Journal of Data Science **4**(2006), 247-255

Bifactorial Design Applied to Recombinant Protein Expression

Martinez-Luaces, V., Guineo-Cobs, G., Velazquez, B.,
Chabalgoity, A. and Massaldi, H.
Universidad de la República

Abstract: We have studied the effect of several factors that influence recombinant protein production, by using the expression of recombinant streptolysin-O as our model. This protein, produced by *Streptococcus pyogenes*, is important in the biotechnological industry, where it is used to produce immunodiagnostic reagents. In order to improve the yield of this protein, we tried an alternative production method using strains of *Escherichia coli* and recombinant DNA technology. We have evaluated this method at the laboratory scale, taking into account factors such as inductor concentration, temperature of induction, proportion of culture medium volume to total flask volume, and strain of *Escherichia coli* used. To this end we applied techniques of experimental design, particularly a "fixed-effects bifactorial design", with the expression level of recombinant streptolysin-O in *E. coli* being the response to the factors. All the effects studied were found to be significant and relevant to the economics of the protein production.

Key words: Experimental design, protein expression, recombinant DNA.

1. Introduction

Native streptolysin-O (SLO) is an exotoxin produced by *Streptococcus pyogenes*, obtained from the supernatant after cultivation of this microorganism (Boreland, Thompson and Fenning, 1987). Purification of the protein from this medium is arduous and the yield is low (Altahus and Merle, 1993). The biotechnology industry uses SLO to manufacture immunodiagnostic reagents. It was therefore considered convenient to investigate alternatives to improve the production yield of SLO to achieve a more profitable production system.

An alternative approach is the production of recombinant streptolysin-O (rSLO) in *Escherichia coli*, appropriately transformed (Weller, Muller, Messner, Palmer, Valeva, Tranum-Jensen, Agrawal, Bjermann, Dobereiner, Kehoe and Bhakdi, 1996). Recombinant protein production usually makes use of an "induction" step in order to promote the desired protein expression during the cultivation in the heterologous host. This is carried out by a chemical agent that acts by inhibiting

the repressor gene of the host thereby allowing expression of the heterologous protein. We have chosen rSLO expression induced with IPTG (see below), as the response to evaluate the effect of several parameters during fermentation in a batch system.

A general methodology for investigating a process implies following a sequential design, that is, not to attempt to embrace all factors at all levels in a single study, but on the contrary, to carry out the experiments in stages. The first stage is to analyse the greatest possible number of factors or parameters that may influence the process, at only two levels. On the basis of first stage results, the most significant factors can be selected to carry out a more detailed analysis in a second stage of the experimental design. Also, if one wishes to further optimise the process on the basis of the selected factors, methodologies that use surface response curves can be applied in a subsequent stage. This type of sequential experimental design leads to obtaining results in a shorter time and at a lower cost.

Here we report the application of "fixed-effects, bifactorial experimental design", as it was used for the second stage, aimed at analysing the culture conditions of transformed *E. coli* strains for optimum production of rSLO.

2. Materials and Methods

2.1 The strategy

The experimental design was applied to analyse the effect of culture conditions during the induction phase of rSLO production. On the basis of first stage results (not described here), inductor concentration, temperature of induction and volume of culture medium as a proportion of total flask volume, at various levels, were the factors tested for different strains of *E. coli* with the rSLO production as the response variable. Since we wanted to evaluate up to 5 levels of certain factors, a complete 5^4 factorial design was not practical, nor was it possible to use randomized blocks (i.e., latin or greek-latin squares) as a means to reduce the number of experiments. Instead, we decided to use various replicated bifactorial experiments because they enable each of the factors to be studied at several levels (Montgomery, 1991), and provide information on possible interactions. In this way, we were also able to estimate the random error, in order to complete the ANOVA tests.

We used a "fixed-effects bifactorial design" (Montgomery, 1991). First, we studied the effect of different proportions (5%, 10%, 20%, 50% and 75%) of air/culture medium and two temperatures, 28 or 37°C, using induction with 1 mM IPTG and *E. coli* strain TG2. Then we tested the effect of inductor (IPTG) concentration at four levels: 0.0, 0.01, 0.1 and 1 mM, and induction temperatures

of 28 and 37°C in flasks with 50% culture medium, using *E. coli* strain TG2. The expression of rSLO was also determined with other strains of *E. coli* (XL, JM109, DH1 and RB791) and different proportions of culture medium (10, 20 and 50%), induced with 0.01 mM IPTG. Another bifactorial experiment was carried out using all strains of *E. coli* (TG2, XL, JM109, DH1 and RB791) grown in flasks with 50% volume of culture medium and induced with 0.01 mM IPTG at different temperatures (28°C or 37°C).

2.2 Methods for analysing and monitoring rSLO production

The proteins expressed by transformed *E. coli* were analysed in bacterial lysates by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (Towbin, Staehelin and Gordon, 1979). Different strains of *E. coli* were grown in Luria Bertani (LB) culture medium with ampicillin in 100 mL flasks. At appropriate optical density (OD) measured at 600 nm, cultures were induced with isopropyl-beta-D-thiogalactoside (IPTG). Cells were harvested by centrifugation, and the pellets were resuspended in phosphate-buffered saline (PBS) and then sonicated. Samples were then incubated with 1% Triton X-100 and centrifuged. The supernatants obtained were analysed by SDS-PAGE. rSLO was purified from the supernatant by affinity chromatography on glutathione-sepharose (Smith, 1993) and total protein content was measured by the bicinchoninic acid assay (Smith, Krohn, Hermanson, Mallia, Gartner, Provenzano, Fujimoto, Goeke, Olson and Klenk, 1985).

3. Results

In the following, experimental data are presented for the various pairs of variables tested. A summary of the corresponding ANOVA table, which shows the results of comparing the sum of squares ratio for each factor (experimental Fisher value) with the critical values of the Fisher distribution at the significance levels of 0.95 and 0.99 is also shown subsequently.

3.1 Expression of rSLO in cultures of E. coli strain TG2 in flasks with different volume proportion of culture medium, induced with 1 mM IPTG at different temperatures.

Table 1 shows the experimental data and ANOVA Table for different levels of the factors: *induction temperature* (factor A) and *percentage of culture volume* (factor B). In accordance with these results, the null hypothesis that induction temperature and the percentage of culture medium make no difference was rejected, as was the assumption that there was no interaction between the two

factors. It is important to note that the mean squares quotient for interaction between the factors is statistically significant at both critical values of 0.95 and 0.99, but is nevertheless 10 times lower than each of the mean squares ratio of the independent factors.

Table 1: Production of rSLO for factors induction temperature (factor A) and percentage of culture medium (factor B), with 1 mM IPTG and strain TG2

1.1 Experimental data

Treatments	Duplicates (mg/mL of rSLO)	
28°C, 5%	0.000	0.000
28°C, 10%	0.000	0.000
28°C, 20%	0.011	0.012
28°C, 50%	0.211	0.245
28°C, 75%	0.202	0.209
37°C, 5%	0.028	0.032
37°C, 10%	0.056	0.063
37°C, 20%	0.179	0.187
37°C, 50%	0.514	0.567
37°C, 75%	0.440	0.403

1.2 Summary ANOVA

Null hypothesis	d.f.	Expiremental F	$F_{0.95}$	$F_{0.99}$
A homogenous	1	452.46	4.96	10.04
B homogenous	4	418.95	3.48	5.99
No AB interaction	4	48.53	3.48	5.99

3.2 Expression of rSLO in flasks with 50% culture medium, using *E. Coli* strain TG2, at different induction temperatures and inductor concentrations

The effect of induction temperature (factor A) was analysed at 28 and 37°C and IPTG concentration (factor B) at 0, 0.01, 01, and 1 mM. Table 2 shows the results of rSLO expression obtained for these factors. Statistical analysis indicates that induction temperature, inductor concentration and the interaction between the two, were all highly significant even when compared to critical values for $F_{0.99}$, and so the null hypothesis was rejected in every case. Note that the most highly significant factor was induction temperature, followed by inductor concentration and, thirdly, the interaction between them (Table 2.2).

Table 2: Production of rSLO for induction temperature (Factor A) and inductor concentration (Factor B), with 50% culture medium and *E. Coli* strain TG2

2.1 Experimental data Treatments

Treatments	Duplicates (mg/mL of rSLO)	
28°C 0 mM	0.044	0.029
28°C 0.01 mM	0.177	0.169
28°C 0.1 mM	0.154	0.163
28°C 1 mM	0.150	0.201
37°C 0 mM	0.000	0.000
37°C 0.01 mM	0.740	0.700
37°C 0.1 mM	0.560	0.570
37°C 1 mM	0.660	0.620

2.2 Summary ANOVA

Null hypothesis	d.f.	Expiremental F	$F_{0.95}$	$F_{0.99}$
A homogenous	1	1217.38	5.32	11.26
B homogenous	3	394.54	4.07	7.59
No AB interaction	3	173.84	4.07	7.59

3.3 Expression of rSLO in different strains of *E. coli* and different volumes of culture medium, induced with 0.01 mM IPTG at 37°C

Table 3 shows the experimental data and ANOVA summary obtained for factors strain (factor A) and *percentage of culture medium* (factor B). The statistical analysis indicates that the null hypothesis that *strain* and *percentage of culture medium* have no effect on rSLO expression, should be rejected. Interaction between these factors was also significant at the critical values of F analysed in Table 2. Similar results were obtained when the induction was carried out at 28°C (results not shown).

3.4 Expression of rSLO in different strains of *E. coli* induced at different temperatures in a fixed 50% volume of culture medium and 0.01 mM IPTG concentration

Table 4 shows the experimental data obtained from the factors *induction temperature* (factor A) and *strains* (factor B). Table 4.2 shows the corresponding ANOVA summary. Both factors had a significant effect on rSLO production; therefore the null hypothesis that these factors made no difference was rejected. An important point to be noted is that the interaction between both factors was even more highly significant than their independent effects.

Table 3: Production of rSLO for strains of E. Coli (Factor A) and percentage of culture medium (Factor B), induced with 0.01 mM IPTG at 37°C.

3.1 Experimental data Treatments Duplicates (mg/mL of rSLO)

Treatments	Duplicates (mg/mL of rSLO)	
XL, 10%	0.000	0.000
XL, 20%	0.000	0.000
XL, 50%	0.000	0.000
JM109, 10 %	0.000	0.000
JM109, 20%	0.333	0.349
JM109, 50%	0.705	0.609
DH1, 10%	0.000	0.000
DH1, 20%	0.220	0.320
DH1, 50%	0.000	0.000
RB791, 10%	0.000	0.000
RB791, 20%	0.301	0.385
RB791, 50%	0.320	0.298

3.2 Summary ANOVA

Null hypothesis	d.f.	Expiremental F	$F_{0.95}$	$F_{0.99}$
A homogenous	3	113.02	3.49	5.95
B homogenous	2	136.49	3.89	6.93
No AB interaction	6	54.00	3.00	4.82

4. Discussion

The process of developing a bacterial culture (bacterial load or biomass content) involves the propagation of a bacterial population from an initial stock to a final culture. The final culture is typically used for the production of biomass or derived products. The aim of cultivation, therefore, is to minimise loss of viability during recovery from the latent phase of the stock culture, to grow genetically identical copies of the original population, to increase biomass and to grow the culture up to an appropriate physiological condition.

In the laboratory, 100 mL flasks are a useful tool for studying the effect of different factors that influence the development of a culture before scaling up to a pilot level, as their size permits ease of handling and sterilisation (Drew, 1981) with minimal cost. The design of experiments used here allowed the analysis of different factors that affect rSLO production. Within a general analytical methodology, this study represents the second stage in the search for the major factors that influence this process. In our study, the culture medium volume (as

Table 4: Production of rSLO for *induction temperature* (Factor A) and *strains* of *E. coli* (Factor B) in flasks with 50% volume proportion of culture medium, induced with 0.01 mM IPTG.

4.1 Experimental data

Treatments	Duplicates (mg/mL of rSLO)	
28°C, TG2	0.175	0.180
28°C, XL	0.000	0.000
28°C, JM109	0.168	0.153
28°C, DH1	0.840	0.900
28°C, RB791	0.169	0.150
37°C, TG2	0.698	0.710
37°C, XL	0.000	0.000
37°C, JM109	0.705	0.609
37°C DH1	0.000	0.000
37°C, RB791	0.320	0.298

4.2 Summary ANOVA

Null hypothesis	d.f.	Expiremental F	$F_{0.95}$	$F_{0.99}$
A homogenous	1	26.04	4.96	10.04
B homogenous	4	205.07	3.48	5.99
No AB interaction	4	457.01	3.48	5.99

percentage of total flask volume) was found to be significant, which is consistent with oxygen being an important nutrient which is limiting for bacterial growth (Banks, 1980) and production of rSLO. Our results also showed that the lower the transfer of dissolved oxygen (higher culture medium volumes relative to total flask volume), the higher the expression of rSLO observed (data not shown).

Induction temperature is an important factor in metabolic processes affecting the velocity of protein transcription-translation, which explains its statistical significance. The *percentage of culture medium* and induction temperature are also important factors on the production of rSLO in adequate three-dimensional form, which would explain why the interaction between them is significant. With respect to the influence of inductor concentration on expression of rSLO, in strain TG2 we observed that production of rSLO was increased at lower concentrations of IPTG, whether induction was carried out at 28°C or at 37°C. This finding is of economic interest, given the typically high cost of IPTG.

The statistical analysis of expression of rSLO in different strains of *E. coli* showed that type of strain was not indifferent (null hypothesis rejected), as was also shown for induction temperature and volume proportion of culture medium in the flask. Our experiments showed that the highest yield of rSLO was produced

by strains TG2 and JM109, both at 37°C and 28°C, compared with that of the other strains analysed.

In conclusion, our results showed that expression of rSLO is influenced by the type of strain of *E. coli* used, the induction temperature, the inductor concentration and the percentage of medium culture (oxygen transfer capacity of the system), for which we had previous evidence of their importance. In these experiments we found the following conditions as the most convenient to be explored in a third stage: strains TG2 and JM109, 0.01 mM IPTG, 37°C and 50% volume proportion of culture medium.

Acknowledgement

H. A. Massaldi is a member of CONICET, Argentina.

References

Altahus, H. and Merle, P. (1993). Process for the purification of streptolysin O, intact streptolysin O obtainable by this process. United States Patent. US005258504A. Patent number, 5.258.504.

Banks, G. T. (1980). Scale-up of fermentation process. *Topics Enzyme Ferment. Biotechnol* **3**, 170-266.

Boreland, P. C., Thompson, E. A. and Fenning, G. (1987). Evaluation of a latex agglutination screening test for the determination of anti-streptolysin O (ASO) antibodies. Serodiagn. *Immnother* **1**,113-116.

Calam, C. T. and Russell, D. W. (1973). Microbial aspects of fermentation process development. *J. Appl. Chem. Biotechnol* **23**, 225-237.

Drew, S. W. (1981). Liquid culture. In it (ed.), Manual of methods in general bacteriology (Edited by P. Gerhardt), 151-178. American Society for Microbiology.

Montgomery, D. C. (1991). *Diseno y Analisis de Experimentos.* Grupo Editorial Iberoamerica, Mexico.

Smith P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. and Klenk, D. C. (1985). Measurement of protein using bicinchoninic acid. *Anal. Biochem* **150**, 76-85.

Smith, D. B. (1993). Purification of glutathione-S-transferase fusion proteins. *Methods in Molecular and Cellular Biology* **4**, 220-229.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of protein from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350-4354.

Weller, U., Muller, L., Messner, M, Palmer, M., Valeva, A., Tranum-Jensen, J., Agrawal, P., Bjermann, C., Dobereiner, A., Kehoe, M. and Bhakdi, S. (1996). Expression of active streptolysin O in Escherichia coli as a maltose-binding-protein-streptolysin-O fusion protein. The N-terminal 70 amino acids are not required for haemolytic activity. *Eur. J. Biochem.* **236**, 34-39.

Received July 4, 2004; accepted May 11, 2005.

V. Martinez-Luaces
Department of Statistics
School of Engineering
Universidad de la Republica,
Montevideo, Uruguay
victorml@fing.edu.uy

G. Guineo-Cobs
Department of Statistics
School of Engineering
Universidad de la Republica
Montevideo, Uruguay
gegctrini@mixmail.com

B. Velazquez
Department of Biotechnology Development
School of Medicine
Universidad de la República
Montevideo, Uruguay.
bvela@higiene.edu.uy

A. Chabalgoity
Department of Biotechnology Development
School of Medicine
Universidad de la República
Montevideo, Uruguay
jachabal@higiene.edu.uy

Hugo Massaldi (corresponding author)
Departamento de Desarrollo Biotecnol ogico
Instituto de Higiene, Montevideo, Uruguay
Also at
Instituto Malbrán, Buenos Aires, Argentina
hmassaldi@hotmail.com