J. theor. Biol. (1981) **93,** 179–195

Accuracy of Proofreading with Zero Energy Cost

MICHAEL A. SAVAGEAU

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

(*Received* 25 *November* 1980, *and in revised form* 4 *June* 1981)

Accuracy of biological discrimination at the molecular level is known in some systems to involve kinetic proofreading mechanisms. Hopfield and Ninio were the first to propose simple specific kinetic mechanisms for such proofreading and to demonstrate that an energy cost accompanies their improvement in accuracy. Savageau and Freter subsequently derived the explicit cost–accuracy relationship for a broad class of proofreading mechanisms, including the conventional Hopfield–Ninio mechanism just referred to. In other systems, the presence of proofreading mechanisms is in question because the diagnostic features of conventional kinetic proofreading are absent. However, Hopfield has recently proposed an alternative "energy-relay" mechanism, which lacks the characteristic features of conventional proofreading and yet is capable of improving accuracy. In this paper, I use the general cost–accuracy relationship that we have previously derived to examine the energy cost and accuracy of proofreading mechanisms involving an energy relay. The principal findings are the following. First, such mechanisms improve accuracy with a *zero cost of proofreading*, when "proofreading cost," defined as the cost due specifically to proofreading, is separated from the costs of putting material through the system. Second, the basic energy-relay mechanism discussed by Hopfield has only a modest improvement in accuracy, but a comparable improvement by a conventional proofreading mechanism would have a cost of about 0·0352 (moles ATP per mole of total product output). Third, accuracy can be increased somewhat if multiple stages of conventional kinetic proofreading precede the energy-relay mechanism. The cost for this improvement is zero while a comparable increase in accuracy achieved by conventional proofreading alone has a cost of about 0·0385. Finally, I propose an alternative arrangement of energy-relay mechanisms that is capable of increasing accuracy still further. The maximum accuracy achieved by this scheme at zero energy cost is comparable to that achieved by an infinite expenditure of energy in a single stage of conventional proofreading.

1. Introduction

The astonishing fidelity of biological replication processes was originally attributed to simple "lock-and-key" specificity of the enzymes that catalyze

0022–5193/81/210179+17 $02.00

these processes. It is now clear that in some of these processes the overall fidelity is the result of proofreading mechanisms that augment the "structural" specificity of enzymes (e.g. see the recent review in Hopfield, 1980). Hopfield (1974) and Ninio (1975) were the first to propose simple explicit kinetic mechanisms for proofreading and demonstrate their capacity to improve the accuracy of biological discrimination. They showed that the *maximum* accuracy was the product of the maximum values of the discrimination factors for the initial discrimination event and subsequent proofreading discrimination events. They also showed that such proofreading mechanisms have to be driven away from thermodynamic equilibrium and have an associated energy cost. Questions concerning the degree of accuracy actually realized by such systems and the amount of the associated energy cost remained to be answered.

The behavior of proofreading mechanisms, when operating under physiological conditions in the cell, can be analyzed by classical steady state enzyme kinetic methods (in which all reactant and modifier concentrations are constant). Their behavior also can be analyzed as a network of fluxes occurring in response to constant potentials (concentrations) at the boundaries. We have used this latter approach as the starting point for the development of a general theory of proofreading. We then used this approach, which has the advantage of yielding simple general constraints among the macroscopic variables of the system, to derive the explicit relationship between accuracy and the associated energy cost of proofreading (Savageau & Freter, 1979*a*). Our results showed that the maximum accuracy is never actually achieved, but only approached asymptotically at the expense of an infinite expenditure of energy. We demonstrated that the energy cost to achieve biologically acceptable levels of discrimination between structurally similar substrates with a single stage of proofreading was considerably higher than was indicated by earlier experimental studies (Savageau & Freter, 1979*a*,*b*) and, by a straightforward generalization, that proofreading effort can be distributed among multiple stages in an optimal way to reduce the total cost required and still achieve the same degree of accuracy (Savageau & Freter, 1979*a*; Freter & Savageau, 1980).

The details of our analysis of multiple stage proofreading reveal five macroscopic factors that affect the performance of these proofreading mechanisms. (1) The initial discrimination factor, (2) the number of proofreading stages, (3) the proofreading discrimination at each stage, (4) the distribution of proofreading effort among stages, and (5) the total energy expended for proofreading. The effectiveness of each of these factors in improving accuracy is subject to physical limitations and the optimal design for such systems is dependent upon their concerted action. It is important

to note that these conclusions concerning macroscopic properties apply to a wider class of proofreading mechanisms than the simple kinetic mechanisms originally proposed by Hopfield and Ninio and subsequently analyzed by other investigators (e.g., See Bennett, 1979; Savageau & Lapointe, 1981).

An example of an alternative proofreading mechanism that differs fundamentally from the simple schemes originally proposed has recently been described by Hopfield (1980) (see Fig. 1). In step γ a fraction of the energy

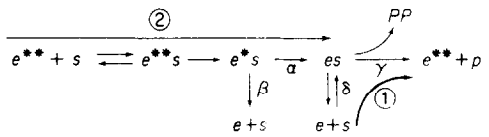

FIG. 1. Schematic representation of an energy-relay mechanism exhibiting dynamic cooperativity. *s*, substrate; *p*, product; *e* and e^{**}, different forms of free enzyme; $e^{**}s$, $e^{*}s$ and *es*, enzyme–substrate complexes; *PP*, pyrophosphate (after Hopfield, 1980). A similar diagram for incorrect substrate and product (s' and p') could also be drawn. See text for discussion.

of the substrate *s* is transferred to the enzyme *e* yielding the activated form e^{**}. A subsequent molecule of the same substrate is then selected by this form of the enzyme and proofread with a certain probability at step β while traversing pathway ② to *es*. Substrate molecules leaving by step β must re-enter via step δ to maintain a steady state; they are selected but *not* proofread while traversing pathway ①. In Hopfield's terms, this new mechanism uses an "energy-relay" from previous substrate molecules to affect proofreading of current substrate molecules and exhibits a form of "dynamic co-operativity". He has demonstrated that this mechanism is capable of improving accuracy beyond that obtainable by Michaelis–Menten binding alone when it is operating sufficiently far from thermodynamic equilibrium to ensure the undirectional behavior outlined above.

In this paper I shall use our previous results for general multiple stage proofreading to examine the accuracy and energy cost of proofreading mechanisms involving an "energy relay". These mechanisms are shown to have a relatively modest improvement in accuracy. However, the associated proofreading cost for this increase is zero. If multiple stages of conventional proofreading occur before the proofread molecules re-enter the main pathway, the accuracy can be improved somewhat, again at *zero* cost. However, if proofreading and re-entry of substrate molecules alternate along a pathway, then substantial improvements in accuracy can occur at *zero* cost. The limit in accuracy with an increasing number of proofreading

re-entry modules is identical to that achieved by a single stage of conventional kinetic proofreading with an infinite expenditure of energy.

2. Cost of Proofreading with Flux Fed Forward

In conventional kinetic proofreading a fraction of the entering flux is recycled or fed back to the beginning of the pathway during proofreading. According to the mechanism involving dynamic co-operativity, flux always moves forward through the pathway. The proofread flux is *not* recycled or fed back to the beginning of the pathway but is fed *forward* and re-enters the main pathway at a more distal point.

The scheme of Hopfield shown in Fig. 1 can be represented more abstractly by the branching diagram in Fig. 2. In comparing these figures

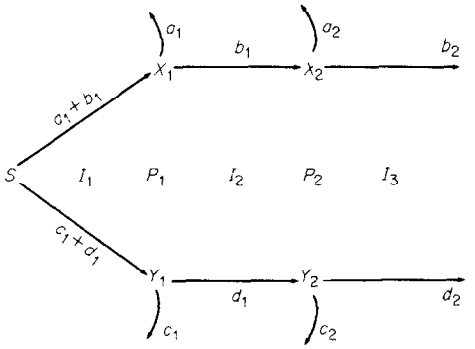

FIG. 2. Abstract branched diagram of an otherwise arbitrary network of reactions. S represents a selection system that interacts with correct and incorrect substrates. The X's represent sets of intermediate complexes involving correct substrate and the selection system. The Y's represent the same sets but with incorrect substrate. The arrows represent net flux between sets. The independent fluxes are designated a_1, a_2, b_2, c_1, c_2, and d_2. P_i is the macroscopic proofreading discrimination ratio and I_i is the macroscopic input discrimination ratio for the ith stage. I_3 is the output discrimination ratio. The macroscopic parameters are related to the elementary fluxes as follows: $P_i = b_i c_i/(a_i d_i)$, $I_1 = (a_1 + b_1)/(c_1 + d_1)$, $I_i = b_{i-1}/d_{i-1}$, the cost of proofreading for the ith stage $C_{i2} = (a_i + c_i)/(b_2 + d_2)$, and the total cost of proofreading $C_2 = C_{12} + C_{22}$. See text for further discussion.

one can make the following associations: S is e^{**}, X_1 is $e^{*}s$, X_2 is es, Y_1 is $e^{*}s'$, Y_2 is es', where s and s' refer to the correct and incorrect substrate, respectively. We have previously used such diagrams to represent the steady state pattern of macroscopic fluxes in a general proofreading system (Freter & Savageau, 1980). For this particular case, the cost of proofreading at the first stage $[C_{12} = (a_1 + c_1)/(b_2 + d_2)]$ is positive because the exit fluxes

a_1 and c_1 are positive. In contrast, the fluxes represented by a_2 and c_2 are negative, since they represent re-entry fluxes. The cost of this "second stage" of proofreading $[C_{22} = (a_2 + c_2)/(b_2 + d_2)]$ is therefore negative. Furthermore, in steady state

$$(a_1 + c_1) = -(a_2 + c_2) \tag{1}$$

so that the costs of the two stages are equal in magnitude but opposite in sign and the total cost of proofreading C_2 is given by

$$C_2 = C_{12} + C_{22} = 0. \tag{2}$$

Because the total cost of proofreading is zero and no flux is lost in passing through the system, one can easily show that the cost of putting flux through the system is identical to that for a conventional Michaelis–Menten selection mechanism with no proofreading; namely,

$$G = (1 - E)RT \ln (K_{eq}s/p) + ERT \ln (K'_{eq}s'/p') \tag{3}$$

where G is the Gibbs free energy per mole of total product output, $E = d_n'/(b_n + d_n)$ is the final net error or the incorrect fraction of total output flux, R is the gas constant, T is absolute temperature, and the primes signify concentrations and equilibrium constants associated with the incorrect substrate. Equation (3) may be rewritten as

$$G = RT \ln (K_{eq}s/p) + ERT \ln [(K'_{eq}s'p)/(K_{eq}sp')]$$

or

$$G = RT \ln \Gamma_p + (I + 1)^{-1} RT \ln (I/I_{eq}), \tag{4}$$

where, by definition, $\Gamma_p = K_{eq}s/p$, $I = (1 - E)/E$ is the final accuracy, $I = p/p'$ when the environment removes products non-selectively from the system, and $I_{eq} = K'_{eq}s'/(K_{eq}s)$ is the value of I at equilibrium.

In steady state with constant concentrations (or potentials) at the boundaries Γ_p will be constant. Thus, the energy to put flux through the system is given by a term dependent only on the potential difference $\ln \Gamma_p$ and one dependent upon the accuracy I. The dependence upon the latter term has a particularly simple form; it clearly has a single maximum beyond which the energy to put flux through the system decreases with increasing accuracy (see Fig. 3). For reasonable values of Γ_p, I_{eq} and I the contribution of this second term is negligible.

These results apply to systems with an arbitrary pattern of proofreading and re-entry stages since no flux is lost in passing through such systems.

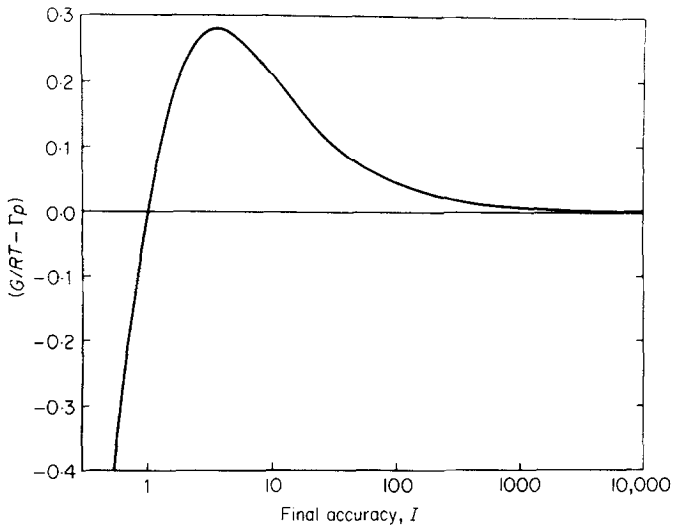

FIG. 3. Energy cost of putting flux through a proofreading system in which the cost of proofreading is zero. G is the Gibbs free energy per mole of total product output, R is the gas constant, T is absolute temperature, $\Gamma_p = K_{eq}s/p$, $I = I_{n+1} = b_n/d_n$ is the final accuracy, and I_{eq} is the value of I at equilibrium. In this case I_{eq} has a value of 1. See text for further discussion.

3. Accuracy of a Single Proofreading Stage with Flux Fed Forward

For convenience let us redraw Fig. 2 as shown in Fig. 4 in order to emphasize the equivalence of exit and re-entry fluxes. Hereafter, one can assume that the displacement from thermodynamic equilibrium is sufficiently large to ensure that the pattern of net flux corresponds to that indicated by the unidirectional arrows in the accompanying diagrams. Furthermore, the discrimination occurring upon re-entry is equivalent to another "initial discrimination event" (Hopfield, 1980) and the ratio of correct to incorrect flux re-entering the system can be represented by I_2^*.

$$I_2^* = a_2/c_2. \tag{5}$$

This macroscopic parameter is related to the proofreading discrimination ratio of the second stage $[P_2 = (c_2/d_2)/(a_2/b_2)]$ and the input discrimination ratio for the next stage or output discrimination ratio $[I_3 = b_2/d_2]$ by the simple relationship

$$P_2 = I_3/I_2^*. \tag{6}$$

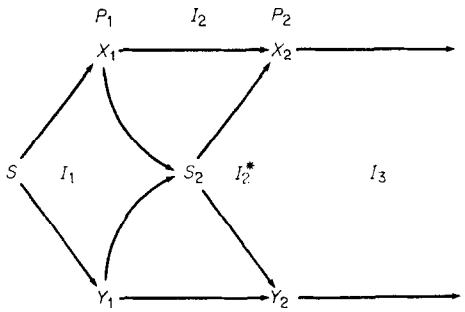

FIG. 4. Abstract representation of an energy-relay mechanism consisting of one module with a single stage of conventional proofreading and exit flux re-entering at a second stage (redrawn from Fig. 2). The exit or proofread fluxes $[a_1+c_1]$ are equal in magnitude to the re-entry fluxes $[-(a_2+c_2)]$. The ratio of correct to incorrect flux re-entering the system is represented by the input discrimination ratio to the second stage $I_2^* = a_2/c_2$. S and S_2 represent different forms of the selection system. Otherwise, the representation of the system is identical to that in Fig. 2. See text for further discussion.

The general cost–accuracy relationship previously derived for this two stage system (Savageau & Freter, 1979*a*) is

$$C_2 = \frac{(I_1+1)(P_1-1)(P_2-1)}{(I_3+1)\left(\frac{I_1}{I_2}P_1-1\right)\left(\frac{I_2}{I_3}P_2-1\right)} - 1. \tag{7}$$

By using the result in equation (6), together with the fact that $C_2 = 0$, one can rewrite equation (7) as

$$I_3 = \frac{I_2(I_2^*+1)+I_2^*C_1(I_2+1)}{(I_2^*+1)+C_1(I_2+1)}, \tag{8}$$

where C_1 is the cost of proofreading measured after the first stage

$$C_1 = \frac{(I_1+1)(P_1-1)}{(I_2+1)\left(\frac{I_1}{I_2}P_1-1\right)} - 1. \tag{9}$$

For fixed values of the macroscopic parameters I_1, I_2^* and P_1, what distribution of proofreading effort (or equivalently what value of I_2) will lead to the maximum value for the output discrimination I_3? This distribution can be determined by differentiating equation (8) with respect to I_2 and setting the equation equal to zero. The result is

$$\frac{\partial C_1}{\partial I_2} = \frac{(I_2^*+1)(C_1+1)}{(I_2+1)(I_2-I_2^*)} \tag{10}$$

but this derivative also can be evaluated directly from equation (9)

$$\frac{\partial C_1}{\partial I_2} = \frac{(C_1+1)}{I_2(I_2+1)}\left[1+\frac{(I_2+1)}{\left(\frac{I_1}{I_2}P_1-1\right)}\right]. \tag{11}$$

Equating the expressions in equations (10) and (11) yields the optimum value for I_2

$$I_2 = \sqrt{I_1 I_2^* P_1}. \tag{12}$$

The macroscopic discrimination at any step provides a concise representation for the net effect of local displacements from thermodynamic equilibrium and microscopic discrimination factors. Explicit relationships between macroscopic and microscopic parameters are quite complex in most cases. In general, though, macroscopic discrimination is less than related microscopic discrimination. They tend to approach one another under certain conditions, such as when displacement from thermodynamic equilibrium is increased. The two types of parameters are equal only under extremes of such conditions; otherwise, they each have a clear but distinct meaning and should not be equated. A general discussion contrasting macroscopic and microscopic parameters (Freter & Savageau, 1980) and a general technique for deriving explicit relationships involving these two types of parameters (Savageau & Lapointe, 1981) are given elsewhere.

Any real system operating in a given steady state is characterized by a specific set of values for the macroscopic parameters. By estimating these values, it is easy to determine the extent to which the system's performance approaches the optimum indicated above.

The results in this section also can be visualized by plotting I_3 as a function of I_2 directly from equations (8) and (9). An example is shown in Fig. 5*a*. For purposes of comparison, the final accuracy I_3 is plotted as a function of the intermediate accuracy I_2 in Fig. 5*b* for the conventional kinetic proofreading mechanism of Hopfield and Ninio having a fixed cost of proofreading. In each case, the maximum accuracy is achieved by an appropriate distribution of proofreading effort. However, the cost for the energy-relay mechanism is zero while that for the conventional kinetic proofreading mechanism is 0·0352 moles of ATP per mole of product output.

The numerical values for the macroscopic parameters in this example have been chosen equal to 100, and to each other, simply for convenience. One might wish to consider these as approximately equal to the corresponding microscopic parameters. This would be the case when the systems are

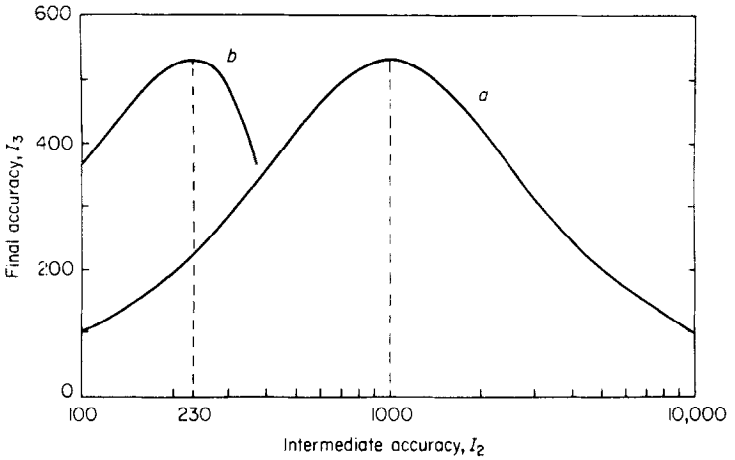

FIG. 5. Optimal distribution of proofreading effort and a comparison of the costs of proofreading for different mechanisms. *a* Results for the single module energy-relay mechanism represented in Fig. 4, *b* results for a conventional kinetic proofreading scheme with two stages of proofreading. The final accuracy I_3 in each case is determined as a function of the intermediate accuracy I_2 while the other macroscopic parameters are fixed. In case *a*, $I_1 = I_2^* = P_1 = 100$ and $C_2 = 0$. In case *b*, $I_1 = P_1 = P_2 = 100$ and $C_2 = 0{\cdot}0352$ (moles ATP per mole of total product output). The cost of proofreading in case *b* has been selected to yield an optimal final accuracy that is identical to that in case *a*. See text for further discussion.

greatly displaced from thermodynamic equilibrium, as is generally true under physiological conditions in the cell. Alternatively, one might wish to consider the displacements less and the microscopic parameter values greater, say 1000 as Hopfield (1980) assumed in his examples. In any case, one could use actual values for these macroscopic parameters and illustrate the same point.

4. Accuracy of Multiple Stage Proofreading with Flux Fed Forward

In the preceding section we saw that the energy-relay mechanisms proposed by Hopfield can achieve a modest increase in accuracy over that due to Michaelis–Menten binding. Because this increase can be had for zero cost it is attractive and we are led to ask whether the enhancement in accuracy might not be greater if multiple stages of proofreading were used.

Figure 6 is the abstract branched diagram that we have previously used to describe the macroscopic flux patterns in multiple stage proofreading systems. This is redrawn in Fig. 7 to emphasize that the exit flux is fed forward during proofreading to become the re-entry flux with its own

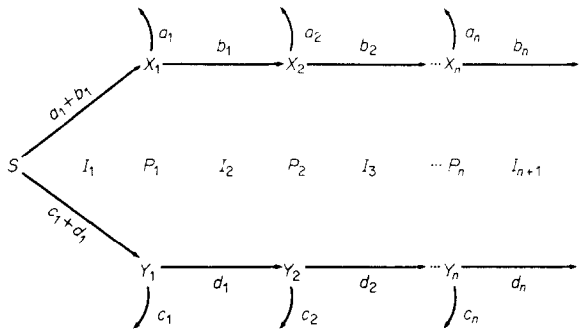

FIG. 6. Abstract branched diagram of an otherwise arbitrary network of reactions. S represents a selection system that interacts with correct and incorrect substrates. The X's represent sets of intermediate complexes involving correct substrate and the selection system. The Y's represent the same sets but with incorrect substrate. The arrows represent net flux between sets. The independent fluxes are $a_1, a_2, \ldots, a_n, b_n, c_1, c_2, \ldots, c_n$, and d_n. 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. These macroscopic parameters are related to the elementary fluxes as follows: $P_i = b_i c_i/(a_i d_i)$, $I_1 = (a_1+b_1)/(c_1+d_1)$, $I_i = b_{i-1}/d_{i-1}$, the cost of proofreading for the ith stage $C_{in} = (a_i+c_i)/(b_n+d_n)$, and the total cost of proofreading $C_n = C_{1n} + C_{2n} + \ldots + C_{nn}$. See text for further discussion.

"initial discrimination ratio" in the final stage. As in the previous section, the ratio of correct to incorrect flux reentering the system can be represented by I_n^* where

$$P_n = I_{n+1}/I_n^*. \tag{13}$$

The general cost–accuracy relationship previously derived for this n stage system (Freter & Savageau, 1980) is

$$C_n = \frac{(I_1+1)(P_1-1)\ldots(P_n-1)}{(I_{n+1}+1)\left(\frac{I_1}{I_2}P_1-1\right)\ldots\left(\frac{I_n}{I_{n+1}}P_n-1\right)} - 1. \tag{14}$$

By using the result in equation (13) and the fact that $C_n = 0$, one can rewrite this relationship as

$$I_{n+1} = \frac{I_n(I_n^*+1) + I_n^* C_{n-1}(I_n+1)}{(I_n^*+1) + C_{n-1}(I_n+1)}, \tag{15}$$

where C_{n-1} is the cost of proofreading determined after the first $(n-1)$ stages

$$C_{n-1} = \frac{(I_1+1)(P_1-1)\ldots(P_{n-1}-1)}{(I_n+1)\left(\frac{I_1}{I_2}P_1-1\right)\ldots\left(\frac{I_{n-1}}{I_n}P_{n-1}-1\right)} - 1. \tag{16}$$

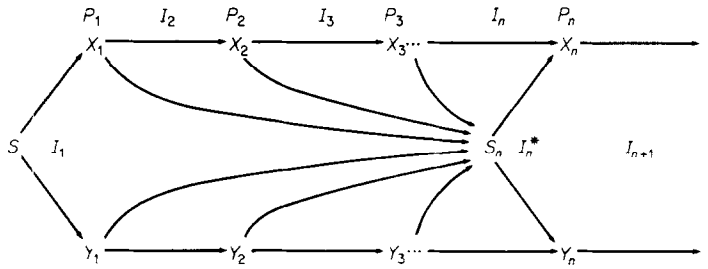

FIG. 7. Abstract representation of an energy-relay mechanism consisting of one module with n stages of conventional proofreading and exit fluxes re-entering at the nth stage (redrawn from Fig. 6). The exit or proofread fluxes $[a_1+a_2+\ldots+a_{n-1}+c_1+c_2+\ldots+c_{n-1}]$ are equal in magnitude to the re-entry fluxes $[-(a_n+c_n)]$. The ratio of correct to incorrect fluxes re-entering the system is represented by the input discrimination ratio to the nth stage $I_n^* = a_n/c_n$. S and S_n represent different forms of the selection system. Otherwise the representation of the system is identical to that in Fig. 6. See text for further discussion.

It is clear from equation (15) that, for a given value of I_n, I_{n+1} will be maximized if C_{n-1} is minimized. The problem of minimizing C_{n-1} for a given I_n has been treated elsewhere (Freter & Savageau, 1980); the unconstrained minimum occurs when the proofreading effort is optimally distributed among the stages and the P values are equal at their maximum value. Under these conditions equations (15) and (16) can be written

$$I_{n+1} = \frac{I_n(I_n^*+1)+I_n^* C_{n-1}^{\min}(I_n+1)}{(I_n^*+1)+C_{n-1}^{\min}(I_n+1)} \tag{17}$$

and

$$C_{n-1}^{\min} = \frac{(I_1+1)(P-1)^{n-1}}{(I_n+1)[(I_1/I_n)^{1/n-1}P-1]^{n-1}} - 1. \tag{18}$$

The optimum value of I_n can be determined by differentiating equation (17) with respect to I_n and setting the equation equal to zero. The result is

$$\frac{\partial C_{n-1}^{\min}}{\partial I_n} = \frac{(I_n^*+1)(C_{n-1}^{\min}+1)}{(I_n+1)(I_n-I_n^*)} \tag{19}$$

but, as before, this derivative also can be evaluated directly from equation (18) to give

$$\frac{\partial C_{n-1}^{\min}}{\partial I_n} = \frac{(C_{n-1}^{\min}+1)}{I_n(I_n+1)}\left\{1+\frac{(I_n+1)}{[(I_1/I_n)^{1/n-1}P-1]}\right\}. \tag{20}$$

Equating the expressions in equations (19) and (20) yields the optimum value for I_n

$$I_n = [I_1(I_n^*P)^{n-1}]^{1/n}. \tag{21}$$

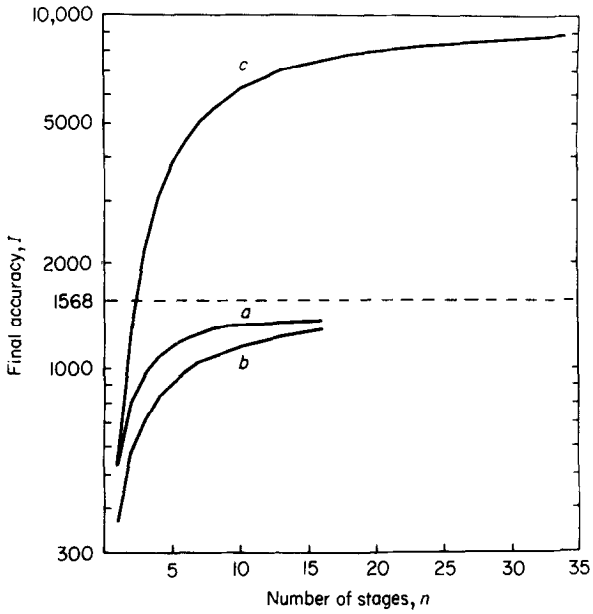

FIG. 8. Optimal accuracy as a function of the number of conventional proofreading stages, n. *a* Results for the single module energy-relay mechanism represented in Fig. 7; *b* results for a conventional kinetic proofreading scheme with no exit fluxes re-entering the system; *c* results for the m module energy-relay mechanism represented in Fig. 9. The final accuracy in each case is determined by optimization with respect to the intermediate accuracy values while the other macroscopic parameters are fixed. In case *a*, $I_1 = I_n^* = P_1 = P_2 = \ldots = P_{n-1} = P = 100$ and $C_n = 0$. In case *b*, $I_1 = P_1 = P_2 = \ldots = P_n = P = 100$ and $C_n = 0{\cdot}0385$ (moles ATP per mole of total product output). In case *c*, $I_1 = I_2^* = I_4^* = \ldots = I_{2m}^* = I^* = P_1 = P_2 = \ldots = P_{2m-1} = P = 100$ and $C_{2m+1} = 0$. The cost of proofreading in case *b* has been selected to yield an asymptotic value (1568) for final accuracy that is identical to that in case *a*. The asymptotic value for final accuracy in case *c* is 10^4. See text for further discussion.

In Fig. 8*a*, final accuracy I_{n+1} is plotted as a function of the number of proofreading stages n for fixed values of the macroscopic parameters I_1, I_n^*, and P. This curve is readily generated from equations (17), (18) and (21).

If n is allowed to increase without bound, then, from equation (21),

$$I_n \to I_n^* P = \hat{I}_n \tag{22}$$

from equation (40) in Freter & Savageau (1980),

$$C_{n-1}^{\min} \to \frac{(I_1+1)}{(I_n^*P+1)}\left[\frac{I_n^*P}{I_1}\right]^{P/(P-1)} - 1 = \hat{C} \tag{23}$$

and, from equation (17),

$$I_{n+1} = \frac{I_n^* P(I_n^* + 1) + I_n^* \hat{C}(I_n^* P + 1)}{(I_n^* + 1) + \hat{C}(I_n^* P + 1)} = \hat{I}_{n+1}. \tag{24}$$

This last expression, with $I_1 = I_n^*$, provides the asymptotic value for curve *a* in Fig. 8.

For comparison, final accuracy I_{n+1} is plotted as a function of n in Fig. 8*b* for the conventional kinetic proofreading mechanism of Hopfield–Ninio having a fixed cost of proofreading. The cost (0·0385 moles of ATP per mole of total product output) has been selected to give the same accuracy asymptote as in curve *a*. At every finite n the accuracy of the energy-relay mechanism is greater than that of the conventional Hopfield–Ninio mechanism, and the cost in the former case is zero. Again, the numerical values for the macroscopic parameters have been chosen for convenience, but one could use other values and illustrate the same point.

5. An Alternative Multiple Stage Proofreading Mechanism

In the preceding section, accuracy for the mechanism in Fig. 7 was shown to increase with increasing number of conventional proofreading stages. Although the cost remained zero, the enhancement of accuracy was still relatively modest (maximum accuracy for the example in Fig. 8*a* is 1568). Perhaps other configurations might be superior.

Figure 9 is a schematic representation of an alternative multiple stage proofreading mechanism involving a concatination of two stage modules as analyzed in section 3. From the results presented there one can write a set of recursive relations describing the m module case.

$$I_{2i} = \sqrt{I_{2i-1} I_{2i}^* P_{2i-1}}, \tag{25}$$

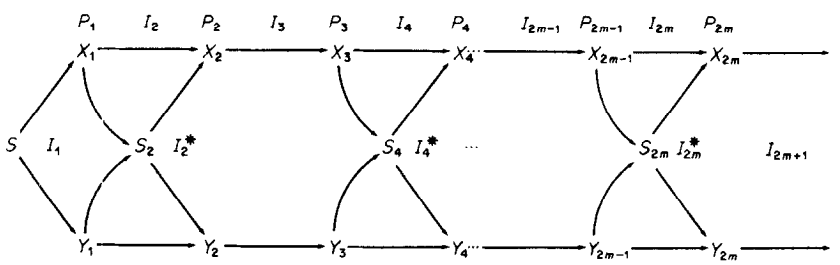

FIG. 9. Abstract representation of an energy-relay mechanism consisting of m modules each with a single stage of conventional proofreading and exit flux re-entry. The details of the representation for each module are given in Fig. 4. See text for further discussion.

$$C_{2i-1} = \frac{(I_{2i-1}+1)(P_{2i-1}-1)}{(I_{2i}+1)\left(\frac{I_{2i-1}}{I_{2i}}P_{2i-1}-1\right)} - 1, \tag{26}$$

$$I_{2i+1} = \frac{I_{2i}(I_{2i}^*+1) + I_{2i}^* C_{2i-1}(I_{2i}+1)}{(I_{2i}^*+1) + C_{2i-1}(I_{2i}+1)} \qquad i = 1, 2, \ldots, m. \tag{27}$$

Substituting equation (25) into equations (26) and (27), and the resulting equation (26) into equation (27) yields the following relationship

$$I_{2i+1} = \frac{I_{2i-1}[(I_{2i}^*+1)(P_{2i-1}-1)] + [\sqrt{I_{2i}^* P_{2i-1}} - \sqrt{I_{2i-1}}]^2}{[(I_{2i}^*+1)(P_{2i-1}-1)] - [\sqrt{I_{2i}^* P_{2i-1}} - \sqrt{I_{2i-1}}]^2} \qquad i = 1, 2, \ldots, m. \tag{28}$$

If it is assumed that the modules are identical and thus that $I_{2i}^* = I^*$ and $P_{2i-1} = P$, then the last of equations (28) can be written

$$I_{2m+1} = \frac{I_{2m-1}[(I^*+1)(P-1)] + [\sqrt{I^* P} - \sqrt{I_{2m-1}}]^2}{[(I^*+1)(P-1)] - [\sqrt{I^* P} - \sqrt{I_{2m-1}}]^2} \tag{29}$$

and in the limit as $m \to \infty$

$$I_{2m+1} \to I_{2m-1} \to I^* P. \tag{30}$$

Thus, the maximum accuracy achieved under these conditions with zero cost is identical to the accuracy achieved by a comparable single stage of conventional kinetic proofreading with *infinite* cost (see Savageau & Freter, 1979*a*).

Figure 8*c* is a plot of the final accuracy I_{2m+1} as a function of the number of modules m, which in this case also is equal to the total number of conventional proofreading stages n. This result is obtained by using equation (29) with fixed values of the macroscopic parameters I_1, I^* and P. It is interesting that while the values on this curve for small n are dependent upon the value of the initial discrimination I_1, the asymptotic value for large n is independent of I_1. This is different from the results in section 4 where it was found that the asymptotic value of final accuracy I_{n+1} is a function of I_1.

If multiple stages of conventional proofreading (rather than a single stage) are associated with each module of the mechanism sketched in Fig. 9, perhaps the accuracy could be enhanced still further. By using equations (17), (18) and (21) repeatedly (i.e. letting the final accuracy I_{n+1} of one stage be the initial accuracy I_1 of the next) one can plot the final accuracy as a function of the number of modules m, each with n stages of conventional proofreading (Fig. 10). Moreover, if n is allowed to increase

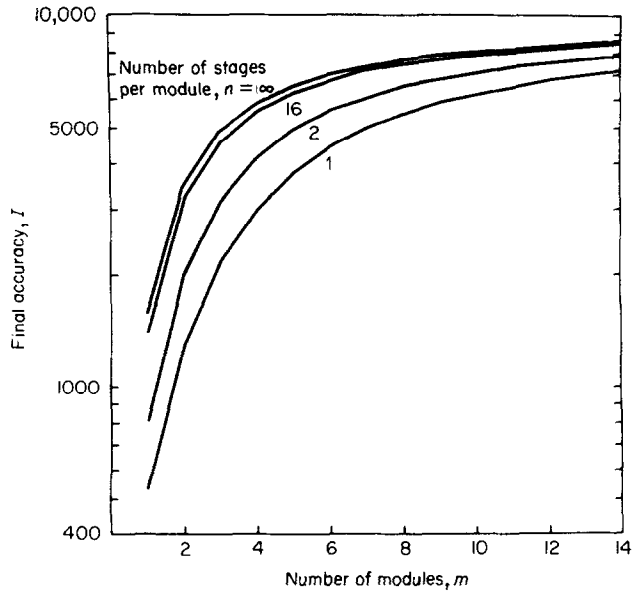

FIG. 10. Optimal accuracy of an energy-relay mechanism as a function of m the number of modules each having n stages of conventional proofreading. The accuracy of each module is optimized with respect to the intermediate accuracy values of the module while all other macroscopic parameters are fixed. $I_1 = 100$, $I^* = 100$ for all modules, $P = 100$ for all conventional proofreading stages within each module, and cost = 0. The maximum accuracy under these conditions is $I^*P = 10^4$. See text for further discussion.

without bound for each module, then equations (22), (23) and (24) can be used repeatedly to obtain the limiting curve in Fig. 10. The asymptote of these curves is given by I^*P, which can be verified readily by substitution into equations (17), (18) and (21). Thus, the maximum accuracy cannot be enhanced further by increasing the number of stages of conventional proofreading per module. Again, other numerical values for the parameters could be used but the point of the example remains the same.

If the total number of stages of conventional proofreading is fixed, then accuracy is maximized by distributing them one to a module.

6. Discussion

The general cost–accuracy relationship and macroscopic analysis previously described (Savageau & Freter, 1979*a*; Freter & Savageau, 1980) applies to a broader class of proofreading mechanisms than that originally

proposed by Hopfield (1974) and Ninio (1975). The energy-relay mechanism recently proposed by Hopfield (1980), and analyzed in this paper, is just one example of alternative proofreading mechanisms that can be described by this formalism.

It is important to emphasize that the optimal distribution of proofreading effort described in the preceding sections is a general or unconstrained optimum, and, as discussed in detail elsewhere (Freter & Savageau, 1980; Savageau & Lapointe, 1981), this optimum may not be reached in specific cases. The particular constraints associated with specific mechanisms may prevent the unconstrained optimum from being reached. A general matrix method of obtaining the constraint relationship for particular mechanisms has been developed (Savageau & Lapointe, 1981) and can be readily applied to the mechanisms analyzed in this paper.

Hopfield (1980) has pointed out several diagnostic features of the energy-relay mechanism and discussed its potential biological significance. I need only discuss the cost–accuracy relations that emerged from the analysis in the preceding sections. The most dramatic result, of course, is that accuracy can be enhanced by proofreading without an accompanying energy cost.

The degree of accuracy enhancement for a single module is rather modest (e.g. in Fig. 5*a*, a maximum of 530 vs. 100 for simple Michaelis–Menten binding). Systems with multiple stages of conventional proofreading followed by an energy-relay mechanism exhibit somewhat greater accuracy (e.g. in Fig. 8*a*, a maximum increase of 1568 for an infinite number of stages (in practice about 30) vs. 530 for a module with a single stage of conventional proofreading).

However, an alternative mechanism involving a concatination of modules each with a single conventional proofreading stage is capable of increasing accuracy still further (e.g. in Fig. 8*c*, a maximum increase of 10^4 for an infinite number of modules (in practice about 30) vs. 530 for a single module) at zero cost. This figure of 10^4 is identical to the accuracy achieved by a comparable single stage of conventional proofreading with an infinite expenditure of energy (Savageau & Freter, 1979*a*). This degree of accuracy is sufficient to account for that observed when one amino acid is accidentally substituted for another in proteins of organisms as diverse as mammals (Loftfield & Eigner, 1966; Loftfield & Vanderjagt, 1972) and bacteria (Edelmann & Gallant, 1977; Gallant & Foley, 1979). The maximum achievable accuracy cannot be increased further by additional stages of conventional kinetic proofreading within each module.

This work was supported in part by a grant from the National Science Foundation.

REFERENCES

BENNETT, C. H. (1979). *Biosystems* **11,** 85.
EDELMANN, P. & GALLANT, J. (1977). *Proc. natn. Acad. Sci. U.S.A.* **74,** 3396.
FRETER, R. R. & SAVAGEAU, M. A. (1980). *J. theor. Biol.* **85,** 99.
GALLANT, J. A. & FOLEY, D. (1979). In: *Regulation of Macromolecular Synthesis by Low Molecular Weight Mediators* (G. Koch and D. Richter eds). New York: Academic Press. pp. 5–23.
HOPFIELD, J. J. (1974). *Proc. natn. Acad. Sci. U.S.A.* **71,** 4135.
HOPFIELD, J. J. (1980). *Proc. natn. Acad. Sci. U.S.A.* **77,** 5248.
LOFTFIELD, R. B. & EIGNER, E. A. (1966). *Biochim. biophys. Acta* **130,** 426.
LOFTFIELD, R. B. & VANDERJAGT, D. (1972). *Biochem. J.* **128,** 1353.
NINIO, J. (1975). *Biochimie* **57,** 587.
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.
SAVAGEAU, M. A. & LAPOINTE, D. S. (1981). *J. theor. Biol.* **93,** 157.