

STOCHASTIC MODELS FOR AN OPEN BIOCHEMICAL SYSTEM*

SANDRA HASSTEDT**

Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48109, U.S.A.

(Received June 1st, 1978)

This paper uses the theory of Markov processes to derive stochastic models for a single open biochemical system at steady state under 3 sets of assumptions. The system is a one substrate, one product reaction. Each set of assumptions results in a separate solution for the probability functions. A system of linear equations in the probability function as well as an equivalent differential equation in its generating function are derived. The assumption of no flux leads to the first (exact) solution of the linear equations. The form agrees with that of the closed systems. Making assumptions that simplify the system to model active transport results in the second (exact) solution to the linear equations. Assuming the presence of a large number of molecules in the system facilitates obtaining the third (approximate) solution to the differential equations.

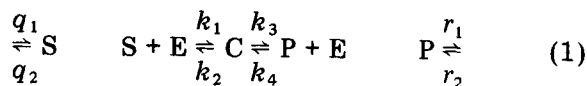
1. Introduction

Biochemical systems, often studied as closed systems in the lab, exist as open systems in nature. The validity of the inference about in vivo enzyme action from closed system studies depends on the degree of equivalence between the systems. In vivo and in vitro systems do not always behave identically (Stuart and Branscomb, 1971). More realistic inference results from using in vivo data and an open system model.

Deterministic models predominate among studies of open systems, as well as closed systems. Complexities in their development lead to approaches using systems theory (Mesarovic, 1968; Savageau, 1969a, b) and computer models (Garfinkel, 1965) to predict quantitative behavior. Crossover plots (Chance et al., 1958; Williamson, 1970) provide a method for locating regulatory sites in a pathway. None of these approaches incorporate random fluctuations and variation in the concentrations of the reactants. Only stochastic models allow for variation.

This paper presents a stochastic treatment

of the open biochemical system defined by the model:



S, E, C and P represent free substrate, free enzyme, enzyme-substrate (or enzyme-product) transition complex and free product, respectively. Substrate and product enter the system at constant rates (q_1 and r_2); they leave at rates proportional to the concentration of each (q_2s and r_1p). The small letters, s , p , c and e , symbolize concentrations of the reactants represented by corresponding large letters. The system disallows the removal of enzyme and complex. Therefore,

$$e + c = E_t \quad (2)$$

where E_t is the total quantity of enzyme present. Within the system, a reversible Michaelis-Menten mechanism operates with rate constants k_1 , k_2 , k_3 and k_4 .

Other investigators consider stochastic models of open biochemical systems (Smith, 1971; Smeach and Smith, 1973; Smeach and Gold, 1975a, b). Focusing on the problem as presented here, Smeach and Smith (1973) develop a model for active transport of substrate into the cell for comparison with a

*Supported by NIH training grant 5T01-GM-711.

**Present address: Department of Medical Biophysics and Computing, University of Utah Medical Center, Salt Lake City, Utah 84143, U.S.A.

deterministic model. They also employ the theory of Markov processes (Feller, 1968).

The treatment of the system presented here derives the exact probability function for 2 sets of restrictions on the parameters and an approximate probability density function for the case without restrictions. It appears very difficult, or impossible, to obtain a solution without restrictions. The first exact solution assumes no net flux through the system ($q_1 r_1 k_1 k_3 = q_2 r_2 k_2 k_4$). The result takes the same form as the probability function for the closed system at equilibrium (Hasstedt, 1978). The assumptions for the second exact solution, irreversibility and only 1 enzyme molecule, simplify the active transport model considered by Smeach and Smith (1973). The third case, assuming the presence of a large number of molecules in the system, shows the generating function to be approximately that of the multivariate normal distribution. The basic set of equations used for each of the 3 cases is derived in the next section.

2. Derivation of the equations

Deriving the joint probability function for

Conversion	Probability
$(s-1, p, c) \rightarrow (s, p, c)$	$q_1 \Delta t + o(\Delta t)$
$(s+1, p, c) \rightarrow (s, p, c)$	$q_2 (s+1) \Delta t + o(\Delta t)$
$(s+1, p, c-1) \rightarrow (s, p, c)$	$k_1 (s+1) (E_t - c + 1) \Delta t + o(\Delta t)$
$(s-1, p, c+1) \rightarrow (s, p, c)$	$k_2 (c+1) \Delta t + o(\Delta t)$
$(s, p-1, c+1) \rightarrow (s, p, c)$	$k_3 (c+1) \Delta t + o(\Delta t)$
$(s, p+1, c-1) \rightarrow (s, p, c)$	$k_4 (p+1) (E_t - c + 1) \Delta t + o(\Delta t)$
$(s, p+1, c) \rightarrow (s, p, c)$	$r_1 (p+1) \Delta t + o(\Delta t)$
$(s, p-1, c) \rightarrow (s, p, c)$	$r_2 \Delta t + o(\Delta t)$
$(s, p, c) \rightarrow (s, p, c)$	$1 - (q_1 + q_2 s + k_1 s (E_t - c) + k_2 c + k_3 c + k_4 p (E_t - c) + r_1 p + r_2) \Delta t + o(\Delta t)$

the concentrations of the reactants in system (1) for each of the 3 cases assumes:

(1) s , p , c and e take on integer values only (reactant concentrations are expressed in molecules).

(2) Any future state of the system (set of values for s , p , c and e) depends only on the present state and not on the history of the system (the Markov property).

(3) The probability of a conversion from one reactant or set of reactants to another is proportional to the concentrations of the reactants making the conversion.

(4) The probability of more than 1 conversion in a time

unit Δt is $o(\Delta t)$ where $\frac{o(\Delta t)}{\Delta t} \rightarrow 0$ as $\Delta t \rightarrow 0$

Equation (2) may be rewritten to express e in terms of c . Thus, the 3 concentrations, s , p and c , alone describe the state of the system. Each arrow in system (1) represents a conversion from one reactant or set of reactants to another, and, consequently, from one state to another. Using the assumptions above, all possible conversions to the state (s, p, c) and their respective probabilities are:

From the list one obtains the forward Kolmogorov equations. (See Hasstedt, 1978 for details.) During the steady state phase of the reaction ($t = \infty$), the probabilities do not change with time. Setting the time derivative to zero leaves the system of linear equations in the unknown probabilities:

$$\begin{aligned}
 & (q_1 + q_2s + k_1s(E_t - c) + k_2c + k_3c \\
 & + k_4p(E_t - c) + r_1p + r_2) P(s, p, c) \\
 & = (1 - \delta_{s, 0})q_1 P(s - 1, p, c) + q_2 (s + 1) \\
 & \times P(s + 1, p, c) + (1 - \delta_{c, 0})k_1 (s + 1) \\
 & \times (E_t - c + 1) P(s + 1, p, c - 1) \\
 & + (1 - \delta_{c, E_t}) (1 - \delta_{s, 0})k_2 (c + 1) \\
 & \times P(s - 1, p, c + 1) + (1 - \delta_{c, E_t}) \quad (4) \\
 & \times (1 - \delta_{p, 0})k_3 (c + 1) P(s, p - 1, c + 1) \\
 & + (1 - \delta_{c, 0})k_4 (p + 1) (E_t - c + 1) \\
 & \times P(s, p + 1, c - 1) + r_1 (p + 1) \\
 & \times P(s, p + 1, c) + (1 - \delta_{p, 0})r_2 \\
 & \times P(s, p - 1, c) \quad (s, p, c) \in \Omega
 \end{aligned}$$

where $\delta_{i, j}$ is the Kronecker delta,

$$\delta_{i, j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \quad (5)$$

$$\Omega = \{(s, p, c): s = 0, 1, \dots, \\ p = 0, 1, \dots, c = 0, 1, \dots, E_t\} \quad (6)$$

and t no longer enters into the probabilities.

A probability function, describing the concentrations of the reactants in system (1), satisfies the equations in (4). Each unknown probability corresponds to one equation. Although one dependency occurs, replacing one equation with the requirement that the

probabilities sum to one results in a linearly independent system of equations.

An alternative to solving the system of linear equations is solving a differential equation in the probability generating function (*pgf*):

$$\begin{aligned}
 \phi(x, y, z) &= \sum_{s=0}^{\infty} \sum_{p=0}^{\infty} \sum_{c=0}^{E_t} x^s y^p z^c P(s, p, c) \\
 &= E(x^s y^p z^c) \quad (7)
 \end{aligned}$$

where E indicates expectation. Multiplying each term in the equations of (4) by $x^s y^p z^c$ and summing over Ω (given in (6)), one obtains an equation in ϕ and various derivatives of ϕ :

$$\begin{aligned}
 [q_1 (1 - x) + r_2 (1 - y)]\phi &= [q_2 (1 - x) \\
 + E_t k_1 (z - x)] \frac{\partial \phi}{\partial x} \\
 + [r_1 (1 - y) + E_t k_4 (z - y)] \frac{\partial \phi}{\partial y} \quad (8) \\
 + [k_2 (x - z) + k_3 (y - z)] \frac{\partial \phi}{\partial z} \\
 + k_1 z (x - z) \frac{\partial^2 \phi}{\partial x \partial z} + k_4 z (y - z) \frac{\partial^2 \phi}{\partial y \partial z}.
 \end{aligned}$$

The *pgf*, ϕ , satisfying this equation is the *pgf* for the probability function satisfying the linear equations in (4).

The *pgf* (also called the factorial moment generating function) has the property that its $(i + j + m)$ th derivative when $x = y = z = 1$ is given by

$$\begin{aligned}
 \frac{\partial^{i+j+m} \phi}{\partial x^i \partial y^j \partial z^m} \Big|_{x=y=z=1} \\
 = E[s^{[i]} p^{[j]} c^{[m]}] \quad (9)
 \end{aligned}$$

$$i = 0, 1, \dots, j = 0, 1, \dots, m = 0, 1, \dots$$

where the expectation of the right is a factorial moment with

$$x^{[n]} = x(x - 1) \dots (x - n + 1), n = 1, 2, \dots \quad (10)$$

for any variable x . Taking derivatives of (8), setting $x = y = z = 1$ in the resulting equations and using (9) to specify terms containing ϕ , one obtains equations relating the parameters to the moments. The first and second derivatives yield 9 equations in the means (μ_s, μ_p, μ_c), the variances ($\sigma_s^2, \sigma_p^2, \sigma_c^2$), the covariances ($\rho_{sc}\sigma_s\sigma_c, \rho_{pc}\sigma_p\sigma_c, \rho_{sp}\sigma_s\sigma_p$) and five third moments. The equations resulting from the first derivatives are:

$$q_1 + k_2\mu_c = (q_2 + k_1(E_t - \mu_c))\mu_s - k_1\rho_{sc}\sigma_s\sigma_c$$

$$r_2 + k_3\mu_c = (r_1 + k_4(E_t - \mu_c))\mu_p - k_4\rho_{pc}\sigma_p\sigma_c$$

$$\times (k_1\mu_s + k_4\mu_p)(E_t - \mu_c) \quad (11)$$

$$= (k_2 + k_3)\mu_c + k_1\rho_{sc}\sigma_s\sigma_c + k_4\rho_{pc}\sigma_p\sigma_c$$

The equations resulting from the single variable second derivatives are:

$$(q_2 + k_1(E_t - \mu_c))(\sigma_s^2 - \mu_s)$$

$$= (k_1(\mu_s - 1) + k_2)\rho_{sc}\sigma_s\sigma_c$$

$$+ k_1 E[(s - \mu_s)^2(c - \mu_c)]$$

$$(r_1 + k_4(E_t - \mu_c))$$

$$\times (\sigma_p^2 - \mu_p) = (k_4(\mu_p - 1) + k_3)\rho_{pc}\sigma_p\sigma_c$$

$$+ k_4 E[(p - \mu_p)^2(c - \mu_c)] \quad (12)$$

$$(k_2 + k_3)(\sigma_c^2 - \mu_c) + (k_1\mu_s + k_4\mu_p)\sigma_c^2$$

$$= (k_1\rho_{sc}\sigma_s\sigma_c + k_4\rho_{pc}\sigma_p\sigma_c)(E_t - \mu_c)$$

$$- k_1 E[(s - \mu_s)(c - \mu_c)^2]$$

$$- k_4 E[(p - \mu_p)(c - \mu_c)^2]$$

And the equations resulting from the mixed second derivatives are:

$$(q_2 + r_1 + (k_1 + k_4)(E_t - \mu_c))\rho_{sp}\sigma_s\sigma_p$$

$$= (k_1\mu_s + k_2)\rho_{pc}\sigma_p\sigma_c + (k_4\mu_p + k_3)\rho_{sc}\sigma_s\sigma_c$$

$$+ (k_1 + k_4) E[(s - \mu_s)(p - \mu_p)(c - \mu_c)]$$

$$(q_2 + k_1(E_t - \mu_c) + k_1(\mu_s - 1) + k_2$$

$$+ k_3 + k_4\mu_p)\rho_{sc}\sigma_s\sigma_c$$

$$= k_1(E_t - \mu_c)(\sigma_s^2 - \mu_s) + k_1\mu_s\sigma_c^2$$

$$+ k_2(\sigma_c^2 - \mu_c) + k_4(E_t - \mu_c)\rho_{sp}\sigma_s\sigma_p$$

$$- k_1 E[(s - \mu_s)^2(c - \mu_c)] + k_1 E \quad (13)$$

$$\times [(s - \mu_s)(c - \mu_c)^2] - k_4 E[(s - \mu_s)$$

$$\times (p - \mu_p)(c - \mu_c)]$$

$$(r_1 + k_4(E_t - \mu_c))$$

$$+ k_4(\mu_p - 1) + k_2 + k_3 + k_1\mu_s)\rho_{pc}\sigma_p\sigma_c$$

$$= k_4(E_t - \mu_c)(\sigma_p^2 - \mu_p) + k_4\mu_p\sigma_c^2$$

$$+ k_3(\sigma_c^2 - \mu_c) + k_1(E_t - \mu_c)\rho_{sp}\sigma_s\sigma_p$$

$$- k_4 E[(p - \mu_p)^2(c - \mu_c)] + k_4$$

$$\times E[(p - \mu_p)(c - \mu_c)^2]$$

$$- k_1 E[(s - \mu_s)(p - \mu_p)(c - \mu_c)]$$

Flux in an open system expresses the equivalent of velocity in a closed system. The expected flux through each individual step in the system is the difference between the average forward and backward rates. At steady state, each step has the same expected flux:

$$\text{flux} = q_1 - q_2\mu_s$$

$$= k_1[\mu_s(E_t - \mu_c) - \rho_{sc}\sigma_s\sigma_c] - k_2\mu_c$$

$$= k_3\mu_c - k_4[\mu_p(E_t - \mu_c) - \rho_{pc}\sigma_p\sigma_c]$$

$$= r_1\mu_p - r_2 \quad (14)$$

The 3 equations in (11) confirm these equivalences. The flux for the whole reaction cannot exceed the maximum forward rate of any one step. These bounds are finite for the first and third steps, q_1 and k_3E_t , respectively. Therefore,

$$\text{flux} \leq \min(q_1, k_3E_t). \quad (15)$$

The second bound, $k_3 E_t$, equals V_{\max} , the maximum velocity of a closed system.

The next 3 sections use the results derived in this section. They solve the linear equations in (4) or the differential equation in (8) to obtain a probability function for each of 3 sets of restrictions imposed on the parameters.

3. Absence of flux

The first exact probability function applies when

$$q_1 r_1 k_1 k_3 = q_2 r_2 k_2 k_4. \tag{16}$$

This condition results in flux equal to zero. This situation may never occur in nature. It is presented here only for comparison with the other 2 solutions and as the link between closed systems (Hasstedt, 1978) and open systems.

The probability function which satisfies the linear eqns. (4) assuming condition (16) is

$$\begin{aligned}
 P(s,p,c) = \exp & \left[\frac{q_1}{q_2} - \frac{q_1 k_1 k_3}{q_2 k_2 k_4} \right] \left(\frac{E_t}{c} \right) \frac{1}{s! p!} \\
 & \times \left(\frac{q_1}{q_2} \right)^s \left(\frac{q_1 k_1 k_3}{q_2 k_2 k_4} \right)^p \left(\frac{q_1 k_1}{q_1 k_1 + q_2 k_2} \right)^c \\
 & \times \left(\frac{q_2 k_2}{q_1 k_1 + q_2 k_2} \right)^{E_t - c} \quad (s,p,c) \in \Omega \tag{17}
 \end{aligned}$$

where Ω is given in (6). The random variables s and p enter the joint probability function as Poisson functions with moments:

$$\mu_s = \sigma_s^2 = \frac{q_1}{q_2} \quad \text{and} \quad \mu_p = \sigma_p^2 = \frac{q_1 k_1 k_3}{q_2 k_2 k_4} \tag{18}$$

and c enters as a binomial function with moments:

$$\mu_c = \frac{E_t q_1 k_1}{q_1 k_1 + q_2 k_2} \quad \text{and} \quad \sigma_c^2 = \frac{E_t q_1 q_2 k_1 k_2}{(q_1 k_1 + q_2 k_2)^2} \tag{19}$$

The random variables, s , p and c , are stochastically independent and therefore uncorrelated. The equivalent *pgf*, satisfying the differential equation in (8) when (16) holds, is

$$\begin{aligned}
 \phi(x,y,z) = \exp & \left[\frac{q_1}{q_2} (x - 1) + \frac{q_1 k_1 k_3}{q_2 k_2 k_4} (y - 1) \right] \\
 & \times \left(\frac{q_1 k_1 z + q_2 k_2}{q_1 k_1 + q_2 k_2} \right)^{E_t} \tag{20}
 \end{aligned}$$

As written, r_1 and r_2 do not appear in the probability function (17) or *pgf* (20); each function can be rewritten, however, using the relationship in (16), to include r_1 and r_2 and eliminate a different pair of parameters, Flux is found to equal zero, computing it from definition (14) and the moments in (18) and (19)

4. Active transport

The assumptions for the second solution simplify the active transport model considered by Smeach and Smith (1973). Their model corresponds to system (1) with $E_t = 1$ and $k_4 = 0$. P represents substrate within the cell; they assume it cannot leave once it enters ($k_4 = 0$). They model only one compartment of the cell membrane, containing a single enzyme molecule ($E_t = 1$). Here, the additional assumption, $q_2 = k_2 = 0$, makes the system completely irreversible. The joint probability function for s and c is derived under these conditions; p is summed out to simplify the derivation.

Letting $q_2 = k_2 = k_4 = 0$ and $E_t = 1$ in eqns. (4) and summing p out, the system of linear equations reduces to

$$\begin{aligned}
 & (q_1 + k_1 s(1 - c) + k_3 c) P(s,c) \\
 & = (1 - \delta_{s,0}) q_1 P(s - 1, c) \\
 & + (1 - \delta_{c,0}) k_1 (s + 1) P(s + 1, c - 1) \tag{21} \\
 & + (1 - \delta_{c,1}) k_3 P(s, c + 1) \\
 & s = 0, 1, \dots, c = 0, 1,
 \end{aligned}$$

where $\delta_{i,j}$ is given in (5). This simplification facilitates recursive solution of the equations. Using this method shows the joint probability function for s and c to be

$$\begin{aligned}
 P(s,c) = & \exp \left[-\frac{q_1}{k_1} \right] \left(\frac{k_3 - q_1}{k_3} \right)^{(q_1/k_1) + 1} \\
 & \times \left(\frac{q_1}{k_3} \right)^c \left(\frac{q_1}{k_1} \right)^s \sum_{i=0}^s \frac{1}{i! (s-i)!} \\
 & \times \left(\frac{k_1}{k_3} \right)^j \left(\frac{q_1}{k_1} + i + c - 1 \right)^{[i]} \quad (22) \\
 & s = 0, 1, \dots, c = 0, 1,
 \end{aligned}$$

where the factorial term $x^{[n]}$ is given in (10). The 2 random variables are not stochastically independent for this case. The marginal for c is again binomial (or Bernoulli since $E_t = 1$) with parameter $q_1/k_3 \leq 1$,

$$P(c) = \left(\frac{q_1}{k_3} \right)^c \left(\frac{k_3 - q_1}{k_3} \right)^{1-c}, \quad c = 0, 1. \quad (23)$$

However, the marginal for s differs from the last case and is not, a standard form. The moments of the probability function are:

$$\begin{aligned}
 \mu_s &= \frac{q_1 (q_1 k_1 + k_3^2)}{k_1 k_3 (k_3 - q_1)}, \quad \mu_c = \frac{q_1}{k_3}, \\
 \rho_{sc} \sigma_s \sigma_c &= \left(\frac{q_1}{k_3} \right)^2 \\
 \sigma_s^2 &= \frac{q_1 (q_1 k_1 + k_3^2)}{k_1 k_3 (k_3 - q_1)} - \left(\frac{q_1}{k_3} \right)^2, \quad (24) \\
 \sigma_c^2 &= \frac{q_1 (k_3 - q_1)}{k_3^2}.
 \end{aligned}$$

The *pgf* which satisfies (8) when $y = 1$ (p summed out) and $q_2 = k_2 = k_4 = 0$ and $E_t = 1$ (the conditions for this case) is

$$\begin{aligned}
 \phi(x, 1, z) = & \exp \left[\frac{q_1 (x - 1)}{k_1} \right] \left(\frac{k_3 + q_1 z - q_1 x}{k_3} \right) \\
 & \times \left(\frac{k_3 - q_1}{k_3 - q_1 x} \right)^{(q_1/k_1) + 1} \quad (25)
 \end{aligned}$$

The first part of the function is the Poisson *pgf* for s ; the second is the Bernoulli *pgf* for c if $x = 1$; the third part is not a standard *pgf*. (Johnson and Kotz [1969] discuss the Poisson and Bernoulli distributions.)

The flux equals q_1 (from equation (14) with $q_2 = 0$). It must be less than or equal to k_3 . Equation (15) states that k_3 is an upper bound on the flux and the relationship $q_1/k_3 \leq 1$ is required for the Bernoulli parameter. The previous investigators derive the rate of transport of substrate into the cell (flux) for their model. Since their result contains q_2 in the denominator, it cannot be evaluated when $q_2 = 0$ to compare with the flux here.

5. Limiting case

This section derives a continuous approximation to the discrete probability function describing system (1) It accomplishes this by showing that the generating function for a trivariate normal distribution satisfies the differential equation in the limit. The limiting form depends on assumptions about the magnitudes of the parameters and the moments.

Converting equation (8), the differential equation in the *pgf*,

$$\phi(x, y, z) = E[x^s y^p z^c], \quad (26)$$

to a differential equation in the standardized moment generating function,

$$\psi(t, v, w) = \mathbb{E} \left[\exp \left\{ t \frac{s - \mu_s}{\sigma_s} + v \frac{p - \mu_p}{\sigma_p} + w \frac{c - \mu_c}{\sigma_c} \right\} \right] \quad (27)$$

and using the relationships in the 3 equations in (11) one obtains

$$\begin{aligned} & [q_2 \mu_s (1 - e^{-t/\sigma_s}) (1 - e^{t/\sigma_s}) + k_2 \mu_c \\ & \times (1 - e^{-w/\sigma_c}) (e^{t/\sigma_s} - e^{w/\sigma_p}) \\ & + k_1 \mu_s (E_t - \mu_c) (1 - e^{-t/\sigma_s}) \\ & \times (e^{w/\sigma_c} - e^{t/\sigma_s}) + k_1 \rho_{sc} \sigma_s \sigma_c \\ & \times (e^{t/\sigma_s} - e^{w/\sigma_c}) + k_4 \rho_{pc} \sigma_p \sigma_c \\ & \times (e^{v/\sigma_p} - e^{w/\sigma_p}) + k_4 \mu_p (E_t - \mu_c) \\ & \times (1 - e^{-v/\sigma_p}) (e^{w/\sigma_c} - e^{v/\sigma_p}) \\ & + r_1 \mu_p (1 - e^{-v/\sigma_p}) (1 - e^{v/\sigma_p}) \\ & + k_3 \mu_c (1 - e^{-w/\sigma_c}) (e^{v/\sigma_p} - e^{w/\sigma_c})] \psi \\ = & \sigma_s e^{-t/\sigma_s} [q_2 (1 - e^{t/\sigma_s}) + k_1 \\ & \times (E_t - \mu_c) (e^{w/\sigma_c} - e^{t/\sigma_s})] \frac{\partial \psi}{\partial t} \\ & + \sigma_p e^{-v/\sigma_p} [r_1 (1 - e^{v/\sigma_p}) \\ & + k_4 (E_t - \mu_c) (e^{w/\sigma_c} - e^{v/\sigma_p})] \frac{\partial \psi}{\partial v} \quad (28) \\ & + \sigma_c [k_1 \mu_s e^{-t/\sigma_s} (e^{t/\sigma_s} - e^{w/\sigma_c}) \\ & + k_2 e^{-w/\sigma_c} (e^{t/\sigma_s} - e^{w/\sigma_c}) \\ & + k_3 e^{-w/\sigma_c} (e^{v/\sigma_p} - e^{w/\sigma_c}) \\ & + k_4 \mu_p e^{-v/\sigma_p} (e^{v/\sigma_p} - e^{w/\sigma_c})] \frac{\partial \psi}{\partial w} \\ & + k_1 \sigma_s \sigma_c e^{-t/\sigma_s} (e^{t/\sigma_s} - e^{w/\sigma_c}) \frac{\partial^2 \psi}{\partial t \partial w} \\ & + k_4 \sigma_p \sigma_c e^{-v/\sigma_p} (e^{v/\sigma_p} - e^{w/\sigma_c}) \frac{\partial^2 \psi}{\partial v \partial w} \end{aligned}$$

This is still exact; the solution is the moment generating function for the probability function satisfying (4) after standardization.

The approximation assumes a system containing a large number of molecules. That is, s , p and E_t must be large. To attain this, the assumption is made that M , some measure of the number of molecules, goes to ∞ . Then, as $M \rightarrow \infty$, the zero order rate constants (q_1 , r_2) and E_t go to ∞ , the first order rate constants (q_2 , k_2 , k_3 , r_1) stay constant and the second order rate constants (k_1 , k_4) converge to zero. This implies that the terms which are bounded as $M \rightarrow \infty$ are

$$q_1/M, q_2, k_1 M, k_2, k_3, k_4 M, r_1,$$

$$r_2/M \text{ and } E_t/M.$$

These assumptions indicate that all rates of individual steps in the system are of the same order of magnitude. The means and variances are assumed to increase with M and the third moments to increase more slowly than M^2 , analogous to diffusion approximations. The assumptions for the moments are

$$\mu_s/M, \mu_p/M, \mu_c/M, \sigma_s^2/M,$$

$$\sigma_p^2/M \text{ and } \sigma_c^2/M$$

converge to constants as $M \rightarrow \infty$ and

$$\mathbb{E}[(s - \mu_s)^2 (c - \mu_c)]/M^2,$$

$$\mathbb{E}[(p - \mu_p)^2 (c - \mu_c)]/M^2,$$

$$\mathbb{E}[(s - \mu_s) (c - \mu_c)^2]/M^2,$$

$$\mathbb{E}[(p - \mu_p) (c - \mu_c)^2]/M^2$$

and

$$\mathbb{E}[(s - \mu_s) (p - \mu_p) (c - \mu_c)]/M^2$$

converge to zero as $M \rightarrow \infty$.

Taking the limit in the differential equation given in (28) as $M \rightarrow \infty$, results in the reduced

differential equation:

$$\begin{aligned}
 &[-q_2\mu_s t^2/\sigma_s^2 + k_1\mu_s (E_t - \mu_c)t/\sigma_s \\
 &\quad \times (w/\sigma_c - t/\sigma_s) \\
 &\quad + k_2\mu_c w/\sigma_c (t/\sigma_s - w/\sigma_c) \\
 &\quad + k_3\mu_c w/\sigma_c (v/\sigma_p - w/\sigma_c) \\
 &\quad + k_4\mu_p (E_t - \mu_c)v/\sigma_p (w/\sigma_c - v/\sigma_p) \\
 &\quad - r_1\mu_p v^2/\sigma_p^2] \psi \tag{29} \\
 = &\sigma_s[-q_2 t/\sigma_s + k_1 (E_t - \mu_c) \\
 &\quad \times (w/\sigma_c - t/\sigma_s) \frac{\partial \psi}{\partial t} \\
 &\quad + \sigma_p[-r_1 v/\sigma_p + k_4 (E_t - \mu_c) \\
 &\quad \times (w/\sigma_c - v/\sigma_p)] \frac{\partial \psi}{\partial v} \\
 &\quad + \sigma_c[k_1\mu_s (t/\sigma_s - w/\sigma_c) \\
 &\quad + k_2 (t/\sigma_s - w/\sigma_c) + k_3 (v/\sigma_p - w/\sigma_c) \\
 &\quad + k_4\mu_p (v/\sigma_p - w/\sigma_c)] \frac{\partial \psi}{\partial w}
 \end{aligned}$$

The moment generating function for a trivariate normal distribution,

$$\begin{aligned}
 \psi(t, v, w) = &\exp[t^2/2 + v^2/2 + w^2/2 \\
 &\quad + \rho_{sp}tv + \rho_{sc}tw + \rho_{pc}vw], \tag{30}
 \end{aligned}$$

satisfies this equation if

$$\begin{aligned}
 &(q_2 + k_1 (E_t - \mu_c)) (\sigma_s^2 - \mu_s) \\
 &= (k_1\mu_s + k_2) \rho_{sc}\sigma_s\sigma_c \\
 &(r_1 + k_4 (E_t - \mu_c)) (\sigma_p^2 - \mu_p) \tag{31} \\
 &= (k_4\mu_p + k_3)\rho_{pc}\sigma_p\sigma_c \\
 &(k_1\mu_s + k_2 + k_3 + k_4\mu_p)\sigma_c^2 \\
 &= (k_2 + k_3)\mu_c + (k_1\rho_{sc}\sigma_s\sigma_c + k_4\rho_{pc}\sigma_p\sigma_c) \\
 &\quad \times (E_t - \mu_c)
 \end{aligned}$$

and

$$\begin{aligned}
 &(q_2 + k_1 (E_t - \mu_c) + k_1\mu_s + k_2 + k_3 + k_4\mu_p) \\
 &\quad \times \rho_{sc}\sigma_s\sigma_c = k_1 (E_t - \mu_c) (\sigma_s^2 - \mu_s) \\
 &\quad + (k_1\mu_s + k_2)\sigma_c^2 - k_2\mu_c \\
 &\quad + k_4 (E_t - \mu_c)\rho_{sp}\sigma_s\sigma_p \\
 &(r_1 + k_4 (E_t - \mu_c) + k_1\mu_s + k_2 + k_3 + k_4\mu_p) \\
 &\quad \times \rho_{pc}\sigma_p\sigma_c = k_4 (E_t - \mu_c) (\sigma_p^2 - \mu_p) \tag{32} \\
 &\quad + (k_4\mu_p + k_3)\sigma_c^2 - k_3\mu_c \\
 &\quad + k_1 (E_t - \mu_c) \rho_{sp}\sigma_s\sigma_p \\
 &(q_2 + r_1 + (k_1 + k_4) (E_t - \mu_c))\rho_{sp}\sigma_s\sigma_p \\
 &= (k_1\mu_s + k_2)\rho_{pc}\sigma_p\sigma_c + (k_4\mu_p + k_3)\rho_{sc}\sigma_s\sigma_c
 \end{aligned}$$

The equations in (31) and (32) are the limits of the equations in (12) and (13). Taking the limit in the 3 equations of (11) finishes specification of the relationships between the moments as

$$\begin{aligned}
 q_1 + k_2\mu_c &= (q_2 + k_1 (E_t - \mu_c))\mu_s \\
 r_2 + k_3\mu_c &= (r_1 + k_4 (E_t - \mu_c))\mu_p \tag{33} \\
 (k_1\mu_s + k_4\mu_p) (E_t - \mu_c) &= (k_2 + k_3)\mu_c
 \end{aligned}$$

The 9 equations in (31), (32) and (33) contain 9 unknowns. The functions of the parameters that satisfy these equations are the moments of the trivariate normal distribution. The equations in the means (33) correspond exactly to a deterministic model of system (1). Rolleston (1972) defines a property of an open system, a mass action ratio, as

$$\Gamma = \frac{\mu_p}{\mu_s} \leq K_{eq} = \frac{k_1 k_3}{k_2 k_4} \tag{34}$$

(expressed here as its stochastic model equivalent). Using this inequality in the equations shows that s and c are positively correlated and p and c are negatively correlated. ρ_{sp} can be either positive or negative. Also, from the equations,

$$\sigma_s^2 \geq \mu_s \quad \text{and} \quad \sigma_p^2 \leq \mu_p.$$

The variance equals the mean in the Poisson distribution (the solution for the no flux case). σ_c^2 exceeds or is less than the binomial variance depending on the magnitudes of the correlations.

6. Discussion

Much of the standard statistical theory rests on the assumption that the random variables are normally distributed. Accepting this assumption for a biochemical system validates the utilization of known procedures for estimation of parameters and testing of hypotheses. Therefore, the conclusion that the reactants in system (1) approximate normality provides the theoretical justification for using standard statistical procedures when the assumptions are met. Assuming normality also simplifies the incorporation of error terms (also assumed to be normal).

The approximate normality of biochemical systems applies generally. Although generalization of the method isn't apparent, one simply duplicates for other biochemical systems the procedures followed in approximating this simple system. Then each individual case approximates normality, dependent upon similar assumptions.

The magnitudes of the rate constants and concentrations found in vivo obey the assumptions made in the derivation. The concentrations and zero order rate constants are always large, first order rates intermediate and second order rate constants small (all when measured in molecules).

The moments for this system are extremely complicated functions of the rate constants

and enzyme concentration. Unfortunately, for more complicated biochemical systems, the complexity of the moments can only be expected to increase. Simplifying approximations are needed to derive estimation and testing procedures.

Acknowledgements

I would like to thank Dr. Peter Smouse, Dr. Michael Savageau, Dr. George Brewer, Dr. Michael Woodroffe and especially Dr. Charles Sing for reviewing and offering valuable, timely suggestions which improved this manuscript. I also wish to thank Dr. Jeffrey Weisberg for proposing needed revisions to increase the clarity of the language.

References

- Chance, B., Holmes, W., Higgins, J. and Connelly, C.M., 1958, Localization of interaction sites in multi-component transfer systems: theorems derived from analogues. *Nature* 182, 1190–1193.
- Feller, W., 1968, An introduction to probability theory and its application, Vol. 1. (John Wiley, New York).
- Garfinkel, D., 1965, Computer simulation in biochemistry and ecology, in: *Theoretical and Mathematical Biology*, eds. T.H. Waterman and H.J. Marowitz, (Blaisdell Publishing Co., New York) p. 292–310.
- Hasstedt, S., 1978, A Stochastic model for closed biochemical system at equilibrium. *J. Theor. Biol.* 70, 199–212.
- Johnson, N.L. and Kotz, S., 1969, *Discrete distributions*, Vol. 1 (Houghton Mifflin, Boston).
- Mesarovic, M.D., ed., 1968, *Systems Theory and Biology*. (Springer-Verlag New York).
- Rolleston, F.S., 1972, A theoretical background to the use of measured concentrations of intermediates in study of the control of intermediary metabolism, in: *Current Topics in Cellular Regulation*, Vol. 5, eds. B.L. Horecker and E.R. Stadtman, (Academic Press, New York), pp. 47–75.
- Savageau, M.A., 1969a, Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions. *J. Theor. Biol.* 25, 365–369.
- Savageau, M.A., 1969b, Biochemical system analysis. II. The steady-state solutions for an n-pool system

- using a power-law approximation. *J. Theor. Biol.* 25, 370–379.
- Smeach, S.C. and Gold, H.J., 1975a, Stochastic and deterministic models for the kinetic behavior of certain structured enzyme systems. I. One enzyme-substrate systems. *J. Theor. Biol.* 51, 59–78.
- Smeach, S.C. and Gold, H.J., 1975b, Stochastic and deterministic models for the kinetic behavior of certain structured enzyme systems. II. Consecutive two enzyme systems. *J. Theor. Biol.* 51, 79–96.
- Smeach, S.C. and Smith, W., 1973, A comparison of stochastic and deterministic models for cell membrane transport. *J. Theor. Biol.* 42, 157–167.
- Smith, W., 1971, Stochastic models for an enzyme reaction in an open linear system. *Bull. Math. Biophys.* 33, 97–115.
- Stuart, R.H. and Branscomb, E.W., 1971, Quantitative theory of in vivo lac regulation: significance of repressor packaging I. Equilibrium considerations. *J. Theor. Biol.* 31, 313–329.
- Williamson, J.R., 1970, General features of metabolic control as applied to the erythrocyte. in: *Red Cell Metabolism and Function*, ed. G.J. Brewer, (Plenum Press, New York) pp. 117–136.