Journal of Applied Microbiology ISSN 1364-5072

ORIGINAL ARTICLE

Modelling and optimization of fermentation factors for enhancement of alkaline protease production by isolated *Bacillus circulans* using feed-forward neural network and genetic algorithm

Ch. Subba Rao[1], T. Sathish[1], M. Mahalaxmi[1], G. Suvarna Laxmi[1], R. Sreenivas Rao[2] and R.S. Prakasham[1]

1 Bioengineering and Environmental Centre, Indian Institute of Chemical Technology, Hyderabad, Andhra Pradesh, India
2 Institute of Grassland and Environmental Research, North Wyke Research Station, Okehampton, Devon, UK

Keywords
alkaline protease, artificial neural networks, *Bacillus circulans,* fermentation, genetic algorithm, optimization.

Correspondence
Reddy Shetty Prakasham, Bioengineering and Environmental Centre, Indian Institute of Chemical Technology, Hyderabad 500 007, Andhra Pradesh, India.
E-mail: prakasam@iict.res.in

2007/0813: received 24 May 2007, revised 6 September 2007 and accepted 6 September 2007

doi:10.1111/j.1365-2672.2007.03605.x

Abstract

Aim: Modelling and optimization of fermentation factors and evaluation for enhanced alkaline protease production by *Bacillus circulans*.

Methods and Results: A hybrid system of feed-forward neural network (FFNN) and genetic algorithm (GA) was used to optimize the fermentation conditions to enhance the alkaline protease production by *B. circulans*. Different microbial metabolism regulating fermentation factors (incubation temperature, medium pH, inoculum level, medium volume, carbon and nitrogen sources) were used to construct a '6-13-1' topology of the FFNN for identifying the nonlinear relationship between fermentation factors and enzyme yield. FFNN predicted values were further optimized for alkaline protease production using GA. The overall mean absolute predictive error and the mean square errors were observed to be 0·0048, 27·9, 0·001128 and 22·45 U ml^{-1} for training and testing, respectively. The goodness of the neural network prediction (coefficient of R^2) was found to be 0·9993.

Conclusions: Four different optimum fermentation conditions revealed maximum enzyme production out of 500 simulated data. Concentration-dependent carbon and nitrogen sources, showed major impact on bacterial metabolism mediated alkaline protease production. Improved enzyme yield could be achieved by this microbial strain in wide nutrient concentration range and each selected factor concentration depends on rest of the factors concentration. The usage of FFNN–GA hybrid methodology has resulted in a significant improvement (>2·5-fold) in the alkaline protease yield.

Significance and Impact of the Study: The present study helps to optimize enzyme production and its regulation pattern by combinatorial influence of different fermentation factors. Further, the information obtained in this study signifies its importance during scale-up studies.

Introduction

Proteases, especially alkaline, constitute 60–65% of the global industrial enzyme market (Banerjee *et al.* 1999; Genckal and Tari 2006). In fact, it was reported that the global proteolytic enzyme demand will increase dramatically to 22 billon US dollars by 2009 (Turk 2006) because of their application potential in several industrial sectors especially food, meat tenderization, peptide synthesis, infant food preparations, baking and brewing, pharmaceuticals and medical diagnosis, detergent industry (additive), as well as in textiles and dehairing (Joo *et al.* 2003).

The application web of these enzymes is increasing in production of high nutritional value fish hydrolysate using *Bacillus subtilis* protease (Rebecca *et al.* 1991; Turk 2006), upgrading of lean meat waste to edible products (O'Meara and Munro 1984) and in the enzymatic modification of zein to produce a nonbitter peptide fraction with high Fischer ratio for patients with hepatic encephalopathy (Tanimoto *et al.* 1991).

Among all protease-producing microbial systems, microbes belonging to *Bacillus* genus gained importance because of extracellular enzyme production under submerged fermentation conditions (Kumar and Takagi 1999; Prakasham *et al.* 2006). Isolation of novel proteases producing species with specific characteristics will be of great value to the enzyme industry for different applications. It was well documented that each microbial strain differs in its product production character, which mainly depends on their fermentation, nutritional, physiological and genetic nature (Gupta *et al.* 2002; Prakasham *et al.* 2005; Rao *et al.* 2006). Exploration and exploitation of such characteristics would offer a competitive advantage over existing products. Hence, use of appropriate fermentation parameters was of critical importance as medium composition, product concentration, yield, and volumetric productivity influences the product productivity (Akhnazarova and Kafarov 1982; Prakasham *et al.* 1999; Sreenivas Rao *et al.* 2006). Hence, it is essential that the isolated strain should be characterized for its growth and optimal protease production. The commonly used conventional optimization method is 'one at a time' method (Prakasham *et al.* 2006), which ignores interactions among the different components even after performing of a large number of experiments. To overcome these problems, statistical methods such as response surface methodology (RSM) (Himabindu *et al.* 2006), Taguchi methodology (Rao *et al.* 2004; Prakasham *et al.* 2007a), etc were used. RSM provides polynomial models with universal approximation properties for any number of input and output variables with desired accuracy. In practice, only two or three levels are applied, given by L^N (N factor at L levels). The level of orthogonal array design or uniform design is also limited by this factor (Fang *et al.* 2003); hence, artificial neural networks (ANN) and genetic algorithms (GAs) are employed.

The ANN has a similarity like human decision-making process and used to solve the nonlinear problems in various fields, such as fermentation, optimization, and pattern recognition in biotechnology and pharmaceutical technology as well as in the bioreactor controlling and online optimization (Montague and Morris 1994; Kamimura *et al.* 1996; Subramanian *et al.* 2004; Arulsudar *et al.* 2005). Many authors compared the ANN with the statistical design and showed that the ANN results are much better than the statistical ones (Liu *et al.* 1999; Nagata and Chu 2003; Kulkarni *et al.* 2004; Subramanian *et al.* 2004; Arulsudar *et al.* 2005).

GAs are based on unorthodox search and optimization algorithms, which help in searching for a solution to the problem by mimicking some of the process of natural evolution. GA performs direct random searches through a given set of alternatives with the aim of finding the best alternative with respect to the given criteria of goodness of fit, which is expressed in terms of an objective function (also referred to fitness function). Freyer *et al.* (1992), Weuster-Botz and Wandrey (1995) and Zuzek *et al.* (1996) optimized 12–14 variables at a time using GA.

Presently, hybrid GA–ANN is becoming popular for fermentation parameter optimization. Hanai *et al.* (1999) optimized 21 variables for koji fermentation process while Hongwen *et al.* (2005) optimized 1,3 propanediol production with total 29 experiments, whereas Fang *et al.* (2003) optimized the xylitol production using GA-coupled neural networks. Nagata and Chu (2003) analysed RSM data [presented by Achary *et al.* (1997)] using neural network and GA to predict optimum conditions and reported that feed-forward neural network and GA (FFNN–GA) is the better optimization method. In the present study, authors used FFNN for modelling alkaline protease production experiments and output FFNN data variables were further optimized using the GA.

Materials and methods

Micro-organism and culture conditions

A laboratory bacterial isolate *Bacillus circulans* (MTCC 6811) was used in this study. This microbial strain was grown in 250-ml conical flasks containing 100 ml of medium consisting of yeast extract 7·5, peptone 7·5 and glucose 10 (in g l^{-1}) at pH 9·0 by incubating at 33°C and at 150 rpm in an orbital shaker (LabTech LSI – 3016 R; Daihan Labtech Co., Ltd, Namyangju-city, Kyonggi-Do, Korea) for 24 h unless otherwise stated. Cell-free broth was collected after centrifugation at 10,000 rpm at 4°C for 10 min and used as enzyme source. The organism was maintained on the above agar (20 g l^{-1}) based medium slants by subculturing at monthly intervals and stored at 4°C till further use. Six different fermentation parameters mentioned in Table 1 were selected in this study.

Measurement of enzyme activity

Alkaline protease activity was determined using modified Auson–Hagihara method (Hagihara *et al.* 1958) in the 24-h grown cell-free broth samples. In this, 1 ml of the

Table 1 Selected factors and their minimum and maximum concentration used for alkaline protease production by isolated *Bacillus circulans* under submerged fermentation

Coded variable	Variable	Lower	Upper
x1	Glucose, (gram per100 ml)	0·25	2·5
x2	Soya bean meal, (gram per 100 ml	0·25	2·5
x3	Temperature (°C)	27	39
x4	pH	8	12
x5	Volume (ml)	25	125
x6	Inoculum size (ml)	1	5

diluted clear fermentation broth was added to 1 ml casein solution (1%, w/v casein solution prepared in 50 mmol l^{-1} glycine–NaOH buffer, pH 11·0) and incubated at 70°C for 20 min. Two millilitres of 10% trichloroacetic acid were added to terminate the reaction and the contents were filtered through a Whatman No. 1 filter paper. The absorbance of the filtrate was read at 280 nm using UV–Visible spectrophotometer (Perkin-Elmer λ25; Perkin-Elmer, Shelton, CT) and the protease activity was calculated using tyrosine standard curve. One unit of alkaline protease activity was defined as 1 μg of tyrosine liberated per millilitre under the assay conditions. All experiments and each enzyme assay performed in triplicate and the average values were reported in this study.

Modelling and optimization of enzyme production

An FFNN together with the backpropagation is used as FFNN paradigm for nonlinear modelling in this study to reduce the experimental error and subsequent optimization of enzyme production using GA. Various steps involved in this study were represented schematically in flowchart (Fig. 1). FFNN consists of three layers namely input, hidden and output consisting of processing nodes (neurons). The network also contains an additional node, known as the bias, in the input and hidden layers. All nodes present in each layer are connected to subsequent layer nodes. The connection between each layer is termed as weight. The tan sigmoid activation function for hidden layer and linear transfer function for output nodes were used in the present study. The input data vectors were scaled to code values as −1 to +1 according to Zhu *et al.* (1996) and Coleman *et al.* (2003).

The FFNN is a nonlinear function-mapping device that determines the N-dimensional nonlinear function vector, f, where f: $X \rightarrow Y$, where X is a set of 'n' number of input vectors (i.e. $X = \{x_n\}$; $n = 1,2,\ldots,N$ and $x = [x_1,x_2,\ldots,x_n]^T$) and Y is a set of the corresponding output vector (protease production) (i.e. $Y = \{y_n\}$; $n = 1,2,\ldots,N$ and $y = [y_1, y_2,\ldots, y_N]$.

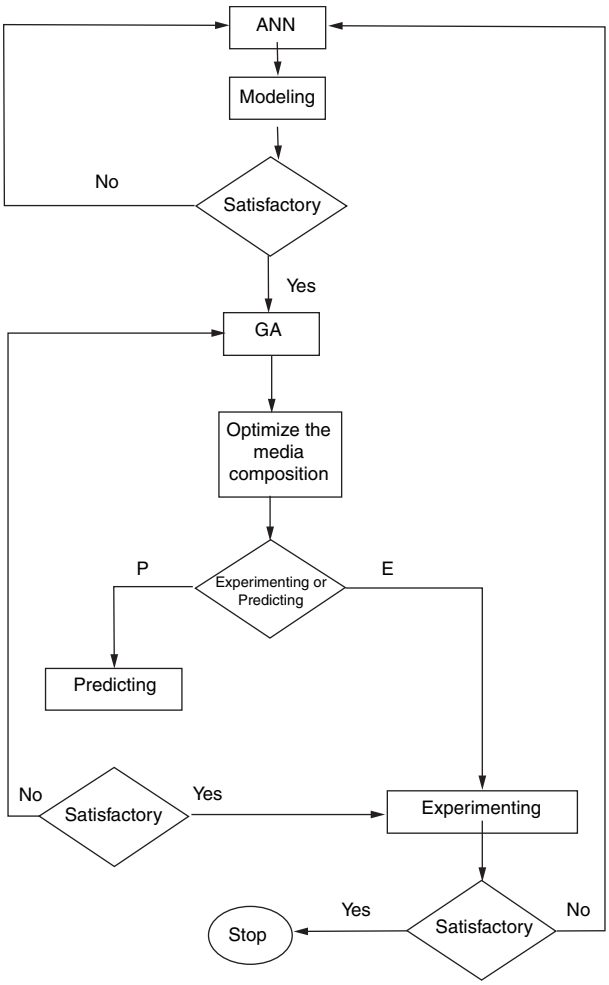

Figure 1 Schematic representation of a hybrid feed-forward neural network (FFNN)–genetic algorithm (GA).

In the present study, six microbial growth and metabolism related fermentation factors such as glucose, soya bean meal concentrations, incubation temperature, medium pH, medium volume (ml) in 250-ml flask and inoculum size (Table 1) were selected based on preliminary experimental results and used.

The precise form of f is determined by:

i Network topology

ii Choice of the activation function used for computing the outputs of the hidden and output nodes and

iii Network weight matrices ($weight^H$ and $weight^O$) and bias ($bias^I$ and $bias^H$) (where $weight^H$ is weight on connections between input and hidden nodes; $weight^O$ is weight on connections between hidden and output nodes; $bias^I$ is input bias; $bias^H$ is hidden layer bias).

The following equation is the outcome of the neural network training relating the input to the output variable, given in terms of weights and biases.

$$Y_{\text{output}} = \text{Weights}^{O} \times \left(\frac{2}{1 + e^{(-2 \times \text{weights}^{H} \times \text{input vector} + \text{input bias}(b^{1}))}} - 1 \right) + \text{Hidden layer bias}(b^{H})$$

The hidden neurons were increased step by step till the best correlation was achieved (limited to 13 in the present study, because increase in the number of hidden neurons usually results in a better learning performance, although there is a practical upper limit because too many hidden neurons may result in problems such as including the process noise, which is also known as overlearning). In total, 34 experimental runs were performed in this study, 28 runs (approx. 80%) were selected for training and remaining 6 runs (approx. 20%; the data shown in bold letters in Table 2) were used for testing. The goodness of fit was determined by the coefficient R^2, which describes the extent of variance in the modelled variables. The error was calculated based on the difference between experimental and predicted values. Mean absolute percentage error (MAPE) and mean square error (MSE) were computed using simulated and experimental data according to Zhang and Fang (2006). The percent contribution of each selected parameter was calculated using factor fitness scores, which were generated during training of the network and the contribution of each factor was calculated by using the below formula.

$$\text{Percentage contribution of each factor} = \frac{\text{Factor fitness score}}{\text{Total factors score}} \times 100$$

Table 2 Alkaline protease production experimental set up, protease activity (experimental and predicted) and error predicted

S.No	x_1	x_2	x_3	x_4	x_5	x_6	Activity	Simulated values	Error
1	0·5	0·5	30	9	50	2	1560	1566·3	-6·3
2	1·5	0·5	30	9	100	2	5236	5229·7	6·3
3	**0·5**	**1·5**	**30**	**9**	**100**	**4**	**5500**	**5498**	**2**
4	1·5	1·5	30	9	50	4	5586	5586	0
5	0·5	0·5	36	9	100	4	4948	4945·3	2·7
6	1·5	0·5	36	9	50	4	5750	5749·9	0·1
7	**0·5**	**1·5**	**36**	**9**	**50**	**2**	**4600**	**4592·3**	**7·7**
8	1·5	1·5	36	9	100	2	5318	5318	0
9	0·5	0·5	30	11	50	4	5500	5495·2	4·8
10	1·5	0·5	30	11	100	4	5660	5588·8	71·2
11	0·5	1·5	30	11	100	2	4527	4527	0
12	1·5	1·5	30	11	50	2	4200	4200·1	-0·1
13	**0·5**	**0·5**	**36**	**11**	**100**	**2**	**4682**	**4681·9**	**0·1**
14	1·5	0·5	36	11	50	2	5460	5460	0
15	0·5	1·5	36	11	50	4	5500	5500	0
16	1·5	1·5	36	11	100	4	2757	2757	0
17	0·25	1·5	33	10	75	3	5265	5265	0
18	**2·5**	**1·5**	**33**	**10**	**75**	**3**	**6100**	**6097·5**	**2·5**
19	1	0·25	33	10	75	3	4950	4950	0
20	1	2·5	33	10	75	3	5600	5599·9	0·1
21	1	1·5	27	10	75	3	4970	4970	0
22	1	1·5	39	10	75	3	3428	3428·3	−0·3
23	**1**	**1·5**	**33**	**8**	**75**	**3**	**5027**	4926·3	**100·7**
24	1	1·5	33	12	75	3	3900	3900	0
25	1	1·5	33	10	25	3	6500	6493·8	6·2
26	1	1·5	33	10	125	3	4797	4796·6	0·4
27	1	1·5	33	10	75	1	4423	4423·1	−0·1
28	1	1·5	33	10	75	5	6547	6547	0
29	1	1·5	33	10	75	3	6760	6827·5	−67·5
30	1	1·5	33	10	75	3	6901	6827·5	73·5
31	**1**	**1·5**	**33**	**10**	**75**	**3**	**6890**	**6827·5**	**62·5**
32	1	1·5	33	10	75	3	6802	6827·5	−25·5
33	1	1·5	33	10	75	3	6790	6827·5	−37·5
34	1	1·5	33	10	75	3	6896	6827·5	68·5

GA optimization

GAs are employed to search and optimize the maximum of a function over some domain space. Points in the domain space of the search (usually real numbers) are encoded as bit strings, called chromosomes. Each bit position in the string is called a gene and each starting solution (here selected fermentation parameter) is known as population. Every population is evaluated for their *fitness* while solving the problem. From the initial population of chromosomes, a new population is generated using three genetic operations: reproduction (generation of new population), crossover (exchange of bits in a pair of chromosome) and mutation (flipping of bits in offspring).

GA optimization was performed using FFNN output values (weights and bias) using same fitness function. Each GA output simulation was used to search in different subspace and to locate the global maximum on the objective function surface. The underlying optimization objective is to find the L-dimensional optimal decision variables by fixing the lower and upper bounds (Table 1) ($x_l^L < x_l < x_l^U$ l = 1, 2, 3,…, L). x denotes the fermentation operating conditions ($L = 6$) and x_l^L and x_l^U represent the lower and upper bounds on x_l. In this study, different parameters of GA optimization such as chromosome length (l_{chr}) as 60, population size (N_{pop}) as 60, crossover probability as 0·9 and mutation probability (P_{mut}) as 0·01 were assumed based on literature reports (Fang *et al.* 2003). Optimum conditions were selected after evaluation of GA for 500 generations ($N_g^{max} = 500$) to achieve fine-tuned fermentation conditions in the given range of input parameters. Neural networks and GA toolboxes of MATLAB 7.0 (The Mathworks, Inc., Natick, MA, USA) were used in modelling studies.

Results

Alkaline protease production by any microbial strain depends on various fermentation, environmental and growth conditions. Our preliminary studies indicated that the enzyme production by isolated *B. circulans* is greatly regulated by carbon (glucose 1%) and nitrogen (soya bean meal 1%) sources in addition to incubation temperature (33°C), pH of the medium (pH 10·0), initial inoculum concentration (3·0% of 24 h grown 0·8 absorbance culture at 600 nm) and aeration (75 ml medium in 250-ml conical flask) (results not shown). Maximum enzyme production of 3245 U ml^{-1} was observed under the above conditions. Hence, further experiments were carried to evaluate the impact of each above parameter and its concentration on alkaline protease production. Each selected parameter's range at which enzyme production was noticed was considered to determine the lower and upper limit (Table 1) and further each parameter was divided into different levels by taking the consensus of research team. Experimental lay out was prepared using modified fractional factorial central composite design (MATLAB 7.0) and the data were analysed using FFNN (Table 2). Desai *et al.* (2006) reported similar experimental study for production of exopolysaccharide (EPS) using *Lactobacillus plantarum*.

The protease production values were varied depending on the experimental conditions (Table 2). The FFNN program was initially tested using Levenberg–Marquardt backpropagation (Rumelhart *et al.* 1986), Bayesian regularization backpropagation (MacKay 1992; Foresee and Hagan 1997), Conjugate gradient backpropagation with Powell–Beale restarts (Powell 1977) and scaled conjugate gradient backpropagation (Moller 1993). Among all these, Levenberg–Marquardt backpropagation showed a better correlation between experimental (1560–6896 Units) and simulated values (1566–6826 Units). Initially, weight and bias values were taken randomly during network training and were further optimized to minimize the error. The final optimized weight and bias values were presented in Table 3. The average error (average difference between software predicted and experimental value) was observed to be approx. 0·95% (Table 2).

At optimum correlated conditions, 13 neurons were observed in the hidden layer suggesting '6-13-1' FFNN topology for this experiment (Fig. 2). Figure 3 depicted the quality of the network for all data points of training and testing. The simulated data showed the excellent correlation with the observed data. The overall MAPE and MSEs were observed to be 0·0048, 27·9, 0·001128 and 22·45 enzyme activity (U ml^{-1}) for training and testing, respectively. The smallest value of MSE suggested that the FFNN possesses good approximation and generalization characteristics for production of alkaline protease by this bacterial strain. The goodness of the neural network prediction was analysed by calculating the coefficient of R^2, which was found to be 0·9993. Figure 3 also shows that the predictions were concentrated near the diagonal line on the graph and almost no scattering points were found.

The FFNN output enzyme production data were further optimized to get the best fermentation parameters using GA. Because, in general, all algorithms give a local optimized solution for the nonlinear problems, whereas the GA gives a global solution. Table 4 depicts the best

Table 3 The weight and bias values of nonlinear function at optimum conditions

Weight on connection between input and hidden nodes											
x1	x2	x3	x4	x5	x6		Bias I	Bias H			
1·8613	0·0712	−0·1622	−0·0509	−0·1703	1·6471		−4·6664	466·4825			
2·7218	1·2877	6·0177	5·6588	0·7556	1·7617		−3·8032				
−9·4915	−15·1343	12·9777	6·82	2·1501	−7·6765		−9·1134				
−1·8613	−0·0712	0·1622	0·0509	0·1703	−1·6471		4·6664				
3·3253	6·4304	5·8927	5·5338	0·7056	1·5218		−4·0408				
4·058	6·4796	5·9829	0·1389	5·2076	1·4019		−3·8612				
−0·833	−1·3201	7·2998	6·7578	6·1707	10·73		11·3593				
−2·7218	−1·2877	−6·0177	−5·6588	−0·7556	−1·7617		3·8032				
−7·9977	−5·8886	-0·7707	−12·241	7·1586	1·7331		−6·1434				
−1·8613	−0·0712	0·1622	0·051	0·1704	−1·6471		4·6664				
−1·6485	−2·6794	−8·9557	11·2918	7·8368	3·0848		−3·3074				
1·8613	0·0712	-0·1622	−0·051	−0·1704	1·6471		−4·6664				
−3·3034	−0·4969	−4·7561	−0·279	−3·4922	−16·0113		−13·7485				
−1·8613	−0·0712	0·1622	0·051	0·1704	−1·6471		4·6664				
Weight on connection between hidden and output nodes											
−527·72	466·4774	527·7182	−466·477	527·6857	466·5878	527·6328	466·4774	−466·581	527·603	−466·567	−466·584	−466·735

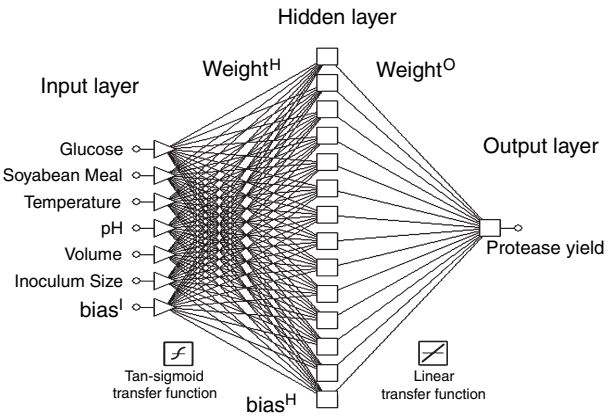

Figure 2 Feed-forward neural network architecture used for alkaline protease production fermentation factors optimization.

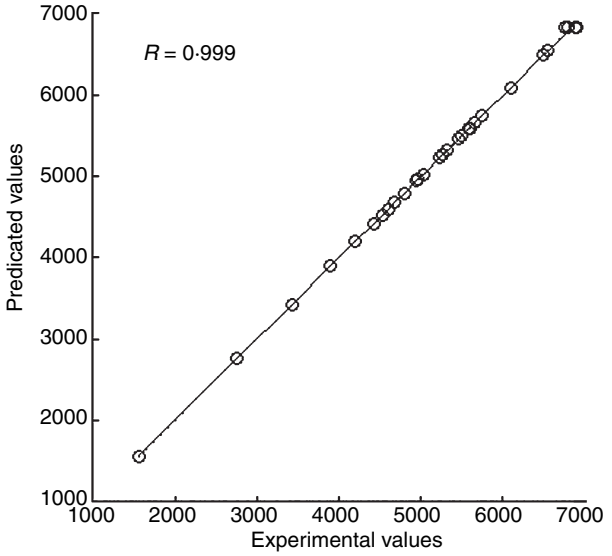

Figure 3 Correlation chart for experimental and FFNN predicted alkaline protease production data. (○) Data points; (——) Best line fit; (……) E = P.

possible optimal conditions obtained after performing various GA trails and only the best four conditions among all were selected for further verification. From the given experimental data, it could be seen that the maximum protease yield was 6896 U ml^{-1} before and increased to 8320 U ml^{-1} after the FFNN–GA hybrid optimization indicating an increase of approximately 20%. These results were in accordance with Dutta *et al.* (2004), where an increase of protease yield from 56 to 58·5 U ml^{-1} was reported by optimizing the pH, temperature and inoculum size with RSM and ANN models while Desai *et al.* (2006) showed increased EPS yield with the help of PB–ANN–GA methods.

Based on the above experimental results, surface contours were generated to understand the impact of one factor on the other using fitness function with the help of MATLAB 7.0. Maximum production surfaces observed are symmetric and flat near the optimum environment (Fig. 4). The protease production varied with variation of glucose and soya bean meal concentration ratio in the fermentation medium (Fig. 4a). For example, maximum enzyme production can be possible in any given carbon source range of 0·3–2·5% (w/v) by altering the nitrogen source (soya bean meal) concentration in the range of 0·25–2·5% (w/v) in the medium (Fig. 4a) indicating the carbon and nitrogen source ratio regulate the protease production in this bacterial strain. Analysis of glucose interaction with incubation temperature (Fig. 4b) indicated that optimal protease production could be possible in the incubation temperature range of 28°–38°C with the supplementation of glucose in the range of 0·5–1·8% (w/v) in fermentation medium. Further increase in glucose concentration caused reduced enzyme production (Fig. 4b). Similar trend was noticed with volume of the medium (Fig. 4f). Selected glucose concentration range also showed positive effect on alkaline protease production at lower level of medium pH (7–10 pH range; Fig. 4c). Comparative evaluation between incubation temperature and inoculum level suggested that 2–4% (ml) inoculum was effective for maximum enzyme production in the temperature range of 28°–37°C (Fig. 4h). Similar enzyme production trend was observed with incubation temperature *vs* medium volume (Fig. 4e). It was interesting to note that 33°C was effective for optimum enzyme production with 2–4% inoculum range and 50–100 ml of volume in 250-ml conical flask (Fig. 4e,h).

Figure 5 represents the each fermentation factor contribution on the overall alkaline protease production by this *B. circulans*. Maximum impact was observed with glucose

Table 4 The best possible selected fermentation conditions predicted and verified enzyme yields

S.No	Glucose	Soya bean meal	Temperature	pH	Volume	Inoculum size	GA-optimized yield	Experimental yield
1	1·12	1·21	34·5	9·5	85	4·5	8210	8228
2	1·2	1·21	34·58	9·65	86	5·1	8283	8320
3	1·22	1·12	34·6	9·66	86	5·3	8308	8240
4	1·23	1·08	34·7	9·66	87	5·8	8236	8256

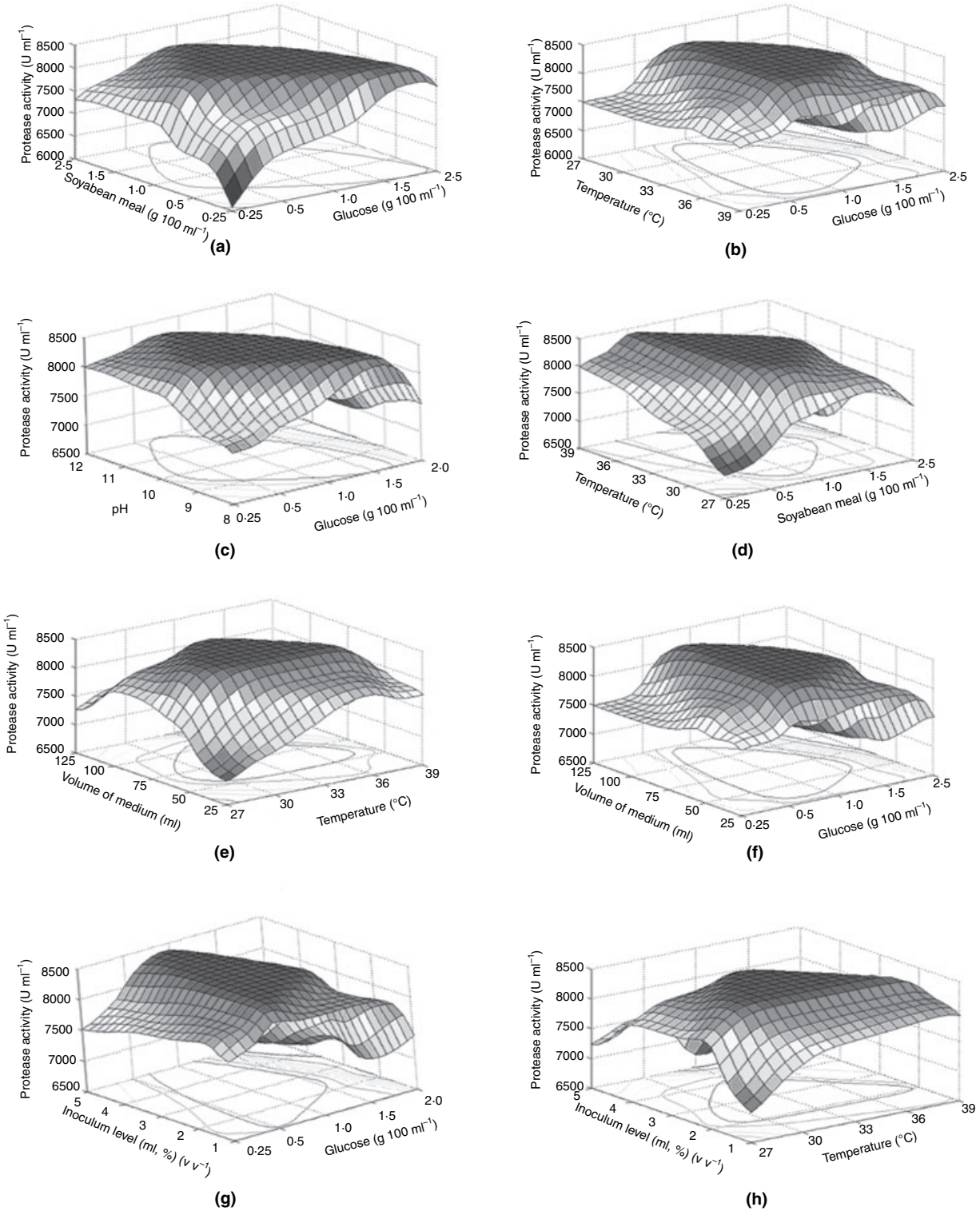

Figure 4 Effect of selected fermentation factors interactions on alkaline protease production by isolated *Bacillus circulans.* (a) Glucose *vs* Soya bean meal (b) Glucose *vs* temperature (c) Glucose *vs* pH (d) Soya bean meal *vs* temperature (e) Temperature *vs* volume of the medium (f) Glucose *vs* volume of the medium (g) Glucose *vs* inoculum level (h) Temperature *vs* inoculum level. [Note: In all the above figures, only two factors mentioned on *x* and *y* axes were varied and other four factors values were as follows: Glucose: 12·0 (g l^{-1}, w/v), Soya bean meal: 12·1 (g l^{-1}, w/v), Temperature: 34·6°C, pH: 9·65, Volume: 86 ml per 250-ml flask (v/v) and Inoculum level: 51 (ml l^{-1}, v/v)].

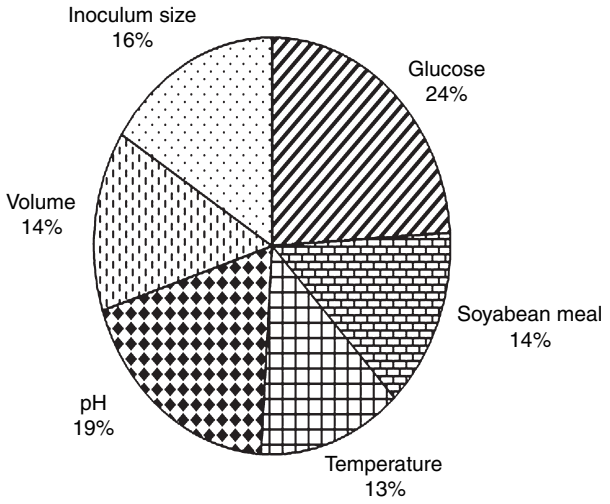

Figure 5 Contribution of each fermentation factor on alkaline protease production by *Bacillus circulans*.

(24%) followed by pH of the medium (19%) and the least was noticed with incubation temperature (13%) under FFNN–GA optimized environment.

Discussion

In this study, six different bacterial growth-associated factors were studied in terms of their individual influence on alkaline protease production. Table 2 clearly indicates the influence of above selected fermentation factors on alkaline protease production by this microbial strain, where minimum and maximum enzyme production was noticed to be 1560 and 6900 U ml^{-1} depending on the fermentation factors' concentrations/values, respectively. Validation of FFNN–GA results revealed an increase in enzyme production from 3245 to 8320 denoting >2·5-fold increase in enzyme yield with the medium (pH 9·65) consisting of glucose 12·0 and soya bean meal 12·1 (g l^{-1}) at 34·58°C using 86 ml of fermentation medium in 250-ml conical flask with inoculum level of 5·1 ml (Table 4). Such optimization mediated enzyme yield improvement was also reported with bacterial and fungal strains (Prakasham *et al.* 2007a,b). These results suggest that fermentation parameters greatly influence the extracellular production of proteases in this microbial strain and their interaction plays an important role in the synthesis of this enzyme because of their regulatory role in the induction or repression of the enzyme production. A similar catabolic control mechanism for extracellular enzyme production has been described for *Pseudomonas maltophilia* (Boethling 1975), *Yersinia ruckeri* (Secades and Guijarro 1999) and *Pseudoalteromonas* sp. SM9913 (Li *et al.* 2004).

The enzyme production was regulated by carbon and nitrogen sources ratio. An effective protease production could be achieved by varying one of these two sources by keeping the other nutrient concentration constant (Fig. 4). One percent glucose supplementation was effective for optimum protease production in the studied soya bean meal concentration range of 0·25–2·5% (g, w/v), at a temperature range of 28–39°C, at varied medium pH of 8–12, at different inoculum levels (1·0–5·0% (ml, v/v) and in different volumes (25–125 ml in 250-ml conical flask; Fig. 4) indicating among all the factors, glucose had the maximum influence followed by pH of the medium (Fig. 5) on bacterial metabolism in alkaline protease production. Though the influence of increased inoculum on enhanced enzyme production was well reported in the literature (Kumar and Takagi 1999; Prakasham *et al.* 2006); however, soya bean meal concentration regulated productivity of protease as observed in this study was not documented.

Interactive influence of selected fermentation factors was analysed using three-dimensional surface plots simulated by network output. Each contour curve represents an infinite number of combinations of two test variables with the other maintained at their respective optimized levels. These surface contours revealed symmetric and flat surface near the optimum environment (Fig. 4) unlike other RSM contours reported in the literature (Himabindu *et al.* 2006) suggesting that the optimum protease production by this microbial strain could be achieved in wide concentration ranges. The above data further supported that interaction among fermentation parameters was one of the important aspects in achieving optimum productivity of any metabolite and further improvement in enzyme production is possible with regulation of interactive influences between selected fermentation parameters. This data helps in simulation of economic fermentation medium especially, with respect to selection of carbon and nitrogen source concentrations in scale-up studies to produce maximum alkaline protease production using this *B. circulans*.

Overall, alkaline protease production by *B. circulans* was optimized using hybrid of FFNN–GA, selecting the physical parameters and medium composition. The FFNN model was constructed on the basis of data from 34 fermentation experiments. This FFNN model showed excellent prediction accuracy and generalization ability. The alkaline protease yield obtained in the validation experiments was 8320 Units, which were in close agreement with the GA, optimized yield of 8283 Units. It can, thus, be seen that the usage of FFNN–GA hybrid methodology has resulted in a significant improvement in the alkaline protease yield (>2·5-fold). The approach presented in this paper is sufficiently general and thus can also be

employed for modelling and optimization of other bioprocesses.

Acknowledgement

Authors thank Department of Science and Technology, New Delhi, Council of Scientific and Industrial Research, New Delhi for financial support.

References

Achary, A., Hariharan, K.A., Bandhyopadhyaya, S., Ramachandran, R. and Jayaraman, K. (1997) Application of numerical modeling for the development of optimized complex medium for D-hydantoinase production from *Agrobacterium radiobacter* NRRL B11291. *Biotechnol Bioeng* **55**, 148–154.

Akhnazarova, S. and Kafarov, V. (1982) *Experiment optimization in chemistry and chemical engineering*. Moscow: Mir Publications.

Arulsudar, N., Subramanian, N. and Murthy, R.S.R. (2005) Composition of artificial neural network and multiple linear regression in the optimization of formulation parameters of leuprolide acetate loaded liposomes. *J Pharm Pharm Sci* **8**, 243–258.

Banerjee, C.U., Sani, R.K., Azmi, W. and Soni, R. (1999) Thermostable alkaline protease from *Bacillus brevis* and its characterization as a laundry detergent additive. *Process Biochem* **35**, 213–219.

Boethling, R.S. (1975) Regulation of extracellular protease secretion in *Pseudomonas maltophilia*. *J Biotechnol* **123**, 954–961.

Coleman, M.C., Buck, K.K.S. and Block, D.E. (2003) An integrated approach to optimization of *Escherichia coli* fermentations using historical data. *Biotechnol Bioeng* **84**, 274–285.

Desai, K.M., Akolkar, S.K., Badhe, Y.P., Tambe, S.S. and Lele, S.S. (2006) Optimization of fermentation media for exopolysaccharide production from *Lactobacillus plantarum* using artificial intelligence based technique. *Process Biochem* **41**, 1842–1848.

Dutta, J.R., Dutta, P.K. and Banerjee, R. (2004) Optimization of culture parameters for extracellular protease production from a newly isolated *Pseudomonas* sp. using response surface and artificial neural network models. *Process Biochem* **39**, 2193–2198.

Fang, B., Hongwen, C., Xiolan, X., Ning, W. and Zongding, H. (2003) Using genetic algorithms coupling neural networks in a study of xylitol production: medium optimization. *Process Biochem* **38**, 979–985.

Foresee, F.D. and Hagan, M.T. (1997) Gauss–Newton approximation to Bayesian regularization. In *Proceedings of the 1997 International Joint Conference on Neural Networks*, Houston, Texas, USA, June 9–12, pp. 1930–1935.

Freyer, S., Weuster-Botz, D. and Wandery, C. (1992) Medium optimization using genetic algorithms. *Bioeng* **8**, 16–25.

Genckal, H. and Tari, C. (2006) Alkaline protease production from alkalophilic *Bacillus* sp. isolated from natural habitats. *Enzyme Microb Technol* **39**, 703–710.

Gupta, R., Beg, Q.K. and Lorenz, P. (2002) Bacterial alkaline proteases: molecular approaches and industrial applications. *Appl Microbiol Biotechnol* **59**, 15–32.

Hagihara, B., Matsubara, H., Nakai, M. and Okunuki, K. (1958) Crystalline bacterial proteinase. I. Preparation of crystalline proteinase of Bacillus subtilis. *J Biochem* **45**, 185–194.

Hanai, T., Honda, H., Ohkusu, E., Ohki, T., Tohayama, H., Muramastu, T. and Kobayashi, T. (1999) Application of an artificial neural network and genetic algorithm for determination of process orbits in the koji making process. *J Biosci Bioeng* **87**, 507–512.

Himabindu, M., Ravichandra, P., Vishalakshi, K. and Annapurna, J. (2006) Optimization of critical medium components for the maximal production of gentamicin by *Micromonospora echinospora* ATCC 15838 using response surface methodology. *Appl Biochem Biotechnol* **134**, 143–154.

Hongwen, C., Baishen, F. and Zongding, H. (2005) Optimization of process parameters for key enzymes accumulation of 1,3,propanediol production from *Klebsilla pneumoniae*. *Biochem Eng J* **25**, 47–53.

Joo, H.S., Kumar, C.G., Park, G.C., Paik, S.R. and Chang, C.S. (2003) Oxidant and SDS stable alkaline protease from *Bacillus clausii* I-52: production and some properties. *J Appl Microbiol* **95**, 267–272.

Kamimura, R., Kostantinov, K. and Stephanopolus, G. (1996) Knowledge-based systems, artificial neural networks and pattern recognition: application in biotechnological processes. *Curr Opin Biotechnol* **7**, 231–234.

Kulkarni, S.G., Choudary, A.K., Nandi, S., Tambe, S. and Kulkarni, B.D. (2004) Modeling and monitoring of batch process using principal component analysis (PCA) assisted generalized regression neural networks (GRNN). *Biochem Eng J* **18**, 193–210.

Kumar, C.G. and Takagi, H. (1999) Microbial alkaline proteases: from a bioindustrial viewpoints. *Biotechnol Adv* **17**, 561–594.

Li, W.J., Pan, J., He, H.L., Chen, X.L., Zhang, Y.Z. and Gao, P.J. (2004) Optimization of cold adapted protease fermentation by benthal psychrotrophic *Pseudoaltermonas* sp. SM9913. *J Shangdong Univ* **39**, 104–108.

Liu, C.H., Hwang, C.F. and Liao, C.C. (1999) Medium optimization for glutathione production by *Saccharomyces cerevisiae*. *Process Biochem* **34**, 17–23.

MacKay, D.J.C. (1992) Bayesian interpolation. *Neural Comput* **4**, 415–447.

Moller, M.F. (1993) A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw* **6**, 525–533.

Montague, G. and Morris, J. (1994) Neural-network contributions in biotechnology. *Trends Biotechnol* **12**, 312–324.

Nagata, Y. and Chu, K.H. (2003) Optimization of a fermentation medium using neural networks and genetic algorithms. *Biotechnol Lett* **25**, 1837–1842.

O'Meara, G.M. and Munro, P.A. (1984) Selection of a proteolytic enzyme to solubilize lean beef tissue. *Enzyme Microb Technol* **6**, 181–185.

Powell, M.J.D. (1977) Restart procedures for the conjugate gradient method. *Math Prog* **12**, 241–254.

Prakasham, R.S., Kuriakose, B. and Ramakrishna, S.V. (1999) The influence of inert solids on ethanol production by *Saccharomyces cerevisiae*. *Appl Biochem Biotechnol* **82**, 127–134.

Prakasham, R.S., Subba Rao, Ch., Sreenivas Rao, R. and Sarma, P.N. (2005) Alkaline protease production by an isolated *Bacillus circulans* under solid-state fermentation using agroindustrial waste: process parameters optimization. *Biotechnol Prog* **21**, 1380–1388.

Prakasham, R.S., Subba Rao, Ch. and Sarma, P.N. (2006) Green gram husk: an inexpensive substrate for alkaline protease production by *Bacillus* sp. in solid-state fermentation. *Bioresour Technol* **97**, 1449–1454.

Prakasham, R.S., Rao, Ch.S., Rao, R.S., Lakshmi, G.S. and Sarma, P.N. (2007a) L-asparaginase production by isolated *Staphylococcus* sp. - 6A: design of experiment considering interaction effect for process parameter optimization. *J Appl Microbiol* **102**, 1382–1391.

Prakasham, R.S., Subba Rao, Ch., Sreenivas Rao, R. and Sarma, P.N. (2007b) Enhancement of acid amylase production by an isolated *Aspergillus awamori*. *J Appl Microbiol* **102**, 204–211.

Rao, R.S., Prakasham, R.S., Prasad, K.K., Rajesham, S., Sarma, P.N. and Rao, L.V. (2004) Xylitol production by Candida sp.: parameter optimization using Taguchi approach. *Process Biochem* **39**, 951–956.

Rao, R.S., Jyothi, C.P., Prakasham, R.S., Rao, Ch.S., Sarma, P.N. and Rao, L.V. (2006) Strain improvement of *Candida tropicalis* for the production of xylitol: biochemical and physiological characteristics of wild-type and mutant strain CT-OMV5. *J. Microbiol* **44**, 113–120.

Rebecca, B.D., Pena-Vera, M.T. and Diaz-Castaneda, M. (1991) Production of fish protein hydrolysates with bacterial proteases, yield and nutritional value. *J Food Sci* **56**, 309–314.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) Learning representations by back-propagating errors. *Nature* **323**, 533–536.

Secades, P. and Guijarro, J.A. (1999) Purification and characterization of an extracellular protease from the fish pathogen *Yersinia ruckeri* and effect of culture conditions on production. *Appl Environ Microbiol* **65**, 3969–3975.

Sreenivas Rao, R., Pavana Jyothi, Ch., Prakasham, R.S., Sarma, P.N. and Venkateswar Rao, L. (2006) Xylitol production from corn fiber and sugarcane bagasse hydrolysates by *Candida tropicalis*. *Bioresour Technol* **97**, 1974–1978.

Subramanian, N., Yajnik, A. and Murthy, R.S.R. (2004) Artificial neural network as a alternative to multiple regression analysis in optimizing formulation parameters of cytrabine liposomes. *AAPS PharmSciTech* **5**, 1–9.

Tanimoto, S.-Y., Tanabe, S., Watanabe, M. and Arai, S. (1991) Enzymatic modification of zein to produce a non-bitter peptide fraction with a very Fisher ratio for patients with hepatic encephalopathy. *Agric Biol Chem* **55**, 1119–1123.

Turk, B. (2006) Targeting proteases: successes, failures and future prospects. *Nat Rev Drug Discov* **5**, 785–798.

Weuster-Botz, D. and Wandrey, C. (1995) Medium optimization by genetic algorithm for continuous production of formate dehydrogenase. *Process Biochem* **30**, 563–571.

Zhang, G. and Fang, B. (2006) A uniform design-based back propagation neural network model for amino acid composition and optimal pH in G/11 xylanase. *J Chem Technol Biotechnol* **81**, 1185–1189.

Zhu, Y.H., Rajalahti, T. and Linko, S. (1996) Application of neural networks to lysine production. *Chem Eng J* **62**, 207–214.

Zuzek, M., Friedrich, J., Cestnik, B., Karalic, A. and Cimerman, A. (1996) Optimisation of fermentation medium by a modified method of genetic algorithms. *Biotechnol Tech* **10**, 991–996.