

Hybrid Neural Network based Optimization of Process parameters and application of Taguchi method for Lipase production from Co-culture of *Lactobacillus brevis* and *Lactobacillus plantarum*

Sita Ramyasree Uppada¹, Aditya Balu², Amit Kumar Gupta²,
Jayati Ray Dutta^{1*}

¹ Biological Sciences Department, Birla Institute of Technology and Science, Pilani-Hyderabad campus Hyderabad, Andhra Pradesh, 500078, India.

² Department of Mechanical Engineering, Birla Institute of Technology and Science, Pilani Hyderabad campus, Hyderabad, Andhra Pradesh, 500078, India.

*Corres.author: jayati@hyderabad.bits-pilani.ac.in, jraydutta2002@yahoo.co.in

Abstract : Different hybrid models for optimization such as Artificial Neural Network enhanced Genetic Algorithm (ANN-GA) and Bayesian Regularized Artificial Neural Network enhanced Genetic Algorithm (BR-ANN-GA) were explored in order to produce lipase economically from co-culture of *Lactobacillus brevis* and *Lactobacillus plantarum*. Taguchi method has been applied to enhance the effectiveness of optimization process and it is seen that the significance of pH and inoculum volume are very less. Hence keeping these factors constant, maximum lipase activity of 74.05 U/ml was obtained.

Key words: Extracellular lipase; *Lactobacillus brevis*; *Lactobacillus plantarum*; Artificial neural network; Genetic algorithm; Taguchi method.

Introduction

Lipases are versatile enzymes capable of undergoing esterification, trans-esterification, hydrolysis, acidolysis and aminolysis reactions. This property made them to use in food, flavoring, dairy, detergent, pharmaceutical, degumming and leather industries¹. Different esters, amides and industrial intermediates are obtained by using lipase as a catalyst. Mainly in field of clinical sciences, lipases play a vital role in production of esters like Fatty acid ethyl esters (FAEEs) which are used widely as a biomarker to identify different diseases². Though lipases are available from animals like porcine, pigs, goat, sheep, calves, plant seedlings like Brassica, mustard, castor bean, papaya and oat seed, microbial enzymes derived from bacteria, yeast and fungi are widely used due to rapid microbial growth, ability of microorganism to with stand high pH and temperature, easy cultivation with cheap substrates and more product formation. Improving the conditions of fermentation will result in obtaining more desired metabolic product or enzyme levels in the microbial system. Most of the industrial enzymes are produced by single microbial strains. Co-culture technique plays important role in production of different enzymes³, biogas, probiotics etc. The production is more when compared to single cultures. It is well known that extracellular lipase production in microorganism is greatly influenced by physical factors like pH, temperature, incubation period, substrate volume and inoculum volume. Classical single variable optimization method for optimization of lipase production is time consuming⁴ hence recently number of statistical experimental designs has been employed for optimizing enzyme production from microorganisms⁵. Since, lipase production is dependent on multiple parameters; the modeling of lipase yield using the statistical techniques such as Response Surface Methodology (RSM) is not very accurate. Hence soft computing techniques such as Artificial Neural Network (ANN) and Bayesian Regulation (BR-ANN) are used

here. Considering the many industrial applications of lipase, this paper proposes a novel methodology of ANN-GA and BR-ANN-GA for optimization of extracellular lipase production from co-culture of *Lactobacillus brevis* and *Lactobacillus plantarum*. Further the significance of each parameter is studied using the Taguchi design to improve the effectiveness of the optimization process.

Materials and Experimentation

Microorganism

For the present investigation, the co-culture of *Lactobacillus brevis* and *Lactobacillus plantarum* were used for producing extracellular lipase. The strains were procured from MTCC and maintained on LB agar slants at 4⁰C.

Composition of medium

LB agar was used to maintain mother culture with a composition of Tryptone (10g), yeast extract (5g), NaCl(10g), agar (15 g) and Distilled water(1lt).

Lipolytic Activity

For preliminary screening of lipase producing bacteria, tributyrin agar was used. All the isolated cultures were inoculated into tributyrin agar plates and kept for incubation at 37°C for 24 hours and observed for zone formation. A clear zone around the colonies indicated the production of lipase.

Enzyme assay

The lipase assay was performed spectrophotometrically using p-nitro phenyl palmitate as substrate. P-nitro phenol was liberated from p-nitro phenyl palmitate by lipase mediated hydrolysis⁶. One unit (U) of lipase was defined as the amount of enzyme that liberates one micromole of p-nitro phenol per minute under the assay conditions⁷.

Experimental design and Lipase production

Lipase assay medium was used for growth and production of lipase from two *Lactobacillus* species. The optimum levels for extracellular lipase production by the *L. brevis* and *L. plantarum* strains with respect to incubation period, temperature, pH and inoculum volume, substrate volume, were obtained by single factor optimization by conducting the experiments in 250 ml Erlenmeyer flasks containing 50 ml of medium comprising of peptone (0.5%), yeast extract (0.3%), NaCl (0.25%), MgSO₄ (0.05%) with olive oil as a substrate were inoculated with the freshly prepared bacterial suspension at 35⁰C. After incubation, the cell-free supernatant was obtained by centrifugation at 7860 rpm for 20 minutes and the extracellular lipase activity of the fermented broth was determined.

In the next stage, ANN along with GA and Taguchi were used to study the interactive effects of the five variables, i.e. pH, temperature, inoculum volume, incubation period, substrate volume for improving total lipase production. Experiments were conducted in triplicate and the results were the average of these three independent trials.

Artificial Neural Network Model

One of the useful tools of the field of the Machine Learning is the Artificial Neural networks (ANN). Predictive modeling of process parameter, such as production of lipase is highly complex. In literature it has been found that feed forward back propagation neural networks are being used to train the neural network which is very reliable. However, the normal back-propagation algorithm suffers the problem for not able to give better generalization. The generalization is given better by using Bayesian regularization (BR-ANN)⁸. In this article, a comparison is made between models trained with back-propagation algorithm and also the back-propagation algorithm enhanced with Bayesian regularization. The parameters of the neural network architecture are generally set by trial and error and reserved test data is used to assess its generalization ability (or more sophisticated cross-validation). These parameters change the effective learning model. However, the parameters are set in this article by the algorithm mentioned below in the next section.

The neural network toolbox of MATLAB software package is used for training and testing the given data with both normal back propagation algorithm and Bayesian regularization function. The selection of Neural Network architecture is one of the great challenges in the modeling of Neural Network. The algorithm for finding the best architecture for both the plane network (ANN) as well as Bayesian Regularized Neural Network(BR-ANN) is been explained in the pseudo code given in the further sections.

Optimization of Yield Using Genetic Algorithm

Genetic algorithm (GA) is a stochastic optimization technique that finds application in bioinformatics, phylogenetics, computational science, engineering, economics, chemistry, manufacturing, mathematics, physics, pharmacometrics and other fields⁹. The major advantage of genetic algorithms over other conventional optimization techniques is the flexibility it provides in giving the objective function and constraints¹⁰.

In GA, a population of candidate solutions (called chromosome) to an optimization problem is evolved towards fitter solutions in an iterative process. Each candidate solution is mutated and altered; traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The selection of chromosomes for the next generation is called reproduction, which is determined by the fitness of an individual. Different selection procedures are used in GA depending on the fitness values, of which proportional selection, tournament selection and ranking are the most popular procedures.

Introduction of random changes at the gene level is done by the mutation operator. Mutation plays a key role in reintroducing genetic diversity back into the population and assists the search to escape from local optima. In this study, the settings for GA in MATLAB are as follows: Population: 100; Generations: 100; Elite Function: 2; Crossover fraction: 0.8; selection procedure: tournament selection; crossover function: Simulated Binary Cross-over function.

Genetic algorithm is a more flexible tool used here for reducing reaction time and increasing the product concentration.

Taguchi Design

The levels of the factors affecting the enzyme yield were optimized via the Taguchi method¹¹. The parameters pH, temperature, incubation period, substrate and inoculum volume are chosen in the experiments using the domain knowledge. The Domain Knowledge although has a significant impact in choosing the modeling parameters, it must also statistically validate its significance. Several approaches for design of experiments are proposed by “Genichi Taguchi” called Taguchi method. This method is now widely used in different fields like biotechnology, engineering, marketing¹². The method utilizes a fractional factorial design in form of an orthogonal array containing a representative set of all possible combinations of experimental conditions.

Table 1 shows the chosen process parameters and their levels used. The appropriated Taguchi orthogonal array for 5 parameters is L25. Table 2 represents the L25 array, the first column represents the experiment number and the subsequent columns represent the parameters and the rows represent the parameter settings in the different experiments. Taguchi recommended a logarithmic transformation of MSE (Signal to noise ratio – S/N) in order to increase the robustness of design against noises and to accommodate wide range data for the analysis of the results. This method helps in proper analysis of results and consistency in the predicted performance. Table 2 shows the S/N ratios for each parameter and the levels are found using the statistical tool Minitab v16.

Table 1 Process parameters and levels of the experiment

Level	pH	Temperature (°C)	Incubation Period (hours)	Inoculum volume(ml)	Substrate volume(ml)
1	5	25	12	0.5	0.5
2	5.5	30	24	1	1
3	6	35	36	1.5	1.5
4	6.5	40	48	2	2
5	-	-	60	-	-

Table 2 L25 array and S/N ratio values obtained by Taguchi method

pH	Temperature (°c)	Incubation Period (hrs)	Inoculum (ml)	Substrate (ml)	Yield	S/N ratios
5	25	12	0.5	0.5	0.112	19.01564
5	30	24	1	1	0.316	10.00626
5	35	36	1.5	1.5	0.243	12.28787
5	40	48	2	2	0.087	21.20961
5	25	60	0.5	0.5	0.255	11.8692
5.5	25	24	1.5	2	0.153	16.30617
5.5	30	36	2	1	0.172	15.28943
5.5	35	48	1	0.5	0.308	10.22899
5.5	40	60	0.5	0.5	0.186	14.60974
5.5	30	12	1	1.5	0.134	17.4579
6	25	36	1.5	1	0.201	13.93608
6	30	48	0.5	1.5	0.19	14.42493
6	35	60	1	2	0.187	14.56317
6	40	12	1.5	1.5	0.181	14.84643
6	35	24	2	0.5	0.222	13.07294
6.5	25	48	1	2	0.115	18.78604
6.5	30	60	1.5	0.5	0.158	16.02686
6.5	35	12	2	1	0.189	14.47076
6.5	40	24	2	1.5	0.187	14.56317
6.5	40	36	0.5	2	0.206	13.72266
5	25	60	2	1.5	0.206	13.72266
5.5	30	12	1	2	0.132	17.58852
6	35	24	0.5	0.5	0.229	12.80329
6.5	40	36	1	0.5	0.209	13.59707
6.5	35	48	1.5	1	0.254	11.90333

Results and discussion

The ANN was trained with plane back-propagation network as well as with Bayesian Regularization. The correlation is one of the statistical measures used to judge the goodness of fit of a model. Correlation between two variables X and Y is measured using the Pearson product-moment coefficient, which takes the value between -1 and +1 inclusive. It is defined by the formula:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

The ideal prediction is supposed to give a value of r which is equal to one. Consequently, the Ideal prediction leads to a straight line with slope 1, as the X-axis and Y-axis represent the experimental and predicted values by each of the methods employed. The correlation plot of the ANN model for the 5 parameter is shown in Figure 1 and the architecture is shown in Figure 2.

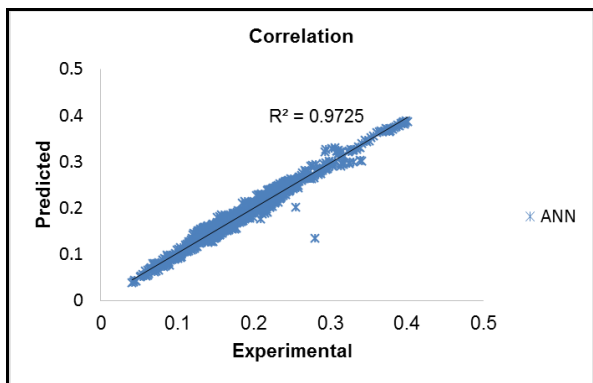


Figure1. Correlation plot of ANN for 5- parameters

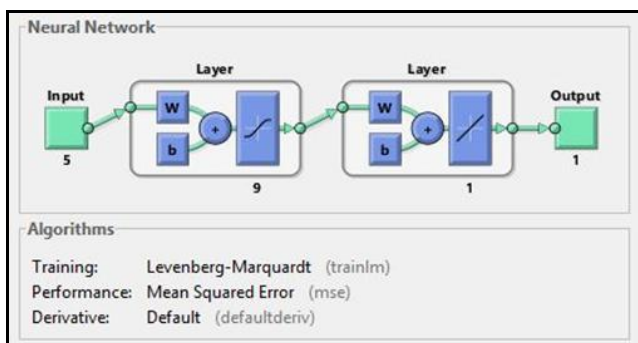


Figure 2. The architecture of ANN for 5-parameter

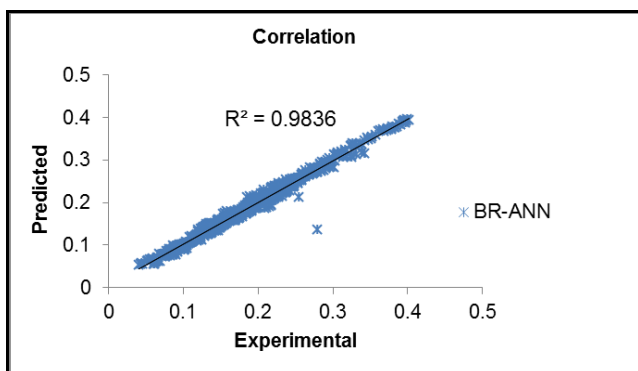


Figure 3. Correlation plot of BR-ANN for 5-parameters

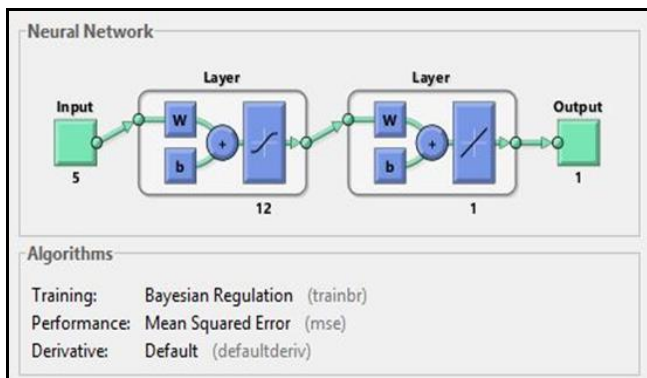


Figure 4. The architecture of BR-ANN for 5-parameter

Similarly the correlation plot for Bayesian regulation used in addition to back propagation was shown in Figure 3. The architecture of BR-ANN is shown in Figure 4.

The results showed that there is satisfactory goodness of fit for the model constructed. On applying the optimization technique of genetic algorithm, the parameters and the predicted yield is obtained for both the Neural Network Models as shown in Table 3. The final column represents the deviation in the yield from the experimental value found out by

$$\% deviation = \frac{|experimental\ yield - predicted\ yield|}{predicted\ yield} \times 100$$

The % deviation in the value is less than 10% for all the models obtained. Further it can be noted that by using the Bayesian Regularized ANN, the % deviation is also reduced drastically which emphasizes the point the BR-ANN is good in generalization capability.

Table 3 Settings of parameters obtained from GA (Genetic algorithm) tool box and predicted yield

Model	pH	Temp	LP	I	S	Yield	Experimental yield	% variation
Single layer ANN	5.00253	28.4171	49.0408	1.45082	1.98817	0.41421	0.451	8.8824938
Single layer ANN-with Bayesian Regularization	5.1402	25.0044	53.0863	1.45556	1.99665	0.41484	0.422	1.7259666

Taguchi method is applied for the parameters. The percentage contribution of each parameter is calculated and shown in Table 4. According to these data, temperature, substrate volume and incubation period were more effective than the other factors, and pH of the medium and inoculum volume showed the least impact factor. Pooling of factors was carried out to understand whether leaving out the factors with little impact can change lipase production significantly or not. Hence pooling of the values of pH 5 and inoculum volume 1.45 ml by keeping them constant without adjustment, the ANN and BR-ANN models were built with only 3 input parameters and one output. The correlation for the models is calculated and the predicted versus the actual values plot for ANN is shown in the Figure 5. The Architecture of ANN for 3parameteris shown in Figure 6.

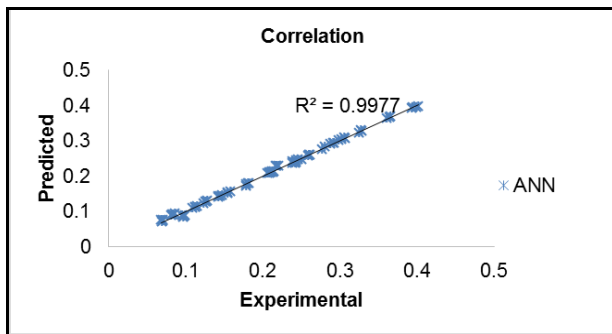


Figure 5. Correlation plot of ANN for 3- parameters

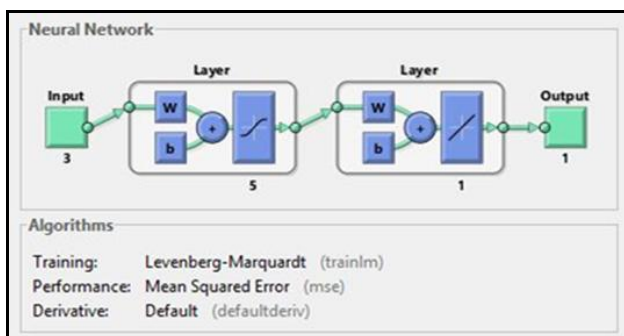


Figure 6. The architecture of ANN for 3-parameter

Table 4 Percentage contribution of each parameter

Parameter	%significance
pH	3.88939
Temperature	23.87197
Incubation Period	32.58957
Inoculum	3.56124
Substrate	36.08783

Similarly correlation plot obtained by BR-ANN is shown in Figure 7 and the architecture was shown in Figure 8.

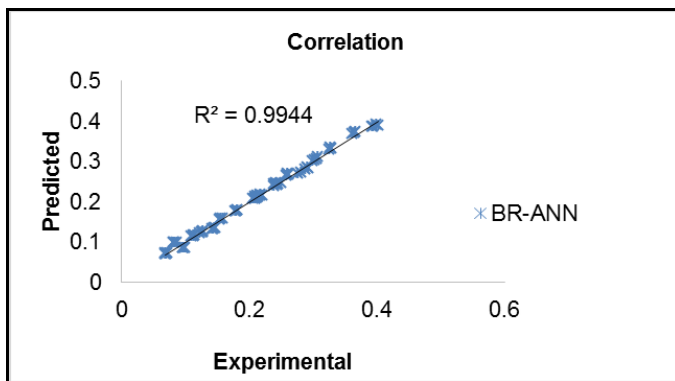


Figure 7. Correlation plot of BR-ANN for 3-parameters

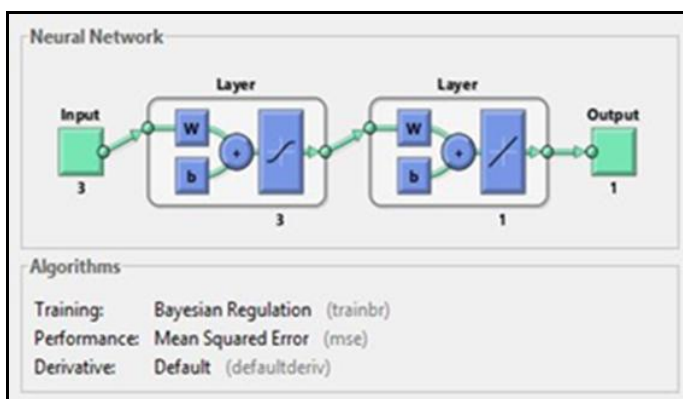


Figure 8. The architecture of BR-ANN for 3-parameter

Table 5 Representing predicted experimental yield and percent deviation from predicted value.

Model	parameters			yield	Yield from experiment	% variation
Single layer ANN	29.0957	45.2212	2	0.40003	0.469	17.24062075
Single layer ANN-with Bayesian Regularization	28.8248	44.1116	1.99989	0.37752	0.471	24.76028014

On applying the optimization technique of genetic algorithm, the parameters and the predicted yield is obtained for both the Neural Network Models. Table 5 represents the parameters and the predicted yield, experimental yield and the % deviation from the predicted value. It is been observed that the experimental yield has increased drastically, when compared to one variable at a time approach which suggests that the removal of insignificant parameters and training the network has improved in the performance of the network. It is also observed that % deviation has also increased. This may be due to neglecting the effect of pH and Inoculum in the modeling of neural network. It is further seen that the yield obtained by the setting of BR-ANN is better than

the ANN model. Through classical OFAT (one factor at a time method) the co-culture of *Lactobacillus brevis* and *Lactobacillus plantarum* produced activity of 36.98 U/ml¹³. Where as using selected optimal conditions, experiments were designed for lipase synthesis and one fold increase in lipase activity (74.05 U/ml) was obtained experimentally and this was closer to the predicted value 73.59 U/ml as shown in Table 6.

Table 6 Validation of model showing lipase production at optimum level of all parameters

Parameters	Optimal conditions	Predicted activity	observed activity
pH	5.14	73.59 U/ml	74.05 U/ml
Temperature (°C)	39.0		
Incubation period(hours)	48.0		
Inoculum volume(ml)	1.45		
Substrate volume(ml)	1.99		

Conclusion

Determining optimal process parameter settings critically influences productivity, quality, and cost of production in the field of enzyme production. Previously, either trial-and-error or elimination methods are used to optimize the process parameters for enzyme production. However, these methods are unsuitable in present because of the requirement of multi-response quality characteristics of the product. Therefore, various hybrid optimization models are constructed to find the optimal yield. It is found that BR-ANN is better than the ANN model in terms of generalization capability, and BR-ANN-GA model gives better results than ANN-GA. Further, Taguchi design has been applied to improve the effectiveness of the optimization process. The optimum yield is obtained after applying the taguchi design by neglecting the effects of pH and Inoculum.

Hence it can be concluded that hybrid integration models such as BR-ANN-GA and ANN-GA are used for optimization of the yield of lipase. Further it can be seen that application of taguchi design enhances the optimization process but has considerable deviation because of neglecting the parameters. The lipase obtained will be used as a catalyst to produce different esters which can be further applied in different industries.

References:

1. Ray A., Application of lipase in industry, Asian J. Pharm. Tech., 2012, 2,33-37.
2. Zelner I, Shor S, Lynn H, Roukema H, Lum L, Eisinga K and Koren G., Clinical use of Meconium fatty acid ethyl esters for identifying children at risk for Alcohol-related disabilities, J. Popul Ther. Clin Pharmacol., 2012, 19, 26-31.
3. Chan Cupul W, Heredia Abarca G, Martínez Carrera D and Rodríguez Vázquez R., Enhancement of ligninolytic enzyme activities in a *Trametes maxima*–*Paecilomyces carneus* co-culture: key factors revealed after screening using a Plackett–Burman experimental design, Electron J. Biotechnol., 2014, 17, 114–121.
4. Stergiou P.Y., Foukis A, Theodorou L, Papagianni M and Papamichael E., Optimization of the production of extracellular α -amylase by *Kluyveromyces marxianus* IF0 0288 by response surface methodology, Braz. Arch. Boil. Technol., 2014, 57, 421-426.
5. Ghosh S, Murthy S, Govindasamy S and Chandrasekaran M., Optimization of L-asparaginase production by *Serratia marcescens* (NCIM 2919) under solid state fermentation using coconut oil cake, Sustain. Chem. Process., 2013, 1, 9.
6. Garlapati V.K. and Banerjee R., Solvent-free synthesis of flavor esters through immobilized lipase mediated transesterification, Enzyme. Res., 2013, 2013, 6.
7. Khayati G, Gilani H.G. and Kazemi M., The effect of olive cake types on lipase production by isolated *Rhizopus* sp. and process statistical optimization, J. BioSci. Biotech., 2013, 2, 45-55.
8. Okut H, Wu X. L., Rosa G. J. M., Bauck S, Woodward B. W., Schnabel R. D., Taylor J.F. and Gianola D., Predicting expected progeny difference for marbling score in Angus cattle using Artificial neural networks and Bayesian regression models, Genet. Sel. Evol., 2013, 45, 34.
9. Berberler M. E., Guler A and Nuriyev U. G., A Genetic algorithm to solve the multidimensional knapsack problem, Math. Comput. Appl., 2013, 18, 486-494.

10. Uppada S. R., Gupta A. K., and Dutta J. R., Statistical optimization of culture parameters for lipase production from *Lactococcuslactis* and its application in detergent industry, Int. J.Chem Tech. Res.,2012, 4, 1509-1517.
11. LópezL. L., Ortiz G. L., Martínez C. J., andSánchez H. H., Sequentially integrated optimization of the conditions to obtain a high-protein and low-anti nutritional factors protein isolate from edible *Jatropha curcas* seed cake, ISRN Biotechnol., 2013, 2013 7.
12. Santangelo S, Lanza M, Piperopoulos E, Galvagno S andMilone C., Optimization of CVD growth of CNT-based hybrids using the Taguchi method, Mater. Res. Bull., 2012, 47, 595-601.
13. Ramyasree S and Dutta J. R., The effect of process parameters in enhancementof lipase production by co-culture oflactic acid bacteria and their mutagenesisstudy, Biocatal.AgricBiotechnol., 2013, 2, 393–398.
