2006

Optimization of Fed-Batch Saccharomyces cerevisiae Fermentation Using Dynamic Flux Balance Models

Jared L. Hjersted
Michael A Henson, *University of Massachusetts - Amherst*

ARTICLES

Optimization of Fed-Batch *Saccharomyces cerevisiae* Fermentation Using Dynamic Flux Balance Models

Jared L. Hjersted and Michael A. Henson*

Department of Chemical Engineering, University Massachusetts, Amherst, Massachusetts 01003-3110

We developed a dynamic flux balance model for fed-batch *Saccharomyces cerevisiae* fermentation that couples a detailed steady-state description of primary carbon metabolism with dynamic mass balances on key extracellular species. Model-based dynamic optimization is performed to determine fed-batch operating policies that maximize ethanol productivity and/or ethanol yield on glucose. The initial volume and glucose concentrations, the feed flow rate and dissolved oxygen concentration profiles, and the final batch time are treated as decision variables in the dynamic optimization problem. Optimal solutions are generated to analyze the tradeoff between maximal productivity and yield objectives. We find that for both cases the prediction of a microaerobic region is significant. The optimization results are sensitive to network model parameters for the growth associated maintenance and P/O ratio. The results of our computational study motivate continued development of dynamic flux balance models and further exploration of their application to productivity optimization in biochemical reactors.

Introduction

Large-scale production of biotechnological products is performed in biochemical reactors operated in batch, fed-batch, and continuous mode. An important advantage of fed-batch operation is that nutrient levels can be varied to achieve favorable growth conditions. Fed-batch yeast fermentation is a powerful technology for producing metabolic products such as ethanol (*1*−*3*) and therapeutic proteins such as human interferon, hepatitis B surface antigen, and insulin (*4*). The primary operational challenges associated with fed-batch operation are the determination of the initial nutrient concentrations and liquid volume, the feeding policies of the nutrients throughout the batch, and the final batch time. Fed-batch performance can be highly sensitive to these variables as a result of their complex effects on cellular metabolism. Therefore, model-based optimization is an essential tool for determining fed-batch operating strategies.

The transient nature of fed-batch fermentation requires that the optimal operating policy be determined by solving a dynamic optimization problem in which a final time objective (e.g., productivity and yield) is maximized, subject to constraints imposed by dynamic model equations (*5*, *6*). A number of computational algorithms have been proposed to solve such dynamic optimization problems (*7*). Sequential solution methods involve repeated iterations between a dynamic simulation code that integrates the model equations given a candidate feeding policy and a nonlinear programming code that determines an improved feeding policy given the dynamic simulation results. Simultaneous solution methods based on temporal discretization of the dynamic model equations have proven to be more effective as a result of their ability to handle state-dependent constraints and their applicability to large optimal control problems (*8*, *9*). The formulation and solution of dynamic optimization problems for maximizing ethanol productivity in fed-batch yeast fermenters have been extensively investigated (*10*−*15*). These studies were based on simple unstructured models with phenomenological descriptions of cell growth and constant yield coefficients. Unstructured models cannot be expected to provide accurate predictions over the wide range of transient conditions observed in fed-batch culture. Therefore, dynamic models based on more detailed representations of cellular metabolism are desirable. The challenge not only involves the development of more accurate models, but also on the application of more sophisticated dynamic optimization techniques demanded by such models.

Dynamic flux balance modeling provides a practical approach for the construction of detailed metabolic models in the absence of enzyme kinetic data and substantial information about cellular regulatory processes. The method is based on the reasonable assumption that metabolite concentrations rapidly equilibrate in response to extracellular perturbations (*16*). Dynamic mass balances on extracellular species are coupled to a stoichiometric model of intracellular metabolism through substrate uptake kinetics and product secretion rates. The stoichiometric model is comprised of a linear system of flux balance equations that relate metabolic species to their intracellular fluxes through a reaction network (*16*). Typically there are more fluxes than intracellular species, and the linear system is underdetermined. The fluxes can be resolved by specifying a cellular objective such as maximization of cell growth and solving the resulting linear programming problem (*17*, *18*). A flux balance description of intracellular metabolism combined with dynamic mass balances on extracellular substrates and products allows the prediction of cellular behavior as the extracellular environment

* To whom correspondence should be addressed. E-mail: henson@ecs.umass.edu.

10.1021/bp060059v CCC: $33.50
Published on Web 08/04/2006

changes with time. Dynamic flux balance models have been used to analyze the dynamic behavior of bacterial (*19*, *20*) and yeast (*21*) cultures operated in batch mode.

Conventional steady-state flux models predict secretion rates for all nonaccumulating metabolic products present in a network. A yeast metabolic network was shown to predict growth phenotype with an ≈70−80% success rate for various gene knockouts (*22*) by constraining fluxes associated with these genes. Therefore, dynamic flux balance models may be used to generate dynamic predictions of extracellular metabolite profiles and to predict the dynamic effects of gene knockouts on the productivity of fed-batch cultures. Unlike unstructured representations, a dynamic flux balance model efficiently captures carbon partitioning between all the represented intracellular pathways as the extracellular environment varies with time. The pathways are not modeled separately; they are instead captured by the distribution of fluxes within a single integrated metabolic network.

Recently, dynamic flux balance analysis has been extended to allow the determination of genetic alterations that optimize metabolite production in batch culture (*23*). The problem was posed as a bilevel optimization problem consisting of an outer nonlinear program in which metabolite production was maximized by dynamically manipulating predetermined fluxes and an inner linear program that maximized the growth rate by distributing fluxes among the unaltered reactions. The bilevel optimization problem was solved using a sequential strategy in which a nonlinear programming code repeatedly called a subroutine that evaluated the inner linear program, integrated the extracellular balance equations, and returned the value of the objective function. Although not investigated in (*23*), an alternative approach involves simultaneous solution by reformulation of the bilevel optimization problem into a single level nonlinear program through the replacement of the inner linear program with a set of equivalent algebraic constraints. Potential advantages compared to the sequential solution strategy include improved computational efficiency and superior robustness for large models. Two formulations have been proposed for handling the inner linear program. Duality theory has been used to infer cellular objective functions (*24*) and to predict optimal knockouts (*25*, *26*) with steady-state flux balance models. We believe that an alternative formulation based on the replacement of the inner linear program with its associated first-order optimality conditions to generate a mathematical program with equilibrium constraints (*27*) is more promising for large-scale problems resulting from dynamic optimization of fed-batch bioreactors. This method has been used previously to perform data reconciliation and parameter estimation with steady-state flux balance models (*28*).

The objective of this paper is to develop a general methodology for utilizing dynamic flux balance models to optimize metabolite production in fed-batch fermentation. We use ethanol production with the yeast *Saccharomyces cerevisiae* as a model system to illustrate our computational framework. Although other metabolites could be investigated, ethanol is a readily measurable product of central metabolism and serves as a convenient target metabolite for the evaluation of predicted trends and future experimental model verification. The dynamic flux balance model is constructed by coupling dynamic balances on extracellular glucose and ethanol, with a detailed stoichiometric model of primary carbon metabolism. Assuming direct manipulation of the dissolved oxygen concentration, oxygen and glucose uptake kinetics are modeled with standard empirical expressions to provide an integrated description of aerobic and anaerobic metabolism. We employ a simultaneous solution strategy to determine fed-batch operating policies that maximize an objective function that is a weighted sum of ethanol productivity and ethanol yield on glucose. Optimization tests are performed to investigate the impact of objective function weighting, model parameters, and modeling errors on achievable performance.

Dynamic Flux Balance Model. The modeling objective is to predict cellular growth and ethanol production for anaerobic and aerobic growth on glucose in fed-batch culture. The dynamic balances for the extracellular environment are described by the following equations:

$$\frac{dV}{dt} = F \tag{1}$$

$$\frac{d(VX)}{dt} = \mu VX \tag{2}$$

$$\frac{d(VG)}{dt} = FG_f - \upsilon_g VX \tag{3}$$

$$\frac{d(VE)}{dt} = \upsilon_e VX \tag{4}$$

where V is the liquid volume, G_f is the glucose feed concentration, F is the glucose feed flow rate, μ is the cellular growth rate, υ_g and υ_e are the membrane exchange fluxes for glucose and ethanol, respectively, and X, G, and E are the extracellular concentrations of biomass, glucose, and ethanol, respectively. The growth rate (μ) and the exchange flux of ethanol (υ_e) are resolved by solution of the inner flux balance model. The exchange fluxes for glucose (υ_g) and oxygen (υ_o) are modeled as:

$$\upsilon_g = \upsilon_{g,\,max}\frac{G}{K_g + G + (G^2/K_{ig})}\frac{1}{1 + (E/K_{ie})} \tag{5}$$

$$\upsilon_o = \upsilon_{o,\,max}\frac{O}{K_o + O} \tag{6}$$

where O is the dissolved oxygen concentration, K_g and K_o are saturation constants, K_{ig} and K_{ie} are inhibition constants that reduce glucose uptake in the presence of high glucose or ethanol concentrations, respectively, and $\upsilon_{g,\,max}$ and $\upsilon_{o,\,max}$ are the maximum uptake rates. The glucose uptake follows Michaelis−Menten kinetics with additional inhibitory terms to capture regulatory effects due to high glucose and ethanol concentrations.

Direct dissolved oxygen manipulation was assumed for all simulations, and, therefore, oxygen balances were not included in the dynamic flux balance model (eqs 1−4). Instead, fed-batch optimization was performed with a maximum value of the dissolved oxygen concentration, DO_{max} ($DO = O/O^*$, where O^* is the saturation concentration). This simplification was deemed reasonable because optimal ethanol production leads to near anaerobic growth conditions at higher cell densities. Therefore, the dissolved oxygen concentration was treated directly as a decision variable under the assumption that the computed trajectory could be rapidly tracked with a suitably designed feedback controller. The uptake of ethanol as a carbon source was excluded from the model because diauxic growth is not observed until glucose is almost completely exhausted, which occurs at the final batch time. A gas-phase ethanol balance was not included under the assumption of negligible ethanol holdup in the vapor space of the bioreactor.

A metabolic network that describes primary carbon metabolism and the formation of cellular biomass in the yeast *Saccharomyces cerevisiae* was used for this study (*29*, *30*). Construction of the flux model required exclusion of some fluxes and determination of species that should be allowed to accumulate. Aerobic and anaerobic growth on glucose was captured with 82 fluxes and 98 biochemical species, whereas the original flux model referenced contained 99 fluxes. Sixteen fluxes pertaining to alternate carbon substrates or organisms other than *S. cerevisiae* (fluxes 13−14, 25, 37−42, 52−57, and 65 in the original model) were removed. An additional anaerobic-only flux (flux 24) was removed because it produced false predictions for aerobic conditions and had a minimal effect under anaerobic conditions. The numbering of the fluxes in the original model was preserved for the sake of clarity.

A series of steps were followed to determine the intracellular species that accumulate. We initially allowed all of the species to accumulate to account for dead-end pathways and the incomplete nature of the metabolic model. Solution of the linear program for this case yielded an optimal distribution of fluxes with very few accumulating species. The flux model was then reformulated to only allow the accumulation of species located at the end of a pathway or metabolites for which accumulation has been experimentally demonstrated (*31*). The latter group consisted of acetate, fumarate, malate, pyruvate, and succinate. The remaining 68 species were constrained to be nonaccumulating. Complete specification of the metabolic network required values for the growth-associated maintenance coefficient (k) and the P/O ratio (PO). The flux model constructed allowed predictions for fully aerobic, mixed aerobic (constrained oxygen uptake), microaerobic, and anaerobic conditions. Cell densities encountered in batch and fed-batch cultivations are such that fully aerobic conditions are rarely achieved as a result of transport limitation. Consequently, we liberally use the term *aerobic* to describe *mixed aerobic* conditions in this study.

The inner linear program (LP) for maximization of the cellular growth rate is posed as:

$$\max_{v,b} \mu = w^{\mathrm{T}} v \tag{7}$$

subject to:

$$Av = b$$
$$v_{\min} \le v \le v_{\max}$$
$$b_i = b_i,\ i \in E$$
$$b_i = 0,\ i \in N$$
$$b_{i,min} \le b_{\mathrm{i}} \le b_{i,max} \quad i \in M$$

where E is the set of intracellular species with externally determined exchange fluxes, N is the set of species that do not accumulate, M is the set of species that are allowed to accumulate, A is the stoichiometric matrix for the metabolic network, v is a vector of reaction fluxes, and b is a vector of accumulation and exchange rates. The cellular growth rate (μ) is determined from the fluxes producing biomass precursors, where the weights (w) are determined from the amount of each precursor necessary for biomass formation (*29*). The metabolic network (*29*) used in this study includes a single flux to represent biomass formation (flux 99) that simplifies the expression for the growth rate.

Nominal model parameter values are presented in Table 1. A significant advantage of the proposed methodology is that only a small number of parameters must be specified to construct a detailed dynamic model of the fermentation process. The maximum uptake rates ($v_{g,\,max}$ and $v_{o,\,max}$) and the saturation constants (K_g and K_o) were obtained from literature values (*32*) after basic unit conversions. Values for the inhibition constants (K_{ig} and K_{ie}) were chosen to yield reasonable model predictions. The saturation concentration of liquid-phase oxygen (O^*) was determined from Henry's law using a peak oxygen concentration of 0.21 atm in the reactor headspace and a liquid temperature of 30 °C. A constant glucose feed concentration was used throughout this study to avoid the situation where the maximum concentration is always selected by the optimizer. The metabolic network parameters k and PO were obtained from the original reference where the steady-state flux balance model was formulated (*29*). A conversion of 25.593 31 g of biomass per C-mol of biomass was used (*30*). The model parameter t_{ss} represents the time required for equipment maintenance between fed-batch runs.

Table 1. Nominal Model Parameter Values

variable	value	variable	value
$v_{g,max}$	0.02 mol g^{-1} h^{-1}	K_g	0.5 g L^{-1}
$v_{o,max}$	0.008 mol g^{-1} h^{-1}	K_o	3×10^{-6} mol L^{-1}
K_{ig}	10 g L^{-1}	K_{ie}	10 g L^{-1}
O^*	2.53×10^{-4} mol L^{-1}	G_f	50 g L^{-1}
DO_{max}	50%	t_{ss}	6 h
k	1.37 mol $Cmol^{-1}$	PO	1.20 mol $atom^{-1}$

We used the MATLAB interface to the LP code MOSEK to solve eq 7 given a set of glucose and oxygen uptake rates. Figure 1 shows the growth rate and ethanol production rate surfaces obtained when the LP was repeatedly solved over a representative grid of glucose and oxygen uptake rates. Although alternative optimal solutions are a well-known problem with classical flux balance analysis (*33*), we did not encounter this issue in the computational studies shown in this paper. The surfaces are nontrivial functions of the substrate uptake rates, as demonstrated by the maximum in the ethanol production rate for microaerobic growth conditions. These surfaces suggest that the development of a comparable unstructured model would minimally require the specification of different metabolic

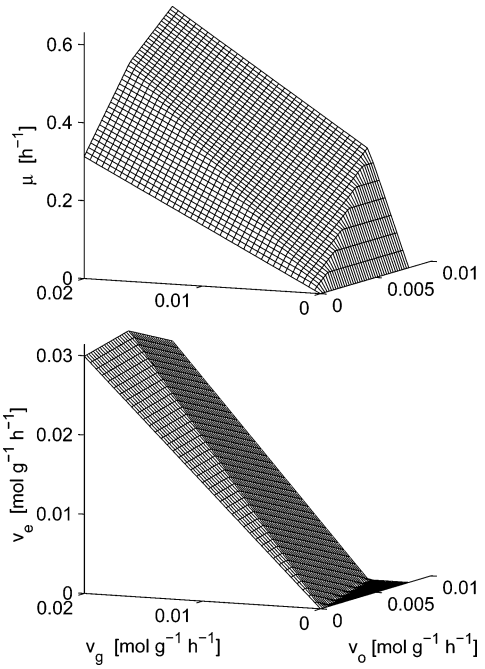

Figure 1. Surfaces for the optimal growth rate (μ) and corresponding ethanol production rate (v_e) obtained by repeatedly solving the linear program (eq 7) for a range of glucose (v_g) and oxygen (v_o) uptake rates.

parameters for anaerobic, microaerobic, and aerobic growth. Considerably more complicated behavior can be observed, as recently demonstrated by the existence of seven metabolic phenotypes in a genome-scale flux model of *S. cerevisiae* metabolism (*34*). While conceptually possible, the development of unstructured models that attempt to reproduce multiple metabolic phenotypes quickly becomes unwieldy. Given the development of highly efficient dynamic simulation and optimization schemes, as reported in this paper, there is little motivation to develop unstructured models when a dynamic flux balance model is available.

Dynamic simulations were performed in MATLAB using the routine ODE15s to integrate the extracellular mass balance equations. A fed-batch simulation for a constant glucose feed flow rate and an aerobic (50% *DO*) to anaerobic (0% *DO*) switch is shown in Figure 2. The biomass concentration increased rapidly under aerobic growth conditions. The switch to anaerobic growth at 10 h resulted in a substantially increased ethanol production rate at the expense of the biomass growth rate. These types of dynamic simulations have been performed with previously developed dynamic flux balance models and have shown good agreement with batch experimental data (*20*, *21*). The primary contribution of this computational study is the use of a dynamic flux balance model for fed-batch optimization.

Fed-Batch Optimization Strategy. The objective function maximized was the weighted sum of ethanol productivity and ethanol yield on glucose. This dual objective allowed the tradeoff between high production rates and efficient substrate usage to be examined. The initial volume $V(0)$ and glucose concentration $G(0)$, the feed flow rate $F(t)$ and dissolved oxygen concentration $DO(t)$ profiles, and the final batch time t_f were treated as decision variables. Therefore, the dynamic optimization problem had the form

$$\max_{V(0),G(0),F(t)DO(t)t_f} \quad c_p P(t_f) + c_y Y(t_f) \tag{8}$$

subject to: extracellular balances (eqs 1−4)

uptake kinetics (eqs 5 and 6)

flux balance LP (eq 7)

$V(0) \geq 0.5$ L, $V(t_f) \leq 1.2$ L

$0 \leq G(0) \leq 50$ g/L

$0 \leq DO(t) \leq 50\%$

$F(t) \geq 0$ L/h

1 h $\leq t_f \leq$ 36 h

$X(t), G(t), E(t) \geq 0$ g/L

The ethanol productivity P and the ethanol yield on glucose Y at the final batch time were defined as:

$$P(t_f) = \frac{V(t_f)E(t_f)}{t_f + t_{ss}} \tag{9}$$

$$Y(t_f) = \frac{V(t_f)E(t_f)}{V(0)G(0) + \int_0^{t_f} G_f F(t)\mathrm{d}t} \tag{10}$$

where t_{ss} is the turnover time for startup and shutdown of a fed-batch run. The parameters c_p and c_y are weights for the productivity and yield objectives, respectively.

The variable bounds were primarily specified to ensure a physically realistic solution. The bounds on the initial and final volumes were chosen for consistency with our experimental system. Lower and upper bounds on the final batch time were included to confine the solution space, but they had no effect on the optimal solutions generated. The nominal values for the glucose feed concentration G_f and fed-batch turnover time t_{ss} listed in Table 1 were used in the optimization studies. Dissolved oxygen *DO* appeared directly as a decision variable in the dynamic optimization problem. As discussed earlier, the upper bound DO_{max} was chosen to avoid the calculation of high dissolved oxygen concentrations that might be unachievable in practice due to oxygen mass transfer limitations. This formulation also avoided problems caused by the insensitivity of the objective function to oxygen transport variable, such as the air flow rate and the mass transfer coefficient.

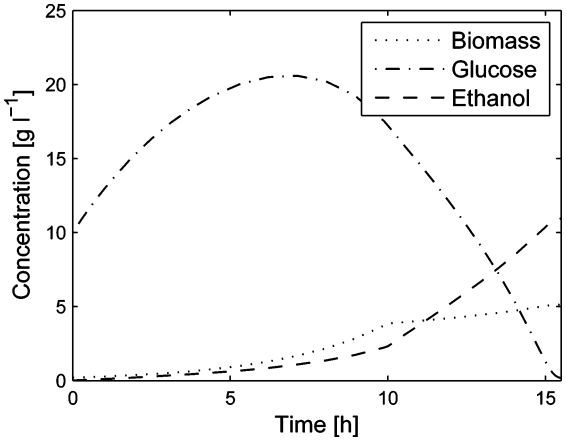

Figure 2. Fed-batch simulation with a constant glucose feed flow rate and an aerobic to anaerobic switch at 10 h.

Our initial attempts to solve the dynamic optimization problem using a sequential method (*23*) failed. The method proved to be intractable due to the size of the problem and suffered from the absence of analytic Jacobian or Hessian information. Therefore, we employed a simultaneous solution method in which the bilevel dynamic optimization problem (eq 8) was reformulated as a single level nonlinear program. The procedure required temporal discretization of the extracellular balances (eqs 1−4) and replacement of the inner LP (eq 7) with its associated first-order optimality conditions (*28*). Discretization was performed with Radau collocation on finite elements using a monomial basis representation (*35*) with 61 finite elements and 2 internal collocation points per element for a total of 184 nodes. The linear program (eq 7) was only enforced at the beginning of each finite element to reduce the overall problem size. As shown below, comparisons with direct simulation results showed that this approximation was adequate provided that a sufficient number of finite elements was employed. The decision variables $DO(t)$ and $F(t)$ were restricted to change only at the element boundaries.

The single level nonlinear program consisted of only algebraic constraints. The discretized dynamic balances for biomass (*VX*), glucose (*VG*), ethanol (*VE*), and volume (*V*) had the form

$$Z_{i,j} = Z_i^0 + h_i \sum_{k=1}^{n_{cp}+1} D_{k,j} \frac{\mathrm{d}(Z_{i,k})}{\mathrm{d}t} \tag{11}$$

where Z is the appropriate variable, the superscript 0 indicates a value at the beginning of an element, i is the finite element index, j is the collocation point index, h is the width of an element, n_{cp} is the number of internal collocation points, and D is the Radau collocation matrix. More details on this collocation strategy and monomial basis representations can be found in

Biegler et al. (*8*). Karush−Kuhn−Tucker (KKT) optimality conditions and complementarity constraints resulted from replacement of the linear program with equivalent optimality conditions. The KKT conditions enforced were

$$w + A^{\mathrm{T}}\lambda + \alpha_i^L - \alpha_i^U = 0 \tag{12}$$

$$\lambda_{i,m} - \eta_{i,m}^L + \eta_{i,m}^U = 0, \qquad m \in N \tag{13}$$

where λ, α, and η are multipliers, the subscripts L and U correspond to lower and upper bounds, respectively, m represents the species index, and the index i indicates that the constraints are imposed at the beginning of each element. The complimentarity constraints had the following form

$$v_{i,n} - v_n^L \geq 0 \tag{14}$$

$$(v_{i,n} - v_n^L)\alpha_{i,n}^L = 0 \tag{15}$$

where n indexes the fluxes. These relations represent one of four sets of complementarity constraints, which were applied to the variables v and b with respect to the lower and upper bounds. More details on the representation of the optimality conditions are available in Raghunathan et al. (*28*).

The dynamic optimization problem was solved through the AMPL interface to the nonlinear program solver CONOPT. AMPL is a mathematical programming language that provides analytic Jacobian and Hessian information to the solver through integrated automatic differentiation. CONOPT is a feasible path, multi-method nonlinear program solver based on the generalized reduced gradient method. The optimization model consisted of 36 049 decision variables and 30 496 constraints. The computation time varied from 126 to 221 s, depending on the initialization and objective function used. Subsequent solutions for small changes in the objective function or constraints required only a small fraction of the initial computation time. All computations were performed on a 3.0 GHz Pentium 4 CPU.

Results and Discussion

The first set of simulation results were generated from solution of the dynamic optimization problem for maximization of ethanol productivity: $c_y = 0$ in eq 8. Figure 3 shows the optimal control profiles for the feed flowrate and the dissolved oxygen. The calculated optimal state profiles and simulated profiles obtained from direct simulation of the optimal control profiles are also displayed in Figure 3. Slight differences between the optimal and simulated profiles originated from the approximation of constant fluxes across finite elements used in the optimization problem. The optimal control policy produced an initial glucose concentration (14.6 g/L) well below its upper bound and no initial glucose feed. The glucose concentration declined until feeding began at $t = 7.0$ h. Then the glucose feed flowrate increased over time such that the glucose concentration remained approximately constant until the final volume constraint was encountered at $t = 13.4$ h. Analysis of eq 5 revealed that this constant glucose concentration resided very close to the relatively flat maximum in the glucose uptake rate. Despite the approximately constant glucose concentration observed in the final half of the batch, the high initial glucose concentration followed by the transient decrease during the first half of the batch suggests that direct substrate concentration control (*36*) may not be suitable for ethanol productivity maximization. A sudden switch in the dissolved oxygen from the initial maximum to a final value near zero was observed at

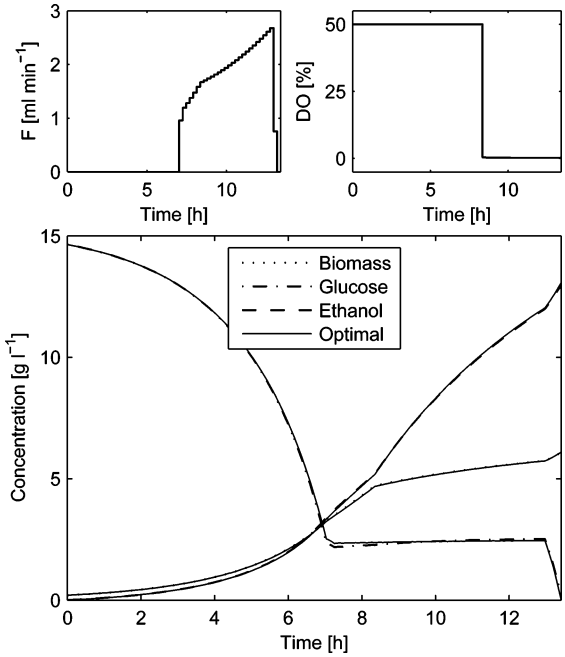

Figure 3. Optimal glucose feed (top left) and dissolved oxygen (top right) profiles for optimization of ethanol productivity and the corresponding simulated and optimal profiles of biomass, glucose, and ethanol (bottom).

$t = 8.4$ h. This switch divided an initial aerobic phase of high cell growth followed by a microaerobic phase of high ethanol production.

Snapshots of the metabolic network fluxes at two time points in the optimal productivity profiles are shown in Figure 4. The time points were chosen to demonstrate the difference in the flux distribution for the aerobic ($t = 5$ h) and microaerobic ($t = 10$ h) operating regions. A higher growth rate and more activity in oxidative phosphorylation were observed under aerobic conditions. Increased ethanol production, lower growth, and less TCA cycle utilization were observed under microaerobic conditions. Complete dynamic profiles of selected exchange and intracellular fluxes are shown in Figure 5. The onset of glucose feeding at 7.0 h and the switch in dissolved oxygen at 8.4 h were distinctly visible in the flux profiles. The influence of inhibition was also observed in the increasing rates of growth and glucose uptake as extracellular glucose concentrations decreased in the preliminary batch phase and in the decreasing rates of growth and glucose uptake as extracellular ethanol concentrations increased in the final batch stage. These results demonstrate the detailed dynamic information that can be extracted from dynamic flux balance models that is unavailable in unstructured models.

Sensitivity analysis of the optimal solution to various bounds and model parameters is shown in Figure 6. Repeated solution of the dynamic optimization problem for different parameter values and bounds was performed to generate the results. The square indicates the nominal value and corresponds to the optimal solution shown in Figure 3. The productivity showed a sharp decrease when *DO* was below 10% of the saturation value. Because there were only slight performance improvements above this point, an upper bound of 50% dissolved oxygen was used for all other optimizations. The productivity increased linearly with the glucose feed concentration, except at low concentrations where a sharp decline was observed. The effect of the maximum oxygen uptake parameter on the productivity was linear, except for large parameter values where oxygen was in stoichiometric excess compared to glucose and very small

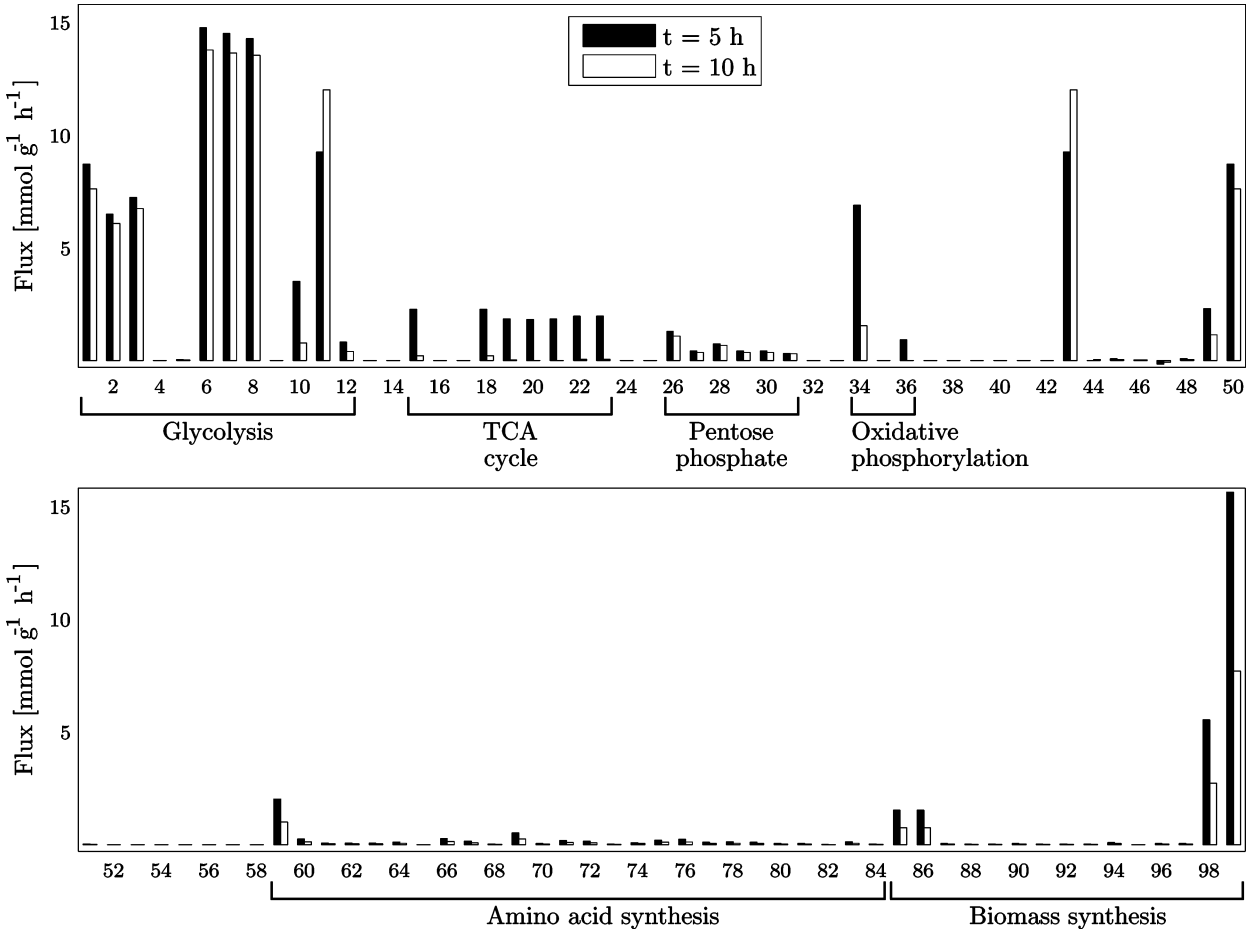

Figure 4. Flux distributions at two time points from the optimal productivity profiles in Figure 3, with key pathways indicated. The first time point ($t = 5$ h) is during the aerobic phase characterized by higher growth (flux 99), whereas the second time point ($t = 10$ h) represents the microaerobic phase where higher ethanol production (flux 43) is observed. The primary fluxes from pyruvate (9−12) are included in glycolysis. Biomass synthesis includes formation of precursors (i.e., protein, RNA, nucleotides, polysaccharides, and so on) and the synthetic flux representing cellular growth rate (flux 99) in m Cmol g^{-1} h^{-1}. The unlabeled fluxes are: 13 and 14 (growth on ethanol), 24 (anaerobic only), 25, 37 and 38 (unused for *S. cerevisiae*), 32 and 33 (glyoxylate shunt), 39−44 (alternate carbon substrates), 45−47 (1-carbon compound transfer), 48−57 (transport), and 58 (H^+ATPase). The flux numbers and directions are consistent with the network published in van Gulik et al. (*29*) and Vanrolleghem et al. (*30*), with the exception of flux 43 (negative as written in the references), which has been reversed for visual purposes.

values where diminished oxygen transport produced a sharp decrease. The ethanol inhibition constant displayed the expected trend where inhibition had a large impact on productivity only for small parameter values.

The dynamic optimization problem formulated in eq 8 contains a dual objective for ethanol productivity and ethanol yield. The optimization results presented in Figures 3−6 were generated from a productivity only objective ($c_y = 0$). A parametric sensitivity analysis was performed to examine the tradeoff between productivity and yield. The analysis involved repeated solution of the optimization problem (eq 8) with a constant value for the productivity weight ($c_p = 0.81^{-1}$) and a wide range of values for the yield weight ($0 \le c_y \le 60$). A nonzero value of the productivity weight was used to avoid solutions lying at the maximum bound of the final time constraint and exhibiting a dramatic decline in the productivity for an insignificant increase in the yield. Therefore, this strategy produced optimal policies for maximization of ethanol yield where the overall productivity loss was minimized. Figure 7 shows that increasing yields were achieved at the expense of decreasing productivities and longer batch times. The productivity versus yield curve represents the locus of achievable optimums for the dual objective where the entire area above the curve is unachievable.

During calculation of the yield−productivity tradeoff curve, the ethanol yield eventually saturated with respect to increasing values of the yield weight (c_y). This indicated that the yield was at its overall maximum and the productivity was at its maximum with respect to this yield. Figure 8 shows the optimal feeding policy that generated this point (circle in Figure 7). The calculated optimal state profiles and the simulated profiles obtained from direct simulation of the optimal control profiles are also shown in Figure 8. The results obtained were markedly different from the maximum productivity results (Figure 3). While the glucose concentration decreased until feeding began such that a relatively constant glucose concentration was maintained, the combined objective produced a lower initial glucose concentration to increase yield and earlier glucose feeding to achieve the glucose concentration that maximized uptake. Consequently, direct control of the extracellular glucose concentration would be expected to provide satisfactory performance for this case. The dissolved oxygen concentration in Figure 8 showed that microaerobic growth conditions were utilized throughout the batch. This effect was evident in the relatively constant rates of biomass and ethanol increase, which suggested that a high ethanol yield was maintained throughout the batch.

Both the productivity and combined yield−productivity objectives generated optimal solutions that suggest microaerobic growth is beneficial compared to purely anaerobic growth. These results are consistent with continuous culture experiments (*37*,

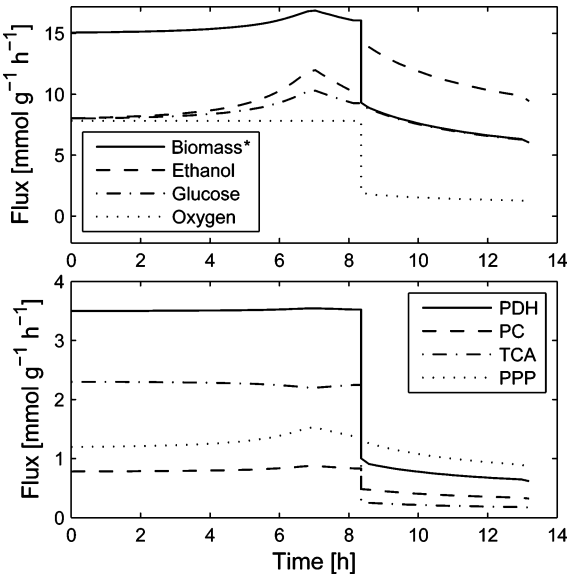

Figure 5. Dynamic profiles of exchange (top) and intracellular (bottom) fluxes corresponding to the optimal productivity case in Figure 3. The intracellular fluxes correspond to the reactions catalyzed by the enzymes pyruvate dehydrogenase (PDH, flux 10), pyruvate carboxylase (PC, flux 12), citrate synthase (CS, flux 15), and glucose-6-phosphate dehydrogenase (G6PDH, flux 26). The initiation of glucose feeding at 7.0 h and the switch in dissolved oxygen at 8.4 h were observed in the flux profiles. *Biomass formation (flux 99) has units on a carbon mole basis (m Cmol g^{-1} h^{-1}).

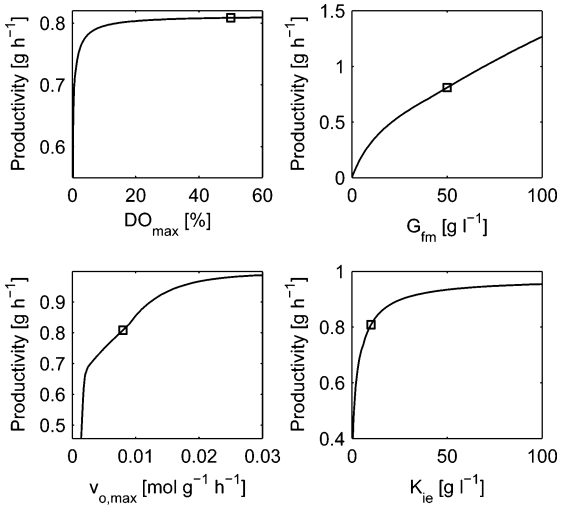

Figure 6. Sensitivity of the ethanol productivity obtained with the optimal profiles in Figure 3 to the upper bound on dissolved oxygen (top left), upper bound on glucose feed concentration (top right), maximum oxygen uptake rate (bottom left), and inhibition constant for glucose uptake due to ethanol (bottom right). The square indicates the productivity obtained with the nominal parameter values.

38) and conventional flux balance analysis (*34*). We performed additional simulations to quantify the significance of the microaerobic region on optimal fed-batch performance. The optimal control profiles previously obtained for the two objectives were modified by constraining dissolved oxygen concentrations less than 1.0% *DO* to be identically zero. Dynamic simulations with these modified control profiles showed a 19% decrease in productivity for the optimal productivity case and a 76% decrease in yield for the combined yield−productivity case. The large yield decrease was mainly due to the presence of substantial residual glucose at the final batch time, although higher glucose concentrations throughout the batch also led to glucose uptake inhibition, as per eq 5. The simulation tests were repeated, with flux 24 included in the network to ensure that

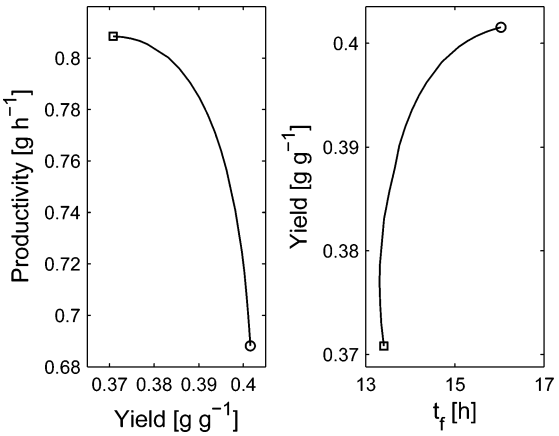

Figure 7. Tradeoff between ethanol productivity and ethanol yield on glucose (left), and the relationship between ethanol yield and the batch time (right). The square and the circle correspond to the optimization results shown in Figures 3 and 8, respectively.

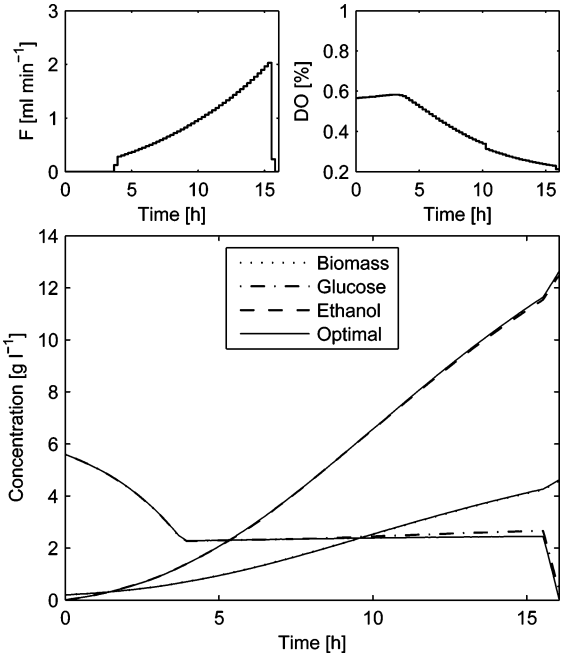

Figure 8. Optimal glucose feed (top left) and dissolved oxygen (top right) profiles for a combined yield−productivity objective, where yield was most heavily weighted, and the corresponding simulated and optimal profiles of biomass, glucose, and ethanol (bottom).

the microaerobic result had not been partially achieved by elimination of this anaerobic-only flux, which allows conversion between FADH2 and NADH in the absence of oxidative phosphorylation. While this flux was active when included, the effect was apparent only in oxidative phosphorylation, and no impact on the overall predictions was observed. The significance of the microaerobic region motivates reassessment of the perfect control assumption for dissolved oxygen. In practice, tracking of very low dissolved oxygen concentrations that characterize the microaerobic growth region might be problematic due to *DO* measurement limitations. Possible solutions include reformulation of the dynamic optimization problem such that low concentrations are avoided, development of a system-dependent oxygen transport model, or implementation of closed-loop dynamic optimization with measurement feedback. The latter option might be preferred due to the added advantage of model correction and disturbance rejection.

The previous optimization results were generated with the values of the maintenance coefficient *k* and the P/O ratio *PO*

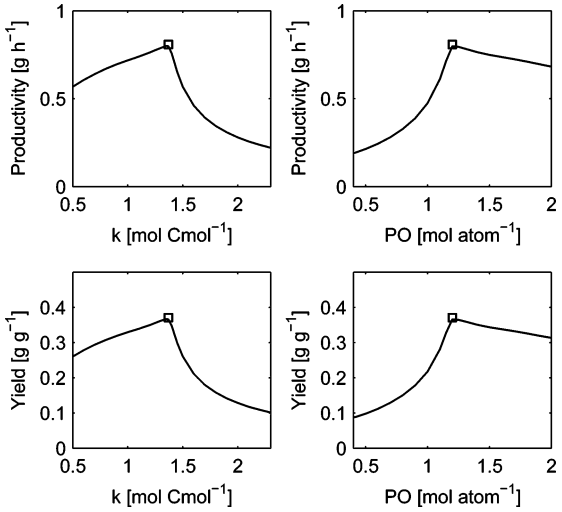

Figure 9. Sensitivity of the ethanol productivity and yield on glucose to the maintenance coefficient and P/O ratio. The square indicates the predicted optimal for the nominal parameter values and corresponds to Figure 3.

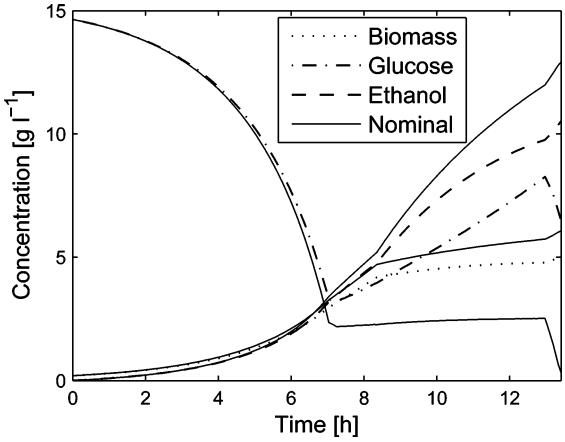

Figure 10. Comparison of simulated profiles obtained with the optimal inputs in Figure 3 for nominal (solid lines) and altered (dashed lines) values of two metabolic network parameters. The growth associated maintenance (k) was increased from 1.37 to 1.4 mol Cmol^{-1}, and the oxidative phosphorylation ratio (PO) was decreased from 1.2 to 1.15 mol atom^{-1}.

listed in Table 1. Because there is considerable uncertainty associated with these nominal values (*30*), we examined the sensitivity of the optimal solution obtained for the productivity objective to these two metabolic network parameters. The optimal control profiles obtained with the nominal parameter values (Figure 3) were applied to dynamic flux balance models with different k and PO values. The results are shown in Figure 9. As expected, the largest productivity and yield were obtained for the nominal parameter values (indicated by the squares). Fed-batch performance was most sensitive to underestimation of the maintenance coefficient (large process k values) and overestimation of the P/O ratio (small process P/O values). Larger k values indicate increased ATP consumption for growth associated maintenance, while smaller PO values represent decreased ATP production via oxidative phosphorylation.

Figure 10 shows simulation profiles obtained with the optimal input profiles in Figure 3 for the nominal values of k and PO (Table 1) and altered values chosen to evaluate the effects of modeling error. The second set of results were generated with the growth associated maintenance coefficient increased by 2.1% (1.37 to 1.4 mol Cmol^{-1}) and the oxidative phosphorylation ratio decreased by 4.35% (1.2 to 1.15 mol atom^{-1}). This comparison mimics a scenario where the optimal profiles were generated offline with an available model and then implemented online to an actual fermenter. A 23% decrease in the overall productivity was observed despite the relatively small changes in the two network parameters. The profiles for the altered case closely tracked the nominal profiles during the initial phase of the batch, but the profiles quickly diverged once glucose feeding began at 7.0 h. Both changes to the network parameters had a negative effect on cellular energy generation and, therefore, led to the phenotype of a slower growing cell. While the biomass profiles appeared to show the closest agreement, the difference resulted in an insufficient cell density to consume the supplied glucose and accumulation of glucose during feeding. The need for feedback in online implementation was evident from the sensitivity of the system to the metabolic network parameters. Feedback tracking of the optimal profiles could provide improved performance, but optimal performance would require a closed-loop formulation of the full optimization problem.

Conclusions and Future Work

A dynamic flux balance model for batch and fed-batch yeast fermentation was developed by embedding a steady-state flux balance description of intracellular metabolism within a system of dynamic mass balances on the extracellular environment. This framework is based on the assumption that intracellular metabolite concentrations rapidly equilibrate to extracellular perturbations and takes full advantage of flux balance models available for various microorganisms. The dynamic flux balance model was incorporated within a dynamic optimization framework to compute fed-batch operating policies that maximized ethanol productivity and/or ethanol yield on glucose. Decision variables were the initial volume and glucose concentration, the feed flow rate and dissolved oxygen concentration profiles, and the final batch time. The resulting bilevel optimization problem was formulated as a single level nonlinear program by temporal discretization of the dynamic balance equations and replacement of the inner linear program with its associated first-order optimality conditions.

Optimal solutions for the productivity objective were characterized by an initial phase of aerobic growth for maximal biomass production, followed by a switch to microaerobic growth for increased ethanol production. Glucose feeding was initiated near the end of the aerobic growth phase, such that the glucose uptake rate was maximized until the final volume constraint was encountered. By contrast, the yield objective produced optimal solutions characterized by microaerobic growth throughout the batch and much earlier initiation of glucose feeding. Prediction of the microaerobic growth region was found to be particularly important for the yield objective, where a 76% decrease in ethanol yield was observed when the optimal feeding profile was modified, by constraining dissolved oxygen concentrations less than 1.0% to be identically zero.

We believe dynamic flux balance modeling offers numerous advantages over typical unstructured models for analysis and optimization of batch and fed-batch cultivations. The incorporation of a detailed metabolic model alleviates the need for constant yield coefficients and parametrization of cellular growth into distinct pathways. The dynamic flux balance model formulation requires uptake kinetics for growth limiting substrates, which we believe is preferable to empirical modeling of the entire cellular growth process. While dynamic simulation of a dynamic flux balance model is relatively straightforward, an inner linear program for determination of intracellular fluxes causes the dynamic optimization problem to become a bilevel nonlinear programming problem. We have demonstrated that the dynamic optimization problem can be successfully refor-

mulated into a single level nonlinear program that is efficiently solved with existing software.

This initial study demonstrates that dynamic flux balance models hold considerable promise for the simulation and optimization of yeast fed-batch cultures. While our study focused on ethanol optimization with a wild-type strain of *Saccharomyces cerevisiae*, the proposed methodology is equally applicable to other microorganisms and metabolic products for which steady-state flux balance models have been developed. The increasing availability of genome-scale flux balance models for bacteria (*39*, *40*) and yeast (*41*) opens the possibility of utilizing molecular level detail for dynamic simulation and optimization of microbial cell culture systems. A limitation of the current dynamic flux balance methodology is the lack of intracellular regulation. Our future work will focus on the inclusion of regulatory mechanisms, incorporation of genome-scale metabolic networks, optimization of other high value metabolites, analysis of gene knockouts and overexpression, and experimental evaluation of the dynamic flux balance modeling and optimization framework. For example, we have been exploring the incorporation of yeast genome-scale metabolic networks within a dynamic flux balance model to investigate the effects of candidate gene knockouts on metabolite production in batch and fed-batch cultures. This recent work will be presented in a future publication.

Notation

DO	dissolved oxygen concentration [%]
E	ethanol concentration [g/L]
F	feed flowrate [L/h]
G	glucose concentration [g/L]
G_f	feed glucose concentration [g/L]
K_{ie}	glucose uptake inhibition constant with respect to ethanol [g/L]
K_{ig}	glucose uptake inhibition constant with respect to glucose [g/L]
K_g	glucose uptake saturation constant [g/L]
K_o	oxygen uptake saturation constant [mol/L]
O	liquid oxygen concentration [mol/L]
O^*	saturation liquid oxygen concentration [mol/L]
PO	P/O ratio [mol-ATP/atom-O]
V	liquid volume [L]
X	biomass concentration [g/L]
k	growth associated maintenance coefficient [mol-ATP/Cmol-biomass]
t_{ss}	batch startup and shutdown time [h]
υ_e	ethanol exchange rate [mol/g/h]
υ_g	glucose uptake rate [mol/g/h]
υ_o	oxygen uptake rate [mol/g/h]
μ	specific growth rate [h^{-1}]

References and Notes

(1) Alfenore, S.; Cameleyre, X.; Benbadis, L.; Bideaux, C.; Uribelarra, J.-L.; Goma, G.; Molina-Jouve, C.; Guillouet, S. E. Aeration Strategy: A Need for Very High Ethanol Performance in *Saccharomyces cerevisiae* Fed-Batch Process. *Appl. Microbiol. Biotechnol.* **2004**, *63*, 537−542.

(2) Converti, A.; Arni, S.; Sato, S.; de Carvalho, J. C.; Aquarone, E. Simplified Modeling of Fed-Batch Alcoholic Fermentation of Sugarcane Blackstrap Molasses. *Biotechnol. Bioeng.* **2003**, *84*, 88−95.

(3) Nilssen, A.; Taherzadeh, M. J.; Linden, G. Use of Dynamic Step Response for Control of Fed-Batch Conversion of Lignocellulosic Hydrolyzates to Ethanol. *J. Biotechnol.* **2001**, *89*, 41−53.

(4) Glick, B. R.; Pasternak, J. J. *Molecular Biotechnology: Principles and Applications of Recombinant DNA*; American Society for Microbiology: Washington, DC, 1998.

(5) Johnson, A. The Control of Fed-batch Fermentation − A Survey. *Automatica* **1987**, *23*, 691−705.

(6) Lubbert, A.; Jorgensen, S. B. Bioreactor Performance: A More Scientific Approach for Practice. *J. Biotechnol.* **2001**, *85*, 187−212.

(7) Biegler, L. T.; Grossmann, I. E. Retrospective on Optimization. *Comput. Chem. Eng.* **2004**, *28*, 1169−1192.

(8) Biegler, L. T.; Cervantes, A. M.; Wachter, A. Advances in Simultaneous Strategies for Dynamic Process Optimization. *Chem. Eng. Sci.* **2002**, *57*, 575−593.

(9) Cuthrell, J. E.; Biegler, L. T. Simultaneous Optimization and Solution Methods for Batch Reactor Control Problems. *Comput. Chem. Eng.* **1987**, *13*, 49−62.

(10) Banga, J. R.; Alonso, A. A.; Singh, R. P. Stochastic Dynamic Optimization of Batch and Semicontinuous Bioprocesses. *Biotechnol. Prog.* **1997**, *13*, 326−335.

(11) Kookos, I. K. Optimization of Batch and Fed-Batch Bioreactors using Simulated Annealing. *Biotechnol. Prog.* **2004**, *20*, 1285−1288.

(12) Lee, J.-H. Comparison of Various Optimization Approaches for Fed-Batch Ethanol Production. *Appl. Biochem. Biotechnol.* **1999**, *81*, 91−106.

(13) Luus, R. Application of Dynamic Programming to Differential Algebraic Process Systems. *Comput. Chem. Eng.* **1993**, *17*, 373−377.

(14) Vera, J.; de Atauri, P.; Cascante, M.; Torres, N. V. Multicriteria Optimization of Biochemical Systems by Linear Programming: Application to Production of Ethanol by *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **2003**, *83*, 335−343.

(15) Wang, F. S.; Shyu, C. S. Optimal Feed Policy for Fed-Batch Fermentation of Ethanol Production by *Zymomous mobilis*. *Bioprocess Eng.* **1997**, *17*, 63−68.

(16) Stephanopoulos, G. N.; Aristidou, A. A.; Nielsen, J. *Metabolic Engineering: Principles and Methodologies*; Academic Press: New York, 1998.

(17) Kauffman, K. J.; Prakash, P.; Edwards, J. S. Advances in Metabolic Flux Analysis. *Curr. Opin. Biotechnol.* **2003**, *14*, 491−496.

(18) Segre, D.; Vitkup, D.; Church, G. M. Analysis of Optimality in Natural and Perturbed Metabolic Networks. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15112−15117.

(19) Mahadevan, R.; Edwards, J. S.; Doyle, F. J. Dynamic Flux Balance Analysis of Diauxic Growth in *Escherichia coli*. *Biophys. J.* **2002**, *83*, 1331−1340.

(20) Varma, A.; Palsson, B. O. Stoichiometric Flux Balance Models Quantitatively Predict Growth and Metabolic By-Product Secretion in Wild-Type *Escherichia coli*. *Appl. Environ. Microbiol.* **1994**, *60*, 3724−3731.

(21) Sainz, J.; Perez-Correa, F. R.; Agosin, E. Modeling of Yeast Metabolism and Process Dynamics in Batch Fermentation. *Biotechnol. Bioeng.* **2003**, *81*, 818−828.

(22) Famili, I.; Forster, J.; Nielsen, J.; Palsson, B. O. *Saccharomyces cerevisiae* Phenotypes can be Predicted by using Constraint-Based Analysis of a Genome-Scale Reconstructed Metabolic Network. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13134−13139.

(23) Gadkar, K. P.; Doyle, F. J.; Edwards, J. S.; Mahadevan, R. Estimating Optimal Profiles of Genetic Alterations using Constraint-Based Models. *Biotechnol. Bioeng.* **2004**, *89*, 243−251.

(24) Burgard, A. P.; Maranas, C. D. Optimization-Based Framework for Inferring and Testing Hypothesized Metabolic Objective Functions. *Biotechnol. Bioeng.* **2003**, *82*, 670−677.

(25) Burgard, A. P.; Pharkya, P.; Maranas, C. D. OptKnock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization. *Biotechnol. Bioeng.* **2003**, *84*, 647−657.

(26) Pharkya, P.; Burgard, A. P.; Maranas, C. D. Exploring the Overproduction of Amino Acids using the Bilevel Optimization Framework OptKnock. *Biotechnol. Bioeng.* **2003**, *84*, 887−899.

(27) Raghunathan, A. U.; Biegler, L. T. Mathematical Programs with Equilibrium Constraints (MPECs) in Process Engineering. *Comput. Chem. Eng.* **2003**, *27*, 1381−1392.

(28) Raghunathan, A. U.; Perez-Correa, J. R.; Biegler, L. T. Data Reconciliation and Parameter Estimation in Flux Balance Analysis. *Biotechnol. Bioeng.* **2003**, *84*, 700−709.

(29) van Gulik, W. M.; Heijnen, J. J. A. Metabolic Network Stoichiometry Analysis for Microbial Growth and Product Formation. *Biotechnol. Bioeng.* **1995**, *48*, 681−698.
(30) Vanrolleghem, P. A.; de Jong-Gubbels, P.; van Gulik, W. M.; Pronk, J. T.; van Dijken, J. P.; Heijnen, S. Validation of a Metabolic Network for *Saccharomyces cerevisiae* using Mixed Substrate Studies. *Biotechnol. Prog.* **1996**, *12*, 434−448.
(31) Nissen, T. L.; Schulze, U.; Nielsen, J.; Villadsen, J. Flux Distributions in Anaerobic, Glucose-Limited Continuous Cultures of *Saccharomyces cerevisiae*. *Microbiology* **1997**, *143*, 203-218.
(32) Sonnleitner, B.; Kappeli, O. Growth of *Saccharomyces cerevisiae* is Controlled by its Limited Respiratory Capacity: Formulation and Verification of a Hypothesis. *Biotechnol. Bioeng.* **1986**, *28*, 927−937.
(33) Phalakornkule, C.; Lee, S.; Zhu, T.; Koepsel, R.; Ataai, M. M.; Grossman, I. E.; Domach, M. M. A MILP−Based Flux Alternative Generation and NMR Experimental Design Strategy for Metabolic Engineering. *Metab. Eng.* **2001**, *3*, 124−137.
(34) Duarte, N. C.; Palsson, B. O.; Fu, P. Integrated Analysis of Metabolic Phenotypes in *Saccharomyces cerevisiae*. *BMC Genomics* **2004**, *5*, 63−73.
(35) Bader, G.; Ascher, U. A New Basis Implementation for a Mixed Order Boundary Value ODE Solver. *SIAM J. Sci. Comput.* **1987**, *8*, 483−500.
(36) Smets, I. Y.; Claes, J. E.; November, E. J.; Bastin, G. P.; Impe, J. F. V. Optimal Adaptive Control of (Bio)chemical Reactors: Past, Present and Future. *J. Process Control* **2004**, *14*, 795−805.
(37) Grosz, R.; Stephanopoulos, G. Physiological, Biochemical, and Mathematical Studies of Micro-Aerobic Continuous Ethanol Fermentation by *Saccharomyces cerevisiae*. I: Hysteresis, Oscillations, and Maximum Specific Ethanol Productivities in Chemostat Culture. *Biotechnol. Bioeng.* **1990**, *36*, 1006−1019.
(38) Hoppe, G. K.; Hansford, G. S. The Effect of Micro-Aerobic Conditions on Continuous Ethanol-Production by *Saccharomyces cerevisiae*. *Biotechnol. Lett.* **1984**, *6*, 681−686.
(39) Edwards, J. S.; Palsson, B. O. The *Escherichia coli* MG1655 *in silico* Metabolic Genotype: Its Definition, Characteristics and Capabilities. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5528−5533.
(40) Schilling, C. H.; Covert, M. W.; Famili, I.; Church, G. M.; Edwards, J. S.; Palsson, B. O. Genome-Scale Metabolic Model of *Helicobacter pylori* 26695. *J. Bacteriol.* **2002**, *184*, 4582−4593.
(41) Forster, J.; Famili, I.; Fu, P.; Palsson, B. O.; Nielsen, J. Genome-Scale Reconstruction of the *Saccharomyces cerevisiae* Metabolic Network. *Genome Res.* **2003**, *13*, 244−253.

Received February 28, 2006. Accepted for publication June 16, 2006.

BP060059V