Optimal Production of Secreted Protein in Fed-Batch Reactors

The optimal control policy for the maximization of the secreted heterologous protein in a fed-batch bioreactor was obtained using SEY2102 as a model host yeast and SUC2-s2 as a model secretory protein. A dynamic model for host cell growth, gene expression, and the secretion of expressed polypeptides was formulated. The optimal system trajectory contains multiple singular arcs that are distinct from one another. Optimal control requires transitions between these singular arcs. Optimal transitions between multiple singular arcs with bounded controls are uniquely located in observance with the Minimum Principle of Pontryagin. An iterative numerical search strategy for determining the optimal control showed successful convergence properties.

Seujeung Park and W. Fred Ramirez
Department of Chemical Engineering
University of Colorado
Boulder, CO 80309

Introduction

The ability of yeast cells to secrete proteins with post-translational modifications has led to a number of commercially useful heterologous proteins produced in baker's yeast and secreted as their mature forms (Das and Shultz, 1987). The secretion pathway and the major macromolecular events were identified by Schekman and his associates (Novick et al., 1981). The close analogy of the secretion machinery between yeast cells and mammalian cells makes the baker's yeast a potential organism for mass production of useful mammalian proteins that need sophisticated post-transitional modifications such as glycosylation and phosphorylation. The differences between yeast and mammalian cells are that the secretory organelles in yeast cells are smaller (Schekman et al., 1982) and the structures of the carbohydrates attached to the glycosylation sites are much simpler than that of mammalian cells (Lehle et al., 1979). The lack of pathogenity of baker's yeast with humans makes it an attractive choice for producing therapeutic enzymes and neuropeptides, for which purification and removal of toxins are important. For this category of protein production, yeast secretion can improve the efficiency of product recovery considerably.

Even for proteins that do not need a post-translational change, advantages of initial separation and ease of downstream processing can be obtained. Unlike cytoplasmically produced proteins, secretory proteins need additional processing after translation to be in the final product form. The rate of secretion processing depends strongly on the metabolic activity of the host yeast cell. A dynamic model for the secretion of heterologous polypeptides in the yeast cell has been formulated based on the major macromolecular interactions between the polypeptides and the host cell's secretion machinery (Park and Ramirez, 1988). Post-translational processing, which converts the nascent form of polypeptides to the final mature form, offers an additional factor that needs to be considered in optimal operation of protein-producing fermentations. In this work, using a mutant secretory invertase as a model heterologous protein, optimal control theory is used to find the best operating condition for maximizing the production of the secreted protein in a fed-batch reactor.

Application of optimization techniques to bioreactors have been reported frequently. The problem of maximum harvest of bacteria in a continuous reactor was solved by D'Ans et al. (1972). A singular unconstrained control for maximum harvest of baker's yeast in a fed-batch reactor was obtained by applying Jacobson's ϵ-method by Menawat et al. (1987). Optimal temperature and pH control were determined for batch pennicilin fermentation (Constantinides et al., 1970). Also a predesigned suboptimal temperature profile for maximization of pennicilin level in a batch reactor was tried (King et al., 1974) by transforming the suboptimization to a problem of parameter identification. Optimal start-up and operation of a continuous amino acid fermentation were solved by Takamatsu et al. (1975). Modak et al. (1986) discussed characteristics of optimal feed rate profiles for various fed-batch fermentation processes. Hasegawa et al. (1987) attempted an optimal periodic operation for a bioreactor that produces metabolites that follow the Leudeking-Piret model.

Material and System Equations

The yeast strain used is SEY2102-s2I that contains the SUC2-s2 gene which is constituted by one codon substitution of yeast secretory invertase and contains no other linked SUC gene

family (Schauer et al., 1985). The heterology created by this one codon mutation is regarded to simulate well a foreign protein cloned in the yeast cell in terms of secretion. A detailed description of the mathematical model and experimental system used is given in our previous work (Park and Ramirez, 1988). The yeast secretion dynamic model has the following structure.

$$\frac{dP_M}{dt} = \Phi(\mu_x)P_I \quad \text{where } \Phi(\mu_x) = \frac{\alpha\mu_x}{\beta + \mu_x} \tag{1}$$

Here the secretion rate constant Φ depends on the host cell's specific growth rate in a saturation manner. For the product of the SUC2-s2 gene at 27°C, the model parameters were identified as $\alpha = 4.75$ and $\beta = 0.12$. The secretion dynamic expression is the result of modelling the major macromolecular events in each secretory organelle as compartmental reactions followed by vesicular transport to the next destination. The secretion model structure and parameters for SUC2-s2 were confirmed experimentally with step input dynamic response tests using cycloheximide. Solid phase enzyme immunoassay was used for the quantitation of the protein concentration.

The transcription of SUC2-s2 by SUC2-promoter is under catabolic repression (Carlson and Botstein, 1982) and becomes depressed at low glucose concentration—below about 0.7 g/L. The dependency of the SUC2-s2 expression rate, f_p, of an average cell at 27°C in the Whickerham's minimal media with glucose as the sole carbon source is formulated as:

$$\frac{1}{X}\frac{dP_T}{dt} = f_p(S) \quad \text{where} \quad f_p(S) = \frac{Se^{-5.0S}}{0.1 + S} \tag{2}$$

The specific growth rate of the host cell, SEY2102 at 27°C in the minimal media was formulated using the Monod-Haldane model structure with parameters of:

$$\mu_x(S) = \frac{21.87S}{(S + 0.4)(S + 62.5)} \tag{3}$$

The kinetics of ethanol inhibition on yeast growth were investigated by Aiba et al. (1968). They reported that the accumulation of ethanol in the culture reduces the specific growth rate of the baker's yeast by the empirically determined factor of $\exp(-kE)$, where the value of k is reported to be about 0.02. We discarded the effect of accumulated ethanol on the cell growth rate to avoid additional complexity since this does not change the control strategy for this problem. Substrate inhibition on cell growth is observed when the medium is highly concentrated with sugar. The Monod-Haldane model accounts for this by multiplying the Monod expression by an additional factor. This additional factor makes the problem more practical by bounding the sugar concentration. We have chosen the glucose concentration that supports the maximum growth rate as 5.0 g/L, above which an increase in glucose concentration does not contribute significantly to the cell growth rate. The model structure of the protein expression rate and the cell growth rate change with a different choice of the promoter in the transcriptional control region of the cloned gene and the choice of the substrate in the media. Accordingly, the optimal control strategy will be different with these changes in the model structure. Here, glucose was chosen as the substrate and the dependency of those rates on the glucose were formulated as mentioned earlier.

With the feed flow rate into the reactor as the only control action, the state variables in a fed-batch reactor can be described by the following set of equations. The yield of cell mass/glucose is assumed to be constant and the glucose concentration of the feed is constant. The physically realizable feed flow rate is bounded. The lower limit of control action is zero and the upper limit is determined by the choice of feed system. Under isothermal operation, the rate functions $\Phi(\mu_x)$, f_p, and μ_x were approximated as unique functions of the glucose concentration in the culture. The system equations with the initial conditions can be expressed as

$$\dot{P}_M = A(S)(P_T - P_M) - \frac{q}{V}P_M \qquad P_M(t_o) = 0.0 \tag{4}$$

$$\dot{P}_T = B(S)X - \frac{q}{V}P_T \qquad P_T(t_o) = 0.0 \tag{5}$$

$$\dot{X} = C(S)X - \frac{q}{V}X \qquad X(t_o) = X_o \tag{6}$$

$$\dot{S} = -Y\,C(S)X + \frac{q}{V}(m - S) \qquad S(t_o) = S_o \tag{7}$$

$$\dot{V} = q \quad \text{where } q_{\min} \le q \le q_{\max} \qquad V(t_o) = V_o \tag{8}$$

with

$$A(S) = \Phi(\mu_x) = \frac{4.75C(S)}{0.12 + C(S)} \tag{9}$$

$$B(S) = f_p(S) = \frac{Se^{-5.0S}}{0.1 + S} \tag{10}$$

$$C(S) = \mu_x(S) = \frac{21.87S}{(S + 0.4)(S + 62.5)} \tag{11}$$

Problem Formulation

The objective is to maximize the total secreted SUC2-s2 product in the reactor at the end of each fed-batch running for a specified final time. Thus, the objective function to be minimized is formulated as

$$J = -P_M(t_f)V(t_f) \tag{12}$$

The necessary conditions for optimization of this system is determined by the Pontryagin's Minimum Principle (Pontryagin et al., 1964). The associated Hamiltonian H can be represented by

$$H = \lambda_1\left[A(S)(P_T - P_M) - \frac{q}{V}P_M\right] + \lambda_2\left[B(S)X - \frac{q}{V}P_T\right] + \lambda_3\left[C(S)X - \frac{q}{V}X\right]$$

$$+ \lambda_4\left[-YC(S)X + \frac{q}{V}m - \frac{q}{V}S\right] + \lambda_5 q \quad (13)$$

$$= \lambda_1 A(S)(P_T - P_M) + \lambda_2 B(S)X + \lambda_3 C(S)X - \lambda_4 YC(S)X + \frac{q}{V}[-\lambda_1 P_M - \lambda_2 P_T - \lambda_3 X + \lambda_4(m - S) + \lambda_5 V] \quad (14)$$

The minimum principle holds along an optimal trajectory, for all admissible controls, $q(t)$:

$$H(x, q_{\text{optimal}}, \lambda) \leq H(x, q, \lambda) \quad (15)$$

The costate variables are defined by the following differential equations. The boundary conditions are given by the final transversality conditions:

$$\dot{\lambda}_1 = \lambda_1 A(S) + \lambda_1 \frac{q}{V} \qquad \lambda_1(t_f) = -V(t_f) \quad (16)$$

$$\dot{\lambda}_2 = -\lambda_1 A(S) + \lambda_2 \frac{q}{V} \qquad \lambda_2(t_f) = 0.0 \quad (17)$$

$$\dot{\lambda}_3 = -\lambda_2 B(S) - \lambda_3 C(S) + Y\lambda_4 C(S) + \lambda_3 \frac{q}{V} \qquad \lambda_3(t_f) = 0.0 \quad (18)$$

$$\dot{\lambda}_4 = -\lambda_1 A'(S)(P_T - P_M) - \lambda_2 B'(S)X - \lambda_3 C'(S)X + Y\lambda_4 C'(S)X + \lambda_4 \frac{q}{V} \qquad \lambda_4(t_f) = 0.0 \quad (19)$$

$$\dot{\lambda}_5 = -\frac{q}{V^2}(\lambda_1 P_M + \lambda_2 P_T + \lambda_3 X + \lambda_4 S - \lambda_4 m) \qquad \lambda_5(t_f) = -P_M(t_f) \quad (20)$$

The Hamiltonian is linear with respect to the control function, and can be represented in the following form.

$$H(x, \lambda, q) = \Psi(x, \lambda) + qH_q(x, \lambda) \quad (21)$$

where

$$\Psi = \lambda_1 A(S)(P_T - P_M) + \lambda_2 B(S)X + \lambda_3 C(S)X - \lambda_4 YC(S)X \quad (22)$$

$$H_q = \frac{1}{V}[-\lambda_1 P_M - \lambda_2 P_T - \lambda_3 X + \lambda_4(m - S) + \lambda_5 V] \quad (23)$$

By the minimum principle, the optimal control function profile is determined by the following conditions:

if H_q stays at zero,	then $q = q_{\text{singular}}$
if $H_q \geq 0$,	then $q = q_{\min}$
if $H_q \leq 0$,	then $q = q_{\max}$

Analytical Approach

Before computing the optimal control action by a numerical solution of this system of five-state variables and five-costate variables, it will be helpful to analyze the behavior of each variable. The usual numerical iterative algorithm for determining the optimal control finds or updates the control action while integrating costates backward from the final time. Qualitative information about the system behavior can be obtained in the same manner. The final transversality conditions for the costates makes H_q zero at $t = t_f$. If there are no physical limits on the permissible control action, singular control can be achieved for the whole span of time during the fermentation. The optimal control profile is then the solution of the following equation.

$$H_q = 0 \quad \text{for } q \in [-\infty, \infty], t \in [0, t_f] \quad (24)$$

Since this does not contain any expression of control action, we evaluate the time derivative to get an explicit relationship for the control action. Since $H_q(t_f) = 0$, the following condition holds along a singular arc.

$$\frac{dH_q}{d(-t)} = 0 \quad (25)$$

Using the expressions for state and costate dynamics, we have

$$\frac{dH_q}{d(-t)} = \frac{1}{V}[\lambda_1 A'(S)(P_T - P_M) + \lambda_2 B'(S)X + (\lambda_3 - Y\lambda_4)C'(S)X](m - S) \quad (26)$$

where

$$A'(S) = \frac{\partial A(S)}{\partial S}, \quad B'(S) = \frac{\partial B(S)}{\partial S}, \quad \text{and} \quad C'(S) = \frac{\partial C(S)}{\partial S} \quad (27)$$

Since $(m - S) \neq 0$, the singular arc will be the solution of the following equation.

$$\lambda_1 A'(S)(P_T - P_M) + \lambda_2 B'(S)X + (\lambda_3 - Y\lambda_4)C'(S)X = 0 \quad (28)$$

The dependencies of rate functions $A'(S)$, $B'(S)$, and $C'(S)$ on sugar concentration are shown in Figure 1. The sugar concentration along the singular arc will be determined by the relative magnitudes of the three coefficients, $\lambda_1(P_T - P_M)$, $\lambda_2 X$, and $(\lambda_3 - Y\lambda_4)X$. Depending on the magnitudes of these coefficients, there is one or three roots of sugar concentration that satisfies Eq. 28. By physical insight into the dynamics of the state and costate variables represented in Eqs. 4–8 and 16–20, we can determine the changes in these coefficients and accordingly the shape of the singular arc during fermentation. Since the costate equations are homogeneous and linear, their behavior is quite predictable.

First coefficient, $\lambda_1(P_T - P_M)$. From the secretion dynamic model, the quantity $(P_T - P_M)$ cannot vanish once the cell starts to produce protein, and this term becomes a positive small value as the final time is reached. Due to the negative eigenvalue of

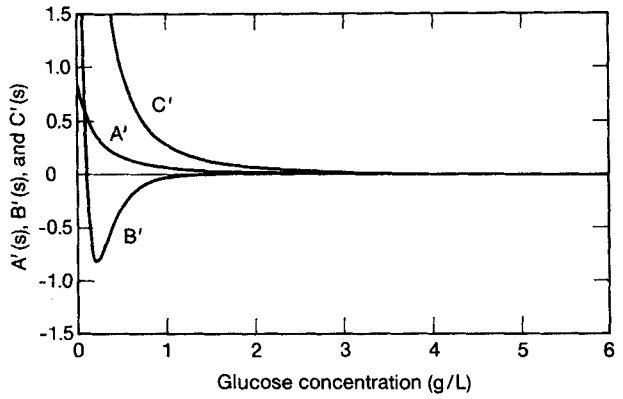


Figure 1. Rate functions $A'(S)$, $B'(S)$, and $C'(S)$ vs. glucose concentration.

the equation, the costate λ_1, from its final value of $-V(t_f)$, will approach zero asymptotically as it travels backwards.

$$\frac{d\lambda_1}{d(-t)} = -\lambda_1\left[A(S) + \frac{q}{V}\right] \tag{29}$$

Thus, this first coefficient stays negative and its magnitude is expected to shrink as t differs from t_f.

Second coefficient, $\lambda_2 X$. The costate λ_2 also stays negative.

$$\frac{d\lambda_2}{d(-t)} = \lambda_1 A(S) - \lambda_2\frac{q}{V} \tag{30}$$

It starts from zero at the final time, and the magnitude increases by λ_1. However, it also will approach zero asymptotically after λ_1 vanishes because of its negative eigenvalue.

Third coefficient, $(\lambda_3 - Y\lambda_4)X$. If we follow the singular arc, then costate λ_4 stays at zero according to Eq. 19. Along the singular arc represented as Eq. 28, the third coefficient has the following dynamics.

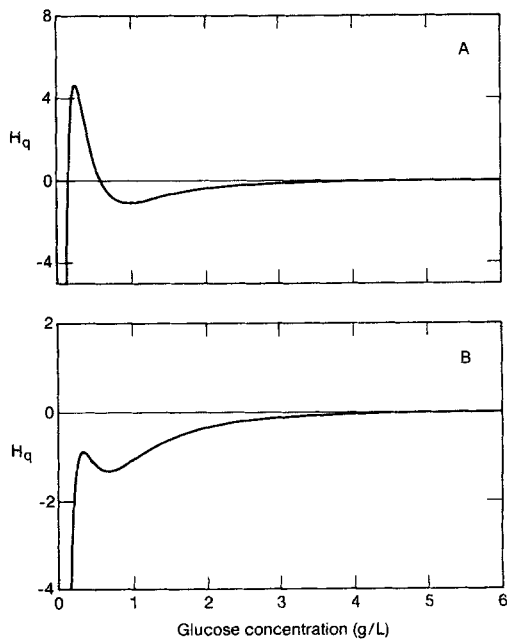
$$\frac{d(\lambda_3 - Y\lambda_4)}{d(-t)} = \lambda_2 B(S) + \left[C(S) - \frac{q}{V}\right](\lambda_3 - Y\lambda_4) \tag{31}$$

This coefficient begins with a negative sign due to the λ_2 term and stays negative. For a feed fresh media of high enough sugar concentration, $[C(S) - q/V]$ has a positive value that makes Eq. 31 increase exponentially as time progresses backward. The numerical approach shows that the magnitude of the costate λ_3 increases as it travels backward along the optimal trajectory, as will be discussed later. The heuristically predicted behavior of the three coefficients is confirmed by the numerical approach.

From this analysis, we get the general shape of these three coefficients as the costates evolve backward. Close to the final time, the first coefficient has a dominant magnitude over the other two. As time proceeds backward, the first coefficient asymptotically vanishes and the second coefficient becomes dominant for a while and then becomes asymptotically small. Finally, the third coefficient increases in magnitude and becomes dominant over the first and second coefficients. The physical meaning of this consecutive change in the dominance of these three coefficients can be stated. In our formulation of the model, $A(S)$, $B(S)$, and $C(S)$ characterize the dynamics of the protein secretion rate, the protein synthesis rate, and the host cell's growth rate, respectively. The solutions of the derivatives $A'(S) = 0$, $B'(S) = 0$, and $C'(S) = 0$ indicate the sugar concentration levels that support the maximum secretion rate, maximum protein synthesis rate, and the maximum cell growth rate, respectively. The solution of the sugar concentration that makes $H_q = 0$ is determined by the relative magnitudes of these three coefficients. The magnitudes of these coefficients change so that the solution of Eq. 28 favors first the protein secretion, second protein synthesis, third the host cell growth, as time proceeds backward from the final time. This means that the sugar concentration changes sequentially so that it favors the increase of cell mass for the first phase of culture, then favors the protein synthesis for the second phase, and finally favors secretion for the third phase close to the final time, as we operate the reactor.

There will be either one or three solutions for the sugar concentrations from Eq. 28, depending on the magnitudes of the three coefficients, Figure 2. During the time that there is only one solution, we need to choose that solution. However, during the time when there are multiple solutions, we have to select the one which is truly the optimal sugar concentration. To eliminate a nonoptimal solution or solutions, we apply additional necessary conditions, the generalized Legendre-Clebsch convexity condition. This was derived by using a special scalar control variation by Kelly (1964), and Kopp and Moyer (1965). The convexity condition requires that the optimal trajectory follow the relationship.

$$-\frac{\partial}{\partial q}\frac{d^2}{dt^2}\frac{\partial H}{\partial q} \geq 0 \tag{32}$$



A. Existence of multiple solutions of glucose concentration for $H_q = 0$.
B. Existence of a single solution of glucose concentration for $H_q = 0$.

Figure 2. H_q vs. glucose concentration.

After derivation, we have

$$-\frac{\partial}{\partial q}\frac{d^2}{dt^2}\frac{\partial H}{\partial q} = \left(\frac{m-S}{V}\right)^2 [\lambda_1 A''(S)(P_T - P_M) + \lambda_2 B''(S)X + (\lambda_3 - Y\lambda_4)C''(S)X] \geq 0 \quad (33)$$

where

$$A''(S) = \frac{\partial^2 A(S)}{\partial S^2}, \quad B''(S) = \frac{\partial^2 B(S)}{\partial S^2},$$

$$\text{and} \quad C''(S) = \frac{\partial^2 C(S)}{\partial S^2} \quad (34)$$

Since $[(m - S)/V]^2 \geq 0$, this condition becomes

$$[\lambda_1 A''(S)(P_T - P_M) + \lambda_2 B''(S)X + (\lambda_3 - Y\lambda_4)C''(S)X] \geq 0 \quad (35)$$

By comparing with Eq. 26, Eq. 35 is equivalent to

$$\frac{\partial}{\partial S}\frac{d}{dt}H_q \leq 0 \quad (36)$$

This condition eliminates the middle local extremal point in Figure 2 when there are three solutions to Eq. 28. As we will see later, this is the solution that results in a local maximum of the Hamiltonian instead of a minimum. Along the singular arc represented by Eq. 24, the Hamiltonian can be expressed as

$$H = \Psi \quad (37)$$

$$= \lambda_1 A(S)(P_T - P_M) + \lambda_2 B(S)X + (\lambda_3 - Y\lambda_4)C(S)X \quad (38)$$

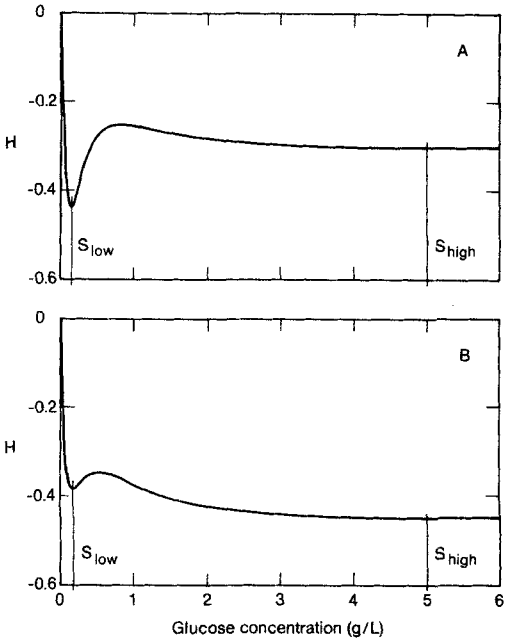
The singular arc, which is constituted by the control action so that $H_q = 0$, is the trajectory formed by the following equivalent condition.

$$\frac{\partial H}{\partial S} = 0 \quad (39)$$

And the convexity condition can be expressed equivalently as

$$\frac{\partial^2 H}{\partial S^2} \geq 0 \quad (40)$$

The optimal sugar concentration profile is the one that maintains the Hamiltonian at its minimum. As we have seen, the Hamiltonian has one or three local extremal points with respect to sugar concentration, depending on the magnitudes of the three coefficients. In Figure 3, the shape of the Hamiltonian is described. When there are multiple local minimum, the choice must be the sugar concentration that has the lowest value for the Hamiltonian. The generalized Legendre-Clebsch convexity condition simply eliminates the local maximum point from the possible candidates for the singular arc. The Hamiltonian changes its shape as it proceeds backwards from t_f. The location of the sugar concentration that maintains the minimum suddenly switches, once from S_{high} to S_{low} and later from S_{low} to S_{high} as shown in Figure 3. Due to the bounded permissible control



A. S_{low} is the optimum glucose concentration.
B. S_{high} is the optimum glucose concentration.

Figure 3. Hamiltonian vs. glucose concentration.

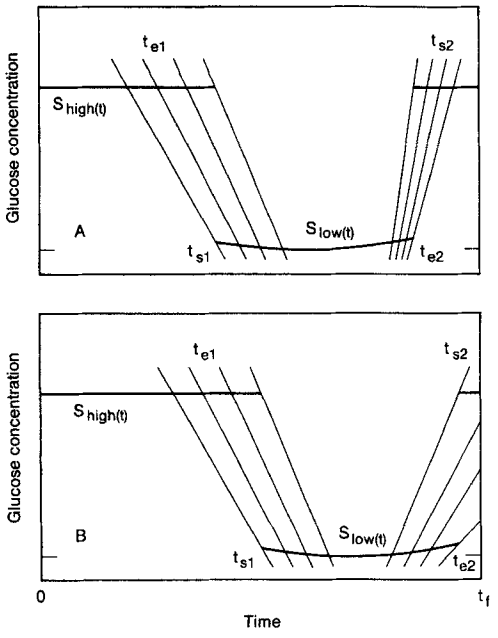
action in the range $[q_{min}, q_{max}]$, abrupt transitions in the culture sugar concentration are not physically possible. For an abrupt change in the sugar concentration from S_{low} to S_{high}, the system needs infinitely high feed flow rates which are not feasible. Also, it is impossible to achieve a sudden transition in the culture sugar concentration from S_{high} to S_{low} since we cannot selectively take only sugar from the culture. In this case, the only thing that can be done is to close the feed valve and to wait until the sugar is consumed by the cells.

The boundedness of the permissible control action forces the optimal trajectory to deviate from the singular arc during transitions in culture sugar concentration. Therefore, H_q will deviate from staying at zero. As we proceed backward from the final time, the Hamiltonian and its derivative with respect to the control function at each time can be represented as

$$H(x, \lambda, q) = \Psi(x, \lambda) + qH_q(x, \lambda) \quad (41)$$

$$H_q(t) = \int_t^{t_f} \frac{d}{d(-t)} H_q dt \quad (42)$$

As discussed before, $H_q(t_f)$ is always zero. When permissible, control action that allows the system to stay on the singular arc, keeping the $H_q = 0$, is the best policy. When the maintenance of $H_q = 0$ violates the permissible controls, then H_q begins to have values other than zero. The condition of minimization of the Hamiltonian requires bang-bang control, either $q = q_{max}$ or $q = q_{min}$ according to the sign of H_q. Since abrupt transitions between the two singular solutions of sugar concentration $S_{high}(t)$ and $S_{low}(t)$ are not feasible, the two singular arcs are connected by the arcs constituted by bang-bang control, Figure 4. There are an infinite number of possible transitions between the singular arcs through the bang-bang control mode, and the optimal trajectory will be the one that strictly observes the mini-



A. Complete transition from S_{low} to S_{high} occurs.
B. Complete transition from S_{low} to S_{high} may not occur.

Figure 4. Singular arc glucose concentration (heavy lines) and trajectories constituted by the bang-bang control during the transitions (light lines).

mum principle during the transition. Because $H_q(t) = 0$ along both singular arcs, the optimal transition between time $[t_1, t_2]$ will have the following properties:

$$\int_{t_1}^{t_2} \frac{d}{d(-t)} H_q \, dt = 0 \tag{43}$$

and, for any t between $[t_1, t_2]$,

$$q[t_1, t_2] = q_{max}, \quad \text{if} \int_t^{t_2} \frac{d}{d(-t)} H_q \, dt \le 0 \tag{44}$$

$$q[t_1, t_2] = q_{min}, \quad \text{if} \int_t^{t_2} \frac{d}{d(-t)} H_q \, dt \ge 0 \tag{45}$$

Numerical Search for the Optimal Trajectory

With the initial condition of state variables, feed stream glucose concentration, and the feed stream flow rate being specified, a numerical search for the optimal control policy is needed with the final time specified. The following algorithm is used.

Step 1. For an initial assumed control policy, the state equations are integrated forward in time until the final time is reached.

Step 2. The costate equations are then integrated backward in time from t_f and the singular arc sugar concentration that maintains the Hamiltonian at its minimum is calculated from

$$H(x, \lambda, q) = \Psi(x, \lambda) + qH_q(x, \lambda) \tag{46}$$

with

$$\Psi(x, \lambda) = \lambda_1 A(S)(P_T - P_M) + \lambda_2 B(S)X + \lambda_3 C(S)X - \lambda_4 YC(S)X \tag{47}$$

$$H_q(t) = \int_t^{t_f} \frac{d}{d(-t)} H_q \, dt \tag{48}$$

The singular arc sugar concentration profile is the solution of $\partial\Psi/\partial S = 0$. The singular arc sugar concentration at the final time is the solution of $A'(S_f) = 0$. When multiple solutions exist, the one that makes Ψ a minimum is chosen. The corresponding control action is calculated by using the fourth state equation.

$$\frac{dS}{d(-t)} = YC(S)X - \frac{q}{V}(m - S) \tag{49}$$

Since costate integration requires the control q, and Ψ is dependent upon the costates which are used to calculate the singular arc sugar concentration, the solution of S and the corresponding q are obtained by the simultaneous solution of Eqs. 16–20, 28 and 49. When the singular arc control action exceeds the control constraints, the actual control is constrained at the boundary. This bang-bang control mode also occurs when the singular trajectory requires a sudden transition of sugar concentration as the Hamiltonian switches location for the lowest local minimum. As discussed previously, there are infinite number of paths connecting distinct singular arcs by the bang-bang action as shown in Figure 4.

Step 3. During backward integration, the integral of the quantity $dH_q/d(-t)$ is calculated and used to find the optimal transition between singular arcs by the bang-bang action that strictly observes the minimum principle. To locate the optimal transition times t_{s1} and t_{s2} in Figure 4, a gradient method is used.

$$t_{s2,updated} = t_{s2,old} + K_2 \int_{t_{e2}}^{t_{s2}} \frac{d}{d(-t)} H_q \, dt \quad K_2 \ge 0 \tag{50}$$

$$t_{s1,updated} = t_{s1,old} + K_1 \int_{t_{e1}}^{t_{s1}} \frac{d}{d(-t)} H_q \, dt \quad K_1 \le 0 \tag{51}$$

Step 4. If t_{s2} exceeds t_f, bang-bang control begins right from the final time and the value of $S_{high}(t_f)$ which is the singular arc sugar concentration at the final time is no longer pertinent. However, $H_q(t_f)$ is still zero. To find the optimal final sugar concentration which satisfies the minimum principle, the following gradient method is used to locate the $S(t_f)$.

$$S(t_f)_{updated} = S(t_f)_{old} + K_S \int_{t_{e2}}^{t_f} \frac{d}{d(-t)} H_q \, dt \quad K_S \le 0 \tag{52}$$

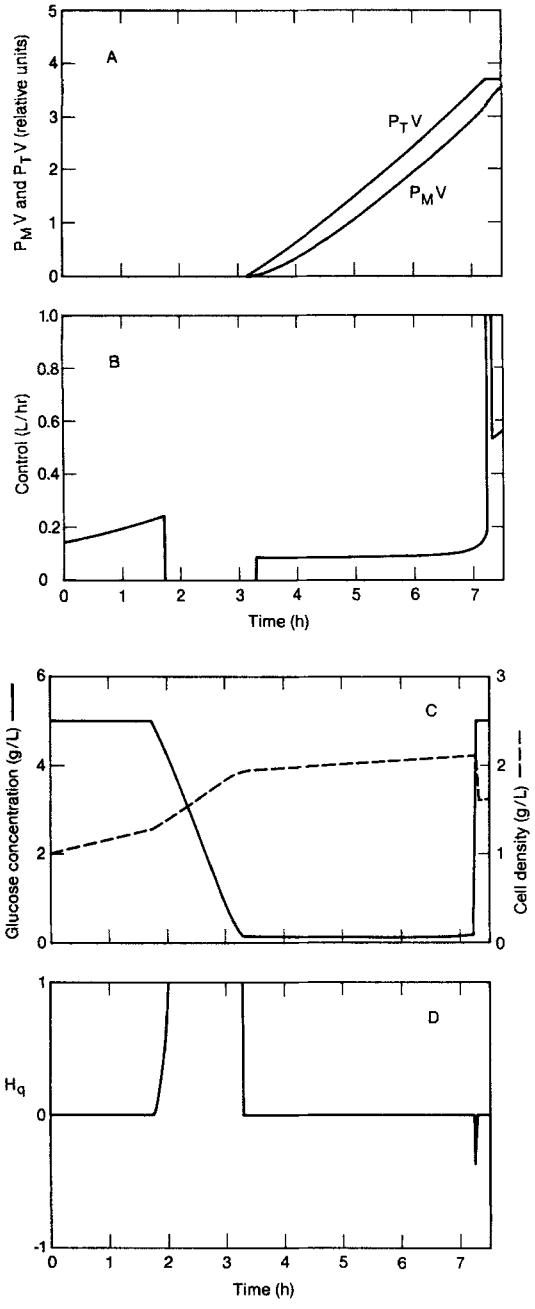
Step 5. With an updated control function, steps 1–4 are repeated until convergence is obtained.

Let us consider the following data for initial conditions of state variables and other system parameters.

$$P_M(t_o) = 0.0, \qquad P_T(t_o) = 0.0$$
$$X(t_o) = 1.0 \text{ g/L}, \qquad S(t_o) = 5.0 \text{ g/L}$$
$$V(t_o) = 1.0 \text{ L}, \qquad m = 20.0 \text{ g/L}$$
$$10.0 \ge q \ge 0.0 \text{ L/h} \qquad t_f = 7.5 \text{ and } 15.0 \text{ h}$$

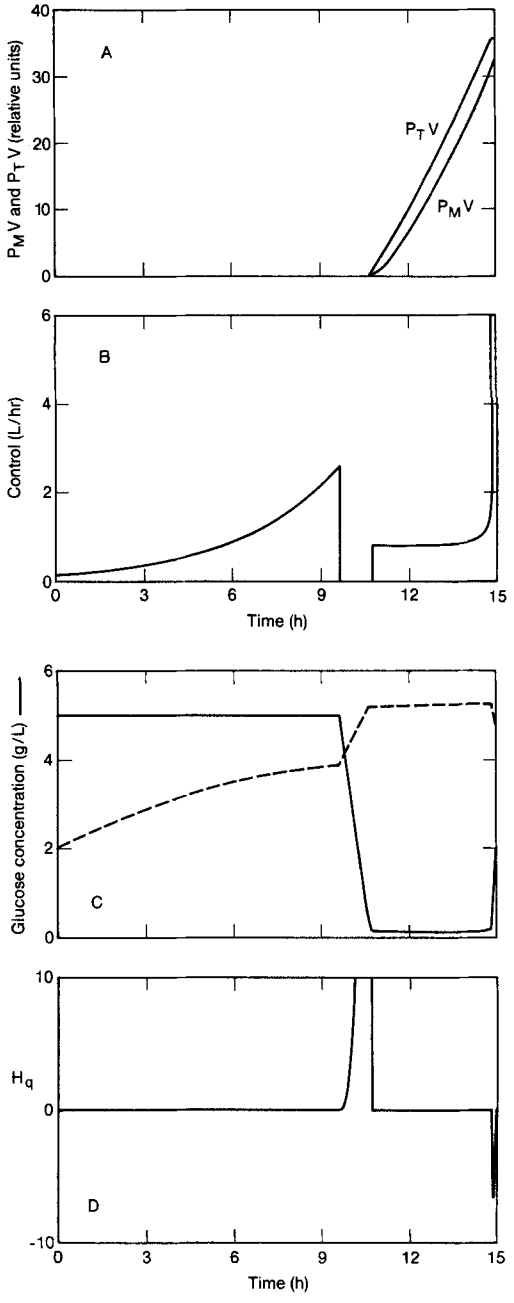
The iterative algorithm successfully converged to the optimal trajectory. The two switching times, t_{s1} and t_{s2}, are uniquely

located to satisfy the minimum principle. The optimal solutions with t_f = 7.5 h and t_f = 15.0 h are shown in Figures 5 and 6, respectively. For this system, the optimal control divides the whole fed-batch running into three consecutive phases which is clearly seen in the optimal sugar concentration profile of Figure 5C. The first phase, during which the glucose level is maintained at the value supporting maximum specific growth rate, is dedicated just for growing the host cell mass in the culture. The importance of this phase is that it is in the host cell where the



A. Optimal total expressed SUC2-s2 level, $P_T(t)$, and optimal secreted SUC2-s2 level, $P_M(t)$.
B. Optimal control function (feed rate) with q_{max} = 10.0 L/h and q_{min} = 0.0 L/h.
C. Optimal glucose concentration and cell density.
D. $H_q(t)$ vs. time.

Figure 5. Optimal trajectories with final time of 7.5 h.



A. Optimal total expressed SUC2-s2 level, $P_T(t)$, and optimal secreted SUC2-s2 level, $P_M(t)$.
B. Optimal control function (feed rate) with q_{max} = 10.0 L/h and q_{min} = 0.0 L/h.
C. Optimal glucose concentration and cell density.
D. $H_q(t)$ vs. time.

Figure 6. Optimal trajectories with final time of 15.0 h.

protein is expressed and the secretion processing occurs. The secreted protein production rate is proportional to the culture host cell mass. Since the promoter used in this system is under catabolic repression, no SUC2-s2 is expressed during this phase as seen in Figure 5A. The next phase is for the expression of SUC2-s2 gene, which requires relatively low glucose concentration. The transition between the two singular arc glucose levels by a bang-bang control of q_{min}, as seen in Figure 5B. This transition strictly observes the minimum principle as seen the positive

value of H_q during the transition in Figure 5D. During expression, secretion is also taking place, leading to the fast increase in the secreted protein level in the culture media as seen in Figure 5A. The quantity, $P_T - P_M$, indicates the accumulation of the expressed polypeptides in the secretory organelles. Close to the final time, the last singular arc phase, the optimal control accelerates the secretion of intracellularly accumulated polypeptides by increasing the glucose concentration back to a high level. The faster increase in the level of secreted product in culture, $P_M V$ of which maximization is the objective, close to t_f is observed. Again, during this upward transition, the control function is constrained at q_{max} and H_q has a negative value during the transition as seen Figure 5D. The last singular arc phase becomes more significant as the expressed polypeptides is less compatible with the host cell's secretion machinery, which is reflected by a smaller value of α and a larger value of β in the secretion dynamic model (Park and Ramirez, 1988). This results in delayed secretion dynamics and a higher level of accumulated polypeptides in the organelles, $P_T - P_M$, at the end of the second phase.

The optimal trajectories of the control function and state variables for $t_f = 15$ h have similar shapes to those for $t_f = 7.5$ h. In this case, more than half of the fermentation time is assigned for the harvesting of the host cell. Due to the higher cell mass accumulated during the longer cell growth phase, the downward transition is faster than that of $t_f = 7.5$ h. Unlike the case of $t_f = 7.5$, the last phase of maximum secretion rate has shrunk, and the glucose level does not reach to that of the singular arc as seen in Figure 6C. A complete upward transition between the two singular arcs would take a longer time compared to that of $t_f = 7.5$ due to the much higher culture cell mass and volume with the twice longer fermentation. This is interpreted as that, for the given operating time, the optimal control assigns more time to the second phase during which both the expression and the secretion occur simultaneously, instead of spending time on the transition which would take longer. The final glucose concentration is located in observance with the minimum principle by Eq. 52. As shown in Figures 6B and 6C, the control function profile is consistent with the behavior of H_q. Figures 5C and 6C give the optimal cell density profiles. Each specified final time results in a unique final culture volume. The final culture volume increases exponentially with the extended final time, such that the final culture volumes with $t_f = 7.5$ h and $t_f = 15$ h are 2.35 L and 14.35 L, respectively. For a given reactor volume, the final time, for which optimal operation can utilize the full capacity of the reactor, can be uniquely determined.

Conclusions

From this work, it was observed how secretion dynamics affects the control policy for protein production in a bioreactor when the protein requires post-translational modification. The objective function to be maximized was chosen as the total secreted protein in the culture at the final time. The general solution shows operation on three singular arcs connected by bang-bang operation. A pumping up of the fresh media at the maximum rate close to the final time significantly increases the total amount of secreted protein in the culture for a short time. Also, the importance of the high cell density in the protein-producing reactor is reflected by the initial singular arc operating condition. In this system where the favorable glucose concentrations for the protein expression, for the host cell growth, and the protein secretion are decoupled, the glucose concentration changes sequentially to optimize in the order of host cell growth, protein expression, and secretion of the expressed polypeptides. The system contains multiple singular arcs, and the optimal operating condition requires switchings between these distinct singular arcs. The best assignment of the whole fed-batch running time on each apart singular arcs are determined by optimal control theory. The optimal transitions between these singular arcs with a constrained control function are uniquely located by applying the minimum principle.

The characteristics of the cloned heterologous protein, the promoter used in the transcriptional control region, and the culture media are reflected in the formulation of the three dynamic rate functions, Φ, f_p, and μ_x. As shown in the singular arc equation of Eq. 28, the structures of these rate functions together with their parameters play an important role in determining optimal operating conditions. All factors that can affect the production rate and the achievable product concentration, such as host cell growth rate, dependency of the promoter activity on the culture condition, secretion rate in the host, and dilution by the fresh media, should be considered in determining optimal operating conditions.

Acknowledgment

This research was supported by the National Science Foundation under grant ECE-8611305.

Notation

$A(S)$ = function form of the scretion rate with respect to S
$B(S)$ = function form of the expression rate with respect to S
$C(S)$ = function form of the specific growth rate with respect to S
E = culture ethanol level, g/L
f_p = protein expression rate
H, H_q = Hamiltonian and its derivative with respect to q
J = objective function
m = glucose concentration of feed stream, g/L
P_I = level of intracellular SUC2-s2 in secretory organelles in culture, arbitrary units/L
P_M = level of secreted SUC2-s2 in culture, arbitrary units/L
P_T = level of total SUC2-s2 in culture, arbitrary units/L
q = feed flow rate, L/h
q_{max}, q_{min} = upper and lower bound of feed flow rate, L/h
S = culture glucose level, g/L
S_{high}, S_{low} = singular arcs of high and low glucose concentration
t_s, t_e = times at which bang-bang control trajectory starts and ends
t_f = final time, h
V = culture volume, L
X = culture cell density, g/L
x = system state variables
Y = yield of glucose(g)/cell mass(g), 7.3

Greek letters

α, β = parameters in secretion rate constant, L/h
λ = system costate variables
μ_x = specific growth rate of host cell, L/h
Φ = protein secretion rate, L/h
Ψ = part of Hamiltonian
ω = weighting factor

Literature Cited

Aiba, S., M. Shoda, and M. Nagatani, "Kinetics of Product Inhibition in Alcohol Fermentation," *Biotech. Bioeng.*, **10**, 845 (1968).

Carlson, M., and D. Botstein, "Two Differently Regulated mRNAs with Different 5' Ends Encode Secreted and Intracellular Forms of Yeast Invertase," *Cell*, **28**, 145 (1982).

D'Ans, G., D. Gottlieb, and P. Kokotovic, "Optimal Control of Bacterial Growth," *Automatica,* **8,** 729 (1972).

Das, R., and J. L. Shultz, "Secretion of Heterlogous Protein from *Saccharomyces Cerevisiae,*" *Biotech. Prog.,* 3(1), 43 (1987).

King, R. E., J. Aragona, and A. Constantinides, "Specific Optimal Control of a Batch Fermenter," *Int. J. Cont.,* **20**(5), 869 (1974).

Hasegawa, S., M. Matsubara, and K. Shimizu, "Noninferior Periodic Operation of the Metabolic Producing Biological Reactor," *Biotech. Bioeng.,* **30,** 703 (1987).

Kelly, H. J., "A Second Variation Test for Singular Extremals," *AIAA J.,* **2**(8), 1380 (1964).

Kopp, R. E., and H. G. Moyer, "Necessary Conditions for Singular Extremals," *AIAA J.,* 3(8), 1439 (1965).

Lehle, L., R. E. Cohen, and C. E. Ballou, "Carbohydrate Structure of Yeast Invertase," *J. Bio. Chem.,* **254**(23), 12209 (1979).

Menawat, A., R. Mutharasan, and D. R. Coughanowr, "Singular Optimal Control Strategy for a Fed-Batch Bioreactor: Numerical Approach," *AIChE J.,* **33**(5), 776 (1987).

Modak, J. M., H. C. Lim, and Y. J. Tayeb, "General Characteristics of Optimal Feed Rate Profiles for Various Fed-Batch Fermentation Processes," *Biotech. Bioeng.,* **28,** 1396 (1986).

Novick, P., S. Ferro, and R. Schekman, "Order of Events in the Yeast Secretory Pathway," *Cell,* **25,** 461 (1981).

Park, S., and W. F. Ramirez, "Dynamics of Heterlogous Protein Secretion from *Saccharomyces Cerevisiae,*" *Biotech. Bioeng.,* **33** (1988).

Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes,* Pergamon Press, New York (1964).

Schekman, R., P. Novick, S. Ferro-Novick, B. Esmond, W. Hansen, T. Etcheverry, and T. Stevens, "Protein Secretion and Organelle Assembly in Yeast," *em Rec. Adv. Yeast Mol. Biol.,* **1,** 143 (1982).

Takamatsu, T., I. Hashimoto, S. Shioya, K. Mizuhara, T. Koike, and H. Ohno, "Theory and Practice of Optimal Control in Continuous Fermentation Process," *Automat.,* **11,** 141 (1975).

Manuscript received Mar. 8, 1988, and revision received May 16, 1988.