The Catalytic Power of Pyruvate Decarboxylase. A Stochastic Model for the Molecular Evolution of Enzymes#

David W. Huhta, Thomas Heckenthaler, Francisco J. Alvarez, Joachim Ermer, Gerhard Hübner, Alfred Schellenberger and Richard L. Schowen

Departments of Chemistry and Biochemistry and Higuchi Biosciences Center, University of Kansas, Lawrence, Kansas 66045-0046 USA and the Institute for Biochemistry, Martin Luther University of Halle-Wittenberg, Weinbergweg 16a, O-4050 Halle/Saale, Germany

Huhta, D. W., Heckenthaler, T., Alvarez, F. J., Ermer, J., Hübner, G., Schellenberger, A. and Schowen, R. L., 1992. The Catalytic Power of Pyruvate Decarboxylase. A Stochastic Model for the Molecular Evolution of Enzymes. – Acta Chem. Scand. 46: 778–788.

Pyruvate decarboxylase (PDC) catalyzes the decarboxylation of pyruvate anion by a factor of around 10^{12}, compared with the non-enzymic decarboxylation by thiamine, under standard state conditions of 1 mM pyruvate and thiamine diphosphate (TDP), pH 6.2. Free-energy diagrams constructed on the basis of earlier measurements for the enzymic and non-enzymic reactions give some information on catalysis by PDC. PDC stabilizes the reactant state preceding TDP addition to pyruvate by 76 kJ mol^{-1} and the transition state for the addition by 83 kJ mol^{-1}. PDC stabilizes the reactant state preceding decarboxylation (presumably α-lactyl-TDP) by 27 kJ mol^{-1} and the transition state for decarboxylation by 68 kJ mol^{-1}. In addition, the free-energy diagrams reveal a leveling of reactant-state free energies in the enzymic reaction compared with the non-enzymic reaction, in that the former are nearly equal to each other. The enzyme-bound transition-state energies are similarly leveled. The energetic leveling of reactant states has been noted by Albery, Knowles and their coworkers in many enzymic reactions and termed 'matched internal thermodynamics.' They showed that the result would arise naturally (and inevitably) in the 'evolution to perfection' of enzymes, when the evolutionary process was treated by a deterministic model. The critical assumption of this model was the validity of a Marcus-type or Brønsted-type linear free-energy relationship between rate and equilibrium constants for reactions occurring wholly within enzyme complexes. Here a completely stochastic simulation of molecular evolution, with no deterministic assumptions, is shown to reproduce both 'matched internal thermodynamics' and the 'matched internal kinetics' or leveling of transition-state energies noted here. The Albery–Knowles result is thus more general than might have been supposed.

Enzyme catalysis and free-energy diagrams

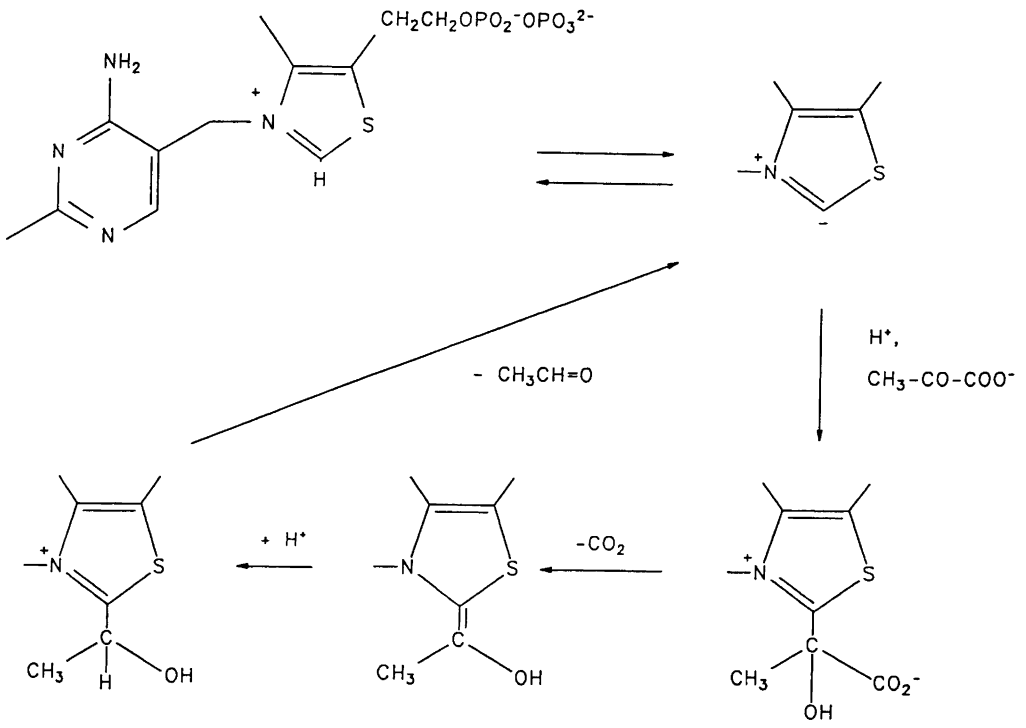
The essence of enzymic catalysis is the selective stabilization of transition states with minimal stabilization of reactant states.[1,2] To appreciate the origins of the catalytic power of a particular enzyme, a useful procedure is the construction of and comparison of free-energy diagrams for the enzymic reaction and a closely related non-enzymic reaction; the free-energy difference for individual states is then graphically apparent. Furthermore, the comparison of such free-energy diagrams should be revealing about molecular evolution. The non-enzymic reaction constitutes a model for a very primitive stage of molecular evolution in which an enzyme produces very little catalytic acceleration. The enzymic diagram then shows the changes that have come about in the course of molecular evolution.

Pyruvate decarboxylation by thiamine and by pyruvate decarboxylase

Information is available for this procedure for the decarboxylation of pyruvate anion catalyzed by the thiamine-diphosphate (TDP)-dependent enzyme yeast pyruvate decarboxylase (PDC).[3-7] The essential chemistry is depicted in Scheme 1, taken from Ref. 7. Thiamine itself will promote the decarboxylation of pyruvate, but under conditions of 1 mM pyruvate and pH 6.2, 25 °C, the reaction has a first-order rate constant of only 2×10^{-11} s^{-1}. Under similar conditions, the first-order rate constant for PDC action is around 60 s^{-1}, so that the enzymic acceleration is a factor larger than 10^{12}. The average net transition-state stabilization by PDC is thus around 69 kJ mol^{-1}.

In this paper, we consider both the specific origins of this catalytic power for PDC, and some general features of the free-energy barriers for PDC catalysis that are seen in other reactions as well. Finally, we propose a stochastic model of molecular evolution to account for the general features of enzyme catalysis.

Presented as a plenary lecture at the third Symposium on Organic Reactivity in Göteborg, Sweden, July 7–12, 1991.



Scheme 1. Decarboxylation of pyruvic acid by thiamine diphosphate.

Non-enzymic decarboxylation of pyruvate by thiamine

The non-enzymic, thiamine-promoted decarboxylation of pyruvate anion is understood from studies of Kluger, Chin and Smyth[3] and Washabaugh and Jencks.[4-6] In the following discussion, we will assume that the relative free energies of species along the reaction paths for both the non-enzymic and the enzymic reaction paths will be independent of whether the thiamine is phosphorylated. Thus we can compare the thiamine-promoted non-enzymic reaction with the TDP-dependent enzymic reaction. Washabaugh and Jencks determined the rate constants for generation of the thiamine C_2-anion from the reactant thiamine, while Kluger, Chin and Smyth synthesized the adduct of pyruvate with thiamine (a reasonable model for the pyruvate-TDP adduct) and studied the kinetics of its formation, reversion to pyruvate and thiamine, and its decarboxylation. The results of these studies permit a free-energy diagram to be constructed for the non-enzymic reaction stages through decarboxylation. The rate constants and free energies for the non-enzymic reaction are given in the first part of Table 1.

Enzymic decarboxylation of pyruvate anion

Alvarez *et al.*[7] deduced the microscopic rate constants for the corresponding PDC-catalyzed reactions from ^{13}C and β-^2H isotope effects. PDC exhibits a complex mechanism of hysteretic regulation, in which a molecule of the substrate, pyruvate anion or S, first binds reversibly into a regulatory site of the enzyme E to give a complex SE. SE then undergoes a unimolecular activating transition, or 'switching' to generate the activated enzyme SE*. SE* then binds pyruvate into its catalytic site and decarboxylates an average of 10000 molecules of pyruvate (at full saturation) before an 'off-switching' transition occurs. Decarboxylation and product-release events occur with about equal rate con-

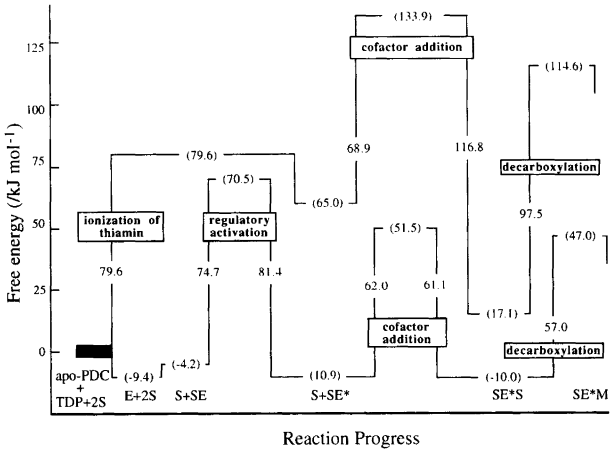


Fig. 1. Free-energy diagrams for the non-enzymic decarboxylation of pyruvate by thiamine (upper curve) and for the pyruvate-decarboxylase (PDC) catalyzed decarboxylation of pyruvate with thiamine diphosphate (TDP) as cofactor (lower curve). Both curves are for standard-state concentrations of 1 mM for pyruvate and TDP, pH 6.2 and 298 K. The rate constants are given in Table 1.

Table 1. Rate constants, standard-state[a] free energies of activation, free energies of reactant and transition states relative to the apoenzyme assembly, and free energies of reactant and transition state stabilization for the enzymic and non-enzymic decarboxylation of pyruvate at 298 K and pH 6.2.

Observed rate constants	Standard-state rate constants/s^{-1}	$\Delta G^{\ddagger}$/kJ mol^{-1}	G_{rel}/kJ mol^{-1} [b]	ΔG_{stab}/kJ mol^{-1} [c]
Non-enzymic reactions				
Deprotonation of TDP[d] 4.3×10^6 M^{-1} s^{-1}	6.8×10^{-2}	79.6		
Transition state for deprotonation			79.6	
Protonation of anion[e] 1.7×10^{10} s^{-1}	1.7×10^{10}	14.6		
Thiamine anion			65.0	
Specific-base catalyzed addition of thiamine to pyruvate 1.3 M^{-2} s^{-1}	2.1×10^{-11}	133.9		
Transition state for addition			133.9	
Specific-base catalyzed reversion of adduct 1.3 M^{-1} s^{-1}	2.1×10^{-8}	116.8		
α-Lactylthiamine			17.1	
Decarboxylation of α-lactylthiamine 5.0×10^{-5} s^{-1}	5.0×10^{-5}	97.5		
Transition state for decarboxylation			114.6	
Enzymic reaction				
Binding of TDP (K_d = 23 μM)				
E:TDP assembly			−9.4	
Reversible binding of pyruvate into the regulatory site (K_d = 8 mM)				
SE			−4.2	
'On-switching' or activation of SE to SE* 0.49 s^{-1}	0.49	74.7		
Transition state for activation			70.5	
'Off-switching' or deactivation from SE* to SE 0.033 s^{-1}	0.033	81.4		
Activated enzyme, SE*			−10.9	75.9
Overall binding of pyruvate into the catalytic site 8.2×10^4 M^{-1} s^{-1}	82	62.0		
Transition state for binding[f]			51.1	82.8
Overall release of pyruvate from the catalytic site 120 s^{-1}	120	61.1		
E:α-lactyl-TDP assembly			−10.0	27.1
Decarboxylation 640 s^{-1}	640	57.0		
Transition state for decarboxylation			47.0	67.6

[a]The standard-state concentrations of pyruvate and TDP are taken as 1 mM. [b]Gibbs free energy relative to apoenzyme + TDP + 2 pyruvate; differences in non-enzymic reactivity between thiamine and TDP are taken as negligible. [c]Gibbs free energy of stabilization of each state by binding to the enzyme. [d]In 1 mM citrate buffers, the fastest route of deprotonation at pH 6.2 is by reaction with hydroxide with the indicated rate constant. [e]Reprotonation of anion by reaction with water. [f]The rate constant is taken to reflect overall binding by pyruvate into the catalytic site and addition of the anion to pyruvate to give α-lactyl TDP in the catalytic site.

stants. Table 1 gives the numerical values for the PDC rate constants under standard-state conditions, along with the corresponding free energies.

Fig. 1 portrays the two standard-state free-energy diagrams for the non-enzymic and enzymic reactions.

Transition-state stabilization and reactant-state stabilization by pyruvate decarboxylase

Thiamine ionization. Kluger, Chin and Smyth[3] showed that thiamine addition to pyruvate is specific-base catalyzed, requiring the conjugate base of thiamine to be formed before the onset of addition to the pyruvate carbonyl group. The molecular assembly at $G = 65.0$ kJ mol^{-1} on the non-enzymic profile of Fig. 1 therefore consists of two pyruvate molecules, apo-PDC, and the conjugate base of TDP (TDP-cb). The assembly at -10.9 kJ mol^{-1} on the enzymic profile consists of activated enzyme SE*, with TDP in the active site, and one molecule of free pyruvate. Thus reaction (1) has $\Delta G^\circ = -65.0 - 10.9 = -75.9$ kJ

$$S + \text{apo-PDC} + \text{TDP-cb} = SE^* \qquad (1)$$

mol^{-1} or an equilibrium ratio of 10^{13} for [SE*]/[apo-PDC] at 1 mM standard state levels of S and TDP-cb. Note that reaction (1) is a thermodynamic proposition, *not* a mechanistic suggestion.

It is not known what ionization state the thiamine moiety of TDP is in when the enzyme is in the activated form SE*, although one hypothesis is that the regulatory activation generates the conjugate base of TDP. If TDP is in fact in the form of its conjugate base in SE*, then enzyme has stabilized the conjugate base by about 76 kJ mol^{-1}, in effect altering the pK_a from around 18 in free solution[4] to around 5 in the PDC catalytic site.

Cofactor addition. The barrier succeeding the state SE* is labeled 'cofactor addition' on the enzymic profile. This corresponds to the supposition that the stable state beyond the barrier, SE*S, contains the adduct α-lactyl-TDP. It is not certain that this supposition is correct, and it is clear that the barrier succeeding SE* describes several successive events. At a minimum, the entry of pyruvate into the catalytic site must be involved. For our discussion in this paper, we shall adopt the hypothesis that entry of pyruvate from solution to the catalytic site, any attendant protein reorganization, and the covalent addition of the TDP-conjugate base to pyruvate all occur in the process represented by the barrier following SE*.

It seems notable that the height of this barrier for the enzymic reaction (62.0 kJ mol^{-1}) is scarcely different from the height of the corresponding barrier for the non-enzymic reaction (68.9 kJ mol^{-1}). Otherwise stated, PDC stabilizes the transition state for TDP addition to the keto-group of pyruvate by 82.4 kJ mol^{-1} or only slightly more than the 75.9 kJ mol^{-1} by which it stabilizes the reactant state for this process. This could result from a reactant-like structure for the addition transition state, in which case the enzyme might be expected to stabilize the reactant state and a reactant-like transition state by similar amounts of energy. Kluger and Brandtl[8] have measured the β-^{2}H secondary isotope effect for the non-enzymic reaction. Decomposition of the adduct to thiamine and pyruvate yields an effect of 1.10, while a pyruvate–thiamine anion-like transition state should produce an effect[9] of 1.20 (three deuteriums in both cases). This evidence indicates about 50 % bond formation in the transition state for the non-enzymic reaction, and offers no easy explanation as to why the enzyme should stabilize reactant state and transition state to an equal degree. An alternative proposition is that the stabilization of TDP-cb is achieved primarily by placement of positive charge near the anionic site of TDP-cb (adding to the effect of the positive nitrogen of the thiazolium nucleus). If this charge is equally effective in stabilizing the oxyanion being formed as carbonyl addition proceeds, then the stabilization free energy might be independent of transition-state advancement and about equal for reactants, transition state and products of carbonyl addition.

Decarboxylation. For the decarboxylation step, the situation is quite different. Here the reactant state (at -10.0 kJ mol^{-1} for the enzymic reaction and 17.1 kJ mol^{-1} for the non-enzymic reaction) is stabilized by a modest 27.1 kJ mol^{-1}, while the transition state is stabilized by 114.6–47.0 = 67.6 kJ mol^{-1}. The net free energy expressed in catalytic acceleration is then 67.6–27.1 = 40.5 kJ mol^{-1}, corresponding to a factor of 10^7. This is entirely consistent with the view of Lienhard and his coworkers[10,11] that the enzymic acceleration of thiamine-promoted decarboxylation derives in part from the hydrophobic character of the enzyme active site. The binding of α-lactyl-TDP, on this model, is accompanied by development of strong binding interactions to the enzyme, but these are opposed by the unfavorable free energy of incorporation of the zwitterionic adduct into the hydrophobic active site. The net stabilization, the algebraic sum of these opposing contributions, is around 27 kJ mol^{-1}. In the transition state, the binding interactions are preserved, but the partial discharge of the zwitterionic dipole means the free energy of extraction of the transition state from water into the hydrophobic active site is less unfavorable than was true for the reactant state. Thus the 'intrinsic' binding energy of the transition state (in Jencks' terminology[12]) is more completely expressed in stabilization of the transition state to the extent of about 40 kJ mol^{-1}.

General features of the free-energy diagrams

Two general aspects of the free-energy diagrams seem particularly worthy of note. These emerge as features not unique to PDC action.

Consider the dispersion of reactant-state free energies in the non-enzymic reaction and in the enzymic reaction. For the non-enzymic reaction, the reactant-state free energies

are (left to right) 0, 65, and 17 kJ mol^{-1}: a mean of 27 kJ mol^{-1} with a mean deviation of 25 kJ mol^{-1}. To express the dispersion in relative terms, we consider its relation to RT (2.5 kJ mol^{-1}), which determines how it will affect the reaction rate or relative populations of intermediates. The dispersion of the non-enzymic reactant-state free energies (as measured by the mean deviation from the mean) is thus $10RT$. In the enzymic reaction, the reactant free energies are −0, −9, −4, −11, and −10 kJ mol^{-1}: the mean is −7 kJ mol^{-1} with mean deviation of 4 kJ mol^{-1} or around $1.5RT$. Thus in both absolute and relative terms, the enzymic reaction exhibits a smaller dispersion of reactant-state free energies, contracting a dispersion of 25 kJ mol^{-1} ($10RT$) to 4 kJ mol^{-1} ($1.5RT$). We conclude that the *enzymic reaction exhibits a leveling of reactant-state free energies.*

Consider the transition-state free energies. For the non-enzymic reaction, we have 80, 134 and 115 kJ mol^{-1} for a mean of 110 kJ mol^{-1} with mean deviation of 20 kJ mol^{-1} ($8RT$). For the enzymic reaction, the values are 71, 52 and 47 kJ mol^{-1}: the mean is 57 kJ mol^{-1}, mean deviation 9 kJ mol^{-1} (ca. $4RT$). Again, in both absolute and relative terms, the enzymic reaction contracts the dispersion of transition-state free energies from 20 kJ mol^{-1} ($8RT$) to 9 kJ mol^{-1} ($4RT$). We conclude that the *enzymic reaction exhibits a leveling of transition-state free energies.*

The conclusion that enzymes produce a leveling of reactant-state free energies has been previously drawn by Albery, Knowles and their coworkers.[13–15] The phenomenon has been termed 'matched internal thermodynamics.' As Albery and Knowles have shown, it is predicted for highly evolved enzymes by a straightforward model of biological evolution at the molecular level.

The Albery–Knowles model of molecular evolution

The fundamentals of the widely known and appreciated Albery–Knowles model of the evolution of enzymes 'to perfection' have recently been reviewed.[15] The model is cast in terms of the simplest of enzyme mechanisms: one substrate goes on to E to give EA, EA is converted into EP ('chemical step') with an equilibrium constant K_{int}, and EP dissociates. The model therefore has six rate constants, a forward and reverse constant for each of the three steps.

The aim of the Albery–Knowles approach is to determine the relative values of the six rate constants that will lead to a maximum flux per molecule of enzyme through the enzymic reaction. Such a system will require the least amount of enzyme for a given required level of product metabolite; it will therefore, to the largest possible degree, spare the host organism the energy expenditure of protein synthesis. This in turn should confer a selective advantage on the host organism, so that highly evolved organisms should have many enzymes adapted to high flux.

The method used was algebraic. It was desired to obtain the optimum values of the rate constants by setting d(flux)/d(rate constant) = 0 for the six constants and solving the resulting system of equations. Constraints were required, so the following were adopted.

(1) The equilibrium constant provided one constraint (the ratio of the product of forward rate constants to the product of reverse rate constants must equal the equilibrium constant).

(2) The rate constants for the E + A and E + P recombinations were set to diffusional values, since these will obviously yield maximum flux, providing two more constraints.

(3) The concept of 'uniform binding' was introduced. For primitive enzymes with minimal enzyme–ligand interaction, it was imagined that all reactant and transition states would bind with equal affinity. Thus the ratio of all rate constants to the rate constant for EA dissociation were made constant and the flux was optimized with respect to this dissociation rate constant. The result provided the fourth of the six constraints.

(4) The stage of evolution in which differential binding of states occurs was now considered. A further constraint was obtained by assuming eqn. (2) to be valid, where β = ca.

$$\log k_2 = \beta \log (k_2/k_{-2}) + \log k_2^{\circ} \qquad (2)$$

0.5. Here k_2 is the forward rate constant of the chemical step, k_{-2} is the reverse rate constant for the chemical step and $(k_2/k_{-2}) = K_{int}$. The relationship is the well-known Brønsted relation and the value of 0.5 for the coefficient is rather a consensus value. The quantity log k_2°, the final unconstrained quantity, defines the so-called intrinsic barrier for the step. It is assumed to be changed in later stages of evolution.

(5) The flux was optimized with respect to K_{int} to obtain the result for the differential-binding stage of evolution. The treatment yields the simple finding, $K_{int} = 1$, or 'matched internal thermodynamics.'

Success of the Albery–Knowles model

The prediction of matched internal thermodynamics has been spectacularly validated. Burbaum and Knowles[15] collected around 20 cases of enzymic reactions where both internal and external equilibrium constants were known. The external equilibrium constants range from 10^{-7} to 10^{8} and the internal equilibrium constants range from 0.6 to 17, with one outlier of 1600. It is fair to say that the predictions of the Albery–Knowles model have been shown to be strongly congruent with experimental observations. The model has correspondingly achieved wide and merited acceptance.

Critique of the Albery–Knowles model

We wish at this point to raise a question about the validity of one assumption in the Albery–Knowles treatment, namely the assumption that the reactions of enzyme-bound species are governed by a linear free-energy relationship between their rate and equilibrium constants, and that this

relationship applies to changes in rate and equilibrium constants generated by mutations in enzyme structure in the course of biological evolution. A mutation in enzyme structure can be considered a random event in which accidental changes in the DNA code for the enzyme give rise to a new structure with altered amino-acid composition in at least one site. Consider the effect of such a mutation on the free energies of three species which occur in succession along the reaction path: EA, T_{12}, and EM (or any other two enzyme-bound reactant states and the transition state connecting them: see Fig. 2 and Chart 1 below). The free energy of EA will be altered by the mutation because the structure of the native enzyme E changes to the structure of the mutant enzyme E′, thus EA mutates to E′A, and the interactions between E and EA will be energetically different from those between E′ and A in E′A. The sign and magnitude of the change in free energy as EA mutates to E′A will be determined by the change in structure from E to E′, the nature of structure A, and the strength and character of the E–A and E′–A interactions. The signs and magnitudes of a long series of such changes in free energy, corresponding to many successive events in molecular evolution, will tend to have a random character because of the large number of different ways the enzyme structure can be mutated and the large number of ways in which A can interact with E and E′. It seems very likely then that as a long series of random mutations occurs in evolution, the attendant changes in free energy at the EA stage will be random too.

An exactly similar argument can be made that changes in the free energy of the ET_{12} species and EM species will also be random. Furthermore, and most importantly, *it is likely that there will be no correlation between changes in the free energies at the EA, the ET_{12}, and the EM stages.* If this is correct, the linear free-energy relationship assumed in the Albery–Knowles approach will be incorrect, for it assumes that changes in $(G_{T12}–G_{EA})$ will always be a constant fraction β of changes in $(G_{EM}–G_{EA})$.

The argument that the changes in free energy of enzyme interaction with reactant-state, transition-state and product-state structures for reactions of enzyme-bound species will be random and uncorrelated has two parts.

First, the structures of the three species may be quite different: consider an S_N2 reaction catalyzed by a methyltransferase, for example. The large structural differences between the three successive states (tetrahedral methyl bonded to donor; approximately planar methyl bonded to both donor and acceptor; tetrahedral methyl bonded to acceptor) in the reaction then dictate that the interactions with enzyme will be quite different and will be affected by enzyme mutations in very different ways.

The second part of the argument addresses the fact that, in spite of these structural differences between successive states, their interactions with various solvents, for example, may nevertheless be governed by just the kind of linear free-energy relationship assumed by Albery and Knowles. If a series of solvents, why not a series of mutant enzymes? In the case of solute transfer from solvent to solvent, the surrounding solvent may relax to accommodate various solute structures. The varying structures of reactant, transition state and product may then, in effect, dictate solute–solvent interactions in such a way that reaction rate and equilibrium constants from solvent to solvent are correlated. In enzyme catalysis, in contrast, the essential fact is that the enzyme structure governs rigorously and unforgivingly the free energies of bound species, often tolerating little or no change in substrate, transition-state or product characteristics. It therefore seems unlikely in the enzymic context that any correlation of free energies can be dictated solely by characteristics of the molecules undergoing catalyzed reaction.

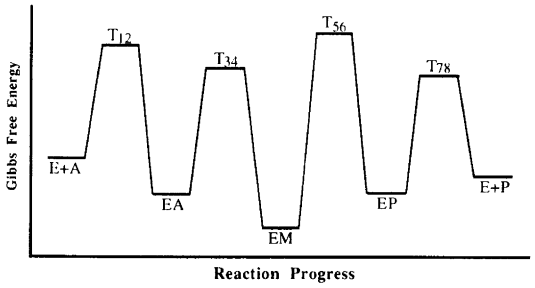
In any case, it seems desirable to explore whether the leveling of reactant-state energies, a prediction of the Albery–Knowles theory which is known to be confirmed by experience, in any way depends on the assumption of a linear free-energy relationship for reactions of bound species.

An alternative, stochastic model of molecular evolution

If we are to construct a model that avoids the assumption of a linear free-energy relationship, we lose constraints necessary for the Albery–Knowles program of optimizing the enzymic flux. Therefore, instead of solving the optimization problem analytically, we have chosen to simulate molecular evolution computationally in such a way that the mutational change in free energy of every state along the reaction pathway is random and independent, within only the constraints imposed by basic principles of thermodynamics and kinetics.

The result is a completely *stochastic model* for molecular evolution. If it fails to reproduce the observed leveling of reactant-state free energies in evolved enzymes, that will indicate that the linear free-energy relationships assumed by Albery and Knowles not only do hold – in spite of the objections offered above – but in addition are critical contributors to the outcome of molecular evolution. It would of course in such a case be possible to modify the stochastic simulation by introduction of contingencies that will simulate the required linear free-energy relationships. On the other hand, if the purely stochastic model reproduces 'matched internal thermodynamics,' it will mean that the observations can be accounted for without any assumption of linear free-energy relationships, which then become an unnecessary physical feature of the Albery–Knowles model, although mathematically convenient.

Fig. 2 shows a schematic free-energy diagram for the kind of system we will subject to simulation. Since the simulation is freed of the necessity for algebraic constraints, any mechanism can be considered. We begin with a Uni-Uni reaction with two 'chemical steps.' Chart 1 shows the corresponding mechanism with assignment of rate constants and gives the relationship between the flux φ



$$E + A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} EA \underset{k_4}{\overset{k_3}{\rightleftharpoons}} EM \underset{k_6}{\overset{k_5}{\rightleftharpoons}} EP \underset{k_8}{\overset{k_7}{\rightleftharpoons}} E + P$$

Fig. 2. Schematic free-energy diagram for an enzymic reaction which is to be modeled stochastically. See Chart 1 for the corresponding mathematical expressions.

and these rate constants and between the rate constants and the free energies of species of species along the reaction path. Note that the thermodynamic constraint imposed by the equilibrium constant of the overall reaction, which is not accessible to evolution, is automatically included.

In Fig. 3, a flow chart of the simulation program MUTATION is presented.[16] The simulation begins with an as-

$$E + A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} EA \underset{k_4}{\overset{k_3}{\rightleftharpoons}} EM \underset{k_6}{\overset{k_5}{\rightleftharpoons}} EP \underset{k_8}{\overset{k_7}{\rightleftharpoons}} E + P$$

$$\phi = [(k_a/K_a)A - (k_p/K_p)P] / [1 + (A/K_a) + (P/K_p)]$$

$$\frac{K_a}{k_a} = \frac{1}{k_1} + \frac{k_2}{k_1k_3} + \frac{k_2k_4}{k_1k_3k_5} + \frac{k_2k_4k_6}{k_1k_3k_5k_7}$$

$$\frac{1}{k_a} = \frac{1}{k_3} + \frac{k_4}{k_3k_5} + \frac{k_4k_6}{k_3k_5k_7} + \frac{1}{k_5} + \frac{k_6}{k_5k_7} + \frac{1}{k_7}$$

$$\frac{K_p}{k_p} = \frac{1}{k_8} + \frac{k_7}{k_6k_4} + \frac{k_7k_5}{k_8k_6k_4} + \frac{k_7k_5k_3}{k_8k_6k_4k_2}$$

$$\frac{1}{k_p} = \frac{1}{k_6} + \frac{k_5}{k_6k_4} + \frac{k_5k_3}{k_6k_4k_2} + \frac{1}{k_4} + \frac{k_3}{k_4k_2} + \frac{1}{k_2}$$

$$k_1 = \frac{kT}{h} e^{-(G_{T12} - G_E - G_A^0)/RT} = \frac{kT}{h} e^{-(G_{T12}/RT)} \quad \text{if } G_E + G_A^0 = 0$$

$$k_2 = \frac{kT}{h} e^{-(G_{T12} - G_{EA})/RT}$$

$$\vdots$$

$$k_8 = \frac{kT}{h} e^{-(G_{T78} - \Delta G^0)/RT} \quad \text{if } G_E + G_A^0 = 0$$

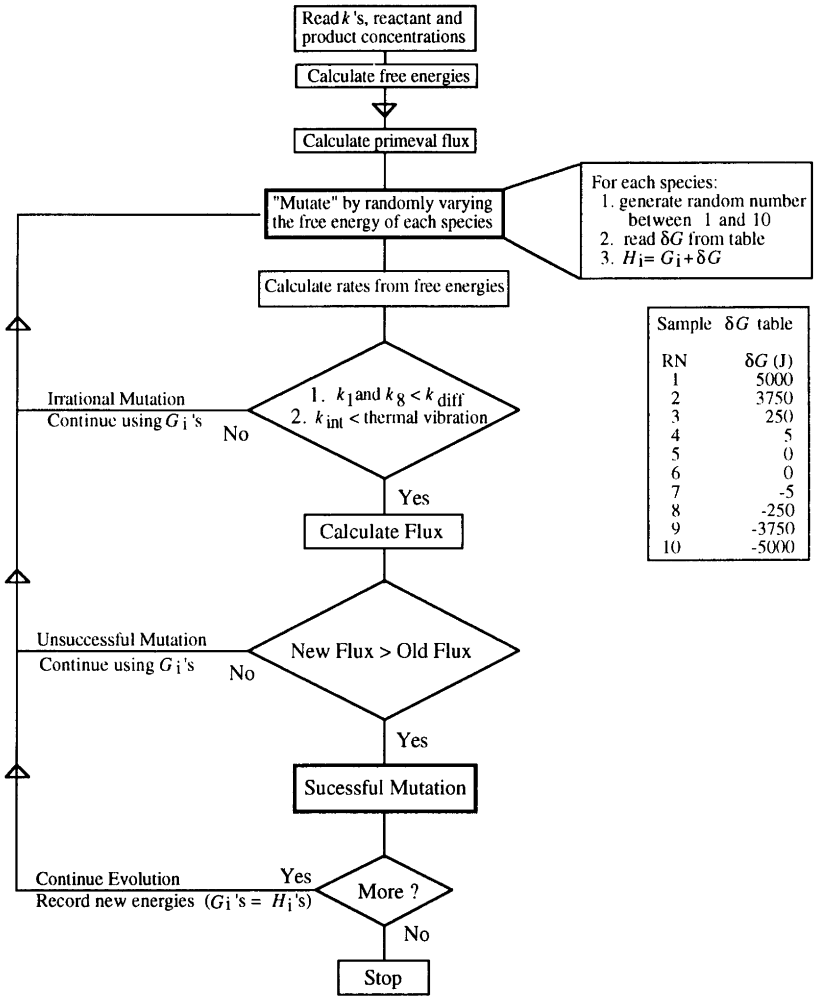


Fig. 3. Flow chart of program MUTATION, which simulates molecular evolution stochastically.

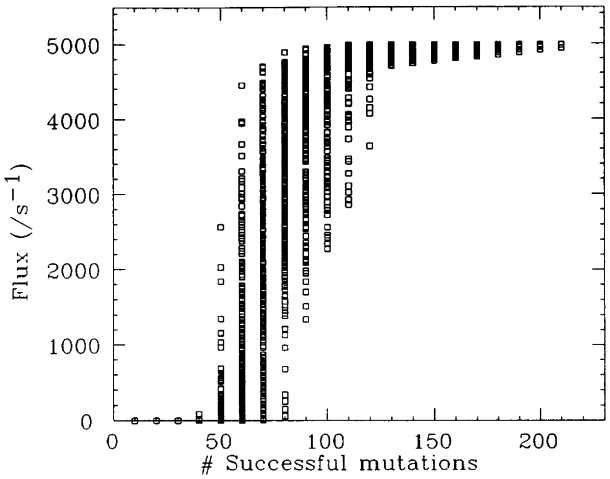


Fig. 4. Flux through the enzymic reaction for 300 evolutionary trajectories, shown as a function of the number of successful mutations. All simulations in this set began with the following primeval rate constants (s^{-1} unless otherwise indicated; see Chart 1): k_1 = 10 m^{-1} s^{-1}; k_2 = 100; k_3 = 0.0001; k_4 = 100; k_5 = 10^7; k_6 = 0.01; k_7 = 0.001; k_8 = 0.01 M^{-1} s^{-1}. The equilibrium constant is therefore unity. Standard-state concentrations were 1 mM (A) and 1 μM (P). The primeval flux was 10^{-8} s^{-1}. The maximum second-order rate constant was set to 10^7 M^{-1} s^{-1} so the maximum flux is around 5000 s^{-1} [= ca. (1/2)(10^7 M^{-1} s^{-1})(1 mM)]. The maximum first-order rate constant was set to 10^{12} s^{-1}. Simulations converged to values of k_1 = k_8 = 10^7 M^{-1} s^{-1} in all cases. The internal rate constants k_2 to k_7 had various values in the neighborhood of 10^{12} s^{-1} (see Fig. 5).

sumed mechanism and set of 'primeval' rate constants supplied by the user. The program contains a random-number generator[17] and a table of arbitrary translations, supplied by the user, of random numbers into energy increments. By use of random numbers and the table, the program alters the free energy of each enzyme-bound state by a random increment and then calculates new microscopic rate constants from the free energies of activation thus generated (see Chart 1). Each of the rate constants is checked for consistency with these two fundamental kinetic constraints.

(1) Second-order rate constants must be equal to or smaller than the diffusional second-order rate constant.

(2) First-order rate constants must be equal to or smaller than a molecular vibrational frequency.

If either constraint is violated, the event is classified as an irrational mutation and discarded; the program returns to the original set of free energies and begins again. If neither constraint is violated, the mutation is declared rational. For

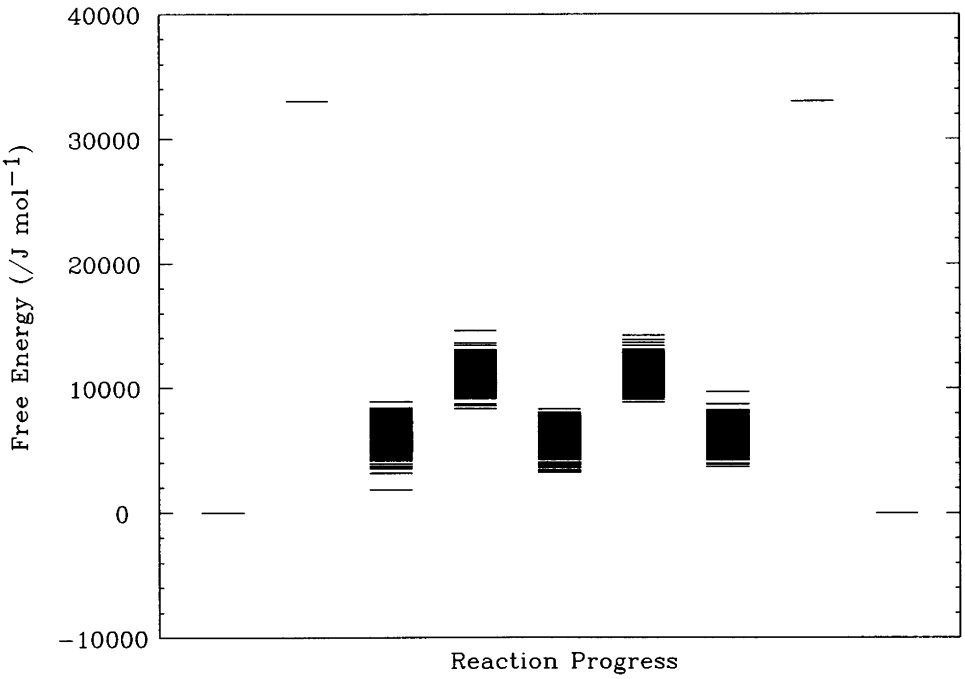


Fig. 5. Superimposed free-energy diagrams for 300 evolutionary trajectories from the starting point given in the caption to Fig. 4. Note that the critical 'on' and 'off' steps have sharp transition-state free energies corresponding to the maximum flux of 'perfect' (diffusion controlled) enzymes. All the internal steps have approached the maximum rate of a molecular vibration frequency, but because the flux is little dependent on the exact values of the rate constants, both reactant-state and transition-state free energies exhibit 'genetic drift.' Nevertheless, reactant-state free energies are substantially equal to each other ('matched internal thermodynamics') and the transition states are also equal in free energy to each other ('matched internal kinetics'). For these diagrams, reactant and product standard states have been adjusted to 1 M, although values of 1 mM (A) and 1 μM (P) were employed in the simulations.

rational mutations, the flux is calculated and compared with the 'old flux' for the preceding enzyme. If the old flux is larger, the mutation is declared unsuccessful, the new free energies are discarded and the old ones retained for the next round of variation. If the new flux is larger, the mutation is declared successful, and the new free energies replace the old ones. The new free energies are then the starting point for the next round of mutations.

Results of the stochastic simulation

Fig. 4 shows the result of 300 simulations, all starting with the same primeval rate constants indicated in the caption. Since the starting flux was the same small value for each simulation, there is a single common starting point at zero successful mutations. The flux is of course required by the program to increase as the number of successful mutations grows, but Fig. 4 shows how dramatically differently this occurs along each of the individual simulations. Thus after a number of successful mutations that produce on the average about half the maximum flux, about 50–70 successful mutations in Fig. 4, a very great diversity is exhibited: the microscopic history of each evolutionary sequence is quite different, so that at 60 successful mutations, there is a dispersion of fluxes covering nearly the entire range from very small values up to around 5000 s^{-1}. As evolution proceeds, this diversity, created by the *chance* of random mutation, is eliminated by the *necessity* of natural selection for higher flux. The range of flux values narrows until the maximum value of about 5000 s^{-1} is achieved along every evolutionary trajectory. This is in fact the maximum physically achievable flux on the model chosen, corresponding to diffusion-controlled reaction of the enzyme with both substrate and product, so that all the enzymes along each

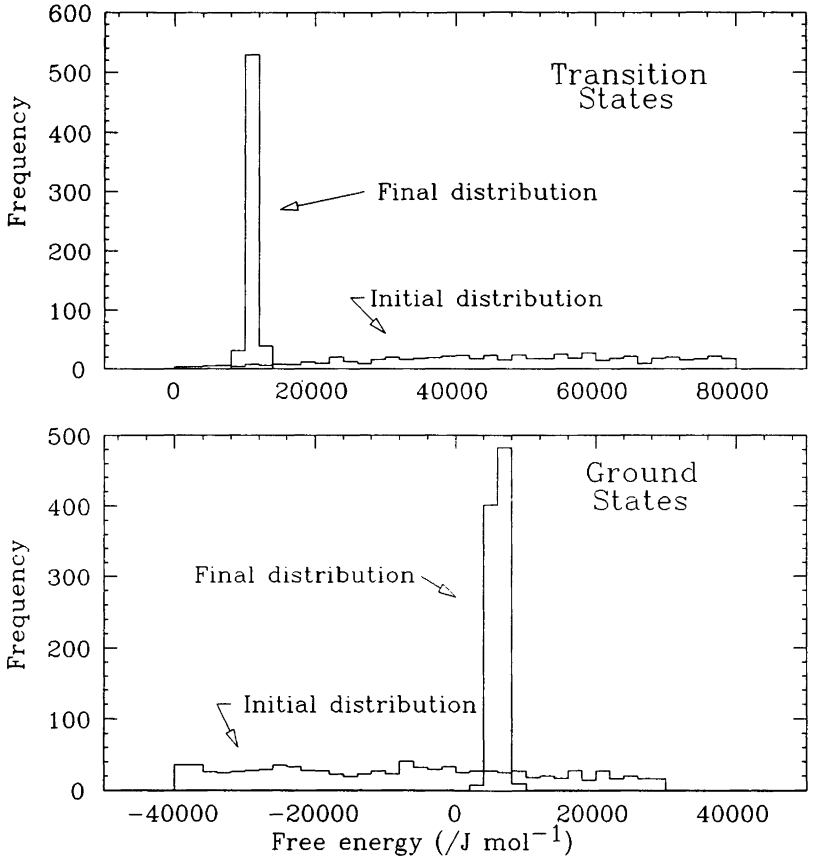


Fig. 6. Initial and final free-energy distributions for 300 evolutionary trajectories. For each trajectory, a starting-point set of reactant-state and transition-state free energies was generated randomly. The free energies of the 'on' and 'off' transition states were initially set at 50 kJ mol^{-1}. The upper and lower bounds for internal transition-state free energies were 0 and 80 kJ mol^{-1}, respectively. The upper and lower bounds for internal reactant-state free energies were −40 and +30 kJ mol^{-1}, respectively. Free-energy profiles were rejected which had free energies of activation smaller than vibrational-frequency values for internal steps. This had the effect of truncating the lower branch of the internal transition-state free-energy distribution and the upper branch of the internal reactant-state free-energy distribution (as can be seen by inspection of the figure). The program MUTATION then subjected each of the starting-point enzymes to evolution until perfection was attained. The final distribution of reactant-state energies is (a) sharp, corresponding to 'matched internal thermodynamics,' and (b) centered at a higher average free energy than the initial distribution, showing that evolution leads to higher reactant-state energies. The final distribution of transition-state free energies is (a) sharp, corresponding to 'matched internal kinetics,' and (b) centered at a lower average free energy than the initial distribution, showing that evolution leads to lower transition-state energies.

evolutionary trajectory have 'evolved to perfection' in the phrase of Albery and Knowles. In the series of 300 simulations shown in Fig. 4, the primeval flux was 10^{-8} s^{-1} and the 'perfect' flux $10^{3.7}$ so that the enzymes evolved an increased catalytic power of 11.7 orders of magnitude.

Fig. 5 presents the results in the form of superimposed free-energy diagrams for the 300 'perfect enzymes.' The energy barriers for the 'on' reaction of substrate and the 'off' reaction of product exhibit transition-state free energies that are very sharply equal for all evolutionary trajectories (corresponding to diffusion control). On the other hand, the free energies of the reactant and transition states for the relatively fast internal reactions to EA to EM and EM to EP show considerable 'genetic drift' because the flux is not much dependent on the exact values of these free energies as long as the steps remain very fast compared with the 'on' and 'off' reactions. Note, however, that the *reactant-state free energies from one internal species to another have become essentially equal to one another, so that the model reproduces the experimental observations of 'matched internal thermodynamics.'* Furthermore, for the internal reactions, the transition-state energies also tend to become equal to each other so that, in addition to values of unity for the internal equilibrium constants, *the model predicts equal rate constants for all internal reactions – 'matched internal kinetics.'*

These points are confirmed by the further simulations described by Fig. 6. Here 300 randomly chosen sets of initial, primeval rate constants were generated. In the upper part of the figure, one sees the broad spectrum of free energies of the corresponding transition states; in the lower part of the figure, a similar spectrum of reactant state free energies. Then, at the end of the simulation, each of the 300 different initial kinetic manifolds had evolved to a similar, convergent manifold, so that the free energies of all enzyme-bound transition states then lay within a narrow band and the free energies of all enzyme-bound reactant states also lie within their own narrow band.

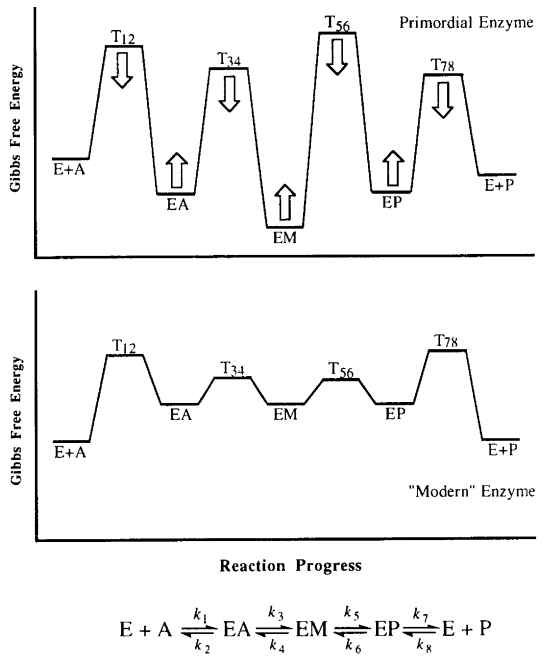


Fig. 7. Natural selection for higher flux exerts a downward pressure on the free energies of transition states and an upward pressure on the free energies of reactant states. The results are 'matched internal thermodynamics' and 'matched internal kinetics.'

Conclusions

These purely stochastic simulations show that the original Albery–Knowles finding of 'matched internal thermodynamics' is more general than might have been supposed: the result is in no way dependent upon the assumption of a linear free-energy relationship for internal reactions. Conversely, the experimental observations of matched internal thermodynamics do not in any way support a supposition of the existence of such linear free-energy relationships.

In addition, the simulation suggests 'matched internal kinetics.' The rate constants for enzyme-bound species will tend to equality as evolution proceeds. This is in general agreement with the observation that multiple steps often contribute to rate limitation for enzymes.

In fact, both characteristics of evolved enzymes, leveled reactant-state free energies and leveled transition-state free energies are predicted by very simple considerations (Fig. 7). An evolutionary process in which enzyme selection pivots on high flux will always exert a downward pressure on the free energies of transition states and an upward pressure on the free energies of reactant states: low transition-state free energies and high reactant-state free energies make for low barriers, for enzymes largely in the free and thus reactive form, and therefore for high flux. At the same time, it is always the highest free-energy transition state and the lowest free-energy reactant state that primarily determine the rate. Thus enzymes will tend to be selected in which the free energy of the primevally highest free-energy transition state is steadily lowered until the next lower transition state is encountered; then it will become the focus of evolutionary attention. Similarly, the lowest free-energy reactant state will be steadily raised until the next higher state is encountered. The eventual result will inevitably be leveling of all transition states and leveling of all reactant states.

Acknowledgements. This work was supported by the US National Institute of General Medical Sciences under research grant No. GM-20198. D.W.H. is grateful for support under a Medicinal Chemistry and Pharmacology Training Grant (GM-07775) from the US National Institute of General Medical Sciences.

References

1. Schowen, R. L., In: Gandour, R. D. and Schowen, R. L., Eds., *Transition States of Biochemical Processes,* Plenum Press, New York 1978, p. 77.
2. Kraut, J. *Science (Washington, DC) 242* (1988) 533.
3. Kluger, R., Chin, J. and Smyth, T. *J. Am. Chem. Soc. 103* (1981) 884.
4. Washabaugh, M. and Jencks, W. P. *J. Am. Chem. Soc. 111* (1989) 674.
5. Washabaugh, M. and Jencks, W. P. *J. Am. Chem. Soc. 111* (1989) 683.
6. Washabaugh, M. and Jencks, W. P. *Biochemistry 27* (1988) 5044.
7. Alvarez, F. J., Ermer, J., Hübner, G., Schellenberger, A. and Schowen, R. L. *J. Am. Chem. Soc. 113* (1991) 8402.
8. Kluger, R. and Brandtl, M. *J. Am. Chem. Soc. 108* (1986) 7828.
9. Alvarez, F. J. and Kovach, I. M. *Unpublished observations.*
10. Crosby, J., Stone, R. and Lienhard, G. E. *J. Am. Chem. Soc. 92* (1970) 2891.
11. Crosby, J. and Lienhard, G. E. *J. Am. Chem. Soc. 92* (1970) 5707.
12. Jencks, W. P. *Adv. Enzymol. 43* (1975) 219.
13. Knowles, J. R. and Albery, W. J. *Biochemistry 15* (1976) 5631.
14. Burbaum, J. J., Raines, R. T., Albery, W. J. and Knowles, J. R. *Biochemistry 28* (1989) 9293.
15. Burbaum, J. J. and Knowles, J. R. *Biochemistry 28* (1989) 9306.
16. The source code for the simulation (mutation.c) is available via anonymous FTP from kuhub.cc.ukans,edu (129.237.1.10) in the (.chemistry) subdirectory.
17. A linear congruential and additive generator [general treatment: Miller, K. W. and Park, S. K. *Communications to the ACM 31* (1988) 1192] was employed. The program is known as XRAND (copyright 1988: Andreas Nowatzyk) and is available via anonymous FTP from the wuarchive.wustl.edu (128.252.135.4) in the /usenet/comp.sources.misc/volume02 directory as xrand. Z. The simulation does not appear to suffer if a system-supplied random-number generator is used.

Received November 12, 1991.