INFORMATION-THEORETIC SIGNIFICANCE OF GIBBS ENERGY SUPPLY TO EDITING MECHANISMS

ALEXANDR KŘEMEN
Institute of Microbiology, Czechoslovak Academy of Sciences, Czechoslovakia

ABSTRACT If a branched multistep editing mechanism is implemented by an enzyme with a single site for the specific substrates, there is no reason to believe that the number of testing steps is fixed and cannot be controlled by some external factors. The paper considers the mechanisms of single- or multistep editing done in response to various factors, particularly to the value of displacement from the reaction equilibrium. To avoid a complicated analysis of a fully specified case, which would be likely to obscure the general significant features, the operating modes of those mechanisms are estimated using the method of minimizing associated information gains. It turns out that sufficiently far from equilibrium the variable mechanisms can essentially operate just as well as the fixed multistep mechanisms.

INTRODUCTION

It is well established that living cells synthesize important macromolecules with a high degree of accuracy (6, 9, 19, 28). Ninio (20, 21), and Hopfield (12) were the first to recognize that the high degree of discrimination between correct and wrong substrates can be achieved by special kinds of dynamic processes. This view contrasted with the attempts to explain high accuracy in macromolecule synthesis in terms of differences in equilibrium properties (binding constants of the substrates to the enzyme, or the like). Since then, a growing number of models of such mechanisms have been studied, all of them involving, in some aspect, the displacement from equilibrium of the process or energy dissipation associated with it. (Unfortunately, the terminology has not been unified yet; Ninio [21] used the term "kinetic amplification of enzyme discrimination," Hopfield [12] called it "kinetic proofreading," Fersht [8], "editing." In the present paper, the terms "editing" and "proofreading" will be used interchangeably.) Already, in the pioneering papers (12, 21), the authors have found it necessary to assume high displacement from equilibrium (provided, for example, by ATP cleavage) in some parts of their branched kinetic schemes. Kurland (18) (see also reference 3) introduced an explicit factor characterizing the displacement from equilibrium between the triphosphate and the cleavage products. Blomberg and Ehrenberg (4, 7) analyzed the constraints imposed on proofreading by energy dissipation. They studied multistep mechanisms, which can achieve much higher accuracy than a single-step mechanism. Savageau and his co-workers (10, 23–26) studied excess ATP consumption in the process and expressed their results in terms of "flows" of the substrates through the direct and discarding branches, both in single-step and in multistep mechanisms; it is meaningful to speak of flows only if the thermodynamic potentials in the branches are sufficiently strong (25).

It would seem that the significance of the displacement from equilibrium in accuracy problems has been fully recognized and is well understood. It is the aim of this paper to show that not all aspects have been dealt with.

In all of the models mentioned so far the complexity of the mechanism was assumed fixed, that is, independent of the degree of displacement from equilibrium or of other external factors (although with the multistep mechanisms the total number of steps was not specified and could be varied). In the present paper we want to show that the displacement from equilibrium can provide not only the necessary drive but also determine the complexity of the mechanism. To be more specific, we consider mechanisms testing a substrate once or twice if the process runs near equilibrium, but many times if the process is highly displaced from equilibrium. For example, such mechanisms could be represented by more complicated versions of that described recently by Hopfield (13) under the name "energy relay." His newly proposed mechanism features memory effects or "dynamic cooperativity." However, our treatment will not be based on any specific kinetic scheme and will thus allow other designs, also possibly leading to the variability in the operating modes of the error-correcting mechanisms.

The idea that editing mechanisms of this kind could exist originates from a simple consideration. With many operating units, such as enzymes or ribosomes, it seems likely that the substrate remains attached to a single site (or at most to a few sites sequentially) throughout the process, until the product is formed or the substrate is discarded; if the substrate is then to be tested many times, there is no reason to believe that the number of testing steps performed at the same site is necessarily fixed without any possibility of control by some agent. Suppose,

for instance, that both the free enzyme and the enzyme-substrate complex can exist in various different states populated with variable probabilities. If there is more than one state from which the testing procedure can start, testing sequences with different numbers of steps will be possible; we may even expect that testing sequences of different lengths will occur with different frequencies and will contribute with appropriate weights to an "average sequence." Both the weights and the length of the average sequence may depend on some external agents as well as on the kind (correct or wrong) of substrate. As already mentioned, in this paper we consider the displacement from equilibrium of the reaction as the external factor responsible for the variations, but this does not mean that other factors cannot be taken into account. For instance, the error-enhancing effect of certain chemicals is a possible candidate.

Besides the motivation described so far, there is another point of interest in the effect of displacement from equilibrium. In his study on dissipation-error tradeoff, Bennett (2) argued that considerable improvement in accuracy can be achieved with surprisingly low dissipation, of the order of 0.1 kT, in a single step. This result tempts us to ask why presumed editing systems dissipate as much energy as is observed (see also the commentary to Bennett's paper by Bremermann). It must be pointed out that in Bennett's kinetic scheme the correcting action is applied only after the product is formed. This contrasts with the branched proofreading schemes cited above; in these schemes, all branching points precede the product-forming reaction and the selection is supposed to proceed with the substrate attached to the same enzyme molecule. Yet the question cannot be explained away by the difference in the kinetic schemes, and we have good reasons to seek an answer. Certainly, one part of the answer is that higher displacements from equilibrium generally support higher rates of synthesis. Another part of the answer is that with the fixed multistep mechanisms the overall accuracy can depend on the value of the displacement (4, 7). The model and the effect of the displacement discussed in the present paper suggest another plausible answer to that question.

From the foregoing it will be clear that our main interest is in the effect the displacement can exert in variable mechanisms, not in an analysis of some particular scheme describing them. The recent note on mechanisms with memory should only make the discussion more suggestive. Although this approach prevents us from obtaining analytical relations between the value of the displacement and the degree of accuracy, it allows us to estimate the effect of the displacement on the operating modes of the variable mechanisms using very simple information-theoretic arguments. The estimates are based on elementary properties common to all branched mechanisms, on the required output, and on an assumption related to the variability of the mechanism. The result is thus sufficiently general; any additional knowledge or requirement concerning details of the kinetics would narrow the validity of the picture, but would also allow one to appreciate which features of the more fully specified model would result from the additional constraints and which from the variability alone. It is therefore useful to study the general case.

In the next section we express the properties of the variable mechanisms, including the required output, and explain the assumption concerning the variability, all in terms of probabilities. Next we describe the estimation method, and in the last section we discuss the results, also presenting numerical examples.

THE MODEL

We consider a process in which the correct substrates (S_c) should be transformed to a product (or incorporated into a product) while the incorrect or wrong substrates (S_w) should be prevented from product formation; one of the most distinguishing features of the process should be very low failure rates in identifying the substrates correctly. The discrimination is performed by an operating unit (for example, by an enzyme or by a ribosome) both in the binding step with the substrate and during the period within which the so-called enzyme-substrate complex exists (we use the term "enzyme-substrate complex" or simply "complex" even if the operating unit is a ribosome or the like). In this paper, we do not consider the binding step and deal only with the activity of the complex. It begins at the moment of formation of the complex and ends either in product formation or in discarding the substrate. "Unspecific" substrates and products can also participate in the reaction. If this is the case, unspecific products can be formed even if the specific substrate (S_c or S_w) is discarded. Usually, the unspecific substrate is a nucleoside triphosphate (e.g., ATP, GTP).

The sequence of kinetic events that may occur in the complex can be described (for our purposes) by a branched scheme of reactions, with at least one discarding branch. The points of this scheme where a substrate can be discarded are called "nodes," each sequence of reactions between adjacent nodes represents one step. At each node the substrate moiety is tested whether it is S_c or S_w; if it is identified as incorrect, it is discarded, otherwise it is passed to the next node or to product formation. Discarding a substrate means decomposition of the complex; the components need not be recovered exactly in the physical states they were in at the beginning of the binding reaction. The variability of the mechanism arises from the circumstance that, in general, the sequence of testing steps may start at any one of the nodes, and also the product-forming reaction may originate at any node not preceding the starting one. If f_i is the probability that the testing sequence begins at the ith node, the reduced product formation flow can be expressed as

$$P = \sum_i f_i P_i \qquad (1)$$

where the summation runs over all nodes. The P_i characterize the product formation associated with the ith node; their form will be discussed in the next paragraph. Actually, we have two sets of the quantities defined so far, one set for the correct substrates and another one for the wrong substrates. If we want to express this fact explicitly, we write the subscripts c or w, respectively. All quantities in Eq. 1 are understood to have subscripts of c or w.

If the properties of the nodes are known, the P_i can be expressed in terms of conditional probabilities characterizing the nodes. Consider, for example, the scheme in Fig. 1. The testing sequence may start with probabilities f_i at any one of the four nodes in the scheme (it might contain any positive number of nodes). Let $p_i q_i$ be the conditional probability of net product formation from the ith node, and let $(1 - p_i) \cdot q_i$ be the conditional probability of passing the complex to the next node. In Fig. 1,

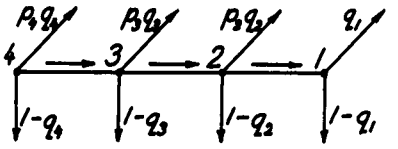

FIGURE 1 Scheme of a branched mechanism. The arrows indicate the direction of flows; the input flows to nodes 1–4 are not shown. p_iq_i is the conditional probability of product formation from the ith node ($p_1 = 1$); $1 - q_i$ is the conditional probability of discarding the substrate from the ith node.

this is the nearest right-hand side node, the $(i - 1)$th. The last node that can be reached at all is the node number 1, with $p_1 = 1$. This way of numbering the nodes has the advantage that we need not renumber them if we add some new nodes to the scheme. Both conditional probabilities introduced so far sum up to q_i, so that $1 - q_i$ is the conditional probability of discarding the substrate from the ith node. Because the direction of flows between the nodes is from left to right (all flows are understood as net flows), the P_i depend on these conditional probabilities as shown by

$$P_i = q_i \cdot [p_i + (1 - p_i) \cdot P_{i-1}], \; i = 2, 3, 4$$
$$P_1 = q_1. \tag{2}$$

Each P_i thus represents that fraction of the flow starting from the ith node (that means not coming from the preceding ones) which is consumed to product formation both directly from the starting node and, with appropriate weights, from all of the subsequent nodes. Note that Eq. 2 reduces to the form $P_i = q_i \cdot q_{i-1} \ldots q_1$, if the product can be formed from the last node only (in that case, $p_2 = p_3 = p_4 = 0$).

The equations in 2 apply to the usual situation when the reactions in the product-forming branches and the discarding branches run as shown by the arrows in Fig. 1. Near equilibrium it may happen, at least in model systems, that these reactions run in the reverse directions (there is then no editing, of course). If we wanted to apply the formalism described in the later sections to such anomalous situations as well, we should have to redefine the meaning of the quantities q_i, p_i, and P_i. Although the formalism could still be quite correct, it would make no sense to apply it, simply because there would be no error correction, therefore also no need to estimate the operating modes of the proofreading mechanism. It will become clear later that the variable mechanisms "degenerate" to one-step mechanisms when operating near equilibrium; if necessary, then, their operating modes can be analyzed directly and more fully by standard kinetic methods.

The probabilities f_i do not appear in Eq. 2. This conforms with an important assumption concerning the variability of the model: These probabilities are not directly related to the properties of the nodes, instead they depend primarily on some properties of the complex as well as on some external factors. For example, suppose we have a mixture of free enzyme molecules in different states so that the complexes formed from them commence their testing procedures at different nodes. Then the probabilities f_i reflect the population fractions of these states, the Gibbs energy differences between them, and the total displacement from equilibrium.

To summarize, the model is generally described by the two sets (P_{ci}), P_{wi}), and by the two probability distributions $F_c = (f_{ci})$, $F_w = (f_{wi})$. These distributions are not uniquely determined by the mechanism itself and may vary with some external factors. With properly designed mechanisms, one of these factors can be the displacement from equilibrium. This is the main point of our interest. The output of the mechanism is described by the reduced product formation flows P_c and P_w defined by Eq. 1.

In the next section, we will study the operation of variable mechanisms of this kind, assuming only the features outlined in the preceding paragraph. No particular dependence of the distributions F on external factors will be specified. We shall consider the sets (P_i) as given, the values of P_c and P_w as required or known, and shall estimate the "best" distributions F_c and F_w. Only in the later section we shall discuss on a general level the relation between the probability distributions and the displacement from equilibrium.

OPERATION OF THE MODEL

If we are given the set (P_i) and the mean P, and if the problem involves a single probability distribution (p_i) (not to be confused with the p_i in Eq. 2), the "best" distribution is that which reflects all our knowledge about the problem (i.e., (P_i) and P), but nothing more. It is called the least prejudiced distribution, because any other distribution gives (perhaps inadvertently) preference to some p_i and thus expresses some piece of additional knowledge, that is, in fact, lacking. An instructive comparison of distributions found by intuitive methods with the least prejudiced distribution in a simple problem with tossing a die is presented by Brostow (5). Jaynes's principle (14) tells that to find the least prejudiced distribution we must maximize the function

$$H = -\sum_i p_i \cdot \log p_i \tag{3}$$

(logarithms are understood to the base e throughout) subject to the constraints that $P = \Sigma_i p_i P_i$ has a prescribed value and $\Sigma_i p_i = 1$. The maximization procedure involves the standard method of Lagrange multipliers. If there are n components in the distribution (p_i) and if the constraint of given P is missing, the maximization procedure yields the absolute maximum of H obtained with

$$p_i = 1/n \quad i = 1, 2, \ldots, n. \tag{4}$$

This property makes the function H as defined by Eq. 3 inapplicable to our problem. In fact, if we only know that the problem involves a distribution $G = (g_i)$ of sequence lengths (the sequence length and the number of nodes in the testing sequence are equivalent), there is no reason to assume that G is a uniform distribution. In a uniform distribution we would have $g_i = 1/N$, where N is some reasonably chosen upper limit of admissible sequence lengths. Instead, without any knowledge indicating that the contrary is true, we expect that, for physical reasons, very long testing sequences are less probable than shorter ones. Equilibrium systems will feature extreme behavior in this respect. In equilibrium, there will be no flows at the input and through the chain of nodes. If some sufficiently large and correctly oriented fluctuation brings the system out of equilibrium for a very short time, the system will hardly perform more than one step; that is, the probability of a testing sequence with a single node will be close to unity, all probabilities of longer sequences summing up to a value much less then unity. This reasoning allows us to realize that elementary knowledge about the physical nature of our mechanism compels us to assume a highly nonuniform equilibrium distribution of sequence lengths; we shall denote it as the prior distribution G. Our conclusion must be changed, however, if we obtain some additional evidence that the actual distribution differs from G; we denote the actual distribution as F. A suitable item of evidence against G may be the value of the actual product formation flow P, if it differs from the value calculated using G. In that case we must estimate the distribution F using not only the original knowledge (yielding the conclusion that G should be the correct distribution) but also all the new item(s) of evidence we have gained. It is another task to discover the physical mechanisms causing the actual distribution to differ from the prior one. There can be different mechanisms leading to the same F with a given G; it may require different (often quite complicated) approaches to study the physical causes. Here we will ignore the underlying physical differences and concentrate our effort on the estimation procedure. The procedure is quite general and therefore may be applied to all cases. The displacement from equilibrium will be discussed only qualitatively as a possible explanation of $F \neq G$.

For the estimation we must therefore use a modified form of Jaynes's principle and look for an information measure that would be the absolute extremum at a nonuniform distribution G (the prior distribution). Infor-

mation theory offers such a function; it has the form

$$I(F \parallel G) = \sum_i f_i \cdot \log f_i/g_i \qquad (5)$$

where G is the prior distribution and F is the unknown distribution to be determined. If in Eq. 5 some $g_k = 0$, f_k must be zero as well, to avoid singularities. Also, it is understood that $0 \cdot \log 0 = 0$. The function I is always nonnegative and vanishes if and only if $F = G$. It can be found under different names in current literature; for our purposes perhaps the best is "information gain" (22). The use of I was discussed quite a few times (see for instance Rényi [22], Kullback [17], chapter 1 in Watanabe [29], or Aczél and Daróczy [1]). The related modification of Jaynes's principle is rather straightforward, and has been analyzed in particular by Hobson (11). Shore and Johnson (27) proved that minimizing I provides a general method of inference about an unknown probability distribution, if the prior G is given and the unknown distribution F satisfies some constraints on mean values. If there is only one mean value as a constraint of the distribution F, the minimizing procedure yields F in the form

$$f_i = g_i \cdot \exp(-aP_i)/\sum_k g_k \cdot \exp(-aP_k),\ i = 1, 2, \ldots \qquad (6)$$

where a is a Lagrange multiplier. The constraints used in deriving Eq. 6 are $\Sigma_i f_i = 1$ and a given value of the mean $P = \Sigma_i f_i P_i$. We also obtain

$$dP/da = \sum_i P_i \frac{df_i}{da} = P^2 - \sum_i f_i P_i^2 \qquad (7)$$

that is the square of the standard deviation. The meaning of the Lagrange multiplier a is suggested by Eq. 7; it may become clearer if we note that a is an analogue of $1/kT$ in statistical thermodynamics (k is the Boltzmann constant, T is the thermodynamic temperature).

Our problem and its solution can now be formulated as follows. We consider a variable multistep editing mechanism represented by a branched scheme. The observed output of this mechanism are the values of the reduced product formation flows P_c and P_w. We also know the sets (P_{ci}) and (P_{wi}) characterizing the product formation flows associated with the individual nodes. We assume that these sets depend only on the properties of the nodes (this assumption is discussed a little later); in contrast, the reduced flows P_c and P_w depend also on some external factors. Variations in these factors change the relative contributions of the individual nodes expressed as the two probability distributions F_c and F_w. Because the observables P_c and P_w are defined by Eq. 1, changes in any of the distributions F are reflected in the corresponding P. For simplicity, we will consider a single factor. Let us express its changes on an appropriate scale (for instance, that of the displacement from equilibrium). We then assume that the physical nature of the entire system allows us to choose one reference point on that scale such that the probability distributions at that point are known (they could be measured, for example). Let us denote them as the reference or prior distributions G_c and G_w. If the system is not in the reference state, our task is to find those distributions that reflect all our knowledge about the system, that is G_c, G_w, (P_{ci}), (P_{wi}), P_c, P_w, but nothing more. The least prejudiced distributions to be found, F_c^+ and F_w^+, are given by Eq. 6, for the correct and wrong substrates, respectively.

We derived these relations by minimizing I_c and I_w defined in Eq. 5 using the method of Lagrange multipliers. This is a simple method, but to be consistent we must assume that the sets (P_{ci}) and (P_{wi}) are fixed and only the distributions F are varied. This can be understood in two ways. First, we may consider a single couple of fixed sets (P_i) for all situations, that is, for any F_c^+ or F_w^+. This condition particularly means that the sets (P_i) do not depend on the factor controlling the distributions F^+. The variability of the model is then due to the distributions F^+ alone. Alternatively, we may assume that each particular situation has its own couple of the sets (P_i); the distributions F^+ then refer to intervals small enough for the sets (P_i) to remain constant. The resulting variability of the model reflects both the variability of the distributions F^+ and the variability of the sets (P_i). It should be stressed that the requirement of fixed sets (P_i) is imposed only by the extremizing procedure chosen merely for its simplicity. In no case is that requirement an inherent feature of the physical nature of the model. Near equilibrium, for instance, the sets (P_i) will vary with the displacement from equilibrium and so will the distributions F^+. The model is applicable to these situations, but the extremizing procedure may not be, if the sets (P_i) cannot be regarded as fixed even in an arbitrarily small interval.

Having found F_c^+ and F_w^+, we can calculate the information gains I_c and I_w using Eq. 5. They are, respectively, functions of a_c and a_w, by virtue of Eq. 6. Also, because the reduced flows P can be expressed as functions of the Lagrange multipliers a, (Eq. 1), we can eliminate the as and obtain relations between I_c and P_c and between I_w and P_w.

Our estimates thus associate with each value of the observed production flow P a least prejudiced distribution F^+ and a value of the information gain I. The distribution F^+ characterizes the relative contributions of the nodes to the flow P; the values of the information gains I then indicate how much the actual distributions F^+ differ from the related prior distributions G. We may say that the information gains measure the change in the complexity of the mechanism between the actual and the reference states.

DISCUSSION

Eq. 2 reveals that increasing, decreasing, or even irregular sequences $P_1, P_2, \ldots$ can be formed. It is then easy to design mechanisms favoring very low wrong product formation while the correct product formation remains high. Consider the scheme in Fig. 1 again. If for the wrong substrate $0 < P_4 < P_3 < P_2 < P_1$ with $P_3 + P_4 << P_1$, and if F^+ is such that $f_3 + f_4$ is near unity (f_1 and f_2 may be neglected), then $P_w \cong f_3P_3 + f_4P_4 << 1$. Of course, P_w cannot be lower than P_4, in this example. Now let the prior distribution G_w have g_1 nearly equal to unity, so that $g_2 + g_3 + g_4 << 1$. Then the information gain I_w (see Eq. 5) achieves a high value.

The following example has been calculated using the values

$$P_1 = 0.1 \quad P_2 = 0.019$$

$$g_1 = 0.99 \quad g_2 = 9 \times 10^{-3}$$

$$P_3 = 2.881 \times 10^{-3} \quad P_4 \cong 3.88 \times 10^{-4}$$

$$g_3 = 9 \times 10^{-4} \quad g_4 = 1 \times 10^{-4}.$$

Then for $P \cong 8.3 \times 10^{-4}$ the least prejudiced distribution is

$$f_1^+ \cong 1.05 \times 10^{-61} \quad f_2^+ \cong 5.56 \times 10^{-11}$$

$$f_3^+ \cong 0.176 \quad f_4^+ \cong 0.824$$

and $I = 8.35774$. The set (P_i) used in this calculation can be obtained using, for instance, the values (see Eq. 2)

$$q_i = 0.1 \quad (\text{all } i)$$

$$p_1 = 1 \quad p_2 = 0.1 \quad p_3 = 10^{-2} \quad p_4 = 10^{-3}.$$

With this choice, the conditional probability of discarding the wrong substrates has the same (rather poor) value $1 - q_i = 0.9$ at all nodes, whereas the total product formation is kept low by starting the testing sequences virtually only at the fourth and, to a lesser extent, at the third node.

A mechanism operating in this way could feature four

states (generally, any number of states higher than one) with sufficiently large Gibbs energy differences between them that near equilibrium virtually only the lowest state would be populated ($g_1 \lesssim 1$). Near equilibrium, the scheme of this mechanism consists of a single node, if we neglect the higher states ($g_2 + g_3 + g_4 << 1$). Far from equilibrium, the mechanism will operate in the higher states, if part of the Gibbs energy (supplied by, say, ATP or GTP cleavage) is used to populate the higher states of the enzyme-substrate complex, and, possibly, if on discarding a substrate the free enzyme remains in one of its higher states long enough to bind with another substrate molecule.

Processing of the correct substrates can now be explained more briefly. Let for the correct substrates, $P_i \cong P_1$ possibly with $P_4 > P_3 > P_2 > P_1$. If P_1 is sufficiently near unity, the differences between the P_i cannot be very large. Then even with a prior distribution G_c such that $g_1 \lesssim 1$, $g_2 + g_3 + g_4 << 1$, the product formation P_c will be high enough with any F_c^+; it may even increase as F_c^+ becomes more different from G_c. Let us consider an example again. With

$$q_1 = 0.9 \quad q_2 = 0.92 \quad q_3 = 0.95 \quad q_4 = 0.97$$

$$p_1 = 1 \quad p_2 = 0.90 \quad p_3 = 0.92 \quad p_4 = 0.95$$

we obtain from Eq. 2 the values

$$P_1 = 0.9 \quad P_2 = 0.9108 \quad P_3 \simeq 0.943 \quad P_4 \simeq 0.967.$$

Then for $P_c = 0.96468$ the least prejudiced distribution is

$$f_1 \simeq 1.31 \times 10^{-2} \quad f_2 \simeq 1.03 \times 10^{-3}$$

$$f_3 \simeq 6.77 \times 10^{-2} \quad f_4 \simeq 0.918$$

and $I_c = 8.61186$ (these figures are obtained with the same prior distribution as in the example for the wrong substrates).

It is clear that with mechanisms of this kind the fraction of the discarded correct substrates can be kept reasonably low in any operating mode. The fraction of the wrong substrates mistakenly processed to product can be diminished sufficiently only if a higher number of nodes is activated. Both kinds of errors, however, remain finite unless the number of nodes becomes infinite; only in that case could the errors be avoided completely. Sufficiently far from equilibrium, then, the variable mechanisms operate essentially equally as well as the fixed multistep mechanisms analyzed by other authors.

There is no universal relation between the supplied amount of Gibbs energy and the values of the information gains I_c and I_w. Of course, we always have I_c and $I_w = 0$ at equilibrium, because of $F^+ = G$; also, both information gains will grow as F^+ become more different from G, that is, as the supplied amount of Gibbs energy increases. The actual forms of those relations will be determined by details concerning the states of the free enzyme and of the enzyme-substrate complex. It is the advantage of the approach adopted in this paper that we can avoid the much more complicated analysis of such fully specified systems and yet obtain satisfactory insight into the operation modes of the mechanisms. In addition, although the physical nature of the mechanisms is specified only qualitatively, it allows the significance of the Gibbs energy supply and its connection with the changes in the distributions and in the associated information gains to appreciate.

It might be questioned whether it is permissible to consider relations between quantities seemingly as different as the supplied amounts of Gibbs energy on the one hand and information gains on the other. This question can be answered positively. In general, the idea that Gibbs energy of a nucleotide pool can be expressed in the form of an information gain has been suggested and used elsewhere (15, 16). In the present problem, the supplied amounts of Gibbs energy are differences of chemical potentials, or affinities. At equilibrium, these differences vanish. The nonequilibrium affinities can be therefore written in the form of differences of one-component information gains. For instance, considering ATP cleavage to AMP and inorganic pyrophosphate we may write

$$I_T = \frac{\mu_T - \mu_M - \mu_{PP}}{kT} = \log \frac{n_T}{n_T^{(o)}} - \log \frac{n_M}{n_M^{(o)}} - \log \frac{n_P}{n_P^{(o)}} \tag{8}$$

where μ_i are the chemical potentials ($i = T$ for ATP, M for AMP, P for inorganic pyrophosphate, n_i are the actual molar fractions, and $n_i^{(o)}$ are the equilibrium molar fractions. The ratio

$$K^{(o)} = n_T^{(o)}/(n_M^{(o)} \cdot n_P^{(o)})$$

appearing in the right-hand side of Eq. 8 is the equilibrium constant of the cleavage reaction. The expressions

$$\log \frac{n_i}{n_i^{(o)}}$$

in Eq. 8 are one-component information gains (22), so that I_T is also an information gain. Because the affinity $\mu_T - \mu_M - \mu_P$ is usually highly positive, so is I_T. The substitution of an information gain for the respective affinity is therefore quite correct, and we may put high I_w and I_c in connection with high I_T (possibly containing additional terms due to the specific substrates and products).

I thank Dr. Clas Blomberg of the Royal Institute of Technology, Stockholm, for useful discussions on accuracy problems.

Received for publication 16 June 1981 and in revised form 3 May 1982.

REFERENCES

1. Aczél, J., and Z. Daróczy. 1975. On measures of information and their characterizations. Academic Press, Inc., New York. 199–208.

2. Bennett, C. H. 1979. Dissipation-error tradeoff in proofreading. *Biosystems.* 11:85–91.
3. Blomberg, C., M. Ehrenberg, and C. G. Kurland. 1980. Free-energy dissipation constraints on the accuracy of enzymic selections. *Q. Rev. Biophys.* 13:231–254.
4. Blomberg, C., and M. Ehrenberg. 1981. Energy considerations for kinetic proofreading in biosynthesis. *J. Theor. Biol.* 88:631–670.
5. Brostow, W. 1979. Science of Materials. John Wiley & Sons, Inc., New York. 19–24.
6. Edelman, P., and J. Gallant. 1977. Mistranslation in *E. coli. Cell.* 10:131–137.
7. Ehrenberg, M., and C. Blomberg. 1980. Thermodynamic constraints on kinetic proofreading in biosynthetic pathways. *Biophys. J.* 31:333–358.
8. Fersht, A. R. 1980. Enzymic editing mechanisms in protein synthesis and DNA replication. *Trends Biochem. Sci.* 5:262.
9. Fowler, R. G., G. E. Degnen, and E. C. Cox. 1974. Mutational specifity of a conditional *Escherichia coli* mutator, mutD5. *Mol. Gen. Genet.* 133:179–191.
10. Freter, R. R., and M. A. Savageau. 1980. Proofreading systems of multiple stages for improved accuracy of biological discrimination. *J. Theor. Biol.* 85:99–123.
11. Hobson, A. 1969. A new theorem of information theory. *J. Statist. Phys.* 1:383–391.
12. Hopfield, J. J. 1974. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. U. S. A.* 71:4135–4139.
13. Hopfield, J. J. 1980. The energy relay: a proofreading scheme based on dynamic cooperativity and lacking all characteristic symptomes of kinetic proofreading in DNA replication and protein synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 77:5248–5252.
14. Jaynes, E. T. 1957. Information theory and statistical mechanics. *Phys. Rev.* 106:620–630.
15. Křemen, A. 1980. Coding of Gibbs function flows from nucleotide pools. *J. Statist. Phys.* 23:483–494.
16. Křemen, A. 1982. Information aspects of Gibbs function output from nucleotide pools and of the adenylate kinase reaction. *J. Theor. Biol.* 96:425–441.
17. Kullback, S. 1959. Information theory and statistics. John Wiley & Sons, Inc., New York. 3–43.
18. Kurland, C. G. 1978. The role of guanine nucleotides in protein biosynthesis. *Biophys. J.* 22:373–392.
19. Loftfield, R. B., and D. Vanderjagt. 1972. The frequency of errors in protein biosynthesis. *Biochem. J.* 128:1353–1356.
20. Ninio, J. 1974. A semi-quantitative treatment of missense and nonsense suppression in the strA and ram ribosomal mutants of *Escherichia coli.* Evaluation of some molecular parameters of translation in vivo. *J. Mol. Biol.* 84:297–313.
21. Ninio, J. 1975. Kinetic amplification of enzyme discrimination. *Biochimie.* 57:587–595.
22. Rényi, A. 1970. Probability theory. Académiai Kiadó, Budapest. 560–563, 569–574.
23. Savageau, M. A., and R. R. Freter. 1979. Energy cost of proofreading to increase fidelity of transfer ribonucleic acid aminoacylation. *Biochemistry.* 18:3486–3493.
24. Savageau, M. A., and R. R. Freter. 1979. On the evolution of accuracy and cost of proofreading tRNA aminoacylation. *Proc. Natl. Acad. Sci. U. S. A.* 76:4507–4510.
25. Savageau, M. A. 1981. Accuracy of proofreading with zero energy cost. *J. Theor. Biol.* 93: 179–196.
26. Savageau, M. A., and D. S. Lapointe. 1981. Optimization of kinetic proofreading: a general method for derivation of the constraint relations and an exploration of a specific case. *J. Theor. Biol.* 93: 157–178.
27. Shore, J. E., and R. W. Johnson. 1980. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory.* IT–26:26–37.
28. Springgate, C. F., and L. A. Loeb. 1975. On the fidelity of transcription by *Escherichia coli* ribonucleic acid polymerase. *J. Mol. Biol.* 97:577–591.
29. Watanabe, S. 1969. Knowing and Guessing. A quantitative study of inference and information. John Wiley & Sons, Inc., New York. 380–388.