Statistical Mechanics of Allosteric Enzymes

Supplementary Information

Tal Einav,[†] Linas Mazutis,[‡] and Rob Phillips[*],[¶]

Department of Physics, California Institute of Technology, Pasadena, California 91125, United States, Institute of Biotechnology, Vilnius University, Vilnius, Lithuania, and Department of Applied Physics and Division of Biology, California Institute of Technology, Pasadena, California 91125, United States

E-mail: phillips@pboc.caltech.edu

---

[*]To whom correspondence should be addressed
[†]Department of Physics, California Institute of Technology, Pasadena, California 91125, United States
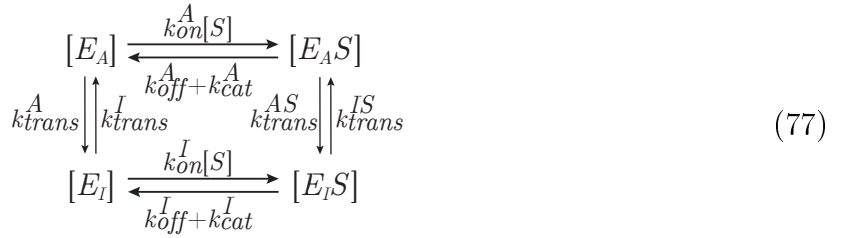[‡]Institute of Biotechnology, Vilnius University, Vilnius, Lithuania
[¶]Department of Applied Physics and Division of Biology, California Institute of Technology, Pasadena, California 91125, United States

A   Validity of Approximations

In section 2.2, we showed the generalization of the Michaelis-Menten model by granting the enzyme access to an active and inactive conformation. We then analyzed this system using two assumptions: the quasi-steady-state approximation Eq (17) and the cycle condition Eq (19). In this section, we will formally determine when these approximations are valid for an MWC enzyme and discuss what happens when we relax these assumptions. It is straightforward to extend these results to the more complicated MWC enzyme models where we introduce allosteric regulators, add competitive inhibitors, and consider enzymes with multiple binding sites.

A.1   Definitions

In section 2.2, we characterized an MWC enzyme using the reaction scheme

$$
\begin{array}{ccc}
[E_A] & \underset{k^A_{off}+k^A_{cat}}{\overset{k^A_{on}[S]}{\rightleftharpoons}} & [E_A S] \\
k^A_{trans} \Big\updownarrow \Big| k^I_{trans} & & k^{AS}_{trans} \Big\updownarrow \Big| k^{IS}_{trans} \\
[E_I] & \underset{k^I_{off}+k^I_{cat}}{\overset{k^I_{on}[S]}{\rightleftharpoons}} & [E_I S]
\end{array}
\tag{77}
$$

which we will now discuss in detail. We will use the following definitions freely:[1]
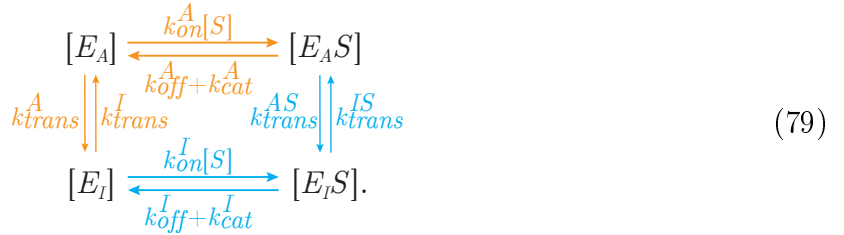
- An *edge* of a reaction scheme denotes the value of an arrow from one enzyme state to another. The edges on the left of (77) are k^A_{trans} (linking $[E_A]$ to $[E_I]$) and k^I_{trans} (linking $[E_I]$ to $[E_A]$).

- A *path* along enzyme states is the product of edges along this path. For example, the path from $[E_I]$ to $[E_A]$ to $[E_A S]$ for the MWC scheme above is given by $k^I_{trans} k^A_{on}[S]$.

- A system is in *steady state* if the concentration of every enzyme conformation does not change over time. For the scheme above this implies $\frac{d[E_A S]}{dt} = \frac{d[E_A]}{dt} = \frac{d[E_I S]}{dt} = \frac{d[E_I]}{dt} = 0$.

- The *cycle condition* states that the product of edges going clockwise around any cycle must equal the product of edges going counterclockwise. For scheme (77), the product of edges clockwise equals $\left(k^A_{on}[S]\right)\left(k^{AS}_{trans}\right)\left(k^I_{off}+k^I_{cat}\right)\left(k^I_{trans}\right)$ and the product of edges moving counter-clockwise equals $\left(k^A_{off}+k^A_{cat}\right)\left(k^A_{trans}\right)\left(k^I_{on}[S]\right)\left(k^{IS}_{trans}\right)$.

- *Detailed balance* implies that the flow between two enzyme states is the same in the forward and backwards direction. For the scheme above, if the flow of enzymes from the $[E_A]$ state to the $[E_A S]$ state (given by $[E_A][S]k^A_{on}$) equals the flow from $[E_A S]$ to $[E_A]$ (given by $[E_A S]\left(k^A_{off}+k^A_{cat}\right)$) then the pair of edges between $[E_A]$ and $[E_A S]$ obeys detailed balance. A reaction scheme is in *equilibrium* if and only if every edge obeys detailed balance which occurs if and only if the system is in steady state and obeys the cycle condition.

A.2 Cycle Condition

In this section we consider why the cycle condition is necessary to ensure that a system in steady state is in equilibrium. Assume the MWC enzyme scheme (77) is in steady state,

$$\frac{d[E_AS]}{dt} = \frac{d[E_A]}{dt} = \frac{d[E_IS]}{dt} = \frac{d[E_I]}{dt} = 0. \tag{78}$$

The cycle condition ensures that equilibrium holds around the cycle in (77) regardless of which path is traversed. For example, suppose the system is in equilibrium and we want to use detailed balance to determine the relation between E_AS and E_I. Detailed balance provides a relation between adjacent vertices (i.e. any two enzyme states connected by arrows) such as E_AS and E_IS or E_IS and E_I. Hence we can find a relation between two non-adjacent edges such as E_AS and E_I by following two different paths,

$$
\begin{array}{ccc}
[E_A] & \xrightleftharpoons[k_{off}^A + k_{cat}^A]{k_{on}^A[S]} & [E_AS] \\
{\scriptstyle k_{trans}^A} \big\downarrow \big\uparrow {\scriptstyle k_{trans}^I} & & {\scriptstyle k_{trans}^{AS}} \big\downarrow \big\uparrow {\scriptstyle k_{trans}^{IS}} \\
[E_I] & \xrightleftharpoons[k_{off}^I + k_{cat}^I]{k_{on}^I[S]} & [E_IS].
\end{array}
\tag{79}
$$

We could travel clockwise and follow the blue path around (79), first using detailed balance between E_AS and E_IS and then between E_IS and E_I,

$$\frac{[E_AS]}{[E_I]} = \frac{[E_AS]}{[E_IS]}\frac{[E_IS]}{[E_I]} = \frac{k_{trans}^{IS}}{k_{trans}^{AS}}\frac{k_{on}^I[S]}{k_{off}^I + k_{cat}^I}. \tag{80}$$
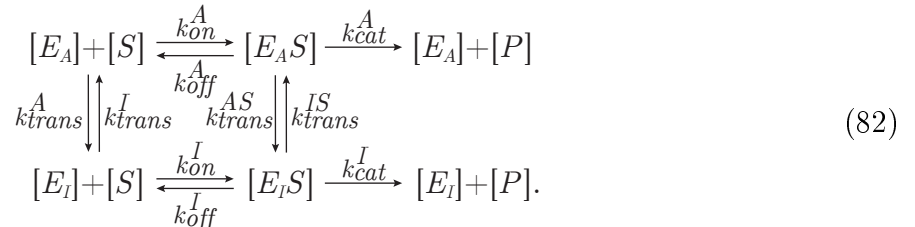
On the other hand, we could have moved counter-clockwise around (79) along the orange path, first using the relationship between E_AS and E_A and then between E_A and E_I,

$$\frac{[E_AS]}{[E_I]} = \frac{[E_AS]}{[E_A]}\frac{[E_A]}{[E_I]} = \frac{k_{on}^A[S]}{k_{off}^A + k_{cat}^A}\frac{k_{trans}^I}{k_{trans}^A}. \tag{81}$$

Setting Eqs (80) and (81) equal to each other yields the cycle condition!

A.3 Quasi-Steady-State Approximation

We will now consider the dynamics of the MWC enzyme,

$$
\begin{array}{ccccc}
[E_A]+[S] & \xrightleftharpoons[k_{off}^A]{k_{on}^A} & [E_AS] & \xrightarrow{k_{cat}^A} & [E_A]+[P] \\
{\scriptstyle k_{trans}^A} \big\downarrow \big\uparrow {\scriptstyle k_{trans}^I} & & {\scriptstyle k_{trans}^{AS}} \big\downarrow \big\uparrow {\scriptstyle k_{trans}^{IS}} & & \\
[E_I]+[S] & \xrightleftharpoons[k_{off}^I]{k_{on}^I} & [E_IS] & \xrightarrow{k_{cat}^I} & [E_I]+[P].
\end{array}
\tag{82}
$$

At time $t = 0$, the enzyme and substrate are mixed together and the rate of product formation is measured over time. The system starts off with all enzymes in the unbound forms E_A or E_I and there are no enzyme-substrate complexes $E_A S$ or $E_I S$.

To gain some intuition into this system, we first consider Figure 15 which shows how this MWC enzyme can behave over time for reasonable parameter values. On the long time scales in Figure 15B, the substrate concentration will appreciably diminish to $1/e$ of its original value after a long time τ_S. On the other hand, Figure 15A shows that within a time $\tau_E \ll \tau_S$ the enzymes reach $1/e$ of what appears to be a "steady state." Of course, this is not a true steady-state, since after a time τ_S the substrate concentration will appreciably decrease and the enzyme conformations will correspondingly change. Instead, we call the situation after one second a quasi-steady-state, meaning that the enzyme conformations have all reached a steady-state value *assuming the current substrate concentration is fixed.*

When τ_E is significantly smaller than τ_S (typically τ_E only needs to be roughly 100 times smaller than τ_S), the dynamics of the enzymes and substrate can be separated. In other words, we can assume that the fast step (where the enzymes equilibrate to the current concentration of substrate) happens instantly when considering the slow dynamics of the substrate concentration diminishing over time. This is the quasi-steady-state approximation that we formally made in Eq (17) of section 2.2. We will next show what relationship between the rate constants must hold so that the quasi-steady-state approximation is valid.
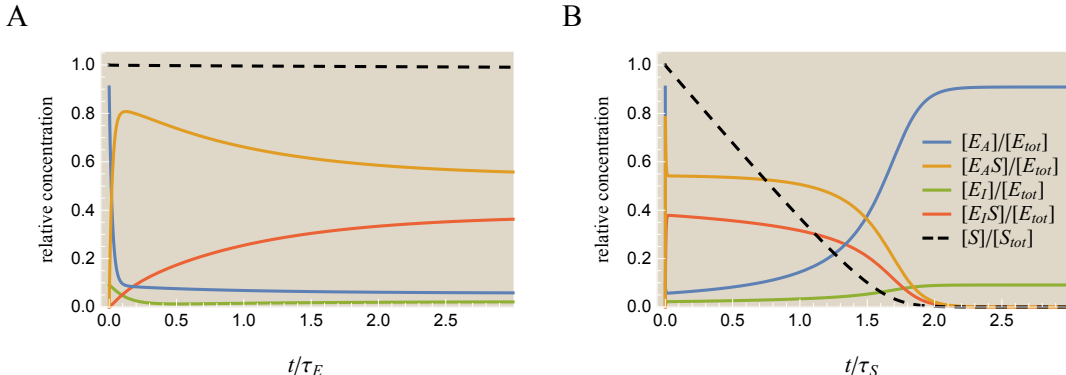

Figure 15: The quasi-steady-state approximation. (A) The fast dynamics of the system in Eq (82) begins by mixing unbound enzymes (E_A and E_I) and substrate. The enzyme conformations quickly reach steady state on a time scale of $\tau_E \approx 0.04\,\text{s}$. During this period, the substrate concentration remains very nearly constant. (B) The substrate changes appreciably over the much longer time scale $\tau_S \approx 11\,\text{s}$. Over this longer time scale, we can assume the quasi-steady-state approximation: the enzymes conformations are always in quasi-steady-state with the slowly diminishing substrate concentration. Concentrations used were $[E_{tot}] = 1\,\mu\text{M}$, $[S_{tot}] = 1\,\text{mM}$, $[E_A S] = [E_I S] = 0$, and $\frac{[E_A]}{[E_I]} = \frac{k_{trans}^I}{k_{trans}^A} \equiv e^{-\beta(\epsilon_A - \epsilon_I)}$. The rate constants used were $k_{on}^A = 1\,\text{s}^{-1}\text{M}^{-1}$, $k_{on}^I = 10^{-1}\,\text{s}^{-1}\text{M}^{-1}$, $k_{off}^A = 1\,\text{s}^{-1}$, $k_{off}^I = 10^{-3}\,\text{s}^{-1}$, $k_{cat}^A = 10^2\,\text{s}^{-1}$, $k_{cat}^I = 10\,\text{s}^{-1}$, $k_{trans}^{AS} = k_{trans}^{IS} = k_{trans}^A = 10\,\text{s}^{-1}$, and $k_{trans}^I = 10^2\,\text{s}^{-1}$.

We first calculate the time scale τ_E for the enzyme conformations to equilibrate. We will assume that the substrate concentration equals the constant value $[S_{tot}]$ throughout this short timescale (which, as shown in Figure 15A, is reasonable) and then invoke a self-consistency

condition to ensure that the actual change in substrate concentration during the period τ_E was negligible.

As a warm up, we first consider the Michaelis-Menten enzyme which we redraw here

$$[E]+[S] \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [ES] \xrightarrow{k_{cat}} [E]+[P]. \tag{83}$$
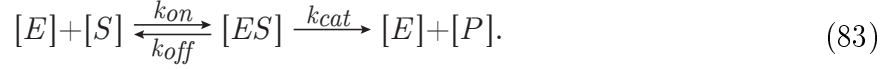
The Michaelis-Menten enzyme is governed by the multiple differential equations

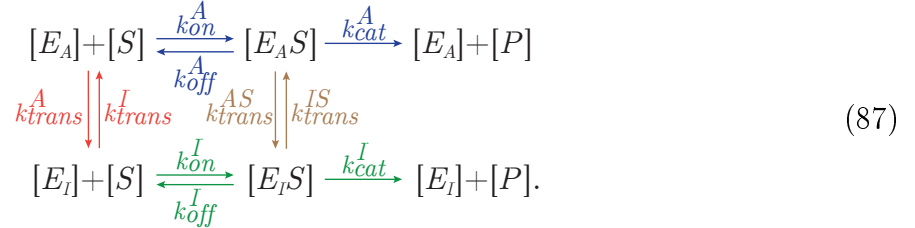$$\frac{d[E]}{dt} = [ES]\left(k_{off} + k_{cat}\right) - [E]\left[S_{tot}\right]k_{on} = -\frac{d[ES]}{dt} \tag{84}$$

and the constraint $[E] + [ES] = [E_{tot}]$. As stated above, we fix the substrate concentration at $[S_{tot}]$ and assume that the system starts off with $[E] = [E_{tot}]$ and $[ES] = 0$. Solving the differential equation Eq (84) yields

$$[E] = [E_{tot}]\frac{K_M + [S_{tot}]e^{-t/\tau}}{K_M + [S_{tot}]} \tag{85}$$

$$[ES] = [E_{tot}][S_{tot}]\frac{1 - e^{-t/\tau}}{K_M + [S_{tot}]} \tag{86}$$

where $\tau = \frac{1}{k_{on}[S_{tot}]+k_{off}+k_{cat}}$ is the time scale for the system to equilibrate. Interestingly, $\frac{1}{\tau}$ equals the sum of all rates between the states $[E]$ and $[ES]$ (i.e. the sum of all time scales in this system). Furthermore, τ does not depend on the initial conditions of the system.

We now turn to the harder case of the MWC enzyme whose kinetics we describe using the scheme

$$\begin{array}{c} [E_A]+[S] \underset{k_{off}^A}{\overset{k_{on}^A}{\rightleftharpoons}} [E_AS] \xrightarrow{k_{cat}^A} [E_A]+[P] \\ {\color{red}k_{trans}^A} \Big\updownarrow {\color{red}k_{trans}^I} \quad {\color{brown}k_{trans}^{AS}} \Big\updownarrow {\color{brown}k_{trans}^{IS}} \\ [E_I]+[S] \underset{k_{off}^I}{\overset{k_{on}^I}{\rightleftharpoons}} [E_IS] \xrightarrow{k_{cat}^I} [E_I]+[P]. \end{array} \tag{87}$$

As we just saw for the Michaelis-Menten enzyme, if we just considered any edge of the MWC enzyme separately, its corresponding time constant would be $\frac{1}{\text{sum of rates along this edge}}$: $\frac{1}{k_{on}^A[S_{tot}]+k_{off}^A+k_{cat}^A}$ between $[E_A]$ and $[E_AS]$ (blue); $\frac{1}{k_{trans}^A+k_{trans}^I}$ between $[E_A]$ and $[E_I]$ (red); $\frac{1}{k_{on}^I[S_{tot}]+k_{off}^I+k_{cat}^I}$ between $[E_I]$ and $[E_IS]$ (green); and $\frac{1}{k_{trans}^{AS}+k_{trans}^{IS}}$ between $[E_AS]$ and $[E_IS]$ (brown). We can approximate the time scale τ_E of this system as the maximum of these four time scales between adjacent edges,

$$\tau_E \approx \max \left(\frac{1}{k_{on}^A[S_{tot}] + k_{off}^A + k_{cat}^A}, \frac{1}{k_{trans}^A + k_{trans}^I}, \frac{1}{k_{on}^I[S_{tot}] + k_{off}^I + k_{cat}^I}, \frac{1}{k_{trans}^{AS} + k_{trans}^{IS}} \right)$$

$$= \frac{1}{\min \left(k_{trans}^A + k_{trans}^I, k_{on}^A [S_{tot}] + k_{cat}^A + k_{off}^A, k_{trans}^{AS} + k_{trans}^{IS}, k_{on}^I [S_{tot}] + k_{cat}^I + k_{off}^I \right)}. \quad (88)$$

This result is very similar (and in fact overestimates) the exact derivation of τ_E discussed in the next section, Appendix A.4.

With this form of τ_E in hand, we could proceed in several ways to determine when the quasi-steady-state approximation holds. For example, we could compute the time scale τ_S for the substrate to diminish and then enforce $\tau_E \ll \tau_S$ as the quasi-steady-state approximation. However, Segel and Slemrod[2] determined a tighter constraint by demanding that the amount of substrate converted into product during the transient period $0 < t < \tau_E$ only amounts to a tiny fraction of the initial substrate concentration. The amount of substrate turned into product $\Delta[S]$ after time τ_E can be overestimated as

$$\Delta[S] \approx \left| \frac{d[S]}{dt} \right|_{\max} \tau_E \quad (89)$$

so that the quasi-steady-state approximation can be written as

$$\frac{\Delta[S]}{[S_{tot}]} \approx \frac{1}{[S_{tot}]} \left| \frac{d[S]}{dt} \right|_{\max} \tau_E \ll 1. \quad (90)$$

From (82), the rate of change of substrate concentration for the MWC enzyme is

$$\frac{d[S]}{dt} = - [E_A] [S] k_{on}^A - [E_I] [S] k_{on}^I + [E_A S] k_{off}^A + [E_I S] k_{off}^I. \quad (91)$$

Recall that at $t = 0$, the system starts off with all enzymes unbound: $[E_A S] = [E_I S] = 0$ and $[E_A] + [E_I] = [E_{tot}]$. Then $\left| \frac{d[S]}{dt} \right|_{\max}$ occurs at $t = 0$ (when $[S] = [S_{tot}]$) and an upper bound is given by

$$\left| \frac{d[S]}{dt} \right|_{\max} = [S_{tot}] \left([E_A] k_{on}^A + [E_I] k_{on}^I \right) \leq [E_{tot}] [S_{tot}] \max \left(k_{on}^A, k_{on}^I \right). \quad (92)$$

Substituting this result and the time scale Eq (88) into Eq (90), we find a sufficient condition for the quasi-steady state approximation to hold for an MWC enzyme:

$$[E_{tot}] \frac{\max \left(k_{on}^A, k_{on}^I \right)}{\min \left(k_{trans}^A + k_{trans}^I, k_{on}^A [S_{tot}] + k_{cat}^A + k_{off}^A, k_{trans}^{AS} + k_{trans}^{IS}, k_{on}^I [S_{tot}] + k_{cat}^I + k_{off}^I \right)} \ll 1. \quad (93)$$

We could repeat this analysis for a Michaelis-Menten enzyme where only the E_A and $E_A S$ states exist. This is equivalent to disregarding all terms except for k_{on}^A, k_{off}^A, and

k_{cat}^A in the max and min of Eq (93), so that the quasi-steady-state conditions reduces to $[E_{tot}] \frac{k_{on}^A}{k_{on}^A [S_{tot}] + k_{cat}^A + k_{off}^A} = \frac{[E_{tot}]}{[S_{tot}] + K_M^A} \ll 1$ which is identical to the condition found by Segel.[2]

A.4 Time Constants for the Quasi-Steady-State Approximation

In this section we derive an exact expression for the time constant for which the MWC enzyme (16) will attain its steady state for each enzyme conformation assuming that the substrate concentration $[S] = [S_{tot}]$ remains fixed. The rate of change of each enzyme conformation can be written in matrix form (with bold denoting vectors and matrices) as

$$\frac{d\boldsymbol{E}}{dt} = \boldsymbol{K}\boldsymbol{E} \tag{94}$$

where

$$\boldsymbol{K} = \begin{pmatrix} k_{cat}^A + k_{off}^A & -k_{on}^A [S_{tot}] - k_{trans}^A & 0 & k_{trans}^I \\ -k_{cat}^A - k_{off}^A - k_{trans}^{AS} & k_{on}^A [S_{tot}] & k_{trans}^{IS} & 0 \\ 0 & k_{trans}^A & k_{cat}^I + k_{off}^I & -k_{trans}^I - k_{on}^I [S_{tot}] \\ k_{trans}^{AS} & 0 & -k_{cat}^I - k_{off}^I - k_{trans}^{IS} & k_{on}^I [S_{tot}] \end{pmatrix}, \boldsymbol{E} = \begin{pmatrix} [E_A S] \\ [E_A] \\ [E_I S] \\ [E_I] \end{pmatrix}. \tag{95}$$

This matrix can be decomposed as

$$\boldsymbol{K} = \boldsymbol{V}^{-1} \boldsymbol{\Lambda} \boldsymbol{V} \tag{96}$$

where $\boldsymbol{V}$'s columns are the eigenvectors of $\boldsymbol{K}$ and $\boldsymbol{\Lambda}$ is a diagonal matrix whose entries are the eigenvalues of $\boldsymbol{K}$. In general, it is known that the eigenvalues of such a matrix $\boldsymbol{K}$ representing the dynamics of any graph such as (16) from the text has one eigenvalue that is 0 while the remaining eigenvalues are non-zero and have negative real parts.[3] (Indeed, because all of the columns of $\boldsymbol{K}$ add up to zero, $\boldsymbol{K}$ is not full rank and hence one of its eigenvalues must be zero.) Defining the vector

$$\tilde{\boldsymbol{E}} \equiv \boldsymbol{V}\boldsymbol{E} = \begin{pmatrix} \tilde{E}_1 \\ \tilde{E}_2 \\ \tilde{E}_3 \\ \tilde{E}_4 \end{pmatrix}, \tag{97}$$

Eq (94) can be rewritten as

$$\frac{d\tilde{\boldsymbol{E}}}{dt} = \boldsymbol{\Lambda}\tilde{\boldsymbol{E}}. \tag{98}$$

If the eigenvalues of Λ are λ_1, λ_2, λ_3, and 0, then $\tilde{E}_j = c_j e^{\lambda_j t}$ for $j = 1, 2, 3$ and $\tilde{E}_4 = c_4$ where the c_j's are constants determined by initial conditions. Since the $\tilde{E}_j$'s are linear combinations of $[E_A S]$, $[E_A]$, $[E_I S]$, and $[E_I]$, this implies that the $\frac{1}{\lambda_1}$, $\frac{1}{\lambda_2}$, and $\frac{1}{\lambda_3}$ are the time scales for the system to come to equilibrium. Therefore, we can compute the overall time scale for the
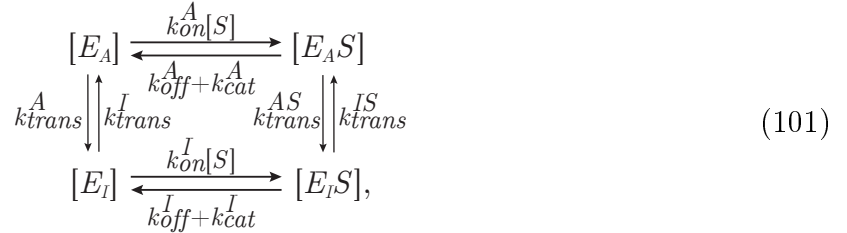
system to come to equilibrium as

$$\tau_E^{(exact)} = \max\left(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \frac{1}{\lambda_3}\right). \tag{99}$$

Although the eigenvalues of this matrix can be calculated in closed form, they are long and complicated expressions that contribute less intuition than the approximation

$$\tau_E = \max\left(\frac{1}{k_{on}^A[S] + k_{off}^A + k_{cat}^A}, \frac{1}{k_{trans}^A + k_{trans}^I}, \frac{1}{k_{on}^I[S] + k_{off}^I + k_{cat}^I}, \frac{1}{k_{trans}^{AS} + k_{trans}^{IS}}\right) \tag{100}$$

used in Eq (88) in the text. However, given the exact form, we can compare how well our approximation Eq (100) matches the exact form Eq (99).

When the four time scales in Eq (100) are comparable to each other, the approximation is very close to the exact form. However, when at least one pair of edges in the MWC enzyme rates diagram,

$$
\begin{array}{ccc}
[E_A] & \xLeftrightarrow[k_{off}^A + k_{cat}^A]{k_{on}^A[S]} & [E_A S] \\
k_{trans}^A \Big\updownarrow k_{trans}^I & & k_{trans}^{AS} \Big\updownarrow k_{trans}^{IS} \\
[E_I] & \xLeftrightarrow[k_{off}^I + k_{cat}^I]{k_{on}^I[S]} & [E_I S],
\end{array}
\tag{101}
$$

is very small the approximation tends to overshoot the exact value of τ_E. For example, if $k_{trans}^A \approx k_{trans}^I \approx 0$), Eq (100) implies $\tau_E \to \infty$ whereas Eq (99) can remain finite.

A.5 Generalizing the Cycle Condition

We now consider what happens if an enzyme does not obey the cycle condition. Provided that the quasi-steady-state approximation holds, then on the long time scales the enzyme conformations quickly equilibrate to the current substrate concentration. From (82), the rate of change of each enzyme species obeys

$$\frac{d[E_A S]}{dt} = 0 = [E_A][S]k_{on}^A - [E_A S]\left(k_{off}^A + k_{cat}^A + k_{trans}^{AS}\right) + [E_I S]k_{trans}^{IS} \tag{102}$$

$$\frac{d[E_A]}{dt} = 0 = [E_A S]\left(k_{off}^A + k_{cat}^A\right) - [E_A][S]k_{on}^A - [E_A]k_{trans}^A + [E_I]k_{trans}^I \tag{103}$$

$$\frac{d[E_I S]}{dt} = 0 = [E_I][S]k_{on}^I - [E_I S]\left(k_{off}^I + k_{cat}^I + k_{trans}^{IS}\right) + [E_A S]k_{trans}^{AS} \tag{104}$$

$$\frac{d[E_I]}{dt} = 0 = [E_I S]\left(k_{off}^I + k_{cat}^I\right) - [E_I][S]k_{on}^I - [E_I]k_{trans}^I + [E_A]k_{trans}^A. \tag{105}$$

This system of equations, together with the conservation of total enzyme, $[E_{tot}] = [E_A S] + [E_I S] + [E_A] + [E_I]$, can be solved to obtain the quasi-steady-state values of each enzyme species. Using the Michaelis constants $K_M^A = \frac{k_{off}^A + k_{cat}^A}{k_{on}^A}$ and $K_M^I = \frac{k_{off}^I + k_{cat}^I}{k_{on}^I}$, we can write the

solutions as the three ratios

$$\frac{[E_A S]}{[E_A]} = \frac{[S]}{K_M^A} \frac{\left(K_M^I k_{on}^I + k_{trans}^I \gamma + [S]k_{on}^I \gamma\right) + k_{trans}^A \alpha\gamma}{\left(K_M^I k_{on}^I + k_{trans}^I \gamma + [S]k_{on}^I \gamma\right) + k_{trans}^A \alpha\beta\delta} \tag{106}$$

$$\frac{[E_I S]}{[E_I]} = \frac{[S]}{K_M^I} \frac{\left(K_M^A k_{on}^A + k_{trans}^A \delta + [S]k_{on}^A \delta\right) + k_{trans}^I \frac{\delta}{\alpha}}{\left(K_M^A k_{on}^A + k_{trans}^A \delta + [S]k_{on}^A \delta\right) + k_{trans}^I \frac{\gamma}{\alpha\beta}} \tag{107}$$

$$\frac{[E_A]}{[E_I]} = \frac{k_{trans}^I}{k_{trans}^A} \frac{\left(K_M^I k_{on}^I + k_{trans}^I \gamma + k_{trans}^A \alpha\beta\delta\right) + [S]k_{on}^I \gamma}{\left(K_M^I k_{on}^I + k_{trans}^I \gamma + k_{trans}^A \alpha\beta\delta\right) + [S]k_{on}^I \beta\delta} \tag{108}$$

where we have defined $\alpha \equiv \frac{k_{on}^I}{k_{on}^A}$, $\beta \equiv \frac{K_M^I}{K_M^A}$, $\gamma \equiv \frac{k_{trans}^{IS}}{k_{trans}^I}$, $\delta \equiv \frac{k_{trans}^{AS}}{k_{trans}^A}$ to simplify the results. Notice that the terms in parenthesis in the numerator and denominator of these three ratios are the same. Indeed, the large fractions in all three equations equal 1 if we set $\gamma = \beta\delta$ so that

$$\frac{[E_A S]}{[E_A]} = \frac{[S]}{K_M^A} \tag{109}$$

$$\frac{[E_I S]}{[E_I]} = \frac{[S]}{K_M^I} \tag{110}$$

$$\frac{[E_A]}{[E_I]} = \frac{k_{trans}^I}{k_{trans}^A}. \tag{111}$$

This fortuitous choice of γ is equivalent to the cycle condition Eq (19), and so it is no surprise that these three ratios match Eqs (20)-(22).

Invoking the cycle condition is a theoretical convenience which greatly simplifies our equations. If the cycle condition does not hold, we can follow our same procedure to turn Eqs (106)-(108) into a more general result for states and weights by only assuming the quasi-steady-state approximation. While this more general procedure is straightforward to implement numerically, it comes at the cost of introducing more parameters into the model (for example, values for k_{on}^I and k_{trans}^I must now be explicitly given whereas before we only needed to determine the ratios $\frac{k_{on}^I}{k_{off}^I + k_{cat}^I}$ and $\frac{k_{trans}^I}{k_{trans}^A}$) and the parameters will now depend upon the substrate concentration.

Finally, we note that the cycle condition need not be invoked if a model does not contain any cycles. In other words, if we instead defined an MWC enzyme using the rates diagram

$$
\begin{array}{c}
[E_A]+[S] \underset{k_{off}^A}{\overset{k_{on}^A}{\rightleftharpoons}} [E_A S] \xrightarrow{k_{cat}^A} [E_A]+[P] \\
k_{trans}^A \Big\updownarrow k_{trans}^I \\
[E_I]+[S] \underset{k_{off}^I}{\overset{k_{on}^I}{\rightleftharpoons}} [E_I S] \xrightarrow{k_{cat}^I} [E_I]+[P],
\end{array}
\tag{112}
$$
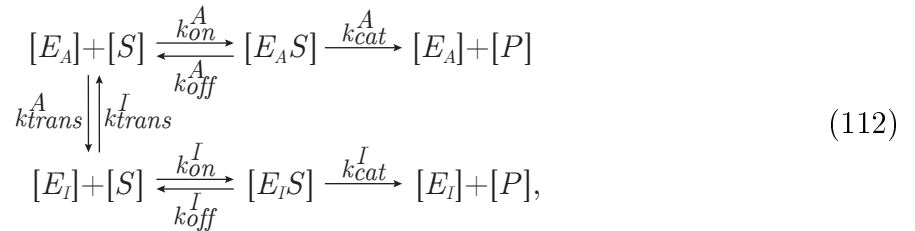
our analysis would proceed identically without needing to invoke the cycle condition. Therefore, the cycle condition ensures that the system (82) has the right value of $\frac{k_{trans}^{AS}}{k_{trans}^{IS}}$ so that it can operate identically to (112).

B   General Enzyme Models

In this section, we discuss the procedure used to fit the experimental enzyme kinetics data to the theoretical framework we have developed for allosteric enzymes. We then discuss the individual fits for each enzyme considered throughout the paper. These fits may also be viewed directly in the supplementary *Mathematica* notebook which contains the code to generate all of the plots in the paper as well as the experimental data for each enzyme.

All fitting was done using nonlinear regression (NonlinearModelFit in *Mathematica*) using the realistic constraints $K_M, C_D, R_D \in [10^{-2}\mu M, 10^6 \mu M]$, $k_{cat} \in [10^{-2}\,s^{-1}, 10^5\,s^{-1}]$, and $e^{-\beta(\epsilon_A - \epsilon_I)} \in [-10, 10]$.[4] Initial conditions for the nonlinear regression were chosen randomly from this parameter space until a sufficiently good fit ($R^2 > 0.99$) was found.

It must be noted that, as with nearly all models, there are serious ambiguities in the best fit values since multiple sets of best fits values yield nearly identical curves. In point of fact, if the nonlinear regression would be performed without any constraints, it nearly always lands outside of the physically relevant parameter space (although the qualitative form of the best fit curves may be nearly indistinguishable from those that we show below). This attribute of models, dubbed as "sloppiness," is well known.[5] One of its implications may be that a biological system can more easily evolve whichever activity profile it requires to maximize fitness, since the system is more likely the stumble across the best possible activity profile if it exists for numerous sets of parameters.

With this in mind, our results below demonstrate that our framework is *sufficient* to describe the complex interactions of allosteric enzymes, but that the individual parameter values (i.e. K_M, C_D, R_D values) are *not tightly determined* by these fits.

B.1   Fitting α-Amylase and Allosteric Regulator Chlorine

Figure 16 shows three activity curves for *A. haloplanctis* α-amylase titrating substrate at different concentrations of the allosteric activator NaCl. This enzyme has one substrate binding site and one allosteric site for binding chlorine ions. As discussed in section 3.1 of the main text, the $[S]/A$ curves are linear in $[S]$,

$$\frac{[S]}{A} = \frac{e^{-\beta\epsilon_A}\left(1 + \frac{[S]}{K_M^A}\right)\left(1 + \frac{[R]}{R_D^A}\right) + e^{-\beta\epsilon_I}\left(1 + \frac{[S]}{K_M^I}\right)\left(1 + \frac{[R]}{R_D^I}\right)}{k_{cat}^A e^{-\beta\epsilon_A}\frac{1}{K_M^A}\left(1 + \frac{[R]}{R_D^A}\right) + k_{cat}^I e^{-\beta\epsilon_I}\frac{1}{K_M^I}\left(1 + \frac{[R]}{R_D^I}\right)}. \tag{113}$$

Note that we are fitting the 7 parameters from this equation into a linear form with 2 parameters (i.e. slope and intercept). Therefore, the individual parameters are not themselves reliable; instead, these fits are intended to show that the MWC model can account for the observed enzyme behavior. One possible set of parameters that matches the data is given by $e^{-\beta(\epsilon_A - \epsilon_I)} = 7.8 \times 10^{-4}$, $K_M^A = 0.6\,mM$, $K_M^I = 0.2\,mM$, $R_D^A = 0.03\,mM$, $R_D^I = 7.9\,mM$, $k_{cat}^A = 14\,s^{-1}$, and $k_{cat}^I = 0.01\,s^{-1}$. To find the value of an individual parameter, we would instead setup an experiment where only that single parameter varies and fit the resulting data.
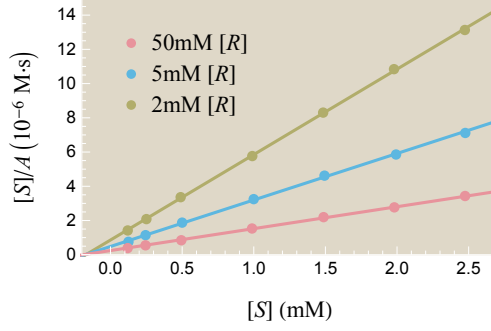
Figure 16: Theoretically and experimentally probing the effects of an allosteric regulator on activity. Data points show experimentally measured activity from Feller et al. for the enzyme α-amylase using substrate analog $[S]$ (EPS) and allosteric activator $[R]$ (NaCl), overlaid by theoretical curves of the form given in Eq (113).[6] Reproduced from Figure 10 in the main text.

B.2 Fitting α-Amylase and Competitive Inhibitor Isoacarbose

Figure 17 shows three activity curves of human pancreatic α-amylase titrating competitive inhibitor at different substrate concentrations. This enzyme has one active site which the substrate or competitive inhibitor can bind to. As discussed in section 3.1 of the main text, the activity curves all take the form

$$\left(\frac{d[P]}{dt}\right)^{-1} = \frac{1}{[E_{tot}]} \frac{e^{-\beta\epsilon_A}\left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right) + e^{-\beta\epsilon_I}\left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)}{k_{cat}^A e^{-\beta\epsilon_A}\frac{[S]}{K_M^A} + k_{cat}^I e^{-\beta\epsilon_I}\frac{[S]}{K_M^I}} \tag{114}$$

which is linear in $[C]$.

As noted above, in fitting 6 parameters to a linear form, the best fit parameter values are not reliable, but are only intended to show that the MWC model can account for the observed enzyme behavior. One possible set of parameters that matches the data is given by $e^{-\beta(\epsilon_A - \epsilon_I)} = 36$, $K_M^A = 0.9\,\text{mM}$, $K_M^I = 2.6\,\text{mM}$, $C_D^A = 12\,\text{nM}$, $C_D^I = 260\,\text{nM}$, and $\frac{k_{cat}^A}{k_{cat}^I} = 1.4$. Because units for activity were not included in original data, we instead fit the dimensionless quantity $[E_{tot}]k_{cat}^A\left(\frac{d[P]}{dt}\right)^{-1}$ which rescales the y-axis but does not change the form of the activity curves.[7]

B.3 Fitting Acetylcholinesterase Data

The acetylcholinesterase data in Figure 18 was taken from *Torpedo marmorata*.[8] Using our framework from section 2.6, activity is given by

$$A = N \frac{e^{-\beta(\epsilon_A - \epsilon_I)}k_{cat}^A\frac{[S]}{K_M^A}\left(1 + \frac{[S]}{K_M^A}\right)^{N-1} + k_{cat}^I\frac{[S]}{K_M^I}\left(1 + \frac{[S]}{K_M^I}\right)^{N-1}}{e^{-\beta(\epsilon_A - \epsilon_I)}\left(1 + \frac{[S]}{K_M^A}\right)^N + \left(1 + \frac{[S]}{K_M^I}\right)^N} \tag{115}$$
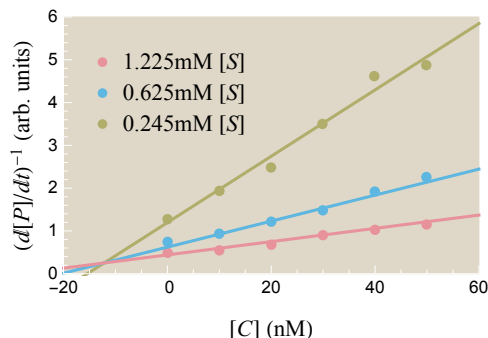
Figure 17: Theoretically and experimentally probing the effects of a competitive inhibitor on activity. Data points show experimentally measured activity in arbitrary units from Li et al. for the enzyme α-amylase using substrate analog $[S]$ (α-maltotriosyl fluoride) and competitive inhibitor $[C]$ (isoacarbose), overlaid by theoretical curves of the form given by Eq (114).[7] Best fit theoretical curves described by the inverse of Eq (65) are overlaid on the data. Reproduced from Figure 11(A) in the main text.
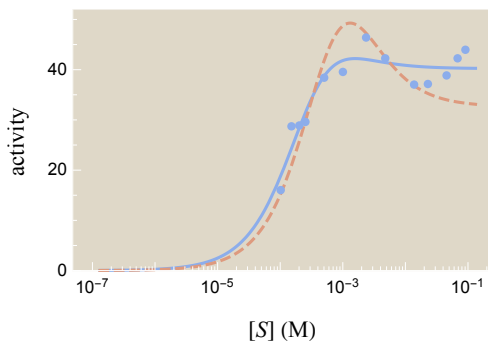
Figure 18: The activity of acetylcholinesterase exhibits a peak. Activity for acetylcholinesterase is shown in units of (nanomoles product)$\cdot$min$^{-1}\cdot$(mL enzyme)$^{-1}$.[8] The theoretical best-fit curve is shown (light blue) together with another theory curve which ignores the last three data points but better captures the height of the peak in the data (dashed, red).

where $N = 2$ is the number of active sites.

Activity is shown in units of (nanomoles product) $\cdot$ min^{-1} $\cdot$ (mL enzyme)$^{-1}$. Using the density $3.6\,\frac{\text{mg}}{\text{mL}}$ and molecular weight $2.3 \times 10^5\,\frac{\text{g}}{\text{mol}}$ of the enzyme,[8] $1\,\text{mL}$ enzyme $= 1.6 \times 10^{-8}\,\text{mol}$. Therefore, 1 unit on the y-axis of the figure corresponds to $10^{-3}\,\text{sec}^{-1}$.

The best fit parameters (light blue curve in Figure 18) were $e^{-\beta(\epsilon_A - \epsilon_I)} = 0.5$, $K_M^A = 6.1 \times 10^{-3}\,\text{M}$, $K_M^I = 2.8 \times 10^{-4}\,\text{M}$, $k_{cat}^A = 3.1\,\text{s}^{-1}$, and $k_{cat}^I = 3.7 \times 10^{-2}\,\text{s}^{-1}$. The fitting is made difficult by two factors. First, the data points are not evenly spaced, and the three data points clumped together near $[S] = 2 \times 10^{-4}\,\text{M}$ have more weight on the fit than other points. Second, we suspect that the final three data points in this figure have a significant amount of error and should not curve back up - indeed, none of the other acetylcholinesterase substrate inhibition curves from the same source exhibit this feature.[8] To that end, we also show another theoretical curve (dashed, red) in order to exemplify that

the MWC model can capture the height of the peak in the data. This latter curve has the parameters $e^{-\beta(\epsilon_A - \epsilon_I)} = 0.7$, $K_M^A = 7.4 \times 10^{-3}$ M, $K_M^I = 5.9 \times 10^{-4}$ M, $k_{cat}^A = 2.9\,\text{s}^{-1}$, and $k_{cat}^I = 2.0 \times 10^{-2}\,\text{s}^{-1}$.

B.4   Further Example Data

In this section, we present data on ATCase (not discussed in the main text) which provides an excellent opportunity to combine all of the molecular players and enzyme features we have analyzed - allosteric regulators, competitive inhibitors, multiple substrate binding sites - in one complete model.
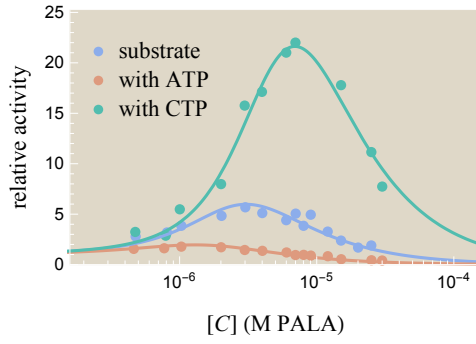
Figure 19: Inhibitor activation in aspartate carbamoyltransferase (ATCase). Activity curves from *E. coli* ATCase are shown in the absence (blue circles) and the presence of allosteric effectors, either the activator ATP (yellow squares) or the inhibitor CTP (green diamonds) as a function of the competitive inhibitor N-(phosphonacetyl)-L-aspartate (PALA). Data reproduced from Wales et al. and fit to an MWC model.[9]

The ATCase data in Figure 19 was taken from *Escherichia coli*.[9] ATCase is an allosteric enzyme with 6 active sites and 6 allosteric regulator sites. A competitive inhibitor PALA is titrated, and the experiment is then repeated in the presence of the allosteric activator ATP and the allosteric repressor CTP. Using our framework from section 2.6, the rate of product formation equals

$$\frac{d[P]}{dt} = N[E_{tot}] \frac{e^{-\beta(\epsilon_A - \epsilon_I)} k_{cat}^A \frac{[S]}{K_M^A} \left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right)^{N-1} + k_{cat}^I \frac{[S]}{K_M^I} \left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)^{N-1}}{e^{-\beta(\epsilon_A - \epsilon_I)} \left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right)^{N} + \left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)^{N}} \quad (116)$$

where $N = 6$ is the number of active sites. The plot in Figure 19 shows relative activity, which is defined as

$$\text{relative activity} = \frac{\frac{d[P]}{dt}}{\left(\frac{d[P]}{dt}\right)_{[C] \to 0}}. \quad (117)$$

All three curves were carried out at a substrate concentration $[S] = 5\,\text{mM}$ of aspartate. In the absence of allosteric effectors (blue curve), the best fit parameters were $e^{-\beta(\epsilon_A - \epsilon_I)} = 0.005$, $K_M^A = 1.1\,\text{mM}$, $K_M^I = 1.8\,\text{mM}$, $k_{cat}^A = 400\,\text{s}^{-1}$, $k_{cat}^I = 0.02\,\text{s}^{-1}$, $C_D^A = 0.3\,\mu\text{M}$, and $C_D^I =$

$1.8\,\mu\mathrm{M}$. As per the theoretical framework developed in section 2.3, an allosteric regulator such as ATP or CTP can be modeled by changing $e^{-\beta(\epsilon_A - \epsilon_I)} \rightarrow e^{-\beta(\epsilon_A - \epsilon_I)} \left(\dfrac{1 + \frac{[R]}{R_D^A}}{1 + \frac{[R]}{R_D^I}} \right)^N$ in Eq (116). From,[9] the concentrations of ATP (gold curve) and CTP (green curve) were $[R] = 2\,\mathrm{mM}$. Using the same MWC parameters as in the blue curve, the best fit parameters for the allosteric activator ATP were $R_D^A = 0.07\,\mathrm{mM}$ and $R_D^I = 0.10\,\mathrm{mM}$; the best fit parameters for the allosteric inhibitor CTP were $R_D^A = 0.14\,\mathrm{mM}$ and $R_D^I = 0.10\,\mathrm{mM}$.

C Data Collapse

In this section, we analyze the concept of data collapse, which allows us to map the result of multiple activity curves onto a single curve using natural parameters of the system. In section C.1, we start by reviewing the simplest case (presented in the main text) of an MWC enzyme with one active site in the presence of a competitive inhibitor. We show that such an enzyme admits a data collapse using a single parameter, so that all activity curves can be collapsed onto a single curve. In section C.2, we next consider the simplest MWC enzyme in the presence of an allosteric regulator, with one active site and one allosteric site. This case requires two parameters for a data collapse, and we show the resulting collapse onto a sheet. We end with a general discussion of data collapse theory in section C.3 which enables us to extend these results to more complex enzymes (e.g. enzymes with more catalytic sites in the presence of multiple species of allosteric regulators and competitive inhibitors).

C.1 Special Case: Enzyme with 1 Active Site and a Competitive Inhibitor

We start with a recap of the data collapse (discussed in section 3.1) of an enzyme with a single active site in the presence of a competitive inhibitor whose states and weights diagram is redrawn in Figure 20. The activity $A = \frac{1}{[E_{tot}]} \frac{d[P]}{dt}$ for such an enzyme is given by


Figure 20: States and weights for an MWC enzyme with an allosteric regulator. Redrawn from Figure 5 in the main text.

$$A = \frac{k_{cat}^A e^{-\beta\Delta\epsilon} \frac{[S]}{K_M^A} + k_{cat}^I \frac{[S]}{K_M^I}}{e^{-\beta\Delta\epsilon}\left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)} \tag{118}$$

where $e^{-\beta\Delta\epsilon} = e^{-\beta(\epsilon_A - \epsilon_I)}$. Dividing the numerator and denominator by $e^{-\beta\Delta\epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)$,

$$A = \frac{k_{cat}^A \left(\frac{e^{-\beta\Delta\epsilon}\frac{[S]}{K_M^A}}{e^{-\beta\Delta\epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)}\right) + k_{cat}^I \left(\frac{\frac{[S]}{K_M^I}}{e^{-\beta\Delta\epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)}\right)}{\frac{e^{-\beta\Delta\epsilon}\frac{[S]}{K_M^A}}{e^{-\beta\Delta\epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)} + \frac{\frac{[S]}{K_M^I}}{e^{-\beta\Delta\epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)} + 1}$$

$$= \frac{k_{cat}^A e^{-\beta\Delta F_{13}} + k_{cat}^I e^{-\beta\Delta F_{23}}}{e^{-\beta\Delta F_{13}} + e^{-\beta\Delta F_{23}} + 1} \tag{119}$$

where we have defined the two *Bohr parameters*,

$$\Delta F_{13} = -\frac{1}{\beta}\text{Log}\left[\frac{e^{-\beta\Delta\epsilon}\frac{[S]}{K_M^A}}{e^{-\beta\Delta\epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)}\right] \tag{120}$$

$$\Delta F_{23} = -\frac{1}{\beta}\text{Log}\left[\frac{\frac{[S]}{K_M^I}}{e^{-\beta\Delta\epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)}\right]. \tag{121}$$

Because both ΔF_{13} and ΔF_{23} have the exact same dependence on $[S]$ and $[C]$, we can characterize the system by a single natural variable. For example, since

$$e^{-\beta\Delta F_{13}} = e^{-\beta\Delta\epsilon}\frac{K_M^I}{K_M^A}e^{-\beta\Delta F_{23}} \tag{122}$$

we can rewrite Eq (121) using only ΔF_{23},

$$A = \frac{k_{cat}^A e^{-\beta\Delta\epsilon}\frac{K_M^I}{K_M^A}e^{-\beta\Delta F_{23}} + k_{cat}^I e^{-\beta\Delta F_{23}}}{e^{-\beta\Delta\epsilon}\frac{K_M^I}{K_M^A}e^{-\beta\Delta F_{23}} + e^{-\beta\Delta F_{23}} + 1}. \tag{123}$$

For cleanliness, we can group the constants using

$$K \equiv e^{-\beta\Delta\epsilon}\frac{K_M^I}{K_M^A}, \tag{124}$$

so that the activity becomes

$$A = \frac{\left(k_{cat}^A K + k_{cat}^I\right) e^{-\beta \Delta F_{23}}}{(K+1)e^{-\beta \Delta F_{23}} + 1},\tag{125}$$

matching Eq (66) from the text. As discussed in the text, this form allows us to map any number of activity curves onto a single curve of activity A versus the natural variable of the system ΔF_{23}. We redraw such a plot from the main text in Figure 21.
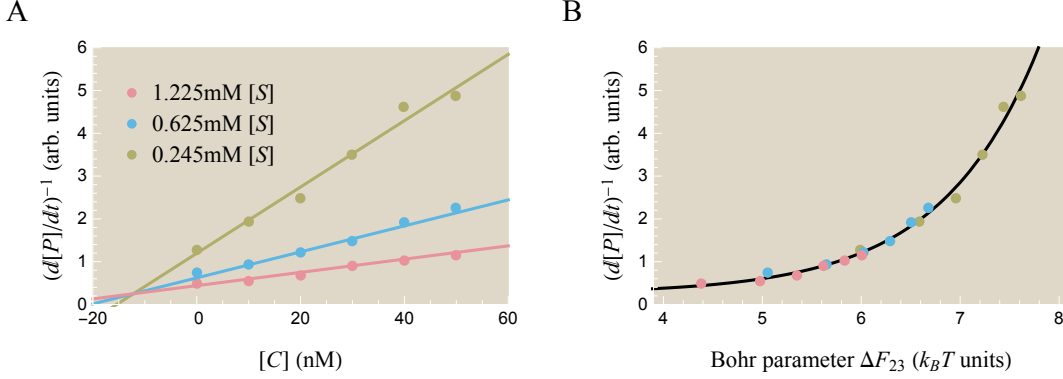

Figure 21: Data from Li et al. showing the effects of a competitive inhibitor C on the rate of product formation $\frac{d[P]}{dt}$. (A) Individual activity curves are shown at various concentrations of the substrate α-maltotriosyl fluoride (αG3F).[7] (B) Curves are all data collapsed onto a single curve using the Bohr parameter ΔF_{23} from Eq (125).

C.2 Special Case: Enzyme with 1 Active Site and an Allosteric Regulator

Consider an enzyme with one active site and one allosteric site in the presence of an allosteric regulator. The states and weights for such an enzyme are redrawn in Figure 22. The activity of such an enzyme is given by

$$A = \frac{k_{cat}^A e^{-\beta \Delta \epsilon} \frac{[S]}{K_M^A}\left(1 + \frac{[R]}{R_D^A}\right) + k_{cat}^I \frac{[S]}{K_M^I}\left(1 + \frac{[R]}{R_D^I}\right)}{e^{-\beta \Delta \epsilon}\left(1 + \frac{[S]}{K_M^A}\right)\left(1 + \frac{[R]}{R_D^A}\right) + \left(1 + \frac{[S]}{K_M^I}\right)\left(1 + \frac{[R]}{R_D^I}\right)}.\tag{126}$$

where $e^{-\beta \Delta \epsilon} = e^{-\beta(\epsilon_A - \epsilon_I)}$. We rewrite the numerator as

$$A = \frac{A_1 e^{-\beta \Delta \epsilon} \frac{[S]}{K_M^A}\left(1 + \frac{[R]}{R_D^A}\right) + A_2 \frac{[S]}{K_M^I}\left(1 + \frac{[R]}{R_D^I}\right)}{e^{-\beta \Delta \epsilon}\left(1 + \frac{[S]}{K_M^A}\right)\left(1 + \frac{[R]}{R_D^A}\right) + \left(1 + \frac{[S]}{K_M^I}\right)\left(1 + \frac{[R]}{R_D^I}\right)}\tag{127}$$

Figure 22: States and weights for an MWC enzyme with a competitive inhibitor. Redrawn from Figure 7 in the main text.

where

$$A_1 = k_{cat}^A \tag{128}$$

$$A_2 = k_{cat}^I. \tag{129}$$

Dividing the numerator and denominator by $\left(1 + \frac{[R]}{R_D^A}\right) + \left(1 + \frac{[R]}{R_D^I}\right)$, we can rewrite the activity using four natural variables,

$$A = \frac{A_1 e^{-\beta \Delta F_{13}} + A_2 e^{-\beta \Delta F_{23}}}{e^{-\beta \Delta F_{13}} + e^{-\beta \Delta F_{23}} + 1} \tag{130}$$

where

$$\Delta F_{13} = -\frac{1}{\beta} \text{Log} \left[\frac{e^{-\beta \Delta \epsilon} \frac{[S]}{K_M^A} \left(1 + \frac{[R]}{R_D^A}\right)}{e^{-\beta \Delta \epsilon} \left(1 + \frac{[R]}{R_D^A}\right) + \left(1 + \frac{[R]}{R_D^I}\right)} \right] \tag{131}$$

$$\Delta F_{23} = -\frac{1}{\beta} \text{Log} \left[\frac{\frac{[S]}{K_M^I} \left(1 + \frac{[R]}{R_D^I}\right)}{e^{-\beta \Delta \epsilon} \left(1 + \frac{[R]}{R_D^A}\right) + \left(1 + \frac{[R]}{R_D^I}\right)} \right]. \tag{132}$$

In this case, the two natural variables have a fundamentally different dependence on $[R]$ and hence cannot be combined as in the case of a competitive inhibitor. With two parameters, any number of activity curves can be collapsed down upon a surface as shown in Figure 23.
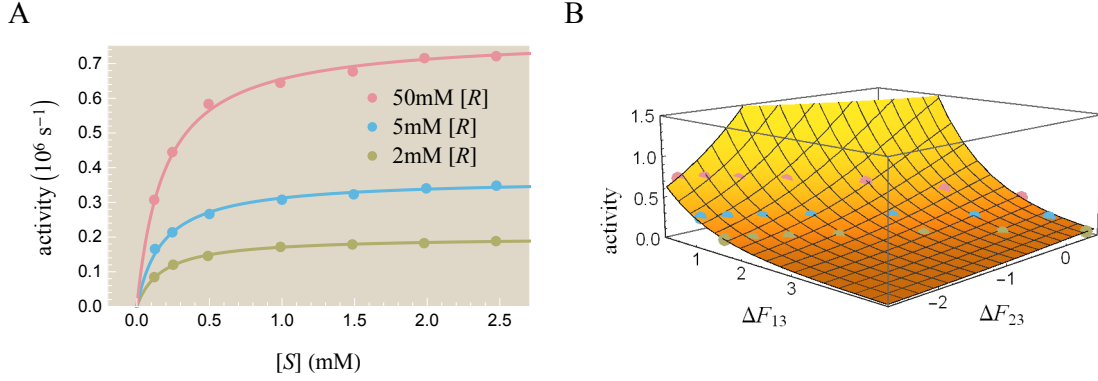
Figure 23: Data from Feller et al. demonstrating the rate of product formation $\frac{d[P]}{dt}$ in the presence of an allosteric activator NaCl. (A) Individual activity curves of α-amylase are shown at various concentrations of a substrate analog (EPS).[6] Curves reproduced from Figure 10 in main text but with the y-axis showing $\frac{d[P]}{dt}$ rather than $\frac{[S]}{d[P]/dt}$. (B) Curves are all data collapsed onto a surface using the Bohr parameters ΔF_{13} and ΔF_{23} from Eqs (131) and (132).

C.3   General Theory

We now abstract the procedure used in the previous sections in order to understand how to obtain a data collapse for any enzyme system. Suppose we enumerate all of the states and weights of an enzyme, and that all of the states pooled together only have three distinct catalytic rates A_1, A_2, and A_3. (It is straightforward to generalize this argument to any number other than three.)

Define S_1, S_2, and S_3 to be the states that have catalytic rates A_1, A_2, and A_3. Then the activity of the enzyme is given by

$$A = \frac{A_1 \sum_{j \in S_1} e^{-\beta E_j} + A_2 \sum_{j \in S_2} e^{-\beta E_j} + A_3 \sum_{j \in S_3} e^{-\beta E_j}}{\sum_{j \in S_1} e^{-\beta E_j} + \sum_{j \in S_2} e^{-\beta E_j} + \sum_{j \in S_3} e^{-\beta E_j}}. \tag{133}$$

Defining the free energies

$$e^{-\beta F_1} \equiv \sum_{j \in S_1} e^{-\beta E_j} \tag{134}$$

$$e^{-\beta F_2} \equiv \sum_{j \in S_2} e^{-\beta E_j} \tag{135}$$

$$e^{-\beta F_3} \equiv \sum_{j \in S_3} e^{-\beta E_j} \tag{136}$$

allows us to rewrite the activity as

$$
\begin{aligned}
A &= \frac{A_1 e^{-\beta F_1} + A_2 e^{-\beta F_2} + A_3 e^{-\beta F_3}}{e^{-\beta F_1} + e^{-\beta F_2} + e^{-\beta F_3}} \\
&= \frac{A_1 e^{-\beta \Delta F_{13}} + A_2 e^{-\beta \Delta F_{23}} + A_3}{e^{-\beta \Delta F_{13}} + e^{-\beta \Delta F_{23}} + 1}
\end{aligned}
\tag{137}
$$

where $\Delta F_{13} \equiv F_1 - F_3$ and $\Delta F_{23} \equiv F_2 - F_3$ are the two minimal parameters defining the system. Here, we see explicitly that each Bohr parameter corresponds to a free energy difference between combinations of states with the same activity (hence the notation ΔF).

For example, in section C.1 above, the activity of an enzyme with one active site in the presence of a competitive inhibitor is given by

$$
A = \frac{k_{cat}^A e^{-\beta \Delta \epsilon} \frac{[S]}{K_M^A} + k_{cat}^I \frac{[S]}{K_M^I}}{e^{-\beta \Delta \epsilon}\left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)}.
\tag{138}
$$

To match the form of Eq (137), we rewrite this equation as

$$
A = \frac{k_{cat}^A \left(e^{-\beta \Delta \epsilon} \frac{[S]}{K_M^A}\right) + k_{cat}^I \left(\frac{[S]}{K_M^I}\right) + 0 \left(e^{-\beta \Delta \epsilon}\left\{1 + \frac{[C]}{C_D^A}\right\} + \left\{1 + \frac{[C]}{C_D^I}\right\}\right)}{e^{-\beta \Delta \epsilon}\left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)},
\tag{139}
$$

with $A_1 = k_{cat}^A$, $A_1 = k_{cat}^I$, and $A_3 = 0$. Dividing the numerator and denominator by $e^{-\beta \Delta \epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)$ yields the data collapse equation

$$
\begin{aligned}
A &= \frac{k_{cat}^A \left(\frac{e^{-\beta \Delta \epsilon} \frac{[S]}{K_M^A}}{e^{-\beta \Delta \epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)}\right) + k_{cat}^I \left(\frac{\frac{[S]}{K_M^I}}{e^{-\beta \Delta \epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)}\right)}{\frac{e^{-\beta \Delta \epsilon} \frac{[S]}{K_M^A}}{e^{-\beta \Delta \epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)} + \frac{\frac{[S]}{K_M^I}}{e^{-\beta \Delta \epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)} + 1} \\
&= \frac{k_{cat}^A e^{-\beta \Delta F_{13}} + k_{cat}^I e^{-\beta \Delta F_{23}}}{e^{-\beta \Delta F_{13}} + e^{-\beta \Delta F_{23}} + 1}
\end{aligned}
\tag{140}
$$

with the two Bohr parameters

$$
\Delta F_{13} = -\frac{1}{\beta} \mathrm{Log}\left[\frac{e^{-\beta \Delta \epsilon} \frac{[S]}{K_M^A}}{e^{-\beta \Delta \epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)}\right]
\tag{141}
$$

$$
\Delta F_{23} = -\frac{1}{\beta} \mathrm{Log}\left[\frac{\frac{[S]}{K_M^I}}{e^{-\beta \Delta \epsilon}\left(1 + \frac{[C]}{C_D^A}\right) + \left(1 + \frac{[C]}{C_D^I}\right)}\right].
\tag{142}
$$

D   Inhibitor Acceleration: ATCase

This section will examine the phenomenon of inhibitor acceleration. The analysis will closely follow section 3.2 in the text. We first demonstrate that inhibitor acceleration (having a peak in activity as a function of competitive inhibitor concentration) cannot occur for any enzyme with one active site and then show that it can occur for an MWC enzyme with two (or more) active sites.

D.1   Inhibitor Acceleration Does Not Occur for an Enzyme with One Active Site

Consider an enzyme with a single active site in the presence of a competitive inhibitor, as in Figure 7. We start by rewriting the activity for such an enzyme from Eq (50),

$$A = \frac{1}{[E_{tot}]} \frac{d[P]}{dt} = \frac{k_{cat}^A e^{-\beta \epsilon_A} \frac{[S]}{K_M^A} + k_{cat}^I e^{-\beta \epsilon_I} \frac{[S]}{K_M^I}}{e^{-\beta \epsilon_A} \left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right) + e^{-\beta \epsilon_I} \left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)}. \tag{143}$$

The derivative of activity with respect to inhibitor concentration $[C]$ is given by

$$\frac{dA}{d[C]} = -\frac{\left(e^{-\beta \epsilon_A} \frac{1}{C_D^A} + e^{-\beta \epsilon_I} \frac{1}{C_D^I}\right) \left(e^{-\beta \epsilon_A} \frac{k_{cat}^A}{K_M^A} + e^{-\beta \epsilon_I} \frac{k_{cat}^I}{K_M^I}\right) [S]}{\left(e^{-\beta \epsilon_A} \left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right) + e^{-\beta \epsilon_I} \left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)\right)^2}. \tag{144}$$

Since the numerator cannot equal zero for any value of $[C]$, a peak cannot occur when the competitive inhibitor is added. Instead, $\frac{dA}{d[C]}$ is negative, indicating that adding more competitive inhibitor will decrease the activity, as is typically expected from an inhibitor.

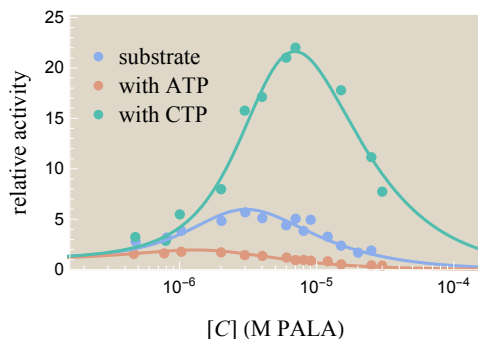## D.2   Inhibitor Acceleration for an Enzyme with Two Active Sites

Figure 24: The activity of aspartate carbamoyltransferase (ATCase) exhibits a peak. Reproduced from Figure 19.

Some allosteric enzymes exhibit an increase in activity when a small amount of competitive inhibitor C is introduced, as shown in Figure 24. The simplest enzyme model which

allows such a peak has two substrate binding sites and includes allostery. For simplicity, we work in the limit $k_{cat}^I = 0$. Combining the results from sections 2.4 and 2.5, the activity for such an enzyme is given by

$$A = k_{cat}^A e^{-\beta \epsilon_A} \frac{2 \frac{[S]}{K_M^A}\left(1 + \frac{[C]}{C_D^A} + \frac{[S]}{K_M^A}\right)}{e^{-\beta \epsilon_A}\left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right)^2 + e^{-\beta \epsilon_I}\left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)^2}. \tag{145}$$

A peak will occur provided that $\frac{dA}{d[C]} = 0$ for a positive value of $[C]$. For now, we skip the details of solving such a root (discussed in Appendix E.2) and move straight to the results. Eq (145) will have a positive root for $[C]$ provided the following relation holds,

$$e^{-\beta(\epsilon_A - \epsilon_I)} < \left(\frac{1 + \frac{[S]}{K_M^I}}{1 + \frac{[S]}{K_M^A}}\right)^2 - 2\frac{C_D^A}{C_D^I}\frac{1 + \frac{[S]}{K_M^I}}{1 + \frac{[S]}{K_M^A}} \qquad (k_{cat}^I = 0). \tag{146}$$

Acceleration by an inhibitor has historically been explained by a competitive inhibitor binding to one active site of an enzyme, forcing it into the active state.[10] This is indeed part of the story. Consider an enzyme that natively favors the inactive state when no inhibitor is present, as shown in the $[C] \ll C_D^A$ region of Figure 25. As $[C]$ increases, many enzymes will bind inhibitor in one active site, leaving the remaining active site free to bind substrate. If the inhibitor favors binding to the active-state enzyme, the ratio of active to inactive enzymes will increase which will generate a peak in activity. When $[C] \gg C_D^A$, the inhibitor will fill nearly all active sites and quash product formation. This story suggests that having a smaller $\frac{C_D^A}{C_D^I}$ value (i.e. having an inhibitor which strongly prefers binding to an active-state enzyme) will increase the likelihood of generating a peak. This is confirmed by the peak condition Eq (146) where decreasing $\frac{C_D^A}{C_D^I}$ increases the right-hand side of the inequality.

However, the complete story behind activation by inhibitor is more nuanced. To gain some intuition, we first consider the limit $\frac{C_D^A}{C_D^I} \approx 0$ where the inhibitor binds exclusively to the active rather than the inactive state. This limit maximizes the right-hand side of Eq (146) which we can rewrite as

$$e^{-\beta \epsilon_A}\left(1 + \frac{[S]}{K_M^A}\right)^2 < e^{-\beta \epsilon_I}\left(1 + \frac{[S]}{K_M^I}\right)^2 \qquad (k_{cat}^I = 0, \frac{C_D^A}{C_D^I} = 0). \tag{147}$$

This inequality tells us about the nature of the enzyme. Let us return momentarily to the states and weights of an allosteric enzyme with two substrate binding sites in the absence of competitive inhibitor which we reproduce here in Figure 26. The total weights of the enzyme being in any active state is given by the sum of the weights in the left column,

$$w_A = e^{-\beta \epsilon_A} + e^{-\beta \epsilon_A}\frac{[S]}{K_M^A} + e^{-\beta \epsilon_A}\frac{[S]}{K_M^A} + e^{-\beta \epsilon_A}\left(\frac{[S]}{K_M^A}\right)^2 = e^{-\beta \epsilon_A}\left(1 + \frac{[S]}{K_M^A}\right)^2. \tag{148}$$
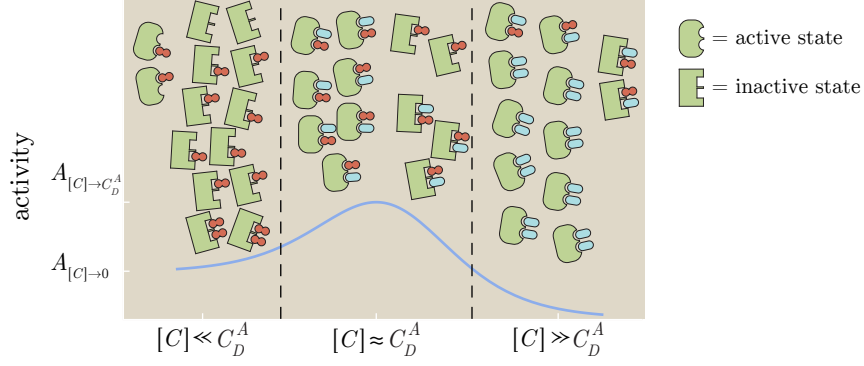
Figure 25: Mechanism underlying peak in activation by a competitive inhibitor C. At low inhibitor concentrations, $[C] \ll C_D^A$, most enzymes are in the inactive form (sharp, green). As the amount of inhibitor increases, it will begin to compete with the substrate for active sites. At medium concentrations, $[C] \gg C_D^A$, some enzymes will have one site filled with a competitive inhibitor which prefers to bind in an active-state (rounded, green) enzyme complex. This increased probability of having active-state enzyme-substrate complexes (albeit with one enzyme site filled with an inhibitor) yields a larger activity compared to the low inhibitor concentrations. At large inhibitor concentrations, $[C] \gg C_D^A$, the inhibitor outcompetes the substrate for active sites and enzyme activity is suppressed.

Figure 26: States and weights for an MWC enzyme with two substrate binding sites. Reproduced from Figure 9.

Similarly, the total weight of the enzyme being in any inactive state is given by

$$w_I = e^{-\beta \epsilon_I} + e^{-\beta \epsilon_I} \frac{[S]}{K_M^I} + e^{-\beta \epsilon_I} \frac{[S]}{K_M^I} + e^{-\beta \epsilon_I} \left(\frac{[S]}{K_M^I} \right)^2 = e^{-\beta \epsilon_I} \left(1 + \frac{[S]}{K_M^I} \right)^2. \qquad (149)$$

Therefore, the relation Eq (147) states that the total weight of the active states is smaller

than the total weight of the inactive states, $w_A < w_I$, or equivalently that the enzyme (in the absence of a competitive inhibitor) is more likely to be in an inactive state.

We now return to the more general case when $\frac{C_D^A}{C_D^I} > 0$. Recall that as $\frac{C_D^A}{C_D^I}$ increases, so does the relative affinity of the competitive inhibitor to the inactive states over the active states. We can rewrite the peak condition when $\frac{C_D^A}{C_D^I} > 0$ from Eq (146) as

$$e^{-\beta\epsilon_A}\left(1 + \frac{[S]}{K_M^A}\right)^2 < e^{-\beta\epsilon_I}\left(1 + \frac{[S]}{K_M^I}\right)^2 - \left\{2e^{-\beta\epsilon_I}\frac{C_D^A}{C_D^I}\left(1 + \frac{[S]}{K_M^I}\right)\left(1 + \frac{[S]}{K_M^A}\right)\right\} \qquad (k_{cat}^I = 0). \tag{150}$$

The term in curly braces $\{\cdots\}$ on the right is positive and increases with $\frac{C_D^A}{C_D^I}$. Compared to the special case $\frac{C_D^A}{C_D^I} = 0$ in Eq (147), an enzyme satisfying Eq (150) must favor the inactive states over the active states to a greater extent. More formally, the maximal ratio $\frac{w_A}{w_I}$ of the active state weights to inactive state weights that permits a peak decreases as $\frac{C_D^A}{C_D^I}$ increases.

Second, consider the limit $C_D^A = C_D^I$ where the competitive inhibitor equally favors the active and inactive states. According to Eq (146), a peak can still occur provided that

$$1 + e^{-\beta(\epsilon_A - \epsilon_I)} < \left(\frac{\frac{[S]}{K_M^A}}{1 + \frac{[S]}{K_M^A}}\right)^2 \left(\frac{K_M^A}{K_M^I} - 1\right)^2 \qquad (k_{cat}^I = 0,\ C_D^A = C_D^I). \tag{151}$$

It may seem surprising that an inhibitor which binds equally well to the active and inactive enzyme states can increase the amount of active state enzymes as per Figure 25. However, Eq (147) shows that any enzyme that exhibits inhibitor acceleration must favor the inactive states more in the absence of inhibitor. Relative to this pool of enzyme which are mostly in the inactive states, the presence of an inhibitor with $C_D^A = C_D^I$ will increase the fraction of enzymes in the active states.

Finally, we consider the case where introducing a competitor keeps the same fraction of enzymes in the active and inactive states, and we expect that this case cannot generate a peak in activity. Drawing on the states and weights in Figure 7 (but recalling that our enzyme has two active sites), the dissociation constants C_D^A and C_D^I of such a competitive inhibitor must satisfy

$$\frac{e^{-\beta\epsilon_A}\left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right)^2}{e^{-\beta\epsilon_I}\left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)^2} = \frac{e^{-\beta\epsilon_A}\left(1 + \frac{[S]}{K_M^A}\right)^2}{e^{-\beta\epsilon_I}\left(1 + \frac{[S]}{K_M^I}\right)^2}. \tag{152}$$

The only solution to this equation occurs when

$$\frac{C_D^A}{C_D^I} = \frac{1 + \frac{[S]}{K_M^I}}{1 + \frac{[S]}{K_M^A}}, \tag{153}$$

which upon substitution into Eq (146) yields the expected result that a peak cannot occur

when the competitive inhibitor does not change the balance between the active and inactive states. One might expect that for all values of $\frac{C_D^A}{C_D^I}$ smaller than this (where the inhibitor does push more enzymes into the active state), a peak could occur. However, Eq (146) indicates that a can only occur provided that a stronger constraint holds, namely

$$2\frac{C_D^A}{C_D^I} < \frac{1 + \frac{[S]}{K_M^I}}{1 + \frac{[S]}{K_M^A}}. \tag{154}$$

A

B
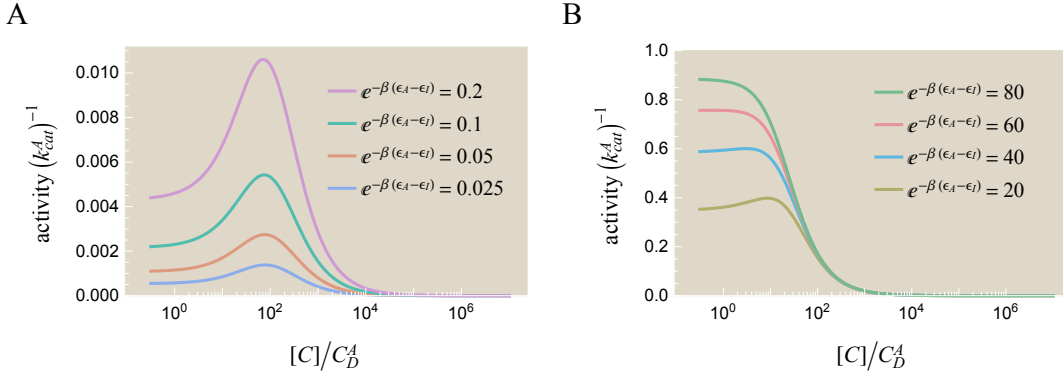

Figure 27: Peak in enzyme activity $A = \frac{1}{E_{tot}}\frac{d[P]}{dt}$ as a function of *competitive inhibitor* concentration $[C]$. As shown in Figure 12B, with Michaelis-Menten kinetics adding a competitive inhibitor can only slow down activity, but an MWC enzyme can be activated by an inhibitor which results in a peak. Peak are shown for (A) small and (B) large ratios of the enzyme's energy in the active versus inactive state, $e^{-\beta(\epsilon_A-\epsilon_I)}$. As in the case of substrate inhibition, the height of the peak increases with $e^{-\beta(\epsilon_A-\epsilon_I)}$. The activity is computed from Eq (145) using the parameters $\frac{[S]}{K_M^A} = 10$, $\frac{C_D^A}{C_D^I} = 10^{-2}$, the parameters from Figure 13, and the different values of $e^{-\beta(\epsilon_A-\epsilon_I)}$ shown. As predicted by Eq (146), for the parameters chosen every value in the range $e^{-\beta(\epsilon_A-\epsilon_I)} < 65$ will yield a peak in activity.

Having analyzed these specific cases, we now turn to some general characteristics of this peak. Having calculated the concentration $[C]_0$ in Appendix E.2 where the peak occurs, it is straightforward to compute the maximum height of the activity curve,

$$A_{peak} = k_{cat}^A \frac{[S]}{K_M^A} \frac{\left(\sqrt{\left(\frac{C_D^A}{C_D^I}\right)^2 + e^{-\beta(\epsilon_A-\epsilon_I)}} - \frac{C_D^A}{C_D^I}\right)}{\left(1 + \frac{[S]}{K_M^I}\right) - \frac{C_D^A}{C_D^I}\left(1 + \frac{[S]}{K_M^A}\right)}. \tag{155}$$

Substituting in the peak condition Eq (146) we obtain

$$A_{peak} < k_{cat}^A \frac{\frac{[S]}{K_M^A}}{1 + \frac{[S]}{K_M^A}}. \tag{156}$$

The enzyme can approach the maximum possible activity k_{cat}^A in the limit $1 \ll \frac{[S]}{K_M^A}$ when the active state enzyme dominates, analogous to the result for substrate inhibition Eq (73). We can also compare the peak height to the activity when no inhibitor is present,

$$A_{[C] \to 0} = 2k_{cat}^A \frac{e^{-\beta(\epsilon_A - \epsilon_I)} \left(\frac{[S]}{K_M^A} + \frac{[S]}{K_M^A}^2 \right)}{e^{-\beta(\epsilon_A - \epsilon_I)} \left(1 + \frac{[S]}{K_M^A}\right)^2 + \left(1 + \frac{[S]}{K_M^I}\right)^2}. \tag{157}$$

Examples of such peaks are shown in Figure 27. As in the case of substrate inhibition, the peak height A_{peak} monotonically increases and the relative peak height $\frac{A_{peak}}{A_{[C] \to 0}}$ monotonically decreases with the energy difference between the active and inactive state, $e^{-\beta(\epsilon_A - \epsilon_I)}$.

The enzyme ATCase offers an example of inhibitor acceleration. ATCase is an allosteric enzyme with 6 active sites and 6 regulatory sites.[11] In the absence of ligand, ATCase exists in an equilibrium between the unbound active and unbound inactive states, the latter being more energetically favorable.[12] When the inhibitor PALA binds to ATCase, it strongly induces a transition from inactive to active state,[13] in line with our theoretical prediction. It has been shown that by adding allosteric regulators, the peak in ATCase activity can be increased or prevented altogether.[9] It would be interesting to undertake the converse experiment and induce inhibitor activation in an enzyme that typically does not show a peak in activity.

E   Derivations

E.1   Substrate Inhibition

We now derive the general peak condition for substrate inhibition without the extra assumption $k_{cat}^I = 0$ used in the text. Recall that we define the active state of an enzyme as the state with the greater catalytic rate so that $k_{cat}^A > k_{cat}^I$. We start by rewriting the full form of the activity equation (70) from section 3.2.2,

$$A = \frac{2k_{cat}^A e^{-\beta \epsilon_A} \frac{[S]}{K_M^A} \left(1 + \frac{[S]}{K_M^A}\right) + 2k_{cat}^I e^{-\beta \epsilon_I} \frac{[S]}{K_M^I} \left(1 + \frac{[S]}{K_M^I}\right)}{e^{-\beta \epsilon_A} \left(1 + \frac{[S]}{K_M^A}\right)^2 + e^{-\beta \epsilon_I} \left(1 + \frac{[S]}{K_M^I}\right)^2}, \tag{158}$$

we derive the peak condition Eq (71). We define the numerator and denominator of the activity as

$$A \equiv \frac{Z_S}{Z_{tot}} \tag{159}$$

where, from states and weights in Figure 9,

$$Z_S = 2k_{cat}^A e^{-\beta \epsilon_A} \frac{[S]}{K_M^A} \left(1 + \frac{[S]}{K_M^A}\right) + 2k_{cat}^I e^{-\beta \epsilon_I} \frac{[S]}{K_M^I} \left(1 + \frac{[S]}{K_M^I}\right) \tag{160}$$

is the sum of all weights multiplied by their rate of product formation and

$$Z_{tot} = e^{-\beta\epsilon_A}\left(1 + \frac{[S]}{K_M^A}\right)^2 + e^{-\beta\epsilon_I}\left(1 + \frac{[S]}{K_M^I}\right)^2 \tag{161}$$

is the sum of all weights. By varying the substrate concentration $[S]$, we find a peak in the activity A provided that

$$\frac{dA}{d[S]} = \frac{\frac{dZ_S}{d[S]}Z_{tot} - Z_S\frac{dZ_{tot}}{d[S]}}{Z_{tot}^2} = 0. \tag{162}$$

Thus, a peak occurs if the numerator $\frac{dZ_S}{d[S]}Z_{tot} - Z_S\frac{dZ_{tot}}{d[S]}$ equals zero. Because Z_S and Z_{tot} are quadratic in $[S]$, the terms $\frac{dZ_S}{d[S]}Z_{tot}$ and $Z_S\frac{dZ_{tot}}{d[S]}$ in the numerator are cubic in $[S]$. However, the cubic terms exactly cancel each other, so that Eq (162) becomes a quadratic equation,

$$0 = \frac{dZ_S}{d[S]}Z_{tot} - Z_S\frac{dZ_{tot}}{d[S]} \equiv 2\left(K_M^A K_M^I\right)^4\left(a[S]^2 + b[S] + c\right), \tag{163}$$

where we have pulled out the prefactor $2\left(K_M^A K_M^I\right)^4$ for convenience and

$$a = \left(e^{-\beta\epsilon_A} + e^{-\beta\epsilon_I}\right)\left(\frac{e^{-\beta\epsilon_A}k_{cat}^A}{\left(K_M^A\right)^3} + \frac{e^{-\beta\epsilon_I}k_{cat}^I}{\left(K_M^I\right)^3}\right) \tag{164}$$

$$-e^{-\beta\epsilon_A}e^{-\beta\epsilon_I}\left(\frac{1}{K_M^A} - \frac{1}{K_M^I}\right)^2\left(\frac{k_{cat}^A}{K_M^A} + \frac{k_{cat}^I}{K_M^I}\right) \tag{165}$$

$$b = 2\left(e^{-\beta\epsilon_A} + e^{-\beta\epsilon_I}\right)\left(\frac{e^{-\beta\epsilon_A}k_{cat}^A}{\left(K_M^A\right)^2} + \frac{e^{-\beta\epsilon_I}k_{cat}^I}{\left(K_M^I\right)^2}\right) \tag{166}$$

$$c = \left(e^{-\beta\epsilon_A} + e^{-\beta\epsilon_I}\right)\left(\frac{e^{-\beta\epsilon_A}k_{cat}^A}{K_M^A} + \frac{e^{-\beta\epsilon_I}k_{cat}^I}{K_M^I}\right). \tag{167}$$

The roots of this equation are given by

$$[S]_0 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{168}$$

Since $b, c > 0$, there will only be a positive real root $[S]_0 > 0$ if

$$a < 0. \tag{169}$$

Writing this inequality out as

$$\left(e^{-\beta\epsilon_A} + e^{-\beta\epsilon_I}\right)\left(\frac{e^{-\beta\epsilon_A}k_{cat}^A}{\left(K_M^A\right)^3} + \frac{e^{-\beta\epsilon_I}k_{cat}^I}{\left(K_M^I\right)^3}\right) < e^{-\beta\epsilon_A}e^{-\beta\epsilon_I}\left(\frac{1}{K_M^I} - \frac{1}{K_M^A}\right)^2\left(\frac{k_{cat}^A}{K_M^A} + \frac{k_{cat}^I}{K_M^I}\right), \tag{170}$$

we multiply by $\frac{\left(K_M^A\right)^3}{e^{-\beta\epsilon_A}e^{-\beta\epsilon_I}k_{cat}^I}$ to obtain

$$\left(1+\frac{e^{-\beta\epsilon_A}}{e^{-\beta\epsilon_I}}\right)\left(\frac{k_{cat}^A}{k_{cat}^I}+\frac{e^{-\beta\epsilon_I}}{e^{-\beta\epsilon_A}}\left(\frac{K_M^A}{K_M^I}\right)^3\right)<\left(\frac{K_M^A}{K_M^I}-1\right)^2\left(\frac{k_{cat}^A}{k_{cat}^I}+\frac{K_M^A}{K_M^I}\right)\qquad(171)$$

and move the $\frac{k_{cat}^A}{k_{cat}^I}$ terms to one side,

$$\left(\frac{K_M^A}{K_M^I}\right)^3\left(\left(1+\frac{e^{-\beta\epsilon_I}}{e^{-\beta\epsilon_A}}\right)-\left(\frac{K_M^I}{K_M^A}-1\right)^2\right)<\frac{k_{cat}^A}{k_{cat}^I}\left(\left(\frac{K_M^A}{K_M^I}-1\right)^2-\left(1+\frac{e^{-\beta\epsilon_A}}{e^{-\beta\epsilon_I}}\right)\right).\qquad(172)$$

There are now two cases to consider. If the term on the right hand side is positive,

$$1+\frac{e^{-\beta\epsilon_A}}{e^{-\beta\epsilon_I}}<\left(\frac{K_M^A}{K_M^I}-1\right)^2,\qquad(173)$$

then we can divide by this term on both sides to obtain the peak condition

$$-\frac{\left(1+\frac{e^{-\beta\epsilon_I}}{e^{-\beta\epsilon_A}}\right)-\left(\frac{K_M^I}{K_M^A}-1\right)^2}{\left(1+\frac{e^{-\beta\epsilon_A}}{e^{-\beta\epsilon_I}}\right)-\left(\frac{K_M^A}{K_M^I}-1\right)^2}\left(\frac{K_M^A}{K_M^I}\right)^3<\frac{k_{cat}^A}{k_{cat}^I}.\qquad(174)$$

On the other hand, if the term on the right-hand side of Eq (172) is negative, then the term on the left-hand side must also be negative,

$$1+\frac{e^{-\beta\epsilon_A}}{e^{-\beta\epsilon_I}}>\left(\frac{K_M^A}{K_M^I}-1\right)^2\qquad(175)$$

$$1+\frac{e^{-\beta\epsilon_I}}{e^{-\beta\epsilon_A}}<\left(\frac{K_M^I}{K_M^A}-1\right)^2,\qquad(176)$$

and because $e^{-\beta\epsilon_A},e^{-\beta\epsilon_I},K_M^A,K_M^I>0$ this implies

$$0<\frac{K_M^A}{K_M^I}<\frac{1}{2}.\qquad(177)$$

Solving Eq (172) for $\frac{k_{cat}^A}{k_{cat}^I}$ (and flipping the sign of the inequality because of Eq (175)) yields the relation

$$-\frac{\left(1+\frac{e^{-\beta\epsilon_I}}{e^{-\beta\epsilon_A}}\right)-\left(\frac{K_M^I}{K_M^A}-1\right)^2}{\left(1+\frac{e^{-\beta\epsilon_A}}{e^{-\beta\epsilon_I}}\right)-\left(\frac{K_M^A}{K_M^I}-1\right)^2}\left(\frac{K_M^A}{K_M^I}\right)^3>\frac{k_{cat}^A}{k_{cat}^I}.\qquad(178)$$

Assuming Eq (177), the term on the left-hand side can be at most $\frac{1}{2}$, so that for an enzyme that satisfies $k_{cat}^A>k_{cat}^I$ Eq (178) can never be satisfied. Hence, for a two substrate binding site enzyme assuming $k_{cat}^I<k_{cat}^A$, a peak in activity as a function of substrate concentration

$[S]$ will occur if and only if

$$\left(1 + \frac{e^{-\beta\epsilon_A}}{e^{-\beta\epsilon_I}}\right) < \left(\frac{K_M^A}{K_M^I} - 1\right)^2 \tag{179}$$

$$-\frac{\left(1 + \frac{e^{-\beta\epsilon_I}}{e^{-\beta\epsilon_A}}\right) - \left(\frac{K_M^I}{K_M^A} - 1\right)^2}{\left(1 + \frac{e^{-\beta\epsilon_A}}{e^{-\beta\epsilon_I}}\right) - \left(\frac{K_M^A}{K_M^I} - 1\right)^2}\left(\frac{K_M^A}{K_M^I}\right)^3 < \frac{k_{cat}^A}{k_{cat}^I}. \tag{180}$$

In the text, we assumed $k_{cat}^I = 0$ so that the second condition Eq (180) is automatically satisfied and Eq (179) became the only necessary condition for a peak. In the general case when k_{cat}^I is not negligible, the second constraint Eq (180) ensures that the contribution of product formation from the inactive state does not destroy the peak which would be formed by the active states alone.

Activity curves that exhibit a peak with a non-zero k_{cat}^I value are shown in Figure 28. Although these curves look very similar to those shown in Figure 13 for the case $k_{cat}^I = 0$, one important difference is that given $K_M^{A,I}$ and $k_{cat}^{A,I}$ values, there is now a *lower* bound for $e^{-\beta(\epsilon_A - \epsilon_I)}$ given by the second peak condition Eq (180).
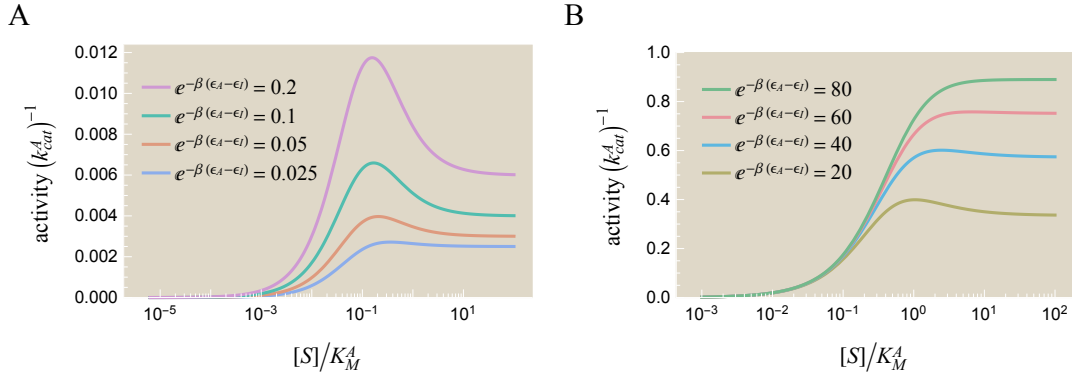

Figure 28: Peak in enzyme activity $A = \frac{1}{E_{tot}}\frac{d[P]}{dt}$ as a function of *substrate* concentration $[S]$. As shown in Figure 12A, with Michaelis-Menten kinetics adding substrate can only increase enzyme activity, but an MWC enzyme can exhibit a peak due to the interactions between the active and inactive state. Peaks are shown for (A) small and (B) large ratios of the enzyme's energy in the active versus inactive state, $e^{-\beta(\epsilon_A - \epsilon_I)}$. The activity is computed from Eq (158) using the same parameter values from Figure 13 except that $\frac{k_{cat}^A}{k_{cat}^I} = 10^3$. The curves with small $e^{-\beta(\epsilon_A - \epsilon_I)}$ values shown in (A) vary appreciably from those in Figure 13 (where $k_{cat}^I = 0$) because the inactive state catalyzes substrate. This changes both the shape and the height of the activity curves.

It is straightforward to substitute the positive root for substrate concentration Eq (168)

into the activity Eq (158) to find the height of the peak, resulting in

$$A_{peak} = \frac{k_{cat}^I K_M^A - k_{cat}^A K_M^I + \sqrt{\left(\frac{1}{e^{-\beta\epsilon_I}} + \frac{1}{e^{-\beta\epsilon_A}}\right)\left(e^{-\beta\epsilon_I}\left(k_{cat}^I K_M^A\right)^2 + e^{-\beta\epsilon_A}\left(k_{cat}^A K_M^I\right)^2\right)}}{K_M^A - K_M^I}. \quad (181)$$

In the limit $k_{cat}^I = 0$ discussed in the text, this simplifies to

$$A_{peak} = k_{cat}^A \frac{K_M^I}{K_M^A - K_M^I}\left(\sqrt{1 + \frac{e^{-\beta\epsilon_A}}{e^{-\beta\epsilon_I}}} - 1\right). \quad (182)$$

Lastly, we note that adding a fixed amount of competitive inhibitor $[C]$ to a system may induce a peak in activity as a function of substrate concentration $[S]$, as shown in Figure 29. In the language of the MWC model (Eqs (52)-(55) in the text), adding the inhibitor tunes the MWC parameters so that the peak conditions Eqs (179) and (180) apply.
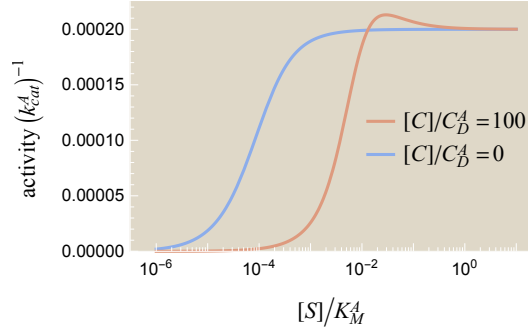

Figure 29: Peaks in activity can be induced by a competitive inhibitor. Adding a competitive inhibitor can induce a peak in activity $\frac{d[P]}{dt}$ versus substrate concentration $[S]$. Curves are shown for an enzyme with two active sites using the parameters $\frac{k_{cat}^A}{k_{cat}^I} = 10^4$, $\frac{K_M^A}{K_M^I} = 10^4$, $\frac{C_D^A}{C_D^I} = 10^{-1}$, and $e^{-\beta(\epsilon_A - \epsilon_I)} = \frac{1}{2}$.

E.2   Inhibitor Acceleration

We now derive the peak condition for Inhibitor Acceleration discussed in Appendix D.2. For an enzyme with two substrate binding sites and a competitive inhibitor C, enzyme activity is given by

$$A = k_{cat}^A \left(2p_{EAS}\right) + k_{cat}^A \left(2p_{EASC}\right) + 2k_{cat}^A \left(p_{EAS^2}\right)$$
$$+ k_{cat}^I \left(2p_{EIS}\right) + k_{cat}^I \left(2p_{EISC}\right) + 2k_{cat}^I \left(p_{EIS^2}\right). \quad (183)$$

Assuming $k_{cat}^I = 0$ for simplicity, this equation takes the form

$$A = 2k_{cat}^A e^{-\beta\epsilon_A} \frac{\frac{[S]}{K_M^A} + \frac{[S]}{K_M^A}\frac{[C]}{C_D^A} + \left(\frac{[S]}{K_M^A}\right)^2}{e^{-\beta\epsilon_A}\left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right)^2 + e^{-\beta\epsilon_I}\left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)^2}$$

$$\equiv 2k_{cat}^A e^{-\beta\epsilon_A} \frac{Z_C}{Z_{tot}} \tag{184}$$

where

$$Z_C = \frac{[S]}{K_M^A} + \frac{[S]}{K_M^A}\frac{[C]}{C_D^A} + \left(\frac{[S]}{K_M^A}\right)^2 \tag{185}$$

$$Z_{tot} = e^{-\beta\epsilon_A}\left(1 + \frac{[S]}{K_M^A} + \frac{[C]}{C_D^A}\right)^2 + e^{-\beta\epsilon_I}\left(1 + \frac{[S]}{K_M^I} + \frac{[C]}{C_D^I}\right)^2. \tag{186}$$

A peak in activity will occur provided that

$$\frac{dA}{d[C]} = 2k_{cat}^A e^{-\beta\epsilon_A} \frac{\frac{dZ_C}{d[C]}Z_{tot} - Z_C\frac{dZ_{tot}}{d[C]}}{Z_{tot}^2} = 0, \tag{187}$$

or equivalently that the numerator $\frac{dZ_C}{d[C]}Z_{tot} - Z_C\frac{dZ_{tot}}{d[C]}$ equals zero. We can rewrite the numerator as

$$0 = \frac{dZ_C}{d[C]}Z_{tot} - Z_C\frac{dZ_{tot}}{d[C]} \equiv \frac{[S]}{K_M^A}\left(a[C]^2 + b[C] + c\right) \tag{188}$$

where

$$a = -\frac{1}{C_D^A}\left(\frac{e^{-\beta\epsilon_A}}{(C_D^A)^2} + \frac{e^{-\beta\epsilon_I}}{(C_D^I)^2}\right) \tag{189}$$

$$b = -2\left(\frac{e^{-\beta\epsilon_A}}{(C_D^A)^2} + \frac{e^{-\beta\epsilon_I}}{(C_D^I)^2}\right)\left(1 + \frac{[S]}{K_M^A}\right)[C] \tag{190}$$

$$c = \frac{e^{-\beta\epsilon_I}}{C_D^A}\left(1 + \frac{[S]}{K_M^I}\right)^2 - \frac{e^{-\beta\epsilon_A}}{C_D^A}\left(1 + \frac{[S]}{K_M^A}\right)^2 - 2\frac{e^{-\beta\epsilon_I}}{C_D^I}\left(1 + \frac{[S]}{K_M^A}\right)\left(1 + \frac{[S]}{K_M^I}\right). \tag{191}$$

The roots of this equation are given by

$$[C]_0 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{192}$$

Since $a, b < 0$, there will only be a positive real root $[C]_0 > 0$ if

$$c > 0. \tag{193}$$

Therefore, the peak condition can be written as

$$2 \frac{C_D^A}{C_D^I} \left(1 + \frac{[S]}{K_M^A}\right) \left(1 + \frac{[S]}{K_M^I}\right) < \left(1 + \frac{[S]}{K_M^I}\right)^2 - \frac{e^{-\beta \epsilon_A}}{e^{-\beta \epsilon_I}} \left(1 + \frac{[S]}{K_M^A}\right)^2 \qquad (194)$$

or equivalently,

$$\frac{e^{-\beta \epsilon_A}}{e^{-\beta \epsilon_I}} < \left(\frac{1 + \frac{[S]}{K_M^I}}{1 + \frac{[S]}{K_M^A}}\right)^2 - 2 \frac{C_D^A}{C_D^I} \frac{1 + \frac{[S]}{K_M^I}}{1 + \frac{[S]}{K_M^A}} \qquad (195)$$

which matches Eq (146), as desired.

E.3   Michaelis-Menten Enzymes Do Not Exhibit Peaks

In this section, we show that a Michaelis-Menten enzyme with an arbitrary number of substrate binding sites cannot exhibit substrate inhibition nor inhibitor acceleration. This implies that the interplay between the active and inactive MWC states were necessary to produce the peaks in activity discussed in section 3.2.2 and Appendix D.2.

Consider a Michaelis-Menten enzyme with N binding sites where either a substrate S or a competitive inhibitor C can bind. Using the general formulation from section 2.6, we will assume that the enzyme only has an active state and drop the A superscripts. Each binding site can be either be empty, occupied by substrate, or occupied by competitor, which would contribute a factor of 1, $\frac{[S]}{K_M}$, or $\frac{[C]}{C_D}$, respectively, to its weight. A state with j bound substrates forms product at a rate of $j k_{cat}$. Therefore, the activity $A = \frac{1}{E_{tot}} \frac{d[P]}{dt}$ equals

$$\begin{aligned}
A &= \frac{\sum_{j=0}^{N} \sum_{k=0}^{N-j} (j k_{cat}) \frac{N!}{j! k! (N-j-k)!} \left(\frac{[S]}{K_M}\right)^j \left(\frac{[C]}{C_D}\right)^k}{\left(1 + \frac{[S]}{K_M} + \frac{[C]}{C_D}\right)^N} \\
&= N k_{cat} \frac{\frac{[S]}{K_M} \left(1 + \frac{[C]}{C_D} + \frac{[S]}{K_M}\right)^{N-1}}{\left(1 + \frac{[S]}{K_M} + \frac{[C]}{C_D}\right)^N} \\
&= N k_{cat} \frac{\frac{[S]}{K_M}}{1 + \frac{[S]}{K_M} + \frac{[C]}{C_D}}.
\end{aligned} \qquad (196)$$

Taking the derivative of the activity with respect to the substrate concentration $[S]$ and the inhibitor concentration $[C]$,

$$\frac{dA}{d[S]} = \frac{N k_{cat}}{K_M} \frac{1 + \frac{[C]}{C_D}}{\left(1 + \frac{[S]}{K_M} + \frac{[C]}{C_D}\right)^2} \qquad (197)$$

and

$$\frac{dA}{d[C]} = -\frac{Nk_{cat}}{C_D} \frac{\frac{[S]}{K_M}}{\left(1 + \frac{[S]}{K_M} + \frac{[C]}{C_D}\right)^2},$$

(198)

we find that neither derivative can be zero. Therefore, inhibitor acceleration cannot occur for a non-MWC enzyme.

References

(1) Gunawardena, J. A Linear Framework for Time-Scale Separation in Nonlinear Biochemical Systems. *PLoS One* **2012**, *7*, e36321.

(2) Segel, L. A.; Marshall, S. The Quasi-Steady-State Assumption: A Case Study in Perturbation. *Soc. Ind. Appl. Math.* **2012**, *31*, 446–477.

(3) Mirzaev, I.; Gunawardena, J. Laplacian Dynamics on General Graphs. *Bull. Math. Biol.* **2013**, *75*, 2118–2149.

(4) Phillips, R.; Milo, R. *Cell Biology by the Numbers*; Garland Science, 2015; Chapter 4.

(5) Transtrum, M. K.; Machta, B. B.; Brown, K. S.; Daniels, B. C.; Myers, C. R.; Sethna, J. P. Perspective: Sloppiness and Emergent Theories in Physics, Biology, and Beyond. *J. Chem. Phys.* **2015**, *143*, 010901.

(6) Bussy, O. l. Structural and Functional Aspects of Chloride Binding to Alteromonas haloplanctis Alpha-Amylase. *J. Biol. Chem.* **1996**, *271*, 23836–23841.

(7) Li, C.; Begum, A.; Numao, S.; Park, K. H.; Withers, S. G.; Brayer, G. D. Acarbose Rearrangement Mechanism Implied by the Kinetic and Structural Analysis of Human Pancreatic Alpha-Amylase in Complex with Analogues and their Elongated Counterparts. *Biochemistry* **2005**, *44*, 3347–3357.

(8) Changeux, J. P. Responses of Acetylcholinesterase from Torpedo marmorata to Salts and Curarizing Drugs. *Mol. Pharmacol.* **1966**, *2*, 369–392.

(9) Wales, M. E.; Madison, L. L.; Glaser, S. S.; Wild, J. R. Divergent Allosteric Patterns Verify the Regulatory Paradigm for Aspartate Transcarbamylase. *J. Mol. Biol.* **1999**, *294*, 1387–1400.

(10) Howlett, G. J.; Blackburn, M. N.; Compton, J. G.; Schachman, H. K. Allosteric Regulation of Aspartate Transcarbamoylase. *Biochemistry* **1977**, *16*, 5091–5099.

(11) Cockrell, G. M.; Zheng, Y.; Guo, W.; Peterson, A. W.; Truong, J. K.; Kantrowitz, E. R. New Paradigm for Allosteric Regulation of Escherichia coli Aspartate Transcarbamoylase. *Biochemistry* **2013**, *52*, 8036–8047.

(12) Fetler, L.; Kantrowitz, E. R.; Vachette, P. Direct Observation in Solution of a Preexisting Structural Equilibrium for a Mutant of the Allosteric Aspartate Transcarbamoylase. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 495–500.

(13) Mendes, K. R.; Kantrowitz, E. R. The Pathway of Product Release from the R State of Aspartate Transcarbamoylase. *J. Mol. Biol.* **2010**, *401*, 940–948.