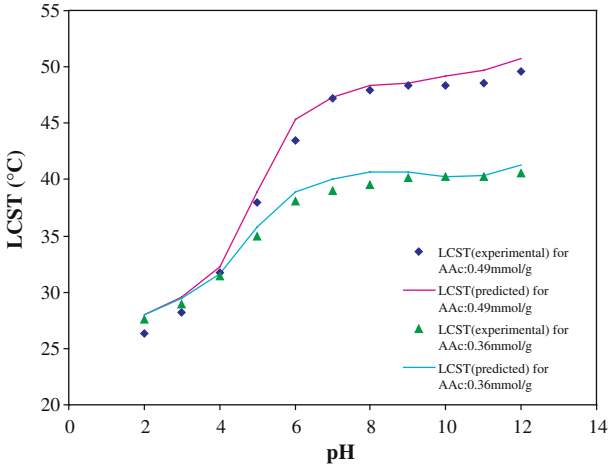

Fig. 5 Validation of network: experimental and predicted LCST vs pH. I.S = 0.5 M, MEA/AIBN = 0/1 mol/mol, C = 0.33% by mass

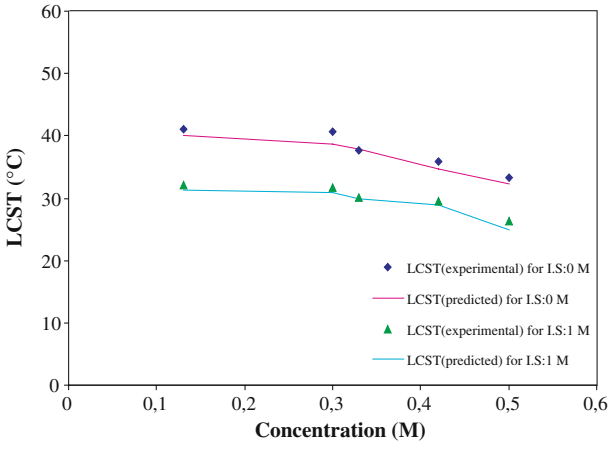

Fig. 8 Validation of network: experimental and predicted LCST vs concentration. AAc = 0.49 mmol/g, I.S = 0 M, MEA/AIBN = 1/2 mol/mol

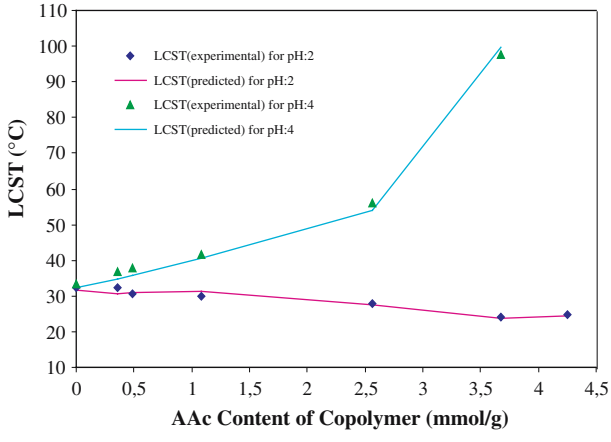

Fig. 6 Validation of network: experimental and predicted LCST vs AAc content. I.S = 0 M, MEA/AIBN = 0/1 mol/mol and C = 0.5% by mass

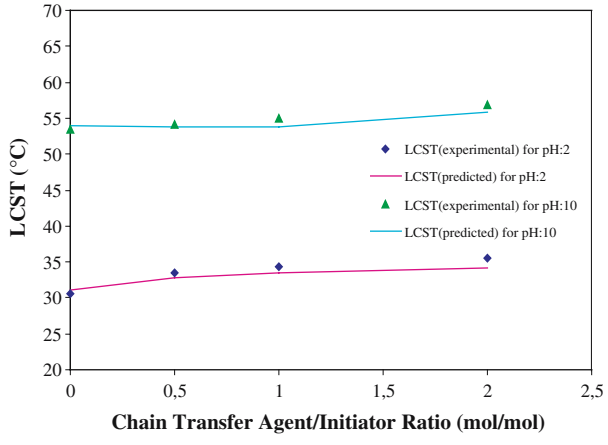

Fig. 9 Validation of network: experimental and predicted LCST vs MEA/AIBN ratio. AAc = 0.49 mmol/g, I.S = 0 M, C = 0.3% by mass

tration and chain transfer agent/initiator ratio were found to be 2.05, 2.87, 2.66, 2.75 and 1.87%, respectively. Plots comparing the predictions of the neural network model with the actual measurements are shown in Figs.5–9. The effects of the five parameters on LCST were mimicked properly by the neural network model developed.

Concluding remarks

Random copolymers of NIPAAm–AAc exhibiting both thermo- and pH-responsive behavior were synthesized by solution copolymerization. Comonomer feed compositions ranging from 0 to 54.5 mol% and the copolymer compositions ranged from 0–4.25 mmol/g copolymer. The primary aim of this work was to predict the LCST value of NIPAAm–AAc copolymer by an ANN model. Since LCST was defined as a function of given five parameters, LCST values of copolymers were measured for changes in these parameters in given ranges. The Levenberg–Marquardt backpropagation algorithm was used to train the network. For the training, the testing and the validation data sets, it was found that a network of one hidden layer with 14 neurons fared well. The average percentage error was found to be 2.53% for the training set, while that for the testing set was 2.38%.

Acknowledgements This work was supported by The Scientific and Technical Research Council of Turkey (TÜB İ TAK, MISAG-242).

References

1. Heskins M, Guillet JE (1968) J Macromol Sci Chem 2:1441–1455
2. Chiklis CK, Grasshoff JM (1970) J Polym Sci A; Polym Chem 8:1617–1626
3. Kujawa P, Winnik FM (2001) Macromolecules 43:4130–4135
4. Elkamel A, Abdul-Wahab S, Bouhamra W, Alper E (2001) Adv Environ Res 5:47–59
5. Baughman DR, Liu YA (1990) Neural networks in bioprocessing and chemical engineering. Academic, New York
6. Agatonovic-Kustrin S, Beresford R (2000) J Pharm Biomed Anal 22:717–727
7. Pollard JF, Broussard MR, Garrison DB, San KY (1992) Comput Chem Eng 16:253–270
8. Churchland PS, Sejnowski TJ (1992) The computational brain. MIT Press, Cambridge
9. Schalkoff RJ (1997) Artificial neural networks. McGraw-Hill, New York
10. Basheer IA, Hajmeer M (2000) J Microbiol Methods 43:3–31
11. Elkamel A, Kargoub M, Gharbi R (1996) Comp Chem Eng 20:515–520
12. Quantrille TE, Liu YA (1991) Artificial intelligence in chemical engineering. Academic, New York
13. Haykin S (1999) Neural networks: a comprehensive foundation. Prentice Hall, Englewoodcliffs
14. Swingler K (1996) Applying neural networks: a practical guide. Academic, New York
15. Kratz K, Hellweg T, Eimer W (2000) Colloids Surf, A 170:137–149
16. Hagan MT, Menhaj M (1994) IEEE Trans Neural Netw 5:989–993
17. Demuth H, Beale M (2003) Neural network toolbox for use with matlab. The Mathworks Inc