Metabolic Engineering from a Cybernetic Perspective. 1. Theoretical Preliminaries

J. Varner[†] and D. Ramkrishna*

School of Chemical Engineering, Purdue University, West Lafayette, Indiana 47907

The theoretical basis of a cybernetic metabolic network design and analysis framework, which has been subsequently successfully applied to predict system response to genetic alteration, is presented. This conceptual methodology consists of three main branches, namely, a model realization framework, a representation of genetic alteration, and lastly, a metabolic design component. These concepts are introduced as a series of postulates that describe the basic tenets of the approach. Each branch is discussed in turn, starting with the cybernetic representation of arbitrarily complex metabolic networks. A set of postulates is put forth that affords the modular construction of cybernetic models of metabolic networks using as a base a library of elementary pathways. This is followed by a discussion of the representation of genetic alterations within the cybernetic framework. It is postulated that the objective of the base network and the altered system are identical (at least on the time scale required for the organism to "learn" new objectives). This implies, with respect to resource allocation, that the base network and its genetically altered counterpart may still be treated as optimal systems; however, the set of competing physiological choices open to the altered network expands or contracts depending upon the nature of the genetic perturbation. Lastly, to add a predictive design aspect to the methodology, we present a set of postulates that outline the application of metabolic control analysis to cybernetic model systems. We postulate that sensitivity coefficients computed from a cybernetic model, although still local in scope, have the added benefit of a systematic representation of regulatory function as described by the cybernetic variables. Thus, information gained from sensitivity measurements stemming from a cybernetic model include the explicit input of metabolic regulation, a component that is lacking in a purely kinetic representation of metabolic function. The sensitivity results can then be employed to develop qualitative strategies for the rational alteration of metabolic function, which can be evaluated by simulation of an appropriately modified cybernetic model of the base network.

1. Introduction

Metabolic networks have evolved, through millions of years of evolutionary pressure, the ability to respond to changes in the environment of the microorganism, ensuring survival. Through a series of seemingly random adaptations to perturbation and competition, elaborate network control systems have evolved to control microbial metabolism. Thus, it is not surprising to learn that often, when perturbed genetically, in vivo metabolic networks are seen to react or respond to perturbation. This response may take many forms, ranging from the simple redirection of carbon flux toward secondary pathways to complete system failure. It is our hypothesis that the very same control systems that have allowed the microorganism to develop tremendous flexibility and adaptive potential hamper our attempts at the redirection of metabolic flux. Thus, at issue is the ability to compensate or design for this contingency.

Certainly no one would argue that the rational design of metabolic pathways has progressed to the point of being a functional engineering discipline (Bailey, 1991). The biological tools are in place for relatively easy genetic manipulation of metabolic network structure. This step in the design process is no longer the limiting factor. Rather, the design guidance that rigorous mathematical tools provide is sorely lacking. This missing component, potentially, constitutes the biggest roadblock to metabolic design. Mathematical tools such metabolic control analysis (MCA) developed by Kascer and Burns (1973) and biochemical systems theory (BST) developed by Savageau (1976) attempt to describe the regulation of metabolic networks through the use of sensitivity coefficients. The results of attempts to a priori rationally design metabolic networks using these tools have met with limited success with some notable exceptions (Hatzimanikatis et al., 1996a,b, 1998). There are a number of reasons why this might be so. Metabolic control analysis is not a modeling framework; rather, it is a set of postulates that allows the systematic computation of network sensitivities to single perturbations in the environmental or network parameters. Thus, its predictive capability is limited by the quality and source of the information that is input

[†] Present address: Institute of Biotechnology, ETH-Honggerberg, Zurich, Switzerland CH-8093. E-mail: varner@biotech.biol.ethz.ch.

* To whom correspondence should be addressed. E-mail: ramkrish@atom.ecn.purdue.edu.

10.1021/bp990017p CCC: $18.00
Published on Web 04/03/1999

into the analysis structure. Surely, an explicit dependence upon experimental input negates any type of a priori predictive potential because the results of the experiment are required for the analysis. If the analysis input stems from a kinetic model of the network, the claim that the sensitivity coefficients reflect the input of metabolic control mechanisms is very strongly suspect because traditional kinetic models lack the description of the regulatory component. In this instance, the sensitivity coefficients do not describe regulatory effects; rather, they reflect the influence upon system behavior of kinetic competition within the network. Biochemical systems theory does contain a modeling aspect, and as such, theoretically it possesses the potential of being predictive in scope. However, it also depends on sensitivity arguments to develop a conception of network regulation. The explicit dependence upon sensitivity coefficients, which in general are valid only locally, limits the predictive scope of the framework.

The application of the cybernetic framework, developed by Ramkrishna and co-workers, for the description and analysis of metabolic networks is a new approach and may offer a distinct advantage over traditional methodologies. The framework hypothesizes that metabolic systems have evolved optimal goal oriented strategies as a result of evolutionary pressures. The optimality hypothesis implies that the system directs the synthesis and activity of network enzymes such that a nutritional objective is achieved in an optimal manner. The inclusion of a goal-oriented regulatory strategy gives the cybernetic description of a metabolic network the key feature of regulatory responsiveness, an element that is missing from all other contemporary metabolic network analysis and modeling frameworks. In other words, the feature that a cybernetic realization of a metabolic network offers, that is absent from all other contemporary approaches, is the ability to predict the manner in which a metabolic network redirects enzymatic expression and activity in the face of environmental perturbations or genetic alterations to network structure or function. Thus, a cybernetic model of a metabolic network has the potential of predicting how network functionality adjusts upon perturbation. The advantage gained from such knowledge is obvious. If we knew, a priori, how a network would respond to perturbation, such an understanding could be incorporated into the design of the genetic alteration, so as to compenate or even take advantage of the network response.

To formulate such a tool, a number of theoretical obstacles must be overcome. First, methodology within the cybernetic framework must be developed for the systematic regulatory realization of arbitrarily complex metabolic networks. Currently, such an understanding does not exist and cybernetic developments, to this point, have been limited to abstracted models of microbial processes. Once this task is accomplished, the tenets of the cybernetic framework must be reexamined in light of the overexpression or deletion of existing pathway enzymes. Certainly, the concept of applying a framework whose conceptual basis is the assumption of optimality to a recombinant system may cause trepidation. Therefore, a thorough discussion of the exact nature of the optimality assumption inherent to cybernetic developments and its relationship to overexpression and deletion of existing pathway enzymes must be undertaken. In short, we take the position that cellular resources are allocated upon the basis of cybernetic goals in place because of the force of evolutionary pressure. Moreover, we postulate that the goal-oriented operation in effect in the base organism is retained, even in the altered network. Thus, we maintain the objectives of the base and altered networks are identical. This perspective follows from the hypothesis that the ramifications of millions of years of selective pressure simply cannot be immediately erased with a small number of genetic alterations. Rather, we hypothesize that the existing metabolic control structure in place for the base network responds to genetic perturbation by trying to maintain the ingrained base objectives as best possible, in the face of the genetic alteration. In terms of resource allocation, the altered network is then treated as optimal, albeit to an expanded or contracted set of physiological alternatives reflecting the genetic change. The satisfaction of these requirements yields a modeling framework that has the potential of describing the time evolution of a base or genetically modified metabolic network. However, as proposed, the framework is limited in its design capabilities. Currently, the selection of sites for possible genetic alteration (which is clearly an integral aspect of the design process) is based solely upon intuition. Perhaps, for simple metabolic networks, this level of analysis coupled with simulation is adequate. In the general case, however, this is clearly not true. Accordingly, we propose to *qualitatively* guide the formation of metabolic engineering strategies, which can then be examined via simulation of the cybernetic model to determine their relative value. The fusion of a cybernetic realization of a metabolic network with MCA allows the sensitivity coefficients computed from the model to truly represent the input of metabolic control and, moreover, lends a design capability to the framework.

Scope of Investigation. The objective of this work is the presentation of the theoretical basis of the extension of the cybernetic framework to include metabolic engineering applications. We begin with the introduction of the tools necessary for the development of cybernetic control systems for arbitrarily complex metabolic networks. The approach presented, termed the modular approach, relies upon the postulate of hierarchical control of metabolic pathways. After the introduction of the pathway modeling aspect of the development, we focus upon the extension of the cybernetic framework to genetically modified systems. We present a series of postulates upon which the cybernetic description of overexpression and deletion of existing pathway enzymes is based. Last, we discuss the application of MCA to the cybernetic realization of a metabolic network.

The postulates presented form the basis of the cybernetic metabolic engineering framework. The methodology has been employed subsequently, for both qualitative and quantitative analysis of metabolic networks; see for example Varner and Ramkrishna (1999a) and Varner et al. (1999). In the second paper of the series we present an analysis of metabolic branch points. In this study we translate the discussion of Stephanopoulos and Vallino (1991) into a rigorous cybernetic realization and, in particular, elucidate the role played by cybernetic resource allocation as the vehicle upon which network response is manifested. This study, although qualitative in scope, provides necessary intellectual precursors for the more elaborate developments that have followed. Thus, it should be considered as a primer upon which the base issues related to the prediction of network response are examined in an elementary system.

2. Modular Cybernetic Regulatory Realization

The extension of the cybernetic framework to metabolic engineering requires a fundamental shift in the vision

of the methodology. Originally, the cybernetic framework was created with the intent of providing a straightforward means of describing macroscopic metabolic regulatory phenomena. The earlier work of Kompala et al. (1986) and Baloo and Ramkrishna (1991) illustrates the incorporation of cybernetic control systems into simple abstracted models of microbial growth. These examples testify to the power and simplicity of the framework, and moreover, they establish the minimalist vision that has guided the evolution of the methodology to date. The later cybernetic developments of Straight and Ramkrishna (1994) are the first attempts at a cybernetic description of metabolic pathway systems. It was hypothesized in this work that a metabolic network can be decomposed into four elementary topological building blocks, namely linear, convergent, divergent, and cyclic processes. Straight derived the functional form of cybernetic regulation for the case in which the pathway element is totally isolated from its surroundings in a regulatory sense. From the minimalist perspective, this allowed the description of biotic phase regulatory dynamics, in particular, the time evolution of abstracted intermediate pools. Using this development, Ramakrishna et al. (1996) proposed a lumped model system for the description of microbial growth on mixed substitutable substrates. This system was capable of the description of both preferential and simultaneous utilization of carbon sources as well as the effect upon growth dynamics of preculturing.

The minimalist theme of the previous developments is centered upon obtaining the simplest possible structure that describes the time evolution of macroscopic variables, such as carbon uptake or biomass formation, etc. The extension of the cybernetic framework to metabolic engineering applications requires a shift in vision away from the minimalist school of thought that has guided the evolution of the methodology to this point.

In direct contrast with the previous paradigm, we seek to incorporate as much biological structural detail as possible. This requirement is not insurmountable, given the large database of reaction information available. However, we seek to make the framework robust so as to be applicable in cases in which specific information on network functionality, for example, the kinetics of individual transformations, is not known. This robustness quality follows from the cybernetic aspects of the system, i.e., the goal-oriented driving force provided by the cybernetic variables. Notwithstanding, the formulation of the cybernetic variables for arbitrarily complex metabolic networks is not, currently, a well-developed endeavor. Accordingly the purpose of this section is the introduction of a conceptual algorithm that can be employed for the derivation of cybernetic control systems for arbitrarily complex pathway topologies. The algorithm is presented as series of postulates and illustrated subsequently by the branch point case study as well as subsequent papers.

2.1. Algorithm Formulation. In what follows we present a set of postulates that form the conceptual basis of cybernetic network realization. The purpose of this section is to simply lay the foundation of the approach and to bring forth the core conceptual aspects. We postpone the discussion of the mathematical significance of the postulates until the next section.

Postulate 2.1. A metabolic network has physiological objectives. There exists limited pools of cellular resources that can be allocated to network function to achieve these objectives. Accordingly, because of the influence of evolutionary pressures, the microbe directs the synthesis and activity of enzymes that mediate network function such that the physiological network objectives are achieved in an optimal manner with respect to resource expenditure.

Postulate 2.2. The expression and activity of the enzymes that catalyze network functionality are regulated by the so-called cybernetic control variables which steer network operation toward the physiological objectives, inherent to the network, in an optimal manner. The functional form of the cybernetic variables can be derived by solving the constrained optimization problem associated with network function.

These postulates are the foundation of the cybernetic approach to metabolic network modeling. In fact, however, they are simply a restatement of the traditional tenets of the framework, i.e., a nutritional objective is postulated which is subject to a constraint upon the level of resources that are available for network function. Because of evolutionary pressures, the objective must be achieved in an optimal manner with respect to resource expenditure. The cybernetic variables steer the operation of the network toward physiological objectives in an optimal manner by regulating the rate of enzyme expression and enzyme activity. The point of departure from the traditional framework is the derivation of the functional forms of the cybernetic variables and, in particular, the assumption of multiple competing objectives. In a deeper mathematical context, these differences manifest themselves in the structure of the constrained optimization problem. The traditional approach to this problem would be to construct network regulation in a top down manner. In other words, solve one optimization problem for the totality of the network. There are several disadvantages to this approach. First, the solution of the constrained optimization problem for an arbitrarily complex metabolic network is a somewhat daunting task. Second, and probably more importantly, each new network topology would require the reexamination of a new constrained optimization problem. Thus, the general applicability of this approach is restricted. We propose an approach that takes a fundamentally different viewpoint. It has at its base a bottom up perspective that follows from the idea of constructing global behavior in a hierarchical manner starting from an elementary base. The algorithm, which is termed the modular approach, bases the construction of the regulation of a metabolic network upon the topological structure of the network.

Postulate 2.3. The geometric structure of metabolic network yields insight into the nature of the regulatory configuration that governs its time evolution. We postulate that certain pathway topologies have ingrained function. Thus, there exists a relationship between the function of a pathway and its topological shape.

Postulate 2.3 goes to the heart of the approach and is central to the development that follows. The theme of the algorithm is the construction of metabolic regulation in a hierarchical manner, starting from a common base of elementary components, each with a distinct topological shape and a behavioral function. The topological structure of a metabolic network can be decomposed into elementary pieces. Thus, given a library of elementary pathways, the topological structure of an arbitrary metabolic network can be formulated by appropriately assembling the elementary pathways, which are members of the library.

Definition 2.1. The topological realization of a metabolic network is defined as the set of elementary pathways that are assembled to create the network structure. Note that topological realizations may not be unique.

Following the development of Straight and Ramkrishna (1994), we postulate that the library of elementary

pathway topologies consists of linear, convergent, divergent, and cyclic segments. An elementary pathway, by definition, is considered to be an isolated, local element of a global metabolic map. Accordingly, an elementary pathway is assumed to be insulated from the direct regulatory action of other portions of network. The only influence felt by an elementary pathway from the remaining portion of the network is via flux connections, i.e., material input and output.

To this point we have been careful not to mention the regulatory structure of a metabolic map. Postulate 2.3 hypothesizes a coupling between topological structure of a metabolic network and the associated regulation that directs its time evolution.

Postulate 2.5. The regulatory structure of a metabolic network is a consequence of its topological realization. Each elementary pathway has a particular regulatory function associated with it stemming from its elementary resource allocation problem. The regulatory realization of the assembled global network is the fusion of these elementary objectives subject to the higher purpose of the microbe.

Remark 2.1. The elementary resource allocation problem follows from the work of Straight and Ramkrishna (1994). They hypothesize that isolated elements of a metabolic network each have a local objective. This objective is subject to a constraint upon the resources available to the microorganism for operation of the elementary pathway. Thus, the elementary allocation problem is the traditional cybernetic problem, i.e., a nutritional objective subject to a resource constraint. In fact, the work of Straight and Ramkrishna is a direct generalization of the original framework described by Kompala et al. (1986).

Remark 2.2. The assumption of elementary pathways as the base of the approach has very deep and profound implications, not only upon the description of system behavior but also upon the evolutionary path traveled by the organism. In direct cybernetic terms, the implication of the assumption of elementary pieces is that all resource allocation is done at the elementary base level. This implies all competition is evaluated by assuming that each elementary pathway is isolated. Of course a metabolic map is not an isolated group of elementary pathways that act independently of one another in a regulatory sense. Rather, we postulate that isolated elementary elements of a pathway form a cohesive coordinated network, where the individual units communicate with one another. Thus, the traditional cybernetic vision of isolated elementary pathways must be modified somewhat in light of the coordination between elements of a metabolic map. We address this issue in the next section by proposing a hierarchical control structure which has as its base the elementary pathway concept.

From a resource allocation viewpoint, the assumption of the realization of the topology of a network from elementary pathways implies the existence of local resource pools from which the elementary pathway receives resources. The impact of this issue upon the evolutionary development of the organism is profound because it stipulates the existence of multiple localized resource pools from which network function is derived. In a biological context, this is equivalent to the organizational role played by operons.

Remark 2.3. The exact mathematical implications of the "fusion" of elementary objectives is brought forth in the next section. The importance stems from its impact upon the resource allocation structure of the network. Note that we postulate that the local "fusion" of objectives is subject to the higher purpose of the microorganism. This follows as a consequence of the hierarchical control structure. In other words, we postulate that the local behavior of the network is subject to global regulatory signals that influence local allocation policy. These issues manifest themselves in the functional form of the cybernetic variables that govern network function.

Remark 2.4. The regulatory realization of an elementary pathway may not necessarily be unique. This follows as a consequence of the nonuniqueness property of the topological realization.

2.2. Algorithm Implementation. Now that we have formulated the basis of the modular approach, let us move a step further and discuss the mathematical specifics of implementation. The key linkage that allows the topological construction of regulatory machinery is the postulate of the hierarchy of metabolic control. This concept is not new in biology. However, incorporation of this stratified understanding into the cybernetic framework is a fresh approach that allows the framework to be utilized to construct the regulation of realistic models of metabolic pathways.

Postulate 2.6. The control of a biological process can be decomposed into a hierarchy of components:

1. The base component is termed the elementary regulatory component. This component is the regulation associated solely with the elementary pathways that constitute the topological realization. This component derives its functionality from the solution of the elementary resource allocation problem. In a biological context, this level of control and functionality is somewhat analogous to that of individual operons.

2. The local regulatory component is constructed from elementary components. This level of regulation governs the interaction of elementary components. In a biological context, this level of control describes how multiple individual operons that yield common functionality interact and compete.

3. The highest level of metabolic control is termed global control. This level of regulatory action consists of both nutritional and topological control signals that govern local regulatory action. Global control acts as a continuous filter or high-level override switch upon local regulatory action. These signals are in operation network wide and coordinate the activity of entire local metabolic pathways, or subnetworks. In a biological context, this class of signal coordinates the functionality of sets of operons.

Remark 2.5. The cybernetic basis of the regulatory hierarchy rests with the formulation of the metabolic regulation of the elementary pathways. The derivation of the cybernetic variables that govern the elementary component of enzyme expression and enzyme activity follows from Straight and Ramkrishna (1994). The functional forms of the elementary cybernetic variables depends on the nature of the elementary pathway. More precisely, it depends on the objective function and resource constraint. These two factors, when coupled with pathway topology, form classes of competition for critical cellular resources. These classes, termed substitutable and complementary, dictate the regulatory characteristics displayed by the elementary pathway.

Definition 2.2. A substitutable elementary pathway is one in which the pathway key enzymes compete for cellular resources to produce a common product. The linear and convergent elementary pathways are substitutable. The elementary cybernetic variable that governs the allocation of critical resources for enzyme synthesis

for a substitutable process follows from the matching law and is given functionally by

$$u_j = r_j / \sum_{i=1}^{k} r_i, \quad j = 1, 2, ..., k \tag{2.1}$$

where r_j denotes the jth specific reaction rate and the index k denotes the number of enzymes competing for resources from the same pool. In a biological context, u_j is akin to an isolated lumped transcriptional/translational efficiency governing the expression of e_j. The elementary cybernetic variable that governs enzyme activity for a substitutable process follows from the proportional law and is given functionally by

$$v_j = r_j / [\max_i(r_i);\ i = 1, 2, ..., k] \quad j = 1, 2, ..., k \tag{2.2}$$

Biologically, v_j is equivalent to the scaled specific activity of e_j measured relative to the other enzymes competing for resources from the same pool.

Definition 2.3. A complementary elementary pathway is one in which the key enzymes that catalyze the pathway reactions compete to maximize the mathematical product of metabolites. The cyclic and divergent elementary pathways are complementary processes. The functional form of the elementary cybernetic variable that governs the allocation of critical resources for a complementary process follows from the matching law and takes the form

$$u_j = \frac{r_j / p_{j+1}}{\sum_{i=1}^{z} r_i / p_{i+1}}, \quad j = 1, 2, ..., z \tag{2.3}$$

where p_{k+1} denotes the specific level of product being produced by r_k and the index z denotes the number of enzymes competing from the same elementary resource pool. The elementary cybernetic variable that governs the activity of jth key enzyme belonging to a complementary elementary pathway follows from the proportional law and is given functionally by

$$v_j = \frac{r_j / p_{j+1}}{[\max_i(r_i / p_{i+1});\ i = 1, 2, ..., z]}, \quad j = 1, 2, ..., z \tag{2.4}$$

Note that the biological meaning of the elementary cybernetic variables u_j and v_j for a complementary elementary pathway is identical to a substitutable elementary pathway; however, their functional forms reflect different objectives.

Remark 2.6. Elementary divergent pathways, as we shall see, are not always complementary processes. Whether an elementary divergent branch point is complementary is dependent upon the choice of objective function and the nature of the resource constraint. This issue is addressed by the subsequent members of this series.

As postulated, the topological realization of a metabolic network involves the assembly of elementary pathway components. These elementary pieces may interconnect in one of two ways. The classification of the types of elementary intersections has a profound influence upon the structure of resource allocation and, hence, the regulatory characteristics of the system.

Definition 2.4. If the end product of one elementary pathway serves as a precursor for the connecting elementary pathway, the two elementary pieces are said to intersect.

Definition 2.5. If at least a single key enzyme can be written as a member of multiple elementary pathways, the elementary units are said to overlap.

Postulate 2.7. Let an arbitrary topological realization of a metabolic network be composed of n elementary pathway units. Each elementary pathway has an independent, distinct pool of key cellular resources that is earmarked for allocation to enzyme synthesis. The elementary cybernetic variable u_j^z controls the allocation of critical resources for the synthesis of the jth enzyme from the zth elementary resource pool.

Postulate 2.8. If a key enzyme is a member of only the jth elementary pathway, allocation of critical resources to this enzyme comes solely from the jth elementary resource pool. If however a key enzyme is a member of both the jth and kth ($k \neq j$) elementary pathways, i.e., is overlapping between the two, the enzyme can receive resources from both elementary resource pools.

Remark 2.7. These two postulates, in a deeper sense, govern the transition from the set of isolated constrained optimization problems associated with the elementary pathways and the interacting regulation of the metabolic network. At issue is the identity of the pool of resources that an arbitrary key enzyme is allocated from.

In the case of an intersecting pathway, the elementary regulatory component is (barring any global effects) not aware of the existence of the remainder of the pathway. In other words, suppose a linear pathway intersects a divergent pathway. The elementary allocation problem of the linear pathway is unaware of the existence of the divergent pathway, and vice versa. However, in the case of overlapping elementary pathways, i.e., enzymes which are members of more than a single elementary pathway, this is not the case. When n elementary pathways overlap, the key enzymes that are members of multiple pathways form a bridge upon which the individual elementary pathways communicate. More precisely, enzymes that are members of multiple elementary pathways compete for key cellular resources from different pools, i.e., they compete for cellular resources from each of the n pools associated with the elementary pathways. It is through this competition that one elementary pathway becomes, in some sense, sensitive to the needs of all the other elementary pathways that it overlaps. This couples the elementary regulation into a local regulatory unit with an objective and resource constraints just as the elementary pathway. The resource constraint in this instance consists of the set of elementary resource constraints associated with each of the elementary pathways used in the realization. The nature of the objective function is discussed subsequently.

The following set of definitions and postulates expand on this discussion:

Definition 2.6. The local cybernetic regulatory component is the level of metabolic control at which the interaction between overlapping elementary pathways is manifested. Let an arbitrary metabolic network be composed of n key enzymes. Furthermore, suppose the jth key enzyme can be written with respect to q elementary pathways. The local cybernetic variable that controls the expression of e_j, $j \leq n$ is given by

$$u_j^l = \prod_{i=1}^{q} u_j^i, \quad j = 1, 2, ..., n \tag{2.5}$$

where the superscript l denotes local and the product limit q denotes the number of elementary pathways of which e_j is a member. The local cybernetic variable that

governs enzyme activity follows by analogy and is given functionally as

$$v_j^l = \prod_{i=1}^{q} v_j^i, \quad j = 1, 2, ..., n \tag{2.6}$$

Remark 2.8. The product of elementary cybernetic variables is a consequence of the possibility of multiple allocation sources, i.e., if e_j is a member of q elementary pathways, then e_j can receive resources from all q elementary resource pools. Accordingly, the product of elementary variables captures the input of the q elementary allocation policies. In other words, the elementary cybernetic variables govern the allocation of resources from each respective elementary pool. Because e_j competes for key cellular resources from q pools, every allocation source must be accounted for.

The local topological structure is cybernetic in the sense that its regulation directs, subject to a constraint upon available resources, the synthesis and activity of key enzymes so as to achieve a specific local objective. The nature of the objective function and resource constraints stem from the particular topological realization, in other words, which combination of elementary pathways was used to construct the network.

Postulate 2.9. Let the realization of an arbitrary metabolic network consist of n elementary pathways whose individual objectives

$$\max(f_1(\mathbf{X}_1)), \max(f_2(\mathbf{X}_2)), ..., \max(f_n(\mathbf{X}_n)) \tag{2.7}$$

are subject to the set of resource constraints

$$\{g_j(\mathbf{X}_j) = G_j\}_{j=1}^{n} \tag{2.8}$$

where $\mathbf{X}_j$ denotes the resource pool earmarked for the operation of the jth elementary pathway. The local objective of the network is then the superimposed goal of the n elementary pathways, subject to $k \leq n$ independent resource constraints, i.e.

$$\max(f_1(\mathbf{X_1})) \text{ and } \max(f_2(\mathbf{X_2})) \text{ and } ... \text{ and } \max(f_n(\mathbf{X}_n)) \tag{2.9}$$

subject to

$$g_1(\mathbf{X}_1) = G_1 \text{ and } g_2(\mathbf{X}_2) = G_2 \text{ and } ... \text{ and } g_n(\mathbf{X}_n) = G_n \tag{2.10}$$

Remark 2.9. Postulate 2.9 defines the cybernetic basis for the local topological realization of a metabolic network. This level of control is termed local because it is unaware of any higher purpose held by the microorganism. In other words, once activated, the local regulatory component controls the function of the metabolic network; however, it is totally unaware of the nutritional factors the lead to its activation or termination. Biologically, this level of control is equivalent to the set of regulatory signals that coordinate the expression and activity of entire sections of metabolic pathways.

Remark 2.10. Postulate 2.9 in the case of overlapping pathways implies that each elementary allocation problem associated with an overlapping key enzyme is solved independently of the others. In other words, because of the assumption of elementary base components, the allocation problem associated with any individual overlapping enzyme is considered in isolation, ignoring the effect of the other elementary allocation problems. It is in this way that the elementary cybernetic variables that govern the allocation from a specific pool of cellular resources can be derived. A more detailed discussion of the mathematical consequences of this aspect of the formulation are given in the Appendix.

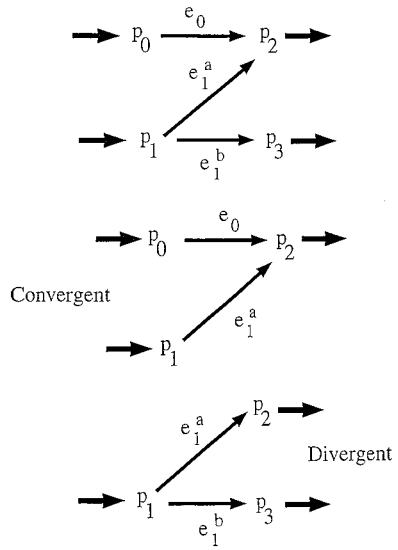

Figure 1. Example of an overlapping topological realization composed of an elementary convergent and divergent pathway. The key enzymes e_0 and e_1^a producing metabolite p_2 are substitutable, whereas key enzymes e_1^b and e_1^a producing metabolite p_3 are complementary. The key enzyme e_1^a is a member of both elementary pathways; accordingly, it competes for resources from both pools. Moreover, e_1^a is sensitive to both elementary pathway objectives.

The next level of control, i.e., the regulation of local metabolic activity is termed global regulation. This class of control signal imparts a higher understanding of the physiological situation of the global network to local sections of metabolic pathways. However, before we embark on a detailed discussion of global control mechanisms, we present a case study to make the local topological realizations postulates more concrete.

Example 2.1. Consider the structure shown in the top portion of Figure 1. This pathway may be decomposed into two overlapping elementary pathways, specifically, convergent and divergent elementary pathway. Accordingly, from postulate 2.7, there exist two elementary resource pools from which resources are allocated to network enzyme synthesis. We define the elementary convergent pathway as pathway a and the elementary divergent pathway as pathway b.

The elementary pathway a is a substitutable process which implies the elementary cybernetic variable(s) that govern the expression and activity of e_0 and e_1^a take the form

$$u_0^a = \frac{r_0}{r_0 + r_1^a}, \qquad u_{1a}^a = \frac{r_1^a}{r_0 + r_1^a} \tag{2.11}$$

$$v_0^a = \frac{r_0}{\max(r_0, r_1^a)}, \qquad v_{1a}^a = \frac{r_1^a}{\max(r_0, r_1^a)} \tag{2.12}$$

The elementary pathway b is a complementary process which implies the elementary cybernetic variable(s) that

govern the expression and activity of e_1^a and e_1^b are of the form

$$u_{1a}^b = \frac{r_1^a/p_2}{r_1^a/p_2 + r_1^b/p_3}, \qquad u_{1b}^b = \frac{r_1^b/p_3}{r_1^a/p_2 + r_1^b/p_3} \tag{2.13}$$

$$v_{1a}^b = \frac{r_1^a/p_2}{\max(r_1^a/p_2, r_1^b/p_3)}, \qquad v_{1b}^b = \frac{r_1^b/p_3}{\max(r_1^a/p_2, r_1^b/p_3)} \tag{2.14}$$

Note that the key enzyme e_1^a can be written with respect to two elementary pathways. From definition 2.6, the local cybernetic regulatory component is synthesized from the elementary cybernetic variables. Because e_0 is a member of only a single elementary pathway, the local cybernetic variable that regulates e_0 expression is given solely by the elementary component, i.e.

$$u_0^l = u_0^a \tag{2.15}$$

The key enzyme e_1^a is a member of both elementary pathways. This implies, by postulate 2.8, that e_1^a can receive critical resources from both elementary resource pools. This overlap is a point of regulatory communication between the elementary pathways. The local cybernetic variable(s) that govern the expression and activity of e_1^a are sensitive to both elementary pathways because of the possibility of joint resource allocation. By postulate 2.6 the local cybernetic variable that governs the expression of e_1^a is given functionally as the product of elementary regulatory components

$$u_{1a}^l = u_{1a}^a\, u_{1b}^b \tag{2.16}$$

The key enzyme e_1^b is a member of only a single elementary pathway which implies from postulate 2.6 that the local cybernetic variable that governs the expression of e_1^b is given by the elementary component

$$u_{1b}^l = u_{1b}^b \tag{2.17}$$

The functional forms of the cybernetic variables that control enzyme activity follow by analogy.

Example 2.1 illustrates the application of the local cybernetic realization postulates. For those interested in simply applying the algorithm no deeper understanding is required. Moreover, the proof of postulate 2.9 and the formal justification of postulate 2.6 is somewhat involved and may be skipped upon a first reading. However, for a deeper understanding of the origin and the associated subtleties of local cybernetic realizations, such a discussion is required. Accordingly, we sketch the foundation of a proof of postulate 2.9 and a formal justification of postulate 2.6 in the Appendix. Additionally, the reader is prompted to consult Varner and Ramkrishna (1998b) for a bifurcation analysis of the structure shown in example 2.1 to understand, at a core level, the interaction and role played by the local level of cybernetic regulation.

The highest postulated level of metabolic control is the global control mechanism. This class of control signal regulates the local regulatory component.

Postulate 2.10. Global control signals constitute the highest level of metabolic control. They act to promote or restrain local physiological functionality. These signals can be subdivided into two subclasses, nutritional and topological global control.

Definition 2.7. The influence of global control mechanisms can be described by the convention

$$u_j^g = \prod_{k=1}^{n} U_j^k, \quad v_j^g = \prod_{k=1}^{n} V_j^k, \quad j = 1, 2, ..., n \tag{2.18}$$

where u_j^g and v_j^g denote the cybernetic variables that describe the global component of the regulation of enzyme expression and activity, respectively.

The global control variable U_j^k ($k = 1, 2, ..., n$) denotes the kth global control signal that influences the expression of e_j. The term V_j^k ($k = 1, 2, ..., n$) denotes the kth global regulatory signal that influences the activity of e_j.

Definition 2.8. The cybernetic variable that reflects both local and global regulatory concerns is termed the complete cybernetic variable. The complete cybernetic variable that governs the expression of the jth key enzyme of an arbitrary metabolic network, denoted as e_j, is given functionally by

$$u_j = u_j^l\, u_j^g, \quad j = 1, 2, ..., n \tag{2.19}$$

where u_j^l and u_j^g denote the local and global cybernetic variables, respectively, and n denotes the number of key enzymes that compose the network. The complete cybernetic variable that governs the activity of e_j follows by analogy and is given functionally by

$$v_j = v_j^l\, v_j^g, \quad j = 1, 2, ..., n \tag{2.20}$$

where v_j^l and v_j^g denote the local and global regulatory component, respectively.

Remark 2.11. The complete cybernetic variable incorporates the totality of the regulatory hierarchy. The local cybernetic variable governs the interaction of the elementary pathways, whereas the global variable reflects the higher physiological purposes of the network. The functional form of the complete cybernetic variable is responsive to both levels of control.

Nutritional global control upon local enzyme expression and activity is better understood as metabolite repression and/or inhibition. Thus, in general, the nutritional global control variable transforms the regulatory significance of nutritional state into a control action which acts to activate or restrain local enzymatic processes. For example, consider the case of storage metabolism. The storage pathway enzymatic machinery is not present in time of nutritional plenty. However, upon the onset of nutritional limitation, for example a nitrogen limitation, the enzymatic machinery that catalyzes the formation of storage polymers would be synthesized. The regulatory action that signals the onset of limitation in this context would be classified as a global nutritional regulatory signal. In other words, once a particular nutritional condition is satisfied, the local regulatory component is allowed to express and activate the storage pathway. However, before the nutritional condition is satisfied, the storage pathway machinery is repressed and/or inhibited. The functional form of global nutritional control variable(s) depends on the nature of the specific nutritional signal. As such the functionality must be derived on a case by case basis.

The second subclass of global control signal, i.e., topological global control, is a direct outgrowth of the geometric shape of the metabolic network. This class of global signal regulates the choice of entire local metabolic

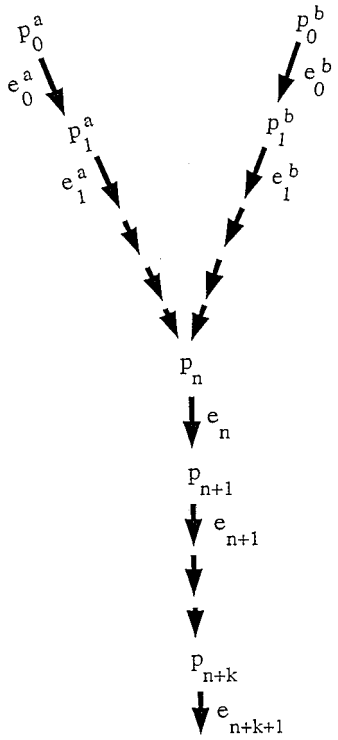

Figure 2. Example of substitutable global topological control. The converging linear pathways converting metabolite p_0^a and p_0^b to p_n are locally unaware of the existence of the opposite branch. However, on a global level, they compete to produce a common metabolite, namely p_n. Thus, key enzymes e_j^k ($j = 0, 1, 2, ..., n - 1$; $k = a, b$) are globally substitutable.

pathways and is best explained by example. Consider the pathway shown in Figure 2.

Assuming the linear segment starting from the intermediate p_n is an independent elementary linear pathway, the converging branches can be modeled as independent elementary linear segments terminating at p_n. Because no key enzyme is a member of more than one elementary pathway, the local regulatory structure is identical to the elementary regulatory component (by postulate 2.6). Locally, because each branch is an elementary pathway, a member of a particular branch is not aware of the existence of any other enzyme outside of its elementary pathway. In other words, e_1^a for example, is not aware of any enzyme belonging to branch *b*. However, notice that both branches compete in a global sense to produce a common product, namely p_n. This represents a wasteful resource allocation structure because no choice of optimum branch is possible. More precisely, suppose the formation of p_n via branch *b* is slow compared to branch *a* production. From a resource allocation viewpoint, all else being equal, the operation of branch *b* represents a poor resource investment on the part of the microorganism. We hypothesize, because of evolutionary pressures, that such situations do not exist, i.e., the organism because of nutritional stress is forced to choose the branch which yields the best return. This choice behavior is manifested via topological global control variable. Clearly, the choice behavior must be described at a global level because locally each branch is unaware of the others existence. The derivation of the form of the global topological control variable, in this case, follows much the same as the elementary cybernetic variables, i.e., we hypothesize a global objective which is subject to a resource constraint. However, in this instance, our resource postulates are not centered around managing resource required for enzyme synthesis because this regulatory decision occurs at a lower level in the regulatory hierarchy. Rather, we hypothesize that topological regulatory decisions require key cellular resources to implement and maintain. Considering the network shown in Figure 2, suppose branch *b* is a poor resource investment, i.e., produces the product p_n at a much slower rate than branch *a*. The global topological control mechanism would act to repress/deactivate the synthesis and activity of the key enzymes which catalyze the branch *b* reactions. The global resource hypothesis stipulates that this regulatory decision requires cellular resource to implement. In other words, the synthesis of branch *b* key enzymes which would certainly require large amounts of cellular resource, especially in the case of a long pathway, is repressed by the microorganism (if branch *b* is a poor investment) only at a resource cost. By spending resources on the global control of branch *b*, the microorganism strategically saves the resource required for branch *b* enzyme activation and expression. As was true with the elementary cybernetic variables, the form of the global topological control variable depends on the geometric shape as well as the competitive structure of the network. Moreover, given the diverse geometric nature of metabolic networks, a general derivation which has application to all structures may not exist. Thus, in this case, as was true with nutritional global control, the functional forms may have to be derived on a case by case basis. We present below two sample derivations so the reader may become familiar with the arguments which underlie the formulation.

Example 2.2. Consider the metabolic network shown in Figure 2. Because the elementary linear branches compete to synthesize a common product, the global topological competition between branches *a* and *b* is substitutable. The derivation of the substitutable global topological control variable follows from the same arguments as the elementary convergent pathway. The objective of the elementary linear branches of the network shown in Figure 2 is the maximization of the end product p_n. Thus, by postulate 2.9, the local goal of the network is the superimposed goal of the elementary pieces, i.e., the maximization of the end product p_n. Furthermore, suppose the choice of synthesis route requires key cellular resources to be implemented, i.e., it costs some resource to repress/deactivate the branch that is a poor investment. Physically, this hypothesis is equivalent to assuming that a certain degree of cellular resource is required to instigate and maintain a global topological decision. These statements formally become the constrained optimization problem

$$\max\left(\sum_{i=1}^{k} p_n^i(R_i)\right) \quad \text{subject to} \quad g = \sum_{i}^{k} R_i = R \qquad (2.21)$$

where p_n^i denotes the production of intermediate p_n by the *i*th route and R_i denotes the resource allocated to instigate global topological control of the *i*th branch. The optimality condition given by

$$\frac{dp_n^1}{dR_1} = \frac{dp_n^2}{dR_2} = \cdots = \frac{dp_n^k}{dR_k} \qquad (2.22)$$

can be rearranged to yield the matching criteria

$$\frac{dp_n^q}{dp_n^q + \sum_{i=1,q}^{k} dp_n^i} = \frac{dR_q}{dR_q + \sum_{i=1,q}^{k} dR_i} \quad (2.23)$$

Equation 2.23 states that the optimum choice of p_n production route occurs when the fractional return from the qth route equals the fractional resource allocation to qth production route. To this point, the derivation is very similar in structure to the elementary convergent pathway derivation. However, in this case, we assume a different form for the return on resource investment. Specifically, we postulate that the return on resource investment for an entire linear metabolic pathway of length n is the average reaction rate, i.e.

$$\frac{1}{n_j}\sum_{i=0}^{n_j-1} r_i^j, \quad j = 1, 2, ..., k \quad (2.24)$$

where r_i^j is the unmodified flux through the ith step of the jth linear elementary pathway. If we assume that the global topological choice is implemented at every instant in time and takes place on time scale dt, the global cybernetic topological control variables that govern the choice of p_n synthesis branch (global signal influencing the wth enzyme of jth branch) are given by

$$U_{w,j} = \frac{\frac{1}{n_j}\sum_{i=0}^{n_j-1} r_i^j}{\frac{1}{n_j}\sum_{i=0}^{n_j-1} r_i^j + \sum_{h=1,j}^{k} \frac{1}{n_h}\sum_{i=0}^{n_h-1} r_i^h} \quad j = 1, 2, ..., k \quad (2.25)$$

Notice if the length of all j linear pathways is unity, global topological control variable 2.25 reduces to the elementary convergent cybernetic u_j variable. Thus, the converging linear pathway aspect of the network is a multiple-step generalization of the elementary convergent pathway.

We reserve comment upon the salient features of the sample derivation until after the completion of the next example.

Example 2.3. Consider the metabolic network shown in Figure 3. In this case the elementary linear branches compete to consume a common substrate, namely p_0, which implies the network is globally complementary. We make the special point of classifying the network as globally complementary, as opposed to elementary or locally complementary, because the elementary pathways that compose the network are linear pathways, which are substitutable. The objective function proposed by Straight and Ramkrishna (1994) for the elementary divergent pathway is the maximization of the mathematical product of the end point metabolites. We assume this objective to be the global regulatory goal and postulate that it is subject to a constraint on the amount of resources which are available to implement the global control action. These statements formally become the constrained optimization problem

$$\max\left(\prod_{j=1}^{k} p_{n_j}^j(R_j)\right) \quad \text{subject to} \quad g = \sum_{q=1}^{k} R_q \quad (2.26)$$

where $p_{n_j}^j$ denotes the endproduct of the jth linear branch of length n_j and R_q denotes the level of critical

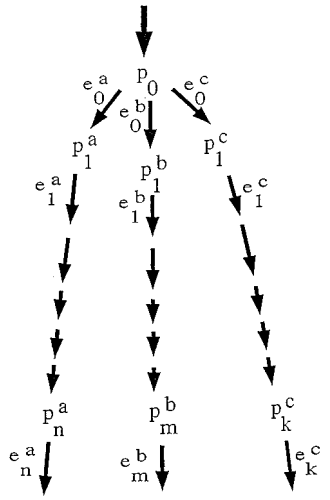

Figure 3. Example of complementary global topological control. This is a generalized divergent pathway composed of three elementary linear branches, *a*, *b*, and *c*. By definition each linear branch contributes to the complementary objective of the maximization of mathematical product of the specific metabolite levels $p_n^a p_m^b p_k^c$. Thus, although branch *a* is locally unaware of branches *b* and *c*, at a global level, the key enzymes catalyzing the reactions of each respective branch are complementary.

resource allocated for the global control of qth linear branch. The optimality condition given by

$$p_{n_2}^2 p_{n_3}^3 \cdots p_{n_k}^k \frac{dp_{n_1}^1}{dR_1} = p_{n_1}^1 p_{n_3}^3 \cdots p_{n_k}^k \frac{dp_{n_2}^2}{dR_2} = \cdots = p_{n_1}^1 p_{n_2}^2 \cdots p_{n_{k-1}}^1 \frac{dp_{n_k}^k}{dR_k} \quad (2.27)$$

can be rearranged to yield the matching criteria

$$\frac{\frac{dp_{n_j}^j}{p_{n_j}^j}}{\frac{dp_{n_j}^j}{p_{n_j}^j} + \sum_{q=1,j}^{k} \frac{dp_{n_q}^q}{p_{n_q}^q}} = \frac{dR_j}{dR_j + \sum_{q=1,j} dR_q} \quad (2.28)$$

Equation 2.28 states that the optimum network operation occurs when the fractional return from the jth branch is equal to the fraction allocation of global control resources to the jth branch. We assume that the return on investment of the entire jth linear pathway is given by the average reaction rate, i.e.

$$\frac{1}{n_j}\sum_{q=1}^{n_j} r_q^j, \quad j = 1, 2, ..., k \quad (2.29)$$

If the global control mechanism is implemented at every instant in time, and takes place in time dt, the global topological control variable which governs the operation of the generalized divergent branch point is given functionally by

$$U_{w,j} = \frac{\frac{1}{n_j p^j_{n_j}} \sum_{q=1}^{n_j} r^j_q}{\frac{1}{n_j p^j_{n_j}} \sum_{q=1}^{n_j} r^j_q + \sum_{l=1,j}^{k} \frac{1}{n_l p^l_{n_l}} \sum_{q=1}^{n_l} r^l_q} \quad (2.30)$$

where $U_{w,j}$ denotes the topological control variable that globally regulates the expression of the wth key enzyme belonging to the jth linear branch. Notice if all the linear elementary pathways are of unit length, eq 2.30 reduces to the regulation of the elementary divergent pathway. Thus, the network shown in Figure 3 is a multiple-step generalization of the elementary divergent pathway.

Remark 2.12. The general features of the derivation of topological global control variables follows the same problem formulation as previous cybernetic investigations. In other words, we postulate an objective for the network and hypothesize a constraint upon the level of resource available to achieve this objective. In the case of global topological control, the concept of the utility of the resource is altered somewhat; however, the basic spirit is preserved. In the general case, multiple-step generalizations of elementary pathways assume the same objective function as the original isolated pathway.

Remark 2.13. Notice that the length of an elementary linear leg is inversely proportional to the measured return on investment. This is consistent with the hypothesis that it requires more resources to operate a lengthy pathway compared to a shorter counterpart. In other words, suppose we have two linear pathways converging to a common product p_n. One linear pathway is length m, while the other is length $n \leq m$. From a resource allocation viewpoint, the shorter pathway is a better investment, i.e., it yields a higher return if the sum of the reaction rates is approximately equal for both pathways. Also note that the choice of optimum synthesis route is dynamic because it is a function of the system reaction rates. Thus, the choice of which route or combination of routes is functioning at any given time is not fixed and is a strong function of the operational environment.

Remark 2.14. The choice of the functional form of the return on investment of an entire linear pathway of length n follows somewhat from the minimalist cybernetic developments of the past. Previously, because of the abstracted nature of the framework, there was no need to consider the "subregulation" associated with individual linear pathway segments, i.e., one set of cybernetic variables was formulated that governed the synthesis and activity of a lumped key enzyme. Given the perspective of the current development, we seek to postulate a framework with a less abstracted stance. Therefore, this is a first postulated step toward this objective. From a mathematical perspective, the underlying basis of the methodology is independent of the functional form of the return on investment in any case. Thus, although the functional form of the control variables will certainly change with a different notion of return, the spirit of the development remains unchanged.

Remark 2.15. In general, topological global control can be rationalized as a high-level override mechanism that regulates function among competing pathways or subnetworks. Thus, unlike traditional cybernetic approaches, the cost versus benefits of operating entire metabolic pathways or subnetworks is weighed. From another perspective, topological global control is the mechanism by which the microorganism switches modes of physiological functionality. More exactly, each of the competitive choices in this instance is a local or elementary pathway, or in the general case an arbitrary subnetwork. Thus, each of these structures has physiological objectives. By weighing the cost versus benefit of operating these structures, global topological control is in effect sanctioning particular functionality.

Remark 2.16. Note that the point where $U_{w,j}$ acts, i.e., which set of enzymes are sensitive to this signal must be handled on a case by case basis. One possibility for the point of action would be the entrance into the pathway. In terms of resource expenditure, such a choice is consistent with the understanding that it would be judicious to regulate network influx, rather than to curtail function somewhere downstream where resources have been expanded to partially process metabolic intermediates.

The last issue that we comment upon is the uniqueness of an arbitrary cybernetic regulatory realization. Given a particular network topology, there may exist a number of degenerate representations, i.e., more than one combination of elementary pathways that can be assembled to form the network. The question that arises in this instance is which combination is the correct representation? To some extent this question can be answered using tools from bifurcation theory to determine the qualitative nature of each elementary pathways and the subsequent effect upon the qualitative characteristics of the system when the units are assembled. Such an undertaking has been done, in a different context, by the authors (Varner and Ramkrishna, 1998b). However, for the most part, this question must be answered by trial an error and experience. The formulation of postulates in this instance is limited to case specific endeavors. Thus, a general undertaking is impossible. However, some basic remarks can be made as to the form of the nonuniqueness.

Remark 2.17. The objective function of an arbitrary network realization is not unique. This follows as a consequence of the nonuniqueness of topological realizations.

Remark 2.18. The resource allocation structure of arbitrary metabolic network is not unique. This follows as a consequence of the nonuniqueness of the topological realization.

2.3. Summary. In this section, we have presented a series of postulates that form the basis of the cybernetic description of the regulation of metabolic networks. This approach, termed the modular approach, can be employed to formulate the cybernetic variables for an arbitrary metabolic network. In the broad scope of the current development, this forms only a single component required for the application of the cybernetic methodology to metabolic engineering.

In the next section, we take another step forward and discuss the extension of the cybernetic framework to describe the regulation of metabolic networks that have been altered genetically. The discussion follows much the same format as the previous development. In particular, we discuss the optimal nature of the cybernetic framework and its ramifications upon the description of genetically altered microorganisms.

3. Extension of the Cybernetic Framework to Recombinant Systems

The purpose of this section is the discussion of the extension of the cybernetic framework to include metabolic pathways that have been altered genetically. We

present a series of postulates that form the basis for the extension. We begin by presenting the basic concepts that underlie the optimality of the cybernetic framework. We then reexamine these postulates in light of overexpression and deletion of existing pathway enzymes.

The core hypothesis of the cybernetic framework is that, because of evolutionary pressure, microorganisms have evolved control machinery that steers the physiological state toward an objective in an optimal manner. The control action takes the form of the regulation of expression and activity of the enzymes that constitute the network. Previous cybernetic investigators have assumed, when formulating abstracted models of microbial processes, the maximization of growth rate as an evolutionary objective. However, in the context of metabolic engineering, although such an objective may be operating at a suitably refined level, the local objective of a metabolic network is postulated to be a nonunique function of the pathway topology, i.e., the network objective follows from the interaction of the individual elementary pathway objectives.

These two schools of thought are not inconsistent, even when they appear to be contradictory, for example in the case of storage metabolism. Often the product of metabolic function is required for growth, i.e., specific intermediates, energetic metabolites, etc. Thus, a postulated objective of a metabolic network is local in the sense that it is a suitable refinement of the macroscopic growth rate postulate. In the case of a seemingly contradictory objectives (nongrowth), it must be remembered that global objectives may be nutritional state dependent (as is true of storage metabolism or maintenance function.) In these instances the objective of nongrowth related metabolic networks is still a function of the topology; however, the nutritional state in some measure influences the network startup, i.e., the network is subject to global nutritional control. Whichever the case may be, the metabolic control system that regulates the network evaluates the existing nutritional alternatives and directs the synthesis of the network enzymes for the set of alternatives that yield the optimum nutritional outcome. We postulate that for a genetically modified network this process is altered slightly, however, in spirit remains unchanged.

Postulate 3.1. The objective of a base network is still in place for a genetically modified version of the same network.

This postulate implies that overexpression or deletion of existing pathway enzymes does nothing to change the objective of the network. In other words, the base network objective is in place because of the force of millions of years of evolutionary pressure and, as such, cannot simply be immediately erased by a small number of perturbations. Overexpression or deletion of enzymatic machinery does, of course, influence the performance of the pathway and, moreover, the choice operational functionality, i.e., which particular native enzymes are expressed and the degree of promotion in each case because now the base network faces a modified set of operational alternatives.

Postulate 3.2. The intact base metabolic control system, in operation in the modified network, must cope with an expanded or contracted set of possible alternatives reflecting overexpression or deletion of network enzymes when regulating the synthesis and activity of existing network enzymes.

In short, it is hypothesized that the objective function of the base network is still in place even in the case of the modified system. However, the available set of physiological alternatives that are open to the genetically modified network is different. Thus, from a resource allocation perspective, the result of genetic modifications to the network (deletion or overexpression) is the contraction or expansion of the set of alternatives that are available to the microorganism, not the alteration of the objective itself. Accordingly, the allocation of critical resources for network operation is still taken to be optimal; however, the set of alternatives that compete for these resources is modified. In simpler terms, these arguments imply that base network control machinery directs metabolic function toward the base objective despite the genetic alteration. Clearly, however, the base case and the altered network are not the same. More precisely, the intact base regulatory machinery still directs function toward the base network objective in an optimal manner; however, it does so in light of the genetic perturbation. In a formal sense, we maintain the following:

Postulate 3.3. Functionality shifts observed in response to genetic perturbation are an attempt on the part of the network control machinery to maintain the base objective as best possible in the face of changes forced upon the system by genetic alteration.

In a cybernetic sense, the deletion of existing pathway enzymes reduces the number of alternatives that compete for key cellular resources. Thus, inclusion of this class of genetic manipulation into the current framework is straightforward.

Postulate 3.4. Let u_j^k denote the elementary cybernetic variable that governs the allocation of critical resources from the kth pool to the jth network enzyme. Now suppose the gene that encodes for the jth enzyme has been deleted. The jth enzyme no longer competes for key cellular resource from the kth resource pool or any other which implies

$$u_j^k \equiv 0 \tag{3.1}$$

Remark 3.1. An entirely equivalent description of the ramifications of gene deletion follows by assuming the rate constant governing the expression of enzyme e_j is defined as

$$\alpha_{e_j} \equiv \begin{Bmatrix} \alpha_{e_j} & \text{gene encoding for } e_j \text{ present} \\ 0 & \text{gene encoding for } e_j \text{ deleted} \end{Bmatrix} \tag{3.2}$$

This representation implies $e_j = 0\ \forall t \rightarrow r_j = 0$ if the gene that encodes for e_j has been deleted; thus, the expression of e_j does not compete for cellular resources because allocation is based upon the rate of return, as measured by reaction rate.

Another possible route of genetic manipulation is the mutation of a gene which encodes for an enzyme such that a normal expression profile is observed; however, the enzyme is capable of only a fraction of its native activity.

Postulate 3.5. Let v_j^k denote the elementary cybernetic variable that regulates the activity of the jth enzyme of the kth elementary pathway. Suppose the gene which encodes for e_j is altered such that e_j is only capable of a fraction of its wild-type activity. The modified elementary cybernetic variable that reflects this contingency, denoted by $\bar{v}_j^k$, is of the form

$$\bar{v}_j^k \equiv \phi_j^k\, v_j^k, \qquad 0 \le \phi_j^k \le 1 \tag{3.3}$$

where ϕ_j^k denotes the fraction of maximum of activity

that the modified e_j is capable of. A value of $\phi_j^k = 1$ denotes the wild-type activity, whereas $\phi_j^k = 0$ indicates a completely blocked enzymatic step.

Remark 3.2. Note that postulate 3.5 assumes a normal expression profile, i.e., the expression of the *j*th gene that encodes for e_j is identical to that of the wild-type network. This implies that the alteration that leads to decreased activity is not in the promoter region. If this is not true, i.e., the expression profile and the activity of e_j are altered, then in addition to postulate 3.5 the rate and saturation constant(s) that govern e_j expression can be altered to reflect the new expression profile.

Remark 3.3. Given the assumption of a normal expression profile, the completely blocked case ($\phi_j^k = 0$) is *not* equivalent to postulate 3.4 or remark 3.1 because we assume e_j is still expressed but has no activity.

The deregulation of existing network enzymes can also be described within the framework, i.e., the removal of metabolite feedback inhibition/repression. Note, within the present context, that these signals are classified as global nutritional signals; thus, we frame the discussion around the more general topic of removal of global regulatory sensitivity.

Postulate 3.6. Suppose the enzyme e_j is a member of the *k*th elementary pathway and is subject to *q* global control signals. The complete cybernetic variable that regulates the expression of e_j is given by

$$u_j = u_j^l(U_1U_2 \ldots U_r \ldots U_q) \tag{3.4}$$

where u_j^l denotes the local regulatory component. Now suppose, as a consequence of an unspecified mutation, e_j is no longer sensitive to the *r*th global control signal. This alteration is reflected by assuming

$$U_r \equiv 1 \tag{3.5}$$

This implies the modified complete cybernetic variable that regulates that expression of e_j is given by

$$u_j = u_j^l(U_1U_2 \ldots 1 \ldots U_q) \tag{3.6}$$

Remark 3.4. The postulate above describes the removal of a global sensitivity connected with enzyme expression. The removal of the global control of enzyme activity follows by analogy.

The representation of the overexpression of existing pathway enzymes is slightly more involved because of the possibility of multiple expression sources.

Postulate 3.7. The overexpression of existing pathway enzymes is a substitutable process. Enzyme expression stemming from the genome and plasmid form a substitutable pair in the cybernetic sense, i.e., both routes compete for key cellular resources to synthesize a common product.

Postulate 3.7 implies the existence of a cybernetic control variable that regulates the allocation of key resources between possible expression sources. Because the native and plasmid expression routes are unaware of the existence of one another, the control action that regulates the choice of synthesis route is classified as a global signal.

Postulate 3.8. There exists a global cybernetic variable, denoted as $U_{e_j}^k$ ($k = g, p$) that regulates the allocation of critical resources to each individual expression route. The global control variable $U_{e_j}^k$ ($k = g, p$), where superscript(s) *g* and *p* denote genome and plasmid, respectively, reflects the substitutable nature of overexpression.

To derive the form of these control variables we pose a problem similar to the elementary convergent pathway. Specifically, there exist two routes for the synthesis of the common product, namely enzyme e_j. Both routes require critical cellular resources to proceed. Accordingly, we postulate that the local objective of enzymatic expression (barring any global interaction, such as nutritional state dependent synthesis) is the maximization of the level of key enzyme e_j subject to a constraint upon the amount of resources available for synthesis of e_j. These statements formally become the constrained optimization problem

$$\max\{\sum_i^n e_j^i(R_i)\} \quad \text{subject to} \quad g = \sum_i^n R_i = R \tag{3.7}$$

where R_i denotes the resource allocated to the *i*th route of synthesis of enzyme e_j, *n* denotes the total number of possible routes, and e_j^i denotes the level of enzyme e_j synthesized via the *i*th route. The optimality condition of the constrained maximization problem given by

$$\frac{\mathrm{d}e_j^1}{\mathrm{d}R_1} = \frac{\mathrm{d}e_j^2}{\mathrm{d}R_2} = \cdots = \frac{\mathrm{d}e_j^{n-1}}{\mathrm{d}R_{n-1}} = \frac{\mathrm{d}e_j^n}{\mathrm{d}R_n} \tag{3.8}$$

can be rearranged to yield the matching criteria

$$\frac{\mathrm{d}e_j^i}{\sum_{k=1}^n \mathrm{d}e_j^k} = \frac{\mathrm{d}R_i}{\sum_{k=1}^n \mathrm{d}R_k}, \qquad \forall i \tag{3.9}$$

Equation 3.9 is a restatement of the matching law criteria put forth by other cybernetic investigators. In short, it stipulates that the optimum allocation policy is one in which the fractional allocation of critical resources is equal to the fractional return on investment. If the allocation policy is implemented at every instant in time and takes place in time d*t*, the return on investment can be measured as the rate of enzyme synthesis. In the case in which the synthesis route is chromosomal, the rate of enzyme expression is given by the regulated rate, i.e.

$$r_{e_j} u_j, \qquad \forall j \tag{3.10}$$

whereas enzyme expression via the plasmid is measured as

$$r_{e_j}^p \tag{3.11}$$

whose form is dependent upon the promoter type. The inclusion of the regulated rate of chromosomal synthesis follows from the postulate that the enzyme synthesis routes are completely functional and independent. Within the present context, we are considering the competition between the chromosomal synthesis route and expression via a plasmid of the key enzyme e_j. Accordingly, it follows from eq 3.9 that the cybernetic variables that govern the allocation of critical resources for the synthesis of key enzyme e_j from the genome and plasmid are given functionally as

$$U^g_{e_j} = \frac{r_{e_j} u_j}{r_{e_j} u_j + r^p_{e_j}}, \qquad U^p_{e_j} = \frac{r^p_{e_j}}{r_{e_j} u_j + r^p_{e_j}} \tag{3.12}$$

where the superscripts g and p denote genome and plasmid, respectively. The overexpression of e_j is accounted for within the model equations by replacing the term $r_{e_j} u_j$ in the e_j balance with

$$r_{e_j} u_j U^g_{e_j} + r^p_{e_j} U^p_{e_j} \tag{3.13}$$

where the second synthesis rate stems from expression via the plasmid. Note the presence of the control variables $U^k_{e_j}$ ($k = g, p$). These act as a global control mechanism upon enzyme synthesis that regulates the source of synthesis of e_j. In simpler terms, $U^k_{e_j}$ is equivalent to the fraction of cellular resources allocated to e_j expression being directed to the kth synthesis source. For example, suppose the native route (in the absence of synthesis of e_j from the plasmid) is highly favored, i.e., $u_j = 1$. Furthermore, suppose the synthesis rate from the plasmid is much greater than chromosomal synthesis rate and under the control of an artificial promoter. Before addition of inducing agent, $r^p_{e_j} = 0$, which implies $U^p_{e_j} = 0$ and $U^g_{e_j} = 1$, i.e., the native expression route is the only active synthesis route for e_j and therefore it consumes all resources allocated to e_j expression. After addition of the inducing agent, because of the assumption $r^p_{e_j} \gg r_{e_j} u_j$, the majority of enzyme synthesis is from the plasmid, i.e., the plasmid route is faster than synthesis from the genome and, accordingly, it is allocated a larger portion of the critical resources earmarked for e_j expression. This allocation imbalance is reflected by the values of the control variables $U^p_{e_j}$ and $U^g_{e_j}$ and expression of e_j is steered away from the chromosomal route. The control variables $U^p_{e_j}$ and $U^g_{e_j}$ acts as a continuous global switch whose purpose is the selection of the particular synthesis route or combination of synthesis routes that yield the maximum level of key enzyme e_j, given a particular level of chemical induction agent.

Remark 3.5. Note that the assumption that the objective of enzyme expression is to maximize the level of e_j is consistent with previous cybernetic developments because the return of the chromosomal route of enzyme expression is measured as the regulated rate. This implies that the local control structure dictates what the exact meaning of "maximum" is relative to the objective of the pathway.

Remark 3.6. As postulated, enzyme expression from a plasmid is not under the direct control of the microorganism. Rather, expression is controlled via the characteristics of the plasmid borne promoter. In effect, this deregulates the expression of the plasmid borne gene. In other words, the microorganism possess no regulatory recourse that could halt expression from the plasmid (barring plasmid structural stability arguments.) As a result of overexpression, the makeup and activity of the network enzymatic machinery is predicted to be modulated with respect to the base pathway makeup because resource distribution is altered. Let us explore this idea further.

Suppose e_j catalyzes r^q_j and is a member of q elementary pathways. It follows from above that e_j competes for cellular resources from each of the q elementary resource pools. The overexpression e_j implies an artificially enhanced level of this key enzyme. Thus, all pathway fluxes catalyzed by this key enzyme are initially increased. Because the elementary allocation policy follows as a consequence of reaction rate, the level of resources allocated to the other enzymes which compete from the same pools must necessarily decrease in the short term. However, the shift in the relative levels of pathway enzymes is the impetus for shifts in the level of network intermediates, and hence, the network reaction rates. Thus, if changes in metabolite level can blunt the effect of overexpression (by altering r_j, the activity of e_j, or fluxes that feed e_j), the qualitative picture of resource allocation is not significantly altered by overexpression. If however, shifts in metabolite level cannot compensate for the artificially high level of flux through the jth network step, the qualitative picture of resource allocation to the elementary pathways may be substantially altered. These classes of system response, namely the initial shift in the jth flux and the ensuing shifts in metabolite level are termed the primary and secondary system responses, respectively.

4. Metabolic Control Analysis of Cybernetic Models

In the previous sections we have presented a formalism by which cybernetic models of metabolic networks can be constructed. Additionally, we formulated arguments as to how a cybernetic perspective can be employed to model metabolic networks, even in the presence of genetic modifications. The objective of this section is to take the last step in the development of a cybernetic metabolic engineering tool and discuss the application of metabolic control analysis (MCA) as described by Reder (1988) to cybernetic models.

MCA is not a modeling framework. In isolation its tenets provide little predictive insight. In the same vein, a cybernetic model, although capable of being predictive in scope, offers little guidance as to particular enzymes that must be overexpressed or deleted to achieve a rational design objective. Such concerns can be investigated by simulation of the model system; however, for complex networks this procedure can become tedious. Moreover, an intuitively based investigation of metabolic networks clearly is not applicable to complicated pathways. The application of MCA to a cybernetic model of the base network, in principle, affords a quantitative measurement of the factors that control the system interaction. MCA, in this respect, offers a systematic methodology for the computation of sensitivity coefficients upon which qualitative metabolic engineering strategies can be based. Once the strategies have been formulated, the appropriate genetic or environmental change can be simulated using the modified cybernetic model to determine the predicted outcome. Thus, using only base network experimental information, the cybernetic model system can be employed as a diagnostic and formulation tool to determine the genetic alterations as well as environmental changes required to achieve the specified objective.

4.1. Mathematical Issues. The application of MCA to a cybernetic model is not straightforward because of the functional form of the elementary cybernetic v variable. Specifically, this variable contains the max(...) function in the denominator which makes the elementary v variable discontinuous. However, this problem is easily overcome by examining the simulations of the wild-type network.

Previously, we defined the local cybernetic variable that governs the activity of the jth key enzyme, denoted by v^l_j, as

$$v_j^l \equiv \prod_q^n v_j^q, \quad j = 0, 1, ..., m \tag{4.1}$$

where the key enzyme e_j is a member of n elementary pathways in a network composed of m elementary units. The elementary variables v_j^q take the form(s)

$$v_k^q = \frac{r_k}{\max(r_0, r_1, ..., r_w)}, \quad k = 0, 1, 2, ..., w \tag{4.2}$$

for a substitutable elementary pathway and

$$v_k^q = \frac{r_k/p_{k+1}}{\max(r_0/p_1, r_2/p_2, ..., r_y/p_{y+1})}, \quad k = 0, 1, 2, ..., y \tag{4.3}$$

for a complementary elementary pathway where p_h denotes the hth metabolite. For a cybernetic system to be applicable for MCA, the max(...) functions must be eliminated. There are a number of ways to achieve this objective, for example it may be possible to approximate the discontinuous v variable, in the region around the discontinuity by a smooth function. However, for the moment, we consider a much less involved approach. Specifically, from the simulation data of the wild-type network, we are able to determine the fastest rate in an elementary pathway, which we denote as

$$r_k \quad \text{or} \quad \frac{r_k}{p_{k+1}} \tag{4.4}$$

depending upon the competitive environment in which e_k competes. Because of the functionality of the cybernetic v variable, this information is all that is required to decompose the regulatory structure. More exactly, given knowledge of the identity of the largest rate in a particular elementary pathway allows the elementary cybernetic variable(s) that govern enzyme activity to be decomposed as

$$v_1^q \simeq \frac{r_1}{r_k}, \quad v_2^q \simeq \frac{r_2}{r_k} \quad \cdots \quad v_k^q \simeq 1 \quad \cdots \quad v_w^q \simeq \frac{r_w}{r_k} \tag{4.5}$$

for a substitutable elementary pathway and

$$v_1^q \simeq \frac{r_1/p_2}{r_k/p_{k+1}}, \quad v_2^q \simeq \frac{r_2/p_3}{r_k/p_{k+1}} \quad \cdots \quad v_k^q \simeq 1 \quad \cdots \quad v_y^q \simeq \frac{r_y/p_{y+1}}{r_k/p_{k+1}} \tag{4.6}$$

for a complementary elementary pathway. This implies the local cybernetic variable that governs the activity of e_j is of the form

$$v_j^l \simeq \prod_q^w \frac{r_q}{r_k}, \qquad v_j^l \simeq \prod_q^y \frac{r_q/p_{q+1}}{r_k/p_{k+1}} \tag{4.7}$$

Note that these forms are the original cybernetic v variable set evaluated at steady state. Thus, we may substitute the set shown above (which are differentiable) for the original set and proceed with the sensitivity analysis.

4.2. Arbitrary Metabolic Network. Suppose the time evolution of a metabolic network is governed by the vector equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{N}v(\mathbf{x},\kappa), \qquad \mathbf{x} \in \mathbf{R}^m \; v \in \mathbf{R}^r \tag{4.8}$$

where $\mathbf{x}$ and v denote the metabolite vector and reaction rate vector, respectively, and $\mathbf{N}$ and k denote the $m \times r$ stoichiometry matrix and parameter vector, respectively. The metabolite vector $\mathbf{x}$ for a cybernetic model of a metabolic network consists of both metabolic intermediates and enzymatic species. Accordingly, in the general case, the metabolite vector takes the banded form

$$\mathbf{x} \equiv \{\mathbf{x}^M, \mathbf{x}^E\}^T, \qquad \mathbf{x}^M \in \mathbf{R}^p \; \mathbf{x}^E \in \mathbf{R}^q \tag{4.9}$$

where $\mathbf{x}^M$ denotes the p-dimensional vector of metabolic intermediates and $\mathbf{x}^E$ denotes the q-dimensional enzymatic state variable vector. The structure of the reaction rate vector, denoted as v, is the biggest point of departure from the traditional application of MCA to model systems. This is primarily true because the system reaction rates are modified by the cybernetic variables. We postulate that, in this context, the sensitivity coefficients computed from a cybernetic model truly reflect a representation of metabolic regulation in addition to the kinetic interactions present within the system. This feature is unique to cybernetic model systems because no other framework affords a systematic representation of metabolic regulation. The structure of the reaction rate vector v is banded and given by

$$v \equiv \{\mathbf{r}_{cat}, \mathbf{r}_{d,x^M}, \mathbf{r}_{e,x^E}, \mathbf{r}_{d,x^E}, \mathbf{r}_{b,x^E}\}^T \tag{4.10}$$

where $\mathbf{r}_{cat}$ and $\mathbf{r}_{d,x^M}$ denote the vector of catalytic rates (including environmental input/output fluxes) and the vector of metabolite dilution terms. The vector(s)

$$\mathbf{r}_{e,x^E}, \; \mathbf{r}_{d,x^E}, \; \mathbf{r}_{b,x^E}$$

denote the enzyme expression vector, the vector of enzymatic dilution/degradation terms and the background/constitutive enzyme synthesis vector, respectively. Functionally, the catalytic reaction rate vector is given by the set

$$\mathbf{r}_{cat} \equiv \{\{r_q v_q\}_q, \{r_{T,k}\}_k\} \tag{4.11}$$

where r_q denotes the qth specific reaction rate and v_q denotes the qth complete cybernetic variable whose functional form is approximated by the steady state component. The second term denotes the set of transport processes that describe interaction with the external environment, for example, diffusion of metabolic intermediates to the abiotic phase where the term $r_{T,k}$ denotes the kth specific rate of metabolite transport. The vector $\mathbf{r}_{d,x^M}$ denotes the vector of metabolite dilution due to growth and is given functionally by the set

$$\mathbf{r}_{d,x^M} \equiv \{r_g x_q^M\}_q \tag{4.12}$$

where r_g denotes the specific growth rate and x_q^M denotes the qth member of the metabolic intermediate vector. The enzymatic expression vector is given by the set

$$\mathbf{r}_{e,x^E} \equiv \{r_{e_q} u_q\}_q \tag{4.13}$$

where r_{e_q} denotes the specific rate of expression of the qth enzyme and u_q denotes the complete cybernetic variable that governs the synthesis of e_q. The enzymatic dilution/degradation vector and the constitutive enzyme

synthesis vector are given functionally by the set(s)

$$\mathbf{r}_{\mathrm{d,x^E}} \equiv \{(r_g + \beta)e_q\}_q \qquad \mathbf{r}_{\mathrm{b,x^E}} \equiv \{r^*_{e_q}\}_q \qquad (4.14)$$

where β denotes the rate constant governing the first-order decay of the qth enzyme and $r^*_{e_q}$ denotes the specific rate of constitutive enzyme synthesis of e_q.

The stoichiometric matrix **N** has special structure depending upon the topological nature of the pathway under consideration. In particular, if certain conservation relationships exist, then certain rows of the stoichiometry matrix can be combined to yield null rows. This is a structural property of the system that is independent of the choice of κ. The number of independent rows of **N** is by definition equal to the rank of the matrix, denoted as m_o. Denote the matrix that consists of the m_o basis rows as the reduced matrix, denoted by $\mathbf{N}_R$. Then it is possible to decompose the stoichiometry matrix **N** as the matrix product

$$\mathbf{N} = \mathbf{L}\mathbf{N}_R \qquad (4.15)$$

where the $m \times m_o$ matrix **L** is termed the link matrix. Note that if **N** has full rank then the link matrix **L** is by definition the $m \times m$ identity matrix.

Let (σ^0,κ^0) denote an asymptotically stable steady state of the network, where σ^0 and κ^0 denote the steady state metabolite vector and parameter vector, respectively. The simple steady state control matrix, denoted as Γ, and the simple steady state flux control matrix, denoted as **C** take the forms

$$\Gamma = -\mathbf{L}(\mathbf{N}_R\mathbf{D}_x v\mathbf{L})^{-1}\mathbf{N}_R \qquad (4.16)$$

$$\mathbf{C} = \mathbf{I}d_r - \mathbf{D}_x v\mathbf{L}(\mathbf{N}_R\mathbf{D}_x v\mathbf{L})^{-1}\mathbf{N}_R \qquad (4.17)$$

where $(\mathbf{N}_R\mathbf{D}_x v\mathbf{L})^{-1}$ denotes the inverse of the Jacobian evaluated at (σ^0,κ^0), which by assumption of asymptotic stability is guaranteed to exist. The elements of the matrix Γ, given by $\gamma_{i,j}$, denote the sensitivity of the ith steady state metabolite level with respect to changes in the jth reaction rate and are given functionally by

$$\gamma_{i,j} = \frac{\partial \sigma_i}{\partial \nu_j}, \quad i = 1, ..., m, \quad j = 1, ..., r \qquad (4.18)$$

The elements of the matrix **C**, given by $c_{i,j}$, denote the sensitivity of the ith steady state flux to changes in the jth reaction rate and are given functionally by

$$c_{i,j} = \frac{\partial J_i}{\partial \nu_j}, \quad i = 1, ..., r, \quad j = 1, ..., r \qquad (4.19)$$

where

$$J_i \equiv \nu_i(\sigma^0,\kappa^0) \qquad (4.20)$$

Remark 4.1. The sensitivity coefficients computed using a cybernetic model of a metabolic network clearly describe the role of metabolic regulation becuase of the inclusion of the cyberneric variables. In the present context we employ the MCA machinery to gain a qualitative understanding of the regulatory fingerprint of the network. An alternative approach, developed in Varner and Ramkrishna (1999a,b) for the visualization of the qualitative properties of a cybernetic model relies upon the classification of the competitive structure that each key enzyme experiences. From this knowledge, the qualitative properties of network evolution can be determined. Although, for small-scale problems, this approach is tractable, the application of such tools may be burdensome for an arbitrary network. However, such an understanding can greatly aid in the visualization of the ramifications of resource allocation structure.

5. Conclusions

We have presented an extension of the cybernetic framework that allows the construction and analysis of cybernetic models of arbitrarily complex metabolic networks. This approach is centered around the formulation of the cybernetic variables using a modular approach in which the regulation of network follows as a consequence of the network topological. Additionally, we discussed the theoretical basis of the application of a cybernetic methodology to networks that have been altered genetically. We postulate, in this instance, that genetic alteration of a metabolic network does nothing to alter the base objectives of the network; rather, modification simply expands or contracts the set of physiological alternatives open to the network. Thus, it follows that a modified network is still taken to be optimal with respect to resource allocation, albeit to the expanded or contracted set of physiological alternatives. The cybernetic model of the metabolic network can then be employed in a predictive design capacity to formulate alteration strategies that are cognizant of the network response to genetic alteration, thereby compensating or even taking advantage of the manner in which the system responds to perturbation. In this regard the systematic computation of sensitivity coefficients afforded by MCA can be used to formulate qualitative strategies for the design of genetic alterations. These strategies can then be simulated using a modified version of the cybernetic model to determine their outcome.

In the subsequent papers of this series, as well as other developments, the framework presented here has been applied to a number of example systems. The second paper of this series is a qualitative examination of branch point control architectures. Specifically, we present the cybernetic representation of the flexible and rigid nodal control architectures discussed by Stephanopoulos and Vallino (1991) as well as consider, on a finer level, the manner in which a cybernetic model predicts network response and, in particular, the role played by resource allocation. This development provides necessary intellectual precursors for the ensuing modeling endeavors. Furthermore, the study of simple branch point structures affords a less involved context (as compared to other cybernetic network models that follow) for the discussion of the deeper significance of the extension of cybernetic concepts to genetically altered networks. Specifically, we contrast the conventional wisdom that enzyme level and activity are fixed parametric quantities with the cybernetic vision that both activity as well as enzyme level are rationally modulated as a result of network control action in the face of genetic alteration. It is our position that this modulation or response by the network is one of the most fundamental elements that is neglected by traditional approaches and, as such, is one of the key features that separates this development from more traditional modeling perspectives. Moreover, we maintain, conceptually, as well as from successful applications of the framework presented here (see for example Varner and Ramkrishna, 1998a; Varner, 1998, Varner et al., 1999), that it is the absence of this key feature that

frustrates the rational mathematical design of metabolic networks using traditional modeling methodologies following from a reaction engineering view of metabolic pathways.

Notation

u_k^j elementary cybernetic variable governing the expression of the kth enzyme in the jth elementary pathway (isolated transcriptional/translation efficiency)

v_k^j elementary cybernetic variable governing the activity of the kth enzyme in the jth elementary pathway (isolated scaled specific activity)

u_k^l local elementary cybernetic variable governing the expression of the kth enzyme (fusion of isolated transcriptional/translation efficiencies)

v_k^l local elementary cybernetic variable governing the activity of the kth enzyme (fusion of isolated scaled specific activities)

$U_{e_k}^j$ global regulatory signal governing the choice between substitutable expression sources for the kth enzyme

V_k^j global regulatory signal influencing the activity of the kth enzyme in the jth elementary pathway

u_k^g global cybernetic variable governing expression of the kth enzyme (product of all global regulatory expression signals)

v_k^g global cybernetic variable governing specific activity of the kth enzyme (product of all global regulatory activity signals)

u_k complete cybernetic variable governing the expression of the kth enzyme (complete transcriptional/translational efficiency reflecting local and global signals)

v_k complete cybernetic variable governing the activity of the kth enzyme (complete scaled specific activity reflecting local and global signals)

r_k^j kth specific flux through the jth elementary pathway (g/gdw-h)

r_{e_k} specific rate of expression of the kth enzyme (g/gdw-h)

$r_{e_k}^p$ specific rate of expression of the kth enzyme from the plasmid (g/gdw-h)

$r_{e_k}^*$ specific rate of constitutive expression of the kth enzyme

e_k specific level of the kth key enzyme (g/gdw)

p_k specific level of the kth metabolite (g/gdw)

N $m \times r$ stoichiometry matrix

$v(\mathbf{x},\kappa)$ r-dimensional vector of specific rates (including enzyme synthesis and dilution due to growth)

X m-dimensional vector of system species (metabolites and enzymes)

κ p-dimensional vector of system parameters

Appendix: Examination of Postulate(s) 2.7 and 2.9

In what follows we examine the postulates 2.7 and 2.9 in a more detailed light. We begin by discussing the implications of the allocation sharing formalism postulated for overlapping key enzymes. We then examine postulate 2.9, which governs the objective function of a hybrid network.

6.1. Justification of Postulate 2.7. If a key enzyme lies in multiple elementary pathways, then this key enzyme competes for resource from multiple pools. The modular formulation has as its base the elementary pathway concept. Thus, when an enzyme lies in multiple allocation regions, in other words, is a member of more than one elementary pathway, the control of the synthesis and activity of this key enzyme has in a larger sense a duality of purpose. At the elementary level, when considered in total isolation, the overlapping enzymes compete for cellular resources to achieve the individual elementary objectives. However, if viewed from a more macroscopic perspective, the regulation of an overlapping key enzyme must be sensitive to the objective of every elementary pathway that it is a member of. Thus, the totality of the local regulatory structure must be composed of the individual elementary regulatory influences. This follows as a consequence of assuming the base library to be isolated elementary pathways.

The form of the local variable, i.e., the product of the elementary variables, follows as a consequence assuming the cybernetic "direction" of a metabolic network is elementary in scope. In other words, the cybernetic aspects of a network constructed using an elementary modular approach stem from the individual isolated elementary allocation problems, not from a constrained optimization of the network. It follows that enzymes which can be written as a member of multiple elementary pathways in an arbitrary network realization feel multiple elementary influences. Thus, the purpose of the local cybernetic variable is to meld these elementary influences into a single local network signal that directs the synthesis and activity of the key enzyme. The postulated functional form is one in which the individual elementary regulatory signals are reflected and allowed to interact, while still maintaining the original postulate that all cybernetic variables must be less than or equal to unity.

6.2. Proof of Postulate 2.9: Intersection Case. To prove postulate 2.9, we must show that the objective of an arbitrary metabolic network is the superimposed goal of the elementary components subject to the elementary resource constraints.

Let an arbitrary metabolic network be composed of n intersecting elementary pathways. These n elementary units have the objective set

$$\{\max(f_j(\mathbf{X}_j))\}_{j=1}^n \tag{6.1}$$

subject to the resource constraint set

$$\{g_1(\mathbf{X}_j) = G_j\}_{j=1}^n \tag{6.2}$$

where $g_k(\mathbf{X}_k) = G_k$ denotes the resource constraint corresponding to the kth elementary objective function and $\mathbf{X}_k$ denotes the pool of resource earmarked for operation of the kth elementary pathway. Note because of the assumption of intersecting elementary pathways that we have no resource allocation overlap, i.e., no key enzyme is a member of more that one elementary pathway. Traditionally, cybernetic investigators assume the resource constrain to take the form

$$g_k \equiv \sum_{q=1}^{h} X_k^q = G_k \tag{6.3}$$

where h denotes the number of key enzymes which are

members of the *k*th elementary pathway. This functionality for the resource constraint implies the optimality condition is given by

$$\frac{df_j}{dX_j^1} = \frac{df_j}{dX_j^2} = \cdots = \frac{df_j}{dX_j^n}, \quad \forall j \tag{6.4}$$

Before we continue some observations are in order:

Remark 6.1. Because the system is composed of intersecting elementary pathways, the *j*th objective is only a function of the resource pool earmarked for operation of the *j*th elementary pathway.

Remark 6.2. Notice the optimization problem of the *j*th elementary pathway is with respect to the *j*th elementary resource pool only. Thus, even if the objective function of the *j*th elementary pathway were a function of resources from the *k*th pool ($k \neq j$), this input would not be reflected, given the current constrained optimization structure.

Remark 6.3. Equation 6.4, after some algebraic simplification and the assumption of a functional form for f_j, becomes the matching criteria from which the cybernetic *u* variable is derived.

Equation 6.4 is valid for every value of *j*, i.e., all elementary pathways that compose the network. Thus, the competition for key cellular resources occurs at the elementary level, as opposed to the local level. Accordingly, the local regulation of the network stems from the *n* separate elementary allocation problems, which implies that the objective of the network is the superimposed objective of the individual allocation problems. Let us consider an example to illustrate this concept.

Example 6.1. Consider the structure shown in Figure 4.

We assume, for the moment, that this framework consists of three intersecting elementary linear pathways (although degenerate representations using different length linear pathways are possible). To illustrate the use of the postulates, postulate 2.9 in particular, we derive the local cybernetic regulation for this network from first principles. We then rederive the cybernetic variables using the modular approach and compare the results of each case. For the sake of simplicity, we present only the derivation of the functional form of the cybernetic *u* variable because the cybernetic *v* variable follows by analogy.

6.2.1. First Principles Derivation. The network shown in Figure 4 is composed into three elementary linear pathways, labeled *a*, *b*, and *c*. Elementary linear pathway *a* consists of the conversion of intermediate p_q^a into p_n ($n-q$ steps), whereas elementary pathway *b* consists of the conversion of p_0 to p_n (n steps). Note the intersection of elementary pathway(s) *a* and *b* at the intermediate p_n. The final elementary linear pathway, denoted as *c*, consists of the conversion of intermediate p_n into the product p_{n+k} (k steps). The functional form of the cybernetic variable that regulates the expression of the *m*th key enzyme belonging to *j*th elementary pathway follows from the solution of the following constrained optimization problem(s): From postulate 2.9, let the objective of the network be the maximization of p_n (a degenerate objective) and p_{n+k} subject to a constraint

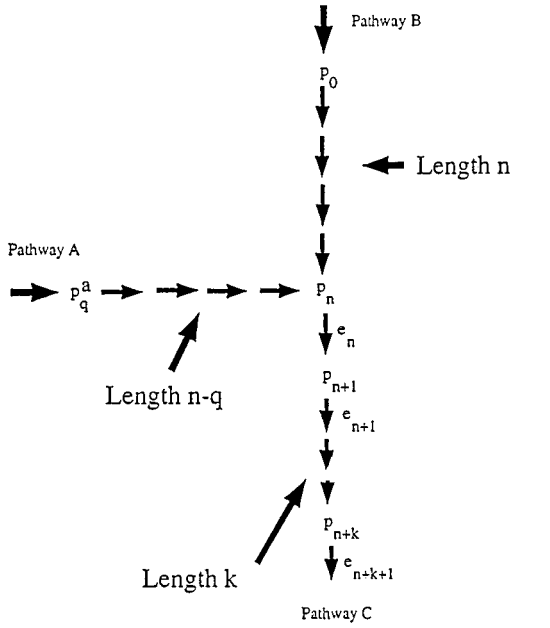

Figure 4. Example system: Hybrid network composed of three intersecting linear pathways. Pathways *a* and *b* intersect at the metabolite p_n, which is then consumed transformed via elementary pathway *c* to form metabolite p_{n+k}. Note that degenerate representations are possible.

upon the level of cellular resources earmarked for elementary pathway operation, i.e.

$$\max(p_n(R_{n-1}^a, p_{n-1}^a(R_{n-2}^a, p_{n-2}^a(\ldots(R_q^a)\ldots)))), \quad \text{subject to} \quad g^a = \sum_{w=q}^{n-q} R_w^a$$

$$\max(p_n(R_{n-1}^b, p_{n-1}^b(R_{n-2}^b, p_{n-2}^b(\ldots(R_q^b)\ldots)))), \quad \text{subject to} \quad g^b = \sum_{w=0}^{n} R_w^b$$

$$\max(p_{n+k}(R_{n+k-1}^c, p_{n+k-1}^c(R_{n+k-2}^c, p_{n+k-2}^c(\ldots(R_n^c)\ldots)))), \quad \text{subject to} \quad g^c = \sum_{w=n}^{k} R_w^c \tag{6.5}$$

In simpler terms, this problem implies the existence of three local objectives, subject to three independent resource constraints. In other words, the objective of the network, because the elementary pathways are intersecting, is the totality of elementary objectives subject to the elementary resource constraints. The optimality criteria given by

$$\frac{dp_n^a}{dR_{n-1}^a} = \frac{dp_n^a}{dp_{n-1}^a}\frac{dp_{n-1}^a}{dR_{n-2}^a} = \cdots = \left(\prod_{k=q+2}^{n}\frac{dp_k^a}{dp_{k-1}^a}\right)\frac{dp_{q+1}^a}{dR_q^a}$$

$$\frac{dp_n^b}{dR_{n-1}^b} = \frac{dp_n^b}{dp_{n-1}^b}\frac{dp_{n-1}^b}{dR_{n-2}^b} = \cdots = \left(\prod_{k=2}^{n}\frac{dp_k^b}{dp_{k-1}^b}\right)\frac{dp_1}{dp_0}$$

$$\frac{dp_{n+K}^c}{dR_{n+k-1}^c} = \frac{dp_{n+k}^c}{dp_{n+k-1}^c}\frac{dp_{n+k-1}^c}{dR_{n+k-2}^c} = \cdots = \left(\prod_{k=n+2}^{n+k}\frac{dp_k^c}{dp_{k-1}^c}\right)$$

can be rearranged to yield the matching criteria

$$\frac{dp_j^a}{dp_j^a + \sum_{x=q,j}^{n-1} dp_x^a} = \frac{dR_j^a}{dR_j^a + \sum_{x=q,j}^{n-1} dR_x^a}$$

$$\frac{dp_j^b}{dp_j^b + \sum_{x=0,j}^{n-1} dp_x^b} = \frac{dR_j^b}{dR_j^b + \sum_{x=0,j}^{n-1} dR_x^b}$$

$$\frac{dp_j^c}{dp_j^c + \sum_{x=n,j}^{n+k-1} dp_x^c} = \frac{dR_j^c}{dR_j^c + \sum_{x=n,j}^{n+k-1} dR_x^c}$$

If we assume the allocation policy described by eq set 6.6 is implemented at every instant in time and allocation takes place in time dt, then the cybernetic variable that regulates the expression of the mth key enzyme belonging to the jth elementary pathway takes the form

$$u_m^j = \frac{r_m}{r_m + \sum_{x=a,m}^{l} r_x} \tag{6.6}$$

where the summation limit l denotes the length of the jth elementary pathway. Because all elementary pathways intersect the mth key enzyme belonging the jth elementary pathway can only receive critical resources from the jth pool. This implies the local regulatory picture is equivalent to the elementary allocation pattern.

6.2.2. Modular Approach. Let us now employ the modular approach put forth to derive the functional forms of the cybernetic variables that govern the allocation of critical resources. Using the same decomposition as the derivation above, the network is postulated to consist of three elementary linear pathways, i.e., the elementary pathways a, b, and c. The cybernetic regulation that controls enzyme expression for the elementary linear pathway, which is a substitutable competitive structure, was derived by Straight and Ramkrishna (1994), assuming the objective function to be the maximization of end product subject to a resource constraint. Accordingly, the cybernetic variables which control enzyme expression take the form

$$u_m^j = \frac{r_m}{r_m + \sum_{x=a,m}^{l} r_x} \tag{6.7}$$

where u_m^j ($j = a, b, c$) denotes the cybernetic variable that governs the expression of the mth key enzyme belonging to the jth linear pathway. Note that eqs 6.6 and 6.7 are identical.

6.3. Proof of Postulate 2.9: Overlapping Case. In the case of overlapping pathways, the arguments follow much the same. However, in this instance, we must invoke the postulate 2.7 to account for allocation to a single enzyme from multiple pools. To illustrate the application of postulate 2.9 in the case of overlapping pathways let us reexamine the network shown in Figure 4. Another realization of the network is two overlapping linear pathways, starting from p_q^a and p_0 terminating at p_{n+k}. The elements from p_n to p_{n+k} are common to both linear pathways, i.e., have the possibility of receiving cellular resources from both elementary pools. We approach this example in much the same fashion as the previous case. First, we derive the appropriate forms of the cybernetic variables from first principles, illustrating the cybernetic basis of the modular postulates. We then, for the sake of comparison, apply the modular construction methodology to the overlapping network and compare the results with the first principles derivation.

6.3.1. First Principle Derivation. The network realization is composed of two overlapping linear pathways with the elements p_n to p_{n+k} being common to each pathway. This implies that the key enzymes e_n through e_{n+k-1} can receive resources from both pools associated with the elementary pathways. The objective of a linear pathway is the maximization of the end product subject to an elementary resource constraint. Accordingly, the objective of the network shown in Figure 4 is the maximization of p_{n+k} from each branch subject to the resource constraints associated with elementary linear pathway(s) a and b. In a formal sense we have the constrained optimization problem for the jth branch

$$\max\{p_{n+k}^j(R_{n+k-1}^j, R_{n+k-1}^q, p_{n+k-1}(R_{n+k-2}^j, R_{n+k-2}^q, \ldots, p_{n+1}(R_n^j, R_n^q, p_n(R_{n-1}^j, \ldots, p_s(R_{s-1}^j))))))\}$$

subject to

$$g_j = \sum_h^{l(j)} R_h^j \tag{6.8}$$

where $q, j = a, b$ ($q \neq j$) and $l(j)$ denotes the length of jth elementary pathway. Notice that the shared elements are postulated to receive resources from both elementary resource pools; however, before p_n, the key enzymes are members of only a single elementary pathway, which implies that they receive resources only from the jth elementary resource pool. In fact this distinction matters little because of the nature of the elementary resource constraints. As discussed in the justification of postulate 2.7 and remark 2.10 the assumption of an elementary isolated pathway implies the resource pools are independent, i.e., we are optimizing the allocation of critical resources with respect to the jth pool independently of the remaining resource pools associated with the network. In other words, we evaluate the jth objective with respect to the members of the jth resource pool only. This implies that, despite the joint resource dependence of intermediates p_n to p_{n+k}, the allocation of resources from the jth pool is optimized with respect to only the objective of the jth elementary pathway. It follows that the jth optimality condition

$$\frac{dp_{n+k}^j}{dR_{n+k-1}^j} = \frac{dp_{n+k}}{dp_{n+k-1}} \frac{dp_{n+k-1}}{dR_{n+k-2}} = \cdots = \frac{dp_{n+k}}{dp_{n+k-1}} \frac{dp_{n+k-1}}{dp_{n+k-2}} \cdots \frac{dp_{s+2}}{dp_{s+1}} \frac{dp_{s+1}}{dR_s} \tag{6.9}$$

can be rearranged to yield the matching criteria

$$u_m^j = \frac{dp_m^j}{dp_m^j + \sum_{q=s,m} dp_q^j} = \frac{dR_m^j}{dR_m^j + \sum_{q=s,m} dR_q}, \quad j = a, b \tag{6.10}$$

where the lower summation limit s denotes the start index of the elementary pathway. If we assume that the allocation policy is implemented at every instant in time and takes place on the time scale of dt, the return on investment from the mth step of the ith elementary pathway can be measured as the reaction rate. It follows that the elementary cybernetic variable that governs the allocation of critical resources from the kth elementary pool to the synthesis of the mth key enzyme, denoted as u_m^k, is given by

$$u_m^k = \frac{r_m^k}{r_m^k + \sum_{q=s,m}^{l(k)} r_q^k} \tag{6.11}$$

Equation 6.11 governs the allocation of critical resources from the elementary resources pools. In the cases in which the key enzyme(s) are members of only a single elementary pathway, the local regulatory element is governed by eq 6.11. If however the key enzyme(s) can be written as a member of multiple elementary pathways, it follows that the expression machinery can receive critical resources from multiple resource pools. In this instance we invoke postulte 2.7 to account for the possibility of multiple allocation. In other words, the local cybernetic variable that regulates the expression of enzymes e_x^a ($x = q, q + 1, ..., n - 1$) and e_x^b ($x = 0, 1, 2, ..., n - 1$) is given soley by the elementary component

$$u_{xj}^l = u_x^j = \frac{r_x^j}{r_x^j + \sum_{q=s,x}^{l(j)} r_q^j}, \quad j = a, b \tag{6.12}$$

The local cybernetic variable that governs the expression of key enzymes e_x ($x = n, n + 1, ..., n + k - 1$) follows from postulate 2.7 and is given functionally as

$$u_x^l = u_x^a\, u_x^b, \quad x = n, n + 1, ..., n + k - 1 \tag{6.13}$$

6.3.2. Modular Approach. We derive the cybernetic regulation for the overlapping realization of the network shown in Figure 4. For the sake of convenience, we show only the derivation of the cybernetic u variable because the functional form of the v variable follows by analogy.

The network consists of two overlapping linear pathways. The key enzymes e_n to e_{n+k-1} are common to both pathways, which implies that they can receive resources from both elementary pools associated with the linear elementary pathways. The cybernetic variables that governs the allocation of critical resources to the jth key enzyme which is a member of the kth elementary pathway is of the form

$$u_j^k = \frac{r_j}{r_j + \sum_{x=d,j}^{l} r_x}, \quad k = a, b \tag{6.14}$$

The key enzymes e_y^a ($y = q, q + 1, q + 2, ..., n - 1$) and e_s^b ($s = 0, 1, ..., n - 1$) are members of isolated linear elementary pathway. It follows that the local cybernetic variable that governs the expression of enzymes e_y^a ($y = q, q + 1, q + 2, ..., n - 1$) and e_s ($s = 0, 1, ..., n - 1$) is identical to the elementary component and is given functionally as

$$u_{ya}^l = u_y^a, \quad y = q, q + 1, q + 2, ..., n - 1 \qquad u_{sb}^l = u_s^b, \quad s = 0, 1, 2, ..., n - 1 \tag{6.15}$$

The key enzyme(s) e_j ($j = n, n + 1, ..., n + k - 1$) are members of both linear pathways a and b. Accordingly, from postulate 2.7, these key enzymes have the possibility of receving cellular resources from both elementary resource pools. It follows that the local cybernetic variable that governs the allocation of critical resources for the synthesis of e_j ($j = n, n + 1, ..., n + k - 1$) is the product of elementary cybernetic variables, i.e.

$$u_j^l = u_j^a\, u_j^b, \quad j = n, n + 1, ..., n + k - 1 \tag{6.16}$$

References and Notes

Bailey, J. E. Towards a science of metabolic engineering. *Science* **1991**, *252*, 1668−1675.

Baloo, S.; Ramkrishna, D. Metabolic Regulation in Bacterial Continuous Cultures- I. *Biotechnol. Bioeng.* **1991**, *20*, 1337−1352.

Hatzimanikatis, V.; Floudas, C. A.; Bailey, J. E. Optimization of regulatory architectures in metabolic reaction engineering. *Biotechnol. Bioeng.* **1996a**, *52*, 485−500.

Hatzimanikatis, V.; Floudas, C. A.; Bailey, J. E. Analysis and design of metabolic networks via mixed-integer linear optimization. *AIChE J.* **1996b**, *42*, 1277−1292.

Hatzimanikatis, V.; Emmerling, M.; Sauer, U.; Bailey, J. E. Application of Mathematical Tools for Metabolic Design of Microbial Ethanol Production. *Biotechnol. Bioeng.* **1998**, *57*, 154−161.

Kacser, H.; Burns, J. A. The control of flux. In *Control of Biological Processes*; Davies, D. D., Ed.; Cambridge: Cambridge University Press: Cambridge, U.K., 1973; pp 65−104.

Kompala, D.; Jansen, N.; Tsao, G.; Ramkrishna, D. Investigation of Bacterial Growth on Mixed Substrates: Experimental Evaluation of Cybernetic Models. *Biotechnol. Bioeng.* **1986**, *28*, 1044−1055.

Ramakrishna, P.; Ramkrishna, D.; Konopka, A. Cybernetic Modeling of Mixed Substitutable Substrate Environments. Simultaneous and Preferential Utilization. *Biotechnol. Bioneg.* **1996**, *52*, 144−154.

Reder, C. Metabolic Control Theory: A Structural Approach. *J. Theor. Biol.* **1988**, *135*, 175−201.

Savageau, M. *Biochemical Systems Analysis*; Addison-Wesley: Reading, MA, 1976.

Shuler, M. *Bioprocess Engineering: Basic Concepts*; Prentice Hall: Englewood Cliffs, NJ, 1992.

Straight, J.; Ramkrishna, D. Cybernetic modeling and regulation of metabolic pathways. Growth on complimentary nutrients. *Biotechnol. Prog.* **1994**, *10*, 574−587.

Vallino, J.; Stephanopoulos, G. Network Rigidity and Metabolic Engineering in Metabolite Overproduction. *Science* **1991**, *252*, 1675−1681.

Varner, J. Metabolic Engineering from A Cybernetic Perspective. Ph.D. Thesis, School of Chemical Engineering, Purdue University, 1997.

Varner, J.; Ramkrishna, D.; Metabolic Engineering from a Cybernetic Perspective. The Aspartate Family of Amino Acids. *Met. Eng.* **1998**, *1*, 88−116.

Varner, J.; Ramkrishna, D.; Nonlinear Analysis of Cybernetic Models: Guidelines for Model Formulation. *J. Biotechnol.* **1999**, in press.

Varner, J.; Ramkrishna, D.; Bailey, J. E. A Self-Optimizing Adaptive Model of Glucose Catabolism in *Escherichia Coli*. Prediction of Network Response to Amplification of Key Glycolytic Enzymes. *Biotechnol. Prog.* **1999**, in press.

Accepted February 4, 1999.

BP990017P