# Artificial neuronal networks (ANN) to model the hydrolysis of goat milk protein by subtilisin and trypsin

**Short title:** ANN modelling of the hydrolysis of goat milk protein

F. Javier Espejo-Carpio, Raúl Pérez-Gálvez*, Antonio Guadix, Emilia M. Guadix

Department of Chemical Engineering, University of Granada, 18071 Granada, Spain

## SUMMARY

The enzymatic hydrolysis of milk proteins yield final products with improved properties and reduced allergenicity. The degree of hydrolysis (DH) influences both technological (e.g., solubility, water binding capacity) and biological (e.g., Angiotensin-converting enzyme (ACE) inhibition, antioxidation) properties of the resulting hydrolysate. Phenomenological models are unable to reproduce the complexity of enzymatic reactions in dairy systems. However, empirical approaches offer high predictability and can be easily transposed to different substrates and enzymes. In this work, the DH of goat milk protein by subtilisin and trypsin was modelled by feedforward artificial neural networks (ANN). To this end, we produced a set of protein hydrolysates, employing various reaction temperatures and enzyme/substrate ratios, based on an experimental design

The time evolution of the DH was monitored and processed to generate the ANN models. Extensive hydrolysis is desirable because a high DH enhances some bioactivities in the final hydrolysate, such as antioxidant or antihypertensive. The optimisation of both ANN models led to a maximal DH of 23.47% at 56.4ºC and enzyme-substrate ratio of 5% for subtilisin, while hydrolysis with trypsin reached a maximum of 21.3% at 35ºC and an enzyme-substrate ratio of 4%.

**Keywords**: enzymatic hydrolysis; goat milk hydrolysates; artificial neural networks; proteases; optimization

* Corresponding author: Tel.: +34 958 240532; Fax: +34 958 248992; Email: rperezga@ugr.es

## 1. INTRODUCTION

Enzymatic hydrolysis of proteins is an important process in the food industry that improves the functional properties of proteins, reduces potential allergenicity and releases peptides displaying a number of biological activities (Tavano, 2013). Food protein hydrolysates present improved properties such as solubility, emulsifying capacity, foaming ability, water or oil holding capacities, related to crude proteins (García-Moreno et al., 2016; Muro Urista et al., 2011). Moreover, these can be incorporated into nutraceutical formulations where they exert certain biological reactions, including antimicrobial, antioxidant and antihypertensive activities (Capriotti et al., 2016). Among a wide range of substrates, cow milk protein hydrolysates have been the subject of extensive research, while goat milk protein hydrolysates have only recently been shown to exhibit functional and bioactive properties (Bernacka, 2011; El-Salam and El-Shibiny, 2013). Serine endopeptidases, such as subtilisin and tripsin, are usually employed in the hydrolysis of food proteins. Particularly, while subtilisin is able to attack a wide range of peptide bonds, trypsin preferentially cleaves at arginine and lysine residues. Both enzymes have been used for producing peptides displaying biological activities such as antioxidant (Pihlanto, 2006), antihypertensive (López-Fandiño et al., 2006) or antimicrobial (Gobbetti et al., 2004). Moreover, these enzymes yield protein hydrolysates with improved technological properties such as solubility, emulsifying and foaming capacity (Severin and Xia, 2006; Van der Ven et al., 2001).

Many functional and biological properties of protein hydrolysates are related to their degree of hydrolysis (DH). For example, emulsifying and foaming capacities present a maximum at a specific degree of hydrolysis, and, if this is exceeded, these properties are reduced (de Castro et al., 2015). An extensive DH exerts a positive effect on antihypertensive activity because most of the active peptides have chain lengths shorter than 12 amino acids (Li et al., 2004; Phelan and Kerins, 2011). Similarly, extensive hydrolysis of milk proteins can reduce allergenicity significantly for use in infant formulas (Duan et al., 2014; Dupont et al., 2015) .

It can be concluded the extent of the hydrolysis reaction is a key parameter which should be controlled and predicted accurately to obtain hydrolysates with specific characteristics. Mechanistic approaches fail to describe the complexity of various proteins present in milk and different reactions that occurs during milk hydrolysis (e.g., product inhibition, enzyme thermal denaturation)(Ba and Boyaci, 2007). In this context, methods based on direct analysis of experimental data using response surfaces or artificial neural networks, are a suitable alternative to those based on phenomenological hypotheses. These empirical methods are applicable to all types of enzymatic reactions and do not require kinetic assumptions (Baş et al., 2007).

In particular, response surface methodology is widely used for modelling and optimisation purposes, for which the response of interest is influenced by several variables. However, this method is limited in most cases by the use of polynomial equations. Instead, ANNs can be used to ensure better data fit and estimation capabilities (Ba and Boyaci, 2007; Fatiha et al., 2013) .

ANN is composed of individual processing elements (i.e., neurons) that transform weighted input variables into an output by means of an activation function. ANNs comprise one or more hidden layers of neurons. A key element of this approach is the training algorithm, which allows to update the weights and biases of the neurons to obtain outputs closer to the targets. This training consists of minimizing the average squared error (MSE) between the calculated values and the experimental data.

The ANNs have been successfully employed for modelling enzymatic reactions. For example, Bryjak et al. (2000) applied ANNs to model starch hydrolysis by glucoamylase, while Bas et al. (2007) studied the reaction rates of maltose hydrolysis by amyloglucosidase. As for the protein hydrolysis, Abakarov et al. (2011) satisfactorily modelled the kinetics of enzymatic hydrolysis of squid protein with subtilisin using the reaction time and the substrate concentration as input variables. Bucinski et al. (2008) and Li et al. (2016) evaluated the variation of DH during the hydrolysis of bovine hemoglobin and pea proteins, respectively. Li et al. (2006) developed a predictive model for the production of antioxidant peptides from fish proteins taking into account a number of input variables such as pH, temperature, hydrolysis time, muscle/water ratio and enzyme/substrate ratio. Regarding milk proteins, Pinto et al. (2007) proposed a hybrid neural-kinetic model for predicting the molecular mass distribution of whey protein hydrolysates.

The aim of this study was to develop two ANN models for the enzymatic hydrolysis of goat milk proteins, employing either subtilisin or trypsin as catalysts. For each model, the DH was modelled as a function of temperature, enzyme-substrate ratio and the reaction time. Firstly, the architecture of the neural network (i.e., number of neurons in the hidden layer) and the training algorithm were chosen to maximise the degree of fitness (i.e., mean squared error) of the model. Both ANN models were then optimised for the maximal DH, which is desirable because it improves ACE inhibitory and antioxidant activities of the resulting hydrolysates.

# 2. MATERIALS AND METHODS

## 2.1. Materials

Commercial UHT goat milk (33 g protein/L) was purchased from local store. The enzymes used for the assays were subtilisin (EC 3.4.21.62) and trypsin (EC 3.4.21.4), both supplied by Novozymes (Denmark).

### 2.2. Enzymatic reaction and determination of the degree of hydrolysis

Before hydrolysis, the milk was skimmed by centrifugation at 4 ºC and 5000 g for 20 min in a Sigma 6k15 centrifuge (Sigma Laborzentrifugen, Germany). Skimmed goat milk (200 mL) was then hydrolysed in a stirred tank reactor for 5 hours. Initially, the pH of the milk was set at pH 8 with 1M NaOH. After reaching the desired temperature, the enzyme was added at different enzyme-substrate ratio. In alkaline medium, the cleavage of peptide bonds releases protons which cause the pH to drop. An automatic titrator was employed (718 Stat Titrino, Metrohm, Switzerland) to keep the pH constant during the reaction by adding NaOH (1M). The degree of hydrolysis (DH), defined as the percentage of available peptide bonds which are cleaved during the reaction, can be related to the amount of base consumed by Eq. 1(Adler Nissen, 1986):

$$DH = n_B / \left( \alpha \cdot m_P \cdot h_{TOT} \right) \quad (1)$$

where DH is the degree of hydrolysis, $n_B$ (mol) is the amount of NaOH consumed to keep the pH constant, $\alpha$ is the average degree of dissociation of $\alpha$-$NH_2$ groups released during hydrolysis, $m_P$ = 6.6 g is the mass of protein in the substrate and $h_{TOT}$ = 0.0082 eq/g is the average number of equivalents of peptide bonds per gram of casein protein.

### 2.3. Experimental design

A total of 60 hydrolysates were produced, broken down into two factorial designs of 30 experiments where subtilisin or trypsin were employed as catalysts. Each hydrolysate was produced at a given combination of reaction temperature (T) and enzyme-substrate ratio (ES), which were the input variables of the factorial designs. The reaction temperature was varied at six levels according to the thermal stability of the enzyme assayed. Subtilisin exhibits wide thermal stability, presenting optimal activity around 50-55ºC. As for trypsin, it presents maximal activity around 40ºC (Adler Nissen, 1986). Therefore, subtilisin was tested at 45, 50, 55, 60, 65 and 70 ºC, while trypsin was at 30, 35, 40, 45, 50 and 55 ºC. The levels assayed for the enzyme to substrate ratio were 1, 2, 3, 4 and 5 % for both enzymes. As for the time of reaction (t), the DH value was recorded every 60 seconds over the course of the reaction (5 hours). This yields an amount of 300 experimental data (T, ES, t, DH) for each hydrolysis curve.

## 2.3. Structure and training of the artificial neuronal network

Two artificial neural network (ANN) models were developed in this work (i.e. subtilisin and trypsin), where DH was related to the reaction temperature (T), the enzyme-substrate ratio (ES) and the time of reaction (t) as input variables. Both ANN models were constructed by means of the Neural Network Toolbox, implemented in Matlab 7.0 (Mathworks, USA).

Both artificial neural networks comprised an input layer, a single hidden layer and an output layer. The input layer comprised three neurons, corresponding to the 3 input variables (T, ES, t). This layer is connected to the hidden layer, whose number of neurons was varied from 1 to 10 neurons. Each neuron *k* of the hidden layer received a weighted signal from the input layer $s_k$, expressed as follows:

$$s_k = \sum_{i=1}^{3} w_{ik} \cdot X_i + b_k \qquad (2)$$

where $w_{ik}$ were the weight factors and $b_k$ was the bias for the neuron *k*. Each neuron of the hidden layer processes the signal $s_k$ by means of a transfer function. The sigmoid function (implemented in Matlab as *logsig*) was selected as transfer function in the hidden layer, which returns a value ranging between 0 and 1 according to Eq. 3:

$$\text{logsig}\ (s_k) = \frac{1}{1 + \exp\left(-s_k\right)} \qquad (3)$$

The *k* responses exiting the hidden layer are combined into a single weighted signal *t*, which is received by an output neuron, which returns the predicted value of DH. The saturated symmetric lineal function was chosen as transfer function for the output layer. This function truncates the weighted signal *t* within the interval [0,1], avoiding either negative DH values or above 1.

Three training algorithms were tested in this work: gradient descent with momentum backpropagation (*traingdm*), resilient backpropagation (*trainrp*) and Levenberg-Marquardt backpropagation (*trainlm*). These algorithms update the weight and bias values in order to minimize the mean squared error (MSE) between observed and predicted DH.

For a fixed number of hidden neurons in the hidden layer and training algorithm (traingdm, trainrp and trainlm), 30 runs were carried, ensuring an appropriate population of predicted data. At the beginning of each run, the dataset was normalized and then randomly divided into three subsets:

training, validation and test. The biggest subset (70 % of the total amount of experimental data) was used for training the network using the algorithm selected. During the training, the error obtained from the validation set (15% of the data) was employed for early stopping (i.e. interruption of the iteration process when over-fitting in the training dataset is detected). In back-propagation methods, over-fitting occurs when an improvement in the fit of the training data is accompanied by larger generalization errors. The number of iterations per training run was limited to 10000. As an early stopping criterion, the training process stopped when the MSE increased for 10 iterations. At this point, the algorithm returned the weights and biases corresponding to the minimal MSE recorded so far. Finally, the remaining data (15%) are employed to compute the test error, which assesses the predictive capability of the network. This error is also useful to know if a good division of the data set (i.e. training, validation and evaluation subset) has been done.

## 2.4. ANN model for DH and optimization procedure

The objective of the ANN procedure was to obtain a predictive model of DH for each of the enzymes employed. Each model allowed the calculation of DH as a function of the experimental conditions of temperature, enzyme-substrate ratio and time of reaction, as expressed by Eq. 4:

$$DH = \sum_{k=1}^{N} \omega_k \cdot logsig\left( \sum_{i=1}^{3} w_{ki} \cdot X_i + b_k \right) + \beta \qquad (4)$$

X (T, eS, t) denotes the vector of input variables (i.e, the experimental conditions for each hydrolysis assay); wki and bk are the weight factors and bias of the input layer, respectively; ω and β are the weights and bias of the hidden layer and the transfer function logsig was defined by Eq. 3.

The training procedure allowed the estimation of the set of parameters $w_k$, $b_k$, ω and β yielding the minimal squared error between predicted and observed DH (i.e., the best fit between the experimental DH and the predictive model).

The goal of the optimisation problem was to find the set of experimental conditions X (T, eS, t), within their experimental range, which maximises DH calculated by Eq. 4. To this end, the Generalized Reduced Gradient (GRG), implemented in the Solver tool of the MS Excel, was chosen for the optimization of both models. GRG is a non-linear optimisation algorithm, which basically evaluates the gradient or slope of the objective function (i.e. predicted DH) as the input values (i.e. experimental conditions Xi in Eq. 4) change and determines that it has reached an optimum solution when the partial derivatives equal zero. Since GRP is a local method, the multistart method was chosen to find a globally optimal solution of the problem. This option consists in operating the GRP algorithm from a set of starting points, reaching different local optimums which are then compared to select a global optimum.

# 3. RESULTS AND DISCUSSION

## 3.1. Architecture and training algorithm of the ANN

The time evolution of the DH was modeled by two artificial neural networks, depending on the enzyme employed for the hydrolysis. Each ANN comprised an input layer of three neurons corresponding to each one of the experimental factors (T, ES, t), connected to a hidden layer with a variable number of neurons between 1 and 10. The hidden layer is connected to an output layer with a single neuron, which returns the predicted value for the degree of hydrolysis. An average of 30 simulations was performed by a combination of three training algorithms (i.e. trainrp, traingdp, trainlm) and a fixed number of neurons in the hidden layer (i.e. 1 to 10). Every training procedure was executed 30 times, starting from different initial values of weights and biases. For each trial, the mean squared errors of the training, validation and test subsets were recorded. Average training and validation errors were in all cases very similar to test error values. Indeed, the differences between these errors were below 1 and 2 % for subtilisin and trypsin networks, respectively. Fig. 1 presents the test error of the networks obtained for the hydrolysis with subtilisin (a) and trypsin (b) as a function of the training algorithm and the number of neurons in the hidden layer. For both ANN models, the Levenberg-Marquardt algorithm showed the best performance, followed by *trainrp* and *traingdm*. Indeed, test errors decreased with an increasing number of neurons in the hidden layer for the *trainlm* algorithm, resulting in final MSE values of and $5 \cdot 10^{-4}$ and $10^{-3}$ at 10 neurons for the subtilisin and tripsin network, respectively. Contrarily, the *traingdp* algorithm presented overfitting above 6 neurons for both ANN models. According to these results, the Levengberg-Marquardt training algorithm was chosen to model the degree of hydrolysis for both enzymes.

The predictability of both ANN models and the *trainlm* algorithm was assessed by the slope and intercept of the linear fit between predicted and observed values of DH (Fig. 2). Ideally, the slope and the intercept should be 1 and 0, respectively. In the case of the subtilisin ANN, the network with 2 neurons in the hidden layer led to an average slope (i.e. mean value from 30 trials) above 0.950, which increased up to 0.996 at 8 neurons. This value remained steady in 9 and 10 neurons. The intercept value for the subtilisin ANN was $6 \cdot 10^{-3}$ at 2 neurons and decreased down to $4 \cdot 10^{-4}$ at 10 neurons. Similarly, the average slope and intercept values for the trypsin ANN increased and decreased, respectively, with the number of neurons in the hidden layer. This model reached a maximal slope of 0.984 and a minimal intercept of $2 \cdot 10^{-3}$ at 10 neurons.

In line of the above, two ANN models were proposed to fit the experimental data employing the Levengberg-Marquardt training algorithm. The number of neurons in the hidden layer was fixed at 8 and 10 for the subtilisin and trypsin ANN, respectively. Under these conditions, the predictability of both models (i.e. MSE and slope values) was acceptable while their complexity and time of computation were limited. This algorithm has been successfully employed to model enzymatic processes. For example, Pinto et al. (2007) modeled the molecular weight distribution of whey protein hydrolysates. Feed-forward ANNs trained by the Levensberg-Marquardt algorithm was used to model the time evolution of DH in the hydrolysis of blood protein (Gálvez et al., 2016) and horse mackerel protein (Morales-Medina et al., 2016) with subtilisin. Abakarov et al. (2011) used gradient descent algorithm to predict the hydrolysis of squid protein using substilisin. Similarly, Buciński et al. (2008) use it for hydrolysis of pea protein employing trypsin.

**3.2. ANN models for the hydrolysis with subtilisin and trypsin**

In all the cases, the time evolution of DH followed the characteristic curve described for enzymatic hydrolysis. As an example, Figures 3a and 3b represent the observed values of DH (point markers) against the time of reaction and enzyme-substrate ratio for subtilisin and trypsin at 50ºC. It can be observed that DH presented a sharp linear increase at the beginning of the hydrolysis, followed by slight reduction to achieve steady state. As the proteolysis progresses, the remaining number of peptide bonds available for enzyme attack decreases and so the reaction rate. (Adler Nissen, 1986; Valencia et al., 2014). Depending on its thermal stability, extensive times of reaction at high temperatures may provoke thermal inactivation of the enzyme. The denaturation of the quaternary structure of the enzyme results in a decrease of its enzymatic activity. Finally, some authors such as Valencia et al. (2014) relate the decrease of the reaction rate to the occurrence of product inhibition. In this case, the peptides released during hydrolysis may inhibit the reaction progress by forming stable complexes with substrate or enzyme.

The hydrolysis curves depicted in Figure 3 show that increasing enzyme-substrate ratios improved the final values of DH for the subtilisin reaction. This trend was not clear for the trypsin reactions, where hydrolysis curves at ES 4% and 5% were very close or even overlapped with each other.

As example, at 50 °C, the final DH values observed for substilisin and trypsin were in the range of 18-24%, and 16-23%, respectively. The solid lines in Figure 3 represent the predicted DH calculated from the ANN models presented above. Figure 3 illustrates the high degree of fitting between the observed values of DH and those calculated by ANN modelling. The determination coefficients of the linear fit between experimental and calculated values of DH were $r^2$=0.996 and $r^2$=0.994 for subtilisin and trypsin, respectively.

### 3.3. Optimization of the degree of hydrolysis

The application of the Levengsberg-Marquardt algorithm allowed estimating the weights and biases of both ANN models (Eq. 2) for a fixed number of neurons in the hidden layer. This set of parameters allowed computing DH as a function of the reaction conditions (i.e. T, ES and t). Furthermore, these model were optimized by an evolutionary algorithm to determine the optimal parameters for maximal DH. Extensive hydrolysis seems to enhance a number of biological activities such as the ACE-inhibitory and the antioxidant activities. Some of the most potent ACE inhibitors identified in milk protein hydrolysates correspond to di and tripeptides (Hernández-Ledesma et al., 2014). Similarly, several short peptides (500-1800 Da) have been identified as potent antioxidants (Ahmed et al., 2015; Moreno-Montoro et al., 2017; Samaranayaka and Li-Chan, 2011). Peptide size is also a crucial factor for bioavailability of bioactive peptides. According to the literature, there are no evidence that peptides bigger than tripeptides can move across the tissues of gastrointestinal tract intact and enter into blood stream in required concentrations (Miner-Williams et al., 2014). The contour plots shown in Figures 4a and 4b represent the calculated values of the final DH (5 h) against the reaction temperature and the enzyme-substrate ratio. Both contour plots confirm the positive effect of increasing enzyme-substrate ratio on DH. This trend was clear for the hydrolysis with subtilisin, regardless the reaction temperature. Increasing ES ratios favoured the prote olysis with subtilisin, obtaining maximum DH of 22 – 23% with 5% ES ratio. Optimisation of the ANN model confirmed that maximum DH (23.47%) can be achieved for goat milk proteins using substilisin at 56.4 °C with 5% ES ratio. The optimal reaction temperature is within the range of maximal activity reported for subtilisin (Adler Nissen, 1986; Ma et al., 2015). According to the contour plot, the final values of DH at 5% ES ratio kept above 22% within the experimental range from 45 to 70ºC. This suggests that this enzyme was not significantly affected by thermal deactivation, and therefore its proteolytic activity remained unaltered. This is in line with previous studies which highlight the high resistance of subtilisin against thermal denaturation (Adler-Nissen, 1986; Nagodawithana and Reed, 2013). Subtilisin-like serine proteases contain a variable number (2 to 7) of $Ca^{2+}$-binding sites. Binding of the calcium ions has been reported to greatly stabilise the protein structure against thermal unfolding (Foophow et al., 2010).

In contrast, the contour plots for the hydrolysis with trypsin (Figure 4b) indicated that the final DH was influenced by the reaction temperature. This may resulted from inactivation of trypsin at higher temperature. The optimal conditions for maximal DH (21.3%) were 35ºC and ES 4%. Higher levels of enzyme-substrate ratio did not improve the extent of the hydrolysis, suggesting the saturation of the peptide bonds available. Trypsin exhibits narrow specificity towards arginine and lysine residues (Olsen et al., 2004) , while subtilisin is a wide-spectrum protease. This fact could explain the saturation of available peptide bonds at ES ratios above 4%. The optimal temperature condition calculated for trypsin was 35ºC, similar to the maximum of 37ºC reported in scientific literature (Adler-Nissen, 1986; Morales-Medina et al., 2016).

For validation purposes, the optimal operating conditions for maximal DH were reproduced experimentally for both enzymes. To this end, the predicted optimum conditions for both substilisin (56.4 ºC and ES 5%) and trypsin (35 ºC and ES 4%) were experimentally evaluated and illustrated in the Figure 5. Some of the observed values of DH were represented by point markers, while the curves predicted from the ANN model were depicted as solid lines. Both curves fitted satisfactorily the observed data. This was confirmed by the coefficients of determination $R^2$ for both models, which were 0.9883 and 0.9929 for subtilisin and trypsin, respectively. Moreover, the average deviation between observed and predicted values was 1.9 ± 1.7 % for the hydrolysis with subtilisin and 1.7± 1.6% for trypsin. The verification hydrolysis with subtilisin led to a maximal DH of 23.26% at experimental conditions 56.4ºC, 5% ES ratio and 5 h of reaction, which was similar to the optimum DH (23.47%, 56.4ºC, 5% ES ratio, 5 h) predicted by the proposed ANN model with 8 neurons in the hidden layer. However, maximum DH obtained for trypsin mediated hydrolysis under experimental conditions (35ºC, 4% ES ratio and 5 h of reaction) was 22.17%, which is slightly higher than the predicted value (21.3%) with optimum conditions (35 ºC, ES 4% and 5 h) using the ANN model with 10 neurons.

## 4. CONCLUSIONS

ANN modelling was successfully employed to predict the degree of hydrolysis of skimmed goat milk proteins with subtilisin and trypsin as a function of the operating conditions, namely the reaction temperature, the enzyme-substrate ratio and the time of hydrolysis. The predictability of both ANN models was improved by testing three training algorithm and a variable number of neurons (i.e. 1 to 10) in the hidden layer. In this regard, two ANN models with 8 and 10 neurons in the hidden layer were selected for subtilisin and trypsin hydrolysis, respectively. As for the training algorithm, the Levengsberg-Marquardt led to the minimal test errors (MSE) for both subtilisin and trypsin with determination coefficients of 0.996 and 0.984, respectively. Furthermore, these models were optimized by an evolutionary algorithm to obtain the combination of operating conditions leading to the maximal DH. Maximum DH (23.47%) was calculated for substilisin at 56.4 ºC, ES 5% and 5 h of reaction, while the máximum DH obtained for trypsin was 21.3% at 35 ºC, ES 4% and 5 h reaction. There was a significant correlation between DH predicted using ANN models and DH obtained under experimental conditions.

ANN arises as an alternative to obtain predictive models for protein hydrolysis, especially when a large volume of data is available. The strength of these models, inspired in human brain, lies on its ability to learn from experimental data by training algorithms. By this approach, the model parameters are updated iteratively, until minimizing the error between predicted and actual data. In the field of biochemical processes, this approach allows obtaining predictive models without needing to have extensive knowledge of the underlying mechanism. This is especially useful in enzymatic reactions where several phenomena (i.e. substrate solubilisation, substrate or product inhibition, thermal inactivation of the enzyme) may occur simultaneously.

**Acknowledgments**

This work was funded by the project P07-TEP-02579 from the Consejería de Economía, Innovación, Ciencia y Empleo of Junta de Andalucía, Spain.

## REFERENCES


Abakarov, A., Teixeira, A., Simpson, R., Pinto, M., Almonacid, S., 2011. Modeling of Squid Protein Hydrolysis: Artificial Neural Network Approach. Journal of Food Process Engineering 34, 2026–2046. doi:10.1111/j.1745-4530.2009.00567.x

Adamson, N.J., Reynolds, E.C., 1996. Characterization of casein phosphopeptides prepared using alcalase: Determination of enzyme specificity. Enzyme and Microbial Technology 19, 202–207. doi:10.1016/0141-0229(95)00232-4

Adler Nissen, 1986. Enzymic hydrolysis of food proteins. Elsevier Applied Science Publishers LTD, London.

Ahmed, A.S., El-Bassiony, T., Elmalt, L.M., Ibrahim, H.R., 2015. Identification of potent antioxidant bioactive peptides from goat milk proteins. Food Research International 74, 80–88. doi:10.1016/j.foodres.2015.04.032

Ba, D., Boyaci, I.H., 2007. Modeling and optimization II: Comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction. Journal of Food Engineering 78, 846–854. doi:10.1016/j.jfoodeng.2005.11.025

Baş, D., Dudak, F.C., Boyacı, İ.H., 2007. Modeling and optimization III: Reaction rate estimation using artificial neural network (ANN) without a kinetic model. Journal of Food Engineering 79, 622–628. doi:10.1016/j.jfoodeng.2006.02.021

Bernacka, H., 2011. Health-promoting properties of goat milk. Medycyna Weterynaryjna 67.

Bryjak, J., Murlikiewicz, K., Zbiciński, I., Stawczyk, J., 2000. Application of artificial neural networks to modelling of starch hydrolysis by glucoamylase. Bioprocess Engineering 23. doi:10.1007/s004499900170

Buciński, A., Karamać, M., Amarowicz, R., Pegg, R.B., 2008. Modeling the tryptic hydrolysis of pea proteins using an artificial neural network. LWT - Food Science and Technology 41, 942–945. doi:10.1016/j.lwt.2007.06.021

Capriotti, A.L., Cavaliere, C., Piovesana, S., Samperi, R., Laganà, A., 2016. Recent trends in the analysis of bioactive peptides in milk and dairy products. Analytical and Bioanalytical Chemistry 408. doi:10.1007/s00216-016-9303-8

de Castro, R.J.S., Bagagli, M.P., Sato, H.H., 2015. Improving the functional properties of milk proteins: Focus on the specificities of proteolytic enzymes. Current Opinion in Food Science 1, 64–69. doi:10.1016/j.cofs.2014.12.004

Duan, C.-C., Yang, L.-J., Li, A.-L., Zhao, R., Huo, G.-C., 2014. Effects of enzymatic hydrolysis on the allergenicity of whey protein concentrates. Iranian Journal of Allergy, Asthma and Immunology 13.

Dupont, C., Hol, J., Nieuwenhuis, E.E.S., De Jongste, J.C., Samsom, J.N., Van Leer, E.H.G., Elink Schuurman, B.E.E., De Ruiter, L.F., Neijens, H.J., Versteegh, F.G.A., Groeneweg, M., Van Veen, L.N., Vaessen-Verberne, A.A., Smit, M.J.M., Vriesman, A.W., Roosen, Y.M., Den Exter, G.L., 2015. An extensively hydrolysed casein-based formula for infants with cows' milk protein allergy: Tolerance/hypo-allergenicity and growth catch-up. British Journal of Nutrition 113. doi:10.1017/S000711451500015X

El-Salam, M.H.A., El-Shibiny, S., 2013. Bioactive Peptides of Buffalo, Camel, Goat, Sheep, Mare, and Yak Milks and Milk Products. Food Reviews International 29, 1–23. doi:10.1080/87559129.2012.692137

Fatiha, B., Sameh, B., Youcef, S., Zeineddine, D., Nacer, R., 2013. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in optimization of the immobilization conditions for lipase from Candida rugosa on amberjet® 4200-Cl. Preparative Biochemistry and Biotechnology 43. doi:10.1080/10826068.2012.693899

Foophow, T., Foophow, T., Tanaka, S., Koga, Y., Takano, K., Kanaya, S., 2010. Subtilisin-like serine protease from hyperthermophilic archaeon Thermococcus kodakaraensis with N- and C-terminal propeptides. Protein Engineering, Design and Selection 23, 347–355. doi:10.1093/protein/gzp092

Gálvez, R.P., Carpio, F.J.E., Guadix, E.M., Guadix, A., 2016. Artificial neural networks to model the production of blood protein hydrolysates for plant fertilisation. Journal of the Science of Food and Agriculture 96. doi:10.1002/jsfa.7083

García-Moreno, P.J., Pérez-Gálvez, R., Espejo-Carpio, F.J., Ruiz-Quesada, C., Pérez-Morilla, A.I., Martínez-Agustín, O., Guadix, A., Guadix, E.M., 2016. Functional, bioactive and antigenicity properties of blue whiting protein hydrolysates: Effect of enzymatic treatment and degree of hydrolysis. Journal of the Science of Food and Agriculture. doi:10.1002/jsfa.7731

Geerlings, A., Villar, I.C., Zarco, F.H., Sánchez, M., Vera, R., Gomez, A.Z., Boza, J., Duarte, J., 2006. Identification and characterization of novel angiotensin-converting enzyme inhibitors obtained from goat milk. Journal of Dairy Science 89.

Gobbetti, M., Minervini, F., Rizzello, C.G., 2004. Angiotensin I-converting-enzyme-inhibitory and antimicrobial bioactive peptides. International Journal of Dairy Technology 57.

doi:10.1111/j.1471-0307.2004.00139.x

Hernández-Ledesma, B., García-Nebot, M.J., Fernández-Tomé, S., Amigo, L., Recio, I., 2014. Dairy protein hydrolysates: Peptides for health benefits. International Dairy Journal 38, 82–100. doi:10.1016/j.idairyj.2013.11.004

Li, G.-H., Le, G.-W., Shi, Y.-H., Shrestha, S., 2004. Angiotensin I–converting enzyme inhibitory peptides derived from food proteins and their physiological and pharmacological effects. Nutrition Research 24, 469–486. doi:10.1016/j.nutres.2003.10.014

Li, L., Wang, J., Zhao, M., Cui, C., Jiang, Y., 2006. Artificial neural network for production of antioxidant peptides derived from bighead carp muscles with alcalase. Food Technology and Biotechnology 44, 441–448.

Li, S., Hu, Y., Hong, Y., Xu, L., Zhou, M., Fu, C., Wang, C., Xu, N., Li, D., 2016. Analysis of the Hydrolytic Capacities of Aspergillus oryzae Proteases on Soybean Protein Using Artificial Neural Networks. Journal of Food Processing and Preservation 40, 918–924. doi:10.1111/jfpp.12670

López-Fandiño, R., Otte, J., van Camp, J., 2006. Physiological, chemical and technological aspects of milk-protein-derived peptides with antihypertensive and ACE-inhibitory activity. International Dairy Journal 16. doi:10.1016/j.idairyj.2006.06.004

Ma, Y., Sun, X., Wang, L., 2015. Study on optimal conditions of alcalase enzymatic hydrolysis of soybean protein isolate. Advance Journal of Food Science and Technology 9, 154–158.

Miner-Williams, W.M., Stevens, B.R., Moughan, P.J., 2014. Are intact peptides absorbed from the healthy gut in the adult human? Nutrition Research Reviews 27, 308–329. doi:10.1017/S0954422414000225

Morales-Medina, R., Pérez-Gálvez, R., Guadix, A., Guadix, E.M., 2016. Artificial neuronal network modeling of the enzymatic hydrolysis of horse mackerel protein using protease mixtures. Biochemical Engineering Journal 105. doi:10.1016/j.bej.2015.10.009

Moreno-Montoro, M., Olalla-Herrera, M., Rufián-Henares, J.Á., Martínez, R.G., Miralles, B., Bergillos, T., Navarro-Alarcón, M., Jauregi, P., 2017. Antioxidant, ACE-inhibitory and antimicrobial activity of fermented goat milk: Activity and physicochemical property relationship of the peptide components. Food and Function 8, 2783–2791. doi:10.1039/c7fo00666g

Muro Urista, C., Álvarez Fernández, R., Riera Rodriguez, F., Arana Cuenca, A., Téllez Jurado, A., 2011. Review: Production and functionality of active peptides from milk. Food Science and

Technology International 17. doi:10.1177/1082013211398801

Nagodawithana, T., Reed, G., 2013. Enzymes in Food Processing. Elsevier.

Olsen, J. V, Ong, S.-E., Mann, M., 2004. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. Molecular & cellular proteomics : MCP 3, 608–14. doi:10.1074/mcp.T400003-MCP200

Phelan, M., Kerins, D., 2011. The potential role of milk-derived peptides in cardiovascular disease. Food and Function 2. doi:10.1039/c1fo10017c

Pihlanto, A., 2006. Antioxidative peptides derived from milk proteins. International Dairy Journal 16, 1306–1314. doi:10.1016/j.idairyj.2006.06.005

Pinto, G.A., Giordano, R.L.C., Giordano, R.C., 2007. Neural Network Inference of Molar Mass Distributions of Peptides during Tailor-Made Enzymatic Hydrolysis of Cheese Whey: Effects of pH and Temperature. Applied Biochemistry and Biotechnology 143, 142–152. doi:10.1007/s12010-007-0039-y

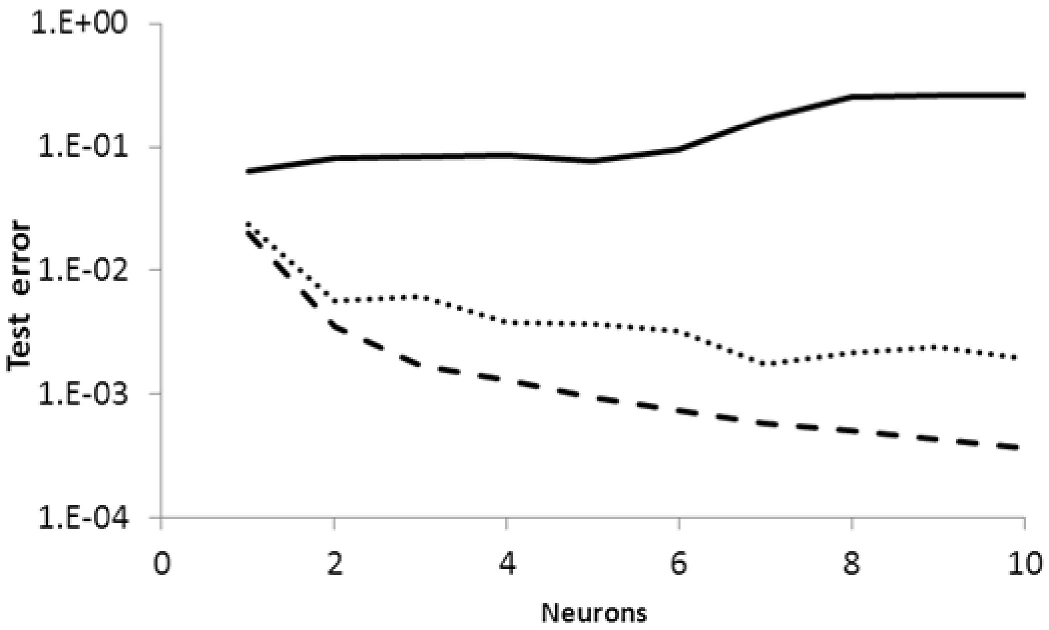
Samaranayaka, A.G.P., Li-Chan, E.C.Y., 2011. Food-derived peptidic antioxidants: A review of their production, assessment, and potential applications. Journal of Functional Foods 3, 229–254. doi:10.1016/j.jff.2011.05.006

Severin, S., Xia, W.S., 2006. Enzymatic hydrolysis of whey proteins by two different proteases and their effect on the functional properties of resulting protein hydrolysates. Journal of Food Biochemistry 30. doi:10.1111/j.1745-4514.2005.00048.x

Tavano, O.L., 2013. Protein hydrolysis using proteases: An important tool for food biotechnology. Journal of Molecular Catalysis B: Enzymatic 90, 1–11. doi:10.1016/j.molcatb.2013.01.011

Van der Ven, C., Gruppen, H., De Bont, D.B.A., Voragen, A.G.J., 2001. Emulsion properties of casein and whey protein hydrolysates and the relation with other hydrolysate characteristics. Journal of Agricultural and Food Chemistry 49, 5005–5012. doi:10.1021/jf010144c

**(a)**




**(b)**

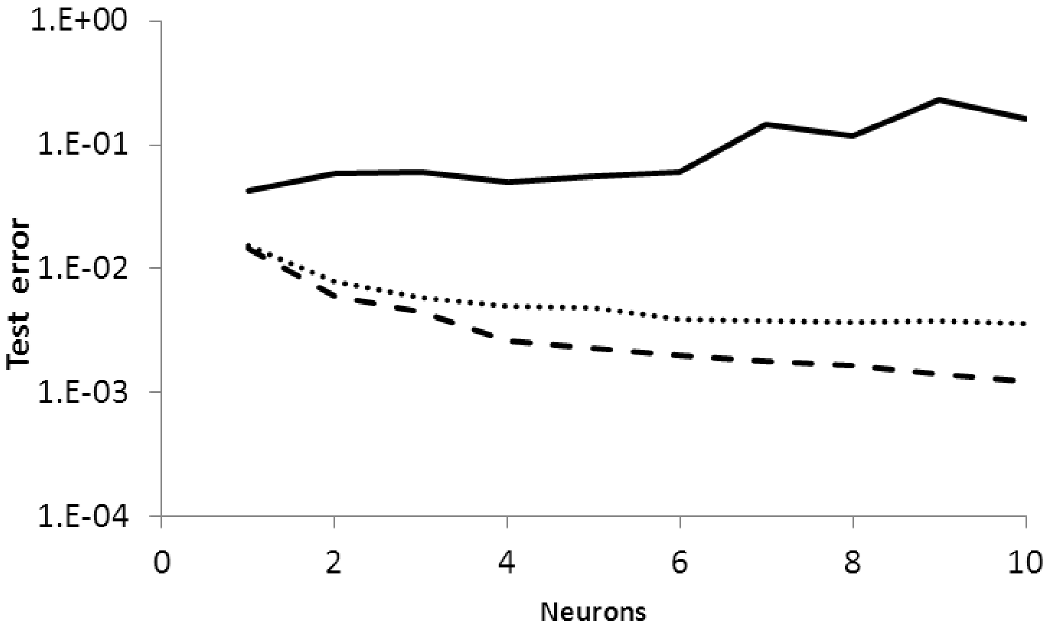
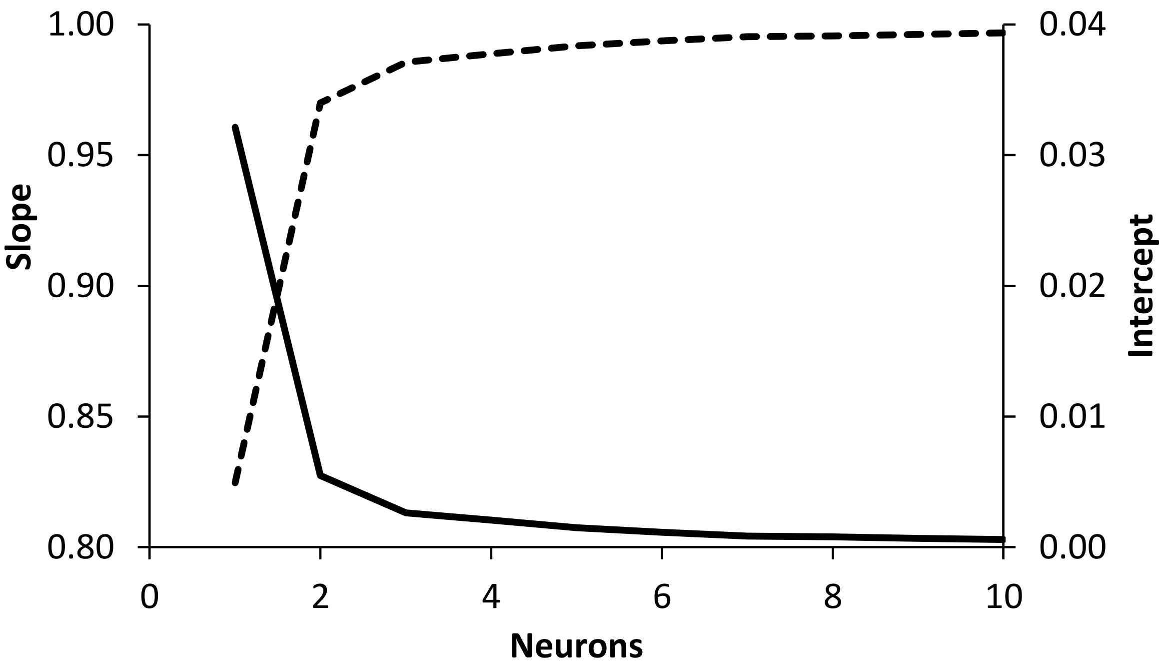



**Figure 1**. Test error as a function of the number of neurons in the hidden layer for (a) subtilisin and (b) trypsin. Test errors reported as average of 30 trials trained by gradient descent with momentum (solid line), resilient backpropagation (dotted line) and Levensberg-Marquardt algorithm (dashed line).

**(a)**




**(b)**

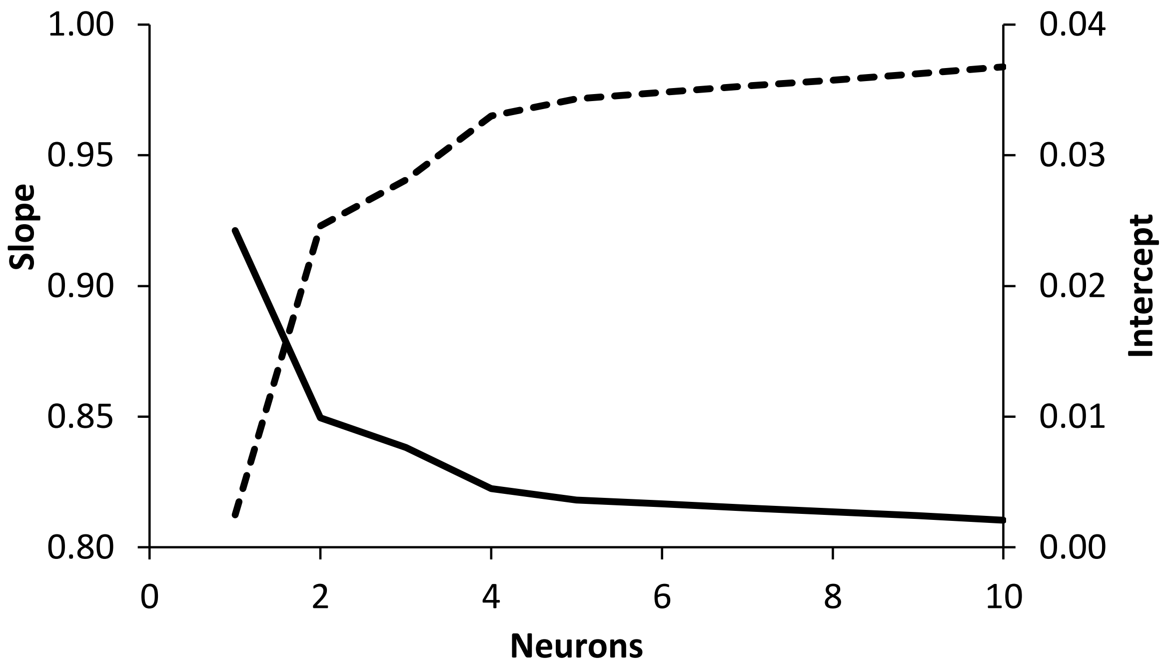
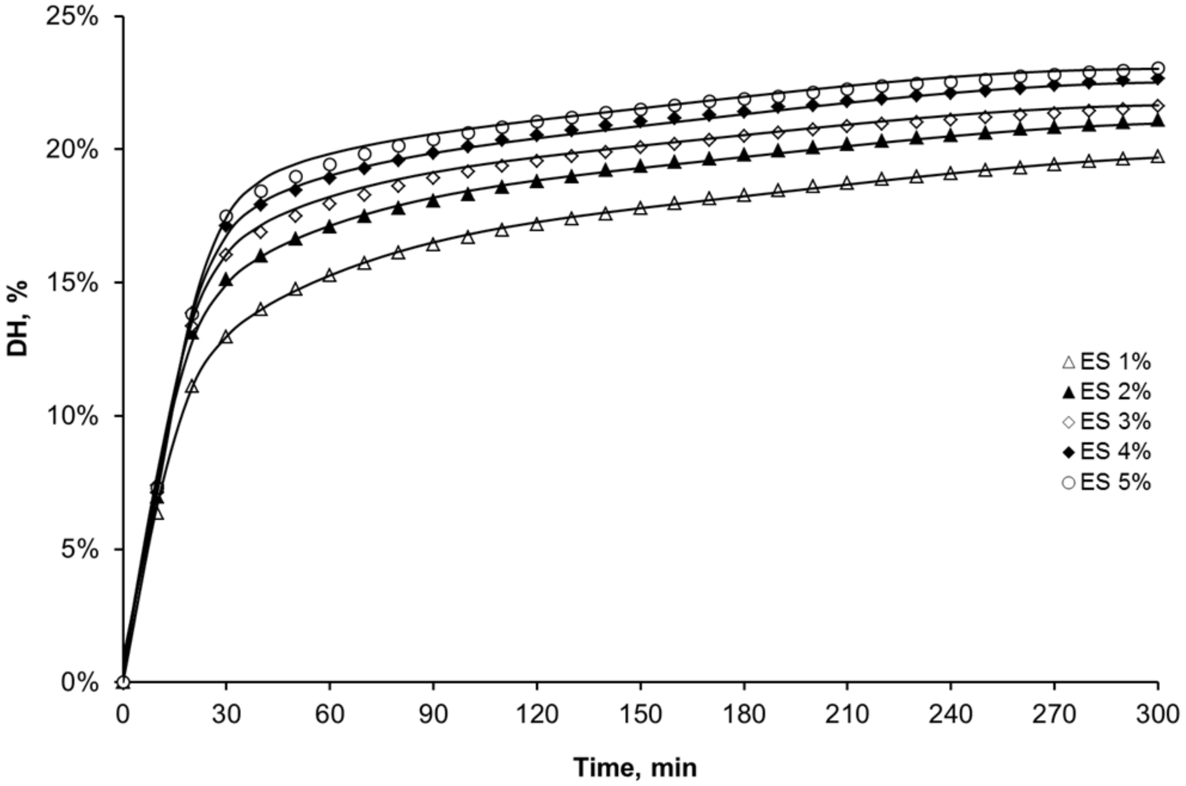



**Figure 2.** Slope (dashed line) and intercept (solid line) of the linear fit of experimental against calculated DH as a function of the number of neurons in the hidden layer for (a) subtilisin and (b) trypsin. Slope and intercept reported as average of 30 trials trained by the Levengsberg-Marquardt algorithm.

**(a)**




**(b)**

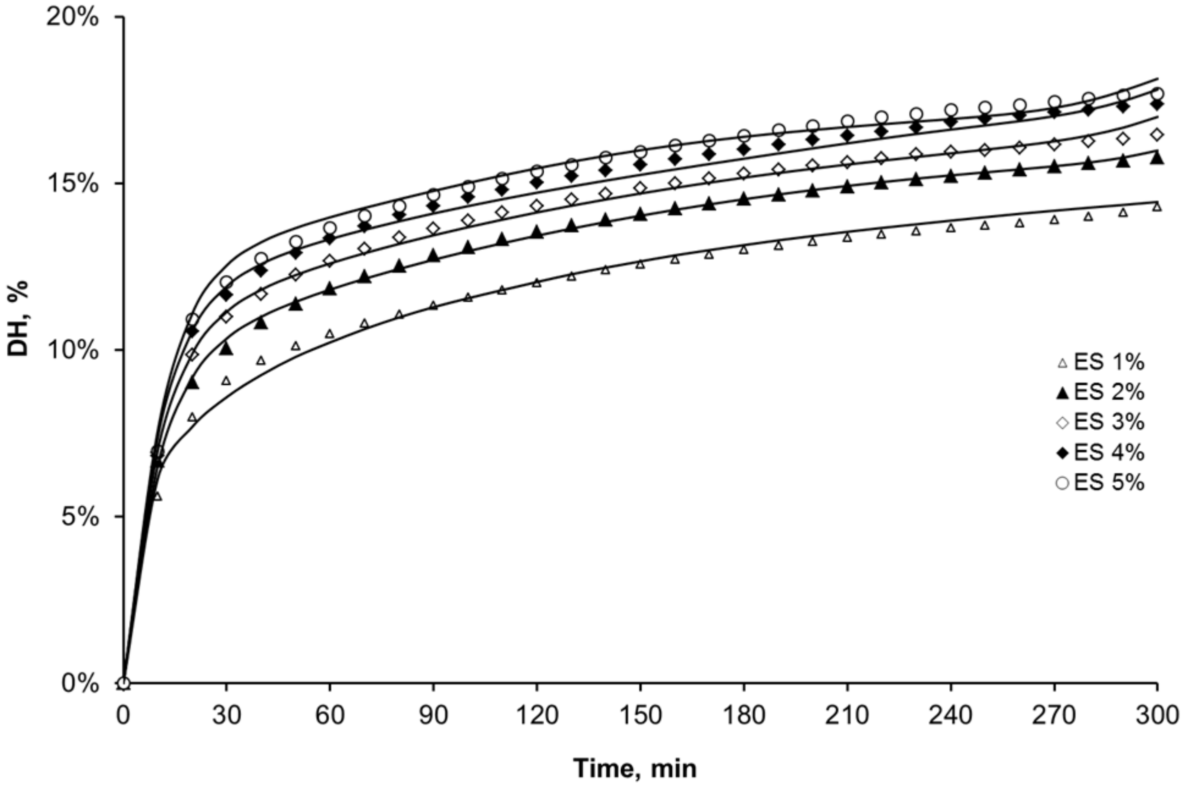
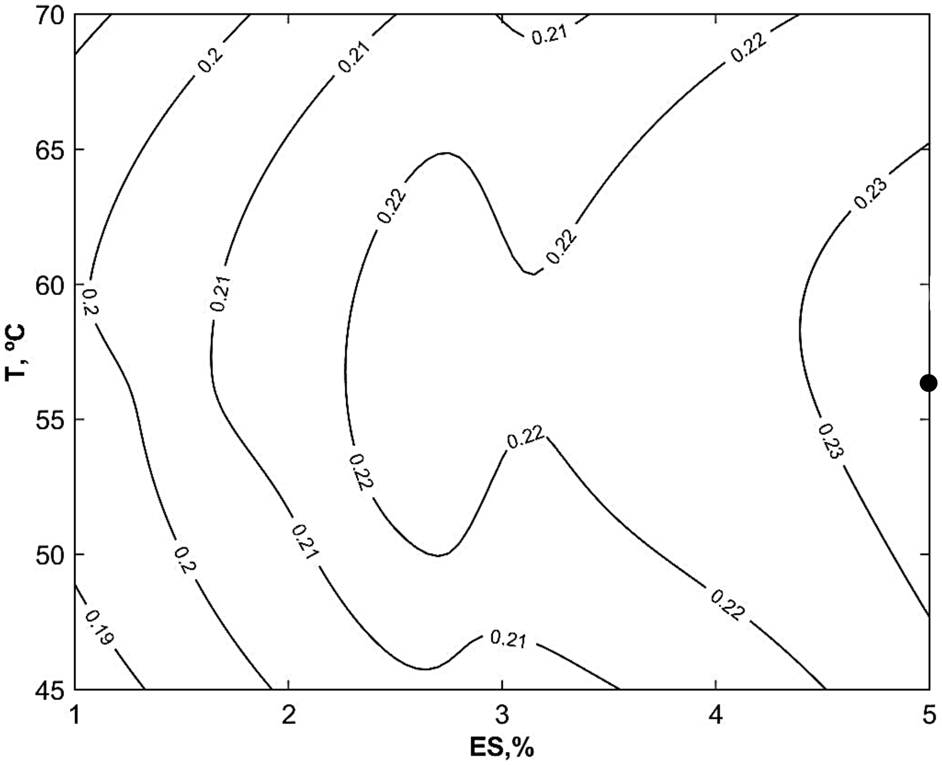



**Figure 3.** Experimental (marker points) and predicted (solid lines) values of DH against time of reaction and enzyme-substrate ratio for (a) subtilisin at 50ºC and (b) trypsin at 50ºC.

**(a)**




**(b)**

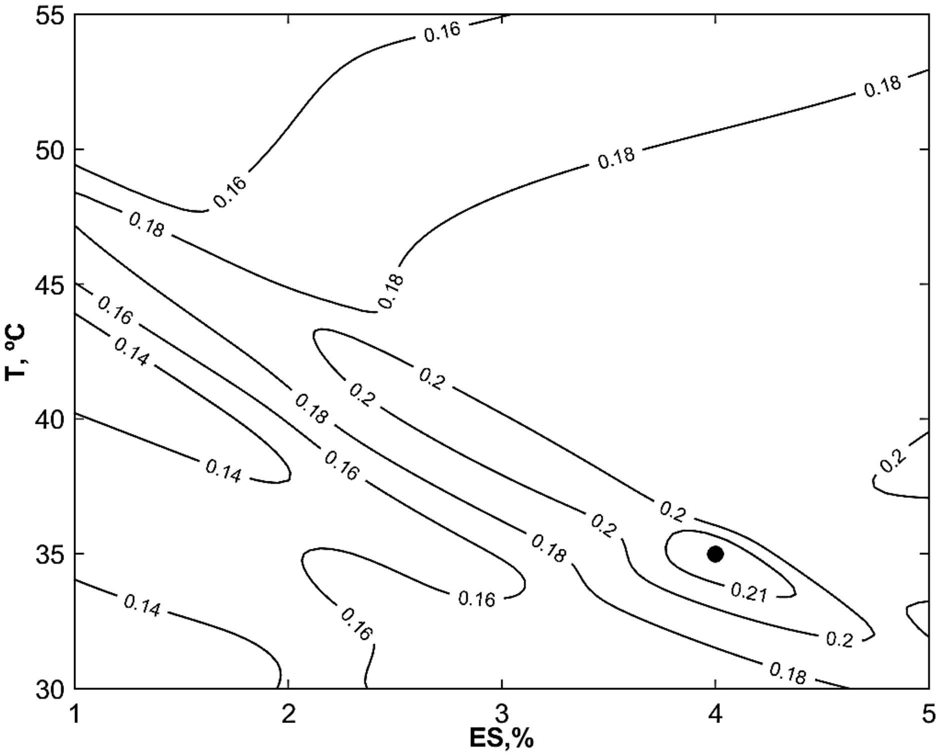



**Figure 4.** Contour plots of final DH (5 h) against enzyme-substrate ratio and reaction temperature for (a) subtilisin ANN model with 8 neurons in the hidden layer and (b) trypsin ANN model with 10 neurons in the hidden layer.

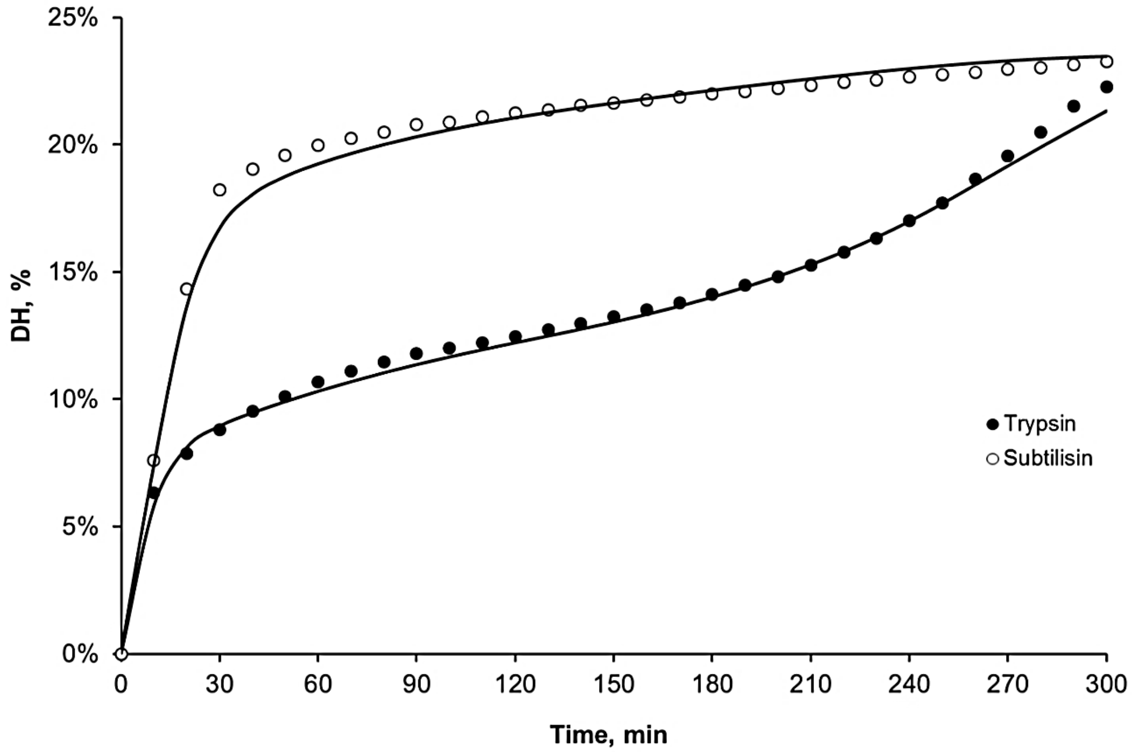



**Figure 5.** Validation experiment: experimental (marker points) and predicted (solid lines) values of DH against time of hydrolysis under optimal conditions for maximising DH (ES 5% , T=56.4ºC and ES 4%, T = 35ºC for subtilisin and trypsin, respectively).