

Proofreading Systems of Multiple Stages for Improved Accuracy of Biological Discrimination

ROLF R. FRETER AND MICHAEL A. SAVAGEAU

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

(Received 20 September 1979, and in revised form 8 February 1980)

The phenomenal accuracy of biological discrimination is due in many cases to specific proofreading mechanisms. We have previously developed a macroscopic theory of such mechanisms and applied it to the case of single-stage proofreading. In this article we apply the theory to systems with multiple stages of proofreading. A specific relationship between improved accuracy due to proofreading and the associated energy cost is given. This is a macroscopic relationship that must be satisfied regardless of the details of the underlying mechanisms. Five factors in the design of such systems are shown to influence their overall accuracy: (1) initial discrimination, (2) number of proofreading stages, (3) proofreading discrimination of each stage, (4) distribution of proofreading effort among the stages, and (5) total energy expended for proofreading. We show that there is an optimal distribution of proofreading effort that, for a given degree of accuracy, minimizes the energy cost of proofreading. We also provide a simple physical interpretation of this minimum condition. These results are used to examine proofreading in two experimental systems for which there is appropriate data available in the literature: the valyl-tRNA synthetase catalyzed misacylation of tRNA^{Val} with threonine and the isoleucyl-tRNA synthetase catalyzed misacylation of tRNA^{Ile} with valine. The correlation between the magnitude of a discrimination factor and the size of the corresponding enzymatic cavity is discussed.

1. Introduction

It has often been customary to think of discrimination in biological systems as occurring with near perfect fidelity. However, it is now known that in many cases discrimination is accomplished with relatively error-prone components and that the relatively error-free overall result is due to specific proofreading mechanisms. For example, such mechanisms were first reported to be associated with DNA replication by Goulian, Lucas & Kornberg (1968) and Brutlag & Kornberg (1972); with RNA transcription by Springgate & Loeb (1975) and Volloch, Rits & Tumerman (1979); with

tRNA aminoacylation by Baldwin & Berg (1966), Yarus (1972), and Schreier & Schimmel (1972); and with ribosomal coding in translation by Gorini (1971), Grosjean, Söll & Crothers (1976), and Thompson & Stone (1977). [See also the review of earlier work by Ninio (1975*a*).]

It was recognized in the earliest of these studies that the accuracy of systems involving a single stage of proofreading is determined by the initial discrimination and the discrimination of the proofreading stage. It was also clear that such proofreading had an energy cost. Many questions remained unanswered, however. At one level ("microscopic") these questions concern the identification and characterization of the detailed enzyme-kinetic mechanisms by which proofreading is physically realized. At another level ("macroscopic") these questions concern the overall or "black-box" behavior of the proofreading system: How much can accuracy be improved by such systems and at what cost to the cell that harbors them?

Several different theoretical approaches to proofreading have been proposed to deal with such questions. These approaches can be classified as microscopic or macroscopic, depending on which level provides their point of departure.

The pioneering theoretical studies of Hopfield (1974) and Ninio (1975) have been the principle stimulus for subsequent theoretical work and much of the recent experimental work as well. They used the microscopic approach, beginning with simple kinetic models, to explore the requirements for proofreading. They showed that such systems must be displaced from thermodynamic equilibrium to achieve an improvement in accuracy. They also showed that the *maximum* accuracy that can be achieved is given by the product of the initial discrimination factor and subsequent proofreading discrimination factors. These factors can be defined in a variety of ways, including ratios of binding or release rates, probability factors and temporal delays required for testing the identity of a molecule.

These early microscopic approaches have been extended in various ways by several authors. In the context of ribosomal accuracy, Blomberg (1977) analyzed in more detail the early models, including specific cases with several proofreading stages. Kurland (1978) refined Hopfield's (1974) treatment of energy driven processes by introducing a parameter that indicates the degree to which the activated compound (e.g. ATP or GTP) is displaced from thermodynamic equilibrium. For the analysis of DNA polymerase accuracy, Galas & Branscomb (1978), Bernardi & Ninio (1978) and Bernardi *et al.* (1979) have developed approaches that combine elements of the earlier microscopic theories with a random-walk process. All of the studies referred to above develop descriptions for the overall behavior of proofreading systems by starting from detailed kinetic or

probability considerations, i.e. these theories are built from the "bottom-up".

In contrast to these microscopic approaches, a macroscopic approach to proofreading was used by Goodman *et al.* (1974) and Galas & Branscomb (1978) and subsequently developed by Savageau & Freter (1979*a*). Its point of departure is the macroscopic pattern of fluxes through the selection system as these exist in the cell under physiological (or experimental) steady-state conditions. No assumptions specifically regarding the detailed mechanisms that generate the particular flux patterns are made, rather these processes are viewed as "black boxes". While this approach, by definition, cannot distinguish between different mechanisms at the microscopic level, it has the unique ability to provide relatively simple, general relationships among the parameters that characterize the system at the macroscopic level. For example, our theory yields an explicit relationship between the accuracy achieved by single-stage proofreading and the corresponding energy cost. Experimental data available in the literature for tRNA aminoacylation have been examined in light of this relationship. The principle conclusions were: (1) High-accuracy proofreading has a high energy cost, as much as 100 times greater than indications from early experimental work. (2) The maximum accuracy derived in early theoretical studies is never actually reached. Instead, lower values, determined by the balance between energy wasted in the cell as a consequence of error and the energy cost of proofreading, appear to be selected (Savageau & Freter, 1979*a*; 1979*b*). Again, by definition, such macroscopic relationships must be satisfied by any specific mechanism. This has been demonstrated with regard to the macroscopic cost-accuracy relationship and the detailed mechanism analyzed by Galas & Branscomb (1978).

In this paper we present our theory as applied to multiple-stage proofreading systems. The general approach and principal conclusions have been outlined elsewhere (Savageau & Freter, 1979*a*). First, we briefly review single-stage proofreading. Second, we show that double-stage proofreading can be understood by repeated application of the results for a single stage. Third, multiple-stage proofreading is considered as an overlapping sequence of double-stage systems. We derive general cost-accuracy relationships for such systems and demonstrate the existence of an optimal distribution of proofreading effort that minimizes the total cost. Fourth, we discuss the macroscopic factors that determine overall accuracy: initial discrimination, number of proofreading stages, proofreading discrimination of each stage, distribution of proofreading effort among the stages and energy expended for proofreading. Finally, we use our results to analyze published experimental data for tRNA aminoacylation.

2. Single-stage Proofreading

For simplicity of exposition, we consider a model defined by the following assumptions (Savageau & Freter, 1979a): (1) A single type of recognition error occurs on a given enzyme. (2) Proofreading takes place in a single stage on the surface of the same enzyme that made the original error. (3) These reactions form part of a symmetrical network. (4) The system operates in a normal steady state. Such a model is represented schematically in Fig. 1. The arrows represent net flux in pathways that are generally reversible, although hydrolytic proofreading reactions, which release the energy of an activated compound (hereafter considered ATP for simplicity), are highly exergonic and essentially irreversible in the cases of interest.

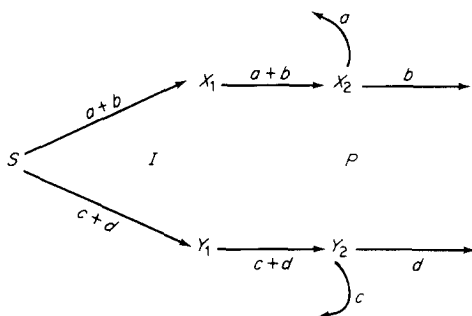


FIG. 1. Schematic model representing a single recognition error and subsequent proofreading stage. The symbols are as follows: S , free enzyme; X_1 and X_2 , enzyme-intermediate complexes following correct substrate recognition; Y_1 and Y_2 , enzyme-intermediate complexes following incorrect substrate recognition. Independent fluxes are represented by a , b , c and d . I is the initial discrimination ratio and P is the proofreading discrimination ratio.

In the more general case, the "nodes" in Fig. 1 represent an arbitrarily complex submechanism consisting of several complexes and reactions. One can think of such a submechanism as surrounded by a boundary (imaginary). Since the fluxes entering and leaving the boundary must satisfy Kirchoff's law in the steady state, the submechanism within will behave as a "node" as far as the macroscopic pattern of fluxes between "nodes" is concerned. The "arrows" in Fig. 1 also can represent arbitrarily complex submechanisms consisting of several complexes and reactions that connect one "node" to another. One can think of an arbitrary plane passing between two "nodes". The net flux represented by the arrow between the two "nodes" is equal to the sum of the fluxes in the submechanism that pass through the plane. Of course the Haldane relationships of the underlying enzymatic mechanisms

must be satisfied, but it is impossible to state these in general because their forms depend upon the specifics of the mechanism.

(A) ELEMENTAL FLUX EQUATIONS

The state of this system is determined at the macroscopic level by the conservation of mass and the values of the four fluxes a , b , c and d . One could pursue the microscopic approach and in principle relate these macroscopic fluxes to the elementary rate constants, which define the system, and the independent concentrations, "correct" as well as "incorrect", that affect it. However, the end result for any realistic mechanism tends to be unwieldy. It is more convenient and general to represent the behavior of the system by four macroscopic parameters defined in terms of the above fluxes: *initial discrimination ratio*, $I = (a + b)/(c + d)$; *proofreading discrimination ratio*, $P = (c/d)/(a/b) = bc/(ad)$; *net error*, $E = d/(b + d)$; and *cost of proofreading*, $C = (a + c)/(b + d)$. In steady state, C differs trivially from the thermodynamic energy cost by a multiplicative constant, namely, the conjugate potential that is fixed in the cell (or, *in vitro*, by the experimentalist). For example, thermodynamic energy cost = GC , where $G = RT \ln \{[ATP] K_{eq}/[AMP][PP]\}$ is a constant in the cell. This is the usual situation of physiological interest (see also Discussion). Therefore, under these conditions one can legitimately consider C as a measure of energy utilization.† C is also related to the ratio R/F , where R is *recycling* ($a + c$) and F is *net flux* ($b + d$).

These macroscopic parameters should not be viewed as rigid constants but as parametric variables. Their magnitude may be considered fixed for given systems under specific conditions. As the systems and/or conditions change, the magnitudes of these parameters may change and thus reflect changes in elementary rate constants (microscopic parameters) or concentrations, including those of high energy compounds that are used to displace the system from equilibrium. It should be clear that these macroscopic parameters are related to but should not be confused with the microscopic parameters used in the microscopic approaches referred to in the Introduction.

(B) COST-ACCURACY RELATIONSHIP

The equations in the previous subsection can be used to obtain an explicit relationship between E , net error or accuracy, and C , cost of proofreading.

† This is analogous to the use of an ammeter to measure energy utilization when the current is provided at a constant (line) voltage.

$$E = \frac{[(P-1) - (1+C) - I(1+PC)] + \sqrt{[(P-1) - (1+C) - I(1+PC)]^2 + 4(P-1)(I+1)(1+C)}}{2(P-1)(I+1)} \quad (1)$$

or, alternatively,

$$C = \frac{[1 - (I+1)E][1 + (P-1)E]}{[(IP+1)E - 1]} \quad (2)$$

As previously discussed (Savageau & Freter, 1979a), these relationships represent a general macroscopic constraint that must be satisfied for any branched network of the form considered, regardless of the particular reaction mechanism and conditions that determine the discrimination at each branch point. We shall assume (unless otherwise stated) that there are no other restrictions on these parameters. Later we shall consider the consequences of relaxing this assumption. A complete representation of this general constraint would require a four-dimensional space, but for convenience we can consider a parametric representation in a two-dimensional plot. For any real system operating in a steady state there will be unique values for the parameters I and P . These in turn define a particular cost-accuracy curve. There are an infinite number of systems in the class that lies along this curve. The specific value of C for the system in question then determines uniquely the accuracy (or E). In this manner, one can determine where on the cost-accuracy curve the particular system operates in relation to the maximum accuracy possible for the class of systems to which it belongs. In other words, one can determine how close the design of the system is to 1 that maximizes accuracy (given I , P and external conditions) irrespective of the cost, C . As $C \rightarrow 0$ the net error approaches the low-cost asymptote $E_0 = 1/(I+1)$; as $C \rightarrow \infty$ the net error approaches the high-cost asymptote $E_\infty = 1/(IP+1)$.† This cost-accuracy relationship is shown graphically in Fig. 2. For a comparison of these relationships and experimental data for tRNA aminoacylation, see Savageau & Freter (1979a).

In general, this approach provides one limit for accuracy, but there may be others. For a given difference in structure between substrates, type of proofreading mechanism and set of external conditions, there will be maximum values for I and P that yield another (less restrictive) high-cost asymptote; there also may be more restrictive limits placed directly upon the allowable ranges of E and C for particular mechanisms. Since these latter limits would be dependent upon the particularities of the mechanism, there is little to conclude in general.

† The condition that $C \rightarrow \infty$ with finite displacements clearly implies that $F \rightarrow 0$, i.e. this maximum accuracy occurs when there is zero net flux through the system.

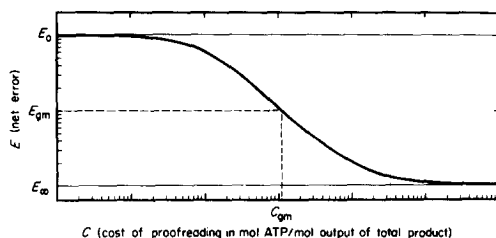


FIG. 2. Cost-accuracy relationship for a system with single-stage proofreading. The low-cost error asymptote is E_0 ; the high-cost error asymptote is E_{∞} ; the geometric mean error is E_{gm} ; and the cost to produce the geometric mean error is C_{gm} . C is given by equation (2). The scales are logarithmic.

3. Double-stage Proofreading

When two stages of proofreading are involved, the same theory can be applied to each stage in succession. The ratio of "correct" to "incorrect" flux emerging from the first stage can be regarded as the "apparent" initial discrimination ratio of the second stage. A model for double-stage proofreading is shown schematically in Fig. 3.

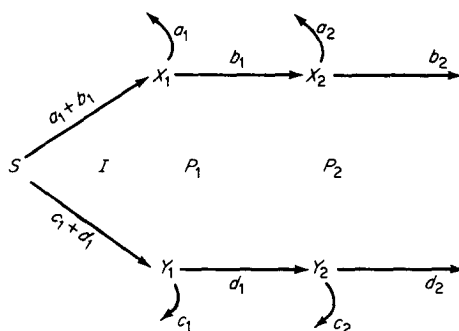


FIG. 3. Schematic model representing a single recognition error and two subsequent stages of proofreading. The independent fluxes are a_1 , a_2 , b_2 , c_1 , c_2 , and d_2 . I is the initial discrimination ratio; P_1 and P_2 are the proofreading discrimination ratios. The model is otherwise identical to that in Fig. 1. See text for further discussion.

(A) ELEMENTAL FLUX EQUATIONS

In this case there are six independent fluxes a_1 , a_2 , b_2 , c_1 , c_2 , and d_2 that can be used to define other parameters of interest.

The *initial discrimination ratio*, I , is defined as the ratio of correct input flux to incorrect input flux,

$$I = \frac{a_1 + b_1}{c_1 + d_1} = \frac{a_1 + a_2 + b_2}{c_1 + c_2 + d_2}. \quad (3)$$

A *proofreading discrimination ratio*, P , is defined for a given stage as the ratio of incorrect flux recycled to that continuing on divided by the ratio of correct flux recycled to that continuing on. For the first stage of proofreading:

$$P_1 = \frac{b_1 c_1}{a_1 d_1} = \frac{(a_2 + b_2) c_1}{a_1 (c_2 + d_2)}. \quad (4)$$

For the second stage

$$P_2 = \frac{b_2 c_2}{a_2 d_2}. \quad (5)$$

A *recycling*, R , is the correct plus incorrect flux leaving the enzyme at a given stage of proofreading. For the first stage

$$R_1 = a_1 + c_1, \quad (6)$$

and for the second stage

$$R_2 = a_2 + c_2. \quad (7)$$

A *net flux*, F , is the output flux, correct plus incorrect, from a given stage of proofreading. The net flux from the first stage is

$$F_1 = b_1 + d_1, \quad (8)$$

and that from the second stage is

$$F_2 = b_2 + d_2. \quad (9)$$

A *net error*, E , is the incorrect flux divided by the net flux from a given stage of proofreading. After the first stage of proofreading

$$E_1 = d_1 / (b_1 + d_1), \quad (10)$$

and after the second stage

$$E_2 = d_2 / (b_2 + d_2). \quad (11)$$

A *cost of proofreading*, C , is defined as the ratio of a recycling flux to the final net flux. The cost of proofreading at the first of two stages is

$$C_{12} = \frac{R_1}{F_2} = \frac{a_1 + c_1}{b_2 + d_2}, \quad (12)$$

at the second of two stages

$$C_{22} = \frac{R_2}{F_2} = \frac{a_2 + c_2}{b_2 + d_2}, \quad (13)$$

and for the total of the two stages

$$C_2 = C_{12} + C_{22} = \frac{a_1 + a_2 + c_1 + c_2}{b_2 + d_2}. \quad (14)$$

An *apparent cost of proofreading* can be defined as the ratio of recycling flux to net flux for an intermediate stage of proofreading. For the first of two stages

$$C_{11} = \frac{R_1}{F_1} = \frac{a_1 + c_1}{b_1 + d_1}. \quad (15)$$

This cost differs from the real cost C_{12} , which is properly normalized with respect to the final net flux. The real cost and the apparent cost of a given stage of proofreading are related as follows:

$$\begin{aligned} C_{12} &= \frac{R_1}{F_2} = \frac{R_1}{F_1} \left(\frac{F_1}{F_2} \right) = C_{11} \left(\frac{R_2 + F_2}{F_2} \right) \\ &= C_{11} \left(\frac{R_2}{F_2} + 1 \right) = C_{11} (C_{22} + 1). \end{aligned} \quad (16)$$

Thus, the *total cost of double-stage proofreading*, C_2 , can be related to the apparent costs as follows:

$$C_2 = (C_{11} + 1)(C_{22} + 1) - 1. \quad (17)$$

These parameters now can be used to obtain an explicit relationship between cost and accuracy of double-stage proofreading.

(B) COST-ACCURACY RELATIONSHIP

The apparent cost for the first of two stages of proofreading, C_{11} , can be obtained directly from the cost-accuracy relationship for single-stage proofreading [see equation (2)].

$$C_{11} = \frac{[1 - (I + 1)E_1][1 + (P_1 - 1)E_1]}{[(IP_1 + 1)E_1 - 1]}. \quad (18)$$

If we consider the ratio of “correct” to “incorrect” flux emerging from the first stage of proofreading as the “apparent” initial discrimination ratio for the second stage (I_2),

$$I_2 = (1 - E_1)/E_1, \quad (19)$$

then the cost of the second stage of proofreading also can be obtained directly from equation (2).

$$C_{22} = \frac{[1 - (1/E_1)E_2][1 + (P_2 - 1)E_2]}{\{(1 - E_1)/E_1\}P_2 + 1}E_2 - 1. \quad (20)$$

Thus, by repeated application of our results for single-stage proofreading and by the use of equation (17) we are able to write the total cost of double-stage proofreading as

$$C_2 = \frac{(I + 1)(P_1 - 1)(P_2 - 1)(1 - E_1)(1 - E_2)E_1E_2}{[(IP_1 + 1)E_1 - 1][(1 - E_1)P_2E_2 - (1 - E_2)E_1]} - 1. \quad (21)$$

A complete representation of this general constraint would require a six-dimensional space. Again, for convenience, we can consider a parametric representation in a two-dimensional cost-accuracy plot. The low-cost asymptote $E_{20} = 1/(I + 1)$ and is approached as $C_2 \rightarrow 0$. The high-cost asymptote $E_{2\infty}$ lies between a maximum value of $1/(IP_1 + 1)$ or $1/(IP_2 + 1)$, whichever is larger, and a minimum value of $1/(IP_1P_2 + 1)$. The precise value in the latter case depends on the distribution of proofreading effort and will be approached as $C_2 \rightarrow \infty$. For a given net error, E_2 , the total cost, C_2 , will vary as a function of the distribution of proofreading effort, and thus as a function of the intermediate net error, E_1 . Although cases with $P_i \leq 1$ can be treated, they are not physiologically interesting because no real proofreading occurs in such stages. In the optimization that follows such cases would receive none of the proofreading effort and the system would be reduced effectively to one with fewer proofreading stages, all of which have $P_i > 1$. Thus, without loss of generality we shall consider only systems with $P_i > 1$.

(C) MINIMUM COST

For any given system operating in a steady state there will be unique values for the parameters I , P_1 , P_2 , and E_2 . These in turn can be thought of as parameters that define a class of systems, each having a different distribution of proofreading effort and hence a different value for the parameter E_1 . For this class of systems, the total cost, C_2 , varies with the distribution of proofreading effort and one can determine which distribution (or value of E_1) yields the minimum cost. One then can compare the system in question, which has a particular distribution of proofreading effort, with the system having the optimal distribution for the class to which they belong. The value of E_1 that minimizes the total cost can be obtained by differentiating C_2 with respect to E_1 and setting the derivative equal to zero.

$$E_{1 \min} = \left[\sqrt{\frac{IP_1(1-E_2)}{P_2 E_2}} + 1 \right]^{-1}. \quad (22)$$

Substituting this value into equation (21) yields the minimum value of the total cost C_2 .

$$C_{2 \min} = \frac{(I+1)(P_1-1)(P_2-1)(1-E_2)E_2}{[\sqrt{IP_1 P_2 E_2} - \sqrt{(1-E_2)}]^2} - 1. \quad (23)$$

The physical meaning of this minimum condition can be obtained from equation (22) by substituting back the flux relationships in equations (3)–(11).

$$\frac{c_1/(c_1+d_1)}{a_1/(a_1+b_1)} = \frac{c_2/(c_2+d_2)}{a_2/(a_2+b_2)}. \quad (24)$$

Hence, minimum cost occurs when the percent recycling of incorrect and correct input fluxes has a common ratio for each stage of proofreading (see also Fig. 3).

With given values for accuracy and external conditions, the most general, or unconstrained, minimum for C_2 occurs when I has its maximum value, P_1 and P_2 have the same maximal value and the mechanism is optimized as shown above. There may be a more restrictive minimum for C_2 if the constraints of a particular mechanism do not permit optimization of design with P_1 and P_2 at this same maximum value. Another more restrictive minimum for C_2 occurs if the constraints of a particular mechanism do not allow both of the parameters P_1 and P_2 to achieve the same maximum value assumed above (even though the design may then be optimized for the different, lower, maximum values of P_1 and P_2). Finally, an even more restrictive minimum for C_2 would be obtained if the constraints of a particular mechanism allowed neither achievement of the same maximum value for P_1 and P_2 nor optimization.

(D) COMPARISON OF SINGLE- AND DOUBLE-STAGE PROOFREADING

One easily can conceive of situations in which single-stage proofreading is superior to double-stage proofreading, as well as situations in which the reverse is true. By superior we mean that proofreading costs to achieve a given degree of accuracy are less, or that the net error resulting from a given expenditure for proofreading is less. In general, for a given recognition error

(identical I values) one stage is better than two stages whenever

$$P > \frac{[(P_1 - 1)(P_2 - 1)(1 - E_1)(1 - E_2)E_1] - [(IP_1 + 1)E_1 - 1][(1 - E_1)P_2E_2 - (1 - E_2)E_1]}{[I(P_1 - 1)(P_2 - 1)(1 - E_1)E_1E_2] - [(IP_1 + 1)E_1 - 1][(1 - E_1)P_2E_2 - (1 - E_2)E_1]} \quad (25)$$

This condition is obtained by comparing the costs of single-stage proofreading [equation (2)] and double-stage proofreading [equation (21)] when their final net errors are identical (i.e. $E = E_2$). However, there are conditions in which two stages always appear to be better than one.

For our first comparison we shall consider the initial discrimination ratios equal and all the proofreading discrimination ratios equal for the two systems. For this condition, double-stage proofreading is always superior.

The cost of double-stage proofreading is given by equation (21) and, when $P_1 = P_2 = P$, it can be rewritten

$$C_2 = \frac{(I + 1)(P - 1)^2(1 - E_1)(1 - E_2)E_1E_2}{[(IP + 1)E_1 - 1][(1 - E_1)PE_2 - (1 - E_2)E_1]} - 1. \quad (26)$$

When the two systems yield the same final net error, $E = E_2$ and the cost of single-stage proofreading given in equation (2) can be rewritten

$$C_1 = \frac{[1 - (I + 1)E_2][1 + (P - 1)E_2]}{[(IP + 1)E_2 - 1]}. \quad (27)$$

For both cases, the errors must remain within the asymptotic limits for the system with single-stage proofreading,[†] i.e.

$$1/(IP + 1) < E_2 < E_1 < 1/(I + 1). \quad (28)$$

The difference between C_1 and C_2 is given by

$$C_1 - C_2 = \frac{(I + 1)(P - 1)P(1 - E_2)E_2(E_1 - E_2)[1 - (I + 1)E_1]}{[(IP + 1)E_1 - 1][(IP + 1)E_2 - 1][(1 - E_1)PE_2 - (1 - E_2)E_1]} \quad (29)$$

and is always positive. Thus for this case, double-stage proofreading is always more economical than single-stage proofreading. This can be seen graphically in Fig. 4. Alternatively, the cost of proofreading always can be distributed among two stages of proofreading to yield a system that is more accurate than one in which the same total amount of energy is expended in a single stage of proofreading.

[†] For some distributions of proofreading effort, E_2 for the double-stage system can have values between $1/(IP + 1)$ and $1/(IP^2 + 1)$, whereas E_2 for the single-stage system cannot. The double-stage system is obviously superior in these cases.

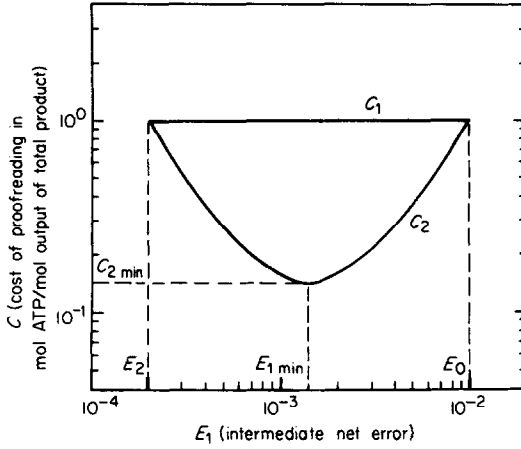


FIG. 4. The cost of single- and double-stage proofreading to achieve a given net error. The cost of single-stage proofreading is a constant C_1 . The cost of double-stage proofreading C_2 varies with the distribution of proofreading effort and, thus, with E_1 , the intermediate net error. Minimum cost for double-stage proofreading occurs at $E_{1\min}$. The curves are drawn by using the following parameter values: $I = P_1 = P_2 = P = 100$ and $E_2 = 2 \times 10^{-4}$.

For our next comparison, we allow the proofreading discrimination ratios of the double-stage system to differ, which is more physiologically realistic. This might occur if the detailed mechanisms of proofreading were different for the two stages. However, if the single- and double-stage systems are indeed comparable, then the single-stage system would be likely to involve one or the other of the proofreading mechanisms found in the double-stage system. Thus, we will assume that one of the two proofreading discrimination ratios of the double stage system is equal to the proofreading discrimination ratio of the single-stage system. For these conditions, the *minimum* cost of double-stage proofreading is always less than the corresponding cost of single-stage proofreading.

The minimum cost of double stage-proofreading is given in equation (23). For the same accuracy, the cost of single-stage proofreading is given by equation (27). The difference in cost is given by

$$C_1 - C_{2\min} = \frac{(I + 1)(1 - E_2)E_2\{[\sqrt{IP_2E_2} - \sqrt{P_1(1 - E_2)}]^2(P - 1) + [1 - (I + 1)E_2](P_1 - 1)(P_2 - P)\}}{[\sqrt{IP_1P_2E_2} - \sqrt{(1 - E_2)}]^2[(IP + 1)E_2 - 1]} \quad (30)$$

which is positive for $P_2 \geq P$. By symmetry† this difference is also positive for

† Expansion of the squared term in the numerator is necessary to reveal the symmetry of the expression in braces.

$P_1 \geq P$. Thus, if at least one of the proofreading discrimination values for the double-stage system is greater than or equal to the P value for the single-stage system, then the minimum cost of double-stage proofreading is always less than the corresponding cost of single-stage proofreading. Conversely, under these conditions the accuracy of the double-stage system will be greater than that of the single-stage system provided the minimum cost of proofreading in two stages is equal to the corresponding cost of proofreading in one stage.

For our last comparison we shall consider the limiting case. As indicated previously, the initial and proofreading discrimination ratios are limited ultimately by the structural differences between the molecules being recognized and the mechanism of proofreading used at each stage (given external conditions, adenylate ratios, etc.). The "correct" and "incorrect" molecules at each stage have the same structural differences. Thus, for situations in which the same mechanism of proofreading is used at each stage of an optimally designed selection system, these discrimination ratios are likely to have the same limiting value. When the discrimination ratios are all equal and have the limiting value, the difference in cost given by equation (30) will be a maximum.

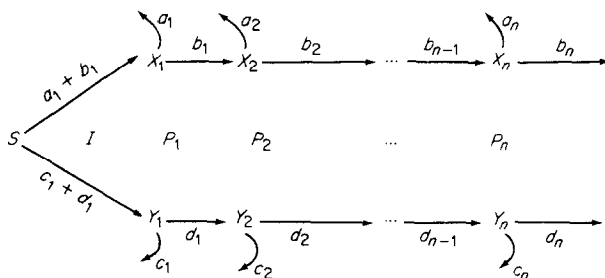


FIG. 5. Schematic model representing a single recognition error and n subsequent stages of proofreading. The independent fluxes are $a_1, a_2, \dots, a_n, b_n, c_1, c_2, \dots, c_n,$ and d_n . I is the initial discrimination ratio. P_1, P_2, \dots, P_n are the proofreading discrimination ratios. The model is otherwise similar to that in Figs 1 and 3. See text for further discussion.

4. Multiple-stage Proofreading

The cost-accuracy relationship for a system involving n stages of proofreading[†] can be developed by repeated application of the results presented in the preceding sections. The model for such a system is shown schematically in Fig. 5.

[†] Multiple stages can be realized biologically by proofreading several different enzyme-bound intermediates along a given pathway. The intermediate-enzyme complexes may differ by chemical and/or conformational changes.

(A) ELEMENTAL FLUX EQUATIONS

In the general case there are $2(n + 1)$ independent fluxes $a_1, \dots, a_n, b_n, c_1, \dots, c_n,$ and d_n that can be used to define the other parameters of interest. By direct extension of the double-stage case: the initial discrimination ratio $I = (a_1 + b_1)/(c_1 + d_1)$; the proofreading discrimination ratio of the i th stage $P_i = b_i c_i / (a_i d_i)$; the recycling in the i th stage $R_i = a_i + c_i$; the net flux from the i th stage $F_i = b_i + d_i$; the net error following the i th stage $E_i = d_i / (b_i + d_i)$; the cost of proofreading for the i th stage $C_{in} = R_i / F_i = (a_i + c_i) / (b_i + d_i)$; the apparent cost of proofreading for the i th stage $C_{ii} = R_i / F_i = (a_i + c_i) / (b_i + d_i)$; and the total cost

$$C_n = \left[\prod_{i=1}^n (C_{ii} + 1) \right] - 1. \tag{31}$$

(B) COST-ACCURACY RELATIONSHIP

The apparent cost for the first stage, C_{11} , is given in equation (18). If we consider the ratio of "correct" to "incorrect" flux emerging from the $(i - 1)$ th stage as the "apparent" initial discrimination ratio for the i th stage (I_i),

$$I_i = (1 - E_{i-1}) / E_{i-1} \tag{32}$$

then the apparent cost of the i th stage of proofreading also can be obtained directly from equation (2).

$$C_{ii} = \frac{(E_{i-1} - E_i)[1 + (P_i - 1)E_i]}{[(1 - E_{i-1})P_i E_i - (1 - E_i)E_{i-1}]}, \quad i = 2, 3, \dots, n. \tag{33}$$

Substituting equations (18) and (33) into the expression for total cost [equation (31)] yields

$$C_n = \frac{(I + 1)(P_1 - 1)(P_2 - 1) \dots (P_n - 1)(1 - E_1)(1 - E_2) \dots (1 - E_n)E_1 E_2 \dots E_n}{[(IP_1 + 1)E_1 - 1][(1 - E_1)P_2 E_2 - (1 - E_2)E_1] \dots [(1 - E_{n-1})P_n E_n - (1 - E_n)E_{n-1}]} - 1. \tag{34}$$

A complete representation of this general constraint would require a $2(n + 1)$ -dimensional space. Again, for convenience, we can consider a parametric representation in a two-dimensional cost-accuracy plot. The low-cost asymptote for this function is $E_{n0} = 1 / (I + 1)$, which is approached as $C_n \rightarrow 0$. The high-cost asymptote $E_{n\infty}$ lies between the maximum of the $1 / (IP_i + 1)$ values and $1 / (IP_1 P_2 \dots P_n + 1)$. The precise value in the latter

case depends on the distribution of proofreading effort and will be approached as $C_n \rightarrow \infty$. For a given net error E_n , the total cost C_n will vary as a function of the distribution of proofreading effort, and thus as a function of the intermediate net errors E_1, E_2, \dots, E_{n-1} . As discussed in the previous section, we shall treat only systems for which $P_i > 1$.

(C) MINIMUM COST

By direct extension of the double-stage case, the values for E_1, E_2, \dots, E_{n-1} that lead to minimum total cost can be obtained by calculating the partial derivatives of C_n with respect to E_i and setting them to zero. In each case, the derivative involves a function of the same form:

$$\frac{\partial}{\partial E_i} \left\{ \frac{(1-E_i)E_i}{[(1-E_{i-1})P_i E_i - (1-E_i)E_{i-1}][(1-E_i)P_{i+1}E_{i+1} - (1-E_{i+1})E_i]} \right\} = 0 \quad (35)$$

and the problem reduces to that of finding the minimum for a series of overlapping double-stage systems, which we have solved in the previous section. Thus, at the minimum

$$E_i = \left[\sqrt{\frac{P_i(1-E_{i-1})(1-E_{i+1})}{P_{i+1}E_{i-1}E_{i+1}}} + 1 \right]^{-1}, \quad i = 1, 2, \dots, n-1 \quad (36)$$

where $I = (1-E_0)/E_0$. These recursive equations can be solved for the E_i in terms of the final net error E_n .

$$E_{i \min} = \left[\frac{IP_1 P_2 \dots P_i}{(IP_1 P_2 \dots P_n)^{i/n}} \left(\frac{1-E_n}{E_n} \right)^{i/n} + 1 \right]^{-1}, \quad i = 1, 2, \dots, n-1. \quad (37)$$

Substituting these values into equation (34) yields the minimum value of the total cost C_n .

$$C_{n \min} = \frac{(I+1)(P_1-1)(P_2-1) \dots (P_n-1)(1-E_n)E_n}{[(IP_1 P_2 \dots P_n E_n)^{1/n} - (1-E_n)^{1/n}]^n} - 1. \quad (38)$$

The physical meaning of this minimum condition can be obtained by repeated application of the results in equation (24) for each double stage in succession. Again, minimum cost occurs when the per cent recycling of incorrect and correct input fluxes has a common ratio for each stage of proofreading (see Fig. 5):

$$\frac{c_1/(c_1+d_1)}{a_1/(a_1+b_1)} = \frac{c_2/(c_2+d_2)}{a_2/(a_2+b_2)} = \dots = \frac{c_n/(c_n+d_n)}{a_n/(a_n+b_n)}. \quad (39)$$

Again, with given values for accuracy and external conditions, the most general or unconstrained minimum for C_n occurs when the discrimination ratios have the same maximum value and the mechanism is optimized as shown above. There may be a more restrictive minimum for C_n if the constraints of a particular mechanism do not permit optimization and/or do not allow all of the discrimination ratios to achieve the same maximum value previously assumed.

(D) COMPARISON OF $(n - 1)$ - AND n -STAGE PROOFREADING

Although there are some situations in which $(n - 1)$ -stage proofreading is superior to n -stage proofreading, under appropriate conditions the reverse is true. This can be demonstrated easily by use of the results in equations (29) and (30).

For systems with identical proofreading discrimination ratios (P), one can replace the i th stage of an $(n - 1)$ -stage system with two stages, each having proofreading discrimination ratio P , that together yield the same net error E_i , thereby constructing a corresponding n -stage system. Since the double-stage subsystem is always more economical than the corresponding single-stage subsystem [see equation (29) and Fig. 4], the n -stage system of proofreading is superior to the $(n - 1)$ -stage system.

For the more physiologically realistic systems in which the proofreading discrimination ratios (P_1, P_2, \dots, P_n) are not identical, one can replace the i th stage of an $(n - 1)$ -stage system with two stages, one having proofreading discrimination ratio P_i , and the other having an arbitrary value, that together yield the same net error E_i at minimum cost. In this manner one constructs a corresponding n -stage system. Since the minimum cost of the double-stage subsystem is always less than that of the corresponding single-stage subsystem [see equation (30)], there is at least one n -stage system of proofreading that is superior to the $(n - 1)$ -stage system.

Even if the proofreading effort should be distributed so as to minimize cost in either of the two $(n - 1)$ -stage systems considered above, the cost of proofreading for the corresponding n -stage system still must be less [i.e. $C_n < C_{(n-1) \min}$]. Thus, the minimum cost of proofreading for the corresponding n -stage system also must be less [i.e. $C_{n \min} < C_{(n-1) \min}$]. This can be seen in Fig. 6, or from the relationship between $C_{n \min}$ for a given net error and the number of identical proofreading stages as depicted in Fig. 7. In the limit as $n \rightarrow \infty$

$$C_{n \min} = (I + 1) \left[\frac{(1 - E_n)^P}{I^P E_n} \right]^{1/(P-1)} - 1. \tag{40}$$

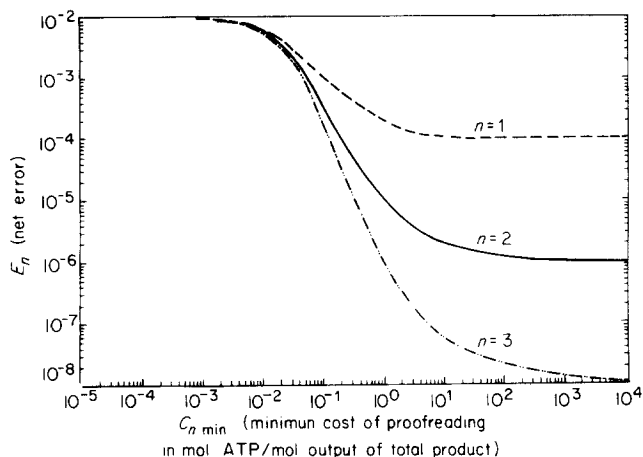


FIG. 6. Minimum cost-accuracy relationships for systems with multiple-stage proofreading. The low-cost error asymptote $E_0 = 1/(I + 1)$ for all systems and the high-cost asymptote $E_{n\infty} = 1/(IP^n + 1)$, where $I = 100$ and $P_i = 100$ for all proofreading stages. n is the number of stages. $C_{n\min}$ is given by equation (38). The scales are logarithmic. See Fig. 2 for additional detail.

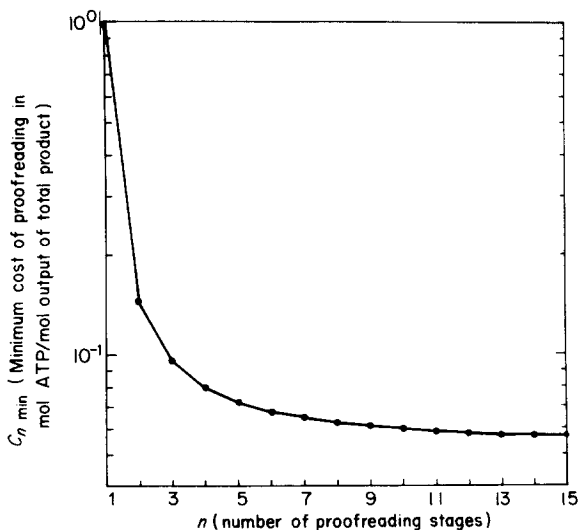


FIG. 7. Minimum cost of multiple-stage proofreading to achieve a given net error. n is the number of stages. Net error $E_n = 2 \times 10^{-4}$. The initial discrimination ratio $I = 100$ and all proofreading discrimination ratios $P_i = 100$. $C_{n\min}$ is given by equation (38). The cost scale is logarithmic. The asymptotic value of $C_{n\min}$ is given by equation (40).

Most of the reduction in cost of proofreading occurs with the increase from one to two or three stages; beyond three stages the reduction in cost is minimal. This expression represents the minimum cost to achieve a given accuracy, or the maximum accuracy that can be achieved with a given energy expenditure. Other than the restriction that the P values be identical, this result is mechanism independent. When P has the maximum value this represents the least restrictive minimum for C_n , independent of n .

5. Discussion

Before discussing our results we shall briefly consider alternative formalisms. In their recent microscopic formalisms Kurland (1979), and Ehrenberg & Blomberg (1980) have stressed the importance of a parameter (the ratio of adenylate or guanylate compounds) that indicates the degree of displacement from thermodynamic equilibrium. Such a parameter is quite useful for experimental studies of accuracy *in vitro* (see Jelenc & Kurland, 1979) because it can be varied independently by the investigator. However, the adenylate or guanylate compounds should not be considered unique in this regard. There are additional microscopic parameters implicit in these formalisms that could be varied independently to influence accuracy (e.g. see also Jelenc & Kurland, 1979). Furthermore, as all of the above authors have shown in their microscopic approaches, there are many microscopic parameters that are necessary and important in influencing accuracy. Hence, the displacement, although necessary and important, should not be accorded a "favored" status over the other microscopic parameters. This is especially true under steady-state physiological conditions for reasons that we shall discuss next.

The microscopic parameters affect the macroscopic pattern of fluxes passing through the selection system; this macroscopic pattern of fluxes ultimately determines the accuracy and cost of proofreading (see previous sections). From the macroscopic point of view the fluxes are the key parameters, but this is not to say that the degree of displacement from equilibrium is unimportant. Obviously displacements, which need not be infinite, must accompany the fluxes and such displacements are implicit in our formalism. We have said this before, but it must be emphasized. Even more important, *in vivo*, the degree of displacement of activated compounds is not an independent variable, but one that is determined by the interaction of a vast number of cellular reactions that generate or consume energy. Therefore, this displacement cannot be adjusted specifically to optimize a

particular process; rather, the process must be optimized by other means subject to the existing displacement.†

Blomberg & Ehrenberg (1980) and Ehrenberg & Blomberg (1980) are the only other authors who have addressed the optimization problem, but they have done so from the microscopic perspective. Their approach has the advantage of yielding detailed information about the particular mechanism analyzed, but the disadvantage that the results are not general. This is complementary to the situation with the macroscopic approach we have used. Because of the complexity of the microscopic problem, their formalism is quite cumbersome. Nevertheless, many of the conclusions they have reached, in so far as they relate to macroscopic properties of the system, agree with our conclusions in the special case of equal P values, which is the only case they considered. This is to be expected because any specific mechanism, such as that treated by Blomberg and Ehrenberg, must conform to the general macroscopic relationships that we have presented [equation (34)]. This correspondence can be seen in easy straight-forward comparisons under special circumstances in which a one-to-one relationship between the parameters in the two approaches is established. This correspondence clearly exists under other circumstances as well, but then the comparisons that demonstrate equivalence are no longer obvious because the parameters, although related to each other, are no longer related in a simple one-to-one fashion.

On the other hand, Ehrenberg & Blomberg (1980) have derived mechanism-dependent results for which there is no obvious correspondence in our general results. For example, they give a number of explicit relationships showing that the accuracy of their selection system increases with the displacement of the adenylate compounds from equilibrium. In general, from a macroscopic point of view, we can show easily that accuracy is related to such a displacement. However, the relationship depends critically on the specifics of the mechanism‡ and, thus, there is no meaningful general conclusion to be made.

In the preceding sections, and in the analysis presented elsewhere (Savageau & Freter, 1979a), we have shown that the accuracy of a biological proofreading system depends on several macroscopic factors. These are the magnitudes of the initial (I) and proofreading (P_i) discrimination ratios, the

† The potential of activated compounds in the cell is analogous to the line voltage in your home. The design of various home appliances is not optimized by having the power company change the line voltage it provides. Instead their design is optimized in other ways subject to the 120 V line voltage at the end of the plug.

‡ One can conceive of mechanisms for which the accuracy will actually decrease with increasing displacement of the adenylate compounds from equilibrium.

total amount of energy (C_n) expended for proofreading, the number (n) of proofreading stages and the distribution of proofreading effort among the stages. Increasing the magnitude of I or P_i , other factors remaining constant, will decrease net error E_n . Alternatively, an increase in I or P_i will decrease the energy cost of proofreading C_n required to achieve a given net error E_n . However, the finite ability of an enzyme to discriminate between structurally similar correct and incorrect substrates limits the accuracy (or cost reduction) that can be achieved in this manner. Increasing the total amount of energy expended for proofreading also will decrease net error. The relationship between net error E_n and total cost C_n is shown in Fig. 6. The lower limit for net error $1/(IP_1P_2 \dots P_n + 1)$ can be approached only as $C_n \rightarrow \infty$. Thus, the degree of accuracy that can be achieved by this method is limited by the finite energy available in the cell and the need for metabolic reactions to proceed with nonzero velocities. Increasing the number of proofreading stages from one to two or three can significantly reduce net error as well. The relationship between minimum cost $C_{n \min}$ and the number of stages n is shown in Fig. 7. This method of increasing accuracy also has limitations. Increasing n beyond small integral values yields only a minimal reduction in cost. The distribution of proofreading effort among the stages also affects net error. The distribution that minimizes net error (or total cost) is given in equation (39). Although each of the factors discussed has the potential to considerably reduce net error, the reduction actually realized is limited by biological constraints. Thus, a system whose errors are particularly critical may require a concerted change in several of these parameters to achieve a biologically acceptable error rate. The extent to which each of the above factors might contribute to the reduction of error in tRNA aminoacylation will now be examined.

We will examine two experimental systems: the valyl-tRNA synthetase catalyzed misacylation of tRNA^{Val} with threonine and the isoleucyl-tRNA synthetase catalyzed misacylation of tRNA^{Ile} with valine. The net errors and proofreading costs for these systems previously have been estimated to be $E = 2.0 \times 10^{-7}$ and $C = 0.053$ for valyl-tRNA synthetase and $E = 2.2 \times 10^{-4}$ and $C = 0.12$ for isoleucyl-tRNA synthetase (Savageau & Freter, 1979a). These values were calculated by combining the data of Mulvey & Fersht (1977) and Yamane & Hopfield (1977). These data involve measurements of net tRNA aminoacylated from a known limiting amount of ATP. These data are therefore independent of the number of proofreading stages and uniquely determine the values of net error and cost. Within these constraints, however, the number of proofreading stages and the magnitudes of the individual proofreading discrimination ratios may vary. The consequence of such variation will now be discussed.

Previously, we assumed a single stage of proofreading for valyl-tRNA synthetase and estimated the value of the proofreading discrimination ratio as 7500 (Savageau & Freter, 1979*a*). We will now assume two stages of proofreading in which the first stage corresponds to proofreading the enzyme-bound aminoacyl-adenylate and the second stage corresponds to proofreading the enzyme-bound aminoacyl-tRNA. Incomplete transfer of threonine from the enzyme-bound threonyl-adenylate complex to tRNA^{Val} has been observed (Fersht & Kaethner, 1976), and may be assumed to result from proofreading of the enzyme-bound threonyl-adenylate. This has been suggested in the similar case of valyl-adenylate bound to isoleucyl-tRNA synthetase (Fersht, 1977). By further assuming that the proofreading discrimination ratios are equal and that the proofreading effort is distributed so as to minimize cost, the proofreading discrimination ratios are calculated to be $P_1 = P_2 = 730$. The theoretical cost-accuracy curves for this valyl-tRNA synthetase system with single- and double-stage proofreading are shown in Fig. 8.

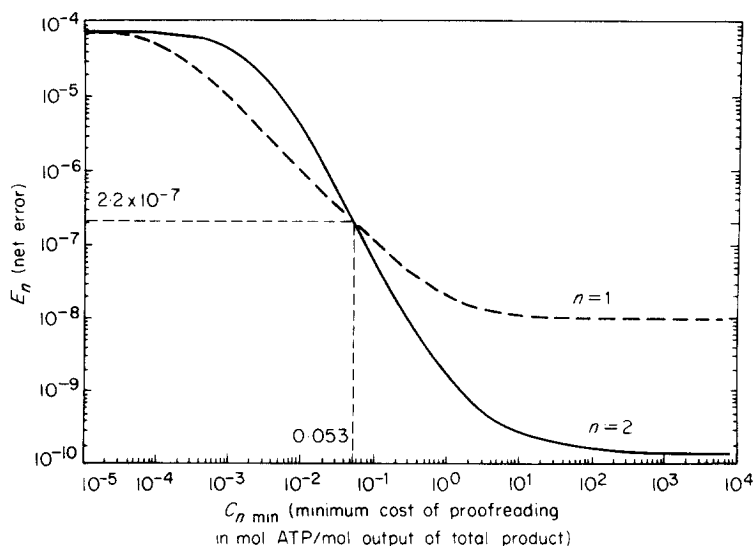


FIG. 8. Minimum cost-accuracy curves for valyl-tRNA synthetase proofreading threonyl-tRNA^{Val}. Data taken from Mulvey & Fersht (1977) and Yamane & Hopfield (1977). Proofreading is assumed to occur in one ($I = 13\,000$ and $P = 7500$) or two ($I = 13\,000$ and $P_1 = P_2 = 730$) stages. The scales are logarithmic. See text for further discussion.

Similarly, we have previously assumed a single stage of proofreading for isoleucyl-tRNA synthetase and estimated the value of the proofreading discrimination ratio as 4200 (Savageau & Freter, 1979*a*). If we assume two

stages of proofreading as suggested by Fersht (1977), we can proceed as in the previous paragraph to calculate proofreading discrimination ratios of $P_1 = P_2 = 490$. As in the previous case, one can increase the number of stages and concomitantly reduce the magnitude of the proofreading discrimination ratios while maintaining the same cost and net error.

According to the assumptions employed elsewhere, the net error in the case of isoleucyl-tRNA synthetase is estimated to be greater than the Loftfield limit (Savageau & Freter, 1979*b*). However, there are at least three ways in which this error can be reduced to the Loftfield limit: (1) A threefold increase in cost C (from 0.12 to 0.32) (see Fig. 2, Savageau & Freter, 1979*b*), (2) a fourfold increase in proofreading discrimination ratio P (from 1100 to 4200) (see Fig. 6, Savageau & Freter, 1979*a*), and (3) an increase in number of stages n (from 1 to 2) and a decrease in the proofreading discrimination ratios (from $P = 1100$ to $P_1 = P_2 = 490$) (see previous paragraph).

An alternative to the assumption of equal P values given above for isoleucyl-tRNA synthetase will now be considered. The value of 100, previously estimated (Savageau & Freter, 1979*a*) from the data of Schreier & Schimmel (1972) for the aminoacyl-tRNA proofreading discrimination ratio, will be assumed to roughly approximate the true value of P_2 . By using the values $I = 18$, $C_{2\text{min}} = 0.12$ and $E_2 = 2.2 \times 10^{-4}$ given earlier, and equation (23), one can calculate the corresponding value of $P_1 = 2000$.[†]

In some cases the magnitude of the proofreading discrimination ratio is indicative of the size of the enzymatic cavity involved in proofreading. For example, in the case of valyl-tRNA synthetase the initial selectivity against isoleucine is at least 10^5 (Fersht, 1979). Presumably the active site of this synthetase is constructed to precisely fit valine. Hence, 10^5 may be considered the *maximum value* for the proofreading discrimination ratio between isoleucine and valine and is related to the energy required for cramming an additional methyl group into an enzymatic cavity that is too small. For the double-stage proofreading mechanism described in the preceding paragraph, the magnitudes of the proofreading discrimination ratios against valine were calculated to be 2000 and 100.[‡] It appears that

[†] This value of P_1 (and the value $P = 4200$ previously given) may be considered unrealistically high from the standpoint of discrimination against valine and for isoleucine. However, these values need not be considered excessive from the standpoint of discrimination against isoleucine and for valine, as pointed out by Crick (1975) and developed more fully by Fersht (1979) in his "double sieve" mechanism of proofreading.

[‡] The first of these values apparently corresponds to a "cavity" somewhat larger than valine. The second value corresponds more closely to a cavity the size of isoleucine. This second cavity, which appears to function differently from the first, could preferentially protect isoleucyl-tRNA from hydrolysis. The associated proofreading ratio of 100 is identical to the initial discrimination ratio for this enzyme (Savageau & Freter, 1979*a*) ($I = 100$ when equal concentrations of amino acid substrates are assumed).

neither proofreading mechanism involves a cavity constructed precisely for valine. In fact, each appears to involve a cavity that is significantly larger than need be, suggesting that isoleucyl-tRNA synthetase may proofread *in vivo* tRNA^{Ile} molecules that have been misacylated with substrates larger than valine.

The above calculations can only be considered suggestive because the steady-state condition was not strictly adhered to in these experiments and because several additional assumptions were necessary to estimate a complete set of macroscopic parameters. Nevertheless, the results do permit comparisons and allow certain limits to be placed upon specific mechanisms.

Very limited data exist on the extent of proofreading at the aminoacyladenylate stage vs. the aminoacyl-tRNA stage. We have assumed that proofreading of the enzyme-bound aminoacyl-adenylate occurs. Yet these enzyme-bound intermediates are stable and can be isolated (Baldwin & Berg, 1966). Hence, one must also assume that addition of cognate tRNA somehow promotes a conformational change in the synthetase that is necessary for proofreading of the aminoacyl-adenylate. To test this last assumption one might be able to treat the cognate tRNA so as to destroy its amino acid accepting ability while retaining its ability to promote the conformational change presumed necessary for hydrolysis of the aminoacyl-adenylate[†]. Addition of this modified tRNA to a solution of enzyme-bound aminoacyl-adenylate would presumably promote the required conformational change and allow the observation of proofreading at the aminoacyl-adenylate stage alone. In this manner more precise estimations for the proofreading discrimination ratios might be obtained.

This work was supported in part by a grant to M.A.S. from the National Science Foundation (PCM 7807635). We thank C. Blomberg, M. Ehrenberg, C. G. Kurland and A. R. Fersht for sending us unpublished manuscripts.

REFERENCES

- BALDWIN, A. N. & BERG, P. (1966). *J. biol. Chem.* **241**, 839.
BERNARDI, F. & NINIO, J. (1978). *Biochimie* **60**, 1083.
BERNARDI, F., SAGHI, M., DORIZZI, M. & NINIO, J. (1979). *J. mol. Biol.* **129**, 93.
BLOMBERG, C. (1977). *J. theor. Biol.* **66**, 307.
BLOMBERG, C. & EHRENBERG, M. (1980). Submitted for publication.
BRUTLAG, D. & KORNBURG, A. (1972). *J. biol. Chem.* **247**, 241.

[†] Baldwin & Berg (1966) found that tRNA^{Ile} oxidized at the 3' terminus would no longer accept valine, but this modification also eliminated the ability of tRNA^{Ile} to promote hydrolysis of valyl-adenylate. However, Fersht & Dingwall (1979) found that tRNA^{Met} reduced at the 3' terminus no longer accepts homocysteine while continuing to promote hydrolysis of homocysteinyl-adenylate.

- CRICK, F. H. C. (1975). *Phil. Trans. R. Soc. Lond. B* **272**, 193.
- EHRENBERG, M. & BLOMBERG, C. (1980). Submitted for publication.
- FERSHT, A. R. (1977). *Biochemistry* **16**, 1025.
- FERSHT, A. R. (1979). In *Transfer RNA*, Part 1 (P. Schimmel, D. Söll & J. Abelson eds), Cold Spring Harbor, New York, Cold Spring Harbor Laboratory.
- FERSHT, A. R. & DINGWALL, C. (1979). *Biochemistry* **18**, 1250.
- FERSHT, A. R. & KAETHNER, M. M. (1976). *Biochemistry* **15**, 3342.
- GALAS, D. J. & BRANSCOMB, E. W. (1978). *J. mol. Biol.* **124**, 653.
- GOODMAN, M. F., GORE, W. C., MUZYCZKA, N. & BESSMAN, M. J. (1974). *J. mol. Biol.* **88**, 423.
- GORINI, L. (1971). *Nature New Biol.* **234**, 261.
- GOULIAN, M., LUCAS, Z. J. & KORNBERG, A. (1968). *J. biol. Chem.* **243**, 627.
- GROSJEAN, H., SÖLL, D. & CROTHERS, D. M. (1976). *J. mol. Biol.* **103**, 499.
- HOPFIELD, J. J. (1974). *Proc. natl. Acad. Sci. U.S.A.* **71**, 4135.
- JELENC, P. C. & KURLAND, C. G. (1979). *Proc. natl. Acad. Sci. U.S.A.* **76**, 3174.
- KURLAND, C. G. (1978). *Biophys. J.* **22**, 373.
- KURLAND, C. G. (1979). In Steenbock Symposium on Ribosomes. (In press).
- MULVEY, R. S. & FERSHT, A. R. (1977). *Biochemistry* **16**, 4731.
- NINIO, J. (1975a). In *L'évolution des Macromolécules Biologiques* (C. Sadron ed.) pp. 51-68. Paris: Editions du C.N.R.S.
- NINIO, J. (1975b). *Biochimie* **57**, 587.
- SAVAGEAU, M. A. & FRETER, R. R. (1979a). *Biochemistry* **18**, 3486.
- SAVAGEAU, M. A. & FRETER, R. R. (1979b). *Proc. natn. Acad. Sci. U.S.A.* **76**, 4507.
- SCHREIER, A. A. & SCHIMMEL, P. R. (1972). *Biochemistry* **11**, 1582.
- SPRINGGATE, C. F. & LOEB, L. A. (1975). *J. mol. Biol.* **97**, 577.
- THOMPSON, R. C. & STONE, P. J. (1977). *Proc. natl. Acad. Sci. U.S.A.* **74**, 198.
- VOLLOCH, V. Z., RITS, S. & TUMERMAN, L. (1979). *Nucleic Acids Res.* **6**, 1535.
- YAMANE, T. & HOPFIELD, J. J. (1977). *Proc. natl. Acad. Sci. U.S.A.* **74**, 2246.
- YARUS, M. (1972). *Proc. natl. Acad. Sci. U.S.A.* **69**, 1915.