Review

Nutrient-response: a long random walk through metabolic pathways

Arthur R. Schulz

Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, USA

An equation has been derived to describe the relationship between response and nutrient intake. This derivation has been based on stochastic principles in which the metabolites involved were treated as discrete states. Derivation involved calculation of the fractional occupancy of each state during a continuous random walk through the graph, which represented the metabolic pathway. The resultant equation is a rational polynomial. A rational polynomial can accommodate the various shapes of nutrient-response curves that have been reported.

Keywords: nutrient-response; metabolic pathways; probability theory

Introduction

Monod[1] demonstrated that bacterial growth in the exponential phase is a rectangular hyperbolic function of the amount of the limiting nutrient. In this regard, the nutrient-response relationship appeared to be analogous to the effect of a substrate on the rate of an enzyme-catalyzed reaction, but Monod advised against the conclusion that a complex biological process, such as growth, might be controlled by the rate of a single enzymic step. Morgan et al.[2] suggested that the response of higher organisms to a nutrient is a rational function of nutrient intake provided that some of the parameters of the rational polynomial were assigned a value of zero. Schulz[3] proposed that the nutrient-response relationship is described better by a general rational polynomial, i.e., a ratio of polynomials, without arbitrary assumptions concerning the magnitude of any of the parameters. Further analysis of the rational polynomial that describes a nutrient-response curve provides a means of partitioning the nutrient into metabolic pathways that are defined mathematically.[4]

It is not surprising that nutrient-response curves might be described by a rational polynomial, for the rates of most biological reactions are described by rational polynomials.[5] In many of these instances, such as the kinetics of enzyme action, the rate equation can be derived on the basis of the mechanism.[6] It has not been possible to derive an equation from comparable considerations for the rate of response of an animal to a given nutrient. However, Bartholomay[7] developed an alternative derivation of the rate equation for an enzyme-catalyzed reaction based on stochastic principles. This approach to enzyme kinetics has been expanded upon by Ninio[8] and by Mazur.[9] Hill has shown, in an excellent series of recent publications, that the application of stochastic principles to enzyme action and biochemical cycle kinetics is facilitated by representing the different enzyme states as directed graphs.[10-15]

The purpose of this publication is to report the utilization of Hill's graphical implementation of stochastic principles for the derivation of a general type of equation to describe nutrient-response relationships. This approach is similar to compartmental analysis[16] except that the present approach will employ graph theory to facilitate the derivation. Compartmental analysis can give rise to the same type of general equation, i.e., a rational polynomial.[17] However, as will be discussed in a later section, the treatment developed here is dedicated to utilize the data generally accessible in nutritional experiments.

Mathematical treatment

Derivation of rate equations for steady state enzyme kinetics based on stochastic principles has been

This work has been conducted with financial assistance from the Grace M. Showalter Residuary Trust and by a grant-in-aid from the Lueker Family Investments.
Address reprint requests to Dr. Arthur R. Schulz at the Department of Biochemistry and Molecular Biology, Indiana University, School of Medicine, Indianapolis, IN 46202-5122, USA.
Received July 31, 1991; accepted November 27, 1991.

accomplished.[6-8] In this discussion, the system under consideration will be the states accessible to a nutrient during its journey through the metabolic pathways, which culminate in the measured response. The stochastic approach to this problem is facilitated by representing the accessible states as vertices on a graph.[9-13] In the context of this discussion, the different states can be different metabolites along the metabolic pathway, but some of the states could represent the same metabolite in different pools. The vertices are connected by edges (lines) that represent the allowable transitions between states. If arrows are added to the edges, the graph becomes a directed graph (digraph), and the indicated transition is assigned a weight, $\alpha_{s,d}$, which is proportional to the probability of the indicated transition. The first subscript to the α identifies the source state, and the second subscript identifies the destination state.

It is not the intention in this discussion to treat the transitions that occur in a metabolic pathway at a fundamental molecular level. Rather, the transitions will be described in terms of rate constants and concentrations. The term rate constant, as employed here, does not imply a true constant, but rather the rate of transition at a given nutrient level. Thus, the rate may reflect steady state enzyme kinetics, first order, or some other order of kinetics.

In light of the foregoing discussion, a metabolic pathway can be viewed as a series of interconnections of discrete states, any one of which might be visited on a journey (walk) through the pathway. To visualize such a walk, it is necessary to identify a discrete starting state and a discrete final state. The starting state could be the nutrient or a compound whose concentration in the body is directly proportional to the level of nutrient intake. The final state, or states, is where the walk through the pathway ends, and it is termed the destination state(s). In his development of random walks on diagrams, Hill[9,13] employed the term absorption rather than termination. One could calculate the probability that any specific state might be visited at a specific time during a very large number of random walks through the pathway. Alternately, if transition to a termination state were to result in an instantaneous transfer to the starting state, one could obtain a similar result by calculating the probability that a given state would be occupied at a given time during a single continuous random walk.[10-13] This transition from the termination state to the starting state is introduced only as a conceptual aid to visualizing a single continuous random walk through the digraph. It should not be misconstrued as a biochemical conversion of the end product directly to the starting metabolite. Rather, the metabolic implication is that this treatment applies to the total synthesis of the end product, and this is measured by comparison of the response of animals on a given nutritional regime compared to animals on a diet devoid of the nutrient in question.

Figure 1 represents a metabolic pathway in which state 1 is the starting state and states 4 and 5 are termination states. In this discussion, it is assumed that the

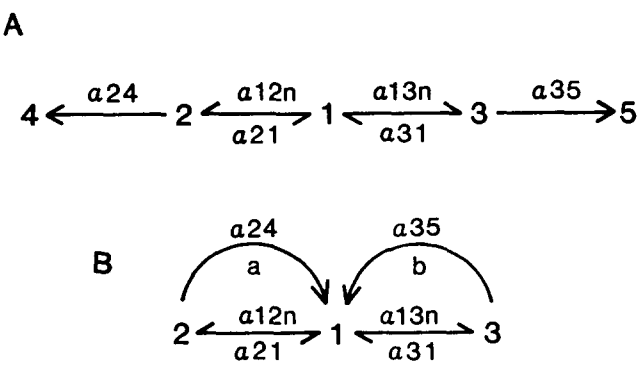

Figure 1 Diagram of a five state metabolic pathway. State 1 is the starting state and states 4 and 5 are termination states. A: open diagram; B: closed diagram.

probability of transition from state 1 to either state 2 or state 3 is proportional to nutrient intake. For a continuous random walk, the open digraph of *Figure 1A* can be converted to the closed digraph of *Figure 1B*. It must be understood that conversion of the open digraph to the closed digraph is to visualize the concept of a single continuous random walk. It should not be interpreted as a closed metabolic cycle. The probability that state i is occupied at any given time is defined as P_i. If the states which comprise the pathway are in steady state, the following system of equations can be written for the pathway in *Figure 1B:*

$$\frac{dP_1}{dt} = -(\alpha_{12} + \alpha_{13})\, nP_1 + (\alpha_{21} + \alpha_{24})\, P_2 + (\alpha_{31} + \alpha_{35}) = 0 \quad (1)$$

$$\frac{dP_2}{dt} = \alpha_{12}\, nP_1 - (\alpha_{21} + \alpha_{24})\, P_2 = 0 \quad (2)$$

$$\frac{dP_3}{dt} = \alpha_{13}\, nP_1 - (\alpha_{31} + \alpha_{35})\, P_3 = 0 \quad (3)$$

In the above expressions, n represents nutrient intake. Only two of the foregoing equations are independent. However, since one of the states will be occupied at any given time:

$$P_1 + P_2 + P_3 = 1 \quad (4)$$

Any one of equations 1–3 can be replaced by equation 4, and the resulting system of equations can be solved for the unknown probabilities. For example,

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha_{12}n & -(\alpha_{21}+\alpha_{24}) & 0 \\ \alpha_{13}n & 0 & -(\alpha_{31}+\alpha_{35}) \end{vmatrix} \begin{vmatrix} P_1 \\ P_2 \\ P_3 \end{vmatrix} = \begin{vmatrix} 1 \\ 0 \\ 0 \end{vmatrix} \quad (5)$$

There are a number of efficient methods available for obtaining a numerical determinant, but Cramer's rule is the most feasible method for obtaining a symbolic determinant. Unfortunately, Cramer's rule is not an efficient procedure for matrix inversion, and for that reason Hill[10-15] used the graphical procedure employed by King and Altman.[18] However, even the King and Altman procedure becomes unwieldy in the

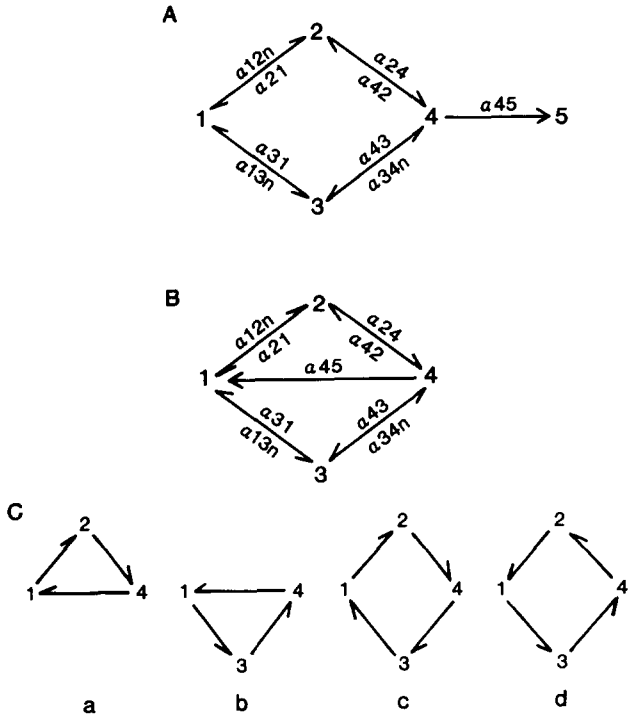

Figure 2 Diagram of a five state, branched metabolic pathway. State 1 is the starting state and state 5 is the termination state. A: open diagram; B: closed diagram; C: cycles formed by closed diagram.

case of branched paths. Representation of a pathway as a connection matrix, as has been done for derivation of rate equations for enzymic reactions,[19] is more concise and has the enormous advantage that it can be incorporated readily into a computer-based algorithm.[20] Derivation of the expressions for the probabilities in *Figure 1B*, using the connection matrix method, is presented in Appendix A.

The derived expressions for the fractional probabilities are:

$$P_1 = (\alpha_{21} + \alpha_{24})(\alpha_{31} + \alpha_{35})/\Sigma \quad (6)$$

$$P_2 = \alpha_{12}(\alpha_{31} + \alpha_{35})\, n/\Sigma \quad (7)$$

$$P_3 = \alpha_{13}(\alpha_{21} + \alpha_{24})\, n/\Sigma \quad (8)$$

where $\Sigma = (\alpha_{21} + \alpha_{24})(\alpha_{31} + \alpha_{35}) + \alpha_{12}(\alpha_{31} + \alpha_{35})\, n + \alpha_{13}(\alpha_{21} + \alpha_{24})\, n$. *Figure 1B* contains two one-way cycles. The flux through each of the cycles is:

$$J_a = \alpha_{24} P_2 \quad (9)$$

$$J_b = \alpha_{35} P_3. \quad (10)$$

Figure 2A portrays a five-state pathway in which state 1 is the starting state and state 5 is the termination state. For the purpose of this discussion, the probability of transitions from state 1 to states 2 and 3 and the transition from state 3 to state 4 are assumed to be proportional to nutrient intake. If a random walk reaches the termination state, it is assumed that the walk will be transferred instantaneously to the starting state, and this is portrayed in the closed digraph of *Figure 2B*. Appendix B gives the derivation of the expressions for the probability of fractional occupancy of the different states. These expressions are:

$$P_1 = [\alpha_{21}\alpha_{31}(\alpha_{42} + \alpha_{43} + \alpha_{45}) + \alpha_{24}\alpha_{31}(\alpha_{43} + \alpha_{45}) + \alpha_{21}\alpha_{34}(\alpha_{42} + \alpha_{45})\, n + \alpha_{24}\alpha_{34}\alpha_{45}\, n]/\Sigma \quad (11)$$

$$P_2 = [\alpha_{12}\alpha_{31}(\alpha_{42} + \alpha_{43} + \alpha_{45})\, n + \alpha_{12}\alpha_{34}(\alpha_{42} + \alpha_{45})\, n^2 + \alpha_{13}\alpha_{34}\alpha_{42}\, n^2]/\Sigma \quad (12)$$

$$P_3 = [\alpha_{13}\alpha_{21}(\alpha_{42} + \alpha_{43} + \alpha_{45})\, n + \alpha_{13}\alpha_{24}(\alpha_{43} + \alpha_{45})\, n + \alpha_{12}\alpha_{24}\alpha_{43})\, n]/\Sigma \quad (13)$$

$$P_4 = [\alpha_{12}\alpha_{24}\alpha_{31}\, n + \alpha_{13}\alpha_{34}(\alpha_{21} + \alpha_{24})\, n^2 + \alpha_{12}\alpha_{24}\alpha_{34}\, n^2]/\Sigma \quad (14)$$

where Σ is the sum of the numerators of the foregoing expressions. There are a total of four one-way cycles described by *Figure 2B*, and these are shown in *Figure 2C*. The total flux through the pathway is the sum of cyles A and B and is given by equation 15.

$$J = J_a + J_b = \alpha_{45} P_4 \quad (15)$$

Discussion

Although it appears that nutrient-response curves can be described by rational polynomials,[1-4] the fitting of these curves to rational polynomials has been empirical. This paper presents mathematical derivation of a general type of equation for nutrient-response based on stochastic principles. This derivation is based on the assumption that the intermediate metabolites are in steady state. The derived expressions are rational polynomials with respect to nutrient intake.

The metabolic pathway in *Figure 1* represents a very simple pathway for the metabolism of a nutrient. One of the termination states, state 4, for example, can be visualized as the state that represents the observed response. State 5, the second termination state, can be visualized as excretion of the nutrient or excretion of a metabolite formed from the nutrient. Thus, flux through cycle A of *Figure 1B* would be the measured response, while cycle B would be an excretory cycle. Equation 9 would express the nutrient-response relationship. In this case, the nutrient-response relationship would be a 1:1 rational polynomial in nutrient intake. That is, the right-hand side of equation 9 is a ratio of polynomials each of which contains nutrient intake raised to the first power. Equation 9 can be reformulated in general terms as

$$r = \frac{\alpha_1 n}{\beta_0 + \beta_1 n} \quad (16)$$

where r is the measured response, and

$$\alpha_1 = \alpha_{12}\alpha_{24}(\alpha_{31} + \alpha_{35}) \quad (17)$$

$$\beta_0 = (\alpha_{21} + \alpha_{24})(\alpha_{31} + \alpha_{35}) \quad (18)$$

$$\beta_1 = \alpha_{12}(\alpha_{31} + \alpha_{35}) + \alpha_{13}(\alpha_{21} + \alpha_{24}) \quad (19)$$

It was postulated previously[4] that, in the case where the response is a 1:1 function of nutrient intake, the

response is proportional to the ratio $\beta_1 n:(\beta_0 + \beta_1 n)$, and that the proportionality factor is $\alpha_1:\beta_1$. This proportionality factor can be expressed as

$$\frac{\alpha_1}{\beta_1} = \frac{\alpha_{24}}{1 + \dfrac{\alpha_{13}(\alpha_{21} + \alpha_{24})}{\alpha_{12}(\alpha_{31} + \alpha_{35})}} \tag{20}$$

The pathway portrayed in *Figure 2* is more complex than that portrayed in *Figure 1*. In this instance, the nutrient, or a metabolite whose concentration is directly proportional to nutrient intake, interacts at two sites in the pathway. This might occur in a number of ways. For example, if the nutrient were dietary protein and the observed response were the net increase in body protein nitrogen, *Figure 2* could represent a very abbreviated pathway for the process. State 1, the starting state, could be visualized as the amino acids derived from dietary protein. Some of these amino acids could be incorporated into body protein by a pathway represented by cycle A of *Figure 2C*. In this path, state 2 might be visualized as amino acyl-AMP, and state 4 as the tRNA amino acid. The probability of transition from state 1 to state 2 would be proportional to dietary protein intake. In an alternate pathway, amino groups from the amino acids arising from dietary protein could be transferred to amino acceptors, so that these amino groups, along with amino groups from other sources, would constitute the total amino pool. Likewise, the resultant alpha keto acids would be one source of an alpha keto acid pool. This alpha keto acid pool can be visualized as state 3 in *Figure 2*. Amino groups could be transferred to alpha keto acids subsequently, and the probability that an amino group arising from dietary protein would be transferred to an alpha keto acid arising from dietary protein is portrayed in cycle B of *Figure 2C*. The pathway shown in *Figure 2* does not contain all of the states that would be involved in the metabolism of dietary protein. It is intended only to portray a minimal pathway to account for the interaction of nutrient at two sites in the pathway.

The total flux through the pathway presented in *Figure 2* is given by equation 15, and gives rise to the following general 2:2 rational polynomial:

$$r = \frac{\alpha_1 n + \alpha_2 n^2}{\beta_0 + \beta_1 n + \beta_2 n^2} \tag{21}$$

where

$$\alpha_1 = \alpha_{12}\alpha_{24}\alpha_{31}\alpha_{45} \tag{22}$$

$$\alpha_2 = \alpha_{34}\alpha_{45}[\alpha_{12}\alpha_{24} + \alpha_{13}(\alpha_{21} + \alpha_{24})] \tag{23}$$

$$\beta_0 = \alpha_{31}[\alpha_{21}(\alpha_{42} + \alpha_{43} + \alpha_{45}) + \alpha_{24}(\alpha_{43} + \alpha_{45})] \tag{24}$$

$$\begin{aligned}\beta_1 = {} & \alpha_{24}[\alpha_{12}(\alpha_{31} + \alpha_{43}) + \alpha_{13}(\alpha_{43} + \alpha_{45}) \\ & + \alpha_{34}\alpha_{45}] + \alpha_{21}[\alpha_{34}(\alpha_{42} + \alpha_{45}) \\ & + \alpha_{13}(\alpha_{42} + \alpha_{43} + \alpha_{45}] \\ & + \alpha_{31}[\alpha_{12}(\alpha_{42} + \alpha_{43} + \alpha_{45}]\end{aligned} \tag{25}$$

$$\begin{aligned}\beta_2 = {} & \alpha_{34}[(\alpha_{12}\alpha_{24} + \alpha_{13}\alpha_{42}) + \alpha_{12}(\alpha_{42} + \alpha_{45}) \\ & + \alpha_{13}(\alpha_{21} + \alpha_{24})]\end{aligned} \tag{26}$$

It was concluded previously[4] that, if the measured response were a 2:2 function of nutrient intake, the proportionality factor that would relate the response to the fraction $\beta_2 n^2/(\beta_0 + \beta_1 n + \beta_2 n^2)$ is $(\alpha_1/n + \alpha_2)/\beta_2$. If a sufficiently high level of nutrient intake can be achieved, the proportionality factor would be approximately α_2/β_2. In terms of the probabilities of the transitions in *Figure 2*, the proportionality factor if the foregoing condition were met is:

$$\frac{\alpha_2}{\beta_2} = \frac{\alpha_{45}}{1 + \dfrac{\alpha_{13}\alpha_{42} + \alpha_{12}(\alpha_{42} + \alpha_{45})}{\alpha_{12}\alpha_{24} + \alpha_{13}(\alpha_{21} + \alpha_{24})}} \tag{27}$$

For the purpose of conciseness, the pathway portrayed in *Figure 2* does not contain a termination state identified as an excretory state. Any number of additional termination states could be included in the pathway.

The metabolic pathways in *Figures 1* and *2* are simple, but they suffice to demonstrate that, by use of probability theory, it is possible to derive an expression relating an observed response to nutrient intake. The simplicity of the pathways considered does not detract from the generality of the derivation presented in this paper as is evidenced by the fact that the treatment developed by Henri for a very simple enzymic reaction is applicable to very complex enzymic models.[6] It appears that this is the first time that an equation has been derived to describe the nutrient-response relationship. The following is the resulting general equation that describes nutrient-response relationships.

$$r = \frac{\sum_{i=1}^{l} \alpha_i n^i}{\sum_{i=0}^{m} \beta_i n^i}, \; m \geq 1, \alpha_i \geq 0, \beta_i \geq 0, \text{ where all } \beta_i \neq 0 \tag{28}$$

The foregoing equation can describe a wide variety of curves, including hyperbolic curves, sigmoidal curves, curves that provide for inhibition of response at high levels of nutrient intake, and even curves with multiple turning points.[3] It is the same type of equation that relates the rate of an enzyme-catalyzed reaction in steady state to substrate concentration.[6] No other equation has been of equal value in elucidating the reaction sequence of enzymes, and equation 28 has equal potential in analysis of nutrient-response relationships. The metabolic fate of the nutrient determines the shape of the curve, and hence it determines the magnitude of the parameters of equation 28. Analysis of the nutrient-response curve provides estimates of the parameters (for example see reference 3). The information thus obtained provides a link between the data accessible to nutritionists working at the level of the whole animal and the current theories of metabolic

control at the molecular level.[21-25] That is, data obtained with the whole animal can be correlated with flux and concentration control coefficients and logarithmic gains.[4]

The present work has provided derivation of a rational polynomial for the nutrient-response such that the fitting of data obtained in nutritional experiments to an equation is not merely empirical, but has a theoretical basis. Brown et al.[26] coined the term "top-down" for the analysis of metabolic control from the level of the organelle and cell down to the level of the metabolic pathway. In the opinion of this author, the science of nutrition is well positioned to pioneer this "top-down" approach to metabolic control from the level of the intact animal.

References

1 Monod, J. (1949). The growth of bacterial cultures. *Ann. Rev. Microbiol.* **3,** 371–394

2 Morgan, P.H., Mercer, L.P., and Flodin, N.W. (1975). General model for nutritional responses of higher animals. *Proc. Natl. Acad. Sci. USA* **72,** 4327–4331

3 Schulz, A.R. (1987). Analysis of nutrient-response relationships. *J. Nutr.* **117,** 1950–1959

4 Schulz, A.R. (1991). Interpretation of nutrient-response relationships. *J. Nutr.* **121,** 1834–1843

5 Savageau, M.A. (1976). In *Biochemical Systems Analysis*, pp. 79–95, Addison-Wesley, Reading, MA

6 Laidler, K.J. and Bunting, P.S. (1973). In *The Chemical Kinetics of Enzyme Action*, 2nd. ed., pp. 68–141, Clarendon Press, Oxford, England

7 Bartholomay, A.F. (1962). A stochastic approach to statistical kinetics with application to enzyme kinetics. *Biochemistry* **1,** 223–230

8 Ninio, J. (1987). Alternative to the steady-state method: derivation of reaction rates from first passage times and pathway probabilities. *Proc. Natl. Acad. Sci. USA* **84,** 663–667

9 Mazur, A.K. (1991). A probabilistic view of steady state enzyme reactions. *J. Theor. Biol.* **148,** 229–242

10 Hill, T.E. (1988). Interrelations between random walks on diagrams (graphs) with and without cycles. *Proc. Natl. Acad. Sci. USA* **85,** 2879–2883

11 Hill, T.E. (1988). Further properties of random walks on diagrams (graphs) with and without cycles. *Proc. Natl. Acad. Sci. USA* **85,** 3271–3275

12 Hill, T.E. (1988). Number of visits to a state in a random walk before absorption and related topics. *Proc. Natl. Acad. Sci. USA* **85,** 4577–4581

13 Hill, T.E. (1988). Discrete-time random walks on diagrams (graphs) with cycles. *Proc. Natl. Acad. Sci. USA* **85,** 5345–5349

14 Hill, T.E. (1989). In *Free Energy Transduction and Biochemical Cycle Kinetics*, pp. 39–88, Springer-Verlag, New York

15 Hill, T.E. (1974). In *Free Energy Transduction in Biology*, pp. 1–32, Academic Press, New York

16 Hearon, J.Z. (1963). Theorems on linear systems. *Ann. N.Y. Acad. Sci.* **108,** 36–68

17 Zhang, L.-Q., Collins, J.C., and King, P.H. (1991). Indistinguishability and identifiability analysis of linear compartmental models. *Math. Biosci.* **103,** 77–95

18 King, E.S. and Altman, C.A. (1956). A systematic method of deriving the rate-laws for enzyme-catalyzed reactions. *J. Phys. Chem.* **60,** 1375–1378

19 Fisher, D.D. and Schulz, A.R. (1969). Connection matrix representation of enzyme reaction sequences. *Math. Biosci.* **4,** 189–200

20 Gibbons, A. (1989). In *Algorithmic Graph Theory*, pp. 39–66, Cambridge University Press, Cambridge, England

21 Savageau, M.A. (1972). The behavior of intact biochemical control systems. *Cur. Topics Cell. Regul.* **6,** 63–130

22 Kacser, H. and Burns, J.A. (1972). The control of flux. *Symp. Soc. Expt. Biol.* **27,** 65–104

23 Heinrich, R., Rapoport, S.M., and Rapoport, T.A. (1979). Metabolic regulation and mathematical models. *Prog. Biophys. Mol. Biol.* **32,** 1–82

24 Crabtree, B. and Newsholme, E.A. (1985). A quantitative approach to metabolic control. *Cur. Topics Cell. Regul.* **25,** 21–76

25 Welch, G.R., Keleti, T., and Vertessey, B. (1988). The control of cell metabolism for homogenous vs. heterogenous enzyme systems. *J. Theor. Biol.* **130,** 407–422

26 Brown, G.C., Hafner, R.P., and Brand, M.D. (1990). A 'top-down' approach to the determination of control coefficients in metabolic control theory. *Eur. J. Biochem.* **188,** 321–325

Appendix A

A connection matrix for the metabolic pathway portrayed in *Figure 1A* is constructed such that the rows of the matrix correspond to the source states of the allowable transitions, while the columns correspond to the destination states of the transitions. Hence, the pathway in *Figure 1A* can be represented by a 3×5 connection matrix. The elements of each row of the connection matrix are determined in accordance with the following rules. (1) If there is a transition from the source state to the destination state, the element is 1 or, if the probability of the transition is proportional to nutrient intake, the element is n. (2) If there is no transition from the source to the destination state, the element is 0. The connection matrix for the pathway in *Figure 1A* is:

$$U = \begin{vmatrix} 0 & n & n & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{vmatrix} \qquad Q = \begin{vmatrix} 2 & 3 \\ 1 & 4 \\ 1 & 5 \end{vmatrix}$$

Analysis of the U matrix is facilitated by a secondary matrix. This is matrix Q, and it is constructed from U such that the elements of each row of Q identify the non-zero elements of the corresponding row of U. The problem is to calculate the probability of occupancy of each of the three states in *Figure 1B* at any given time during the continuous walk through the closed digraph. In each case, the numerator of the probability is given by the sum of all of the directed spanning trees that lead to the state for which the probability is being calculated.[20] The denominator is the sum of all the numerators. A directed spanning tree is a non-cyclic path that leads to the state in question from all of the other states. These directed spanning trees are obtained by constructing a matrix for the probability of occupancy of each state by replacing each element in the row of Q corresponding to the state involved with zeros.[18] The procedure is illustrated by the expressions for the numerators of each probability:

$$Q_1 = \begin{vmatrix} 0 & 0 \\ 1 & 4 \\ 1 & 5 \end{vmatrix} \quad \begin{matrix} (0,1,1) & \equiv \alpha_{21}\alpha_{31} \\ (0,1,5) & \equiv \alpha_{21}\alpha_{35} \\ (0,4,1) & \equiv \alpha_{24}\alpha_{31} \\ (0,4,5) & \equiv \alpha_{24}\alpha_{35} \end{matrix} \quad = P_1$$

$$Q_2 = \begin{vmatrix} 2 & 3 \\ 0 & 0 \\ 1 & 5 \end{vmatrix} \quad \begin{matrix} (2,0,1) & \equiv \alpha_{12}\alpha_{31}n \\ (2,0,5) & \equiv \alpha_{12}\alpha_{35}n \end{matrix} \quad = P_2$$

$$Q_3 = \begin{vmatrix} 2 & 3 \\ 1 & 4 \\ 0 & 0 \end{vmatrix} \quad \begin{matrix} (3,1,0) \equiv \alpha_{13}\alpha_{21}n \\ (3,4,0) \equiv \alpha_{13}\alpha_{24}n \end{matrix} \quad = P_3$$

Each probability is a quotient, the denominator of which is the sum of the above numerators.

Each path, which may or may not be a valid directed spanning tree, is obtained by forming a vector containing one element from each row of the Q_i matrix. The elements of the vector are aligned in accordance with the row of the matrix from which they came. Each element in the vector represents the weight, $\alpha_{s,d}$, of a transition. The source state of the transition is identified by the location of the element in the vector while the destination state is the value of the element. When the element in the vector is 0, no transition originates from the identified state. In the foregoing expressions, the valid spanning trees that comprise the probability of occupancy of each state are expressed both in vector form and as products of transition weights.

The algorithm to distinguish between spanning trees and cyclic paths has been described fully elsewhere.[19] The algorithm was developed for implementation on a computer, but it can be employed manually. The test for validity is conducted on the path in question in the vector form. An index is initially set to 1, that is, the index is set initially to point to the first element in the vector. If the value of the element identified is either 0 or greater than the index, the path is not cyclic to that point, and the index is incremented. It must be borne in mind that, if the element identified has the value of a termination state, the value of the element should be interpreted as that of the starting state. If the value of the identified element is less than that of the index, it is necessary to use that value as a pointer to a new element that is to be compared to the index. If the value of the new element is either equal to 0 or is greater than the index, the path is valid to that point and the index can be incremented. If the value of the new element is less than the index, the foregoing procedure must be repeated. Anytime the value of the identified element is equal to the index, the path is cyclic and therefore, not a valid spanning tree. The described procedure is repeated until the last element in the vector has been tested.

All of the possible paths in P_1 are valid spanning trees. However, in P_2 the third and fourth paths in vector form are (3, 0, 1) and (3, 0, 5), respectively. The first two elements in both of these vectors satisfy the test as valid paths, but the third element in both cases completes a cyclic path, and therefore, the paths do not constitute valid spanning trees. Likewise, the first two paths in P_3 are (2, 1, 0) and (2, 4, 0), and these are cyclic paths.

Appendix B

The connection matrix and the matrix of non-zero elements for the pathway portrayed in *Figure 2A* are:

$$U = \begin{vmatrix} 0 & n & n & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & n & 0 \\ 0 & 1 & 1 & 0 & 1 \end{vmatrix} \qquad Q = \begin{vmatrix} 2 & 3 & 0 \\ 1 & 4 & 0 \\ 1 & 4 & 0 \\ 2 & 3 & 5 \end{vmatrix}$$

The directed spanning trees that constitute the numerator of the probability of occupancy for each state are obtained as follows:

$$Q_1 = \begin{vmatrix} 0 & 0 & 0 \\ 1 & 4 & 0 \\ 1 & 4 & 0 \\ 2 & 3 & 5 \end{vmatrix} \quad \begin{matrix} (0,1,1,2) \equiv \alpha_{21}\alpha_{31}\alpha_{42} \\ (0,1,1,3) \equiv \alpha_{21}\alpha_{31}\alpha_{43} \\ (0,1,1,5) \equiv \alpha_{21}\alpha_{31}\alpha_{45} \\ (0,1,4,2) \equiv \alpha_{21}\alpha_{34}\alpha_{42}n \\ (0,1,4,5) \equiv \alpha_{21}\alpha_{34}\alpha_{45}n \\ (0,4,1,3) \equiv \alpha_{24}\alpha_{31}\alpha_{43} \\ (0,4,1,5) \equiv \alpha_{24}\alpha_{31}\alpha_{45} \\ (0,4,4,5) \equiv \alpha_{24}\alpha_{34}\alpha_{45}n \end{matrix} \quad = P_1$$

$$Q_2 = \begin{vmatrix} 2 & 3 & 0 \\ 0 & 0 & 0 \\ 1 & 4 & 0 \\ 2 & 3 & 5 \end{vmatrix} \quad \begin{matrix} (2,0,1,2) \equiv \alpha_{12}\alpha_{31}\alpha_{42}n \\ (2,0,1,3) \equiv \alpha_{12}\alpha_{31}\alpha_{43}n \\ (2,0,1,5) \equiv \alpha_{12}\alpha_{31}\alpha_{45}n \\ (2,0,4,2) \equiv \alpha_{12}\alpha_{34}\alpha_{42}n^2 \\ (2,0,4,5) \equiv \alpha_{12}\alpha_{34}\alpha_{45}n^2 \\ (3,0,4,2) \equiv \alpha_{13}\alpha_{34}\alpha_{42}n^2 \end{matrix} \quad = P_2$$

$$Q_3 = \begin{vmatrix} 2 & 3 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 0 \\ 2 & 3 & 5 \end{vmatrix} \quad \begin{matrix} (2,4,0,3) \equiv \alpha_{12}\alpha_{24}\alpha_{43}n \\ (3,1,0,2) \equiv \alpha_{13}\alpha_{21}\alpha_{42}n \\ (3,1,0,3) \equiv \alpha_{13}\alpha_{21}\alpha_{43}n \\ (3,1,0,5) \equiv \alpha_{13}\alpha_{21}\alpha_{45}n \\ (3,4,0,3) \equiv \alpha_{13}\alpha_{24}\alpha_{43}n \\ (3,4,0,5) \equiv \alpha_{13}\alpha_{24}\alpha_{45}n \end{matrix} \quad = P_3$$

$$Q_4 = \begin{vmatrix} 2 & 3 \\ 1 & 4 \\ 1 & 4 \\ 0 & 0 \end{vmatrix} \quad \begin{matrix} (2,4,1,0) \equiv \alpha_{12}\alpha_{24}\alpha_{31}n \\ (2,4,4,0) \equiv \alpha_{12}\alpha_{24}\alpha_{34}n^2 \\ (3,1,4,0) \equiv \alpha_{13}\alpha_{21}\alpha_{34}n^2 \\ (3,4,4,0) \equiv \alpha_{13}\alpha_{24}\alpha_{34}n^2 \end{matrix} \quad = P_4$$

The denominator of each probability is the sum of the above numerators.