Electronic Journal of Biotechnology ISSN: 0717-3458
http://www.ejbiotechnology.info
DOI: 10.2225/vol16-issue4-fulltext-2 *RESEARCH ARTICLE*

Analysis of batch and repeated fedbatch productions of *Candida utilis* cell mass using mathematical modeling method

Yutian Lin[1] · Guoli Liu[2,3] · Huibin Lin[4] · Ling Gao[5] ✉ · Jianqiang Lin[2]

1 Zhejiang University, Faculty of Information Technology, Hangzhou, China

2 Shandong University, State Key Laboratory of Microbial Technology, Jinan, China

3 Research Center of Microbial Engineering, Liaoning Academy of Agricultural Sciences, Shenyang, China

4 Shandong Academy of Chinese Medicine, Jinan, China

5 Shandong Normal University, Shandong Provincial Key Lab for Distributed Computer Software Novel Technology, School of Information Science & Engineering, Jinan, China

✉ Corresponding author: gaoling33@hotmail.com
Received October 26, 2012 / Accepted March 21, 2013
Published online: July 15, 2013

Abstract

Background: *Candida utilis* is widely used in bioindustry, and its cell mass needs to be produced in a cost effective way. Process optimization based on the experimental results is the major way to reduce the production cost. However, this process is expensive, time consuming and labor intensive. Mathematical modeling is a useful tool for process analysis and optimization. Furthermore, sufficient information can be obtained with fewer experiments by using the mathematical modeling, and some results can be predicted even without doing experiments.
Results: In the present study, we performed the mathematical modeling and simulation for the cell mass production of *Candida utilis* based on limited batch and repeated fedbatch experiments. The model parameters were optimized using genetic algorithm (GA), and the processes were analyzed.
Conclusions: Taken together, this newly developed method is efficient, labor saving and cost effective.

Keywords: *Candida utilis*, fermentation, genetic algorithm, mathematical model.

INTRODUCTION

Candida utilis is widely used in bioindustry, such as biochemical product production (Liang et al. 2009; Tamakawa et al. 2012), enzyme production (Yokoyama et al. 1988; Sengupta et al. 2011), and waste water treatments (Elmaleh et al. 1999; Zu et al. 2006) and so on. Therefore, its production in an economic way is important for the applications above mentioned. During the growth of *Candida utilis*, nutrients are continuously consumed and their concentrations are dynamically changing, leading to the limitations for the cell growth. The cell growth and yield can be systematically increased through experimental process optimizations. However, these processes are expensive, time-consuming and labor-intensive. Previous studies have proved that mathematical modeling and simulation are useful tools in bioprocess analysis, optimization and automatic control (Lin et al. 2005; Wang et al. 2009), especially under the condition of shortages of the probes for on-line fermentation monitoring. Moreover, sufficient information can be obtained from limited number of experiments by using the mathematical modeling and simulation. In the present study, we performed the mathematical modeling and simulation for the cell mass production of *Candida utilis* based on limited batch and repeated fedbatch experiments. Our data demonstrated the usefulness and efficiency of the mathematical modeling

method in the bioprocess analysis, optimization and prediction for improvements of bioprocess in an efficient and cost effective way.

MATERIALS AND METHODS

Strain, medium and fermentation conditions

Previously stored *Candida utilis* I (a mutant of *C. utilis* CICC1769) was used in the present study. The fermentation medium (1 l) consisted of 100 g glucose, 10 g $(NH_4)_2SO_4$, 0.5 g $MgSO_4$ x $7H_2O$, 1 g KH_2PO_4, 20 mg $CaCl_2$ x $2H_2O$, 20 mg $FeSO_4$ x $7H_2O$, 10.6 mg $ZnSO_4$ x $7H_2O$, 2 mg $MnCl_2$ x $4H_2O$, 0.5 mg $CuSO_4$ x $5H_2O$. For the fedbatch culture, the initial glucose concentration was 90 g/l. Briefly, 0.45 l seed culture was inoculated into the fermentor containing 4.5 l culture medium. The temperature was maintained at 30ºC, pH was controlled at 6 by using NaOH (10 M), aeration rate was set at 1 v/v/min, and agitation speed was 600 r/min.

In the fedbatch culture, concentrations of glucose, NH_4^+, and PO_4^{3-} were on-line analyzed and feedback controlled. Feeding solutions contained the same compositions as the fermentation medium except that the glucose was saturated for the glucose feeding solution, $(NH_4)_2SO_4$ was 500 g/l for the NH_4^+ feeding solution, and KH_2PO_4 was 100 g/l for the PO_4^{3-} feeding solution. The setpoints used in the feedback control of the feeding solutions were as follows: glucose 10 g/l from 15 to 30 hrs and 5 g/l from 30 to 46 hrs, NH_4^+ 70 mmol/l, PO_4^{3-} 4 mmol/l from 15 to 28 hrs and 3 mmol/l from 33 to 46 hrs. The feedings started at 15 hrs of the culture.

Equipments and analyzing methods

Fermentation was carried out using a 5 l BioFlo 310 type fermentor (New Brunswick Scientific Co., USA) equipped with pH, dissolved oxygen (DO) (Mettler Toledo Co., Switzerland) and temperature sensors. BioProfile 300A Biochemical analyzer together with the feedback control box (Nova Biomedical Co., USA) was used for off-line (for batch culture) and on-line glucose, NH_4^+, and PO_4^{3-} concentration analyses and on-line feedback controls. Cell-free sampling probe (Flownamics, USA) was used for on-line sampling. OD_{600} measurement was determined using a type 722 spectrophotometer (Shanghai Yuefeng Instruments & Meters Co., China). Dry cell weight was determined as follows. Two sets of triple samples were prepared. One was centrifuged to remove the supernatant, washed and centrifuged for three times. The sediment was dried at 55ºC until its weight became constant. Another set of triple samples was diluted for 200, 300, 400, 500, 600 and 800 times, respectively, with the OD_{600} ranging from 0.2~0.8. The relationship between dry cell weights and OD_{600} value was investigated by the linear regression analysis.

Mathematical modeling

Following dynamic equations were used in the present study:

$$\mu = \frac{\mu_m \cdot S}{k_S + S} \cdot (1 - \frac{X}{X_m})$$

[Equation 1]

$$q_S = \frac{\mu}{Y_{X/S}} = \frac{\mu}{Y_G} + m_S$$

[Equation 2]

$$q_N = \frac{\mu}{Y_{X/N}}$$

[Equation 3]

$$q_P = \frac{\mu}{Y_{X/P}}$$

[Equation 4]

Where, μ is the specific growth rate (1/h); μ_m is the maximum specific growth rate (1/h); S is the glucose concentration (g/l); X is the cell concentration (g/l); X_m is the maximum cell concentration (g/l); k_S is the saturation constant (g/l); q_S (g/g/h), q_N (mmol/g/h) and q_P (mmol/g/h) are the specific consumption rates of glucose, NH_4^+, and PO_4^{3-}; m_S is the maintenance coefficient defined by the amount of glucose consumed per gram of cells per hour for maintenance (g/g/h); $Y_{X/S}$ is the cell yield from glucose (g/g); Y_G is the theoretical cell yield from glucose (g/g); $Y_{X/N}$ is the cell yield from NH_4^+ (g/mmol); $Y_{X/P}$ is the cell yield from PO_4^{3-} (g/mmol).

μ was modeled by the equation with the combination of Monod type equation and Logistic equation, which can simulate the exponential growth and the limited growth controlled by both glucose and cell concentrations. In modeling glucose consumption, both cell growth and maintenance were considered.

Mass balance equations were used as follows:

$$\frac{dX}{dt} = \mu \cdot X = \frac{\mu_m \cdot S}{k_S + S} \cdot (1 - \frac{X}{X_m}) \cdot X$$ **[Equation 5]**

$$\frac{dS}{dt} = -q_S \cdot X$$ **[Equation 6]**

$$\frac{dN}{dt} = -q_N \cdot X$$ **[Equation 7]**

$$\frac{dP}{dt} = -q_P \cdot X$$ **[Equation 8]**

Where, N and P are the concentrations of NH_4^+ and PO_4^{3-} (mmol/l), respectively.

Parameter optimization

The initial parameter values of μ_m, X_m, Y_G, $Y_{X/N}$, and $Y_{X/P}$ were calculated from the experimental data. The initial values of k_S and m_S were estimated to be 1 and theoretically calculated as shown in next section, respectively. Subsequently, genetic algorithm (GA) (Goldberg, 1989) was used to refine the parameter values. Small spans were given around the above parameter values, within which GA was used to search the optimal parameter values with the least sum of square errors (*Err*) as described in Equation [9], between the model prediction and the experimental data.

$$\min\left\{ Err = \sum\left[\left(\frac{X_i - X_i'}{X_i'} \right)^2 + \left(\frac{S_i - S_i'}{S_i'} \right)^2 + \left(\frac{N_i - N_i'}{N_i'} \right)^2 + \left(\frac{P_i - P_i'}{P_i'} \right)^2 \right] \right\}$$ **[Equation 9]**

Where, X_i', S_i', N_i' and P_i' are the measured ith concentrations of the cells, glucose, NH_4^+ and PO_4^{3-}, respectively; X_i, S_i, N_i and P_i are the model predicted values corresponding to X_i', S_i', N_i' and P_i', respectively. The denominator of each item in Equation [9] ensures an equal contribution from each item to *Err*.

In application of GA in parameter optimization (Figure 1), a population containing n individuals was prepared. One individual was represented by one chromosome, which was constituted by a binary vector. One chromosome contained m genes (m = 7, the number of parameters) and one gene encoded one parameter. Each gene was constituted by a 10-bit binary vector, so one chromosome

was constituted by an *m* x 10 bits binary vector. At first, “initialization” was performed to the initial population by randomly assigning initial values of “0” or “1” to each bit of the *n* chromosomes (binary vectors). Then, genetic operations of “hybridization” and “mutation” were performed to the population at the probabilities defined by hybridization rate and mutation rate, respectively. “Hybridization” randomly occurred between two randomly selected individuals (chromosomes) by exchanging the corresponding parts of the two selected binary vectors (chromosomes) at a randomly selected position between 1 and (*m* x 10 - 1). “Mutation” randomly occurred to any chromosome at any position by turnover the selected bit from “0” to “1” or vice versa. After the genetic operations, the encoded binary value of each gene was decoded, converted to decimal and then scaled to real. Each chromosome contained one set of *m* model parameters, by which the mathematical model can be solved and the *Err* of this chromosome can be calculated. The chromosome (individual) with smaller *Err* was defined to have larger fitness and vice versa. Next, “natural selection” was performed to the population using a gambling roulette to randomly select the chromosomes (individuals), the larger fitness the larger possibility being selected. After the selection operation, a new generation was produced, which had larger averaged fitness than the old generation. The calculation was cycling until Equation [7] was less than the desired value. Figure 1 shows the diagram of GA used in this study. The population size was 50, and the hybridization and mutation rates were 0.2 and 0.05, respectively. In programming, the population of GA was represented by a *n* x (*m* x 10) two-dimensional array, and the differential equations of the mathematical model were solved by using the forth rank Runge-Kutta method. The software was programmed using Microsoft Visual Basic (license no. 61007978). Table 1 shows the optimized model parameters for batch culture.

Table 1. The parameter values of batch culture mathematical model.

μ_m	k_S	m_S	X_m	Y_G	$Y_{X/N}$	$Y_{X/P}$
0.43	5	0.013	27	0.32	0.13	2.15

RESULTS AND DISCUSSION

Batch culture simulation and analysis

Figure 2 shows that the model prediction was quite accurate, confirming the GA efficiency in the model parameter identification. GA is an optimization algorithm developed by imitating the evolution of a biological population, and it is efficient especially in optimizing nonlinear and sophisticated systems (Goldberg 1989). The perfect fits of the model prediction with the experimental data also indicated that the developed mathematical model was reasonable (Figure 2), and it can be used in the analysis of the fermentation process.

Table 1 shows that the values of m_S and Y_G optimized using GA are 0.013 g/g/h and 0.32 g/g, respectively. According to Equation [1], the averaged value of μ was 0.28 1/h. Equation [10] can be obtained from Equation [2]. Accordingly, $Y_{X/S}$ was calculated to be 0.29 g/g from Equation [10] using the above values.

$$\frac{1}{Y_{X/S}} = \frac{1}{Y_G} + \frac{m_S}{\mu} \qquad \textbf{[Equation 10]}$$

$$Y_G = \frac{\Delta X}{\Delta S_{ATP} + \Delta S_{Cell}} \qquad \textbf{[Equation 11]}$$

Equation [10] indicates that the consumed glucose is composed of two parts; one part is used in the real growth, and the other one part is used in the maintenance. The typical values of maintenance coefficient for ATP (m_{ATP}, moles of ATP consumed per C-mole of cells per hour for maintenance) are in the range of 0.04~0.1 mole/mole/h (Shuler and Kargi 2002), equivalent to m_S in the range of 0.01~0.02

g/g/h for eukaryotic cells in the case of aerobic catabolism of glucose through tricarboxylic acid (TCA) cycle with P/O of 3. Using the values of m_S and averaged value of μ, the item of m_s/μ in Equation [10] was calculated between 0.036 and 0.071 g/g. The consumed glucose in the real growth also includes two parts. One part is converted to cellular building blocks (ΔS_{Cell}), and the other part is used to produce the energy (ΔS_{ATP}) for polymerization of the building blocks into macromolecules. For the glucose converted to ΔS_{Cell}, about 91.3% of the glucose can be converted for the decarboxylation occurred in the pathways of ΔS_{Cell} synthesis (Heijnen 1992). An empirical molecular formula for *Candida utilis* cell composition, $CH_{1.84}O_{0.56}N_{0.2}P_{0.01}$ with the molecular weight (M_{w_cell}) of 25.8, was modified besed on the previous study (Heijnen 1992) and used in the calculations. The total amount of the produced cells was 26.2 g cells/l, containing 1.02 mole carbon/l (26.2/M_{w_cell}), and it required 33.52 g glucose/l (1.02/91.3% x M_{w_gluc}/6, M_{w_gluc} the molecular weight of glucose) for ΔS_{Cell} in syntheses of cellular building blocks. On the other hand, energy is required in the polymerization of cellular building blocks into macromolecules, which is defined by Y_{ATP} with the typical value of 10 g cells /mole ATP. The produced 26.2 g cells/l required 2.62 mol ATP/l. When glucose is metabolized by eukaryotic cells under the aerobic condition through TCA cycle, 36 mole ATP is produced when consuming 1 mole glucose. The synthesis of 26.2 g cells/l consumed 13.1 g glucose/l for the ATP production (ΔS_{ATP}) ((2.62/36) x M_{w_gluc}). Totally, 46.62 g glucose/l (ΔS_{ATP} + ΔS_{Cell}) was needed for the real growth of 26.2 g cells/l. Therefore, Y_G was calculated to be 0.56 g/g according to Equation [11]. Then, $Y_{X/S}$ can be theoretically calculated between 0.54 and 0.55 g/g, which is much larger than 0.29 g/g obtained from the model predicted parameters using Equation [10]. The reason is that the above assumption that the glucose consumed for ΔS_{ATP} and m_S were fully oxidized under the aerobic condition through TCA cycle is not true. In fact, NaOH was added to control the pH during the fermentation, indicating that the organic acid was produced and the glucose was partially metabolized through the fermentation. The results showed that the fermentation process could be further improved in order to increase $Y_{X/S}$. For example, for a more efficient use of glucose, glucose metabolism can be directed to TCA cycle through glucose concentration control and DO control.

Repeated fedbatch culture and simulation

Repeated fedbatch culture was performed as described in the section of Materials and Methods. Fedbatch culture is the most commonly used method in substrate concentration control. If high concentrated feeding solution is used and the volume change of the fermentation broth can be neglected or compensated by evaporation, the mathematical model developed in last section can also be used in the repeated fedbatch culture without any change. The simulation of repeated fedbatch culture was performed, fulfilled by resetting the initial values of the differential equations to the step changed concentrations at the start of the pulse feeding and solve the equations from this time point until the start of the next pulse feeding, like a new batch culture after another. This method made the simulation of the repeated fedbatch culture easier.

The measured time course data were used to re-optimize the model parameter values of the fedbatch culture using GA. The optimized values were X_m 71 g/l, $Y_{X/N}$ 0.11 g/mmol, $Y_{X/P}$ 1.5 g/mmol and the other values were the same with Table 1. By the computer simulation, Figure 3 shows that the glucose concentration was well controlled with small fluctuations, NH_4^+ concentration was well controlled but with larger fluctuations, while PO_4^{3-} was depleted from 10 to 16 hrs resulted from the late in the start of the feeding of PO_4^{3-} solution. The obtained maximum cell concentration was 71 g/l, which was more than doubled compared with that of the batch culture of 26 g/l.

According to Equation [10], the portion of the glucose consumed on maintenance is increased when μ is decreased, leading to the decrease of $Y_{X/S}$. In order to investigate the dynamic change of $Y_{X/S}$ in repeated fedbatch culture, we carried out the computer simulation (Figure 3). The time courses of μ of the batch and repeated fedbatch cultures were compared, and Figure 3 shows that μ of the repeated fedbatch culture was higher than that of the batch culture resulted from the substrate feeding. As a result, $Y_{X/S}$ was maintained at a relatively high level by using repeated fedbatch culture (Figure 3), indicating the usefulness of repeated fedbatch culture in process optimization. Compared with the continuous feeding, repeated fedbatch culture has several advantages, such as ease of operation and less contamination. Therefore, it is often used in industry. The results of fermentation operations can be predicted through using computer simulation with the mathematical model, which is efficient, labor-saving and cost effective.

In the present study, we developed a mathematical model for the production of *Candida utilis* cell mass, and GA was used in the model parameter optimization. This newly developed mathematical

model was useful in analyses of fermentation process and prediction of fermentation operations (batch and repeated fedbatch cultures). Our results showed that the cell mass production was significantly increased and the $Y_{X/S}$ value can be maintained at a high level by using the repeated fedbatch culture.

Financial support: This work was partially supported by the National Key Technology Research and Development Program (2011BAC02B04), the National Program on Key Basic Research Project (2010CB630902), "Twelfth Five-Year Plan" for Science and Technology Research (2011BAI06B01), the National High Technology Research and Development Program (2007AA05Z455) of China.

REFERENCES

ELMALEH, S.; DEFRANCE, M.B. and GHOMMIDH, C. (1999). Organic acids oxidation by *Candida utilis*: Application to industrial waste water treatment. *Process Biochemistry,* vol. 35, no. 5, p. 441-449. **[CrossRef]**

GOLDBERG, D.E. (1989). *Genetic algorithm in search, optimization and machine learning.* Addison Wesley, Chapter 3, 432 p. ISBN 10 0201157675.

HEIJNEN, J.J. (1992). Modeling of microbiological processes and kinetic conversion rates. In: VAN DEN AKKER, H.E.A.; HEIJNEN, J.J. and LEACH, C.K. eds. *Bioprocess Technology: Modeling and transport phenomena,* Butterworth Heinemann, 177-210 p.

LIANG, G.; LIAO, X.; DU, G. and CHEN, J. (2009). A new strategy to enhance glutathione production by multiple H_2O_2-induced oxidative stresses in *Candida utilis*. *Bioresource Technology,* vol. 100, no. 1, p. 350-355. **[CrossRef]**

LIN, J.Q.; LEE, S.M. and KOO, Y.M. (2005). Modeling and simulation of simultaneous saccharification and fermentation of paper mill sludge to lactic acid. *Journal of Microbiology and Biotechnology,* vol. 15, no. 1, p. 40-47.

SENGUPTA, S.; LAHIRI, S.; BANERJEE, S.; BASHISTHA, B. and GHOSH, A.K. (2011). Arginine mediated purification of trehalose-6-phosphate synthase (TPS) from *Candida utilis*: Its characterization and regulation. *Biochimica et Biophysica Acta (BBA) General Subjects,* vol. 1810, no. 12, p. 1346-1354. **[CrossRef]**

SHULER, M.L. and KARGI, F. (2002). *Bioprocess Engineering: Basic concepts, 2nd* ed, Prentice Hall, Englewood Cliffs, N.J., Chapter 7, 207-218 p. ISBN 10 0130819085.

TAMAKAWA, H.; IKUSHIMA, S. and YOSHIDA, S. (2012). Efficient production of l-lactic acid from xylose by a recombinant *Candida utilis* strain. *Journal of Bioscience and Bioengineering,* vol. 113, no. 1, p. 73-75. **[CrossRef]**

WANG, Y.; CHU, J.; ZHUANG, Y.; WANG, Y.; XIA, J. and ZHANG, S. (2009). Industrial bioprocess control and optimization in the context of systems biotechnology. *Biotechnology Advances,* vol. 27, no. 6, p. 989-995. **[CrossRef]**

YOKOYAMA, S.; OGAWA, A. and OBAYASHI, A. (1988). Rapid extraction of uricase from *Candida utilis* cells by use of reducing agent plus surfactant. *Enzyme and Microbial Technology,* vol. 10, no. 1, p. 52-55. **[CrossRef]**

ZU, Y.G.; ZHAO, X.H.; HU, M.S.; REN, Y.; XIAO, P.; ZHU, L.; CAO, Y.J. and ZHANG, Y. (2006). Biosorption effects of copper ions on *Candida utilis* under negative pressure cavitation. *Journal of Environmental Sciences,* vol. 18, no. 6, p. 1254-1259. **[CrossRef]**

How to reference this article:

LIN, Y.; LIU, G.; LIN, H.; GAO, L. and LIN, J. (2013). Analysis of batch and repeated fedbatch productions of *Candida utilis* cell mass using mathematical modeling method. *Electronic Journal of Biotechnology*, vol. 16, no. 4. http://dx.doi.org/10.2225/vol16-issue4-fulltext-2

Figures

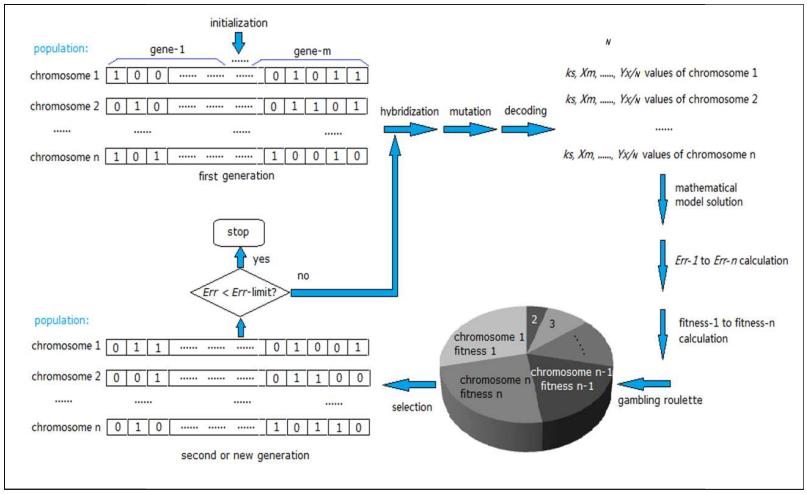

Fig. 1 The diagram of model parameter identification using GA.

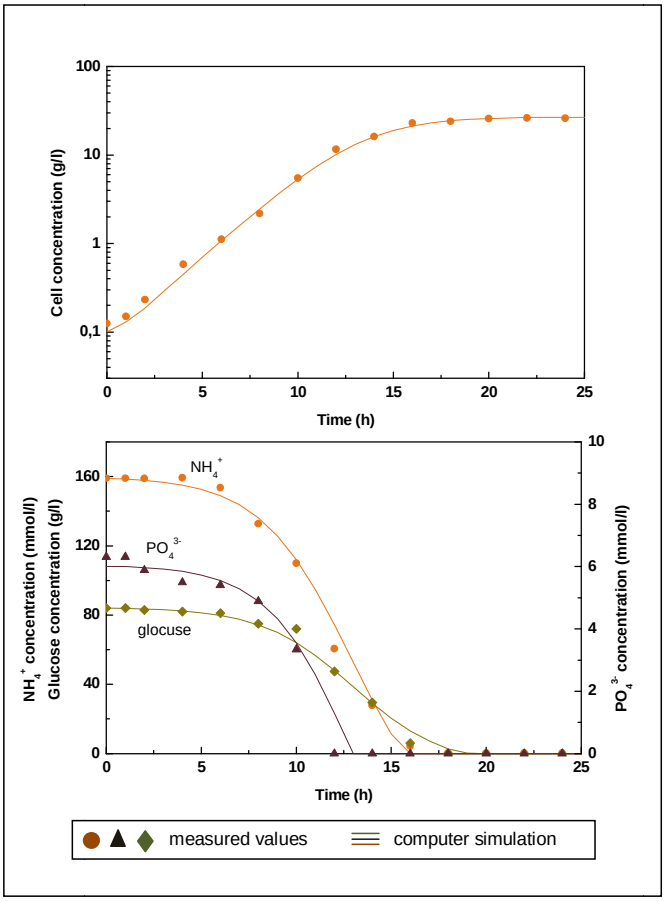

Fig. 2 The experimental and computer simulation results of batch culture for production of *Candida utilis* cell mass.

DOI: 10.2225/vol16-issue4-fulltext-2

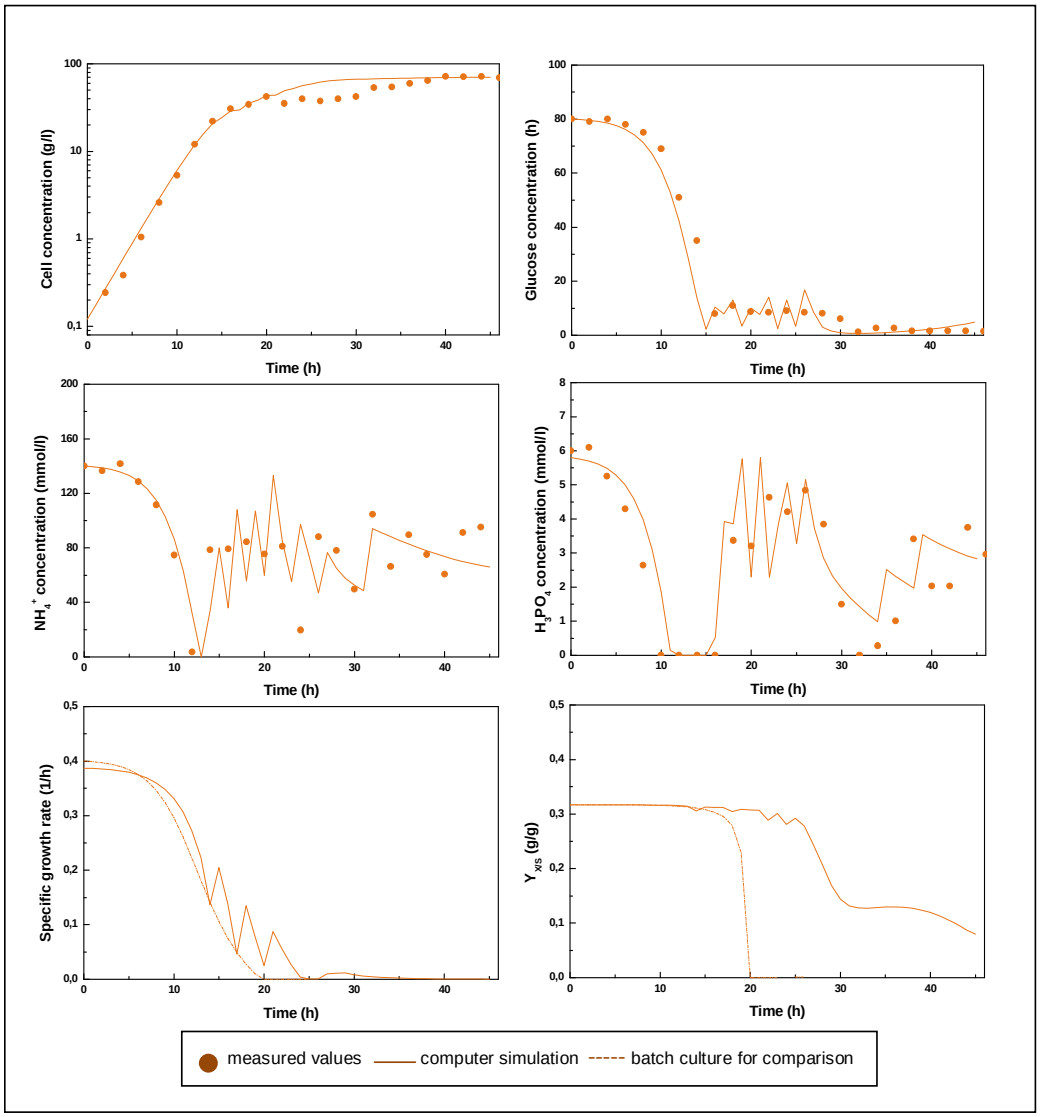

Fig. 3 Computer simulation of repeated fedbatch culture for production of *Candida utilis* cell mass.

DOI: 10.2225/vol16-issue4-fulltext-2