Optimal Fed-Batch Control of Induced Foreign Protein Production by Recombinant Bacteria

Jongdae Lee and W. Fred Ramirez
Dept. of Chemical Engineering, University of Colorado, Boulder, CO 80309

Optimal control strategies for maximizing the production of induced foreign protein by recombinant bacteria were sought by the optimal control theory. Nutrient and inducer feeding rates were selected as key control variables. Since the problem is linear in the control variables, the optimal control is bang-bang or singular. Singular solutions are shown to exist. The optimal control theory showed that the specific growth rate with respect to nutrient concentration must be kept in its maximum phase and that there exist both a cell growth period and a protein production period. The optimal control theory calculates exactly the growth and production periods. The glucose concentration is controlled along a singular arc to give a maximum specific growth rate. The inducer level is controlled along a separate singular arc.

Introduction

Recent work concerning the production of proteins using recombinant bacteria has shown the importance of nutrient concentration and inducer level on foreign protein production. It has been demonstrated that high levels of inducer can result in higher production amounts of protein product; however, the cell growth rate can be reduced significantly with high levels of inducer. The time of induction has also been shown to be an important factor in maximizing protein production. Park et al. (1989) studied operating conditions for recombinant *E. coli* cells in fed-batch and two-stage continuous fermenters. Bentley (1989) experimentally found an optimal inducer injection policy for a specific system. He also showed that inducer levels can affect the specific growth rate and that the specific growth rate changes with time. He strongly suggested that inducer levels could reduce the specific growth rate significantly. Bentley and Kompala (1989) developed a mathematical model for protein production that includes inducer effects as well as nutrient effects. The model is not appropriate for online optimization since it includes many unobservable state variables. Even more complex models have been developed (Shu and Shuler, 1989; Peretti and Bailey, 1986). They, of course, suffer the same problem for control purposes as does the Bentley and Kompala model. Lee and Ramirez (1992) have developed a mathematical model which includes the inducer effect on specific growth rate and foreign protein production, and is suitable for control studies.

A key concern in operating a bioreactor for the production of foreign protein using recombinant bacteria is exactly how to use the two main controls of nutrient feed rate and inducer feed rate to maximize the productivity of the reactor. Most biotech reactor systems are carried out using fed-batch reactors; therefore, the system is transient rather than steady-state. Dynamic optimization techniques of optimal control are therefore required in order to determine optimal operational strategies.

The optimal control theory has been applied to maximize the cell concentration and metabolite yield by controlling nutrient concentration, temperature, and pH (Menawat et al., 1987). Park and Ramirez (1988, 1990) applied optimal control theory to a yeast system where the control variable is the nutrient feeding rate. Multiple singular arcs were shown to exist and were related to biological function. Effort has been placed on the optimization of foreign protein production using recombinant bacteria also (Georgiou, 1988; Zabriskie et al., 1986).

Rather than taking the usual approach to finding optimal operational strategies by intuition and experience based upon experimentation and simulation, we propose to use a system science approach. The class of problems considered in this work is optimal operation of *E. coli* recombinant systems producing foreign proteins when protein induction and nutrient level are both affecting product productivity.

Physical System

The class of problems considered are the fed-batch production of protein using recombinant bacteria where protein induction is required. This class of problems has been mathematically modeled by Lee and Ramirez (1992). Their model essentially involves seven state variables and two control variables. The state variables are the reactor volume (x_1), the cell density (x_2), the nutrient concentration (x_3), the foreign protein concentration (x_4), the inducer concentration (x_5), the inducer shock factor on cell growth rate (x_6), and the inducer recovery factor on cell growth rate (x_7). The two control variables are the glucose feeding rate (u_1) and the inducer feeding rate (u_2). The model is given by the following differential equations:

$$\dot{x}_1 = u_1 + u_2 \tag{1}$$

$$\dot{x}_2 = \mu(x_3, x_5, x_6, x_7)x_2 - \frac{u_1+u_2}{x_1}x_2 \tag{2}$$

$$\dot{x}_3 = \frac{u_1}{x_1}C_n^f - \frac{u_1+u_2}{x_1}x_3 - Y^{-1}\mu(x_3, x_5, x_6, x_7)x_2 \tag{3}$$

$$\dot{x}_4 = R_{fp}(x_3, x_5)x_2 - \frac{u_1+u_2}{x_1}x_4 \tag{4}$$

$$\dot{x}_5 = \frac{u_2}{x_1}C_i^f - \frac{u_1+u_2}{x_1}x_5 \tag{5}$$

$$\dot{x}_6 = -k_1(x_5)x_6 \tag{6}$$

$$\dot{x}_7 = k_2(x_5)(1-x_7) \tag{7}$$

Optimal Control Problem

The objective is, for a specified final time of fed-batch operation, to maximize the profitability of the process. We consider the major elements in profitability to be the difference between the value of the product and the cost of the inducer. Inducers such as isopropylthiogalactoside (IPTG) are quite expensive, costing around \$20/g (Sigma Chemical Catalog, 1991); therefore, they should be included in calculating the net profit. The performance functional that mathematically expresses this objective is:

$$J = -x_4(t_f)x_1(t_f) + Q\int_{t_o}^{t_f} u_2(t)\,dt \tag{8}$$

where Q is the ratio of the cost of inducer to the value of the protein product. The first term of Eq. 8 is the amount of product produced and the second term is the relative cost of using the inducer. We wish to maximize this performance functional. At this stage, it is virtually impossible to have an intuitive optimal solution since too many factors are involved. However, since the problem is linear in the control (Eqs. 1-7 and Eq. 8), the optimal operational policy is known to be that of bang-bang control or that of regulation along a singular arc (Bell and Jacobson, 1975; Ramirez, 1994). The shapes of these singular arcs, together with the complete overall solution, can be determined through model-based process optimization applying Pontryagin's maximum principle (Bryson and Ho, 1975; Sage and White, 1977). In using the maximum principle, the Hamiltonian becomes:

$$\begin{aligned} H = &\left(\lambda_1 - \frac{x_2}{x_1}\lambda_2 + \frac{C_n^f}{x_1}\lambda_3 - \frac{x_3}{x_1}\lambda_3 - \frac{x_4}{x_1}\lambda_4 - \frac{x_5}{x_1}\lambda_5\right)u_1 \\ &+ \left(\lambda_1 - \frac{x_2}{x_1}\lambda_2 - \frac{x_3}{x_1}\lambda_3 - \frac{x_4}{x_1}\lambda_4 + \frac{C_i^f}{x_1}\lambda_5 - \frac{x_5}{x_1}\lambda_5 + Q\right)u_2 \\ &+ \mu(x_3, x_5, x_6, x_7)x_2\lambda_2 - Y^{-1}\mu(x_3, x_5, x_6, x_7)x_2\lambda_3 \\ &+ R_{fp}(x_3, x_5)x_2\lambda_4 - k_1(x_5)x_6\lambda_6 + k_2(x_5)(1-x_7)\lambda_7 \end{aligned} \tag{9}$$

The costate equations are:

$$\begin{aligned} \dot{\lambda}_1 = -\frac{\partial H}{\partial x_1} = &-\frac{x_2}{x_1^2}\lambda_2 u_1 + \frac{C_n^f}{x_1^2}\lambda_3 u_1 - \frac{x_3}{x_1^2}\lambda_3 u_1 \\ &- \frac{x_4}{x_1^2}\lambda_4 u_1 - \frac{x_5}{x_1^2}\lambda_5 u_1 - \frac{x_2}{x_1^2}\lambda_2 u_2 - \frac{x_3}{x_1^2}\lambda_3 u_2 - \frac{x_4}{x_1^2}\lambda_4 u_2 \\ &+ \frac{C_i^f}{x_1^2}\lambda_5 u_2 - \frac{x_5}{x_1^2}\lambda_5 u_2 \end{aligned} \tag{10}$$

$$\begin{aligned} \dot{\lambda}_2 = -\frac{\partial H}{\partial x_2} = &\frac{\lambda_2}{x_1}u_1 + \frac{\lambda_2}{x_1}u_2 - \mu(x_3, x_5, x_6, x_7)\lambda_2 \\ &+ Y^{-1}\mu(x_3, x_5, x_6, x_7)\lambda_3 - R_{fp}(x_3, x_5)\lambda_4 \end{aligned} \tag{11}$$

$$\begin{aligned} \dot{\lambda}_3 = -\frac{\partial H}{\partial x_3} = &\frac{\lambda_3}{x_1}u_1 + \frac{\lambda_3}{x_1}u_2 - \frac{\partial \mu(x_3, x_5, x_6, x_7)}{\partial x_3}x_2\lambda_2 \\ &+ \frac{\partial\{Y^{-1}\mu(x_3, x_5, x_6, x_7)\}}{\partial x_3}x_2\lambda_3 - \frac{\partial R_{fp}(x_3, x_5)}{\partial x_3}x_2\lambda_4 \end{aligned} \tag{12}$$

$$\dot{\lambda}_4 = -\frac{\partial H}{\partial x_4} = \frac{\lambda_4}{x_1}u_1 + \frac{\lambda_4}{x_1}u_2 \tag{13}$$

$$\begin{aligned} \dot{\lambda}_5 = -\frac{\partial H}{\partial x_5} = &\frac{\lambda_5}{x_1}u_1 + \frac{\lambda_5}{x_1}u_2 - \frac{\partial \mu(x_3, x_5, x_6, x_7)}{\partial x_5}x_2\lambda_2 \\ &+ \frac{\partial\{Y^{-1}\mu(x_3, x_5, x_6, x_7)\}}{\partial x_5}x_2\lambda_3 - \frac{\partial R_{fp}(x_3, x_5)}{\partial x_5}x_2\lambda_4 \\ &+ \frac{\partial k_1(x_5)}{\partial x_5}x_6\lambda_6 - \frac{\partial k_2(x_5)}{\partial x_5}(1-x_7)\lambda_7 \end{aligned} \tag{14}$$

$$\begin{aligned} \dot{\lambda}_6 = -\frac{\partial H}{\partial x_6} = &k_1(x_5)\lambda_6 - \frac{\partial \mu(x_3, x_5, x_6, x_7)}{\partial x_6}x_2\lambda_2 \\ &+ x_2\lambda_3\frac{\partial\{Y^{-1}\mu(x_3, x_5, x_6, x_7)\}}{\partial x_6} \end{aligned} \tag{15}$$

$$\begin{aligned} \dot{\lambda}_7 = -\frac{\partial H}{\partial x_7} = &k_2(x_5)\lambda_7 - \frac{\partial \mu(x_3, x_5, x_6, x_7)}{\partial x_7}x_2\lambda_2 \\ &+ x_2\lambda_3\frac{\partial\{Y^{-1}\mu(x_3, x_5, x_6, x_7)\}}{\partial x_7} \end{aligned} \tag{16}$$

From the transversality conditions, the final costate values are:

$$\lambda_1(t_f) = -x_4(t_f) \tag{17}$$

$$\lambda_2(t_f) = 0 \tag{18}$$

$$\lambda_3(t_f) = 0 \tag{19}$$

$$\lambda_4(t_f) = -x_1(t_f) \tag{20}$$

$$\lambda_5(t_f) = 0 \tag{21}$$

$$\lambda_6(t_f) = 0 \tag{22}$$

$$\lambda_7(t_f) = 0 \tag{23}$$

We can describe the Hamiltonian in a compact form as:

$$H = H_{u_1}u_1 + H_{u_2}u_2 + \mu(x_3, x_5, x_6, x_7)x_2\lambda_2 - Y^{-1}\mu(x_3, x_5, x_6, x_7)x_2\lambda_3 + R_{fp}(x_3, x_5)x_2\lambda_4 - k_1(x_5)x_6\lambda_6 + k_2(x_5)(1-x_7)\lambda_7 \tag{24}$$

where

$$H_{u_1} = \lambda_1 - \frac{x_2}{x_1}\lambda_2 + \frac{C_n^f}{x_1}\lambda_3 - \frac{x_3}{x_1}\lambda_3 - \frac{x_4}{x_1}\lambda_4 - \frac{x_5}{x_1}\lambda_5 \tag{25}$$

$$H_{u_2} = \lambda_1 - \frac{x_2}{x_1}\lambda_2 - \frac{x_3}{x_1}\lambda_3 - \frac{x_4}{x_1}\lambda_4 + \frac{C_i^f}{x_1}\lambda_5 - \frac{x_5}{x_1}\lambda_5 + Q \tag{26}$$

To get optimal control laws, we must minimize the Hamiltonian with respect to each control. From the maximum principle, the control laws must satisfy the following conditions:

- If $H_{u_i} \geq 0$, then $u_i = u_{i_{\min}}$.
- If $H_{u_i} \leq 0$, then $u_i = u_{i_{\max}}$.
- If H_{u_i} stays zero, then $u_i = u_{i_{\text{singular}}}$.

Since the Hamiltonian is linear in both control variables, it is not possible to get the singular control laws directly from the development to date. For a singular solution to exist for control u_1, the time derivative of H_{u_1} must be zero:

$$\frac{\partial H_{u_1}}{\partial t} = \frac{x_2}{x_1}\left(-\frac{\partial \mu}{\partial x_3}\lambda_2 C_n^f + \frac{\partial (Y^{-1}\mu)}{\partial x_3}\lambda_3 C_n^f - \frac{\partial R_{fp}}{\partial x_3}\lambda_4 C_n^f + \frac{\partial \mu}{\partial x_3}\lambda_2 x_3 - \frac{\partial (Y^{-1}\mu)}{\partial x_3}\lambda_3 x_3 + \frac{\partial R_{fp}}{\partial x_3}\lambda_4 x_3 + \frac{\partial \mu}{\partial x_5}\lambda_2 x_5 - \frac{\partial (Y^{-1}\mu)}{\partial x_5}\lambda_3 x_5 + \frac{\partial R_{fp}}{\partial x_5}\lambda_4 x_5 \right) - \frac{x_5}{x_1}\left\{ \frac{\partial k_1}{\partial x_5}x_6\lambda_6 - \frac{\partial k_2}{\partial x_5}(1-x_7)\lambda_7 \right\} = 0 \tag{27}$$

The condition that must exist for Eq. 27 to hold is (Lee, 1992):

$$\frac{\partial \mu}{\partial x_3}\lambda_2 - \frac{\partial (Y^{-1}\mu)}{\partial x_3}\lambda_3 + \frac{\partial R_{fp}}{\partial x_3}\lambda_4 = 0 \tag{28}$$

For a singular solution to exist for control u_2, the time derivative of H_{u_2} must be zero:

$$\frac{\partial H_{u_2}}{\partial t} = \frac{x_2}{x_1}\left(+ \frac{\partial \mu}{\partial x_3}\lambda_2 x_3 - \frac{\partial (Y^{-1}\mu)}{\partial x_3}\lambda_3 x_3 + \frac{\partial R_{fp}}{\partial x_3}\lambda_4 x_3 - \frac{\partial \mu}{\partial x_5}\lambda_3 C_i^f + \frac{\partial (Y^{-1}\mu)}{\partial x_5}\lambda_3 C_i^f - \frac{\partial R_{fp}}{\partial x_5}\lambda_4 C_i^f + \frac{\partial \mu}{\partial x_5}\lambda_2 x_5 - \frac{\partial (Y^{-1}\mu)}{\partial x_5}\lambda_3 x_5 + \frac{\partial R_{fp}}{\partial x_5}\lambda_4 x_5 \right) - \frac{(x_5 - C_i^f)}{x_1}\left\{ \frac{\partial k_1}{\partial x_5}x_6\lambda_6 - \frac{\partial k_2}{\partial x_5}(1-x_7)\lambda_7 \right\} = 0 \tag{29}$$

The condition that must exist for Eq. 28 to hold is:

$$\left(\frac{\partial \mu}{\partial x_5}\lambda_2 - \frac{\partial (Y^{-1}\mu)}{\partial x_5}\lambda_3 + \frac{\partial R_{fp}}{\partial x_5}\lambda_4 \right) x_2 - \left\{ \frac{\partial k_1}{\partial x_5}x_6\lambda_6 - \frac{\partial k_2}{\partial x_5}(1-x_7)\lambda_7 \right\} = 0 \tag{30}$$

To calculate the optimal glucose and inducer concentrations from Eqs. 28 and 30, we use the following iterative method. After integrating the state equations forward and integrating costate equations backward with arbitrary controls, we calculate the singular glucose concentration x_3^{opt}, which satisfies Eq. 28 and the singular inducer concentration x_5^{opt}, which satisfies Eq. 30 during the backward integration of the costate equations. The resulting singular glucose and inducer profiles are followed at the next forward integration of the state equations if the controls, u_1 and u_2 calculated from Eqs. 3 and 5 do not violate their constraints. This procedure is repeated until the objective function does not change significantly from iteration to iteration.

The control domain is divided into four subdomains by the two hyperplanes, $H_{u_1} = 0$ and $H_{u_2} = 0$. Figure 1 shows a typical plot of time vs. these variables. The controls, u_1 and u_2, are determined by the maximum principle. The singular arcs for each control, $H_{u_i} = 0$, stay on appropriate hyperplanes. Any point off the hyperplanes is in a region of extremal control defined by $H_{u_i} > 0$ or $H_{u_i} < 0$. The trajectory H_{u_i} by an optimal

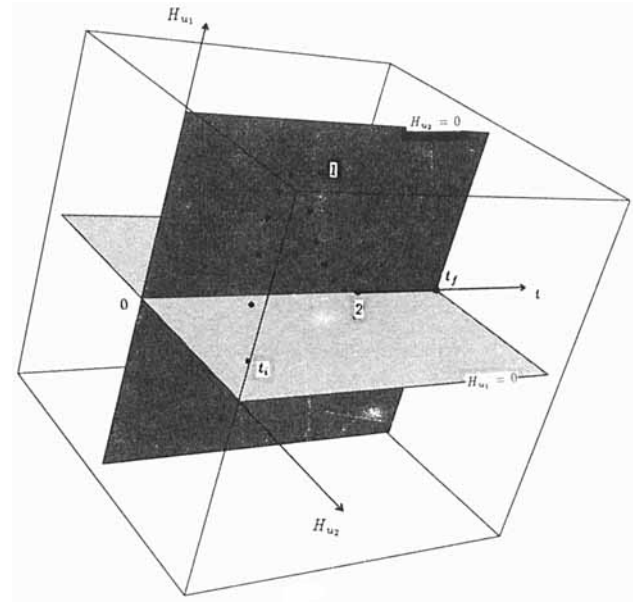

Figure 1. Optimal control domains.

control action can lie outside both hyperplanes, on a hyperplane or on both. Those cases are illustrated in Figure 1. The initial points between t_i and point 1 are in a region of $H_{u_1}>0$ and $H_{u_2}>0$. Therefore, minimum effort control is applied to both controls during that period. For the next period up to point 2, H_{u_2} stays at zero and $H_{u_1}>0$. A singular arc control for u_2 and minimum effort control for u_1 are applied during this period. The singular arcs for both controls exist at the final period represented as a straight line along $H_{u_1}=H_{u_2}=0$.

Based upon the work of Lee and Ramirez (1992), general functional forms for the specific growth rate, μ, the foreign protein production rate, R_{fp}, and shock and recovery parameters, k_1 and k_2, for recombinant bacteria systems are as follows. For the specific growth rate:

$$\mu=\frac{\mu_{\max}x_3}{K_{C_N}+x_3+\dfrac{x_3^2}{K_S}}\{x_6+x_7R_R(x_5)\} \tag{31}$$

where $\mu_{\max}$ and K_{C_N} are Monod type constants and K_S is a substrate inhibition constant. The recovery ratio $R_R(x_5)$ is defined as:

$$R_R(x_5)=\frac{K_{C_I}}{K_{C_I}+x_5} \tag{32}$$

The foreign protein production ratio is described as:

$$R_{fp}=\left(\frac{f_{\max}x_3}{K_{C_N}+x_3+\dfrac{x_2^3}{K_S}}\right)\left(\frac{f_{I_O}+x_5}{K_I+x_5}\right) \tag{33}$$

where $f_{\max}$ and K_I are Monod type constants and f_{I_O} is a constant.

The shock and recovery parameters are:

$$k_1=\frac{k_{11}x_5}{K_{IX}+x_5} \tag{34}$$

$$k_2=\frac{k_{22}x_5}{K_{IX}+x_5} \tag{35}$$

where k_{11}, k_{22}, and K_{IX} are constants.

For this class of problems, the nutrient singular arc relation of Eq. 28 becomes:

$$\frac{\left(K_{C_N}-\dfrac{x_3^2}{K_S}\right)}{\left(K_{C_N}+x_3+\dfrac{x_3^2}{K_S}\right)^2}\left[\mu_{\max}\{x_6+x_7R_R(x_5)\}(\lambda_2-Y^{-1}\lambda_3)+f_{\max}\frac{f_{IO}+x_5}{K_I+x_5}\lambda_4\right]=0 \tag{36}$$

which yields a singular nutrient concentration specified by:

$$x_3^{\text{opt}}=\sqrt{K_SK_{C_N}} \tag{37}$$

The nutrient feeding rate along this singular arc is:

$$u_1=\frac{x_3^{\text{opt}}}{C_n^f-x_3^{\text{opt}}}u_2+\frac{Y^{-1}\mu(x_3^{\text{opt}},\,x_5,\,x_6,\,x_7)x_1x_2}{C_n^f-x_3^{\text{opt}}} \tag{38}$$

The optimal concentration of inducer along its singular arc (Lee, 1992) is described by:

$$\frac{a_1}{(K_{C_I}+x_5)^2}+\frac{a_2}{(K_I+x_5)^2}+\frac{a_3}{(K_{IX}+x_5)^2}=0 \tag{39}$$

with

$$a_1=-x_7(\lambda_2-Y^{-1}\lambda_3)\frac{\mu_{\max}x_3}{K_{C_N}+x_3+\dfrac{x_3^2}{K_S}}K_{C_I},$$

$$a_2=x_2\lambda_4\frac{f_{\max}x_3}{K_{C_N}+x_3+\dfrac{x_3^2}{K_S}}(K_I-f_{IO}),$$

and

$$a_3=-x_6\lambda_6k_{11}K_{IX}+(1-x_7)\lambda_7k_{22}K_{IX}.$$

Using Eqs. 38 and 29, we compute the singular controls in terms of the singular arc nutrient concentration, x_3^{opt}, and the singular arc inducer concentration, x_5^{opt}, for protein production by recombinant bacteria to be:

$$u_1=\frac{Y^{-1}\mu(x_3^{\text{opt}},\,x_5^{\text{opt}},\,x_6,\,x_7)x_1x_2(C_i^f-x_5^{\text{opt}})+x_1\dot{x}_5^{\text{opt}}x_3^{\text{opt}}}{C_n^fC_i^f-C_n^fx_5^{\text{opt}}-C_i^fx_3^{\text{opt}}} \tag{40}$$

$$u_2=\frac{Y^{-1}\mu(x_3^{\text{opt}},\,x_5^{\text{opt}},\,x_6,\,x_7)x_1x_2x_5^{\text{opt}}+x_1\dot{x}_5^{\text{opt}}(C_n^f-x_3^{\text{opt}})}{C_n^fC_i^f-C_n^fx_5^{\text{opt}}-C_i^fx_3^{\text{opt}}} \tag{41}$$

These are important new quantitative results that are certainly not obvious from the original problem statement. They apply to all problems described by the assumed functional forms for protein production by recombinant bacteria that have been shown valid for one system by Lee and Ramirez (1992).

Existence of singular arcs

The optimal concentration of nutrient, x_3^{opt}, can exit over the whole period of operation (Eq. 37). Therefore, the nutrient feed singular arc (Eq. 28) is satisfied as long as the nutrient level is kept at a specified level. By substituting Eq. 28 into the costate relation (Eq. 12), we have:

$$\lambda_3=e^{\int_0^t\frac{u_1+u_2}{x_1}dt}\lambda_3(0) \tag{42}$$

By combining this with the transversality condition (Eq. 19), we have:

$$\lambda_3(t)=0,\quad \text{where } t_i\le t\le t_f. \tag{43}$$

Table 1. System Parameters

Parameter	Value	Parameter	Value
μ_{max}	0.407 h^{-1}	K_{C_N}	0.108 g/L
K_{C_I}	0.22 g/L	k_{11}	0.09 h^{-1}
k_{22}	0.09 h^{-1}	K_{IX}	0.034 g/L
K_S	14,814.8 g^2/L^2	f_{max}	0.095 h^{-1}
f_{IO}	0.0005 g/L	K_I	0.022 g/L
C_n^f	100 g/L	C_i^f	4 g/L
Y	0.51		

For both controls to follow singular arcs, H_{u_1} and H_{u_2} must be both equal to zero; therefore:

$$x_1 = -\frac{C_i^f}{Q}\lambda_5, \quad \text{where } t_i \le t \le t_f \text{ and } Q \neq 0. \tag{44}$$

From the fact that λ_5 should be equal to zero at the final time (transversality condition, Eq. 21), then the reactor volume, x_1, would be equal to zero at the final time if Q is not zero. Because this statement cannot be true, the singular arcs for both controls cannot exist over the entire period of operation when the relative constant parameter Q is not zero. Both singular arcs could exist simultaneously over a partial period of operation. For the case when we do not consider the cost of the inducer, $Q=0$, then the singular arcs for both controls can exit over the whole period of operation. Under these conditions, λ_5 is always zero.

Optimal production of Beta-galactosidase

Let us consider the specific problem studied by Lee and Ramirez (1992), that of the production of β-galactosidase by *E. coli* D1210 and plasmid pSD8. Isopropylthiogalactoside (IPTG) was used as the inducer. All the system parameters for this bacteria system are given in Table 1. The initial conditions for the system are:

$$x_1(0) = 1 \tag{45}$$

$$x_2(0) = 0.1 \tag{46}$$

$$x_3(0) = 40 \tag{47}$$

$$x_4(0) = 0 \tag{48}$$

$$x_5(0) = 0 \tag{49}$$

$$x_6(0) = 1 \tag{50}$$

$$x_7(0) = 0 \tag{51}$$

The numerical procedure used for the calculation of the optimal operational policy is as follows:

(1) Assume a control law for the inducer feeding rate.

(2) Integrate the state equations forward in time with the assumed inducer control keeping the nutrient concentration at the value that maximizes the nominal specific growth rate.

(3) Integrate the costate equations backwards in time while calculating the optimal inducer concentration profile by Eq. 39.

(4) Integrate the state equations forward while calculating the optimal control laws by Eqs. 40 and 41 to maintain singular arc concentrations of inducer and nutrient. If the control laws are not feasible (violate upper and lower limits), then bang-bang saturated values are used.

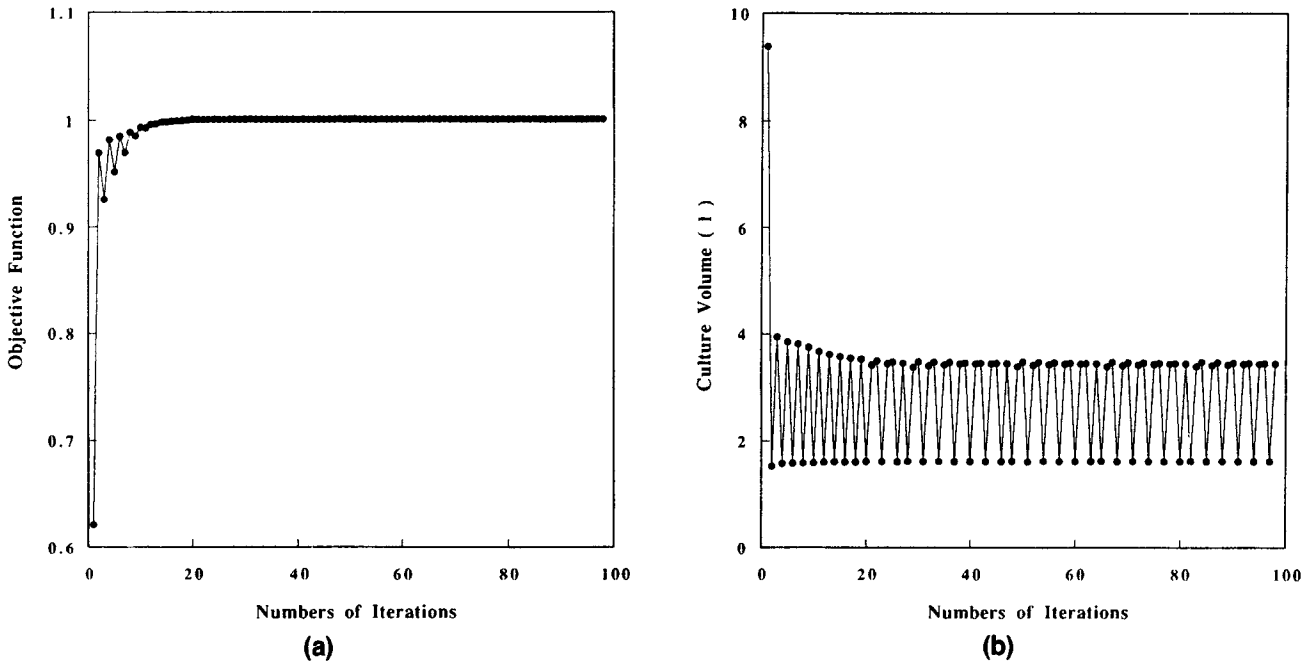

Figure 2. Optimal control convergence.

(a) Convergence of algorithm for the case when the relative cost parameter Q is zero; (b) plot of culture volume state (x_1) vs. iteration number shows that convergence has not been obtained.

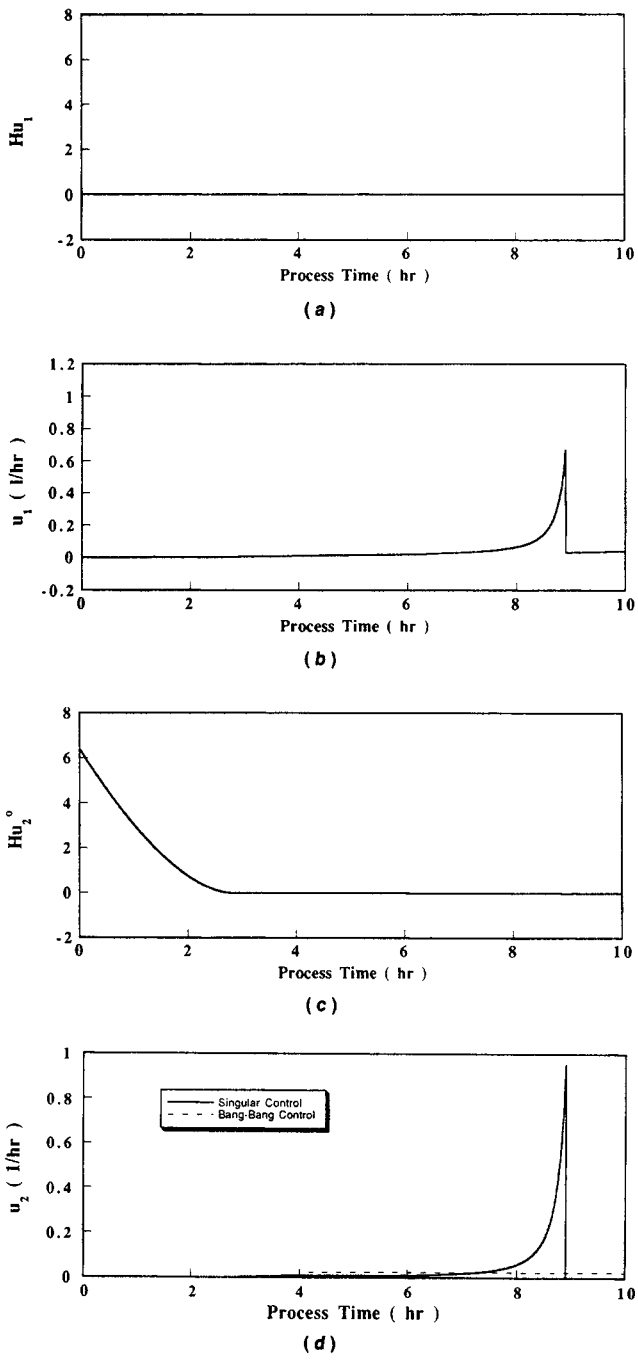

Figure 3. Optimal control policies ($Q=0$).

(a,b) Nutrient feed policy; (c,d) inducer feed policy.

(5) Iterate by returning to step 3 until the control policy converges.

Figure 2a gives the convergence of the algorithm for the case when the relative cost parameter Q is zero. The solution seems to converge after 20 iterations. However, a plot of the culture volume state (x_1) vs. iteration number (Figure 2b) shows that convergence has not been obtained. There is a persistent oscillation which causes large changes in the culture volume although the objective function is nearly constant. The reason for this lack of convergence was determined by studying the control strategies causing the culture volume to oscillate. There were two different shapes to the inducer control policies. The policies were similar in that they are composed of three regions which are an initial bang-bang region, a singular arc region, and a final bang-bang region. It is the last bang-bang period that caused the oscillations and lack of convergence. For one case, the sign of $H_{u_2}(t)$ is positive in this region, which results in maximum effort control. For the other case, the sign is negative, which results in minimum effort control. To resolve this convergence problem, only one extremal control (maximum or minimum) is allowed for the inducer during the last period regardless of sign changes in H_{u_2}. Convergence was easily obtained for both of these problems. For minimum inducer control along this final bang-bang period, the converged maximum objective functional was 1.0012 with an inducer input volume of 1.63 L. For maximum inducer control along this period, the objective functional was 1.0011, and the inducer input volume was 3.42 L. There is not a large difference in the two objective functionals even though the minimum effort control is preferred. This is especially true when considering plant operational cost associated with smaller culture volumes. The reason for this lack of objective functional sensitivity is that the foreign protein production rate and the cell growth inhibition rate were almost saturated with a high concentration of inducer before the last bang-bang control period. This specific convergence problem does not exist when we include the cost of inducer in the objective functional ($Q \neq 0$).

For the case when the cost of inducer can be neglected ($Q=0$), Figure 3 shows the optimal control policies for both nutrient feed and inducer feed rates. The nutrient feed policy (Figure 3b) always stayed on its singular arc, as shown by the fact that H_{u_1} always stayed zero (Figure 3a). The inducer policy is composed of three distinct periods (Figure 3d). For the initial period (0 to 2.88 h), there is minimum control effort. During this period, H_{u_2} was positive (Figure 3c). Physically, this period allows the number of cells to grow at their maximum specific growth rate without the adverse affect of the inducer.

During the middle of the reactor operation (2.88 to 8.9 h), both inducer and nutrient controls were operated along singular arcs. Along the singular arc, inducer was added in an exponential fashion as long as the rate does not exceed the maximum value, $u_{2_{max}}$. During this singular region, the control law maximized the foreign protein production rate by regulating both the number of cells and the foreign protein amount within the cells. In the early stages of singular arc operation, the control action increased the number of cells. During the later stage of singular arc operation, the control action increased the amount of foreign protein within cells.

The last inducer control period is one of minimum effort control (8.9 to 10 h). The fact that H_{u_2} is almost zero (Figure 3c) during this period shows the insensitivity of the objective functional to the two possible extremal control actions. No further inducer is required during this period since the cells are saturated with foreign protein.

Figures 4a and 4b provide the inducer control results when the relative cost of inducer to the value of product is 5 ($Q=5$). The nutrient control always stayed on its singular arc. The inducer control again exhibits three distinct periods, which are an initial period of not feeding the inducer, followed by a singular arc period, and finally a terminal period of not feeding the inducer. In comparing these results to the results with $Q=0$ (Figure 3d), we see that there are significant differences in both

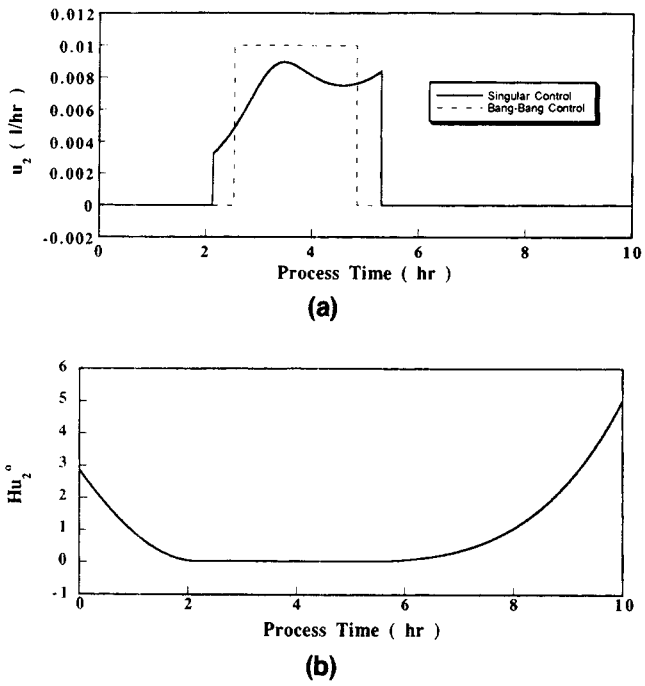

Figure 4a,b. Optimal control policies ($Q=5$).

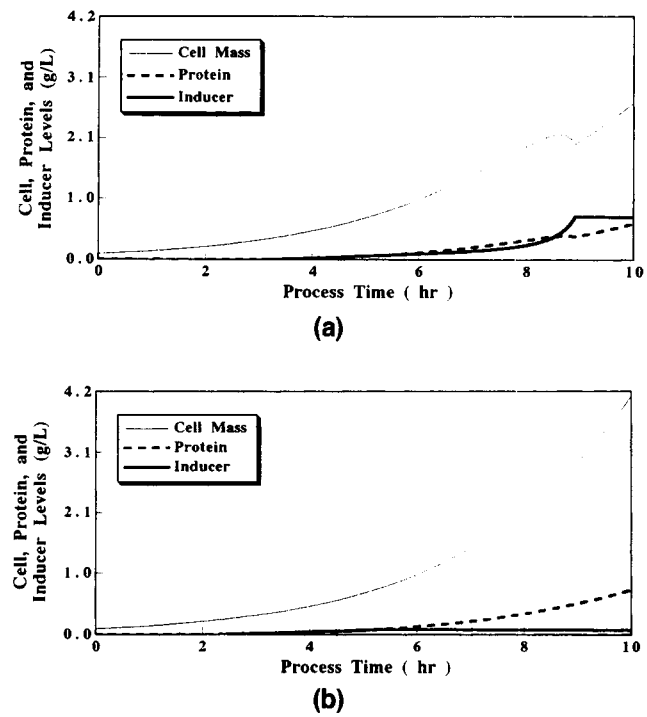

Figure 5a,b. Process variables under optimal control.

the time for singular control as well as the singular feeding policy itself.

Again, convergence to the optimal solution for the case when $Q \neq 0$ is somewhat difficult. When the sign of H_{u_2} is changed from plus to minus, the singular inducer control is first implemented. To find the time to end this singular arc, the amount of inducer injected was considered. If the total input volume of inducer exceeds an assumed total inducer volume, the singular arc is stopped. By a trial and error method of increasing or decreasing the assumed input inducer volume in the direction to increase the objective functional, we obtained our final optimal results.

Figures 5a and 5b show the major process variables for optimal operation when the cost of inducer is negligible ($Q=0$) and when there is a significant cost associated with the inducer ($Q=5$). There is nine times more inducer used for the optimal $Q=0$ case than for the optimal $Q=5$ case. There is a 10% reduction in the biomass for the $Q=0$ case which is due to the reduction in the growth rate when large amounts of inducer are used. The increased usage of inducer results in an 18% increase in the total protein product that was produced under optimal $Q=0$ conditions as compared to optimal $Q=5$ conditions.

Suboptimal bang-bang inducer control

Since it is sometimes difficult to regulate along a singular arc, we investigate the possibility of adding inducer using only extremal full on-off control of the inducer feeding rate. The glucose nutrient concentration was maintained constant at its optimal value. The computational algorithm we used to determine best inducer bang-bang policy is:

(1) Assume the following control law for the inducer feeding rate, where the number of subintervals are N: $u_2 = u_{2_{max}}$ for the subintervals located between j and $(j+m)$ with $(0<j, m)$ and $(j+m<N)$; $u_2 = u_{2_{min}}$ at the other subintervals.

(2) Give variations to both sides of the assumed control by increasing or decreasing the mesh locations by one. The result of the variations is nine control laws for the inducer feeding rate.

(3) Integrate the state equations forward nine times by using the nine control laws.

(4) Choose the best control law by comparing objective function values.

(5) Go to 2.

Table 2. Suboptimal Bang-Bang Control Policies for Various Maximum Inducer Flow Rates ($Q=0$)

$u_{2_{max}}$ (L/h)	Time Period When $u_2 = u_{2_{max}}$ (h)	Objective Function	IPTG Conc. at Final Time (g/L)	Glucose Input Volume (L)	IPTG Input Volume (L)
1	4.94-5.06	0.9740	0.3879	0.2289	0.1300
0.5	4.84-5.09	0.9761	0.3855	0.2281	0.1300
0.1	4.87-9.99	0.9833	1.0313	0.4786	0.5130
0.01	2.82-9.99	0.9811	0.2270	0.1940	0.0718
0.001	0-9.99	0.7109	0.0340	0.1810	0.0100

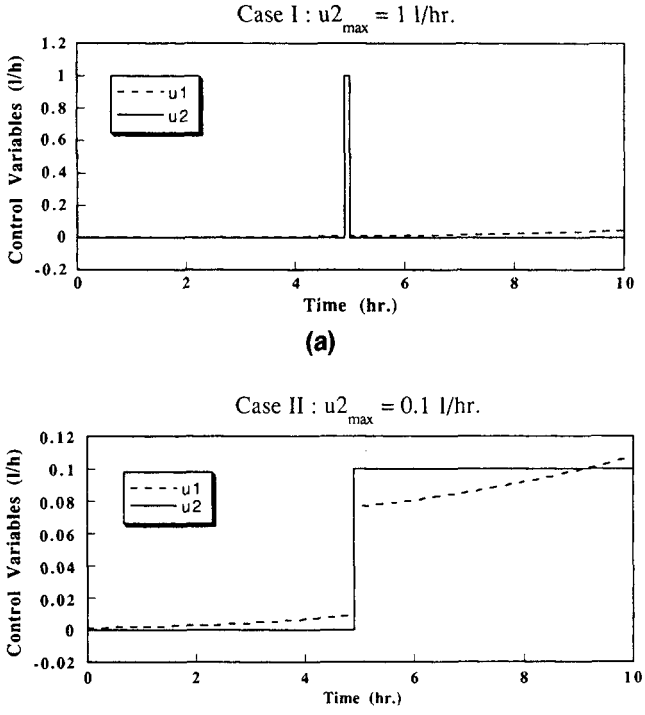

Figure 6. Bang-bang control for various minimum induction feed rates.

Case I (6a)—$u_{2_{max}} = 1.0$ L/h; Case II (6b)—$u_{2_{max}} = 0.1$ L/h.

Table 2 provides the results of investigating the effect of various maximum inducer feeding rates, $u_{2_{max}}$, on the best bang-bang suboptimal policies when the relative cost factor $Q = 0$. For high values of $u_{2_{max}}$, induction was delayed approximately five hours. This allowed for a fairly high biomass concentration before large quantities of inducer were fed to the system. For lower values of $u_{2_{max}}$, the delay period was shortened and the induction period increased. The biomass concentration does not have to be very high before induction with lower values of $u_{2_{max}}$. These two cases are illustrated in Figures 6a and 6b which show the results for $u_{2_{max}}$ of both 1 and 0.1. There exists a best value for the upper feeding rate $u_{2_{max}}$ that results in a maximum value for the objective functional. This value is $u_{2_{max}} = 0.025$ with a corresponding objective functional value of 0.9917. Since the performance does not increase with increasing values of $u_{2_{max}}$, we have clearly shown that bang-bang feeding is not optimal. Optimal control requires singular arc operation. However, the best suboptimal bang-bang control plotted in Figure 3c results in almost equivalent performance as compared with optimal singular arc operation. The objective functional for optimal singular operation is 1.0012 while that of suboptimal bang-bang operation is 0.9917, only a 0.95% decrease in productivity. For this specific bioreactor system, the operator is clearly justified in using the best suboptimal bang-bang inducer injection policy.

Table 3 gives the suboptimal inducer control results for $Q = 5$. The maximum effort control regions were reduced to use less amounts of inducer compared to the results for $Q = 0$. The best value for the maximum inducer feeding rate is $u_{2_{max}} = .01$ with $Q = 5$. The resulting best suboptimal performance functional is 0.7988 which compares favorably to the optimal singular arc value of 0.7899. Again the best suboptimal control, plotted on Figure 4a, is clearly nearly equivalent in performance to the true optimal policy.

Conclusions

Optimal operational strategies for maximizing the production of foreign protein by recombinant bacteria under the control of an inducible promoter have been determined. This class of problems involves two control variables, a glucose feeding policy, and an inducer injection policy. Since the problem is linear in the control variables, the optimal control is bang-bang or singular. Singular solutions are shown to exist and computational algorithms unique to this class of problems are developed for their realization. The singular solution for the glucose feeding policy allows the system to seek its maximum growth rate. The inducer injection policy has three distinct periods. The initial period is one with no induction. This delay in induction is due to growth considerations. Since the inducer affects the growth rate adversely, it is best to allow the culture density to increase before product expression is enhanced by inducible promotion. The second phase is a singular arc solution that allows for product enhancement without too great an adverse effect on cell growth. Finally, when the cells become saturated with inducer, the inducer injection is terminated. The optimal control algorithm developed in this work allows for the computation of optimal solutions for any specific problem that is in this general class and is the most common method of foreign protein expression.

We also present a new algorithm for the computation of best bang-bang inducer control strategies. Bang-bang strategies are easier to implement and preferable when there is not a severe loss in productivity. For the β-galactosidase example considered in the article, best bang-bang control is nearly equivalent to optimal singular operation.

Acknowledgment

This research was supported by the National Science Foundation under grant BCS-8911259.

Table 3. Suboptimal Bang-Bang Control Policies for Various Maximum Inducer Flow Rates ($Q = 5$)

$u_{2_{max}}$ (L/h)	Time Period When $u_2 = u_{2_{max}}$ (h)	Objective Function	IPTG Conc. at Final Time (g/L)	Glucose Input Volume (L)	IPTG Input Volume (L)
1	4.55-4.57	0.7777	0.1006	0.1813	0.0300
0.5	4.54-4.58	0.7820	0.0836	0.1811	0.0250
0.10	4.44-4.68	0.7829	0.0831	0.1809	0.0250
0.01	2.55-4.85	0.7988	0.0773	0.1732	0.0231
0.001	0.00-9.55	0.6620	0.0321	0.1808	0.0097

Notation

A = 2×2 matrix
B = $x \times 1$ vector
C_l^f = inducer level of feed, g/L
C_n^f = nutrient level of feed, g/L
f_{l0} = constant defined in Eq. 33
$f_{\max}$ = maximum protein production rate defined in Eq. 33
H = Hamiltonian defined in Eq. 9
H_{u_1} = function defined in Eq. 25
H_{u_2} = function defined in Eq. 26
$H_{u_2}^o$ = Hamiltonian defined in Eq. 46
k_1 = shock rate constant defined in Eq. 34
k_{11} = constant defined in Eq. 34
k_2 = recovery rate constant defined in Eq. 35
k_{22} = constant defined in Eq. 35
K_{C_l} = constant defined in Eq. 32
K_{C_n} = constant defined in Eq. 36
K_{IX} = constant defined in Eq. 34 and Eq. 35
K_I = constant defined in Eq. 33
K_S = substrate inhibition constant defined in Eq. 31
Q = cost factor defined in Eq. 8
R_{fp} = foreign protein production rate, h^{-1}
R_R = recovery ratio defined in Eq. 32
t_i = initial time, h
t_f = final time, h
u_1 = inducer feeding rate, L/h
$u_{1_{\max}}$ = maximum inducer feeding rate, L/h
u_2 = nutrient feeding rate, L/h
u_2^o = modified nutrient feeding rate defined in Eq. 47, L/h
$u_{2_{\max}}$ = maximum nutrient feeding rate, L/h
x_1 = culture volume, L
x_2 = cell mass, g/L
x_3 = nutrient level insider reactor, g/L
x_3^{opt} = optimal nutrient level inside reactor, g/L
x_4 = protein level inside reactor, g/L
x_5 = inducer level inside reactor, g/L
x_5^{opt} = optimal inducer level inside reactor, g/L
x_6 = shock rate effect
x_7 = recovery rate effect
Y^{-1} = inverse of growth yield coefficient (produced cell mass/consumed nutrient)

Greek letters

λ_i = costate variables defined in Eqs. 10–16
μ = specific growth rate, h^{-1}
$\mu_{\max}$ = maximum specific growth rate, h^{-1}

Literature Cited

Bell, D. J., and D. H. Jacobson, *Singular Optimal Control Problems*, Academic Press, Inc., New York (1975).

Bentley, W. E., "Effect of Plasmid-Mediated Activity on Bacterial Metabolism and Culture Stability," PhD Dissertation, University of Colorado, Boulder (1989).

Bentley, W. E., and D. S. Kompala, "A Novel Structured Kinetic Modeling Approach for the Analysis of Plasmid Instability in Recombinant Bacteria Cultures," *Biotech. Bioeng.*, **33**, 49 (1989).

Bryson, A. E., and Y. C. Ho, *Applied Optimal Control*, Hemisphere Publishing Co., Washington, DC (1975).

Georgiou, G., "Optimizing the Production of Recombinant Proteins in Microorganisms," *AIChE J.*, **34**, 1233 (1988).

Lee, J., "Modeling, State and Parameter Estimation, and Optimal Control of Induced Foreign Protein Production by Recombinant Bacteria in Fed-Batch Reactors," PhD Dissertation, University of Colorado, Boulder (1992).

Lee, J., and W. F. Ramirez, "Mathematical Modeling of Induced Foreign Protein Production by Recombinant Bacteria," *Biotech. Bioeng.*, **39**, 635 (1992).

Menawat, A., R. Mutharasan, and D. R. Coughanowr, "Singular Optimal Control Strategy for a Fed-Batch Bioreactor: Numerical Approach," *AIChE J.*, **33**, 776 (1987).

Park, S., and W. F. Ramirez, "Optimal Production of Secreted Protein in Fed-Batch Reactors," *AIChE J.*, **34**, 1550 (1988).

Park, S., and W. F. Ramirez, "Effect of Transcription Promoters on the Optimal Production of Secreted Protein in Fed-Batch Reactors," *Biotech. Prog.*, **6**, 311 (1990).

Park, T. H., J. Seo, and H. C. Lim, "Optimization of Fermentation Processes Using Recombinant *Escherichia coli* with the Cloned *trp* Operon," *Biotech. Bioeng.*, **34**, 1167 (1989).

Peretti, S. W., and J. E. Bailey, "Mechanistically Detailed Model of Cellular Metabolism for Glucose-Limited Growth of *Escherichia coli* B/r-A," *Biotechn. Bioeng.*, **28**, 1672 (1986).

Ramirez, W. F., *Process Control and Identification*, Academic Press (1994).

Sage, A. P., and C. C. White, *Optimum Systems Control*, Prentice Hall, Englewood Cliffs, NJ (1977).

Shu, J., and M. L. Shuler, "A Mathematical Model for the Growth of a Single Cell of *E. coli* on a Glucose/Glutamine/Ammonium Medium," *Biotech. Bioeng.*, **33**, 1117 (1989).

Zabriskie, D. W., and E. J. Arcuri, "Factors Influencing Productivity of Fermentations Employing Recombinant Microorganisms," *Enzyme Microb. Technol.*, **8**, 706 (1986).

Manuscript received Feb. 26, 1992, and revision received July 15, 1993.