Jurnal Teknologi, 45(F) Dis. 2006: 19 – 30

OPTIMIZATION OF CRITICAL MEDIUM COMPONENTS FOR THE EXPRESSION OF RECOMBINANT HUMAN TRANSFERRIN IN INSECT CELLS BACULOVIRUS SYSTEM

CLARENCE M. ONGKUDON[1], BADARULHISAM ABDUL RAHMAN[2] & AZILA ABDUL AZIZ[3*]

Abstract. Human Transferrin (hTf) plays a big role in providing bacteriostatic functions as well as to transport iron from the storage part to all proliferating cells by receptor mediated endocytosis. Insect cells baculovirus expression system has been widely used as an alternative expression system for the production of recombinant human Transferrin (rhTf). This work focused mainly on the optimization of glutamine, glucose and lipid mixtures 1000x to increase rhTf yield. An experimental design involving 17 central composite design (CCD) experiments was employed and results were analyzed by Statistica (Statsoft v. 5.0). The response surface methodology (RSM) had identified the optimum values where glutamine=2211.20 mg/L, glucose=1291.95 mg/L, and lipid mixtures 1000x=0.64 %v/v. Using the optimized parameters, the studies demonstrated an increase in the rhTf yield by three-fold from 19.89 µg/ml to 65.12 µg/ml.

Keywords: Human transferrin; insect cells baculovirus; experimental design; central composite design; response surface methodology

Abstrak. Transferin manusia (hTf) memainkan peranan yang penting dalam fungsi bakteriostatik dan pengangkutan ferum dari bahagian penyimpanan ke sel-sel yang membiak melalui proses endositosis janaan reseptor. Sistem ekspresi bakulovirus sel serangga telah dipakai secara meluas sebagai sistem alternatif dalam penghasilan Transferin manusia rekombinan (rhTf). Kajian ini ditumpukan ke atas pengoptimuman glutamina, glukosa dan campuran lipid 1000x yang dapat meningkatkan penghasilan rhTf. Reka bentuk eksperimen yang melibatkan 17 eksperimen reka bentuk komposit berpusat (CCD) telah digunakan dan hasil kajian dianalisis oleh Statistika (Statsoft v. 5.0). Metodologi permukaan tindak balas (RSM) telah mengenalpasti nilai optimum parameter-parameter yang dikaji iaitu glutamina=2211.20 mg/L, glukosa=1291.95 mg/L, dan campuran lipid 1000x=0.64 %v/v. Hasil optimasi menunjukkan peningkatan hasil rhTf sebanyak tiga kali ganda, iaitu daripada 19.89 µg/ml kepada 65.12 µg/ml.

Kata kunci: Transferin manusia; bakulovirus sel serangga; reka bentuk eksperimen; reka bentuk komposit berpusat; metodologi permukaan tindak balas

[1 & 3]Department of Bioprocess Engineering, Faculty of Chemical and Natural Resources Engineering, Universiti Teknologi Malaysia, 81310, Skudai, Johor Darul Ta'zim, Malaysia
[2] Inno Biologics Sdn. Bhd., W 303, 3rd Floor, Block 2200, Enterprise Building 3, Persiaran Apec, 63000, Cyberjaya, Selangor Darul Ehsan, Malaysia
*Phone: +6075535525, Email: r-azila@utm.my

1.0 INTRODUCTION

Human transferrin is a glycoprotein that transports iron and provide bacteriostatic functions in a wide variety of physiological fluids in vertebrates [1, 2]. Recently, many efforts have been demonstrated in expressing recombinant human transferrin (rhTf) and mutants for fundamental, structural and functional studies [3]. The insect cells baculovirus expression system is an alternative method for the expression of recombinant mammalian proteins [4]. Besides the simplicity and cost-effectiveness of this method, insect cells own many of the protein processing and folding mechanisms of mammalian cells [5].

Research efforts have been focused mainly towards evaluating the effect of various carbon and nitrogen as nutrient substrates on the yield of recombinant proteins in the culture medium. The optimization of environmental parameters such as pH, temperature, aeration, and agitation was also conducted [6]. However, no defined medium has ever been established for the optimum production of rhTf in the insect cells baculovirus expression system. It is important to note that each organism or strain has its own requirements for optimum growth and protein production.

Optimization of medium by the conventional method involves changing one independent medium component while retaining all others at a fixed level. Experimental design and Response Surface Methodology (RSM) are useful tools for studying the effect of several factors influencing the cells growth and product optimization. In this research, critical medium components of the SF900-II medium were optimized for the production of high yield recombinant human transferrin in insect cells baculovirus system. The study was based on the preliminary screening of medium components, lipid mixture 1000x, glutamine and glucose, which were critical to rhTf production.

2.0 MATERIALS AND METHODS

2.1 Cell Lines, Recombinant Baculovirus and Medium

Sf9 cells were obtained from American Tissue Culture Collection (ATCC). Sf-900 II SFM culture medium was obtained from GIBCO^{TM}. Autographa Californica Multiple Nuclear Polyhydrosis Virus encoding for Human Transferrin gene (rhTf-AcMNPV) was a gift from Professor Michael Betenbaugh of Johns Hopkins University USA.

2.2 Sf9 Cells Cultivation in Spinner Flasks

The magnetic stirrer was maintained at 90 – 100 rpm, 27°C ± 0.5°C, nonhumidified, non-CO_2 equilibrated, in a temperature controlled biosafety cabinet. Sf9 cells were cultured at seeding densities of 5×10^5 viable cells/ml as a 30 ml serum free culture in a 100 ml spinner flask. The cultures were incubated until they reached 3×10^6 to 4×10^6 viable cells/ml. Sf9 cells were subcultured twice weekly. Once every three weeks, the cell suspension was gently centrifuged at 100 g for five minutes. The cell pellet was

resuspended in fresh medium to reduce the accumulation of cell debris and metabolic byproducts. The Sf9 cells number and viability were determined using haemocytometer and Trypan Blue solution.

2.3 Amplification and Titration of Baculovirus

The method used for quantifying virus concentration (pfu) was End Point Dilution Method which was described by Reed and Muench [7]. Virus stock was diluted by serial dilution in the range of $10^{-1} - 10^{-8}$. 10 µl of each virus dilution was mixed with 90 µl of 0.5×10^6 cells/ml in each well (one row per dilution) of the 96-well plate and incubated in a humidified environment for 4 – 7 days. The aim was to dilute the virus such that, when exposed to the Sf9 cells culture, 50% or less of the culture was infected. This represented the End Point Dilution or Tissue Culture Infectious Dose ($TCID_{50}$).

Baculovirus stock was amplified in a 75 cm^2 T-flask seeded with 10 ml of 5×10^5 viable cells/ml Sf9 cells and infected with recombinant virus stock at 0.1 – 1 Multiplicity of Infection (MOI). Infection was carried out at 27 + 0.5°C for 5 – 7 days until the cells were well infected. Viruses were harvested by centrifuging the infected medium at 250 g (1800 rpm) for five minutes and stored at 4°C.

2.4 RhTf Expression

For optimization and cost effective purpose, rhTf expression was done in a standard 24-well plate. Based on earlier studies and also studies by Bahia *et al.*, [8], suspension culture could be done in a 24-well plate. The plate was placed on a shaker and rotated at 125 – 130 rpm in a 27°C ± 0.5°C biosafety cabinet. For verification of optimum condition, rhTf expression was done as a 30 ml culture in a shaker flask using the same condition as the 24-well plate. Sf9 cells were seeded at 1.6×10^6 viable cells/ml and infected after two days by direct addition of 10 MOI of high titer virus inoculums into the culture. Samples were harvested after eight days post infection and analyzed using Enzyme Linked Immunosorbent Assay (ELISA).

2.5 Experimental Design and Optimization

Response surface methodology was used to optimize the levels of lipid mixtures 1000x, glucose and glutamine. A series of 17 central composite design (CCD) matrix experiments were conducted which incorporated eight 2-level factorial experiments, six extreme level experiments, two experiments at the center point and one control. Experiments were done in duplicates to obtain the error regions for rhTf concentration.

Table 1 gives the actual and the coded values of the variables tested. The coding of variables was done using the following equation:

$$x_j = 4\left(\frac{X_j - X_{cp}}{\Delta X_j}\right), \; j = 1, 2, 3, \ldots, n \qquad \ldots 1$$

Table 1 Central composite design for the optimization of glutamine, glucose and lipid mixtures 1000x

Test no.	Coded values			Real values mg/L	mg/L	%v/v	RhTf yield (mg/ml)		
	Gln	Gluc	Lip	Gln	Gluc	Lip	Actual	Predicted	Residual
1	–1	–1	–1	2500	2500	0.4	37.89	37.17	0.72
2	–1	–1	1	2500	2500	1.1	25.77	27.55	–1.78
3	–1	1	–1	2500	7500	0.4	32.65	30.20	2.45
4	–1	1	1	2500	7500	1.1	25.76	31.73	–5.97
5	1	–1	–1	7500	2500	0.4	62.28	58.06	4.22
6	1	–1	1	7500	2500	1.1	28.46	32.66	–4.20
7	1	1	–1	7500	7500	0.4	28.62	28.59	0.03
8	1	1	1	7500	7500	1.1	15.47	14.33	1.14
9	–1.7	0	0	796	5000	0.8	49.41	45.97	3.44
10	1.7	0	0	9204	5000	0.8	47.93	48.95	–1.02
11	0	–1.7	0	5000	796	0.8	56.38	55.03	1.35
12	0	1.7	0	5000	9204	0.8	34.60	33.53	1.07
13	0	0	–1.7	5000	5000	0.1	12.66	16.29	–3.63
14	0	0	1.7	5000	5000	1.4	2.05	–4.00	6.05
15	0	0	0	5000	5000	0.8	36.84	35.02	1.82
16	0	0	0	5000	5000	0.8	30.22	35.02	–4.79
17	–2	–2	–2	0	0	0	19.89	20.79	–0.90

where x_j is the coded level, X_j is the real value, X_{cp} is the real value of middle point, and ΔX_j is the range of values for each variable [6].

The responses obtained were subjected to multiple non-linear regression analysis to determine the coefficients. Coefficients with levels higher than 95% ($P < 0.05$) significance were included in the final models. The significance of the models was determined to check their efficiency. RhTf concentration can thus be expressed as a function of lipid mixtures 1000x, glucose and glutamine concentration by the second-order polynomial model:

$$Y = \beta_0 + \beta_j x_j + \beta_{jj} x_j x_j + \beta_{jk} x_j x_k \quad \ldots 2$$

where Y is the response in terms of rhTf concentration, β_0 is the intercept, β_j, β_{jj}, β_{jk} are linear, quadratic and interactive coefficients, respectively [9]. All analyses were done using the software Statistica (Statsoft, v. 5.0).

In general the model was considered to be efficient and workable if it had a significant F-value, an insignificant lack-of-fit F-value and a good R^2 (multiple correlation coefficient) [10, 11]. The optimum values for lipid mixtures 1000x, glucose and glutamine concentration were predicted using the steepest ascent method of RSM.

2.6 RhTf Assay

RhTf concentrations were determined using direct-sandwich type ELISA in a 96-well plate. The primary antibody was Goat-anti-human transferrin and the secondary antibody was Goat-anti-human transferrin-HRP conjugated. Incubation period for each step was 30 – 60 minutes. The enzyme substrate used was Tetramethylbenzydine (TMB). Enzyme reaction was stopped using 200 µl of 2M Sulphuric Acid (H_2SO_4) and absorbance was read at 450 nm.

3.0 RESULTS AND DISCUSSIONS

3.1 Regression Model

The results of the optimization experiments are shown in Figure 1 and Table 1. In control experiment (test no. 17), where there was no nutrient feeding, the rhTf concentration is 19.89 µg/ml. The maximum rhTf yield is in test no. 5 with 62.28 mg/ml. This indicates that the nutrients feeding had successfully increased the rhTf yield.

A multiple regression analysis was conducted on the experimental data to understand the relationship among the nutrients and the rhTf concentration. The results are given in Table 2. The parameters' coefficients were used to construct the second-

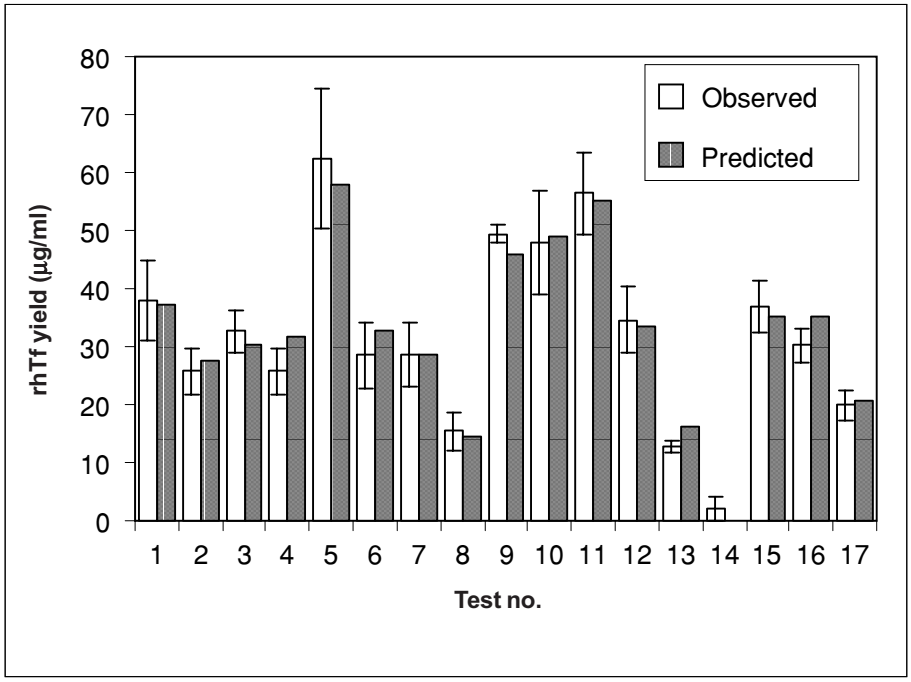

Figure 1 Observed and predicted experimental data

Table 2 Analysis of variance (ANOVA)

Regression statistics	
Multiple R	0.98
R Square	0.96
Adjusted R Square	0.90
Standard error	5.00
Observations	17

ANOVA

	Df	**SS**	**MS**	**F**	**Significance F**
Regression	9	3754.50	417.17	16.71	0.00
Residual	7	174.75	24.96		
Total	16	3929.24			

Variables	**Coeff**	**Std error**	**t Stat**	**P-value**
Intercept	35.02	2.87	12.22	0.00
Gln	0.87	1.33	0.66	0.53
Gln x Gln	4.31	1.39	3.09	0.02
Gluc	–6.32	1.33	–4.75	0.00
Gluc x Gluc	3.21	1.39	2.30	0.05
Lip	–5.97	1.33	–4.48	0.00
Lip x Lip	–9.99	1.39	–7.17	0.00
Gln x Gluc	–5.63	1.63	–3.45	0.01
Gln x Lip	–3.95	1.63	–2.42	0.05
Gluc x Lip	2.79	1.63	1.71	0.13

order polynomial model which explained the correlation of each nutrient and their second-order interactions with the rhTf production. The equation is given as:

$$Y = 35.02 + 0.87x_1 - 6.32x_2 - 5.97x_3 - 5.63x_1x_2 - 3.95x_1x_3 + 2.79x_2x_3 + 4.31_1{}^2 + 3.21x_2{}^2 - 9.99x_3{}^2 \quad \ldots 3$$

where Y is the rhTf response in µg/ml, x_1 is the coded value of glutamine, x_2 is the coded value of glucose and x_3 is the coded value of lipid mixtures 1000x. The quadratic model in Equation 3 with nine terms contains three linear terms, three quadratic terms and three, two-factor interactions. All terms are included in the model to give the optimum fit of the experimental data. This equation was used to predict the output of rhTf concentration with planned parameters and compared with observed values. The observed and predicted experimental values are given in Table 1.

Analysis of variance (ANOVA) was done using Statistica (Statsoft v. 5.0) and the results are given in Table 2. The fisher *F*-test value signifies how greater the mean square of the regressed model as compared to mean square of the residuals (errors). The *F* value for this case is 16.71. The greater the *F* value is, the more efficient the

model is. The significance of F value or sometimes referred to as P value is the probability to get larger F value by chance alone. A very low probability ($P_{model} > F = 0.00001$) demonstrates a very high significance for the regression model. This shows that F value is too significant to have arisen by chance alone. The lack of fit of the model was checked by the determination coefficient (R^2) which is the ratio of $SS_{regression}$ to SS_{total}. R^2 is a measure of the amount of reduction in the variability of y obtained by using the parameters' coefficients in the model. In this case, the value of the determination coefficient ($R^2 = 0.96$) indicates that only 6.00% of the total variations are not explained by Equation 3. The value of the adjusted determination coefficient (Adj. $R^2 = 0.90$) is also very high, which indicates a high significance of the model. The correlation coefficient ($R = 0.98$) shows a significant correlation between the independent variables and the rhTf response.

The significance of each coefficient was determined by student's t-test and P values, which are listed in Table 2. The larger the magnitude of the t-value and the smaller the P-value, the more significant the corresponding coefficient is. As a rule of thumb, coefficients with $P < 0.05$ are considered significant [12]. Almost all effects are significant except for first order effect of glutamine and two-way effect of glutamine and lipid mixtures 1000x. The quadratic effects of glutamine and glucose are both positive which indicate that there are minimum values for their concentrations. Meanwhile, the quadratic effect of lipid mixtures 1000x signifies that there is an optimum value for its concentration. The effects of linear, quadratic and two-way interaction can be arranged according to their ascending order of P value. Generally, the quadratic effect of lipid mixtures 1000x (x_3) is the most significant as is evident from the respective P-values ($Px_3^2 = 0.00001 > Px_1^2 = 0.0200 > Px_2^2 = 0.0500$) with the first order main effects ($Px_3 = 0.00001 > Px_2 = 0.0001 > Px_1 = 0.5300$) and two-way main effects ($Px_1x_2 = 0.0100 > Px_1x_3 = 0.0500 > Px_2x_3 = 0.13$). All of these values suggest that the concentration of glutamine, glucose and lipid mixtures 1000x have a direct correlation on the expression of rhTf. The magnitudes of the coefficients are evenly large which indicate that all of the coefficients have significantly contributed to the rhTf concentration. The comparison of the predicted and observed experimental data gives a standard deviation, Se = 3.3049, which signifies that none of the residuals exceed twice the magnitude of Se. Thus, all of the above considerations should be able to give excellent adequacy of the regression model.

3.2 Nutrient Interactions

To study the effect of nutrient interactions on rhTf expression, three surface plots involving two nutrients as X-axis and Y-axis while rhTf as Z-axis were constructed. The third nutrient was held at its center point. The results of the surface plots are shown in Figures 2, 3 and 4.

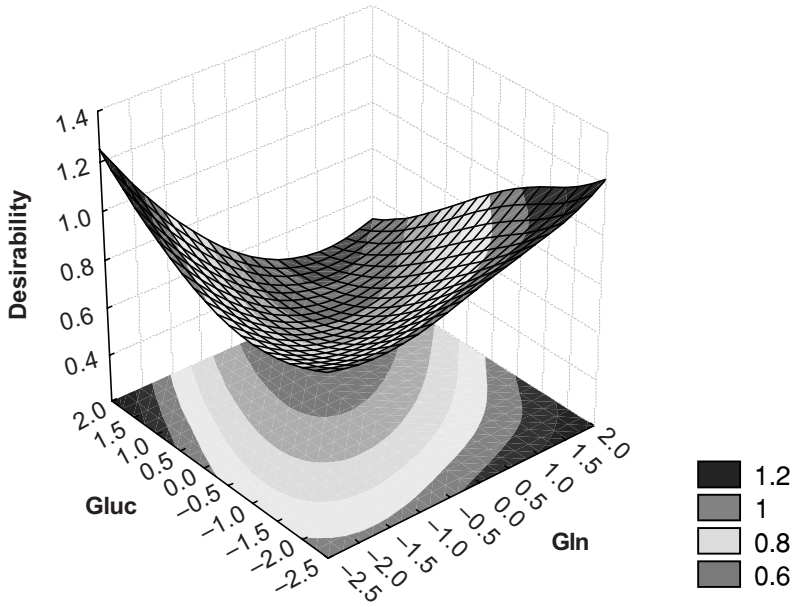

Figure 2 Glutamine (Gln) *vs* glucose (Gluc) *vs* rhTf

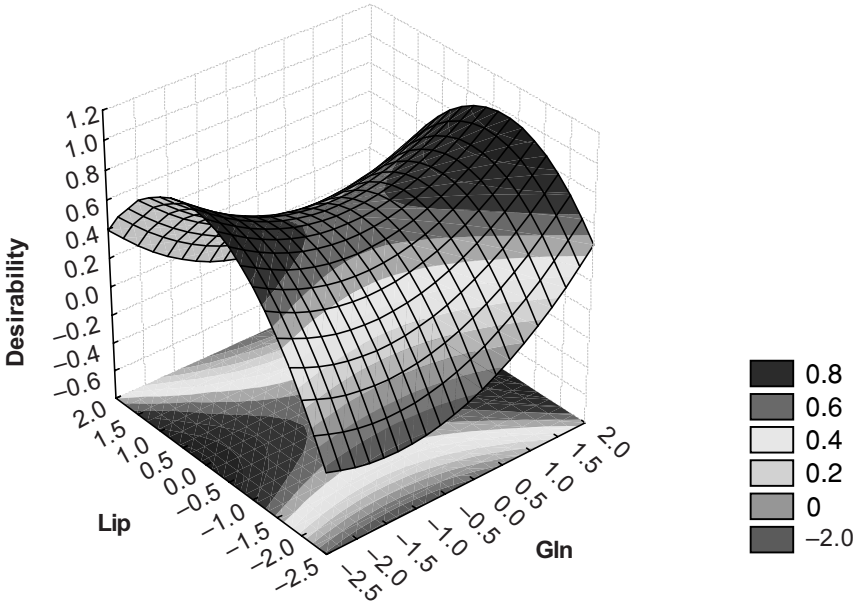

Figure 3 Glutamine (Gln) *vs* lipid mixtures 1000x (Lip) *vs* rhTf

Glutamine and glucose interactions are shown in Figure 2 using the regressed equation. It can be seen that at a lower glucose concentration (coded value = –2.5), an increase in glutamine concentration will result in increased rhTf yield. At a higher glucose concentration however, an increase in glutamine concentration will result in decreased rhTf yield. These interactions signify that rhTf yield is improved when using lower concentration of glucose and higher concentration of glutamine (glucose = –2.5, glutamine = 2.0) and *vice versa* (glucose = 2.0, glutamine = –2.5).

Glutamine and lipid mixtures 1000x interactions are shown in Figure 3. It seems that these two nutrients have less significant interactions compared to Figure 2. Each nutrient tends to follow its own patterns regardless of what concentration the other nutrient has. For example in Figure 3, an increase in lipid mixtures 1000x concentration will improve the rhTf yield until at a certain point where the rhTf yield will start to decrease. These patterns are observed in all region of the glutamine concentration. The quadratic effect of lipid mixtures 1000x is also more pronounce than the quadratic effect of glutamine. This gives an optimum value of lipid mixtures 1000x at around the middle value (coded value = 0). For glutamine, there are two rhTf peaks observed. The first peak is at lower concentration and the second peak is at higher concentration of glutamine. For cost effective purpose, the lower concentration of glutamine is more preferable.

Glucose and lipid mixtures 1000x interactions are shown in Figure 4. These nutrients also have insignificant interaction as evident by its *P*-value in the ANOVA. Each nutrient has the same patterns over the concentration range of the other nutrient. For

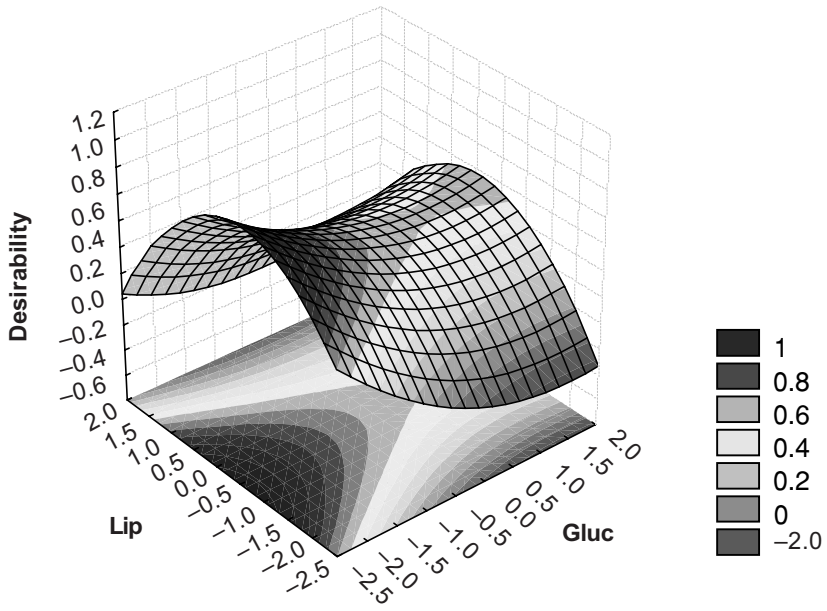

Figure 4 Glucose (Gluc) *vs* lipid mixtures 1000x (Lip) *vs* rhTf

example in Figure 4, the quadratic effect of lipid mixtures 1000x is very significant as compared to Figure 3. These patterns again, are observed in all region of the glucose concentration. This also gives an optimum value of lipid mixtures 1000x at around its middle value (coded value = 0). For glucose, one rhTf peak is observed in the region of its lower concentration.

Optimum values for glucose and glutamine have been observed in the lower concentration region. It is however observed in Figure 2 that the rhTf will improve when using glucose > glutamine or glutamine > glucose. Based on these considerations, the optimum values for glucose and glutamine are presumably in the lower concentration region where concentration of glutamine is higher than glucose concentration. In addition to that, the optimum value for lipid mixtures 1000x is in the center point region of its coded value.

Response Surface Methodology (RSM) based on the method of steepest ascent was carried out to hunt for actual optimum point of rhTf yield using the regression model. The optimum values of the test variables in coded values are $x_1 = -1.1155$, $x_2 = -1.4832$, and $x_3 = -0.2933$ with the corresponding response $Y = 47.33$. The real values of the test variables are glutamine = 2211.20 mg/L, glucose = 1291.95 mg/L, and lipid mixtures 1000x = 0.64 %v/v. The predicted rhTf yield using the optimized concentration of the nutrients is 47.33 µg/ml. The predicted value was verified by conducting an experiment using the optimized parameters. Result shows that rhTf yield was increased to 65.12 µg/ml. This result therefore verified the predicted value and the effectiveness as well as the usefulness of the model towards achieving the optimization.

4.0 CONCLUSIONS

The use of Central Composite Design (CCD) and Response Surface Methodology (RSM) had been demonstrated to be beneficial in optimizing an output of a biological process. The effect of the test variables could be studied simultaneously thus maximized the amount of information gathered for limited time and number of experiments. The regression model obtained in this work was highly effective and the nutrients had significant effects on the rhTf production. This work had successfully increased the rhTf yield by three-fold from 19.89 µg/ml to 65.12 µg/ml.

ACKNOWLEDGEMENTS

The authors would like to thank the Malaysian Ministry of Science, Technology and Innovation (MOSTI) for funding this project (vote no. 74192) and Prof. Michael Betenbaugh of Johns Hopkins University USA for providing the recombinant baculovirus.

REFERENCES

[1] Aisen, P., and I. Listowsky. 1980. Iron Transport and Storage Proteins. *Annu. Rev. Biochem*. 49: 357 – 393.

[2] Huebers, H. A., and C. A. Finch. 1987. The Physiology of Transferrin and Transferrin Receptors. *Physiol. Rev.* 67(2): 520 – 82.

[3] Ali, S. A., H. C. Joao, R. Csonga, F. Hammerschmind, and A. Steinkasserer. 1996. High Yield Production of Functionally Active Human Serum Transferrin Using a Baculovirus Expression System and Its Structural Characterization. *Biochem. J*. 319: 191 – 195.

[4] O'Reilly, D. R., L. K. Miller, and V. A. Luckow. 1994. *Baculovirus Expression Vectors: A Laboratory Manual*. U.S.A.: Oxford University Press.

[5] O'Reilly, D., L. Miller, and V. Luckow. 1992. *In: Baculovirus Expression Vectors*. New York: W. H. Freeman and Co. 109 – 215.

[6] Adinarayana, K., and P. Ellaiah. 2002. Response Surface Optimization of the Critical Medium Components for the Production of Alkaline Protease by a Newly Isolated *Bacillus sp. J. Pharm. Pharm*aceut. Sci. 5(3): 272 – 278.

[7] Reed, L., and H. Muench. 1938. A Simple Method for Estimating Fifty Percent Endpoints. *Am. J. Hyg*. 27: 493 – 497.

[8] Bahia, D., R. Cheung, M. Buchs, S. Geisse, and I. Hunt. 2004. Optimisation of Insect Cell Growth in Deep-Well Blocks: Development of a High Throughput Insect Cell Expression Screen. *Protein Expr. Purif*. 39: 61 – 70.

[9] Naveena, B. J., M. Altaf, K. Bhadrayya, S. S. Madhavendra, and G. Reddy. 2004. Direct Fermentation of Starch to L(+) Lactic Acid in SSF by Lactobacillus amylophilus GV6 Using Wheat Bran as Support and Substrate: Medium Optimization Using RSM. *Process Biochem*. 40(2): 681 – 690.

[10] Hasmann, F. A., D.V. Cortez, A. P. Junior, and I. C. Roberto. 2003. Optimization of Beta-Xyloxidase Recovery by Reversed Micelles Using Response Surface Methodology. *Electronic Journal of Biotechnology*. 6(2): 20 – 27.

[11] Mahat, M. K., R. M. Illias, R. A. Rahman, N. A. A. Rashid, N. A. N. Mahmood, O. Hassan, S. A. Aziz, and K. Kamaruddin. 2004. Production of Cyclodextrin Glucanotransferase (Cgtase) From Alkalophilic Bacillus Sp. TS1-1: Media Optimization Using Experimental Design. *Enzyme Microb. Technol*. 35: 467 – 473.

[12] Kalil, S. J., F. Maugeri, and M. I. Rodrigues. 1999. Response Surface Analysis and Simulation as a Tool for Bioprocess Design and Optimization. *Process Biochem*. 35(6): 539 – 550.