JOURNAL OF PURE AND APPLIED MICROBIOLOGY, May 2014. Vol. **8**(Spl. Edn. 1), p. 741-750

Screening of Important Autoinduction Medium Composition for High Biomass Production of *E. coli* Expressing Recombinant Bromelain

Mohd Jamil Aizat Jamaluddin[1], Azura Amid[1,2]*, Azlin Suhaida Azmi[1,2] and Muhd. EzzaFaiez Othman[1]

[1]Department of Biotechnology Engineering, Faculty of Engineering, International Islamic University Malaysia, Jalan Gombak 53100 Kuala Lumpur, Malaysia.
[2]Bioprocess and Biomolecular Engineering Research Unit (BPMERU), Department of Biotechnology Engineering, Faculty of Engineering, International Islamic University Malaysia, Jalan Gombak 53100 Kuala Lumpur, Malaysia.

(Received: 08 January 2014; accepted: 24 March 2014)

Bromelain, a naturally available therapeutic protease enzyme from pineapple stem, was expressed in *Escherichia coli* using autoinduction-based fermentation medium. There are numerous recombinant proteins functionally expressed in *E. coli* using this autoinduction approach. Preliminary media trials have shown that this medium is capable of producing a high cell density batch culture as claimed. Since the level of soluble expression is both protein and host/strain specific, further screening to identify significant media components affecting biomass production is therefore required. Currently, screening of full autoinduction medium components has yet to be examined elsewhere in the literature due to its tedious nature. Hence, the application of a fractional factorial design for identifying significant components in the autoinduction formulation was reported. Statistical analysis showed that glucose, glycerol and L-arabinose were most significant components influencing the biomass production. However, their effects were demoting rather than promoting the biomass production at elevated concentrations in shake flask culture. The highest biomass production (7.3 g/l) was achieved at low levels (-1). This represents 0.05 % (w/v) glucose, 0.5 % (w/v) glycerol, without the need of additional L-arabinose inducer. For a low number of experimental runs, this statistical approach has been proven efficient for screening vital medium components in comparison to conventional method.

Key words: Cysteine protease, fractional factorial design, Studier formulation, T7 expression system.

In general, optimization of various variables affecting growth during fermentation is necessary for efficient production of microbial biomass expressing a recombinant protein. The goal is to achieve a high growth rate of the protein host and production of the target protein while maintaining a low overall production cost[1-8]. Also, the growth of the host strain largely depends on the composition and nutrients in the growth medium.

* To whom all correspondence should be addressed.
Tel.: +603-61964000;
E-mail: azuraamid@iium.edu.my

Reports from Amid, Ismail *et al*.[9], Bala *et al*.[10] and Muntari *et al*.[6], on recombinant bromelain, only discussed molecular cloning and optimization of the shake flask culture process conditions. Under the reported experimental conditions, their maximum small scale production of purified recombinant Bromelain in shake flask using different enzyme assay were 1.2 U/mg-protein, 228.6 U/mg-protein and 9.6 U/mg-protein, respectively. However, none reported on productivity of biomass. Hence, this research study is intended to optimize cultivation medium compositions aiming at improving small scale

volumetric production of recombinant bromelain on the basis of its host biomass productivity.

Fermentation medium design is intrinsically linked to the host strain being used. Previous production of recombinant bromelain incorporated the *E. coli* BL21-AI strain with pDEST17/Bromelain vector[9]. The strain is a derivative of BL21 which supplies T7 RNA polymerase by transcription from the arabinose-inducible pBAD promoter in the chromosome. Determining the best combination of medium components and concentrations suited for this particular expression system would be both difficult and time-consuming. However, recent promising advancement in fermentation media reported in literature may serve as a starting point for identifying the most suitable formulation.

The autoinduction medium formulated by Studier may potentially be used to express recombinant bromelain in *E. coli* BL21-AI under control of the T7lac promoter with possible omission of the expensive induction agents, such as arabinose or IPTG[11]. This medium depends on the preferential diauxic metabolic mechanisms which bacteria use to regulate the uptake of multiple carbon sources such as glucose, glycerol and lactose, in the growth medium. As glucose is depleted, glycerol is preferentially consumed, leading to biomass production. This is followed by lactose which will be metabolized to allolactose leading to induction and subsequent expression. Using this medium, cultures in batch production in well-aerated baffled shake flask are reportedly capable of reaching final OD600 of 10 – 20, which is relatively high compared to most standard batch cultivation which uses other types of media[12-16]. Moreover, since unintended induction is greatly minimized by catabolite repression due to presence of glucose, the medium is also ideal for the expression of potentially toxic proteins[11,16].

In this study, developing for high-level expression of recombinant bromelain in *E. coli* clearly highlighted a need for fine-tuning the autoinduction medium composition for maximization of biomass production. The effects of medium components on biomass production were investigated using fractional factorial design to identify which candidates influence biomass productivity the most.

MATERIALS AND METHODS

Bacterial strain and plasmid

E. coli strain BL21-AI harboring pDEST17/Bromelain vector was used[9]. In this study, expression was induced in the presence of L-arabinose. Production seeds were stored in MDG broth[11] with 20% glycerol at -80°C.

Media and chemicals

Reagents and chemicals were purchased from Merck Bioscience (Darmstadt, Germany), Sigma-Aldrich (St Louis, MO, USA) and DuchefaBiochemie (Sparks, MD, USA).

Inoculum preparation

A frozen glycerol stock of recombinant cells was sufficiently scrapped using sterile pipette tips and immediately dipped into a 10 ml of MDG broth supplemented with 100 µg/ml ampicillin, in a 250-ml Erlenmeyer ûask. The inoculum ûask was incubated until OD600 reached approximately ~10-11 (12-14 hrs) on a rotary shaker at 250 rpm and, 37°C.

Experimental design

In this work, 22 factors from Studier autoinduction media components including antibiotic were screened for their effects on CDW (the response) using minimum run equireplicated resolution IV of factorial design. Table 1 shows the factors and levels applied in the design; the delimitation of experimental region for each factor was determined from published data. A total of 52 experimental runs were required to analyse the effect of each component. The statistical signiûcance of each individual factor and their combinations at 5% signiûcance level were evaluated using the DesignExpert v8.0 (StatEase Inc., Mn, USA) software.

Experimental procedures

All the 52 experimental runs were performed simultaneously and in triplicate. This was made possible using a 50-ml centrifuge tube instead of conventional Erlenmeyer flask. An inoculum (50 µl) was added to 5 ml of each media as tabulated in Table 2 in separate 50-ml centrifuge tubes incubated at 250 rpm and, 37°C. After 3 hrs, the incubation temperature was then reduced to 25°C for an additional 8 hrs. The cells were harvested by centrifugation at 4°C 4,696 x g for 15 min and the supernatant was discarded. The cells were kept at -80°C until further analysis.

Determination of cell dry weight (CDW)

Biomass concentration expressed in g/l was determined by aliquoting a 1-ml sample from the main culture followed by centrifugation as described in cell lysate preparation section and subjected to drying at 105°C for 12 hours in a pre-weighted aluminium container.

RESULTS AND DISCUSSION

Productivity in recombinant protein production depends on cultivation of biomass of the host in fermentation. It is generally believed that high productivity of recombinant protein in fermentation results from high productivity in biomass and, vice versa. Although this has been demonstrated in prior studies, it is not always the case, because other problems associated with metabolic burden such as plasmid instability and recombinant protein toxicity are common[11]. These problems are mainly caused by unintended induction of recombinant protein[11]. The autoinduction medium reported here was originally formulated in such a way as to minimize or eliminate these cell stress effects[11]. Therefore, the former correlation between recombinant protein and biomass productivity was assumed to hold true using the medium. In this study, the medium components' optimization on the basis of recombinant biomass productivity was highlighted, assuming that it reflected the productivity of recombinant bromelain itself using the investigated cultivation medium.

Table 2 shows the design matrix used under resolution IV of fractional factorial design (equireplicated minimum run), along with the cell dry weight (CDW in g/l) represented biomass productivity as the response measured in each run. All experiments (52 runs) were conducted in one block of measurements and the experimental sequence (Std Order) was randomized in order to minimize the effects of biased factors. The CDW was found to range from 3.1 to 7.3 g/l and the significant effect of each factor on CDW was evaluated by a normal probability plot of standardized effects, a Pareto chart, main effects

Table 1. Range and level of components individually tested in fractional factorial design for biomass production of recombinant *E. coli* BL21-AI

Screening	Factors		Low -1	Center 0	High +1
X1	Tryptone	% (w/v)	1	2.5	4
X2	Yeast Extract	% (w/v)	0.5	1.25	2
X3	Na_2HPO_4	mM	25	37.5	50
X4	KH_2PO_4	mM	25	37.5	50
X5	NH_4Cl	mM	50	75	100
X6	Na_2SO_4	mM	5	8	10
X7	Glycerol	% (w/v)	0.5	2.75	5
X8	Glucose	% (w/v)	0.05	0.525	1
X9	Lactose	% (w/v)	0	0.1	0.2
X10	$MgSO_4$	mM	0	1	2
X11	$CoCl_2$	µM	0	0.2	0.4
X12	$CuCl_2$	µM	0	0.2	0.4
X13	$NiCl_2$	µM	0	0.2	0.4
X14	Na_2MoO_4	µM	0	0.2	0.4
X15	Na_2SeO_3	µM	0	0.2	0.4
X16	$FeCl_3$	µM	0	5	10
X17	$CaCl_2$	µM	0	2	4
X18	$MnCl_2$	µM	0	1	2
X19	$ZnSO_4$	µM	0	1	2
X20	H_3BO_3	µM	0	0.2	0.4
X21	L-arabinose	% (w/v)	0	0.025	0.05
X22	Ampicillin	µg/ml	100	150	200

Table 2 . Fractional factorial design matrix for biomass production of recombinant *E. coli* BL21-AI. Shaded rows are runs with high CDW.

Stdrun	X1 % (v/v)	X2 % (v/v)	X3 mM	X4 mM	X5 mM	X6 mM	X7 % (w/v)	X8 % (w/v)	X9 % (w/v)	X10 mM	X11 uM	X12 uM	X13 uM	X14 uM	X15 uM	X16 uM	X17 uM	X18 uM	X19 uM	X20 uM	X21 %	X22 ug/ml (w/v)	CDW g/l
1	1	-1	1	1	-1	-1	-1	1	-1	1	1	-1	-1	-1	-1	1	1	-1	1	1	-1	-1	5.1
2	-1	1	-1	1	-1	-1	-1	-1	1	1	1	1	1	-1	1	-1	-1	1	1	1	-1	1	7.3
3	1	1	1	-1	-1	-1	1	1	1	1	-1	-1	1	1	1	-1	1	1	1	-1	-1	-1	4.5
4	1	1	-1	1	1	-1	1	1	-1	-1	-1	1	-1	-1	-1	-1	-1	1	1	1	1	-1	2.8
5	1	1	-1	-1	-1	1	-1	1	-1	-1	-1	-1	1	-1	1	1	-1	1	-1	1	-1	-1	5.3
6	-1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	5.1
7	-1	1	-1	-1	1	1	-1	1	-1	1	1	1	-1	1	1	-1	-1	-1	1	-1	-1	-1	5.1
8	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	-1	1	1	1	-1	-1	1	-1	-1	1	5.5
9	-1	1	-1	-1	1	-1	1	-1	1	-1	1	1	-1	-1	1	1	1	-1	-1	-1	-1	-1	5.7
10	1	1	1	-1	-1	-1	-1	1	-1	-1	1	1	1	1	1	-1	1	-1	-1	1	1	1	3.9
11	1	-1	1	1	1	1	1	-1	-1	1	-1	1	1	-1	1	-1	1	-1	-1	1	-1	-1	6.0
12	-1	1	1	1	1	1	-1	1	-1	-1	1	1	1	-1	-1	1	1	1	1	-1	-1	1	5.3
13	-1	1	1	-1	1	-1	-1	1	1	1	-1	1	1	-1	-1	-1	-1	-1	-1	-1	1	-1	4.2
14	1	1	1	-1	1	1	-1	-1	1	-1	-1	-1	-1	-1	-1	-1	1	-1	1	1	-1	1	7.0
15	1	-1	-1	-1	1	1	-1	-1	1	1	1	1	1	1	-1	1	1	1	1	1	1	-1	5.9
16	-1	-1	1	-1	-1	1	-1	1	1	-1	-1	1	-1	-1	1	-1	1	1	1	1	1	-1	4.5
17	1	-1	-1	-1	-1	1	1	1	1	-1	1	1	1	-1	-1	-1	-1	-1	1	-1	-1	1	4.1
18	-1	-1	1	-1	1	1	1	1	-1	1	1	-1	-1	1	-1	-1	-1	1	-1	1	-1	1	3.9
19	-1	1	-1	1	-1	1	1	1	1	1	-1	1	-1	1	-1	1	1	-1	-1	1	-1	1	4.5
20	-1	1	-1	1	-1	1	-1	-1	-1	-1	-1	-1	1	1	-1	-1	1	-1	1	-1	1	-1	5.9
21	-1	-1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	1	-1	1	1	1	1	2.9
22	1	-1	-1	-1	-1	-1	-1	-1	-1	1	-1	1	-1	-1	-1	-1	1	1	-1	-1	-1	1	6.8
23	1	1	-1	1	1	1	-1	1	1	1	1	-1	-1	-1	1	-1	1	1	-1	-1	1	1	4.3
24	-1	1	-1	-1	1	1	1	-1	1	-1	-1	1	1	1	1	-1	-1	1	-1	-1	1	1	4.5
25	1	-1	1	-1	1	1	1	1	-1	-1	-1	-1	-1	1	-1	1	1	-1	-1	-1	1	-1	3.1
26	-1	-1	-1	1	1	1	-1	-1	-1	-1	1	1	-1	-1	-1	1	-1	-1	-1	1	1	1	6.0

Table 2 . Cont...

Stdrun	X1 % (v/v)	X2 % (v/v)	X3 mM	X4 mM	X5 mM	X6 mM	X7 % (w/v)	X8 % (w/v)	X9 % (w/v)	X10 mM	X11 uM	X12 uM	X13 uM	X14 uM	X15 uM	X16 uM	X17 uM	X18 uM	X19 uM	X20 uM	X21 %	X22 ug/ml (w/v)	CDW g/l
27	-1	-1	1	-1	-1	1	-1	-1	1	1	1	-1	1	1	1	1	1	-1	-1	-1	-1	1	7.2
28	-1	-1	1	1	1	-1	1	-1	1	1	1	1	-1	1	-1	-1	1	-1	1	-1	1	1	4.8
29	1	1	1	-1	-1	1	1	-1	-1	1	-1	1	-1	-1	1	1	-1	-1	1	-1	1	1	4.4
30	1	-1	1	1	-1	-1	1	-1	1	-1	-1	-1	1	-1	-1	1	1	1	-1	1	1	1	4.5
31	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	-1	1	1	-1	1	1	-1	4.5
32	1	-1	1	1	-1	1	-1	1	-1	1	-1	-1	1	1	-1	-1	-1	1	1	1	1	1	3.8
33	-1	-1	-1	1	1	1	1	-1	1	1	-1	-1	-1	-1	-1	1	-1	1	1	-1	-1	-1	5.7
34	-1	1	-1	-1	-1	-1	-1	1	1	-1	1	-1	-1	1	-1	1	-1	1	1	-1	1	1	4.1
35	1	-1	-1	-1	-1	-1	1	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	-1	1	1	-1	4.6
36	1	-1	-1	1	-1	1	1	-1	-1	-1	1	-1	-1	1	1	1	1	1	1	1	-1	1	5.9
37	-1	-1	-1	1	-1	-1	1	1	-1	1	1	1	1	1	1	1	-1	1	-1	-1	1	-1	3.1
38	-1	1	1	1	-1	-1	1	1	-1	-1	-1	-1	-1	-1	1	-1	-1	-1	-1	-1	-1	1	4.6
39	1	1	-1	1	1	-1	1	-1	-1	1	1	-1	1	1	-1	1	-1	-1	-1	-1	-1	1	5.9
40	-1	1	1	1	1	-1	-1	-1	-1	1	-1	-1	-1	1	1	1	1	1	-1	1	1	-1	5.8
41	1	1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	-1	1	1	1	-1	1	-1	1	-1	6.0
42	1	-1	1	-1	1	-1	-1	-1	-1	-1	1	-1	1	-1	1	-1	-1	1	1	-1	1	-1	5.7
43	1	-1	1	-1	1	-1	1	1	1	1	1	1	-1	-1	1	1	-1	1	-1	1	-1	1	4.3
44	1	1	1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	-1	1	-1	-1	-1	-1	6.7
45	-1	1	1	1	1	1	1	1	1	-1	1	-1	1	1	1	1	-1	-1	1	1	1	-1	3.1
46	-1	-1	1	-1	-1	-1	1	-1	-1	-1	-1	1	1	1	-1	1	-1	-1	1	1	-1	-1	5.7
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.2
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.3
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.4
51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.3
52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.3

Table 3. Original ANOVA of main effects for CDW: Their estimated coefficients and significance

Screening Factors			Coefficient (Code value)	p-value
X1	Tryptone	% (w/v)	4.91	0.56
X2	Yeast Extract	% (w/v)	-0.034	0.8849
X3	Na_2HPO_4	mM	-8.78E-03	0.9474
X4	KH_2PO_4	mM	3.89E-03	0.3701
X5	NH_4Cl	mM	0.053	0.8044
X6	Na_2SO_4	mM	-0.014	0.8372
X7	Glycerol	% (w/v)	0.012	< 0.0001
X8	Glucose	% (w/v)	-0.57	< 0.0001
X9	Lactose	% (w/v)	-0.81	0.3287
X10	$MgSO_4$	mM	0.057	0.9038
X11	$CoCl_2$	µM	7.01E-03	0.6982
X12	$CuCl_2$	µM	-0.023	0.9376
X13	$NiCl_2$	µM	4.63E-03	0.9347
X14	Na_2MoO_4	µM	4.75E-03	0.8047
X15	Na_2SeO_3	µM	-0.015	0.2692
X16	$FeCl_3$	µM	0.065	0.8479
X17	$CaCl_2$	µM	0.011	0.575
X18	$MnCl_2$	µM	0.033	0.6229
X19	$ZnSO_4$	µM	-0.029	0.9412
X20	H_3BO_3	µM	-4.26E-03	0.811
X21	L-arabinose	% (w/v)	-0.014	< 0.0001
X22	Ampicillin	µg/ml	-0.6	0.8359
R-Squared	0.9344			
Adj R-Squared		0.8846		
Pred R-Squared		0.8638		

plotted at 5% signiûcance level using the DesignExpert v8.0 (StatEase Inc., Mn, USA) software.

The estimates of main effects of the factors are shown on a half-normal probability plot of effects (Figure 1). All insignificant effects are normally distributed with a mean of zero and tend to fall along a straight line in the plot. In contrast, significant effects have non-zero means and are located further away from the straight line. The larger the signiûcant effects, the further away they are from the straight line. Signiûcant effects that emerge from this analysis are the main effects of glucose (represented in the software as H), L-arabinose (V) and glycerol (G).

Their apparent significance may be attributed to negative rather than positive effects, as shown in Figure 2(a). It has been perceived that pure carbon sources such as glucose and glycerol promote rapid growth of *E. coli* cells which is a

Table 4. Final ANOVA from reduced regression model of CDW

Source	Sum of Square	df	Mean square	F value	p-value prob> F	
Model	57.27	3	19.09	194.82	< 0.0001	significant
X7-Glycerol	14.92	1	14.92	152.27	< 0.0001	significant term
X8-Glucose	30.83	1	30.83	314.59	< 0.0001	significant term
X21-L-arabinose	16.48	1	16.48	168.23	< 0.0001	significant term
R-Squared	0.9241					
Adj R-Squared	0.9194					
Pred R-Squared	0.9161					

contradictory with the findings. One feasible cause is that while glucose and glycerol support favourably cell growth, at increasing concentrations or prolonged cultivation especially in shaking culture, they may pose growth inhibitory effects due to unmonitored pH fluctuations and limited oxygen mass transfer. This tendency imposed crab-tree effects due to accumulation of acidic TCA cycle products, especially acetate, during cell cultivation[11,13,15,17]. The pH fluctuations assumption were confirmed in 2-liter bioreactor production (data not shown), as online monitoring showed decreasing in pH when glucose and glycerol were consumed and then started increasing as the cells consumed solely on nitrogen source to grow. As described in previous section, each experimental run was performed in a standard 50-ml tube. Despite attempts to minimize oxygen limiting effects by keeping the culture-to-container volume ratio about 10% in the tube and shaking at high revolution (rpm), the oxygen limiting condition may have been unavoidable in non-aerated cultivation as the cell density increased until saturation.

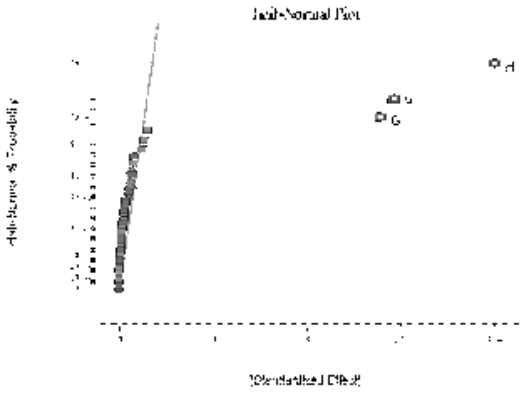

Fig. 1. Half-normal plots of standardized effects for CDW (in g/l)

The upper rank of positive main effects with respect to CDW was found to be Na_2SeO_3, lactose and KH_2PO_4, accordingly, as shown in Figure 2(a). While these main effects evidently provided growth support, their overall degree of significance (<1%) was actually largely overshadowed by the extent of glucose, L-arabinose and glycerol (>90%), as shown in Figure 2(b). As a result, there was relatively insignificant overall contribution to the standardized effects. Specific mechanisms on how Na_2SeO_3 and KH_2PO_4 are crucially involved in cell growth are relatively unknown and the current statistical analysis does not support deduction of concrete possibilities until further validation experiments are carried out. Lactose, in addition to being a metabolizable inducer, is by itself a carbon source for cell growth. Unlike glucose and glycerol, lactose is not prone to crab-tree effect because the metabolism was much slower and the cells shift from fermentative

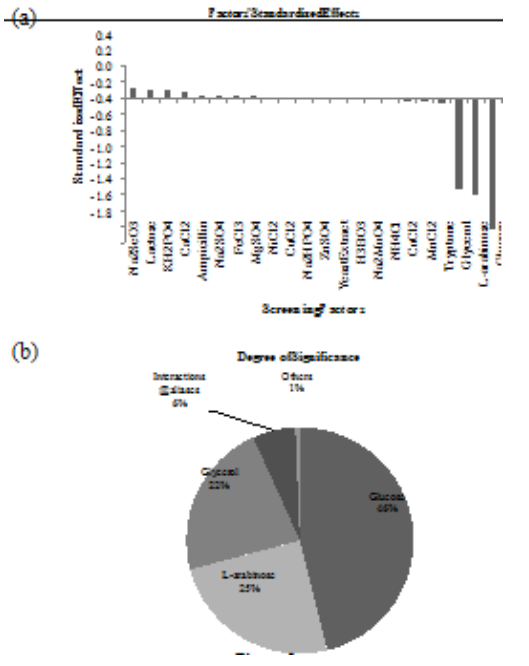

Fig. 2. Factors' standardized effects (a) and their relative degree of significance (b)

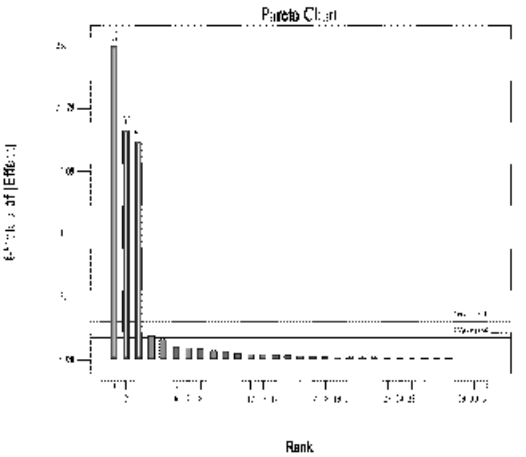

Fig. 3. Pareto chart of standardized effects for CDW productivity

to recombinant protein expressive modes when being consumed.

It should be noted that terms of interaction were purposely excluded in the model terms selection because they were statistically aliases due to the nature of minimal run in resolution IV of factorial design as reported by the software. Inclusion of the aliased terms would probably produce misleading regression results. To truly understand, such interactions found would require further work or additional experimental runs. However, given that not enough information was gathered to properly analyze and dealias the interactions, reliable conclusions in the findings were restricted to main effects model only. It can be concluded that glucose, L-arabinose and glycerol are the only important effects that stand out in this screening step.

The results obtained from the half-normal and normal probability plot of effects (Figure 1(a) and 1(b)) were conûrmed with a Pareto chart as shown in Figure 3. The lower horizontal line (T-value limit) across the Pareto chart indicates the minimum level where the effect is possibly statistically significant while the upper line (Bonferroni limit) indicates the minimum level where the effects are almost certainly significant. The vertical column lengths are proportional to the degree of signiûcance for each main effect. In Figure 3, the sequence of the signiûcant main effects with respect to increasing inûuence on CDW was in agreement with that obtained from the normal probability plot of standardized effects, which were KH_2PO_4, lactose and Na_2SeO_3, accordingly.

The estimates of main effects of the factors were also visually examined in the main effects plots (Figure 4). The steep effect lines between the low and high levels obtained for the main effects of glycerol, L-arabinose, and glucose

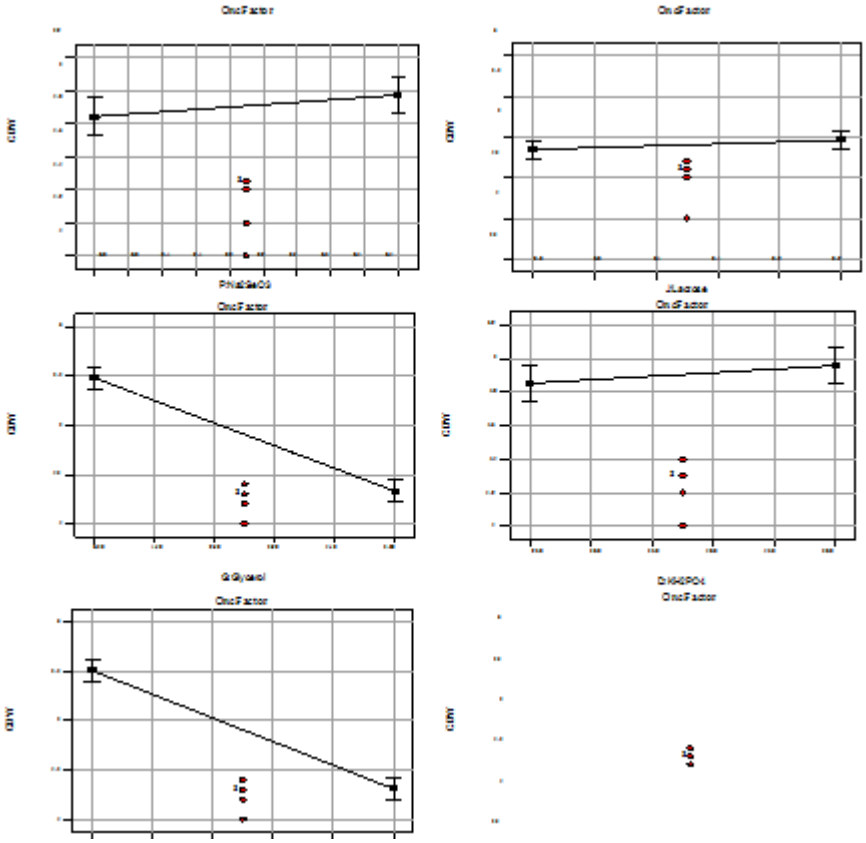

Fig. 4. Visual inspection on main effects plots for CDW. The Y-axis represents CDW values while X-axis represents main effects code between low and high levels. From upper left to right side:Na_2SeO_3, lactose, KH_2PO_4, glycerol, L-arabinose, and glucose

denote that they significantly affected CDW. The ûat effect lines attained for the rest of the main effects, on the other hand, reveal their insigniûcant effects on CDW (Figure 4). The plots which also resemble one-factor-at-time graph will be used as indicators to determine suitable optimal range for the next optimization steps.

Table 3 presents the analysis of variance (ANOVA) for the main effects obtained from the least squares method using the DesignExpert software. An initial first-order model in coded unit (-1, 0 or +1) which correlates CDW was proposed and given by equation (1):

$$CDW = +4.91 - 0.034X1 - 8.782\times10^{-3}X2 + 3.892\times10^{-3}X3 + 0.053X4 - 0.014X5 + 0.012X6 - 0.57X7 - 0.81\times X8 + 0.057X9 + 7.007\times10^{-3}X10 - 0.023X11 + 4.626\times10^{-3}X12 + 4.750\times10^{-3}X13 - 0.015X14 + 0.065X15 + 0.011X16 + 0.033X17 - 0.029X18 - 4.264\times10^{-3}X19 - 0.014X20 - 0.60X21 + 0.012X22 \quad ...(1)$$

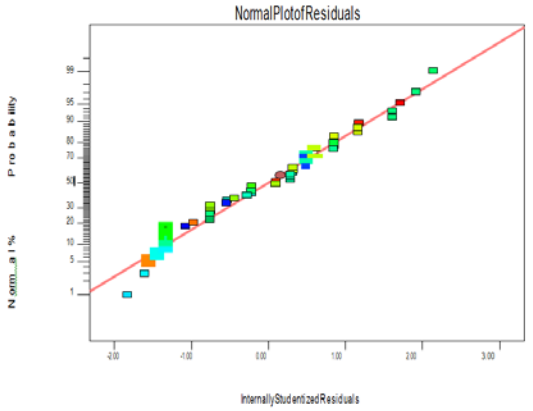

Fig. 5. The plot shows ANOVA assumptions were satisfied upon visual inspection

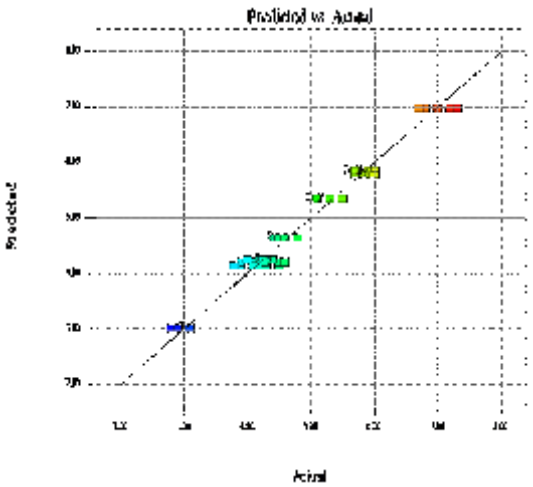

Fig. 6. Comparison of predicted and experimental CDW

Where each regression coefûcient was computed according to the least square method and the results are as tabulated in Table 3. After eliminating the insigniûcant terms of P-values more than 0.05 (Table 3), a reduced model of Eq. (2) was produced:

$$CDW = +4.91 - 0.57X7 - 0.82\times X8 - 0.060X21 \quad ...(2)$$

The adequacy or goodness of ût of the reduced regression model for CDW (equation (2)) was analyzed by ANOVA (Table 4). From Table 4, the high F-values and low P-values of the main effects (glycerol, L-arabinose, and glucose) suggest that they contribute signiûcantly to the response CDW.

The R^2 value of the model obtained is 0.9241 (Table 4). This denotes that only 7.59% of the total variability is not explained by the regression in the model. The high R^2 value signifies that the model is able to give a reasonably good estimate of response for the system in the range studied. This finding reliability is supported by a normal probability plot of standardized residuals (Fig. 5). From Figure 5, all points lie reasonably along the straight line without apparent s-shaped pattern, lending the support that the ANOVA assumptions were satisfied and thus the analysis was reliable.

A relatively small difference between the R^2 and adjusted-R^2 values at 4.98%, implies that there is a lower chance that non-signiûcant terms have been included in the model. The predicted CDW from the model was also compared to the experimentally measured CDW (Figure 6). Hence, the reduced model (Eq. (2)) can be used as a predictive tool to obtain *E. coli* BL21-AI biomass production over the entire uncertainty range of glycerol, L-arabinose, and glucose studied.

CONCLUSIONS

Screening of 22 factors of Studier autoinduction media affecting the CDW using minimum run equireplicated resolution IV of fractional factorial design reveals that only glycerol, L-arabinose, and glucose are inûuential statistically. A reduced regression model for CDW was developed and its R^2 (92.41%), adjusted-R^2 (91.94%) and predicted-R^2 (91.61%) values were determined. A high R^2 indicates that the model obtained is able to give a reasonably good estimate

of response for a system in the studied range. The results of this study indicate the suitability of fractional factorial design for evaluating the effect of a large number of variables with a minimal number of experiments.

ACKNOWLEDGEMENTS

The authors acknowledged the Ministry of Agriculture Malaysia (MOA) for providing the research fund (Technofund 10-01F046) and the Biotechnology Engineering Department, IIUM for research facilities. Authors also acknowledge the Ministry of Education Malaysia (MOE) for providing MSc tuition fees under MyBrain scheme for MohdJamilAizatJamaluddin.

REFERENCES

1. Peti, W., Page, R. Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expres Purif*. 2007; **51**: 1-10.
2. Kim, M., Elvin, C., Brownlee, A., Lyons, R. High yield expression of recombinant pro-resilin: Lactose-induced fermentation in *E. coli* and facile purification. *Protein Expres Purif*. 2007; **52**: 230-6.
3. Maiti, S.K., Lantz, A.E., Bhushan, M., Wangikar, P.P. Multi-objective optimization of glycopeptide antibiotic production in batch and fed batch processes. *Bioresource Technol*. 2011; **102**: 6951-8.
4. Dubey, K.K., Behera, B.K. Statistical optimization of process variables for the production of an anticancer drug (colchicine derivatives) through fermentation: at scale-up level. *New Biotechnol*. 2011; **28**: 79-85.
5. Wang, Z., Li, J., Cheong, S., Bhaskar, U., Akihiro, O., Zhang, F., *et al*. Response surface optimization of the heparosan N-deacetylation in producing bioengineered heparin.*JBiotechnol*. 2011; **156**: 188-96.
6. Muntari, B., Amid, A., Mel, M., Jami, M., Salleh, H. Recombinant bromelain production in *Escherichia coli*: process optimization in shake flask culture by response surface methodology. *AMB Expr*. 2012; **2** :1-9.
7. Yari, K., Fatemi, S.-A., Tavallaei, M. High level expression of recombinant BoNT/A-Hc by high cell density cultivation of *Escherichia coli*. *Biopro and Biosyst Eng*. 2012; **35**: 407-14.
8. Sarduy, E.S., Muñoz, A.C., Trejo, S.A., Chavéz Planes, M.d.l.A. High-level expression of Falcipain-2 in *Escherichia coli* by codon optimization and auto-induction.*Protein Expres Purif*. 2012; **83**: 59-69.
9. Amid, A., Ismail, N.A., Yusof, F., Salleh, H.M. Expression, purification, and characterization of a recombinant stem bromelain from Ananas comosus. *Process Biochem*. 2011; **46**: 2232-9.
10. Bala, M., Salleh, H., Amid, A., Mel, M., Jami, M. Recovery of recombinant bromelain from *Escherichia coli* BL 21-AI. *Afr J of Biotechnol*. 2011; **10**: 18829-32.
11. Studier, F.W. Protein production by auto-induction in high density shaking cultures.*Protein Expres Purif*.. 2005; **41**: 207-34.
12. Tripathi, N.K., Shrivastva, A., Biswal, K.C., Rao, P.L. METHODS: Optimization of culture medium for production of recombinant dengue protein in *Escherichia coli*. *Industrial Biotechnology*. 2009; **5**: 179-83.
13. Ikram, N., Naz, S., Rajoka, M.I., Sadaf, S., Akhtar, M.W. Enhanced production of subtilisin of Pyrococcus furiosus expressed in *Escherichia coli* using autoinducing medium.*Afr J of Biotechnol*. 2009; **8**.
14. Voulgaridou, G.-P., Mantso, T., Chlichlia, K., Panayiotidis, M.I., Pappa, A. Efficient *E. coli* Expression Strategies for Production of Soluble Human Crystallin ALDH3A1. *PloS one*. 2013; 8: e56582.
15. Blommel, P.G., Becker, K.J., Duvnjak, P., Fox, B.G. Enhanced Bacterial Protein Expression During Auto Induction Obtained by Alteration of Lac Repressor Dosage and Medium Composition. *Biotechnology progress*. 2007; **23**: 585-98.
16. Lee, S.K., Keasling, J.D. Heterologous protein production in *Escherichia coli* using the propionate-inducible pPro system by conventional and auto-induction methods.*Protein Expres Purif*. 2008; **61**: 197-203.
17. Losen, M., Frölich, B., Pohl, M., Büchs, J. Effect of Oxygen Limitation and Medium Composition on *Escherichia coli* Fermentation in Shake Flask Cultures. *Biotechnology progress*. 2004; **20**: 1062-8.