Stability of an Active Enzyme

Robert Marsland
8.592 Final Project
(Dated: May 16, 2013)

Psychrophilic enzymes have much higher catalytic activity than close mesophilic homologues, but their native conformations are less stable against unfolding. An explicitly non-equilibrium mechanism may reduce this stability still further *in vivo* than the measured reduction at equilibrium. This mechanism is illustrated first using a simplified four-state model. Then a more general relationship between stability and reaction rate is derived based on a Jarzynski-style fluctuation theorem, which highlights the strong dependence of the stability of the native state on extremely rare fluctuations in the reaction rate.

Recent thermodynamic characterizations of extremophilic enzymes reveal a trade-off between stability and activity [3]. Psychrophilic enzymes have remarkably high catalytic activity at the low temperatures where they are found, but their native states are also less stable than those of mesophilic homologues.

Studies of this effect so far have mainly focused on determining the differences in conformational free energy landscapes between the homologues that might contribute to this effect (cf. [2]). In general there will also be another, dynamic, contribution, due to the fact that the substrate-bound state of the enzyme does not have the same thermodynamic stability as the free state. In such a case, changes in the relative steady-state populations of the substrate-bound and free sub-states of the native-configuration enzyme will modify the overall ratio between the probabilities of the native state and the unfolded state.

It is straightforward to derive the dependence of this modification on the transition rates of a Markov model of the system, but recent results in non-equilibrium statistical mechanics suggest that an analysis in terms of the statistics of entropy production may be more physically enlightening [6]. After presenting the results of the standard kinetic calculation, we will show explicitly how to write the steady-state probabilities in terms of averages of exponentials of entropy production, in the spirit of [4]. We will conclude by discussing the physical insight one might hope to gain from this, and the difficulties involved in obtaining it.

We will begin by carefully setting up the key pieces of our model, so that the derivation of the entropy-based formula for the steady state probabilities can be as transparent as possible.

SETUP

Consider a single enzyme in a solution, whose dynamics are described by a Markov-state model with $\mathcal{N}$ states. The system has constant energy E and is isolated from the rest of the world, but the container of solution is large enough to act as a heat and chemical bath with constant temperature and chemical potentials over the time scales of interest. Each of the $\mathcal{N}$ states will be composed of many sub-states, both because of the necessary coarse-graining over internal conformational states of the enzyme and because each conformational state is compatible with many different microstates of the water and solute molecules. In the limit where the stochasticisty of the dynamics comes mainly from thermal fluctuations in the bath and not from quantum fluctuations, we can treat the enzyme, water molecules and atoms of the solute molecules as classical objects, measuring the number of states by computing classical phase space volumes and dividing by $N_j!$ for each set of j identical particles. We will assume that for all the phase-space points corresponding to a coarse-grained state i the enzyme has the same internal energy ϵ_i (and thus the solution has energy $E-\epsilon_i$), and that the relaxation time for the sub-states is sufficiently faster than the relaxation time of the coarse-grained states that the distribution over the sub-states can safely be assumed to always be uniform. Thus by measuring the phase space volume Ω_i within state i, we can obtain the entropy $S_i = k_B \ln \Omega_i$.

The enzyme catalyzes a reaction in which various solute molecules X_l are converted into one another. The total numbers N_l of solute molecules of kind l in the solution is assumed to be large enough that their concentrations $[X_l]$ can be regarded as fixed over the time scales of interest, even when they are far from equilibrium. When one or more substrate molecules are bound to the enzyme, they will be treated as part of the enzyme, with their internal energy included in the state energy ϵ_i.

The phase space volume Ω_i of a coarse-grained enzyme conformation i is thus a function of the energy $E_s = E-\epsilon_i$ in the solution and of the number of molecules N_l in the solution, so we have:

$$S_i = S_i(E_s, N_l). \tag{1}$$

An arbitrary transition from state i to state j will there-

fore be accompanied by a change of entropy

$$\Delta S_{ij} = (S_j - S_i)_{E_s,N_l} - \left(\frac{\partial S_i}{\partial E_s}\right)_{N_l} \Delta\epsilon_{ij} + \left(\frac{\partial S_i}{\partial N_l}\right)_{N_{m\neq l},E_s} (\Delta N_l)_{ij} \quad (2)$$

where a summation over l is understood in every term where l appears twice. The combination of discrete and continuous variables we have here would generate some ambiguity if $\partial S_i/\partial E_s$ and $\partial S_i/\partial N_l$ depended on i, because we would have to decide whether to take derivatives of S_i or S_j. But our assumptions about the large amount of water and solute guarantee that these derivatives take on constant, i-independent values

$$\frac{\partial S_i}{\partial E_s} \equiv \frac{1}{T} \qquad \frac{\partial S_i}{\partial N_l} \equiv \frac{\mu_l}{T} \quad (3)$$

where these relations define T and μ_l. The entropy change thus becomes

$$\Delta S_{ij} = \frac{1}{T}(\Delta F_{ij} + \mu_l(\Delta N_l)_{ij}) \quad (4)$$

where $\Delta F_{ij} = \Delta\epsilon_{ij} - T(S_j - S_i)$ with E_s and N_l held fixed.

The state of an ensemble of these enzymes at time t is described by a probability distribution $p_i(t)$ over all the configurational states $i = 1, 2, \ldots N$ of the enzyme/solution system. The time-evolution of this distribution is governed by the Master equation

$$\frac{dp_i(t)}{dt} = \sum_{j=1}^{N} W_{ij} p_j(t) \quad (5)$$

where W_{ij} is the matrix of instantaneous transition rates from states j to state i for $j \neq i$, and $W_{ii} \equiv -\sum_{j\neq i} W_{ji}$ is minus the total instantaneous rate for transitions away from i.

We will be considering the case where the concentrations $[X_l]$ are far from equilibrium, and detailed balance is violated. Nevertheless, we can still write a general relation between the forward and reverse rates W_{ji} and W_{ij} by invoking the time-reversibility of the microscopic dynamics (cf. [5, p.548-553] for a more detailed treatment without any assumptions of classicality, in the special case of no particle exchange). Although our system is stochastic, if we consider the aforementioned limit where thermal fluctuations dominate and quantum fluctuations are unimportant, then on short time scales there will be a deterministic Hamiltonian evolution from each phase space point. We can therefore compute the phase space volume $\Omega(i \to j)$ of microstates of state i that deterministically evolve into microstates within state j in time dt, and divide this by the total phase space volume of state i to find the transition probability

$$W_{ji}dt = \frac{\Omega(i \to j)}{\Omega_i} \quad (6)$$

Since the evolution is Hamiltonian over this time scale, the time-symmetry of the equations of motion combined with Liouville's theorem guarantees a one-to-one correspondence between trajectories from i to j and trajectories from j to i (as long as the coarse-grained states i and j are themselves time-symmetric, so that a reversal of all the momenta does not generate a distinct state):

$$\Omega(i \to j) = \Omega(j \to i). \quad (7)$$

Thus we can conclude that

$$\frac{W_{ij}}{W_{ji}} = \frac{\Omega(j \to i)\Omega(i)}{\Omega(j)\Omega(i \to j)} = \frac{\Omega(i)}{\Omega(j)} = e^{-\Delta S_{ij}/k_B} \quad (8)$$

FOUR-STATE MODEL

Before presenting our fluctuation-theorem analysis, we will compute the steady-state probabilities directly in a simplified model where $\mathcal{N} = 4$. In three of these states, the enzyme is in its native 'well-folded' conformation (f): free enzyme ready to receive reactant ($f0$), enzyme bound to reactant (fA), and enzyme bound to product (fB). The fourth state is the 'unfolded' state (u). We allow reactions among all three folded states, and between f and u, but not between u and the substrate-bound states. This is an extreme version of the mechanism we are investigating, where the substrate-bound enzyme states fold and unfold at a rate negligibly small compared to that of the free state. Of course the opposite case is possible as well, where the substrate-bound states fold and unfold faster than the free state. The psychrophilic and mesophilic enzymes tested in [2], however, have been found to be much more stable against unfolding when bound to an analogue of the transition state of their substrate. It seems more likely, therefore, that they are also more stable rather than less stable when bound to the initial and final substrate states.

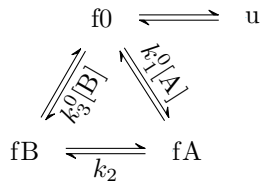

FIG. 1. Four-state model of enzyme reaction with unfolding.

From this model, we can now compute the steady state probabilities for finding the enzyme in each of the four states. One way of doing this is to first demand that the probability currents on each leg of the triangle be equal in steady state, so that no net probability flows into or

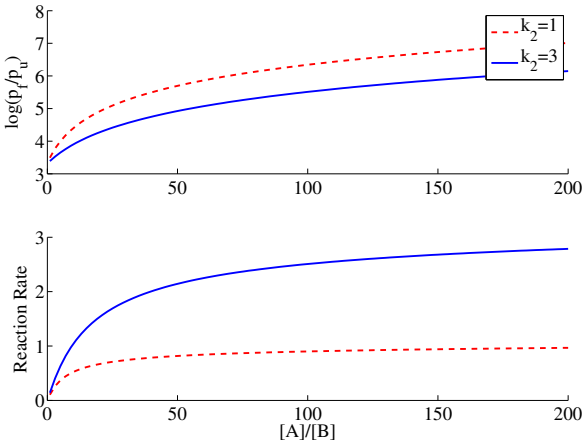

FIG. 2. Stability and activity versus reactant concentration for reaction depicted in Fig. 1. [B] is held fixed, and all rates are measured in units where $k_1^0 = 1/[B]$. We have set $k_3 \gg k_2$ to represent a reaction with a single rate-limiting step. The reactant molecules are assumed to have internal energy $\Delta\epsilon_i = 10k_BT$ greater than the product molecules, and to bind to the protein with binding energy $\Delta\epsilon_b = 1k_BT$.

out of the two bottom nodes of the triangle. This will fix the relative probabilities of the three folded states. The probability current into the unfolded state must be zero, because there is nowhere for the probability to flow out; this fixes the relative probabilities of the states u and $f0$. Thus if a change in kinetic parameters decreases the probability of fA and fB relative to $f0$, it will also decrease the overall probability of f relative to u.

Figure 2 shows how the relative probabilities depend on the ratio of concentrations of species A and B in the solution. We see that as the concentration of A is increased, the probability of f relative to u increases, as more probability is concentrated in the substrate-bound states. This is plotted for two different values of $k_2 \approx k_{cat}$ (we have set $k_3 \gg k_2$, so k_2 is the limiting rate), the faster one representing a 'psychrophilic' enzyme that processes substrates at a higher rate than a 'mesophilic' homologue operating at the same low temperature. The free energies of all four states are the same for the two curves - only the absolute rate of the reaction step is changed. The ratio of rates was chosen based on the measured k_{cat} ratio between psychrophilic and mesophilic α-amylase homologues in [2]. Also plotted for comparison is the standard saturation curve for the reaction rate. We see that at the concentration where the reaction rate is half its maximum value, the effective free energy gap $k_BT\ln(p^s s_f/p^s s_u)$ for the mesophile has already increased by $\sim k_BT = 0.6$ kcal/mol. The difference between the effective free energy gaps for the two homologues at that concentration is $\sim k_BT/2 = 0.3$ kcal/mol. These are significant modifications for a biological situation, where typical energy scales are on the order of k_BT. Our extreme modeling assumption of no unfolding from the substrate-bound state has exaggerated the effect, however, and so it is possible that it is negligible after all for a real enzyme.

TRAJECTORY AVERAGES

We now show how to write steady-state probabilities in terms of the fluctuations in the reaction rate for a single enzyme, in a way that is applicable to an arbitrarily complicated Markov model of the enzymatic reaction. A very similar result is derived in [1] under more general conditions, and applied to a simplified enzyme reaction in [4]. The derivation given here uses more restrictive assumptions, but is intended to be easier to follow and more immediately applicable to the physics at hand.

We start by discretizing trajectories of duration t into M small chunks $dt = t/M$, giving

$$\langle e^{-\Delta S/k_B}\rangle_{i\to j,t} = \frac{\sum_{\{\alpha_k\}}(\delta_{j\alpha_M} + W_{j\alpha_M}dt)e^{-\Delta S_{j\alpha_M}}(\delta_{\alpha_M\alpha_{M-1}} + W_{\alpha_M\alpha_{M-1}}dt)e^{-\Delta S_{\alpha_M\alpha_{M-1}}}\dots(\delta_{\alpha_1 i} + W_{\alpha_1 i}dt)e^{-\Delta S_{\alpha_1 i}}}{p(i\to j)}$$
$$= \frac{1}{p(i\to j)}\sum_{\{\alpha_k\}}(\delta_{\alpha_M j} + W_{\alpha_M j}dt)(\delta_{\alpha_{M-1}\alpha_M} + W_{\alpha_{M-1}\alpha_M}dt)\dots(\delta_{i\alpha_1} + W_{i\alpha_1}dt) \quad (9)$$

This equation multiplies all the transition probabilities $W_{\alpha_{i+1}\alpha_i}dt$ together for each trajectory, together with the exponential of $-\Delta S_{\alpha_i\alpha_{i+1}}$ for each transition, then sums over all trajectories. For transitions from a state i to the same state, the probabilities are given by

$$1 - \sum_{j\neq i} W_{ji}dt = 1 + W_{ii}dt.$$

In the second line we used (8), to eliminate the exponential terms and switch the order of the indices. We have normalized by dividing by the total probability $p(i\to j)$ for a transition from i to j. We can put this in matrix

form as

$$\langle e^{-\Delta S/k_B}\rangle_{i\to j,t} = \frac{\langle i|(1+Wdt)^{t/dt}|j\rangle}{p(i\to j)} \quad (10)$$

$$= \frac{\langle i|\exp(Wt)|j\rangle}{p(i\to j)} \quad (11)$$

where $|i\rangle$ and $\langle j|$ are abstract column and row vectors representing states i and j, respectively (normalized so $\langle i|j\rangle = \delta_{ij}$), and we have taken the limit as $\lambda dt \ll 1$ for the eigenvalue λ of W with smallest nonzero absolute value.

We can break the entropy change into contributions from changes in conformation and changes in numbers of solute molecules: $T\Delta S = -\Delta F + \mu_l \Delta N_l$. The former is path-independent, and so we can take it outside the average to find

$$\langle e^{-\frac{\mu_l \Delta N_l}{k_B T}}\rangle_{i\to j,t} = \frac{\langle i|\exp(Wt)|j\rangle}{p(i\to j)} e^{-\Delta F_{ij}/k_B T} \quad (12)$$

To average over all trajectories of duration t that start in i, regardless of ending state, we multiply by $p(i\to j)$ and sum over j to get

$$e^{-F_i/k_B T}\langle e^{-\frac{\mu_l \Delta N_l}{k_B T}}\rangle_{i,t} = \sum_j \langle i|\exp(Wt)|j\rangle e^{-F_j/k_B T}. \quad (13)$$

If we let the duration t of the trajectories become large enough that $|\lambda t| \gg 1$ for all nonzero eigenvalue λ of W, then since all the nonzero eigenvalues are negative, the matrix $\exp(Wt)$ will transform any normalized probability vector $|j\rangle$ into the steady-state probability distribution $|p^{ss}\rangle = \sum_j p_j^{ss}|j\rangle$. Thus we obtain

$$p_i^{ss} = \frac{e^{-F_i/k_B T}}{Z}\langle e^{-\frac{\mu_l \Delta N_l}{k_B T}}\rangle_{i,t} \quad (14)$$

where $Z = \sum_i e^{-F_i/k_B T}$.

In the limit where the differences between the μ_l are very small, and the system is close to equilibrium, we can derive some helpful intuition from equation (14) by taking the cumulant expansion and terminating at the second cumulant. For the simple case of one reactant and one product species with a chemical potential difference $\Delta\mu \ll k_B T$, this gives

$$k_B T \ln p_i^{ss} = -F_i - k_B T \ln Z - \Delta\mu\langle\Delta N\rangle_{i,t} + \frac{\Delta\mu^2}{2k_B T}\langle\Delta N^2\rangle_{i,t}^c + \mathcal{O}\left(\frac{\Delta\mu^2}{(k_B T)^2}\right) \quad (15)$$

where $\langle\Delta N^2\rangle_{i,t}^c$ is the variance in the total number of reactant molecules ΔN converted to products over the time t. This equation tells us that the steady-state probability of the native conformation will always be reduced in faster enzymes, unless the standard deviation in the number of cycles increases by at least a factor of $\sqrt{2k_B T/\Delta\mu}$ more than the mean number of cycles. In this limit, then, we see that increasing enzyme speed necessarily incurs a cost either in the conformational stability or in the reaction rate stability.

Most enzyme-catalyzed reactions in living cells are highly irreversible, however, occupying the opposite limit where $\Delta\mu \gg k_B T$, and it becomes much more difficult to obtain an intuitive interpretation of (14). Each individual reaction on a trajectory suppresses that trajectory's contribution to $\langle e^{-\mu_l \Delta N_l/k_B T}\rangle_{i,t}$ by a factor of e^{-10}, so in the large t limit the very rare trajectories with few or negative net forward reaction cycles can have a significant impact on the equilibrium probabilities.

In the four-state model presented above, termination of the series after two cumulants results in a RHS for equation (15) greater than $k_B T$, which is impossible for a properly normalized p^{ss}.

CONCLUSION

We have presented a simple model of steady-state enzyme dynamics in which faster enzymes in a nonequilibrium chemical environment are less stable than slower enzymes with the same conformational free energy landscape in the same environment. This indicates that the loss of stability in fast, psychrophilic enzymes measured in an equilibrium solution compared to slower, mesophilic enzymes is further enhanced when the enzymes are actively processing molecules in their native environment. We showed how the relationship of stability to the reaction rate could be presented in a more general way by writing the steady state probabilities in terms of the statistics of the reaction rate. In the case of a small chemical gradient, this equation showed that the stability of the native state becomes a simple function of the balance between the mean number of forward cycles of the reaction in a long time t and the width of the distribution of the number of cycles. Real biochemical reactions tend to be very far from this limit, with the steady state probability values depending crucially on the probabilities of extremely rare trajectories; this made it difficult to understand qualitatively what the relationship between chemical reaction rate statistics and stability should look like *in vivo*. Further investigation is required to understand the counterintuitively sensitive dependence of the stability of the native enzyme conformation on extremely rare fluctuations in the mean reaction rate.

[1] G.E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60:2721, 1999.

[2] S. D'Amico, J.C. Marx, C. Gerday, and G. Feller. Activity-stability relationships in extremophilic enzymes. *J. Biol. Chem.*, 278:7891–7896, 2003.

[3] C. Gerday. Psychrophily and catalysis. *Biology*, 2:719–741, 2013.

[4] H. Qian. Nonequilibrium potential function of chemically driven single macromolecules via jarzynski-type log-mean-exponential heat. *J. Phys. Chem. B*, 109:23624–23628, 2005.

[5] F. Reif. *Fundamentals of Statistical and Thermal Physics*. Waveland Press, Long Grove, Illinois, 1965.

[6] U. Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.*, 75:126001, 2012.