J. theor. Biol. (1981) **93**, 157–177

Optimization of Kinetic Proofreading: A General Method for Derivation of the Constraint Relations and an Exploration of a Specific Case

MICHAEL A. SAVAGEAU† AND DAVID S. LAPOINTE

Department of Microbiology and Immunology, The University of Michigan, Ann Arbor, Michigan 48109, *U.S.A.*

(*Received* 17 *November* 1980, *and in revised form* 30 *April* 1981)

We have previously presented a general cost–accuracy relationship for a broad class of kinetic proofreading mechanisms. In this paper we present a general matrix method, based upon classical enzyme kinetics, for the derivation of the constraint relation that characterizes specific proofreading mechanisms. For purpose of illustration we present the method in the context of a conventional Michaelis–Menten mechanism with side reactions. We then explore optimization of the general cost function under a variety of different constraints that may exist for such a mechanism. In this way we are able to contrast different perspectives on the optimization of enzyme design.

1. Introduction

The specificity traditionally associated with enzyme catalysis is largely responsible for the overall accuracy of the recognition processes that occur at the molecular level in biological systems. However, there are physical limitations to the accuracy that can be achieved in this manner. In some cases, this structural specificity is not sufficient to account for the overall accuracy (Pauling, 1958; Loftfield & Eigner, 1966; Loftfield & Vanderjagt, 1972) and other accuracy enhancing mechanisms must be postulated.

Hopfield (1974) and Ninio (1975) were the first to propose specific kinetic mechanisms for proofreading that are capable of enhancing accuracy with an expenditure of energy. More recently, we have developed a general macroscopic theory of proofreading which, when applied to such kinetic proofreading mechanisms, yields an explicit relationship between accuracy improvement and energy cost (Savageau & Freter, 1979*a*; Freter & Savageau, 1980). We have show that the cost to achieve a given degree of accuracy in a multiple stage mechanism depends in large part upon the distribution of proofreading effort and proofreading discrimination among the stages, and we derived the general (or unconstrained) distribution that minimizes this cost (Freter & Savageau, 1980).

† To whom all correspondence should be addressed.

0022–5193/81/210157+21 $02.00/0

Specific proofreading mechanisms must operate within specific constraints as well as the general cost-accuracy relation described above. Thus, the optimum distribution of proofreading effort for a specific mechanism may differ from the general or unconstrained optimum referred to above, and the minimum cost of proofreading will be correspondingly greater than the general or unconstrained minimum. To go beyond these generalizations, each specific mechanism must be examined individually, and for this task it will be important to have a general approach that can be applied to a wide variety of specific cases.

In this paper we present a rather general matrix method for obtaining the specific constraint relations that characterize particular proofreading mechanisms operating in a steady state. Then, for purposes of illustration, we analyze the specific law or constraint associated with the kinetic mechanism originally proposed by Hopfield and Ninio. This is one of the most important specific cases to examine because it has become the model most used by others in the field (e.g. Ehrenberg & Blomberg, 1980). There are, however, other more general models that now can be examined with the techniques presented here.

2. General Description and Cost–Accuracy Relations of the System

A proofreading system is one that discriminates initially between structurally similar (correct and incorrect) substrates and then checks (repeatedly) in subsequent stages, which again discriminate between correct and incorrect intermediates, to improve the accuracy of the output produced. Such systems, which might consist of an arbitrary network of reactions, can be represented abstractly in the steady state by a branching diagram of the form shown in Fig. 1. The state of the system is characterized

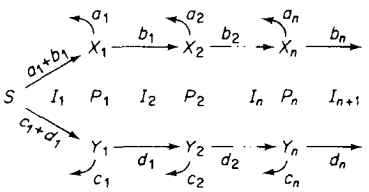

FIG. 1. A proofreading system represented as an abstact branched diagram of an otherwise arbitrary network of reactions. S represents a selection system that interacts with correct and incorrect substrates. The X's represent sets of intermediate complexes involving correct substrate and the selection system. The Y's represent the same sets but with incorrect substrate. The arrows represent net flux between sets. P_i is the macroscopic proofreading discrimination ratio and I_i is the macroscopic input discrimination ratio for the ith stage. I_{n+1} is the output discrimination ratio. See text for further discussion.

by macroscopic fluxes $a_1, \ldots, a_n, b_n, c_1, \ldots, c_n, d_n$. These are net fluxes taken to be positive in the direction of the arrows. These fluxes also define other macroscopic parameters of interest: the initial or input discrimination ratio (I_1) is the ratio of correct to incorrect flux entering the system, $I_1 = (a_1 + b_1)/(c_1 + d_1)$; the proofreading discrimination ratio of the ith stage (P_i) is the fraction of incorrect flux rejected divided by the fraction of correct flux rejected in proofreading at the ith stage, $P_i = (c_i/d_i)/(a_i/b_i)$; the net error following the ith stage of proofreading (E_i) is the incorrect fraction of total flux leaving the ith stage, $E_i = d_i/(b_i + d_i)$; the cost of proofreading for the ith stage (C_{ni}) is the ratio of net flux rejected in the ith stage of proofreading to net flux leaving the system, $C_{ni} = (a_i + c_i)/(b_n + d_n)$. From these definitions it is clear that the input discrimination ratio to the ith stage $I_i = (a_i + b_i)/(c_i + b_i) = b_{i-1}/d_{i-1}$ is directly related to the net error following the $(i-1)$st stage

$$I_i = (1 - E_{i-1})/E_{i-1} \tag{1}$$

and that the total cost of proofreading C_n is given by

$$C_n = \sum_{i=1}^{n} C_{ni}. \tag{2}$$

All these macroscopic parameters of the system are clearly dimensionless quantities. The cost of proofreading, however, can often be interpreted in "units" of moles ATP for proofreading per mole of output produced (e.g. Savageau & Freter, 1979*a*). Other definitions of discrimination and cost could be used, but those given above are conventional ones recognized by most authors and they allow results to be expressed in relatively simple form.

By the use of Kirchhoff's laws we have derived an explicit relationship between cost and accuracy; namely,

$$C_n = \frac{(I_1+1)(P_1-1)\ldots(P_n-1)(1-E_1)\ldots(1-E_n)E_1\ldots E_n}{[(I_1P_1+1)E_1-1][(1-E_1)P_2E_2-(1-E_2)E_1]\ldots[(1-E_{n-1})P_nE_n-(1-E_n)E_{n-1}]} - 1 \tag{3}$$

or

$$C_n = \frac{(I_1+1)(P_1-1)\ldots(P_n-1)}{(I_{n+1}+1)[(I_1P_1/I_2)-1]\ldots[(I_nP_n/I_{n+1})-1]} - 1. \tag{4}$$

When the proofreading effort is optimally distributed

$$I_i = \sqrt{I_{i-1}I_{i+1}P_{i-1}/P_i} \tag{5}$$

which is equivalent to

$$\frac{c_1/(c_1+d_1)}{a_1/(a_1+b_1)} = \ldots = \frac{c_n/(c_n+d_n)}{a_n/(a_n+b_n)}. \tag{6}$$

Under these conditions the cost of proofreading is

$$C_n = \frac{(I_1+1)(P_1-1)(P_2-1)\ldots(P_n-1)}{(I_{n+1}+1)[(I_1/I_{n+1})^{1/n}(P_1P_2\ldots P_n)^{1/n}-1]^n} - 1. \tag{7}$$

Furthermore, since $\partial C_n/\partial P_i < 0$ and $\partial C_n/\partial I_1 < 0$, the cost will be minimum when the P values are all at the same maximum and I_1 is at its maximum.

$$C_{n\min} = \frac{(I_1+1)(P-1)^n}{(I_{n+1}+1)[(I_1/I_{n+1})^{1/n}P-1]^n} - 1. \tag{8}$$

(See Freter & Savageau, 1980; for additional details concerning this section.)

This general or unconstrained minimum might not be achieved in specific cases if the corresponding specific constraints prevent it. In sections 6 and 7, we shall see the effect of permissive and restrictive constraints on the cost of achieving a given degree of accuracy, but first we will develop the specific formulation of the Hopfield–Ninio mechanism in terms of classical enzyme kinetics.

3. Specific Enzyme-kinetic Description of the Hopfield–Ninio System

The n stage version of the Hopfield–Ninio mechanism can be represented schematically as in Fig. 2. The determination of fluxes in terms of elementary rate constants and concentrations of reactants in steady state is simply a problem in classical steady state enzyme kinetics. At this point we need only describe the procedure in outline and for that portion of the mechanism involving the correct substrate s. (For a more complete discussion of these methods and their generality, see Chapter 3 Savageau, 1976.) There will be one equation for each intermediate complex:

$$\begin{aligned}
\dot{X}_1 &= k_1 e.s.\ \mathrm{ATP} + k_{-2} e.s.\mathrm{AMP}\ .\ \mathrm{PP} + k_{-3}X_2 - (k_{-1}+k_2+k_3)X_1 \\
\dot{X}_i &= k_{2i-1}X_{i-1} + k_{-2i} e.s.\mathrm{AMP}\ .\ \mathrm{PP} + k_{-(2i+1)}X_{i+1} - (k_{-(2i-1)}+k_{2i}+k_{2i+1})X_i \\
\dot{X}_n &= k_{2n-1}X_{n-1} + k_{-2n} e.s.\mathrm{AMP}\ .\ \mathrm{PP} + k_{-(2n+1)} e.p.\mathrm{AMP}\ .\ \mathrm{PP} \\
&\quad - (k_{-(2n-1)}+k_{2n}+k_{2n+1})X_n.
\end{aligned} \tag{9}$$

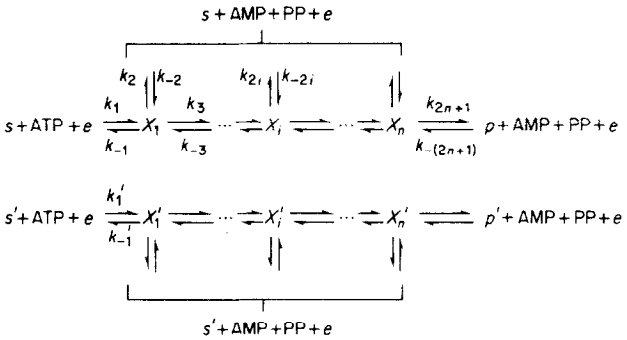

FIG. 2. Schematic representation of an adenylate coupled enzymatic reaction involving a Michaelis–Menten mechanism with side reactions for proofreading. The symbols are defined as follows: ATP, adenosine triphosphate; AMP, adenosine monophosphate; PP, pyrophosphate; e, free enzyme. For the portion of the mechanism involving correct substrate and product: s, substrate; p, product; X_i, intermediate enzyme-substrate complexes; k_{2i+1} and $k_{-(2i+1)}$, elementary rate constants for the forward and reverse components of the main reactions; k_{2i} and k_{-2i}, elementary rate constants for the forward and reverse components of the side reactions. For the portion of the mechanism involving incorrect substrate and product, all of these latter symbols are primed. See text for further discussion.

In steady state, the time derivatives of the intermediates are zero and the resulting equations can be written

$$[A]X] = b] \tag{10}$$

where

$$b] = \begin{bmatrix} (k_1 e.s.\mathrm{ATP} + k_{-2} e.s.\mathrm{AMP}\,.\,\mathrm{PP}) \\ \vdots \\ (k_{-2i} e.s.\mathrm{AMP}\,.\,\mathrm{PP}) \\ \vdots \\ (k_{-2n} e.s.\mathrm{AMP}\,.\,\mathrm{PP} + k_{-(2n+1)} e.p.\mathrm{AMP}\,.\,\mathrm{PP}) \end{bmatrix} ; \quad X] = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_n \end{bmatrix}$$

and

$$[A] = \begin{bmatrix} (k_{-1}+k_2+k_3) & -k_{-3} & 0 & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 \cdots & -k_{2i-1} & (k_{-(2i-1)}+k_{2i}+k_{2i+1}) & & \\ \vdots & & & -k_{-(2i+1)} & \cdots & 0 \\ 0 & \cdots & 0 & \ddots & & \vdots \\ & & & -k_{2n-1} & (k_{-(2n-1)}+k_{2n}+k_{2n+1}) \end{bmatrix}.$$

The concentration of each intermediate in steady state can be obtained by Cramer's rule from elementary algebra

$$X_i = k_1 e.s.\mathrm{ATP}\Delta_i/\Delta \tag{11}$$

where Δ is the determinant of the matrix $[A]$ and Δ_i is identical to Δ except that the ith column has been replaced by $(k_1 e.s.\mathrm{ATP})^{-1}\mathrm{b}]$.

The dependencies among the elementary rate constants are related to the equilibrium constants of the mechanism by the Haldane relations; namely,

$$\frac{k_{-1}k_{-3}\ldots k_{-(2n+1)}}{k_1 k_3 \ldots k_{2n+1}} = K_1 K_3 \ldots K_{2n+1} = K_a/K_s \tag{12}$$

and

$$\frac{k_{-1}k_{-3}\ldots k_{-(2i-1)}k_{-2i}}{k_1 k_3 \ldots k_{2i-1}k_{2i}} = K_1 K_3 \ldots K_{2i-1}K_{2i} = K_a; \qquad i = 1, 2, \ldots, n \tag{13}$$

where

$$K_a = \mathrm{ATP}_{eq}/(\mathrm{AMP}_{eq} \cdot \mathrm{PP}_{eq})$$

$$K_s = p_{eq}/s_{eq}.$$

The corresponding expressions for the portion of the mechanism involving incorrect substrate are identical except that all symbols (other than ATP, AMP, PP and e) are primed. When one solves for the intermediate concentrations in each half of the mechanism they will be a function of e, the concentration of free enzyme, which then cancels in all subsequent ratios (see section 4).

The following definitions will be used to relate the two portions of the mechanism:

ratio of equilibrium constants,	$L = K_s/K_{s'}$
ratio of substrate concentrations,	$\sigma = s/s'$
ratio of product concentrations,	$\pi = p/p'$
ratio of forward rate constants,	$f_i = k_i/k'_i$
ratio of reverse rate constants,	$r_i = k_{-i}/k'_{-i}$.

(14)

4. Specific Constraint Relation for the Hopfield–Ninio System

A comparison of Figs 1 and 2 allows one to relate the general and specific descriptions in a straightforward manner.

$$I_1 = \frac{k_1 e.s.\mathrm{ATP} - k_1 K_1 X_1}{k_1' e.s'.\mathrm{ATP} - k_1' K_1' X_1'}$$

$$= \sigma f_1 \frac{\Delta'[\Delta - k_{-1}\Delta_1]}{\Delta[\Delta' - k_{-1}'\Delta_1']}, \tag{15}$$

$$I_i = \frac{k_{2i-1} X_{i-1} - k_{2i-1} K_{2i-1} X_i}{k_{2i-1}' X_{i-1}' - k_{2i-1}' K_{2i-1}' X_i'}$$

$$= \sigma f_1 f_{2i-1} \frac{\Delta'[\Delta_{i-1} - K_{2i-1}\Delta_i]}{\Delta[\Delta_{i-1}' - K_{2i-1}'\Delta_i']}, \qquad i = 2, 3, \ldots, n, \tag{16}$$

$$I_{n+1} = \frac{k_{2n+1} X_n - k_{2n+1} K_{2n+1} e.p.\mathrm{AMP}\,.\,\mathrm{PP}}{k_{2n+1}' X_n' - k_{2n+1}' K_{2n+1}' e.p'.\mathrm{AMP}\,.\,\mathrm{PP}}$$

$$= \frac{f_1 f_{2n+1}\Delta'[k_{-1} s.\mathrm{ATP}\Delta_n - K_1 K_{2n+1} p.\mathrm{AMP}\,.\,\mathrm{PP}\Delta]}{r_1 \Delta[k_{-1}' s'.\mathrm{ATP}\Delta_n' - K_1' K_{2n+1}' p'.\mathrm{AMP}\,.\,\mathrm{PP}\Delta']}, \tag{17}$$

$$P_i = \frac{[k_{2i+1} X_i - k_{2i+1} K_{2i+1} X_{i+1}][k_{2i}' X_i' - k_{2i}' K_{2i}' e.s'.\mathrm{AMP}\,.\,\mathrm{PP}]}{[k_{2i} X_i - k_{2i} K_{2i} e.s.\mathrm{AMP}\,.\,\mathrm{PP}][k_{2i+1}' X_i' - k_{2i+1}' K_{2i+1}' X_{i+1}']}$$

$$= \frac{r_1 I_{i+1}\Delta[k_{-1}'\,\mathrm{ATP}\,\Delta_i' - K_1' K_{2i}'\mathrm{AMP}\,.\,\mathrm{PP}\,\Delta']}{\sigma f_1 f_{2i}\Delta'[k_{-1}\mathrm{ATP}\,\Delta_i - K_1 K_{2i}\mathrm{AMP}\,.\,\mathrm{PP}\,\Delta]}, \qquad i = 1, 2, \ldots, n. \tag{18}$$

The input and output discrimination ratios and the proofreading discrimination ratio of a given stage can be related to one another as follows. For the first stage:

$$P_1 = \frac{r_1 I_2}{\sigma f_1 f_2}\,\frac{B_1'}{B_1}. \tag{19}$$

Where B_1 and B_1' are proportional to and have the same sign as the exit fluxes a_1 and c_1 in Fig. 1.

From equations (15), (18) and (19) one can see that

$$B_1 = (\mathrm{ATP} - K_a \mathrm{AMP}\,.\,\mathrm{PP}) - \mathrm{ATP}(\Delta - k_{-1}\Delta_1)/\Delta \tag{20}$$

so that equation (15) can be rewritten

$$I_1 = \sigma f_1 \frac{[(\mathrm{ATP} - K_a \mathrm{AMP}\,.\,\mathrm{PP}) - B_1]}{[(\mathrm{ATP} - K_a \mathrm{AMP}\,.\,\mathrm{PP}) - B_1']}. \tag{21}$$

One can solve equation (19) for B_1', substitute this value into equation

(21), and rearrange the resulting expression to give

$$B_1 = (\mathrm{ATP} - K_a \mathrm{AMP} \,.\, \mathrm{PP}) \frac{\left[1 - \dfrac{I_1}{\sigma f_1}\right]}{\left[1 - \dfrac{f_2 I_1 P_1}{r_1 I_2}\right]}. \tag{22}$$

In a similar fashion, one can obtain the corresponding relationship for the ith stage. From equation (18)

$$P_i = \frac{r_1 I_{i+1} B_i'}{\sigma f_1 f_{2i} B_i}, \qquad i = 2, 3, \ldots, n \tag{23}$$

where

$$B_i = (k_{-1}\mathrm{ATP}\Delta_i - K_1 K_{2i} \mathrm{AMP} \,.\, \mathrm{PP}\Delta)/\Delta, \qquad i = 2, 3, \ldots, n. \tag{24}$$

Again, B_i is proportional to the exit flux a_i. Equation (16) then can be rewritten as

$$I_i = \frac{\sigma f_1 f_{2i-1}}{r_1} \frac{(B_{i-1} - K_{2i-1} B_i)}{(B_{i-1}' - K_{2i-1}' B_i')}, \qquad i = 2, 3, \ldots, n. \tag{25}$$

Again, one can solve equation (23) for B_i', substitute this value into equation (25), and rearrange the resulting expression to give

$$B_i = B_{i-1} \frac{\left[1 - \dfrac{f_{2i-2}}{f_{2i-1}} P_{i-1}\right]}{K_{2i-1}\left[1 - \dfrac{f_{2i}}{r_{2i-1}} \dfrac{I_i}{I_{i+1}} P_i\right]}, \qquad i = 2, 3, \ldots, n \tag{26}$$

and thus

$$B_i = B_1 \prod_{j=2}^{i} \left\{ \frac{\left[1 - \dfrac{f_{2j-2}}{f_{2j-1}} P_{j-1}\right]}{K_{2j-1}\left[1 - \dfrac{f_{2j}}{r_{2j-1}} \dfrac{I_j}{I_{j+1}} P_j\right]} \right\}, \qquad i = 2, 3, \ldots, n. \tag{27}$$

The output discrimination ratio from the nth stage, which was given in equation (17), also can be rewritten by using the relationships in equations (23) and (24):

$$I_{n+1} = \frac{\sigma f_1 f_{2n+1}}{r_1} \frac{\{B_n + [1 - p/(sK_s)] K_1 K_{2n} \mathrm{AMP} \,.\, \mathrm{PP}\}}{\{B_n' + [1 - p'/(s'K_s')] K_1' K_{2n}' \mathrm{AMP} \,.\, \mathrm{PP}\}}. \tag{28}$$

Again, one can solve equation (23) for B_n', substitute this value into equation

(28) and rearrange the resulting expression to give

$$B_n = \frac{\dfrac{K_a \mathrm{AMP}\,.\,\mathrm{PP}}{K_3 K_5 \ldots K_{2n-1}} \left[\left(\dfrac{I_{n+1}}{\sigma L r_{2n+1}} - 1 \right) - \dfrac{p}{sK_s} \left(\dfrac{I_{n+1}}{\pi r_{2n+1}} - 1 \right) \right]}{\left[1 - \dfrac{f_{2n}}{f_{2n+1}} P_n \right]} \tag{29}$$

where σL may be considered the (limiting) output discrimination ratio at equilibrium.

Finally, equating the expressions for B_n in equations (27) and (29) yields the constraint relation $H = 0$, where

$$H = \frac{\left[1 - \dfrac{I_1}{\sigma f_1} \right] \left[1 - \dfrac{f_2}{f_3} P_1 \right] \ldots \left[1 - \dfrac{f_{2n}}{f_{2n+1}} P_n \right]}{\left[1 - \dfrac{f_2}{r_1} \dfrac{I_1}{I_2} P_1 \right] \ldots \left[1 - \dfrac{f_{2n} I_n}{r_{2n-1} I_{n+1}} P_n \right]} - W \tag{30}$$

and

$$W = \frac{K_a \mathrm{AMP}\,.\,\mathrm{PP}}{\mathrm{ATP} - K_a \mathrm{AMP}\,.\,\mathrm{PP}} \left[\left(\frac{I_{n+1}}{\sigma L r_{2n+1}} - 1 \right) - \frac{p}{sK_s} \left(\frac{I_{n+1}}{\pi r_{2n+1}} - 1 \right) \right]. \tag{31}$$

W contains all the boundary conditions and is fixed by the environment and the nature of the reactants, except for the output discrimination ratio I_{n+1} and the ratio of rate constants r_{2n+1}.

A similar relationship has been developed independently using different methods by Ehrenberg & Blomberg (1980).

5. Optimization: General Considerations

The minimization of cost is to be achieved subject to the specific constraint that $H = 0$. This can be obtained by finding appropriate values of I_i and P_i and of the r and f ratios. A number of different methods can be used to find stationary states of C_n in the region of allowable parameter values, but optima also may exist on the boundaries of this region, particularly with regard to the r and f ratios. For our purposes the Lagrange method is appropriate. Thus, we seek values for I_i and P_i that will minimize the function

$$T = C_n + \lambda H, \tag{32}$$

where λ is the conventional Lagrange multiplier. The stationary states are given by the solutions of the following equations:

$$\partial T/\partial I_i = 0, \qquad \partial T/\partial P_i = 0 \quad \text{and} \quad H = 0. \tag{33}$$

These solutions are in general a function of the r and f ratios. By choosing boundary values for these ratios one can obtain the minimum cost.

The first of equations (33), which can be written

$$\frac{\partial T}{\partial I_1} = -\frac{(C_n+1)}{(I_1+1)} \frac{\left(\frac{P_1}{I_2}+1\right)}{\left(\frac{I_1 P_1}{I_2}-1\right)} + \frac{\lambda W\left(\frac{f_2}{r_1}\frac{P_1}{I_2} - \frac{1}{\sigma f_1}\right)}{\left(1-\frac{I_1}{\sigma f_1}\right)\left(1-\frac{f_2}{r_1}\frac{I_1}{I_2}P_1\right)} = 0 \tag{34}$$

and

$$\frac{\partial T}{\partial I_i} = \frac{(C_n+1)}{I_i} \frac{\left(\frac{I_i P_i}{I_{i+1}} - \frac{I_{i-1}P_{i-1}}{I_i}\right)}{\left(\frac{I_{i-1}P_{i-1}}{I_i}-1\right)\left(\frac{I_i P_i}{I_{i+1}}-1\right)}$$

$$+\frac{\lambda W}{I_i} \frac{\left(\frac{f_{2i}}{r_{2i-1}}\frac{I_i P_i}{I_{i+1}} - \frac{f_{2i-2}}{r_{2i-3}}\frac{I_{i-1}P_{i-1}}{I_i}\right)}{\left(1-\frac{f_{2i-2}}{r_{2i-3}}\frac{I_{i-1}P_{i-1}}{I_i}\right)\left(1-\frac{f_{2i}}{r_{2i-1}}\frac{I_i P_i}{I_{i+1}}\right)} = 0 \qquad i = 2, 3, \ldots, n \tag{35}$$

indicate that the general or unconstrained optimal distribution of proofreading effort (equation (5)) can only be achieved if

$$f_{2i}/r_{2i-1} = f_{2i-2}/r_{2i-3} \quad \text{for all } i. \tag{36}$$

The second of equations (33) can be written

$$\frac{\partial T}{\partial P_i} = \frac{(C_n+1)}{(P_i-1)} \frac{\left(\frac{I_i}{I_{i+1}}-1\right)}{\left(\frac{I_i P_i}{I_{i+1}}-1\right)} + \frac{\lambda W\left(\frac{f_{2i}}{r_{2i-1}}\frac{I_i}{I_{i+1}} - \frac{f_{2i}}{f_{2i+1}}\right)}{\left(1-\frac{f_{2i}P_i}{f_{2i+1}}\right)\left(1-\frac{f_{2i}}{r_{2i-1}}\frac{I_i P_i}{I_{i+1}}\right)} = 0$$

$$i = 1, 2, \ldots, n. \tag{37}$$

These $2n$ equations, together with the constraint $H = 0$, provide $2n+1$ relations among the $2n+1$ variables (I_i, P_i and λ). Although the simultaneous solution of these equations is not obvious, straightforward iterative methods can be used successfully to obtain solutions.

In the following sections we shall assume that $\sigma = 1$ and $\pi = I_{n+1}$, i.e. that the environment maintains equal substrate concentrations and removes

products *non*-selectively†. The selectivity under these conditions is due to the system alone. Furthermore, in the cases of interest the substrate and product molecules being distinguished are closely related so that $L \doteq 1$, i.e. $K_s = K_{s'}$. More detailed treatments of the optimization procedure will be given for appropriate cases in sections 6 and 7.

6. Optimization under Permissive Conditions

As we have discussed elsewhere (Freter & Savageau, 1980), the specific constraints imposed by a given system may prevent the general or unconstrained optimum from being reached. This can occur if the specific constraints (1) prevent the unconstrained optimum distribution of proofreading effort from being established, (2) prevent the P values from achieving the same maximum value, or (3) interfere with the optimization by a combination of these two limitations.

There are several ways that one may view the variation of kinetic design for an enzyme. These different perspectives generally reflect a choice made on the basis of computational ease or conceptual emphasis and have little if anything to do with the manner in which a specific design may be achieved in nature. In this section we will consider two examples that illustrate this point.

First, let us consider the situation in which the proofreading and initial discrimination ratios have fixed values but the r and f ratios may be varied. The optimal distribution of proofreading effort can be obtained by selecting the optimum values for I_i subject to the specific constraint that $H = 0$. From equation (30) one can see that varying I_i directly with r_{2i-1} and inversely with r_{2i-3} will leave H unchanged. Therefore, one can vary the I_i in this manner until

$$I_i = \sqrt{I_{i-1} I_{i+1} P_{i-1} / P_i}. \tag{38}$$

This minimizes cost.

In some cases, the allowable range of the ratios r_{2i-1} may prevent the unconstrained optimum from being reached. These cases will be considered further in section 7.

Now let us consider the variation in enzyme design from another perspective; namely, the situation in which the r and f ratios are fixed and the P and I values are varied.

† If p and p' are removed non-selectively from the common terminal pool, then the ratio of their rates of removal will be proportional to the ratio of their concentrations. In steady state these concentrations cannot be changing with time and therefore the input rate and the removal rate must be identical for each product. Thus, $\pi = p/p' = I_{n+1}$ regardless of the removal mechanism. In particular, the reactions removing product need not be first order.

It can be shown that Δ is always positive†. Similarly, if ATP is sufficiently greater than K_aAMP . PP, which is the usual case of interest, it can be shown that $\Delta - K_{-1}\Delta_1$, $\Delta_{i-1} - K_{2i-1}\Delta_i$, $K_{-1}s.\mathrm{ATP}\Delta_n - K_1K_{2n+1}p.\mathrm{AMP}$. PPΔ, all the B_i and the proofreading costs at each stage are non-negative. From these results and equations (15–18) we conclude that all of the macroscopic parameters I_i and P_i are non-negative; from equations (20) and (25) we see that $B_1 \leq \mathrm{ATP} - K_a\mathrm{AMP}$. PP and $K_{2i-1}B_i \leq B_{i-1}$. Equations (22) and (26) then imply that

$$0 \leq \frac{\left(1 - \dfrac{I_1}{\sigma f_1}\right)}{\left(1 - \dfrac{f_2}{r_1}\dfrac{I_1 P_1}{I_2}\right)} \leq 1 \tag{39}$$

and

$$0 \leq \frac{\left(1 - \dfrac{f_{2i-2}P_{i-1}}{f_{2i-1}}\right)}{\left(1 - \dfrac{f_{2i}}{r_{2i-1}}\dfrac{I_i P_i}{I_{i+1}}\right)} \leq 1, \qquad i = 2, 3, \ldots, n. \tag{40}$$

Furthermore, if $\pi r_{2n+1} > I_{n+1} > \sigma L r_{2n+1}$, which again is the usual case of interest, then equation (29) implies that

$$0 < \left(1 - \frac{f_{2n}}{f_{2n+1}} P_n\right) < 1. \tag{41}$$

For $H = 0$ in equation (30) we see that equations (39) through (41) imply

$$0 \leq W < 1 \tag{42}$$

† Rewrite Δ as follows. First, add the elements of the first row to the corresponding elements of the second row. Second, add the elements of the new second row to the corresponding elements of the third row. Continue in this manner until the nth row has been rewritten. These operations leave the value of Δ unchanged. The resulting determinant has the following form:

$$\begin{vmatrix} + & - & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ + & + & - & 0 & 0 & & & & 0 \\ + & + & + & - & 0 & & & & 0 \\ \cdot & & & \cdot & \cdot & \cdot & & & \cdot \\ \cdot & & & & & & \cdot & + & - & 0 \\ \cdot & & & & & & & + & + & - \\ + & + & + & \cdot & \cdot & \cdot & + & + & + \end{vmatrix}$$

A determinant with this pattern of signs is always positive, as can be shown readily by induction (see p. 342 in Savageau, 1976).

and, thus, that

$$\frac{p}{sK_s} \geq \frac{\left(1 - \frac{I_{n+1}}{\sigma L r_{2n+1}}\right)}{\left(1 - \frac{I_{n+1}}{\pi r_{2n+1}}\right)} \tag{43}$$

and

$$\frac{\text{ATP}}{K_a \text{AMP} \cdot \text{PP}} > \frac{I_{n+1}}{\sigma L r_{2n+1}} + \frac{p}{sK_s}\left(1 - \frac{I_{n+1}}{\pi r_{2n+1}}\right). \tag{44}$$

For each stage there are two possibilities; either the ratio of factors in equations (39) or (40) is a ratio of two positive factors or it is a ratio of two negative factors. Only the positive factors lead to a self-consistent optimum†. Therefore, maximum limits for I_1 and P_i are given by

$$I_1 \leq \sigma f_1 \quad \text{and} \quad P_i \leq f_{2i+1}/f_{2i}. \tag{45}$$

The least restrictive conditions will be to fix these f ratios at their most extreme value so as to maximize the range of I_1 and P_i. In other words,

$$f_{2i+1} = \rho, \qquad f_{2i} = 1/\rho \quad \text{and} \quad f_1 = \rho \qquad i = 1, 2, \ldots, n \tag{46}$$

and, since $L = 1$,

$$r_{2i-1} = \rho \qquad i = 1, 2, \ldots, n+1, \tag{47}$$

where ρ is the maximum value for the particular substrates and intermediates being considered. Under these conditions equations (4) and (30) can be rewritten as

$$C_n = \frac{(I_1 + 1)(P_1 - 1) \ldots (P_n - 1)}{(I_{n+1} + 1)\left(\frac{I_1 P_1}{I_2} - 1\right) \ldots \left(\frac{I_n P_n}{I_{n+1}} - 1\right)} - 1 \tag{48}$$

and

$$H = \frac{\left(1 - \frac{I_1}{\sigma\rho}\right)\left(1 - \frac{P_1}{\rho^2}\right) \ldots \left(1 - \frac{P_n}{\rho^2}\right)}{\left(1 - \frac{I_1}{I_2}\frac{P_1}{\rho^2}\right) \ldots \left(1 - \frac{I_n}{I_{n+1}}\frac{P_n}{\rho^2}\right)} - W. \tag{49}$$

† This can be seen by examining the signs of the terms in equations (34) and (37) under the conditions in equations (39) through (41) and the condition that all $P_i > 1$. This latter condition is required because an optimum with any $P_i < 1$ and $C_{ii} > 0$ would not be a true optimum. One could always increase accuracy and decrease the cost of proofreading by eliminating the proofreading at such a stage.

The apparent cost (Savageau & Freter, 1979*a*; Freter & Savageau, 1980) for the ith stage is

$$C_{ii} = \frac{(I_i+1)}{(I_{i+1}+1)} \frac{P_i-1)}{\left(\frac{I_iP_i}{I_{i+1}}-1\right)} - 1. \tag{50}$$

With $B_i \geq 0$ for all i, which implies positive or zero exit fluxes due to proofreading at stage i, this cost (and the true cost as well) will be positive or zero. The latter condition implies $I_{i+1} = I_i$ and in effect the ith stage is eliminated from consideration. Thus, there will be $2^n - 1$ optima† for equations (48) and (49) corresponding to the 2^n combinations of zero and positive proofreading costs for the n stages.

Therefore, without loss of generality, we can consider these optima as solutions to equations (48) and (49) when $n = 1, 2, \ldots$ and every stage has a positive cost of proofreading. For this case, equations (34), (35) and (37) become

$$\frac{(C_n+1)}{(I_1+1)} \frac{\left(\frac{P_1}{I_2}+1\right)}{\left(\frac{I_1P_1}{I_2}-1\right)} + \frac{\lambda W\left(1-\frac{P_1}{\rho I_2}\right)}{\rho\left(1-\frac{I_1}{\rho}\right)\left(1-\frac{I_1}{I_2}\frac{P_1}{\rho^2}\right)} = 0 \tag{51}$$

$$\frac{(C_n+1)}{I_i} \frac{\left(\frac{I_iP_i}{I_{i+1}}-\frac{I_{i-1}P_{i-1}}{I_i}\right)}{\left(\frac{I_{i-1}P_{i-1}}{I_i}-1\right)\left(\frac{I_iP_i}{I_{i+1}}-1\right)} + \frac{\lambda W}{\rho^2 I_i} \frac{\left(\frac{I_iP_i}{I_{i+1}}-\frac{I_{i-1}P_{i-1}}{I_i}\right)}{\left(1-\frac{I_{i-1}P_{i-1}}{I_i\rho^2}\right)\left(1-\frac{I_i}{I_{i+1}}\frac{P_i}{\rho^2}\right)} = 0 \tag{52}$$

$$\frac{(C_n+1)}{(P_i-1)\left(\frac{I_iP_i}{I_{i+1}}-1\right)} + \frac{\lambda W}{\rho^2\left(1-\frac{P_i}{\rho^2}\right)\left(1-\frac{I_iP_i}{I_{i+1}\rho^2}\right)} = 0. \tag{53}$$

If $I_iP_i/I_{i+1} \neq I_{i-1}P_{i-1}/I_i$ for any i then equations (52) and (53) together imply that $I_{i-1} = I_{i+1}$. This is impossible under the present conditions because $I_{i-1} \neq I_i$ and $I_i \neq I_{i+1}$, and therefore either $I_i > I_{i+1}$ or $I_{i-1} > I_i$. In general, the condition $I_i > I_{i+1}$ implies that either $B_i < 0$ or $P_i < 1$, both of which are forbidden under the conditions of our analysis.

† The optimum corresponding to the condition $I_1 = I_2 = \ldots = I_{n+1}$, which would have zero cost, is forbidden under the conditions we are considering here because $H = 0$ implies $[1-(I_{n+1}/\sigma\rho)] = W$. This last condition implies that $I_{n+1} < \sigma\rho$, but we already have stated that $I_{n+1} > \sigma\rho$ (sentence preceding equations (41)).

Thus, under the present conditions, the only optimum occurs when

$$\frac{I_i}{I_{i+1}} P_i = \frac{I_{i-1}}{I_i} P_{i-1} \quad \text{for all } i \tag{54}$$

and from equations (53), this condition implies

$$P_i = P_{i-1} = P \quad \text{for all } i. \tag{55}$$

Solving the recursive relations in equations (54) yields

$$I_i/I_{i+1} = (I_1/I_{n+1})^{1/n}. \tag{56}$$

By combining equations (51) and the first of equations (53) with equations (55) and (56) we obtain a quadratic function in P

$$\left[\frac{I_1}{\rho}\left(1+\frac{1}{\rho}\right)-\left(1-\frac{1}{\rho^2}\right)\right]\left(\frac{I_1}{I_{n+1}}\right)^{1/n} P^2 - \left[1+\left(\frac{I_1}{I_{n+1}}\right)^{1/n}\right]\left[\frac{I_1}{\rho}(\rho+1)\right] P$$

$$+ I_1\left[\left(1-\frac{I_1}{\rho}\right)+\rho(I_1+1)\right] = 0. \tag{57}$$

This equation has one positive solution for P when the coefficient of the P^2 term is negative and two positive solutions when this coefficient is positive. However, in the latter case the larger value of P is forbidden because it exceeds the allowable range for this parameter. The relevant solution in either case is

$$P = (-B - \sqrt{B^2 - 4AC})/2A \tag{58}$$

where A, B, and C are the coefficients of the polynomial in equation (57).

Finally, inserting this value of P (as a function of I_1) into equation (49) and setting $H = 0$ gives an equation only in I_1. Solving for the appropriate value of I_1 and back substituting into the other equations determines the optimum with minimum cost.

The solutions in this section correspond to the optimum when the specific constraint is least restrictive. In each case the optimum occurs when the proofreading effort has the distribution of the unconstrained optimum. Now let us examine situations in which the specific constraint is more restrictive.

7. Optimization under Restrictive Conditions

Return now to the optimization described first in section 6. If I_1 and P values are at their maximum and fixed, the unconstrained optimal distribution of proofreading effort yields the minimum cost. However, in adjusting

the I_i to achieve this optimum it is necessary that the corresponding ratios r_{2i-1} and r_{2i-3} have a sufficiently wide range of allowable variation. For example, if I_i is below its optimum value and r_{2i-1} is below its maximum value, they can both be increased (and r_{2i-3} decreased) to lower cost, while maintaining $H = 0$. However, if r_{2i-1} reaches its maximum value *before* I_i reaches its optimum value, then the unconstrained optimum distribution of proofreading effort is not obtained and the corresponding cost will be higher. This is illustrated in Fig. 3, where the increase in cost is associated with progressively lower maximum values for the appropriate r ratio. This example illustrates how the specific constraints may prevent the general or unconstrained minimum from being reached because the constraint does not permit the optimal distribution of proofreading effort.

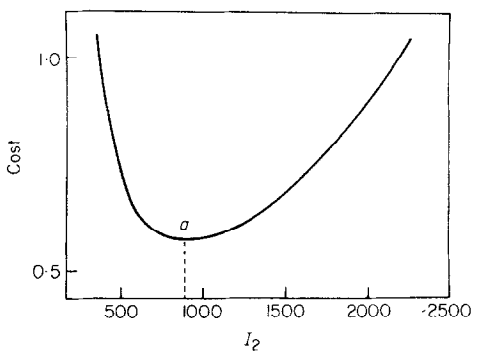

FIG. 3. Optimization with restrictions upon the allowable range of values for r_{2i-1} and r_{2i-3}. Under conditions in which I_1, I_{n+1}, and the P_i are fixed, the minimum cost of proofreading is achieved by varying the odd r ratios and the I_i to achieve the optimal distribution of proofreading effort. This is the point marked *a*. When the maximum value of r_{2i-1} or r_{2i-3} is restricted the minimum cost of proofreading increases. In this example the particular parameter values at the permissive optimum are: $K_a\text{AMP} \cdot \text{PP}/\text{ATP} = 10^{-10}$, $K_s s/p = 10^{-8}$, $n = 2$, $P_1 = P_2 = 310$, $I_1 = 16$, $I_2 = 903$, $I_3 = 50{,}000$, $r_1 = r_3 = r_5 = f_1 = f_3 = f_5 = 20$, $r_2 = r_4 = f_2 = f_4 = 0{\cdot}05$. For the restrictive optima, all parameters are identical except for r_1, r_3 and I_2. They become $r_1 = 20\alpha$, $r_3 = 20/\alpha$, and $I_2 = 903/\alpha$, where α is a factor indicating the degree to which the maximum value of $r_1 (\alpha < 1)$ or $r_3 (\alpha > 1)$ is restricted. Cost is measured in moles ATP for proofreading per mole of product formed. See text for further discussion.

Now consider the optimization from the alternative perspective, that of fixed r and f ratios. In this case, the I and P values are adjusted to give minimum cost. As we saw in the previous section, this involves the usual Lagrange approach to obtaining stationary states within the allowable range of values for the I and P variables. However, there also may be optima corresponding to solutions on the boundary of this allowable range. For example, let us consider a restriction on the maximum value of P_i. One

maximum value for P_i is given in equation (45). However, the constellation of k values for the specific mechanism, equation (18), may determine another maximum value for P_i that is lower than that given in equation (45). If the lower maximum is less than the value of P_i previously required to achieve the optimum, then a new optimum with higher cost will be achieved on the boundary. This is shown in Fig. 4*a*. As long as the maximum P value is less than f_{2i+1}/f_{2i}, but greater than the P value that gives lowest cost under permissive conditions, the same optimum is achieved. As the maximum value of P becomes progressively lower, the cost corresponding to optimization under the restrictive condition increases. This example shows how specific constraints can prevent the unconstrained optimum from being reached because the P values are forbidden to have the same maximum value.

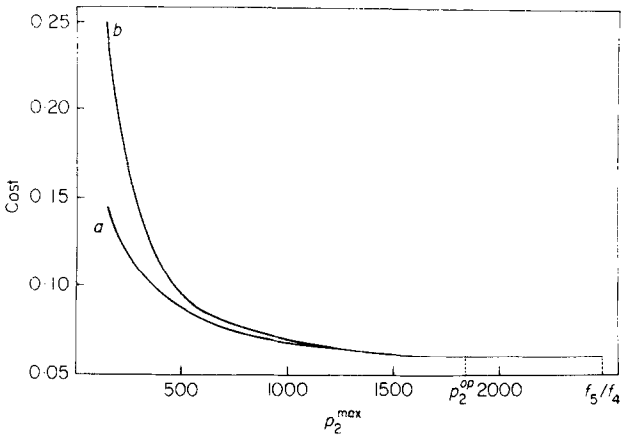

FIG. 4. Optimization with restrictions upon the maximum value of P_i and I_i under conditions in which the r and f ratios are fixed. Curve a: if the maximum value of P_i is f_{2i+1}/f_{2i} and the system is optimized under permissive conditions an optimum value of P_i^{op} (less than f_{2i+1}/f_{2i}) is determined. If the maximum value of P_i is greater than this optimum value but less than f_{2i+1}/f_{2i} an identical optimum is obtained. However, if the maximum value of P_i is less than P_i^{op}, then the cost of proofreading increases. For these restricted optima the value of I_i will increase so as to maintain the unrestricted optimal distribution of proofreading effort, i.e. I_i will continue to satisfy the relationship $I_i = \sqrt{I_{i-1}I_{i+1}P_{i-1}/P_i}$. Curve b: if in addition to the restriction on the maximum value of P_i there is a restriction on the maximum value of I_i that prevents it from re-establishing the above relationship, then the cost of proofreading increases at a greater rate. For these examples the parameter values at the permissive optima are: K_aAMP . PP/ATP $= 10^{-10}$, $K_s s/p = 10^{-8}$, $n = 2$, $P_1 = P_2 = 1844$, $I_1 = 43$, $I_2 = 1470$, $I_3 = 50{,}000$, $r_1 = r_3 = r_5 = f_1 = f_3 = f_5 = 50$, $r_2 = r_4 = f_2 = f_4 = 0{\cdot}02$. For the restrictive optima in a P_1, I_1 and I_2 vary according to the maximum value of P_2 and $I_2 = \sqrt{I_1 I_3 P_1/P_2}$; in b P_1 and I_1 vary according to the maximum values of P_2 and I_2 and $I_2 \neq \sqrt{I_1 I_3 P_1/P_2}$. Cost is measured in moles ATP for proofreading per mole of product formed.

In this last example, the values of the I_i at the optimum under restrictive conditions are adjusted to provide the optimal distribution of proofreading effort. This requires that the other P values change and that the corresponding intermediate discrimination I_i increase. Thus, if there is also a restriction imposed by the mechanism, equation (16), on the maximum value of I_i, then the optimum distribution of proofreading effort may be forbidden as well. The effect of this double restriction on cost is shown in Fig. 4*b*.

A second example illustrating a restriction on P values and a restriction on the optimal distribution of proofreading effort is the following. Consider a mechanism for which f_{2i} is restricted to a value greater than the other f_{2j} ratios. In comparison to the permissive conditions, this restriction implies that P_i will have a value lower than the other P_j values because P_j is proportional to $1/f_{2j}$. The optimal distribution of proofreading effort also will be forbidden because $f_{2i}/r_{2i-1} \neq f_{2i+2}/r_{2i+1}$ for some i. In Fig. 5 the costs at the restrictive optima are shown as a function of the ratio f_{2i}.

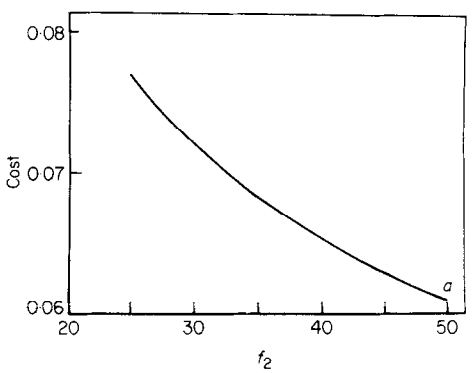

FIG. 5. Optimization with restrictions upon the allowable range of values for f_{2i} (and r_{2i}) under conditions in which the other r and f ratios are fixed. When $f_{2i} = 0{\cdot}02$ the usual permissive optimum is reached as indicated by the point marked *a*. As f_{2i} (and r_{2i}) is progressively raised and the system reoptimized, the cost of proofreading increases. This increase is due to the combined effect of lowering P_i (which is proportional to $1/f_{2i}$) and failure to establish the unconstrained distribution of proofreading effort (i.e. $I_i \neq \sqrt{I_{i-1}I_{i+1}P_{i-1}/P_i}$). For this example the parameter values at the permissive optimum are identical to those listed in the caption of Fig. 4. Cost is measured in moles ATP for proofreading per mole of product formed. See text for further discussion.

8. Discussion

The two principal purposes of this paper are (1) to present a general matrix method for the derivation of the steady state constraint relations that characterize specific proofreading mechanisms and (2) to show the

variety of constraints that a specific mechanism might impose upon the optimization.

Many of the biological systems in which proofreading has been implicated are among the most complex enzymatic machinery of the cell, e.g. DNA dependent DNA polymerases, DNA dependent RNA polymerases and ribosomes. At this time these systems are poorly understood in terms of the underlying kinetic mechanisms and it is undoubtedly overly optimistic to assume that these mechanisms will turn out to be as simple as the "ladder" networks analyzed by most authors. For this reason we have found it desirable to develop sufficiently general methods that can be used to analyze these systems regardless of the detailed nature of the mechanisms that eventually emerge.

Previously we have derived a general cost–accuracy relationship for proofreading systems that can be represented by an abstract branching diagram as shown in Fig. 1, but are otherwise mechanism independent (Savageau & Freter, 1979*a*; Freter & Savageau, 1980). In this paper we have presented a general matrix method, based upon classical enzyme kinetics, for deriving the constraint relation for specific mechanisms. Thus, what is general in this paper is the cost–accuracy relationship, equation (4) (which is equivalent to the "cost" or "objective" function in conventional optimization theory), and the method of deriving the specific constraint relation; what is specific is the "ladder" network we have analyzed—with specific constraint relation ($H = 0$) and boundary values for its parameters. This specific ladder network, used here only for the purposes of illustrating the method, is just one example of the types of proofreading systems that can be analyzed by the methods presented here.

The matrix method of deriving the H constraint has two principal advantages. First, it is a straightforward development based upon classical enzyme kinetics and is therefore familiar to most investigators. Second, it is quite generally applicable to steady state mechanisms whether they involve the simple "ladder" structure of the specific mechanism examined in this paper or enzymatic mechanisms that involve "bridges" (e.g. bimolecular random order mechanisms) or "non-planar" structures (e.g. trimolecular random order mechanisms).

Ehrenberg & Blomberg (1980) have analyzed the ladder network in some detail with an approach that involves formulation of an electrical analog of the enzymatic mechanism and the use of well-established methods for the series-parallel reduction of the network. Although these methods are not well-known among biologists, they were originally developed in the 1920's, particularly by Cauer in Germany, and can be found in most elementary texts dealing with network analysis (e.g. see Van Valkenberg,

1960). These methods are appropriate for ladder networks, but they cannot be used for certain networks such as those involving "bridges" and in general those that are "non-planar" (Van Valkenberg, 1960).

Because our original studies dealt only with the general or unconstrained optimization and those of Ehrenberg and Blomberg dealt only with the optimization of the ladder network under the least restrictive conditions, we have explored several different types of constraints and their effects upon the optimization process. The purpose is to show the limitations of the perspective from which one views the optimization of enzyme design and to show that the specific nature of the assumptions can dramatically affect the results that one obtains.

The primary sequence of amino acids determines the tertiary structure and functional properties of an enzyme. The details of the theory that will relate changes in the primary sequence, which must be discrete, to changes in the tertiary structure and function of an enzyme remain to be revealed. These structural changes in general affect a number of the "elementary" rate constants, k's, used to define its mechanism. Although in most cases the k's are undoubtedly dependent parameters, the real constraints among them are unknown. For conceptual convenience one often treats these k's as independent. From this perspective, however, one can equally well consider ratios of k's, or other configurations of the k's that might make physiological sense, as the independent parameters. (For a discussion of an analogous situation, see pp. 157–159 in Savageau, 1976.)

We have discussed elsewhere how various molecular design features affect the physiological performance of proofreading systems (Savageau & Freter, 1979*a*; Freter & Savageau, 1980). We have also examined the relationship between proofreading systems and their "environment" (which includes the remainder of the cell) and made the following conjecture concerning the evolution of accuracy and proofreading cost: the cost of proofreading will decrease until the energy saved by an additional decrement in proofreading is just equal to the increment in energy waste resulting from the concomitant increase in net error (Savageau & Freter, 1979*b*). Experimental data concerning proofreading costs (Hopfield *et al.*, 1976; Mulvey & Fersht, 1977) and accuracy (Edelmann & Gallant, 1977) were used to test this conjecture and the results show remarkable agreement.

These considerations are important because ultimately it is the physiological performance of the enzyme within the cell that determines whether or not a given kinetic design will have selective value and flourish. For this reason we have attempted to express the behavior of the enzyme in terms of macroscopic or physiological variables insofar as possible. We have done this by converting a microscopic description involving k's into a macroscopic

description involving r's, f's, I's and P's. We have also illustrated the effects of a variety of constraints upon the magnitudes of these parameters. Some constraints relate to important ratios of elementary rate constants (r's and f's), others relate to important functions of the elementary rate constants (I's and P's). In each instance, optimization under the more restrictive conditions leads to a cost of proofreading that is greater than that found under the less restrictive conditions.

The important point here is that in no case is the general cost–accuracy relation changed or violated; what determines the different costs are the specific conditions or constraints under which the optimization is performed. We have explored a number of these specific constraints to show that the degree to which the cost of proofreading approaches the general or unconstrained minimum depends upon the severity of the constraints imposed.

This work was supported in part by a grant to M.A.S. from the National Science Foundation. D.S.L. were supported in part by an NIH Postdoctoral Training Grant (5 T32 GM 07 123 07) to the University of Michigan.

REFERENCES

EDELMANN, P. & GALLANT, J. (1977). *Proc. natl. Acad. Sci. U.S.A.* **74,** 3396.
EHRENBERG, M. & BLOMBERG, C. (1980). *J. Biophys.* **31,** 333.
FRETER, R. R. & SAVAGEAU, M. A. (1980). *J. theor. Biol.* **85,** 99.
HOPFIELD, J. J. (1974). *Proc. natn. Acad. Sci. U.S.A.* **71,** 4135.
HOPFIELD, J. J., YAMANE, T., YUE, V. & COUTTS, S. M. (1976). *Proc. natn. Acad. Sci. U.S.A.* **73,** 1164.
LOFTFIELD, R. B. & EIGNER, E. A. (1966). *Biochim. biophys. Acta* **130,** 426.
LOFTFIELD, R. B. & VANDERJAGT, D. (1972). *Biochem. J.* **128,** 1353.
MULVEY, R. S. & FERSHT, A. R. (1977). *Biochemistry* **16,** 4731.
NINIO, J. (1975). *Biochimie* **57,** 587.
PAULING, L. (1958). In *Festschrift Arthur Stoll* pp. 597–602, Basel, Switzerland: Birkhaeuser.
SAVAGEAU, M. A. (1976). *Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology.* Reading, Massachusetts: Addison-Wesley.
SAVAGEAU, M. A. & FRETER, R. R. (1979*a*). *Biochemistry* **18,** 3486.
SAVAGEAU, M. A. & FRETER, R. R. (1979*b*). *Proc. natn. Acad. Sci. U.S.A.* **76,** 4507.
VAN VALKENBERG, M. E. (1960). *Modern Network Synthesis.* New York: Wiley.