ARTICLES

On the Matching and Proportional Laws of Cybernetic Models

Jamey D. Young† and Doraiswami Ramkrishna*

Forney Hall of Chemical Engineering, Purdue University, 480 Stadium Mall Drive, West Lafayette, Indiana 47907

The Matching and Proportional Laws are heuristic control policies that have found widespread use in cybernetic models of biological systems. Within this context, the laws serve as optimization surrogates for predicting the response of metabolic control circuits that modulate enzyme levels and activities. The key result of the current contribution is to demonstrate clearly the optimality properties of these laws and the assumptions that underlie their development. In doing so, we arrive at generalized versions of the Matching and Proportional Laws that are shown to collapse to the forms originally derived by Kompala et al. (*Biotechnol. Bioeng.* **1986**, *28*, 1044−1055) when certain simplifications are applied. As a further line of investigation, we show how Kompala et al.'s cybernetic laws compare with alternative control policies in their ability to describe diauxic growth behavior of microbial cultures. We find that Kompala et al.'s model describes the experimental observations more accurately than other limiting-case models that are either too aggressive or too passive in capturing the mixed-substrate growth rates and intermediate lag periods. Monte Carlo analysis of computational growth experiments in which strains obeying different regulatory policies directly compete for available nutrients reveals that the Matching and Proportional Law policy does not maximize the average growth rate of the culture. However, it allocates metabolic resources more frugally than other policies that outperform it and may be more realistic in reflecting the cell's true fitness-to-cost tradeoff as judged by its agreement with experimental growth data.

1. Introduction

Biological systems are characterized by their ability to grow and reproduce under diverse environmental conditions. This resiliency is owed to a knack for regulating their metabolic machinery in ways that ultimately promote survival in one way or another. Such behavior has led to the frequent use of teleology in explanations of biological phenomena, which has imposed a simultaneously detracting influence on the "scientific" essence of biology. Scientific laws are generally meant to spell the effect of a cause rather than to sense the future in any way. Mayr (*1*), in an eloquent support of the need for a grossly different perspective of biological systems, recommends the term "teleonomy" as a more respectable substitute for teleology. The implication of this suggestion is of course considerably deeper than an act of rechristening, however. It proposes that during evolution there has developed a genetic program providing guidelines for optimal response of the organism to a wide range of ecological pressures. This capacity for optimal behavior is based on a record of its past experience together with progressive development of a machinery for implementing such guidelines and is chiefly acquired through the process of natural selection. In the jargon of modern information technology, we might regard this as the organism acquiring through its learning experience both a software for survival strategies and the hardware necessary for implementing them. Viewed from the foregoing perspective, this "cybernetic" treatment of biological systems is eminently within the realm of rational science.

The cybernetic modeling framework of Ramkrishna (*2*) builds upon the idea that an organism's nutritional goals are realized entirely within the domain of chemical kinetics through judicious utilization of metabolic capabilities. Metabolism, the term used to denote the totality of chemical reactions in an organism, is enabled by a vast collection of enzymes with highly specific catalytic functions. Metabolic regulation is responsible for controlling the amounts and activities of these enzymes. The rates (or fluxes) of metabolic reactions are thus subject to rigorous control aimed at protecting the interests of the organism. It is essential that any valid treatment of biochemical reaction systems is able to incorporate the effects of metabolic regulation for it to be an effective scientific endeavor. Furthermore, in view of the essentially speculative nature of the goals of regulation, the need for a rational methodology that can be progressively refined through experimental observation cannot be overemphasized.

The strategic response of an organism to its environment is enacted by regulatory programs in which heuristics play a focal role. We use the term "heuristics" to connote a set of rules that a system, capable of learning, has acquired through extended experience. Further, we deem the mathematical representation of these heuristics to be time-invariant, at least on a time scale (much) smaller than that of evolution. This experience of the

* To whom correspondence should be addressed. Tel: +1-765-494-4066. Fax: +1-765-494-0805. Email: ramkrish@ecn.purdue.edu.

† Present address: Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue 56-439, Cambridge, MA 02139.

10.1021/bp060176q CCC: $37.00
Published on Web 12/07/2006

past is held as a tool for the organism's envisioning and fashioning of its future in the interest of survival. In mathematical terms, the heuristics serve as a basis for computing optimal strategies. We assume that the concept of optimality is tied up with the economic utilization of the organism's internal resources. It should then become apparent that the cybernetic modeling approach has a potential role in extracting the relationship between the metabolic function of an organism and its genomic background. Although the issue of relating genotype to phenotype has been the subject of intense recent interest, a rational methodology for the same has been elusive.

The chief objective of this article is to revisit the cybernetic modeling framework of Ramkrishna (*2*) from a fresh viewpoint. The exercise is significant in elucidating the exact nature of the optimality implied by previous cybernetic control laws. Two optimal laws have appeared in the foregoing cybernetic modeling literature. The first of these, concerned with regulation of enzyme synthesis, has been referred to as the Matching Law, and the second, associated with control of enzyme activity, has been termed the Proportional Law. Kompala et al. (*3*, *4*) derived these laws using relatively crude arguments, which we now attempt to refine. Besides providing clarity, the current development is prerequisite to the offshoot of generalizations leading to new cybernetic laws or to extended applicability of the existing laws. Moreover, this treatment enables a systematic comparison of alternate control policies that result from differing notions of optimality, thereby providing a context in which the unique properties of the Matching and Proportional Laws can be fully discerned and appreciated.

2. Background

Cybernetic modeling has been extensively applied over the past 20+ years to analyze and predict the dynamics of biological systems. It was the first modeling approach to emphasize the implied optimal behavior that results from selection through evolution and to substitute optimal control heuristics for a detailed description of regulatory mechanisms. This is a core virtue of the cybernetic modeling paradigm, since other approaches are severely hampered by their inability to describe regulatory effects whenever the underlying mechanisms are poorly understood. Even in cases where a great deal is known about the native regulatory pathways, system disturbances brought about by genetic and environmental perturbations can routinely elicit unforeseen responses. Any model that takes a purely mechanistic view of regulation will cope poorly with incomplete regulatory information and will produce unreliable predictions in situations where these missing details become important. In fact, because such a model often lacks the inherent robustness of the biological system it describes, it may fail completely under conditions very different from those that it was expressly built to handle. This will prevent the model from being effectively deployed as a predictive tool to discover novel phenotypes resulting from targeted manipulations.

Since cybernetic modeling was first introduced in the early 1980s, the notion of supplementing biological models with optimization heuristics has become well established in the literature, but its application has been largely restricted to steady-state models. Flux Balance Analysis (FBA) is one such approach that has met with considerable success in predicting experimentally determined flux changes resulting from gene knockouts or altered carbon source availability (reviewed in ref 5). FBA relies upon linear programming to compute the steady-state flux distribution that optimizes a postulated metabolic objective. Its main advantage is that it does not require estimation of kinetic parameters and therefore can be applied to genome-scale metabolic networks. A significant shortcoming of FBA is that because it invokes pseudo-steady-state approximations on intracellular metabolites, it cannot be extended to analyze the dynamics of intracellular processes. Hence, FBA models are inadequate for probing into some of the more complex aspects of metabolism, such as multistability, oscillatory behavior, transient phenomena, and the effects of reaction reversibilities or futile cycling. A second and more serious objection is that the computed fluxes are decoupled from the intracellular metabolite and enzyme levels, so there is no way to incorporate the kinetic and regulatory effects of these molecules into an FBA model. The same objections apply to Minimization of Metabolic Adjustment (MOMA) (*6*) and Regulatory On/Off Minimization (ROOM) (*7*), which are offshoots of FBA that impose different types of objective functions and/or constraints on the steady-state flux distributions.

Metabolic Control Analysis (MCA) represents a separate body of theory that has greatly influenced the way metabolic regulation is understood. MCA enables the systematic measurement and interpretation of indices describing the control of biochemical pathways (*8*, *9*). These so-called control coefficients reflect the sensitivities of the steady-state fluxes and metabolite levels within the pathway to variations in parameters such as enzyme levels or effector concentrations. MCA has become a popular framework for discovery and analysis of metabolic engineering strategies. However, it lacks the capability to predict system behavior under conditions very far removed from the steady state under which the control indices were derived. Since metabolic engineering usually involves large perturbations to the physiological state of the cell, it is doubtful that the control coefficients alone can serve as reliable guides for designing recombinant organisms.

In addition to steady-state modeling approaches, there has been a handful of attempts to construct detailed kinetic models of cellular metabolism. A few notable examples are found in the work of Domach et al. (*10*) and Chassagnole et al. (*11*) to model *E. coli* metabolism and that of Rizzi et al. (*12*) to model yeast. The difficulty in identifying suitable rate expressions, estimating parameters, and accounting for regulation has limited the scope of kinetic models to relatively small- to medium-sized networks of well-studied biochemical reactions. Most investigations are restricted to model strains of bacteria and yeast for which the important kinetic and regulatory features are thoroughly understood. Biochemical Systems Theory (BST) is a kinetic modeling approach that attempts to reduce the burden of model identification by using canonical power-law rate expressions (*13*, *14*). The major drawback of this approach, however, is that all regulation must be captured in the power-law exponents. As a result, identifying the regulatory structure of the model boils down to a high-dimensional parameter estimation problem, which can become quite challenging to solve. Furthermore, the resulting parameter values can hardly be considered to represent the system-level control policies of the organism. Much like the MCA coefficients, they merely reflect the local regulatory behavior in the neighborhood of the physiological state for which they were determined. There is no reason to assume that they should remain constant over a large swath of phenotypic space.

The goal of cybernetic modeling is to bridge the gap between steady-state approaches such as FBA, which have been successful in describing metabolic regulation using optimization

heuristics, and kinetic models that provide a more complete picture of cellular dynamics and are amenable to a wider range of analysis techniques. Unlike other dynamic modeling approaches, cybernetic modeling views regulation as a distinct subprocess of the cell, separate from the underlying kinetic properties of the biochemical network. In doing so, it attempts to capture the systemic features of regulation, those that must be present for the proper functioning of the system at large. This shift in focus from the mechanistic to the systemic results in improved model robustness and a greater range of model applicability. However, it has prompted some authors to classify cybernetic modeling as a "black box" approach. This is a completely unwarranted characterization since cybernetic control laws can be readily extended to detailed models of arbitrary size and complexity, as evidenced in the fairly elaborate model of *E. coli* metabolism presented by Varner (*15*). In reality, the network structure and kinetic expressions can be made as sophisticated as desired, and cybernetic controls can still be applied to enhance model robustness and improve the reliability of predictions. Like other dynamic models, however, the main difficulty in constructing cybernetic models lies in the formulation of kinetic expressions and the estimation of rate parameters. Therefore, cybernetic models should not be expanded indiscriminately but instead should be tailored to the problem at hand, taking into account the amount of data available for model identification and validation.

3. Matching Law

3.1. An Optimal Control Problem. Let $\mathbf{x}$ represent the state vector of the metabolic system of interest. The metabolic state is described by the vector of metabolite concentrations $\mathbf{y}$, the vector of enzyme levels $\mathbf{e}$ and the biomass concentration c, so it is convenient to partition $\mathbf{x}$ into

$$\mathbf{x} = \begin{bmatrix} \mathbf{y} \\ \mathbf{e} \\ c \end{bmatrix} \tag{1}$$

The dimension of $\mathbf{y}$ is denoted N_Y, and the dimension of $\mathbf{e}$ is likewise denoted N_E. Therefore, the dimension of $\mathbf{x}$ is $N_X = N_Y + N_E + 1$.

The system is subject to regulatory control inputs enacted at the transcriptional and translational levels that determine the enzyme synthesis rates. We account for these inputs by introducing the control vector $\mathbf{u}$, which specifies how transcriptional and translational resources are to be allocated among the various enzyme synthesis alternatives. More specifically, we assume that there exists a single key resource that limits the rate of enzyme synthesis, and u_i represents the fraction of this resource earmarked for enzyme E_i (*16*). Ignoring other possible regulatory inputs and system dependencies, the time evolution of this dynamical system is described by the differential equation model

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \tag{2}$$

The organism is presumed to regulate $\mathbf{u}$ in such a way that its metabolic performance index J is maximized, thus leading to optimal allocation of metabolic resources. Hence, we can view the organism's regulatory machinery as an analog computer that has evolved heuristic strategies for solving the optimal control problem

$$\begin{aligned} &\max J \\ &\text{s.t.} \quad \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \end{aligned} \tag{3}$$

From a modeling standpoint, our aim is to approximate the solution of eq 3 using a digital computer and thereby provide a system-level approach for predicting and analyzing the dynamics of metabolic regulation.

Before venturing further, there are a few philosophical concerns that warrant discussion. First, there is the problem of identifying a suitable performance index for the system of interest. The "true" performance index is, of course, unknowable to the modeler because it is a consequence of the evolutionary history of the organism. Therefore, some guesswork is implicit in our eventual choice of J. Despite this apparent ill-posedness, we are confined in our choice to functionals that (1) are consistent with a general understanding of evolutionary biology and reflect the organism's basic requirements for growth and survival and (2) are amenable to mathematical analysis. A second concern lies in the fact that even if we possess the "true" form of J, the problem of rigorously computing the optimal control in eq 3 requires numerical solution of a two-point boundary value problem. Such an approach is conceptually difficult and computationally expensive, which undermines our minimalist intentions for seeking out a cybernetic representation of regulatory mechanisms in the first place. The most tractable instance of eq 3 results when the performance index is quadratic in the control variable and the state equation is linear, so it is to our distinct advantage if we can re-express the relevant control problem in this form.

In light of the preceding discussion, we choose to simplify eq 3 by first assuming that regulatory decisions are made at each instant based on the projected system response over a (short) time interval of length Δt. We will refer to this time interval as the system's planning window. We further assume that the system response at time $t + \tau$ where $\tau \in [0, \Delta t]$ can be approximated by linearizing eq 2 about the current state $\mathbf{x}(t)$ and the reference control input $\mathbf{u}^\circ$ to give

$$\Delta\dot{\mathbf{x}} = \mathbf{A}\Delta\mathbf{x} + \mathbf{B}\Delta\mathbf{u} + \mathbf{f}(\mathbf{x}(t), \mathbf{u}^\circ)$$

$$\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}(t), \mathbf{u}^\circ), \quad \mathbf{B} = \frac{\partial \mathbf{f}}{\partial \mathbf{u}}(\mathbf{x}(t), \mathbf{u}^\circ)$$

$$\Delta\mathbf{x}(t + \tau) = \mathbf{x}(t + \tau) - \mathbf{x}(t) \tag{4}$$

$$\Delta\dot{\mathbf{x}} = \frac{\mathrm{d}}{\mathrm{d}\tau}\Delta\mathbf{x}$$

$$\Delta\mathbf{u}(t + \tau) = \mathbf{u}(t + \tau) - \mathbf{u}^\circ$$

It has proven useful to associate the reference input with the unregulated state of the system, e.g., $u_i^\circ = 1/N_E \ \forall i$, but this choice does not directly impact the present development and will not concern us here. Ultimately, it is left to the interpretation of the modeler. Note, however, that all cybernetic models examined to date have been linear in the control inputs, and so the matrix $\mathbf{B}$ would not be influenced by the choice of $\mathbf{u}^\circ$—only $\mathbf{A}$ is affected. It should also be pointed out that linearizing about $\mathbf{u}(t)$ is problematic because $\mathbf{u}(0)$ would then need to be specified along with the initial conditions, but the initial values of the control variables are generally unknown. We avoid this unnecessary complication by selecting a constant $\mathbf{u}^\circ$ about which to linearize the system.

Next, we model the change in performance index over the time interval $[t, t + \Delta t]$ as

$$\Delta J = \mathbf{q}^{\mathrm{T}}\Delta\mathbf{x}(t+\Delta t) - \frac{\sigma}{2}\int_t^{t+\Delta t}\mathbf{u}^{\mathrm{T}}\mathbf{u}\,\mathrm{d}\tau$$
$$\mathbf{q} = \frac{\partial\phi}{\partial\mathbf{x}}(\mathbf{x}(t)), \quad \Delta J = J(t+\Delta t) - J(t) \tag{5}$$

where the function $\phi(\mathbf{x})$ represents the metabolic objective of the system and σ is a parameter that scales the cost associated with resource investment. If we examine the two terms that comprise ΔJ separately, the first represents the accrued benefit that is derived from resources invested during the planning window, and the second represents the cost or penalty ascribed to these investments. The form of eq 5 is chosen to reflect the microeconomic principle of diminishing marginal utility per unit of resource input. The quadratic penalty term ensures that the per-unit cost of resource increases as its rate of consumption increases, which is dictated by the Laws of Supply and Demand. Taken as a whole, this performance index captures the inherent tradeoff that is required to meet nutritional requirements while simultaneously conserving scarce internal resources.

Instead of solving the full optimal control problem of eq 3 over the time course of the simulation, we now endeavor to compute the trajectory $\mathbf{u}(t)$ as the solution to a sequence of linear quadratic regulator (LQR) problems of the form

$$\max \Delta J = \mathbf{q}^{\mathrm{T}}\Delta\mathbf{x}(t+\Delta t) - \frac{\sigma}{2}\int_t^{t+\Delta t}\mathbf{u}^{\mathrm{T}}\mathbf{u}\,\mathrm{d}\tau$$
$$\text{s.t. } \Delta\dot{\mathbf{x}} = \mathbf{A}\Delta\mathbf{x} + \mathbf{B}\Delta\mathbf{u} + \mathbf{f}(\mathbf{x}(t), \mathbf{u}^\circ) \tag{6}$$

The LQR problem is solved at each time instant to determine the optimal open-loop control over the planning window $[t, t + \Delta t]$. The prescribed control action is injected at time t, which advances the system to the next time instant. From an algorithmic standpoint, this is analogous to the procedure that is implemented in receding horizon control of nonlinear engineering systems (*17*). Borrowing from the language of Kompala et al. (*3*), it is apparent that the control policy derived in this manner maintains a short-term perspective, in that only an assessment of the current state $\mathbf{x}(t)$ is required to compute the system parameters $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{q}$ upon which the optimal control depends. No foreknowledge of future events is implied, which is a major source of inconsistency in models that take a long-term perspective, such as the one presented in Dhurjati et al. (*18*).

The optimization problem summarized in eq 6 is readily solved by appealing to classical approaches in the area of optimal control theory (*19*). First, we define the Hamiltonian function

$$H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) = -\frac{\sigma}{2}\mathbf{u}^{\mathrm{T}}\mathbf{u} + \boldsymbol{\lambda}^{\mathrm{T}}[\mathbf{A}\Delta\mathbf{x} + \mathbf{B}\Delta\mathbf{u} + \mathbf{f}(\mathbf{x}(t), \mathbf{u}^\circ)] \tag{7}$$

where $\boldsymbol{\lambda}$ is the N_X-dimensional costate vector. The stationarity condition gives

$$0 = \frac{\partial H}{\partial\mathbf{u}} = -\sigma\mathbf{u} + \mathbf{B}^{\mathrm{T}}\boldsymbol{\lambda} \tag{8}$$

which can be rearranged to express the optimal control in terms of the costate

$$\mathbf{u} = \frac{1}{\sigma}\mathbf{B}^{\mathrm{T}}\boldsymbol{\lambda} \tag{9}$$

Next, the costate equation

$$-\dot{\boldsymbol{\lambda}} = \frac{\partial H}{\partial\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\boldsymbol{\lambda} \tag{10}$$

must be solved subject to the boundary condition

$$\boldsymbol{\lambda}(t+\Delta t) = \mathbf{q} \tag{11}$$

The solution to this linear differential equation is readily obtained as

$$\boldsymbol{\lambda}(t+\tau) = \mathbf{e}^{\mathbf{A}^{\mathrm{T}}(\Delta t-\tau)}\mathbf{q}, \quad 0 \le \tau \le \Delta t \tag{12}$$

Finally, substitution of eq 12 into eq 9 gives the result

$$\mathbf{u}(t+\tau) = \frac{1}{\sigma}\mathbf{B}^{\mathrm{T}}\mathbf{e}^{\mathbf{A}^{\mathrm{T}}(\Delta t-\tau)}\mathbf{q}, \quad 0 \le \tau \le \Delta t \tag{13}$$

Since we are ultimately interested in determining the optimal control input at the current time t, we set $\tau = 0$ in eq 13 to arrive at a heuristic expression for $\mathbf{u}(t)$,

$$\mathbf{u}(t) = \frac{1}{\sigma}\mathbf{B}^{\mathrm{T}}\mathbf{e}^{\mathbf{A}^{\mathrm{T}}\Delta t}\mathbf{q} \tag{14}$$

Let us now express this relationship in a more convenient form. Denoting the ith column of $\mathbf{B}$ as $\mathbf{b_i}$, the "return-on-investment" for resource allocated to the ith enzyme is defined as

$$p_i(t) \equiv \mathbf{b_i}^{\mathrm{T}}\mathbf{e}^{\mathbf{A}^{\mathrm{T}}\Delta t}\mathbf{q} = \mathbf{q}^{\mathrm{T}}\mathbf{e}^{\mathbf{A}\Delta t}\mathbf{b_i} \tag{15}$$

The return-on-investment, p_i, reflects the expected metabolic benefit derived from synthesis of enzyme E_i, which subsequently catalyzes reaction R_i. In other words, it represents the extent to which the metabolic objective $\phi(\mathbf{x})$ can be expected to improve over the future time interval of length Δt due to control actions implemented at the current time t. We can now re-express u_i in terms of the return-on-investment by substituting eq 15 into eq 14 to give the result

$$u_i = \frac{p_i}{\sigma} \tag{16}$$

3.2. Constrained Inputs. In our attempt to focus on the essential features of the Matching Law development, we have thus far neglected to discuss the effects of input constraints on the optimal control policy. However, it is apparent that the constraint $u_i \ge 0$ must be applied to each element of the control vector $\mathbf{u}$, thereby restricting resource investments to be nonnegative. Furthermore, the total resource investment is bounded by the resource availability, giving rise to a constraint of the form $\sum_{i=1}^{N_R} u_i \le 1$, where N_R is the number of reactions in the metabolic network. Pontryagin's maximum principle can be used to derive a more general form of the stationarity condition (previously given by eq 8) that is applicable to constrained input problems (*20*),

$$\mathbf{u}^* = \arg\max_{\mathbf{u}\in D} H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) \tag{17}$$

Here we have used $\mathbf{u}^*$ to denote the optimal control. For the present case, the admissible region is described by $D = \{\mathbf{u} : \sum_{i=1}^{N_R} u_i \le 1, u_i \ge 0\}$. Because H is a concave function of $\mathbf{u}$ and because D is circumscribed by linear constraints, the maximization in eq 17 represents a convex optimization problem. Therefore, any solution that satisfies the Karush−Kuhn−Tucker (KKT) conditions is a global optimum (*21*).

We can simplify eq 17 by removing those terms in H that do not involve $\mathbf{u}$ to arrive at the result

$$\mathbf{u}^* = \arg\max_{\mathbf{u}\in D}\left(-\frac{\sigma}{2}\mathbf{u}^{\mathrm{T}}\mathbf{u} + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{B}\mathbf{u}\right) \tag{18}$$

Applying the KKT conditions to eq 18 gives the following set of constraints that characterize the optimal control policy at time t:

$$-\sigma u_i + \underbrace{\mathbf{b_i}^{\mathrm{T}}\boldsymbol{\lambda}}_{p_i} = \eta - \nu_i \tag{19}$$

$$\nu_i u_i = 0 \tag{20}$$

$$\eta\left(1 - \sum_{i=1}^{N_R} u_i\right) = 0 \tag{21}$$

$$u_i \geq 0 \tag{22}$$

$$\sum_{i=1}^{N_R} u_i \leq 1 \tag{23}$$

$$\nu_i \geq 0 \tag{24}$$

$$\eta \geq 0 \tag{25}$$

where η is the Lagrange multiplier associated with the total resource constraint (eq 23) and ν_i is the Lagrange multiplier associated with the ith non-negativity constraint (eq 22). The solution that simultaneously satisfies these conditions is expressed as

$$u_i = \max\left(\frac{p_i - \eta}{\sigma}, 0\right) \tag{26}$$

The Lagrange multiplier η describes the impact of the total resource constraint (eq 23) on the optimal control. Its value must satisfy the non-negativity constraint $\eta \geq 0$ and the equality constraint that comes from substituting eq 26 into eq 21:

$$\eta\left(1 - \frac{1}{\sigma}\sum_{i=1}^{N_R}\max(p_i - \eta, 0)\right) = 0 \tag{27}$$

When $\sigma > \sum_{i=1}^{N_R}\max(p_i, 0)$, this equation implies that $\eta = 0$ and the constraint expressed in eq 23 is inactive. Under such a condition, the transcriptional resource is underutilized because the fractional allocations sum to less than unity. If we consider the extreme limit of $\sigma \rightarrow \infty$, the investment penalty is so severe that all resource is withheld, resulting in $\mathbf{u} = \mathbf{0}$. On the other hand, if $\sigma < \sum_{i=1}^{N_R}\max(p_i, 0)$, we have $\eta > 0$ and resource is fully utilized. In the extreme case when $\sigma = 0$, the optimal control is a Bang-Bang policy in which resource is allocated exclusively to the reaction with the greatest return-on-investment. Neither of the extreme cases $\sigma \rightarrow \infty$ or $\sigma = 0$ represents an economically efficient and robust system. If the cost parameter σ is too large there will be wastage of resource, yet if σ is too small the organism will fail to maintain a diverse set of enzymes that are needed to respond to changing conditions, the ecological equivalent of "putting all its eggs into a single basket". In order to strike a heuristic balance between these two extremes, we postulate that $\sigma \approx \sum_{i=1}^{N_R}\max(p_i, 0)$ and therefore $\eta \approx 0$. Equation 26 then simplifies to

$$u_i = \frac{\max(p_i, 0)}{\sum_{n=1}^{N_R}\max(p_n, 0)} \tag{28}$$

which is the generalized form of the Matching Law. This result contrasts with the original form derived by Kompala et al. (*3*), which can be simply stated as

$$u_i = \frac{p_i}{\sum_{n=1}^{N_R} p_n} \tag{29}$$

but the model considered in that work lacked the possibility of negative returns-on-investment. The modified form expressed in eq 28 should be used whenever some of the reactions in the network have the potential to detract from the metabolic objective and hence generate negative returns. The choice of σ that gives rise to eq 28 corresponds to the threshold value at which the resource constraint eq 23 just becomes active. Except in the singular case when none of the network reactions offer a positive return-on-investment, the available resource is fully utilized and the optimal control will always satisfy $\sum_{i=1}^{N_R} u_i = 1$.

3.3. System Identification. In addition to the cost parameter σ discussed in the previous section, there are a number of other system parameters that must be identified before the optimal control policy expressed in eqs 15 and 28 can be readily computed. These include the state transition matrix **A**, the input matrix **B**, the objective gradient **q**, and the length of the planning window Δt. The state transition matrix **A** and the input matrix **B** are directly obtained by linearizing **f** in eq 2. It may be desirable to simplify the expressions for **A** and **B** obtained in this manner to produce expressions that are cheaper to evaluate computationally or can be more easily manipulated to arrive at analytical results. Since we are primarily interested in deriving a heuristic control policy that is minimalist in nature, we should openly embrace such simplifications when they lead to a corresponding reduction in complexity. However, we must simultaneously guard against oversimplifications that degrade controller performance because the resulting system matrices fail to adequately represent the underlying system dynamics.

Once a metabolic objective function $\phi(\mathbf{x})$ has been identified for the system of interest, the gradient vector $\mathbf{q} = \partial\phi/\partial\mathbf{x}(\mathbf{x}(t))$ is readily computed at each time instant. It is beyond the scope of the current discussion to outline specific strategies for identifying the functional form of ϕ, although several examples can be found in the cybernetic modeling literature. Suffice it to say that the objective functions applied in prior investigations have been selected to reflect supposed evolutionary outcomes or key physiological requirements of the organism. For instance, early lumped models relied upon objective functions that involved the biomass concentration or external substrate concentrations in such a way that growth rate (*3*) or substrate uptake (*22*) was maximized. On the other hand, later work with structured models attempted to develop objective functions that directly characterized the material and energy requirements of the organism (*15*, *23*). In some cases, empirical knowledge of how the flux distribution within the network responds to targeted perturbations has been used to infer objective function candidates (*24*). It should be noted, however, that there may exist multiple objective functions that give rise to the same qualitative model behavior and cannot be distinguished on the basis of the available experimental data. In this case, any member of the

class of similar objective functions can be employed to approximate the observed regulatory effects, and further model refinement must await the generation of discriminatory data sets. There is a wide and diverse literature on optimal experiment design and model discrimination, which describes methods for selecting experimental inputs that lead to maximal discernment between candidate models (*25*, *26*). In addition, Burgard and Maranas (*27*) present a novel approach for inferring and testing metabolic objective functions using experimentally derived flux distributions obtained, for example, from ^{13}C tracer studies.

The interval Δt represents the characteristic time scale over which the regulatory circuit attempts to reckon the consequences of its present control actions. There is no logical inconsistency in proposing that organisms have evolved mechanisms to optimize the expected future benefit of their current actions, even when these actions are involuntary and occur without conscious deliberation. For instance, whenever a cell initiates transcription of a particular gene, a definite time must pass before the nascent mRNA strand is completely transcribed, and even more time must elapse before this message is subsequently translated into a functional protein. Therefore, regulatory programs must allocate transcriptional resources in a way that, on average, will best meet the anticipated needs of the cell at some later time, and the length of this anticipatory period is symbolized by the parameter Δt. If this time scale is too short, the cell may refuse to allocate available resources whenever the resulting benefits are not reaped immediately. On the other hand if the time scale is too long, system disturbances and other unanticipated effects degrade the reliability of the control circuit and result in poor regulatory performance. Obviously, an appropriate value of Δt must be commensurate with the characteristic times that are inherent in the system dynamics, but in general there is a variety of disparate time scales associated with eq 2 and with its linearized analogue eq 4. One logical choice is to associate Δt with the minimum time scale of the linearized system eq 4,

$$\Delta t = \frac{1}{\mu(\mathbf{A})} \tag{30}$$

$$\mu(\mathbf{A}) = \max\{\lambda : \lambda \in \sigma((\mathbf{A} + \mathbf{A}^{\mathrm{T}})/2)\} \tag{31}$$

where $\sigma(\mathbf{A})$ is the spectrum of $\mathbf{A}$, viz., the set of all eigenvalues of $\mathbf{A}$. The scalar $\mu(\mathbf{A})$ is the logarithmic norm of $\mathbf{A}$ and has several interesting properties related to the growth of solutions to ODEs (*28*). In particular, $\mu(\mathbf{A}) \geq \max\{|\mathrm{Re}(\lambda)| : \lambda \in \sigma(\mathbf{A})\}$. The resulting value of p_i computed from eq 15 therefore reflects the impact that regulatory actions have upon the fastest system dynamics. This choice of Δt is not only appealing from a theoretical standpoint, but it has computational advantages as well. It may turn out that $\mathbf{A}$ becomes temporarily unstable at certain points in the system trajectory, viz., it could have eigenvalues with positive real parts. When this occurs, the evaluation of the exponential matrix $\mathbf{e}^{\mathbf{A}\Delta t}$ becomes difficult or impossible for large values of Δt due to numerical overflow. This problem cannot arise if the planning window is at least as short as that described by eq 30.

In some instances, only the immediate consequences of the injected control actions need to be considered when evaluating $\mathbf{u}$, and we can safely set $\Delta t = 0$ without compromising the integrity of the model. This leads to a drastic reduction in computational effort, since eq 15 simplifies to

$$p_i(t) = \mathbf{q}^{\mathrm{T}}\mathbf{b}_{\mathrm{i}} \tag{32}$$

and we are no longer required to evaluate $\mathbf{A}$ or its associated eigenvalues and matrix exponential. In fact, section 6 describes how the returns-on-investment identified in the work of Kompala et al. (*3*) can be considered a special case of eq 32. However, Young (*29*) presents other situations in which setting $\Delta t = 0$ leads to undesirable (and unstable) model behaviors. For example, a linear sequence of reactions with short-time-scale intermediates is destabilized by using eq 32 to compute returns-on-investment, leading to unrealistic model behaviors such as severe accumulation or depletion of intermediate metabolites.

4. Proportional Law

4.1. Revisiting the Optimal Control Problem. In contrast to the Matching Law, which provides a heuristic policy for describing the transcriptional controls contained in the vector $\mathbf{u}$, the Proportional Law addresses allosteric and covalent regulatory controls that modulate enzyme activities. We therefore introduce the control vector $\mathbf{v}$ and define v_i as the relative activity of the enzyme catalyzing the ith reaction. By "relative activity" we mean the activity relative to the maximum possible for a given amount of enzyme. Once again appealing to the premise that biological control circuits have evolved optimal strategies for regulating the available metabolic machinery, we flesh out the Proportional Law statement by following essentially the same development as in section 3. In doing so, we momentarily ignore the effects of transcriptional and translational control and reinterpret the optimal control problem of eq 6 as

$$\begin{gathered}\max \Delta J = \mathbf{q}^{\mathrm{T}}\Delta\mathbf{x}(t + \Delta t) - \frac{\sigma}{2}\int_t^{t+\Delta t}\mathbf{v}^{\mathrm{T}}\mathbf{v}\,\mathrm{d}\tau \\ \text{s.t. } \Delta\dot{\mathbf{x}} = \mathbf{A}\Delta\mathbf{x} + \mathbf{B}\Delta\mathbf{v} + \mathbf{f}(\mathbf{x}(t), \mathbf{v}^{\circ})\end{gathered} \tag{33}$$

where we have simply substituted the control $\mathbf{v}$ in place of $\mathbf{u}$. Note that eq 33 implies that there is a quadratically increasing cost associated with the activation of enzymes, which reflects the notion of diminishing marginal returns as discussed in section 3. Also note that the $\mathbf{B}$ matrix involved in eq 33 necessarily differs from that of the previous section because the dependence of $\dot{\mathbf{x}}$ upon $\mathbf{v}$ is different from its dependence upon $\mathbf{u}$. While keeping this subtle yet important detail in mind, it is apparent that the structure of eq 33 is identical to that of eq 6, and the optimal control is therefore expressed as

$$v_i = \frac{p_i}{\sigma} \tag{34}$$

where the form of p_i is given by eq 15.

4.2. Constrained Inputs. As before, we have neglected the effects of input constraints in arriving at eq 34, and we will now endeavor to remedy this apparent shortcoming. First, each element of the control vector $\mathbf{v}$ must satisfy $v_i \geq 0$ to avoid negative activities, which have no physical interpretation. An upper bound constraint must also be applied to the elements of $\mathbf{v}$ to avoid unrealistically large activities. In the seminal work of Kompala et al. (*3*), a summation constraint similar to eq 23 was applied to the cybernetic $\mathbf{v}$-variable, but this constraint was later found to be inconsistent with experimental observations (*4*). It was then modified to a simple bound constraint $v_i \leq 1$ reflecting the fact that, by definition, the relative activity cannot exceed unity. Applying Pontryagin's maximum principle to the Hamiltonian function that results from eq 33 gives rise to a convex optimization problem. The KKT conditions describing the global optimum are

$$-\sigma v_i + p_i = \eta_i - \nu_i \tag{35}$$

$$\nu_i v_i = 0 \tag{36}$$

$$\eta_i(1 - v_i) = 0 \tag{37}$$

$$v_i \geq 0 \tag{38}$$

$$v_i \leq 1 \tag{39}$$

$$\nu_i \geq 0 \tag{40}$$

$$\eta_i \geq 0 \tag{41}$$

where η_i and ν_i are the Lagrange multipliers associated with the ith upper and lower bound constraints, respectively. The solution that simultaneously satisfies these conditions is expressed as

$$v_i = \max\left(\frac{p_i - \eta_i}{\sigma}, 0\right) \tag{42}$$

To identify the cost parameter σ, we need to make one additional assumption about the nature of **v**. Following the work of Kompala et al. (*4*), we stipulate that the reaction(s) with the highest return-on-investment shall be fully activated, viz.,

$$v_i = 1, \text{ if } p_i = \max_{n \in \{1, ..., N_R\}}(p_n) \tag{43}$$

Except in the singular case when all returns-on-investment are zero or negative, eqs 41−43 together imply that (*30*)

$$\sigma \leq \max_n(p_n) \tag{44}$$

When σ takes on the maximum value allowed by this inequality, all reactions with $p_i < \max_n(p_n)$ have relative activities $v_i < 1$. Because the constraint given by eq 39 is inactive for these reactions, the Lagrange multiplier η_i must equal zero. We then obtain the control law

$$v_i = \frac{\max(p_i, 0)}{\max_n(p_n)} \tag{45}$$

Choosing σ to be less than the maximum value of $\max_n(p_n)$ leads to a situation where some reactions may become fully activated even when their returns-on-investment are less than the maximum available. In the limit of $\sigma = 0$, we reach a condition in which all reactions with positive returns-on-investment are fully activated. This obviously results in an unrealistic control law because no preference is given to reactions that yield higher metabolic benefits, a result that is tantamount to removing all activity regulation from the network. In light of these problems, we select $\sigma = \max_n(p_n)$ and settle upon eq 45 as the generalized statement of the Proportional Law. This result can be simplified to the form used by Kompala et al. (*4*),

$$v_i = \frac{p_i}{\max_n(p_n)} \tag{46}$$

whenever there is no possibility of p_i becoming less than zero.

Historically, the Matching and Proportional Laws have been treated as distinct control policies, with separate justifications given for each one. The current development reveals that both are actually manifestations of the same optimal control law, albeit with different input constraints applied. In fact, both are forms of proportional control in that the resulting expressions for u_i and v_i are proportional to the associated return-on-investment p_i. However, each expression is scaled by a different normalization factor, which is determined by the value of the cost parameter σ.

5. Tandem Application of Matching and Proportional Laws

As we are ultimately interested in studying the simultaneous effects of transcriptional/translational control and enzyme activity control upon the phenotype of the organism, we should combine the results of sections 3 and 4 to arrive at a full description of the regulatory dynamics. The ODE model $\dot{\mathbf{x}} = \mathbf{f(x, u, v)}$ is linearized to give

$$\Delta\dot{\mathbf{x}} = \mathbf{A}\Delta\mathbf{x} + \mathbf{B_u}\Delta\mathbf{u} + \mathbf{B_v}\Delta\mathbf{v} + \mathbf{f}(\mathbf{x}(t), \mathbf{u}^\circ, \mathbf{v}^\circ)$$

$$\mathbf{A} = \frac{\partial\mathbf{f}}{\partial\mathbf{x}}(\mathbf{x}(t), \mathbf{u}^\circ, \mathbf{v}^\circ)$$

$$\mathbf{B_u} = \frac{\partial\mathbf{f}}{\partial\mathbf{u}}(\mathbf{x}(t), \mathbf{u}^\circ, \mathbf{v}^\circ) \tag{47}$$

$$\mathbf{B_v} = \frac{\partial\mathbf{f}}{\partial\mathbf{v}}(\mathbf{x}(t), \mathbf{u}^\circ, \mathbf{v}^\circ)$$

The optimal control problem that applies during the planning window $[t, t + \Delta t]$ is formulated as

$$\max \Delta J = \mathbf{q}^{\mathrm{T}}\Delta\mathbf{x}(t + \Delta t) - \frac{1}{2}\int_t^{t+\Delta t}[\sigma_u\mathbf{u}^{\mathrm{T}}\mathbf{u} + \sigma_v\mathbf{v}^{\mathrm{T}}\mathbf{v}]\,\mathrm{d}\tau$$
$$\text{s.t. } \Delta\dot{\mathbf{x}} = \mathbf{A}\Delta\mathbf{x} + \mathbf{B_u}\Delta\mathbf{u} + \mathbf{B_v}\Delta\mathbf{v} + \mathbf{f}(\mathbf{x}(t), \mathbf{u}^\circ, \mathbf{v}^\circ) \tag{48}$$

Stated in this form, the tasks of determining optimal policies for **u** and **v** are decoupled from each other, and we can apply the results of previous sections to compute the control inputs

$$\mathbf{u} = \frac{\mathbf{p_u^+}}{||\mathbf{p_u^+}||_1} \tag{49}$$

$$\mathbf{v} = \frac{\mathbf{p_v^+}}{||\mathbf{p_v^+}||_\infty} \tag{50}$$

where $||\mathbf{p_u^+}||_1$ and $||\mathbf{p_v^+}||_\infty$ represent the l_1 and l_∞ vector norms, respectively. The modified returns-on-investment that appear in these equations are defined to reflect only positive contributions to the metabolic objective,

$$p_{ij}^+ = \max(p_{ij}, 0), \quad (i, j) \in \{1, ..., N_R\} \times \{u, v\} \tag{51}$$

which allows eqs 49 and 50 to be expressed more compactly. They are computed readily once the returns-on-investment

$$\mathbf{p_j} = \mathbf{B_j}^{\mathrm{T}}\mathbf{e}^{\mathbf{A}^{\mathrm{T}}\Delta t}\mathbf{q}, \mathbf{j} \in \{\mathbf{u}, \mathbf{v}\} \tag{52}$$

have been evaluated.

6. Reconciliation with Past Cybernetic Modeling Approaches

The first cybernetic model to make use of the Matching and Proportional Laws was that of Kompala et al. (*4*). This model describes bacterial growth on mixtures of substitutable sub-

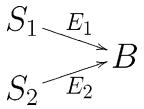

Figure 1. Example network for growth on two substitutable substrates.

strates, especially under conditions that give rise to diauxie. The aim of our current discussion is to show that the expressions for **u** and **v** employed in Kompala et al.'s model can result as a special case of eqs 49−52, thereby establishing that the treatment of section 5 is in fact a generalization and extension of prior work in the cybernetic modeling area. Let us focus on the simplest case of two substitutable substrates, S_1 and S_2, which react to form biomass B according to the network in Figure 1.The state vector is

$$\mathbf{x} = \begin{bmatrix} \mathbf{s} \\ \mathbf{e} \\ c \end{bmatrix} \tag{53}$$

where $\mathbf{s} = [s_1\ s_2]^T$ and $\mathbf{e} = [e_1\ e_2]^T$. Assuming that the enzymes are degraded according to a first-order process, the system dynamics are modeled according to

$$\frac{d\mathbf{s}}{dt} = \mathbf{S}\,\mathrm{diag}(\mathbf{v})\,\mathbf{r}c \tag{54}$$

$$\frac{d\mathbf{e}}{dt} = \mathrm{diag}(\mathbf{u})\mathbf{r}_\mathbf{E} - \mathrm{diag}(\boldsymbol{\beta})\mathbf{e} - \mu\mathbf{e} \tag{55}$$

$$\frac{dc}{dt} = \mu c \tag{56}$$

where the "diag" operator forms a diagonal matrix from its argument and $\boldsymbol{\beta} = [\beta_1\ \beta_2]^T$ contains the enzyme degradation rate constants. The stoichiometric matrix **S** contains the biomass yields of the two substrates,

$$\mathbf{S} = \begin{bmatrix} -\frac{1}{Y_1} & 0 \\ 0 & -\frac{1}{Y_2} \end{bmatrix} \tag{57}$$

The elements of the reaction rate vector **r** and the enzyme synthesis rate vector $\mathbf{r_E}$ are computed using standard kinetic expressions:

$$r_i = k_i e_i \frac{s_i}{K_i + s_i}\ \forall i \in \{1, 2\} \tag{58}$$

$$r_{E_i} = \alpha_i \frac{s_i}{K_i + s_i}\ \forall i \in \{1, 2\} \tag{59}$$

The rate r_i represents the growth rate that would be derived from consumption of S_i if the catabolic enzyme E_i were fully activated. The actual growth rate μ involves contributions from each element of **r**, weighted by the relative activity of the corresponding enzyme:

$$\mu = \mathbf{v}^T\mathbf{r} \tag{60}$$

To complete the model formulation, the following parameters should be identified from experimental data: the biomass yields Y_1 and Y_2, the reaction rate constants k_1 and k_2, the saturation constants K_1 and K_2, the enzyme synthesis rate constants α_1 and α_2, and the enzyme degradation rate constants β_1 and β_2.

The objective function chosen by Kompala et al. was inspired by the observation that organisms undergoing diauxic growth nearly always consume the substrate that gives the highest growth rate in preference to other available substrates. Thus the metabolic objective was postulated as $\phi(\mathbf{x}) = x_5 = c$, reflecting the hypothesis that regulatory control actions should drive the system toward maximum growth. The authors then relied upon qualitative arguments in concluding that the corresponding expressions for **u** and **v** should be written as

$$\mathbf{u} = \frac{\mathbf{r}}{||\mathbf{r}||_1} \tag{61}$$

$$\mathbf{v} = \frac{\mathbf{r}}{||\mathbf{r}||_\infty} \tag{62}$$

Comparing these expressions with eqs 49−51 allows us to conclude that the two results are equivalent if $\mathbf{p_u} = \mathbf{p_v} = \mathbf{r}$. Our goal is now to determine whether these returns-on-investment are consistent with eq 52.

First of all, we recognize that Kompala et al.'s approach does not consider any indirect effects of the injected control inputs (because the matrix exponential does not factor into the evaluation of returns-on-investment). Only the instantaneous effects of transcriptional/translational resource allocation and enzyme activation are included in the computation of $\mathbf{p_u}$ and $\mathbf{p_v}$, which implies setting $\Delta t = 0$ in eq 52. Computing the gradient of ϕ gives $\mathbf{q} = [0\ 0\ 0\ 0\ 1]^T$, and therefore eq 52 reduces to

$$p_{ij} = b_{5ij}, \quad (i, j) \in \{1, 2\} \times \{u, v\} \tag{63}$$

where b_{5ij} is the (5, i)-element of the $\mathbf{B_j}$ matrix. The relevant input matrix for the cybernetic **v**-variable is

$$\mathbf{B_v} = \begin{bmatrix} -r_1 c/Y_1 & 0 \\ 0 & -r_2 c/Y_2 \\ -r_1 e_1 & -r_2 e_1 \\ -r_1 e_2 & -r_2 e_2 \\ r_1 c & r_2 c \end{bmatrix} \tag{64}$$

Plugging the last row of this matrix into eq 63 implies that $\mathbf{p_v} = \mathbf{r}c$, but the factor of c will cancel from the numerator and denominator of eq 62, allowing us to simply write $\mathbf{p_v} = \mathbf{r}$ as desired. Thus, we have reconciled Kompala et al.'s rendering of the Proportional Law with the treatment of section 5.

The task of reconciling eq 61 with eq 49 is not as straightforward, however. Because the **u**-variable only influences the enzyme synthesis rates and does not directly impact the reaction rates, the input matrix in this case becomes

$$\mathbf{B_u} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ r_{E_1} & 0 \\ 0 & r_{E_2} \\ 0 & 0 \end{bmatrix} \tag{65}$$

There are no entries in row 5 of $\mathbf{B_u}$, so we must conclude from eq 63 that $\mathbf{p_u} = \mathbf{0}$ and consequently $\mathbf{u} = \mathbf{0}$. This is obviously not the result we are seeking. In fact, this comes as a direct result of setting $\Delta t = 0$ in arriving at eq 63. Because transcriptional/translational resource investment does not directly increase the biomass level, it is considered fruitless as a result of the overly myopic planning window. However, our intuition disagrees with this result, because we know that resource

investment is required to synthesize enzymes, and these enzymes are needed to catalyze biomass formation at times beyond the present. To circumvent this problem, we pretend that the enzyme levels are at quasi-steady state when writing the expression for $\mathbf{B_u}$. It is important to note that this quasi-steady-state assumption does not change the balance equations that describe the system dynamics but only affects the form of $\mathbf{B_u}$ used to compute the return-on-investment vector $\mathbf{p_u}$. Setting $d\mathbf{e}/dt = 0$ in eq 55 and solving for the steady-state value of $\mathbf{e}$ leads to

$$e_i^{ss} = \frac{r_{E_i} u_i}{\beta_i + \mu} \equiv e_i^\circ u_i \tag{66}$$

Here we have defined the reference enzyme level e_i° as $r_{E_i}/(\beta_i + \mu)$, which corresponds to the steady-state level of the fully induced enzyme under the current conditions. Replacing e_i with e_i^{ss} in the kinetic expression eq 58 leads to the following expression for $\mathbf{B_u}$:

$$\mathbf{B_u} = \begin{bmatrix} -\dfrac{r_1 v_1^\circ e_1^\circ c}{e_1 Y_1} & 0 \\ 0 & -\dfrac{r_2 v_2^\circ e_2^\circ c}{e_2 Y_2} \\ r_{E_1} & 0 \\ 0 & r_{E_2} \\ \dfrac{r_1 v_1^\circ e_1^\circ c}{e_1} & \dfrac{r_2 v_2^\circ e_2^\circ c}{e_2} \end{bmatrix} \tag{67}$$

Assuming the reference control inputs v_1° and v_2° correspond to the unregulated network state, viz., $v_i^\circ = 1\ \forall i$, leads us to conclude that the elements of the return-on-investment vector $\mathbf{p_u}$ should be written as

$$p_{u_i} = r_i \left(\frac{e_i^\circ}{e_i} \right) \tag{68}$$

Although this provides a nontrivial expression for $\mathbf{u}$, it is still inconsistent with Kompala et al.'s result $p_{u_i} = r_i$. To remove this discrepancy, we need to scale each return-on-investment by the factor e_i/e_i°. Introducing this additional factor in the definition of p_{u_i} causes the unwanted terms in eq 68 to cancel off, leaving only $\mathbf{p_u} = \mathbf{r}$.

What is the implication of scaling each return-on-investment by e_i/e_i° when evaluating the Matching Law expression? From a theoretical standpoint, this is equivalent to adjusting the penalty term in eq 48 to produce the modified performance index:

$$\Delta J = \mathbf{q}^{\mathrm{T}} \Delta \mathbf{x}(t + \Delta t) - \frac{1}{2} \int_t^{t+\Delta t} [\sigma_u \mathbf{u}^{\mathrm{T}} \mathbf{R} \mathbf{u} + \sigma_v \mathbf{v}^{\mathrm{T}} \mathbf{v}]\, d\tau \tag{69}$$

where the weighting matrix $\mathbf{R}$ is given by

$$\mathrm{R} = \begin{bmatrix} e_1^\circ/e_1 & 0 \\ 0 & e_2^\circ/e_2 \end{bmatrix} \tag{70}$$

The transcriptional resource investment penalty in this updated performance index is inversely proportional to the relative enzyme level. Therefore, investment in enzyme E_i is penalized more severely when e_i is small relative to e_i°. As a result, the kinetics of enzyme synthesis are made autocatalytic as a result of the dependence of u_i upon e_i. Narang et al. (*31*) have shown that the autocatalytic nature of enzyme synthesis in Kompala et al.'s model is primarily responsible for its ability to describe diauxic behavior. Without this feature, the model would not admit solutions characteristic of diauxie. As discussed in Narang et al. (*32*), autocatalysis in the kinetics of gene expression gives rise to global enzyme dynamics that are qualitatively similar to those exhibited by species populations in the Lotka−Volterra model of multispecies competition. In this context, the enzymes play the role of species that compete for a common resource pool. The enzyme with the highest return-on-investment is induced to the greatest extent according to the Matching Law, and this enzyme dominates the other competing enzymes. Unless an environmental perturbation leads to a shift in dominance, the less preferred enzymes are depleted by dilution and degradation until they become virtually extinct.

Inclusion of the $\mathbf{R}$ weighting matrix in eq 69 confers an inertial quality to the enzyme dynamics by promoting synthesis of enzymes already at high levels over those at low levels. As shown by Namjoshi and Ramkrishna (*33*), this inertia leads to bistable switching behavior and hysteresis effects in models of continuous cell cultures. Switching behavior of this sort is not an unusual occurrence in biological systems. In fact, it arises in a wide variety of contexts including cell signaling pathways, regulation of cell-cycle events, cellular differentiation and, of course, microbial sugar uptake by the phosphotransferase system (*34*, *35*). Observations of multiple steady states in hybridoma cell cultures reported by Hu et al. (*36*) and Follstad et al. (*37*) provide further examples of multistable metabolic systems that derive their interesting features from nonlinear regulatory phenomena. Thattai and Shraiman (*38*) suggest that a bistable "winner-take-all" phenotype, qualitatively similar to the Matching Law result, is more robust than other switching phenotypes because it manifests over a broader range of parameter space. This relative insensitivity to system parameters allows bistable switches to resist disturbances that may arise from mutation or stochastic variations inside the cell. These observations lend support to the notion that the Matching and Proportional Laws, although originally intended to describe a fairly isolated set of behaviors attributed to diauxie, could in fact reflect a ubiquitous motif of biological regulation.

7. Computational Assessment of Cybernetic Control Laws

7.1. A Few Limiting-Case Policies. Several assumptions were made in arriving at the Matching and Proportional Laws of sections 3 and 4. These assumptions were expressly intended to reflect the efficient and robust behavior observed in biological systems while simultaneously providing simple, closed-form control laws that avoid numerical inconveniences such as discontinuities and philosophical inconsistencies such as noncausality. However, it is important to recognize that these laws represent but one possible route that evolution may have taken to optimize regulatory performance. To establish additional confidence in the Matching and Proportional Laws, it is important to assess their predictions in comparison to alternative control policies and to experimental growth data.

First, we need to introduce some nomenclature to help us differentiate between the various control laws we now intend to investigate. The control policy dictated by eqs 49 and 50 will henceforth be referred to as the MP policy, which stands for Matching and Proportional Laws. Another possible control law is the Indifferent (IN) policy, which results when each enzyme is considered equally important and $\mathbf{p_u^+} = \mathbf{p_v^+} = \mathbf{1}$ is substituted into eqs 49 and 50. The Zero Cost (ZC) policy is obtained by setting $\sigma_u = \sigma_v = 0$ in eq 48, which results in the discontinuous control law:

$$u_i = \begin{cases} 1, & \text{if } p_{u_i} = \max(\mathbf{p_u}) \\ 0, & \text{otherwise} \end{cases} \quad (71)$$

$$v_i = \begin{cases} 1, & \text{if } p_{v_i} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (72)$$

The policy for **u** represents an extreme example of aggressive control, in that all transcriptional and translational resources are allocated exclusively to the enzyme with the highest return-on-investment. However, the policy for **v** is wholly passive in that any enzyme with a positive return-on-investment is fully activated. The difference can be attributed to the less restrictive upper bound constraint (eq 39) applied to the elements of **v** in contrast to the summation constraint (eq 23) that applies to **u**. A related control law can be derived by forcing the elements of **v** to satisfy a summation constraint analogous to eq 23 while once again ignoring the input penalties. The resulting control law for the **u**-variable is the same as in eq 71, but the **v**-variable is described by

$$v_i = \begin{cases} 1, & \text{if } p_{v_i} = \max(\mathbf{p_v}) \\ 0, & \text{otherwise} \end{cases} \quad (73)$$

We refer to this policy as the Bang-Bang (BB) policy because both the **u** and **v** control inputs exhibit discontinuous jumps that typically result in a single dominant enzyme becoming fully induced and activated at the expense of all others. Table 1 summarizes the expressions for **u** and **v** control variables implied by the MP, IN, ZC, and BB policies.

We must also distinguish between different policies for evaluating $\mathbf{p_u}$ and $\mathbf{p_v}$. The policy embodied by eq 52 will be referred to as the Unweighted Temperate (UT) policy. Including the multiplicative weighting factor e_i/e_i^o in the evaluation of p_{u_i} results in the Weighted Temperate (WT) policy. Setting $\Delta t = 0$ in the UT or WT policy leads to Unweighted or Weighted versions of the Greedy policy (UG or WG policies), respectively. Table 2 summarizes the expressions used to compute the returns-on-investment $\mathbf{p_u}$ and $\mathbf{p_v}$ under each of the aforementioned policies. In all cases, we evaluate $\mathbf{B_u}$ after making the substitution $\mathbf{e} = \mathbf{e}^{ss}$ (cf. eq 66) and $d\mathbf{e}/dt = 0$ in **f**. This ensures continuity of the present cybernetic laws with those introduced by Kompala et al. (*4*). To fully specify the control policy implemented in a particular cybernetic model, we need to declare both the policy used to compute $\mathbf{p_u}$ and $\mathbf{p_v}$ as well as the policy used to evaluate **u** and **v** in terms of the computed returns-on-investment. For instance, the policy used by Kompala et al. would be classified as WG/MP, indicating a Weighted Greedy policy for computing $\mathbf{p_u}$ and $\mathbf{p_v}$ coupled with a Matching and Proportional Law policy for evaluating **u** and **v**. As mentioned in section 6, a Weighted policy is required to produce diauxie with an MP policy; otherwise the MP policy will lead to stable coexistence of enzymes.

7.2. Model of Kompala et al. A classic example of biological control is found in the diauxic growth behavior of *E. coli* and other microbes. Originally discovered by Monod (*39*), diauxie is a growth pattern that often arises when multiple growth-supporting substrates are simultaneously present in the external environment. Instead of consuming the available nutrients in proportion to their abundance, the bacterial cells actively control substrate uptake by regulating their internal repertoire of enzymes. Culturing on a mixture of two substitutable substrates produces two distinct exponential growth phases separated by an intermediate lag phase in which growth is stagnant. The first growth phase involves exclusive consumption of the preferred substrate, which often confers a higher nutritional benefit to the cells as evidenced by a faster growth rate. After the preferred substrate has been completely exhausted, the cells undergo a lag period during which they switch to synthesizing the enzymes needed to metabolize the less preferred substrate. The second growth phase promptly commences once these enzymes have been amassed to sufficient levels. From a cybernetic viewpoint, the diauxie phenomenon is the outward manifestation of a biological control circuit that has evolved a capacity for optimal decision-making, which in this case serves to optimize the growth rate of the culture.

Table 1. Policies for Computing Cybernetic Controls[a]

	MP	IN	ZC	BB
u	$\dfrac{\mathbf{p_u^+}}{\|\mathbf{p_u^+}\|_1}$	$\dfrac{\mathbf{1}}{N_R}$	eq 71	eq 71
v	$\dfrac{\mathbf{p_v^+}}{\|\mathbf{p_v^+}\|_\infty}$	**1**	eq 72	eq 73

[a] Refer to eq 51 for the definitions of $\mathbf{p_u^+}$ and $\mathbf{p_v^+}$. N_R denotes the number of reactions in the network.

Table 2. Policies for Computing Returns-on-Investment[a]

	UT	WT	UG	WG
$\mathbf{p_u}$	$\mathbf{B_u^T e^{A^T \Delta t} q}$	$\mathbf{R^{-1} B_u^T e^{A^T \Delta t} q}$	$\mathbf{B_u^T q}$	$\mathbf{R^{-1} B_u^T q}$
$\mathbf{p_v}$	$\mathbf{B_v^T e^{A^T \Delta t} q}$	$\mathbf{B_v^T e^{A^T \Delta t} q}$	$\mathbf{B_v^T q}$	$\mathbf{B_v^T q}$

[a] The weighting matrix **R** has the form given by eq 70.

The ability to model the mixed-substrate growth dynamics of microbial cultures for purposes of designing and optimizing fermentation processes is of interest in the biotechnology industry. Kompala et al. (*4*) employed an objective of growth maximization within a cybernetic model of *Klebsiella oxytoca* to simulate diauxic growth on a number of substrate mixtures involving glucose, xylose, arabinose, lactose and fructose sugars. *K. oxytoca* is an enteric, Gram-negative bacterium that, like *E. coli*, is able to metabolize a wide range of sugar monomers via its phosphoenolpyruvate-dependent phosphotransferase (PTS) system. Unlike *E. coli*, however, *K. oxytoca* can efficiently metabolize a wide variety of oligosaccharides derived from incomplete digestion of lignocellulosic biomass including cellobiose, cellotriose, xylobiose, xylotriose and arabinosides (*40*). This feature makes recombinant strains of *Klebsiella* attractive biocatalysts for the conversion of waste biomass into useful products such as fuel ethanol. The significant outcome of Kompala et al.'s study was the formulation of a cybernetic model that could accurately portray the observed mixed-substrate growth behavior of *K. oxytoca* based solely on experimental results obtained on single substrates. The model correctly predicted the order in which the substrates would be consumed and the overall growth dynamics of the culture using kinetic parameters determined from single-substrate experiments, without the need for additional model parameters to capture the mixed-substrate interactions. This impressive predictive capability flows directly out of the imposed objective of growth maximization as mediated by the Matching and Proportional Laws.

Kompala et al.'s model of growth on substitutable substrates provides a natural context within which to evaluate opposing control policies while avoiding unnecessary complexity. Despite its apparent simplicity, this model is capable of producing a rich assortment of possible dynamic behaviors depending upon the control strategy applied. We focus on the particular case of triauxic growth on glucose, xylose, and lactose because these substrates exhibit a wide diversity of kinetic constants. Figure 2 depicts the topology of this simple biochemical network, and

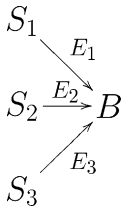

Figure 2. Biochemical network representing growth on three substitutable substrates: glucose (S_1), xylose (S_2), and lactose (S_3).

the relevant model equations are summarized in eqs 53−60 where $\mathbf{s} = [s_1\ s_2\ s_3]^T$, $\mathbf{e} = [e_1\ e_2\ e_3]^T$, and

$$\mathbf{S} = \begin{bmatrix} -\frac{1}{Y_1} & 0 & 0 \\ 0 & -\frac{1}{Y_2} & 0 \\ 0 & 0 & -\frac{1}{Y_3} \end{bmatrix} \tag{74}$$

The parameter values for each substrate are listed in Table 3, which are equivalent to those used by Kompala et al. It is important to note that these values were derived from single substrate growth data and were not adjusted to accommodate mixed substrate growth situations. In the single-substrate case, all valid control policies must collapse to give $u = v = 1$, indicating full activation and induction of the sole pathway for substrate utilization. Therefore, the parameters should not be expected to vary according to the choice of control policy because all policies are equivalent under the conditions for which the parameters were estimated. Consequently, there is no bias introduced by using the same parameter set to evaluate each different control policy.

7.3. Prediction of Growth Phenotype. As a first appraisal, it is instructive to examine simulation results derived from the various control policies side-by-side with comparison to experimental growth data. One feature that quickly stands out is that the Temperate and Greedy return-on-investment policies yield essentially identical results (not shown). This is because the so-called indirect effects of enzyme activation and induction are negligible under the chosen conditions. We use the term "indirect" to describe any consequence of an injected control move that is not realized immediately but instead manifests itself gradually over time. For instance, the direct effect of activating enzyme E_i is to increase the rate of biomass production via reaction R_i. However, this will necessarily lead to faster depletion of substrate S_i, which may result in lower overall biomass production whenever S_i is present at levels near or below the saturation constant K_i. This reduction in growth rate is considered an indirect effect because it is mediated through the dynamics of substrate disappearance. The matrix exponential appearing in eq 52 accounts explicitly for indirect effects in the return-on-investment calculation, at least when they occur on a time scale commensurate with Δt. In the current scenario, the saturation constants for glucose and xylose are sufficiently small and the substrate dynamics sufficiently slow that indirect effects of the control inputs do not play a significant role in determining their associated returns-on-investment. However, this conclusion does not apply to more complex reaction networks, especially those that involve sequences of reactions with short-time-scale intermediates (*29*).

A second notable feature is that the Weighted and Unweighted return-on-investment policies exhibit considerable disagreement. As shown in Figure 3, the WG/MP policy produces almost an exact trace of the experimental growth data obtained by Kompala et al. (*4*), whereas the UG/MP policy is off by about as much as the Indifferent policy. Both the UG/MP and IN policies significantly underestimate the growth rate during the first growth phase because they predict stable coexistence of all three enzyme systems E_1−E_3 rather than dominance of E_1 (cf. Figure 4). The UG/MP policy exhibits a moderate preference for E_1, which is intermediate between that of the IN and WG/MP policies. Although the UG/MP policy allows the less preferred enzymes to persist throughout the first exponential phase, Figure 5 reveals that they are inhibited in the presence of glucose and are rendered essentially useless. As a result, the UG/MP policy predicts sequential uptake of sugars similar to the WG/MP policy and in contrast to the simultaneous uptake pattern predicted by the IN policy. Neither the IN nor UG/MP policy predicts significant diauxic lags because the competition for transcriptional/translational resource is not fierce enough to drive the less preferred enzymes toward extinction. Therefore, these enzymes are ready to come on-line as soon as the preferred substrate is exhausted and their inhibition is lifted. The WG/MP policy, on the other hand, regulates gene expression more aggressively and provides a more accurate representation of diauxie.

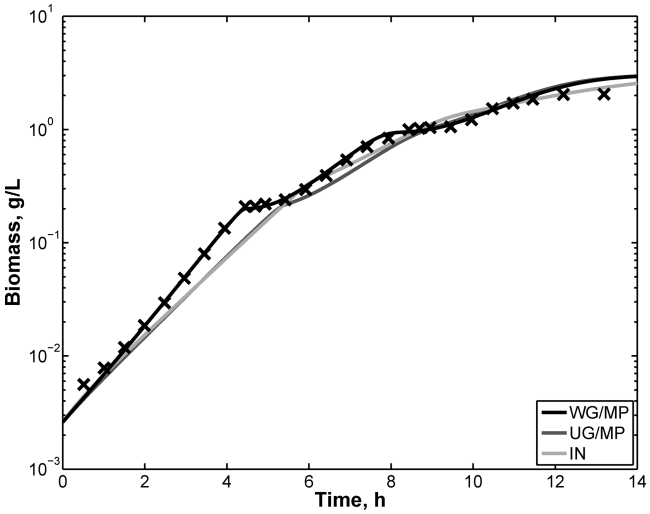

Figure 3. Simulated growth curves for WG/MP, UG/MP, and IN policies. Initial concentrations are $\mathbf{s}(0) = [0.38\ 1.5\ 4.7]^T$ g/L, $\mathbf{e}(0) = [0.90\ 0.17\ 0.20]^T$, and $c(0) = 0.0026$ g/L. Data for *Klebsiella oxytoca* growth (×) on a mixture of glucose, xylose, and lactose are from Kompala et al. (*4*).

The UG/ZC and UG/BB policies both ignore the costs inherent to enzyme induction and activation. As a result, they take whatever action is necessary to maximize the instantaneous growth rate of the culture, subject to the imposed input constraints. Figure 6 shows that these control policies overestimate the observed growth rate during each phase of growth and significantly underpredict the lengths of the diauxic lag periods. Both the UG/ZC and UG/BB policies switch to consuming xylose and lactose earlier than the WG/MP policy. This premature switching is attributable to a lack of autocatalysis in the enzyme synthesis kinetics. The switching dynamics are further illustrated in Figures 7 and 8, which highlight the discontinuous nature of the ZC and BB policies. One noticeable feature is that the UG/BB policy always leaves a small amount of the preferred substrate unconsumed when it switches to the less preferred substrate and must eventually "double back" to consume this remainder. The UG/ZC policy does not display this feature because the accumulated enzymes remain active even after the transcriptional switch is thrown.

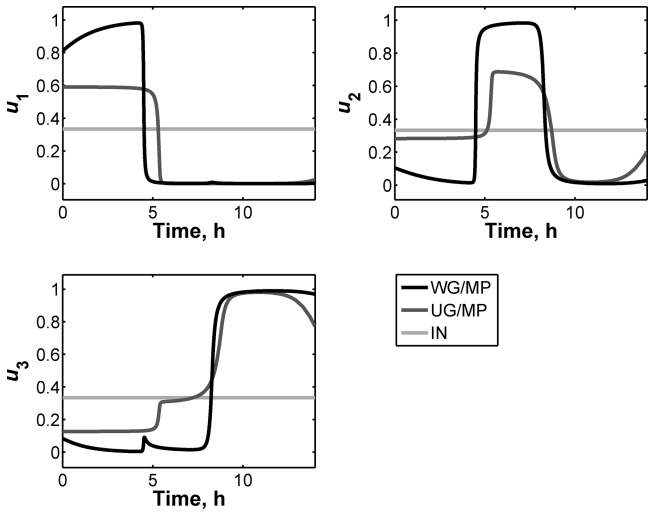

Figure 4. Cybernetic *u*-variables for WG/MP, UG/MP, and IN policies under conditions described in Figure 3.

Table 3. Stoichiometric and Kinetic Parameters Used To Simulate Triauxic Growth.

sugar	i	k_i (h^{-1})	K_i (g/L)	Y_i (gB/gS_i)	α_i (h^{-1})	β_i (h^{-1})
glucose	1	1.08	0.01	0.52	1.13	0.05
xylose	2	0.82	0.2	0.5	0.87	0.05
lactose	3	0.95	4.5	0.45	1.00	0.05

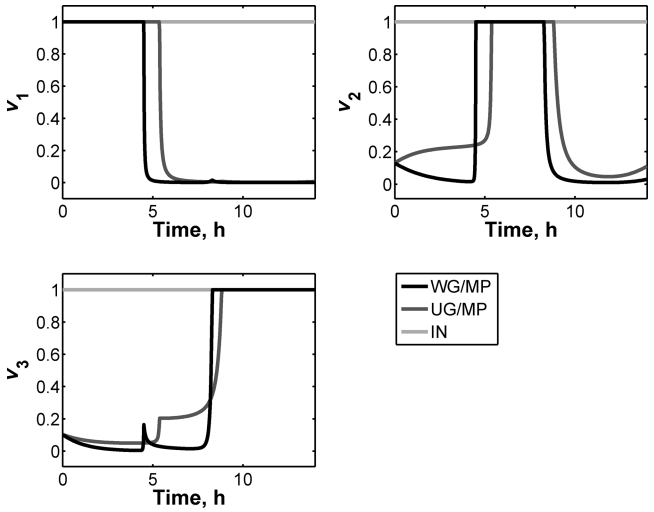

Figure 5. Cybernetic *v*-variables for WG/MP, UG/MP, and IN policies under conditions described in Figure 3.

The salient conclusion that can be drawn from an inspection of Figure 6 is that *Klebsiella* has not evolved a strategy of unbridled growth maximization that ignores the presence of resource limitations. Although higher biomass productivity can be achieved by following the discontinuous UG/ZC or UG/BB strategies, we find that the less aggressive WG/MP policy does a better job of describing the experimental observations. Indeed, it is only logical to assume that the optimal policy should reflect internal resource limitations, since they impose real constraints on the performance of the system. Such reasoning flies in the face of the FBA approach, which assumes that any stoichiometrically feasible flux distribution can be achieved regardless of the physiological state of the cell. To the contrary, it is quite implausible to suggest that reactions can be turned arbitrarily on or off without concern for resource or enzyme availability. This is a crucial distinction that should not be overlooked when comparing the optimality criteria of cybernetic models to those of FBA models and other constraint-based methods.

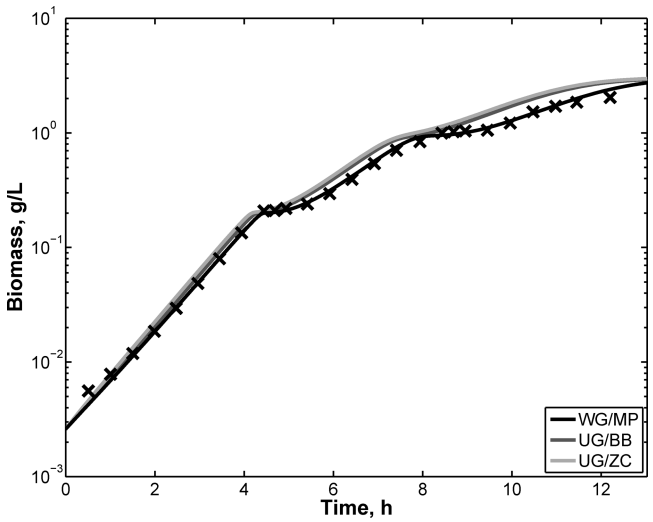

Figure 6. Simulated growth curves for WG/MP, UG/BB, and UG/ZC policies. Initial concentrations are $\mathbf{s}(0) = [0.38\ 1.5\ 4.7]^T$ g/L, $\mathbf{e}(0) = [0.90\ 0.17\ 0.20]^T$, and $c(0) = 0.0026$ g/L. Data for *Klebsiella oxytoca* growth (×) on a mixture of glucose, xylose, and lactose are from Kompala et al. (*4*).

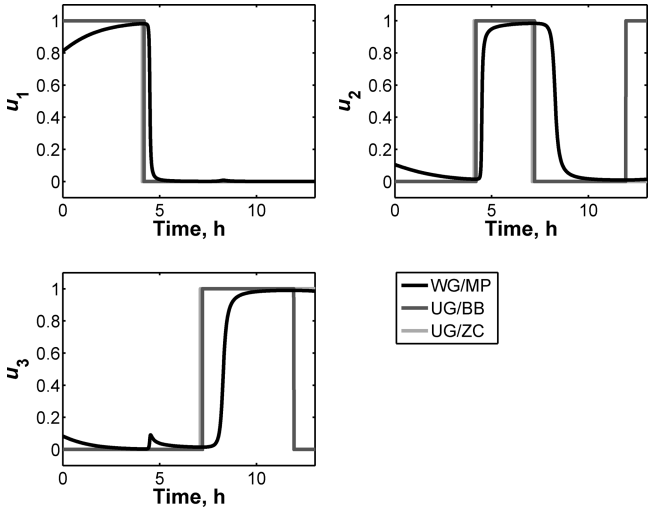

Figure 7. Cybernetic *u*-variables for WG/MP, UG/BB, and UG/ZC policies under conditions described in Figure 6.

The WG/ZC and WG/BB control policies oftentimes produce dynamics similar to the WG/MP policy. In many ways, the WG/MP policy can be considered the continuous analogue of these discontinuous policies. Figure 9 reveals the triauxic growth curves to be nearly identical for all three policies, with the ZC and BB laws giving similar growth rates and lags in comparison to the MP law. However, the all-or-none nature of the BB and ZC policies coupled with a Weighted policy for computing returns-on-investment translates into a pronounced history effect that can lead to erroneous predictions. For instance, growth on a glucose/xylose mixture should evoke the same preference for glucose regardless of how the cells were precultured. The WG/MP policy correctly describes this growth behavior along with the decreased intermediate lag that results from xylose preculturing. The WG/BB and WG/ZC policies, on the other hand, predict that cells precultured on xylose will exhibit a reverse preference for xylose over glucose. Consequently, they predict an incorrect growth curve and an anomalous sequence of substrate utilization as shown in Figure 10.

7.4. Competition Studies. In the previous section, we have been concerned with assessing the isolated performance of a variety of biological control laws. From an evolutionary

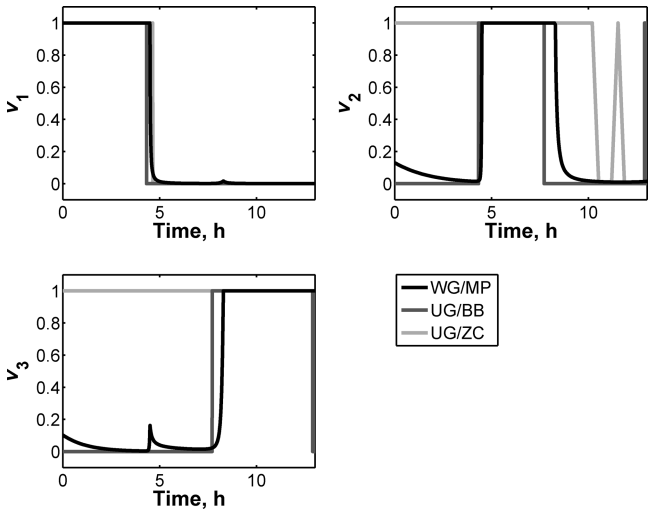

Figure 8. Cybernetic υ-variables for WG/MP, UG/BB, and UG/ZC policies under conditions described in Figure 6.

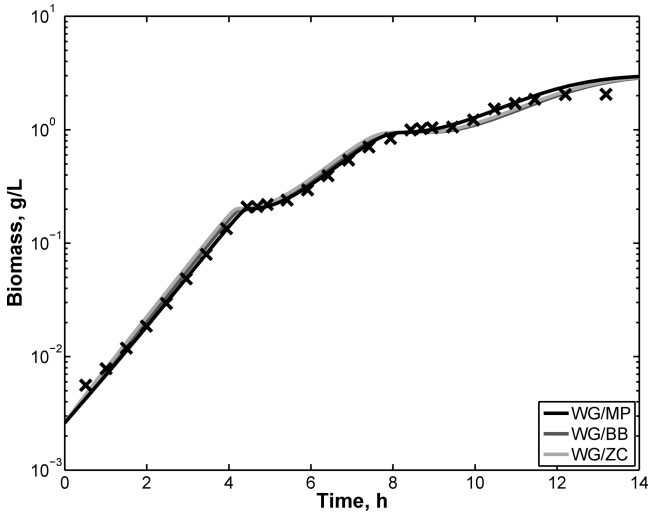

Figure 9. Simulated growth curves for WG/MP, WG/BB, and WG/ZC policies. Initial concentrations are $\mathbf{s}(0) = [0.38\ 1.5\ 4.7]^T$ g/L, $\mathbf{e}(0) = [0.90\ 0.17\ 0.20]^T$, and $c(0) = 0.0026$ g/L. Data for *Klebsiella oxytoca* growth (×) on a mixture of glucose, xylose, and lactose are from Kompala et al. (*4*).

standpoint, a more interesting line of questioning surrounds the outcome of direct competition between microbial species that implement opposing regulatory strategies. We now wish to apply the mixed substrate growth models of section 7.3 to describe competition among communities of differentially regulated species rather than pure cultures in isolation. The model equations are identical, except that the substrate balance equation must account for the specific uptake of each bacterial species present in the culture. As a first example, let us investigate the competition between the various control policies when the culture is provided a mixture of glucose, xylose and lactose. Using the same initial enzyme concentrations as in Figure 3 while setting the initial substrate levels to $s_i(0) = 10$ g/L $\forall i$ and the initial biomass levels to $c_i(0) = 10^{-3}$ g/L $\forall i$ gives the growth curves of Figure 11. The UG/ZC policy results in the highest species abundance throughout the course of the batch, with the UG/BB, WG/ZC, WG/BB, and WG/MP species not far behind. The less aggressive IN and UG/MP policies give slower growth and are ultimately dominated by the other policies. The increased growth rates provided by the best performing policies do not come without a price, however. By examining the cumulative input penalties associated with the weighting matrix of eq 70, we find that the more aggressive policies accrue markedly higher enzyme synthesis penalties over the course of the batch (cf. Figure 12). Overall, the analysis suggests exactly what our intuition could have already told us: that increased growth rate comes at the expense of placing an increased strain on internal resources.

However, the story does not end there. The foregoing analysis presupposes a single set of initial conditions that determine the substrate, enzyme, and species levels at the start of the batch. In their native habitats, however, microbes face continually changing environmental conditions, and their regulatory programs must be sufficiently robust and flexible to ensure survival across the entire ecological landscape. For instance, the intestinal flora of humans contains a diverse community of microbes whose dietary regimens are set by the feeding preferences of their hosts. This means that their metabolic machinery must retool every few hours in response to a new influx of food. To test the robustness of each policy defined in section 7.1, Monte Carlo simulations were performed by repeatedly integrating the batch growth equations for many different realizations of $\mathbf{s}(0)$ and $\mathbf{e}(0)$. The initial substrate combination for each trial was

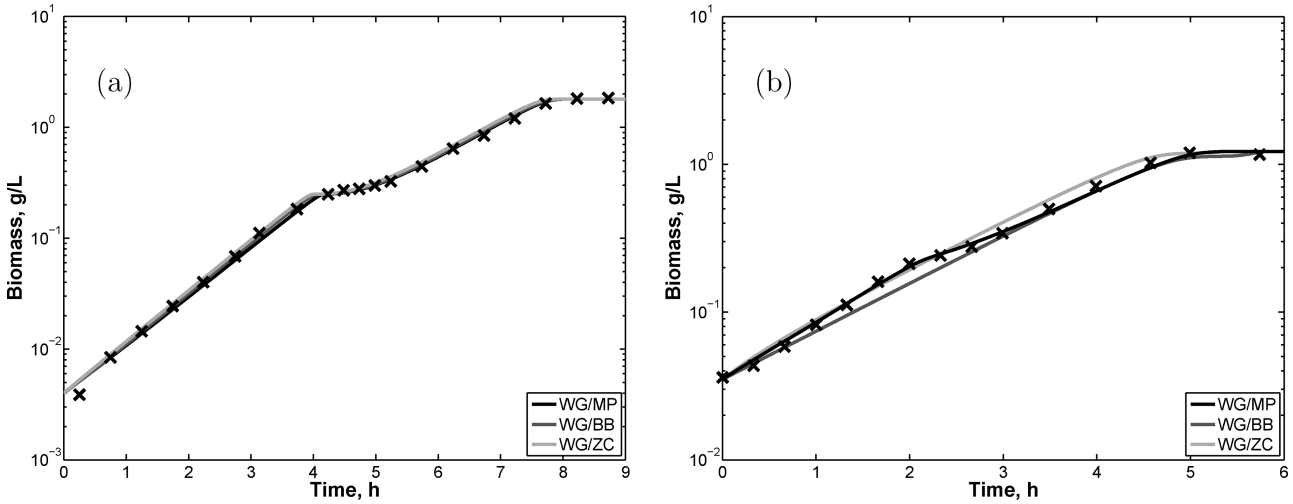

Figure 10. Simulated growth curves for WG/MP, WG/BB, and WG/ZC policies for glucose/xylose mixture with (a) glucose or (b) xylose preculturing. Initial conditions following glucose preculturing are $\mathbf{s}(0) = [0.47\ 3.1]^T$ g/L, $\mathbf{e}(0) = [0.90\ 0.18]^T$, and $c(0) = 0.004$ g/L. Initial conditions following xylose preculturing are $\mathbf{s}(0) = [0.17\ 2.2]^T$ g/L, $\mathbf{e}(0) = [0.38\ 1.0]^T$, and $c(0) = 0.035$ g/L. Data for *Klebsiella oxytoca* growth (×) are from Kompala et al. (*4*).

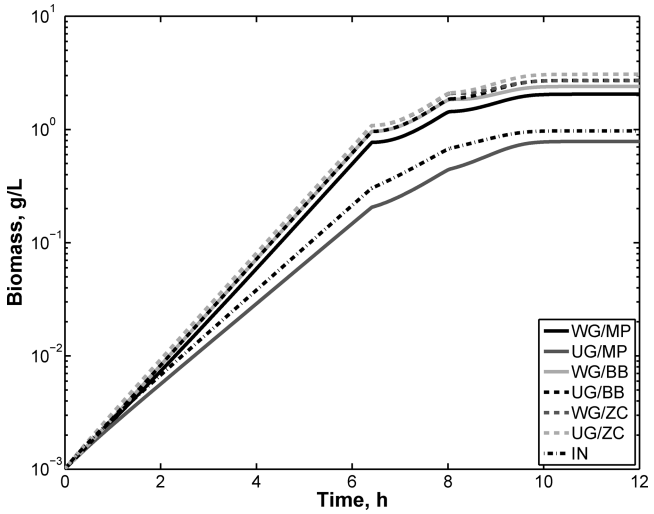

Figure 11. Competitive growth of species obeying the WG/MP, UG/MP, WG/BB, UG/BB, WG/ZC, UG/ZC, and IN policies on a glucose/xylose/lactose mixture.

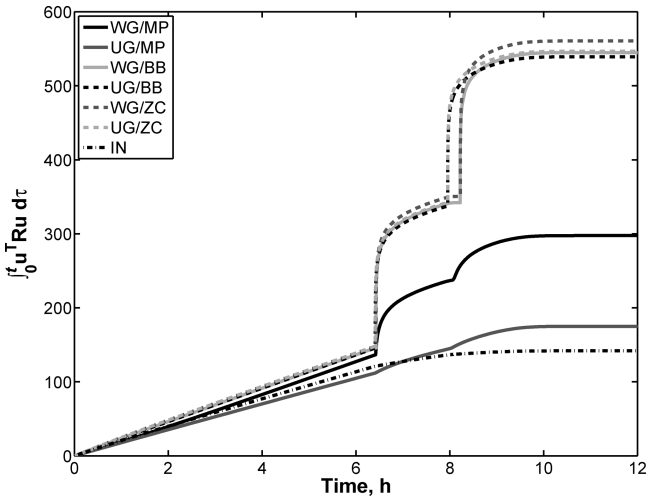

Figure 12. Cumulative enzyme synthesis penalties for the WG/MP, UG/MP, WG/BB, UG/BB, WG/ZC, UG/ZC, and IN policies during competitive growth on glucose, xylose, and lactose.

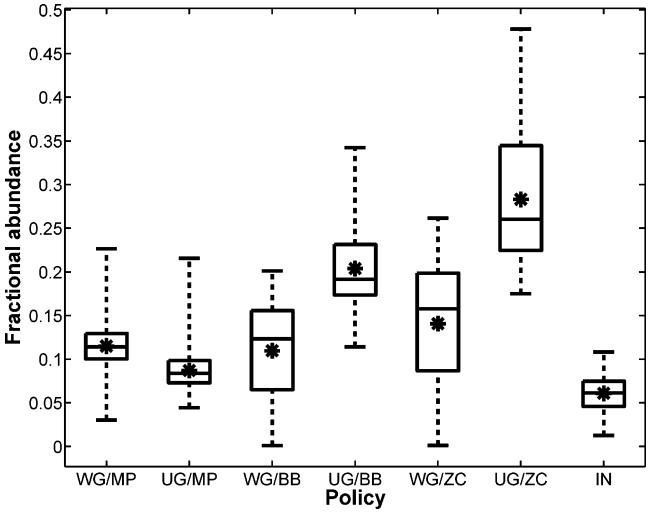

Figure 13. End-of-batch fractional abundance of each species averaged over 1000 trials of competitive growth on glucose, xylose, and lactose. The boxes have lines at the lower quartile, median, and upper quartile values. The whiskers show the extent of the rest of the data. The mean value for each policy is represented by an asterisk.

chosen by randomly sampling from a uniform distribution within the interval 0−10 g/L. The initial enzyme levels were similarly chosen from the uniform distribution defined over [0, 1]. To be clear, each element of the **s**(0) and **e**(0) vectors was set to a different realization from this sampling process, so that each element was uniquely determined, but the different species were assigned the same values of **e**(0) and $c(0)$.

Figure 13 summarizes the results from 1000 trials of competitive growth in a box and whisker plot. The end-of-batch fractional abundance of each species is computed by dividing its final biomass concentration by the total amount of biomass present. As Figure 13 shows, the UG/ZC policy gives the best average-case performance followed by the UG/BB policy. The WG/MP policy outperforms the WG/BB, UG/MP, and IN policies on average but does not fare so well against the more aggressive UG/ZC, UG/BB, and WG/ZC policies. It is also interesting to note that the WG/MP, UG/MP, and IN policies are less sensitive to the choice of initial conditions as revealed by their small interquartile ranges. On the other hand, the WG/BB and WG/ZC policies are highly sensitive to the initial conditions and exhibit very poor worst-case performance.

Apparently, the Monte Carlo results depict a tradeoff between fitness and cost. The more aggressive policies provide faster growth, while the less aggressive policies limit the penalties associated with enzyme synthesis and activation. How can we go about determining which policies are most efficient? This question is difficult to answer a priori because it requires that we establish suitable metrics for the availability of transcriptional and translational resources inside the cell and also for the cell's ability to simultaneously activate multiple competing pathways. From an optimality standpoint, such factors influence the weightings that appear in the system performance index and determine how severely regulatory control actions should be penalized. Of course, we can always compare simulation results produced by the various control policies with experimental growth data to discover which policies best reflect the observations, as we have already discussed in the previous section. However, are there any conclusions that can be drawn about the relative efficiencies of these control laws apart from their ability to describe experimental results? One possible answer is obtained by plotting some measure of system fitness versus some measure of the investment needed to achieve that fitness, as shown in Figure 14. Here we have used the fractional species abundance as the fitness measure and a normalized investment penalty χ defined for the ith species by

$$\chi_i = \frac{1}{2}\left[\frac{\zeta_i}{\|\boldsymbol{\zeta}\|_\infty} + \frac{\eta_i}{\|\boldsymbol{\eta}\|_\infty}\right] \text{ where} \tag{75}$$

$$\zeta_i = \int_0^T \mathbf{u_i}^T \mathbf{R_i} \mathbf{u_i}\, d\tau \tag{76}$$

$$\eta_i = \int_0^T \mathbf{v_i}^T \mathbf{v_i}\, d\tau \tag{77}$$

as the measure of input cost. The batch time T is defined as the time when all substrates are depleted to within a tolerance of

$$\max_i \left(\frac{s_i}{K_i + s_i}\right) < 10^{-3} \tag{78}$$

The form of χ was chosen because it places both enzyme synthesis and activation penalties on equal footing.

Figure 14 reveals that not all policies are equally efficient. Those policies that lie to the upper-left of the plot return the highest species abundance for a given amount of resource

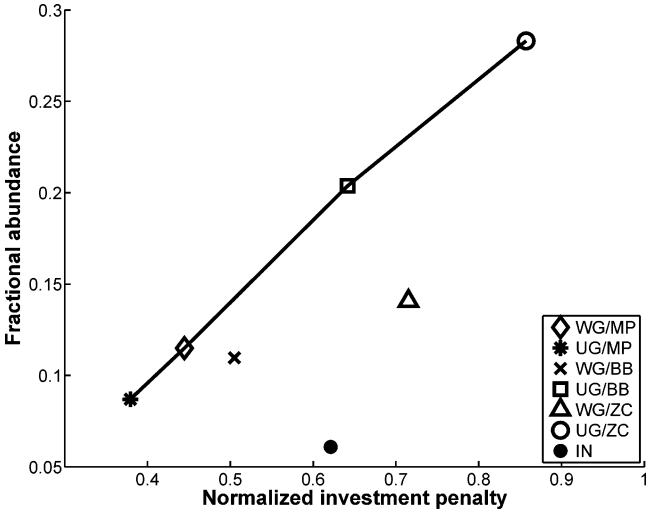

Figure 14. Tradeoff between fitness and cost for the WG/MP (◇), UG/MP (∗), WG/BB (×), UG/BB (□), WG/ZC (Δ), UG/ZC (○), and IN (●) policies. Results are averaged over 1000 trials of competitive growth on glucose, xylose and lactose.

investment. The solid curve connecting the UG/MP, WG/MP, UG/BB, and UG/ZC data points represents the "efficient frontier". This is the locus of points for which no other policy provides better fitness while simultaneously achieving a lower value of χ. Although we cannot know a priori where the system lies along this tradeoff curve without knowing the "true" system performance index, it stands to reason that its evolution should have converged to some point along the efficient frontier. Otherwise, it would be possible to achieve improved fitness for an equal or lesser cost. This analysis implies that the IN, WG/BB, and WG/ZC policies are inefficient because they lie to the lower-right of the efficient frontier. However, the remaining policies are equally efficient in the sense that no other policy achieves a better fitness with lower resource investment.

7.5. Further Empirical and Practical Considerations. Kompala et al.'s original model predicts that batch growth on mixed substrates will always result in diauxie. It should be noted, however, that many instances of batch growth do not display sequential uptake of substrates. The observed behavior generally falls into one of three categories: (1) sequential growth pattern independent of preculturing, (2) simultaneous growth pattern independent of preculturing, or (3) growth pattern dependent on preculturing (*41*). Sequential growth may be either diauxic or biphasic, with the latter term denoting absence of an intermediate lag period. Kompala et al. (*4*) has shown how biphasic growth can be easily accommodated within the cybernetic framework by assuming that the two competing pathways are catalyzed by the same enzyme system. Later work by Straight and Ramkrishna (*42*) and Ramakrishna et al. (*43*) has shown that it is possible to describe all of the aforementioned growth patterns ranging from simultaneous to sequential uptake by augmenting the reaction network or by altering the cybernetic objective function ϕ. In the modular approach to cybernetic modeling (*44*), one also has discretion in choosing which enzymes should be placed in competition depending on the level of interaction observed between parallel metabolic branches. Therefore, these laws can indeed form the basis for a general treatment of regulation that is adaptable to many different metabolic scenarios.

In addition to the features already discussed, there are other practical reasons for invoking the Matching and Proportional Laws. One key advantage is that they are described by relatively simple and well behaved functions. A discontinuous control policy such as the ZC or BB policy, on the other hand, creates many analytical and numerical problems. In situations where the rates of two competing reactions are nearly equal, e.g., in a chemostat under steady-state conditions, frequent discontinuous switching leads to severe slowing and eventual failure of the numerical integrator. This problem does not occur with the Matching and Proportional Laws, which readily describe the smooth transition from preferential uptake at high dilution rates to simultaneous uptake at low dilution rates that is observed in chemostat cultures (*45*). Furthermore, the application of techniques such as bifurcation analysis or sensitivity analysis to discontinuous models is complicated by the presence of non-differentiable terms in the closed-loop system equations. All of these reasons make a continuous control law such as the MP policy preferable to a discontinuous law in practice.

8. Conclusion

The Matching and Proportional Laws are the central tenets of cybernetic modeling. They have been successfully applied to numerous modeling studies, ranging from lumped models of diauxic growth to highly structured models of biochemical networks. Until now, the optimality background of these laws was partially shrouded due to lack of a rigorous mathematical development leading to their derivation. The main contribution of this article, therefore, is to establish how generalized versions of the Matching and Proportional Laws are obtained as the solution to a well-defined optimal control problem. Moreover, the assumptions needed to reconcile these results with previous treatments in the cybernetic modeling literature are fully enumerated and discussed. These results are not only interesting in their own right but have led to the outgrowth of a novel cybernetic modeling approach that utilizes elementary mode decomposition of biochemical networks to infer regulatory interactions and resource competitions within the resulting models (*29*). To maintain simplicity in the present exposition and continuity with previous work, we have chosen to illustrate the expanded Matching and Proportional Law treatment using a lumped model of diauxic growth that does not require network decomposition methods of this sort. A detailed description of how the current results can be extended to more complex networks is left to future publications. Suffice it to say, however, that an optimal control problem similar to the one discussed in section 5 is solved for each elementary flux mode to determine local cybernetic control variables that maximize the composite mode flux. These local controls are combined with global cybernetic variables that determine the extent to which each mode should be activated in order to best achieve the organism's nutritional objectives.

Computational results were presented that illustrate the relative merits of several different control laws in describing the growth dynamics of mixed-substrate bacterial cultures. The WG/MP policy provides superior predictive capabilities in comparison to the UG/MP and IN policies, which are too passive, and the UG/BB and UG/ZC policies, which are too aggressive. The WG/BB and WG/ZC policies give results similar to the those of WG/MP policy in some cases but may predict the wrong order of substrate preference in others. Monte Carlo simulations designed to probe the response of each control law under varying ecological conditions were also presented. These results shed light on the robustness and efficiency properties exhibited by each law when placed in direct competition with multiple opposing policies. Although some laws are clearly less efficient than others, a handful of laws including the WG/MP policy lie along an efficient frontier and are

equivalent from the standpoint that no other policy achieves a better fitness with lower resource investment. However, empirical results suggest that the diauxic switching behavior of *Klebsiella oxytoca* has not evolved toward a policy of unbridled growth maximization but instead exhibits a tendency to weigh the potential benefits of alternative control actions against the costs required to implement those actions. The Matching and Proportional Laws mimic this behavior by recognizing that optimality is related to the economic utilization of scarce internal resources, which cannot be liberally allocated without incurring diminishing marginal returns.

Acknowledgment

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship (J.D.Y.) and a National Science Foundation GOALI grant (Award no. BES-0000961).

References and Notes

(1) Mayr, E. *Toward a New Philosophy of Biology: Observations of an Evolutionist*; Harvard University Press: Cambridge, 1988.
(2) Ramkrishna, D. A Cybernetic Perspective of Microbial Growth. In *Foundations of Biochemical Engineering: Kinetics and Thermodynamics in Biological Systems*; Papoutsakis, E., Stephanopoulos, G. N., Blanch, H. W., Eds.; American Chemical Society: Washington, DC, 1982.
(3) Kompala, D. S.; Ramkrishna, D.; Tsao, G. T. Cybernetic Modeling of Microbial Growth on Multiple Substrates. *Biotechnol. Bioeng.* **1984**, *26*, 1272−1281.
(4) Kompala, D. S.; Ramkrishna, D.; Jansen, N. B.; Tsao, G. T. Investigation of Bacterial Growth on Mixed Substrates: Experimental Evaluation of Cybernetic Models. *Biotechnol. Bioeng.* **1986**, *28*, 1044−1055.
(5) Reed, J. L.; Palsson, B. O. Thirteen Years of Building Constraint-Based In Silico Models of *Escherichia coli*. *J. Bacteriol.* **2003**, *185*, 2692−2699.
(6) Segrè, 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.
(7) Shlomi, T.; Berkman, O.; Ruppin, E. Regulatory On/Off Minimization of Metabolic Flux Changes after Genetic Perturbations. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7695−7700.
(8) Fell, D. *Understanding the Control of Metabolism*; Portland Press: London, 1997.
(9) Stephanopoulos, G. N.; Aristidou, A. A.; Nielsen, J. *Metabolic Engineering: Principles and Methodologies*; Academic Press: San Diego, 1998.
(10) Domach, M. M.; Leung, S. K.; Cahn, R. E.; Cocks, G. G.; Shuler, M. L. Computer Model for Glucose-Limited Growth of a Single Cell of *Eshcerichia coli* B/r-A. *Biotechnol. Bioeng.* **1984**, *26*, 203−216.
(11) Chassagnole, C.; Noisommit-Rizzi, N.; Schmid, J. W.; Mauch, K.; Reuss, M. Dynamic Modeling of the Central Carbon Metabolism of *Escherichia coli*. *Biotechnol. Bioeng.* **2002**, *79*, 53−73.
(12) Rizzi, M.; Baltes, M.; Theobald, U.; Reuss, M. In Vivo Analysis of Metabolite Dynamics in *Saccharomyces cerevisiae*. II. Mathematical Model. *Biotechnol. Bioeng.* **1997**, *55*, 592−608.
(13) Savageau, M. A. *Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology*; Addison-Wesley: Reading, MA, 1976.
(14) Voit, E. O. *Canonical Nonlinear Modeling: S-System Approach to Understanding Complexity*; Van Nostrand Reinhold: New York, 1991.
(15) Varner, J. D. Large-Scale Prediction of Phenotype: Concept. *Biotechnol. Bioeng.* **2000**, *69*, 664−678.
(16) No further attempt is made to elucidate the properties of the key resource, as it is not relevant to the current discussion. However, it has been variously identified with the pool of available RNA polymerase proteins (*3*), the ribosome content of the cell (*22*), the pool of free amino acids (*18*), and the available ATP supply (*3*). It should be noted that ribosome content exhibits the most growth-rate dependence of any cellular property (*46*), which argues strongly for its role as a limiting resource in enzyme synthesis.
(17) Mayne, D. Q.; Michalska, H. Receding Horizon Control of Nonlinear Systems. *IEEE Trans. Automat. Contr.* **1990**, *35*, 814−824.
(18) Dhurjati, P.; Ramkrishna, D.; Flickinger, M.; Tsao, G. A Cybernetic View of Microbial Growth: Modeling of Cells as Optimal Strategists. *Biotechnol. Bioeng.* **1985**, *27*, 1−9.
(19) Lewis, F. L.; Syrmos, V. L. *Optimal Control*, 2nd ed.; Wiley: New York, 1995.
(20) Żak, S. H. *Systems and Control*; Oxford University Press: New York, 2003.
(21) Rardin, R. L. *Optimization in Operations Research*; Prentice Hall: Upper Saddle River, NJ, 1998.
(22) Baloo, S.; Ramkrishna, D. Metabolic Regulation in Bacterial Continuous Cultures: I. *Biotechnol. Bioeng.* **1991**, *38*, 1337−1352.
(23) Straight, J. V.; Ramkrishna, D. Modeling of Bacterial Growth under Multiply-Limiting Conditions. Experiments under Carbon- or/and Nitrogen-Limiting Conditions. *Biotechnol. Prog.* **1994**, *10*, 588−605.
(24) Namjoshi, A. A.; Hu, W.-S.; Ramkrishna, D. Unveiling Steady-State Multiplicity in Hybridoma Cultures: The Cybernetic Approach. *Biotechnol. Bioeng.* **2003**, *81*, 80−91.
(25) Kremling, A.; Fischer, S.; Gadkar, K.; Doyle, F. J.; Sauter, T.; Bullinger, E.; Allgöwer, F.; Gilles, E. D. A Benchmark for Methods in Reverse Engineering and Model Discrimination: Problem Formulation and Solutions. *Genome Res.* **2004**, *14*, 1773−1785.
(26) Stewart, W. E.; Shon, Y.; Box, G. E. P. Discrimination and Goodness of Fit of Multiresponse Mechanistic Models. *AIChE J.* **1998**, *44*, 1404−1412.
(27) Burgard, A. P.; Maranas, C. D. Optimization-Based Framework for Inferring and Testing Hypothesized Metabolic Objective Functions. *Biotechnol. Bioeng.* **2003**, *82*, 670−677.
(28) Ström, T. On Logarithmic Norms. *SIAM J. Numer. Anal.* **1975**, *12*, 741−753.
(29) Young, J. D. A System-Level Mathematical Description of Metabolic Regulation Combining Aspects of Elementary Mode Analysis with Cybernetic Control Laws, Thesis, Purdue University, 2005.
(30) From this point forward, the notation $\max_n$ replaces $\max_{n\in\{1,\ \ldots,\ N_R\}}$ so as to maintain compact notation.
(31) Narang, A.; Konopka, A.; Ramkrishna, D. Dynamic Analysis of the Cybernetic Model for Diauxic Growth. *Chem. Eng. Sci.* **1997**, *52*, 2567−2578.
(32) Narang, A.; Konopka, A.; Ramkrishna, D. The Dynamics of Microbial Growth on Mixtures of Substrates in Batch Reactors. *J. Theor. Biol.* **1997**, *184*, 301−317.
(33) Namjoshi, A. A.; Ramkrishna, D. Multiplicity and Stability of Steady States in Continuous Bioreactors: Dissection of Cybernetic Models. *Chem. Eng. Sci.* **2001**, *56*, 5593−5607.
(34) Ozbudak, E. M.; Thattai, M.; Lim, H. N.; Shraiman, B. I.; van Oudenaarden, A. Multistability in the lactose utilization network of *Escherichia coli*. *Nature* **2004**, *427*, 737−740.
(35) Ferrell, J. E., Jr. Self-Perpetuating States in Signal Transduction: Positive Feedback, Double- Negative Feedback and Bistability. *Curr. Opin. Cell Biol.* **2002**, *14*, 140−148.
(36) Hu, W.-S.; Zhou, W.; Europa, L. F. Controlling Mammalian Cell Metabolism in Bioreactors. *J. Microbiol. Biotechnol.* **1998**, *8*, 8−13.
(37) Follstad, B. D.; Balcarcel, R. R.; Stephanopoulos, G.; Wang, D. I. C. Metabolic Flux Analysis of Hybridoma Continuous Culture Steady State Multiplicity. *Biotechnol. Bioeng.* **1999**, *63*, 675−683.
(38) Thattai, M.; Shraiman, B. I. Metabolic Switching in the Sugar Phophotransferase System of *Escherichia coli*. *Biophys. J.* **2003**, *85*, 744−754.
(39) Monod, J. Recherches sur la croissance des cultures bactériennes (Studies on the Growth of Bacterial Cultures). Hermann and Cie: Paris, 1942.
(40) Wood, B. E.; Yomano, L. P.; York, S. W.; Ingram, L. O. Development of Industrial Medium Required Elimination of the 2,3-Butanediol Fermentation Pathway To Maintain Ethanol Yield in an

Ethanologenic Strain of *Klebsiella oxytoca*. *Biotechnol. Prog.* **2005**, *21*, 1366−1372.

(41) Narang, A.; Konopka, A.; Ramkrishna, D. New Patterns of Mixed-Substrate Utilization During Batch Growth of *Escherichia coli* K12. *Biotechnol. Bioeng.* **1997**, *55*, 747−757.

(42) Straight, J. V.; Ramkrishna, D. Cybernetic Modeling and Regulation of Metabolic Pathways. Growth on Complementary Nutrients. *Biotechnol. Prog.* **1994**, *10*, 574−587.

(43) Ramakrishna, R.; Ramkrishna, D.; Konopka, A. E. Cybernetic Modeling of Growth in Mixed, Substitutable Substrate Environments: Preferential and Simultaneous Utilization. *Biotechnol. Bioeng.* **1996**, *52*, 141−151.

(44) Varner, J.; Ramkrishna, D. Metabolic Engineering from a Cybernetic Perspective. 1. Theoretical Preliminaries. *Biotechnol. Prog.* **1999**, *15*, 407−425.

(45) Ramkrishna, D.; Kompala, D. S.; Tsao, G. T. Are Microbes Optimal Strategists? *Biotechnol. Prog.* **1987**, *3*, 121−126.

(46) Neidhardt, F. C.; Ingraham, J. L.; Schaechter, M. *Physiology of the Bacterial Cell: A Molecular Approach*; Sinauer Associates: Sunderland, MA, 1990.

Received June 9, 2006. Accepted October 17, 2006.

BP060176Q