Appl Biochem Biotechnol (2009) 158:25–40
DOI 10.1007/s12010-008-8315-z

# Constitutive Expression and Optimization of Nutrients for Streptokinase Production by *Pichia pastoris* Using Statistical Methods

**Ravi Nagaraj Vellanki · Ravichandra Potumarthi · Lakshmi Narasu Mangamoori**

Received: 4 April 2008 / Accepted: 2 July 2008 /
Published online: 25 July 2008


**Abstract** The *Pichia pastoris* clone producing streptokinase (SK) was optimized for its nutritional requirements to improve intracellular expression using statistical experimental designs and response surface methodology. The *skc* gene was ligated downstream of the native glyceraldehyde 3-phosphate dehydrogenase promoter and cloned in *P. pastoris*. Toxicity to the host was not observed by SK expression using YPD medium. The transformant producing SK at level of 1,120 IU/ml was selected, and the medium composition was investigated with the aim of achieving high expression levels. The effect of various carbon and nitrogen sources on SK production was tested by using Plackett–Burman statistical design and it was found that dextrose and peptone are the effective carbon and nitrogen sources among all the tested. The optimum conditions of selected production medium parameters were predicted using response surface methodology and the maximum predicted SK production of 2,136.23 IU/ml could be achieved with the production medium conditions of dextrose ($x1$), 2.90%; peptone ($x2$), 2.49%; pH, 7.2 ($x3$), and temperature, 30.4 ($x4$). Validation studies showed a 95% increase in SK production as compared to that before optimization at 2,089 IU/ml. SK produced by constitutive expression was found to be functionally active by plasminogen activation assay and fibrin clot lysis assay. The current recombinant expression system and medium composition may enable maximum production of recombinant streptokinase at bioreactor level.

R. N. Vellanki · L. N. Mangamoori
Centre for Biotechnology, Jawaharlal Nehru Technological University,
Kukatpally, Hyderabad 500085, India

R. N. Vellanki
e-mail: ravivellanki2004@gmail.com

R. Potumarthi
Bioengineering and Environmental Centre, Indian Institute of Chemical Technology (CSIR),
Tarnaka, Hyderabad 500607, India
e-mail: pravichandra@gmail.com

L. N. Mangamoori (✉)
Centre for Biotechnology, Institute of Science and Technology,
Jawaharlal Nehru Technological University, Kukatpally, Hyderabad 500085, India
e-mail: mangamoori@jntuap.ac.in
e-mail: mangamoori@rediffmail.com

**Keywords** *Pichia pastoris* · Streptokinase · Constitutive expression · Placket–Burman design · Response surface methodology (RSM) · Optimization

## Introduction

Streptokinase (EC 3.4.99.22) (SK) is secreted as a 47 kDa protein comprising 414 amino acids by a 26 amino acid signal peptide in β-hemolytic streptococci groups A, C, and G [1]. SK acts as an activator of human fibrinolytic system by stoichiometrically binding to circulatory and thrombus bound plasminogen to generate SK–plasminogen activator complex. Cleavage of plasminogen in zymogen form at an Arg-Val bond generates plasmin, an active enzyme that degrades the fibrin component of thrombin [2].

Commercially employed SK is extracted from group C streptococci and the gene encoding SK '*skc*' of *Streptococcus equisimilis* H46A is being used for heterologous expression. Low yields of SK production and association of other antigenic molecules like Streptodornase led to expression of SK in heterologous host systems. Prokaryotic heterologous host systems *Escherichia coli*, *Bacillus subtilis*, *Proteus mirabilis, Streptococcus sanguis*, and eukaryotic host systems *Pichia pastoris* and *Schizosaccharomyces pombe* were successfully used to over express SK [3–9].

In prokaryotic hosts post-translational proteolysis of mature 47 kDa SK at its C terminus generated a truncated 44 kDa SK that retained the plasminogen activation property. The N and C termini of SK are not involved in the biological function of the molecule and the exact role is not known. Expression of *skc* gene in *P. pastoris* intracellularly generated mature 47 kDa SK and in secretory mode a glycosylated SK was expressed. Both the molecules demonstrated the plasminogen activation property. In *S. pombe* secretory expression of mature SK was achieved with thiamine-based induction system.

The AOX1-promoter-based heterologous protein expression is being employed in *P. pastoris*. With the identification and cloning of genes involved in metabolism of *P. pastoris*, promoters like Glyceraldehyde 3-phosphate dehydrogenase (GAP), Formaldehyde dehydrogenase (FLD1), Isocitrate Lyase (ICL) etc. were successfully applied for foreign gene expression. GAP promoter is constitutively active irrespective of carbon source, with best expression in presence of glucose [10]. GAP-promoter-based vectors pGAPZ and pGAPZα are commercially available [11].

The GAP-based *P. pastoris* expression system for SK was successfully applied. In this study, we report the construction of an expression vector with the mature streptokinase expressed intracellularly under the control of *GAP* promoter in *P. pastoris* and also optimization of nutrients for maximum production.

A two level design like Plackett–Burman was used to screen the different carbon sources for maximal production SK. Plackett–Burman will be of choice for screening large number of variables as this design screens $k$ variables in just $k+1$ experiments [12].

Significantly, peptone and dextrose have shown the maximal SK production. Further, these components of production medium were optimized by response surface methodology (RSM). We have obtained a level of expression considerable to that achieved in other yeast systems.

The classical method of experimental optimization involves changing one variable at a time while keeping the others constant. In addition, it is not practical to carry out experiments with every possible factorial combination of the test variables because of the large number of experiments required [13, 14]. This does not consider the effect of different interactions of various parameters. Besides this, it is a tedious, cumbersome, and time-

consuming process especially when a large number of parameters are taken into account. An alternative and more efficient approach is the use of statistical method.

Response surface methodology has been widely used to evaluate and understand the interactions between different process parameters [15]. RSM was applied successfully for optimizing process parameters for various processes in biotechnology [16–20].

In the present study, based on the results obtained by the classical approach, parameters found significantly affecting SK production by *P. pastoris* were taken into account. Two level Placket and Burman design was used for the screening of carbon sources. A $2^4$ full factorial central composite design (CCD) and RSM were used for optimization of immobilization conditions and regression analysis was performed to obtain the optimum conditions for maximal production of SK.

This is the first report for expression of SK constitutively and optimization of production medium components by statistical method.

## Materials and Methods

### Materials

All enzymes were purchased from New England Biolabs (Beverly, MA, USA), and *pfu* DNA polymerase was purchased from Stratagene (La Jolla, CA, USA). Luria Bertani (LB) medium, yeast extract peptone dextrose (YPD) components are from Himedia (Mumbai, India). Polymerase chain reaction (PCR) was performed with Eppendorf Mastercycler® via *pfu* DNA polymerase. *E. coli* DH5α from Gibco BRL (Gaithersburg, MD, USA) was employed for plasmid maintenance. *P. pastoris* GS115 (*his4*) and pGAPZA from Invitrogen (Carlsband, CA, USA) were used in expression studies; human plasminogen from Calbiochem (La Jolla, CA, USA), S-2251 from Fluka (Buchs/Switzerland), Human thrombin, and Human Fibrinogen from Sigma (St.Louis, MO, USA).

### Growth Media and Conditions

Low salt Luria broth (1% tryptone, 0.5% yeast extract, and 1.5% NaCl, pH 7.5) with Zeocin (25 μg $ml^{-1}$) was used for culture of *E. coli* DH5α transformants at 37 °C. The YPD medium (1% yeast extract, 2% Bacto peptone and 2% dextrose) containing Zeocin (0.1 mg $ml^{-1}$) was used for selection and growth of *P. pastoris* transformants at 30 °C. Protein expression studies on *P. pastoris* clones were performed in YPD medium at 30 °C.

### Construction of Recombinant Expression Vector

The *skc* gene expressing mature SK was commercially synthesized and amplified by PCR using two synthetic primers S1—CAGCAG**GAATTC**ATTGCTGGACCTGAG TGG and S2—TCCC**GCGGCCGC**TTATTTGTCGTTAGGGTTATC. Amplified *skc* gene was inserted at *EcoRI*, *NotI* sites in pGAPZA vector downstream of GAP promoter. Orientation of *skc* gene flanked by promoter and terminator in the expression vector was confirmed by restriction digestion analysis and by nucleotide sequence analysis using sequencing primers S3—GTCCCTATTTCAATCAATTGAA and S4—GCAAATGGCAT TCTGACATCC. The recombinant expression vector pGAPZA-SK (Fig. 1a) was maintained in *E. coli* DH5 α under Zeocin selection. pGAPZA-SK was used as an integrative expression vector in *P. pastoris*.

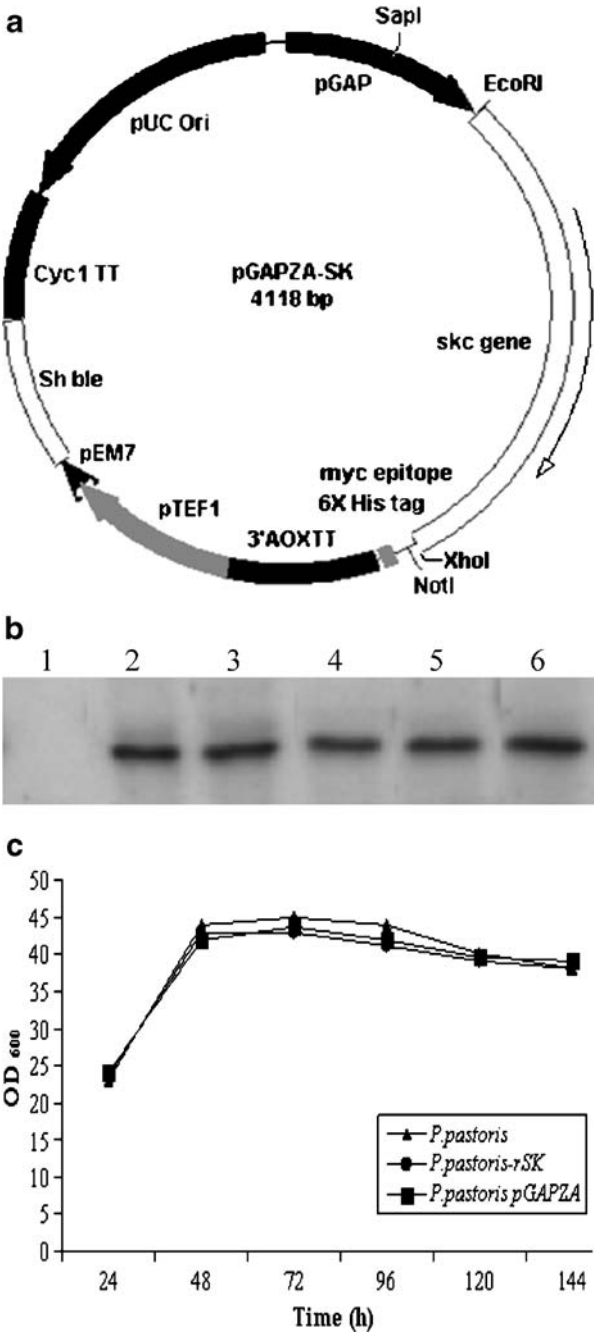


**Fig. 1** **a** Plasmid map of expression vector pGAPZASK constructed to express streptokinase (with out signal sequence) in *P. pastoris*. The position of *skc* gene and its orientation are designated. Appropriate restriction endonuclease sites are indicated. **b** Southern analyses of *P. pastoris*-SK clone. Hybridization with a $\alpha^{32}$P-labelled 1.25 Kbp *skc* gene was performed to recombinant clone genomic DNA. *P. pastoris*-SK clone genomic DNA extracted from shake flask culture at 24, 48, 72, 96, 120 h (Lanes *2–6*), Untransformed *P. pastoris* genomic DNA (Lane *1*). **c** Expression of SK has no toxic effect on the growth rate of the *P. pastoris* host. *P. pastoris* GS115 cells were transformed with pGAPZA parent plasmid and recombinant expression vector pGAPZA-SK and untransformed *P. pastoris* was used as control. Cells were cultured in YPD broth and aliquots were withdrawn at indicated time points up to 144 h, $OD_{600}$ recorded and plotted

## Transformation of *P. pastoris*

*P. pastoris* was transformed with *SapI* linearized pGPZA-SK (12 μg) vector by lithium chloride transformation method. Briefly, *P. pastoris* culture of 0.3 OD was inoculated in

30 ml of YPD medium and incubated at 30 °C to an of OD 0.8–1.0. The cells were pelleted and washed with 100 mM LiCl twice. Cells were suspended in 400 μl of 100 mM LiCl to obtain competent cells. To 60 μl competent cells 240 μl PEG-3350, 36 μl 1M LiCl, 50 μl (2 mg $l^{-1}$) of herring sperm DNA and 50 μl linearized pGPZA-SK DNA were mixed by vigorous vortexing. Competent cell/DNA mixture was incubated at 30 °C for 45 min followed by heat shock at 42 °C for 30 min. The cells were then plated on YPDZ agar and incubated at 30 °C for 3 to 4 days.

## Screening of Pichia Transformants

Transformants harboring the recombinant vector integrated into the chromosome via homologous recombination were screened for Zeocin resistance on YPDZ medium. Zeocin resistant transformants were checked for *skc* gene insertion by colony PCR using S3 and S4 primers. Transformants that were screened for *skc* gene by colony PCR were confirmed by genomic DNA PCR.

## Growth Conditions for Expression of SK

Recombinant *P. pastoris* integrated with vector pGPZA-SK was inoculated in 10-ml YPD medium and incubated at 30 °C for 36 h at 180 rpm. 1.0 OD of culture was inoculated in 30-ml YPD medium in a 150-ml flask. A time course study was done from 24–120 h to evaluate the growth and expression profile in comparison with *P. pastoris* harboring parent plasmid and untransformed strain.

*Screening of Carbon and Organic Nitrogen Sources by Plackett–Burman Design* The present study was aimed at screening of the important carbon and organic nitrogen sources with respect to their main effects and not the interaction effects between various medium constituents and hence, Plackett–Burman design was used. The Plackett–Burman design was used for the screening of different carbon sources viz. dextrose, galactose, fructose, maltose, sucrose, lactose, and glycerol and nitrogen sources viz. yeast extract, tryptone, peptone, casamino acids, beef extract corn steep liquor, and polypeptone. A total of seven carbon and nitrogen sources (variables, $k=7$), were selected for the study with each variable represented at two levels, high concentration (+1) and low concentration (−1) were selected and tested for SK production.

The +1 and −1 concentration for carbon and nitrogen sources represents 1% and 2% *w*/*v* respectively. The design is orthogonal in nature and thus gives pure effect of each variable not confounded with interactions among variables. All these experiments were carried in Erlenmeyer flasks of 250 ml containing 100 ml of medium. Different carbon sources were screened for SK production at 1% yeast extract, 2% peptone, pH 7.0, and 200 rpm. Different nitrogen sources were screened for SK production at 2% dextrose, pH 7.0, and 200 rpm. The Plackett–Burman for the screening of carbon and nitrogen sources resulted in eight experiments. Tables 1 and 2 shows the design matrix for the screening of carbon and nitrogen sources. The number of positive signs and negative signs per trial are $(k+1)/2$ and $(k-1)/2$, respectively. Each column should contain equal number of positive and negative signs. Thus, each row represents a trial run and each column represents an independent variable. The effect of each variable was determined by following equation:

$$E(x_i) = 2\left(\sum M_i^+ + M_i^-\right)\Big/N$$

**Table 1** Two level Plackett-Burman design table for the screening of carbon sources.

| Variable codes<br>Carbon sources | C1<br>Fructose | C2<br>Sucrose | C3<br>Maltose | C4<br>Dextrose | C5<br>Glycerol | C6<br>Lactose | C7<br>Galactose |
|---|---|---|---|---|---|---|---|
| **1** | +1 | +1 | −1 | +1 | −1 | −1 | −1 |
| **2** | −1 | +1 | −1 | −1 | +1 | −1 | +1 |
| **3** | −1 | −1 | +1 | +1 | −1 | −1 | +1 |
| **4** | −1 | +1 | +1 | −1 | −1 | +1 | −1 |
| **5** | +1 | +1 | +1 | +1 | +1 | +1 | +1 |
| **6** | +1 | −1 | +1 | −1 | +1 | −1 | −1 |
| **7** | +1 | −1 | −1 | −1 | −1 | +1 | +1 |
| **8** | −1 | −1 | −1 | +1 | +1 | +1 | −1 |

+1 is for higher concentration and −1 for lower concentration

where $E(x_i)$ is the concentration effect of the tested variable. $M_i^+$ and $M_i^-$ from the trials where the variable ($x_i$) measured was present at high and low concentrations, respectively; and $N$ is the number of trials. STATISTICA 6.0 (Stat Soft, Inc, Tulsa, OK, USA) software was used for regression and graphical analysis of the data obtained.

## Experimental Design and Optimization of SK Production Medium by RSM [21–23]

The optimum levels of parameters of SK production medium were determined by means of RSM. The RSM consists of a group of empirical techniques devoted to the evaluation of relationships existing between a cluster of controlled experimental factors and measured responses according to one or more selected criteria. According to this design, the total number of treatment combinations was $2\text{^}k+2\times k+n0$ where $k$ is the number of independent variables and $n0$ is the number of repetitions of the experiments at the center point. Based on the best results of one at a time approach, four critical parameters of SK production medium were selected and further evaluated for their interactive behaviors by using a statistical approach. The levels of four variables viz. peptone, 2% ($x1$); dextrose, 2% ($x2$); pH, 6.9 ($x3$), and temperature, 30 °C ($x4$) and each of the variables was coded at five levels −2, −1, 0, 1, and 2 by using Eq. 1. For statistical calculations, the variables $Xi$ were coded as $xi$ according to the following transformation.

**Table 2** Two level Plackett-Burman design table for the screening of organic nitrogen sources.

| Variable codes<br>Organic nitrogen sources | N1<br>Tryptone | N2<br>Casamino acids | N3<br>Peptone | N4<br>Corn steep liquor | N5<br>Polypeptone | N6<br>Yeast extract | N7<br>Beef extract |
|---|---|---|---|---|---|---|---|
| **1** | +1 | +1 | −1 | +1 | −1 | −1 | −1 |
| **2** | −1 | +1 | −1 | −1 | +1 | −1 | +1 |
| **3** | −1 | −1 | +1 | +1 | −1 | −1 | +1 |
| **4** | −1 | +1 | +1 | −1 | −1 | +1 | −1 |
| **5** | +1 | +1 | +1 | +1 | +1 | +1 | +1 |
| **6** | +1 | −1 | +1 | −1 | +1 | −1 | −1 |
| **7** | +1 | −1 | −1 | −1 | −1 | +1 | +1 |
| **8** | −1 | −1 | −1 | +1 | +1 | +1 | −1 |

+1 is for higher concentration and −1 for lower concentration

The range and levels of the variables in coded units for RSM studies are given in Table 3.

$$x_i = X_i - X_0/\Delta X \tag{1}$$

where $xi$ is the dimensionless coded value of the variable $X_i$, $X_0$ the value of the $X_i$ at the center point, and $\Delta X$ the step change.

The behavior of the system was explained by the following quadratic model 2.

$$Y = \beta_0 + \sum \beta_i \times x_i + \sum \beta_{ii} \times x_i^2 + \sum \beta_{ij} \times x_{ij} \tag{2}$$

where $Y$ is the predicted response, $\beta_0$ the intercept term, $\beta_i$ the linear effect, $\beta_{ii}$ the squared effect, and $\beta ij$ the interaction effect. The full quadratic equation for four factors is given by model 3.

$$\begin{aligned} Y = {} & \beta_0 + \beta_1 x1 + \beta_2 x2 + \beta_4 x3 + \beta_4 x4 + \beta_{11} x1 \times x1 + \beta_{12} x1 \times x2 + \beta_{13} x1 \times x3 + \beta_{14} x1 \\ & \times x4 + \beta_{22} x2 \times x2 + \beta_{23} x2 \times x3 + \beta_{24} x2 \times x4 + \beta_{33} x3 \times x3 + \beta_{34} x3 \times x4 + \beta_{44} x4 \times x4 \end{aligned} \tag{3}$$

Several experimental designs have been considered for studying such models and central composite design (CCD) was selected. For this study, a $2^4$ full factorial central composite design with eight points and six replicates at the central points were employed to fit the second order polynomial model, which indicated that 30 experiments were required for this procedure. STATISTICA 6.0 (Stat Soft, Inc, Tulsa, OK, USA) software was used for regression and graphical analysis of the data obtained.

In order to search for the optimum combination of major critical parameters of the production medium for the production of SK, experiments were performed according to the CCD experimental plan (Table 6). The results of CCD experiments for studying the effect of four independent variables are presented along with the mean predicted and observed responses in Table 6. The regression equations obtained after the analysis of variance (ANOVA) gave the level of SK production as a function of the initial values of peptone, dextrose, pH, and temperature.

## Preparation of Yeast Cell Extracts

Constitutive intracellular SK expression in recombinant *P. pastoris* was investigated as follows. Briefly, Single colony was inoculated and cultured in 10-ml YPD medium for 48 h at 30 °C, 200 rpm. The culture was centrifuged at 2,000 ×*g* for 5 min and washed once with sterile distilled water and twice with lysis buffer (50 mM potassium phosphate pH 7.5, 2 mM EDTA, 1 mM PMSF).The cells were resuspended in 0.1 ml lysis buffer, and equal amount of acid washed glass beads (0.5 mm) were added and vortexed by ten 1 min bursts with 1–2 min cooling between bursts. Cell lysate was then centrifuged at 12,000 ×*g* for

**Table 3** Range and levels of the variables in coded units for rsm studies.

| Variables | −2 | −1 | 0 | 1 | 2 | ΔX[a] |
|---|---|---|---|---|---|---|
| Peptone, *x*1, % (*w*/*v*) | 1 | 1.5 | 2 | 2.5 | 3 | 0.5 |
| Dextrose, *x*2, % (*w*/*v*) | 1 | 1.5 | 2 | 2.5 | 3 | 0.5 |
| pH, *x*3 | 6.1 | 6.5 | 6.9 | 7.3 | 7.7 | 0.4 |
| Temperature *x*4, °C | 26 | 28 | 30 | 32 | 34 | 2 |

[a] ΔX step increment in the input variable values

15 min. The cell supernatant was analyzed for expression by chromogenic assay and clot lysis assay using purified rSK as a standard. All fractions were stored at −80 °C.

### Plasminogen Activation Assay [24]

ChromogenicS-2251 substrate was used in quantification of SK plasminogen activation.

Briefly human plasminogen was activated with SK at 37 °C for 15 min and the formation of plasmin mediated by streptokinase was measured by its hydrolysis of the chromogenic H-D-Val-Leu-Lys-pNA substrate S-2251. The reaction was performed in a total of 0.2 ml containing 0.1 M Tris-HCl (pH 7.4), 0.75 cu/ml plasminogen, and 0.6 mM S-2251. The reaction was monitored at 405 nm in a Multiskan Spectrum (Thermo) spectrophotometer.

### Fibrin Clot Lysis Assay [25]

A 1.0-ml mixture of 20 μl plasminogen (120 IU/ml) and 1.0 ml of fibrinogen (2 mg/ml) were added to 0.5 ml of thrombin (33 IU/ml). Subsequently 0.5 ml of streptokinase mixture was added and incubated at 37 °C. A visual turbid clot formed within 30 s is followed by formation of bubbles within the clot. Clot lysis time calibration is recorded with commercial streptokinase. Used as diluting buffer was 0.5 ml of 0.1 M Citro phosphate buffer, pH 7.4.

## Results and Discussion

This is the first report of a *GAP* promoter operable, constitutive intracellular expression and application of statistical methods for optimization of production medium of recombinant streptokinase in *P. pastoris*.

### Construction of Expression vector

Commercially synthesized *skc* gene was amplified by PCR as outlined above and was inserted downstream of GAP promoter in pGAPZA vector, generating recombinant pGAPZA-SK vector (Fig. 1a). pGAPZA-SK vector thus constructed was extracted, digested to completion with *EcoRI*, *NotI*, and *SapI* individually. Digestion of the recombinant vector yielded ~4.0 Kbp linearized product. Insertion of *skc* gene was also confirmed by PCR amplification using primers S1, S2 and S3, S4.

pGAPZA-SK vector constitutes pGAP-*skc*-Aox TT expression cassette and Zeocin selection cassette and is used for linearized integration in to host genome and for dominant selection of *Pichia* transformants. Zeocin is also used for selection of *E.coli* DH5α transformants.

### Generation of Constitutive Pichia GS115-SK Clones

pGAPZA-SK vector linearized with *SapI* was used in transformation of *P. pastoris* GS115. The recombinant vector has unique *SapI* site in GAP promoter region. Integration of the recombinant vector at GAP promoter region of *P. pastoris* generated transformants that were resistant to zeocin. *Pichia* transformants harboring the constitutive *skc* expression cassette were selected by plating on YPDZ plates and incubated for 3–4 days at 30 °C. The presence of *skc* insert in these 64 transformants was confirmed by colony PCR using S3, S4

primers. Expression of SK by these transformants was investigated by assaying the cell lysis supernatant for plasminogen activation efficiency.

## Growth and Expression Profile

Fibrin clot lysis assay and plasminogen activation assay were used for qualitative and quantitative estimation of rSK activity at different time intervals of growth. Constitutive intracellular streptokinase expression levels were checked for the *P. pastoris* GS115-SK clone at shake flask level to establish a suitable medium to accomplish maximum production applying statistical methods. *P. pastoris* GS115-SK clone was cultivated in media containing various kinds of sugars. Table 1 and Fig. 2 indicate the SK activity in the cell lysis supernatant. Dextrose was most effective for SK production, followed by galactose and maltose. The cell growth showed a constant value at over 2% dextrose, with maximum productivity at 3.5%. SK expression was repressed at higher glucose concentrations. Dextrose was therefore used as a carbon source for subsequent experiments. Next, the concentration of each medium component was investigated. The cell growth and the SK expression increased with increasing yeast extract and peptone concentrations, and the maximum production was observed at a concentration over 2.5% and 2%. These results emphasize the type and concentration of carbon source and concentration of yeast extract are essential requisites for large scale production of SK by *P. pastoris*. The results indicate that expression of rSK is observed from day 1 and is maintained till day 5 with plasmid stability (Fig. 1b). SDS-PAGE and western analysis indicated a 47 kDa protein with no glycosylation (data not shown).

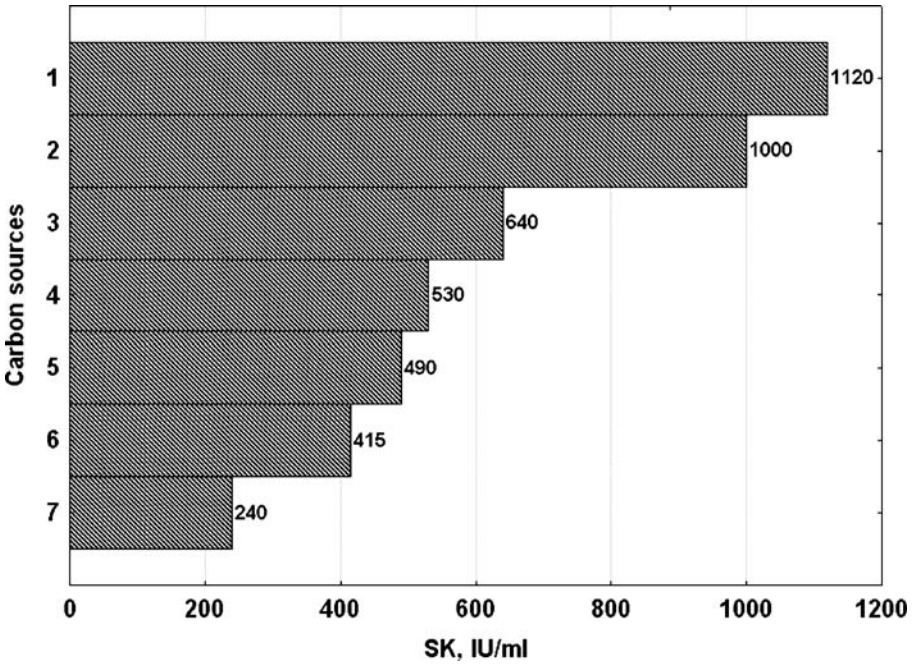


**Fig. 2** Pareto chart for the estimation of influence of different carbon sources on SK production (IU/ml) by *P. pastoris*. (*1*) dextrose (*2*) galactose (*3*) maltose (*4*) sucrose (*5*) lactose (*6*) fructose (*7*) glycerol

To investigate the toxic effects of expression of SK on host cells, cell growth of *P. pastoris* transformed with pGAPZA, pGAPZA-SK and untransformed strain were compared (Fig. 1c). Cell growth was exponential and similar in all the three strains indicating the expression of SK is not toxic to host.

*Screening of Carbon and Nitrogen Sources by Plackett–Burman Design* The effect of various carbon and nitrogen sources on SK production was tested by using Plackett–Burman statistical design. The Plackett–Burman design was used for the screening of different Carbon sources viz. dextrose, sucrose, fructose, maltose, galactose, lactose, and glycerol for SK production.

Figure 2 shows the Pareto chart that has been used to graphically summarize and display the relative importance of the differences between different carbon sources for SK production. From Fig. 2 it is evident that dextrose is the most significant carbon source with a maximum SK production of 1,120 IU/ml and glycerol is the least significant with reference to less SK production of 240 IU/ml. The difference in most significant to least significant is too high for SK production. Further it was decided to optimize the levels of dextrose in production medium by RSM.

Pareto chart in Fig. 3 shows the effect of different organic nitrogen sources on the SK production. It is clear from the figure that peptone had significant effect of SK production, which resulted in 1,080 IU/ml of SK production, followed by 1,025 and 1,015 IU/ml for yeast extract and polypeptone. Tables 4 and 5 shows the statistical analysis from the results of Plackett–Burman design for screening of carbon and nitrogen sources. The confidence

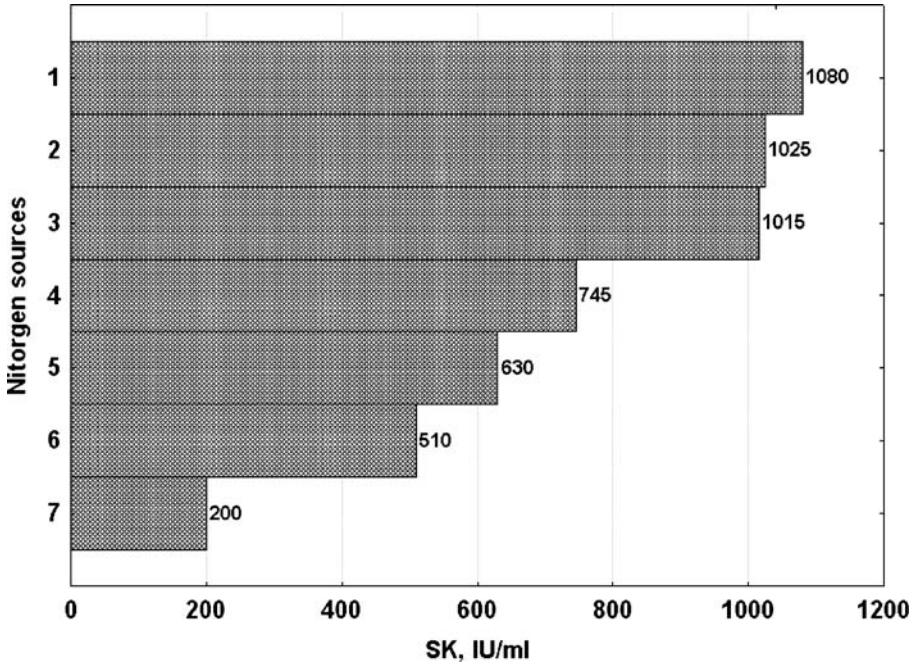


**Fig. 3** Pareto chart for the estimation of influence of different nitrogen sources on SK production (IU/ml) by *P. pastoris*. (*1*) peptone (*2*) yeast extract (*3*) polypeptone (*4*) beef extract (*5*) tryptone (*6*) casamino acids (*7*) corn steep liquor

**Table 4** Statistical analysis from the results of Plackett–Burman design for screening of carbon sources.

| Factors | Medium components | Effect | *T* | *P* value | Confidence level (%) |
|---|---|---|---|---|---|
| C1 | Fructose | −2.35 | 2.12 | 0.0982 | 62.10 |
| C2 | Sucrose | 1.89 | −5.65 | 0.1301 | 58.40 |
| C3 | Maltose | −3.65 | 4.39 | 1.0458 | 39.90 |
| C4 | Dextrose | 4.12 | −3.24 | **0.0089** | 99.20 |
| C5 | Glycerol | 2.84 | −1.25 | 2.3662 | 27.60 |
| C6 | Lactose | −1.12 | 2.41 | 0.3568 | 56.40 |
| C7 | Galactose | −1.94 | −4.36 | **0.0168** | 97.80 |

limits below 95% were considered as insignificant. The significant values for dextrose and galactose are 0.0089 and 0.0168 respectively. Similarly the significant values for peptone, yeast extract, and polypeptone are 0.0023, 0.0041, and 0.0056 respectively. From the results it is clear that the Plackett–Burman design is a powerful tool for identifying carbon and nitrogen sources, which had significant influence on SK production. The exact optimal values of the individual were determined by the subsequent central composite design and response surface methodology.

*Optimization of SK Production Medium Conditions by Design of Experiments and RSM* Table 6 shows the SK production in IU/ml under different experimental conditions designed by $2^4$ full factorial central composite designs with eight star points and six replicates at the central points. Experiment 17, 15, and 2 (Table 6) generated the maximum production of SK with 1,736, 1,592, and 1,564 IU/ml respectively among all the combinations tested. The RSM is an effective sequential and stepwise procedure and defines the effect of the independent variables, alone or in combination, on the process. The lead objective of the RSM was to run rapidly and efficiently along the path of improvement towards the general vicinity of the optimum. It is appropriate when the optimal region for running the process has been identified. The four critical independent variables, peptone, dextrose, pH, and temperature were chosen to optimize the production of SK by *P. pastoris*. Experiments were performed according to the CCD experimental design given in Table 3 in order to search for the optimum combination of components of the medium. The coefficient of determination ($R^2$) was calculated as 0.848 for SK production (model summary, Table 7), indicating that the statistical model can explain 84.8% of variability in the response. The $R^2$ value is always between 0 and 1. The closer the $R^2$ is to 1.0, the stronger the model and the better it predicts the response. In this case, the value of the determination coefficient ($R^2$= 0.848) indicates that the model does not explain only 15.2% of the total variations. The

**Table 5** Statistical analysis from the results of Plackett–Burman design for screening of nitrogen sources.

| Factors | Medium components | Effect | *t* | *p* | Confidence level (%) |
|---|---|---|---|---|---|
| N1 | Tryptone | 1.89 | 1.28 | 1.2562 | 61.90 |
| N2 | Casamino acids | −3.45 | 2.85 | 2.0125 | 56.40 |
| N3 | Peptone | −1.65 | −6.25 | **0.0023** | 99.80 |
| N4 | Corn steep liquor | 2.69 | 1.69 | 2.9842 | 24.30 |
| N5 | Polypeptone | 1.59 | 4.25 | **0.0056** | 98.10 |
| N6 | Yeast extract | −4.41 | −4.58 | **0.0041** | 98.90 |
| N7 | Beef extract | −1.61 | 4.23 | 0.0985 | 74.50 |

**Table 6** Design of experiments by central composite design (CCD) for RSM studies.

| Run number | *x*1 | *x*2 | *x*3 | *x*4 | Coefficients assessed by | SK production measured (IU/ml) | SK production predicted (IU/ml) |
|---|---|---|---|---|---|---|---|
| 1 | −1 | −1 | −1 | −1 | Fractional factorial $2^4$ design (16 expts) | 1,325 | 1,366 |
| 2 | 1 | −1 | −1 | −1 | | 1,564 | 1,558 |
| 3 | −1 | −1 | −1 | −1 | | 1,323 | 1,230 |
| 4 | 1 | 1 | −1 | −1 | | 1,377 | 1,574 |
| 5 | −1 | −1 | 1 | −1 | | 980 | 1,163 |
| 6 | 1 | −1 | 1 | −1 | | 1,432 | 1,215 |
| 7 | −1 | 1 | 1 | −1 | | 1,369 | 1,259 |
| 8 | 1 | 1 | 1 | −1 | | 1,468 | 1,463 |
| 9 | −1 | −1 | −1 | 1 | | 1,057 | 1,044 |
| 10 | 1 | −1 | −1 | 1 | | 1,123 | 1,095 |
| 11 | −1 | 1 | −1 | 1 | | 1,275 | 1,062 |
| 12 | 1 | 1 | −1 | 1 | | 1,174 | 1,266 |
| 13 | −1 | −1 | 1 | 1 | | 1,287 | 1,288 |
| 14 | 1 | −1 | 1 | 1 | | 1,460 | 1,199 |
| 15 | −1 | 1 | 1 | 1 | | 1,592 | 1,537 |
| 16 | 1 | 1 | 1 | 1 | | 1,736 | 1,601 |
| 17 | −2 | 0 | 0 | 0 | 8 star points | 1,275 | 1,134 |
| 18 | 2 | 0 | 0 | 0 | | 1,478 | 1,389 |
| 19 | 0 | −2 | 0 | 0 | | 1,169 | 1,177 |
| 20 | 0 | 2 | 0 | 0 | | 1,423 | 1,442 |
| 21 | 0 | 0 | −2 | 0 | | 1,466 | 1,336 |
| 22 | 0 | 0 | 2 | 0 | | 1,311 | 1,468 |
| 23 | 0 | 0 | 0 | −2 | | 1,403 | 1,348 |
| 24 | 0 | 0 | 0 | 2 | | 1,083 | 1,165 |
| 25 | 0 | 0 | 0 | 0 | 6 central points | 1,469 | 1,445 |
| 26 | 0 | 0 | 0 | 0 | | 1,381 | 1,445 |
| 27 | 0 | 0 | 0 | 0 | | 1,498 | 1,445 |
| 28 | 0 | 0 | 0 | 0 | | 1,441 | 1,445 |
| 29 | 0 | 0 | 0 | 0 | | 1,402 | 1,445 |
| 30 | 0 | 0 | 0 | 0 | | 1,478 | 1,445 |

adjusted $R^2$ value corrects the $R^2$ value for the sample size and for the number of terms in the model. The value of the adjusted determination coefficient (Adj $R^2$=0.706) is also high to advocate for a high significance of the model [21, 22]. If there are many terms in the model and the sample size is not very large, the adjusted $R^2$ may be noticeably smaller than the $R^2$. Here in this case the adjusted $R^2$ value is 0.706, which is lesser than the $R^2$ value of 0.848. At the same time, a relatively lower value of the coefficient of variation (CV=

**Table 7** Model summary and analysis of variance (ANOVA) for the quadratic model.

| Model estimates | Sum of squares | Degrees of freedom | Mean square | *F* value | Sig. |
|---|---|---|---|---|---|
| Regression | 689,871.800 | 14 | 49,276.557 | 5.975 | 0.001 |
| Residual | 123,701.167 | 15 | 8,246.744 | | |
| Total | 813,572.967 | 29 | | | |

***R***, 0.921; $\boldsymbol{R^2}$, 0.848; Adjusted $\boldsymbol{R^2}$, 0.706; **Standard error of the estimate, 90.818; CV**, 9.12%

9.12%) indicates a better precision and reliability of the experiments carried out. By applying multiple regression analysis on the experimental data, the experimental results of the CCD design were fitted with a second order full polynomial equation. The empirical relationship between SK production (Y) and the four test variables in coded units obtained by the application of RSM is given by equation 4.

$$\begin{aligned} Y = {} & 1444.833 + 63.833 \times x1 + 66.417 \times x2 + 33.167 \times x3 - 32.250 \times x4 \\ & -13.687 \times x1x1 - 45.875 \times x1x2 + 38.125 \times x1x3 - 35.125 \times x1x4 - 33.812 \\ & \times x2x2 + 57.875 \times x2x3 + 38.375 \times x2x4 - 10.687 \times x3x3 + 111.625 \times x3x4 \\ & -47.062 \times x4x4 \end{aligned} \quad (4)$$

where Y is SK production in IU/ml, is response and $x1$, $x2$, $x3$, and $x4$ are the coded values of the test variables, peptone, 2% ($x1$); dextrose, 2% ($x2$); pH, 6.9 ($x3$); and temperature, 30 °C ($x4$).

The ANOVA was conducted for the second order response surface model and the results are given in Tables 7 and 8. The significance of each coefficient was determined by Student's $t$-test and $p$ values, which are listed in Tables 7 and 8. The larger the magnitude of the $t$ value and smaller the $p$ value, the more significant is the corresponding coefficient. This implies that the linear effects of medium variables on SK production were, linear effects of dextrose ($p<0.004$) and peptone ($p<0.003$), quadratic effects of temperature ($p<0.016$) are significant. Similarly, interactive effects of peptone and pH ($p<0.022$), and pH and temperature ($p<0.016$) are significant. The regression model developed can be represented in the form of contour plots. The SK production for different levels of production medium conditions can also be predicted from the respective contour plots as shown in Fig. 4a–f. Each contour curve represents an infinite number of combinations of two test variables with the other two maintained at their respective 0 level. The contour plots described by the regression model were drawn to illustrate the effects of the independent variables, and combined effects of each independent variable upon the response variable.

**Table 8** Model coefficients estimated by multiple linear regressions (significance of regression coefficients).

| Variables | Beta | Standard error | $T$ | $P$ value (Sig. at <0.05) |
|---|---|---|---|---|
| (Constant) | 1,444.833 | 37.074 | 38.972 | 0.000 |
| $x1$ | 63.833 | 18.537 | 3.444 | **0.004** |
| $x2$ | 66.417 | 18.537 | 3.583 | **0.003** |
| $x3$ | 33.167 | 18.537 | 1.789 | 0.094 |
| $x4$ | −32.250 | 18.537 | −1.740 | 0.102 |
| $x1x1$ | −13.687 | 17.340 | −0.789 | 0.442 |
| $x1x2$ | −45.875 | 22.703 | −2.021 | 0.062 |
| $x1x3$ | 38.125 | 22.703 | 1.679 | 0.114 |
| $x1x4$ | −35.125 | 22.703 | −1.547 | 0.143 |
| $x2x2$ | −33.812 | 17.340 | −1.950 | 0.070 |
| $x2x3$ | 57.875 | 22.703 | 2.549 | **0.022** |
| $x2x4$ | 38.375 | 22.703 | 1.690 | 0.112 |
| $x3x3$ | −10.687 | 17.340 | −0.616 | 0.547 |
| $x3x4$ | 111.625 | 22.703 | 4.917 | **0.008** |
| $x4x4$ | −47.062 | 17.340 | −2.714 | **0.016** |

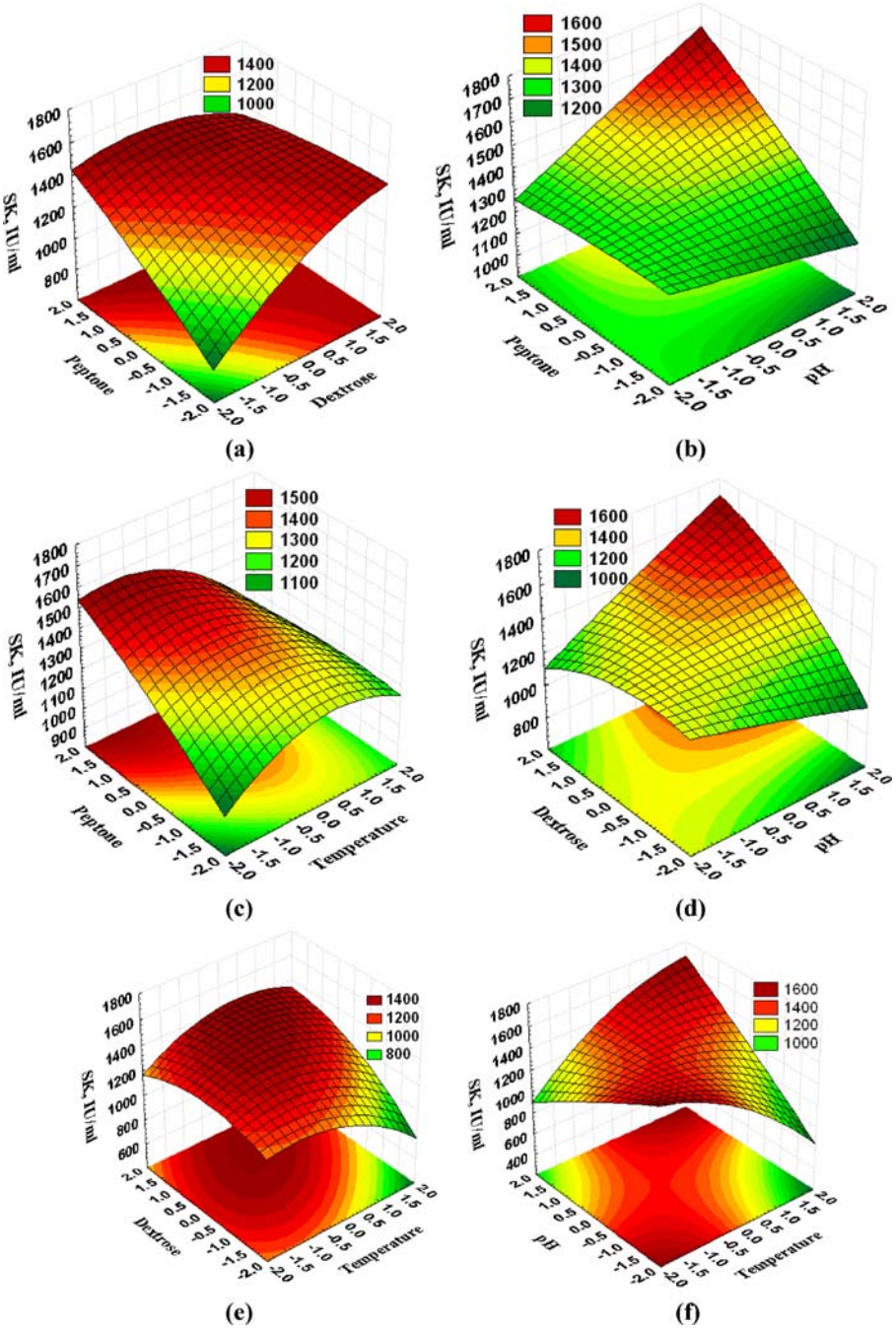


**Fig. 4** **a**–**f** 3-D surface and contour plot of SK production by *Pichia pastoris* (IU/ml): the effect of two variables while the other two are held at 0 levels

Figure 4a–c shows the trend of SK production in IU/ml with the variation dextrose with respect to peptone, pH, and temperature respectively. The maximum predicted SK production of 1,400, 1,600, and 1,500 IU/ml were observed from Fig. 4a–c respectively. Figure 4d–e shows the trend of SK production in IU/ml with the variation in peptone with respect to pH and temperature respectively. The maximum predicted SK production of 1,600 and 1,400 IU/ml was observed from Fig. 4d and e respectively. Similarly, Fig. 4f shows the maximum of 1,600 IU/ml SK production respectively as was predicted with the variation in pH and temperature.

A numerical method given by [22, 23] was used to solve the regression Eq. 4. The optimal values of the four test variables in coded units were observed to be $x1=-0.98$, $x2=0.562$, $x3=6.5$, and $x4=-0.785$. The predicted value of Y (SK production, IU/ml) at these values of $x$'s was 2,136.23 IU/ml. The real values of the four test variables were obtained by substituting the respective coded values in Eq. 1 and found to be dextrose ($x1$), 2.90%; peptone ($x2$), 2.49%; pH, 7.2 ($x3$), and temperature, 30.4 ($x4$).

The optimum conditions of selected production medium parameters were predicted using RSM and the maximum predicted SK production of 2,136.23 IU/ml could be achieved with the optimized production medium conditions. The validation of the optimized production medium conditions was done in a 250-ml EM flask containing 100 ml of production medium. The experiments were conducted in three similar flasks for the reproducibility of the SK production using optimized conditions. The SK production of 2,089 IU/ml (average of three experiments) obtained by using optimized conditions show that an increase of almost 95% was achieved than that obtained before optimizing the experimental conditions. The experimental value of the SK production was almost equal if we consider 95% of the confidence limits for the prediction of $Y$ value at optimized conditions with shake flask results. The level of SK activity achieved by constitutive intracellular expression in *P. pastoris* is 2,089 IU/ml as in compared with that of *S. pombe* (2,450 IU/ml) [9], *P. pastoris* (3,200 IU/ml) [8]. The optimization of nutrients at shake flask level are to be extrapolated to bioreactor level to obtain large scale production of rSK. Continuous constitutive bioreactor culture is being investigated to achieve rSK expression at industrial scale.

## Conclusion

Constitutive intracellular recombinant streptokinase production was achieved by glyceraldehyde 3-phosphate dehydrogenase promoter in *P. pastoris* expression system. Plasminogen activation of rSK protein was confirmed by chromogenic and fibrin clot lysis assay. Optimization of nutrients in YPD medium was performed to achieve high production of rSK by response surface methodology. The effect of various sugars on SK production was tested by using Plackett–Burman statistical design and found that dextrose is the effective carbon source among all the tested. The optimum conditions of selected production medium parameters were predicted using RSM and the maximum predicted SK production of 2,136.23 IU/ml could be achieved with the production medium conditions of dextrose ($x1$), 2.90%; peptone ($x2$), 2.49%; pH, 7.2 ($x3$), and temperature, 30.4 ($x4$). Validation studies showed a 95% increase in SK production as compared to initial levels. SK produced by constitutive expression was found to be functionally active by plasminogen activation assay and fibrin clot lysis assay. The current recombinant expression system and medium composition may enable maximum production of recombinant streptokinase at bioreactor level.

**Acknowledgements** This study was supported by a funding under the Technology Education Quality Improvement Program (TEQIP) by World Bank to Centre for Biotechnology, JNTU, Hyderabad. RNV acknowledges Council of Scientific and Industrial Research (CSIR) India for research fellowship.

## References

1. Tillett, W. S., & Garner, R. L. (1933). *The Journal of Experimental Medicine*, *58*, 485–502. doi:10.1084/jem.58.4.485.
2. Banerjee, A., Chisti, Y., & Banerjee, U. C. (2004). *Biotechnology Advances*, *22*, 287–307. doi:10.1016/j.biotechadv.2003.09.004.
3. Malke, H., & Ferretti, J. J. (1984). *PNAS USA*, *81*, 3557–3561. doi:10.1073/pnas.81.11.3557.
4. Malke, H., Gerlach, D., Kohler, W., & Ferretti, J. J. (1984). *Molecular & General Genetics*, *196*, 360–363. doi:10.1007/BF00328072.
5. Laplace, F., Muller, J., Gumpert, J., & Malke, H. (1989). *FEMS Microbiology Letters*, *53*, 89–94. doi:10.1111/j.1574-6968.1989.tb03602.x.
6. Hagenson, M. J., Holden, K. A., Parker, K. A., Wood, P. J., Cruze, J. A., et al. (1989). *Enzyme and Microbial Technology*, *11*, 650–656. doi:10.1016/0141-0229(89)90003-3.
7. Wong, S. L., Ruiqiong, Y., & Nathoo, S. (1994). *Applied and Environmental Microbiology*, *60*, 517–523.
8. Pratap, J., Rajamohan, G., & Dikshit, K. L. (2000). *Applied Microbiology and Biotechnology*, *53*, 469–475. doi:10.1007/s002530051643.
9. Kumar, R., & Singh, J. (2004). *Yeast (Chichester, England)*, *2*, 1343–1358. doi:10.1002/yea.1184.
10. Patrick, S. M., Fazenda, M. L., Mcneil, B., & Harvey, L. M. (2005). *Yeast (Chichester, England)*, *22*, 249–270. doi:10.1002/yea.1208.
11. Invitrogen Instruction Manual, A manual of methods for expression of recombinant proteins in *Pichia pastoris*. 2002, Catalog no. K1710–01.
12. Plackett, R. L., & Burman, J. P. (1946). *Biometrika*, *33*, 305–325. doi:10.1093/biomet/33.4.305.
13. Akhnazarova, S., & Kafarov, V. (1982). *Experiment optimization in chemistry and chemical engineering*. Moscow: Mir Publications.
14. Myers, R. H., & Montgomery, D. C. (1995). Response Surface Methodology. In Process and Product Optimization Using Designed Experiments. New York: Wiley-Interscience.
15. Khuri, A. I., & Cornell, J. A. (1987). *Response surfaces: design and analysis*. New York: Marcel Dekker Inc.
16. Potumarthi, R., Gopal, M., & Jetty, A. (2008). *Applied Biochemistry and Biotechnology*. doi:10.1007/s12010-008-8229-9.
17. Dutta, J. R., Dutta, P. K., & Banerjee, R. (2004). *Process Biochemistry*, *39*, 2193–2198. doi:10.1016/j.procbio.2003.11.009.
18. Ravichandra, P., Subhakar, C., Pavani, A., & Jetty, A. (2008). *Bioresource Technology*, *99*, 1776–1786. doi:10.1016/j.biortech.2007.03.041.
19. Radhika, T., Kiran Kumar, D., Ravichandra, P., & Lakshmi Narasu, M. (2007). *Applied Biochemistry and Biotechnology*, *141*, 187–201. doi:10.1007/BF02729061.
20. Himabindu, M., Ravichandra, P., Vishalakshi, K., & Jetty, A. (2006). *Applied Biochemistry and Biotechnology*, *134*, 143–154. doi:10.1385/ABAB:134:2:143.
21. Box, G. E. P., Hunter, W. G., & Hunter, J. S. (1978). *Statistics for experimenters*. New York: Wiley.
22. Cochran, W. G., & Cox, G. M. (Eds.) (1957). In Experimental design. New York: Wiley.
23. Montgomery, D. (2001). *Design and analysis of experiments*. New York: Wiley.
24. Sands, D., Whitton, C. M., & Longstaff, C. (2004). *Journal of Thrombosis and Haemostasis*, *2*(8), 1411–1415. doi:10.1111/j.1538-7836.2004.00814.x.
25. Couto, L. T., Donato, J. L., & de Nucci, G. (2004). *Brazilian Journal of Medical and Biological Research*, *37*(12), 1889–1894. doi:10.1590/S0100-879X2004001200015.