3.5 Cooperativity

The study of Hemoglobin in the beginning of the 20th century motivated study of proteins with multiple binding sites, and the resulting cooperative effects. The experimental results (below) indicated that the binding curves of oxygen to hemoglobin had a sigmoidal, inconsistent with expectations from a single binding site. In 1910, Hill introduced a model to explain these results which is the subject of this section.

Hemoglobin is a fully α-helical protein consisting of four chains, bound together. Each chain forms a compact domain. The four domains fall into two types, denoted by the Greek letters α and β. Each domain binds to a bulky molecule of *heme,* which has an iron atom in the center of a porphyrin ring.

The purpose of hemoglobin is to carry oxygen. In the lungs, hemoglobin in the blood picks up O_2 from inhaled air. The blood carries HbO_2 to the muscles, where the O_2 is transferred to molecules of *myoglobin,* a protein similar to Hb but having only one domain. (We will see later why this difference is important.) In the muscles, respiration takes up oxygen from the myoglobin, releasing CO_2.

3.5.1 Simple Binding

Imagine that we have a molecule of hemoglobin. Originally, it was not known that one Hb molecule could bind four O_2 molecules. If so, the binding of a single molecule satisfies the equation

$$\mathrm{Hb} + \mathrm{O_2} \Leftrightarrow \mathrm{HbO_2}, \tag{3.65}$$

with an equilibrium constant K_{eq} given by

$$K_{\mathrm{eq}} = \frac{[\mathrm{Hb}][\mathrm{O_2}]}{[\mathrm{HbO_2}]}. \tag{3.66}$$

The fraction of Hb molecules which are oxygenated, which we shall label Y, is given by

$$Y \equiv \frac{[\mathrm{HbO_2}]}{[\mathrm{Hb}][\mathrm{HbO_2}]}, \tag{3.67}$$

which in terms of the equilibrium constant is

$$Y = \frac{[\mathrm{O_2}]}{[\mathrm{O_2}] + K_{\mathrm{eq}}}. \tag{3.68}$$

The plot of Y as a function of the oxygen concentration starts linearly, and gently curves to its asymptotic value.

In 1904, Christian Bohr (the father of physicist Niels Bohr) measured the fraction of oxygenated protein molecules as a function of oxygen concentration. For myoglobin, the result matched Eq. (3.68)'s prediction. However, the results for hemoglobin where startlingly different, following instead a sigmoidal (later termed "coöperative") curve.

The physiological significance of the different shape is the following: Due the sharply sigmodical curve, Hb molecules will release significant amounts of O_2 under small changes in oxygen pressure. Such rapid release is relevant to function of Hb in blood, whereas the function of Mb, storing oxygen in muscle, is quite different.

3.5.2 Cooperative Binding

The first explanation of the sigmoidal Hb curve was proposed by Archibald Hill in 1913. He hypothesized that several molecules of oxygen simultaneously bind to the same Hb unit, i.e.

$$\mathrm{Hb} + n\mathrm{O_2} \Leftrightarrow \mathrm{Hb}(\mathrm{O_2})_n, \tag{3.69}$$

implying that the equilibrium constant is

$$K_{\mathrm{eq}} = \frac{[\mathrm{Hb}][\mathrm{O_2}]^n}{[(\mathrm{O_2})_n]}. \tag{3.70}$$

If we are interested in the fraction Y of oxygenated Hb, we can use the same reasoning as before to obtain

$$Y = \frac{[\mathrm{O_2}]^n}{[\mathrm{O_2}]^n + K_{\mathrm{eq}}}. \tag{3.71}$$

As a means of measuring the parameter n, Hill suggested recording the logarithm of $Y/(1-Y)$, which from Eq. (3.71) takes the form

$$\log\left(\frac{Y}{1-Y}\right) = n\log[\mathrm{O_2}] - \log K_{\mathrm{eq}}. \tag{3.72}$$

A ("Hill") plot of the logarithm of $Y/(1-Y)$ against $\log[\mathrm{O_2}]$, in the simplest case, should yield a straight line of slope n (the "Hill coefficient"). This is what is seem for myoglobin, but for hemoglobin the result is more complex: the slope begins at 1, changes abruptly to 3, and then drops again to 1. Based on these results in 1913, there was the suggestion that Hb contains *at least* 3 subunits, with additional mechanisms necessary to explain the full behavior.

Note that in writing Eq. (3.69), Hill implicitly makes the assumption that each Hb molecule is found in either the completely oxygenated, or completely deoxygenated, state. In 1924, Adair proposed that the intermediate states with oxygen occupancies between 0 and 4 must be relevant to the shape of the curve. (By this point, it was known that Hb consisted of four subunits.) Adair's phenomenological model is more general than the specific case Hb situation, and as such we shall speak of enzymes E which can bind several molecules of substrate S. The set of reaction equations for $n = 4$ can then be written as

$$E + S \Leftrightarrow ES \Leftrightarrow ES_2 \Leftrightarrow ES_3 \Leftrightarrow ES_4, \tag{3.73}$$

where each reaction has a macroscopic equilibrium constant,

$$K_i = \frac{[S][ES_{i-1}]}{[ES_i]}. \tag{3.74}$$

Adair's idea was that the binding strength of each site changes with the number of substrates adsorbed. Microscopic equilibrium constants are different from macroscopic constants, because the concentration of *sites* is greater than the concentration of *molecules.* At the microscopic level,

$$k_i = \frac{[S][ES_{i-1}](n-i+1)}{[ES_i]} i. \tag{3.75}$$

The binding occupancy measured in experiment, Y, is

$$Y = \frac{[ES] + 2[ES_2] + 3[ES_3] + 4[ES_4]}{4([ES] + [ES_2] + [ES_3] + [ES_4] + [E])}, \tag{3.76}$$

which using Eq. (3.76) can be cast as

$$Y = \frac{S/k_1 + 3S^2/(k_1k_2) + 3S^3/(k_1k_2k_3) + S^4/(k_1k_2k_3k_4)}{1 + 4S/k_1 + 6S^2/(k_1k_2) + 4S^3/(k_1k_2k_3) + S^4/(k_1k_2k_3k_4)}. \tag{3.77}$$

Using this relation, known as the Adair equation, one can estimate $\{k_i\}$ by fitting measured values of Y versus S. The table of values of $\{k_i\}$ (in torr) for Hb appears below.

k_1	k_2	k_3	k_4
41	13	12	0.14

In 1965, Monod, Wynman and Changeux proposed a simpler model which demonstrated how coöperativity can be achieved with identical subunits. This model postulates a protein whose subunits can exist in two states, T (for tense) and R (for relaxed). The central assumption is that all four subunits change from T to R or vice versa *at the same time.*

$$T_0 \Leftrightarrow R_0 \tag{3.78}$$

$$\begin{aligned} T_0 + S &\Leftrightarrow T_1 \\ T_1 + S &\Leftrightarrow T_2 \\ &\vdots \\ T_{n-1} + S &\Leftrightarrow T_n \end{aligned}$$

$$\begin{aligned} R_0 + S &\Leftrightarrow R_1 \\ R_1 + S &\Leftrightarrow R_2 \\ &\vdots \\ R_{n-1} + S &\Leftrightarrow R_n \end{aligned}$$

We define the useful quantities

$$L = \frac{T_0}{R_0},\ \alpha = \frac{[S]}{k_R},\ c = \frac{k_R}{k_T}. \tag{3.79}$$

Here, L stands for the bias between the tense and relaxed states, and α is the normalized concentration of substrate (for example, oxygen). c is the ratio of binding constants to the two states. If c is near zero, the substrate binds preferentially to R.

The microscopic equilibrium constant for the relaxed state is

$$k_R = \frac{n-i+1}{i}\frac{R_{i-1}[S]}{R_i}. \tag{3.80}$$

$$Y = \frac{R_1 + 2R_2 + \cdots + nR_n + T_1 + 2T_2 + \cdots + nT_n}{n(R_0 + R_1 + \cdots + R_n + T_0 + T_1 + \cdots + T_n)}. \tag{3.81}$$

First, let's find an expression for the sum in the numerator involving R_i.

$$\begin{aligned} R_1 + 2R_2 + \cdots + nR_n &= R_0\left(n\alpha + \frac{2n(n-1)\alpha^2}{2} + \frac{3n(n-1)(n-2)\alpha^3}{2\cdot 3} + \cdots + \frac{nn!}{n!}\alpha^n\right) \\ &= R_0 n\left(1 + (n-1)\alpha + \frac{(n-1)(n-2)}{2}\alpha^2 + \cdots + \alpha^{n-1}\right) \\ &= R_0\alpha n(1+\alpha)^{n-1}. \end{aligned}$$

A similar expression will hold for the sum in the denominator of Eq. (3.81):

$$\begin{aligned} R_0 + R_1 + \cdots + R_n &= R_0\left(1 + n\alpha + \frac{n(n-1)}{2}\alpha^2 + \cdots + \alpha^n\right) \\ &= R_0(1+\alpha)^n. \end{aligned}$$

Recalling that $T_0 = LR_0$,

$$Y = \frac{R_0\alpha(1+\alpha)^{n-1}R_0 Ln\alpha c(1+\alpha c)^{n-1}}{n\left(R_0(1+\alpha)^n + R_0 L(1+\alpha c)^n\right)}. \tag{3.82}$$

We can eliminate various factors of R_0 and n, leaving

$$\boxed{Y = \frac{\alpha(1+\alpha)^{n-1} + L\alpha c(1+\alpha c)^{n-1}}{(1+\alpha)^n + L(1+\alpha c)^n}.} \tag{3.83}$$

For best coöperativity, we want c close to zero and L large. That is, we want the protein to spend most of its time in the T state, but we want the ligand to bind preferentially to R!

At the beginning, we promised you population genetics, and now we can deliver, thanks to the disease called *sickle-cell anemia.* Patients with this disease have erythrocytes (red blood cells) which take on sickle shapes, as opposed to the double-dimpled disc shapes of

regular erythrocytes. In the 1940s, it was suggested that the difference in cell shape was related to a difference in hemoglobin composition. It turns out that the difference is due to a mutation in a single amino acid in the hemoglobin gene. This mutation produces a place where a valine can form a close association with a leucine on another subunit, changing the protein's conformation and allowing deoxygentated Hb molecules to form fibrils. These fibrils stretch the erythrocytes, making them unable to fit through capillaries, sometimes straining them to bursting. Surprisingly, this mutation ("Hemoglobin S") occurs with high frequency in some African regions. In some areas, more than 14% of the population carries the sickle-cell allele.

It is important to know that HbS is recessive, . *i.e.,* that only homozygotes have full symptoms of the disease. From our work earlier, we recall that if the frequency of HbS is x, a simple population genetics model predicts the only stable states to be $x = 0$ and $x = 1$. Other stable states are possible, if the fitness of the heterozygote is larger than the fitness of either homozygote.

As it turns out, possessing one HbS allele gives the owner some resistance to malaria.