

A “random steady state” model for the pyruvate dehydrogenase and alpha-ketoglutarate dehydrogenase enzyme complexes

T S Najdi^{1,3}, G W Hatfield¹ and E D Mjolsness^{1,2}

¹Institute for Genomics and Bioinformatics, ²Donald Bren School of Information and Computer Science, ³Department of Biological Chemistry, College of Medicine, University of California, Irvine, Irvine, CA 92697, USA.

Email: tnajdi@uci.edu

Abstract. The multienzyme complexes, pyruvate dehydrogenase and alpha-ketoglutarate dehydrogenase, involved in the central metabolism of *Escherichia Coli* consist of multiple copies of three different enzymes, E1, E2 and E3 that cooperate to channel substrate intermediates between their active sites. The E2 components form the core of the complex while a mixture of E1 and E3 components bind to the core. We present a Random Steady State model to describe catalysis by such multienzyme complexes. At a fast time scale, the model describes the enzyme catalytic mechanisms of substrate channeling at steady state, by polynomially approximating the analytic solution of a biochemical Master Equation. At a slower time scale, the structural organization of the different enzymes in the complex and their random binding/unbinding to the core is modeled using methods from equilibrium statistical mechanics. Biologically, the model describes the optimization of catalytic activity and substrate sharing over that entire enzyme complex. The resulting enzymatic models illustrate the random steady state (RSS) for modeling multienzyme complexes in metabolic pathways.

MSC classification codes: 92C05, 92B05, 92C45.

Keywords: 2-oxo acid dehydrogenase, partition function, enzyme complexes, random steady state

Submitted to: Physical Biology.

1. Introduction

Multienzyme complexes are multimeric structures made up of several non-covalently interacting enzymes that catalyze successive steps of a metabolic pathway. Such structures can enhance enzyme activity by allowing the limited diffusion of substrate intermediates and possibly protecting any unstable catalytic intermediates [1-2]. Two such multienzyme complexes in *Escherichia coli* are pyruvate dehydrogenase (PDH) and α -keto glutarate dehydrogenase (KGDH) which are notable examples of 2-oxo acid dehydrogenase complexes. The PDH and KGDH complexes function at strategic points in the TCA cycle to catalyze the conversion of pyruvate and α -ketoglutarate into acetyl-CoA and succinyl-CoA, respectively.

The PDH complex consists of multiple copies of three different enzymes: pyruvate dehydrogenase / decarboxylase (E1); lipoyl acetyl-transferase (E2); and dihydrolipoyl dehydrogenase (E3). E1 catalyzes the oxidative decarboxylation of pyruvate and this product is transferred to a lipoyl domain of E2 to form acetyl-lipoic acid (figure 1A). E2 then catalyzes the transfer of this acetyl group to CoA. The resulting dihydrolipoyl group of E2 is regenerated (oxidized) by E3. The PDH complex and KGDH complex are very closely related. They have a similar structural organization and catalyze similar reactions. They only differ by their E1 and E2 components which are encoded by different genes.

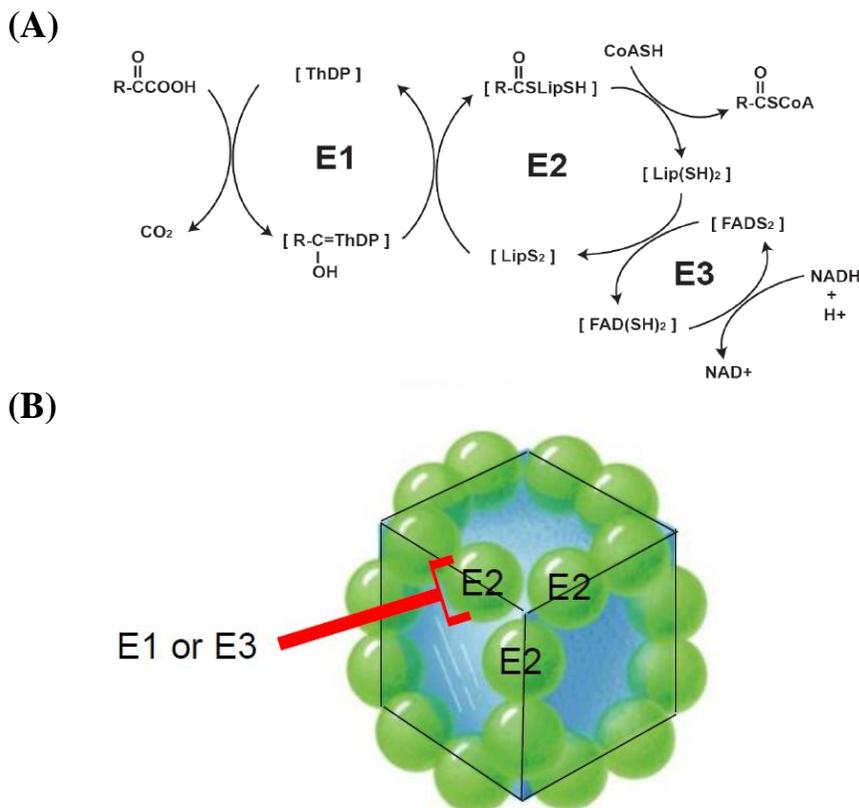


Figure 1. Schematic diagrams of the reaction mechanisms (Panel A) and the structural organization (Panel B) of the enzyme components of the PDH or KGDH complexes are shown. The E2s are organized into an octahedral core. The structure can be pictured as a cube with a trimer of E2 occupying each vertex. Each E2 within a trimer can either bind an E1 dimer or an E3 dimer. The lipoyl domains of E2 act as swinging arms to move the substrates between active sites. The abbreviations used are: ThDP, thiamin diphosphate; Lip, lipoic acid.

The structure of the PDH and KGDH multienzyme complexes and the spatial distribution of the E2 lipoyl moieties have been studied by scanning transmission electron microscopy and electron cryotomography [3-4]. The core of the PDH and KGDH enzyme complexes consist of 24 E2 copies organized into an octahedral symmetry (figure 1B). This can be pictured as a cube with 3 E2s occupying each of its eight vertices. Each vertex of the cube has three neighboring vertices. The E2 core of PDH can accommodate up to 24 dimers of either E1 alone or E3 alone (3 dimers of E1 or E3 at each vertex subcomplex), and these compete for binding when mixed [5]. The lipoyl groups of E2 act as swinging arms that carry the substrate intermediates to successive active sites. The E1:E2:E3 ratio has been measured at different stoichiometric ratios [6], such as 2:2:1 [5] and 4:3:2 [5-7], and this is likely because of the different methods of preparation and possibly growth conditions[8-9]. A recent study by Murphy *et al.* [4] demonstrated that the entire complex is inherently non-homogenous, asymmetric, and flexible and suggest that a ratio of 4:3:2 consisting of 16 E1 dimers, 24 E2 monomers and 8 E3 dimers is the optimal

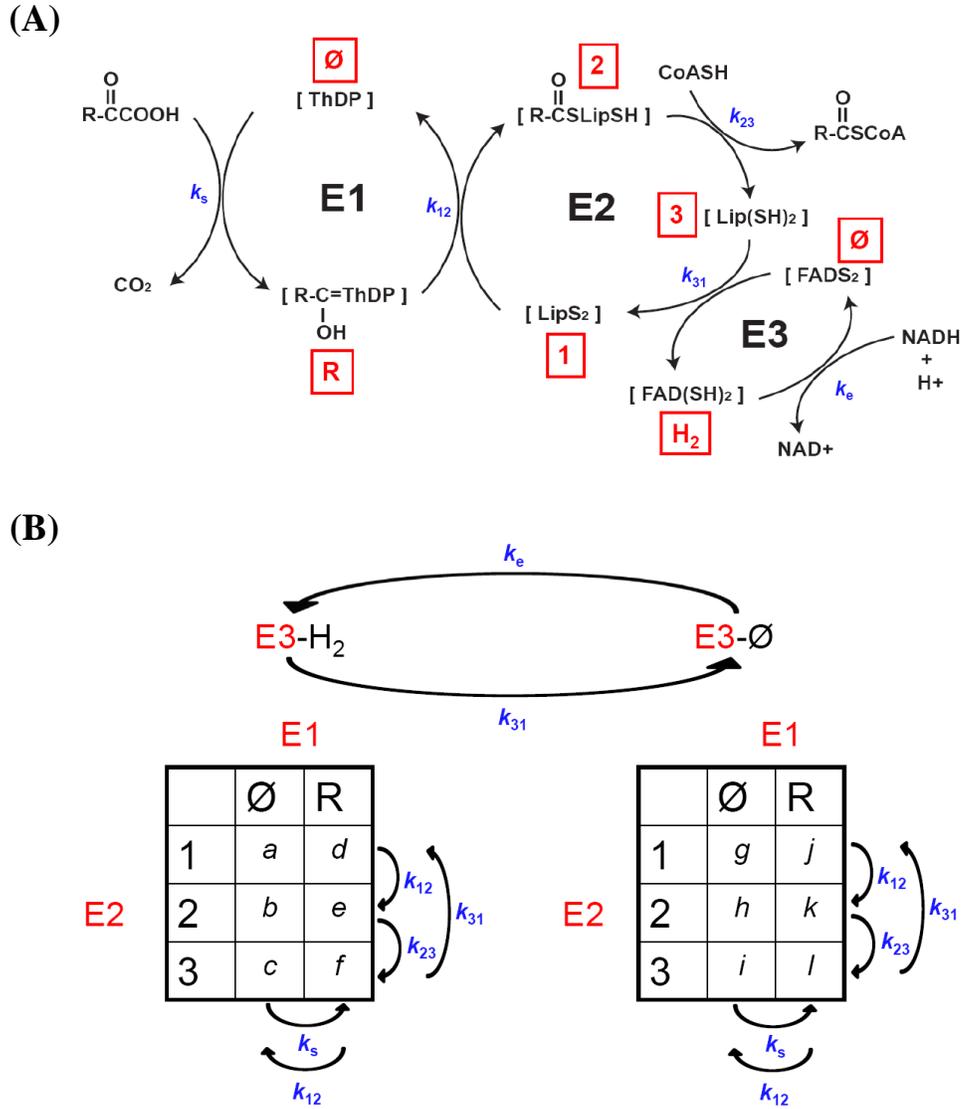
stoichiometry for catalysis. This ratio corresponds to a 2:1 E1:E3 catalytic units at each vertex subcomplex and over the whole enzyme complex.

kMech, a mathematical modeling tool for complex enzyme mechanisms, was previously described [10] and applied to the biosynthesis of the branched chain amino acids, L-isoleucine, L-valine, and L-leucine in *E. coli* [11] and to L-threonine biosynthesis, along with the development of a new model, the generalized Monod, Wyman, Changeux (GMWC) model, for regulation by allosteric enzymes [12]. Cellerator, a Mathematica™ package designed for the generation of chemical reaction networks [13], converts fundamental association-dissociation reactions generated by kMech into a system of ordinary differential equations (ODEs), which are numerically solved by Mathematica™. The kMech/Cellerator models have been expanded to include the pathways of glycolysis and the TCA cycle in order to generate a combined simulation of central metabolism and the pyruvate and aspartate family amino acids biosynthesis in *E. coli*.

To simulate and analyze the complicated catalytic mechanisms of the PDH and KGDH complexes in the central metabolic pathways, new models that describe the unique properties of enzyme complexes and expand the capabilities of the kMech enzyme models are developed. Here, we propose a Random Steady State (RSS) model to describe the behavior of multienzyme complexes. This model comprises dynamics on two separate time scales. At a fast time scale, the model incorporates the enzyme catalytic mechanisms of substrate channeling. At this fast time scale, the model assumes that a steady state of the substrate kinetics is reached, holding constant the organization of the enzymatic complex. At a slower time scale, the structural organization of the different enzymes in the complex and their random binding/unbinding to the core is modeled to allow for the optimization of catalytic activity and possible substrate sharing over that whole complex. At this slower time scale, binding and unbinding events randomly change the set of enzymes present in the complex. We seek a production rate over the resulting probability distribution on the composition and structure of the enzymatic complex. This combination of time scales is an inversion of the usual quasi-equilibrium which is stochastic and in thermal equilibrium at a fast time scale and evolves under a deterministic differential equation model at a slower time scale [14-15]. The model for this system is implemented as a Mathematica™ package along with kMech and Cellerator, for building a general framework for enzyme complex models.

2. The RSS model: derivation and simulations

To achieve a complete cycle of catalysis, the E1, E2 and E3 components cooperate to channel substrate intermediates between their active sites. The kinetics for the individual enzyme reactions involved have been described before [16-19]. To generate a rate equation that describes catalysis by the enzyme complex, the steady state reaction rate for a subcomplex of E1, E2, and E3 is derived by solving a system of linear equations representing a cycle of enzyme-catalyzed reactions (figure 2A). This linear system is the steady-state solution to the Master Equation [20] describing the flow of probability in state space as a single molecule of substrate successively invokes the enzymatic reactions catalyzed by E1, E2 and E3. For this subcomplex, n_1 , n_2 and n_3 are, respectively, the numbers of E1, E2 and E3 at a vertex subcomplex of the cube or any other polyhedral structure of an enzyme complex. In the case of PDH (and KGDH), $n_2 = 3 = n_1 + n_3$, then $n_3 = 3 - n_1$ and so n_1 can be the only variable in the derivation. The enzyme states in the cycle are chosen to represent all possible binding states of intermediates. E1 has two states depending on whether the substrate intermediate is bound; E2 has three states depending on whether the lipoyl group is oxidized or reduced or the substrate intermediate is bound; E3 has two states depending on whether the cofactor is oxidized or reduced. The rate constants k_s , k_{12} , k_{23} , k_{31} and k_e for each step are also shown (figure 2A). The steady state transition, consistent with the reaction scheme of figure 2A, is shown in 12 dimensions in figure 2B. All 12 possible states of a complex of E1, E2 and E3 are denoted alphabetically with the letter “a” through “l”.



The resulting 12 x 12 matrix, $\hat{\omega}$, is:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & k_{31} & 0 & 0 & 0 \\ 0 & 0 & 0 & k_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & k_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ k_s & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & k_{31} \\ 0 & k_s & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_s & 0 & k_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ k_e & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & k_e & 0 & 0 & 0 & 0 & 0 & 0 & 0 & k_{12} & 0 & 0 \\ 0 & 0 & k_e & 0 & 0 & 0 & 0 & k_{23} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & k_e & 0 & 0 & k_s & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & k_e & 0 & 0 & k_s & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & k_e & 0 & 0 & k_s & 0 & k_{23} & 0 \end{bmatrix}$$

The rate for the steady state problem is solved in MathematicaTM as a function of the number of E1's and E3's at each vertex as follows:

$$\hat{1} = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$$

$$\omega = \hat{\omega} - \text{diag}(1 \cdot \hat{\omega})$$

$$A = (a, b, c, d, e, f, g, h, i, j, k, l)$$

and the Master Equation is $\frac{dA}{dt} = \omega \cdot A$, which conserves probability since $1 \cdot \omega = 0$. The steady-state solution of the Master Equation is achieved by solving:

$$\begin{cases} \omega \cdot A = 0 \\ 1 \cdot A = 1 \end{cases}$$

and then computing the rate:

$$\rho = k_{23} (b + e + h + k)$$

The resulting rate law is $\rho(k_s, k_{12}, k_{23}, k_{31}, k_e)$. The rate constants in the resulting rate equation are substituted with their respective rate variable parameters and constants from experiments [6, 16-17]; such that $k_s = \frac{k s_1}{k_m + s_1}$; $k_{23} = \frac{k s_2}{k_m + s_2}$; $k_{12} = \varepsilon n_1 k$; $k_{31} = k (3 - n_1)$ and $k_e = \frac{k (3 - n_1)}{S_f}$ where n_1 is the number of E1s per vertex of the cube in our subcomplex, s_1 and s_2 are the substrate concentrations and k , k_m and ε are estimated from experiments. ε is used to estimate the rate limiting step k_{12} , which is 5 to 10 times slower than the other rate steps [16].

To simplify the function ρ , a Taylor series was generated to the third order in S_f , where $S_f = k_{31}/k_e$ is a small quantity estimated at 0.1 or less since the oxidation of the lipoic acid group of E3 is very fast relative to the rest of the reactions in the complex [17]. Large cancellations are observed in the expansion, which required the use of exact numerics as is supported in a computer algebra environment. The first and second order terms vanish exactly to zero, so the zeroth order (constant) term is expected to be an excellent approximation with errors of order 10^{-3} or less, by the remainder term in the Taylor expansion, and this was verified numerically. The resulting function is simple and unimodal, as shown in figure 3. Consequently, the zeroth order term is taken as the value of the rate $\rho(n_1, s_1, s_2, k, k_m, \varepsilon)$ for the (n_1, n_2, n_3) subcomplex:

$$\begin{aligned}
& \rho(n_1, s_1, s_2, k, k_m, \varepsilon) \\
&= \frac{k n_1 s_1 s_2 \varepsilon (n_1 - 3) (k_m (n_1 - 3) + s_1 (n_1 - 4)) (2 s_1 s_2 + k_m (s_1 + s_2))}{(s_1 s_2 (n_1 - 3) (k_m (n_1 - 3) + s_1 (n_1 - 4)) (2 s_1 s_2 + k_m (s_1 + s_2)) \\
&\quad + n_1 \varepsilon (s_1^3 s_2^2 (n_1 (3n_1 - 22) + 41) \\
&\quad + k_m s_1^2 s_2 (s_1 (n_1 - 4) (3 n_1 - 10) + s_2 (6n_1^2 - 40n_1 + 67)) \\
&\quad + k_m^3 (n_1 - 3)^2 (s_1^2 + s_1 s_2 + s_2^2) \\
&\quad + k_m^2 s_1 (n_1 - 3) (s_1^2 (n_1 - 4) + 2 s_1 s_2 (2n_1 - 7) + s_2^2 (4n_1 - 13)))}
\end{aligned} \tag{1}$$

and used in subsequent polynomial approximations described in equation 4 below. Figure 3 shows a plot of the rate $\rho(n_1)$ as a function of n_1 after substituting for the experimentally fitted values of $k = 2300 \text{ sec}^{-1}$, $k_m = 1300 \mu\text{M}$, $\varepsilon = 0.15$ and saturating substrate concentrations [16]. From the plot, it can be seen that the optimal values can be approximated for a 2:1 ratio of E1 to E3 at each vertex, as suggested before for the optimal catalytic stoichiometry [6]. Also, it can be noted that the rate drops dramatically as n_1 , the number of E1 at a vertex, approaches 0 or 3 as would be expected if the vertices of the cube are isolated and either E1 or E3 are absent at a vertex.

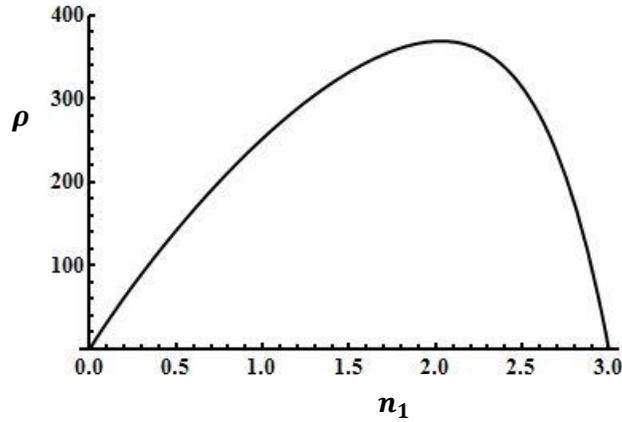


Figure 3. The reaction rate of pyruvate dehydrogenase as a function of the number of E1 per vertex of the enzyme complex cube structure is shown. The reaction rate $\rho(n_1)$ is in sec^{-1} . Each vertex of the cube can accommodate n_1 (number of E1s) ranging from 0 to 3.

However, the overall rate of the reactions cannot be simply described by three single enzymes. As mentioned in the earlier section, PDH and KGDH form a complex made up of as many as 48 copies of E1, E2 and E3, in which E2s form the octahedral core and each E2 can only bind either an E1 or an E3. Therefore, such structural organization and the random nature of E1 and E3 binding at each vertex have to be taken into consideration to correctly model enzyme catalysis by this complex.

For this reason and in addition to the solution for the cycle of enzymatic reactions, random E1 and E3 binding was incorporated to derive a more complete model of the vertex subcomplex for the RSS model. Using the multiplication and composition principles for partition functions in statistical mechanics [21], we calculate:

$$Z_{trimer}(z_1, z_3) = (\omega_1 z_1 + \omega_3 z_3)^3 = \sum_{n_1=0}^3 \binom{3}{n_1} (\omega_1 z_1)^{n_1} (\omega_3 z_3)^{3-n_1} \tag{2}$$

where z_1 and z_3 , are "fugacities" for occupying the tether binding site with an E1 or an E3. Here, they are proportional to the concentrations of E1 and E3 [21]. Assuming each E1/E3 binding site is independent, random binding of E1 and E3 to the E2s at a vertex subcomplex ($n_2 = 3$) was modeled by a binomial distribution of accessible E1s and E3s multiplied by the rate calculated above:

$$\rho' = \langle \rho(n_1) \rangle_{n_1} = \sum_{n_1=0}^3 \binom{3}{n_1} q^{n_1} (1-q)^{3-n_1} \rho(n_1) \quad (3)$$

where $\langle \rangle_{n_1}$ is the average over n_1 and $q = \frac{[E1]k_1}{[E1]k_1 + [E3]k_3}$ is dependent on the concentrations of E1 and E3 and their binding affinities to E2. Assuming that the binding affinities of E1 and E3 to E2 are similar as suggested before [6, 19], the rate ρ' (with random binding of E1/E3) as a function of q is plotted in figure 4. By comparing the optimal value from this plot with that of figure 3, it can be noted that the achievable rate from random E1/E3 occupancy is smaller than the optimal rate achieved earlier.

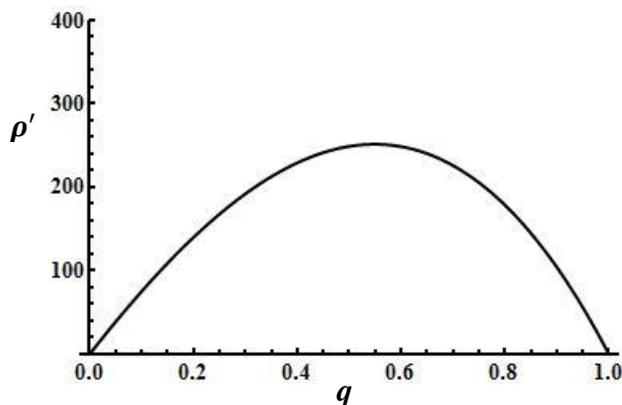


Figure 4. The reaction rate of pyruvate dehydrogenase with random binding of E1 and E3 to E2 is shown. The reaction rate ρ' (sec^{-1}), derived with random binding of E1 and E3 to the E2 core (see text), is plotted as a function of q .

Sharing of E1/E3 capacity among E2 trimers on neighboring vertices can minimize the chances of zero production when n_1 approaches 0 or 3. It can also reduce the difference between the achievable rate, when random binding at the trimers is taken into account, and the optimal rate of the reaction. Such sharing could explain the formation of an octahedral structure of the enzyme complex *in vivo* rather than trimers alone [6]. For these reasons, the rate is next modeled to allow for sharing at each vertex with three neighboring vertices (figure 1B) and throughputs are compared with the isolated vertex results, and reported experiments.

Before modeling the reaction rate with sharing, the rate function $\rho(n_1)$ is first approximated with a low-degree polynomial:

$$\tilde{\rho}(n_1) = \left(1 - \frac{n_1}{3}\right) \sum_{p=1}^{p_{max}} \lambda_p \frac{(p+1)^{p+1}}{p^p} \left(\frac{n_1}{3}\right)^p \quad (4)$$

This form is chosen so $\tilde{\rho}(n_1)$ is a linear combination of basis functions with the same maximal value, 1, and which are zero for $n_1 = 0$ or 3. The values of λ in the polynomial are next optimized so that $\tilde{\rho}(n_1)$ matches the reaction rate $\rho(n_1)$. We take $p_{max} = 6$ for a good fit. Both functions are plotted in figure 5.

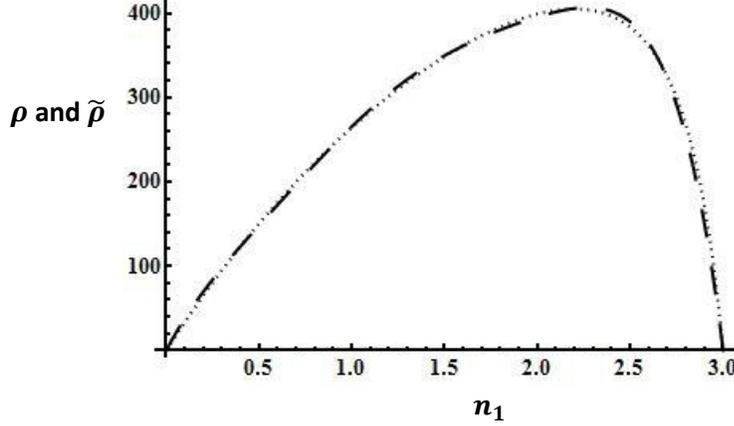


Figure 5. The reaction rate of pyruvate dehydrogenase and its approximating function are shown. The reaction rate $\rho(\mathbf{n}_1)$ (dotted) is plotted with its approximating function $\tilde{\rho}(\mathbf{n}_1)$ (dashed) as a function of \mathbf{n}_1 , the number of E1 per vertex of the enzyme complex cube structure. The rates ρ and $\tilde{\rho}$ are in sec^{-1} .

The resulting approximation $\tilde{\rho}(\mathbf{n}_1)$ will be used below (equation 5) to approximate the multi-vertex partition function: $Z_{\text{complex}} = \prod_v Z(n_{1,v}, n_{3,v})$, where $n_{1,v}$ and $n_{3,v}$ are n_1 and n_3 at a vertex of the polyhedral complex. To allow sharing, n_1 and $n_3 = 3 - n_1$ are replicated by the number of vertices and a sharing factor $\alpha \in [0,1]$ is introduced where α is the fraction of time an E1 from a neighboring vertex is shared and $(1 - \alpha)$ is the fraction of time an E1 from the same vertex participates in the reaction. Hence, the average with sharing is calculated by evaluating:

$$\tilde{\rho}[n_{1,v}(\text{effective})] = \tilde{\rho} \left[(1 - \alpha)n_{1,v} + \frac{\alpha}{|\text{nbrs}(v)|} \sum_{w \in \text{nbrs}(v)} n_{1,w} \right] \quad (5)$$

where $\text{nbrs}(v)$ is the set of neighbors of vertex v , and $|\text{nbrs}(v)|$ is the cardinality of that set, i.e. the number of neighbors of v . In this expression, each monomial in $n_{1,v}$ averages separately, by the linearity of averaging, and each monomial factor $n_{1,v}^k$ averages separately at its own vertex v , by independence of the vertices. If vertex-vertex interactions were known, they could be added in at the expense of a more difficult equilibrium calculation. Then globally averaging polynomials in $n_{1,v}$ is multilinear:

$$\begin{aligned} \langle \sum_{\{k\}} c(k_1, \dots, k_m) (n_{1,v_1})^{k_1} \dots (n_{1,v_m})^{k_m} \rangle \\ = \sum_{\{k\}} c(k_1, \dots, k_m) \langle (n_{1,v_1})^{k_1} \rangle \dots \langle (n_{1,v_m})^{k_m} \rangle \end{aligned} \quad (6)$$

However, what is easiest to compute using partition functions is not $\langle n_{1,v}^k \rangle$ but the falling factorial, $\langle (n_{1,v})_k \rangle = \langle \frac{n_{1,v}!}{(n_{1,v}-k)!} \rangle$. Since the averages of $(n_{1,v})_{(k)}$ are k^{th} derivatives of the vertex partition function, then:

$$\langle (n_{1,v})_k \rangle \equiv \langle \frac{n_{1,v}!}{(n_{1,v}-k)!} \rangle = \frac{1}{Z(z_{1,v}, \dots)} \frac{\partial^k Z(z_{1,v}, \dots)}{\partial z_{1,v}^k} = \frac{1}{Z_{\text{complex}}} \frac{\partial^k Z_{\text{complex}}}{\partial z_{1,v}^k}$$

where $Z(z_{1,v}, \dots) = (z_{1,v} + z_{3,v})^3$ since each of three sites at an E2 is linked either to an E1 or an E3 (equation 2). Lacking knowledge of possible interactions between the three binding states at a vertex, we

again assume independence. For a trimer, only $k = 1, 2, 3$ are nonzero, so for our cubic partition function, in fact, the falling factorials evaluate to zero for $k > 3$. For $k = 1, 2, 3$, the averages of $(n_{1,v})_{(k)}$ evaluate to:

$$\text{For } k = 1: \frac{1}{Z} \frac{\partial}{\partial z_{1,v}} Z = \frac{z_{1,v}}{Z} \left[\frac{\partial}{\partial z_{1,v}} (z_{1,v} + z_{3,v})^3 \right] = 3 \frac{z_{1,v}}{(z_{1,v} + z_{3,v})} = 3q$$

$$\text{For } k = 2: \frac{1}{Z} \frac{\partial^2}{\partial z_{1,v}^2} Z = \frac{z_{1,v}^2}{Z} \left[\frac{\partial}{\partial z_{1,v}} 3(z_{1,v} + z_{3,v})^2 \right] = 6 \frac{z_{1,v}^2}{(z_{1,v} + z_{3,v})^2} = 6q^2$$

$$\text{For } k = 3: \frac{1}{Z} \frac{\partial^3}{\partial z_{1,v}^3} Z = \frac{z_{1,v}^3}{Z} \left[\frac{\partial}{\partial z_{1,v}} 6(z_{1,v} + z_{3,v}) \right] = 6 \frac{z_{1,v}^3}{(z_{1,v} + z_{3,v})^3} = 6q^3$$

$$\text{For } k > 3: \frac{1}{Z} \frac{\partial^k}{\partial z_{1,v}^k} Z = 0$$

The vertex averages $\langle n_{1,v}^k \rangle$ are related to those of the falling factorial by Stirling numbers of the first kind $\left[\begin{smallmatrix} l \\ m \end{smallmatrix} \right]$ and the second kind $\left\{ \begin{smallmatrix} k \\ l \end{smallmatrix} \right\}$ by a change of polynomial basis between n^k and $(n)_k$ [22]:

$$n^k = \sum_{l=0}^k \left\{ \begin{smallmatrix} k \\ l \end{smallmatrix} \right\} (n)_l \quad \text{and} \quad (n)_l = \sum_{m=0}^l \left[\begin{smallmatrix} l \\ m \end{smallmatrix} \right] n^m$$

We can interconvert between powers and falling factorials using the Stirling numbers and vertex averages $\langle n^k \rangle$ computed by recurrence using:

$$\langle n^k \rangle = \langle (n)_k \rangle - \sum_{l=0}^{k-1} \left[\begin{smallmatrix} k \\ l \end{smallmatrix} \right] \langle n^l \rangle$$

So using Stirling numbers, the averages were calculated resulting in:

$$\begin{aligned} \langle n_1 \rangle &= 3q \\ \langle n_1^2 \rangle &= 6q^2 + \langle n_1 \rangle = 3(q + 2q^2) \\ \langle n_1^3 \rangle &= 6q^3 + 3 \langle n_1^2 \rangle - 2 \langle n_1 \rangle = 3(q + 6q^2 + 2q^3) \\ \langle n_1^4 \rangle &= 0 + 6 \langle n_1^3 \rangle - 11 \langle n_1^2 \rangle + 6 \langle n_1 \rangle = 3(q + 14q^2 + 12q^3) \\ \langle n_1^5 \rangle &= 0 + 10 \langle n_1^4 \rangle - 35 \langle n_1^3 \rangle + 50 \langle n_1^2 \rangle - 24 \langle n_1 \rangle = 3(q + 30q^2 + 50q^3) \\ \langle n_1^6 \rangle &= 0 + 15 \langle n_1^5 \rangle - 85 \langle n_1^4 \rangle + 225 \langle n_1^3 \rangle - 274 \langle n_1^2 \rangle + 120 \langle n_1 \rangle \\ &= 3(q + 62q^2 + 180q^3) \\ &\vdots \\ &\text{etc} \end{aligned} \tag{7}$$

Using multilinearity, the new rate, $\rho_s(q, \alpha)$, for the RSS model was defined as:

$$\begin{aligned} \rho_s(q, \alpha) &= \left\langle \tilde{\rho} \left((1 - \alpha) n_{1,0} + \alpha \left(\frac{n_{1,1} + n_{1,2} + n_{1,3}}{3} \right) \right) \right\rangle \\ &= \sum_{\{k\}} c(k_1, \dots, k_m) \prod_i \langle n_{1,v}^{k_i} \rangle \end{aligned} \tag{8}$$

where $k_i = 1, 2, \dots, p_{max} = 6$, which, upon substituting $n_{1,v}^P$ from equation 7, results in a polynomial function of q and α .

To investigate the range of values of q and α that result in optimal values for the rate, ρ_s , a contour plot is generated (figure 6). The best values for the sharing factor, α , can be observed around 0.75 which represents a completely equal division of work between and the vertex in question and its three neighboring vertices on the enzyme complex. Also, the best values for q are observed around 0.66 which corresponds to a concentration E1 about twice the concentration of E3 which is consistent with an optimal ratio of 2:1 E1 to E3 in the enzyme complex. The values for q and α for a maximal value of ρ_s are 0.684 and 0.75, respectively.

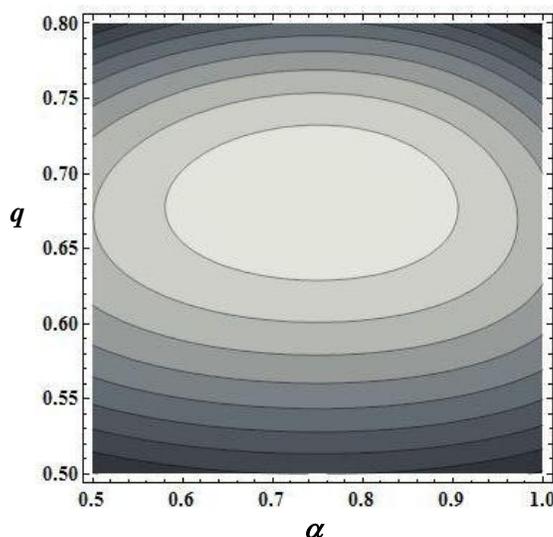


Figure 6. Values of q and α for an optimal reaction rate with enzyme sharing over the vertices of the PDH complex are shown. Here, the best values of q and α are 0.684 and 0.75, respectively.

The rate ρ_s (equation 8) offers a significant improvement due to sharing over the enzyme complex when compared to the rate ρ' (equation 3) as can be seen when both rates are plotted as a function of q (figure 7A). The value of α for ρ_s in this case is 0.75, which allows for optimal sharing. Gain in reaction rate due to sharing peaks at 62% for $\alpha = 0.75$ and drops with distance from that optimal value of sharing with neighbors (figure 7B). Furthermore, for an $\alpha = 0.75$, the reaction rate ρ_s as a function of q is very comparable with the optimal rate from figure 3 and is also in agreement with data of rate as a function of enzyme composition of PDH from experiments (figure 7A) [19].

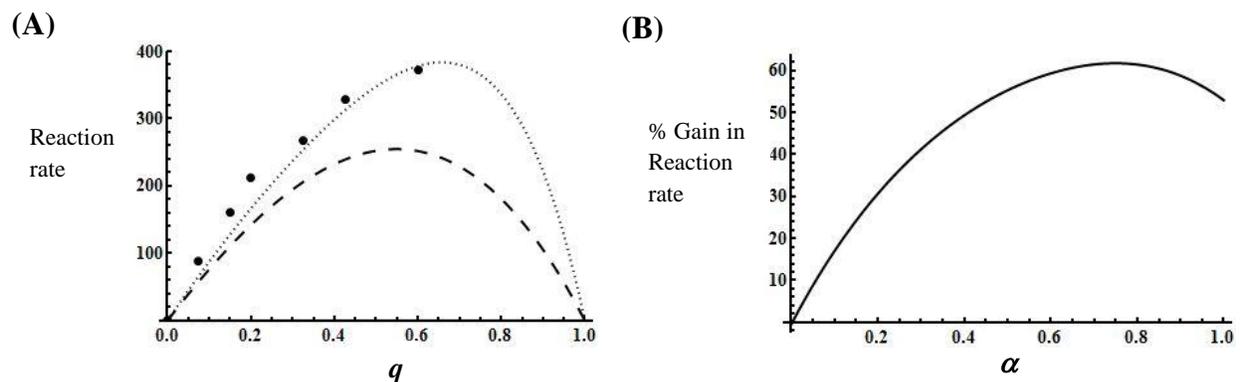


Figure 7. A comparison of reaction rates with and without enzyme sharing between the vertices of the PDH complex are shown in these plots. Panel A shows a plot of the reaction rates ρ_s (dotted) (equation 8) and ρ' (dashed) (equation 3) as a function of q_s , which depends on the relative concentrations of E1 and E3. For ρ_s , the value of the sharing factor α is 0.75. Experimental values for rate as a function of relative enzyme composition (filled circles) [19] are also plotted. The rate is $v_0/[E_T]$ (sec^{-1}), where v_0 is the initial velocity and $[E_T]$ is the total enzyme concentration. In panel B, the % gain in reaction rate of ρ_s over ρ' as a function of the sharing factor α is plotted. Here, for α between 0 and 1, the optimal values for ρ_s were recalculated and gain in reaction rate compared to ρ' .

3. Application of the RSS model for the simulation of metabolic pathways in a computer algebra system

The RSS model is developed in a MathematicaTM package that facilitates the generation of rate equations for enzyme complexes. As described in the derivation section, in the package, the rate ρ (equation 1) is approximated by the function $\tilde{\rho}$ (equation 5). Then the rate with enzyme sharing over the complex ρ_s (equation 8) is derived. In addition, an interpolation table of rate values that correspond to ten substrate concentrations over a range of concentrations from zero to complete enzyme saturation is used. For any combination of selected substrate concentrations in the interpolation table, the approximating function is generated and the final rate equation is derived. The values for rate and substrate concentrations from the interpolation table are fitted into a rate function with the substrate concentrations as variables. Therefore, for given values of enzyme concentration and rate constants for an enzyme complex-catalyzed reaction, the differential equations describing the rate of change of substrates and products are generated.

To simulate large metabolic networks, the rate equations resulting from the running the RSS package are integrated into the ordinary differential equations from kMech and Cellerator described earlier [10, 13]. This allows for the simultaneous simulation of the models of the enzyme complexes with the kMech/Cellerator enzyme models of metabolic pathways. The RSS package and MathematicaTM notebooks for the application of the RSS model in the simulation of the central metabolic pathways and pyruvate and aspartate family amino acids are available upon request.

4. Conclusion

We have defined a Random Steady State (RSS) model to describe catalysis by large multienzyme complexes and applied it to the PDH and KGDH complexes of *Escherichia coli*. Unlike the usual quasi-equilibrium approximation which is stochastic and in thermal equilibrium at a fast time scale and deterministic at a slower time scale, the RSS model (1) describes a steady state process such as substrate channeling between active sites at a fast time scale by solving the Master Equation for the steady state of a small system analytically and then approximating the solution with a polynomial; and also (2) describes an equilibrium process such as binding/unbinding of enzyme subunits to the core, at a slower time scale. The model is further shown to agree with experimental data in these molecular complexes when sharing of substrates over the enzyme complex is incorporated. Further improvements in the RSS method would allow for nontrivial vertex-vertex interactions at the coarse time scale modeled by equilibrium partition functions.

In addition to the PDH and KGDH complexes, other closely related 2-oxo acid dehydrogenase complexes catalyze reactions located at strategic points in metabolic pathways. These complexes exhibit similar mechanisms of catalysis. They include the branched-chain dehydrogenase complex, the glycine dehydrogenase complex, and the acetoin dehydrogenase complex. The RSS model could in principle also be applied to describe enzymatic catalysis by these complexes and potentially many other cellular enzyme complexes with multiple active sites that channel metabolic intermediates. The model is also expandable to 2-oxo acid dehydrogenase complexes in different species which differ in the number and structural organization of enzymes in their complexes. The RSS model, therefore, demonstrates a method for

modeling multienzyme complexes that can help explore the role of such mechanisms in regulating metabolic pathways.

References

- [1] Reed L J 1973 Multienzyme Complexes *Acc Chem Res* **7** 40-6
- [2] Perham R N 1975 Self-assembly of biological macromolecules *Philos Trans R Soc Lond B Biol Sci* **272** 123-36
- [3] Yang Y S, Datta A, Hainfeld J F, Furuya F R, Wall J S, Frey P A 1994 Mapping the lipoyl groups of the pyruvate dehydrogenase complex by use of gold cluster labels and scanning transmission electron microscopy *Biochemistry* **33** 9428-37
- [4] Murphy G E, Jensen G J 2005 Electron cryotomography of the E. coli pyruvate and 2-oxoglutarate dehydrogenase complexes *Structure* **13** 1765-73
- [5] Reed L J, Pettit F H, Eley M H, Hamilton L, Collins J H, Oliver R M 1975 Reconstitution of the Escherichia coli pyruvate dehydrogenase complex *Proc Natl Acad Sci U S A* **72** 3068-72
- [6] de Kok A, Hengeveld A F, Martin A, Westphal A H 1998 The pyruvate dehydrogenase multi-enzyme complex from Gram-negative bacteria *Biochim Biophys Acta* **1385** 353-66
- [7] Bosma H J, de Kok A, Westphal A H, Veeger C 1984 The composition of the pyruvate dehydrogenase complex from Azotobacter vinelandii. Does a unifying model exist for the complexes from gram-negative bacteria? *Eur J Biochem* **142** 541-9
- [8] Hale G, Perham R N 1979 Polypeptide-chain stoichiometry and lipoic acid content of the pyruvate dehydrogenase complex of Escherichia coli *Biochem J* **177** 129-36
- [9] Dave E, Guest J R, Attwood M M 1995 Metabolic engineering in Escherichia coli: lowering the lipoyl domain content of the pyruvate dehydrogenase complex adversely affects the growth rate and yield *Microbiology* **141** 1839-49
- [10] Yang C R, Shapiro B E, Mjolsness E D, Hatfield G W 2005 An enzyme mechanism language for the mathematical modeling of metabolic pathways *Bioinformatics* **21** 774-80
- [11] Yang C R, Shapiro B E, Hung S P, Mjolsness E D, Hatfield G W 2005 A mathematical model for the branched chain amino acid biosynthetic pathways of Escherichia coli K12 *J Biol Chem* **280** 11224-32
- [12] Najdi T S, Yang C R, Shapiro B E, Hatfield G W, Mjolsness E D 2006 Application of a generalized MWC model for the mathematical simulation of metabolic pathways regulated by allosteric enzymes *J Bioinform Comput Biol* **4** 335-55
- [13] Shapiro B E, Levchenko A, Meyerowitz E M, Wold B J, Mjolsness E D 2003 Cellerator: extending a computer algebra system to include biochemical arrows for signal transduction simulations *Bioinformatics* **19** 677-8
- [14] Mjolsness E 2007 On cooperative quasi-equilibrium models of transcriptional regulation *J Bioinform Comput Biol* **5** 467-90
- [15] Kholodenko B N, Schuster S, Garcia J, Westerhoff H V, Cascante M 1998 Control analysis of metabolic systems involving quasi-equilibrium reactions *Biochim Biophys Acta* **1379** 337-52
- [16] Akiyama S K, Hammes G G 1980 Elementary steps in the reaction mechanism of the pyruvate dehydrogenase multienzyme complex from Escherichia coli: kinetics of acetylation and deacetylation *Biochemistry* **19** 4208-13
- [17] Sahlman L, Williams C H, Jr. 1989 Lipoamide dehydrogenase from Escherichia coli. Steady-state kinetics of the physiological reaction *J Biol Chem* **264** 8039-45
- [18] Graham L D, Packman L C, Perham R N 1989 Kinetics and specificity of reductive acylation of lipoyl domains from 2-oxo acid dehydrogenase multienzyme complexes *Biochemistry* **28** 1574-81

- [19] Danson M J, Fersht A R, Perham R N 1978 Rapid intramolecular coupling of active sites in the pyruvate dehydrogenase complex of *Escherichia coli*: mechanism for rate enhancement in a multimeric structure *Proc Natl Acad Sci U S A* **75** 5386-90
- [20] van Kampen N G. 1981 *Stochastic Processes in Physics and Chemistry*. Amsterdam: North-Holland
- [21] Mjolsness E 2007 Towards a calculus of biomolecular complexes at equilibrium *Brief Bioinform* **8** 226-33
- [22] Graham R L, Knuth D E, Patashnik O. 1994 *Concrete Mathematics: A Foundation for Computer Science*. 2nd ed. Reading, Massachusetts: Addison-Wesley Professional