Bioresource Technology 101 (2010) 2367–2374

Contents lists available at ScienceDirect

Bioresource Technology

journal homepage: www.elsevier.com/locate/biortech

Improvement of a two-stage fermentation process for docosahexaenoic acid production by *Aurantiochytrium limacinum* SR21 applying statistical experimental designs and data analysis

Silvina Mariana Rosa [a,b], Marcelo Abel Soria [c], Carlos Guillermo Vélez [b], Miguel Angel Galvagno [a,d,*]

[a] Instituto de Investigaciones Biotecnológicas, IIB-CONICET, Universidad Nacional de San Martín, Av. Colectora General Paz 5445, (1650) Buenos Aires, Argentina
[b] Departamento de Biodiversidad y Biología Experimental, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, (1428) Buenos Aires, Argentina
[c] Cátedra de Microbiología Agrícola, Facultad de Agronomía, Universidad de Buenos Aires, Av. San Martín 4453, (1417) Buenos Aires, Argentina
[d] Departamento de Ingeniería Química, Facultad de Ingeniería, Universidad de Buenos Aires, Pabellón de Industrias, Ciudad Universitaria, (1428) Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 4 February 2008
Received in revised form 20 February 2009
Accepted 24 October 2009
Available online 16 December 2009

Keywords:
Docosahexaenoic acid
Aurantiochytrium
Two-stage fermentation
Statistical designs
Artificial neural networks

ABSTRACT

Statistical screening experimental designs were applied to identify the significant culture variables for biomass production of *Aurantiochytrium limacinum* SR21 and their optimal levels were found using a combination of Artificial Neural Networks, genetic algorithms and graphical analysis. The biomass value obtained (40.3 g cell dry weight l^{-1}) employing the selected culture conditions agreed with that predicted by the model. Subsequently, two significant culture conditions for docosahexaenoic acid (DHA) production were determined, finding that an inoculum of 10% (v/v), obtained from the previous (statistically optimized) stage, should be used in a DHA production medium having a molar C:N ratio of 55:1, to reach a production of 7.8 g DHA l^{-1} d^{-1}. The production step was thereafter scaled in a 3.5 l bioreactor, and DHA productivity of 3.7 g l^{-1} d^{-1} was obtained. This two-stage strategy: statistically optimized inoculum production (fist step) and a DHA production step, is presented for the first time to optimize a bioprocess conducive to the obtention of microbial DHA.

1. Introduction

Docosahexaenoic acid (22:6$n-3$) (DHA) is a long-chain polyunsaturated fatty acid (PUFA) that belongs to the omega-3 group. In recent years, DHA has attracted much attention because of its health related aspects (Ratledge, 2004; Sijtsma and de Swaaf, 2004), including its importance for proper development of the brain and eye in infants (Lauritzen et al., 2001) and for inhibiting risk factors involved in cardiovascular diseases in adults (Nordoy et al., 2001). As the typical "Western" diet provides low levels of omega-3 PUFAs, the market for these fatty acids has expanded last years and their usage as dietary supplements is recommended by many milestone advisories nowadays (Ward and Singh, 2005). The traditional main sources of omega-3 fatty acids have been fatty fish, but they have some disadvantages such as variable quality, contamination by environmental pollution, unpleasant smell and taste and expensive purification (Gunstone, 2001). In this context, exploration of alternative sources, like microbial oils (single cell oils, SCO) is currently being developed (Ratledge, 2004) because they have interesting advantages: possibility to obtain oils with high and specific PUFAs content, with higher oxidative stability, production from sustainable raw materials, lower purification costs, the absence of environmental man-made pollutants (Sijtsma and de Swaaf, 2004) and a more constant product quality is possible. Among the oleaginous microorganisms, thraustochytrids – common marine microheterotrophs – can be cultured to produce high biomass, containing substantial amounts of lipids rich in PUFAs (Lewis et al., 1999). The high DHA yields obtained with the thraustochytrid *Schizochytrium* resulted in a production of low-cost oil that is used as a nutraceutical in the food and feed market (Ward and Singh, 2005).

The improvement of the bioprocess and the reductions of its cost are crucial for the expansion of the SCO market. Several factors influence the yield, productivity and total amounts of PUFAs produced. Since interactions between those process factors are expected, the application of experimental design strategies appears as an appropriate strategy for optimization (Thiry and Cingolani, 2002). Methods for experimental design can be broadly divided in two categories (Montgomery, 1991). The first one includes all those designs used for process screening, that is, the exploration of a potentially large number of input variables to discover those that are statistically significant and estimate their magnitude; fractional factorial (FFr) and Plackett–Burman (PB) designs are examples of these methods (Standbury et al., 1986).

* Corresponding author. Address: Instituto de Investigaciones Biotecnológicas, IIB-CONICET, Universidad Nacional de San Martín, Av. Colectora General Paz 5445, (1650) Buenos Aires, Argentina. Tel.: +54 11 4576 3300x222; fax: +54 11 4576 3341.
E-mail address: mag@di.fcen.uba.ar (M.A. Galvagno).

0960-8524/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.biortech.2009.11.056

Methods in the second category aim to optimize a process given a reduced number of variables. A typical scenario is the use of Box-Behnken or Central Composite designs and response surface analysis on linear or quadratic models. Additionally to traditional statistical experimental designs, other screening strategies can be employed, such as artificial neural networks (ANNs) (Kennedy and Krouse, 1999; Mutihac and Mutihac, 2008), which are models that mimic the learning ability of the brain. They are trained with sets of inputs and outputs and "learn" how to reproduce an output from the input. The main advantage of ANNs is that they can model non-linear situations without any previous knowledge of the system dynamics (Soria et al., 2004). The drawback of ANNs is that they are "black-box" models, that is, the inspection of its parameters does not yield an understanding of the influence of the input variables on the output. For this reason, the standard techniques of differentiation applied to parametric models to find maxima or minima do not apply for ANNs optimization. Instead, a number of search methods were developed, among them, the family of genetic algorithms (GA) are the most frequently used. The name of this method relates to the way it proceeds. First, the output values are calculated using the ANN model for a number of "individuals", each containing a chromosome with as many "genes" as input variables and initial random values assigned to them. Those individuals with the highest or lowest output values, depending whether the objective is finding a maximum or minimum respectively are kept for the next generation and new individuals are created by crossing over the successful chromosomes and by randomly mutating individuals from the previous generation. The process is repeated for a given number of generation or until the population reaches a maximum (or minimum). The combination of ANN and GA for process optimization has been successfully applied in biotechnology and industrial chemistry (Mutihac and Mutihac, 2008; Desai et al., 2008).

In order to achieve lipid accumulation in a microorganism, it needs to be grown in a medium with excess of carbon substrate and a limiting amount of other nutrients, usually nitrogen, (Ratledge, 2004). However, different nutritional conditions are more likely to be required for growth; for example, concentration of N source is a determining factor for cell division. In this context, the aim of the present study was to optimize DHA production by a two-stage (growth and production) fermentation with *Aurantiochytrium* (=*Schizochytrium*) (Yokoyama and Honda, 2007) *limacinum* SR21. First, statistical experimental designs and ANNs were used to define the optimal growth medium (first stage), and then different conditions for lipid accumulation in production media (second stage) were studied. The optimal conditions obtained for the production bioprocess in flask cultures were scaled up in a 3.5 l bioreactor.

2. Methods

2.1. Microorganism and maintenance

Aurantiochytrium limacinum SR21 used in this study was provided by the Institute of Fermentation of Osaka (Japan, strain number IFO 32693). The microorganism was maintained in at −70 °C in medium containing glycerol 10% v/v. Working cultures were carried out in GPY medium (% w/v: 2.0, glucose; 1.0, peptone; 0.5, yeast extract – YE; 1.75, artificial sea salt; 1.5, agar; pH 5.5) at 28 °C in an orbital shaker at 250 rpm.

2.2. Statistical experimental designs, data analysis and ANN

Plackett–Burman and fractional factorial designs were applied to screen culture variables that significantly effected biomass accumulation. These screening designs were set up for three or four factors (see Section 3) with two coded levels (−1 and +1) (Montgomery, 1991) covering different ranges, as indicated in Table 3, to evaluate their linear effect on dry cell weight (DCW). Three centre points were run for each experiment. The results were fitted with a first-order model, estimating the coefficient (slope) for each factor and its level of significance.

2.2.1. Clustering of experiments and group validation

A two-step procedure was used to determine whether the runs from the screening experiments could be pooled for optimization. In the first step individual runs from the five screening experiments (Table 3) were clustered using the K-means algorithm according to the values of the significant factors, i.e., inoculum size, glucose and corn steep liquor concentration (Hartigan and Wong, 1979). The K-means algorithm distributes samples, runs in our case, among a given number of clusters. If that number, *n*, is not know *a priori*, several values of *n* are evaluated and the cumulative within-group sums of squares penalized by *n* are calculated for each. That *n* with the lowest penalized sum of squares is selected. Due to the partial overlap of sample spaces most or all the groups resulting from the K-means algorithm, should contain runs from different experiments. Table 4 shows it was the case, four out of five experiments had their runs allocated in at least three clusters, a five out of six clusters had runs from at least two experiments. The second step of the analysis was to determine whether there was a significant difference of mean DCW values across clusters obtained by K-means. This assertion is a direct consequence of clustering, because it was done on the significant factors. However, the expected differences in DCW could be blurred if the uncontrolled variability among screening experiment was large. To test this point we performed a one-factor analysis of variance of runs using the cluster membership as the classifier and the DCW values as the dependent variable.

2.2.2. Model building and analysis

For the optimization stage, the dataset consisting of the pooled runs of five screening experiments were split into two subsets, one for model fitting or training and the other for validation. Three different models were fitted to the training dataset: linear with interactions, quadratic with interactions and ANN. The models were used to make predictions on the validation subset and for comparison of their performance, the squared root of the mean sum of squares of the errors (mse) was calculated for each model:

$$mse = \sqrt{\frac{1}{n}\sum_{i}^{n}(\text{observed}_i - \text{expected}_i)^2}$$

The statistical and numerical analyses of data were performed with R (R Development Core Team, 2007).

ANN design, training and predictions were performed as described by Soria et al. (2004) using the Neural Network Toolbox included in MATLAB ver. 6.0 (The MathWorks, Natick, MA, USA). The architecture of the ANN consisted of a feed-forward network with three layers: one input layer with three inputs (inoculum size and corn steep liquor-CSL- and glucose concentrations), one hidden layer with four neurons and one output layer that rendered the predicted DCW value. The transfer function of the neurons in the hidden layer was the hyperbolic tangent sigmoid transfer function (tansig) and the neuron in the output layer had a linear transfer function. The Bayesian regularization back-propagation (trainbr) method was used for training (Foresse and Hagan, 1997; MacKay, 1992). This training method searches the number of neurons in the hidden layer that produce good predictions while avoiding overfitting.

A genetic algorithm was applied to find a combination of glucose, CSL concentrations and inoculum size that yielded a maximum of biomass production. The Genetic Algorithm and Direct Search Toolbox of MATLAB was used with the following parameters: population type, double vector; population size, 40 individuals; elite count (number of fit individuals that survive to the next generation), 10; cross-over fraction (fraction of the 30 remaining individuals that are generated by recombination), 0.95; scale value for the variance of the first generation, 0.8. Other parameters were left to their default values. The possible values that could take the individuals of the first generation were restricted to glucose 8–12 g l^{-1}, CSL 11–13% v/v and inoculum size 5.0–6.0 $\times$ 10^6 cells ml^{-1}. These parameter and initial values were selected to thoroughly search a region that an initial exploratory analysis showed to contain a maximum. The fitness function was the trained ANN with a constraint to prevent it for searching concentrations of CSL higher than 13% and inoculum sizes higher that 6.0 $\times$ 10^6 cells ml^{-1}, because any maximum found beyond those limits would not be realizable in practice.

2.3. Culture conditions

Statistical experimental designs were carried out with five-day cultures grown in media with different nutrient concentration (see Table 3). To investigate the effect of C:N ratio on biomass and lipid production, cultures were grown in production basal medium PBM (% w/v): 10.0, glucose; 0.4, KH_2PO_4; 0.1, YE; 1.75 artificial sea salt; variable concentrations of ammonium acetate ranging from 0.17% to 2.15% (see results); pH 5.5. For optimization of DHA production by a two-stage fermentation, cells were firstly incubated in the optimized growth medium, centrifuged at 7000 rpm for 15 min and washed with 1.75% artificial sea water three times, transferred to the same volume of PBM-derived media with different C:N ratios and incubated during three days at 28 °C in an orbital shaker at 250 rpm. In all cases, cultures were carried out in 100 ml Erlenmeyer flasks containing 20 ml of culture medium.

Fermentor experiments were performed in a 5.6-l BioFlo 110 bioreactor (New Brunswick Scientific, Edison, NJ, USA). The fermentor was equipped with controllers for pH, temperature, agitation and dissolve oxygen concentration (DOC). Batch cultures were carried out in 3.5 l of production medium (composition based on results of previous shaken flask experiments). Temperature was maintained at 28 °C and the agitation speed automatically varied (from 300 to 500 rpm) at a fixed air flow rate of 1.71 vvm to maintain the DOC at 20% air saturation. To control foam formation, 30 μl antifoam l^{-1} (Antifoam289, Sigma, Saint Louis, MO, USA) was added at the beginning of the run. Samples (15 ml) for off-line determination of biomass and glucose, lipids and DHA concentration were withdrawn every 8 h until the end of the fermentation (96 h).

2.4. Analytical determinations

For biomass production estimations as dry cell weight (DCW), samples were washed with distilled water and dried at 90 °C for 1 day. Total lipids were extracted and determined gravimetrically from washed cells as described by Kates (1998). Samples for DHA measurement were prepared as described by Yokochi et al. (1998), and subjected to gas chromatography/flame ionization detector (GC/FID) (HP 5890; Hewlett–Packard Co., Palo Alto, CA, USA) on a SB five capillary column (Supelco; Bellefonte, Pennsylvania, USA) with temperature programming (180–225 °C at 5 °C min^{-1}). Residual glucose was estimated with a commercial enzymatic kit based on the glucose oxidase–peroxidase method (Wiener Lab., Argentina), residual ammonia by the indophenol method (Scheiner, 1975) and proteins concentration according to Lowry et al. (1951).

3. Results

3.1. Screening designs

The concentration of carbon and nitrogen source and their interaction are crucial for *A. limacinum* SR21 growth and DHA production, and their effect are not expected to be the same on both parameters. As previous reports found optimal conditions for DHA production in a one-stage process (Yaguchi et al., 1997; Yokochi et al., 1998), we investigated in more detail the significant variables on *A. limacinum* SR21 growth as a first step to optimize DHA production by a two-stage process.

To assess the relative importance of nutrient concentration, environmental conditions and their interaction on *A. limacinum* SR21 growth, the effect of glucose (C source), CSL (N source), pH and salinity were firstly evaluated modulating the variables according to a FFr. Since values of cell number and DCW were not coincident along the experiment, the later was chosen as growth parameter because it is more significant from a biotechnological point of view. Table 1 shows one of the FFr used, the coded and the actual levels of the independent variables and the response attained, DCW, which widely varied from 4.1 to 38.5 g l^{-1} depending on the conditions tested. Regression coefficients for codified variables and their interactions are shown in Table 2. Only the concentration of glucose and CSL and their interaction had significant effect on biomass production ($P < 0.05$). Positive coefficients for this variables means that the highest levels of biomass are produced with high concentration of glucose and CSL in the culture medium.

As other factors could affect the production of biomass too, like the inoculum size or the addition of phosphate or vitamins to the culture medium, it seemed worthwhile to include their study during the screening stage. Several "small" screening experimental designs (PB and FFr) were carried out in order to test the effect of different ranges of nutrients (C, N and P sources and vitamins), environmental conditions (pH and salinity) and inoculum size on *A. limacinum* SR21 growth (Table 3). Our results revealed that under the conditions assayed, artificial sea salt (0.35–3.5%) and pH (4–7) did not significantly affect biomass production, so they were maintained at 1.75% concentration and 5.5, respectively in the following experiments. Within the range of tested concentrations for YE (0.1–0.8%) and KH_2PO_4 (0.1–1%), as supplements of vitamins or phosphate, respectively, no significant differences were observed by including these variables, so they were not investigated further. On the other hand, concentration of glucose and CSL, covering 0.5–13.5% and 0.5–15% ranges, respectively, in the different experiments performed, had both a positive significant effect on biomass production, except for glucose in FFr2 that showed a negative effect, possibly caused by a hyperosmotic effect. Positive significance of glucose-CSL interaction suggested an effect of the C:N ratio on biomass accumulation. Inoculum size was maintained at approximately 10^6 cells ml^{-1} in all experiments, except in FFr2, in which it varied between 10^3 and 10^6 cells ml^{-1} and presented a significant and positive effect.

3.2. Data analysis for process optimization

When the screening stage of a process development consists of one or two screening designs, the recommended course of action for the optimization stage is to apply other types of experiments, like central composite and Box-Behnken designs. In our case, we had performed five experiments in which the range of at least

Table 1
Fractional factorial screening design (FFr1) and response (DCW) for biomass production by *A. limacinum* SR21.

Trial	Factor level*								DCW
	Glucose (X_1)		pH (X_2)		CSL (X_3)		Artificial sea salt (X_4)		
1	+1	9.0	+1	7.0	−1	1.1	−1	0.35	6.30
2	+1	9.0	+1	7.0	+1	11.0	−1	3.50	38.36
3	−1	2.0	+1	7.0	−1	1.1	−1	0.35	5.34
4	0	5.5	0	5.7	0	6.0	0	1.92	21.66
5	−1	2.0	+1	7.0	+1	11.0	−1	0.35	6.40
6	+1	9.0	+1	7.0	+1	11.0	+1	3.50	38.56
7	−1	2.0	−1	4.5	+1	11.0	+1	3.50	9.20
8	+1	9.0	−1	4.5	+1	11.0	−1	0.35	32.40
9	+1	9.0	−1	4.5	−1	1.1	−1	0.35	5.16
10	−1	2.0	+1	7.0	+1	11.0	+1	3.50	8.28
11	−1	2.0	−1	4.5	−1	1.1	−1	0.35	5.04
12	−1	2.0	−1	4.5	−1	1.1	+1	3.50	4.14
13	+1	9.0	−1	4.5	+1	11.0	+1	3.50	32.28
14	0	5.5	0	5.7	0	6.0	0	1.92	20.70
15	+1	9.0	−1	4.5	−1	1.1	+1	3.50	8.94
16	−1	2.0	−1	4.5	+1	11.0	−1	0.35	6.12
17	−1	2.0	+1	7.0	−1	1.1	+1	3.50	4.18
18	0	5.5	0	5.7	0	6.0	0	1.92	21.44
19	+1	9.0	+1	7.0	−1	1.1	+1	3.50	9.10

For each factor, numbers in the first column are the coded values.
* Glucose and artificial sea salt in% (w/v), CSL in% (v/v) and DCW in g l^{-1}. About 3.5% artificial sea salt is equivalent to the mean sea salt concentration.

Table 2
Analysis of a Fractional Factorial screening design (FFr1) for biomass production (DCW) by *A. limacinum* SR21.

Term	Regression coefficient	Standard error	*P*-value
Intercept	14.926	1.588	0.00255**
X_1	7.650	1.730	0.02149*
X_2	0.827	1.730	0.66518
X_3	7.712	1.730	0.02102*
X_4	0.597	1.730	0.75266
X_1: X_2	0.865	1.730	0.65151
X_1: X_3	6.300	1.730	0.03572*
X_2: X_3	0.622	1.730	0.74287
X_1: X_4	0.235	1.730	0.90057
X_2: X_4	−0.132	1.730	0.94378
X_3: X_4	0.032	1.730	0.98619
X_1: X_2:X_3	0.745	1.730	0.69586
X_1: X_2:X_4	0.050	1.730	0.97876
X_1: X_3:X_4	−0.845	1.730	0.65878
X_2: X_3:X_4	0.022	1.730	0.99044
X_1: X_2:X_3:X_4	0.140	1.730	0.94061

Codified values for: X_1, glucose; X_2, pH; X_3, CSL; X_4, artificial sea salt; ":" means interaction among factors.
Residual standard error: 6.921 on three degrees of freedom.
Multiple R-squared: 0.9472; adjusted R-squared: 0.6834.
* Significance code: $P < 0.05$.
** Significance code: $P < 0.005$.

one-factor overlapped to some extent with the range of the same factor in another experiment. As a consequence, we already counted with a relatively large number of data points covering the experimental space that could be pooled together to fit to a model and search for optimal points. However, the variations among experiments and the fact that the overlapping of sampling spaces was only partial could preclude the pooling of the data. To test this point a two-step procedure was carried out that combined a K-means clustering with an analysis of variance (for details, see Section 2). The analysis showed that runs from different screening experiments with similar values of inoculum size, glucose and CSL concentrations could be grouped into a single cluster and that the differences in DCW among clusters were statistically different ($P < 10^{-6}$), indicating that we could safely pool the runs from the screening designs into a single dataset for process optimization (Table 4).

To find the actual optimal values of inoculum size, concentration of glucose and concentration of CSL, we split the dataset into a training subset and a validation subset to fit linear and quadratic regression and ANN models on the former subset and made predictions on the latter (see Section 2). The mse values for the linear, quadratic and ANN were 5.8, 4.2 and 2.6, respectively. In consequence, the trained ANN model was used to simulate the biomass produced by 5172 combinations of glucose and CSL concentrations and inoculum sizes. An exploratory analysis showed that biomass

Table 3
Experiments to screen the effect of different variables on biomass production by *A. limacinum* SR21.

Experiment code	Significant factors						Other assayed non-significant factors
	Glucose		CSL		Log (inoculum)		
	Range (%)	Slope	Range (%)	Slope	Log (cells ml^{-1})	Slope	
PB1	0.5–3.5	3.181***	0.5–2.0	−0.028	6.2	–	Artificial sea salt (0.35–3.15%), YE (0.1–0.8%)
PB2	2–9	1.128*	0.5–5.0	0.07**	5.9	–	Artificial sea salt (0.70–3.50%), pH (4–7)
PB3	2–10	10.201*	3–14	7.451*	6.3	–	Artificial sea salt (0.70–3.50%)
FFr1	2–9	7.650*	1–11	7.712*	6.3	–	Artificial sea salt (0.35–3.50%), pH (4.5–7.0)
FFr2	6.5–13.5	−2.778*	2.5–15.0**	4.941**	3.0–6.0	3.247*	KH_2PO_4 (0.1–1.0%)

Screening experiments with Plackett–Burman (PB) or factorial fractional (FFr) designs were run to test the effect of several variables on biomass accumulation. Ranges of variation are shown for tested factors, and fixed values otherwise.
* Significance code: $P < 0.05$.
** Significance code: $P < 0.005$.
*** Significance code: $P < 0.0005$.

Table 4
Pooling of data from screening experiments.

		Cluster					
		1	2	3	4	5	6
Runs per cluster		8	10	26	13	12	4
Number of runs contributed by experiment	PB1			13			
	PB2		3	6	4		
	PB3	4	3	3		3	
	FFr1	4		4	5	4	
	FFr2		4		4	5	4
Mean DCW (SE)		14.0 (3.0)	25.3 (2.0)	8.4 (0.8)	13.6 (2.5)	34.8 (2.6)	38.4 (1.9)

To determine the validity of merging data points obtained from different experiments, runs from the screening designs were clustered in six groups using the K-means algorithm. Two first rows of the table shows the number of treatments discriminated by experiment that were allocated to each cluster. The last row shows the mean dry cell weight (DCW) for the cluster and its SE. Differences in DCW were significant at $P < 10^{-6}$.

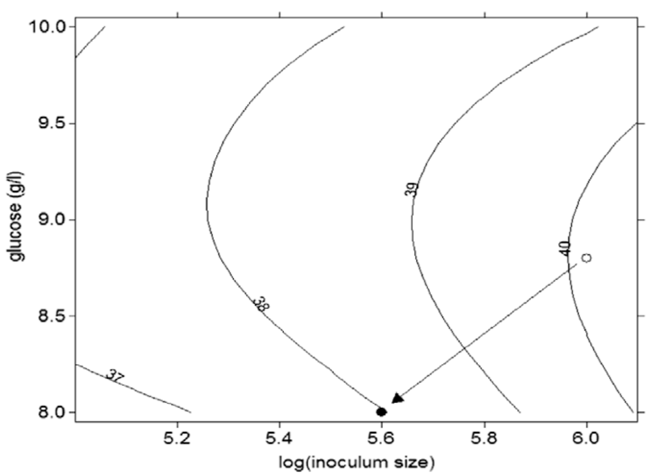

Fig. 1. Culture media optimization. The contour plot was produced with a subset of 336 predictions with CSL concentration fixed at 12% v/v out of the total 5172 simulations done the trained ANN model out. The white circle shows the point of maximum biomass determined by the genetic algorithm and the black circle is the operating point chosen after taking into account the feasibility of preparing highly concentrated inocula.

production was strongly and positively correlated with CSL concentration. However, on one hand, preparing media with high concentration of CSL was difficult. On the other, our previous experiments showed that increasing CSL levels did indeed led to an increase in biomass, but with a reduced content of lipids, which would negatively affect the second phase of omega acids accumulation. So, a concentration of 12% CSL was set as the upper limit for the search step.

To find a combination of critical media components that would yield a maximum value of biomass we used a GA coupled to the trained ANN (for computing details, see materials and methods). The GA predicted a maximum biomass production of 42.1 g l^{-1} with 8.8 g l^{-1} glucose, 12% CSL and a inoculum of 1×10^6 cells ml^{-1} (white circle in Fig. 1). For routine work, a lower inoculum size around 4×10^5 cells ml^{-1} is easier to attain, the contour plot in Fig. 1 shows that this lower inoculum lies on the isoline of 38 g l^{-1} biomass. The plot also indicates that the glucose concentration should be decreased to 8 g l^{-1}.

3.3. Validation

Fig. 2 depicts the growth profile in the optimized Growth Medium (GM; composition: 8.0% (w/v) glucose, 12.0% (v/v) CSL, 1.75% (w/v) artificial sea salt; 6:1 C:N ratio). Taking into account cell division as a measurement of growth, the stationary phase was reached at 40 h with a specific growth rate (μ_{max}) of 0.11 h^{-1} although cell biomass continued increasing up to c.a. 120 h incubation. High glucose consumption for accumulation of high levels of biomass and lipids were found from 100 h culture onwards and 40.33 g l^{-1} DCW was obtained at 120 h, a value very close to the 38 g l^{-1} prediction and given the variability between replicates, comparable to the predicted maximum, 42.1 g l^{-1}. Lipid accumulation was coincident with ammonium depletion and high glucose concentration in the culture medium (high C:N ratio).

3.4. Conditions for DHA production by two-stage fermentation

In the following two-stage experiments, cultures were incubated in GM (first stage), washed and then transferred to production medium to accumulate DHA (second stage). Because of the metabolic status of inoculum could be another significant variable on production, the influence of first stage cells at different growth stages (exponential or stationary) was investigated. For this purpose, two-day and five-day GM-grown cultures were washed and transferred to PBM-based media varying the ammonium acetate concentrations to reach different (10:1 to 100:1) C:N ratios an incubated for 3 days (Table 5). Previous experiments had revealed a similar effect of 100:1 and 150:1 C:N culture media ratio and 3 or 5 day incubation time (data not shown). A centre point (three and a half-day GM culture as inoculum and production medium with 55:1 C:N ratio) was run for statistical analysis. Table 5 shows that whereas inoculum age did not have any effect on biomass, lipids nor DHA production, the C:N ratio of production medium had a significant and positive effect ($P < 0.005$) on the three assayed responses. The maximum DHA concentration, 7 g l^{-1}, was obtained in PBM with C:N 55:1 (Production Medium – PM).

To eliminate GM medium that could cause an inhibitory effect on DHA production in PM, GM-grown cells were centrifuged, washed three times with sterile artificial sea water and transferred to the original volume of medium PM in previous assays. Since this process would not be the most convenient at high fermentation scale, alternative transfer processes were investigated. Washing of centrifuged GM-grown cells before resuspension in the same volume of PM (complete transference) had no significant effect on biomass, lipids or DHA production compared to non-washed cells (Table 6), indicating that removal of GM was enough to get high DHA concentration (c.a. 9.5 g l^{-1}). Another alternative lower consuming energy process to avoid the possible inhibitory effect of GM could be the dilution of GM culture in PM. Our results showed that although this partial transference (10%) of GM-grown cells to PM resulted in slightly lower biomass and DHA production than a full transference (100%) (Table 6), DHA concentration was still high (7.75 g l^{-1}). As this procedure was far economic and easier to perform than the centrifugation step, it was chosen for next experiments.

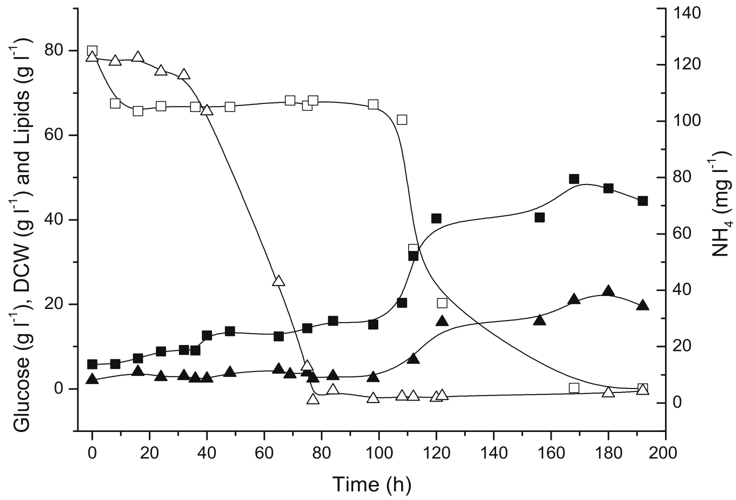

Fig. 2. Growth profile of *A. limacinum* SR21 in 500 ml Erlenmeyer flasks with 100 ml of Growth Medium. Time variation is shown for residual glucose (□) and ammonium (NH_4) (△) and for biomass (measured as DCW) (■) and lipids (▲) production.

Table 5
Effect of the inoculum age and C:N ratio of PM on biomass (DCW), lipid and DHA production by *A. limacinum* SR21.

C:N ratio	Inoculum age (d)	DCW (g l^{-1})	Lipids (g l^{-1})	DHA (g l^{-1})
10:1	2	5.20 ± 0.01	1.10 ± 0.42	0.14 ± 0.01
10:1	5	9.65 ± 0.49	0.75 ± 0.21	0.29 ± 0.03
55:1	3.5	45.80 ± 0.85	29.25 ± 2.47	7.00 ± 0.50
100:1	2	42.05 ± 0.92	29.40 ± 0.71	5.00 ± 0.50
100:1	5	44.05 ± 1.20	28.50 ± 0.75	6.50 ± 0.50

Mean values and SD for DCW, lipid and DHA concentrations obtained with different C:N ratio media and inoculum age. PM contained 10.0% glucose and a variable concentration of ammonium acetate (2.15%, 0.39% and 0.17%) to obtained the desire C:N ratio (10:1, 55:1 and 100:1, respectively).

Optimized flask conditions were conducted in a stirred tank fermentor. For this, 350 ml of a three-day growing GM culture were transferred to 10 times volume of PM in a bioreactor. Biomass and lipids accumulated from 24 h up to 72 h incubation, reaching values of 51.95 and 35.4 g l^{-1}, respectively, while glucose was completely consumed by this time (Fig. 3). DHA concentration increased up to 72 h and maintained up to 96 h culture, resulting in a maximal DHA productivity of 3.7 g l^{-1} d^{-1}. As we could not find a significant increase in cell number (data not shown), DCW and lipid production were attributed to cell content increase.

4. Discussion

Large-scale application of microbial DHA in human nutrition depends on the quality and the production costs of the SCO (Sijtsma and de Swaaf, 2004). Thraustochytrids are a new and potentially competitive player in the DHA rich products market, but to achieve this aim, some key questions, among others, need to be addressed: the screening of high DHA-producing strains, improvement of downstream processes and manipulation of culture conditions to optimize PUFA production (Lewis et al., 1999). Although there are several works in which nutrient levels (mainly C and N sources) and culture conditions (temperature, pH and salinity) were manipulated to optimize DHA production in *Schizochytrium* and *Aurantiochytrium* spp. (Bowles et al., 1999; Iida et al., 1996; Perveen et al., 2006; Wu and Lin, 2003; Yokochi et al., 1998) the effect of these variables was only evaluated on DHA accumulation, not discriminating their differential influence on growth. Considering that different nutritional requirement for growth and lipids production are expected (Anderson and Wynn, 2001), we have investigated separately the factors that influence both processes in the strain *A. limacinum* SR21, and designed a two-stage fermentation to accumulate high DHA levels.

Our results confirm that the concentrations of C and N-sources and their interactions are significant variables for biomass and lipid accumulation. The C:N ratio influence on growth and lipid accumulation can be inferred from the fermentation profile in GM (Fig. 1): cells divided actively and enlarged at low C:N ratios, and only after the N source was depleted, and the ratio increased, lipids were accumulated. As stimulation of DHA production in thraustochytrids by high C:N ratio was suggested previously by other authors (Bowles et al., 1999; Wu and Lin, 2003; Yokochi et al., 1998), the effect of this relationship in media with high C source concentration was thoroughly tested, concluding that a C:N ratio of at least 55:1 was suitable to obtain high DHA yields. Additionally, use of CSL as N-source, a low-cost substrate, avoided

Table 6
Effect of the GM culture treatment and size before transfer to PM on biomass (DCW), lipid and DHA production by *A. limacinum* SR21.

Treatment before transference to PM Centrifugation	 Washes	DCW (g l^{-1})	Lipids (g l^{-1})	DHA (g l^{-1})
Full (100%) transference				
Yes	Yes	43.70 ± 0.28	26.60 ± 0.71	9.50 ± 1
Yes	No	44.95 ± 0.78	26.90 ± 0.42	9.75 ± 1
Partial (10%) transference				
No	No	37.80 ± 0.01	25.65 ± 0.07	7.75 ± 1

Mean values and SD for DCW, lipid and DHA concentrations obtained in two-stage cultures, transferring 100% or 10% of GM-grown cells to PM under different transference treatments.

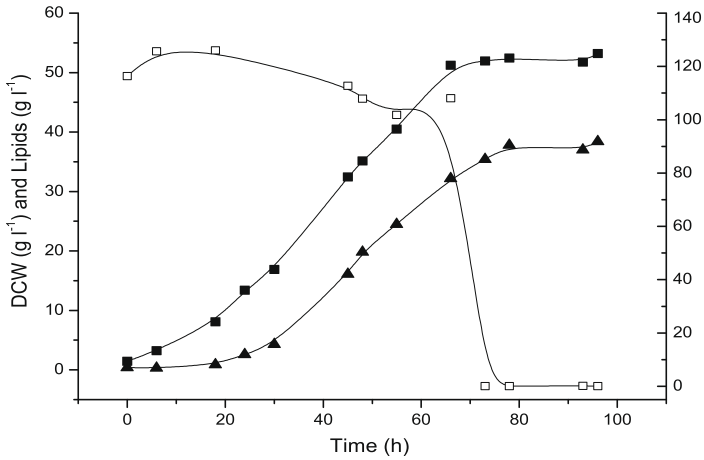

Fig. 3. Growth profile of *A. limacinum* SR21 in a 3.5 l bioreactor with Production Medium. Time variation is shown for residual glucose (□) and for biomass (measured as DCW) (■) and lipids (▲) production.

the addition of a source of vitamins (YE) to GM, lowering the production costs. On the other hand, physicochemical conditions in the ranges assayed in this work did not have any significant influence on biomass production, allowing working with low salt concentration levels and a wide range of pH. It is not surprising considering that *A. limacinum* SR21 was isolated from an estuarine environment, that is, a biome that experiences widely varying conditions of temperature, salinity and in the concentrations of a wide variety of chemicals (Constanza et al., 1993).

Inoculum preparation was the other variable which resulted fundamental for DHA production by *A. limacinum* SR21. We found that inoculum size had a significant effect on biomass production. Direct inoculation of PM with cells that were not previously propagated in GM produced very low DCW (i.e. lipid concentration), because they failed to divide (results not shown). In contrast, high lipid accumulation rates were obtained with three-day GM cultures inoculated at 10% v/v in PM (see Table 6), resulting in higher levels of DHA (11 g l^{-1}) in a shorter production period of time (three days), thus increasing DHA productivity.

Statistical experimental designs had been applied for DHA production improvement with the strain *Schizochytrium* sp S31 previously (Wu and Lin, 2003). Screening and optimization designs performed in that work suggested that conditions for maximizing DHA production (0.516 g l^{-1}) were glucose 27.98 g l^{-1}, YE 4.52 g l^{-1}, sodium chloride 24.82 g l^{-1}, pH 6.96 and four-day incubation period. Although we have studied the effect of similar variables on biomass in *A. limacinum* SR21 applying screening designs firstly, other alternative statistical tools were employed to follow the analysis. After validating the pooling of data obtained from experimental designs, we used ANN, linear and quadratic models to fit the data and, as expected from the literature and our previous experience, the best fit was achieved with ANNs. Probably, the clustering was satisfactory because in spite of using data from different experiments with their expected biological variability, several factors contributed to reduce overall variability. The application of ANNs allowed us to predict very accurately optimal conditions without carrying out optimization experiments, saving effort and time, and supporting the high predictable value of this methodology for medium design when large amounts of data are available.

A previous study in shake flask cultures showed that *A. limacinum* SR21 reached a maximum productivity of 0.84 g DHA l^{-1} d^{-1} with 2% CSL, 1.75% artificial sea salt and 9% glucose, or 12% glycerol media (Yokochi et al., 1998). Our results demonstrated that a productivity of 2.58 g DHA l^{-1} d^{-1} could be obtained if the same strain was first grown for three days in GM, a medium with low (10:1) C:N ratio, transferred to ten times volume of PM with higher (55:1) C:N ratio and then incubated three days more. This approach, lowering incubation time in PM by preparing a convenient inoculum, tripled DHA productivity, which is comparable with the highest value reported for strain 12B, 2.8 ± 0.7 DHA g l^{-1} d^{-1}, a thraustochytrid-like microorganism with a higher growth rate (μ_{max} of 0.38 h^{-1}, Perveen et al., 2006, vs. 0.11 h^{-1} in this study) than *A. limacinum* SR21. Validation of optimized shake flask culture conditions in a bioreactor resulted in a maximum productivity of 3.7 g DHA l^{-1} d^{-1}, which was similar to the 3.3 g DHA l^{-1} d^{-1} obtained by Yaguchi et al. (1997) with 3 l fermentor cultures. Although as general statement shake flask conditions can be improved in bioreactor (Unagul et al., 2007), it required knowing and manipulating significant variables at this scale. Bailey et al. (2003) identified oxygen concentration as a critical factor for high DHA productivity from high density thraustochytrids cultures. In that sense, our results are encouraging and they could be improved by investigating other potentially important factors operating at higher production scale, like oxygen availability during the production stage.

5. Conclusions

Optimization of DHA production by the strain *A. limacinum* SR21 was achieved by a two-stage bioprocess. A combination of statistical screening experimental designs and non-linear data analysis techniques (including ANN, genetic algorithms and graphical analysis) were successfully applied for the first time to this strain as an effective tool to find the optimal culture conditions for the first, or growth, stage. This approach is very useful for a bioprocess in which several number of cultures variables must be taken into account in order to obtain optimal microbial biomass in culture. Under the selected conditions, 8% (w/v) glucose, 12% (v/v) Corn Steep Liquor (CSL) and an inoculum of 4×10^5 cells ml^{-1}, a dry cell weight of 40.3 g l^{-} was obtained, after five days incubation; this experimental value did not differ significantly (less than 10%) with the predicted by the model used. The improved flask culture conditions for the second, or DHA production, stage were: 10% v/v inoculum (obtained from the first stage) pitched in a production medium containing glucose and ammonium acetate at a 55:1 C:N ratio (i.e. more than five times higher than in the growth medium). This later stage was run in a 3.5 l bioreactor, and a DHA productivity of 3.7 g l^{-1} d^{-1} was reached. This study demonstrates by the first time that DHA productivity by *A. limacinum* SR21 could be improved by a two-stage bioprocess, i.e. firstly statistically optimizing the production of cell biomass by the combination of two methods, and then using thE biomass obtained as inoculum for PUFAs production. At this second stage, inoculation conditions (concentration and pre-treatment of inoculum culture) and C:N ratio in the medium resulted to be the most significant variables studied for DHA production in thraustochytrids.

Acknowledgements

We are grateful to Lic. Ruben O. Fernandez (CONEA-Argentina) for helping us with DHA samples preparation and measurement by GC/FID, and to Dr. Pablo I. Nikel for collaborating in biorreactor assays. SMR is graduate fellow of CONICET and MAG is a CONICET researcher.

References

Anderson, A., Wynn, J., 2001. Microbial polyhydroxyalkanoates, polysaccharides and lipids. In: Ratledge, C., Kristiansen, B. (Eds.), Basic Biotechnology. Cambridge University Press, Cambridge, pp. 325–348.

Bailey, R.B., DiMasi, D., Hansen, J.M., Mirrasoul, P.J., Ruecker, C.M., Veeder, G.T., Kaneko, T., Barclay, W.R., 2003. In: "United States Patent". 6607900, USA.

Bowles, R.D., Hunt, A.E., Bremer, G.B., Duchars, M.G., Eaton, R.A., 1999. Long chain $n-3$ polyunsaturated fatty acid production by members of the marine protistan group the thraustochytrids: screening of isolates and optimisation of docosahexaenoic acid production. J. Biotechnol. 70, 193–202.

Constanza, R., Kemp, M., Boynton, W., 1993. Predictability, scale and biodiversity in coastal and estuarine ecosystems: implication for management. Ambio 22, 88–96.

Desai, K.M., Survase, S.A., Saudagar, P.S., Lele, S.S., Singhal, R.S., 2008. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: case study of fermentative production of scleroglucan. Biochem. Eng. J. 41, 266–273.

Foresse, D.F., Hagan, M.T., 1997. Gauss-Newton approximation for Bayesian learning. In: 1997 Proc. Int. Joint Conf. Neural Networks, pp. 1930–1935.

Gunstone, F.D., 2001. Basic oleochemicals, oleochemical products and new industrial oils. In: Gunstone, F.D., Hamilton, R.J. (Eds.), Oleochemical Manufacture and Application. Sheffield Academic Press, Sheffield, pp. 1–22.

Hartigan, J.A., Wong, M.A., 1979. A K-means clustering algorithm. Appl. Stat. 28, 100–108.

Iida, I., Nakahara, T., Yocochi, T., Kamisaka, Y., Yagi, H., Yamaoka, M., Suzuki, O., 1996. Improvement of docosahexaenoic acid production in a culture of *Thraustochytrium aureum* by medium optimization. J. Ferment. Bioeng. 81, 76–78.

Kates, M., 1998. Techniques of lipidology. Isolation, analysis and identification of lipids. In: Burdon, R.H., van Knippenberg, P.H. (Eds.), Laboratory techniques in biochemistry and molecular biology. Elsevier, Amsterdam, pp. 100–111.

Kennedy, M., Krouse, D., 1999. Strategies for improving fermentation medium performance. a review. J. Ind. Microbiol. Biotechnol. 23, 456–475.

Lauritzen, L., Hansen, H.S., Jorgensen, M.H., Michaelsen, K.F., 2001. The essentiality of long chain $n-3$ fatty acids in relation to development and function of the brain and retina. Prog. Lipid Res. 40, 1–94.

Lewis, T.E., Nichols, P.D., McMeekin, T.A., 1999. The biotechnological potential of thraustochytrids. Mar. Biotechnol. 1, 580–587.

Lowry, O.H., Rosbrough, N.J., Farr, A.L., Randall, R.J., 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. 193, 265–275.

MacKay, D., 1992. Bayesian Interpolation. Neural Comput. 4, 415–447.

Montgomery, D.C., 1991. Design and analysis of experiments. John Wiley, New York.

Mutihac, L., Mutihac, R., 2008. Mining in chemometrics. Anal. Chim. Acta 612, 1–18.

Nordoy, A., Marchioli, R., Arnesen, H., Videbaek, J., 2001. $n-3$ polyunsaturated fatty acids and cardiovascular diseases. Lipids 36, 127–129.

Perveen, Z., Ando, H., Ueno, A., Ito, Y., Yamamoto, Y., Yamada, Y., Takagi, T., Kaneko, T., Kogame, K., Okuyama, H., 2006. Isolation and characterization of a novel thraustochytrid-like microorganism that efficiently produces docosahexaenoic acid. Biotechnol. Lett. 28, 197–202.

Ratledge, C., 2004. Fatty acid biosynthesis in microorganisms being used for single cell oil production. Biochimie 86, 807–815.

Scheiner, D., 1975. Determination of ammonia and kjeldhal nitrogen by indophenol method. Water Res. 10, 31–36.

Sijtsma, L., de Swaaf, M.E., 2004. Biotechnological production and applications of the omega-3 polyunsaturated fatty acid docosahexaenoic acid. Appl. Microbiol. Biotechnol. 64, 146–153.

Soria, A.M., Gonzalez Funes, J.L., Garcia, A.F., 2004. A simulation study comparing the impact of experimental error on the performance of experimental designs and artificial neural networks used for process screening. J. Ind. Microbiol. Biotechnol. 31, 469–474.

Standbury, P.F., Whitaker, A., Hall, S.J., 1986. Media for industrial fermentations. In: Standbury, P.F., Whitaker, A., Hall, S.J. (Eds.), Principles of Fermentation technology. Pergamon, Oxford, pp. 93–122.

The MathWorks, Inc., 2004. Genetic Algorithm and Direct Search Toolbox User's Guide.

Thiry, M., Cingolani, D., 2002. Optimizing scale-up fermentation processes. Trends Biotechnol. 20, 103–105.

Unagul, P., Assantachai, C., Phadungruengluij, S., Suphantharika, M., Tanticharoen, M., Verduyn, C., 2007. Coconut water as a medium additive for the production of docosahexaenoic acid (C22:6 n3) by *Schizochytrium mangrovei* Sk-02. Biores. Technol. 98, 281–287.

Ward, O.P., Singh, A., 2005. Omega-3/6 fatty acids: alternative sources of production. Process Biochem. 40, 3627–3652.

Wu, S.-T., Lin, L.-P., 2003. Application of response surface methodology to optimize docosahexaenoic acid production by *Schizochytrium* sp. S31.. J. Food Biochem. 27, 127–139.

Yaguchi, T., Tanaka, S., Yocochi, T., Nakahara, T., Higashihara, T., 1997. Production of high yields of docosahexaenoic acid by *Schizochytrium* sp. strain SR21. J. Am. Oil Chem. Soc 74, 1431–1434.

Yokochi, T., Honda, D., Higasshihara, T., Nakahara, T., 1998. Optimization of docosahexaenoic acid production by *Schizochytrium limacinum* SR21. Appl. Microbiol. Biotechnol. 49, 72–77.

Yokoyama, R., Honda, D., 2007. Taxonomic rearrangement of the genus *Schizochytrium* sensu latu based on morphology, chemotaxonomic characteristics, and 18S rRNA gene phylogeny (Thraustochytriaceae, Labyrinthulomycetes): emendation for *Schizochytrium* and erection of *Aurantiochytrium* and *Oblongichytrium* gen. Mycoscience 48, 199–211.