J. Microbiol. Biotechnol. (2004), 14(6), 1163–1169

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

Model Development for Lactic Acid Fermentation and Parameter Optimization Using Genetic Algorithm

LIN, JIAN-QIANG[1], SANG-MOK LEE[2], AND YOON-MO KOO[2]*

[1]*State Key Lab of Microbial Technology, School of Life Science, Shandong University, Jinan 250100, China*
[2]*Department of Biological Engineering, ERC for Advanced Bioseparation Technology, Inha University, Incheon 402-751, Korea*

Received: November 8, 2003
Accepted: March 12, 2004

Abstract An unstructured mathematical model is presented for lactic acid fermentation based on the energy balance. The proposed model reflects the energy metabolic state and then predicts the cell growth, lactic acid production, and glucose consumption rates by relating the above rates with the energy metabolic rate. Fermentation experiments were conducted under various initial lactic acid concentrations of 0, 30, 50, 70, and 90 g/l. Also, a genetic algorithm was used for further optimization of the model parameters and included the operations of coding, initialization, hybridization, mutation, decoding, fitness calculation, selection, and reproduction exerted on individuals (or chromosomes) in a population. The simulation results showed a good fit between the model prediction and the experimental data. The genetic algorithm proved to be useful for model parameter optimization, suggesting wider applications in the field of biological engineering.

Key words: Model, lactic acid, fermentation, genetic algorithm, optimization

Lactic acid has extensive applications in the food and pharmaceutical industries. In addition, its potential use as a source for polylactate polymers in making biodegradable plastics has an even bigger impact. Currently, about half of the world production of lactic acid is based on fermentation, while the remainder is created synthetically [10].

Modeling and simulation are powerful tools for the feasibility evaluation, design, and optimization of a fermentation system. The most popular mathematical model for lactic acid production is the Luedeking-Piret model [4], along with several modified models [1, 6, 7, 9]. The main advantages of these models are simplicity and fewer parameters. Yet, since they only relate lactic acid production to growth-associated and nongrowth-associated terms, these mathematical models have limited physiological value. Although a structured model is more rational and meaningful [7], it is too sophisticated for practical applications and has too many parameters that are difficult to determine. Accordingly, the current study uses an unstructured model to represent lactic acid production as a function of the energy production through the glycolysis pathway, thereby providing more physiological meaning, while maintaining simplicity for practical application. In addition, a genetic algorithm (GA), which is efficient in optimizing nonlinear and sophisticated systems, is used to optimize the model parameters [2].

*Corresponding author
Phone: 82-32-860-7513; Fax: 82-32-875-0827;
E-mail: ymkoo@inha.ac.kr

Materials and Methods

Strain, Medium, and Cultivation Conditions

The strain used was *Lactobacillus casei* subsp. *rhamnosus* (ATCC 10863). The medium contained yeast extract 15 g/l, sodium acetate 1 g/l, K_2HPO_4 0.3 g/l, KH_2PO_4 0.15 g/l, $MgSO_4 \cdot 7H_2O$ 0.15 g/l, and glucose 90 g/l. Four percent of $CaCO_3$ was also added to buffer the pH. Initial lactic acid concentrations were 0, 30, 50, 70, and 90 g/l.

The cultivation was performed in 100-ml flasks, containing 50 ml of the medium and 5% inoculum. The mixture was cultivated at 42°C, while shaking at 150 rpm in a rotary shaking incubator (Model KMC-8480 SF, Vision Scientific Co., Korea).

Analytical Procedures

A 1.5 ml sample was taken from each flask and centrifuged, then the supernatant was stored in a refrigerator for a subsequent analysis of the glucose and lactic acid. The glucose and lactic acid concentrations were measured based on the glucose oxidase-peroxidase and lactic acid oxidase-peroxidase methods, respectively, using an autoanalyzer (Biochemistry Analyzer 2700; YSI, Ohio, U.S.A.). A 0.5 ml

samples was also taken from each flask and the cell concentration measured according to the light absorbance method using a spectrophotometer (Spectronic Instruments Co., U.S.A.) at a wavelength of 600 nm. The samples were diluted from 5 to 50 times to keep the light absorbance value below 0.7. A 0.5 M HCl solution was used to dilute the samples so as to dissolve the solid $CaCO_3$ and avoid any interference in the light absorbance measurement.

RESULTS AND DISCUSSION

Yields from Lactic Acid Fermentation

The fermentation experiments were performed with initial lactic acid concentrations of 0, 30, 50, 70, and 90 g/l, as described above. The time course of the cell mass, glucose and lactic acid concentrations were measured. The yield of lactic acid from glucose, $Y_{P/S}$, was calculated, as shown in Fig. 1, and found to be slightly more than 80% (in the case of initial lactic acid concentrations of 70 and 90 g/l, $Y_{P/S}$ was not calculated, as the glucose consumption and lactic acid production were both very low, thereby potentially leading to large relative errors).

In a minimum medium, a carbon source is used for energy production, product and byproduct synthesis, and cell synthesis; while in a complex medium, a carbon source is only used for energy production, and product and byproduct synthesis, but not for cell synthesis. This was previously confirmed when using a complex medium for *Lactobacillus*, where only a little glucose was used in anabolism and the amount of glucose consumption for cell synthesis could be omitted in the mass balance calculations [8, 12]. Thus, if all the glucose is metabolized through the glycolysis pathway to produce two moles of lactic acid based on one mole of glucose, the theoretical value of $Y_{P/S}$ should be 1 g/g[$Y_{P/S}=(2\times M_{Lac}/M_{Gluc})=1$]. The results in Fig. 1

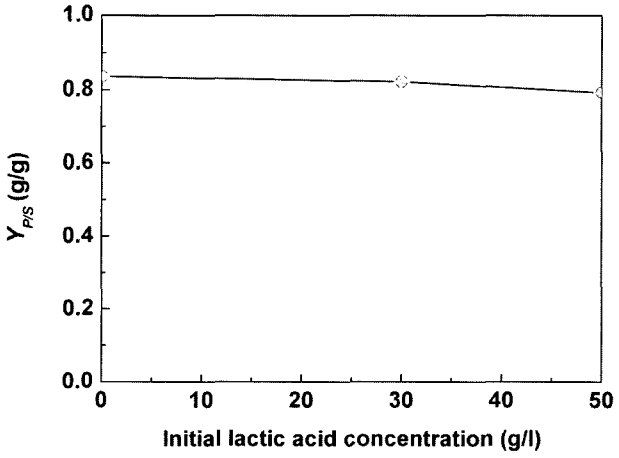

Fig. 1. $Y_{P/S}$ with various initial lactic acid concentrations.

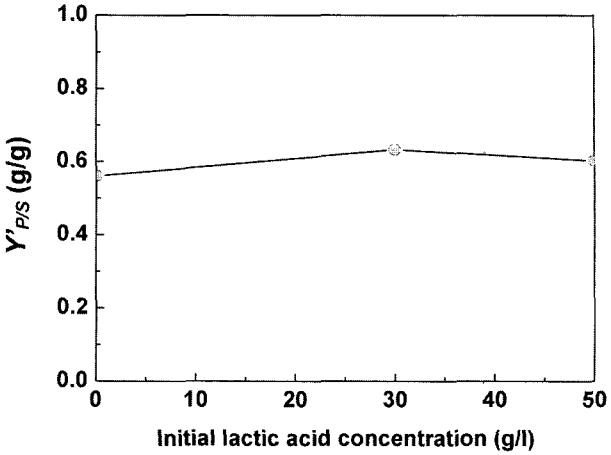

Fig. 2. $Y_{P/S}$ with various initial lactic acid concentrations.

show that nearly 20% of the glucose was not metabolized through the glycolysis pathway to produce energy and lactic acid. Instead, this portion of glucose, G'_S, was used in the synthesis of the secreted byproducts. In addition to lactic acid, various oligosaccharides and other kinds of organic acids were also detected in the fermentation broth using HPLC (data not shown). As such, G'_S can be calculated based on the total glucose consumption, G_S, minus the glucose consumption for lactic acid production, G_{Lac}, as follows:

$$G'_S=G_S-G_{Lac}=G_S-\frac{W_{Lac}\cdot M_{Gluc}}{M_{Lac}} \quad (1)$$

where, W_{Lac} is the weight of lactic acid produced, and M_{Gluc} and M_{Lac} are the molecular weights of glucose and lactic acid, respectively. The yield of byproducts from G'_S, and $Y'_{P/S}$, is shown in Fig. 2, which was about 0.6 g/g under the current experimental conditions. $Y'_{P/S}$ was confirmed to be significantly influenced by the cultivation conditions and the value of $Y_{P/S}$ when cultivation using a jar fermenter was 0.3 g/g.

Model Development

The specific glucose consumption rate was modeled using a Monod-type equation together with a lactic acid inhibition term and modified Logistic equation term as follows:

$$q_S=\left(\frac{q_{S,max}\cdot S}{k_S+S}\right)\cdot\left(1-\frac{P}{P_{cri}}\right)^m\cdot\left(1-\frac{X}{X_{max}}\right)^n \quad (2)$$

The specific byproduct production rate, q'_P was found to be growth related, therefore, the specific consumption rate of G'_S, q'_S, could be expressed as follows:

$$q'_S=\frac{q'_P}{Y'_{P/S}}=\frac{\gamma\cdot\mu}{Y'_{P/S}}=\frac{1}{\delta}\cdot\mu \quad (3)$$

Then, the specific lactic acid production rate could be calculated using Equation (4):

$$q_P=2\cdot(q_S-q_S')\cdot\frac{M_{Lac}}{M_{Gluc}} \quad (4)$$

The specific ATP production rate and specific ATP consumption rate were described using Equations (5) and (6), respectively:

$$q_{ATP,P}=\frac{q_P}{M_{Lac}} \quad (5)$$

$$q_{ATP,C}=\frac{\mu}{Y^*_{ATP}}+m_{ATP} \quad (6)$$

The quasi-steady state was assumed, so that the specific ATP production rate was equal to the specific ATP consumption rate:

$$q_{ATP,P}=q_{ATP,C}=\frac{q_P}{M_{Lac}}=\frac{\mu}{Y^*_{ATP}}+m_{ATP} \quad (7)$$

The specific growth rate could be obtained using Equations (2)-(4), and (7):

$$\mu=\frac{(q\cdot q_S-M_{Gluc}\cdot m_{ATP})\cdot Y^*_{ATP}\cdot\delta}{M_{Gluc}\cdot\delta+2\cdot Y^*_{ATP}} \quad (8)$$

The equations for the mass balance were as follows:

$$\frac{dX}{dt}=\mu\cdot X \quad (9)$$

$$\frac{dS}{dt}=-q_S\cdot X \quad (10)$$

$$\frac{dP}{dt}=q_P\cdot X \quad (11)$$

This model can also be used in modeling cell death. According to Equation (7), μ is minus when $(q_P/M_{Lac})<m_{ATP}$. Yet, it was assumed that $\mu\geq 0$ in the following simulations, as no death phase was observed in the experiments. Accordingly, the above model has more physiological meaning compared to the Luedeking-Piret model [3] or other kinds of unstructured mathematical models.

Parameter Value Determination and Optimization Using Genetic Algorithm

The parameter value of $Y_{P/S}$ was calculated from the theory described above, the value of X_{max} was measured directly, and the value of P_{cri} was obtained from a plot of μ versus the initial lactic acid concentration. The value of δ was calculated from the mass balance of the carbon source. The initial values of Y^*_{ATP} and m_{ATP} were obtained from other research and calculated using the energy balance. Meanwhile, the initial values of $q_{S,max}$, and k_S were obtained using linear plots from experimental data, and the values of m and n were obtained using nonlinear fitting from experimental data. Thereafter, the parameter values of Y^*_{ATP}, m_{ATP}, $q_{S,max}$, k_S, m, and n were refined using parameter optimization, which was performed by adjusting the parameter values within respective predetermined small ranges to minimize the objective function J of Equation (13) (the sum of the relative squared errors between model predictions and the raw experimental data) based on a GA using a computer.

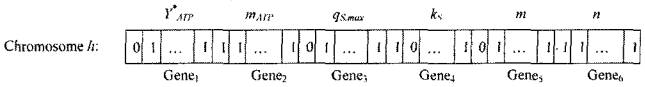

Fig. 3. Chromosome structure in GA.

A standard GA was used in the optimization, which included the operations of coding, initialization, hybridization, mutation, decoding, fitness calculation, selection, and reproduction exerted on individuals (or chromosomes) in a population. The potential solution (individual) was coded as a binary vector, called a chromosome, the elements of which were called genes and situated in predefined positions, indicated as alleles (Fig. 3). One chromosome consisted of six genes, coding for the six parameters, respectively. One gene was coded by a binary string of 10 bits. Then, every possible solution of the six parameters was simply represented by a binary string of 60 bits (Fig. 3). Initialization was performed to randomly assign initial values of "0" or "1" to each bit of the binary vectors (chromosomes). Hybridization was applied to two chromosomes called parents and two new chromosomes created by the exchange of one or more parts of the parent chromosomes, which occurred based on a probability called the hybridization rate. Mutation was applied by turning over one or more randomly selected bits in the binary chromosome vectors from "0" to "1" or from "1" to "0", which occurred based on a probability called the mutation rate. Decoding was performed to transform the binary number of each gene in the chromosome to a decimal system using the following equation:

$$D_{hi}=D_{min}+\frac{b_{hi}}{2^z-1}\cdot(D_{max}-D_{min}) \quad (12)$$

where h is the ordinal rank of the chromosome, i is the ordinal rank of the gene in the chromosome, z is the digits of the binary vector for one gene, b is the binary value coded by the corresponding gene, D is the decimal value transformed from b, and D_{max} and D_{min} are the decimal values of the maximum and minimum limits, respectively, of the corresponding parameter.

Selection and reproduction were performed to select individuals (chromosomes) to create a mating pool for reproducing offspring. The selection procedure was stochastic with the fitted chromosomes, judged by fitness (or objective function J), indicating a better chance of being selected. A low value of J was defined as a high fitness, with J defined as follows:

$$J=\frac{1}{R\cdot K\cdot L}\cdot\sum_r^R\sum_k^K\sum_l^L\left(\frac{\hat{Y}_{rkl}-Y_{rkl}}{Y_{rkl}}\right) \quad (13)$$

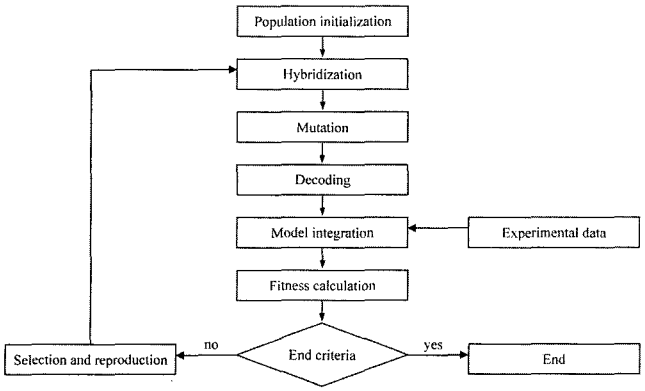

Fig. 4. GA flow diagram in parameter optimization.

where R is the number of experiments, K is the number of measured variables, L represents the measurement points of one variable, $\hat{Y}$ is the model output, and Y is the measured value in experiments. To give each variable the same importance in the optimization, the relative error $(\hat{Y}_{rkl}-Y_{rkl}/Y_{rkl})$ was used so that J had no unit. Another advantage to using the relative error instead of the absolute error $(\hat{Y}_{rkl}-Y_{rkl})$ in calculating J was to allow the data points of one variable at different sampling times to have the same importance in the optimization. A flow diagram of the GA used for the model parameter optimization is shown in Fig. 4. The program was self-programmed using Visual Basic and applied the fourth order Runge-Kutta method to solve the differential Equations and the GA to minimize J in Equation (13). A population size of 20, hybridization rate of 0.2, and mutation rate of 0.05 were used in the GA for the optimization. The J values of the predictions for the 20 chromosomes in the first 50 generations are shown in Fig. 5, where the largest value of J became larger as the number of generations increased. This resulted from the hybridization and mutation operations, which enlarged the search span and helped to avoid reaching the local optimum value. Meanwhile, the lowest value of J decreased as the number of generations increased (Fig. 5), showing an increased fitness. Very small values of J (<0.15) were obtained in the early generations (Fig. 5), indicating the efficacy of the GA in optimizing. The values of the optimized parameters (Table 1) also produced a good fit between the simulation and the experimental data (Fig. 6), thereby reconfirming the effectiveness of a GA for parameter optimization.

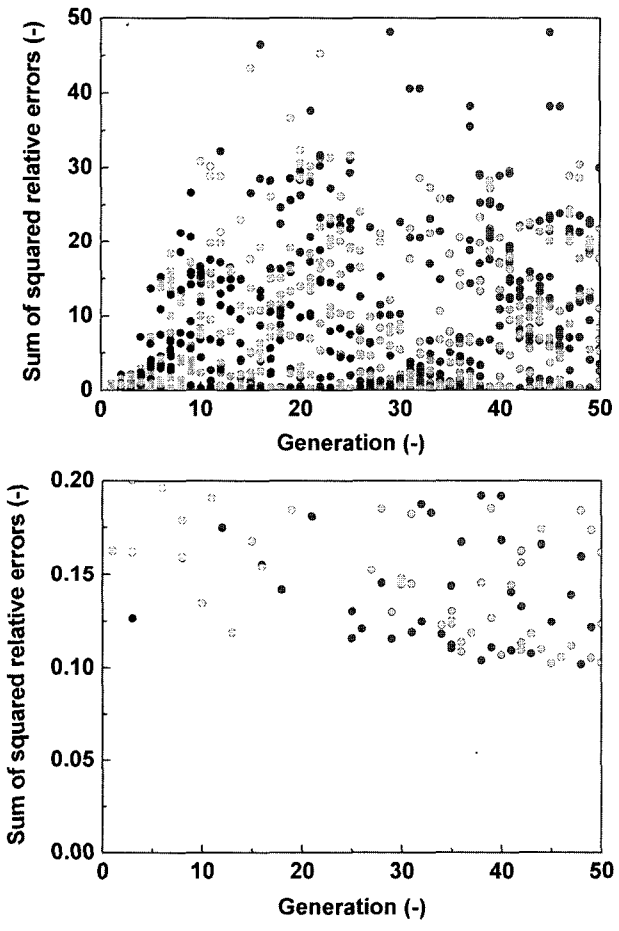

Fig. 5. Changing trend of sum of squared relative errors with generations in parameter optimization using GA.

Table 1. Parameter values for lactic acid fermentation model.

Parameter	Value	Unit
$q_{S,max}$	3	g/g/h
k_S	0.3	g/l
P_{cri}	70	g/l
m	0.7	-
n	4.5	-
m_{ATP}	0.002	mol/g/h
δ	0.6	g/g
Y^*_{ATP}	20	g/mol
$Y_{P/S}$	1.0	g/g

Model Simulation and Comparison with Experimental Data

The experimental data and model simulations of cell growth, lactic acid production, and glucose consumption under various initial lactic acid concentrations are shown in Fig. 6. The model simulation was performed by solving the differential equations of the mathematical model using the fourth order Runge-Kutta method with a self-programmed Visual Basic (Microsoft Co., U.S.A.) program running on an IBM compatible computer with Windows ME (Microsoft Co., U.S.A.). The values of the model parameters used are shown in Table 1. The simulation results in Fig. 6 show that the model matched the experimental data satisfactorily. The simulation of the specific growth rate, specific glucose consumption rate, and specific ATP production rate using

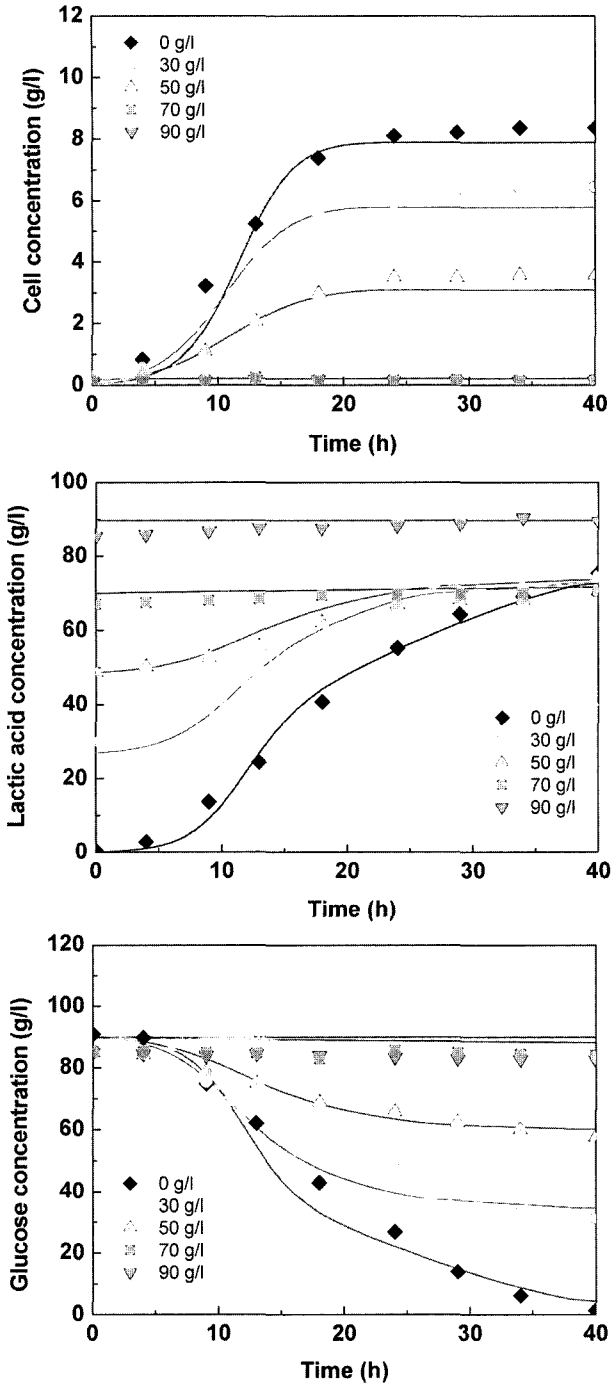

Fig. 6. Lactic acid fermentation with various initial lactic acid concentrations.

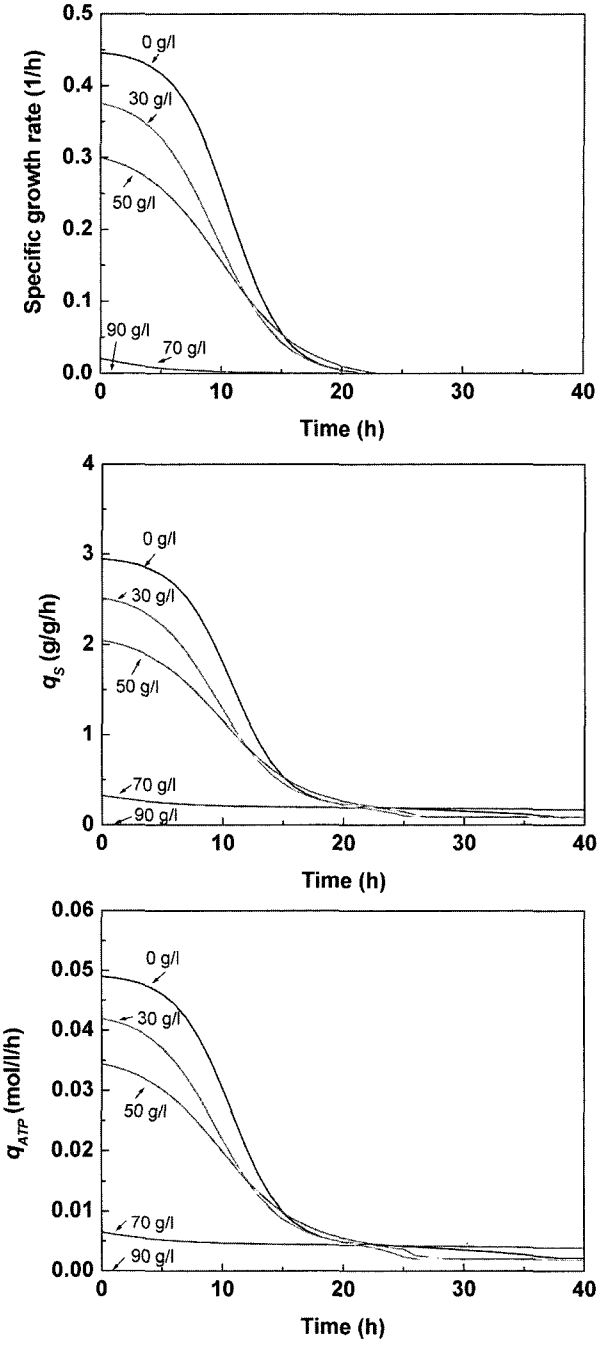

Fig. 7. Simulation of specific growth rate, glucose consumption rate, and ATP production rate with various initial lactic acid concentrations.

Equations (8), (2), and (7), respectively, is shown in Fig. 7, and the simulations of $Y_{X/ATP}$ ($=m/q_{ATP}$) and $Y_{ATP/S}$ ($=q_{ATP}/q_S$) are shown in Figs. 8 and 9, respectively.

The maximum values for the specific rates of cell growth, glucose consumption, and ATP production clearly decreased with an increase in the initial lactic acid concentrations. The rates dropped rapidly to low values after about 15 h without glucose depletion when the lactic acid and/or cell concentrations increased to high levels, exhibiting strong inhibitory effects (Figs. 6 and 7). The maximum values of $Y_{X/ATP}$ decreased with an increase in the initial lactic acid concentration (Fig. 8); as μ is low, a larger portion of ATP is used in maintenance, according to the following Equation (14):

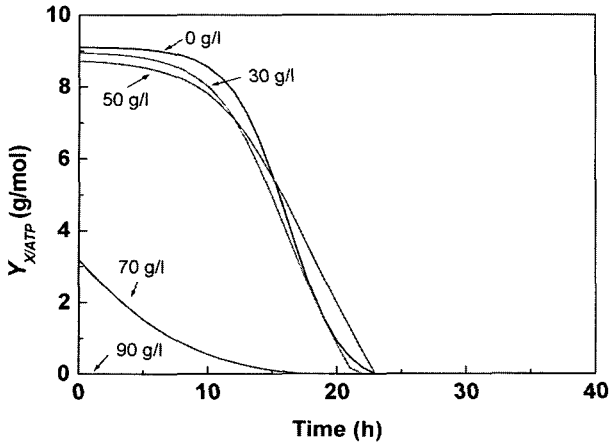

Fig. 8. Simulation of cell yield from ATP with various initial lactic acid concentrations.

$$\frac{1}{Y_{X/ATP}} = \frac{1}{Y^*_{ATP}} + \frac{m_{ATP}}{\mu} \tag{14}$$

$Y_{ATP/S}$ was noticeably larger in the case of an initial lactic acid concentration of 70 g/l with an extremely small value of μ (Figs. 8 and 9), as a small portion of glucose was incorporated into the growth-associated byproducts when the cell growth rate was low. When the cell growth stopped, $Y_{ATP/S}$ increased to the maximum values (Fig. 9). The case of an initial lactic acid concentration of 90 g/l was not simulated, as almost no cell growth and glucose consumption were found.

In conclusion, the proposed mathematical model was found to effectively reflect the energy metabolic state, yet still be simple enough for the purposes of process analysis, design, and optimization. A GA was also proven to be useful in model parameter optimization, suggesting wider applications in the field of bioengineering [11].

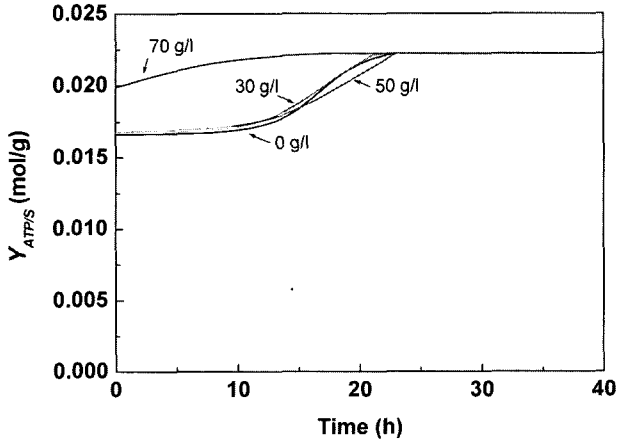

Fig. 9. Simulation of ATP yield from glucose with various initial lactic acid concentrations.

Acknowledgment

This work was supported by the ERC for Advanced Bioseparation Technology, KOSEF.

NOMENCLATURE

b : value of binary number coding for one gene (-)
D : decimal value of b (-)
G_S : amount of glucose consumed totally (g)
G'_S : amount of glucose used for cell synthesis *etc.* (g)
G_{Lac} : amount of glucose converted to lactic acid (g)
h : ordinal rank of chromosome (-)
H : chromosome number (-)
i : ordinal rank of gene in chromosome (-)
I : gene number (-)
J : objective function (-)
K : number of measurable variables in model (-)
k_S : glucose saturation constant in Monod-type equation (g/l)
L : number of measurements for one variable (-)
m : constant (-)
m_{ATP} : rate of ATP consumption for maintenance energy (mol/g/h)
M_{Gluc} : molecular weight of glucose (g/mol)
M_{Lac} : molecular weight of lactic acid (g/mol)
n : constant (-)
P : lactic acid concentration (g/l)
P_{cri} : critical lactic acid concentration for inhibition of glucose consumption (g/l)
q_{ATP} : specific metabolic rate of ATP (mol/g/h)
$q_{ATP,C}$: specific ATP consumption rate (mol/g/h)
$q_{ATP,P}$: specific ATP production rate (mol/g/h)
q'_P : specific byproduct production rate (g/g/h)
q_S : specific glucose consumption rate (g/g/h)
q'_S : specific rate of G'_S (g/g/h)
$q_{S,max}$: maximum specific glucose consumption rate (g/g/h)
q_P : specific lactic acid production rate (g/g/h)
R : experimental number (-)
S : glucose concentration (g/l)
W_{Lac} : weight of lactic acid produced (g)
X : cell concentration (g/l)
X_{max} : experimental value of maximum cell concentration (g/l)
$Y_{X/ATP}$: "apparent" cell yield from ATP (g/mol)
Y^*_{ATP} : maximum cell yield from ATP (g/mol)
$Y_{P/S}$: lactic acid yield from glucose (g/g)
$Y'_{X/S}$: cell yield from glucose (g/g)
Y : measured experimental data (g/l)
$\hat{Y}$: outputs of mathematical model (g/l)
z : digits of binary vector for one gene (-)

Greek Symbols

δ : constant (g/g)
γ : constant (g/g)
μ : specific growth rate (1/h)

REFERENCES

1. Boonmee, M., N. Leksawasdi, W. Bridge, and P. L. Rogers. 2003. Batch and continuous culture of *Lactococcus lactis* NZ133: Experimental data and model development. *Biochem. Eng. J.* **14:** 127-135.
2. Goldberg, D. E. 1989. *Genetic Algorithm in Search, Optimization and Machine Learning*. Addison Wesley, U.S.A.
3. Lin, J. Q., S. M. Lee, H. J. Lee, and Y. M. Koo. 2000. Modeling of typical microbial cell growth in batch culture. *Biotechnol. Bioprocess Eng.* **5:** 382-385.
4. Luedeking, R. and E. L. Piret. 1959. A kinetic study of the lactic acid fermentation. *J. Biochem. Microbiol.* **1:** 393-412.
5. Keller, A. K. and P. Gerhardt. 1975. Continuous lactic acid fermentation of whey to produce a ruminant feed supplement high in crude protein. *Biotechnol. Bioeng.* **17:** 997-1018.
6. Kim, N.-J., I. S. Suh, B.-K. Hur, and C.-G. Lee. 2002. Simple monodimensional model for linear growth rate of photosynthetic microorganisms in flat-plate photobioreactors. *J. Microbiol. Biotechnol.* **12(6):** 962-972.
7. Nielsen, J., K. Nikolajsen, and J. Villadsen. 1991. Structured modeling of a microbial system, part 1, 2. *Biotechnol. Bioeng.* **38:** 1-10, 11-23.
8. Olmos-Dichara, A., F. Ampe, J. Uribelarrea, A. Pareilleux, and G. Goma. 1997. Growth and lactic acid production by *Lactobacillus casei ssp. Rhamnosus* in batch and membrane bioreactor. *Biotechnol. Lett.* **8:** 709-714.
9. Suh, I. S. and S. B. Lee. 2003. Optimization of radiator position in an internally radiating photobioreactor: a model simulation study. *J. Microbiol. Biotechnol.* **13(5):** 789-793.
10. Vickroy, T. B. 1985. Lactic acid, pp. 761-776. *In*: Blanch, H. W., Drew, S. and Wang, D. I. C. (eds.), *The Practice of Biotechnology: Commodity Products*. Pergamon Press, Elmsford, NY, U.S.A.
11. Wang, D., L. Gao, J. Q. Lin, Y. B. Qu, J. Q. Lin, and S. Y. Yu. 2002. Optimization of substrate feeding trajectory for fed-batch culture using genetic algorithm. *Industr. Microbiol.* **32:** 40-43. (in Chinese).
12. Youssef, C. B., V. Guillou, and A. Olmos-Dichara. 2000. Modeling and adaptive control strategy in a lactic acid fermentation process. *Contr. Eng. Pract.* **8:** 1297-13.