2

CHEMICAL KINETICS I: BASICS

[c1] In the previous chapter, we [c2]discussed some the fundamental biological processes [c3]undertaken by cells such as transcription and translation. The purpose of this chapter is to introduce the basic concepts needed to model the dynamics of such processes. [c4]These next two chapters explain how to describe chemical reactions mathematically, both at a deterministic level where stochastic effects are ignored, as well as probabilistically where stochasticity is explicitly incorporated.

2.1 Law of mass action

Consider a reaction where two kinds of molecules, [c5] A and B, irreversibly react to produce a third kind [c6]of molecule, C. Schematically, such a reaction is represented as

$$\mathrm{A} + \mathrm{B} \xrightarrow{k} \mathrm{C}. \tag{2.1}$$

The parameter k is the rate of the reaction. [c7]In general, kinetic parameters such as k depend on the environment through thermodynamic quantities such as the pressure and temperature. However, since cells often operate in environments where these quantities do not vary much, for simplicity, we will neglect these dependencies in what follows. According to the law of mass action,the rate of increase of the concentration of the product is given by

$$\frac{d[\mathrm{C}]}{dt} = k[\mathrm{A}][\mathrm{B}] \tag{2.2}$$

where we follow the standard convention in chemistry texts: the concentration of the chemical X is represented by [X]. Note that the accompanied decrease of the concentrations of A and B is given by the same expression: k[A][B]. Namely,

$$\frac{d[\mathrm{A}]}{dt} = \frac{d[\mathrm{B}]}{dt} = -k[\mathrm{A}][\mathrm{B}] \tag{2.3}$$

[c1] *Pankaj*: This is a test
[c2] *Pankaj*: ~~have introduced some of the basic~~
[c3] *Pankaj*: *Text added.*
[c4] *Pankaj*: *Text added.*
[c5] *Pankaj*: *Text added.*
[c6] *Pankaj*: *Text added.*
[c7] *Pankaj*: ~~in general, depends upon the temperature~~

The law of mass action applies to elementary reactions where molecules of type A and of type B collide, with the collision event giving rise to the product C with a probability related to the rate constant k. [c8] The basic intuition behind the law of mass action is that the probability per unit time of collision is proportional to the concentrations of A and B. The crucial [c9]assumption underlying these equations is that, for most practical purposes, the formation of C can be considered as a one step process rather than a process with multiple intermediates.

Since the reaction 2.1 is irreversible, eventually, either molecules of type A or of type B will be completely depleted and the reaction will stop. Each time this reaction takes place, we reduce the number of A and B molecules by one each. Therefore, if we start with a certain amount of A and B, [c1] the species with the smaller number of molecules will be depleted first. Let us see how [c2]this works mathematically by explicitly solving the [c3]corresponding differential equation.

Initially, at time $t = 0$, assume $[\mathrm{A}] = A_0$, $[\mathrm{B}] = B_0$ and $[\mathrm{C}] = 0$. Furthermore, without loss of generality, we assume $B_0 > A_0$. For convenience, we introduce the following notation for the time dependent concentrations of A, B and C [c4]at time t: $[\mathrm{A}] = a(t), [\mathrm{B}] = b(t), [\mathrm{C}] = c(t)$ [c5]. Notice that the equations 2.2 and 2.3 imply that the conservation laws

$$a(t) + c(t) = A_0, \tag{2.4}$$

$$b(t) + c(t) = B_0. \tag{2.5}$$

[c6]These conservation laws reflect the fact that in order to produce a single molecule of C, one has to consume a molecule of A and a molecule of B. A direct consequence of the equations above is that

$$b(t) - a(t) = B_0 - A_0 \equiv \Delta \tag{2.6}$$

is also a constant.

We can now use the conservation laws to write equation 2.2 as a differential equation constraining a single variable.

$$\frac{d}{dt}(A_0 - a(t)) = \frac{dc(t)}{dt} = ka(t)b(t) = ka(t)(\Delta + a(t)) \tag{2.7}$$

or

[c8] *Pankaj*: *Text added.*

[c9] *Pankaj*: ~~aspect of such reactions~~

[c1] *Pankaj*: ~~which will be ultimately depleted depends upon which had the smaller number of molecules in the beginning~~

[c2] *Pankaj*: ~~that works out~~

[c3] *Pankaj*: *Text added.*

[c4] *Pankaj*: *Text added.*

[c5] *Pankaj*: ~~for arbitrary time t~~

[c6] *Pankaj*: *Text added.*

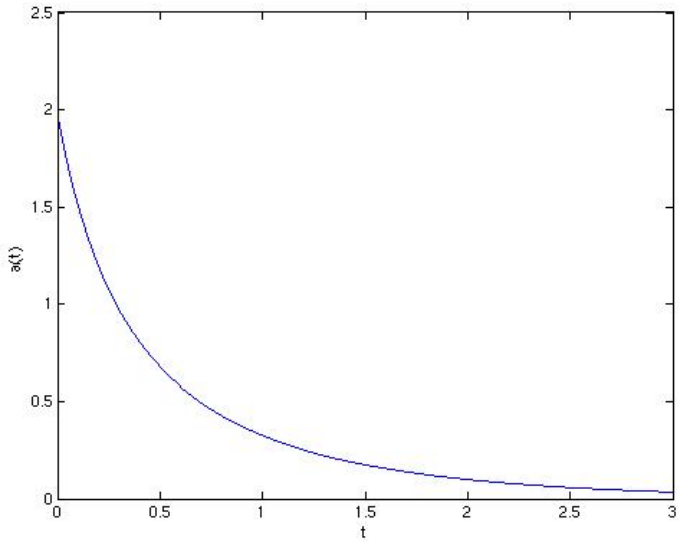

FIG. 2.1. Depletion of A with time.

$$\frac{da(t)}{dt} = -ka(t)(\Delta + a(t)) \tag{2.8}$$

We now solve this equation by the method of separation of variable:

$$\int_{A_0}^{a(t)} \frac{da}{a(\Delta + a)} = -k \int_0^t dt' \tag{2.9}$$

implying

$$\Big[\ln \frac{a}{\Delta + a} \Big]_{A_0}^{a(t)} = -\Delta k t. \tag{2.10}$$

which, after some algebraic manipulations, can be rewritten as

$$a(t) = \frac{A_0 \Delta}{B_0 e^{\Delta k t} - A_0}. \tag{2.11}$$

[c1]. Since the time dependence in eqn 2.11 comes in only through the combination Δkt, the problem has a characteristic time scale given by $(\Delta k)^{-1}$ To understand the significance of this time scale, consider the the asymptotic dependence of $a(t)$ on t. For large t, $a(t) \approx$ const. $\exp(-\Delta kt)$. Thus, $(\Delta k)^{-1}$ sets the characteristic time scale for the decay of particles due to depletion of the reactant

[c1] *Pankaj*: I significantly reworked paragraph and added clarifications

molecules. This is depicted graphically in Fig. 2.1 where we have plotted the solution to the equations above for the initial conditions $A_0 = 2, B_0 = 3$ and $k = 1$.

[c2]. **Exercise:** As an aside, one could ask, what happens when $\Delta = 0$. A realistic way to be in this situation is have a dimerization reaction, 2A→C. Show that, in that case, $a(t)$ goes to zero as $1/t$ (and not as an exponential) in the long time limit.

2.2 Reversible reactions

Now consider the reversible reaction where A and B combine to make C but C can also dissociate back into A and B.

$$\mathrm{A} + \mathrm{B} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{C}, \tag{2.12}$$

[c1]with k_+ and k_- the rate constants for the forward and backward reactions, respectively. In the case of the irreversible reaction discussed in the previous section, one of the two reactants gets completely depleted as time goes by. For a reversible reaction, an equilibrium with nonzero concentrations of all three species of chemicals is reached. [c2] In particular, eqn 2.2 must be modified to include the reverse reaction and becomes

$$\frac{d[\mathrm{C}]}{dt} = k_+[\mathrm{A}][\mathrm{B}] - k_-[\mathrm{C}] \tag{2.13}$$

[c3]By definition, at equilibrium, the concentration of C does not change with time and $d[\mathrm{C}]/dt = 0$, implying

$$k_+[\mathrm{A}][\mathrm{B}] = k_-[\mathrm{C}]. \tag{2.14}$$

Assume, once more that initially at time $t = 0$, $[\mathrm{A}] = A_0, [\mathrm{B}] = B_0$ and $[\mathrm{C}] = 0$. The conservation laws (eqns 2.4, 2.5 and 2.6) [c4]derived for the irreversible case are are valid even in the reversible case as well and imply that at equilibrium $[\mathrm{A}] + [\mathrm{C}] = A_0$. [c5]We begin by rewriting equation 2.14 as

$$\frac{[\mathrm{C}]}{[\mathrm{A}]} = \frac{[\mathrm{B}]}{K_{eq}}, \tag{2.15}$$

[c2] *Pankaj*: I think we should more explicit about exercises and mark them
[c1] *Pankaj*: *Text added.*
[c2] *Pankaj*: Changed wording
[c3] *Pankaj*: *Text added.*
[c4] *Pankaj*: *Text added.*
[c5] *Pankaj*: *Text added.*

where [c6]we have defined the equilibrium constant, $K_{eq} = k_-/k_+$. Using the equations above, it is easy to see that

$$\frac{[\mathrm{C}]}{A_0} = \frac{[\mathrm{C}]}{[\mathrm{C}]+[\mathrm{A}]} = \frac{[\mathrm{B}]}{[\mathrm{B}]+K_{eq}} \quad \text{or} \quad [\mathrm{C}] = \frac{A_0[\mathrm{B}]}{[\mathrm{B}]+K_{eq}}, \tag{2.16}$$

[c1]where in writing the first equality we have used the conservation law 2.4.

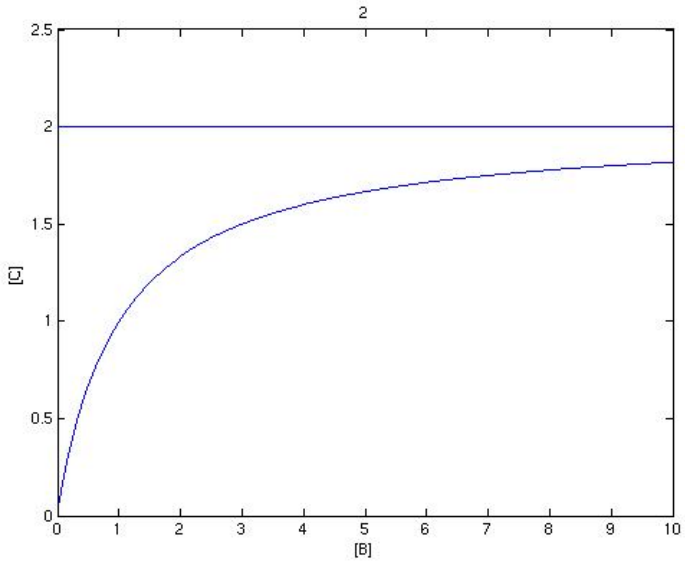

FIG. 2.2. Saturation as a function of the concentration of reactant B for a reversible reaction.

Let us take a moment to ponder what equation 2.16 is telling us. It gives us a relation between the amount of C and the amount of *free* B (as opposed to initial amount of B, B_0) at equilibrium. When $[\mathrm{B}] << K_{eq}$, we have an approximately linear relation between [C] and [B]. On the other hand, when $[\mathrm{B}] >> K_{eq}$, we have [C] approaches A_0, as expected. The crossover between the two limits is depicted in Fig. 2.2.

[c2] In practice, it is often useful to also have an expression for the equilibrium value of C as a function of the initial concentration of B_0 and not just the concentration of *free* B, [B]. That expression is a solution of a quadratic equation and a bit more complicated than the rational functions in eqn 2.16. However,

[c6] *Pankaj*: *Text added.*
[c1] *Pankaj*: *Text added.*
[c2] *Pankaj*: reworded this paragraph and added excercise

there are many situations where there is much more B than A, [B] $\gg$ [A]. In this case, the difference between [B] and B_0 is negligible and we can replace [B] with B_0.

Exercise: Derive an exact expression for the concentration of the product C as a function of B_0. Show that when [B] $\gg$ [A], replacing [B] with B_0 in equation 2.16 is a good approximation.

[c1] Now consider a more complicated reaction scheme:

$$\mathrm{A} + \mathrm{B} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{C} \overset{k}{\rightarrow} \mathrm{A} + \mathrm{D} \tag{2.17}$$

When $k = 0$, the equation reduces to the reversible reaction considered above. In this case, we know that A, B and C would reach equilibrium with the equilibrium concentration of C given by equation 2.2. Now consider the case where k is much smaller than k_+ and k_-. In this case, A,B, and C will still quickly equilibrate since the production of D from C is slow compared to the reversible reaction. Hence, a good approximation for the kinetics when k is nonzero but still small ($k \ll k_\pm$) is to model the the production rate of D as k times the equilibrium value of [C].

$$\frac{d[\mathrm{D}]}{dt} \approx k[\mathrm{C}] \tag{2.18}$$

When C is in equilibrium, we know that

$$(k_- + k)[\mathrm{C}] = k_+[A][B]. \tag{2.19}$$

This equation is identical to equation 2.14 except for the replacement $k_- \rightarrow k_- + k$. Hence, we know that the equilibrium value of [C] is given by Equation 2.16 except now $K_{eq} = (k_- + k)/k_+$. Thus, we can approximate the kinetics as

$$\frac{d[\mathrm{D}]}{dt} \approx k[\mathrm{C}] = \frac{kA_0[\mathrm{B}]}{[\mathrm{B}] + K_{eq}}. \tag{2.20}$$

The approximation employed above is often referred to as the *quasi-equilibrium* approximation and will be discussed more below.

It is worth comparing equation 2.20 with equation 2.2 for the mass action kinetics of an irreversible process with *no* intermediate steps. Notice that where as the production rate of the product in equation 2.2 is linear in the concentration [B], the production rate of the product in equation 2.20 saturates as a function of [B] at large concentrations. This saturation effect arises because the production of D is a multistep process and lies at the heart of Michaelis Menten kinetics discussed below. [c2] When $[B] \gg [A]$, all the A molecules are quickly bound by the excess B molecules and the kinetics is limited only by the concentration of $[A]$.

[c1] *Pankaj*: Completely reworded the rest of the section and added more clarifications, in particular clarified what is meant by K_eq

[c2] *Pankaj*: *Text added.*

2.3 Michaelis Menten kinetics

The reaction scheme 2.17 is nothing other than a generic catalytic reaction where B is converted into D by the catalyst A. A special class of catalytic reactions of great importance in biology are enzymes acting on their substrates. Written in more conventional notation, eqn 2.17 becomes

$$\mathrm{E} + \mathrm{S} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{ES} \overset{k_{\mathrm{cat}}}{\rightarrow} \mathrm{E} + \mathrm{P}. \tag{2.21}$$

The enzyme E captures the substrate S and makes a complex ES reversibly. Occasionally the complex gives rise to the product P, a modification of the substrate, and releases the enzyme for further action. The rate of product formation is often expressed as

$$\frac{d[\mathrm{P}]}{dt} = \frac{k_{\mathrm{cat}}[\mathrm{E}]_{\mathrm{total}}[\mathrm{S}]}{[\mathrm{S}] + K_M}, \tag{2.22}$$

[c1] with $K_M = (k_- + k_{\mathrm{cat}})/k$.

In the conventional derivation of equation 2.22 from the reaction 2.21, one makes some implicit assumptions:

- For calculating the rate of product formation, often called the initial velocity, the concentration of product can be neglected compared to that of the substrate (i.e $[\mathrm{P}] << [\mathrm{S}]$).
- The amount of the intermediate ES is approximately at equilibrium (the quasi-equilibrium condition).

The last assumption says that the loss of ES [c2] due to disassociation is balanced by the formation of new complexes..

$$(k_- + k_{\mathrm{cat}})[\mathrm{ES}] \approx k_+[\mathrm{E}][\mathrm{S}]. \tag{2.23}$$

This equation in conjunction with the exact conservation condition

$$[\mathrm{E}] + [\mathrm{ES}] = [\mathrm{E}]_{\mathrm{total}} \tag{2.24}$$

immediately imply that

$$\frac{[\mathrm{ES}]}{[\mathrm{E}]_{\mathrm{total}}} = \frac{1}{1 + \frac{k_- + k_{\mathrm{cat}}}{k_+[\mathrm{S}]}} \tag{2.25}$$

Since $d[\mathrm{P}]/dt = k_{\mathrm{cat}}[\mathrm{ES}]$, Michaelis Menten equation follows, once we identify the Michaelis constant K_M to be $(k_- + k_{\mathrm{cat}})/k_+$. In practical applications, it is often assumed, in addition to conditions mentioned above, that the concentration

[c1] *Pankaj: Text added.*
[c2] *Pankaj:* ~~compensates the gain~~

of substrate is in large excess over that of the enzyme (i.e. [E] << [S]), allowing us to ignore the difference between the total amount of substrate molecules and the amount that is free.

[c3] **Exercise:** This exercise examines the validity of the quasi-equilibrium condition.
a) Derive a differential equation for $\frac{d[S]}{dt}$ as a function of [S] within the quasi-equilibrium approximation.
b) Explicitly solve the differential equation for $[S]$ as a function of time for the initial conditions [P] = [ES] = 0 and [S] = S_0 at $t = 0$.
c) Show that this solution does not satisfy the initial condition [ES] = 0 at $t = 0$.
d) Discuss what goes wrong with the quasi-equilibrium assumption. For a detailed explanation see Chapter 6 of **Murray**.

Much ink has been spent on the conditions under which this standard derivation holds, or on how to derive it better. Instead of delving into these details, we will think of this formula as a phenomenological description, capturing the intuitive expectations in the limits of small or of large amount of substrates. For a small amount of substrates, the bottleneck is the enzyme and the substrate finding each other. Hence the rate is proportional to the product of $[\mathrm{E}]_{\mathrm{total}}$ and [S]. In the other limit, there is so much substrate that almost all the enzymes are in the complex ES. Hence the rate is just $k_{\mathrm{cat}}[\mathrm{E}]_{\mathrm{total}}$. The Michelis constant, K_M can be thought of, operationally, as the substrate concentration where the rate is half of the maximum value.

For a fixed amount of enzyme, $[\mathrm{E}]_{\mathrm{total}}$, the velocity, $v([\mathrm{S}]) = d[\mathrm{P}]/dt$, satisfies

$$\frac{1}{v([\mathrm{S}])} = \frac{1}{v_{\mathrm{max}}} + \frac{K_M}{v_{\mathrm{max}}}\frac{1}{[\mathrm{S}]}. \tag{2.26}$$

with $v_{\mathrm{max}} = k_{\mathrm{cat}}[\mathrm{E}]_{\mathrm{total}}$. Thus, the plotting inverse of velocity against inverse of substrate concentration is expected to produce a straight line. This plot, known as Lineweaver-Burk plot, is often used in enzyme kinetics. The slope and the intercept of the straight line fit to the data can be used to extract parameters like v_{max} and K_M.

The tables 2.1 and 2.2 show that the values for the parameters K_M and k_{cat} vary widely from molecule to molecule. Note that when the concentration of the substrate is low compared to K_M, the product formation rate is [c1]

$$\frac{k_{\mathrm{cat}}}{K_M}[\mathrm{E}]_{\mathrm{total}}[\mathrm{S}] = k_+[\mathrm{E}]_{\mathrm{total}}[\mathrm{S}] \times \left(\frac{k_{\mathrm{cat}}}{k_{\mathrm{cat}} + k_-}\right). \tag{2.27}$$

[c3] *Pankaj*: added excercise
[c1] *Pankaj*: reworked paragraph below to separate efficiency from collisions more explicitly

Enzyme	Substrate	$K_M(\mu\text{M})$
Chymotrypsin	Acetyl-L-tryptophanamide	5000
Lysozyme	Hexa-N-acetylglucosamine	6
β-Galactosidase	Lactose	4000
Threonine deaminase	Threonine	5000
Carbonic anhydrase	CO_2	8000
Penicillinase	Benzylpenicillin	50
Pyruvate carboxylase	Pyruvate	400S
	HCO_3^-	1000
	ATP	60
Arginine-tRNA synthetase	Arginine	3
	tRNA	0.4
	ATP	300

Table 2.1 *K_M values for some enzymes and substrates (based on* **Biochemistry by Berg, Tymoczko and Stryer ******)

[c2]This equation has a simple interpretation as the rate of collisions between enzyme and substrate, $k_+[\text{E}]_{\text{total}}[\text{S}]$, times the probability that a collision gives rise to a product. Thus, we see that the rate of the reaction is bounded above by the rate at which the substrate collides with the reactive pocket of the enzyme. The chance of collision, on the other hand, is limited by the diffusion rate (see exercise below). The highest observed values (k_{cat}/K_M) turns out to be in the range $10^8 - 10^9 \text{s}^{-1}\text{M}^{-1}$. [c3]These numbers are the same order of magnitude expected from diffusion limited kinetics. These enzymes are believed to have achieved kinetic perfection, in the sense that every encounter with a substrate is highly likely to lead to the product. We will return to diffusion-limited rates in the next chapter (also see exercise below).

[c1] **Excercise:** This problem explores diffusion limited fluxes. Consider a spherical cell of radius a immersed in a medium that contains molecules of a species X in a law concentration with diffusion constant D. Furthermore, assume that the cell is a perfect sink. Show that the steady-state current of molecules into the cell is given by

$$J = 4\pi a D c_\infty, \tag{2.28}$$

where c_∞ is the concentration far from the cell assumed to be maintained at steady state. (Hint: Think about the analogy between the time-independent diffusion equation and Laplaces equation and use Gauss's law.) Use the expression above to estimate the diffusion limited value of k_{cat}/K_M for a small molecule substrate hitting the reactive pocket of an enzyme.

[c2] *Pankaj: Text added.*
[c3] *Pankaj: Text added.*
[c1] *Pankaj:* added excercise

Enzyme	$k_{\rm cat}(\rm s^{-1})$
Carbonic anhydrase	600,000
3-Ketosteroid isomerase	280,000
Acetylcholinesterase	25,000
Penicillinase	2,000
Lactate dehydrogenase	1,000
Chymotrypsin	100
DNA polymerase I	15
Tryptophan synthetase	2
Lysozyme	0.5

Table 2.2 *k_{cat} values for some enzymes (based on* **Biochemistry by Berg, Tymoczko and Stryer ******)

2.4 Cooperativity

In Michaelis Menten equation, the velocity rises approximately linearly with substrate abundance, till it reaches K_M. For some reactions, the dependence of the rate on the substrate is strongly sigmoidal. A well known example, displaying such cooperativity, is th O_2 binding of hemoglobin. [c2]. Cooperativity arises in proteins with multiple binding sites where the binding of ligand to binding site increases the affinity of the remaining binding sites for the substrate (see Exercise below on the MWC model). Regardless of the details of how cooperativity arises, the resulting sigmoidal behavior is often phenomenologically described using a Hill Function. We now give a brief derivation of the Hill equation and discuss how its used in chemical kinetics.

Consider a reaction scheme where the protein has to bind multiple substrate molecules before being productive.

$$\mathrm{E} \underset{k_{1-}}{\overset{k_{1+}[\mathrm{S}]}{\rightleftharpoons}} \mathrm{ES} \underset{k_{2-}}{\overset{k_{2+}[\mathrm{S}]}{\rightleftharpoons}} \mathrm{ES}_2 \underset{k_{3-}}{\overset{k_{3+}[\mathrm{S}]}{\rightleftharpoons}} \cdots \underset{k_{p-}}{\overset{k_{p+}[\mathrm{S}]}{\rightleftharpoons}} \mathrm{ES}_p \overset{k_{\rm cat}}{\rightarrow} \mathrm{E} + \mathrm{P} + \text{other things}. \quad (2.29)$$

To solve of this problem in the quasi-equilibrium approach, we find that

$$\frac{[\mathrm{ES}_l]}{[\mathrm{ES}_{l-1}]} = \frac{k_{l+}[\mathrm{S}]}{k_{l-}} \Rightarrow [\mathrm{ES}_l] = [\mathrm{E}]([\mathrm{S}]/K_l)^l \quad (2.30)$$

where,

$$K_l = \left(\prod_{i=1}^{l} \frac{k_{i-}}{k_{i+}} \right)^{1/l}. \quad (2.31)$$

Since the number of enzymes $[\mathrm{E}]_{\rm total}$ is conserved, this yields

[c2] *Pankaj*: changed this around

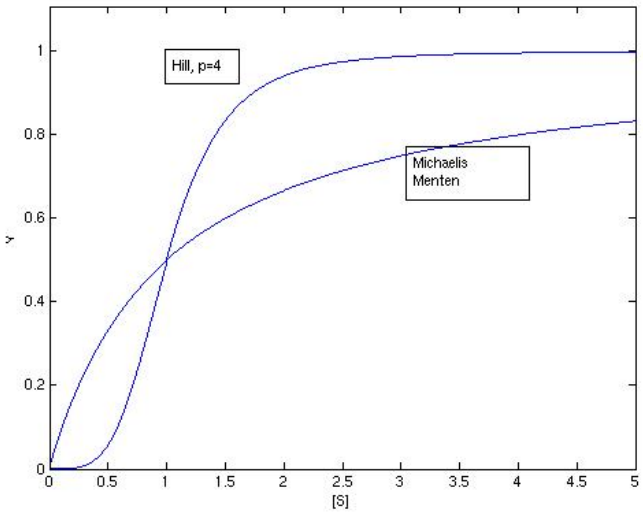

FIG. 2.3. Velocity as a function of substrate concentration [S]: Hill equation compared to Michaelis Menten equation

$$\frac{[\mathrm{ES}_p]}{[\mathrm{E}]_{\mathrm{total}}} = \frac{[\mathrm{ES}_p]}{[\mathrm{E}] + [\mathrm{ES}] + [\mathrm{ES}_2] + \cdots + [\mathrm{ES}_{p-1}] + [\mathrm{ES}_p]}$$
$$= \frac{\big([\mathrm{S}]/K_p\big)^p}{1 + ([\mathrm{S}]/K_1) + ([\mathrm{S}]/K_2)^2 + \cdots + ([\mathrm{S}]/K_{p-1})^{p-1} + ([\mathrm{S}]/K_p)^p} \tag{2.32}$$

In the extreme case of cooperativity, binding of single substrate is unlikely, but once achieved, helps binding of additional substrates, which, in turn, makes further substrate binding easier. Thus, the protein spends most of its time in one of two states: E or ES_p. In this case, the denominator in equation 2.32 is dominated by the constant term and the term going as $[\mathrm{S}]^p$ and the velocity of product formation is given by

$$v = \frac{d[\mathrm{P}]}{dt} = k_{\mathrm{cat}}[\mathrm{ES}_p] = \frac{v_{\max}\big([\mathrm{S}]/K_p\big)^p}{1 + ([\mathrm{S}]/K_p)^p}, \tag{2.33}$$

with $v_{\max} = k_{\mathrm{cat}}[\mathrm{E}]_{\mathrm{total}}$. Equation 2.33 is called the Hill equation and p the Hill coefficient. K_p, like Michaelis constant, can be defined as the substrate concentration corresponding to the rate that is half of the maximum possible. Figure 2.3 shows how having a Hill coefficient of 4 gives a much more sigmoidal response when compared to the Michaelis Menten form.

In practice, the Hill coefficient is used as an extra parameter that is used provide better phenomenological description of a reaction rate. Therefore you

should not be surprised to see, say, fractional Hill coefficients. These coefficients are often detemined from the slope of the Hill plot: $\ln[v/(v_{\max}-v)]$ versus $\ln[\mathrm{S}]$. If the system is described by Hill equation, then

$$\ln \frac{v}{v_{\max}-v} = \ln(\frac{[\mathrm{S}]}{K_p})^p = p\ln[\mathrm{S}] - p\ln K_p. \tag{2.34}$$

Unlike the Lineweaver-Burk plot, Hill plot requires one to estimate $v_{\max}$, making fitting data a slightly more involved exercise.

Cooperative effects in chemical reactions, like those described above, have important consequence for the dynamics of the whole network. As we will see later, biomolecular networks with strong cooperative effects can, sometimes, show switch like behavior, a feature biological systems use to accomplish certain goals. The Hill equation will play an important role in description of such systems.

[c1] **Exercise:** This exercise introduces the Monod-Wyman-Changeux (MWC) model of allosteric interactions. The MWC model was first proposed to explain the sigmoidal response of hemoglobin to oxygen and has since become one of the canonical models allostery in biochemistry and biophysics. The main idea of the model is that an enzyme or protein can exist in multiple, interconvertible conformations with the probability that the enzyme is in a given confirmation determined by thermal equilibrium. The presence of ligands biases the enzyme towards one of these confirmations by shifting the relative free energies of the underlying protein confirmations.

In this exercise, we will derive the main results of the MWC model from simple thermodynamic and statistical mechanical arguments. This problem assumes knowledge of partition functions.

a) Consider an protein with a single conformational state that can bind a ligand [L] from the environment. In thermal equilibrium, show that the free energy difference, ΔF, between the bound and unbound state is given by

$$\Delta F = -\log \frac{[\mathrm{L}]}{K_D}, \tag{2.35}$$

with $K_D = k_-/k_+$, k_+ the ligand binding rate, and k_- is the ligand unbinding rate. K_D is called the binding affinity of the protein

b) Now consider a protein that can exist in two states, an active state A, and an inactive state I. In the absence of ligand, the free energy of the active state is ϵ_A and the inactive state is ϵ_I. Furthermore, denote the binding affinity of the protein in the active state by K_D^A and the binding affinity in the inactive state K_D^I. Calculate the probability that the protein is in the active state. Show that in the limit where ligand binding strongly favors the active state $K_D^I \gg [\mathrm{L}] \gg K_D^A$, this expression reduces to a form similar to the Michaelis Menten equation. Briefly

[c1] *Pankaj*: Added a problem on MWC model because I think every biophysicist should know this model

discuss the meaning of K_M and the relationship to the Michaelis Metin equation.
c) Generalize the calculation in **b)** to the case when the protein is composed on 2 independent, identical subunits each of which can bind ligand. For this case, there are 8 total possible states: the protein can be active or inactive with 0,1, or 2 ligands bound to the protein. Show that when $K_D^I \gg [\mathrm{L}] \gg K_D^A$, your expression reduces to a form similar to the Hill equation with a Hill coefficient of 2. Discuss the relationship of the MWC model to the derivation in the main text.

d) Repeat the calculation in **c)** for the case a protein is composed of p independent subunits each of which can bind a ligand molecule.
For an interesting application of the MWC model to describe quantitative data on bacterial chemotaxis, see the papers by **Wingreen**.

[c2] **Exercise:** In this exercise, we will explore the concept of kinetic proof reading. Kinetic proofreading is an error-correction mechanism introduced by John Hopfield to understand the high fidelity of translation. The original paper **Hopfield 1974** is considered a classic and deserves a close reading.

Kinetic proofreading allows enzyme to discriminate between two substrates with a small free energy differences with higher specificity than would be expected by simple thermodynamic arguments. The basic idea behind kinetic proofreading is to introduce extra "irreversible" steps leading to the formation of the product. Since at each step the true substrate is much less likely to disassociate from the complex than the wrong substrate, the addition of extra intermediate steps leads to increased specificity. In particular, for each extra step, the specificity can be increased by a factor proportional to the ratio of the disassociation constants of the two substrates.

To see how this works, consider an enzyme E that can bind two substrates A and B found in equal concentrations with kinetic constants $k_{\pm}^{\mathrm{A}}$ and $k_{\pm}^{\mathrm{B}}$, with A the "correct" substrate and B, the "incorrect" substrate:

$$\mathrm{E} + \mathrm{A} \underset{k_-^A}{\overset{k_+^A}{\rightleftharpoons}} \mathrm{EA} \xrightarrow{k} \mathrm{AP} + \mathrm{E} \tag{2.36}$$

$$\mathrm{E} + \mathrm{B} \underset{k_-^B}{\overset{k_+^B}{\rightleftharpoons}} \mathrm{EB} \xrightarrow{k} \mathrm{BP} + \mathrm{E}. \tag{2.37}$$

In general, since the forward rates are often diffusion limited, it is reasonable to assume $k_+^{\mathrm{A}} \approx k_+^{\mathrm{B}}$ and the difference in specificity between the two substrates comes from a difference in disassociation constants, $k_-^{\mathrm{B}} > k_-^{\mathrm{A}}$. This gives a lower-bound on the error rate.
a) Use the quasi-equilibrium approximation to calculate the error rate, F. In

[c2] *Pankaj*: Added a problem on Kinetic proofreading because it is a good application and really important idea

particular, show that $F = F_0 = K_A/K_B$ where K_A and K_B are the equilibrium constants for the two reactions.
b) Now consider a reaction scheme where one forms an irreversible intermediate. In practice, this is often done by explicitly consuming energy through phosphorylating the intermediate. The reaction schemes now take the form

$$\begin{array}{c} \overset{k_+^A}{\underset{k_-^A}{\rightleftharpoons}} \\ \mathrm{E} + \mathrm{A} \end{array} \begin{array}{c} \mathrm{EB} \\ \end{array} \begin{array}{c} \overset{m}{\rightarrow} \\ \end{array} \begin{array}{c} \mathrm{EA}^* \\ \downarrow k_-^A \\ E + A \end{array} \begin{array}{c} \overset{k}{\rightarrow} \\ \end{array} \mathrm{AP} + \mathrm{E}, \tag{2.38}$$

with an analogous scheme for B. Again, assuming a quasi-equilibrium approximation for the intermediates, show that the error rate is now given by $F = F_0^2$. What is the error rate when the intermediate is formed by a reversible reaction?
c) How does the answer generalize for the case on m high-energy intermediates?