Enzyme Kinetics and Molecular Evolution

STEVEN A. BENNER

Laboratory for Organic Chemistry, ETH-Zentrum, CH-8092 Zürich, Switzerland

Received May 10, 1988 (Revised Manuscript Received March 6, 1989)

Contents

Steven Benner received a B.S.–M.S. degree in Molecular Biophysics and Biochemistry at Yale in 1976 and a Ph.D. degree in Chemistry in 1979 at Harvard under the joint tutelage of F. H. Westheimer and R. B. Woodward. He then stayed at Harvard as a Xerox Fellow, a Junior Fellow of the Harvard Society of Fellows, and an Assistant Professor until 1986, before taking the Chair in Bioorganic Chemistry at the E.T.H. in Zürich. He is especially interested in developing ways in which biological information can be used by organic chemists to direct experiments designing and manipulating macromolecules. From this interest has come the first synthetic genes designed for protein engineering, designed polypeptides that fold in aqueous solution and catalyze reactions, several ancient ribonucleases from extinct organisms (prepared by recombinant DNA methods), synthetic bases that expand the genetic alphabet in DNA and RNA, and protocols for predicting the tertiary structure of natural proteins from sequence data.

I. Introduction

Reactions in the active sites of enzymes are typically 10^{10}–10^{14} times faster than analogous reactions in solution. In many cases, rates of enzymatic reactions are astonishingly high by chemical standards. For example, mammalian carbonic anhydrases turn over 36 million molecules of substrate per minute at saturation.[1] Steroid isomerase from *Pseudomonas testosteroni* turns over substrate 6 million times per minute.[2]

Nothing in biology makes sense except in the light ov evolution,[3] and this is no less true for the catalytic behavior of enzymes as it is for more "biological" traits of living systems. The catalytic power of natural enzymes reflects two contrasting evolutionary processes, natural selection and neutral drift. Natural selection controls behaviors of enzymes that influence the ability of a host organism to survive and reproduce. Enzymatic traits that do not have an impact on the survival of the host will drift as the structure of the protein drifts as random mutations are accumulated.[4]

Distinguishing between selected and nonselected behaviors in enzymology has proven to be challenging.[5-7] Yet the challenge must be met to understand enzymatic behavior at a biologically fundamental level. One approach to make this distinction assumes that functional behavior should convergently evolve in nonhomologous enzymes catalyzing analogous reactions in analogous environments and should be conserved as homologous proteins divergently evolve to yield different enzymes performing analogous functions in different organisms. Thus, comparison of the behaviors of different enzymes that perform analogous functions in different organisms should (at least in principle) allow the biochemist to identify functional traits.

This analysis is more complicated with enzymatic traits that depend on many structural features in the active site (such as stereospecificity and reaction mechanism).[6,7] Such traits might not drift rapidly even if they serve no selectable purpose. For example, a ribonucleotide reductase dependent on vitamin B_{12} as a cofactor may have no selective advantage over one using iron. However, it should be difficult for an enzyme using the first mechanism to evolve to use the other without passing through structural intermediates with no catalytic activity whatsoever. Thus, mechanism would not necessarily drift even if an alternative mechanism has equal survival value, meaning that conservation of mechanism in homologous enzymes, even widely divergent ones, does not necessarily imply that the mechanistic choice is an optimal one.

For kinetic behaviors, this complicating consideration almost certainly does not apply. Essentially all of the

0009-2665/89/0789-0789$06.50/0

catalytic power of an enzyme can be destroyed by a single critically placed point mutation. Michaelis constants, kinetic order, regulatory properties, and inhibition patterns of an enzyme can all be changed simply by changing by a small number of amino acids.[5] Thus, if the details of catalytic behavior are conserved in homologous enzymes from different organisms, these details must serve some directly selected function.

For example, the kinetic parameters (k_{cat}/K_M in particular) of the triose phosphate isomerases from rabbit and from coelacanth, a lobe-finned fish, are quite similar.[8] The two organisms diverged 400 million years ago; in the time since their divergence, many point mutations might have altered the kinetic behavior of the protein. The fact that the behavior has not been altered must reflect continued selection pressure on both enzymes to maintain these parameters; coelacanths and rabbits unfortunate enough to sustain mutations that make their isomerases slower (and there must have been many in the last 400 million years) simply must have failed to survive and reproduce.

Therefore, kinetic behavior in enzymes offers one of the best opportunities to examine the ability of natural selection to optimize enzymatic behavior independent of complicating historical factors. Further, as kinetic parameters can often be quantitatively measured, such measurements may provide an estimate of the ability of natural selection to "fine tune" enzymatic behavior. Thus, it is not surprising that the kinetic behavior of enzymes has been a focus of discussions of functional adaptation in enzymes for many years.

II. "Catalytic Efficiency" as a Goal of Natural Selection

Most biochemists have little difficulty accepting the proposition that natural selection prevents the erosion through "drift" of the high catalytic power of enzymes. This is especially true if the enzyme plays a key role in an "important" pathway. For example, triose phosphate isomerase catalyzes a reaction central to glycolysis, important as an organism fights or flees in a life-threatening situation.[9]

However, this proposition is not as obvious as it seems. There are several reasons why a rabbit (or coelacanth) with a mutant triose phosphate isomerase that is half as fast as the wild type protein might not run (or swim) half as fast.

First, and most trivially, triose phosphate isomerase might not be a "rate-limiting" enzyme in a biological process that is limiting for survival. However, even if we assume that triose phosphate isomerase *is* rate limiting in glycolysis, a slower enzyme does not necessarily imply a slower physiological flux through glycolysis. A mutation that makes an enzyme half as fast may be compensated by a second mutation that increases the level of the expression of the gene by a factor of 2 (Figure 1). The net flux through glycolysis is the same; the only "cost" associated with the slower enzyme is the cost to synthesize twice as much of it.

Such a "cost" is a small fraction of the total metabolic energy expended by an organism,[10-12] and the associated selective disadvantage of synthesizing more protein is far smaller than the selective disadvantage associated with being able to run only half as fast. In an animal with 10 000 proteins, the energy cost of synthesizing another batch of the mutant protein should be on the order of 0.01% of the total energy expenditure. It is far from obvious that such a small additional outlay of energy would be selectively disadvantageous in a multicellular organism.

Number of enzyme	k_{cat}/K_M / diffusion	total flux (relative)	size of each enzyme	biosynthetic cost (relative)
100	1.0	1	10000	1
1000	0.1	1	1000	1
10000	0.01	1	100	1
100000	0.001	1	10	1

Figure 1. The same net flux can be produced at the same total biosynthetic cost with different enzymes with different microscopic kinetic parameters with different sizes.

What appears to be certain (at least in 1989) are the following:

(a) Protein expression is regulated throughout biology; genes that serve no immediate purpose are turned off. This strongly suggests that expression of excess protein is selectively disadvantageous.

(b) The cost of synthesizing excess proteins can be estimated, although quantitative aspects of these estimates remain problematical.[11-16]

(c) The fact that enzymes are not slower, despite the ready availability of mutations that would make them slower, forces the conclusion that substantially slower enzymes are selectively disadvantageous.

(d) It remains possible that faster enzymes are also accessible by point mutation from native enzymes; indeed, a few point mutants made by recombinant DNA techniques have been reported to display "improved" kinetic parameters.[17] The fact that these mutations are not themselves selected in the wild strongly suggests that improvement in kinetic behavior (at least of the type described) does not confer selectable advantage. Alternatively, the "improved" catalytic efficiency in these mutants might be at the expense of other desirable enzymatic traits.

(e) However, available data do not allow the conclusion that faster enzymes are *generally* accessible by point mutation from native enzymes. Thus, it is possible that enzymes with catalytic efficiency higher than those found naturally might exist, be selectively advantageous, and yet be so structurally distant from the enzymes presently available they cannot easily arise by point mutation of existing enzymes.[5] In this view, the catalytic efficiency of existing enzymes reflects only a local optimization of protein structure.[5]

These points are potentially paradoxical. Enzymes are fast, despite the ready availability of mutations that make them slower. Thus, catalytic efficiency of some sort must be selected. However, the apparent selective disadvantage of a slower enzyme, estimated by the cost of additional enzyme to maintain the net flux through a pathway, seems to be too small to critically influence the survival of a multicellular organism. Further, to some, enzymes seem to be "too big". While this impression reflects primarily the perceptions of the biochemist (in fact, enzymes are not especially large when judged by the standards of everyday life;[111] the linear dimension of a protein "box" increases only with the cube root of its molecular weight; therefore, for the box to contain a substrate, it must have a molecular weight much larger than that of the substrate), it has caused many biochemists to think that the biosynthesis of

extra protein matter has only slight selective disadvantage.

Conversely, many enzymes could be faster. That they are not might reflect a preference for more, slower enzyme molecules. However, it could also indicate that catalytic efficiency is sacrificed for nonkinetic properties or that the enzyme is trapped in an evolutionary local optimum.

This review attempts to resolve these paradoxes. First, specific models that aim to correlate specific kinetic parameters with those expected for a catalytically optimal enzyme are discussed. Predictions made by these models are compared with experimental data to distinguish selected and nonselected kinetic behavior in enzymes. Finally, these data are used to assess quantitatively the extent to which natural selection can "fine tune" the behavior of enzymes in general.

The most significant conclusion that arises from this discussion is that the logical foundations of many commonly held views concerning kinetic adaptation in proteins are flawed. While these flaws do not alter the conclusion that natural selection is capable of fine tuning kinetic behavior in proteins, they do force a reevaluation of the view that catalytic efficiency, as commonly defined, is a target of natural selection.

A. The Magnitude of K_M

A correlation between the value of the Michaelis constant (K_M) for an enzyme and the physiological concentration of substrate was mentioned over two decades ago. Cleland noted as a "general rule" that enzymes "will operate with reactant concentrations in the region of their apparent Michaelis constants, or above, when the pathway they are a part of is operating at full capacity. Otherwise the catalytic potential of the enzyme is wasted".[18]

Fersht, Atkinson, Crowley, and Albery and Knowles presented different views.[9,19-21] Noting that tight binding of substrate implies a stable, and therefore unreactive, substrate, Fersht argued that K_M should be high, even as high as 10 times the concentration of substrate under physiological conditions. Crowley presented a similar conclusion, but on different grounds. In his view, K_M should be higher than the concentration of substrate so that fluctuations in the concentration of substrate could be damped.[21] While agreeing that enzymes metabolically responsible for maintaining constant concentrations of intermediates might have K_M values approximately the same as in vivo substrate concentrations, Albery and Knowles suggested that "digestive" enzymes, mediating a constant flux of substrate, might always work at saturation.[9] Unfortunately, digestive enzymes often act physiologically on a heterogeneous mixture of substrates under variable conditions; thus, the substrate and conditions appropriate for measuring physiologically relevant rates are difficult to know.

The K_M values for many enzymes are indeed equal to or slightly above the physiological concentration of substrates in pathways that operate at full capacity; the glycolytic pathway again provides a good illustration.[22] However, even here, there are some noteworthy exceptions. The reported K_M for aldolase is 12 μM, while the physiological concentration of the substrate, fructose 1,6-diphosphate, ranges between 27 and 216 μM under a variety of conditions. The reported K_M for glyceraldehyde-3 phosphate dehydrogenase is 44 μM, while the reported physiological concentration of glyceraldehyde 3-phosphate is 1–3 μM.

This variation of K_M within a factor of 10 of the substrate concentration may reflect drift, subtle adaptation, or experimental error. However, a comparison of the K_M values of isozymes that have evolved under slightly different metabolic conditions suggests that at least some of this variation is adaptive. For example, isozymes of lactate dehydrogenase and glycerol-3-phosphate dehydrogenase are found in liver and muscle tissue in chicken. The K_M values for lactate and glycerol 3-phosphate for the two isozymes from muscle are 10-fold higher than for the two isozymes from liver.[23] This difference is consistent with K_M values adapted to higher physiological concentrations of the two substrates in muscle. This suggests that K_M values for the two isozymes are "tuned" with a tolerance less than a factor of 10.

Precise definitions of the optimal value for K_M are needed to decide whether K_M is yet more finely tuned. The literature mentioned above suggests that no definition is generally accepted. Further, concentrations of substrates fluctuate widely in many pathways, in particular, glycolysis. While arguments that K_M should be higher than the concentration of substrate to maintain homeostasis are applicable only where the substrate concentrations fluctuate, fluctuations obviously make it difficult to determine concentrations that are relevant for natural selection. In turn, this makes it difficult to determine with precision whether experimentally measured values of K_M conform quantitatively to criteria defining an optimum.

One example that argues persuasively that natural selection can finely tune K_M comes from studies by Shaw on chloramphenicol acetyltransferase.[24] The enzyme, found in many bacteria, confers resistance to chloramphenicol, a naturally occurring antibiotic. The K_M for chloramphenicol is about 15 μM, which is also the minimum concentration of chloramphenicol able to inhibit the growth of *E. coli*. Values for K_M much below this level would not be positively selected for; values for K_M much above this level would be decidedly disadvantageous. Here, K_M seems to be optimized to within a factor of 2.

Another example of an adjustment of K_M that appears to be adaptive comes from studies of tumor cells. Two enzymes are known to phosphorylate glucose in the first step of glycolysis. Hexokinase has a rather low K_M and is presumed to be the most important catalyst for this step at low glucose concentrations. Glucokinase has a much higher K_M. It contributes a significant proportion of the total flux only at high concentrations of glucose. Weber has noted that upon transformation, the relative amounts of different isozymes are often shifted to favor the isozyme with the lower K_M.[25] This appears to be the case for hexokinase and glucokinase in hepatoma cells. Weber has argued persuasively that this change in isozyme distribution is an adaptive response to the needs of transformed cells.[25] Indeed, it appears that transformed cells with special phenotypes are selected as the disease progresses. Such progressions may be instructive about the detailed kinetic behaviors that confer survival advantage in the peculiar

environment of tumor cells.

B. k_{cat}/K_M and the Diffusion Limit

Focusing on K_M to the exclusion of other kinetic parameters overlooks much of the power of enzymes as catalysts.[26,27] Especially important as a measure of kinetic power is the ratio of k_{cat} to K_M. This ratio has the same units as the second-order diffusion rate constant and is frequently compared to the diffusion rate constant as a measure of catalytic efficiency. If the k_{cat}/K_M value for an enzyme is similar in magnitude to the diffusion rate constant, the enzyme is viewed as catalytically optimal.

This conclusion is consistent with the fact that at concentrations of substrate lower than K_M, the rate of turnover of an enzyme with a k_{cat}/K_M value at the diffusion limit will be the rate at which substrate diffuses to the enzyme. This is an extrinsic rate that cannot be altered by changing the structure of the enzyme itself and therefore represents an upper limit to the kinetic performance of the enzyme.

More than a decade ago Cleland noted that the k_{cat}/K_M values of many enzymes approach the diffusion limit and are therefore "perfect" by this criterion.[28] For example, the k_{cat}/K_M for carbonic anhydrase is 8.3×10^7 M^{-1} s^{-1}.[1] For acetylcholinesterase, the value is 1.6×10^8 M^{-1} s^{-1}.[29] For triose phosphate isomerase, the value is 4×10^8 M^{-1} s^{-1}. Fumarase is reported to have a value of 1.6×10^8 M^{-1} s^{-1}.[30]

Extending this analysis, Albery and Knowles defined an "efficiency function" for an enzyme that equals the ratio of the net flux through the enzyme to the flux through an idealized diffusion-limited enzyme catalyzing the same reaction under the same circumstances. A ratio of 1.0 indicated a "perfect" enzyme.[9] When applied to the enzyme triose phosphate isomerase,[9] the enzyme was found to have an efficiency function equal to 0.6. The enzyme appeared to have evolved to the edge of a "plateau of perfection",[31] a level of kinetic performance beyond which further adaptation does not offer sufficient additional selective advantage to be worthwhile.

Such conclusions depend on an accurate assessment of the second-order diffusion rate constant, a value that is quite variable in the bioorganic literature. Occasionally it is quoted as being as high as 10^{11} M^{-1} s^{-1}. At the other extreme, enzymes with k_{cat}/K_M ratios of 10^5 are regarded as being "nearly perfect" because their k_{cat}/K_M ratios "approach the diffusion limit". In fact, values of 10^{10}–10^{11} M^{-1} s^{-1} should be reversed for the rate of diffusion of a proton in water.[32] For the diffusion of other molecules in water, values of 10^8 M^{-1} s^{-1} are the maximum that have been observed.[33] For example, a value of 9×10^7 M^{-1} s^{-1} has recently been reported for the association of chymotrypsin with a low molecular weight substrate.[34] The diffusion-limited rate of association of *p*-nitrobenzoic acid with horseradish peroxidase is reported as 1.3×10^8 M^{-1} s^{-1}.[35]

Further, there has been some confusion regarding the constraint on enzymic behavior that the diffusion limit imposes. At one extreme, a k_{cat}/K_M ratio of 10^7 M^{-1} s^{-1} has been interpreted as requiring that the enzyme turnover 10^7 substrate molecules to product per second,[36] an incorrect conclusion unless the K_M is 1 M. Further, it is occasionally argued that the enzyme with such a high value of k_{cat}/K_M is diffusion limited under all circumstances. In fact, the enzyme is only diffusion limited when [S] < K_M.

What fraction of enzymes have k_{cat}/K_M ratios that approach the value 10^8 M^{-1} s^{-1}? The literature must be reviewed with two caveats in mind. First, to be relevant, k_{cat}/K_M values must be measured under "physiological" conditions, conditions that are rarely reproduced in vitro. Further, the enzymes studied must be both pure and "whole". A certain fraction of the data inevitably will be "unphysiological" because an essential cofactor was missing or because of unexpected complications arising from incomplete reproduction of physiological conditions in vitro.

C. The Diffusion Limit and Viscosity

The very high values of k_{cat}/K_M mentioned above are often cited, and the general impression seems to be that all enzymes conform to this high standard of performance. This is not the case; many enzymes have k_{cat}/K_M values well below the diffusion limit. For example, proline racemase, catalyzing the interconversion of D- and L-proline, is a dimeric protein with a subunit molecular weight of 38 000, a k_{cat} of approximately 1000 s^{-1}, and a K_M of approximately 3 mM.[37] The k_{cat}/K_M ratio is thus approximately 300 000. This number is high, but it is still 2.5 orders of magnitude below 10^8 M^{-1} s^{-1}, the diffusion limit as defined above. Similarly, chorismate mutase from *E. coli* has a k_{cat} of approximately 50 s^{-1} and a K_M of approximately 45 μM. The k_{cat}/K_M is approximately 10^6 M^{-1} s^{-1},[38,39] 2 orders of magnitude below the second-order diffusion limit.

There are several possible explanations for failures of k_{cat}/K_M to conform to the criterion defined above for catalytic perfection. These enzymes may not in fact have been perfected by natural selection. A diffusion-limited proline racemase may exist but be structurally so distant from wild-type proline racemases that selective pressures have not been able to find it.[5] Polypeptides able to catalyze the racemization of proline with $k_{cat}/K_M = 10^8$ M^{-1} s^{-1} may simply not exist. Catalytic perfection may not be the goal of natural selection for these enzymes. Kinetic optimality may have been sacrificed for other properties in the polypeptide chain, e.g., stability, regulation, or specificity.

Finally, there might be something wrong with the theory. Kinetic optimality may be incorrectly defined by a comparison of k_{cat}/K_M values with the second-order diffusion rate constant, or the magnitude of the second-order diffusion rate constant may have been incorrectly estimated.

Which explanation is correct influences our view of proteins and evolution in general. If the first explanation is correct, enzymatic behavior must reflect ancient historical accidents more than adaptation, as these accidents produced the first polypeptide sequences that later were optimized. If the second is correct, polypeptide catalysts must be intrinsically limited in what they can do. Certain behaviors could be optimized only at the expense of others, and the "mix" of behaviors observed in modern proteins would reflect "tradeoffs" between different behaviors in the protein.

For example, the K_M for proline racemase from *Clostridium sticklandii* (3 mM) is rather high considering the likely concentration of proline in the mi-

croorganism.[37] While the intracellular concentrations of proline in *Clostridium sticklandii* are not known, 30 μM is a reasonable guess. Thus, one might attribute the low k_{cat}/K_M to a high K_M and suggest that the enzyme may have been unable to evolve to bind proline tighter. This implies certain limitations on polypeptide catalysts that might manifest themselves elsewhere in different classes of enzymes.

However, it is most likely that the diffusion limit may be substantially lower than 10^8 M^{-1} s^{-1}, at least under the relevant physiological conditions. The second-order diffusion rate constant is dependent on several environmental factors, most notably viscosity. The value of 10^8 M^{-1} s^{-1} applies only in pure water, and biological water shows very little resemblance to pure water.[40-42] For example, the viscosity of blood (a relatively dilute biological fluid) is approximately 5 times that of pure water.[43] The viscosity of a solution of hemoglobin (30 mg/mL) is 8 times that of water.[44] The bulk viscosity or "macroviscosity" of cytoplasm has been measured in a variety of cells, including eggs from Spisula (sea clam) (4.3 cP)[45] and amoeba (3–11 cP).[46] These values are considerably higher than that of pure water (approximately 1 cP).

However, these measurements refer to "macroviscosity", the resistance to motion felt by macroscopic particles in a fluid. For the rate at which substrate molecules find an enzyme, a more appropriate measure is "microviscosity". Microviscosity is also higher in biological fluids, cytoplasm in particular, than in pure water. For example, the self-diffusion of water protons in yeast cytoplasm is 3–4-fold slower than in water.[47] As the diffusion of protons involves the "handing on" of protons from one water molecule to the next, and not the physical movement of protons, diffusion of moderately sized substrates inside cells is likely to be much slower.

Indeed, estimates of microviscosity based on correlation times of electron spin resonance probes show that microviscosity inside cells is considerably higher than in pure water.[48] In human embryonic lung cells, pole bean root tips, and Chlamydomonas, correlation times were approximately 100 times longer than in pure water. In yeast and *E. coli*, correlation times were 10 times longer.[48]

Thus, the second-order diffusion rate constant describing the encounter between enzymes and their substrates is expected to be 1–2 orders of magnitude smaller inside cells than in pure water. Thus, the physiologically relevant diffusion rate constant will also be lower in cells than in pure water. This implies that the diffusion limit on k_{cat}/K_M values of enzymes operating in these media should be on the order of 10^6–10^7 M^{-1} s^{-1}, rather than 10^8 M^{-1} s^{-1}. These lower values are consistent with k_{cat}/K_M values of "slow" enzymes such as proline racemase and chorismate mutase. In this view, these "low" values do not indicate that these enzymes are kinetically suboptimal but rather that they are optimized for viscous biological media.

This explanation creates another apparent paradox. To the extent that lower values for the diffusion rate constant in biological fluids are consistent with the *low* k_{cat}/K_M values for enzymes such as proline racemase and chorismate mutase, they are inconsistent with the very *high* values for k_{cat}/K_M for carbonic anhydrase, acetylcholinesterase, triose phosphate isomerase, and fumarase. Carbonic anhydrase and acetylcholinesterase operate extracellularly in rather dilute biological fluids. Here, rates of diffusion are high and high values of k_{cat}/K_M might be expected. However, both triose phosphate isomerase and fumarase are intracellular, operating in media with high viscosity and therefore low diffusion rates. Their k_{cat}/K_M values of 10^8 M^{-1} s^{-1} are now 2 orders of magnitude larger than the physiologically relevant diffusion rate constant.

These abnormally high values of k_{cat}/K_M for these enzymes might be explained by the assumption that enzymes catalyzing glycolytic reactions occur physiologically as multienzyme aggregates. The rate of transfer of substrate molecules within these particles is higher than diffusion in viscous solutions.[49,50] While this is an attractive hypothesis, multienzyme aggregates are not often well documented experimentally, except perhaps in trypanosomes.[51]

Nevertheless, these considerations suggest a general hypothesis. Enzymes that operate in more dilute biological solutions (e.g., blood or other intercellular media) and enzymes that act in multienzyme aggregates by direct transfer of substrates are expected to have values of k_{cat}/K_M that are higher than enzymes that act on "free" substrates inside cells. The literature suggests that this hypothesis is not obviously false. Further, enzymes with in vitro k_{cat}/K_M values higher than 10^7 M^{-1} s^{-1} are expected to occur in vivo as multienzyme aggregates. Such a working hypothesis would have its greatest impact if it led to the discovery of direct evidence for physiologically important multienzyme aggregates.

Even when enzymes catalyzing a multienzyme sequence occur in a complex, the first substrate must diffuse freely to the complex through viscous medium. In glycolysis, that substrate is often glucose. The k_{cat}/K_M value for glucose (assuming saturating ATP) is well below the "diffusion" limit, consistent with this picture. The slowness of regulated enzymes at the beginning of major metabolic pathways is occasionally interpreted as evidence that regulation and speed are intrinsically incompatible. This discussion suggests an alternative interpretation: the first step in a pathway must obtain its substrate via free diffusion in a viscous medium rather than via intracomplex transfer.

This hypothesis is presented to encourage further experimentation. Whether valid or not, the fact remains that enzymes with k_{cat}/K_M values in the range of 10^6–10^7 M^{-1} s^{-1} are not necessarily catalytically suboptimal. Indeed, as discussed below, other work suggests that enzymes that are suboptimal by the k_{cat}/K_M criterion nevertheless display finely tuned kinetic behavior.[5,6,52] However, because of uncertainty in the magnitude of the relevant second-order diffusion rate constant, we are unable to draw conclusions here about the ultimate degree to which k_{cat}/K_M values are finely tuned. The level of uncertainty again is about an order of magnitude.

D. Analysis of Microscopic Rate Constants

The kinetic parameters k_{cat} and K_M are algebraic aggregates of microscopic rate constants associated with microscopic reaction steps. Kinetic constants describe the rates of binding of substrates, the rates of release

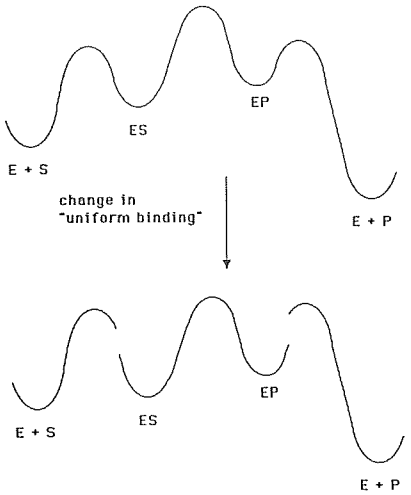

Figure 2. Changes in "uniform binding" reflect a perturbation in the interaction between the enzyme and a region of the substrate that is not undergoing reaction. Thus, the free energies of all bound states move "uniformly" up or down.

of products, the rates of steps in which chemical reactions occur, the rates of conformational changes, and the rates for binding or releasing protons and solvent molecules. In enzymes acting on multiple substrates, kinetic behavior can be quite complicated. The order in which different substrates bind is variable, with ordered and random mechanisms being only limiting extremes in a range of mixed mechanisms. Finally, in regulated enzymes or in enzymes with multiple subunits, microscopic rate constants can reflect the extent to which an allosteric effector occupies the regulatory site and the occupancy of other active sites in a multisubunit enzyme.

Knowles, Albery, and their co-workers made one of the first attempts to address the issue of selectable function at the level of microscopic kinetic parameters in their classic studies on triose phosphate isomerase.[9] First, microscopic rate constants for individual steps in the enzyme-catalyzed isomerization reaction were measured. Then, in a seminal discussion, Albery and Knowles interpreted the detailed kinetic behavior in terms of selected function.[53]

The Albery–Knowles analysis started with the "obvious position" that an optimal enzyme is the one that mediates the fastest flux of substrates to products. An algebraic expression that described part of this flux in terms of microscopic rate constants was proposed as an "efficiency function". The efficiency function was then evaluated and compared with the diffusion limited rate. This comparison provided a value of 0.6 as a quantitative measure of the efficiency of triose phosphate isomerase.

The Albery–Knowles analysis was the first to systematically address ways in which mutations in the protein sequence might alter the free energy profile for the reaction catalyzed to increase the efficiency of the enzyme. The evolution of catalytic efficiency of an enzyme was viewed as proceeding in three steps, each involving mutations of the protein that alter only some of the microscopic rate constants. The "easiest" way to improve efficiency was argued to be the accumulation of mutations that move the free energies of all bound states uniformly. These were called changes in "uniform binding" (Figure 2).

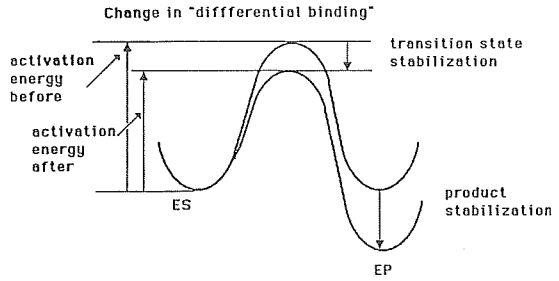

Figure 3. According to the "differential binding" mechanism, an enzyme can increase the rate of a chemical step by changing the internal equilibrium constant between two bound species. By lowering the free energy of the EP state, the free energy of the transition state is lowered relative to the free energy of the ES state. Hence, the reaction goes faster. The relationship between the rate of the reaction and the equilibrium constant for the reaction is described by a linear free energy relationship: k = constant $\times K_{eq}^{\beta}$.

The next easiest way to improve catalytic performance is by accumulating mutations that alter the relative free energies of bound products and bound reactants. These were called "differential binding" mutations (Figure 3). The influence of differential binding on the activation energy for the reaction interconverting bound substrate and product was assumed to be a linear function of the free energy difference between bound substrates and bound reactants, formally represented by a linear free energy relationship (LFER).

Finally, the most difficult method for increasing catalytic efficiency involved mutations that influenced only the energy of the transition state for the reaction relative to the energy of bound reactants and products. These mutations produced "catalysis of the elementary step".

This view of evolution is an approximation, with the concepts of "easy" and "difficult" referring to the fraction of mutations in a particular protein that influence the overall kinetic behavior of the protein favorably in each of the three ways. Few argue that the actual evolution of the kinetic behavior of proteins can always be separated cleanly into these different steps. In fact, Hall, Sinnott, and their co-workers in their study of the evolution of β-galactosidase found essentially no separation of the three processes.[54-56] However, the picture does not seem to be a misleading one. For example, Fersht and his co-workers[57a] in their work modifying the structure of a fragment of aminoacyl-tRNA synthetase found a set of mutations with free energy profiles altered approximately as expected using a differential binding model, with the activation energy for the chemical step linearly dependent on the free energy difference between ES and EP complex. Further, a second set of mutations appeared to influence only the energy of the catalytic step. Analogously, Strauss et al.,[58] working with triose phosphate isomerase, found a mutation that altered primarily the energy of the transition state.

If one assumes that these workers made mutations approximately "randomly", one might conclude that differential binding mutants arise approximately twice as frequently as mutants catalyzing the elementary step. Thus changes in "catalysis of the elementary step" would be viewed as only slightly less "easy" than changes in "differential binding". Of course, these mutants were not made randomly, and the ratio of

mutants of the two types produced randomly could be quite different. Indeed, the different conclusions drawn from the experiments of Li et al., Fersht et al., and Strauss et al. may simply reflect the fact that the last two deliberately chose mutations with the goal of finding some that influenced the free energy profile by "differential binding" and by "catalysis of the elementary step". If so, this makes a quite positive statement about the ability of chemists to anticipate how deliberately introduced mutations will influence enzymatic behavior.

In this context, it is important to comment on the recent exchange between Estell and Fersht over statistical methods used to determine whether kinetic data from mutant proteins conform to a LFER.[57b] Normally, the logarithms of the rate constants for mutant enzymes are plotted against the logarithms of the corresponding equilibrium constants. If the points fall along a straight line, a LFER is proposed.

It is well-known that such logarithmic plots can make data appear linearly correlated even when the correlation is quite poor. This is especially true for LFERs when variation in the forward rate constant (the one plotted) is large compared to variation in the reverse rate constant, a fact that is apparent by inspection of the linear free energy equation:

$$\log k_1 = \beta \log K_{eq} + Q$$

which can be rewritten

$$\log k_1 = \beta \log k_1 - \beta \log k_{-1} + Q$$

If variation in k_1 is greater than variation in k_{-1}, variation in the term containing k_1 will dominate the right-hand side of the equation, and a line (with a slope of 1) will be achieved, even with random data. This problem is particularly acute in arguments for or against LFERs in enzymes, as more scatter and lower correlation coefficients are normally tolerated here than in other kinetic discussions.

Estell's comment raised the scientific question: Do the data reported by Fersht obey a LFER with acceptable statistics?[57b] This question was addressed by Fersht by replotting log k_1 vs log k_{-1}. These values should also correlate linearly if the data conform to a linear free energy relationship, and the correlation is significant if the data include enzymes with a range of equilibrium constants. This replot arises from a simple transformation of the linear free energy equation itself:

$$\log k_1 = Q/(1 - \beta) - \beta(\log k_{-1})/(1 - \beta)$$

The data collected by Fersht (with one exception) lie on a line in the replot.

Much of this discussion has precedent in the physical organic literature.[57c,d] Indeed, it seems to be the fate of bioorganic chemists to reexplore most of the controversies of classical physical organic chemistry and reinvent most of the solutions developed by workers in this field several decades ago. A reader hoping to cut short this process should consider several paradoxes involved in linear free energy relationships (including "isokinetic points", problems of correlating enthalpy and entropy components of the free energy terms, and the fact that LFERs over wide ranges of K_{eq} are expected to be curved)[57c] and excellent articles explaining the conditions under which LFERs are expected to arise.[57d]

E. Matched Internal Thermodynamics

One widely cited conclusion from the analysis of Albery and Knowles is that the "internal equilibrium constant" (K_{int}, the equilibrium constant between enzyme-bound substrates and products, ES and EP) for enzymes optimized with respect to uniform and differential binding is unity, regardless of the magnitude of the "external equilibrium constant" (that between substrates and products in solution). In simpler terms, "perfect" enzymes, those that have evolved to have the highest rate of turnover, bind substrate and product so that the bound complexes have equal free energies.[9]

Although this criterion for perfection was originally intended to apply only to enzymes that had been optimized with respect to "uniform binding" and "differential binding", an internal equilibrium constant of unity is now widely regarded as a general criterion for kinetic perfection.[59-63] In its evolved form, Knowles and his co-workers stated the criterion simply and elegantly: "On theoretical grounds, an enzyme catalyst [will] be most efficient if 'the kinetically significant transition state is flanked by kinetically significant intermediates of equal free energy.'"[64]) For some observers, the suggestion has attained the status of principle.[65,66]

In the original argument, matched internal thermodynamics was regarded as most likely in enzymes not performing regulatory functions or other roles that might require sacrifice of kinetic optimality. Nevertheless, the idea suggested another more subtle possibility for assessing the kinetic optimality of an enzyme. To the extent that an internal equilibrium constant can be measured, it can be compared with the number 1. The closer to 1 the measured equilibrium constant is, the more optimal the enzyme is.

The internal equilibrium constants of many enzymes are close to unity, but only a few are very close.[64] For example, the internal equilibrium constant of triose phosphate isomerase is reported to be approximately 1.[9] Cohn, Rose, and their co-workers reported internal equilibrium constants close to unity in several phosphoryl transferases, including creatine kinase and hexokinase.[67-69] The second case is especially interesting, as it is an enzyme that catalyzes a reaction that is far "downhill" energetically and that operates physiologically far from equilibrium.

These predictive successes prompted the use of the notion of "matched internal thermodynamics" to formulate functional hypotheses explaining the stereoselectivity of enzymatic reactions[70,71] and to analyze the structure of proteins involved in electron transfer.[72]

While the energies of enzyme–substrate (ES) and enzyme–product (EP) complexes are often more similar than the energies of substrate (S) and product (P) free in solution for many enzymes, many internal equilibrium constants are not unity within reported experimental error. For example, the internal equilibrium constant for pyruvate kinase was reported as 10:1 to 15:1 in favor of the enzyme–ATP–pyruvate complex.[73] The internal equilibrium constant for fructose bis-

phosphatase appears to be far from unity.[74] The internal equilibrium constant for various dehydrogenases ranges from 1:1 to 5:1.[75-77]

Again, there are several possible explanations for the failure of these enzymes to meet this criterion for optimality. These particular enzymes may not have been perfected by natural selection. Catalytic optimality may be sacrificed in these enzymes for other desirable enzymatic properties. There may be something wrong with the theory. Matched internal states might be optimal only for certain enzymes. Matched internal thermodynamics may not in fact be the optimal free energy profile for catalytically optimal enzymes.

Recent experimental data suggest that the last explanation is most likely. The notion that "matched internal thermodynamics" are expected as the consequence of optimization of an enzyme for catalytic efficiency was first questioned,[70] then challenged,[78] then defended,[72] and ultimately revised.[73] To replace it, the "decreasing staircase" model was proposed.[73] According to this model, the internal equilibrium constant of a kinetically optimized enzyme should reflect a portion of the external drop in chemical potential driving the reaction catalyzed by the enzyme, under the physiological conditions for which the enzyme is adapted.[73] In simpler language, the internal equilibrium constant should be "downhill" in the direction of physiological flux. A detailed algebraic argument description of the descending staircase model was recently presented.[79]

The predictions of this model are quantitatively complex, as the precise magnitude of the internal equilibrium constant for a kinetically optimized enzyme reflects the free energies of substrates and products, the degree to which substrates and products are out of equilibrium under physiological conditions, the extent to which the chemical step is rate determining, and the nature of the linear free energy relationship correlating the rates and equilibrium constants of internal chemical reactions. The closer to equilibrium the metabolic step is under physiological conditions, the less rate limiting the chemical step is, and the smaller the constant β describing the linear free energy relationship between the energy of activation of the reaction and the internal equilibrium constant is, the closer the optimal internal equilibrium constant will be to unity.[79]

Experimental data appear to be qualitatively consistent with predictions made by the model. For example, the experimental fact that the internal equilibrium constant is close to unity for many enzymes is explained as a consequence of two general facts: (a) the step involving the chemical reaction is not fully rate limiting in many enzymes, and (b) many enzymes operate physiologically under conditions where their substrates and products are close to equilibrium. Conversely, enzymes catalyzing reactions far from equilibrium (e.g., fructose bisphosphatase) and enzymes having chemical steps that are rate limiting (e.g., pyruvate kinase) should have internal equilibrium constants different from unity, with the internal equilibrium constant favoring the enzyme–product complex.[79]

Here again, quantitatively precise predictions require data for kinetic parameters that are rarely known with precision. This again frustrates most efforts to determine the degree to which the internal equilibrium constant has been "fine tuned" by natural selection. However, the model makes an interesting qualitative prediction that, upon experimental examination, suggests that the internal equilibrium constant is tuned to within a factor of 2 of the value predicted by the model to be optimal.

The prediction concerns isozymes that have evolved to catalyze the same reaction in opposite directions. As the descending staircase model qualitatively predicts that the internal equilibrium constant will be downhill in the direction of metabolic flux, the model predicts that one substrate–enzyme complex will be favored in one isozyme, while the other will be favored in the other isozyme.[79]

This expectation is consistent with experimental results obtained on isozymes of lactate dehydrogenase. Lactate dehydrogenase from muscle catalyzes the conversion of pyruvate to lactate; lactate dehydrogenase from heart catalyzes the reverse reaction, the conversion of lactate to pyruvate.[80] Consistent with a qualitative expectation based on the model, the lactate dehydrogenase from muscle is reported to have an internal equilibrium constant of approximately 4 favoring the enzyme–lactate–NAD^+ ternary complex.[76] In contrast, the lactate dehydrogenase from heart is reported to have an internal equilibrium constant for lactate dehydrogenase from approximately 3 in favor of the enzyme–pyruvate–NADH ternary complex.[77] While reservations exist regarding experimental error in these studies, qualitatively similar results are obtained with different methods that allow direct comparison of the values measured for the two isozymes.[81]

The isozymes of lactate dehydrogenase are homologous.[82] For example, in pig, the sequences of the two isozymes are 75% identical.[83] The difference in the reported equilibrium constants (a factor of 9) is less than a kilocalorie per mole; the impact of the difference on the overall kinetic performance of the isozymes is correspondingly less. However, the fact that the difference is anticipated by a functional model of enzymatic behavior suggests that this difference is selected, not drifting. While more cases must be examined to produce a statistically valid argument, these results suggest that natural selection can detect a mutation that alters a kinetic parameter with an energetic consequence of less than a kilocalorie per mole, even when operating in complex organisms such as mammals.

Thus, both theory and experiment suggest that an internal equilibrium constant of unity is not required for a kinetically optimized enzyme. Instead, the optimal internal equilibrium constant is expected to reflect some of the drop in chemical potential between substrate and products under physiological conditions, with the overall free energy profile resembling a descending staircase. Predictions made by the descending staircase model are consistent with the experimental data available so far, and it appears that the internal equilibrium constants in enzymes are optimized well within a factor of 10.

III. Evidence for Fine Tuning

Despite this progress toward defining the degree to which natural selection can tune the kinetic behavior of enzymes, many biochemists continue to doubt that the enterprise will be successful. Some biochemists simply refuse to believe that natural selection can

TABLE I. Sequences of Alcohol Dehydrogenase Isozymes, Hybrids, and Point Mutants[a]

enzyme	residue 8	9	15	20	30	58	127	147	168	173	204	211	213	229	232	236	242	259	265	270	277	283	324	338
Expressed from Adh 1 Promoter																								
ADH I	G	V	H	Y	A	V	Q	Q	M	V	E	R	I	G	L	D	V	V	T	M	C	Q	T	V
R211T	G	V	H	Y	A	V	Q	Q	M	V	E	R	I	G	L	D	V	V	T	M	C	Q	T	V
E204	G	V	H	Y	A	V	Q	Q	M	V	E	R	I	G	L	D	V	V	T	M	C	Q	T	V
i242/259ii	G	V	H	Y	A	V	Q	Q	M	V	E	R	I	G	L	D	V	C	V	L	S	H	S	A
i173/204ii	G	V	H	Y	A	V	Q	Q	M	V	P	T	L	S	V	N	I	C	V	L	S	H	S	A
Expressed from Adh 2 Promoter																								
ADH II	A	I	N	H	P	T	E	E	R	A	P	T	L	S	V	N	I	C	V	L	S	H	S	A
ii259/265i	A	I	N	H	P	T	E	E	R	A	P	T	L	S	V	N	I	C	T	M	C	Q	T	V
ii204/211i	A	I	N	H	P	T	E	E	R	A	P	R	I	G	L	D	V	V	T	M	C	Q	T	V

[a] Only those residues where Adh 1 and Adh 2 vary are indicated; the residues at other positions in the sequence are the same in all of the proteins.

control the fine kinetic behavior of enzymes, especially in large organisms. This view is supported by the fact that in many pathways, the concentrations of enzymes are quite large relative to the concentration of their substrates. This is taken as a suggestion that small perturbations in the kinetic performance of the enzyme have no impact on performance at the level of physiology.

Others might not object to the proposition that natural selection finely tunes kinetic behavior but instead argue that kinetic experiments in vitro can never demonstrate this fact, as they can never be performed under conditions that resemble those inside a living cell sufficiently to be interpretable. While these points are well taken, three sets of experimental results are worth mentioning as counterpoint to this pessimism.

A. Temperature Optima of Enzymes

The temperature at which an enzyme displays its maximal rate reflects the temperature dependence of kinetic parameters (k_{cat}, K_M) that are rather complex aggregates of many microscopic kinetic parameters, each of which itself varies with temperature. Thus, temperature optima certainly cannot, at present, be directly related to protein structure. Nevertheless, the temperature optimum of an enzyme is quite possibly a trait that directly influences survival and therefore may be finely tuned.

Several particularly elegant studies suggest that this is the case. For example, Place and Powers studied the lactate dehydrogenases from the fish *Fundulus heteroclitus*.[52] The enzyme is polymorphic. Two structural forms are found in fish, and the temperature optima of the two forms differ by 3 °C, as determined by kinetic studies in vitro. A survey of fish living at different latitudes off the North American Atlantic coast found that the relative abundance of the two isozymes of lactate dehydrogenase correlates with the temperature of the water in which the fish is found. This correlation suggests that the different temperature optima of the two isozymes measured in vitro have physiological significance and reflects adaptation to different environments.

Similar experiments have found that the temperature optima (measured in vitro) of lactate dehydrogenases from cow and fish correlate with the temperature in the environment from which the enzyme was isolated.[84] The reader is referred to other examples where similar arguments have been developed in detail.[85-90]

B. Failure To Observe Polymorphism at the Level of the Protein

Should natural selection be unable to distinguish between proteins with similar behaviors and/or structures, one expects to find many different structural forms of the protein in a population. This structural variation is termed "polymorphism" and is viewed as the first step in "neutral drift".

Kreitman recently searched for such polymorphism in the alcohol dehydrogenase from *Drosophila melanogaster*.[91] Several genes for this enzyme were cloned and sequenced from individual flies collected worldwide. A high level of polymorphism in the genes was observed; over 40 variants were observed in the coding region or intervening sequences of the gene. However, virtually all of the polymorphism at the level of the gene was silent at the level of the expressed protein (appearing in the third position of codons or in untranslated regions). Only a single variant protein was observed, a variant that is well-known and appears from its distribution in a population to be adaptive.

A statistical analysis suggests that a dozen or so variant proteins are expected in the population displaying the polymorphism observed at the level of the gene. The fact that these variant proteins were not observed suggests that they must have been removed by natural selection. This argues that natural selection is more capable of distinguishing between structural variation at the level of a protein than at the level of a gene. Further, the data suggest that essentially all point variation in alcohol dehydrogenase in *Drosophila* is selectively disadvantageous.

Unfortunately, it is difficult to obtain information about the kinetic behavior of mutant alcohol dehydrogenases that natural selection has removed from a population of *Drosophila*. However, the behaviors of the variant proteins, not the structures of the proteins per se, are the targets of natural selection. To actually obtain these proteins to determine the kinetic variability that was not tolerated, one must obtain the variants that are not present naturally. This requires the use of recombinant DNA technology.

C. Site-Directed Mutagenesis Studies

Saccharomyces cerevisiae contains two isozymes of alcohol dehydrogenase (Adh). The first (Adh I) is constitutive and catalyzes primarily the reduction of acetaldehyde to ethanol during fermentation. The second (Adh II) is induced in the presence of oxygen

TABLE II. Kinetic Properties, Expression Levels, Stability, and Influence upon Growth of Adh 1, Adh 2, Point Mutations, and Hybrids Introduced into a Strain of Yeast Lacking Chromosomal Adh 1 and Adh 2[a]

enzyme	promoter	spec act., U/mg	K_M, μM NAD+	K_M, μM NADH	$V_{max,red}/V_{max,ox}$	expression level, U/L of cells	act. after 12 h, %	dividing time, min
ADH I	Adh I	476	224 (12)	135 (9)	3.0	450	75	160 ± 4 (4)
R211T	Adh I	501	245 (9)	157 (15)	2.9	830	75	256 ± 6 (6)
E204	Adh I	471	178 (7)	147 (12)	3.3	340	40	280 ± 31 (3)
i242/259ii	Adh I	523	371 (38)	118 (4)	2.6	280	85	218 ± 13 (5)
i173/204ii	Adh I	410	426 (66)	137 (21)	2.6	250	100	223 ± 5 (6)
ADH II	Adh II	158	165 (9)	56 (2)	3.7	220	100	224 ± 10 (5)
ii259/265i	Adh II	146	148 (7)	67 (4)	4.1	185	100	236 ± 6 (5)
ii204/211i	Adh II	129	66 (5)	73 (6)	6.3	570	90	215 ± 6 (5)

[a] Specific activities are for the oxidation of ethanol (330 mM) at pH 8.4 and 20 °C and are accurate to ±5%. K_M determinations were made by Lineweaver–Burk plot, and standard errors are given in parentheses. Expression levels were determined from enzymatic activities in crude extracts of 100 mL of cells grown to OD 0.9–1.0 in minimal medium with histidine, lysine, and arginine as supplements. Activity was measured in initial cell lysate again after 12 h at 4 °C to estimate stability in vivo. Dividing times were for growth in the same medium at 30 °C (10-L culture, 2 L min^{-1} of air, 100 rpm agitation). Numbers in parentheses are the number of data points used in the determination.

and catalyzes the oxidation of ethanol. The amino acid sequences of Adh I and II are different at only 24 of the 375 amino acid positions (Table I).

Most of the amino acid differences between Adh I and II lie on the surface of the protein and are "conservative". Naively, these appear to be the products of neutral drift. However, the two isozymes have somewhat different kinetic behaviors (Table II), and at least some of these differences are expected to be adaptive in two proteins catalyzing one reaction in opposite directions (vide supra). Therefore, some of the substitutions must be responsible for the differences in kinetic behavior and should also be adaptive.

To estimate what fraction of these structural variations are likely to be adaptive, Ellington constructed genes for eight hybrids of Adh I and II using recombinant DNA methods.[81] These were of two types. In four mutants, a single amino acid substitution was introduced. In four others, hybrids were constructed that were part Adh I and part Adh II. The constructed genes were cloned via a plasmid into a yeast strain in which the genes for both isozymes of Adh had been disrupted.

Ellington then expressed, purified, and subjected to in vitro kinetic analysis the hybrid proteins. Further, he estimated the stability and level of expression of the proteins. Finally, the doubling times of the eight strains of yeast containing plasmids with the mutant Adh genes as the only alcohol dehydrogenase genes available were measured (Table II).

As is apparent from Table II, native Adh I conferred the greatest survival value; yeast containing all non-wild type Adh genes grew slower. Surprisingly, this included yeast with genes for Adh proteins that were largely indistinguishable from wild type Adh by standard kinetic and physical analysis.[81] Subject to caveats (the environment in which the yeast was grown was artificial, the mutant genes were not chromosomal, and the rate of growth is highly sensitive to precise conditions), it appears that natural selection was better at detecting subtle variations in the behavior of proteins than was the biochemist using standard in vitro assay methods.

Of course, the differential impact on growth rate of the various mutants proteins might arise from an unnoticed (and undesired) impact of the mutation on an unexamined property of the protein. Nevertheless, the influence of subtle structural variation in a single protein on the rate of growth of yeast is worthy of note, as similarly subtle perturbations in behavior are expected in the mutants of Adh from *Drosophila melanogaster* mentioned above that were removed from a population by natural selection.

D. Conclusions

The quantitative nature of kinetic measurements makes them useful as criteria for evaluating the extent of functional adaptation of enzymes. The experimental biochemist now has several such criteria. Theoretical and experimental ambiguities notwithstanding, direct application of these criteria to data obtained from natural enzymes suggests that natural selection is capable of tuning kinetic behavior quite finely, perhaps to within a factor of 2, when judged against standards provided by theoretical models of kinetic optimality.

IV. Enzyme Economics

There remain other difficulties with simple models of "catalytic efficiency" as the starting point for experiments to evaluate the evolutionary optimization of enzymes. Such models are based on the assumption (often unstated) that enzymes with faster microscopic velocities are selectively better than enzymes with slower microscopic velocities. Albery and Knowles perhaps stated most clearly that "we start from the obvious position that an efficient enzyme will mediate a high flux of substrate to product".[9]

This position is not obvious. Indeed, the common focus of the biochemist on external constraints (diffusion limits and substrate concentration in particular) in the evolution of enzymes has meant that other factors intrinsic in the design and synthesis of enzymes have been overlooked. A more complex argument involving the cost of the catalysts themselves forces the conclusion that an enzyme that mediates a high flux of substrate to product is not necessarily better than one with a lower flux, even if we assume that there are no other desirable enzymatic traits (e.g., regulation, specificity, and stability; vide infra) for which catalytic optimality must be sacrificed.

A. Catalytic Efficiency and the Minimization of Biosynthetic Cost

We have noted (vide supra) that a mutation that lowers the catalytic activity of an enzyme can be compensated by a regulatory mutant that increases the

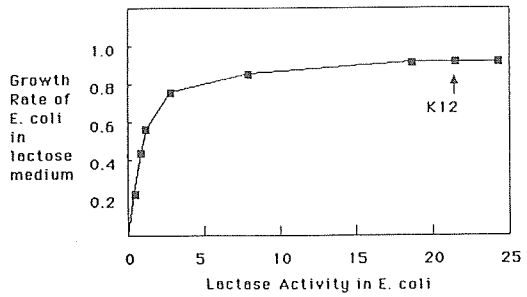

Figure 4. The growth rate of *E. coli* when lactose is present as a sole carbon source appears to be a function of the total lactase activity rather than a function of the microscopic rate constants of individual β-galactosidase enzyme molecules.

number of enzyme molecules. Selective pressures are almost certainly sensitive to the total flux of substrate to product at the level of the metabolic step, not the microscopic flux through a single enzyme molecule. This presumption is supported by a variety of data. For example (Figure 4), the rate of growth of *E. coli* under conditions where lactose is the sole carbon source is a function of the total β-galactosidase activity, not of the microscopic catalytic efficiency of the enzyme.[92]

Consider a biological system wishing to catalyze the conversion of 1000 molecules per second of dihydroxyacetone phosphate to glyceraldehyde phosphate. It must select a triose phosphate isomerase that will perform the task optimally. The organism is prepared to dedicate 10 000 of its amino acids to this endeavor. The organism must chose between two options.

In the first option, the organism assembles the amino acids into 10 enzyme molecules, each with 1000 amino acids and each turning over 100 substrate molecules per second, a velocity corresponding to the diffusion limit under physiological conditions. In the second, all things are equal, except that the organism assembles 100 enzyme molecules, each with 100 amino acids and each turning over 10 substrate molecules per second. These enzymes therefore catalyze the same reaction, but at one-tenth the diffusion limit.

The bioorganic chemist, assuming that the better enzyme is the one with the value for k_{cat}/K_M nearer the diffusion limit, would conclude that the first option provides enzymes nearer to evolutionary perfection by a factor of 10. Yet these two options produce the same total flux of substrate to product at the same apparent "cost", at least in terms of total amino acids required for biosynthesis. From a physiological point of view, it is hard to see any reason why natural selection would prefer one option over the other. In language borrowed from economics, the organism is "indifferent" to the choice.

Of course, if the enzymes with velocities of 10 s^{-1} could be made with only 90 amino acids instead of 100, 100 of these slower enzymes would produce the same net flux, but the organism would have 1000 amino acid molecules left over to put to some other use. The organism would no longer be "indifferent"; it would prefer slower enzymes to those that catalyze the reaction at the diffusion limit.

This analysis permits an important conclusion. Given the fact that different enzymes with different sizes have different biosynthetic costs, it is conceivable that a kinetically optimal enzyme could mediate a high flux, an intermediary flux, or a slow flux and still be the

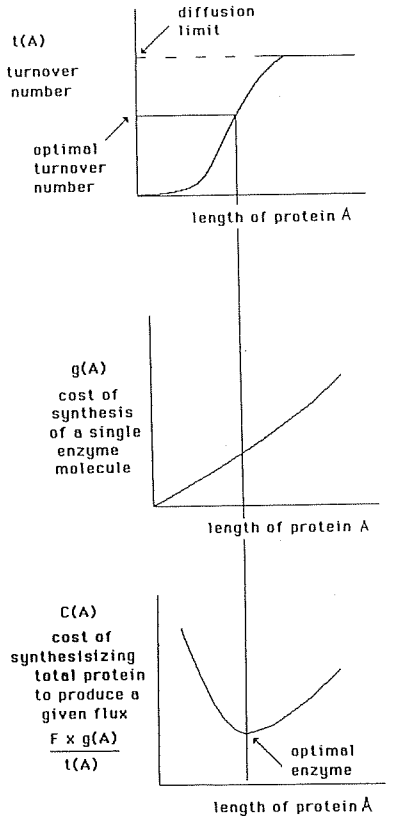

Figure 5. Turnover number ($t(A)$) is expected to be a function of the structure of a protein. This is a behavioral surface in n dimensions, where n is the number of amino acid residues in the protein. To simplify the analysis, we show only two dimensions in plot a; turnover number is shown as an increasing function of the length of the protein until the point where the diffusion limit is reached. Likewise, the cost of biosynthesis of a protein molecule is a function of the length of the protein (plot b). Convoluting these two curves gives plot c, where the cost of synthesizing the protein needed to produce the desired flux is shown as a function of the length of the protein, A. The optimal enzyme is the one that produces the desired flux at the minimal cost. In this case, this protein has a turnover number approximately half of the diffusion limit.

optimal solution to a catalytic problem. Enzymes catalyzing reactions at rates far slower than the diffusion limit need not be viewed as kinetically suboptimal.

Some formalism makes this point more clearly. Figure 5a shows a plot (without the axes scaled) of the catalytic power (k_{cat}/K_M) of a hypothetical enzyme with n amino acids as a function of n. Of course, each protein of n amino acids has a complete surface in n dimensions relating structure to velocity under physiological conditions. Figure 5a merely plots the value of the velocity under physiological conditions of the fastest enzyme on that surface. This relationship is a function, turnover vs length, $t(A)$.

We do not know the scale on the axes nor the precise shape of the curve (vide infra). However, the general shape of the curve expected is clear. For short polypeptides, turnover numbers should be small. For example, the best "*p*-nitrophenyl esterase" with only 1 amino acid ($n = 1$) might well have the structure $(His)_1$. However, with longer chain length, the catalytic power is expected to increase. At some point, it would reach the diffusion limit and would become no faster even with increasing size.

However, larger enzymes are more "expensive", considering both biosynthesis at the ribosome and the

biosynthesis of the constituent amino acids. Figure 5b is a plot (again without scales) of the cost of synthesizing a single enzyme molecule as a function of its length, $g(A)$. Again, the scale is difficult to predict, but the general shape of the curve is not. The cost of synthesizing a protein rises with its length.

These two curves can be combined. The number of enzyme molecules E required to produce the desired flux F is the flux divided by the flux per enzyme molecule.

$$E = F/t(A)$$

The cost C of producing the number of enzyme molecules necessary to produce this flux is the product of the cost per enzyme molecule and the number of enzyme molecules needed:

$$C = Eg(A) = (Fg(A))/t(A)$$

This equation is plotted in Figure 5c. The abscissa of the minimum, where the first derivative

$$dC/dA = 0$$

represents the length of the peptide that catalyzes the desired flux at the minimum cost, is the optimum sought by natural selection. Analytically, this minimum is given by the expression

$$t(A)\frac{d}{dA}[g(A)] = g(A)\frac{d}{dA}[t(A)]$$

This analysis, except for subject matter, could be taken from an economics textbook. We do not know the precise shape and scales of the curves, nor are they important. All that is important is the recognition that the position of the stationary point in Figure 5c determines the optimal size and velocity for an enzyme. The shapes of the various functions in turn determine the position of this point. For functions $t(A)$ that rise more steeply, the optimal enzyme will be shorter and have higher microscopic rate constants than for enzymes optimized when the function $t(A)$ rises less steeply. Of course, with sufficiently steep functions $t(A)$, the microscopic velocity, the kinetically optimal enzyme may be at the diffusion limit.

We may now return to the premise that the evolutionary optimal enzyme is the one that has a k_{cat}/K_M value approaching the diffusion limit. This is in fact a premise not about kinetics but about the relationship between the shapes of the plots in Figure 6. In an analysis that concerns only the cost of the initial synthesis of a protein and the catalytic power of an enzyme, the following statement is true: *Assuming that the best enzymes are diffusion limited is logically equivalent to assuming that the slope of the $t(A)$ curve is steep relative to the slope of the $C(A)$ curve.* Conversely, if enzymes are found generally to be diffusion limited, we must conclude that diffusion-limited enzymes are generally not much more expensive than slower enzymes.

This point of view has been overlooked in the literature seeking to justify on evolutionary grounds the catalytic power of enzymes. Indeed, in the context of this argument, the focus on diffusion-limited enzymes seems rather parochial. Of greater relevance are how much protein material is required to construct a performing protein and what the marginal cost is for improving the enzyme's performance.

This picture provides some interesting insights into some details of enzymology. For example, methionine

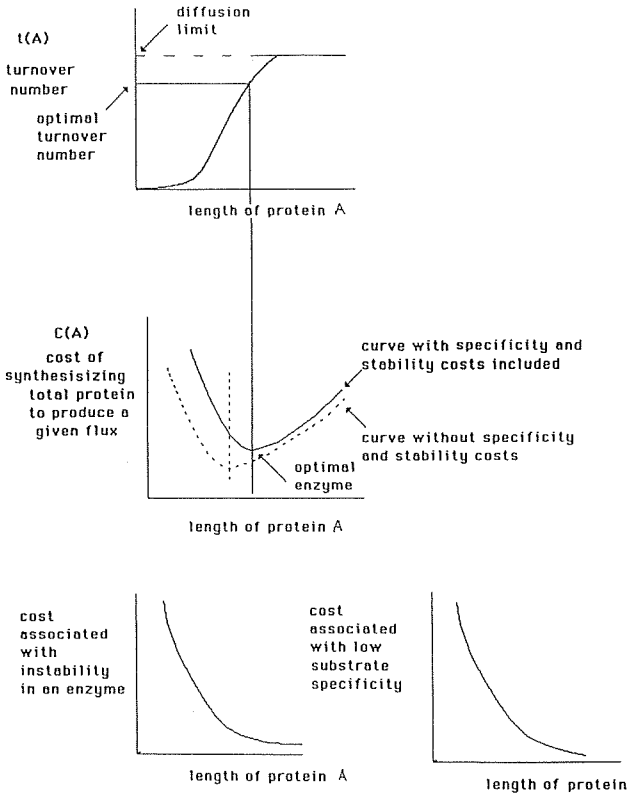

Figure 6. The total cost curve (Figure 5c) can be convoluted with any number of additional cost curves. We convolute it here with two, the "instability cost" and the "specificity cost". Small enzymes are expected to be less stable than large ones. Thus, the shape of the curve presumes that the cost of maintaining a small enzyme to produce a desired flux is more expensive than maintaining a larger enzyme, as the smaller enzyme must be resynthesized more often. Likewise, the shape of the curve suggests that a small enzyme that is nonspecific implies a cost as well. In both cases, the effect of the convolution is to shift the total cost curve so that the optimal protein has a higher molecular weight and a higher turnover. Of course, convolutions of other functions could have different effects on the nature of the kinetically optimal enzyme. Again, stability and specificity as a function of structure are described by behavioral surfaces in n dimensions. The representations here are simplified for the sake of clarity.

is synthesized in *E. coli* by the transfer of a methyl group from 5-methyltetrahydrofolate derivatives to homocysteine. Two enzymes capable of catalyzing this reaction are known in *E. coli*. One (E.C. 2.1.1.13) catalyzes the reaction with the aid of a vitamin B_{12} cofactor (and via a B_{12}–methyl intermediate). The second (E.C. 2.1.1.14) is induced by *E. coli* when grown in the absence of B_{12} and effects the same transformation without a B_{12} cofactor.

The fact that the B_{12} enzyme is used when *E. coli* has access to the vitamin implies that the B_{12} enzyme is the cheaper way to obtain the desired metabolic flux.[93-95] Indeed, the non-B_{12} enzyme is turned off in the presence of B_{12}. However, the B_{12} enzyme is larger and therefore costs more to synthesize. The molecular weights of the B_{12} and non-B_{12} enzymes appear to be 130 000 and 99 000, respectively.[94] Further, the B_{12} enzyme has a turnover number (in vitro under saturating conditions) 100 times larger than the non-B_{12} enzyme. Thus, on a cost basis, the B_{12} enzyme is 50 times better as a catalyst.

The methionine synthetase not dependent on B_{12} illustrates problems in biochemical adaptation faced by a relatively impotent enzyme. The non-B_{12} enzyme is

2 orders of magnitude slower than the B_{12} enzyme. To achieve the desired flux, the enzyme must comprise 5% of the soluble protein in *E. coli*.[93] Is this enzyme suboptimal? We simply cannot say. A polypeptide chain that catalyzes this reaction any faster may be impossible to construct. As 5-methyltetrahydrofolate is a poor alkylating reagent, this possibility might be chemically plausible. Natural selection may not (yet) have found the mutations in this protein that will make it a more efficient catalyst. Faster non-B_{12} enzymes might exist but may be much larger and therefore much more expensive, making the small, slower enzyme that is observed a more economical choice.

How then are we to evaluate experimentally the proposition that a single enzyme is kinetically optimized or, for that matter, the proposition that enzymes in general are kinetically optimized? There is no direct answer, as we cannot rule out the possibility with slow enzymes that they arose because they are cheaper unless we know the shape of cost curves sketched in Figure 5.

There are two ways in which the "difficulty" of a particular reaction can influence the shape of the $t(A)$ function and therefore the characteristics (size and kinetics) of an optimal enzyme. First, the reaction type can dictate the size that is needed before the curve describing k_{cat} as a function of size begins to rise. Enzymes catalyzing reactions on larger substrates or reactions requiring many functional groups are expected to need a larger protein "critical mass" before catalysis at a level expected for an enzyme is observed. In contrast, enzymes acting only on small substrates or requiring only one functional group are expected to require a smaller critical mass. Thus, the size of enzymes catalyzing a particular reaction might be expected to correlate with reaction type.

Such a picture is consistent with the fact that enzymes transferring phosphoryl groups between substrates with similar structures generally catalyze reactions via a phosphoryl enzyme intermediate and display ping-pong kinetics. Such a mechanism allows the enzyme to use a single binding site to catalyze the transfer between two substrates and thus economizes on protein.[97]

Inspection of data collected for enzymes catalyzing many reaction types allows an estimate of the size of a protein that is needed to catalyze a reaction of a particular type. Further, the shape of a plot $t(A)$ is expected to be a characteristic of the reaction type. It will be different for different reactions on different substrates, and its shape reflects the "difficulty" of catalyzing the reaction. For example, reactions that require complex dynamic movements to achieve, or where refinement of stereochemical positioning can be advantageous, may have a smaller slope. Enzymes catalyzing reactions that have stringent stereoelectronic requirements (e.g., methyltransferases) should have $t(A)$ functions with smaller slopes than enzymes catalyzing reactions with no stringent stereoelectronic requirements (e.g., removal of protons).

Thus, the plausibility of an argument that an enzyme is slow because faster enzymes are more expensive can often be evaluated. In particular, if smaller enzymes catalyzing similar reaction types are faster, the argument is not plausible. Unfortunately, such data are not available for the non-B_{12} methyltransferase mentioned above. However, the fact that many enzymes have k_{cat}/K_M values similar to the physiologically relevant diffusion rate constant (vide supra) means that in most cases, the $t(A)$ curve rises steeply after a critical mass of protein is obtained.

B. Other Constraints

Only one equation has been used so far to describe the cost of an enzyme as a function of length. Additional curves can be considered to describe the relationship between other enzymatic traits and the length (and expense) of the enzyme. Convoluting these new cost functions shifts the curve $C(A)$, and therefore shifts the optimal value of A, either to smaller (and slower) enzymes or to larger (and faster) enzymes.

These additional functions may also be represented as plots (again without scaled axes). Figure 6 shows a plot of "specificity cost" as a function of length. Presumably smaller enzymes with a given catalytic power are less specific than larger ones. This nonspecificity creates cost by creating nonspecific catalytic events. These must be corrected with the expenditure of energy. Presumably, this cost can be avoided by constructing a more expensive enzyme.

Figure 6 also shows a plot of the cost associated with stability as a function of length. Presumably, less stable enzymes must be resynthesized more frequently, and more expensive enzymes are more stable. Convoluting these functions shifts the curve to the right, favoring enzymes that are larger. Further, if specificity and stability are independent of rate, enzymes more closely approaching the diffusion limit are also favored by this optimization.

This discussion assumes that different enzymatic traits, including turnover number, stability, regulatability, and specificity, can vary independently. The only constraints that have been considered are the diffusion limit and the cost of biosynthesis of the protein. Should in fact all enzymes have rate constants approaching the diffusion limit, the notion that kinetic parameters are independent of other enzymatic behaviors would be supported.

C. Tradeoffs

So far, only external constraints (the second-order diffusion rate constant, physiological substrate concentrations) have been considered as possible limits on the catalytic power of enzymes. There remains the possibility, noted above and elsewhere,[9,21] that internal factors might also constrain the optimization of kinetic behavior. In particular, a single enzyme may be unable to be both fast and stable, fast and highly specific, or fast and regulatable. If this is the case, the behavior of a modern enzyme will reflect tradeoffs between different traits.

Evaluating functional and nonfunctional aspects of enzymatic behavior becomes more difficult if the traits being examined cannot vary independently, that is, if optimizing one requires sacrificing another. The notion that traits of enzymes may not be independently variable is an interesting one with widespread implications. At the very least, it implies that the job of the "enzyme engineer" will be more difficult. However, it fundamentally perturbs our view of polypeptides as powerful

molecules capable of remarkable things. Any suggestion that polypeptides are inherently limited is important, and a substantial body of data now provides evidence concerning such tradeoffs.

The literature contains more suggestions that catalytic performance is sacrificed for other selectable traits than can be reviewed here. Three sets of data serve to illustrate general difficulties that accompany such suggestions and provide support for the tentative conclusion that they are not generally correct.

1. Specificity vs Catalytic Efficiency

Tradeoffs between specificity and rate are often implied, and the literature contains two conflicting views. On one hand, high specificity and large values of k_{cat}/K_M are often regarded as being incompatible.[98] In this view, high specificity implies a tight binding of substrate, and tight binding of substrate is viewed as incompatible with rapid reaction rates. To some, this is because tight binding of substrate implies an interaction that stabilizes the reactant with respect to the transition state for a reaction. To others, an enzyme that binds substrate tightly also is likely to bind product tightly; the slow step then becomes product release. An analysis along these lines was developed by Cleland to argue that there was an upper limit on the forward rate of a reaction.[28]

On the other hand, high k_{cat}/K_M values are often regarded as incompatible with low specificity. According to this exactly contrary view, a high value of k_{cat} requires a precise alignment of binding and functional groups in the active site with respect to substrate and an active site can be finely "tuned" to only a single substrate. An enzyme that must act on several substrates must have an arrangement of functional groups in the active site that is a compromise among the optimal arrangements for each individual substrate. An enzyme that is a "Jack-of-all-trades" cannot be a master of any single one.

Data generally seem to be more consistent with the latter point of view. For example, the broadly specific alcohol dehydrogenase from horse liver has a considerably lower k_{cat}/K_M with all primary alcohols as substrates than the highly specific alcohol dehydrogenases from yeast with ethanol as a substrate. It is tempting to argue that the enzyme from liver has sacrificed kinetic power with one specific substrate in return for obtaining catalytic power against a wide range of substrates. Indeed, it is a somewhat general (but by no means universal) fact that in matched pairs of enzymes, the faster enzyme is also the more specific.

Such generalizations have many pitfalls. In many nonspecific enzymes (including liver alcohol dehydrogenase), the physiological role of the enzyme is not known, making it conceivable that the enzyme is a highly efficient catalyst for another (unknown) reaction. Also, enzymes from multicellular organisms (e.g., horse) generally have less catalytic power than the analogous enzymes from unicellular organism (e.g., yeast), making interspecies comparisons risky.

If substrate specificity is independent of other parameters, a hypothesis can be formulated regarding substrate specificities in general. Substrate specificity is expected to drift unless it is constrained by function. One apparent functional constraint arises from the fact that an enzyme should not catalyze a transformation that is metabolically undesired on a substrate that is physiologically present. In contrast, there is no selective pressure to prevent an enzyme from acting on a substrate that is not physiologically present.

Thus, selective pressures should constrain the substrate specificity of a galactose dehydrogenase from drifting to act on D-glucose. D-Glucose is present physiologically, and an enzyme that cannot oxidize galactose without also oxidizing glucose should cause metabolic confusion. However, natural selection has no reason to prevent the substrate specificity of galactose dehydrogenase from drifting to include D-talose. This sugar is not present metabolically; there would be no selectable disadvantage to a mutant galactose dehydrogenase that acted on this substrate.

Likewise, an L-threonine aldolase that operates physiologically to synthesize threonine from glycine and acetaldehyde almost certainly would synthesize L-threonine and not L-allothreonine. However, an L-threonine aldolase that operates physiologically to degrade threonine to form glycine and acetaldehyde may not be physiologically constrained so that it does not act on L-allothreonine. The threonine aldolases that have been studied in nature are probably of the latter type; they show some activity with allothreonine as a substrate.[99]

Ribulose diphosphate carboxylase (RuBisCo) is one enzyme where a strong case can be made that optimal catalytic efficiency and optimal substrate specificity are incompatible, and the properties of natural enzymes reflect a tradeoff between the two. In addition to catalyzing the reaction of carbon dioxide with ribulose diphosphate in carbon fixation in plants, RuBisCo catalyzes the reaction of oxygen with ribulose diphosphate. This side reaction destroys a molecular of ribulose diphosphate and is presumed to be metabolically wasteful.[100] Nevertheless, it appears as if the side reaction cannot be removed without making enzyme a less efficient catalyst of the primary reaction.

2. Stability vs Catalytic Efficiency

Suggestions of a tradeoff between stability and catalytic power in enzymes originated in observations that enzymes from thermophiles have k_{cat} values lower than those of analogous enzymes from mesophiles under similar assay conditions.[84-88] The increased stability of thermophilic enzymes has been assigned to increased numbers of internal hydrogen bonds, which in turn are presumed to decrease the flexibility of the enzyme. Flexibility is presumed to be important for catalytic power; hence the conclusion that stable enzymes are poorer catalysts.

If k_{cat} is taken to be an indicator of catalytic efficiency, then lactate dehydrogenases from thermophiles and mesophiles seem to confirm this suggestion. However, k_{cat}/K_M is a better criterion for catalytic efficiency. If this criterion is used, the enzymes from thermophiles have comparable (or higher) catalytic efficiencies, and the experimental evidence for a tradeoff between speed and stability disappears.[89]

Finally, in the highly homologous isozymes of yeast ethanol dehydrogenase, the stability of the enzyme from mitochondria to thermal inactivation is considerably greater than that for Adh I. Yet the k_{cat}/K_M values of

these enzymes are not substantially different.[101]

If stability is viewed as an independent behavior in an enzyme, a functional theory must be based on the impact of stability (or instability) on the survival of an organism. One functional theory for enzyme stability focuses on the physiological lifetime of the enzyme. Intracellular constitutive enzymes are expected to be more stable than intracellular inducible enzymes simply because constitutive enzymes are used for a longer time. An unstable constitutive enzyme must be resynthesized frequently; this is wasteful of metabolic energy and presumably implies selective disadvantage (vide supra). In contrast, inducible enzymes are likely to be degraded as soon as the inducing stimulus is removed to permit the constituent amino acids to be recycled.

Therefore, natural selection will deliberately select for enzymes that are unstable under physiological conditions, the degree of instability determining the rate at which the enzyme will be degraded. To the extent that thermal stability in vitro is correlated with stability in vivo, constitutive enzymes are expected to be thermally more stable than inducible ones. Of course, extracellular enzymes are expected to be the most stable; once excreted, the organism may lose the option to recycle their constituent amino acids.

The complexities of in vivo degradation of proteins make it unlikely that any simple relationship will exist between in vitro and in vivo stability. Nevertheless, data are sometimes consistent with expectations. For example, inducible and constitutive transaminases are both known in *E. coli*. The inducible enzyme is rapidly denatured at moderate temperatures; the constitutive enzyme is quite stable to thermal inactivation.[102] Several fumarases from obligate aerobes, including mammals, are considerably more stable than fumarases from facultative aerobes.[103] In higher organisms, stability and in vivo turnover appear correlated, and unstable enzymes seem to be selectively disadvantageous. For example, hemolytic anemias and certain neuromuscular dysfunction in humans appear to be associated with a thermolabile triose phosphate isomerase.[104]

Nevertheless, this generalization has many counterexamples. For example, the enzymes induced by yeast to degrade allantoin are reported to be extremely stable.[105] We cannot say at this point whether such exceptions suggest that the underlying premise is incorrect, that the enzymes in question are not limiting for survival, or merely reflect the fact that thermal stability is itself not a direct measure of the stability of an enzyme against degradation under physiological conditions.

3. Regulation vs Catalytic Efficiency

Theories regarding kinetic perfection have often been coupled to suggestions that regulation and kinetic perfection are incompatible in the same protein.[9] An interesting study along these lines looked for evidence that kinetic optimality was sacrificed in the interest of metabolic control in the enzyme methylenetetrahydrofolate reductase from pig liver (Figure 7).[106] This enzyme has a noncovalently bound flavin group, which is reduced by NADPH. In a second step, the enzyme-bound reduced flavin reduces methylenetetrahydrofolate to 5-methyltetrahydrofolate. Ping-pong Bi Bi

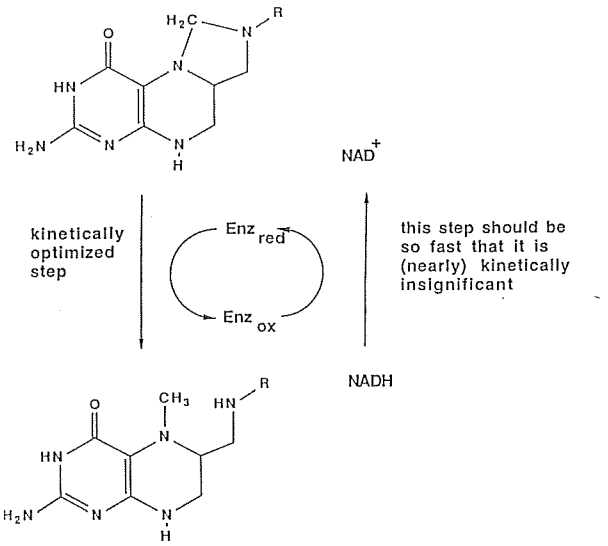

Figure 7. The reaction catalyzed by methylenetetrahydrofolate reductase proceeds in two steps. If the evolutionary goal is to maximize the rate of reduction of methylenetetrahydrofolate, the rate of the rereduction of the enzyme by NADH should be so fast as to be kinetically insignificant. This appears to be the case.

kinetics are observed, corresponding to the binding of NADPH and the release of $NADP^+$ in the first cycle that converts Enz_{ox} to Enz_{red} (where the subscripts indicate the redox state of the bound flavin), followed by the binding of methylenetetrahydrofolate and release of 5-methyltetrahydrofolate to reconvert Enz_{red} back to Enz_{ox}.

A detailed study of the microscopic rate constants for the reduction of the enzyme by NADPH suggested that the rate of turnover of Enz_{ox}–NADPH to Enz_{red}–$NADP^+$ was at least 29-fold slower than the rate at which NADPH was expected to diffuse to the active site at the estimated physiological concentrations of NADPH. The fact that the catalytic step was slower than the diffusion step led the authors to conclude that "the reductive half-reaction catalyzed by methylenetetrahydrofolate is not optimized for catalytic efficiency". It was further proposed that this inefficiency may reflect the fact that the enzyme is a point of regulation; *S*-adenosylmethionine is a feedback inhibitor.

The conclusion is worth examining in some detail. The theoretical analysis is again somewhat encumbered by the limitations noted in the discussion above, which applies to single substrate–single product enzymes. For ping-pong Bi Bi enzymes, where there are effectively two separate reactions, the rates of both reactions are critical. The Albery–Knowles analysis (vide supra) argues that the rates of each of the consecutive steps under physiological conditions are approximately equal. However, for this enzyme, it is more likely that the rate reduction of methylenetetrahydrofolate is under selective pressure, and it is the rate of this step that must be optimized. This evolutionary goal is best achieved by making the rate of the other step, the rereduction of the enzyme, so fast as to be kinetically insignificant.

The kinetics of the enzyme suggest nothing inconsistent with this analysis. The k_{cat} for the reduction of methylenetetrahydrofolate by Enz_{red} is approximately 50 s^{-1}, while K_M is approximately 2×10^{-5} M. Thus, the k_{cat}/K_M value for the "critical" reaction, 2.5×10^6 M^{-1} s^{-1} is within the range expected for a kinetically

optimized enzyme considering the discussion above.[107]

The k_{cat} for the conversion of Enz_{ox} to Enz_{red} by NADPH is approximately 160 s^{-1}. At the presumed physiological concentrations of NADPH, this step is considerably faster than the rate of reduction of methylenetetrahydrofolate. Indeed, it is so much faster that the rate of rereduction of the enzyme has nearly no impact on the rate of the overall reaction (Figure 7). If the rate of rereduction were infinite, the overall rate would only increase by a factor of approximately 1.4. This corresponds to a statement that the kinetics of reduction of the enzyme by NADPH are nearly optimized, assuming that it is the rate reduction of methylenetetrahydrofolate that is being optimized.

Vanoni and Matthews[106] recognized that the rereduction of the enzyme was much faster than the critical reaction, the reduction of methylenetetrahydrofolate. They wrote that "kinetic insulation" of some steps in ping-pong Bi Bi enzymes may occur generally. This is most likely correct. In a ping-pong Bi Bi reaction where only one of the steps is physiologically important, the other step will have a rate that is so fast as to have a small impact on the rate of the important step. Here, methylenetetrahydrofolate reductase is almost certainly not optimized to maximize the flux for the oxidation of NADPH. Indeed, the kinetics of oxidation of NADPH are on the borderline of optimality. One might argue that Vanoni and Matthews have found the limits of natural selection, here a factor of 1.4.

Thus, the data on methylenetetrahydrofolate reductase do not compel the conclusion that kinetic optimality is sacrificed in this enzyme to obtain satisfactory regulatory properties. However, they do support the suggestion that catalytic behavior can drift down when not functionally constrained. This example illustrates the value of careful and thorough kinetic studies in laying the ground for analysis of kinetic perfection. Because of this work, methylenetetrahydrofolate reductase should be an important enzyme for testing new ideas of catalytic perfection as they emerge.

V. Site-Directed Mutagenesis

Mutant enzymes prepared by recombinant DNA technology are becoming the focus of biochemical study in many laboratories. Such mutants are normally made to test ideas relating structure and behavior in proteins. However, if kinetic data on the mutants are thoroughly collected, they may provide insight into the role of natural selection in optimizing the behavior of the native enzyme. In particular, the distribution of the kinetic behaviors of the mutants relative to those of the native enzyme (wild type) can suggest conclusions about whether a behavior was maximized, optimized but not maximized, or drifting in the native enzyme.

For example, if k_{cat} is maximized, k_{cat} values of enzymes containing single amino acid substitutions should be distributed to lower values (Figure 8). The distribution of k_{cat} values of double mutations is expected to be more skewed in the direction of lower values. Indeed, as more mutations are introduced, the k_{cat} values of the mutants will approach those of a randomly constituted protein, most likely a distribution that can be approximated by a simple exponential curve (Figure 8).

These expectations are most easily understood by considering the surface in n-dimensional space that

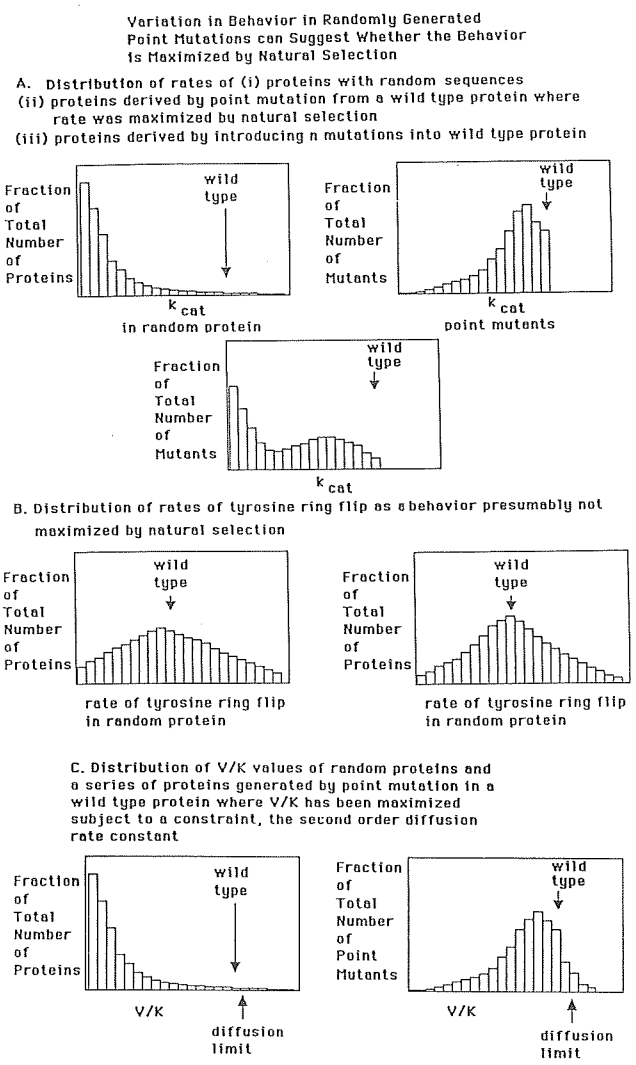

Figure 8. Introduction of random mutations into a protein will create a set of mutants displaying a distribution in behavior tending toward the distribution of that behavior in a protein with random sequence. For example (top pair of diagrams), the distribution of the rate of catalysis observed in a set of proteins with a random sequence should resemble a simple decaying exponential. Thus, introduction of point mutations into a protein with the catalytic rate optimized to the right end of the distribution will create a set of mutants with a rate distribution skewed to the left. In contrast, the rate of tyrosine ring flip in a random polypeptide is distributed around some norm characteristic for proteins in general. If not selected, this rate in mutants will be distributed more evenly (second pair). If the catalytic rate is maximized with respect to an external constraint (e.g., V/K and the diffusion limit), the distribution of this value in mutants shown in pair 3 is expected.

describes a function relating k_{cat} to amino acid sequence in the protein.[5] For behaviors that are optimized (either locally or globally), natural selection has found a point on that surface where small perturbations in structure always diminish k_{cat}.

Kinetic parameters such as K_M that are presumably optimized, but where the optimum is not a maximum, behave differently. The K_M values of random mutants are expected to be uniformly scattered on both sides of the K_M of the native protein. Similarly nonskewed distributions of mutant behaviors are expected for traits that are not selected, and traits that are maximized subject to external constraint.

Of the many proteins studied using site-directed mutagenesis techniques, only mutants of tyrosine

aminoacyl-tRNA synthetase have been studied in sufficient detail to support this type of analysis.[108] Mutants of this protein prepared by Fersht and his coworkers were designed to test hypotheses regarding catalytic activity. They were not intended to test evolutionary questions, and therefore so may be regarded as "semirandom" from our point of view. As expected for a maximized trait, the distribution of k_{cat} values of the mutants are skewed toward lower values. Likewise, values of K_M for these mutants are more evenly distributed around the wild type, suggesting that this parameter, if optimized, is not maximized. Here, the analysis is incomplete; in particular, it is complicated by the fact that the protein being studied is only a fragment of the physiologically relevant protein.

Evolutionary considerations of this sort are very important as guides to site-directed mutagenesis experiments. In general, structural "tinkering"[109] by a biochemist is likely to provide mutant proteins with improved properties (in the opinion of the biochemist) only to the extent that the goal of natural selection is not the same as the goal of the biochemist. If the biochemist wishes to improve on a behavior that is already optimized by natural selection, mutants (both deliberately and randomly created) are unlikely to yield enzymes with improved performances. Fortunately, in many cases (stability, substrate binding, and inhibition in particular), the goals of the biochemist and natural selection are not congruent. In these areas, "enzyme engineering" is likely to yield mutant proteins with more desirable characteristics.[110]

VI. Conclusions

Models that begin by assuming that the optimal enzymes are diffusion limited must undergo substantial revision to accommodate the low second-order diffusion rate constant under relevant biological conditions and to correct incomplete theories that predict optimal values for microscopic kinetic parameters (e.g., K_{int}). The revised models are physiologically more reasonable and algebraically more correct and can be experimentally tested. Where these models have been subjected to experimental examination, their predictions have matched the data to within an order of magnitude in nearly all cases and to within a factor of 2 in some.

Thus, enzymes appear to display optimal kinetic behavior, as defined by these models, to a reasonable degree of precision. This permits the theory to serve as the basis for more broadly based experimental efforts. For example, if the models are presumed to be true, they suggest that extremely high values of k_{cat}/K_M may be diagnostic of enzymes that act in multienzyme aggregates.

There remains very little evidence to suggest that there exist tradeoffs between kinetic behavior and other behaviors, such as stability, substrate specificity, or regulation. Only in the case of ribulose diphosphate carboxylase does it appear that evolution has been unable to produce in a single polypeptide a good solution to both a catalytic problem and a substrate specificity problem. In contrast, the apparent tradeoff between stability and speed suggested for enzymes from thermophiles disappears after inclusion of K_M in the analysis. Substrate specificity may be best understood as varying independently. Finally, there is no clear evidence that regulation and kinetic optimization are incompatible.

However, it is clear that theories that attempt to define kinetic optimality in enzymes by comparisons with diffusion rate constants are fundamentally paradoxical. As both fast and slow enzymes can be optimal catalysts depending on the relative metabolic cost of the two, attention must be focused on largely unexplored questions in enzymatic reactions: (a) How are catalytic power and biosynthetic cost related in enzymes? (b) Do stereoelectronic and functional group requirements for different reaction types make some reactions easier to catalyze and others more difficult? (c) Can pairs of enzymes evolving in similar environments (isozymes, enzymes using different mechanisms to catalyze functionally identical reactions) provide critical tests of particular theories regarding functional adaptation in enzymes? Attempts to answer these questions experimentally should provide much exciting biochemical research in the next decade.

Acknowledgments. I am indebted to the Swiss National Science Foundation, Sandoz, Hoffmann-La Roche, Ciba-Geigy, and the Searle Foundation for support of my laboratory during the time when this review was prepared, to supporting co-workers whose work is cited within, and to many colleagues (and referees) who read early versions of the manuscript.

References

(1) Kernohan, J. C. *Biochim. Biophys. Acta* **1964**, *81*, 346.
(2) Weintraub, H.; Baulieu, E.-E.; Alfsen, A. *Biol. J.* **1980**, *185*, 723.
(3) Dobzhansky, T.; Ayala, F. J.; Stebbins, G. L.; Valentine, J. W., Eds. *Evolution*; W. H. Freeman: San Francisco, 1977; p 19.
(4) Kimura, M. *Molecular Evolution, Protein Polymorphism, and the Neutral Theory*; Springer-Verlag: Berlin, 1982.
(5) Benner, S. A.; Ellington, A. D. *CRC Crit. Rev. Biochem.* **1988**, *23*, 369.
(6) Benner, S. A. *Top. Stereochem.*, in press.
(7) Benner, S. A. In *Redesigning the Molecules of Life*; Benner, S. A., Ed.; Springer-Verlag: Berlin, 1988, p 115.
(8) Kolb, E.; Harris, J. I. *Biochem. J.* **1972**, *130*, 26.
(9) Albery, W. J.; Knowles, J. R. *Biochemistry* **1976**, *15*, 5631.
(10) Koch, A. L. *Adv. Microb. Physiol.* **1971**, *6*, 147.
(11) Novick, A.; Weiner, M. *Proc. Natl. Acad. Sci. U.S.A.* **1957**, *43*, 553.
(12) Andrews, K. J.; Hegeman, G. D. *J. Mol. Evol.* **1976**, *8*, 317.
(13) Diamond, J. M. *Nature* **1986**, *321*, 565.
(14) Dykhuizen, D. E. *Evolution* **1978**, *32*, 125.
(15) Koch, A. L. *J. Mol. Evol.* **1983**, *19*, 455.
(16) Li, W.-H. *Mol. Biol. Evol.* **1984**, *1*, 213.
(17) Wilkinson, A. J.; Fersht, A. J.; Blow, D. M.; Carter, P.; Winter, G. *Nature* **1984**, *307*, 187.
(18) Cleland, W. W. *Annu. Rev. Biochem.* **1967**, *36*, 77.
(19) Atkinson, D. E. *Curr. Top. Cell. Regul.* **1969**, *1*, 29.
(20) Fersht, A. R. *Proc. R. Soc. London, B* **1974**, *187*, 397.
(21) Crowley, P. H. J. *Theor. Biol.* **1975**, *50*, 461.
(22) Lowry, O. .; Passonneau, J. V. *J. Biol. Chem.* **1964**, *239*, 31.
(23) Senbeil, E.; White, H. B., III. *J. Mol. Evol.* **1978**, *11*, 57.
(24) Shaw, W. V. *CRC Crit. Rev. Biochem.* **1983**, **1983**, *14*, 1.
(25) Weber, G. *Cancer Res.* **1983**, *43*, 3466.
(26) Fersht, A. R. *Enzyme Structure and Mechanism*; W. H. Freeman: New York, 1977.
(27) Cornish-Bowden, A. *J. Mol. Biol.* **1976**, *101*, 1.
(28) Cleland, W. W. *Acc. Chem. Res.* **1975**, *8*, 145.
(29) Rosenberry, T. I. *Adv. Enzymol.* **1975**, *43*, 103.
(30) Teipel, J. W.; Hass, G. M.; Hill, R. L. *J. Biol. Chem.* **1968**, *243*, 5684.
(31) Albery, W. J.; Knowles, J. R. *Angew. Chem.* **1977**, *16*, 285.
(32) Eigen, M.; Hammes, G. G. *Adv. Enzymol.* **1963**, *25*, 1.
(33) Hammes, G. G.; Schimmel, P. R. *The Enzymes (3rd Ed.)* **1970**, *2*, 67.
(34) Brouwer, A. C.; Kirsch, J. F. *Biochemistry* **1982**, *21*, 1302.
(35) Nakatani, H.; Dunford, H. B. *J. Phys. Chem.* **1979**, *83*, 2662.
(36) Fink, A. L.; Cartwright, S. J. *CRC Crit. Rev. Biochem.* **1981**, *11*, 145.

(37) Rudnik, G.; Abeles, R. H. *Biochemistry* **1975**, *14*, 4515.
(38) Llewellyn, D. J.; Daday, A.; Smith, G. D. *J. Biol. Chem.* **1980**, *255*, 2077.
(39) Andrews, P. R.; Smith, G. D.; Young, I. G. *Biochemistry* **1973**, *12*, 3492.
(40) Sachs, F. *Magnetic Resonance in Colloid and Interface Science*; Resing, H. A., Wade, C. G., Eds.; ACS Symposium Series *34*; American Chemical Society: Washington, DC, 1976.
(41) Cerofolini, G. F.; Cerofolini, M. *Speculations Sci. Technol.* **1980**, *3*, 149.
(42) Wiggins, P. M. *J. Theor. Biol.* **1971**, *32*, 131.
(43) Schmidt-Nielsen, K.; Taylor, C. R. *Science* **1968**, *162*, 274.
(44) Cokelet, G. R.; Meiselman, H. J. *Science* **1968**, *162*, 275.
(45) Wilson, W. L.; Heilbrun, L. V. *Q. J. Microsc. Sci.* **1960**, *101*, 95.
(46) Ashton, F. *Biol. Bull.* (*Woods Hole, Mass.*) **1957**, *113*, 319.
(47) Tanner, J. E.; Stejskal, E. O. *J. Chem. Phys.* **1968**, *49*, 1768.
(48) Keith, A. D.; Snipes, W. *Science* **1974**, *183*, 666.
(49) Friedrich, P. *Supramolecular Enzyme Organization. Quaternary Structure and Beyond*; Pergamon Press: Oxford, 1984.
(50) Ginsburg, A.; Stadtman, E. R. *Annu. Rev. Biochem.* **1970**, *39*, 429.
(51) Srivastava, D. K.; Bernhard, S. A. *Curr. Top. Cell. Regul.* **1986**, *28*, 1. Robinson, J. B., Jr.; Srere, P. A. *J. Biol. Chem.* **1985**, *260*, 10800.
(52) Place, A. R.; Powers, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 2354.
(53) Knowles, J. R.; Albery, W. J. *Acc. Chem. Res.* **1977**, *10*, 105.
(54) Burton, J.; Sinnott, M. L. *J. Chem. Soc., Perkin Trans. 2* **1983**, 359.
(55) Li, B. F. L.; Osborne, S.; Sinnott, M. L. *J. Chem. Soc., Perkin Trans. 2* **1983**, 365.
(56) Hall, B. G.; Murray, M.; Osborne, S.; Sinnott, M. L. *J. Chem. Soc., Perkin Trans. 2* **1983**, 1595.
(57) (a) Fersht, A. R.; Leatherbarrow, R. J.; Wells, T. N. C. *Nature* **1986**, *322*, 284. (b) Estell, D. A. *Protein Eng.* **1987**, *1*, 441. Fersht, A. R. *Protein Eng.* **1987**, *1*, 445. (c) Chapman, N. B.; Shorter, J., Eds. *Advances in Linear Free Energy Relationships*; Plenum Press: London, 1972. Leffler, J. E.; Grunwald, E. *Rates and Equilibria of Organic Reactions*; Wiley: New York, 1983. (d) Chen, M. Y.; Murdoch, J. R. *J. Am. Chem. Soc.* **1984**, *106*, 4735. Murdoch, J. R. *J. Am. Chem. Soc.* **1972**, *94*, 4410.
(58) Strauss, D. R.; Raines, R. T.; Kawashima, E.; Knowles, J. R.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2272. Raines, R. T.; Sutton, E. L.; Strauss, D. R.; Gilbert, W.; Knowles, J. R. *Biochemistry* **1986**, *25*, 7142. Blacklow, S. C.; Raines, R. T.; Lim, W. A.; Zamore, P. D.; Knowles, J. R. *Biochemistry* **1988**, *27*, 1158.
(59) Lolkema, J. S.; Hoeve-Duurkens, R. H.; Robillard, G. T. *Eur. J. Biochem.* **1986**, *154*, 387.
(60) Myers, J. A.; Boyer, P. D. *Biochemistry* **1984**, *23*, 1269.
(61) Cross, R. L.; Cunningham, D.; Tamura, J. K. *Curr. Top. Cell. Regul.* **1984**, *24*, 335.
(62) Ashton, A. R.; Hatch, M. D. *Arch. Biochem. Biophys.* **1983**, *227*, 406.
(63) Cross, R. L., *Annu. Rev. Biochem.* **1981**, *50*, 681.
(64) Hassett, A.; Blaettler, W.; Knowles, J. R. *Biochemistry* **1982**, *21*, 6335.
(65) Kenyon, G. L.; Reed, G. H. *Adv. Enzymol.* **1983**, *54*, 367.
(66) Cook, P. F.; Cleland, W. W. *Biochemistry* **1981**, *20*, 1807.
(67) Nageswara Rao, B. D.; Cohn, M. *J. Biol. Chem.* **1981**, *256*, 1716.
(68) Nageswara Rao, B. D.; Kayne, F. J., Jr.; Cohn, M. *J. Biol. Chem.* **1979**, *254*, 2689.
(69) Wilkinson, K. D.; Rose, I. A. *J. Biol. Chem.* **1979**, *254*, 12567.
(70) Benner, S. A. *Experientia* **1982**, *38*, 633.
(71) Benner, S. A.; Stackhouse, J. In *Studies in Organic Chemistry*; Green, B. S., Ashani, Y., Chipman, D., Eds.; Elsevier: Amsterdam, 1982; Vol. 10, p 32.
(72) Rees, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 3082.
(73) Stackhouse, J.; Nambiar, K. P.; Burbaum, J.; Stauffer, D. M.; Benner, S. A. *J. Am. Chem. Soc.* **1985**, *107*, 2757.
(74) Rahil, J. F.; de Maine, M. M.; Benkovic, S. J. *Biochemistry* **1985**, *21*, 3358.
(75) Nambiar, K. P.; Stauffer, D. M.; Kolodziej, P. A.; Benner, S. A. *J. Am. Chem. Soc.* **1983**, *105*, 5886.
(76) Gutfreund, H. *Prog. Biophys. Mol. Biol.* **1975**, *29*, 161.
(77) Schwert, G. W.; Miller, B. R.; Peanasky, R. J. *J. Biol. Chem.* **1967**, *242*, 3245.
(78) Chin, J. *J. Am. Chem. Soc.* **1983**, *105*, 6502.
(79) Ellington, A. D.; Benner, S. A. *J. Theor. Biol.*, in press.
(80) Kaplan, N. O. *Evolving Genes and Proteins*; Bryson, V., Vogel, H. J., Eds.; Academic Press: New York, 1968; p 190.
(81) Ellington, A. Thesis, Harvard University, 1988.
(82) Eventoff, W.; Rossmann, M. G.; Taylor, S. S.; Torff, H.-J.; Meyer, H.; Keil, W.; Kiltz, H.-H. *Proc. Natl. Acad. Sci. U. S.A.* **1977**, *74*, 2677.
(83) Kiltz, H.-H.; Keil, W.; Griesbach, M.; Petry, K.; Meyer, H. *Hoppe-Seyler's Z. Physiol. Chem.* **1977**, *358*, 123.
(84) Low, P. S.; Bada, J. L.; Somero, G. N. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 430.
(85) Hochachka, P. W.; Somero, G. N. *Biochemical Adaptation*; Princeton University Press: Princeton, NJ, 1984.
(86) Somero, G. N. *Annu. Rev. Ecol. Syst.* **1978**, *9*, 1.
(87) Somero, G. N. *J. Exp. Zool.* **1975**, *194*, 175.
(88) Somero, G. N.; Yancey, P. H. *Symp. Biol. Chem.* **1978**, *249*, 4571.
(89) Hall, J. G. *Mol. Biol. Evol.* **1985**, *2*, 251.
(90) Bennetzen, J. L.; Hall, B. D. *J. Biol. Chem.* **1982**, *257*, 3018.
(91) Kreitman, M. *Nature* **1983**, *304*, 412.
(92) Hartl, D. L.; Dykhuizen, D. E.; Dean, A. M. *Genetics* **1985**, *111*, 655.
(93) Whitfield, C. D.; Steers, E. J.; Weissbach, H. *J. Biol. Chem.* **1970**, *245*, 390.
(94) Old, I. G.; Hunter, M. G.; Wilson, D. T. R.; Knight, S. M.; Weatherstone, C. A.; Glass, R. E. *Mol. Gen. Genet.* **1988**, *211*, 78.
(95) Plamann, L. S.; Stauffer, G. V. *J. Bacteriol.* **1987**, *169*, 3932.
(96) Taylor, R. T.; Weissbach, H. *Methods Enzymol.* **1971**, *17B*, 379.
(97) Sheu, K. R.; Richard, J. P.; Frey, P. A. *Biochemistry* **1979**, *18*, 5548.
(98) Eigen, M. *Quantum Statistical Mechanics in the Natural Sciences*; Plenum Press: New York, 1974; p 37.
(99) Stoecklein, W.; Schmidt, H.-L. *Biochem. J.* **1985**, *232*, 621.
(100) Miziorko, H. M.; Lorimer, G. H. *Annu. Rev. Biochem.* **1983**, *52*, 507.
(101) Wiesenfeld, M.; Schimpfessel, L.; Crokaert, R. *Biochim. Biophys. Acta* **1976**, *405*, 500.
(102) Chesne, S.; Monnier, N.; Pelmont, J. *Biochimie* **1978**, *60*, 403.
(103) Hill, R. L.; Bradshaw, R. A. *Methods Enzymol.* **1969**, *13*, 91.
(104) Daar, I. O.; Artymiuk, P. J.; Phillips, D. C.; Maquat, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 7903.
(105) Cooper, T. G. *TIBS* **1980**, 332.
(106) Vanoni, M. A.; Matthews, R. G. *Biochemistry* **1984**, *23*, 5272.
(107) Kutzbach, C.; Stokstad, E. L. R. *Biochim. Biophys. Acta* **1971**, *250*, 459.
(108) Fersht, A. R.; Leatherbarrow, R. J.; Wells, T. C. *Biochemistry* **1987**, *26*, 6030.
(109) Knowles, J. R. *Science* **1987**, *236*, 1252.
(110) Benner, S. A. *Bioorg. Chem. Frontiers*, in press.
(111) Cornish-Bowden, A. *Trends Biochem. Sci.* **1986**, *11*, 286.