CHEMICAL ENGINEERING TRANSACTIONS Volume 17, 2009
Editor Sauro Pierucci
 ISBN 978-88-95608-01-3 ISSN 1974-9791
DOI: 10.3303/CET0917279

An Artificial Neural Network For The Ultrafiltration Of B-Lactoglobulin In Different Chemical Enviroments

R. Ibáñez, A. Guadix, M.C. Almécija, E.M. Guadix*

Department of Chemical Engineering, University of Granada, 18071 Granada, Spain
eguadix@ugr.es

In this research work, experimental data were collected by carrying out the ultrafiltration of β-lactoglobulin through a 300 kDa tubular ceramic membrane at different values of pH and ionic strength. Then, it was developed a feedforward neural network model that receives as inputs the operational time (0 - 60 min), pH (3 - 12) and ionic strength (0 - 20 mM) and returns as outputs the filtrate flow and the protein transmission. Excellent fitting was obtained by using the Levenberg-Marquardt training algorithm with a network including 10 neurons in the hidden layer.

1. Introduction

Protein fractionation is an important task for the bioprocess engineer, so the development of effective isolation processes is of high scientific and industrial interest. However, traditional methods used at laboratory scale find a difficult implementation in the industry ought to their complex scale-up and high cost. As a solution to this problem, the use of membranes for protein fractionation has generated a great expectation during the last years because they present significant advantages with respect to other techniques, namely, use of mild operation conditions, absence of contaminants in the final product and easy scale-up. Since the development of the high performance tangential flow filtration (Van Reis et al., 1997), appropriate strategies like an accurate selection of pH and ionic strength allow the fractionation of proteins even with similar molecular weights. In this technique, the differences in hydrodynamic volume of the proteins are maximized, while the electrostatic interactions between protein and membrane are controlled. In order to build a knowledge base prior to the filtration of multicomponent protein mixtures, much of the research effort has been focused on the filtration of individual protein solutions, as studied by Rabiller-Baudry et al. (2000), de la Casa et al. (2007) or Ibáñez et al. (2007).

Apart from phenomenological approaches to modeling, artificial neural networks have been useful recently for the empirical discussion of membrane operations, mostly motivated by the availability of a large amount of experimental data but the absence of a complete mechanistic understanding due to the complexity of the process considered. Since the work by Wessling et al. (1994), a number of authors have investigated the applicability of neural networks to describe membrane processes, for instance, Bowen et

Please cite this article as: Ibanez R., Guadix A.M., Almecija M.C. and Guadix E.M., (2009), An artificial neural network for the ultrafiltration of b-lactoglobulin in different chemical enviroments, Chemical Engineering Transactions, 17, 1669-1674 DOI: 10.3303/CET0917279

al. (1998), Chellam (2005), Curcio et al. (2005) and Fu et al. (2005). In this context, the purpose of this research work was to develop an artificial neural network that predicts the time evolution of observed transmission and filtrate flow in the ultrafiltration of a single protein, β-lactoglobulin, through a 300 kDa membrane at different pH and ionic strength values.

2. Materials and methods

2.1 Experimental configuration

The protein selected for this research work was β-lactoglobulin (monomer molecular weight 18.3 kDa, isoelectric point 5.2) obtained from Sigma (St. Louis MO, USA). The ceramic membrane employed was an Inside Céram 3-channel module from Tami Industries (Nyons, France). This membrane is tubular with a length of 25 cm, a surface area of 94 cm^2 and a molecular weight cut off of 300 kDa. The support is made of aluminum, titanium and zirconium oxides, whereas titanium oxide is used for the filtering layer.

The experimental rig consisted of a 1 L feed tank immersed in a thermostatic bath at 30 ºC, a precision positive displacement recirculation pump (Procon, Murfreesboro TN, USA), a membrane housing, one back-pressure valve, two manometers located before and after the membrane, a flowmeter (Badger Meter, Milwaukee WI, USA) and a temperature probe. Prior to operation, the membrane was conditioned by recirculation of Milli-Q water at 30 ºC and working pH for 10 min. Then, 1 L of protein solution was prepared with a total protein concentration of 0.25 g/L using Milli-Q water. In order to study the influence of pH, this variable was adjusted adding HCl or NaOH in the range 3-12. To determine the effect of the ionic strength, the pH was set to the isoelectric point and the ionic strength was adjusted adding NaCl in the range 0-20 mM. The cross-flow ultrafiltration experiments were performed at 30 ºC in the total recycle mode (both retentate and permeate were returned to the feed tank), at a transmembrane pressure of 1 bar and a cross-flow velocity of 3.3 m/s. Permeate flow was monitored along the experiments, which were followed for 60 min. Samples from filtrate and retentate were analyzed by UV spectrophotometry with detection at 280 nm in order to determine protein concentration. The membranes were regenerated after use by performing a cleaning procedure as follows: (a) initial rinse with demineralized water, (b) total recirculation of a solution of 20 g/L NaOH + 2 g/L SDS at 50 ºC for 60 min, (c) final rinse with demineralized water until neutrality.

2.2 Neural network modeling

The implementation of the artificial neural network modelling was done by employing the Matlab 6.5 Neural Network Toolbox (Demuth et al., 2008). The system studied comprised 3 input variables (pH, ionic strength and filtration time) and 2 outputs (observed transmission and filtrate flow). Feedforward networks with a single hidden layer including a number of neurons between 5 and 60 were tested. The sigmoid function was selected as transfer function in the hidden layer, while a saturated symmetric lineal function was used for the output layer. In order to minimize the mean squared error (MSE) between experimental and predicted values, the networks were trained with the Levenberg-Marquardt algorithm, allowing a maximum of 10000

iterations. Early stopping was employed for improving generalization. After normalization, the available dataset was randomly divided into three subsets: training (70 %), validation (15 %) and test set (15 %). The training set is used for computing the gradient and updating the network weights and biases. The validation error is used to avoid overfitting because when this fact takes place, this error begins to rise: when the validation error increases for a specified number of iterations (10), the training is stopped, and the weights and biases at the minimum of the validation error are returned. The test set error is not used during the training, but it is used to evaluate the predictive capability of the network. This error is also useful to know if a good division of the data set has been done.

3. Results and discussion

In spite of the deterministic nature of the Levenberg-Marquardt algorithm, the initial values of weights and biases are assigned randomly which suggests that a network has to be trained a number of times high enough to obtain a representative population. Therefore, 30 training cycles were performed for each network. As a result of each training cycle, relevant parameters were obtained such as the training, validation and test errors, as well as those from a linear fit of calculated against experimental data (slope and y-intercept, -ideally, equal to the 1 and 0, respectively). Finally, the average value of each parameter was obtained after removing the outliers, which were defined as those values that were numerically distant from the original mean more than 3 standard deviations. In Fig. 1, it is represented the evolution of the average errors with respect to an increasing number of neurons in the hidden layer. It can be seen that the training error decreased from 0.01 at 5 neurons to 0.003 at 20 neurons, where it reached its minimum. Then, an increase was observed until 30 neurons, where a stable value around 0.004 was achieved. Regarding the validation and test error, their minima are obtained at 10 and 5 neurons, respectively, showing an increasing trend afterwards, which eventually multiplies the minimum value by a factor of 5.

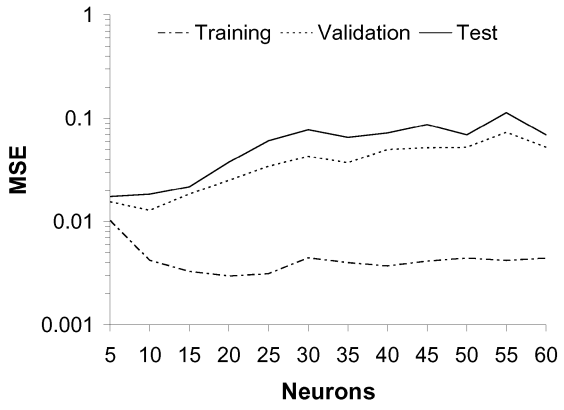

Fig. 1. Average values of training, validation and test errors as a function of the number of neurons in the hidden layer.

Taking into account the parameters obtained from the linear fits (Fig. 2), it can be said that satisfactory results were obtained for both output variables. The slope in the case of permeate flow, which is more affected by a change in the number of neurons than the transmission, presented its best values between 10 and 25 neurons. For transmission, slope values above 0.98 were achieved for hidden layers with up to 30 neurons. The behavior of the intercept is quite similar, since the fit of the permeate flow is again more sensitive to the network size and optimal values were located in equivalent ranges.

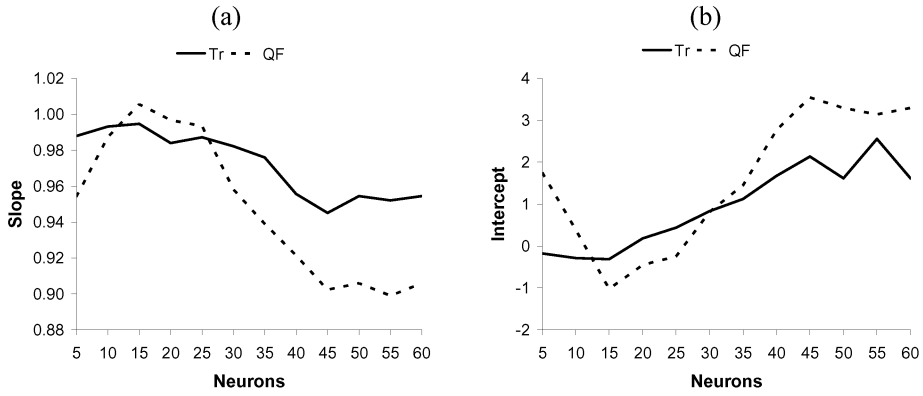

Fig. 2. Influence of the number of neurons in the hidden layer on the average values of the (a) slope and (b) intercept of the linear function that fits calculated vs experimental data for transmission (Tr) and filtrate flow (QF).

In order to select an optimal network architecture (i.e. the number of neurons in the hidden layer) a weighted sum of errors (Fig. 3) was defined by multiplying the training, validation and test errors by the respective fraction of the dataset.

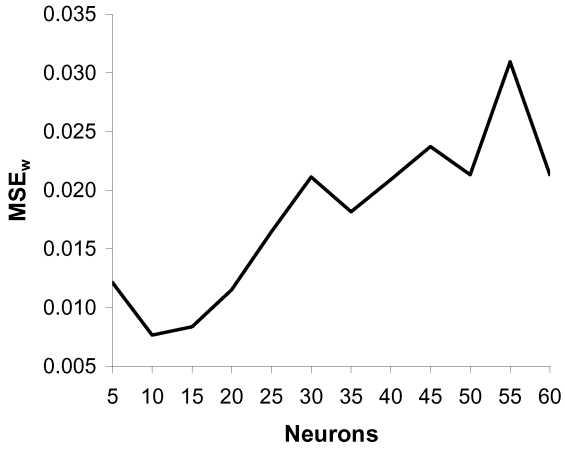

Fig. 3. Weighted average error as a function of the number of neurons in the hidden layer.

A minimum value of the weighted error equal to 0.0076 resulted for a network with 10 neurons, while the worst value (4-fold of the minimum) was for the 55-neuron network. As an illustrative example, it is shown in Fig. 4 the evolution of the errors during one single application of the Levenberg-Marquardt algorithm to a 10-neuron network. Final values of 0.0036, 0.0063 and 0.0047 were calculated for training, validation and test errors, respectively, after 63 epochs or iterations.

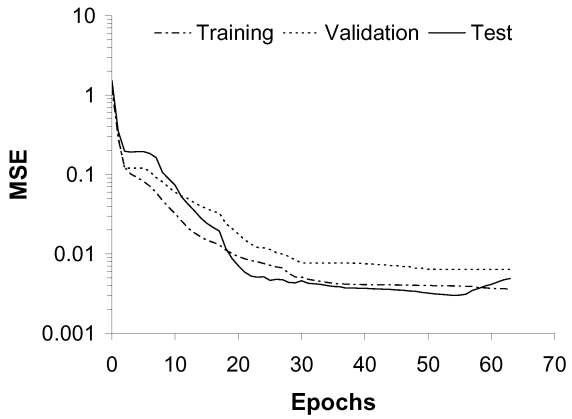

Fig. 4. Evolution of the training, validation and test errors as a function of the number of epochs for a neural network with 10 neurons in the hidden layer.

For the same single application, the linear fit of calculated vs experimental is plotted in Fig. 5, in which the dashed lines represent deviations of ±10 %. The excellent correlation obtained, especially in the case of permeate flow, demonstrates the applicability of the approach proposed.

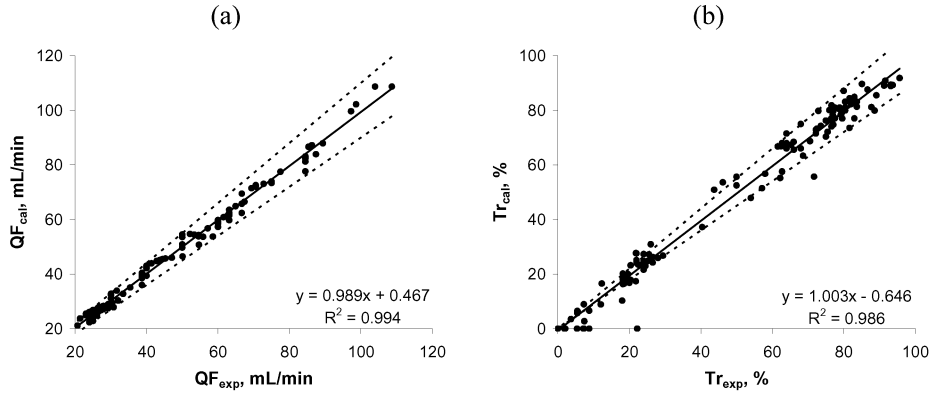

Fig. 5. Calculated vs experimental values of (a) filtrate flow and (b) transmission for a neural network with 10 neurons in the hidden layer.

4. Conclusion

The fractionation of proteins by membrane ultrafiltration is a useful but complex method for the biotechnology industry. The lack of a comprehensive mechanistic model that explains the effect of the chemical environment in the evolution of the process variables, suggests the employment of empirical modeling techniques such as artificial neural networks. The network proposed in this work was aimed at the dynamic simulation of permeate flow and protein transmission of β-lactoglobulin through a 300 kDa membrane, given the values of solution pH and ionic strength. Matlab Neural Network Toolbox, which incorporates the Levenberg-Marquardt algorithm, was employed as computational software. As a main result, when the size of the hidden layer was 10 neurons, an optimal agreement between calculated and experimental data was found.

References

Bowen, W.R., Jones, M.G. and Yousef, H.N.S., 1998, Dynamic ultrafiltration of proteins - A neural network approach, J. Membr. Sci. 146, 225-235.

Chellam, S., 2005, Artificial neural network model for transient crossflow microfiltration of polydispersed suspensions, J. Membr. Sci. 258, 35-42.

Curcio, S., Scilingo, G., Calabro, V. and Iorio, G., 2005, Ultrafiltration of BSA in pulsating conditions: An artificial neural networks approach, J. Membr. Sci. 246, 235-247.

de la Casa, E.J., Guadix, A., Ibáñez, R. and Guadix, E.M., 2007, Influence of pH and salt concentration on the cross-flow microfiltration of BSA through a ceramic membrane, Biochem. Eng. J. 33, 110-115.

Demuth, H., Beale, M. and Hagan, M., 2008, Matlab Neural Network Toolbox User's Guide. MathWorks, Natick MA, USA.

Fu, R.Q., Xu, T.W. and Pan, Z.X., 2005, Modelling of the adsorption of bovine serum albumin on porous polyethylene membrane by back-propagation artificial neural network, J. Membr. Sci. 251, 137-144.

Ibáñez, R., Almécija, M.C., Guadix, A. and Guadix, E.M., 2007, Dynamics of the ceramic ultrafiltration of model proteins with different isoelectric point: Comparison of β-lactoglobulin and lysozyme, Sep. Purif. Technol. 57, 314-320.

Rabiller-Baudry, M., Chaufer, B., Aimar, P., Bariou, B. and Lucas, D., 2000, Application of a convection-diffusion-electrophoretic migration model to ultrafiltration of lysozyme at different pH values and ionic strengths, J. Membrane Sci. 179, 163-174.

Van Reis, R., Gadam, S., Frautschy, L.N., Orlando, S., Goodrich, E.M., Saksena, S., Kuriyel, R., Simpson, C.M., Pearl, S. and Zydney, A.L., 1997, High performance tangential flow filtration, Biotechnol. Bioeng. 56, 71-82.

Wessling, M., Mulder, M.H.V., Bos, A., Van Der Linden, M., Bos, M. and Van Der Linden, W.E., 1994, Modelling the permeability of polymers: A neural network approach, J. Membr. Sci. 86, 193-198.