7 Engineering faster enzymes

Summary. Engineering faster catalysis enables new substrates to react and reduces the amount of enzyme needed. Most enzyme-catalyzed reactions show saturation kinetics where the reaction rate initially increases with increasing substrate concentration, but later levels out and does not increase further. Two kinetic constants define this curve: V_{max}, the maximum rate of reaction at high substrate concentration, and K_M, the concentration of substrate at half of V_{max}. The definition of these constants differs for different reaction mechanisms. The simplest reaction mechanism is the Michaelis-Menten model, which proposes that the substrate first reversibly forms an enzyme-substrate complex, then reacts to form product. In most cases, the physical steps involved in enzyme catalysis are more complex. Engineering faster enzymes requires speeding up the physical steps that limit catalysis. Engineering faster enzymes may involve improving binding of the substrate to the enzyme, reorienting the substrate, adjusting the pK_a of acids and bases, relieving product inhibition or other steps in the reaction mechanism. Distant residues contribute to catalytic activity in part by altering movements that position active site residues.

Key learning goals

- Most enzymes follow saturation kinetics defined by two kinetic constants: V_{max} and K_M. The physical steps that contribute to these constants differ when the enzyme reaction mechanisms differ.
- The Michaelis-Menten model for enzyme catalysis involves with formation of the enzyme· substrate complex (E· S) followed by the chemical step that converts sub-

strate to product. Enzymes speed up this chemical step by stabilizing the transition state for the reaction.

- Two limiting cases of enzyme-catalyzed reactions are low and high substrate concentration. Here low and high are relative to the K_M of the enzyme· substrate complex. At low substrate concentration, the reaction rate can be improved either by improved binding (lower K_M), which increases the fraction of enzyme that contains a bound substrate, or by a faster chemical step (k_{cat}). At high substrate concentration all of the enzyme molecules already contain a bound substrate, so only a faster chemical step (k_{cat}) can increase the reaction rate.
- Improved binding requires increasing shape and interaction complementarity between substrate and enzyme. Some reaction steps may also contribute to K_M.
- Speeding up the conversion of the enzyme-substrate complex to product requires increased stabilization of the transition state. Depending on the reaction mechanism details, this stabilization can involve 1) pre-organizing the geometry of the E· S complex for reaction, 2) ensuring the correct protonation state of any needed acids and bases, 3) stabilizing the charge and shape of the transition state, and 4) breaking up difficult mechanistic steps into several easier ones.

7.1 Introduction

Faster enzymes are a common goal in protein engineering since faster enzymes require less protein for the application. Faster may refer to an existing substrate for the enzyme, and it may also refer to expanding the substrate range of an enzyme to new, previously unreactive, substrates.

Chemical reactions are processes that break and make chemical bonds. Catalysts can not change the Gibbs energies of the reactants and products, nor the equilibrium constant between them, Fig. 7.1. A catalyzed reaction reaches the same equilibrium state as an uncatalyzed reaction, but reaches this equilibrium more rapidly. Catalysts increase the rates of both the forward and reverse reactions.

The highest energy structure along the path from substrate to product is the transition state, Fig. 7.1. The energy gap between the substrate and transition state limits the rate of the reaction. The size of this gap, $\Delta G^{\ddagger}$, is the activation energy for the reaction. The transition state is a peak in the energy diagram, so it has a fleeting existence. Only structures that correspond to valleys on the energy diagram can exist for a finite time.

Catalysts act by stabilizing the transition state and lowering the activation energy. For example, bond breaking or making in a transition state often creates partial charges. A catalyst can orient the substrate for reaction and provide compensating charges, thereby stabilizing the transition state, lowering the barrier from substrate to product and speeding up the reaction.

Transition state theory connects the rate constant for the reaction, k, to the Gibbs energy of the transition state, $\Delta G^{\ddagger}$, eq. 7.1. Transition state theory assumes a quasi-equilibrium* between the substrate and transition state. The constants in eq. 7.1 account

*It is not a true equilibrium because some of the molecules that reach transition state continue along

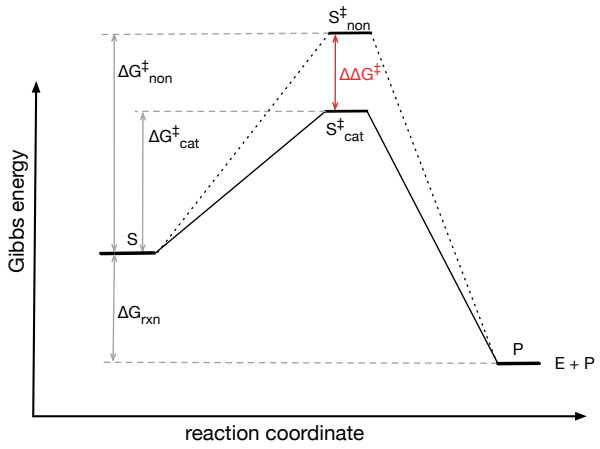

Figure 7.1. Gibbs energy changes associated with a chemical reaction. The Gibbs energy change for the transformation of substrate S to product P, ΔG_{rxn}, is negative indicating that the reaction is favorable. The barrier between the starting material and product limits the rate of the reaction. The top of this barrier is the highest energy structure along the bond-making and bond-breaking path of the reaction and is called the transition state. $\Delta G^{\ddagger}_{non}$ and $\Delta G^{\ddagger}_{cat}$ are both positive because the transition states, $S^{\ddagger}_{non}$ and $S^{\ddagger}_{cat}$, are higher in energy than S. The catalyst stabilizes the $S^{\ddagger}_{non}$ transition state by an amount $\Delta\Delta G^{\ddagger}$, which lowers the barrier to product and speeds up the reaction.

for the fraction of molecules that react once they reach the top of the barrier. The activation Gibbs energy is a positive value because the value of $\frac{k}{\text{constants}}$ is <1, which makes $\ln\left(\frac{k}{\text{constants}}\right)$ a negative value.

$$\Delta G^{\ddagger} = -RT \ln\left(\frac{k}{\text{constants}}\right) \tag{7.1}$$

A lower activation energy corresponds to a lower barrier and a faster reaction (larger k). A faster rate, or larger k, decreases the absolute value of $\ln \frac{k}{\text{constants}}$, which leads to a smaller activation energy. For example, the $\ln(0.001)$ is –6.9, while the $\ln(0.1)$ is –2.3. The decrease in activation energy for a catalyzed reaction as compared to a non-catalyzed reaction is proportional to the natural logarithm of the ratio of the rate constants for the catalyzed and non catalyzed reactions, eq. 7.2.

$$\begin{aligned} \Delta G^{\ddagger}_{cat-non} &= -RT \ln\left(\frac{k_{cat}}{\text{constants}}\right) - \left(-RT \ln\left(\frac{k_{non}}{\text{constants}}\right)\right) \\ &= -RT \ln\left(\frac{k_{cat}}{k_{non}}\right) \end{aligned} \tag{7.2}$$

Transition state stabilization accounts for most of the catalytic effect of enzymes.[1] Experimental evidence that enzymes stabilize transition states includes the finding that transition state analogs are potent enzyme inhibitors and that antibodies to transition state analogs can catalyze reactions, albeit slowly.

7.2 Measuring reaction rates

Consider the rate of a unimolecular chemical reaction, eq. 7.3.

$$[S] \rightarrow [P] \tag{7.3}$$

The rate of reaction, V, is the disappearance of substrate over time, $-d[S]/dt$, and has units of concentration per unit time. Researchers measured the initial rate of reaction using different concentrations of substrate, [S], and found that the rate increases linearly as [S] increases, Fig. 7.2.

The proportionality constant, k, for this rate increase is the rate constant and has units of time^{-1}, eq. 7.4.

$$V = -\frac{d[S]}{dt} = k \cdot [S] \tag{7.4}$$

the reaction coordinate to form product. In a true equilibrium, they would always return to the substrate species.

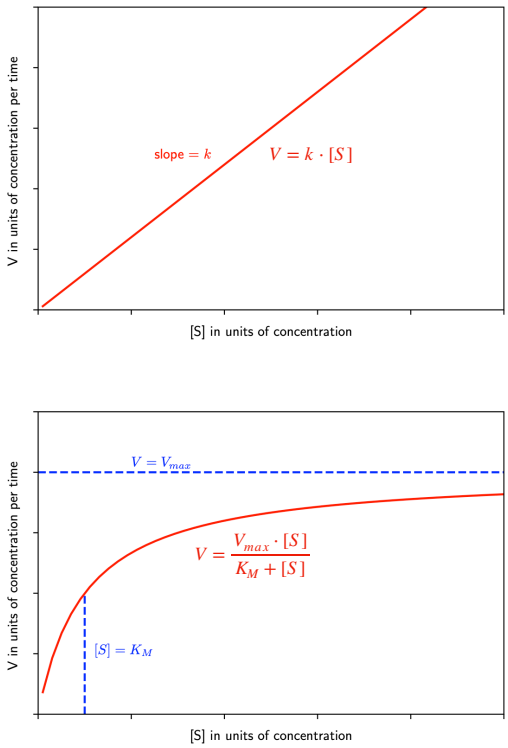

Figure 7.2. The rate of reaction for a unimolecular chemical reaction (top) and a single substrate enzyme-catalyzed reaction (bottom). In both cases the rate of reaction, V, increases with increasing substrate concentration, $[S]$, but this increase flattens for the enzyme-catalyzed reaction. The slope of the line, k, describes the speed of a unimolecular chemical reaction, while two constants are needed to describe the speed of the enzyme-catalyzed reaction. The constant K_M in units of concentration is the concentration at which the reaction rate reaches half of the maximum rate. K_M also indicates the substrate concentration at which the curve starts to flatten. The constant V_{max} in units of concentration/time indicates the maximum rate.

This rate constant reveals the inherent propensity of the substrate to react. If another reaction has a larger rate constant, then it is faster. It will convert more molecules into product over the same time at the same concentration of substrate.

The rate of an enzyme-catalyzed reaction, eq. 7.5, also varies with the concentration of substrate, but the relationship between V and [S] is not linear, but curved, see Fig. 7.2 above.

$$[S] \xrightarrow{\text{enzyme}} [P] \tag{7.5}$$

The equation the describes this curve, eq. 7.6, contains two constants: V_{max} and K_M. These constants reveal the inherent speed of an enzyme-catalyzed reaction.

$$V = \frac{V_{max} \cdot [S]}{K_M + [S]} \tag{7.6}$$

V_{max} is the reaction rate at high substrate concentration, while K_M is the substrate concentration at which the reaction rate is half of V_{max}. For chemical process, the physical meaning of the constant k is the inherent rate of spontaneous reaction of the substrate. For the enzymatic process, the physical meanings of the constants V_{max} and K_M depend on the the mechanism of the enzyme-catalyzed reaction and will be discussed below. At this time, they are simply kinetic constants.

Consider the two limiting cases of low substrate concentration and high substrate concentration. Here low and high are in comparison to K_M for the reaction. At low substrate concentration, [S] is much smaller than K_M, so the denominator of eq. 7.6 simplifies to K_M. Equation 7.6 simplifies to eq. 7.7.

$$V_{low[S]} = \frac{V_{max}}{K_M} \cdot [S] \tag{7.7}$$

This equation resembles the equation for a chemical reaction, eq. 7.4 above, where the ratio V_{max}/K_M corresponds to the rate constant for the reaction. Thus, at low substrate concentration, one compares the rates of enzyme-catalyzed reactions by comparing V_{max}/K_M. Higher values correspond to faster reactions.

At high substrate concentration, K_M is much smaller than [S], so the denominator of eq. 7.6 simplifies to [S]. Next, [S] cancels out since it appears as a factor in both the numerator and denominator. Equation 7.6 simplifies to eq. 7.8.

$$V_{high[S]} = V_{max} \tag{7.8}$$

At high substrate concentration, the reaction rate is independent of substrate concentration. To compare enzymes at high substrate concentrations, one compares V_{max} values. Higher values correspond to faster reactions.

The comparison of reaction rates described in this section is steady-state kinetics. The steady-state refers to constant conditions over the time of the experiment. The temperature, pH and substrate concentration are all constant over the course of the experiment. To keep the substrate concentration constant, researchers measure the initial rates of reaction only while the first few percent of the substrate react, that is, the initial rate of reaction. There is a large excess of substrate as compared to enzyme. Each enzyme molecule completes many catalytic cycles over the course of the measurement. The large excess of substrate also ensures that formation of an enzyme-substrate complex does not significantly alter the concentration of free substrate.

To measure the steady-state kinetic parameters for a reaction, K_M and V_{max}, one measures the initial rate of reaction at varying substrate concentrations and find the values of K_M and V_{max} for equation 7.6 that best fit the data. Huitema & Horsman (2018) provide scripts for this fitting using the statistical program R. The the supporting information of this chapter provide a Python script to find the best values of V_{max} and K_M.[†] For a good fit one needs data points that define both the low [S] and high [S] regions of the curve; that is, measurements at substrate concentrations both above and below K_S.

7.2.1 The Michaelis-Menten model

The simplest model of the physical steps within an enzyme-catalyzed reaction is the Michaelis-Menten model, Fig. 7.3. The substrate, S, binds reversibly to the enzyme, E, to form an enzyme-substrate complex, E·S. The two physical step involved in are association of enzyme and substrate with rate k_1 and dissociation of the enzyme-substrate complex with rate k_{-1}. Next, the enzyme substrate complex reacts irreversibly to form and release the product with rate k_2.

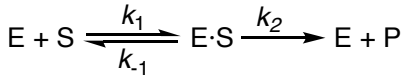

Figure 7.3. The Michaelis-Menten model of an enzyme-catalyzed reaction involves two physical steps: reversible equilibration of free substrate and enzyme with the enzyme-substrate complex, E·S, followed by reaction of $E \cdot S$ to form and release the product, P.

Starting from this simple model, one can derive an equation with the same form as eq. 7.6. This derivation will assign physical meanings to the constants V_{max} and K_M.

In the Michaelis-Menten model, the rate of product formation, V, depends on the amount of enzyme-substrate complex, $[E \cdot S]$, and the rate at which it is converted to product, k_2, eq. 7.9.

[†]An older approach used a rearranged form of equation 7.6 to the fit the experimental data. This rearrangement allowed fitting a line to the data instead of a curve. This Lineweaver-Burke plot was more convenient when computers to fit data to a curve were not available, but has the disadvantage that it overweights data at low substrate concentration, which can lead to small errors. See supporting information for details.

$$V = k_2 \cdot [E \cdot S] \tag{7.9}$$

The amount of $[E \cdot S]$ is unknown, so we express $[E \cdot S]$ in terms of known quantities. One known quantity is the amount of added enzyme, $[E]_0$. The total of $[E \cdot S]$ and $[E]$ must be equal to the amount of enzyme added, $[E]_0$.

$$[E]_0 = [E] + [E \cdot S] \tag{7.10}$$

Second, we assumed that $[E]$ and $[S]$ are in equilibrium with $[E \cdot S]$, so we may write the following equation.

$$\frac{k_1}{k_{-1}} = \frac{[E \cdot S]}{[E] \cdot [S]} \tag{7.11}$$

Combining these two equations and solving for $[E \cdot S]$ yields eq. 7.12,

$$[E \cdot S] = \frac{[E]_0 \cdot [S]}{\frac{k_{-1}}{k_1} + [S]} \tag{7.12}$$

Combining eq. 7.9 and eq. 7.12 yields the Michaelis-Menten equation, eq. 7.13 where $K_D = \frac{k_{-1}}{k_1}$

$$V = \frac{k_2 \cdot [E]_0 \cdot [S]}{K_D + [S]} \tag{7.13}$$

This equation assigns the following physical meaning to V_{max} and K_M.

$$V_{max} = k_2 \cdot [E]_0 \tag{7.14}$$

$$K_M = K_D = \frac{k_{-1}}{k_1} \tag{7.15}$$

V_{max} corresponds to the product of k_2, the inherent reactivity of the $[E \cdot S]$ complex, and the amount of of added enzyme. K_M corresponds to the equilibrium dissociation constant of the $[E \cdot S]$ complex, K_D. We can now draw a physical picture of what occurs at different substrate concentrations in an enzyme-catalyzed reaction, Fig. 7.4.

The Michaelis-Menten model simplifies enzyme-catalyzed reactions to a binding step described by K_M or K_D that forms the enzyme-substrate complex and a reaction step described by $k_2 \cdot [E]_0$ that converts the substrate to product. Improving enzymes catalysis may require improvements to either to both steps. When the substrate does not bind to the enzyme or binds poorly, then improved binding can increase the enzyme-catalyzed reaction rate. When the substrate binds, but reacts slowly, then the chemical reaction step must be improved by stabilizing the transition state.

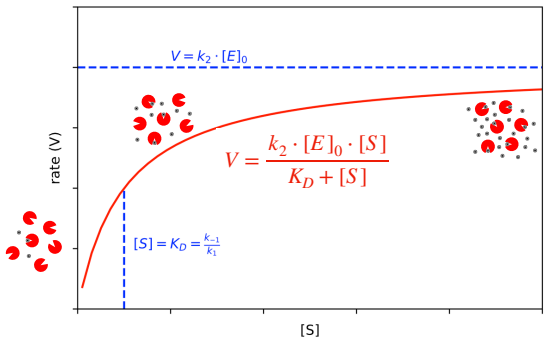

Figure 7.4. The Michaelis-Menten model adds a physical interpretation to the observed variation in rate of an enzyme-catalyzed reaction with substrate concentration. At low substrate concentration, the rate is slow because only a few enzyme molecule contain bound substrate (black dots). The rate increases as the substrate concentration increases because an increasing fraction of enzyme molecules contain substrate. At high substrate concentration, all the enzyme molecules contain a bound substrate and the enzyme-catalyzed reaction rate reaches a maximum and no longer varies with substrate concentration.

7.2.2 Gibbs energy diagrams for enzyme-catalyzed reactions

Michaelis-Menten model. Enzymes catalyze reactions by first binding the substrate and then transforming the substrate to product, so a Gibbs energy diagram includes these steps to show how K_M and k_{cat} affect an enzyme-catalyzed reaction,[2] Figure 7.5. First, the Gibbs energy of the transition state for the uncatalyzed reaction, $S^{\ddagger}$, is higher than that for the enzyme-catalyzed reaction, $E \cdot S^{\ddagger}$ indicating that the enzyme-catalyzed reaction is faster. The formation of the $E \cdot S$ complex is favorable by an amount ΔG_{K_M}; then transformation of $E \cdot S$ to the transition state $E \cdot S^{\ddagger}$, is uphill by $\Delta G_{k_{cat}}$.

More complex reaction mechanisms. The Gibbs energy diagrams include additional states for more complex reaction mechanisms. For example, a reaction that includes an enzyme intermediate includes a valley that represents that intermediate, Figure 7.6.

Even if the Gibbs energy profile is complex, the largest barrier (from valley to peak) is the step that determines k_{cat} and the highest point measured from the starting $E + S$ determines k_{cat}/K_M.

Simplification for limiting cases. Consider two limiting cases of enzyme-catalyzed reactions: high and low substrate concentration, Figure 7.7. At high substrate concentration most of the enzyme exists as the E· S complex, while at low substrate concentration most of the enzyme exists as free enzyme.

At high substrate concentration, most of the enzyme exists as the $E \cdot S$ complex, so the starting state is this complex. The barrier from the starting state to product is determined

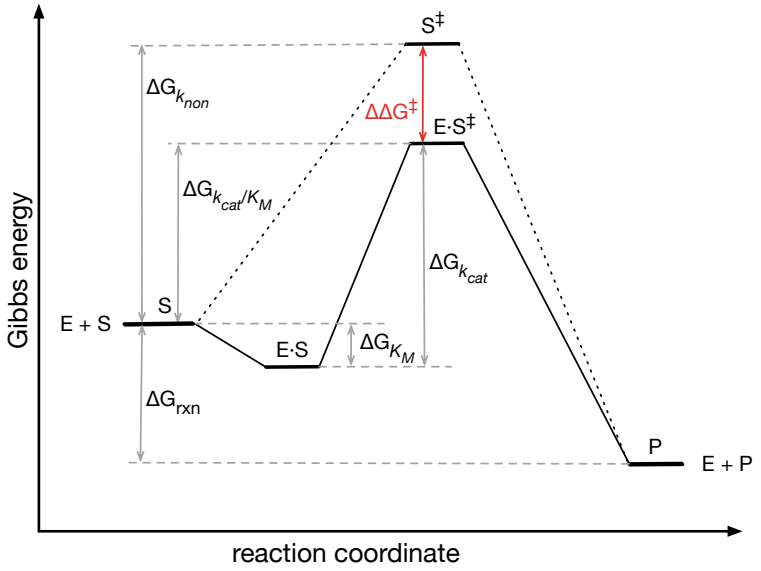

Figure 7.5. Gibbs energy diagram for an uncatalyzed and an enzyme-catalyzed reaction at standard state. The uncatalyzed reaction proceeds via transition state $S^{\ddagger}$, which lies $\Delta G_{k_{non}}$ above the energy of S. The enzyme-catalyzed reaction is faster because the energy distance from $E + S$ to the transition state, $E \cdot S^{\ddagger}$, $\Delta G_{k_{cat}/K_M}$, is lower by an amount $\Delta\Delta G^{\ddagger}$. The path from from $E + S$ to $E \cdot S^{\ddagger}$ involves binding to form $E \cdot S$ (decrease Gibbs energy by ΔG_{K_M}), then an increase to $E \cdot S^{\ddagger}$ by an amount $\Delta G_{k_{cat}}$.

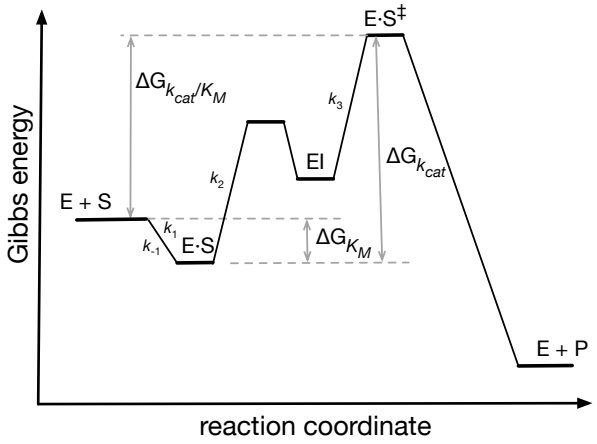

Figure 7.6. Gibbs energy diagram for an enzyme-catalyzed reaction that includes an enzyme intermediate as shown in Figure 7.8 above. k_{cat} depends on both k_2 and k_3.

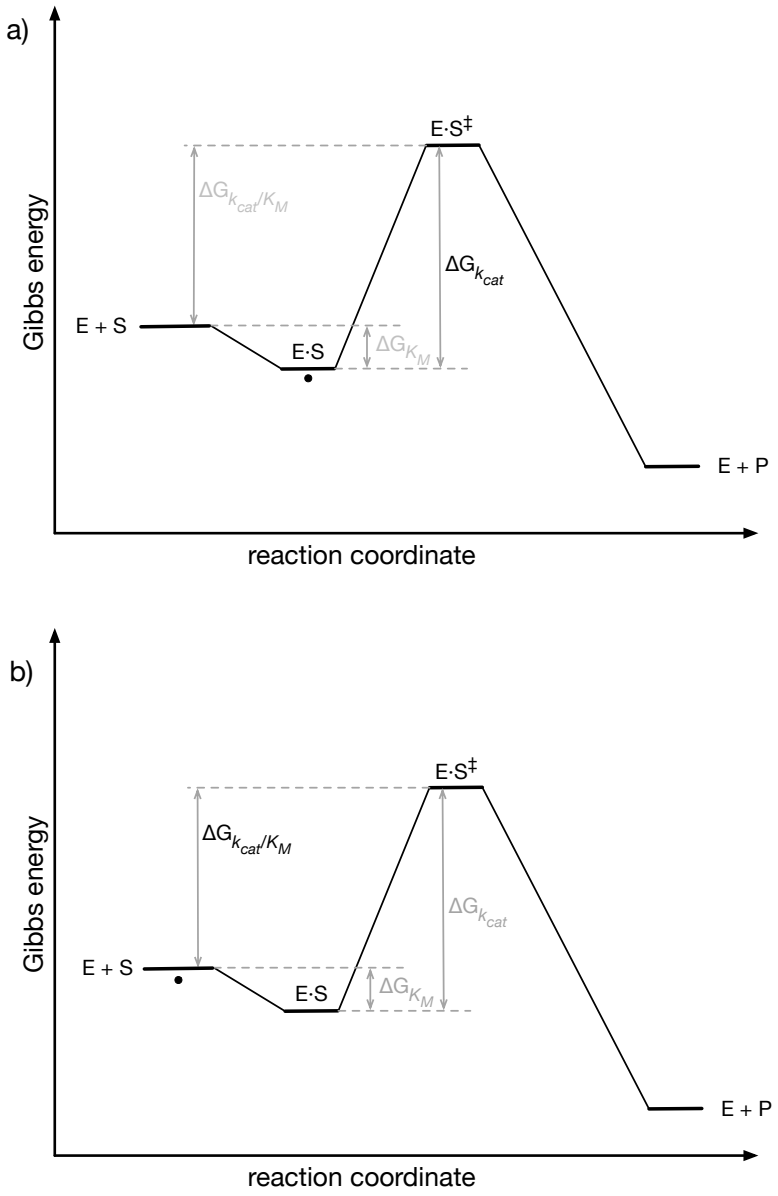

Figure 7.7. Limiting cases for enzyme-catalyzed reactions. a) At high substrate concentration, most of the enzyme contains bound substrate as indicated by the black dot under that state. The reaction rate under these conditions depends only on k_{cat}. b) At low substrate concentration, most of the substrate and enzyme are in the free state as indicated by the black dot under that state. The reaction rate depends on substrate concentration and k_{cat}/K_M at low substrate concentration. The substrate concentration and K_M determine the small fraction of enzyme that forms $E \cdot S$ while k_{cat} determine how fast it is converted to product.

Table 7.1. Assumptions within the Michaelis-Menten model of enzyme catalysis and how they can be removed.

Assumption	Removing assumption
single substrate	additional experiments
equilibrium between E, S and $E \cdot S$	redefine K_M
single step reaction	redefine k_{cat} and K_M
product release is fast	additional experiments

by k_{cat} and the only way to increase the rate of reaction is to increase k_{cat}. Raising the energy of the $E \cdot S$ complex (weakening binding) will increase k_{cat} and speed up the reaction up to a point. Once the binding becomes so weak that the reaction is no longer under 'high substrate concentration' conditions and most of the enzyme no longer exists as the $E \cdot S$ complex, it will slow down and the limiting case of high substrate concentration no longer applies. Increasing binding (lowering K_M) slows down the reaction when the substrate concentration is high because it increases the distance between the $E \cdot S$ complex and the transition state. Most biocatalysis applications use high substrate concentrations so the rates are limited by k_{cat}.

At low substrate concentration, most of the enzyme exists as $E + S$. The distance to the barrier to product is determined by k_{cat}/K_M. Selectively stabilizing $E \cdot S^{\ddagger}$ will lower the energy barrier and speed up the reaction as it did in the high substrate concentration case. However, stabilizing $E \cdot S$ (decreasing K_M) in the low substrate concentration case also increases the reaction rate by increasing the proportion of enzyme molecules bound to substrate. This approach is limited since once $E \cdot S$ is so stable that all the enzyme is in the $E \cdot S$ state, then the high substrate concentration case applies and only stabilizing $E \cdot S^{\ddagger}$ can increase the reaction rate. For example, the anti-cancer activity of asparaginase[3] (mentioned in Chapter 1) depends on its ability to reduce the concentration asparagine in the blood so that it is not available for cancer cells. The blood contains 2-25 μM asparagine, so lowering the K_M of an asparaginase to this level or slightly below would make the reaction faster, but further decreases in K_M would not.

Explain: The reaction rate of an enzyme-catalyzed reaction was faster with variant A than with variant B under one substrate concentration, but was slower under a different substrate concentration. How this is possible?

7.3 More complex models of enzyme-catalyzed reactions

The simple Michael-Menten model makes several simplifying assumptions, Table 7.1, but these can be removed either by additional experiments or redefining the meanings of k_{cat} and K_M.

The Michaelis-Menten model includes only one substrate, but it can be be extended to two or more substrates by repeating the measurements under different limiting conditions. To extend the model to two-substrate enzyme-catalyzed reactions one first adds a

large excess of one substrate. The other substrate now limits the reaction rate according eq. 7.6, so varying its concentration reveals its steady-state kinetic parameters. Next, a second set of experiments with saturating amounts of the other substrate allows measurement of the kinetic parameters of the first substrate.

The Michael-Menten model assumed that the free enzyme and substrate are in equilibrium with the enzyme substrate complex in eq. 7.11 above. This assumption was needed in order to define the concentration of $E \cdot S$. The Brigg-Haldane approach replaces this assumption with a more relaxed assumption, that the concentration of $E \cdot S$ is constant. The assumption, known as the steady-state approximation, is less restrictive and leads to a different physical meaning for K_M.

The steady state assumption means that the rate of formation of $E \cdot S$ (k_1 term in eq. 7.16) must be equal to the rate of destruction of $E \cdot S$ ($k_{-1} + k_2$ term in eq. 7.16).

$$k_1 \cdot [E] \cdot [S] = (k_{-1} + k_2)[E \cdot S] \tag{7.16}$$

This equation is used in place of eq. 7.10 above to express $[E \cdot S]$ in terms of known quantities. Combining eq. 7.16 with eq. 7.10 and solving for $[E \cdot S]$ yields the equation below.

$$[E \cdot S] = \frac{[E]_0 \cdot [S]}{\frac{k_{-1}+k_2}{k_1} + [S]} \tag{7.17}$$

Combining this equation with eq. 7.9 above yields equation 7.18 below.

$$\begin{aligned} V &= \frac{k_2 \cdot [E]_0 \cdot [S]}{K_M + [S]} \\ \text{where } K_M &= \frac{k_{-1} + k_2}{k_1} \end{aligned} \tag{7.18}$$

The form of the equation and the definition of V_{max} is the same as for the Michaelis-Menten model, but the definition of the constant K_M includes an additional physical step. It is no longer purely a dissociation constant since it contains the rate constant for the chemical step, k_2, in the numerator. Nevertheless, it still represents an amount of $[E \cdot S]$ since the rate constants in the numerator are for processes that destroy $[E \cdot S]$, while the rate constant in the denominator is for formation of $[E \cdot S]$. If k_2 is slow, the K_M is a true dissociation constant.

Many enzyme-catalyzed reactions involve multiple steps, not just the single step in the Michaelis-Menten model. For example, many enzyme-catalyzed reaction form an intermediate, EI, Fig. 7.8. Subtilsin-catalyzed hydrolysis of esters involves an acyl-enzyme intermediate. Later in this chapter, Fig. 7.25 shows a detailed mechanism.

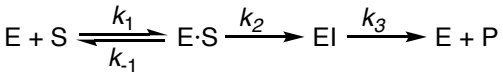

Figure 7.8. Many enzyme-catalyzed reactions involve an enzyme-bound intermediate. These reactions have two steps that involve bond-making and bond-breaking, k_2 and k_3, instead of just one in the Michaelis-Menten model. These reactions still follow the curve in eq. 7.6, but multiple physical steps contribute to V_{max} and K_M.

The steady-state enzyme kinetics for subtilisin still follow the Michaelis-Menten equation, but both K_M and k_{cat} are combinations of multiple physical steps:

$$K_M = \frac{k_3}{k_2 + k_3} \cdot \frac{k_{-1} + k_2}{k_1}$$

$$k_{cat} = \frac{k_2 \cdot k_3}{k_2 + k_3}$$

Measuring steady state kinetics still yields values for K_M and k_{cat}, but their meaning cannot be assigned to a simple step in the catalytic mechanism. In the special case where the deacylation step (k_3) is fast as compared to the acylation step (k_2), then the equations simplify to those for a single step reaction. In this case, k_{cat} represents the slow step of catalysis, k_2, the formation of the acyl-enzyme intermediate.

$$K_M = \frac{k_2 + k_{-1}}{k_1}$$

$$k_{cat} = k_2$$

Finally, the Michaelis-Menten model assumes that product release is fast, but in practice product inhibition may slow the rate of enzyme-catalyzed reactions. Addition experiments in the presence of varying amounts of product can identify product inhibition.

In this text, we will use the Michaelis-Menten model to discuss most enzyme-catalyzed reactions. The constant K_M reflects the amount of $[E \cdot S]$ that forms and k_{cat} represents the conversion of $[E \cdot S]$ into product. When the mechanisms are more complex, additional physical steps contribute these processes and the Gibbs energy diagrams are more complex. Despite this increased complexity, the physical steps that contribute to K_M are those that define the amount of enzyme-substrate complex that is present and the physical steps that contribute to V_{max} are those that define how fast the chemical transformation of the enzyme-substrate complex to product occurs.

7.3.1 Which enzyme is faster?

The metric used to identify the faster enzyme depends on the application. At low substrate concentrations, the ratio k_{cat}/K_M, often called the catalytic efficiency, is a good metric. It measures the ability of the enzyme to both bind the substrate and to catalyze the chemical step. In vivo applications such as therapeutic enzymes typically act at low substrate concentrations, so catalytic efficiency is a good measure. Typical values of k_{cat}/K_M for enzyme-catalyzed reactions are $\sim 10^5 M^{-1}s^{-1}$ and the limit is $\sim 10^9 M^{-1}s^{-1}$. This limit comes from the requirement for two species (E and S) to combine to make the $E \cdot S$ complex. This rate cannot be faster than diffusion.

In many biocatalysis applications, the users control the substrate concentrations and usually choose high concentrations so the reaction rate is high and the reaction vessel is small. At high substrate concentrations, k_{cat} determines the reaction rate, so the enzyme with the higher k_{cat} value is the faster one. For example, the isomerization of glucose to fructose is a continuous process in glucose syrup as solvent. The glucose concentration (~5 M) is much higher than the K_M of glucose for the isomerase (0.09-0.9 M).[4] The value of k_{cat} determines the reaction rate.

In chemical manufacturing using batch processes, the substrate concentration decreases while product concentration increases as the reaction proceeds leading to changes in reaction rate. The value of k_{cat} may control the reaction rate initially, but later k_{cat}/K_M controls the reaction. Substrate and product inhibition, not discussed above, may also influence the reaction rate. Fox and Clay[5] proposed an average velocity metric that compares how long it takes for a reaction to reach completion. An example where this approach may be useful is the manufacture of a pharmaceutical or pharmaceutical intermediate.

The following sections examine molecular strategies to alter k_{cat} and K_M speed up enzyme-catalyzed reactions. Identifying whether k_{cat} or K_M or both need improvement is a good first step, but it is not enough. Both binding and the reaction steps may include multiple physical steps for different catalytic mechanisms. Engineering requires identifying which physical steps limit catalysis and then speed up these steps. For example, a substrate may bind poorly to an enzyme, which could indicate a mismatch in the shapes of the substrate and active site. When a substrate reacts slowly, there are many possible causes. The substrate may orient in a non-reactive orientation, the required acids or bases may be in the incorrect protonation state, the charges generated during bond-breaking and bond making may not be stabilized, a required conformational change of a protein loop may be slow as well as other causes.

7.4 Increasing reaction rates by engineering tighter binding

The optimum binding of the substrate and enzyme will be just strong enough (K_M value just low enough) so that most of the enzyme contains bound substrate under the reaction conditions. Filling most enzyme active sites with substrate creates the opportunity for each enzyme molecule to contribute to catalysis. Once most of the enzyme molecules contain bound substrate, further stabilization has the undesired effect of slowing down the reaction. Further lowering of the energy of the $E \cdot S$ state in Figure 7.5 above in-

creases the energy difference between the $E \cdot S$ state and transition state ($E \cdot S^{\ddagger}$), which slows down the reaction (lowers k_{cat}).

In many cases, the target substrate is not the natural substrate of the enzyme and the reaction is slow because little $E \cdot S$ forms. Adjusting the complementarity of binding site to match the target substrate improves binding and increases the reaction rate. The engineering approaches are the same as those described in the previous chapter on binding: preorganization, matching the size and shape and the interactions between the substrate and the binding site.

Matching substrate size. Substrates that are larger than the natural substrate bind poorly when they cannot fit in the active site. For example, alcoholic fermentation by *Lactococcus lactis*, involves the reduction of acetaldehyde to ethanol by NADH catalyzed by alcohol dehydrogenase, Fig 7.9 , which shows good catalytic efficiency: K_M = 0.4 mM, k_{cat} = 35 s^{-1}, k_{cat}/K_M ~ 90,000 M^{-1} s^{-1}.

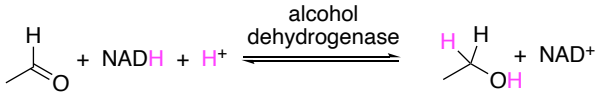

Figure 7.9. Wild-type alcohol dehydrogenase efficiently catalyzes reduction of the natural substrate, acetaldehyde.

As part of a project on biofuels, researchers sought to use this enzyme to reduce isobutyraldehyde to isobutanol, Fig 7.10, where the larger isobutyraldehyde replaces acetaldehyde.

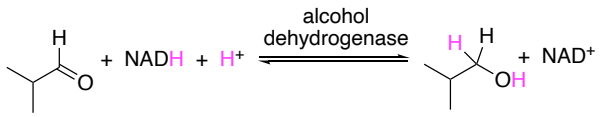

Figure 7.10. Wild-type alcohol dehydrogenase also catalyzes reduction of the substrate analog, isobutyraldehyde, but ~30-fold less efficiently.

The catalytic efficiency of this *Lactococcus* alcohol dehydrogenase is ~30-fold lower toward isobutyraldehyde, mainly due to looser binding of the substrate: K_M = 12 mM, k_{cat} = 30 s^{-1}, k_{cat}/K_M ~ 3,000 M^{-1} s^{-1}. The researchers identified several improved variants by random mutagenesis and screening.[6] The structure of one variant showed an expanded binding site caused by substituting Leu264 with valine to remove a methylene group and substituting Tyr50 with phenylalanine to remove an oxygen atom.[7] The modified enzyme bound the larger substrate more tightly (K_M = 1.6 mM) and showed a ten-fold higher catalytic efficiency (k_{cat}/K_M ~ 32,000 M^{-1} s^{-1}). These substitutions made the substrate-binding site larger by removing two non-hydrogen atoms (a carbon and an oxygen) to match the new substrate, which is larger by two non-hydrogen atoms (two carbons).

In an industrial example, Savile and coworkers engineered a transaminase to accept a

precursor for a diabetes drug by expanding the substrate-binding site with multiple substitutions to fit the larger substrate.[8]

Although smaller substrates fit in the active site, they may bind poorly because only part of the smaller substrate contacts the surface of the binding site. Decreasing the size of the binding site will increase the contact surface and strengthen the binding interaction.

Matching interaction complementarity. Engineering a suitable binding site often requires adjusting the non-covalent interactions as well as the size. For example, accommodating the new substrate, a glutamine residue, in a site that normally accepts lysine or arginine required adjusting the size, the charge and the hydrogen bonding partners of the binding site. Gluten, a protein found in wheat, rye, and barley, contains a high proportion of proline and glutamine. Digestive proteases do not cleave peptides after a Pro-Gln sequence so oligopeptides enriched in the dipeptide sequence proline-glutamine accumulate upon digestion of gluten. In celiac disease patients, these oligopeptides are believed to trigger an autoimmune response, which causes diarrhea and other gastrointestinal problems. One potential solution is adding a protease that can degrade Pro-Gln containing peptides to the meals of celiac patients. Engineering such a protease started with kumamolisin, which catalyzes hydrolysis after a Pro-Arg/Lys sequence in a peptide.[9] The goal was to increase the reaction rate for hydrolysis after a Pro-Gln sequence, Fig 7.11.

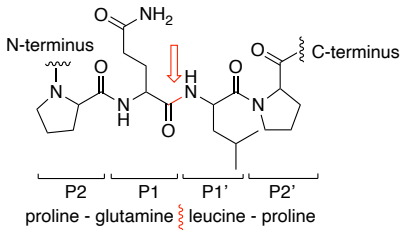

Figure 7.11. Target tetrapeptide used for computational design of a protease that cleaves immunogenic gluten peptides. Cleavage occurs between the P1 (Gln) and P1' (Leu) amino acid residues as marked by the red arrow and wavy line. The starting protease, kumamolisin, favors hydrolysis after a Pro-Arg/Lys sequence. The engineering goal was to increase catalysis of hydrolysis after a Pro-Gln sequence such as the one shown.

The changes involved replacing two negatively-charged Glu residues in the binding site with neutral residues to favor binding of the neutral Gln instead of the positively charged Arg or Lys. In addition, a glycine in binding site was replaced by serine to decrease the size of the binding site because the side chain of Gln is smaller than the side chain of Arg or Lys. The side chain of the newly-added serine can also make a hydrogen bond with the side chain of Gln. Several other substitutions were also introduced outside the binding site. The rate of hydrolysis (k_{cat}/K_M) of a test peptide (R-ProGlnProGln~LeuPro-R') where the '~' marks the cleavage site, increased 120-fold from 4.9 to 570 M^{-1} s^{-1}. In this example, creating a complementary binding site required changing the size of the binding site, the electrostatic character of the binding site, and the hydrogen bonding

partners in the binding site. A further-engineered variant of this protease is currently undergoing clinical trials.[10]

Slow binding due to tunnels or flexible loops. In some cases, the rate of association of substrate and enzyme is slow; this slow rate can limit the rate of reaction. In this case improving binding will speed up the reaction.

The most likely cause of slow binding is a deeply-buried active site. Reaching the active site may require the substrate to diffuse through a tunnel or require a flexible loop to move out of the way to allow the substrate into the active site. This slower binding hinders formation of the the $E \cdot S$ complex and thus raises its energy. Increasing the rate of formation of the $E \cdot S$ complex can increase the amount of complex that forms. For example, to increase the ability of a transaminase to accept larger substrates up to 120-fold, Pavlidis and coworkers[11] both expanded the active site with two substitutions of Tyr with Phe and increased the flexibility at the opening of the binding site with a Pro to His substitution to allow larger substrates to enter.

Tunnels and flexible loops can also contribute to the selectivity of enzymes.[12] Tunnels can prevent access to large substrates or to substrates that interact unfavorably with the tunnel (hydrophilic substrate may not pass through a hydrophobic tunnel) or substrates that interact too favorably (hydrophobic substrate may bind strongly to a hydrophobic tunnel leading to non-productive binding). Flexible loops can also exclude solvent from the active site to prevent undesirable side reactions.

7.5 Stabilizing the transition state to speed up reactions

After binding, the enzyme must stabilize the transition state for the reaction. This stabilization promotes the bond-making and bond-breaking steps. Enzymes use different strategies for this stabilization depending on what is required by the reaction, Table 7.2.

Table 7.2. Contributions to the stabilization of transition states by enzymes.

catalytic principle	rationale	implementation
preorganize $E \cdot S$ geometry for reaction	$E \cdot S$ complex adopts catalytically productive conformation	-model near attack complex -remove non-productive binding
optimize acids & bases	acids & bases usually needed for catalysis	-adjust pK_a of acids & bases
stabilize charge & shape of transition state	bond rearrangements alter charge & shape of substrate	-add complementary charges -adjust shape of active site
enable new mechanistic steps	enzymes can create new pathways for reaction	-chemical reasoning -introduce new functional groups

The first requirement is a reactive orientation of the substrate, which is called preorganization. This preorganization is a binding orientation where the substrate adopts a ready-to-react geometry close to the catalytic residues of the enzyme. In contrast, holding the substrate in a non-reactive orientation slows down the reaction. Preorganization may require the substrate to adopt a conformation different from its low energy conformation in solution.

Once the substrate is oriented for reaction, additional interactions can stabilize the bond-breaking and bond-making process. If the reaction involves proton transfers, the required acids must be protonated and the bases deprotonated for catalysis. The enzyme can further stabilize the new molecular shapes and charges as electrons reorganize in the transition state. In some cases, enzymes create new mechanisms as compared to the uncatalyzed reaction. These new mechanisms break up a high energy path into several low energy steps.

Individual enzymes vary in the contributions that they use for catalysis because reactions differ in the reasons that they are slow. For example, chorismate mutase stabilizes the transition state mainly by folding the substrate into a reactive conformation (preorganization), while a designed eliminase catalyzes the Kemp elimination reaction by providing a base and stabilizing the charges formed in the transition state. Enzymes must provide whichever catalytic contributions are needed to stabilize the transition state for each reaction. A cooking analogy is that adding salt sometimes improves the taste of a dish, but other times adding something sour or spicy is needed. Each dish differs on what improves it.

To qualitatively predict the transition state for an enzyme-catalyzed reaction, start with an arrow-pushing mechanism to identify the bond-making and bond-breaking steps. Identify the distances and angles that define a catalytically productive conformation. Draw a shape with partial bonds that lies in between the shapes of the substrate and product. Label any partial charges that form in the transition state. Identify any acids and bases required for the mechanism. Quantitative prediction of a transition state requires quantum-chemical modeling.

To identify how an enzyme accelerates the rate of reaction, compare the enzyme-catalyzed mechanism and transition state to a similar analysis of the non-catalyzed reaction. Does the enzyme preorganize the reacting atoms for the reaction? Does the enzyme supply acids and bases as needed for any proton transfer steps? Does the enzyme stabilize the unique shape of the transition state and any partial chargest that form in the transition state? If the enzyme-catalyzed reaction includes covalent intermediates, then the enzyme also changes the mechanism of the reaction as compared to the non-catalyzed reaction.

7.5.1 Preorganization of $E \cdot S$ for reaction (binding correctly)

Preorganization is the bringing of atoms into the correct orientation to react. Preorganization lowers the Gibbs energy of the transition state, $E \cdot S^{\ddagger}$. The preorganized state (sometimes called a near-attack complex) involves both the conformation of the substrate and the orientation of substrate relative to the enzyme. This preorganized state

is like a turnstile through which the substrate must pass to reach the transition state. Preorganization involves not just binding, but catalytically productive binding. In an efficient enzyme, formation of the $E \cdot S$ complex also organizes the complex for reaction such that binding and preorganization are the same processes. K_M can be improved by any type of binding, but k_{cat} can only be improved by catalytically productive binding. Preorganization may occur within the $E \cdot S$ state or along the path to the transition state, $E \cdot S^{\ddagger}$.

Holding reactive groups close to each other in the correct orientation for the reaction enormously speeds up reactions as compared to freely diffusing molecules containing the same groups. Even when the reacting groups are in the same molecule, reducing flexibility to favor a reactive orientation speeds up the reaction rate tremendously. For example, the intramolecular ring closure reaction in Fig 7.12 varies 50,000-fold by varying the number of rotatable bonds between the reacting groups.[13] The chemical steps are the same in all cases, but the faster ones orient the substrate better for reaction than the slower ones.

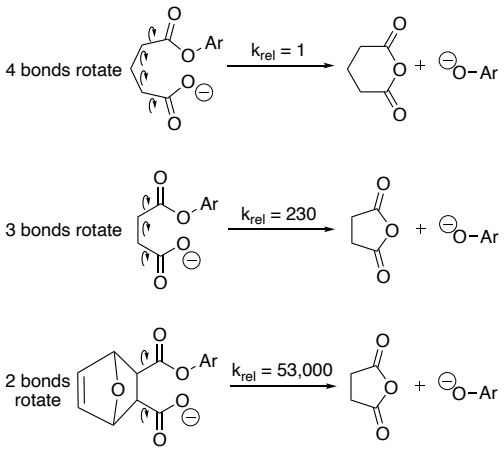

Figure 7.12. The rates of a ring closure reaction increase dramatically as the connecting fragments restrict possible rotations to favor the reactive conformation. The ring closure reaction requires attack of the carboxylate oxygen at the ester carbonyl carbon (Ar = 4-bromophenyl). Similarly, an active site that positions substrate and enzyme functional groups in a reactive orientation is a better catalyst than one that positions the groups in a non-reacting orientation. k_{rel} = relative rate constant.

In solution, substrates are usually flexible and adopt many different conformations by rotations along single bonds. Only some conformations orient the reactive groups such that reaction can occur. This reactive conformation might be rare in solution. The changes in structure force the carboxylate and aryl ester groups close to each other in Fig 7.12 above. Binding to the enzyme can similarly shift the balance toward a reactive conformation. The balance shifts because the enzyme restricts the available space for the substrate to move and interacts with it using hydrogen bonds, the hydrophobic effect,

and electrostatic interactions.

For example, the enzyme chorismate mutase catalyzes the rearrangement of chorismate by binding it in a conformation that rarely exists in solution. Chorismate mutase catalyzes a Claisen rearrangement, which rearranges bonds within a cyclic transition state, Fig 7.13. The transition state requires the two carbons that form a bond in the product to be near each other.

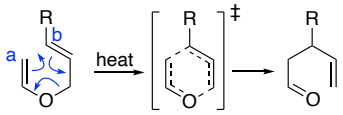

Figure 7.13. The Claisen rearrangement requires the carbon atoms marked a and b to be nearby each other in the transition state (marked ‡) so that a new bond can form between them.

The favored conformation of chorismate in solution places the two carboxylate groups far from each other, Fig 7.14. This separation of the carboxylates minimizes electrostatic repulsion and allows water to solvate each carboxylate. The Claisen rearrangement forms a bond between carbon 1 and 9 of chorismate, so these carbon atoms must be nearby in the transition state. Chorismate mutase holds chorismate in a folded, reactive conformation using two positively charged arginine side chains to position the carboxylates and two hydrophobic side chains to restrict their motion.[14] Molecular dynamics simulations indicate that chorismate adopts the correct preorganized conformation only 0.00007% of the time in solution, but 32% of the time while bound to the enzyme. This preorganization is the primary way by which chorismate mutase stabilizes the transition state and catalyzes this rearrangement.

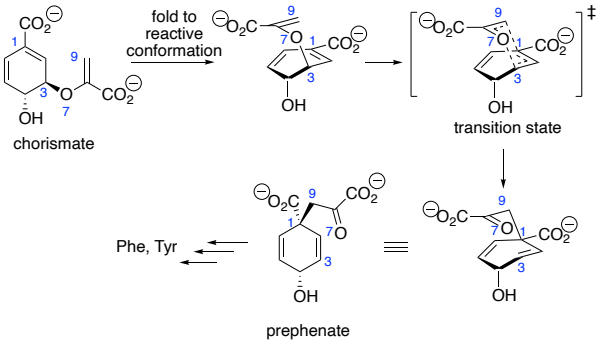

Figure 7.14. Chorismate mutase catalyzes a Claisen rearrangement that converts chorismate to prephenate in the pathway to make aromatic amino acids. In the transition state, the carbon atoms marked 1 and 9 must be nearby, but in solution, these carbon atoms lie far from each other to minimize electrostatic repulsion between the two carboxylate groups and to maximize their solvation. Chorismate mutase stabilizes the transition state for reaction by binding the substrate in a folded conformation that places carbon atoms 1 and 9 nearby for reaction.

Many enzyme-catalyzed reactions involve proton transfers. For these reactions, preorganization involves positioning the donor and acceptor atoms to make a hydrogen bond. Once the hydrogen bond is formed, the hydrogen is ready to transfer from the donor to acceptor atom.

An example of a reaction involving proton transfers in the Kemp elimination. (Figure 7.29 later in this chapter shows the Kemp elimination reaction.) Researchers designed many enzymes to catalyze this reaction, but testing revealed that most of them were inactive.[15] Additional molecular dynamics calculations revealed that most of the active enzymes were preorganized for reaction, while most of the inactive enzymes were not,[16] Figure 7.15. Preorganization was defined by the presence of a hydrogen bond between substrate and catalytic histidine. Researchers suggested that future enzyme designs could include molecular dynamics simulations to more accurately predict preorganization.

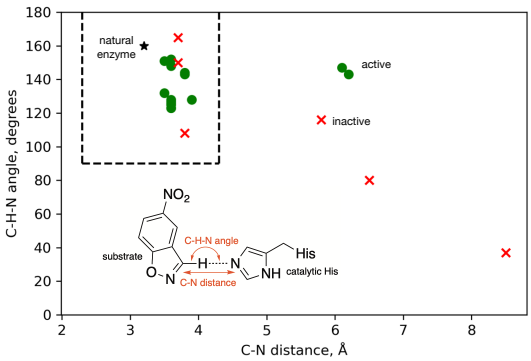

Figure 7.15. Preorganization of the substrate and catalytic histidine for proton transfer correlates with catalytic activity for the the Kemp elimination. The reaction involves the catalytic histidine deprotonating the carbon atom indicated. Preorganization involves positioning the donor C–H of the substrate and acceptor N of the catalytic histidine at a suitable distance (x-axis) and angle (y-axis) for a hydrogen bond (boxed region). The angles and distances come from molecular dynamics simulations of the enzyme-substrate complex. The star marks the orientation of donor and acceptor atoms between catalytic groups in the natural enzyme cathepsin.

The opposite of preorganization is non-productive binding where the enzyme binds the substrate in an unreactive mode at the active site, Fig 7.16. Non-productive binding increases the distance from the $E \cdot S$ complex to the transition state and therefore slows down the reaction. The $E \cdot S$ complex must first reorient into reactive orientation before it can react.

Non-productive binding slows the thermolysin-catalyzed coupling of *N*-benzyloxycarbonyl-L-aspartate (Cbz-L-Asp) with L-phenylalanine methyl ester (L-Phe-OMe), Fig 7.17. This reaction is a condensation to form an amide link, not the usual hydrolysis of

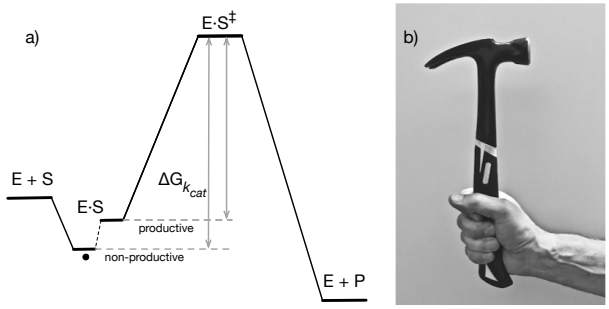

Figure 7.16. Non-productive binding slows down catalysis. This hypothetical Gibbs energy diagram for an enzyme-catalyzed reaction shows the $E \cdot S$ complex as an equilibrating mixture of two $E \cdot S$ species. The more stable one is non-productive because it lies earlier along the reaction coordinate. If most of the $E \cdot S$ complex exists in the non-productive form (indicated by the black dot), then the reaction rate is slower ($\Delta G_{k_{cat}}$ increases) than if it existed in the productive form. This case is similar to over stabilization of the $E \cdot S$ complex mentioned previously. b) A hammer is ineffective in the wrong orientation and will not drive nails. An active site that positions the substrate in a nonreactive orientation is also ineffective.

a peptide. Precipitation of the product dipeptide drives this reaction in the otherwise unfavorable direction.[17] This condensation is a critical step in the manufacture of aspartame (L-aspartyl-L-phenylalanine methyl ester), a low-calorie sweetener.

X-ray structures of substrates bound to thermolysin revealed the reason for this slow reaction, Fig 7.17.[18] The active site of thermolysin contains a catalytic zinc ion as well as separate regions to bind the carboxyl and amine donors. The carboxyl donor (Cbz-L-Asp) binds to the wrong site, the amine donor site, instead of the carboxyl donor site. The rotation of the Cbz-L-Asp substrate into the correct site is slow, but required before the amine donor substrate (L-Phe-OMe) can bind to make the catalytically productive complex. Replacement of a histidine residue in the amine binding region with aspartate improved the binding of the amine donor substrate approximately 2.5-fold, likely due to favorable electrostatic interactions, and increased the yield of aspartame derivative approximately 2-fold.[19]

Substrates smaller than the natural one often react poorly due to non-productive binding. Although the substrate fits in the binding site, it adopts a non-productive orientation as it adjusts within the larger active site to find favorable non-covalent interactions. The difficulty in protein engineering is that one rarely knows the nature of this non-productive orientation and how to prevent it. For example, expanding the active site with a Ile303Ala substitution expanded the substrate range of a ligase from the native substrate 4-chlorobenzoate to a larger substrate 3,4-chlorobenzoate, Fig 7.18. However, a smaller regioisomer, 3-chlorobenzoate, reacted about 110-fold slower. The researchers hypothesized that the smaller 3-chlorobenzoate must fit in the active site, but bind in a non-productive orientation. Additional substitutions to close the space for the 4-chloro

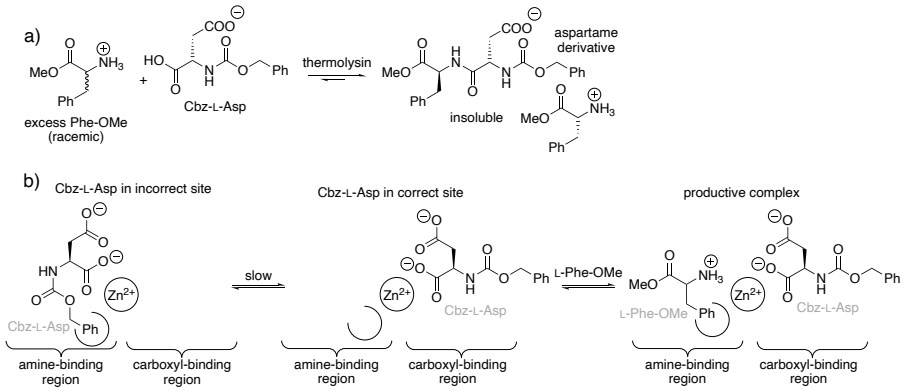

Figure 7.17. Thermolysin-catalyzed coupling yields an aspartame derivative. a) Thermolysin catalyzes the coupling of L-phenylalanine methyl ester (L-Phe-OMe) with Cbz-L-Asp to yield an aspartame derivative. This derivative precipitates from the aqueous reaction mixture by forming a salt with D-phenylalanine methyl ester. The starting Phe-OMe is racemic. Thermoylsin selects the L-enantiomer for reaction, while the D-enantiomer is the counter ion for salt formation. b) This coupling is slow because the carboxyl donor, Cbz-L-Asp, initially binds in the incorrect location. It binds in the amine donor site, thereby blocking access of L-Phe-OMe. Once the Cbz-L-Asp rotates and moves out of the way, L-Phe-OMe can bind and the reaction proceeds. Cbz = benzyloxycarbonyl.

substituent (Phe184Trp) and orient 3-chlorobenzoate with a hydrogen bond to the 3-chloro substituent (Val209Thr) improved k_{cat}/K_M about 5-fold.

Molecular modelling of preorganization and non-productive binding uses docking and molecular dynamics to model the different orientations of the substrate within the active site. Researchers classify the orientations as either productive or non-productive complexes. Substitutions that favor formation of the near-attack complexes are expected to increase k_{cat}.

7.5.2 Optimize acids & bases

Many reactions involve the transfer of protons, so acids (proton donors) and bases (proton acceptors) must be in the correct protonation state for the reaction to occur. For example, serine esterases require the histidine of the catalytic triad to act as a base in the first chemical step, Fig 7.19. Only the neutral form of the histidine side chain can act as a base. The protonated histidine has already accepted a proton (acted as a base) and cannot accept at second proton. At low pH, this histidine is mostly protonated so only the tiny fraction that is neutral can contribute to catalysis, Fig 7.19. As the pH increases, the fraction of neutral histidine increases, and the observed rate increases. The inflection point of the increase in activity corresponds to the pK_a of the catalytic histidine. Once the pH increases beyond the pK_a of the histidine, the fraction of neutral histidine approaches 100% and the reaction proceeds at the maximum rate. Thus, the reaction rate varies with pH because the fraction of enzyme in the catalytically active form varies.[20]

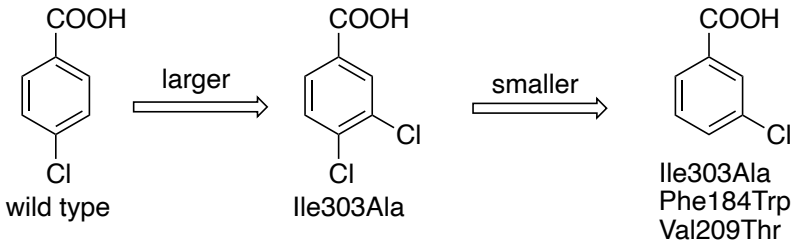

Figure 7.18. Expanding the substrate range of 4-chlorobenzoate:coenzyme A ligase to a larger substrate (3,4-chlorobenzoate) and to a regioisomer (3-chlorobenzoate). Enlarging the active site improved reaction of the larger substrate, but not of the regioisomer, likely due to non-productive binding. Additional adjustments to match the shape of the regioisomer improved its reactivity 5-fold. The reaction rates with the unnatural substrates were 100-1000-fold lower than the natural substrates even with the engineered variants.

The hydrolysis of β-lactam antibiotics by penicillin-binding proteins is slow because the pK_a of Tyr150 is too high thereby preventing it from acting as a base.[21]

A general equation for pH-dependent behavior is

$$k_{obs} = \frac{k_{A^-} \cdot K_a + k_{HA} \cdot [H^+]}{K_a + [H^+]} \tag{7.19}$$

where k_{A^-} is the rate constant of the deprotonated form, k_{HA} is the rate constant for the protonated form and K_a is the ionization constant of the acid. If the reaction requires the basic form, then one sets k_{HA} to zero and equation 7.19 simplifies to equation 7.20 below. The observed rate constant, k_{obs}, is the rate constant for the deprotonated form, k_{A^-}, scaled by the by the fraction of the molecule in the deprotonated form.[‡]

[‡]The fraction deprotonated, $\frac{[A^-]}{[A^-]+[HA]}$ at a given pH depends on the pH and the acid dissociation constant, K_a. The acid dissociation constant is defined as:

$$K_a = \frac{[H^+] \cdot [A^-]}{[HA]}$$

Solving for $[HA]$ yields:

$$[HA] = \frac{[H^+] \cdot [A^-]}{K_a}$$

Substitution into the definition for fraction deprotonated yields:

$$\text{fraction deprotonated} = \frac{[A^-]}{[A^-] + \frac{[H^+] \cdot [A^-]}{K_a}}$$

which simplifies to

$$\text{fraction deprotonated} = \frac{K_a}{K_a + [H^+]}$$

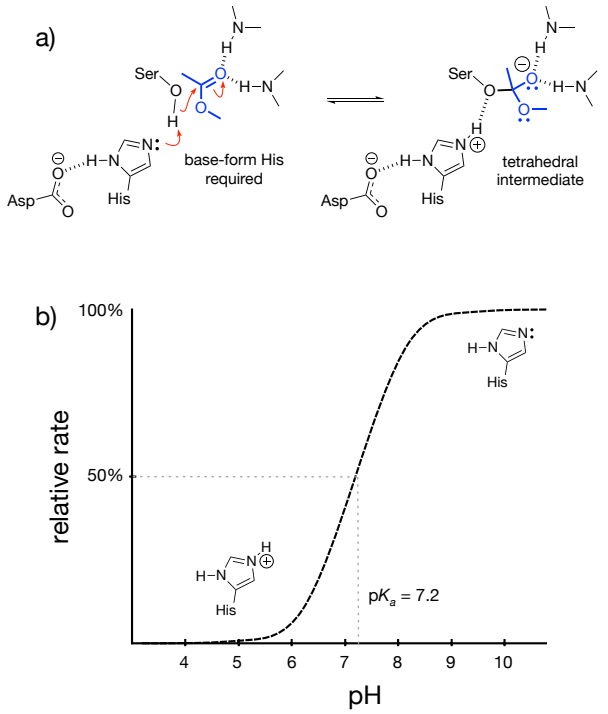

Figure 7.19. The catalytic mechanism of serine esterases requires the deprotonated form of the catalytic histidine. a) The first chemical step of a serine esterase-catalyzed hydrolysis of an ester requires the catalytic histidine to act as a base and deprotonate the catalytic serine. b) The activity of the esterase varies with pH. The esterase reaches full activity when the histidine is fully deprotonated. The inflection point corresponds to the pK_a of the catalytic histidine. Here the pK_a of the catalytic histidine is 7.2, slightly higher than that for the free histidine due to the environment of the active site.

$$k_{obs} = k_{A^-} \cdot \frac{K_a}{K_a + [H^+]} \tag{7.20}$$

The pK_a of a functional group within a folded protein may differ from that in free solution due to changes in solvation, hydrogen bonding and electrostatic interactions with the protein. For example, the phenolic hydroxyl of tyrosine has a pK_a of ~10 in solution, but Tyr80 in the xylanase has a predicted pK_a of ~19. This increase in pK_a corresponds to a billion-fold shift in the equilibrium between phenol and phenoxide toward the phenol. The folded protein environment dramatically destabilizes the charged phenoxide form, in part by surrounding Tyr80 with hydrophobic residues that hinder solvation. The phenol form of Tyr80 may help position a catalytic glutamate.[22]

Besides solvation, hydrogen bonding and electrostatic effects also shift the pK_a of ionizable groups in proteins. When hydrogen bond donors or acceptors interact with ionizable groups and the interaction is unequal for the protonated and deprotonated forms, then their pK_a will shift, Fig 7.20. The pK_a's also shift when nearby charged groups stabilize or destabilize the charges associated with protonation changes. The key to these changes is that the interactions are unequal – either the protonated or deprotonated state interacts more strongly.

PROPKA predicts shifts in pK_a. PROPKA estimates the pK_a of ionizable groups in a protein (p$K_a^{protein}$) by adjusting the pK_a of the groups in solution (p$K_a^{solution}$) for their different environment in the folded protein.[23] The estimate requires a protein structure and assumes that the structure is static. The program uses empirical equations and parameters to account for desolvation (Δp$K_a^{desolvation}$, disfavors charged form), formation of hydrogen bonds (Δp$K_a^{hydrogen\,bonds}$, H-bond donors favor deprotonated form, H-bond acceptors favor protonated form) and electrostatic effects of nearby charged groups (Δp$K_a^{electrostatic}$, oppositely charged groups favor charged form), eq. 7.21. A web interface of PROPKA is available at https://server.poissonboltzmann.org/pdb2pqr.

$$\text{p}K_a^{protein} = \text{p}K_a^{solution} + \Delta\text{p}K_a^{desolvation} + \Delta\text{p}K_a^{hydrogen\,bonds} + \Delta\text{p}K_a^{electrostatic} \tag{7.21}$$

A more detailed look at the predicted altered pK_a of Tyr80 in the active site of xylanase, Fig 7.21, shows that, besides desolvation mentioned above, hydrogen bond differences and Coulombic interactions also contribute. The phenoxide is disfavored because desolvation destabilized the phenoxide as mentioned above, because Glu172 makes a favorable hydrogen bond with the phenol form, and because two nearby negatively-charged glutamates create unfavorable Coulombic interactions with the phenoxide, eq. 7.22. PROPKA predictions typically fit experimentally determined values with an rmsd <1 pK_a unit.

$$\text{p}K_a^{Try80} = 10 + 3.55 + 0.79 + 4.63 = 18.98 \tag{7.22}$$

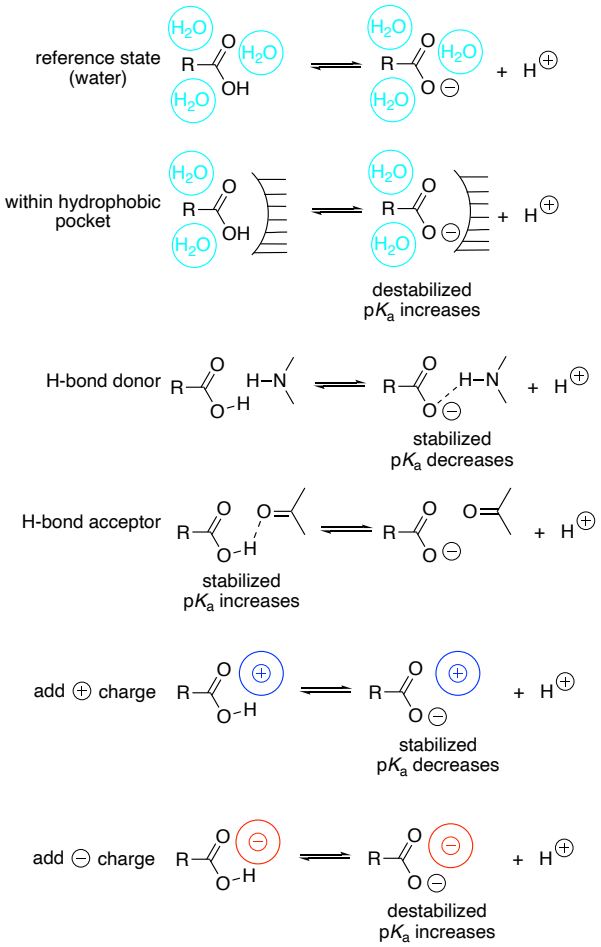

Figure 7.20. The pK_a of a carboxylic acid changes in a folded protein as compared to solution due different in the interactions between the protein and the two protonation states. Placing the carboxylic acid in a hydrophobic pocket destabilizes the charged form more than the neutral form, thereby shifting the equilibrium toward the neutral form. Hydrogen bond donors and acceptors stabilize different protonation forms and shift the pK_a in opposite directions. These hydrogen bond donor and acceptors can also interact with the carboxylic acid/carboxylate in other ways; the interactions shown are those the cause differences in the interactions. Electrostatic effects also shift the pK_a of the carboxylic acid.

RESIDUE	pKa	BURIED	DESOLVATION REGULAR	EFFECTS RE	SIDECHAIN HYDROGEN BOND	BACKBONE HYDROGEN BOND	COULOMBIC INTERACTION
TYR80 A	18.98	100 %	3.55 632	0.00 0	0.79 GLU 172 A	0.00 XXX 0 X	-0.27 ARG 112 A
TYR80 A					0.00 XXX 0 X	0.00 XXX 0 X	0.68 TYR 65 A
TYR80 A					0.00 XXX 0 X	0.00 XXX 0 X	0.16 TYR 166 A
TYR80 A					0.00 XXX 0 X	0.00 XXX 0 X	2.03 GLU 78 A
TYR80 A					0.00 XXX 0 X	0.00 XXX 0 X	2.03 GLU 172 A

Figure 7.21. Part of a PROPKA prediction of the pK_a's of ionizable groups in xylanase from *Bacillus circulans* (pdb id: 1XNB). PROPKA predicts that the pK_a of the phenolic hydroxyl of Tyr80 is 18.98, much higher than the solution value of 10. The side chain is buried disfavoring the charged anion, which raises the pK_a by 3.55 units. This estimate comes from the 632 (under desolvation effects 'regular') non-hydrogen protein atoms within 15.5 Å of the phenolic oxygen, which indicates how deeply the group is buried. The desolvation effects 're' is a small correction (zero in this case) to account for the non-hydrogen-bonding electrostatic effects of replacing water with protein atoms. The Tyr80 phenol hydrogen bonds to the anionic Glu172, but the phenoxide loses this hydrogen bond, which raises the pK_a by 0.79 units. Coulombic interactions with five nearby groups raise the pK_a by 4.63 units. Negative charges on two nearby glutamates strongly disfavor formation of the negatively-charged phenoxide. Smaller contributions come from a more distant positively-charged arginine and partial negative charges on two nearby tyrosines.

Shifting an enzyme's pH-rate profile. When the pH optimum of the enzyme differs from the pH of the application, then engineering a shift in the pH optimum toward the pH of the application will increase the reaction rate. For example, some subtilisins (proteases for laundry applications) reach their maximum activity above pH 10, but some wash applications use pH 8 where the enzyme shows low activity. Phytase, a feed additive enzyme, must be active in digestive tract, which is pH 2-5 for pigs and poultry. Other reasons to shift the pH optimum of an enzyme are to get enzymes with different pH optima to work together in a multi-step pathway, to allow a reaction pH that alters the solubility of substrates or products or partitioning into an organic phase, or to reduce susceptibility of microbial contamination by running process outside of neutral range.[24]

Since solvation, hydrogen bond networks and electrostatic effects all influence the pK_a of the ionizable groups, researchers can, in principle, use any of these to engineer a shift in the pH optimum. However, since the first two involve changes near the ionizable group, there is a danger of disrupting catalytic activity and researchers avoid these approaches. Most engineering to shift the pH optimum uses remote electrostatic effects. Electrostatic effects decrease proportional to $1/r$, so charges even 10 Å away still affect the pK_a of an ionizable group. Being further away from the group makes it less likely that the changes disrupt orientation of the catalytic residues or the shape of the active site. Shifting the pH optimum of an enzyme usually involves changing the type and number of charges outside the active site.

The strategy to alter the pH optimum by changing electrostatics is as follows, Fig 7.22.

- Add $\oplus$ charge or remove $\ominus$ charge = lower pH optimum

- Add ⊖ charge or remove ⊕ charge = raise pH optimum

The example below considers the effect of added charges on the pK_a of histidine. The reader should confirm that the pK_a of a negatively-charged side chain like glutamate also shifts in the same direction.

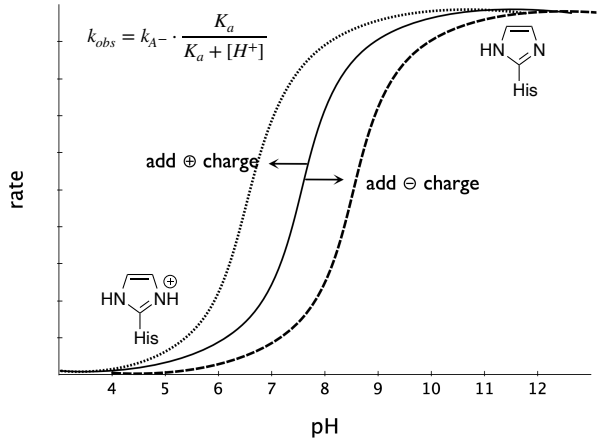

Figure 7.22. Shifting the pK_a of histidine using electrostatic effects. The curve of rate versus pH assumes that catalysis requires the base form of histidine. Adding a nearby positive charge destabilizes the protonated form of histidine so the curve shifts to lower pH. Similarly, adding a nearby negative charge shifts the curve to higher pH.

For example, to increase the activity of a serine hydrolase (subtilisin) at lower pH, the engineering must shift the pK_a of the catalytic histidine to lower pH. Adding positive charges or removing negative charges will destabilize the protonated histidine, so that the inflection point for deprotonation occurs at lower pH. Replacing a distant (~12 Å away) glutamate or a distant aspartate with lysine in shifted this inflection point to lower pH by ~0.5 pH units.[25] The effects were approximately additive, so replacing both residues with lysine decreased the pH of the inflection point by ~1.0 pH units. The lysine destabilized the positively charged (inactive) form of catalytic histidine so the shift to the active from occurred at lower pH. Increasing the ionic strength lowers the effect of surface charges due to their association with counterions. These measurements were made a low ionic strength. Immobilization of enzymes on charged surfaces can similarly shift the pH optimum.

Catalysis often requires several ionizable groups in the correct protonation state. For, example, glycosidases require two acids, one in the acidic form, the other in the basic form, Fig 7.23.

Eq. 7.23 below predicts the fraction of the enzyme that exists in the catalytically active form at different pH. The k_{obs} is the rate constant at the pH in question, while k_{best} is the rate of the catalytically active form and K_{a1} and K_{a2} are the acid ionization constants of the two ionizable groups. The active form of the enzyme require the deprotonated

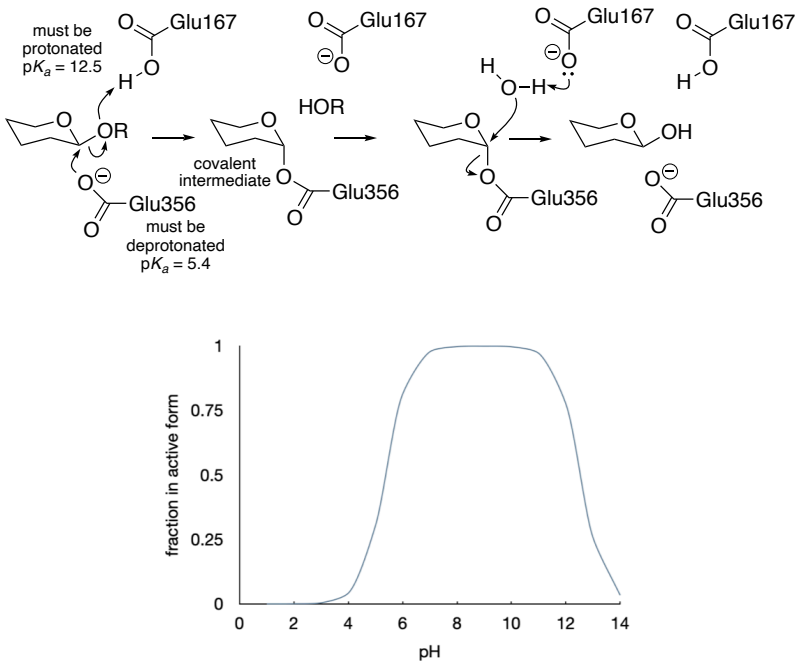

Figure 7.23. A glycosidase require one glutamte in its an acid form and another in its basic form. Their catalytic activity peaks in the range where both are in the correct protonation state. The first step of the catalytic mechanism requires Glu167 to act as an acid and and Glu356 to act as a nucleophile, which requires the basic form. If the $\mathrm{p}K_a^{Glu356} = 5.4$ and $\mathrm{p}K_a^{Glu167} = 12.5$, then the glycosidase is the fully catalytic form between pH 7 and 11. At lower pH Glu356 is not deprotonated, while at higher pH Glu167 is not protonated. The curve is drawn from equation 7.23.

form of group 1 and the protonated form of group 2.

$$k_{obs} = k_{best} \cdot \frac{[H^+]K_{a1}}{[H^+]K_{a1} + [H^+]^2 + K_{a1}K_{a2}} \tag{7.23}$$

If the pK_a's of the two groups are widely separated, e.g., 6.0 and 10.0, then the maximum activity can reach to nearly k_{best}, Fig 7.24. The optimum at pH 8.0 is broad, halfway between the two pK_a's. If the pK_a's of the two groups closer to each other, e.g., 6.0 and 7.3, then the maximum activity can never reach k_{best}. At best 69% of the enzyme molecules are in the correct protonation state. The pH optimum is narrow.

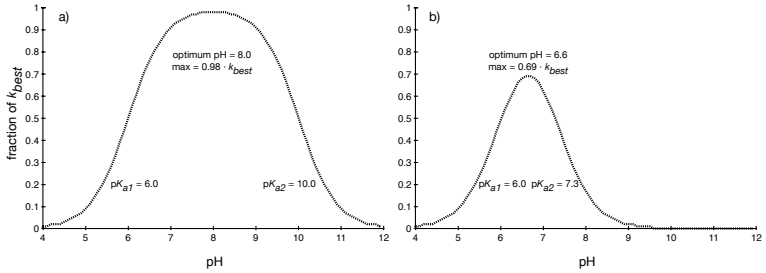

Figure 7.24. Narrowing the pH range for optimum activity reduces the catalytic activity. a) If the pK_a's of the two groups are well separated (6.0 and 10.0), then the pH optimum is broad and the activity reaches the maximum possible. b) If the pK_a's of the two groups are closer to each other (6.0 and 7.3), then the pH optimum is narrow and the activity only reaches part way to the maximum possible because only a fraction of the enzyme can contain the correctly ionized groups.

If protein engineering to shift the pH optimum also narrows the gap between the two pK_a's, then the overall activity may decrease. For example, adding an arginine decreased the pH optimum of the glycosidase xylanase from 6.5 to 5, but unfortunately also decreased the catalytic activity 100-fold.[26] The added arginine was closer to one of the carboxylic acids and strongly shifted its pK_a. The narrowing of the difference between the two pK_a's made the correct protonation state for catalysis rarer thereby slowing the reaction.

> One common confusion is the notion of regions of high pH or low pH. There are no such regions; the pH is constant throughout the solution since proton transfer is fast. The protonation state of all groups is at equilibrium and reflects its pK_a. However the same groups in different environments may differ in their pK_a therefore have different protonation states. A carboxyl group in a non-polar environment may be uncharged while one in a polar environment may be charged. The pK_a of those two carboxyl groups is different, but the pH is identical in both places.

One reason that shifting the pH optimum may fail is that the group added to introduce charge may not be charged as expected at the required pH. For example, consider the goal of increasing the pH of the inflection point of a serine hydrolase from 6.5 to 7.5.

Adding a negative charge near the histidine will increase its pK_a because the negative charge stabilizes the protonated histidine, Fig 7.22 above. Replacing a nearby neutral residue with glutamate or aspartate is expected to introduce a negative charge since the pK_a of the carboxylic side chain is ~5 so it will be negatively charged at pH 6.5-7.5. However, if the added carboxylate side chain lies in a hydrophobic region where the carboxylate is poorly solvated, its pK_a may increase from 5 to 9. In this case, the charged histidine and charged carboxylate never coexist because the charged histidine exists below pH 6.5 and the charged carboxylate exists above pH 9. The charged carboxylate cannot stabilize the charged histidine, so the pK_a of the histidine does not change. Thus, although adding charges shifts the pK_a of groups in the active site by interacting with the charged form of those groups, the 'added charges' must be charged at the same pH as the charged form of the active site groups.

In summary, the effect of proteins on a ionizable groups is predictable, but complex, because all charges interact. Engineering to shift the pK_a of catalytic groups can shift the pH optimum of proteins while maintaining high catalytic activity. These shifts are typically ≤1 pH unit.

7.5.3 Stabilize shape and charge of transition state

To selectively stabilize the transition state over the enzyme-substrate complex, the enzyme must stabilize features that are unique to the transition state and not present in the enzyme-substrate complex. Those features are likely to be the charge and shape of the transition state. For example, consider the transition state for hydrolysis of an ester catalyzed by a serine hydrolase, Fig 7.25. Hydrolysis of an ester involves an acyl-enzyme intermediate. After binding the substrate and formation of the first tetrahedral intermediate, the acyl-enzyme intermediate forms, then a second tetrahedral intermediate forms, and finally the product forms and dissociates. In most cases, the transition state (highest energy species) is the acylation step, which forms the acyl-enzyme intermediate. The structure of the transition state is not shown, but it is assumed to be similar to the first tetrahedral intermediate, Fig 7.26.

The transition state differs in charge from the enzyme substrate complex. First, positive charge on the imidazole ring increases in the transition state as it is protonated. Second, the negative charge on the carbonyl oxygen increases in the transition state as it form an oxyanion. Serine hydrolases stabilize the positive charge on the imidazole ring with the carboxylate side chain of the catalytic aspartate. Serine hydrolases stabilize the the negative-charge on the oxyanion with two hydrogen bonds from main chain N-H's.

The shape of the transition state also differs from the enzyme-substrate complex. The transition state is tetrahedral at the carbonyl carbon, while the enzyme-substrate complex is planar at the carbonyl carbon. If the shape of the active site matches the tetrahedral arrangement better than the planar arrangement, then this shape selectively stabilizes the transition state.

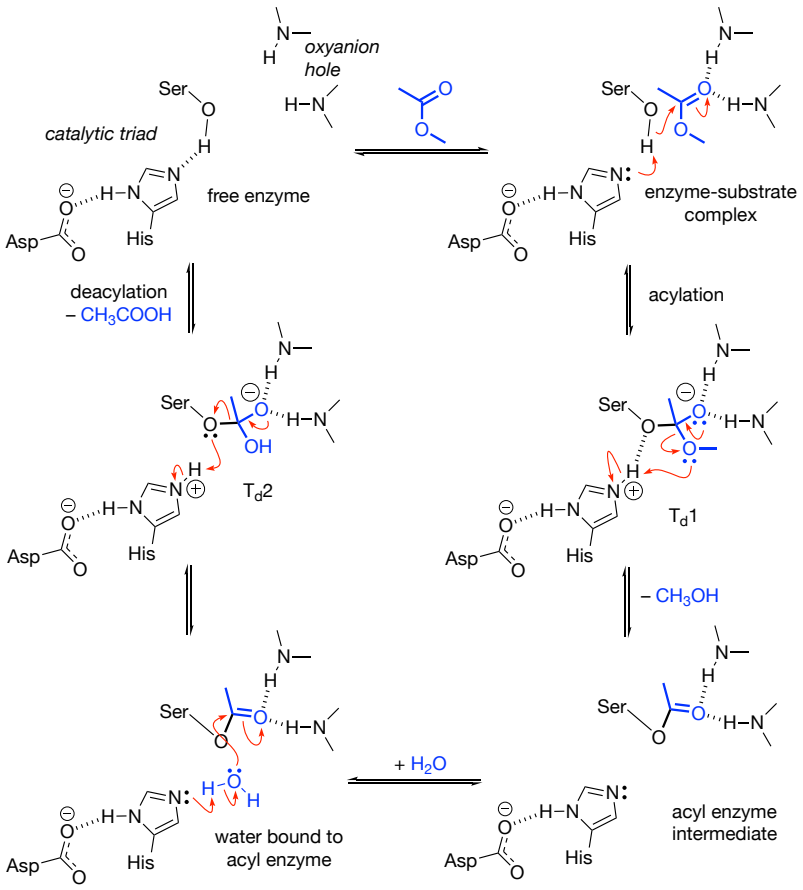

Figure 7.25. Serine esterases use a multi-step mechanism for hydrolysis of esters. Hydrolysis proceeds via an acyl-enzyme intermediate (bottom right structure). After formation of the enzyme-substrate complex (top equilibrium), the catalytic serine γ-oxygen adds to the carbonyl to form a tetrahedral intermediate. Collapse of this intermediate forms the acyl-enzyme. A similar set of steps hydrolyze this acyl enzyme to complete the cycle.

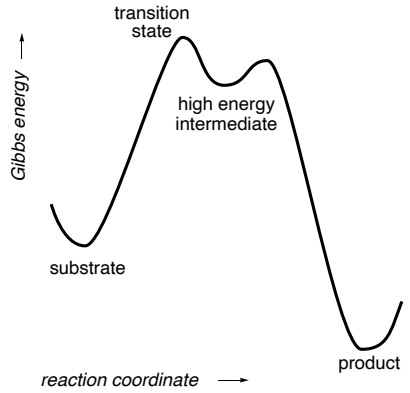

Figure 7.26. Both transition states and high energy intermediates lie in similar regions of the reaction coordinate diagram and therefore resemble each other. Transition states correspond to peaks in the diagram indicating that they have a fleeting existence, while intermediates correspond to valleys indicating that they have a finite lifetime.

7.5.4 Create new mechanistic steps

The rate-acceleration mechanisms above assume that the transition states for the enzyme-catalyzed and the uncatalyzed reactions are the same; that is, that the mechanisms for the two reactions are the same. However, enzymes often change the mechanism of the uncatalyzed reaction. These changes split the high energy step of the uncatalyzed reaction into multiple lower-energy steps, Fig 7.27. These mechanistic changes may involve covalent enzyme intermediates, the participation of cofactors, such as pyridoxal phosphate or metal ions, and acid-base catalysis not available to the uncatalyzed reaction.

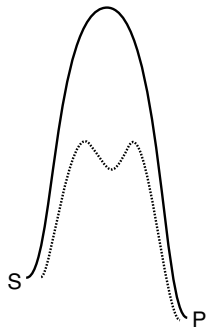

Figure 7.27. One way to lower the energy of a transition state is to break an high energy step into two lower energy steps. This change corresponds to changing the mechanistic steps involved in the reaction.

One type of new mechanism is the formation of a covalent intermediate with the enzyme.

Both glycoside hydrolases, Fig 7.23, and serine esterases Fig 7.25 form covalent intermediates: a glycosyl enzyme link to the catalytic glutamte residue and a acyl enzyme link to the catalytic serine residue, respectively. Instead of water directly attacking the substrate, a catalytic residue within the enzyme attacks the substrate and releases part of it while forming a covalent intermediate. In the next step, water attacks this covalent intermediate to release it from the enzyme. In contrast, the acid catalyzed hydrolysis of glycosides and esters involved direct attack of water on the substrate and formation of a high energy intermediate, Fig 7.28. This mechanism is higher in energy for two reasons. First, simultaneously positioning the water and substrate for reaction is entropically costly. The enzyme uses a stepwise approach: position the substrate in the first step and the water in the second step. Second, the charged intermediate is much less stable than the neutral covalent intermediates formed in the enzyme-catalyzed reaction.

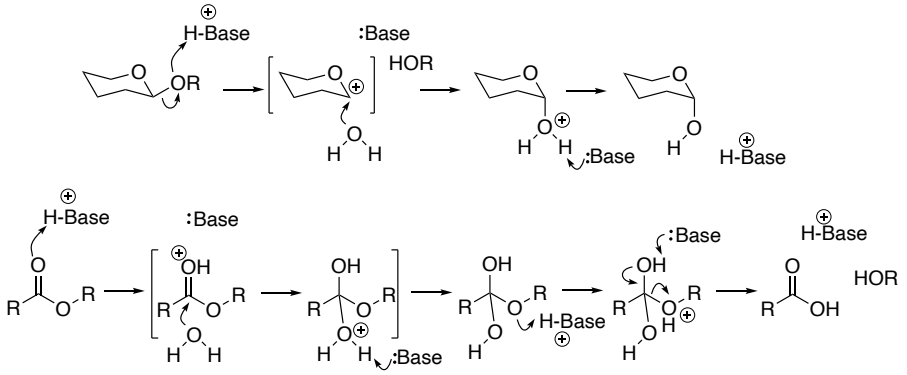

Figure 7.28. Acid-catalyzed hydrolysis of a glycoside (top) or an ester (bottom) involves direct attack of water on the substrate and formation of a high energy charged intermediates, shown in brackets. Enzymes avoid formation of these high energy intermediates by breaking the reaction into two steps where a catalytic residue first attacks the substrate to form a covalent intermediate and then water attacks the covalent intermediate. The neutral covalent intermediate is lower in energy than the charged intermediates shown here. See Fig 7.23 and 7.25 for the enzyme-catalyzed mechanisms.

A comparison of binding strengths to enzyme rate acceleration shows the importance of these mechanistic changes to enzyme catalysis. If binding the transition state were the only way to stabilize the transition state, then transition state stabilization energies would be similar to binding constants. The maximum strength of binding between receptors and drugs, antibodies and antigens, or enzyme and drugs corresponds to K_d ~10^{-13} M, so the maximum rate acceleration by enzymes would be similar. In contrast to this expectation, the maximum rate acceleration by enzyme corresponds to a binding constant a billion times lower: 10^{-22} M.[27] This additional rate acceleration comes from the ability of enzymes to change the catalytic mechanism.

7.6 Challenges in enzyme design

Given the complex interactions and multiple physical steps needed to catalyze chemical reactions, it is not surprising that enzyme design remains challenging. Computational modeling typically chooses one part (e.g., substrate binding, formation of near attack complex, reaction step, etc.) to model, but all of the steps much function for catalysis. The accuracy needed to design the multiple steps and interactions that stabilize transition states is at the edge of the capabilities of current computational methods.[28]

An number of research groups have engineered enzymes for higher catalytic activity toward new substrates.[29] For example, the reactivity (k_{cat}/K_M) of the protease kumamysin toward gluten peptides increased >100-fold.[9] The reactivity (k_{cat}/K_M) of adenosine deaminase increased >4,000-fold toward was a fluorogenic nerve gas analog (7-O-diethylphophorlyl-2-cyano-7-hydroxycoumarin), but the k_{cat} was still very low (~ 0.0002 s^{-1} and required additional directed evolution to improve the enzyme.[30]

The engineering of enzymes with new catalytic activity remains difficult. One of the first examples of computational enzyme design was the design of a Kemp eliminase. The Kemp elimination involves removing a proton from the ring carbon, adding a proton to the ring oxygen, converting the C–N double bond into a triple bond, with the accompanying shift in geometry, and breaking the O–N bond, Fig 7.29. Besides optimizing the pK_a of the acid and base for this reaction as described in the previous section, a catalyst can match the shape and charge of this transition state. The O–N bond lengthens and the orientation of the C≡N changes. The distribution of charge changes to accommodate the new bonding. No natural enzymes evolved to catalyze this unnatural reaction, but Röthlisberger and coworkers[15] designed a protein that catalyzed this reaction by modeling the binding of this transition state within a protein. The choice of a non-flexible substrate and a single-step reaction simplified the design.

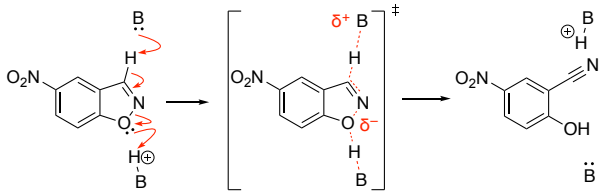

Figure 7.29. he Kemp elimination reaction is a one-step reaction involving proton transfers, breaking of the N–O bond and formation of the C–N triple bond. This reaction proceeds via a single transition state involving a base and an acid. The transition state differs from the substrates in increased partial charges at atoms marked by δ and by altered bond lengths marked by red dashed lines.

The researchers predicted and tested 59 protein catalysts; of these, eight showed measurable activity. The best catalyst followed the Michaelis-Menten equation with $k_{cat} = 0.29$ s^{-1} and K_M = 1.8 mM, which corresponds to a k_{cat}/K_M value of ~160 M^{-1} s^{-1}, far lower than the ~10^5 M^{-1} that is typical for enzymes. Despite the low catalytic efficiency, this designed enzyme accelerates the reaction 2.5 $\times 10^5$-fold over the uncatalyzed reaction. The design created a hydrophobic pocket to orient the substrate for

reaction, including hydrogen bonds to the nitro group. Stacking the aromatic ring of a tryptophan side chain on top of the substrate stabilized the new charges within the transition state. The carboxylate of Glu231 was the catalytic base, but the design did not include a catalytic acid. Presumably, a water molecule served as the proton donor. The authors hypothesize that the omission of flexibility and long-range second shell interactions lowered the precision of the design.

Glossary

item coming soon

References

1. Garcia-Viloca, M., Gao, J., Karplus, M., & Truhlar, D. G. (2004). How enzymes work: Analysis by modern rate theory and computer simulations. *Science*, *303*(5655), 186–195. https://doi.org/10.1126/science.1088172

2. Bozlee, B. J. (2007). Reformulation of the Michaelis–Menten equation: How enzyme-catalyzed reactions depend on Gibbs energy. *J. Chem. Ed.*, *84*(1), 106. https://doi.org/10.1021/ed084p106

3. Cantor, J. R., Yoo, T. H., Dixit, A., Iverson, B. L., Forsthuber, T. G., & Georgiou, G. (2011). Therapeutic enzyme deimmunization by combinatorial T-cell epitope removal using neutral drift. *Proc. Natl. Acad. Sci. U. S. A.*, *108*(4), 1272–1277. https://doi.org/10.1073/pnas.1014739108

4. Bhosale, S. H., Rao, M. B., & Deshpande, V. V. (1996). Molecular and industrial aspects of glucose isomerase. *Microbiol. Rev.*, *60*, 280–300.

5. Fox, R. J., & Clay, M. D. (2009). Catalytic effectiveness, a measure of enzyme proficiency for industrial applications. *Trends Biotechnol.*, *27*(3), 137–140. https://doi.org/10.1016/j.tibtech.2008.12.001

6. Bastian, S., Liu, X., Meyerowitz, J. T., Snow, C. D., Chen, M. M. Y., & Arnold, F. H. (2011). Engineered ketol-acid reductoisomerase and alcohol dehydrogenase enable anaerobic 2-methylpropan-1-ol production at theoretical yield in *Escherichia coli*. *Metab. Eng.*, *13*(3), 345–352. https://doi.org/10.1016/j.ymben.2011.02.004

7. Liu, X., Bastian, S., Snow, C. D., Brustad, E. M., Saleski, T. E., Xu, J.-H., Meinhold, P., & Arnold, F. H. (2012). Structure-guided engineering of *Lactococcus lactis* alcohol dehydrogenase LlAdhA for improved conversion of isobutyraldehyde to isobutanol. *J. Biotechnol.*, *164*(2), 188–195. https://doi.org/10.1016/j.jbiotec.2012.08.008

8. Savile, C. K., Janey, J. M., Mundorff, E. C., Moore, J. C., Tam, S., Jarvis, W. R., Colbeck, J. C., Krebber, A., Fleitz, F. J., Brands, J., Devine, P. N., Huisman, G. W., & Hughes, G. J. (2010). Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science*, *329*(5989), 305–309. https://doi.org/10.1126/science.1188934

9. Gordon, S. R., Stanley, E. J., Wolf, S., Toland, A., Wu, S. J., Hadidi, D., Mills, J. H., Baker, D., Pultz, I. S., & Siegel, J. B. (2012). Computational design of an α-gliadin peptidase. *J. Am. Chem. Soc.*, *134*(50), 20513–20520. https://doi.org/10.1021/ja3094795

10. Pultz, I. S., Hill, M., Vitanza, J. M., Wolf, C., Saaby, L., Liu, T., Winkle, P., & Leffler, D. A. (2021). Gluten degradation, pharmacokinetics, safety, and tolerability of TAK-062, an engineered enzyme to treat celiac disease. *Gastroenterology*, *161*(1), 81–93. https://doi.org/10.1053/j.gastro.2021.03.019

11. Pavlidis, I. V., Weiss, M. S., Genz, M., Spurr, P., Hanlon, S. P., Wirz, B., Iding, H., & Bornscheuer, U. T. (2016). Identification of (*S*)-selective transaminases for the asymmetric synthesis of bulky chiral amines. *Nat. Chem.*, *8*(11), 1076–1082. https://doi.org/10.1038/nchem.2578

12. Gora, A., Brezovsky, J., & Damborsky, J. (2013). Gates of enzymes. *Chem. Rev.*, *113*(8), 5871–5923. https://doi.org/10.1021/cr300384w

13. Bruice, T. C., & Lightstone, F. C. (1999). Ground state and transition state contributions to the rates of intramolecular and enzymatic reactions. *Acc. Chem. Res.*, *32*(2), 127–136. https://doi.org/10.1021/ar960131y

14. Hur, S., & Bruice, T. C. (2003). Comparison of formation of reactive conformers (NACs) for the Claisen rearrangement of chorismate to prephenate in water and in the E. Coli mutase: The efficiency of the enzyme catalysis. *J. Am. Chem. Soc.*, *125*(19), 5964–5972. https://doi.org/10.1021/ja0210648

15. Röthlisberger, D., Khersonsky, O., Wollacott, A. M., Jiang, L., DeChancie, J., Betker, J., Gallaher, J. L., Althoff, E. A., Zanghellini, A., Dym, O., Albeck, S., Houk, K. N., Tawfik, D. S., & Baker, D. (2008). Kemp elimination catalysts by computational enzyme design. *Nature*, *453*(7192), 190–195. https://doi.org/10.1038/nature06879

16. Kiss, G., Röthlisberger, D., Baker, D., & Houk, K. N. (2010). Evaluation and ranking of enzyme designs. *Protein Sci*, *19*(9), 1760–1773. https://doi.org/10.1002/pro.462

17. Isowa, Y., Ohmori, M., Ichikawa, T., Mori, K., Nonaka, Y., Kihara, K., Oyama, K., Satoh, H., & Nishimura, S. (1979). The thermolysin-catalyzed condensation reactions of N-substituted aspartic and glutamic acids with phenylalanine alkyl esters. *Tetrahedron Lett.*, *20*(28), 2611–2612. https://doi.org/10.1016/S0040-4039(01)86363-2

18. Birrane, G., Bhyravbhatla, B., & Navia, M. A. (2014). Synthesis of aspartame by thermolysin: An x-ray structural study. *ACS Med. Chem. Lett.*, *5*(6), 706–710. https://doi.org/10.1021/ml500101z

19. Zhu, F., Jiang, T., Wu, B., & He, B. (2018). Enhancement of Z-aspartame synthesis by rational engineering of metalloprotease. *Food Chem.*, *253*, 30–36. https://doi.org/10.1016/j.foodchem.2018.01.108

20. Tipton, K. F., & Dixon, H. B. F. (1979). Effects of pH on enzymes. *Meth. Enzymol.*, *63*, 183–234. https://doi.org/10.1016/0076-6879(79)63011-2

21. Gherman, B. F., Goldberg, S. D., Cornish, V. W., & Friesner, R. A. (2004). Mixed quantum mechanical/molecular mechanical (QM/MM) study of the deacylation reaction in a penicillin binding protein (PBP) versus in a class C β-lactamase. *J. Am. Chem. Soc.*, *126*(24), 7652–7664. https://doi.org/10.1021/ja036879a

22. Wakarchuk, W. W., Campbell, R. L., Sung, W. L., Davoodi, J., & Yaguchi, M. (1994). Mutational and crystallographic analyses of the active site residues of the *Bacillus {c}irculans* xylanase. *Protein Sci.*, *3*(3), 467–475. https://doi.org/10.1002/pro.5560030312

23. Olsson, M. H. M. (2011). Protein electrostatics and pK_a blind predictions; contribution from empirical predictions of internal ionizable residues. *Proteins*, *79*(12), 3333–3345. https://doi.org/10.1002/prot.23113

24. Tynan-Connolly, B. M., & Nielsen, J. E. (2007). Redesigning protein pK_a values. *Protein Sci.*, *16*(2), 239–249. https://doi.org/10.1110/ps.062538707

25. Russell, A. J., & Fersht, A. R. (1987). Rational modification of enzyme catalysis by engineering surface charge. *Nature*, *328*(6130), 496–500. https://doi.org/10.1038/328496a0

26. Pokhrel, S., Joo, J. C., & Yoo, Y. J. (2013). Shifting the optimum pH of Bacillus circulans xylanase towards acidic side by introducing arginine. *Biotechnol. Bioprocess Eng.*, *18*(1), 35–42. https://doi.org/10.1007/s12257-012-0455-x

27. Zhang, X., & Houk, K. N. (2005). Why enzymes are proficient catalysts: Beyond the Pauling paradigm. *Acc. Chem. Res.*, *38*(5), 379–385. https://doi.org/10.1021/ar040257s

28. Linder, M. (2012). Computational enzyme design: Advances, hurdles and possible ways forward. *Comput. Struct. Biotechnol. J.*, *2*(3). https://doi.org/10.5936/csbj.201209009

29. Wijma, H. J., & Janssen, D. B. (2013). Computational design gains momentum in enzyme catalysis engineering. *FEBS J*, *280*(13), 2948–2960. https://doi.org/10.1111/febs.12324

30. Khare, S. D., Kipnis, Y., Greisen, P. J., Takeuchi, R., Ashani, Y., Goldsmith, M., Song, Y., Gallaher, J. L., Silman, I., Leader, H., Sussman, J. L., Stoddard, B. L., Tawfik, D. S., & Baker, D. (2012). Computational redesign of a mononuclear zinc metalloenzyme for organophosphate hydrolysis. *Nat. Chem. Biol.*, *8*(3), 294–300. https://doi.org/10.1038/nchembio.777

31. Kemmer, G., & Keller, S. (2010). Nonlinear least-squares data fitting in Excel spreadsheets. *Nat. Protoc.*, *5*(2), 267–281. https://doi.org/10.1038/nprot.2009.182

32. Huitema, C., & Horsman, G. (2018). *Analyzing enzyme kinetic data using the powerful statistical capabilities of R* [Preprint]. BioRxiv. https://doi.org/10.1101/316588

Problems

coming soon

Supporting Information

Fitting steady-state kinetic data using Lineweaver-Burke plots (linear fit) yields less accurate estimates of V_{max} and K_M.

The text recommends finding V_{max} and K_M by fitting the Michaelis-Menten equation below, eq. S5.24, directly to the experimental data (measured rates, V, at different substrate concentrations, $[S]$). This equation describes a curve since it has the form $y = \frac{ax}{b+x}$ so fitting this equation to the data requires a non-linear fitting program. Convenient examples include Microsoft Excel,[31] R scripts,[32] or the Python script given below. Non-linear least squares fitting of eq. S5.24 to the data starts with initial guesses for V_{max} and K_M followed by iteration to find the best values.

$$V = \frac{V_{max} \cdot [S]}{K_M + [S]} \tag{S5.24}$$

Before the widespread use of computers, fitting a curve to experimental data was difficult, so an alternative was to rearrange the Michaelis-Menten equation to a form that yields a straight line, eq. S5.25.

$$\frac{1}{v} = \frac{1}{S} \cdot \frac{K_M}{V_{max}} + \frac{1}{V_{max}} \tag{S5.25}$$

Plotting 1/v on the y-axis and 1/[S] on the x-axis yields a straight line with a slope of K_M/V_{max} and y-intercept of $1/V_{max}$. This plot is known as the Lineweaver-Burke plot. Fitting a line to experimental data is mathematically simpler, so it was often used in the past. The disadvantage of fitting the Lineweaver-Burke line to the data is that the estimates of V_{max} and K_M are less accurate when the data are imperfect.

Both methods would yield the same values of K_M and V_{max} if the data were perfect. Typically, errors in measuring reaction rates are approximately constant; for example, one might measure V = 5±1 at low [S] and 50±1 at high [S]. That is, rates measured at low [S] have a higher percentage error (20% for this example) than rates measured at high [S] (2%). When fitting the data to the Michaelis-Menten curve, the deviation of each data point from the curve counts equally. This is the best approach when the data points have equal absolute errors as suggested above.

In contrast, fitting the data to eq. S5.25 minimizes the deviation of each 1/V data point from the line. In our example, the error for the value at low [S], 1/V = 0.20±0.04, is 10-fold higher than the error for the value at high [S], 0.0200±0.0004. Since deviations from either data point count equally, but the uncertainty is higher for data point at low [S], this approach yields estimates of V_{max} and K_M that weight the data points at low

[S] too heavily. These data have high errors, but are treated as having the same errors as values at high [S].

For example, fitting the Michaelis-Menten curve and the Lineweaver-Burke line to the imperfect data below yields different estimates for V_{max} and K_M. This simulated data starts with a perfect fit to the Michaelis-Menten equation and adds Gaussian random error to simulate experimental error.[31]

```
S = (0, 2.5, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100)
     # units = mM
V = (-2, 153, 231, 342, 396, 438, 467, 505, 523, 523, 539,
     548, 555, 554) # units = mM/s
```

Fitting the Michaelis-Menten curve, eq. S5.24, to the data yields $V_{max} = 598 \pm 3 mM/s$, $K_M = 7.6 \pm 0.2\ mM$, while fitting the Lineweaver-Burke line, eq. S5.25, to the same data yields $V_{max} = 591 \pm 7 mM/s$, $K_M = 7.3 \pm 0.2 mM$, Figure S5.1. The V_{max} differs by 1.2%, while K_M differs by 4.5%. The first data point (S = 0, V = -2) was omitted from the linear fit because 1/0 is undefined. The differences are within the estimated errors (standard deviations) for the values.

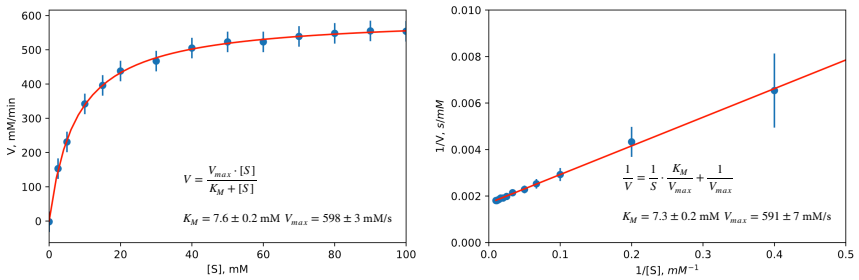

Figure S5.1. Fitting the Michaelis-Menten equation of a curve (left) to experimental data and fitting the Lineweaver-Burke line to the same data (right). The best fit yields slightly different values of K_M and V_{max}. Error bars correspond to a fixed error of ±30 ($\sim 5\%$ of V_{max}) for the values of V.

The code blocks below generate the plots shown in Figure S5.1 so that you can experiment with your own data. Another requirement for accurate estimates of K_M and V_{max} is that the data include data points at substrate concentrations both below and above K_M. Otherwise, the shape of the curve is poorly defined and the result will be inaccurate regardless of which equation is fit the data.

Code Block S5.1 Python script to fit experimental data to the Michaelis-Menten curve. This approach is the preferred method for fitting steady-state kinetic data to extract K_M and V_{max}.

```
"""
This script fits the nonlinear Michaelis-Menten curve to the
experimental data by adjusting the initial guesses for Km and
```

```
Vmax until the calculated values of V best match the experimen-
tal ones. This script is the preferred one to calculate Vmax and
Km from experimental data.
"""
# fit to Michaelis-Menten curve
import numpy as np                #import math functions to use arrays
from scipy import optimize        #import non-linear fit function
import matplotlib.pyplot as plt   #import plotting function

# list of data, assumes units of S are millimolar, units of V are millimolar/min
S = np.array([0, 2.5, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90,
100])
V = np.array([-2, 153, 231, 342, 396, 438, 467, 505, 523, 523,
539, 548, 555, 554])
# set a fixed error for all V measurements
Verr = 30
Km = 5.0 # initial guesses for Km and Vmax
Vmax = 550

# define equation for MM curve to be fit
# This function returns calculated values of V for currrent
# values of Vmax and Km. These calculated values should match
# the experimental data for V best for when Vmax and Km are
# correct.
def MM(S, Km, Vmax):
    return S*Vmax/(Km + S)

# non-linear fit adjusts Km and Vmax so the calculated values
# of V match the experimental values of V
popt, pcov = optimize.curve_fit(MM, S, V, [Km, Vmax])
perr = np.sqrt(np.diag(pcov))
print("Km =", "{0:.3f}".format(popt[0]),"±", "{0:.3f}".format
(perr[0]), "Vmax =", "{0:.3f}".format(popt[1]),"±", "{0:.3f}".format
(perr[1]))

# plot data and best fit curve
plt.scatter(S, V)
xfit = np.linspace(0,100)
plt.errorbar(S, V, yerr=Verr, linestyle = 'None')
plt.plot(xfit, MM(xfit, popt[0], popt[1]), 'r-')
plt.xlim(0, 100)
plt.xlabel('[S], mM')
plt.ylabel('V, mM/min')
plt.show()

# comment line above and uncomment lines below to save plot as
```

```
# eps file to Desktop; adjust file path as needed
# changing file extension (e.g., jpg, pdf) changes format of file

# import os
# os.chdir('/Users/romas/Desktop')
# plt.savefig('mm_plot.eps', dpi=600)
```

Code block S5.2 Python commands to fit experimental data to the Lineweaver-Burke linear form of the Michaelis-Menten equation. This approach is common in older texts, but is no longer preferred.

```
"""
This script fits the Lineweaver-Burke line (a linearized form of
the Michaelis-Menten curve) to the experimental data by adjusting
the initial guesses for Km and Vmax. This approach is less accurate
because it overweights data a low [S].
"""
# fit to LB line
import numpy as np                #import math functions to use arrays
from scipy import optimize        #import non-linear fit function
import matplotlib.pyplot as plt   #import plotting function

# list of data
S = np.array([2.5, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100])
# delete 0 data point to avoid error of 1/0
V = [153, 231, 342, 396, 438, 467, 505, 523, 523, 539, 548, 555, 554]
Verr = 30
# define as list
Vinv = np.array([1/x for x in V]) # y-axis is 1/V for this fit
# calculate error in Vinv
Vinv_err = np.array([(1/(x-Verr))-(1/x) for x in V])

Km = 5.0 # initial guesses for Km and Vmax
Vmax = 550

# define equations for LB line to be fit
def LB(S, Km, Vmax):
    return (1/S) * (Km/Vmax) + 1/Vmax
    # result should equal data for Vinv when Vmax and Km are correct

# fit the data to the LB line
popt, pcov = optimize.curve_fit(LB, S, Vinv, [Km, Vmax])
perr = np.sqrt(np.diag(pcov))
print("Km =", "{0:.3f}".format(popt[0]),"±", "{0:.3f}".format
(perr[0]), "Vmax =", "{0:.3f}".format(popt[1]),"±", "{0:.3f}".format
(perr[1]))
# calculate error in Vinv
```

```
Vinv_err = np.array([(1/(x-Verr))-(1/x) for x in V])

# plot data and best fit curve
plt.scatter(1/S, Vinv)
xfit = np.linspace(1,100)
plt.errorbar(1/S, Vinv, yerr=Vinv_err, linestyle='None')
plt.plot(1/xfit, LB(xfit, popt[0], popt[1]), 'r-')
plt.xlim(0,0.5)
plt.ylim(0, 0.01)
plt.xlabel('1/[S], $mM^{-1}$')
plt.ylabel('1/V, $s/mM$')
plt.show()

# comment line above and uncomment lines below to save plot as
# eps file adjust file path to your computer, adjust file exten-
# sion as needed
# import os
# os.chdir('/Users/romas/Desktop')
# plt.savefig('lb_plot.eps', dpi=600)
```