Optimization of Fermentation Medium for the Production of Recombinant Human Antithrombin III from *Saccharomyces cerevisiae* through Statistical Experimental Designs

Maheswara Reddy Mallu, Sandeep Vemula, Srinivasa Reddy Ronda*
KLE F University, Centre for Bioprocess Technology, Department of Biotechnology, Guntur-522 502, Andhra Pradesh, India.

**ABSTRACT**

This work, for the first time reports the optimization of recombinant human antithrombin (rhAT) production in *Saccharomyces cerevisiae* BY4741 using statistical experimental designs. Most applications of design of experiments (DoE) have concerned optimization of the composition of growth and production culture media. The typical objective is to identify a best selection and quantitative composition of significant medium supplements. In the present study, Plackett-Burman (PB) design followed by central composite design has been employed to evaluate and optimize the suitable culture medium for rhAT production. Influence of raffinose, glutamic acid and vitamin mixture were screened to be significant variables by PB design. The significant nutritional variables were further optimized using central composite design (CCD) for maximum production of rhAT. Central composite design (CCD) has been selected to explain the interaction effect of the three significant variables such as raffinose, glutamic acid, and vitamin mixture. The multiple regression equation ($R^2=0.9967$) was used to optimize significant impact of medium component values to maximize rhAT formation. The optimized values were found to be 23.09765 g/L, 8.01816 g/L and 77.2056 mg/L for raffinose, glutamic acid and vitamin mixture, respectively. The maximum yield of rhAT of 38.97 µg/mL was obtained experimentally using CCD and was very close to the predicted response of 38.93 µg/mL.

**INTRODUCTION**

Antithrombin is a glycolprotein having 432 amino acid residues, 4 glucosamine-based oligosaccharide side units, and 3 disulfide bridges (Franzen *et al.*, 1980 and Petersen *et al.*, 1979). It is produced in the liver and has a plasma level of approximately 150 mg/L (Murano *et al.*, 1980). Recombinant human antithrombin (rhAT) is the most significant inhibitor of thrombin family and activated factor X (FXa). Although, antithrombin acts mainly as the most significant plasma inhibitor of thrombin and antithrombin has also strong inhibitory effects on variety of active serine proteases, including IXa, Xa, Xii, XIIa, and FVIlia factors (Damus *et al.*, 1973 and Travis *et al.*, 1983). This is demonstrated by the fact that individuals with antithrombin deficiency are predisposed to thrombotic events and deficiency of antithrombin is a hereditary disorder that is connected with recurrent thrombophlebitis, aortic thrombosis and thromboembolism (Egeberg 1965). Besides being involved in regulating coagulation, antithrombin has been shown to have anti-inflammatory properties (Wiedermann *et al.*, 2002). rhAT is useful to cure diseases such as Transient Ischemic Attack, Thrombosis, Thrombocytosis and Cerebrovascular diseases are the main causes of death and disability around the globe (De Stefano *et al.*, 2008). Because of these properties, antithrombin is useful as a therapeutic agent to treat diseases and hence rich source of extremely purified and active rhAT might be supportive for clinical and model validation studies (Linhardt *et al.*, 2003). rhAT is achieving increasing attention due to their possible use in clinical and model validation studies.

The use of recombinant proteins at a commercial scale is constrained by the ability to gain the product in higher yields from microbial processes (Freigassner *et al.*, 2009).
The growing demand for therapeutic products calls for robust production hosts, capable expression systems and appropriate cultivation conditions. The limitation is often in terms of obtaining upmost quantities at sufficiently low cost to allow for marketing. Yeast is a single-celled, eukaryotic microbe that can grow speedily in complex or defined media (doubling time is typically 2.5 h in glucose-containing medium) and is simple and economical to use for r-DNA based protein production than insect and other mammalian cell lines (Bill 2001). The main advantage of yeast expression systems is the capacity to perform strict quality control and post-translational modifications (Demain et al., 2009).

To achieve high product yields, media composition is one of the important parameter. The optimal culture medium depends on many factors, including host metabolism, potential inhibitory products, target proteins, etc. and the development of the optimal medium is often a trial and error process (Shojaisadati et al., 2008). On one hand, it is desirable from the cost perspective to make the medium as simple as possible by reducing the amount of non essential carbon and nitrogen sources. It is prerequisite to evaluate and optimize positive impact of medium components in an efficient fermentation process.

Statistical experimental designs are useful tools in bioprocess development to screen out an effective real variable with significance impact (Weuster-Botz, 2000; Mandenius et al., 2008; Box et al., 2005). The application of statistical experimental designs in optimization can give better biomass yields, decreased process variability, closer confirmation of the response to nominal, target requirements, reduced time and overall costs (Brereton et al., 2003).

Usual practice of single factor optimization by maintaining other factors involved an undetermined steady point does not show the pooled effect of all the factors involved (Vaidya et al., 2003).

To make a full factorial search, which would study each possible combination of independent variable at suitable levels could require a big number of experiments and is a time consuming process, which is unreliable. Industrially the aim is to carry out the minimum number of experiments to analyze optimal conditions. These limitations of a single factor optimization process can be removed by PB design to examine more than five variables (Ahuja et al., 2004), the PB design is used to evaluate the most important variables for the better cell growth and product formation.

The next stage of medium optimization would be to find the optimum level of each effective independent variables which has been recognized by the Plackett-Burman design. This may be done by using most popular central composite design (CCD) under response surface methodology (RSM) technique. The CCD is used to optimize the relative significance of several affecting factors even in the presence of complex interactions and to estimate the coefficients of quadratic model (Rao et al., 2000 and Bezerra et al., 2008). In the present study, this communication reports an effort to evaluate and optimize highly effective medium components with significance impact on rhAT production by P-B design followed by RSM.

**MATERIALS AND METHODS**

**Strains and chemicals**

*Saccharomyces cerevisiae* BY4741 with pYES2/CT expression vector was used as the host for production of rhAT (ATCC 404002, USA). Yeast nitrogen base (YNB) without uracil, aspartic acid, Raffinose, Galactose, complete supplement mixture and vitamin mix were procured from Himedia, India. Surose, Starch, Fructose, Glutamic acid, Soy peptone, Peptone and ammonium nitrate were procured from Merck (CA, USA). SDS loading dye was purchased from Thermo Scientific, USA and coomassie brilliant blue (CBB-250) and Bovine serum albumin (BSA) were obtained from Bio-Rad Inc, USA.

**Culture conditions**

For inoculum development, a single colony of yeast was used to grow the overnight seed culture. For optimization experiments, the seed culture at 10% (v/v) was transferred to 100 mL YNB-URA media containing flasks and incubated for 24 h at 30 °C. The cultures were induced with galactose (20%) and raffinose (10%) induction medium.

**SDS-PAGE analysis of rhAT**

Equal cell density was maintained by changing the optical density to 1.0 with the growth media and the pellets were collected through centrifugation. To remove the media traces and membrane surface proteins, cells were resuspended in the sterile distilled water. The cell pellets were then collected through centrifugation at 30,000 g for 15 min at 4°C and stored at -20°C for further analysis.

The cells were treated with SDS loading dye (Thermo Scientific, USA) and lysed with frequent vortexing followed by heating at 100°C. The samples were directly loaded and resolved with the 12% acrylamide sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). The gels were stained with coomassie brilliant blue (CBB-250, Bio-Rad Inc, USA) solution for 8 h. The intensity and rhAT bands were analyzed with ImageJ v1.48 software against the standard protein (Bovine serum albumin, BSA) (Vemula et al., 2015).

**Experimental design and optimization**

**PB Design analysis for rhAT production**

PB Design was employed for screening the most significant medium components affecting rhAT production by *Saccharomyces cerevisiae*. In the present study, a 12 run PB Design have been applied to evaluate 11 variables in which two were dummy variables to obtain an estimate of error. Each independent variable was tested at two levels, low and high, which were denoted by (+) and (-) respectively (Table. 1). Table. 2 shows the details of the experimental design and yields of rhAT.
The relationships of nutritional variables were evaluated by fitting a second order polynomial equation to information of 20 experiments. The quadratic model was established as follows

\[ \hat{Y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \sum_{j=1}^{k} \beta_{ij} x_i x_j \]  

(2)

where \( \hat{Y} \) is the predicted response (rhAT production) \( \beta_0 \) is a constant, \( \beta_i \) linear terms coefficients, \( \beta_{ij} \) quadratic terms coefficient and \( \beta_{ij} \) interaction coefficients. The relationship between coded and uncoded form of the variables is given as follows.

\[ A_i = (Z_i - (Z_i^*)/\Delta Z_i \]  

(3)

where \( A_i \) is the coded value of the variable, \( Z_i \) is the real value of an independent variable (un coded), \((Z_i^*)\) is the center point value and \( \Delta Z_i \) is the step change between the levels.

**RESULTS AND DISCUSSION**

**Screening of significant nutritional variables for rhAT Production using PB design**

Statistical design for media optimization has proved to be a powerful and practical tool for bioprocess engineering. An attempt has been made to improve the composition of the medium formulation by simultaneous comparison between two levels of several factors through PB design. In the present work, a study has been performed to analyze the significant effect of various nutritional variables such as raffinose, sucrose, starch, fructose, glutamic acid, ammonium nitrate, peptone, soy peptone and vitamin mix. on rhAT yield. The effects of these components on rhAT production and maximum rhAT concentration (17.5 \( \mu \)g/mL) were shown in Table 2. Fig.1 shows pareto chart of the most important factors determining the rhAT production and response values of raffinose, glutamic acid and vitamin mixture,
which were above the “Bonferroni limit. From the results, it has been observed that the nutritional variables which show significant effect were considered to have greater effect on rhAT yield.

Table 5 shows Analysis of variance (ANOVA) outcome with significant F-value (30.5) and Prob > F value (less than 0.05). The goodness of the fit of the model was examined by the multiple correlation coefficient ($R^2 = 0.9937$). The adjusted multiple correlation value (adj $R^2 = 0.9656$) is in reasonable correlation with the predicted multiple correlation (pred $R^2 = 0.7750$). Adeq Precision investigates the signal to noise ratio, a ratio of 18.102 indicates an adequate signal. The coefficient of variance (CV = 2.87%), which is comparatively low value, indicates a better reliability and precision. A first order model was used to build predictions about the response to the given level of each factor which is defined by the following equation:

$$\text{rhAT} = 14.88 + 1.02 * A - 0.38 * B - 0.2 * C - 0.2 * D + 1.23 * F + 0.25 * G - 0.42 * J - 0.23 * K + 0.95 * L$$

(4)

where $A =$ raffinose, $B =$ sucrose, $C =$ starch, $D =$ fructose, $F =$ glutamic acid, $G =$ Soy peptone, $J =$ peptone, $K =$ Ammonium nitrate and $L =$ vitamin mixture. From the results it has been observed that, the rhAT yield was mainly due to three significant factors such as raffinose, glutamic acid and vitamin mixture, which were not reported in previous reports of media optimization. These three significant nutritional variables have shown the potentiality for the biomass and rhAT yield. Of the different carbon sources tested, raffinose showed the maximum positive effect when compared to other sources such as sucrose, starch and fructose (Fig.1). On the other hand, glutamic acid (F) showed considerable effect and while other nitrogen sources (peptone, soy peptone and ammonium nitrate) showed no significant effect on rhAT yield (Fig.1). The yield of rhAT has been controlled by the required concentration of nitrogen in the medium. The optimum nitrogen level can simulate rhAT production, whereas high nitrogen level represses it. The supplementation of the nutrient medium with an appropriate vitamin mixture maximized the yield of rhAT. In the present study, vitamin mixture (L) also showed a significant effect (Fig.1) on rhAT yield. The addition of vitamin mixture in the growth medium had an optimistic impact on rhAT production. From the PB design analysis, it can be seen that raffinose, glutamic acid and vitamin mixture were screened to have a significant effect on rhAT production and the same were used for further experiments.

**Estimation of the optimal Concentrations and Interactive Effects of the important factors on rhAT production using CCD**

RSM was adapted to find out a proper direction by increasing or decreasing the concentrations of fermentation variables according to the results of PB design. Table 4 shows full factorial central composite design and their observed values of rhAT production. A regression model having $R^2$ value more than 0.9 was considered to be having a very high correlation. Hence, the present $R^2$ value reflected a very good fit between the observed and predicted responses of rhAT, and implied that the present model is consistent for rhAT production. The $F$-test also shows a high significance for the regression model. Each of the observed value is evaluated with a predicted value (Table. 4) and residuals with the residual variance indicates that none of the individual residues exceed twice the square root of the residual variance. All of these considerations specify a better adequacy of the regression model. Table 6 shows the ANOVA of the regression model with respect to rhAT. The goodness of fit of the model was examined by several criteria. A $P$ value of less than 0.0001 indicates that the model terms are significant. The fitness of the model was tested by the coefficient of determination $R^2$, which was found to be 0.9967 representing that the simple variation of 99.67% was attributed to the variables and only less than 0.33% of the total variance could not be explained by the model.
The rhAT response model was analyzed as follows
\[ Y = 36.62 + 1.73 A_1 + 2.09 A_2 + 2.18 A_1 A_2 - 0.29 A_3 + 0.36 A_1 A_3 - 0.25 A_2 A_3 + 1.41 (A_1)^2 - 0.84 (A_2)^2 - 1.25 (A_3)^2 \]  
(5)

The above response model can be converted into the un-coded unit where
\[ A_1 = (Z_1 - 20) / 5 \]  
\[ A_2 = (Z_2 - 6) / 2 \]  
\[ A_3 = (Z_3 - 60) / 20 \]  
(6)

In order to get a better understanding of significant medium supplements effect on rhAT production, the predicted model was presented as response surface graphs. The interaction between the effective medium components can be seen in the response surface (3D) and counter (2D) plots (Figs 2a-c).

The interaction between nutritional factors indicates that the change in level of one factor affects the level of remaining factor for a fixed level of rhAT yield. Fig 2a-c shows the suitable level of raffinose, glutamic acid and vitamin mixture to get utmost production of rhAT.

### Table 7: Regression coefficients and their importance for response surface model.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Coefficient of estimate</th>
<th>Std. Error</th>
<th>F-Value</th>
<th>p-Value</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>36.62</td>
<td>1</td>
<td>0.12</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>A_rafrafinose</td>
<td>1.73</td>
<td>0.075</td>
<td>528.77</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>A_glutamicacid</td>
<td>2.09</td>
<td>0.075</td>
<td>773.04</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>A_vitaminmix</td>
<td>2.18</td>
<td>0.075</td>
<td>845.93</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>A_2</td>
<td>0.36</td>
<td>0.11</td>
<td>11.16</td>
<td>0.0075</td>
<td></td>
</tr>
<tr>
<td>A_3</td>
<td>-0.25</td>
<td>0.11</td>
<td>5.76</td>
<td>0.0373</td>
<td></td>
</tr>
<tr>
<td>(A_1)^2</td>
<td>-1.41</td>
<td>0.06</td>
<td>555.35</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>(A_2)^2</td>
<td>-0.84</td>
<td>0.06</td>
<td>195.77</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>(A_3)^2</td>
<td>-1.25</td>
<td>0.06</td>
<td>434.03</td>
<td>0.0001</td>
<td></td>
</tr>
</tbody>
</table>

A study has also been conducted to analyze the effect of the maximum and minimum principle of differential calculus on maximization of the equation (5) with respect to individual tested nutritional variables. The values of second order partial differential equations with respect to \( A_1, A_2 \) and \( A_3 \) obtained are -2.82*\( A_1 \), -1.68 and -2.5. The negative values of second order partial differential equations specify the absence of local maximum and applicability of rhAT maximization. First order partial differential model response equation 5 with respect \( A_1, A_2 \) and \( A_3 \) are equated to zero and \( A_1, A_2 \) and \( A_3 \) values were solved with the maximum rhAT response.

\[ 1.73 - 2.82 A_1 - 2.18 A_2 - 1.25 A_3 = 0 \]  
(9)

\[ 2.09 - 2.09 A_1 - 0.84 A_2 - 0.84 A_3 = 0 \]  
(10)

\[ 1.41 - 2.82 A_1 - 1.25 A_2 - 1.25 A_3 = 0 \]  
(11)

The numerical solution to the above equations (9-11) were \( A_1 = 0.61953, A_2 = 1.00908 \) and \( A_3 = 0.86028 \). Uncoded values of \( Z_1 = 23.09765 \) g/L, \( Z_2 = 8.01816 \) g/L and \( Z_3 = 77.2056 \) mg/L were estimated using equation 6-8. At these optimum concentrations, the maximum rhAT yield was estimated to be 38.97 µg/mL. An improvement (2.3 folds) of rhAT from 17.5 µg/mL (PBD) to 38.97 µg/mL was achieved after optimization using CCD.
CONCLUSIONS

The recombinant AT has gained enormous significance in the healthcare sector. Design expert analysis is the functional tool to optimize process variables for the production of recombinant proteins and provide valuable information on nutritional variable interactions.

In the present work, the effective variables on the rhAT production using PB design were screened. Three significant medium components screened from the PB design were raffinose, glutamic acid and vitamin mixture. The suitable concentrations of selected variables were optimized using CCD under response surface methodology.

The maximum predicted rhAT yield (38.97 µg/mL) from the CCD was achieved with raffinose (23.09765 g/L), Glutamic acid (8.01816 g/L) and vitamin mix (77.2056 mg/L). The proper use of media supplements can influence bioprocess economics at a large scale production of commercially value added therapeutic recombinant proteins.

ACKNOWLEDGMENTS

The authors are grateful to the Department of Biotechnology and management of K L E F University for providing the financial assistance and laboratory facilities during execution of the work.

Fig. 2: Three dimension (3D) and two dimension (2D) response surface plots showing the effects of variables on rhAT production by S. cerevisiae. A) Interaction of raffinose (g/L) and glutamic acid (g/L) (A₁A₂); (B) raffinose (g/L) and vitamin mix (mg/L) (A₁A₃); (C) glutamic acid and vitamin mix (A₂A₃).
REFERENCES


Bezerra M A; Santelli R E; Oliveira E P; Villar L S Escaleira L A. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. Talanta, 2008; 76 : 965-977.


De Stefano V; Za T; Rossi E; Vannucchi A M; Ruggeri M; Elli E; Micò C; Tieghi A; Caccia R R Santoro C. Recurrent thrombosis in patients with polycythemia vera and essential thrombocythemia: incidence, risk factors, and effect of treatments. Haematologica, 2008; 93: 372-380.


Friggassner M; Pichler H; Glieder A. Tuning microbial hosts for membrane protein production. Microb Cell Fact, 2009;8:69.


Mandenius C F; Brundin A. Bioprocess optimization using design-of-experiments methodology. Biotechnology Progress, 2008;24: 1191-1203.


Shojaosadati S A; Varedi Kolaei S M; Babaeipour V; Farnoud A M. Recent advances in high cell density cultivation for production of recombinant protein. Iranian J Biotechnol, 2008; 6: 63-84.


How to cite this article: