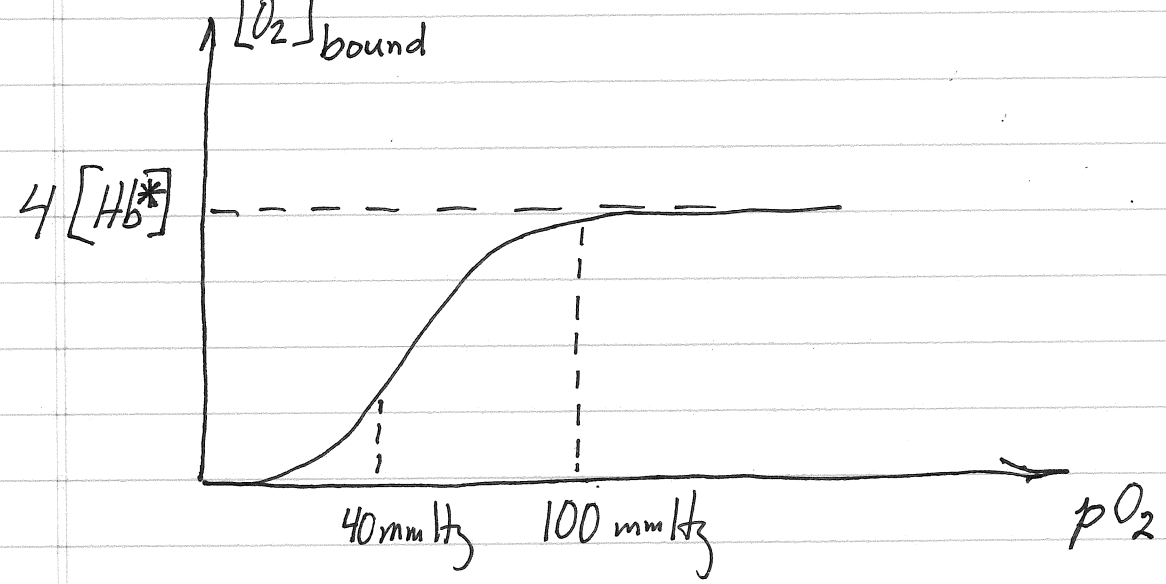
CS Peakm 11/1/2018

Hemeoglobin

Hemeoglobin is a protein that binds O_2. It is found in red blood cells, and its role is to transport O_2 from the lungs to the tissues. Each hemeoglobin molecule has four subunits, and each subunit contains one iron atom in a particular environment that makes this iron atom capable of binding O_2 reversibly, without actually becoming oxidized. Thus, each molecule of hemeoglobin can carry up to four molecules of O_2.

The binding of O_2 to hemeoglobin is described by a curve that looks like this



In this plot, $[O_2]_{bound}$ denotes the concentration of O_2 that is bound to hemeoglobin:

(1) $$[O_2]_{bound} = [Hb\,O_2] + 2[Hb\,2O_2] + 3[Hb\,3O_2] + 4[Hb\,4O_2]$$

where $[Hb\,kO_2]$ is the concentration of hemeoglobin molecules that have exactly k molecules of O_2 bound (and therefore $4-k$ empty sites for O_2 binding). Clearly,

(2) $$[O_2]_{bound} \leq 4[Hb^*]$$

where $[Hb^*]$ is the total concentration of hemeoglobin:

(3) $$[Hb^*] = [Hb] + [Hb\,O_2] + [Hb\,2O_2] + [Hb\,3O_2] + [Hb\,4O_2]$$

The independent variable of the hemoglobin curve is pO_2, the partial pressure of O_2. The concept of partial pressure of a gas in a liquid requires discussion.

First, consider a gas which is a mixture of molecular species, like the atmosphere. Each species exerts its own partial pressure as if the other species were not present, and the total pressure is the sum of the partial pressures. The atmosphere is 1/5 oxygen, and has a total pressure of 760 mmHg, so the partial pressure of O_2 in the atmosphere is 152 mmHg.

The partial pressure of a gas in a liquid is defined to be the partial pressure in a gas with which that liquid would be in equilibrium. For example, if a sample of blood has a partial pressure of oxygen equal to 100 mmHg (a realistic value for arterial blood), and if that sample is put in contact with a gas having a partial pressure of 100 mmHg, then the sample of blood will neither

gain nor lose oxygen, since the blood and the gas are already at equilibrium with respect to oxygen. If, on the other hand, that same sample of blood is brought into contact with the atmosphere, it will gain oxygen, since the partial pressure of oxygen in the atmosphere (152 mm Hg) is greater than the partial pressure of oxygen in the sample of blood (100 mm Hg).

Thus, when two materials are brought into contact so that oxygen can be exchanged between them, it is the difference of their partial pressures (and not the difference of their concentrations) of oxygen that determines which way the oxygen will move.

In addition to the O_2 that is carried by hemoglobin, oxygen can also be dissolved in blood, forming a solution. The concentration of dissolved oxygen will here be denoted $[O_2]$, since this is our usual notation for a free molecular species. The free concentration is linearly related to the partial pressure of oxygen:

(4) $$[O_2] = \sigma p_{O_2}$$

The constant σ is small enough that most of the oxygen in blood is carried by hemeoglobin under normal circumstances. As the partial pressure of oxygen increases, however, $[O_2]_{bound}$ has an upper limit, see (2), but $[O_2]$ as given by (4) does not, so the dissolved oxygen can become physiologically significant. This happens, for example, when breathing pure oxygen and even more so when this is done in a hyperbaric (high-pressure) chamber.

Since $[O_2]$ (which means $[O_2]_{free}$) is linearly related to p_{O_2}, the horizontal axis of the hemeoglobin curve could alternatively be labeled $[O_2]$ instead of p_{O_2}, without changing the shape of the curve. (The numbers on the axis would, of course, change, as would their units.) We shall adopt this point of view when we study the mechanism underlying the hemeoglobin curve.

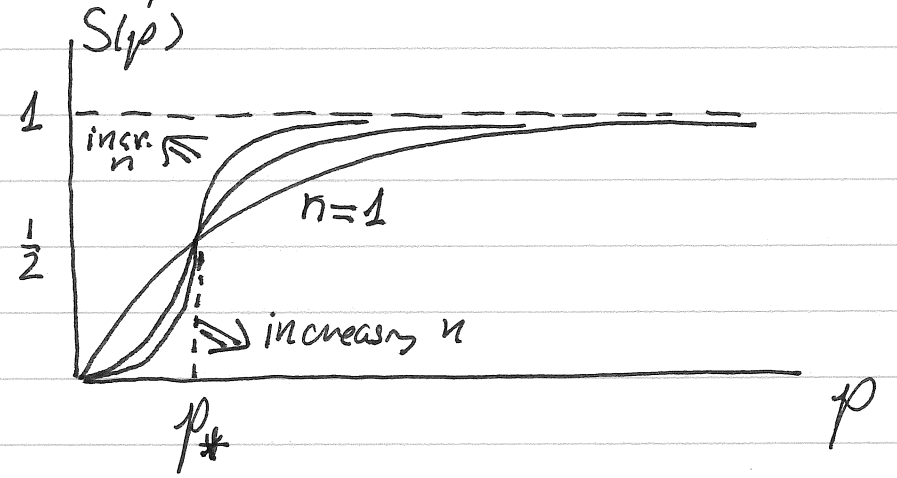
The shape of the hemeoglobin curve suggests that it could be modeled by

$$[O_2]_{bound} = [O_2]^{max}_{bound} \, S_n(p) \tag{5}$$

where

$$S_n(p) = \frac{(p/p_*)^n}{1 + (p/p_*)^n} \tag{6}$$

in which p denotes pO_2, $[O_2]^{max}_{bound}$ is the limiting value of $[O_2]_{bound}$ as $p \to \infty$, and $S(p)$ is called the saturation of hemeoglobin. The parameters of this model, besides $[O_2]^{max}_{bound}$, are the constants p_* and n. As n increases, these curves get steeper and steeper



In fact,

$$(7) \qquad \lim_{n\to\infty} S_n(p) = \begin{cases} 0 & , \quad p < p_* \\ \frac{1}{2} & , \quad p = p_* \\ 1 & , \quad p > p_* \end{cases}$$

This illustrates the mathematical fact that the limit of a sequence of continuous functions does not have to be continuous.

The function $S_n(p)$ can be derived from the following unrealistic model: Suppose each hemoglobin molecule has n sites at which O_2 could be bound, but these sites interact in such a way that the only allowed states are Hb and $Hb\,nO_2$. We leave the derivation of $S_n(p)$ from this model as an exercise for the reader. It fits the hemoglobin curve best with $n \approx 3$. This shows that the model is not really a good one, despite capturing the shape of the curve, since we know that hemoglobin has four oxygen-carrying sites.

The following optimization problem illustrates the benefit of higher n for the function of hemoglobin:

Let $P_1 = 40$ mm Hg be the partial pressure of O_2 in the tissues to which oxygen is being delivered, and let $P_2 = 100$ mm Hg be the partial pressure of oxygen in the lungs. (Note that this is substantially less than the partial pressure of oxygen in the atmosphere, since oxygen is continually being removed from the lungs as it is picked up by the blood for delivery to the tissues.)

For each $n = 1, 2, 3, 4, \ldots$, find the value of P^* that maximizes

(8) $$\Delta S_n = S_n(P_2) - S_n(P_1)$$

Note that ΔS_n is the fraction of the carrying capacity of hemoglobin that is actually being used.

The result, P^*, may depend on n. For each n, use the best value of P^*

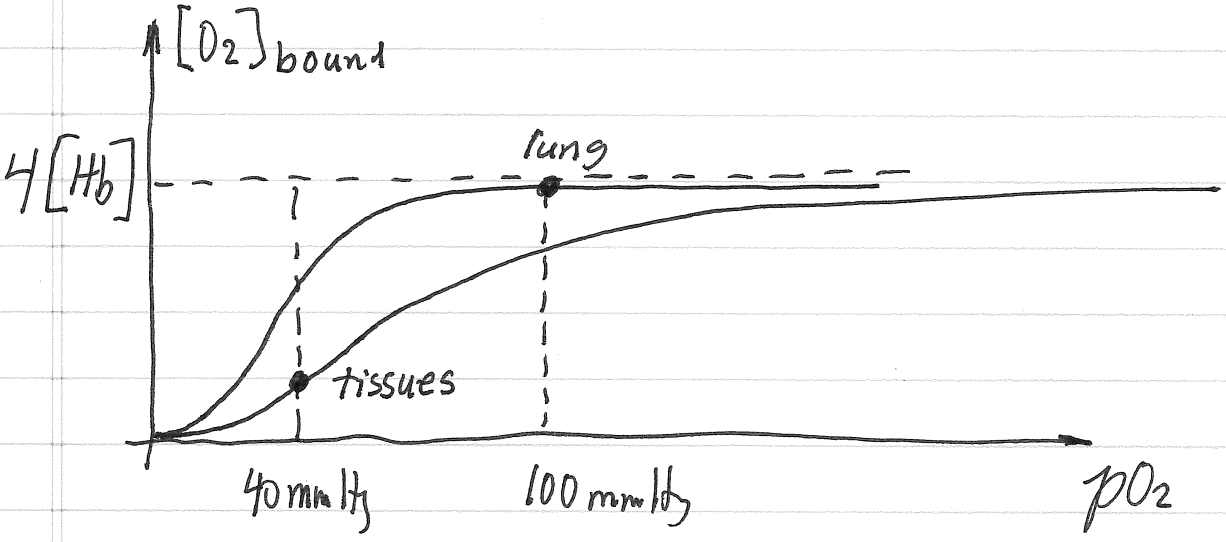
and evaluate ΔS_n^{max}. Plot the result as a function of n.

Even for fixed n, it is interesting that there is an optimal choice of p_*. This means that the affinity of hemeoglobin for oxygen should not be too high or too low. (High affinity corresponds to a small value of p_*, since smaller p_* makes $S_n(p)$ larger for any given p.) If the affinity is too high, hemeoglobin will carry a lot of oxygen but be unable to unload it in the tissues. If the affinity is too low, hemeoglobin will not carry enough oxygen.

What the foregoing exercise shows, however, is that even after optimizing ΔS_n by choosing the best p_* for some fixed n, further improvement can be obtained by increasing n.

Besides optimizing p_* and choosing a shape for the hemeoglobin curve corresponding to $n=3$, nature has one more trick up its sleeve for delivering as much oxygen to the

tissues as possible. This is to shift the whole curve so that the relationship between pO_2 and $[O_2]_{bound}$ is different in the tissues from what it is in the lungs. The direction of the shift is that hemoglobin has a higher affinity for oxygen in the lungs than it has in the tissues, thus facilitating the uptake of O_2 in the lungs and the delivery of O_2 to the tissues. The way this works is that the tissues have a higher concentration of CO_2 and are also more acidic, and both of these conditions lower the affinity of hemoglobin for oxygen, as shown in the figure below.



Allosteric model*

At low pO_2, the slope of the hemoglobin curve increases with increasing pO_2. This suggests that the binding of oxygen to one or more sites increases the affinity of the remaining sites for oxygen. A theory of how this works is called the allosteric ("other shape") model. It is widely applicable to proteins with multiple subunits, including not only hemoglobin but also many enzymes and receptor molecules.

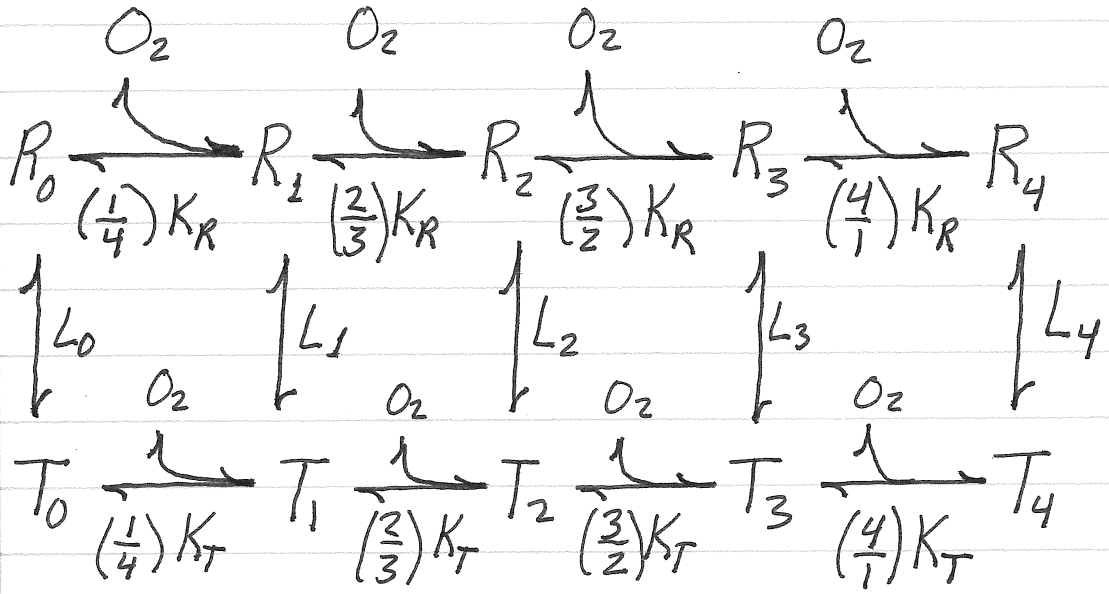
The theory postulates that the hemoglobin molecule, as a whole, can be in one of two states. The transition between these two states involves all of the subunits at once, as they snap from one stable configuration of the whole molecule into another. In each state, the four binding sites for oxygen

---

* Monod J, Wyman J, and Changeux J-P On the nature of allosteric transitions: A plausible model. J. Mol. Biol 12, 88-118, 1965

function independently of each other, but the affinity of each of them for oxygen depends on which state the hemoglobin molecule as a whole is in. It then follows from the principle of detailed balance that the binding of oxygen shifts the equilibrium between the two states of hemoglobin. The binding of other molecules besides oxygen to other sites on the hemoglobin molecule can also shift the equilibrium between the two overall states of hemoglobin, and this provides an important regulatory mechanism (see project suggestion at end of these notes.)

The reaction scheme of the allosteric model is shown at equilibrium in the following figure. The two global states of hemoglobin are called T and R, and a subscript indicates the number of oxygen molecules that are bound. The equilibrium (dissociation) constants for any one site are denoted K_T and K_R.



The numerical factors multiplying K_R and K_T in the above reaction scheme are explained as follows. Consider the reaction

$$R_i + O_2 \underset{(i+1)\gamma_R}{\overset{(4-i)k_R}{\rightleftharpoons}} R_{i+1} \tag{9}$$

Here k_R and γ_R are the single-site rate constants. The rate constant for the forward reaction is $(4-i)k_R$ because there are $(4-i)$ empty sites available to bind oxygen. The rate constant for the reverse reaction is $(i+1)\gamma_R$ since there are $i+1$ molecules of oxygen bound to R_{i+1} and available to dissociate.

At equilibrium

(10) $$(4-i)\, k_R [R_i][O_2] = (i+1)\, \gamma_R [R_{i+1}]$$

Thus

(11) $$\frac{[R_i][O_2]}{[R_{i+1}]} = \frac{(i+1)\gamma_R}{(4-i)\, k_R} = \left(\frac{i+1}{4-i}\right) K_R$$

where

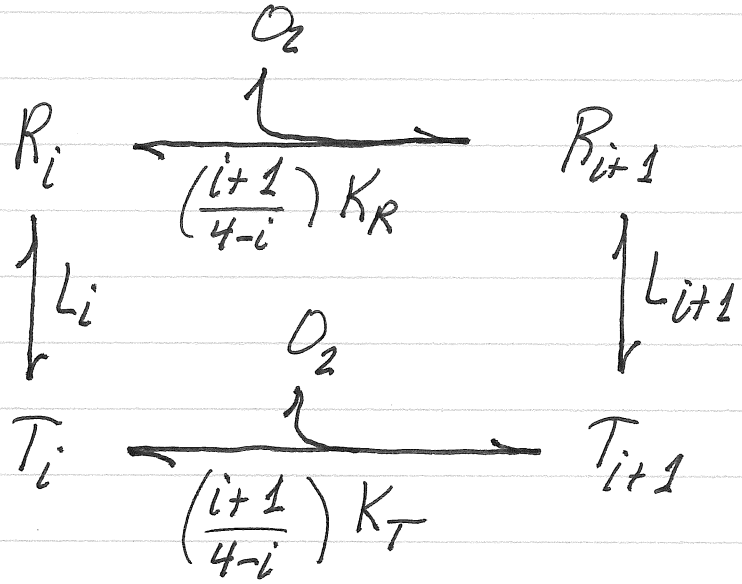
(12) $$K_R = \frac{\gamma_R}{k_R}$$

so that K_R is the single-site equilibrium constant. It follows that the equilibrium (dissociation) constant for the reaction (9) is

(13) $$\left(\frac{i+1}{4-i}\right) K_R$$

Of course, the same considerations apply with R replaced by T, and we get the same numerical factors.

Now we apply the principle of detailed balance to the allosteric reaction scheme for hemeoglobin. There are four loops of the form



for $i = 0, 1, 2, 3$. According to the principle of detailed balance, the product of the equilibrium constants around any loop in a reaction scheme must be equal to 1. Applying this to a counterclockwise circuit of the above loop, and paying attention to the direction in which we a proceeding and the direction for which the equilibrium constant is defined, we get

$$\left(\frac{i+1}{4-i}\right) K_R \, L_i \, \frac{1}{\left(\frac{i+1}{4-i}\right) K_T} \, \frac{1}{L_{i+1}} = 1 \tag{14}$$

which simplifies to

$$\frac{L_{i+1}}{L_i} = \frac{K_R}{K_T} \tag{15}$$

and this immediately implies

$$L_i = \left(\frac{K_R}{K_T}\right)^i L_0 \tag{16}$$

so the L_i form a geometric sequence with ratio K_R/K_T. Thus, the whole allosteric scheme has three parameters (besides the number of subunits $n_s = 4$), which may be taken to be K_R, K_T, and L_0. The parameters K_R and K_T have units of concentration, but L_0 is dimensionless.

Suppose, for example, that $L_0 = 10^4$ and $(K_R/K_T) = 10^{-2}$. This is not far from the actual situation in hemoglobin. Then

$$(L_0 \cdots L_4) = 10^4, 10^2, 10^0, 10^{-2}, 10^{-4} \tag{17}$$

Rearranging equation (11) and lowering i by 1 we get

$$\frac{[R_i]}{[R_{i-1}]} = \frac{4-i+1}{i} \frac{[O_2]}{K_R} \tag{18}$$

and therefore

$$\frac{[R_i]}{[R_0]} = \frac{(4)\cdots(4-i+1)}{1 \cdots i} \left(\frac{[O_2]}{K_R}\right)^i \tag{19}$$

$$= \frac{4!}{(4-i)!\ i!} \left(\frac{[O_2]}{K_R}\right)^i = \binom{4}{i} \left(\frac{[O_2]}{K_R}\right)^i$$

Note that $\binom{4}{i}$ is the number of ways of choosing i sites at which to bind O_2 from the 4 sites on the hemoglobin molecule.

Similarly

$$\frac{[T_i]}{[T_0]} = \binom{4}{i} \left(\frac{[O_2]}{K_T}\right)^i \tag{20}$$

Also

$$\frac{[R_0]}{[T_0]} = \frac{1}{L_0} \tag{21}$$

From (19-21), we can express the concentration of any state of hemeoglobin in terms of the concentration of T_0. In this way, we get

(22) $$[Hb^*] = \sum_{i=0}^{4} [R_i] + \sum_{i=0}^{4} [T_i]$$

$$= [T_0]\left(\frac{1}{L_0} \sum_{i=0}^{4} \binom{4}{i} \left(\frac{[O_2]}{K_R}\right)^i + \sum_{i=0}^{4} \binom{4}{i} \left(\frac{[O_2]}{K_T}\right)^i \right)$$

and also

(23) $$[O_2]_{bound} = \sum_{i=1}^{4} i[R_i] + \sum_{i=1}^{4} i[T_i]$$

$$= [T_0]\left(\frac{1}{L_0} \sum_{i=1}^{4} i\binom{4}{i} \left(\frac{[O_2]}{K_R}\right)^i + \sum_{i=1}^{4} i\binom{4}{i} \left(\frac{[O_2]}{K_T}\right)^i \right)$$

The sums in (22-23) can be evaluated by making use of the binomial theorem. First,

(24) $$\sum_{k=0}^{n} \binom{n}{k} x^k = (1+x)^n$$

Also, by differentiating on both sides with respect to x and then multiplying the results by x, we get

(25) $$\sum_{k=1}^{n} k\binom{n}{k} x^k = nx(1+x)^{n-1}$$

Thus (22.23) become

(26) $$[HbO_2^*] = [T_0]\left(\frac{1}{L_0}\left(1+\frac{[O_2]}{K_R}\right)^4 + \left(1+\frac{[O_2]}{K_T}\right)^4\right)$$

(27) $$[O_2]_{bound} = 4[T_0]\left(\frac{1}{L_0}\left(1+\frac{[O_2]}{K_R}\right)^3 \frac{[O_2]}{K_R} + \left(1+\frac{[O_2]}{K_T}\right)^3 \frac{[O_2]}{K_T}\right)$$

The saturation of hemoglobin's four sites by oxygen is therefore given by

(28) $$S = \frac{[O_2]_{bound}}{4[Hb^*]} =$$

$$\frac{\frac{1}{L_0}\left(1+\frac{[O_2]}{K_R}\right)^3 \frac{[O_2]}{K_R} + \left(1+\frac{[O_2]}{K_T}\right)^3 \frac{[O_2]}{K_T}}{\frac{1}{L_0}\left(1+\frac{[O_2]}{K_R}\right)^4 + \left(1+\frac{[O_2]}{K_T}\right)^4}$$

If we let $L_0 \to \infty$ with K_R and K_T fixed, we get

(29) $$\lim_{L_0 \to \infty} S = \frac{\frac{[O_2]}{K_T}}{1+\frac{[O_2]}{K_T}}$$

Similarly, if we let $L_0 \to 0$ with K_R and K_T fixed, the result is

(30) $$\lim_{L_0 \to 0} S = \frac{\frac{[O_2]}{K_R}}{1 + \frac{[O_2]}{K_R}}$$

Equations (29) & (30) are exactly what we would get if the four oxygen binding sites of hemeoglobin were completely independent of each other, and if each of them had the dissociation constant K_T in the case of (29) or K_R in the case of (30). This illustrates the somewhat paradoxical nature of the allosteric model: if we lock the hemeoglobin molecule as a whole in its R state, then the four oxygen-binding sites function independently, and this is also true if we lock the hemeoglobin molecule in its T state. Nevertheless, when transitions $T \leftrightarrow R$ are allowed, the four sites interact because the binding of oxygen shifts that equilibrium in the direction that favors stronger oxygen binding!

Equation (28) for S is essentially a formula for the hemoglobin curve. [Recall that $[O_2]$ denotes the free concentration of oxygen, and that this is proportional to the partial pressure of oxygen.)

The formula for S involves 3 parameters. We can reduce this to two by choosing $\sqrt{K_R K_T}$ as a reference concentration. Let

$$x = \frac{[O_2]}{\sqrt{K_R K_T}}, \qquad \varepsilon = \sqrt{\frac{K_R}{K_T}} \tag{31}$$

Then

$$S = \frac{\frac{1}{L_0}\left(1+\frac{x}{\varepsilon}\right)^3 \frac{x}{\varepsilon} + (1+\varepsilon x)^3 \varepsilon x}{\frac{1}{L_0}\left(1+\frac{x}{\varepsilon}\right)^4 + (1+\varepsilon x)^4} \tag{32}$$

A particularly nice choice of L_0 ~~(discussed later)~~ ~~is~~ (see remark containing equations (54-57), below) is

(33) $$L_0 = \frac{1}{\varepsilon^4}$$

Note that we then get the example parameters mentioned above (which are fairly realistic for hemoglobin) if we set $\varepsilon = 0.1$, see (17). With this choice of L_0,

(34) $$S = \frac{(\varepsilon + x)^3 x + (1 + \varepsilon x)^3 \varepsilon x}{(\varepsilon + x)^4 + (1 + \varepsilon x)^4}$$

If we set $\varepsilon = 1$ in (34), we get

(35) $$S_{\varepsilon=1} = \frac{x}{1 + x}$$

but if we set $\varepsilon = 0$, (34) reduces to

(36) $$S_{\varepsilon=0} = \frac{x^4}{1 + x^4}$$

Let the right-hand side of (34) be denoted $S_\varepsilon(x)$. We then have

$$S_\varepsilon(1) = \frac{1}{2} \tag{37}$$

for all $\varepsilon \in [0,1]$, and it is interesting to see how $S_\varepsilon'(1)$ depends on ε.

Let

$$D_\varepsilon(x) = (\varepsilon + x)^4 + (1+\varepsilon x)^4 \tag{38}$$

Then

$$D_\varepsilon'(x) = 4(\varepsilon+x)^3 + 4(1+\varepsilon x)^3 \varepsilon \tag{39}$$

$$D_\varepsilon''(x) = 12(\varepsilon+x)^2 + 12(1+\varepsilon x)^2 \varepsilon^2 \tag{40}$$

Therefore

$$D_\varepsilon(1) = 2(1+\varepsilon)^4 \tag{41}$$

$$D_\varepsilon'(1) = 4(1+\varepsilon)^3(1+\varepsilon) = 4(1+\varepsilon)^4 \tag{42}$$

$$D_\varepsilon''(1) = 12(1+\varepsilon)^2(1+\varepsilon^2) \tag{43}$$

From (38) & (39),

$$S_\varepsilon(x) = \frac{x}{4}\,\frac{D_\varepsilon'(x)}{D_\varepsilon(x)} \tag{44}$$

Therefore

$$S_\varepsilon'(x) = \frac{1}{4}\,\frac{D_\varepsilon'(x)}{D_\varepsilon(x)} \tag{45}$$

$$+\frac{x}{4}\,\frac{D_\varepsilon(x)D_\varepsilon''(x) - (D_\varepsilon'(x))^2}{(D_\varepsilon(x))^2}$$

We can evaluate (44) & (45) at $x=1$ by using (41-43). In this way, we get

$$S_\varepsilon(1) = \frac{1}{4}\,\frac{4(1+\varepsilon)^4}{2(1+\varepsilon)^4} = \frac{1}{2} \tag{46}$$

$$S_\varepsilon'(1) = \frac{1}{4}\left(2 + \frac{2(1+\varepsilon)^4\,12(1+\varepsilon)^2(1+\varepsilon^2)}{4(1+\varepsilon)^8} - 4\right) \tag{47}$$

//

$$= \frac{1}{4}\left(\frac{6(1+\varepsilon^2)}{(1+\varepsilon)^2} - 2 \right)$$

$$= \frac{1}{4}\left(\frac{6(1+\varepsilon)^2 - 12\varepsilon}{(1+\varepsilon)^2} - 2 \right)$$

$$= \frac{1}{4}\left(4 - \frac{12\varepsilon}{(1+\varepsilon)^2}\right) = 1 - \frac{3\varepsilon}{(1+\varepsilon)^2}$$

It is easy to check that this formula agrees with the slopes that can be evaluated directly from (35) & (36) at $x = 1$, namely

(48) $$S_1'(1) = \frac{1}{4} \quad , \quad S_0'(1) = 1$$

Also, with our idealized choice of parameters for hemeoglobin, $\varepsilon = 0.1$, $L_0 = \varepsilon^{-4} = 10^4$, we have

(49) $$S_{0.1}'(1) = 1 - \frac{0.3}{(1.1)^2}$$

$$= 1 - \frac{0.3}{1.21} \approx \frac{3}{4}$$

The formula (34) for $S_\varepsilon(x)$ is complicated, and for some purposes it may be useful to have a simpler but approximate formula. A formula that suggests itself is

$$f_n(x) = \frac{x^n}{1+x^n} \tag{50}$$

As we have seen, this gives a perfect match to $S_\varepsilon(x)$ for $\varepsilon = 1$ by setting $n = 1$, and a perfect match for $\varepsilon = 0$ by setting $n = 4$. Also $f_n(1) = \frac{1}{2}$ for all n, just as $S_\varepsilon(x) = 1/2$ for all ε. For $\varepsilon \in (0,1)$, there is no choice of n that makes $f_n(x)$ agree with $S_\varepsilon(x)$ for all x, but we can choose n to match the slopes of these two functions at $x = 1$. By differentiating (50), we get

$$f_n'(x) = \frac{(1+x^n)\,n x^{n-1} - x^n\, n x^{n-1}}{(1+x^n)^2} = \frac{n x^{n-1}}{(1+x^n)^2} \tag{51}$$

$$f_n'(1) = \frac{n}{4} \tag{52}$$

By forcing (52) to agree with (47), we get

$$n = 4\left(1 - \frac{3\varepsilon}{(1+\varepsilon)^2}\right) \tag{53}$$

This indeed gives $(n=1)$ when $(\varepsilon=1)$, and $(n=4)$ when $(\varepsilon=0)$.

When $\varepsilon=0.1$, which is realistic for hemeoglobin, $n \approx 3$. When a function like $f_n(x)$ is used in biochemical kinetics, the parameter n is called a Hill coefficient. As we see from the example of hemeoglobin, the Hill coefficient need not be indicative of the number of sites that are actually involved, and in fact n need not be an integer, since it is just obtained by curve fitting. What a Hill coefficient of n does tell us, however, is that the number of interacting sites must be greater than or equal to n. Even the case of equality is unlikely to occur in nature, since it would require the n sites to operate in an all-or-none manner, and it is hard to imagine how this could work.

Remark on the choice of L_0:

In equation (33), we set $L_0 = \varepsilon^{-4}$, where $\varepsilon = \sqrt{K_R/K_T}$. This is the most symmetrical choice possible, since it makes

$$L_2 = \left(\frac{K_R}{K_T}\right)^2 L_0 = \varepsilon^4 L_0 = 1 \tag{54}$$

This means that when 2 of the 4 sites of hemeoglobin are occupied the R and T states are equally probable, i.e., that

$$[R_2] = [T_2] \tag{55}$$

and it gives the model a kind of skew-symmetry about these middle states, since it implies that

$$\frac{[R_1]}{[T_1]} = \frac{[T_3]}{[R_3]} \tag{58}$$

$$\frac{[R_0]}{[T_0]} = \frac{[T_4]}{[R_4]} \tag{57}$$

The above considerations make it seem likely that this "nice" choice of L_0 is actually optimal in some sense, but we have not stated or solved such an optimization problem.

A possible project that we leave for the reader is to incorporate the regulation of oxygen binding by CO_2 and H_3O^+ into the allosteric model for hemoglobin. This could be done by providing separate binding sites for these species (one or possibly more on each subunit of hemoglobin) with different affinities in the R and T states of the molecule for each of these species at each of its binding sites. As in the case of oxygen, all binding sites for all species should function independently when hemoglobin is locked in either of its two states, so that the sites interact only because of their influence on the $T \leftrightarrow R$ equilibrium. This would model the shift in the hemoglobin curve that occurs between the lungs and the tissues, as discussed above.