Modeling and optimization of fermentation variables for enhanced production of lactase by isolated *Bacillus subtilis* strain VUVD001 using artificial neural networking and response surface methodology

T. C. Venkateswarulu¹ · K. Vidya Prabhakar² · R. Bharath Kumar¹ · S. Krupanidhi¹

Received: 4 January 2017 / Accepted: 11 May 2017 / Published online: 29 June 2017
© Springer-Verlag GmbH Germany 2017

Abstract Modeling and optimization were performed to enhance production of lactase through submerged fermentation by *Bacillus subtilis* VUVD001 using artificial neural networks (ANN) and response surface methodology (RSM). The effect of process parameters namely temperature (°C), pH, and incubation time (h) and their combinational interactions on production was studied in shake flask culture by Box–Behnken design. The model was validated by conducting an experiment at optimized process variables which gave the maximum lactase activity of 91.32 U/ml. Compared to traditional activity, 3.48-folds improved production was obtained after RSM optimization. This study clearly shows that both RSM and ANN models provided desired predictions. However, compared with RSM ($R^2 = 0.9496$), the ANN model ($R^2 = 0.99456$) gave a better prediction for the production of lactase.

Keywords *B. subtilis* strain VUVD001 · Process variables · Shake flask culture · Box–Behnken design

Introduction

Lactase intolerance is the inability to digest milk lactose due to the non-availability of lactase among a few children and adults and is treated with lactase capsules. This

enzyme has commercial value in pharmaceutical and food industries. Particularly, in the dairy industry, it is used for hydrolysis of lactose in milk or other products such as cheese and whey (Gasteiger et al. 2003). It is also used in the production of prebiotic galactooligosaccharides because of its transgalactosylation activity (Iqbal et al. 2010). Lactase was produced by different microorganisms, namely filamentous fungi, bacteria, and yeasts. Traditionally, lactase was obtained from *Aspergillus* sp. and *Kluyveromyces* sp. because of acceptable yields from the cultivation of these organisms (Torsvik et al. 1998; Picard et al. 2005; Kosseva et al. 2009; Bibi et al. 2014). The bacterial species are extensively used for large scale production due to metabolic diversity and rapid growth (Prakaham et al. 2015; Rao et al. 2008). The fermentation process parameters such as growth, temperature, pH, incubation time, agitation, and nutritional source availability affect the yield. Therefore, optimization of factors is the primary step in designing the process for improving production (Mahoney et al. 1975; Kamaran et al. 2015). Modeling and optimization are the significant stages in bioprocess research to enhance the production. The classical optimization method is time consuming and it cannot give a complete picture of independent factors affecting the production and also the combinational interaction of variables on process could not be understood. But in response surface methodology (RSM), we can predict the effect of independent variables and their interactions on yield. RSM is an acquisition of statistical and mathematical techniques which is used to describe the relative effects between process output and variables in process. This model also generates mathematical equation based on the experimental data (Basri et al. 2007; Bas and Boyacı 2007). Artificial neural network (ANN) model is a well established and fashionable tool in analysis and is also used to understand

✉ T. C. Venkateswarulu
venki_biotech327@yahoo.com

¹ Department of Biotechnology, Vignan's Foundation for Science Technology and Research University, Vadlamudi, Andhra Pradesh 522213, India

² Department of Biotechnology, Vikrama Simhapuri University, Nellore, Andhra Pradesh 524320, India

the biotechnology applications such as expression function, functional analysis of genomics and proteomics. Artificial neural network is an extremely connected network structure consisting of many processing elements which are capable of performing parallel computation for data processing (Manohar and Divakar 2005). This work was designed to improve the production through an optimization of three culture conditions which influences lactase yield by *B. subtilis* strain VUVD001 in shake flask fermentation using RSM and ANN.

Materials and methods

Microorganism and culture medium

VUVD001 strain used for the study was isolated from the dairy effluent, Sangam Dairy, Vadlamudi, Guntur Dist., India. The VUVD001 strain was identified as *B. subtilis* through 16S rRNA analysis and the strain was preserved at 4 °C on nutrient agar medium in our laboratory. One loop full of bacterial culture from a 36-h old agar slant was transferred to the lactose broth medium for inoculum preparation and production. The compositions of the medium were: (in g/L) lactose 14.01, yeast extract 10.30, tryptophan 0.43, and magnesium sulfate (MgSO_4) 5.321.

Lactase enzyme assay

The lactase activity was determined using *ortho*-nitrophenyl- β -galactoside (ONPG) as the substrate. The ONPG solution was prepared with phosphate buffer and used for the assay. 0.5 ml of enzyme source was added with 2.0 ml of substrate and incubated for 30 min. The reaction stopped with the addition of 0.5 ml of 1 M Na_2CO_3 and absorbance was recorded at 420 nm. Activity of lactase was determined from ONP standard graph. One unit of activity is defined as an amount of enzyme that liberates 1 micromole of ONP from the substrate per minute under assay conditions (Pulicherla et al. 2013).

Culture conditions

The liquid fermentation was conducted batchwise in 250 ml flask using 100 ml of lactose broth medium and is sterilized at 121 °C for 15 min. The flasks were inoculated with 5% inoculum and the percent of inoculum was selected based on traditional optimization. Then, the flasks were incubated at 37 °C on rotary shaker driven at 150 rpm. The growth condition levels of temperature, pH, and incubation time used in the optimization study, by an application of response surface methodology are given in Table 1.

Table 1 Variable ranges used in experimental design

Symbol	Name of the variable	Range
A	Temperature (°C)	35–40
B	pH	5–8
C	Incubation time (h)	10–50

Prediction of variables' effect on production

Box–Behnken method was applied to predict maximum ranges of the significant factors (temperature, pH and incubation time) and their combinational interaction on production. Design Expert version 7.0 software was utilized for the design of the experiment, data analysis, and development of statistical model. Each experiment was studied in triplicate and average yield obtained was taken as the response (R), while the predicted values of response were obtained from the quadratic model. The variable's impact on response was predicted through regression analysis of data. Three-dimensional surface plots were obtained to understand the variable effect and to determine their optimum levels for production. Response surface model was fitted to response, i.e., activity of lactase (U/ml). The second-order response function for three variables is given by the following equation:

$$R = \beta_0 + \beta_1A + \beta_2B + \beta_3C + \beta_{12}AB + \beta_{13}AC + \beta_{23}BC + \beta_{11}A^2 + \beta_{22}B^2 + \beta_{33}C^2$$

where R is a dependent variable (lactase production) and A , B , and C are independent variables (temperature, pH, and incubation time, respectively), β_0 is an intercept term and β_1 , β_2 , and β_3 are linear coefficients whereas, β_{12} , β_{13} , and β_{23} are the interaction coefficients. β_{11} , β_{22} , and β_{33} are the quadratic coefficients.

Modeling using artificial neural networks

Artificial neural network (ANN) modeling is an alternative tool to RSM for regression analysis of polynomial non-linear systems. ANN architecture is an interlinked complex with elements such as neurons and the connections between the neurons were described by weights (w) and bias (b). The neurons were controlled by transfer and summing functions and general transfer functions include purelin, log sig, and tan sig (Das et al. 2015). In the present study, the predictive model was developed using temperature (°C), pH, and incubation time (h) as input variables and lactase activity (U/ml) as the output for the model. The input layer function is to pass the scaled input values to hidden layer through weights. A back-Propagation algorithm is used with one hidden layer enhanced with Levenberg–Marquardt optimization method (Arcaklioglu et al. 2004).

Results and discussion

Optimization by Box–Behnken design

Statistical model for experimental design is an essential tool in optimizing conditions that may bring severalfold increase in production. The effect of parameters and their interaction on lactase synthesis was determined by conducting 17 experiments given by the model (Table 2). Box–Behnken design provides necessary information about effects of variables on response. The quadratic model equation was obtained by Box–Behnken design, which predicts the variables impact on response:

$$R1 = +86.68 - 12.42 \times A + 8.39 \times B + 9.99 \times C \\ - 10.52 \times A \times B - 0.060 \times A \times C - 2.56 \times B \times C \\ - 30.83 \times A^2 - 25.93 \times B^2 - 19.84 \times C^2.$$

where the response (R1) is lactase production and other variables *A*, *B*, and *C* represent temperature, pH, and incubation time, respectively.

The coefficient R^2 value 0.9496 recommends that the design was significant to predict the effect of variables on production by *B. subtilis* VUVD001. The model *F* value of 14.64 implies that the model is significant. The values of “Prob > *F*” less than 0.0500 illustrate that the model terms are significant. In this case, *A*, *B*, *C*, A^2 , B^2 , and C^2 are significant model terms. Values greater than 0.1000 indicate that the model terms are not significant. The “Lack of Fit *F* value” of 5.87 shows that, the Lack of Fit is not

significant relative to the pure error (Table 3). The three-dimensional response plots gave understandable information to predict the relationship between response and variable range. The graphs show an oval shape curve that suggests optimum conditions and combinational effects on production. It has been observed that maximum production was attained at 38 °C, pH 6.57 and incubation time of 35 h (Fig. 1). In previous reports, several workers achieved highest enzyme activity at 37 °C and 24 h incubation time through submerged fermentation with *Lactobacillus delbrueckii* (Mozumder et al. 2012) and *L. amylophilus* GV6 at 48 h (Mozumder et al. 2012) and *Bacillus* sp. MPTK 121 (Kumar et al. 2012). Kamaran et al. (2015) achieved the highest lactase activity of 62 U/ml from *Bacillus* strain B-2 at optimum conditions of 48 h incubation time and 37 °C. Similarly, in 2015 Carevic et al. has observed optimum activity of 1.01I U/ml at 37 °C, pH 6.5–7.5 and incubation time 48 h in shake flask culture fermentation using *L. acidophilus* ATCC 4356 (Carevic et al. 2015). These results showcase that the bacterium maintained its exponential phase within 24–48 h. However, in the present study, the high activity of 91.32 U/ml was accomplished at optimized conditions with lactose broth medium.

Validation of RSM model

The RSM model is validated by conducting an experiment with the best predicted solution given by RSM for the production of lactase. The enzyme activity reached 91.32 U/ml from *Bacillus subtilis* VUVD001 after

Table 2 Actual data for design of experiments

Run	Temp. (°C)	pH	Incubation time (h)	Activity (U/ml)		
				Experimental	RSM-predicted	ANN-predicted
1	37.5	6.5	30	88.64	86.67	85.59
2	35	8	30	67.35	61.25	67.35
3	37.5	5	50	41.26	45.07	43.13
4	35	6.5	10	28.46	38.37	28.46
5	40	6.5	10	9.57	13.66	12.79
6	37.5	5	10	30.16	19.96	30.16
7	37.5	8	50	46.53	56.72	44.03
8	37.5	6.5	30	81.45	86.67	85.59
9	40	6.5	50	43.44	33.52	41.98
10	37.5	6.5	30	91.04	86.67	85.59
11	37.5	6.5	30	80.12	86.67	85.59
12	40	5	30	13.55	19.64	13.55
13	40	8	30	15.65	15.37	15.69
14	35	5	30	23.15	23.42	23.15
15	37.5	8	10	45.68	41.86	45.68
16	37.5	6.5	30	92.14	86.67	85.59
17	35	6.5	50	62.57	58.47	62.57

Table 3 Analysis of variance of quadratic model for production of Lactase

Source	Sum of squares	Df	Mean squares	F value	P value
Quadratic model	12508.84	9	1389.871	14.63993	0.0009
A-Temperature	1233.058	1	1233.058	12.98816	0.0087
B-pH	562.6335	1	562.6335	5.926386	0.0451
C-incubation time	798.6006	1	798.6006	8.411898	0.0230
AB	443.1025	1	443.1025	4.667331	0.0676
AC	0.0144	1	0.0144	0.000152	0.9905
BC	26.26563	1	26.26563	0.276664	0.6151
A ²	4000.825	1	4000.825	42.14189	0.0003
B ²	2830.519	1	2830.519	29.8147	0.0009
C ²	1657.83	1	1657.83	17.46242	0.0041
Residual	664.5592	7	94.93703		
Lack of fit	541.5099	3	180.5033	5.87	0.0602
Pure error	123.0493	4	30.76232		
Cor total	13173.4	16			

$R^2 = 0.9496$, Adj. $R^2 = 0.8847$, CV = 19.24%, Adeq. precision = 9.770

Fig. 1 The three-dimensional response plots for combinational effects of temperature, pH, and incubation time on lactase production

statistical optimization and this value is almost near to the RSM-predicted value (Table 4). Previously, in 2011 Rashmi and Siddalingamurthy (2011) reported that the lactase activity of 0.31 U/ml was achieved by *Aspergillus*

terreus even after RSM model. Compared with this fungal strain, our bacterium has higher potential for lactase production. Hsu et al. (2007) reported that maximum beta-galactosidase activity of 18.6 U/ml was obtained in

Table 4 RSM optimized process variables for maximum lactase activity

Temp. (°C)	pH	Incubation time (h)	Activity (U/ml)	
			Predicted	Experimental
36.91	6.8	34.77	90.16	91.32

Fig. 2 Output vs. target regression plot

submerged fermentation with *Bifidobacterium longum* CCRC 15708. The process was run at optimum conditions of pH 6.5, temperature 37 °C, and incubation time 16 h in liquid medium containing 4% lactose, 3.5% yeast extract, 0.3% K_2HPO_4 , 0.1% KH_2PO_4 , 0.05% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, and 0.03% L-cysteine. Anisha et al. (2008a, b) reported that the activity reached from 17 to 50 U/ml under the conditions of 33 °C and pH 7 in submerged fermentation by *Streptomyces griseolalbus* in Box–Behnken design. Lee et al. (2013) gained a maximum activity of lactase of 1.06 U/ml in submerged cultivation with *Bacillus* sp. LX-1. In the same way, Edupuganti et al. (2014) reported that the lactase production in submerged fermentation by *Acinetobacter* sp. achieved the activity of 10.2 U/ml after the modeling and genetic algorithm optimization. The process was carried out under the optimum conditions of

temperature 37 °C, pH 7.2, and agitation speed 183 RPM. Similarly Arekal et al. (2013) reported that highest lactase activity of 10.6 U/ml was obtained with *L. plantarum* MTCC 5422 in the soy–whey-based medium at RSM-optimized conditions of incubation time 33 h, temperature 37 °C, and pH 6.6. Based on these observations, it was proved that the predicted model is significant and the optimized process conditions were used in bioindustries for large scale production.

Development of neural network model and result analysis

Training, testing, and validation of neural networks were performed with three input variables and one output using tools, namely feed forward back-propagation network and

Fig. 3 Comparison of experimental data with RSM- and ANN-predicted values

TRAINLM in MATLAB R2013a version. The elevated regression value of 0.9945 was attained from the ANN model. The performance curve was developed using MATLAB R2013a for training, testing, and validation of the data. The regression plot of the output versus target was developed with ten hidden nodes and R^2 value 0.99 was accomplished with validation of the model (Fig. 2).

RSM- and ANN-predicted values vs experimental data

The RSM- and ANN-predicted data were compared with experimental values (Table 1; Fig. 3). It was observed that the ANN prediction is almost similar to experimental values.

Conclusion

The effect of temperature, pH, and incubation period and their combinational interaction on production was predicted using Box–Behnken design. Under the optimized conditions, highest activity of 91.32 U/ml was obtained. ANN model gave the R^2 value of 0.9945 for the RSM-predicted production data. Based on the study, it was confirmed that ANN showed excellent predictable accuracy and can be used in modeling of bioprocess. These findings reveal that *Bacillus subtilis* VUVD001 is also an alternative promising strain for commercial production of lactase.

Acknowledgements Authors acknowledge Vignan Foundation for Science, Technology and Research University, Guntur for providing facilities. Special thanks to Mr. E. R. Reddy, Assistant Professor, Department of Biotechnology, Vignan University, Vadlamudi for his active help in modeling the study.

Compliance with ethical standards

Conflict of interest The authors declared that there are no conflicts of interest on publication of this article.

References

- Anisha GS, Sukumaran RK, Prema P (2008a) Evaluation of α -galactosidase biosynthesis by *Streptomyces griseolobus* in solid-state fermentation using response surface methodology. *Lett Appl Microbiol* 46:338–343
- Anisha GS, Sukumaran RK, Prema P (2008b) Statistical optimization of α -galactosidase production in submerged fermentation by *Streptomyces griseolobus* using response surface methodology. *Food Technol Biotechnol* 46:171–177
- Arcaklioglu E, Cavusoglu A, Erisen A (2004) Thermodynamic analyses of refrigerant mixtures using artificial neural networks. *Appl Energy* 78(2):219–230
- Arekal N, Roopashri AN, Varadaraj MC (2013) Soy whey based medium for optimized phytase activity in *Saccharomyces cerevisiae* MTCC 5421 and α -D-galactosidase and antibacterial activities in *Lactobacillus plantarum* MTCC 5422 by response surface methodology. *J Food Sci Technol* 51:519–526
- Bas D, Boyacı İH (2007) Modeling and optimization II: comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction. *J Food Eng* 78:846–854
- Basri M, Rahman RNZRA, Ebrahimpour A, Salleh AB, Gunawan ER, Rahman MBA (2007) Comparison of estimation capabilities of response surface methodology (RSM) with artificial neural network (ANN) in lipase-catalyzed synthesis of palm-based wax ester. *BMC Biotechnol* 7:1–14
- Bibi Z, Ansari A, Zohra RR, Aman A (2014) Production of xylan degrading endo-1,4- β -xylanase from thermophilic *Geobacillus stearothermophilus* KIBGE-IB29. *J Radiat Res Appl Sci* 7:478–485
- Carevic M, Vukasinovic-Sekulic M, Grbavcic S, Stojanovic M, Mihailovic M, Dimitrijevic A, Bezbradica D (2015) Optimization of β -galactosidase production from lactic acid bacteria. *Hem ind* 69:305–312
- Das S, Bhattacharya A, Haldar S, Ganguly A, Sai Gu YP, Ting PK (2015) Optimization of enzymatic saccharification of water hyacinth biomass for bio-ethanol: comparison between artificial neural network and response surface methodology. *Sustain Mater Technol* 3:17–28
- Edupuganti S, Potumarthi R, Sathish T, Mangamoori L (2014) Role of feed forward neural networks coupled with genetic algorithm in capitalizing of intracellular alpha-galactosidase production by *Acinetobacter* sp. *Biomed Res Int* 2014:1–8
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31:3784–3788

- Hsu CA, Lee YRC, Chou CC (2007) Cultural condition affecting the growth and production of β -galactosidase by *Bifidobacterium longum* CCRC 15708 in a jar fermenter. *Int J Food Microbiol* 116:186–189
- Iqbal S, Nguyen TH, Nguyen TT, Maischberger DH (2010) Beta-galactosidase from *Lactobacillus plantarum* WCFSI, biochemical characterization and formation of prebiotic galacto-oligosaccharides. *Carbohydr Res* 345:1408–1416
- Kamaran A, Bibi Z, Aman A, Qader SAU (2015) Lactose hydrolysis approach: isolation and production of β -galactosidase from newly isolated *Bacillus* strain B-2. *Biotech Agric Biotechnol* 12:1–19
- Kosseva MR, Panesar PS, Kaur G, Kennedy JF (2009) Use of immobilised biocatalysts in the processing of cheese whey. *Int J Biol Macromol* 45:437–447
- Kumar DJM, Sudha M, Devika S, Balakumaran MD, Ravi Kumar M, Kalaichelvan PT (2012) Production and optimization of β -galactosidase by *Bacillus* Sp. MPTK 121, isolated from dairy plant soil. *Ann Biol Res* 3:1712–1718
- Lee J, Park I, Cho S, Kim S, Choi Y, Cho J (2013) Optimization of medium composition for α -galactosidase production by Antarctic bacterial isolate, *Bacillus* sp. LX-1 using response surface methodology. *Afr J Microbiol* 7:3494–3500
- Mahoney RR, Nickerson TA, Whitaker JR (1975) Selection of strain, growth conditions, and extraction procedures for optimum production of lactase from *Kluyveromyces fragilis*. *J Dairy Sci* 58:1620–1629
- Manohar B, Divakar S (2005) An artificial neural network analysis of porcine pancreas lipase catalysed esterification of anthranilic acid with methanol. *Process Biochem* 40:3372–3376
- Mozumder NHMR, Akhtaruzzaman M, Baki MA, Zohra FT (2012) Study on isolation and partial purification of Lactase enzyme from *Lactobacillus* bacteria isolated from yogurt. *J Sci Res* 4:239–249
- Picard C, Fioramonti J, Francois A, Robinson T, Neant F, Matuchansky C (2005) Review article: *bifidobacteria* as probiotic agents-physiological effects and clinical benefits. *Aliment Pharmacol Ther* 22:495–512
- Prakaham RS, Subba Rao Ch, Sreenivas Rao R, Sarma PN (2015) Alkaline protease production by an isolated *Bacillus circulans* under solid state fermentation using agroindustrial waste: process parameters optimization. *Biotechnol Prog* 21:1380–1388
- Pulicherla KK, Suresh Kumar P, Manideep K, Rekha VPB, Ghosh Mrinmoy, Sambasiva Rao KRS (2013) Statistical approach for the enhanced production of coldactive β -galactosidase from *Thalassospira frigidophilosprofundus*: a novel marine psychrophile from deep waters of bay of Bengal. *Prep Biochem Biotechnol* 43:766–780
- Rao Ch, Sathish S, Mahalaxmi T, Laxmi M, Rao GS, Prakasham RS (2008) Modelling and optimization of fermentation factors for enhancement of alkaline protease production by isolated *Bacillus circulans* using feed-forward neural network and genetic algorithm. *J Appl Microbiol* 104:889–898
- Rashmi R, Siddalingamurthy KR (2011) Optimization of β -galactosidase production by response surface methodology. *Int J Biosci* 1:119–127
- Torsvik V, Daae FL, Sandaa RA, Ovreas L (1998) Novel techniques for analysing microbial diversity in natural and perturbed environment. *J Biotechnol* 64:53–62