Evolution of a single gene highlights the complexity underlying molecular descriptions of fitness

Matthew I. Peña, Elizabeth Van Itallie, Matthew R. Bennett, and Yousif Shamoo^{a)}

Department of Biochemistry and Cell Biology, Rice University, 6100 Main St., MS-140, Houston, Texas 77005, USA

(Received 12 February 2010; accepted 23 May 2010; published online 30 June 2010)

Evolution by natural selection is the driving force behind the endless variation we see in nature, yet our understanding of how changes at the molecular level give rise to different phenotypes and altered fitness at the population level remains inadequate. The reproductive fitness of an organism is the most basic metric that describes the chance that an organism will succeed or fail in its environment and it depends upon a complex network of inter- and intramolecular interactions. A deeper understanding of the quantitative relationships relating molecular evolution to adaptation, and consequently fitness, can guide our understanding of important issues in biomedicine such as drug resistance and the engineering of new organisms with applications to biotechnology. We have developed the “weak link” approach to determine how changes in molecular structure and function can relate to fitness and evolutionary outcomes. By replacing *adenylate kinase* (AK), an essential gene, in a thermophile with a homologous AK from a mesophile we have created a maladapted weak link that produces a temperature-sensitive phenotype. The recombinant strain adapts to non-permissive temperatures through point mutations to the weak link that increase both stability and activity of the enzyme AK at higher temperatures. Here, we propose a fitness function relating enzyme activity to growth rate and use it to create a dynamic model of a population of bacterial cells. Using metabolic control analysis we show that the growth rate exhibits thresholdlike behavior, saturating at high enzyme activity as other reactions in the energy metabolism pathway become rate limiting. The dynamic model accurately recapitulates observed evolutionary outcomes. These findings suggest that *in vitro* enzyme kinetic data, in combination with metabolic network analysis, can be used to create fitness functions and dynamic models of evolution within simple metabolic systems. © 2010 American Institute of Physics. [doi:10.1063/1.3453623]

Adaptation is the essence of Darwin’s brilliant conception of natural selection and is the mechanism by which all organisms are shaped by their environment. Adaptation is a fundamental property of all life and is responsible for recent increases in the populations of drug resistant pathogens that directly impact human health. While microorganisms have relatively simpler and smaller genomes than humans, they still contain thousands of genes that interact with each other to create large networks whose response to challenges is remarkably complex and largely unpredictable. The genome defines the ways in which an organism can respond to its environment and is also a record of adaptation that encodes the “current state” of the organism. The genome provides an excellent system for recording the steps of adaptation at a molecular level (molecular evolution). The challenge of understanding molecular evolution can be addressed through the development of robust experimental systems that afford the opportunity to build and validate models for adaptation. Our work has focused on the development of a “weak link” approach that relates organismal fitness to the performance of a single gene and permits the identification of the intermediate “steps” of adaptation. Using

the weak link approach, the scale of the problem is dramatically reduced but retains key aspects of the larger system that facilitates the development of quantitative models for adaptation. Understanding the physical basis for adaptation of an organism is the first step in the development of approaches that can accurately model, and someday predict, the manner in which organisms would respond to new antibiotics and improve upon the current clinical regimens. As adaptation is the driving force for drug resistance in bacteria and cancer, the benefits to biomedicine are potentially enormous.

I. INTRODUCTION

The role of natural selection in shaping the success or failure of species can be observed all around us. Perhaps one of the most common and vexing examples of natural selection is the rise of drug resistant microbes and viruses. Antibiotic resistance provides a clear and pressing case for a better understanding of adaptation. The extensive use of antibiotics since the 1940s has led increasingly to strains of bacteria that are highly resistant to each new generation of antibiotics. Data from 2004 show that methicillin resistant *Staphylococcus aureus* (MRSA) were found in over 60% and vancomycin resistant *Enterococci* in nearly 30% of ICU patients compared to 37% and 14%, respectively, in 1995.^{1,2}

^{a)} Author to whom correspondence should be addressed. Electronic mail: shamoo@rice.edu. Telephone: 713-348-5493. FAX: 713-348-5154.

FIG. 1. (Color online) In the weak link approach the chromosomal copy of the essential adenylate kinase gene of the thermophile *Geobacillus stearothermophilus* (adk_{Gste}) was replaced with a copy from the mesophile *Bacillus subtilis* through homologous recombination (adk_{Bsub}) (upper left). The resulting cell line (NUB3612-R:ThEV) is no longer able to grow at higher temperatures (upper right). As adk is the only gene within the genome no longer adapted for growth at higher temperatures, the force of natural selection favors changes to adk as it is the weak link (Ref. 27). A microbial population carrying the maladapted gene was then subjected to selection with a gradually increasing temperature in a turbidostat (lower left). As the temperature rises a polymorphic population of adaptive mutants (lower center) is produced that can be characterized for their physicochemical properties (lower right) and incorporated in modeling adaptation. Many of the principles derived from this model system are applicable to clinically relevant scenarios such as adaptation to drugs.

The costs to society are staggering and extend well beyond the additional 5 billion dollars in health care costs per year.^{3,4} The Centers for Disease Control and Prevention estimates that there are approximately two million cases of infection within healthcare settings that result in over 80 000 patient deaths.^{1,5–7} Community associated outbreaks of MRSA are also increasing⁸ suggesting that drug resistant strains are making their way into school locker rooms^{9,10} and other public areas.^{11,12} In addition to human-to-human transmission, the widespread use of antibiotics in agriculture^{13,14} and aquaculture^{15,16} promotes the evolution and proliferation of resistant strains. Hospitals and farms can therefore act as breeding grounds for drug resistant strains.^{17–20} At the root of this problem is the highly adaptive nature of microorganisms and life in general. Resistance can be achieved by various mechanisms including chemical inactivation of the antibiotic, modification of the target site, and reduction of the effective *in vivo* concentration by efflux pumps or reduced permeability. Antibiotic resistance has served as model for adaptation, but as a system for studying inter or intramolecular influences on fitness it is not easily tractable as there are so many potential adaptive pathways to increased resistance. In order to understand the fundamental physical principles underlying adaptation, a more tractable experimental system is needed to provide a good starting point for the examination of this important phenomenon. In turn, the principles elucidated and validated in model systems can be extended to more complex adaptation such as that seen to drugs.

We have developed a model system that allows a systematic examination of intramolecular epistasis in a maladapted copy of an essential gene (Fig. 1) within a metabolic network. Epistasis is a general phenomenon describing the dependence of a mutation's effect on the genetic background in which it occurs. Epistasis is commonly applied to intermolecular gene-gene interactions but can also be extended to the intramolecular networks that define the characteristic fold and function of an enzyme within these larger relationships. In principle, intermolecular and intramolecular epistasis represents facets of the same phenomenon: the effect of perturbations to a node is dependent on the structure of the net-

work of interactions. To understand an adaptation, one must know the interconnectedness of the allele (among genes and pathways) and how changes such as substitutions (the smallest unit of evolutionary change) affect activity and specificity while reshaping the landscape for future adaptations.

A. Why is adaptation difficult to predict?

Adaptation to any challenge begins with the available set of genes that determine the fitness of the organism to its current environment (fitness here is defined as the average reproductive success). Even in organisms with smaller genomes, like bacteria, there are typically thousands of genes whose regulation and expression define the fitness of the cell. These genes interact in complex ways to produce the phenotype and adaptive characteristics of the organism. The genome of the organism is its blueprint, and the ability of the organism to adapt is determined by the ability of the cell to respond through existing mechanisms (for example: gene regulation) or through changes to its genome that produce an organism with an altered fitness. Many organisms are able to increase the range of their adaptive response by acquiring new genes from other organisms via horizontal gene transfer (HGT) or sexual reproduction. However it is constituted, the genome defines the starting point for adaptation and delineates the molecular pathways toward increasing fitness to a particular challenge. As each generation passes, those cells that can replicate more rapidly have more success and begin taking over the population until they reach fixation (fixation is the point at which the population is overtaken by a particular strain and all others are effectively eliminated). The process by which adaptation to increased fitness occurs is often referred to as an “adaptive walk.”²¹ In simple models that ignore HGT or large-scale rearrangements of the genome, mutation is an important source of variation within the population that can be acted upon by natural selection. Many mutations occur from errors in DNA replication or repair. While most mutations are deleterious, those that even slightly increase the fitness of the organism are steps toward

a local maximum in the fitness landscape. Step by step, mutation by mutation, the cells become increasingly successful over the course of natural selection.

In the simplest case of adaptation, the change of a single nucleotide within a gene produces a protein with different properties that alters the fitness of the organism. A good example for this effect is found in the adaptation of organisms to antibiotics. One way of easily describing the resistance of a cell to a specific antibiotic is through an estimation of the minimum inhibitory concentration (MIC). The MIC is the minimum concentration of a drug required to prevent growth. *E. coli* carrying the gene *TEM-1* that encodes the enzyme β -lactamase have excellent resistance to the β -lactam antibiotic ampicillin with a MIC in excess of 15 000 $\mu\text{g/ml}$.²² The same cells show little resistance (MIC of ~ 0.09 $\mu\text{g/ml}$) to the newer β -lactam antibiotic cefotaxime (CTX).²³ A single nucleotide change of a G $\rightarrow$ A in *TEM-1* changes the amino acid at position 238 in β -lactamase from a glycine to a serine (denoted as G238S) and increases the MIC for CTX by 16-fold. When combined with a second point mutation leading to the change E104K, the MIC increases over 4000-fold. A single nucleotide change is the smallest increment of change in DNA and demonstrates the power of even these small changes on the fitness of an organism during adaptation.

The typical bacterial genome contains thousands of genes, these genes interact with one another to create vast networks that, when taken as a whole, translate the genotype into phenotype. Therefore, it is nearly impossible to predict what impact, if any, a single point mutation will have on the overall fitness of the organism. To make things even more complicated, the context-dependent manner in which individual proteins fold and function means that mutations that appear similar can have drastically different phenotypic consequences. For example, consider the mutation of the β -lactamase that confers CTX resistance. Wild-type β -lactamase contains 21 glycines throughout its structure. Single point mutations might turn any one of these glycines into a serine, but only the specific mutation G238S significantly increases resistance. Whether a mutation is beneficial or not is therefore entirely dependent on the spatial and stereochemical context of the change.

Another difficulty that complicates our ability to model and predict adaptation is that mutations have a strong tendency to be epistatic, meaning the effects of mutations are nonadditive. To put it another way, two point mutations that individually increase fitness may decrease fitness when jointly present in the genome. In the case of *TEM-1*, Weinreich and co-workers made all 32 (2^5) combinations of five amino acid mutations associated with increased resistance to CTX in a clinical isolate. Natural selection favors mutations providing immediate benefits and not potential benefits, and thus not all mutational pathways are equally probable although the end point is shared. It was shown that of the 120 ($5!$) possible paths toward constructing the five mutant genotype only 18 of these increase the MIC at each sequential step.²³ In addition, the order of mutational events is often critical. So, although smooth adaptive landscapes with local maxima are qualitatively satisfying, realistic landscapes are

the product of a complex hyperdimensional genomic or genotypic fitness functions whose properties are more akin to a “rough” landscape—making it difficult to predict and model.^{24–26}

Adaptations to antibiotics are often manifested as mutations to genes that initially provide a minimal mechanism of resistance, as was the case with *TEM-1* and *CTX*. This observation suggests that the introduction of a maladapted gene required for survival, but lacking optimization to an environmental pressure, can be used as a strategy to focus adaptations to a single gene and thus provide a model in which to study intramolecular epistasis and the relationship between function and phenotype.

B. A tractable model for studying evolutionary dynamics at the molecular level

A central goal of our work has been the development of the weak link approach that relates organismal fitness to the performance of a single gene. Using the weak link approach, the scale of the problem is dramatically reduced but retains key aspects of the larger system. The intermediate “steps” of adaptation can be easily identified and suitable models for adaptation developed within an experimental framework that promotes validation. In the weak link approach, the one chromosomal copy of the essential adenylate kinase (*adk*) gene of the thermophile *Geobacillus stearothermophilus* was replaced with a copy from the mesophile *Bacillus subtilis* through homologous recombination. As *adk* is the only gene within the genome no longer adapted for growth at higher temperatures, the force of natural selection favors changes to *adk* as it is the weak link within the genome of the organism. Adenylate kinase (AK) maintains the crucial balance of adenosine triphosphate (ATP) to diphosphate (ADP) and monophosphate (AMP) adenylates in cells and is therefore central to cellular energy homeostasis. While *G. stearothermophilus* (NUB3621-R) is a moderate thermophile able to grow over a broad temperature range (48–70 °C), the recombinant *G. stearothermophilus* (NUB3621-R:ThEv) expressing the *B. subtilis* *adk* is no longer able to grow at temperatures above ~ 55 °C (Refs. 27 and 28) (Fig. 1). Cell death at non-permissive temperatures results from thermal unfolding of *B. subtilis* adenylate kinase (AK_{BSUB}) that, in turn, disrupts adenylate homeostasis. To follow adaptation, a large population ($\sim 10^{11}$) of continuously growing cells is founded at a permissive temperature (55 °C) and then challenged with increasing temperatures over a long time course to select for those individuals with mutations to *adk* that increase functionality and thus viability of the organism at higher temperatures. During the course of selection the population was sampled and functional intermediates of adaptation observed as mutations to *B. subtilis* *adk*. In our original study, the first mutant to reach fixation was a single mutation AK_{BSUB} Q199R (the glutamine at position 199 replaced with arginine). AK_{BSUB} Q199R was eventually replaced at 62–63 °C by five double mutants that arose nearly simultaneously within the population and share AK_{BSUB} Q199R as their progenitor [Fig. 2(a)]. All five double mutants arose to at least 5% of the population in the temperature range of 62–63 °C as AK_{BSUB} Q199R declined.²⁷ Reconstruction of the identi-

FIG. 2. (Color) (a) The cumulative frequency of alleles obtained by experimental evolution within an exponentially growing population as a function of temperature (Ref. 27). (b) AK activity and folding parameters were used to fit ADP production rates for enzyme variants at 1000 μM ATP as a function of temperature. [Reprinted with permission from M. I. Peña, M. Davlieva, M. R. Bennett, J. S. Olson, and Y. Shamoo, "Evolutionary fates within a microbial population highlight an essential role for protein folding during natural selection," *Mol. Syst. Biol.* (to be published). Copyright ©2010 Nature Publishing Group.]

fied mutants by site-directed mutagenesis to the *B. subtilis* *adk* gene followed by homologous recombination within an unadapted strain recapitulated the phenotype (thermal tolerance) and reflects that these substitutions to *adk* and not potential mutations outside of this allele represent the predominant mechanism of adaptation to high temperatures. Adaptations to the growth environment could have occurred outside of *adk*; however, all adapted strains contained mutations to the *adk* allele indicating strong selection. Changes to AK activity and thermal stability that resulted from mutation had direct consequences for cellular fitness and therefore met our goal for an experimental system that allows us to develop and test models for quantitative molecular evolution. The strong coupling between adaptations to AK and organismal thermal tolerance is due to the essential role of AK in energy metabolism.

C. The physiological role of adenylate kinase

ATP is the predominant energy carrier enabling cellular work to be directed through the hydrolysis of its two phosphoanhydride bonds. AK, in conjunction with ATP synthase, is responsible for ATP regeneration from lower energy moieties (ADP with a single high energy bond and AMP with none). AK catalyzes the reversible reaction that charges AMP with the addition of a phosphate forming an ADP at the expense of one phosphate bond from ATP,

ATP synthase utilizes the proton-motive force across the plasma membrane created by the electron-transport chain to drive the generation of ATP from the lower energy ADP,

Without AK, there is no efficient way of charging the AMP

formed *de novo* or by reactions coupled to the energy releasing hydrolysis of ATP. High energy ATP is a fundamental regulatory effector of energy metabolism pathways (e.g., glycolysis, citric acid cycle, and oxidative phosphorylation), and in some cases acts as an allosteric inhibitor slowing the rate of catabolic regeneration of ATP (e.g., phosphofructokinase within glycolysis). The relative intracellular concentrations of AMP, ADP, and ATP reflect the overall energy level and are often referred to as energy charge (EC) [Eq. (1)],²⁹

$$\text{EC} = ([\text{ATP}] + \frac{1}{2}[\text{ADP}]) / ([\text{ATP}] + [\text{ADP}] + [\text{AMP}]). \quad (1)$$

In the event that AK loses function, EC will decline due to a decrease in flux of ADP and consequently a decrease in ATP and a buildup of AMP. It is the essentiality of AK function that makes the weak link approach a successful experimental framework for the study of adaptation.^{27,28,30}

D. A simple "physiological threshold" can reproduce observed evolutionary dynamics

A combination of biophysical studies was used to evaluate the temperature dependent enzyme activity, folding and aggregation properties of the adaptive AK mutants isolated during experimental evolution.³¹ In principle, the fitness of each cell carrying these adaptive changes should be largely dependent on changes to the physicochemical properties of the mutant proteins. A comparison of observed enzyme activity as a function of temperature and evolutionary outcomes supported the hypothesis that these enzyme variants have very comparable effects on fitness despite diverse molecular mechanisms for adaptation.³² For example, AK_{BSUB} Q199R/Q16L, with the greatest *in vitro* activity at all measured temperatures, could not outcompete the other functional double mutants until they would later succumb to the effects of AK thermal denaturation (Fig. 2). These observa-

FIG. 3. First approximation fitness functions include a simple physiological threshold (dashed) and Hill function (solid). A relative growth rate (y axis) can be calculated using these relations with an input of *in vitro* activity (x axis). Either can be used to reproduce experimental outcomes, but the Hill function has a form more consistent with response of the energy metabolism network to varying enzyme activity.

tions suggested that enzyme activity alone is not a sufficient proxy for fitness. In the evolution experiment, a polymorphic population resulted across temperatures ranging from 62 to 69 °C over approximately 500 generations. Relative fitness differences are reflected in the frequencies of strains over time and temperature. Simulations have shown that fitness differences between variants must be small (less than 5% difference); otherwise, this polymorphism would not have been so persistent.

To determine what parameters best correlate to fitness, fitness functions were constructed relating single parameters or combinations of parameters (e.g., k_{cat} , K_M , and K_{unf}) to growth rates. These growth rates were then defined within an *in silico* population to determine which combination of parameters produced polymorphisms best matching empirical outcomes. Attempts to build a fitness function for these adaptive changes were only successful in reproducing the evolutionary outcomes when a physiological threshold was applied to a model of *in vitro* enzyme activity (Fig. 3).³¹ The hypothesis was that once the physiological demand for adenylate interconversion was met; further enhancements to AK function conferred diminishingly small returns on fitness. As the activity dropped below the threshold there would be a proportional and direct relationship between growth rate and activity.

Although a physiological threshold is simplistic, it was sufficient to reproduce the evolutionary outcomes in simulation.³¹ Such a physiological threshold, while useful, does not capture the central role of AK in the flux of adenylate species within the cell. For example, a loss of AK function below a certain point does not lead to relatively linear decreases in fitness, but rather results in a catastrophic and irreversible event—cell death. Simply, the organism is non-viable below a nonzero value of EC, and hence AK activity. To provide a more robust theoretical framework for future

experimentation, we have constructed a fitness function modeling the response of EC at equilibrium to activity of AK.

A Hill function has qualities similar to our original fitness function with a threshold: increases in activity translate to negligible increases in growth rate far from the transition, and a sharp transition captures the physiological threshold required in our first approach (Fig. 3). Hyperbolic or “concave” fitness functions have been used previously to relate enzyme activity to fitness.^{33–35} These functions, equivalent in form to the familiar Michaelis–Menten equation, represent a special case of the Hill function with a Hill coefficient of unity. A Hill function with a coefficient greater than one was needed to recreate the sustained polymorphism *in silico*, although attempts to match empirical frequencies resulted in a shift in the transition from AK_{BSUB} Q199R to double mutants to higher temperatures (64–65 °C) relative to the experimental shift (62–63 °C). To relate activity-temperature profiles of enzyme variants to competition between the mutant strains within a population, activities were fed into a Hill function to calculate a growth rate that was then used to determine population dynamics [Fig. 4(a)]. A Hill function with a midpoint (m) of 4.5 and a coefficient of 5 (h) was used as the fitness function relating activity (a) to growth rate (w) reproduces evolutionary outcomes in Fig. 4(b) consistent with experimental data seen in Fig. 2(a),

$$w = w_{\max} \frac{\left(\frac{a}{m}\right)^h}{1 + \left(\frac{a}{m}\right)^h}. \quad (2)$$

This fitness function has been determined for one data set but can be used to predict outcomes of competition experiments between these adaptive mutant AKs or novel ones that will be characterized prior to competition. Despite the simplicity of the fitness function, contrasted with the complexity of energy metabolism, we are able to reproduce evolutionary outcomes using only *in vitro* activity data.

E. The Hill function fitness model is consistent with metabolic flux analysis

What is the influence of a single enzyme on the fitness of an organism? In trying to answer this question, Kacser and Burns^{36,37} along with Heinrich and Rapoport^{36–38} established the field of metabolic control analysis whereby the fluxes through metabolic networks are quantitatively determined by identifying the structure and variables defining the network. Bridging flux to fitness provided the connections that allow the molecular attributes of a protein, like AK, to be linked to the properties of the metabolic systems under selection.³⁹ Resource-limited chemostat studies confirmed that growth rates (fitness) were dependent on flux through the metabolic system as determined by the respective enzyme activities.^{40,41} The sensitivity of the flux (J) through a system to small changes in enzyme concentration (e_i) is described by the flux control coefficient (C_i^J),

FIG. 4. (Color) Fitness function modeled as a Hill function can reproduce experimental evolutionary dynamics. (a) Growth rates are calculated using the Hill function (Fig. 3) and *in vitro* enzyme activity between 60 and 70 °C [Fig. 2(b)]. (b) Growth rates in (a) (along with mutation rate and seeding subpopulation sizes) are used in a simulation modeling growth of a clonal population to determine potential evolutionary outcomes, which correspond to experimental observations.

$$\frac{dJ}{J} \bigg/ \frac{de_i}{e_i} = C_i^J. \quad (3)$$

A consequence of C_i^J is a pattern of diminishing returns in flux with increasing enzyme concentration, a finding that has been confirmed in many studies.^{41–43} The dependence of flux on activity or concentration of a single enzyme within a metabolic pathway is identical (e.g., an enzyme variant may have half the activity as wild type, or the system has half the enzyme concentration, the resulting effect on flux is the same regardless). By examining the extremes of the flux through metabolic networks, it is possible to appreciate the relationship of AK activity to cellular fitness. Without AK, the cell cannot maintain a sufficient EC and dies. At the other extreme, increasing activity cannot translate to faster growth rates because a series of metabolic pathways become less sensitive to modulations as enzyme concentration or activity increases ($C_i^J \rightarrow 0$). Simply put, increases to AK activity within the complex metabolic network reach a point of diminishing returns. Understanding the dependence of fitness on activity between these extremes is critical to predicting evolutionary outcomes.

The simplified rate laws described in Fig. 5 demonstrate that the response of ATP flux through the metabolic pathway saturates at high AK activity, resulting in the physiological threshold suggested by our experimental evolution data [Fig. 6(b)]. The differential equations used to describe the response of equilibrium concentrations of adenylate moieties, M, D, and T for AMP, ADP, and ATP, respectively, to changes in activity (a) of AK are

$$0 = \begin{cases} \dot{M} = \beta v_F - av_{AK}MT + av_{-AK}D^2 + v_2T - \gamma M \\ \dot{D} = v_1T + av_{AK}MT - av_{-AK}D^2 - v_5P_iD - \gamma D \\ \dot{T} = -av_{AK}MT + av_{-AK}D^2 - v_1T - v_2T + v_5P_iD - \gamma T. \end{cases} \quad (4)$$

Feedback exists in this system in a variety of forms and has been observed in a number of studies. *In vivo*, a high EC and ATP pool size correspond to growth rates.⁴⁴ ATP serves as a fundamental building block or energy source in nearly every aspect of cellular growth and replication. Faster growth increases utilization of ATP and increases demand for ATP through the ATP synthase pathway. The growth rate dependent demands on AK can be modeled as sinks of adenylate moieties from the system (γ) as well as potentially increasing the flux of AMP into the system (β) with varying dependencies (n_β and n_γ),

$$\beta = n_\beta \frac{EC}{EC + EC_{max}} = n_\beta \frac{EC}{EC + 1} \quad \text{and} \quad \gamma = n_\gamma \frac{EC}{EC + 1}. \quad (5)$$

Temperature will also affect rates, but these cases are ignored with the exception of the temperature dependence of AK activity. Our model suggests that energy metabolism catastrophically fails as enzyme activity becomes sufficiently low [Fig. 6(a)]. This corresponds to an ATP concentration

FIG. 5. An idealized pathway for adenylate generation and use within a cell. AK reaction (v_{AK} , v_{-AK}), ATP synthase reaction (v_S), AMP formation (v_F), and ATP hydrolysis pathways to lower energy moieties (v_1 , v_2) describe the dynamics of energy metabolism.

FIG. 6. (Color online) (a) Concentrations of adenylate moieties and energy charge as a function of enzyme activity. Maximal energy charge would correspond to a system whereby growth rate is not limited by the ATP regeneration pathway. (b) The Hill function (solid) closely approximates the fall of EC as predicted by metabolic control theory (dashed) that would occur as AK failed to maintain adenylate homeostasis.

and EC that approaches zero. The finding that a diminishing EC results in a catastrophic failure of energy metabolism is supported by *in vivo* observations for *E. coli*⁴⁴ as well as in *G. stearothermophilus*.⁴⁵ In *E. coli*, a high EC is maintained in most conditions, except for those in which cellular viability is severely limited.⁴⁴ Thus, EC is a critical indicator of viability. In growing recombinant *G. stearothermophilus*, the EC was experimentally determined to be ~ 0.23 at 55°C .²⁸ A transient increase in temperature from 50 to 65°C resulted in an EC of ~ 0.09 where an EC could not be determined after 30 s of incubation at 65°C due to a high concentration of adenylate moieties in the media. This contrasts to an increase in EC (~ 0.4) after the same incubation period for the wild-type, nonrecombinant strain. Modeling this system with seeding values reveals that the response of the EC in the system to a change in AK activity is indeed well approximated by a Hill function [Fig. 6(b)]. A more realistic model relating activity to flux and flux to fitness will ultimately lead to more accurate predictions of evolutionary outcomes and demonstrates the power of modeling nonlinear systems to understand the complex and ubiquitous phenomenon of natural selection.

II. DISCUSSION

Experimental evolution of microbial populations has been used successfully to study adaptation to various pressures, from the presence of antibiotics to the absence of nutrients.^{21,46} In order for quantitative prediction of evolutionary outcomes to have success, we must have experimental systems that tightly couple the effect of mutations within a cellular population to their evolutionary fates. The weak link approach creates a node within a metabolic network that has a large flux control coefficient ($C_i^J \rightarrow 1$), meaning that improvements to the weak link will have the greatest impact on fitness. This focuses adaptation during selection primarily to this single gene. This has the benefit of simplifying the identification of mutations and allows for the study of

changes in properties of a single enzyme to fitness. Most importantly, adaptation through a single gene can capture many of the complexities and nuances of the larger system, including nonlinear phenomena such as mutational epistasis and saturation effects due to large-scale metabolic networks.

While much has been made of identifying and characterizing mutations as candidates of adaptation, less is known of how these mutations correspond to alterations of protein structure and function that ultimately alter fitness and shape molecular evolution. Metabolic control analysis has been used to determine fitness of organisms over a range of enzyme expression levels but the theory has yet to be combined with models of bacterial growth and evolution to predict a population's response to stress. By correlating fitness to enzyme activity and population dynamics to point mutations, we have begun to understand the multiscale phenomena driving evolution. This same approach can be applied to the study of other evolutionary processes—like the emergence of antibiotic resistance in pathogens or the adaptation of cancer cells to chemotherapy—thereby aiding in the design and implementation of treatment strategies.

The successful implementation of models relating physicochemical properties to evolutionary dynamics permits hypothesis driven ideas to undergo robust evaluation. Biophysical studies of enzyme variants isolated from adapted intermediates during experimental evolution effectively reveal the enzyme properties that are conserved or enhanced by natural selection. Although the adaptive mutations to AK alter stability and activity through complex mechanisms at the molecular level, each isolated variant extends essential activity to higher temperatures which translates to increased range of viability.^{31,32} Flux based models provide general and scalable systems for linking *in vitro* properties of molecules like AK to fitness and evolutionary dynamics thereby bridging intra- and intermolecular influences. Most importantly, these systems afford an opportunity to critically evaluate which of many measurable properties are most critical for

predictive success. In the case of the thermal adaptation of AK, classic biochemical properties such as catalytic rate (k_{cat}) and the Michaelis constant (K_M) must be combined with a thorough evaluation of AK stability and folding for quantitative success to be achieved. The weak link approach provides a critically needed system for honing our models for adaptation and moving toward a more quantitative understanding of natural selection with important implications for biotechnology and biomedicine.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (Grant No. 0641792), the National Institutes of Health (Grant No. 5R01AI080714-02), and The Welch Foundation (Grant No. C1584) to Y.S.

- ¹NNIS System, *Am. J. Infect. Control* **32**, 470 (2004).
- ²L. B. Rice, *Am. J. Med.* **119**, S11 (2006).
- ³Institute of Medicine, "Antimicrobial drug resistance: Issues and options," Workshop Report (National Academy Press, 1998).
- ⁴J. E. McGowan, Jr., *Emerg. Infect. Dis.* **7**, 286 (2001).
- ⁵R. W. Haley, D. H. Culver, J. W. White, W. M. Morgan, and T. G. Emori, *Am. J. Epidemiol.* **121**, 182 (1985).
- ⁶B. Starfield, *JAMA, J. Am. Med. Assoc.* **284**, 483 (2000).
- ⁷R. Gaynes, C. Richards, J. Edwards, T. G. Emori, T. Horan, J. Alonso-Echanove, S. Fridkin, R. Lawton, G. Peavy, J. Tolson, and the National Nosocomial Infections Surveillance (NNIS) System Hospitals, *Emerg. Infect. Dis.* **7**, 295 (2001).
- ⁸S. K. Fridkin, J. C. Hageman, M. Morrison, L. Thomson Sanza, K. Como-Sabetti, J. A. Jernigan, K. Harriman, L. H. Harrison, R. Lynfield, and M. M. Farley, *N. Engl. J. Med.* **352**, 1436 (2005).
- ⁹E. M. Begier, K. Frenette, N. L. Barrett, P. Mshar, S. Petit, D. J. Boxrud, K. Watkins-Colwell, S. Wheeler, E. A. Cebelinski, A. Glennen, D. Nguyen, J. L. Hadler, and the Connecticut Bioterrorism Field Epidemiology Response Team, *Clin. Infect. Dis.* **39**, 1446 (2004).
- ¹⁰E. A. Bancroft, *JAMA, J. Am. Med. Assoc.* **298**, 1803 (2007).
- ¹¹A. L. Chapman, J. M. Greig, and J. A. Innes, *N. Engl. J. Med.* **353**, 530 (2005).
- ¹²J. T. Weber, *Clin. Infect. Dis.* **41**, S269 (2005).
- ¹³J. C. Chee-Sanford, R. I. Aminov, I. J. Krapac, N. Garrigues-Jeanjean, and R. I. Mackie, *Appl. Environ. Microbiol.* **67**, 1494 (2001).
- ¹⁴D. L. Smith, A. D. Harris, J. A. Johnson, E. K. Silbergeld, and J. G. Morris, Jr., *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6434 (2002).
- ¹⁵M. Furushita, T. Shiba, T. Maeda, M. Yahata, A. Kaneoka, Y. Takahashi, K. Torii, T. Hasegawa, and M. Ohta, *Appl. Environ. Microbiol.* **69**, 5336 (2003).
- ¹⁶Y. Agerso and L. Guardabassi, *J. Antimicrob. Chemother.* **55**, 566 (2005).
- ¹⁷S. B. Levy, G. B. FitzGerald, and A. B. Maccone, *N. Engl. J. Med.* **295**, 583 (1976).
- ¹⁸J. Bates, J. Z. Jordens, and D. T. Griffiths, *J. Antimicrob. Chemother.* **34**, 507 (1994).
- ¹⁹H. C. Wegener, F. M. Aarestrup, L. B. Jensen, A. M. Hammerum, and F. Bager, *Emerg. Infect. Dis.* **5**, 329 (1999).
- ²⁰R. Singer, R. Finch, H. Wegener, R. Bywater, J. Walters, and M. Lipsitch, *Lancet Infect. Dis.* **3**, 47 (2003).
- ²¹S. F. Elena and R. E. Lenski, *Nat. Rev. Genet.* **4**, 457 (2003).
- ²²J. E. Mroczkowska and M. Barlow, *Antimicrob. Agents Chemother.* **52**, 2340 (2008).
- ²³D. M. Weinreich, N. F. Delaney, M. A. Depristo, and D. L. Hartl, *Science* **312**, 111 (2006).
- ²⁴J. T. Bridgman, E. A. Ortlund, and J. W. Thornton, *Nature (London)* **461**, 515 (2009).
- ²⁵J. E. Barrick, D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh, D. Schneider, R. E. Lenski, and J. F. Kim, *Nature* **461**, 1243 (2009).
- ²⁶S. P. Miller, M. Lunzer, and A. M. Dean, *Science* **314**, 458 (2006).
- ²⁷R. Couñago, S. Chen, and Y. Shamoo, *Mol. Cell* **22**, 441 (2006).
- ²⁸R. Couñago and Y. Shamoo, *Extremophiles* **9**, 135 (2005).
- ²⁹D. E. Atkinson, *Biochemistry* **7**, 4030 (1968).
- ³⁰R. Couñago, C. J. Wilson, M. I. Peña, P. Wittung-Stafshede, and Y. Shamoo, *Protein Eng. Des. Sel.* **21**, 19 (2008).
- ³¹M. I. Peña, M. Davlieva, M. R. Bennett, J. S. Olson, and Y. Shamoo, "Evolutionary fates within a microbial population highlight an essential role for protein folding during natural selection," *Mol. Syst. Biol.* (to be published).
- ³²C. Miller, M. Davlieva, C. Wilson, K. White, R. Couñago, G. Wu, J. C. Myers, P. Wittung-Stafshede, and Y. Shamoo, "Experimental evolution of adenylate kinase reveals contrasting strategies towards protein thermostability," *Biophys. J.* (in press).
- ³³D. L. Hartl, D. E. Dykhuizen, and A. M. Dean, *Genetics* **111**, 655 (1985).
- ³⁴A. M. Dean, *Genetics* **139**, 19 (1995).
- ³⁵M. Lunzer, S. P. Miller, R. Felsheim, and A. M. Dean, *Science* **310**, 499 (2005).
- ³⁶H. Kacser and J. A. Burns, *Symp. Soc. Exp. Biol.* **27**, 65 (1973).
- ³⁷H. Kacser and J. A. Burns, *Genetics* **97**, 639 (1981).
- ³⁸R. Heinrich and T. A. Rapoport, *Eur. J. Biochem.* **42**, 89 (1974).
- ³⁹C. J. Marx, S. J. Van Dien, and M. E. Lidstrom, *PLoS Biol.* **3**, e16 (2005).
- ⁴⁰A. M. Dean, D. E. Dykhuizen, and D. L. Hartl, *Genet. Res.* **48**, 1 (1986).
- ⁴¹D. E. Dykhuizen, A. M. Dean, and D. L. Hartl, *Genetics* **115**, 25 (1987).
- ⁴²M. Salter, R. G. Knowles, and C. I. Pogson, *Biochem. J.* **234**, 635 (1986).
- ⁴³N. V. Torres, F. Mateo, E. Melendez-Hevia, and H. Kacser, *Biochem. J.* **234**, 169 (1986).
- ⁴⁴C. C. Glembofski, A. G. Chapman, and D. E. Atkinson, *J. Bacteriol.* **145**, 1374 (1981).
- ⁴⁵A. G. Chapman, L. Fall, and D. E. Atkinson, *J. Bacteriol.* **108**, 1072 (1971).
- ⁴⁶A. M. Dean and J. W. Thornton, *Nat. Rev. Genet.* **8**, 675 (2007).