Reversible ligand binding

Andrea Bellelli
Dipartimento di Scienze Biochimiche “A. Rossi Fanelli”
Sapienza Università di Roma

The reversible combination of a protein with a low molecular weight ligand is extremely common and important in biology:

- oxygen transport; transport of poorly soluble substances
- combination of hormones and receptors or neuromediators and receptors
- combination of enzymes with their substrates
- antigen-antibody reactions
- chemotaxis; recognition of smell and flavor
- …

The reversible binding between a protein and a ligand obeys the mass action law:

$$P + X \rightleftharpoons PX \qquad K_a = \frac{[PX]}{[P]\,[X]}$$

oppure

$$PX \rightleftharpoons P + X \qquad K_d = \frac{[P]\,[X]}{[PX]}$$

Often the dissociation constant is preferred because it has as dimension a molarity and hence it can be directly compared with the ligand concentration.

The fraction of iganded protein: $[PX] / [P]_{tot}$

A biologically relevant parameter is the fraction of liganded protein; for example a cell might respond to a hormone when at least, say, 20% or 30% of its receptors are liganded. We define:

$[P]_{tot} = [P] + [PX]$

$K_d = ([P]_{tot} - [PX])\,[X] / [PX]$

$[PX]\,(K_d + [X]) = [P]_{tot}\,[X]$

$[PX] / [P]_{tot} = [X] / (K_d + [X])$

The concept of $[X]_{50}$

From the equation: $[PX] / [P]_{tot} = [X] / (K_d + [X])$ we can derive the ligand concentration required to obtain any degree (fraction) of ligation. Typically the relevant or required parameter is the ligand concentration required to achieve the semisaturation of the ligand binding sites ($[X]_{50}$):

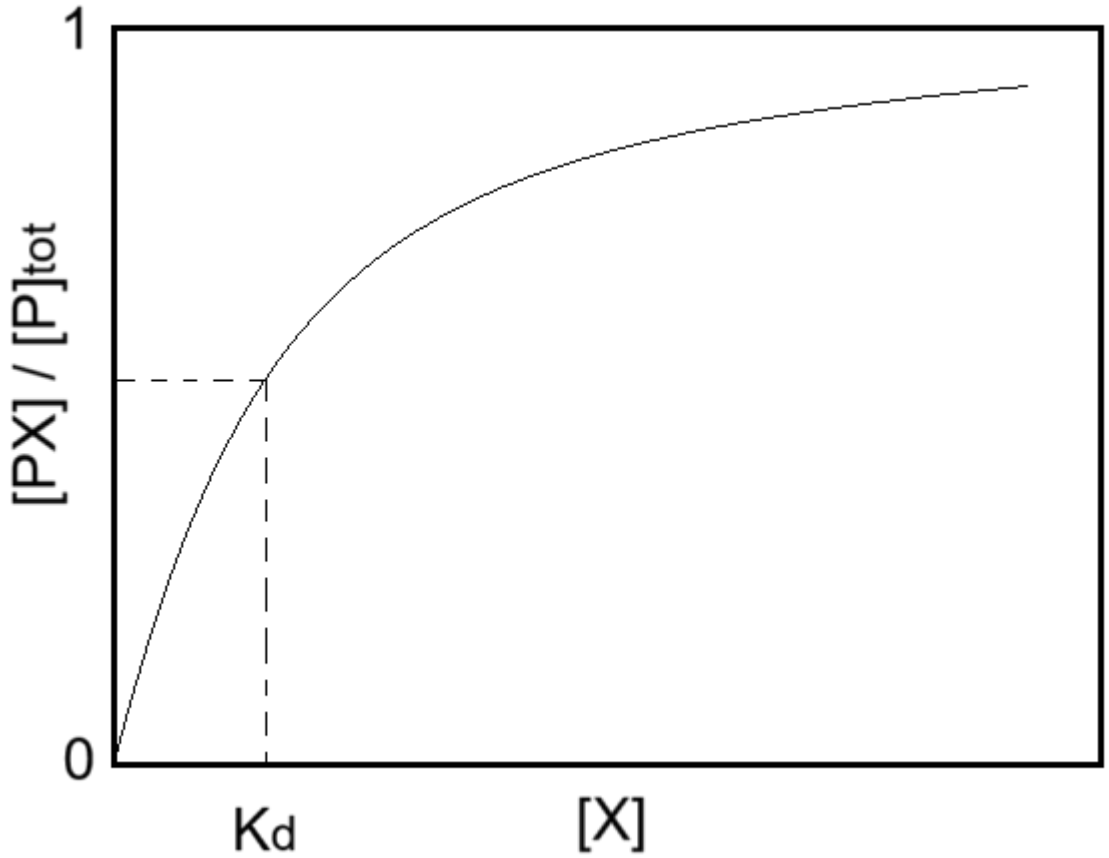
$$[PX] / [P]_{tot} = 0,5 = [X]_{50} / (K_d + [X]_{50})$$

$$2\ [X]_{50} = K_d + [X]_{50}$$

$$[X]_{50} = K_d$$

The ligand concentration required for saturating half the binding sites equals the K_d.

The [PX]/[P]$_{tot}$ versus [X] graph is a hyperbola



How to design the experiment

The measurement of Kd and/or the binding stoichiometry demands one or more experiments. The most common type of experiment relies on the preparation of a protein solution at known concentration to which aliquots of ligand are progressively added.

An instrument capable of detecting a signal proportional to the fraction of liganded protein is required.

Thus we need:

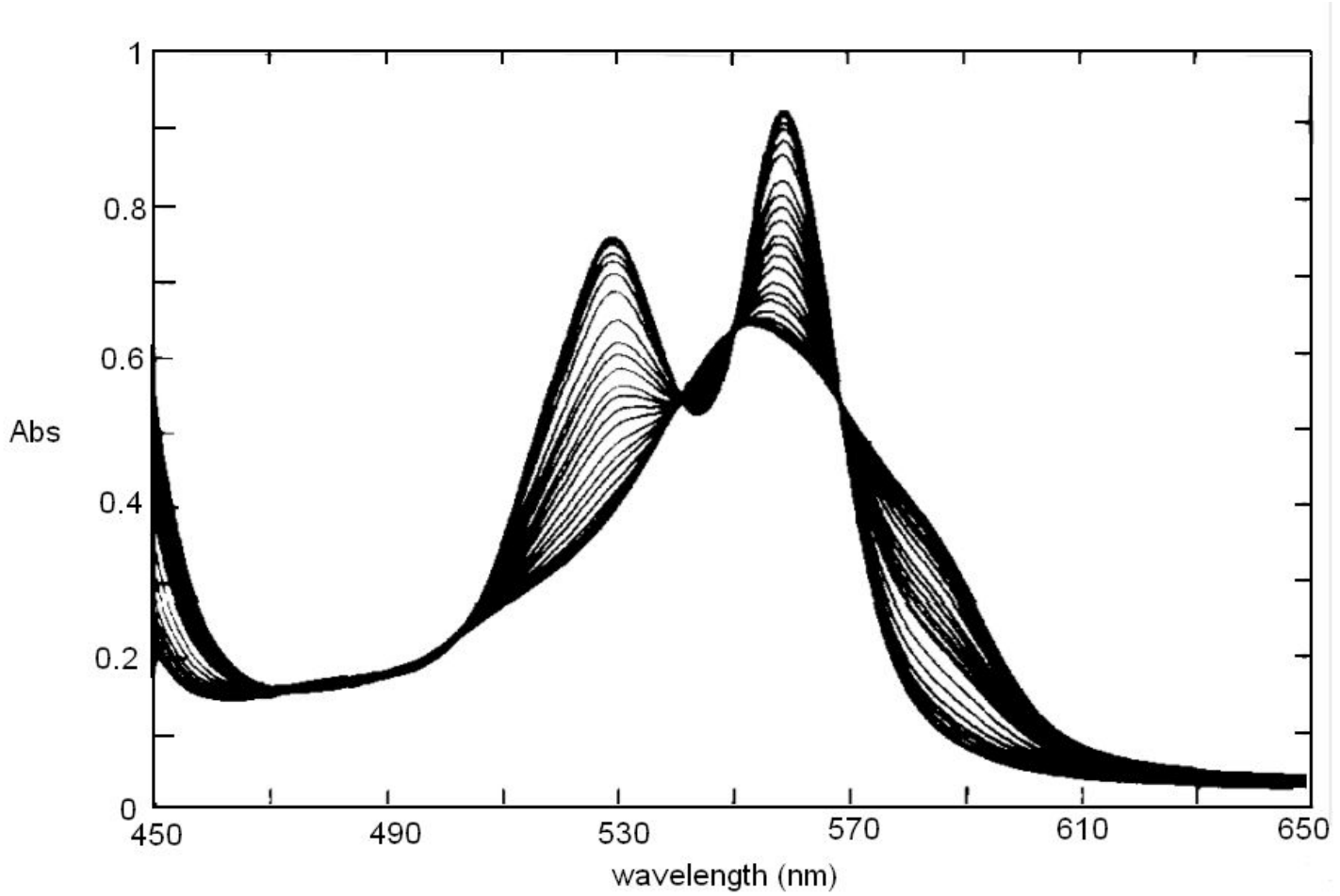
- to decide the concentrations of protein and ligand
- to select the signal we want to detect and the appropriate instrument

Deciding the protein concentration

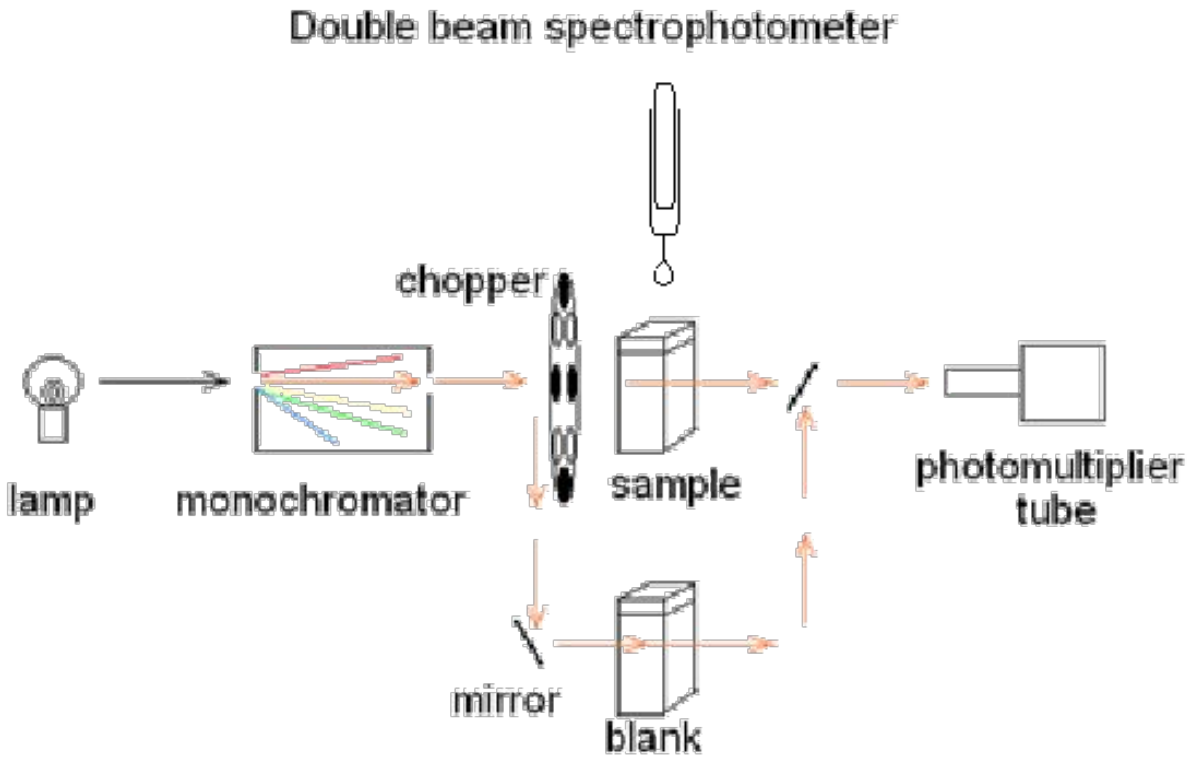
Usually (but not always) the protein is the species that produces the signal; thus the protein concentration is usually determined by the sensitivity of the instrument we use.

For example, if the signal is an absorbance change, and the instrument selected is a spectrophotometer, the optimal concentration should yield a difference in absorbance between 0,2 and 0,8 absorbance units. This, if we use an optical path of 1 cm, is consistent with protein concentrations in the order of 0.1 mM, given the typical molar extinction coefficients of proteins and their cofactors.

Reversible binding of n-propyl isocyanide to human hemoglobin, monitored by absorbance spectroscopy (modified after Reisberg and Olson, J. Biol. Chem. 1980, 255, 4144-4150)



Scheme of a spectrophotometer



Explore a large interval of protein concentration

Because of reasons we are going to discuss later, and given that the ligand concentrations to explore are fixed by its K_d:

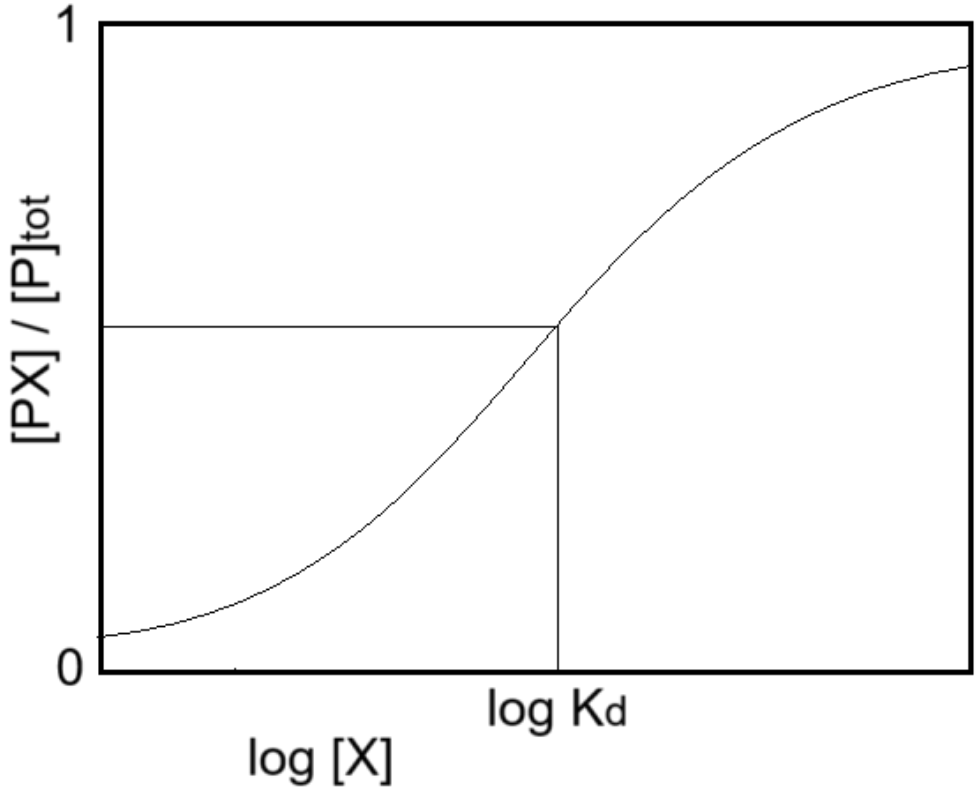
- a low protein concentration is required if we want to measure K_d;
- a high protein concentration is required if we want to determine the reaction stoichiometry;
- using several methods to detect the binding of the ligand allows us to extend the accessible range of protein concentration; for example our reaction could yield an absorbance signal and a fluorescence signal, and if we use both we take advantage of the fact that fluorescence requires a two order of magnitude lower protein concentration than absorbance.

Deciding the ligand concentration

The ligand concentration to be explored depends on K_d. If we plot $[PX] / [P]_{tot}$ versus $\log [X]$ we obtain a symmetric sigmoid curve centred on the point with coordinates:

$$\log [X]_{50} = \log K_d$$

$$[PX] / [P]_{tot} = 0.5$$



The ligand concentration to explore is between 1/20 of K_d (which yields $[PX]/[P]_{tot} = 0,05$) and 20 times the K_d (which yields $[PX]/[P]_{tot} = 0,095$).

Above and below these concentrations we would obtain unreliable values of $[PX]/[P]_{tot}$ because the experimental error would be large compared to the signal change.

Moreover it is advisable to space the ligand concentrations over a logarithmic scale because in this interval the $[PX]/[P]_{tot}$ vs. log [X] plot is nearly linear; for example we could obtain 6 experimental determinations with the following values of ligand concentration: 0,05 K_d; 0,18 K_d; 0,65 K_d; 2,3 K_d; 8,4 K_d; e 20 K_d. Each value, except the last one equals the preceding one multiplied by 3.6.

Obviously, we could have no previous estimate of Kd and could not design the required additions of ligand.

In this case it is advisable to run a preliminary experiment to explore a larger ligand concentration interval, always using logarithmically spaced intervals; for example we could explore an interval from nanomolar to millimolar using 7 ligand additions spaced by a factor of 10 each. Upon observing a large signal change we could also switch to smaller changes of ligand concentration , e.g. spaced by a factor between 2 and 4, in order to obtain more experimental points in the region of interest.

Free ligand versus total ligand

The term [X] that appears in the mass law action:

$$[PX] / [P]_{tot} = [X] / (K_d + [X])$$

represents the concentration of the free ligand; total ligand being:

$$[X]_{tot} = [X] + [PX]$$

Do we know the free ligand concentration or do we know the total ligand concentration? Or both?

Let us consider three possible cases.

1) We know (measure) the free ligand concentration

In some cases we can precisely measure the free ligand concentration (and perhaps also that of the bound ligand), e.g. because we use a radioactive ligand or because we have a suitable electrode.

For example, if we use a radioactive ligand, we can measure the total ligand concentration ($[X]_{tot}$). We then can precipitate the PX complex and measure the free ligand concentration in solution ([X]). Since we know the total protein concentration, we can easily calculate the fractional saturation:

$$[PX] / [P]_{tot} = ([X]_{tot} - [X]) / [P]_{tot}$$

2) In molti casi possiamo calcolare [X]

Se conosciamo soltanto $[P]_{tot}$ e $[X]_{tot}$ possiamo riscrivere la legge di azione delle masse per questi due parametri ed ottenere una equazione di secondo grado:

$$[PX] / [P]_{tot} = ([X]_{tot} - [PX]) / (K_d + ([X]_{tot} - [PX]))$$

$$-[PX]^2 + [PX] (K_d + [X]_{tot}) = [P]_{tot} [X]_{tot} - [P]_{tot} [PX]$$

$$[PX]^2 - [PX] (K_d + [X]_{tot} + [P]_{tot}) - [P]_{tot} [X]_{tot} = 0$$

Questa soluzione è applicabile solo se $[X]_{tot} \geq 2 [PX]$; in caso contrario l'errore sperimentale su $[X]_{tot}$ può portare a risultati inaffidabili.

3) $[P]_{tot}$ è molto bassa rispetto a $[X]_{tot}$

Se $[P]_{tot} \ll [X]_{tot}$ possiamo trascurare [PX] e scrivere direttamente:

$$[X] \approx [X]_{tot}$$

Questa condizione semplifica di molto l'interpretazione dell'esperimento, ma non è sempre applicabile; nelle prossime diapositive vedremo perché.

Relazione tra le concentrazioni di proteina e ligando

In un esperimento nel quale la proteina è mantenuta a concentrazione costante mentre il ligando viene progressivamente aggiunto ed ha quindi una concentrazione variabile:

- la concentrazione di proteina è determinata dal segnale che questa è in grado di generare e dalla sensibilità dello strumento che utilizziamo.
- La concentrazione del ligando è determinata dalla costante di equilibrio della reazione.
- Il rapporto ligando / proteina di per sé non è significativo; però è ovvio che alla fine della titolazione deve verificarsi la condizione: $[X]_{tot} \gg [P]_{tot}$.

Se il nostro scopo è determinare la Kd

Se lo scopo dell'esperimento è determinare la K_d le condizioni ideali sono:

$$[P]_{tot} \ll K_d$$

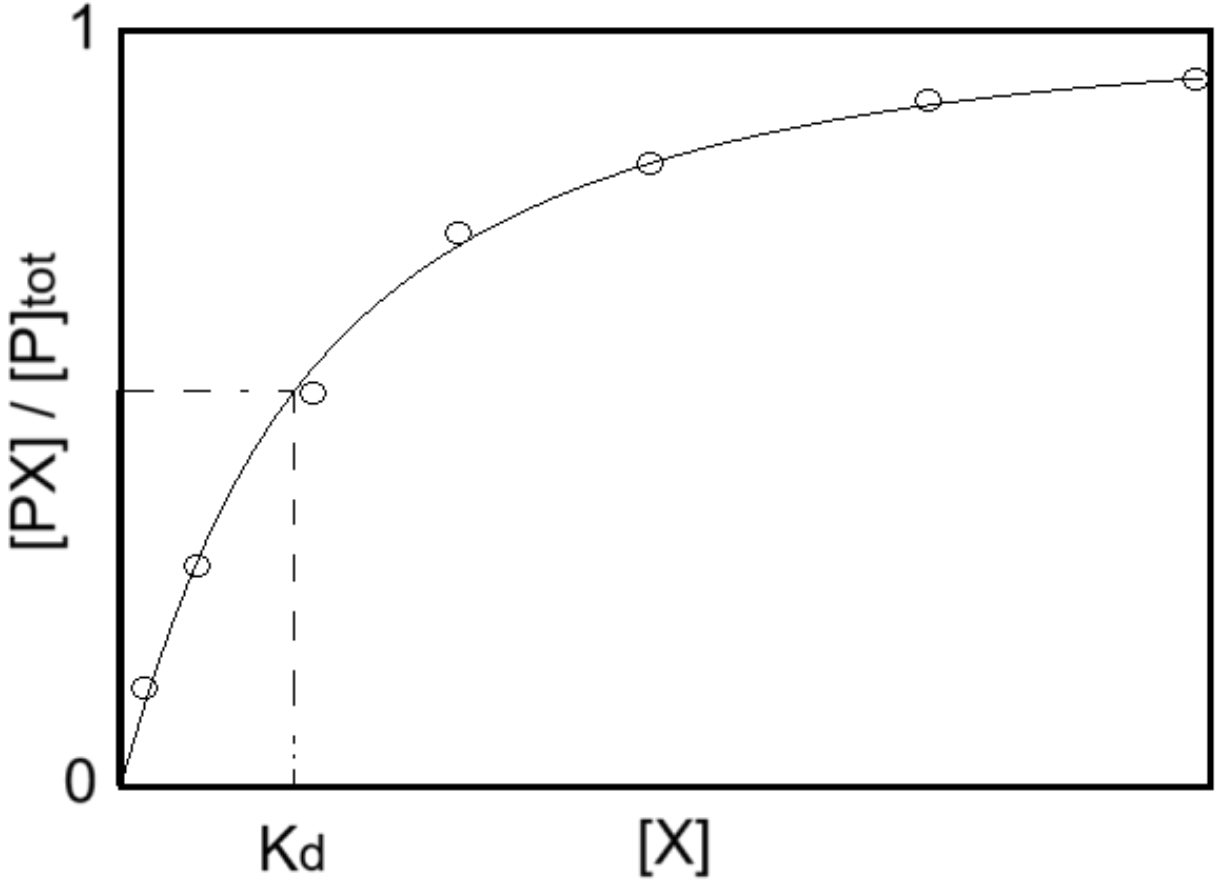
$$0{,}05\ K_d \leq [X]_{tot} \leq 20\ K_d$$

$$[P]_{tot} \ll [X]_{tot}$$

In queste condizioni non possiamo però determinare la stechiometria di reazione: non possiamo cioè distinguere tra:

$$P + X \leftrightarrows PX \quad \text{e} \quad P + 2X \leftrightarrows PX_2$$

L’esperimento verrà in questo modo



Se il nostro scopo è determinare la stechiometria

Se lo scopo dell'esperimento è determinare la K_d le condizioni ideali sono:

$$[P]_{tot} \gg K_d$$

$$0 \leq [X]_{tot} \leq 2\text{-}5\ [P]_{tot}$$

$$[P]_{tot} \approx [X]_{tot}$$

In queste condizioni noi avremo $[X] \ll [PX]$ e potremo assumere

$$[PX] \approx [X]_{tot}$$

Poiché [X] è trascurabile rispetto a $[X]_{tot}$, non sarà possibile determinare la K_d.

L’esperimento verrà in questo modo

Titolazione dell’emoglobina con il CO. Il CO libero è trascurabile. Anderson e Antonini, J. Biol. Chem. 243, 2918-2920.

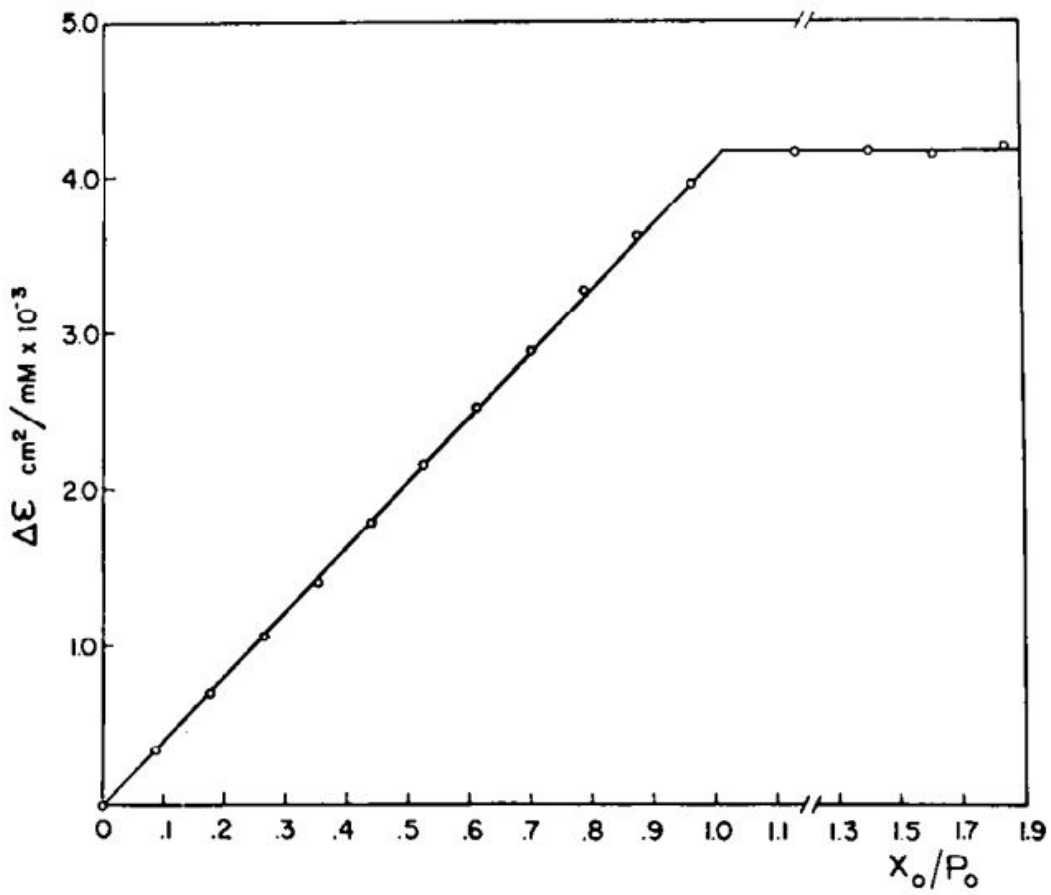


FIG. 2. Stoichiometric titration of human hemoglobin with CO. The changes in molar absorptivity were measured at 538 mμ Conditions: 6.71×10^{-6} M hemoglobin in 0.10 M phosphate, pH 7.0 (25°).

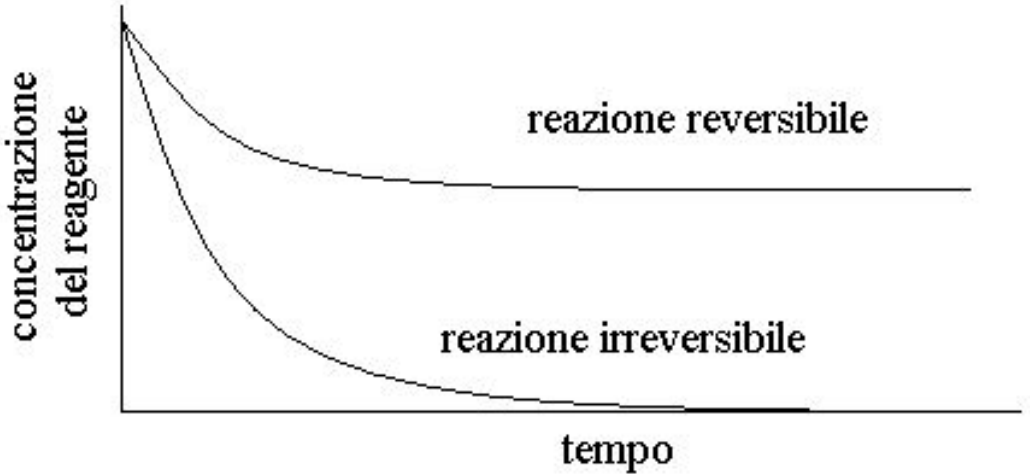
PROBLEMS !

Several problems may become apparent during our carefully planned experiment; or, worse still, several problems may be present and go undetected, leading to wrong estimates of the parameters we intend to measure. Let’s consider:

- Slowly binding ligands
- Low affinity ligands
- High affinity ligands
- Poor signal or signal to noise ratio

Slowly binding ligands

The rate of approach to equilibrium after each addition of ligand to the protein solution under the optimal condition $[X]_{tot} \gg [P]_{tot}$ is exponential or nearly exponential:



How long should I wait to reach equilibrium?

The apparent pseudo-first order rate constant of approach to equilibrium, after each ligand addition and under the condition $[X]_{tot} \gg [P]_{tot}$ is given by the sum of binding and dissociation rate constants:

$$k_{app} = k_{on}[X] + k_{off}$$

This rate “constant” is different with each ligand addition because it contains the variable term [X].

The only way to know that the equilibrium condition has been reached after each addition is to repeat the measurement and check that no further binding is occurring.

Definition of low- and high-affinity ligands

In order to obtain a ligand binding isotherm that can be reliably analyzed, we need to explore a ligand saturation range spanning at least from 0.2 to 0.8, corresponding to ligand concentrations from 0.25 K_d to 4 K_d. If we cannot obtain experimental points over the whole range, the analysis will be plagued by large uncertainties.

Low affinity ligands are those ligands that cannot reach the condition $[X] \geq 4\ K_d$.

High affinity ligands are those for which $[X] \leq 0.2\ K_d$ can be reached only under conditions that negate the requirement $[X]_{tot} \gg [P]_{tot}$.

Remember that $[P]_{tot}$ is dictated by the signal, whereas $[X]_{tot}$ is dictated by K_d.

Low-affinity ligands

Low affinity ligands cannot reach the condition $[X] \geq 4\ K_d$.

There may be several reasons for this problem:

- the concentration required exceeds ligand solubility;
- the concentration required entails significant effects on the solution itself (high ionic strength if the ligand is charged; high viscosity; etc.);
- the concentration required perturbs the solution or the protein (e.g. it denatures the protein).

Overcoming the problem of low-affinity

Actually there are not many ways of overcoming the problem of low affinity, except if the binding reaction has a significant ΔH, either positive or negative.

We remember that:

$$\Delta G = \Delta H - T\,\Delta S = - RT \ln K$$

Thus:

$$\ln K = \Delta S/R - \Delta H/RT$$

Taking advantage of the ΔH

The equation $\ln K = \Delta S/R - \Delta H/RT$ tells us that a plot of $\ln K$ versus $1/T$ yields a straight line with intercept $\Delta S/R$ and slope $-\Delta H/R$ (van t'Hoff plot).

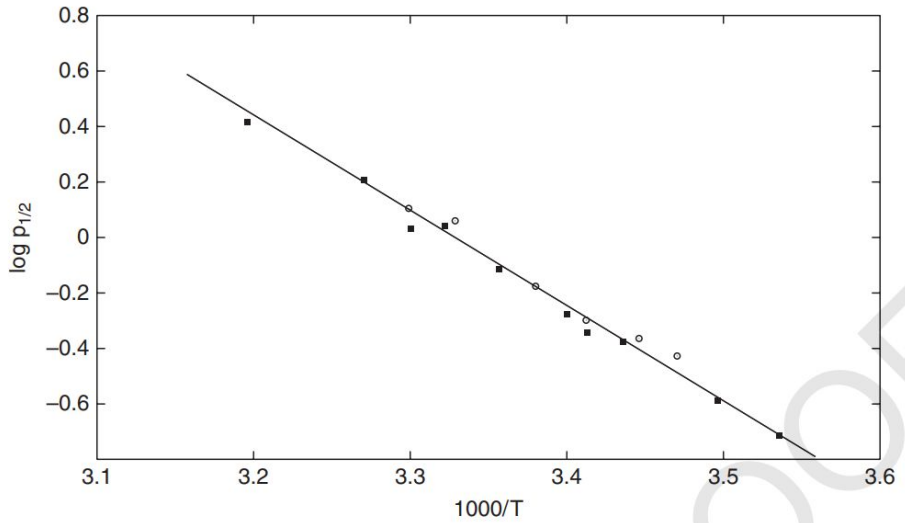


Figure 1.4 van't Hoff plot of the reaction of sperm whale myoglobin with oxygen. The ordinate reports the logarithm of the oxygen partial pressure required to half saturate the protein ($P_{1/2}$) that for a gaseous ligand is equivalent to log $\mathbf{X}_{1/2}$. *Source*: Adapted from Antonini and Brunori, 1971; original data by G. Amiconi.

Extrapolation of K_d

In conclusion, to overcome the problem of low-affinity ligands we can carry out a series of measurements at lower temperatures than the one we are interested in (if $\Delta H > 0$) or at higher temperatures (if $\Delta H < 0$) and linearly extrapolate the value of the K_d at the temperature we are interested in.

High-affinity ligands

In the case of high-affinity ligands, we may adopt the same strategy as for low-affinity ligands and determine the K_d over a temperature range that decreases ligand affinity, followed by extrapolation to the desired temperature.

However, in this case we also have a different trick up of our sleeve: ligand replacement reactions.

Replacement reactions

A ligand replacement reaction takes advantage of the competition between two ligands for the same binding site of the protein:

$$PX + Y \leftrightarrows P + X + Y \leftrightarrows PY + X$$

If at least one of the ligands has high affinity, the concentration of the unliganded protein will be negligible and we can semplify to:

$$PX + Y \leftrightarrows PY + X$$

An obvious example of a replacement reaction is that of CO binding to oxy-hemoglobin:

$$HbO_2 + CO \leftrightarrows HbCO + O_2$$

The ligand partition constant

If we carry out our experiment starting from a complex of the protein with the ligand with lower affinity, which is kept at constant concentration, and gradually add the ligand with higher affinity, we obtain a partition reaction, for which the mass law action dictates:

$$PX + Y \leftrightarrows PY + X$$

$$K_p = [PY]\,[X] / [PX]\,[Y]$$

If Y is the lower affinity ligand, kept at constant concentration, we rewrite:

$$K_p' = K_p\,[Y] = [PY]\,[X] / [PX]$$

From the partition constant to the affinity constant

Since we treat sour systems for dissociation reactions, the equation:

$$K_p' = K_p [Y] = [PY] [X] / [PX]$$

implies that we can increase the value of K_p' until its is within the range we desire by increasing the (fixed) concentration of the lower affinity ligand Y, and collect a series of ligand binding isotherms at different values of [Y]. From a set of K_p' versus [Y] pairs, we can extrapolate the value of K_p. and calculate $K_{d,Y}$ from the relationship:

$$K_p = K_{d,\,Y} / K_{d,\,X}$$

To summarize: high-affinity ligands

In practice, to measure the dissociation constant of a high-affinity ligand using the replacement method we proceed as follows:

- select a competitive ligand (Y) with measurable affinity;
- measure the K_d of this ligand ($K_{d,Y}$);
- run a series of replacement reactions at different concentrations of Y and measure a series of K_p' values;
- extrapolate K_p
- calculate $K_{d,Y} = K_p / K_{d,X}$

Poor signal or signal-to-noise ratio

What if the binding of the selected ligand does not yield a good signal or yields no signal at all?

Obviously we can select an artificial ligand that provides signal, e.g. a radioactively labelled ligand.

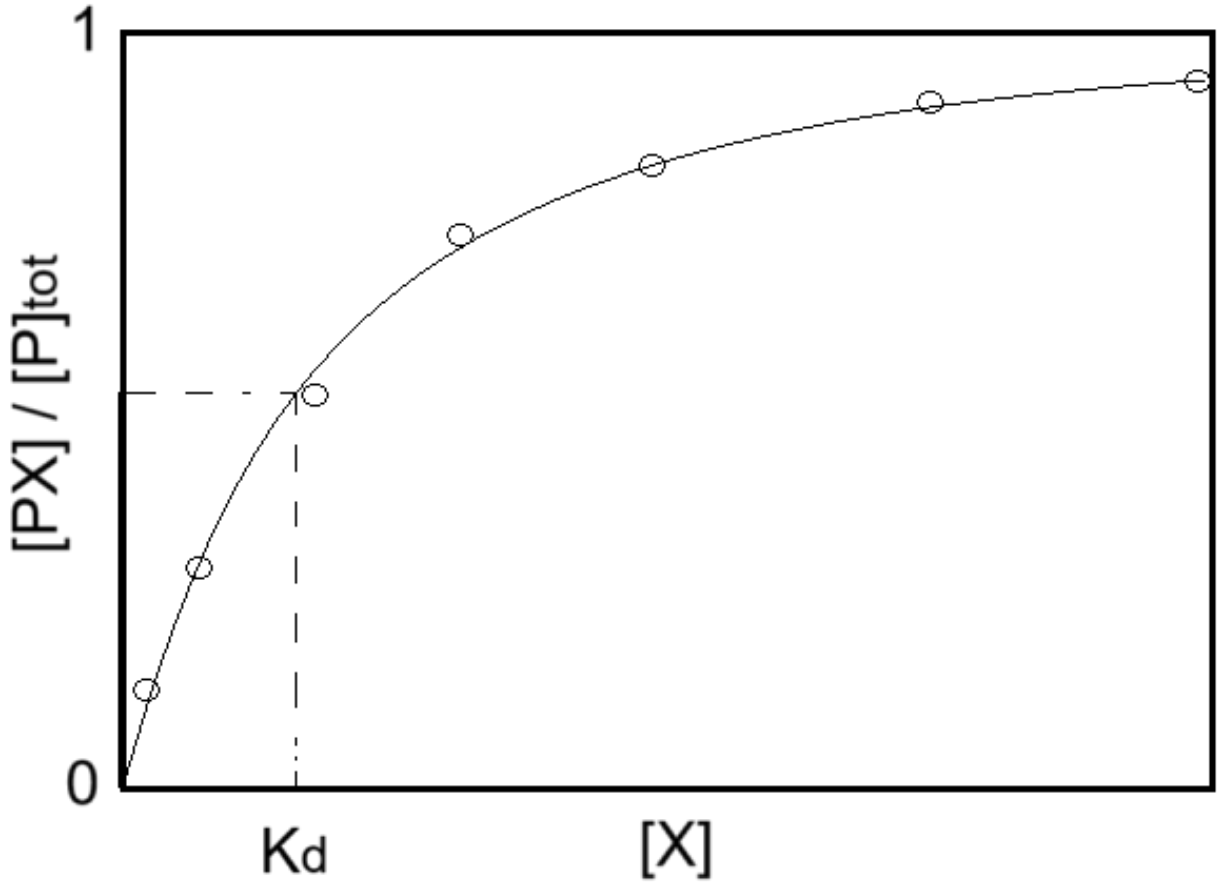
However, we can take advantage of the replacement reactions described previously, provided that we have a competing ligand (Y) that yields a good signal. We proceed as follows:

- record the affinity of Y (measure $K_{d,\,Y}$) taking advantage of the signal provided by the transition $P \rightarrow PY$
- carry out a replacement reaction recording the transition $PY \rightarrow PX$; since P is similar to PX the signal will be due to the dissociation of Y.

Data analysis

Let's suppose that you have collected a set of data; what next?

The first thing one can do is plotting the signal change as a function of $[X]_{tot}$ and draw a tentative curve on the experimental data, to obtain an estimate of K_d.



There must be better methods!

Clearly, drawing a hyperbola by hand is a difficult task; may be the curve we have drawn is not a hyperbola at all! In this case the Kdwe have estimated is a very rough approximation of the “real” value.

Methods have been devised to linearize the hyperbola, under the assumption that we are much better off if trying to draw a straight line through the experimental points, using a ruler.

The simplest linearization is to transform $K_d = ([P]_{tot} - [PX])\,[X] / [PX]$

into $[PX] / ([P]_{tot} - [PX]) = [X] / K_d$

This yields a straight line crossing the origin of axes, with slope $1/K_d$.

The Hill plot

A more refined linearization is due to the Nobel laureate A.V. Hill who proposed a logarithmic transformation of the previous equation:

$$\log ([PX] / ([P]_{tot} - [PX])) = \log [X] - \log K_d$$

The Hill plot has unitary slope, unless the protein is multimeric and cooperative, and intercept $- \log K_d$.

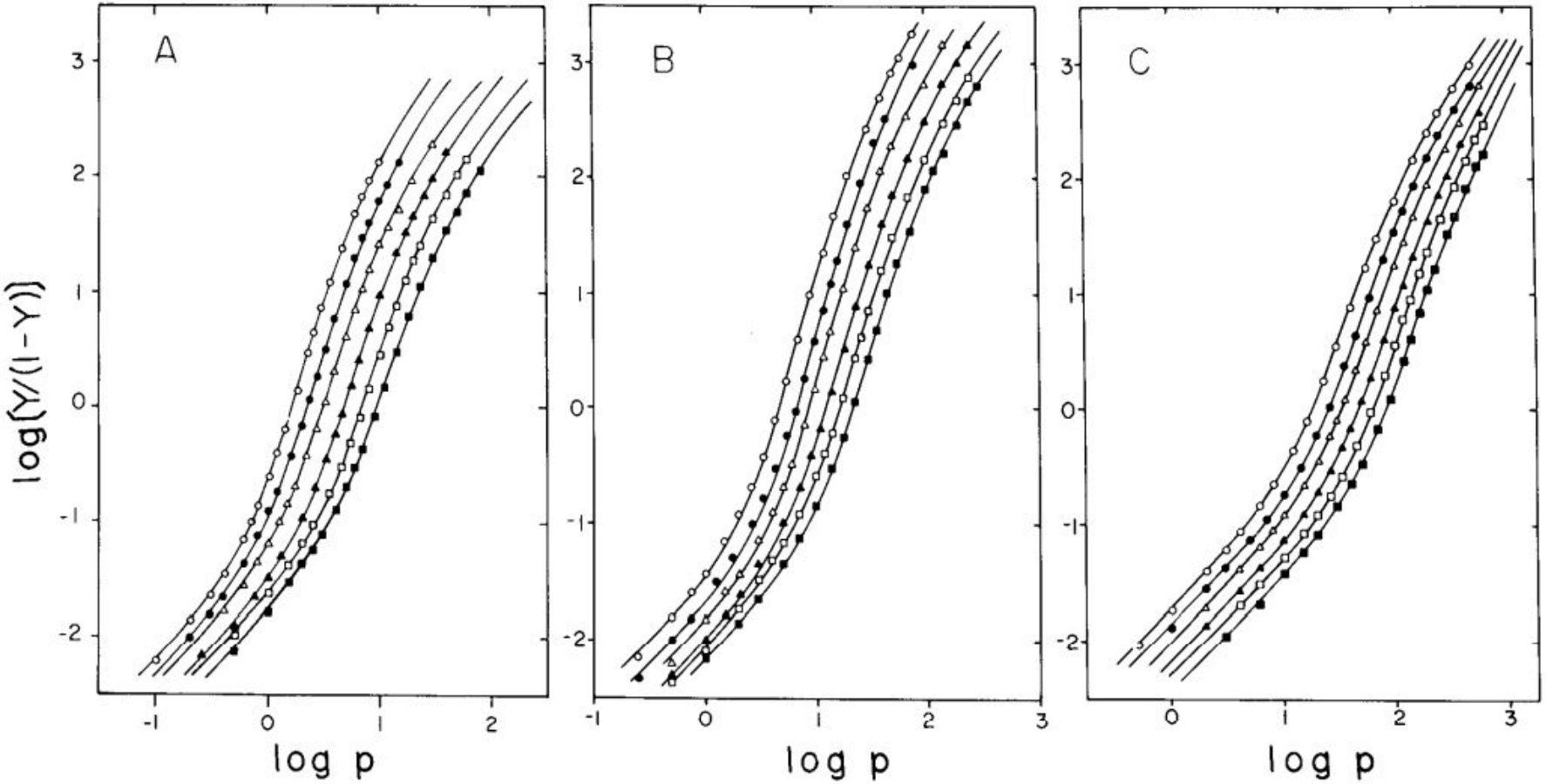


FIG. 1. Hill plots of oxygen binding by hemoglobin. Y, fractional oxygen saturation; p, partial oxygen pressure. Heme concentration, 60 μM; pH 7.4; in 0.05 M bis-tris buffer containing 0.1 M Cl^-. *A*, stripped; *B*, + 2 mM DPG; *C*, + 2 mM IHP. ○, 10°; ●, 15°; △, 20°; ▲, 25°; □, 30°; ■, 35°. Points were experimentally obtained. Lines were calculated from the estimate values of k_i (i = 1,2,3,4) listed in Table I.

The problem of linearizations

Linearizations are not the best way to analyze ones' data. The reason is that the arithmetic transformations required distort the experimental error and the points one obtains do not have all the same precision/error.

Least squares minimization routines

The availability of personal computers has drastically changed the procedures of data analysis. In order to take advantage of computers we need to state the problem we want to solve in the form of an algorithm that the computer can execute. Let's define:

- the experimental points are couples of ([X],signal change) coordinates
- let's call the signal change coordinate Y_{exp}
- we have a ligand binding function capable of calculating a value of Y using [X] and a set of parameters (K_d, total signal change, etc.); let's call this value Y_{calc}
- our task is to find the set of parameters which minimizes the sum of the squared differences between Y_{exp} and Y_{calc}: $\Sigma (Y_{calc} - Y_{exp})^2$

Finding the minimum of the function $\Sigma (Y_{calc} - Y_{exp})^2$ is usually not an exact procedure, i.e. it has no analytical solution; however the computer may calculate this value for a given set of parameters (K_d, total signal change), and vary each parameter slightly to detect in which direction the error function decreases; at the end of the procedure a new set of parameters, yielding a smaller error will be obtained and the procedure can be repeated (iterated), till a minimum is empirically found, or at least a parameter set that changes very little at every successive iteration.