J Ind Microbiol Biotechnol (2011) 38:1045–1054
DOI 10.1007/s10295-010-0879-2

ORIGINAL PAPER

Influence of induction conditions on the expression of carbazole dioxygenase components (CarAa, CarAc, and CarAd) from *Pseudomonas stutzeri* in recombinant *Escherichia coli* using experimental design

Ariane Leites Larentis · Haryana de Cássia Cunha Sampaio · Orlando Bonifácio Martins · Maria Isabel Rodrigues · Tito Lívio Moitinho Alves

Received: 20 March 2010 / Accepted: 15 September 2010 / Published online: 17 October 2010

Abstract Carbazole 1,9a-dioxygenase (CarA), the first enzyme in the carbazole degradation pathway used by *Pseudomonas* sp., was expressed in *E. coli* under different conditions defined by experimental design. This enzyme depends on the coexistence of three components containing [2Fe–2S] clusters: CarAa, CarAc, and CarAd. The catalytic site is present in CarAa. The genes corresponding to components of carbazole 1,9a-dioxygenase from *P. stutzeri* were cloned and expressed by salt induction in *E. coli* BL21-SI (a host that allows the enhancement of overexpressed proteins in the soluble fraction), using the vector pDEST^{TM}14. The expression of these proteins was performed under different induction conditions (cell concentration, temperature, and time), with the help of two-level factorial design. Cell concentration at induction (measured by absorbance at 600 nm) was tested at 0.5 and 0.8. After salt induction, expression was performed at 30 and 37°C, for 4 h and 24 h. Protein expression was evaluated by densitometry analysis. Expression of CarAa was enhanced by induction at a lower cell concentration and temperature and over a longer time, according to the analysis of the experimental design results. The results were validated at $Abs_{ind} = 0.3$, 25°C, and 24 h, at which CarAa expression was three times higher than under the standard condition. The behavior of CarAc and CarAd was the inverse, with the best co-expression condition tested being the standard one ($Abs_{ind} = 0.5$, $T = 37$°C, and $t = 4$ h). The functionality of the proteins expressed in *E. coli* was confirmed by the degradation of 20 ppm carbazole.

Keywords Carbazole · Recombinant protein expression · *Escherichia coli* BL21-SI · Factorial design · Statistical design of experiments

A. L. Larentis (✉) · H. C. C. Sampaio · T. L. M. Alves
Programa de Engenharia Química-COPPE, Laboratório de Bioprocessos, Universidade Federal do Rio de Janeiro–UFRJ, Centro de Tecnologia (CT), G115, Cidade Universitária, Ilha do Fundão, Caixa Postal 68502, Rio de Janeiro, RJ 21945-970, Brazil
e-mail: ariane@peq.coppe.ufrj.br

A. L. Larentis · H. C. C. Sampaio · O. B. Martins
Instituto de Bioquímica Médica, Laboratório de Biologia Molecular, Universidade Federal do Rio de Janeiro–UFRJ, Centro de Ciências da Saúde (CCS), Bloco D, Cidade Universitária, Ilha do Fundão , Caixa Postal 68041, Rio de Janeiro, RJ 21941-590, Brazil

M. I. Rodrigues
Faculdade de Engenharia de Alimentos, Universidade Estadual de Campinas–UNICAMP, Cidade Universitária Zeferino Vaz, Barão Geraldo, Campinas, SP 13083-970, Brazil

Introduction

The degradation of carbazole and its dibenzopyrrole derivatives, commonly found in fossil fuels and associated with nitrogen oxide production during oil combustion, has been widely investigated, in particular in the last decade as environmental policies on gas emissions have become more stringent. The presence of these nitrogenated aromatic compounds, of which carbazole is one of the main species, is a characteristic of Brazilian crude oil [29]. These species are deleterious to the refining process, because they are catalyst inhibitors and alter the quality of petroleum derivatives [2, 42]. However, in nature there are some carbazole-degrading bacteria, such as *Pseudomonas stutzeri*, which are able to assimilate carbazole from oil as a source of nitrogen and carbon [17]. The first enzyme in

their degradation route is carbazole 1,9a-dioxygenase (CarA), which depends on the coexistence of three components: CarAa (catalytic oxygenase component that contains a Rieske [2Fe–2S] cluster and a mononuclear iron domain), CarAd (containing FAD and [2Fe–2S] cluster binding sites) acts as a reductase component and transfers electrons from NADH to a ferredoxin component with a Rieske [2Fe–2S] cluster (CarAc). The catalytic site for oxygen activation is a mononuclear iron domain present in CarAa [24, 44]. The use of these enzymes overexpressed in recombinant microorganisms is a strategy that is attracting great biotechnological interest in the biorefining field because the protein production is higher than when enzymes from wild strains are used, which could improve degradation rates [21, 39, 42].

The Gram-negative bacteria *Escherichia coli* is the most widely used prokaryotic expression host for the high-level production of recombinant proteins, and also on a commercial scale. Of the many systems available for heterologous expression, it remains one of the most attractive due to its ability to grow rapidly and at high cell concentrations on inexpensive substrates, as well as its well-characterized genetics and the availability of an increasingly large number of cloning vectors and mutant host strains [1, 20, 46, 50]. The efficiency of recombinant protein expression is directly associated with: the plasmid, which depends on the promoter and its regulation, initiation, and terminator regions, enhancers, ribosome binding site efficiency, origin of replication (*ori*), resistance markers, plasmid copy number (gene dosage) and size; the target heterologous protein, such as the nucleotide and amino acid sequence/secondary structure, the appropriate codon usage, mRNA and protein stability; and the host strain [1, 20, 31, 46]. Besides exploring the genetic characteristics of the different *E. coli* strains employed in recombinant protein overexpression, it is essential to manipulate the process variables in order to enhance the yield of heterologous proteins and process productivity. As most heterologous proteins are intracellular in the recombinant bacteria, process productivity is proportional to final cell mass and specific protein productivity. By manipulating the process temperature, medium composition (especially carbon sources), pH, aeration/rotation, and seed conditions, it is possible to reduce undesirable effects, such as substrate inhibition, the formation of metabolites that inhibit cell growth, and to prevent and/or reduce the formation of inclusion bodies [10, 20, 31].

Indeed, recombinant protein expression and cell growth both depend on numerous mutually interacting variables. The most common strategy for evaluating the influence of these variables on heterologous protein expression is to change one factor at a time while keeping the others constant. This method assumes that all the variables are independent, which is usually too simplistic an assumption when considering complex biological systems, and is often inefficient because it does not permit an analysis of the interaction between the variables. Furthermore, varying several factors and several conditions for each factor simultaneously leads to the need for a higher number of experiments to encompass the same information than would be necessary using experimental design. The best approach is to use experimental design, which allows the most process information to be gathered at the lowest cost. It also makes it possible to evaluate which variables and interactions influence the expression of recombinant proteins, reducing the total number of experiments that need to be performed to obtain such information [14, 15, 18, 19, 49, 51].

Experimental design has only recently been applied to the screening and optimization of recombinant protein expression in *E. coli*. These techniques are still not usually employed for the analysis of expression variables related to specific target proteins (construct length and expression system), fermentation conditions (culture media and process time), or protein induction conditions in molecular biology practice. To our knowledge, there are some research groups using experimental design for recombinant protein expression in *E. coli*: for evaluating the effects of specific induction variables on heterologous expression, such as induction temperature, post-induction time, cell concentration at induction, inducer concentration, host strain, and plasmid [5, 8, 9, 14–16, 19, 28, 30, 32, 36, 37, 47, 49, 51–53]; for culture medium composition [6, 7, 11, 13–15, 18, 19, 22, 27, 28, 34, 35, 38, 40, 41, 45, 48, 51, 52, 54, 55]; for the concentration of antibiotics [14, 15, 18] used for selective pressure; for seed conditions and inoculum level/density [14, 15, 37, 47]; for bioprocess variables, such as oxygen transfer rate, pH, temperature, fluid hydrodynamics, process time, as well as feeding strategy [11–13, 19, 40, 45, 53].

Given the importance of analyzing the influence of process variables on the expression of recombinant proteins, the aim of this work was to evaluate the effects of induction conditions (cell concentration, temperature, and time) on the expression of CarAa, CarAc, and CarAd, components of carbazole 1,9a-dioxygenase from *P. stutzeri*, in *E. coli* BL21-SI with the help of experimental design. This strain contains an NaCl-inducible T7 system and is protease-deficient to minimize heterologous protein degradation [3]. By employing two-level factorial experimental design, it was possible to identify the interaction between the variables under investigation, which is not common in studies of recombinant protein expression.

Materials and methods

Strain and plasmid

The three components (CarAa, CarAc, and CarAd) of carbazole 1,9a-dioxygenase from *P. stutzeri* ATCC31258 were expressed in *E. coli* BL21-SI [3] with NaCl induction using the plasmid pDEST^{TM}14 (Invitrogen, Carlsbad, CA, USA), without purification tags, in two different cloning constructions (Fig. 1): one containing the *carAa* gene and the other with *carAc* and *carAd* genes, besides the *ORF7* gene, whose encoded protein is not essential to dioxygenase activity. The cloning procedures are described in detail in Larentis et al. [25, 26] for CarB (2′-aminobiphenyl-2,3-diol-1,2-dioxygenase), the second enzyme of the *P. stutzeri* carbazole degradation pathway (Fig. 1).

Experimental design, statistical analysis, and validation

The analysis of the effects of three independent induction variables (cell concentration, temperature, and time) on the expression of components of carbazole 1,9a-dioxygenase in *E. coli* BL21-SI was performed by using two-level factorial design, containing duplicates of all the combinations of the three factors (2^3). The statistical evaluation of the effects of cell concentration, temperature, and time on the induction of proteins was done using the normalized variable at levels −1 (lower of the experimental conditions used) and +1 (higher of the experimental conditions), using STATISTICA 6.0 software (StatSoft, Inc., Tulsa, OK, USA).

The significance of each linear effect and interactions was determined by using Student's *t* test, at a 0.05 probability level (95% confidence level). The *p*-value represents the probability that a given variable has an insignificant effect on the response [19]. The effects were statistically significant when the *p*-value was lower than 0.05.

The effects estimated by the statistical software indicated the strength of the main effects and their interactions. The variable normalization allows for the comparison of each effect and interactions from the most to the least significant (a higher effect indicates a higher influence). A statistically significant effect at the significance level used for the statistical analysis indicates that a variation from the

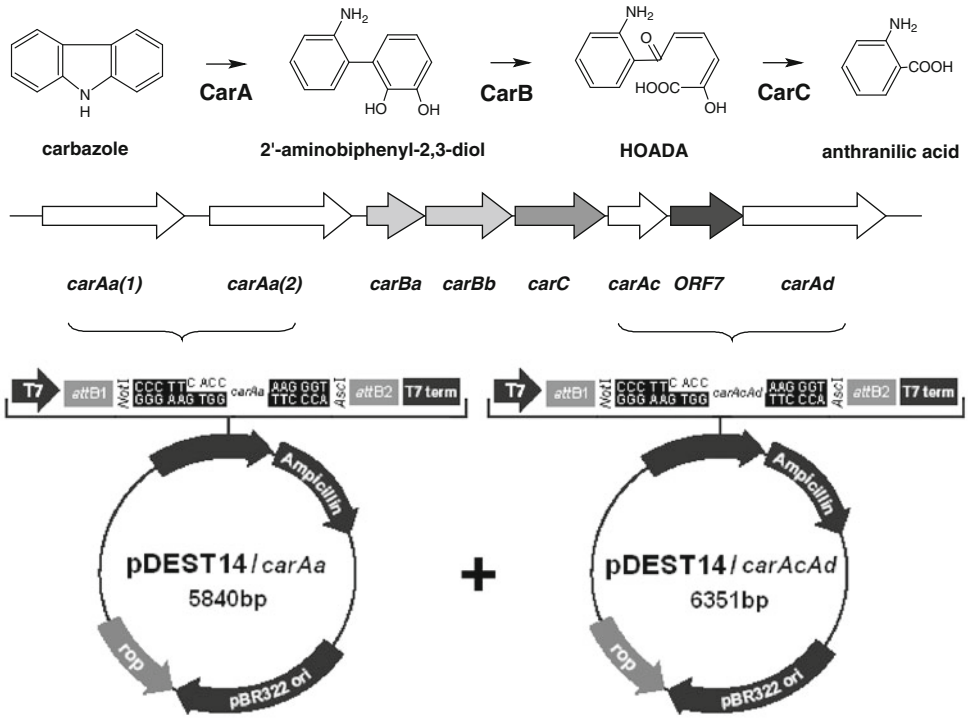

Fig. 1 Enzymes and map of genes from operon *car* required for carbazole degradation pathway to anthranilic acid from *Pseudomonas* sp. [2, 42, 44]. The first enzyme, carbazole 1,9a-dioxygenase (CarA), converts carbazole into 2′-aminobiphenyl-2,3-diol and depends on the coexistence of three components containing [2Fe–2S] clusters, encoded by *carAa* (duplicated in the operon), *carAc* and *carAd* genes, respectively: CarAa, CarAc, and CarAd. The second enzyme, a *meta*-cleavage dioxygenase named 2′-aminobiphenyl-2,3-diol 1,2-dioxygenase (CarB), which is encoded by both *carBa* and *carBb* genes, catalyzes the *meta*-cleavage of the catecholic ring of 2′-aminobiphenyl-2,3-diol to produce 2-hydroxy-6-oxo-6-(2′-aminophenyl)-hexa-2,4-dienoic acid (HOADA). The third enzyme (CarC), encoded by *carC* gene, presents an α/β hydrolase characteristic domain, and catalyzes the hydrolysis of HOADA to produce anthranilic acid. The reaction mechanism is detailed in Larentis et al. [24, 25]. *carAa* (1,100 bp) and *carAcORF7carAd* (1,700 bp) were cloned separately in two different constructions, as indicated in the expression vectors (pDEST14/*carAa* and pDEST14/*carAcAd*)

lower (−1) to the upper level (+1) in the studied variable leads to an enhancement in the response if the effect is positive. A negative effect indicates a reduction in the evaluated response with a variation from (−1) to (+1) levels.

The best experimental conditions for CarAa, CarAc, and CarAd expression were validated at cell concentrations, temperatures, and induction times defined by the statistical analysis of the two-level experimental design results.

Expression

Both *E. coli* BL21-SI harboring expression vectors pDEST^{TM}14 cloned with *carAa* and *carAcAd* genes were cultivated in 10 ml LBON medium (1% bactotryptone and 0.5% yeast extract) in the absence of NaCl, with ampicillin 100 μg/ml at 37°C for 16 h. Then 5% of each culture after 16 h were used to inoculate 10 ml LBON with ampicillin (resulting in an initial absorbance of 0.2 at 600 nm) in 50-ml Falcon tubes, with vigorous shaking at 37°C. The cells were grown at this temperature for 90–120 min until they reached the exponential phase (absorbance 0.5 or 0.8 at 600 nm, depending on the experimental design, as described in the previous section). NaCl was then added until a final concentration of 0.3 M for the expression of recombinant proteins under the control of the osmotically induced promoter of *E. coli* BL21-SI [3]. In the induction phase, protein expression was performed at 30°C and 37°C for 4 h and 24 h, according to the experimental design. Then 1-ml samples (before salt addition and after induction under the different experimental conditions) were harvested and the pellets were stored at −20°C. *E. coli* BL21-SI without any plasmid was used as a negative control. Cells with an empty plasmid were not used as control because the pDEST^{TM}14 vector contains a selection marker (*ccd*B gene), which inhibits growth of *E. coli* [26].

Preparation of cell extract and SDS–PAGE

The pellets from the 1-ml expression samples were resuspended in the appropriate buffer (described below) and disrupted by sonication on ice over five cycles of 10 s pulses (constant cycle, output 3) with a 30 s interval using an ultrasonic cell disruptor Branson Sonifier® 250 (Branson Ultrasonics Corp., Danbury, CT, USA) to obtain the cell extracts.

For the SDS–PAGE analysis, the cell extracts were prepared in 50 mM glucose, 10 mM EDTA, and 25 mM Tris–HCl, and had their total protein concentrations measured as described by Bradford [4], using bovine serum albumin as a standard. The 18% SDS–polyacrylamide gel electrophoresis [23] was performed with 20 μg of each cell extract (obtained from uninduced and induced samples) in a Bio-Rad apparatus. The proteins were stained with Coomassie brilliant blue R-250.

For the carbazole degradation test, the cell extracts were prepared in 50 mM Tris–HCl (pH 7.5).

Protein expression analysis

After deducting background values, the band areas from the SDS–PAGE corresponding to each protein expressed (42 kDa CarAa, 12 kDa CarAc, and 36 kDa CarAd) were analyzed by densitometry with the help of QuantiScan 1.25 program (BIOSOFT, Great Shelford, Cambridge, UK; http://www.biosoft.com/w/quantiscan.htm) and QuantiOne 4.4.1 software (Bio-Rad Laboratories, Hercules, CA, USA). They were then normalized, assuming that Area = 100 is the area obtained in the standard condition for the induction of recombinant proteins in *E. coli* ($Abs_{ind} = 0.5$, $T = 37$°C, and $t = 4$ h). A similar strategy for normalizing the intensity of bands from SDS–PAGE and comparing the expression of the desired protein with the standard condition was also described by Urban et al. [51].

Carbazole degradation test

The functionality of CarAa, CarAc, and CarAd expressed under the conditions selected for statistical data validation was confirmed by the degradation of 20 ppm carbazole (120 μM) with carbazole dioxygenase, using gas chromatography in a Varian CP-3380 GC with FID detector and CP-Sil 5 CB column (Varian, Inc., Palo Alto, CA, USA). NADH (500 μM), FAD (1 μM), 35 μM ammonium ferrous sulfate (Fe(II) source), and 50 μM ascorbic acid (reducing agent) were used in the sonicated cell extract (300 μl) of both recombinant *E. coli* BL21-SI/pDEST^{TM}14 cloned with the genes *carAa* and *carAcAd*. Tween 20 was added to aid the dispersion of carbazole in water. The degradation was verified after 15 h of reaction at 30°C in 50 mM Tris–HCl (pH 7.5).

Results and discussion

The proteins CarAa, CarAc, and CarAd were expressed in *E. coli* BL21-SI by salt induction under different cell concentrations (Abs_{ind}), temperatures (T), and times (t). Through SDS–PAGE analysis, it was possible to identify stronger bands for the induced samples than for the uninduced samples (before salt addition), at the expected protein sizes: about 42 kDa, 12 kDa, and 36 kDa for CarAa, CarAc, and CarAd, respectively [44]. The results indicate the expression of the desired proteins by the induction of recombinant plasmids with salt. The expression levels obtained for the site-specific recombination system and

salt-inducible cells were similar to those obtained for the well-known pUC cloning system with isopropyl β-D-1-thiogalactopyranoside (IPTG) induction [25, 44]. These bands were not detected in the cell extract of *E. coli* BL21-SI without the cloning construction (negative control). The BL21-SI host is derived from the GJ1158 strain, which employs NaCl as the inducer for the overexpression of proteins. In this system, developed by Bhandari and Gowrishankar [3], the synthesis of the RNA polymerase of bacteriophage T7 is placed under the control of the osmotically inducible *E. coli proU* promoter. Salt induction is associated with an increased proportion of overexpressed proteins in the soluble fraction due to the reduced formation of inclusion bodies [3].

Expression was performed at different temperatures after NaCl addition, in order to enhance protein production. The pre-induction phase was conducted at the optimum temperature for *E. coli* growth (37°C) and post-induction was tested at a lower temperature, which led to lower growth rates in order to reduce the formation of inclusion bodies and undesirable metabolites, as discussed below.

The areas of the expression bands obtained by densitometry analysis and normalized for each induction condition are presented in Table 1. Each condition was analyzed in duplicate, and the average error obtained was around 12%.

The expression areas of CarAa, CarAc, and CarAd ranged from 0 (no protein expression) to 250.5 (more than 150% enhancement when compared to the standard condition, i.e., Area = 100), achieved by varying the cell concentration, temperature, and induction time. Densitometry is the technique normally employed to determine protein concentration [9, 11, 22, 27, 28, 32, 34, 36, 40, 45, 51, 53] when it is not possible to evaluate it by means of enzyme activity assays.

Table 1 Area of expression bands for proteins CarAa, CarAc, and CarAd under different cell concentrations at induction (0.5 and 0.8), induction temperatures (30°C and 37°C), and times after induction (4 h and 24 h) with 0.3 M NaCl, calculated in relation to the standard condition (Area = 100 at $Abs_{ind} = 0.5$, $T = 37$°C, and $t = 4$ h, *highlighted in bold*)

Runs	Abs_{ind}	T (°C)	t (h)	CarAa	CarAc	CarAd
1	−1 (0.5)	−1 (30°C)	−1 (4 h)	96.1	28.5	80.3
2	−1 (0.5)	−1 (30°C)	+1 (24 h)	250.5	136.2	46.9
3	−1 (0.5)	+1 (37°C)	−1 (4 h)	**100.0**	**100.0**	**100.0**
4	−1 (0.5)	+1 (37°C)	+1 (24 h)	111.1	40.4	51.9
5	+1 (0.8)	−1 (30°C)	−1 (4 h)	112.7	66.5	144.4
6	+1 (0.8)	−1 (30°C)	+1 (24 h)	151.4	91.2	94.5
7	+1 (0.8)	+1 (37°C)	−1 (4 h)	105.7	156.2	12.0
8	+1 (0.8)	+1 (37°C)	+1 (24 h)	18.5	191.7	0.0

The values of the areas correspond to the average calculated from duplicates (error of around 12%)

Abs_{ind} cell concentration at induction (measured by absorbance at 600 nm), T induction temperature (°C), t induction time (h)

Each protein responded differently to the induction variables analyzed. Higher levels of CarAa expression were obtained for lower temperatures and longer induction times. For all the experiments performed at 4 h induction, the expression level was near the standard condition ($Abs_{ind} = 0.5$, $T = 37$°C, and $t = 4$ h), independent of the cell concentration at induction or the temperature. CarAa expression at $Abs_{ind} = 0.5$ and 37°C is shown in Fig. 2. However, the comparison between the experiments carried out at the same temperature and for 24 h indicated that a lower cell concentration at induction led to higher expression levels. At 30°C, longer induction periods led to higher expression levels (Table 1), and for 24 h there was higher cell growth. Meanwhile, at 37°C and $Abs_{ind} = 0.8$, a likely degradation of CarAa was verified when comparing 4 h of induction with 24 h (Table 1). Thus, in experiments performed over 24 h induction, expression was greater at the lowest temperature and cell concentration (at $Abs_{ind} = 0.8$ and 37°C, CarAa expression was quite low). These experiments showed the interaction between the variables which could not be evaluated by a one-factor-at-a-time analysis.

CarAc and CarAd are present in the same plasmid (as indicated in Fig. 1) and were expressed together in the same flask. CarAd expression (Table 1) was lower at absorbance 0.5 than at 0.8 for 30°C; the higher expression level obtained for a higher cell concentration at induction could be associated with the toxic effect of this heterologous protein on the cell. In contrast, for 37°C the level of expression at absorbance 0.8 was near zero, but was higher at 0.5 (condition shown in Fig. 2). The lower level of CarAd expression under all conditions after 24 h of induction, as compared to 4 h, suggests the possibility of recombinant protein degradation. For CarAc, on the other hand, the effects of the variables under analysis were the opposite of what was found for CarAd expression. The best results for CarAc expression were obtained at $Abs_{ind} = 0.8$ and 37°C, conditions where no CarAd expression was identified (as shown in Fig. 2). The expression levels were similar to those obtained by Sato et al. [44]. The inverse variable effects presented in Fig. 2 indicate the trend for CarAc and CarAd expression to have opposing responses. As discussed elsewhere, protein expression depends on the nucleotide/amino acid sequence and the size of the target, as well as expression vector construct [19, 51].

The levels of expression of carbazole 1,9a-dioxygenase components in *E. coli* are similar to those for other proteins produced in shaking flasks [9, 13, 22, 40, 44]. In this system, concentrations in the order of grams per liter of recombinant product can be obtained, depending on the

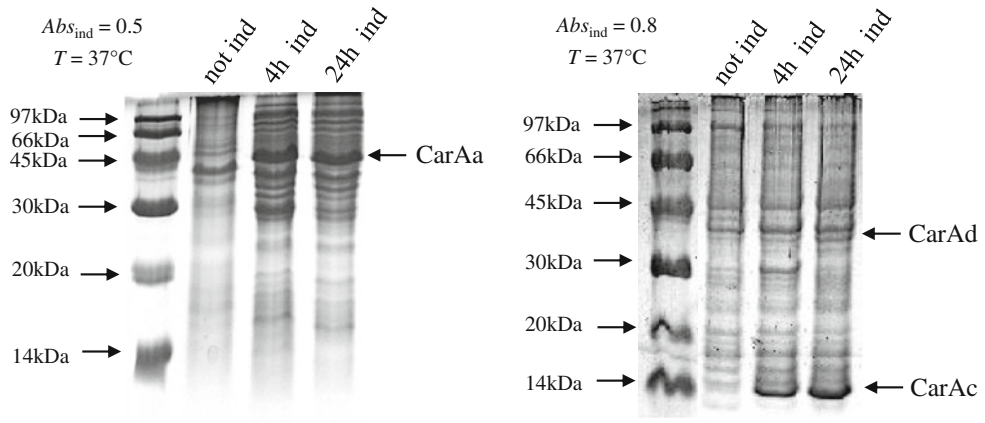

Fig. 2 SDS–PAGE of recombinant *E. coli* BL21-SI/pDEST^{TM}14 with 20 μg total protein in each sample of cell extract of CarAa (42 kDa), CarAc (12 kDa), and CarAd (36 kDa). The protein molecular weight standard is LMW (Amersham Bioscience). The corresponding areas of each protein in the gel are: 100.0 and 111.1 for CarAa expression at $Abs_{ind} = 0.5$, 37°C, and 4 h and 24 h, respectively; 156.2 and 191.7 for CarAc at $Abs_{ind} = 0.8$, 37°C, and 4 h and 24 h, respectively, and near zero for CarAd under the same conditions (opposing responses), as shown in Table 1

heterologous protein, expression system, and the process conditions employed.

Statistical analysis

The statistical evaluation of the effects of cell concentration, temperature, and time on the induction of proteins was performed by using normalized variables at levels (−1) and (+1), which were the lower and higher values from the experimental conditions tested, respectively, using the expression bands presented in Table 1. The effects of each variable can be seen in Table 2. The adopted experimental design presented no correlation between the variables, i.e., it was possible to evaluate each effect independently.

The statistical analysis of CarAa expression served to identify which variables had most influence on the results of the experimental design. It was found that the temperature (T) and the combination of this variable with induction time ($T \times t$) had the greatest influence on CarAa expression, followed by cell concentration at induction and its synergism with time ($Abs_{ind} \times t$ and Abs_{ind}), all having a negative effect on CarAa expression. T and Abs_{ind} were found to have a negative effect, meaning that higher levels of CarAa expression were obtained for lower T and Abs_{ind}. Under these conditions, the expression levels could be enhanced in longer processes, since the protein was not degradated at lower T and Abs_{ind}, as described above. The estimated effects and statistical analysis are presented in Table 2, where all effects were significant ($p < 0.05$, in bold), except $Abs_{ind} \times T$.

For CarAc, the synergism of cell concentration with the temperature ($Abs_{ind} \times T$) and Abs_{ind}, followed by the linear effect of temperature (T) and its combination with time ($T \times t$) were statistically significant, presenting p-values lower than 0.05 (effects shown in bold in Table 2). Abs_{ind}, T, and $Abs_{ind} \times T$ had a positive effect on the expression, indicating that the enhancement of CarAc expression could be obtained by higher cell concentrations and temperatures, while longer induction times had no detrimental effect on protein expression levels. $T \times t$ was the only parameter to have a negative effect on CarAc expression.

The interaction of cell concentration with temperature ($Abs_{ind} \times T$), temperature (T), and induction time (t) had the greatest negative effects on CarAd expression (with $p < 0.05$, in bold), i.e., for CarAd, lower temperatures and induction times can enhance protein expression. The others variables and interactions were not statistically significant, since the p-values associated with their effects were higher than 0.05. As discussed elsewhere, the opposite behaviors of CarAc and CarAd expression under the different induction conditions tested can be seen by the opposite signs in terms of the statistical effects presented in Table 2.

Table 2 Effect estimates for the expression of carbazole dioxygenase subunits CarAa, CarAc, and CarAd in *E. coli* BL21-SI (2^3 factorial design)

Factors	CarAa	CarAc	CarAd
Mean	**118.2**	**101.3**	**66.2**
Abs_{ind}	**−42.4**	**50.1**	−7.0
T (°C)	**−68.9**	**41.5**	**−50.5**
t (h)	**29.3**	27.1	**−35.8**
$Abs_{ind} \times T$ (°C)	−1.1	**53.6**	**−62.9**
$Abs_{ind} \times t$ (h)	**−53.5**	3.0	4.9
T (°C) × t (h)	**−67.3**	**−39.1**	5.8

A negative signal indicates a negative effect of the variable on the response, while a positive signal indicates a positive effect. $p < 0.05$ indicates the effect is statistically significant (*highlighted in bold*)

Influence of cell concentration at induction, temperature, and induction time

Cell concentration at induction could also be determined by the pre-induction time and is related to the growth phase when the inducer is added to the culture for the expression of heterologous protein under the control of a regulated inducible promoter. Manderson et al. [33] tested induction at seeding and showed that this procedure severely limits culture growth and recombinant protein production as a result of metabolic stress: the expression of foreign genes imposes considerable pressure on metabolic systems and limits the available energy for *E. coli* growth. Induction in the stationary phase decreases culture viability and growth rates and can lead to protease production (as well as the degradation of recombinant proteins and yield reduction) brought on by nutrient depletion [10, 52]. It is discussed in the literature that recombinant protein production in *E. coli* is proportional to the specific growth rate at induction if cell metabolic capacity is at its best condition [33]. In view of this observation, induction is usually performed at the exponential phase, but depending on the protein and induction system [9, 19, 30, 32, 47, 53], protein expression can vary with a lower or higher cell concentration prior to induction (i.e., early or late logarithmic phase). In many cases, the best approach is to maximize the number of viable cells prior to the addition of the inducer, which often retards cell growth [19]. Donovan et al. [10] observe that for strains whose growth is reduced after the addition of an inducer, induction in the late logarithmic phase provides high cell densities, increasing heterologous protein production. This strategy of growing cells first to a high concentration (late exponential phase) in the absence of recombinant target expression and then inducing the promoter system is also described by Shin et al. [45] and Manderson et al. [33]. Because the recombinant protein can significantly inhibit host cell growth, presumably due to its toxicity, it is essential to control the time and extent of plasmid gene expression [45]. Retardation of growth may be caused by the metabolic load imposed on the host through heterologous gene expression [27, 33]. When foreign protein expression is low or not significantly influenced by cell growth, the expression yield is maximized by inducing throughout the entire growth cycle [10], with high biomass yield. Cellular responses to induction (and the metabolic burden required for overproduction of a foreign protein) depend on a number of interacting factors, including the host/vector system and the properties of the expressed protein [8, 19, 27, 33, 45], which highlights the importance of using experimental design rather than a simplistic approach that would consider the variables independently.

Temperature is usually described as being one of the most influential variables on the expression of recombinant proteins [9, 30, 45, 49, 51, 52] because of its relationship with bacterial metabolism, as well as protein synthesis and the folding mechanisms associated with the production of inclusion bodies in *E. coli* (the rate of heterologous protein synthesis can be so high that proteins are not properly folded). Temperature is related to modulations in chaperone activity or other accessory proteins involved in protein folding [51], so it is associated with proteins obtained in soluble form, which are useful for further purification [19, 49]. Temperature is also reported to affect plasmid stability in recombinant *E. coli* cultures [45, 53]. Cultivation at lower temperatures can reduce protein aggregation into inclusion bodies, proteolytic degradation, and responses to cell stress conditions, leading to a higher production of soluble recombinant proteins [10, 30, 31, 52, 53]. As discussed by Wang et al. [53], the optimal temperature for the promotion of cell growth can be deleterious to foreign protein expression since a higher growth rate would lead to a greater probability of plasmid loss and mispartition of the expression vector, especially for plasmid-carrying cells expressing toxic proteins. Lower temperatures and correspondingly lower growth rates are usually associated with the enhancement of soluble protein production [51], depending on the characteristics of the heterologous protein and expression system. The most suitable strategy, which was adopted in this study, is to conduct the pre-induction process phase at 37°C, which leads to the highest *E. coli* growth rates, followed by the addition of the inducer at the appropriate cell concentration, then reduce the temperature and the post-induction growth rate so that the factors affecting protein synthesis can be better regulated and recombinant protein expression is correspondingly enhanced. Shin et al. [45] report that no correlation between the pre-induction specific growth rate and specific recombinant protein synthesis was observed for several proteins, while the results of Islam et al. [19] indicated that the temperature in the pre-induction process phase had an insignificant effect on protein expression. Nevertheless, Shin et al. [45] argue that the relationship depends on the characteristics of the particular recombinant system (promoter system, host–vector interaction, toxicity of the recombinant product). Ultimately, the experimental conditions required to obtain an optimal response should take into account both cell growth and genetic factors (mRNA and protein metabolic synthesis and degradation) in order to enhance the expression of the heterologous protein. Researchers [19, 51] have noted that the highest soluble protein yields are obtained under conditions which promote both a slow growth rate during protein synthesis and a high final biomass yield.

Swalley et al. [49] and Pan et al. [37] describe the interdependence of induction temperature and time after the addition of the inducer. Longer induction times at higher temperatures usually had a detrimental effect on protein expression levels, mainly on the obtainment of protein in the soluble fraction because of a tendency to form inclusion bodies. These authors [49] also observe that the interaction of time with temperature tends to mask the effect of time as a main variable. Depending on the time selected to perform the experiments, it could also mask the effects of other variables. This is the reason why Rodrigues and Iemma [43] warn that time should be avoided as an independent variable in batch systems, and that whenever possible, samples should be taken at all times of interest and analyzed one by one. The degradation of heterologous proteins expressed in long cultivation processes is also pointed out by Urban et al. [51], Wang et al. [53], Chen et al. [6], and Islam et al. [19], probably due to the action of some proteinases.

Validation of statistical data

The statistical analysis of the two-level experimental design results should be validated and employed in order to define the best experimental conditions for enhancing protein expression.

The analysis of the influence of cell concentration, temperature, and time on the expression of components of carbazole 1,9a-dioxygenase indicated that the induction of CarAa at the lowest Abs_{ind}, at lower temperatures, and for longer times could significantly enhance the expression of the heterologous protein in comparison with the standard condition. Based on this, the experiment for CarAa expression was performed at Abs_{ind} of about 0.31, allowing the cells to grow for 30 min at 37°C before induction, since culture growth and recombinant protein production are severely limited by induction at seeding [33]. Protein expression after induction was performed at 25°C for 24 h. This condition was selected in order to allow the cells to grow, and the use of temperatures below ambient temperature was avoided. Lower temperatures would lead to higher costs and require longer times to achieve similar growth, which would reduce process productivity. CarAa expression (Area = 300) was three times higher than under the standard condition and similar (slightly enhanced) to the best condition obtained in the two-level experimental design performed (Abs_{ind} = 0.5, 30°C, and 24 h), validating the statistical analysis performed on cell concentration at induction, temperature, and time.

As the *carAc* and *carAd* genes are present in the same plasmid (Fig. 1) and the conditions for CarAc expression enhancement were the opposite of those found for CarAd (the behavior of CarAc expression was the inverse to that of CarAd), the co-expression of both proteins was performed under the standard condition (Abs_{ind} = 0.5, T = 37°C, and t = 4 h). Reasonable expression levels were achieved for both proteins under this condition, given that the active site of carbazole 1,9a-dioxygenase is present in the CarAa subunit and CarAc and CarAd are involved in electron transfer to the active site [24, 44].

The degradation of 20 ppm carbazole by CarA was confirmed by the reaction of in vitro cell extracts of both recombinant *E. coli* (harboring expression vectors pDEST^{TM}14 carrying *carAa* and *carAcAd* genes, respectively, as indicated in Fig. 1), obtained under the respective validated expression conditions. The capacity to degrade carbazole shown by the sonicated cell extract with the addition of cofactors confirmed the functionality of the expressed proteins CarAa, CarAc, and CarAd in *E. coli*. The conditions, concentrations, and ratios of the components involved in the in vitro reaction of carbazole degradation can also be optimized by using experimental design in future studies.

Conclusions

The two-level factorial experimental design was a valuable tool for analyzing the influence of induction variables on the expression of components of carbazole dioxygenase and screening experimental conditions for the improvement of protein expression. The analysis presented in this work could help the design of experiments to enhance expression of the target proteins associated with carbazole degradation. It was observed that the response of each of the three components of carbazole dioxygenase was different for each induction variable analyzed: expression of CarAa was enhanced by induction at a lower cell concentration and temperature, for longer induction times, while the behavior of CarAc and CarAd was the contrary, with the standard condition providing the best expression results. The procedure used for the data validation of CarAa expression could be used to analyze process productivity over time, in order to reduce the time for which the same expression of heterologous protein is kept. The experimental design employed made it possible to evaluate the effect of the expression variables and their interactions in order to choose the best conditions to enhance the production of a recombinant protein, achieving higher process yields and lower costs for bioreactor applications in scaling-up.

Despite the increasing number of published articles making use of experimental design, there is still ample room to explore the enhancement of recombinant protein expression using this technique. It is still usual in molecular biology to evaluate each variable separately, which

may lead to a misinterpretation of the data obtained, because it fails to take account of their interactions, and usually requires more experiments to reach the same conclusions as those reached using experimental design techniques. These powerful tools could be very helpful for screening production conditions and optimizing recombinant protein expression in *E. coli*, as well as other heterologous hosts. The results presented in this work for the analysis of induction variables on the expression of components of carbazole dioxygenase corroborate the applicability of experimental design to the field of molecular biology.

Acknowledgments CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and Petrobras supported this work. We thank to Iuri Bastos M.Sc. (UFRJ) for the densitometry analysis using the QuantiScan 1.25 program, Dr. Márcio Schwaab and Prof. José Carlos Pinto (COPPE/UFRJ) for fruitful discussions, and the reserchers of the Setor de Imunobiológicos laboratory at INCQS/Fiocruz for making available the QuantiOne 4.4.1 program.

References

1. Baneyx F (1999) Recombinant protein expression in *Escherichia coli*. Curr Opin Biotechnol 10:411–421. doi:10.1016/S0958-1669(99)00003-8
2. Benedik MJ, Gibbs PR, Riddle RR, Willson RC (1998) Microbial denitrogenation of fossil fuels. Trends Biotechnol 16:390–395. doi:10.1016/S0167-7799(98)01237-2
3. Bhandari P, Gowrishankar J (1997) An *Escherichia coli* host strain useful for efficient overproduction of cloned gene products with NaCl as the inducer. J Bacteriol 179:4403–4406
4. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254. doi:10.1016/0003-2697(76)90527-3
5. Cao Y, Xia Q, Fang B (2006) Optimization of expression of *dha*T gene encoding 1,3-propanediol oxidoreductase from *Klebsiella pneumoniae* in *Escherichia coli* using the methods of uniform design and regression analysis. J Chem Technol Biotechnol 81:109–112. doi:10.1002/jctb.1360
6. Chen Y, Xing X-H, Ye F, Kuang Y, Luo M (2007) Production of MBP-HepA fusion protein in recombinant *Escherichia coli* by optimization of culture medium. Biochem Eng J 34:114–121. doi:10.1016/j.bej.2006.11.020
7. Choi WC, Oh BC, Kim HK, Lee ES, Oh TK (2002) Medium optimization for phytase production by recombinant *Escherichia coli* using statistical experimental design. J Microbiol Biotechnol 12:490–496
8. Chuan YP, Lua LHL, Middelberg APJ (2008) High-level expression of soluble viral structural protein in *Escherichia coli*. J Biotechnol 134:64–71. doi:10.1016/j.jbiotec.2007.12.004
9. De León A, Jiménez-Islas H, González-Cuevas M, Barba de la Rosa AP (2004) Analysis of the expression of the *Trichoderma harzianum ech42* gene in two isogenic clones of *Escherichia coli* by surface response methodology. Process Biochem 39:2173–2178. doi:10.1016/j.procbio.2003.11.013
10. Donovan RS, Robinson CW, Glick BR (1996) Optimizing inducer and culture conditions for expression of foreign proteins under the control of the *lac* promoter. J Ind Microbiol Biotechnol 16:145–154. doi:10.1007/BF01569997
11. Fontani S, Niccolai A, Kapat A, Olivieri R (2003) Studies on the maximization of recombinant *Helicobacter pylori* neutrophil-activating protein production in *Escherichia coli*: application of Taguchi robust design and response surface methodology for process optimization. World J Microbiol Biotechnol 19:711–717. doi:10.1023/A:1025104119260
12. García-Arrazola R, Dawson P, Buchanan I, Doyle B, Fearn T, Titchener-Hooker N, Baganz F (2005) Evaluation of the effects and interactions of mixing and oxygen transfer on the production of Fab' antibody fragments in *Escherichia coli* fermentation with gas blending. Bioprocess Biosyst Eng 27:365–374. doi:10.1007/s00449-005-0414-4
13. Han GH, Shin H-J, Kim SW (2008) Optimization of bio-indigo production by recombinant *E. coli* harboring fmo gene. Enzyme Microb Technol 42:617–623. doi:10.1016/j.enzmictec.2008.02.004
14. Hao DC, Zhu PH, Yang SL, Yang L (2006) Optimization of recombinant Cytochrome P450 2C9 protein production in *Escherichia coli* DH5α by statistically-based experimental design. World J Microbiol Biotechnol 22:1169–1176. doi:10.1007/s11274-006-9158-9
15. Hao DC, Zhu PH, Yang SL, Yang L (2007) Enhanced production of human Cytochrome P450 2C9 by *Escherichia coli* BL21(DE3)pLysS through the novel use of grey relational analysis and Plackett–Burman design. World J Microbiol Biotechnol 23:71–78. doi:10.1007/s11274-006-9194-5
16. Hernández VEB, Maldonado LMTP, Rivero EM, Barba de la Rosa AP, Acevedo LGO, De León Rodríguez A (2008) Optimization of human interferon gamma production in *Escherichia coli* by response surface methodology. Biotechnol Bioprocess Eng 13:7–13. doi:10.1007/s12257-007-0126-5
17. Hisatsuka K, Sato M (1994) Microbial transformation of carbazole to anthranilic acid by *Pseudomonas stutzeri*. Biosci Biotechnol Biochem 58:213–214. doi:10.1271/bbb.58.213
18. Hounsa CG, Aubry JM, Dubourguier HC, Hornez JP (1996) Application of factorial and Doehlert designs for optimization of pectate lyase production by a recombinant *Escherichia coli*. Appl Microbiol Biotechnol 45:764–770. doi:10.1007/s002530050760
19. Islam RS, Tisi D, Levy MS, Lye GJ (2007) Framework for the rapid optimization of soluble protein expression in *Escherichia coli* combining microscale experiments and statistical experimental design. Biotechnol Prog 23:785–793. doi:10.1021/bp070059a
20. Jana S, Deb JK (2005) Strategies for efficient production of heterologous proteins in *Escherichia coli*. Appl Microbiol Biotechnol 67:289–298. doi:10.1007/s00253-004-1814-0
21. Kilbane II JJ (2004) Petroleum biorefining: the selective removal of sulfur, nitrogen, and metals. In: Vazquez-Duhalt R, Quintero-Ramirez R (eds) Petroleum biotechnology, vol 151: development and perspectives, studies in surface science and catalysis, chap 2. Elsevier, Mexico City
22. Kotik M, Kocanová M, Maresová H, Kyslík P (2004) High-level expression of a fungal pyranose oxidase in high cell-density fed-batch cultivations of *Escherichia coli* using lactose as inducer. Protein Expr Purif 36:61–69. doi:10.1016/j.pep.2004.02.011
23. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685. doi:10.1038/227680a0
24. Larentis AL, Almeida RV, Cardoso AM, Almeida WI, Rössle SC, Bisch PM, Martins OB, Alves TLM (2006) Homology modeling of the oxygenase component of carbazole 1,9a-dioxygenase (CarAa) involved in petroleum denitrogenation pathway of Pseudomonas sp. Braz Arch Biol Technol 49:53–61
25. Larentis AL, Almeida RV, Rössle SC, Cardoso AM, Almeida WI, Bisch PM, Alves TLM, Martins OB (2006) Expression and homology modeling of 2′-aminobiphenyl-2,3-diol 1,2-dioxygenase

(CarB) from *Pseudomonas stutzeri* carbazole degradation pathway. Cell Biochem Biophys 44:530–538. doi:10.1385/CBB:44:3:530
26. Larentis AL, Alves TLM, Martins OB (2005) Cloning and expression of *meta*-cleavage enzyme (CarB) of carbazole degradation pathway from *Pseudomonas stutzeri*. Braz Arch Biol Technol 48:127–134. doi:10.1590/S1516-89132005000400016
27. Lee KM, Rhee CH, Kang CK, Kim JH (2006) Sequential and simultaneous statistical optimization by dynamic design of experiment for peptide overexpression in recombinant *Escherichia coli*. Appl Biochem Biotechnol 135:59–80. doi:10.1385/ABAB:135:1:59
28. Lee KM, Rhee CH, Kang CK, Kim JH (2006) Statistical medium formulation and process modeling by mixture design of experiment for peptide overexpression in recombinant *Escherichia coli*. Appl Biochem Biotechnol 135:81–100. doi:10.1385/ABAB:135:1:81
29. Leite LF, Neto JNN, Bevilaqua JV (2005) Biorefineries and biofuels: current activities and future vision of Petrobras. ACS Div Fuel Chem 50:726–727
30. Lo PK, Hassan O, Ahmad A, Muhammad Mahadi N, Illias RM (2007) Excretory over-expression of *Bacillus* sp. G1 cyclodextrin glucanotransferase (CGTase) in *Escherichia coli*: optimization of the cultivation conditions by response surface methodology. Enzyme Microb Technol 40:1256–1263. doi:10.1016/j.enzmictec.2006.09.020
31. Makrides SC (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. Microbiol Rev 60:512–538
32. Maldonado LMTP, Hernández VEB, Rivero EM, Barba de la Rosa AP, Flores JLF, Acevedo LGO, De León Rodríguez A (2007) Optimization of culture conditions for a synthetic gene expression in *Escherichia coli* using response surface methodology: the case of human interferon beta. Biomol Eng 24:217–222. doi:10.1016/j.bioeng.2006.10.001
33. Manderson D, Dempster R, Chisti Y (2006) A recombinant vaccine against hydatidosis: production of the antigen in *Escherichia coli*. J Ind Microbiol Biotechnol 33:173–182. doi:10.1007/s10295-005-0046-3
34. Niccolai A, Fontani S, Kapat A, Olivieri R (2003) Maximization of recombinant *Helicobacter pylori* neutrophil activating protein production in *Escherichia coli*: improvement of a chemically defined medium using response surface methodology. FEMS Microbiol Lett 221:257–262. doi:10.1016/S0378-1097(03)00184-8
35. Nikerel İE, Toksoy E, Kirdar B, Yildirim R (2005) Optimizing medium composition for *Taq*I endonuclease production by recombinant *Escherichia coli* cells using response surface methodology. Process Biochem 40:1633–1639. doi:10.1016/j.procbio.2004.06.017
36. Oliveira C, Costa S, Teixeira JA, Domingues L (2009) cDNA cloning and functional expression of the α-D-galactose-binding lectin frutalin in *Escherichia coli*. Mol Biotechnol 43:212–220. doi:10.1007/s12033-009-9191-7
37. Pan H, Xie Z, Bao W, Zhang J (2008) Optimization of culture conditions to enhance *cis*-epoxysuccinate hydrolase production in *Escherichia coli* by response surface methodology. Biochem Eng J 42:133–138. doi:10.1016/j.bej.2008.06.007
38. Pan HF, Bao WN, Xie ZP, Zhang JG (2010) Optimization of medium composition for *cis*-epoxysuccinate hydrolase production in *Escherichia coli* by response surface methodology. Afr J Biotechnol 9:1366–1373
39. Park H-S, Kayser KJ, Kwak JH, Kilbane JJ II (2004) Heterologous gene expression in *Thermus thermophilus*: β-galactosidase, dibenzothiophene monooxygenase, PNB carboxy esterase, 2-aminobiphenyl-2,3-diol dioxygenase, and chloramphenicol acetyl transferase. J Ind Microbiol Biotechnol 31:189–197. doi:10.1007/s10295-004-0130-0
40. Pistorino M, Pfeifer BA (2009) Efficient experimental design and micro-scale medium enhancement of 6-deoxyerythronolide B production through *Escherichia coli*. Biotechnol Prog 25:1364–1371. doi:10.1002/btpr.250
41. Ren X, Yu D, Han S, Feng Y (2006) Optimization of recombinant hyperthermophilic esterase production from agricultural waste using response surface methodology. Bioresour Technol 97:2345–2349. doi:10.1016/j.biortech.2005.10.027
42. Riddle RR, Gibbs PR, Willson RC, Benedik MJ (2003) Recombinant carbazole-degrading strains for enhanced petroleum processing. J Ind Microbiol Biotechnol 30:6–12. doi:10.1007/s10295-002-0005-1
43. Rodrigues MI, Iemma AF (2005) Planejamento de experimentos e otimização de processos: uma estratégia seqüencial de planejamentos. Casa do Pão Editora, Campinas
44. Sato SI, Nam J-W, Kasuga K, Nojiri H, Yamane H, Omori T (1997) Identification and characterization of genes encoding carbazole 1,9a-dioxygenase in *Pseudomonas* sp. strain CA10. J Bacteriol 179:4850–4858
45. Shin CS, Hong MS, Bae CS, Lee J (1997) Enhanced production of human mini-proinsulin in fed-batch cultures at high cell density of *Escherichia coli* BL21(DE3)[pET-3aT2M2]. Biotechnol Prog 13:249–257. doi:10.1021/bp970018m
46. Sørensen HP, Mortensen KK (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J Biotechnol 115:113–128. doi:10.1016/j.jbiotec.2004.08.004
47. Sunitha K, Kim Y-O, Lee J-K, Oh T-K (2000) Statistical optimization of seed and induction conditions to enhance phytase production by recombinant *Escherichia coli*. Biochem Eng J 5:51–56. doi:10.1016/S1369-703X(99)00062-5
48. Sunitha K, Lee J-K, Oh T-K (1999) Optimization of medium components for phytase production by *E. coli* using response surface methodology. Bioprocess Biosyst Eng 21:477–481. doi:10.1007/PL00009086
49. Swalley SE, Fulghum JR, Chambers SP (2006) Screening factors effecting a response in soluble protein expression: formalized approach using design of experiments. Anal Biochem 351:122–127. doi:10.1016/j.ab.2005.11.046
50. Terpe K (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 72:211–222. doi:10.1007/s00253-006-0465-8
51. Urban A, Ansmant I, Motorin Y (2003) Optimisation of expression and purification of the recombinant Yol066 (Rib2) protein from *Saccharomyces cerevisiae*. J Chromatogr B 786:187–195. doi:10.1016/S1570-0232(02)00742-0
52. Volontè F, Marinelli F, Gastaldo L, Sacchi S, Pilone MS, Pollegioni L, Molla G (2008) Optimization of glutaryl-7-aminocephalosporanic acid acylase expression in *E. coli*. Protein Expr Purif 61:131–137. doi:10.1016/j.pep.2008.05.010
53. Wang Y-h, Jing C-f, Yang B, Mainda G, Dong M-l, Xu A-l (2005) Production of a new sea anemone neurotoxin by recombinant *Escherichia coli*: optimization of culture conditions using response surface methodology. Process Biochem 40:2721–2728. doi:10.1016/j.procbio.2004.12.024
54. Zhang X, Li Y, Zhuge B, Tang X, Shen W, Rao Z, Fang H, Zhuge J (2006) Optimization of 1,3-propanediol production by novel recombinant *Escherichia coli* using response surface methodology. J Chem Technol Biotechnol 81:1075–1078. doi:10.1002/jctb.1538
55. Zhao J, Wang Y, Chu J, Zhang S, Zhuang Y, Yuan Z (2008) Statistical optimization of medium for the production of pyruvate oxidase by the recombinant *Escherichia coli*. J Ind Microbiol Biotechnol 35:257–262. doi:10.1007/s10295-007-0301-x