Law of mass action for reactions

Salvatore Assenza

November 18, 2021

1 Introduction

Here we dedicate some words to the world of reactions, which in biophysics play a central role due to the ubiquity of enzymatic reactions in the life of a cell. At a general level, chemical reactions take place with certain rates whenever reacting particles meet each other. This means that proper modelling of these processes should take into account both the diffusion of the reactants and their reactions. A possible approach is to consider an extension of the diffusion equation to a *reaction-diffusion equation*, but it goes outside the scope of the present document. More commonly, one tries to identify the most relevant mechanism and considers either diffusion- or reaction-limited reactions, depending on which is the slowest contribution to the process. Below, we make some physical considerations on both cases.

2 Binding of ligand and receptor

2.1 Thermodynamic equilibrium

Let us consider a system where a ligand L can bind to a receptor R, forming a complex C:

$$\mathrm{L} + \mathrm{R} \rightleftharpoons \mathrm{C} \tag{1}$$

Before considering the kinetics, we start by looking at the thermodynamic equilibrium. We set the zero of the interaction energy as corresponding to the unbound case, while upon binding the energy is $-U$ with $U > 0$. We further denote by V the total volume of the box, while a bound ligand is confined within a small volume v_0 around the receptor, which can be estimated approximately as the size of the receptor itself.

In order to form a ligand-receptor complex, there is a competition between the energy gained upon binding and the loss in translational entropy due to the reduced volume v_0 in which the ligand is confined when forming the complex. The probability p_o that a receptor is occupied can thus be written as

$$p_o = \frac{v_0 e^{\frac{U}{k_B T}}}{V + v_0 e^{\frac{U}{k_B T}}} \ , \tag{2}$$

while the probability of being unoccupied is simply $p_u = 1 - p_o$. Denoting by N_L the number of ligands and defining a reference concentration $c_0 \equiv N_L/v_0$, we note that $v_0/V = c_0/c$, where c is the total concentration of ligands that were put in the solution. With this definition, the previous formula can be written as

$$p_o = \frac{\frac{c}{c_0} e^{\frac{U}{k_B T}}}{1 + \frac{c}{c_0} e^{\frac{U}{k_B T}}} \ . \tag{3}$$

A typical way to describe quantitatively the reaction (1) is by means of the *dissociation constant* K_d

$$K_d \equiv \frac{[L][R]}{[C]} \ , \tag{4}$$

where we denoted by $[\cdots]$ the concentration of each species. We note that, by construction K_d has the units of a concentration. Alternatively, one can also define the *association constant* $K_a \equiv 1/K_d$, which has the units of an inverse concentration. By denoting as c and c_R the concentrations of ligands and receptors that were put initially in the box, conservation of particles implies that $c = [L] + [C]$ and $c_R = [R] + [C]$. Assuming thermodynamic equilibrium, we can write

$$\frac{[R]}{[C]} = \frac{p_u}{p_o} = \frac{V}{v_0} e^{-\frac{U}{k_B T}} \ . \tag{5}$$

Assuming that $[L] \gg [C]$, i.e. that most of the ligands are unbound (which can be obtained by considering an abundant amount of ligands, $c \gg c_R$), we have $[L] \simeq c$, so that

$$K_d = c\frac{V}{v_0}e^{-\frac{U}{k_B T}} = c_0\, e^{-\frac{U}{k_B T}} \ . \tag{6}$$

Therefore, apart from a normalizing constant given by the reference concentration, the dissociation constant decreases exponentially with the binding energy. We also note that the

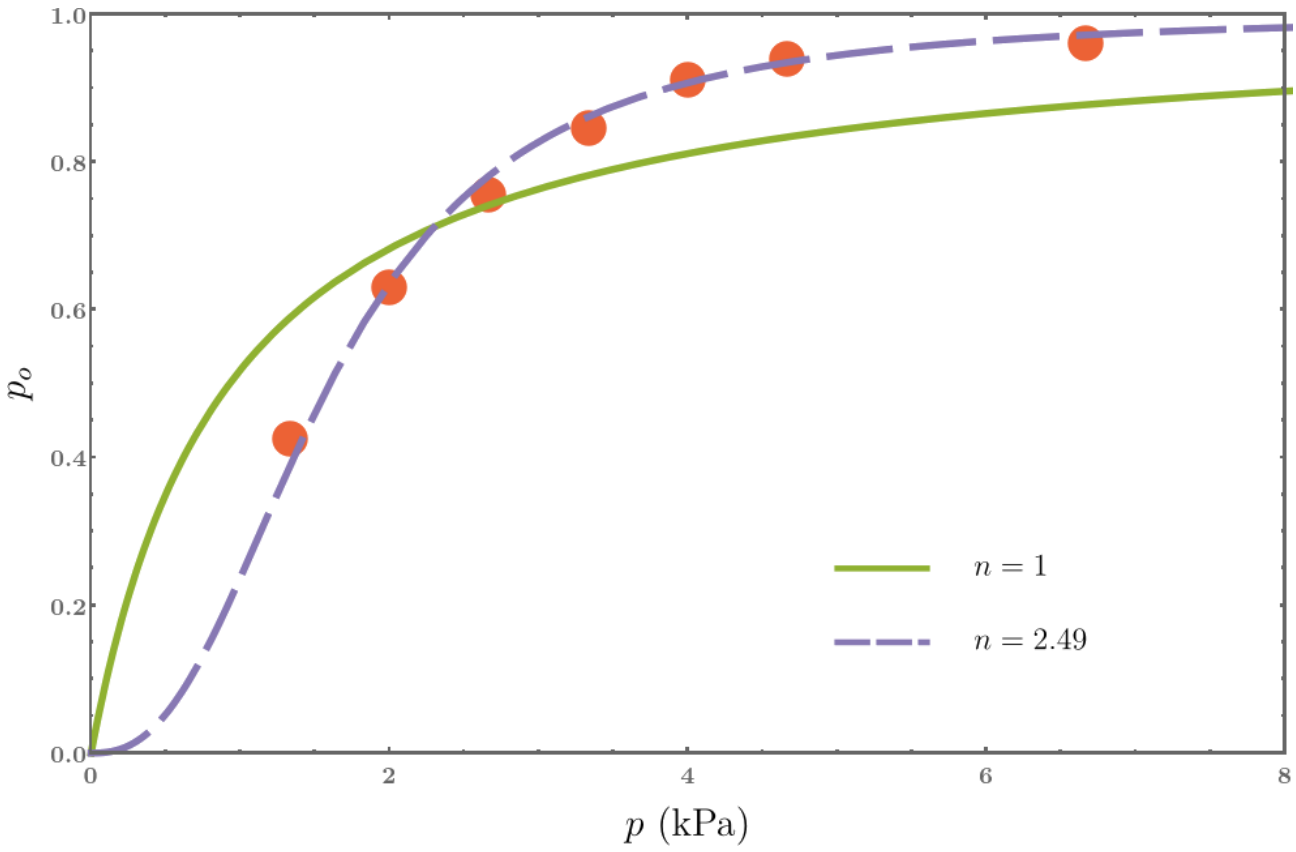

Figure 1: Experimental measurements of occupancy versus pressure for oxygen in haemoglobin (orange circles) cannot be accounted for by a model with one-to-one binding between ligands and receptors (green continuous curve, $n = 1$), but can be fitted by considering a cooperative model (purple dashed line, Hill model with $n = 2.49$). The data were extracted from the wikipedia page https://en.wikipedia.org/wiki/Hill_equation_(biochemistry) and correspond to the original work by Hill, *J. Physiol.* **40**:i-vii (1910). Note that pressure p is used in lieu of the concentration since they are proportional and p is the actual control parameter for the experiment by Hill.

probability that a receptor is occupied can be written as

$$p_o = \frac{c}{c + K_d} \ . \tag{7}$$

This enables giving a *physical interpretation of the dissociation constant as the ligands concentration at which half of the receptors are occupied*, that is for $c = K_d$ one finds $p_o = 1/2$.

In many real cases, this model turns out to be too simplistic. For instance, in Fig.1 we report the occupancy data of oxygen in haemoglobin as a function of oxygen pressure (which is proportional to its concentration). A fit with Eq.(7) does not satisfactorily fit the data (green continuous line). From a structural point of view, a single haemoglobin can bind up to four oxygens. Moreover, binding with a single oxygen causes some internal structural changes which enhance the binding probability of other oxygens, i.e. the binding is *cooperative*. As a crude model of the possibility of multiple binding, one can assume that on average there are n oxygens binding at the same time, which gives the reaction

$$n\cdot\mathrm{L} + \mathrm{R} \rightleftharpoons \mathrm{C} \tag{8}$$

The dissociation constant is generalized as

$$K_d = \left(\frac{[L]^n[R]}{[C]}\right)^{\frac{1}{n}} = c\left(\frac{[R]}{[C]}\right)^{\frac{1}{n}} \ , \tag{9}$$

so that it still has the dimensions of a concentration[1]. Following the same approach as above, the probability of occupancy can be written as

$$p_o = \frac{v_0^n e^{\frac{U}{k_B T}}}{V^n + v_0^n e^{\frac{U_n}{k_B T}}} \ , \tag{10}$$

and $p_u = 1 - p_o$, where U_n is the overall binding energy of the n bound ligands. Therefore, we can write

$$\frac{[R]}{[C]} = \frac{p_u}{p_o} = \frac{V^n}{v_0^n} e^{-\frac{U_n}{k_B T}} \ . \tag{11}$$

However, from the definition of the dissociation constant we also have

$$\frac{[R]}{[C]} = \frac{K_d^n}{c^n} \Rightarrow \frac{V^n}{v_0^n} e^{-\frac{U_n}{k_B T}} = \frac{K_d^n}{c^n} \ , \tag{12}$$

[1]Depending on the source, an alternative definition might be employed, according to which

$$K_d = \frac{[L]^n[R]}{[C]} \ .$$

In this case, K_d has the units of a concentration only when $n = 1$.

hence

$$p_o = \frac{c^n}{c^n + K_d^n} \, . \tag{13}$$

Equation (13) is called *Hill's equation* and n is the *Hill exponent*. Note that the result derived above corresponds to $n = 1$, as it should. Also in this case, the dissociation constant corresponds to the concentration at which half of the receptors are occupied. As for the haemoglobin data, we see that a satisfactory fit can be obtained by setting $n = 2.49$ (dashed purple line in Fig.1). This suggests that cooperativity results on average in two and a half oxygens bound to a single protein. This is compatible with the upper limit imposed by the four sites identified from a structural study.

2.2 Kinetics

From a kinetic point of view, at a general level for any reaction to happen the particles need first to diffusively meet each other. Hence, there are two physical mechanisms at play: diffusion and reaction. If diffusion is fast in comparison to the reaction rates (*reaction-limited processes*), one can consider the *well-mixing* approximation, according to which the concentration of ligands and receptors is instantaneously equilibrated. Based on this, for reaction-limited processes one can estimate the time evolution of each concentration by simple proportionality rules, which go under the name of *law of mass action*. In the simple case considered in reaction (1), one considers a *forward rate* k_f and a *backward rate* k_b, and writes the reaction as

$$\mathrm{L} + \mathrm{R} \underset{k_\mathrm{b}}{\overset{k_\mathrm{f}}{\rightleftharpoons}} \mathrm{C} \, . \tag{14}$$

The law of mass action is implemented by writing the following differential equations:

$$\begin{cases} \dfrac{d[R]}{dt} = \dfrac{d[L]}{dt} = -k_f[L][R] + k_b[C] \, , \\ \dfrac{d[C]}{dt} = k_f[L][R] - k_b[C] \, . \end{cases} \tag{15}$$

The logic is that the number of disassembled complexes per unit time is proportional to their amount, while the probability that a ligand and a receptor meet each other is given by the product of the respective concentrations. Note that k_f has the units of inverse concentration and inverse time, while k_b has the units of an inverse time.

Following the same approach as in the previous section, we assume a large amount of ligands, so that $[L] \simeq c$. As for the receptors, we assume that initially none of them is bound $[R](0) = c_R$. Moreover, particle conservation imposes that $[C] = c_R - [R]$. This results in the equation

$$\frac{d[R]}{dt} = -(k_f c + k_b)[R] + k_b c_R \, . \tag{16}$$

This equation can be solved by observing that, if only the first term was present in the right-hand side, the solution would be an exponential. Hence, we define an auxiliary function $u(t)$ such that

$$[R](t) = e^{-(k_f c + k_b)t} \left[c_R + u(t) \right] \ . \tag{17}$$

Note that by construction $u(0) = 0$. Deriving with respect to time and making use of the original differential equation for $[R]$, we obtain that

$$\frac{d[u]}{dt} = k_b c_R e^{(k_f c + k_b)t} \Rightarrow u(t) = \frac{k_b \, c_R}{k_f c + k_b} \left[e^{(k_f c + k_b)t} - 1 \right] \ . \tag{18}$$

Therefore, we finally find

$$[R](t) = c_R e^{-(k_f c + k_b)t} + c_R \frac{k_b}{k_f c + k_b} \left[1 - e^{-(k_f c + k_b)t} \right] \ . \tag{19}$$

The time evolution of the complex is easily found as

$$[C](t) = c_R - [R](t) = c_R \frac{k_f c}{k_f c + k_b} \left[1 - e^{-(k_f c + k_b)t} \right] \ . \tag{20}$$

At long times, we find

$$[R](\infty) = \frac{k_b}{k_f c + k_b} c_R \tag{21}$$

and

$$[C](\infty) = \frac{k_f c}{k_f c + k_b} c_R \ . \tag{22}$$

Therefore, starting from zero the concentration of the complex reaches exponentially its saturation value with a time constant τ which corresponds to the sum of both rates (suitably normalized): $\tau = 1/(k_f c + k_b)$. Since at long times one expects thermodynamics equilibrium, from the definition of the dissociation constant we obtain

$$K_d = \frac{[L][R]}{[C]} = c \cdot \frac{k_b}{k_f c} = \frac{k_b}{k_f} \ . \tag{23}$$

Note that this results could have been easily obtained from Equation (15) by observing that at long times $d[R]/dt = d[C]/dt = 0$.

3 Michaelis-Menten kinetics for Enzymatic reactions

In enzymatic reactions, a specialized protein (an enzyme) acts as a catalyzer for a given chemical reaction. The presence of enzymes is paramount for the cell, since it enables controlling the amount of various chemical groups and the rate of their production. The evolution has shaped these reactions to be virtually impossible in the absence of the enzymes, since the free-energy barrier is too large. In the simplest case, an enzyme E catalyzes a chemical reaction that generates a product P starting from a substrate S. The reaction scheme is thus

$$\mathrm{E} + \mathrm{S} \underset{k_\mathrm{b}}{\overset{k_\mathrm{f}}{\rightleftharpoons}} \mathrm{C} \xrightarrow{k_\mathrm{cat}} \mathrm{E} + \mathrm{P} \ , \tag{24}$$

where C is the complex formed by the enzyme and the substrate. By the law of mass action, we thus have the following set of differential equations:

$$\begin{cases} \dfrac{d[E]}{dt} = -k_f[E][S] + (k_b + k_{\mathrm{cat}})[C] \ , \\ \dfrac{d[S]}{dt} = -k_f[E][S] + k_b[C] \ , \\ \dfrac{d[C]}{dt} = k_f[E][S] - (k_b + k_{\mathrm{cat}})[C] = -\dfrac{d[E]}{dt} \ , \\ v \equiv \dfrac{d[P]}{dt} = k_{\mathrm{cat}}[C] \ . \end{cases} \tag{25}$$

In a typical experimental assay, one can control the total concentration of enzyme put in the solution $[E]_0$ and the total amount of substrate initially put in the solution $[S]_0$, while the experimental observable is the product concentration $[P]$, which is monitored in time. Note that particle conservation implies that $[E] + [C] = [E]_0$, which is preserved by the condition $d[E]/dt = -d[C]/dt$.

Solving the differential equations is quite cumbersome. Nevertheless, what one really needs in order to interpret the experiments is the evolution of the product concentration $[P]$. To this aim, two different approximations are typically considered, which we now proceed to study.

3.1 Equilibrium approximation

In the equilibrium approximation, one assumes that product formation is very slow compared with the rest of reactions. In reaction (24), this means that one can assume that $\mathrm{E} + \mathrm{S} \rightleftharpoons \mathrm{C}$ is always at equilibrium, i.e. that

$$k_f[E][S] = k_b[C] \tag{26}$$

holds. This approximation is valid as long as $k_{\mathrm{cat}} \ll k_b$. Under this approximation, we can apply Equation (7) by considering S as the ligand and E as the receptor. Hence, observing that in the present case the occupancy probability is $p_o = [C]/[E]_0$, we find

$$\frac{[C]}{[E]_0} = \frac{[S]}{K_d + [S]} \Rightarrow [C] = \frac{[E]_0[S]}{K_d + [S]} \ . \tag{27}$$

The velocity of product generation is thus

$$v = \frac{d[P]}{dt} = k_{\text{cat}}[C] = V_{\max}\frac{[S]}{K_d + [S]} \ , \tag{28}$$

where $V_{\max} = [E]_0 k_{\text{cat}}$ is the maximum speed of product formation.

3.2 Quasi-steady-state approximation

In the quasi-steady-state approximation, one assumes that the amount of complexes does not change in time, i.e. that the amount of complexes lost in product generation is compensated by complex generation from E and S. Mathematically, this is implemented by imposing that $d[C]/dt = 0$, i.e.

$$(k_b + k_{\text{cat}})[C] = k_f[E][S] \ . \tag{29}$$

Remembering that $[E] = [E]_0 - [C]$, the previous condition enables computing $[C]$:

$$(k_b + k_{\text{cat}})[C] = k_f([E]_0 - [C])[S] \Rightarrow [C] = \frac{[E]_0[S]}{K_M + [S]} \ , \tag{30}$$

where

$$K_M \equiv \frac{k_b + k_{\text{cat}}}{k_f} \tag{31}$$

is called the *Michaelis constant*. It can be shown that this approximation is valid under the condition that $[E]_0/(K_M + [S]) \ll 1$ holds. The speed of product formation is

$$v = k_{\text{cat}}[C] = V_{\max}\frac{[S]}{K_M + [S]} \ . \tag{32}$$

3.3 Analysis of experimental results

We note that Equation (28) and Equation (32) have the same analytical formula, although the constants K_d and K_M have a different physical origin. Note that in both cases, they correspond to the concentration at which the speed of product formation is half the value of the maximum $V_{\max}$. In practice, it is often forgotten which approximation is actually being used, and experimental works usually refer to the Michaelis constant K_M. As mentioned above, the typical experimental observable is the time evolution of the product concentration $[P]$. Qualitatively, the function $[P](t)$ behaves as depicted in Fig.2a: the product concentration starts increasing linearly, although it does not keep up the rate of production due to the limited amount of substrates available, so that it slows down until it eventually reaches a plateau. The slope of the linear regime (dashed lines) gives the speed v at the beginning of the reaction study, where $[S] \simeq [S]_0$. Naturally, increasing the amount of substrate impacts the speed, as shown by the comparison between the two continuous lines in Fig.2a. By repeating the

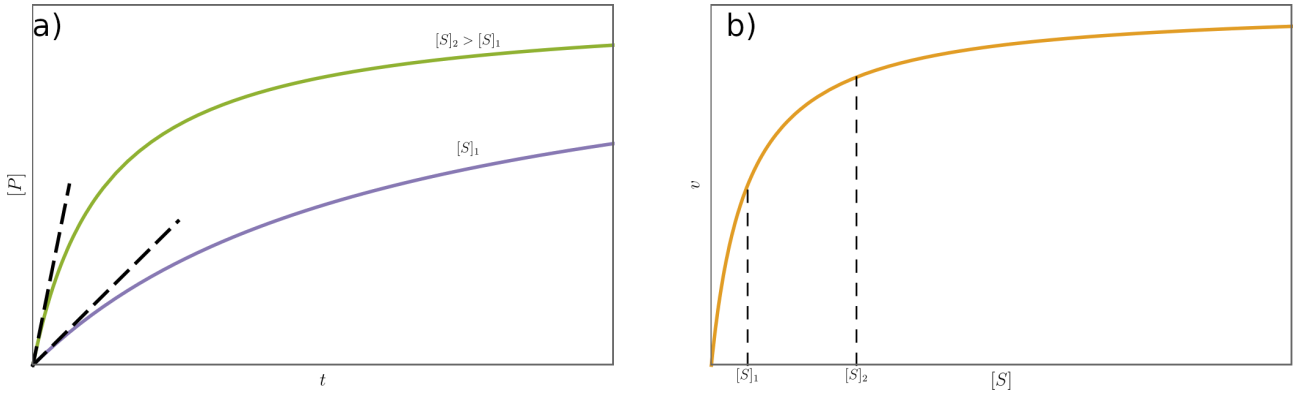

Figure 2: a) Typical time evolution of product formation for two different choices of substrate concentration. $[P](t)$ typically starts off in a linear fashion but then it slows down until it saturates to a plateau value. The initial velocity v corresponds to the slope of the dashed lines. b) Repeating the same experiment as in (a) at different amounts of substrate $[S]$ enables building the plot of v versus $[S]$, which is then fitted by means of Equation (32).

same experiments at different values of $[S]$, one can build the plot reported in Fig.2b, where v is depicted as a function of $[S]$. Fitting by means of Equation (32) enables extracting the Michaelis constant K_M and the maximum speed $V_{\max} = k_{\text{cat}}[E]_0$. Since the total enzyme concentration $[E]_0$ is a known parameter, this finally allows computing the value of k_{cat}. It is customary to quantitatively characterize how fruitful the reaction is by means of the ratio k_{cat}/K_M, which is called *catalytic efficiency* of *turnover*. The reason behind the choice of this ratio is that k_{cat} controls the maximum speed of product formation, $V_{\max}$, while K_M quantifies how fast is the maximum speed reached upon addition of substrate. A lower K_M corresponds to faster achievement of large speeds, so that optimal enzymatic reactions have low K_M and large k_{cat}. As a practical example, we report in Fig.3a the experimental points obtained for product formation of reactions involving ABTS as a substrate and catalyzed by the enzyme Horseradish Peroxidase. In this case, the enzyme and substrate are simply added to a water solution. Fitting by means of Equation (32) gives the continuous line with $K_M = 1.4$ mM, $k_{\text{cat}} = 78.4$ s^{-1} and turnover $k_{\text{cat}}/K_M \simeq 5.6 \cdot 10^5$ M^{-1}s^{-1}. In Fig.3b the same reaction is considered when the enzyme is embedded in a lipidic mesophase with two different nano-geometries. The impact of geometry is evident by comparing the two curves with each other and with the curve in Fig.3a. Even more strikingly, there is a qualitative change in the trend, since now the points follow a sigmoidal curve, which results in bad fitting by means of Equation (32). In this regard, it is customary to make use of the Hill equation, although the physical justification for this is not very solid. The Hill equation reads in this case

$$v = V_{\max} \frac{[S]^n}{K_M^n + [S]^n} \ . \tag{33}$$

The two continuous curves in Fig.3b are obtained by fitting by means of the Hill equation.

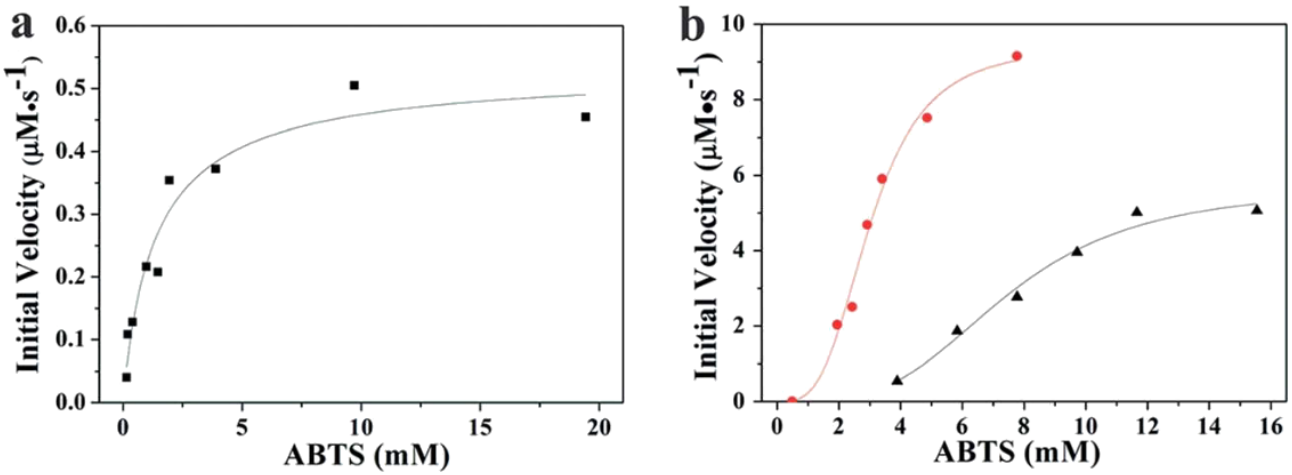

Figure 3: a) Speed of product formation for reactions considering the enzyme Horseradish Peroxidase and the substrate ABTS in water. b) Speed of product formation for the same system as in (a), but with the enzyme embedded in two lipid mesophases with different geometries. Adapted from W. Sun et al., *Nanoscale* **6**:6853 (2014).

a)

Enzyme	K_M (M)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)
Chymotrypsin	1.5×10^{-2}	0.14	9.3
Pepsin	3.0×10^{-4}	0.50	1.7×10^{3}
T-RNA synthetase	9.0×10^{-4}	7.6	8.4×10^{3}
Ribonuclease	7.9×10^{-3}	7.9×10^{2}	1.0×10^{5}
Carbonic anhydrase	2.6×10^{-2}	4.0×10^{5}	1.5×10^{7}
Fumarase	5.0×10^{-6}	8.0×10^{2}	1.6×10^{8}

b)

Kinetics parameters	Pure water	Swollen *Pn3m*	Standard *Pn3m*
K_m (mM)	1.4	3.0	7.5
n	0.9	3.3	3.4
K_{cat} (s^{-1})	78.4	72.2	27.6
$K_{cat}/K_{cat\ water}$	100%	92.1%	35.2%
K_{cat}/K_m ($mM^{-1}\ s^{-1}$)	55.6	23.9	3.7

Figure 4: Typical parameters encountered in experiments. Table (a) is reproduced from https://en.wikipedia.org/wiki/Michaelis-Menten_kinetics. Table (b) is reproduced from W. Sun et al., *Nanoscale* **6**:6853 (2014) and corresponds to the curves reported in Fig.3a ("pure water") and in Fig.3b ("swollen Pn3m" for red circles and "standard Pn3m" for black triangles).

4 Diffusion-limited reactions: Smoluchowskii limit

For a reaction to take place, the two reacting particles have first to move close to each other, which is typically obtained as a diffusion process. In the derivations above, we have used the well-mixing approximation, which assumes that diffusion is much faster than the typical reaction scales. Here we quantify better this statement by estimating the typical rate at which diffusing particles meet each other.

To this aim, we consider the simplest case of a solution with a certain bulk concentration c_0 of substrates of size r_S and a single enzyme of size r_E. We consider a process in which the reaction takes place as soon as a substrate and an enzyme meet, i.e. when their distance is $r_S + r_E$. This can be modelled by studying the profile of substrate concentration in the neighborhood of the enzyme by means of the diffusion equation in polar coordinates

$$\frac{\partial c}{\partial t} = D \cdot \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial c}{\partial r} \right) \ , \tag{34}$$

where $D = D_S + D_E$ is the relative diffusion coefficient, D_S and D_E being the diffusion coefficient of substrate and enzyme, respectively. The boundary conditions to be imposed are that at large distances the bulk concentration is obtained, $c(\infty) = c_0$, and that due to the instantaneous reaction $c(r_S + r_E) = 0$. We are interested in the long-time behavior, when the concentration profile has reached a steady state, so that $\partial c/\partial t = 0$. This enables solving easily the diffusion equation:

$$\frac{\partial}{\partial r} \left(r^2 \frac{\partial c}{\partial r} \right) = 0 \Rightarrow r^2 \frac{\partial c}{\partial r} = A \Rightarrow \frac{\partial c}{\partial r} = \frac{A}{r^2} \Rightarrow c = -\frac{A}{r} + B \ , \tag{35}$$

where A and B are integration constants. The condition $c(\infty) = c_0$ implies that $B = c_0$, while the condition $c(r_S + r_E) = 0$ gives $A = c_0 \cdot (r_S + r_E)$. Hence, the concentration profile is

$$c(r) = c_0 \left(1 - \frac{r_S + r_E}{r} \right) \ . \tag{36}$$

The flux of particles crossing the border $r = r_S + r_E$ is obtained as $J = -D\partial c/\partial r$, thus giving

$$J(r_S + r_E) = -(D_S + D_E) \cdot \frac{c_0}{r_S + r_E} \ . \tag{37}$$

The minus sign indicates that the flux goes towards the inside of the spherical surface. Remember that the flux is the amount of particles per unit surface and unit time, so that the rate (=number of particles per unit time) is obtained by multiplying $|J|$ by the spherical surface with radius $r_S + r_E$:

$$(D_S + D_E) \cdot \frac{c_0}{r_S + r_E} \cdot 4\pi \left(r_S + r_E \right)^2 = 4\pi c_0 (D_S + D_E) \left(r_S + r_E \right) \ . \tag{38}$$

Within this context, this is called the *Smoluchowskii limit* (be careful not to make confusion with the homonimous limit that we saw in the Langevin equation!). In order to compare this

result with the turnover of a reaction-limited enzymatic reaction, it is useful to consider the previous rate divided by c_0 (so that it has the same units as a turnover), thus giving the diffusion rate constant

$$k_{\text{diff}} = 4\pi(D_S + D_E)\,(r_S + r_E) \ . \tag{39}$$

We recall that the diffusion coefficient D of a sphere of size R can be extimated by means of the Stokes-Einstein equation as

$$D = \frac{k_B T}{6\pi\eta R} \ , \tag{40}$$

where η is the viscosity of the solvent. With this formula, the diffusion rate constant can be written as

$$k_{\text{diff}} = \frac{2k_B T}{3\eta} \frac{(r_S + r_E)^2}{r_S\, r_E} \ . \tag{41}$$

Assuming for simplicity that both the substrate and the enzyme have the same size, the previous formula further simplifies to

$$k_{\text{diff}} = \frac{8k_B T}{3\eta} \ . \tag{42}$$

Substituting $k_B = 1.38 \cdot 10^{-23}$, $T = 300K$ and the viscosity of water $\eta \simeq 10^{-3}$ Pa·s, we obtain $k_{\text{diff}} \simeq 2.76 \cdot 10^{-18}$ m^3/s. Remembering that 1 M= 1 mol/l= 10^3 mol/m^3, in order to convert to the proper units we have to multiply k_{diff} by $10^3 N_A$, where $N_A = 6.022 \cdot 10^{23}$ mol^{-1} is Avogadro's constant. Hence, we finally obtain $k_{\text{diff}} \simeq 1.66 \cdot 10^9$ M^{-1}·s^{-1}. The exact numerical value is of course just a rough approximation, but it is important to observe that the order of magnitude is much larger than the values reported in Fig.4, thus justifying *a posteriori* the well-mixing approximation used to extract them.

Further reading

- https://en.wikipedia.org/wiki/Michaelis-Menten_kinetics
- R. Delgado Buscalioni, notes on ligand-receptor binding (you can find them in the Moodle)
- R. Phillips et al., "Physical Biology of the Cell", 2nd Ed., Garland Science (2013), Chapter 13