

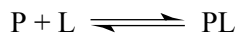
## 2 – PROTEIN-LIGAND ASSOCIATION EQUILIBRIA

A common function of all proteins bestowed with biological activity is the non-covalent binding of specific or non-specific ligands. The association of substrates or diverse effector molecules – inhibitors or activators – to an enzyme, and the association of biologically active molecules to a membrane receptor or to a soluble transporter, are just some of the more familiar examples. The binding of protons,  $H^+$ , to a protein is another particular case of protein-ligand interaction.

The binding of diverse compounds to proteins has been the subject of numerous studies. Such associations lie within the scope of the theory of multiple equilibria developed by VON MURALT in 1930. The studies presented by SCATCHARD (1949), KLOTZ (1953), EDSALL and WYMAN (1958) are among the most classic. Different cases of protein-ligand association equilibria will be discussed, including those for proteins that possess one or more binding sites, either independent or dependent, and equivalent or non-equivalent. These equilibria obey the law of mass action (see Chap. 1).

### 2.1. PROTEINS POSSESSING A SINGLE LIGAND-BINDING SITE

When a protein P possesses only a single binding site for a ligand L, for example a monomeric enzyme having a single substrate-binding site, the association takes place according to a simple equilibrium:



with the association constant: 
$$K_{eq} = \frac{(PL)}{(P)(L)}$$

and the standard free-energy change,  $\Delta G_0$ , for the formation of the complex:

$$\Delta G_0 = -RT \ln K_{eq}$$

In order to determine the equilibrium constant, it is sufficient to measure the concentration of free ligand when equilibrium is reached:

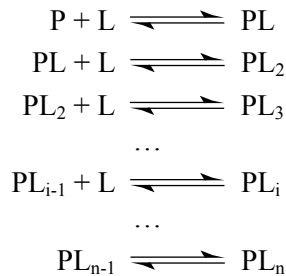
$$K_{eq} = \frac{(\text{bound ligand})}{(\text{protein-bound ligand})(\text{free ligand})}$$

and: 
$$\text{total ligand} = \text{bound ligand} + \text{free ligand}$$

By simply determining the free ligand concentration at equilibrium we can obtain the association constant. The experimental methods used to achieve this are described further in the chapter.

## 2.2. PROTEINS POSSESSING SEVERAL EQUIVALENT AND INDEPENDENT SITES

Now let us consider the most general case where a protein possesses  $n$  sites that are equivalent and independent. This is frequently encountered with oligomeric enzymes having a substrate-binding site on each protomer. There is a system of multiple equilibria, such that:



with the following association constants:

$$K_1 = \frac{(PL)}{(P)(L)}; K_2 = \frac{(PL_2)}{(PL)(L)} \dots K_i = \frac{(PL_i)}{(PL_{i-1})(L)} \dots K_n = \frac{(PL_n)}{(PL_{n-1})(L)}$$

Since all the sites are equivalent and independent, a ligand binding to one site can be defined by a single **microscopic or intrinsic association constant,  $K$** , which is the same for all receptor sites. The successive equilibrium constants,  $K_1, K_2, K_3 \dots K_i \dots K_n$ , only differ from the microscopic constant,  $K$ , by a probability factor; the probability of binding to the first molecule being different from binding to the second and so on. Thus, we have the relationships:

$$\begin{aligned}
 K_1 &= nK \\
 K_2 &= \frac{(n-1)}{2} K \\
 &\dots \\
 K_i &= \frac{[n-(i-1)]}{i} K \\
 &\dots \\
 K_n &= \frac{K}{n}
 \end{aligned}$$

Indeed, there are  $n$  different possibilities for the binding of the first ligand molecule and only one for the dissociation of the complex  $PL$  (from where the probability factor  $n$  for the first equilibrium is derived). There are  $(n - 1)$  ways to bind the second ligand molecule and two ways to dissociate the complex  $PL_2$ , which gives the probability factor  $(n - 1)/2$  for the second equilibrium and so on.

The average number of sites on the protein molecule occupied by the ligand may be defined as:

$$\bar{v} = \frac{\sum_{i=1}^n iPL_i}{\sum_{i=0}^n PL_i}$$

We can expand this expression in terms of the individual species:

$$\bar{v} = \frac{(PL) + 2(PL_2) + 3(PL_3) + \dots + n(PL)_n}{(P) + (PL) + (PL_2) + \dots + (PL_n)}$$

Introducing the different equilibrium constants:

$$\bar{v} = \frac{K_1(L) + 2K_1K_2(L)^2 + \dots + nK_1K_2 \dots K_n(L)^n}{1 + K_1(L) + K_1K_2(L)^2 + \dots + K_1K_2 \dots K_n(L)^n}$$

We then substitute the equilibrium constants for their values, which are functions of the microscopic constant multiplied by the probability factor:

$$\bar{v} = \frac{nK(L) \left[ 1 + (n-1)K(L) + \dots + \frac{(n-1)(n-2) \dots 1 \times K^{n-1}(L)^{n-1}}{2 \times 3 \dots (n-1)} \right]}{1 + nK(L) + \frac{n(n-1)K^2(L)^2}{2} + \dots + \frac{n(n-1)(n-2) \dots 1 \times K^n(L)^n}{2 \times 3 \dots n}}$$

Both numerator and denominator contain binomial series expansions, so:

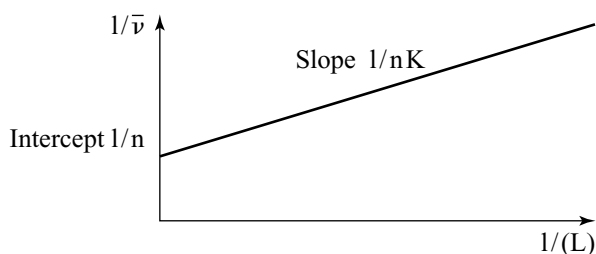
$$\bar{v} = \frac{nK(L)[1 + K(L)]^{n-1}}{[1 + K(L)]^n}$$

giving the relationship:  $\bar{v} = \frac{nK(L)}{1 + K(L)}$

The value  $\bar{v}$  is experimentally measurable. This relationship corresponds to a hyperbola and can therefore be linearised. Two methods of linearisation have been suggested, one by KLOTZ and the other by SCATCHARD. The KLOTZ expression is written thus:

$$\frac{1}{\bar{v}} = \frac{1}{n} + \frac{1}{nK(L)}$$

The KLOTZ or inverse plot involves plotting  $1/\bar{v}$  as a function of  $1/(L)$ . This gives a straight line with slope  $1/nK$  that intersects the y-axis at  $1/n$  (Fig. 2.1). The x-axis intercept gives the value  $-K$ . This representation requires a sufficient number of points corresponding to high ligand concentrations in the zone close to saturation, which is not always experimentally feasible. Furthermore, experimental precision in this zone is not reliable. An imprecise value for the extrapolated point of intersection on the vertical-axis will lead to quite a large error for  $n$  since it is its inverse: the larger the value of  $n$ , the smaller the value of  $1/n$  and the greater will be the risk of error. This representation is practically no longer used these days.

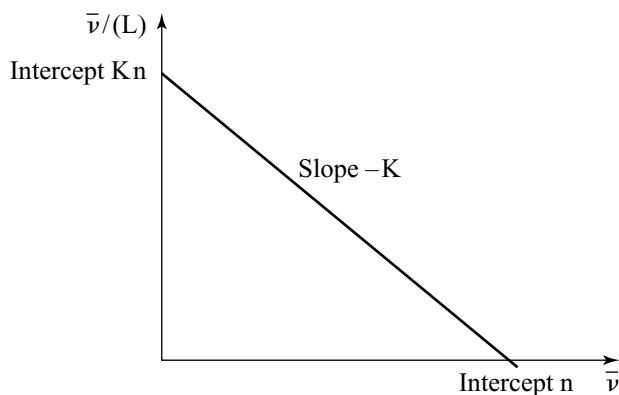


*Fig. 2.1 KLOTZ plot*

The SCATCHARD equation is expressed as follows:

$$\frac{\bar{v}}{(L)} = K(n - \bar{v})$$

The SCATCHARD plot, where  $\bar{v}/(L)$  is plotted as a function of  $\bar{v}$  is frequently used for ligand-binding studies involving soluble or membrane proteins. Figure 2.2 shows the linear relationship between  $\bar{v}/(L)$  and  $\bar{v}$ . The horizontal-axis intercept gives  $n$ , the number of sites, and from the slope we obtain  $K$ , the intrinsic association constant. The vertical-axis intercept is equal to  $nK$ . This diagram enables greater precision in the estimation of the parameters  $n$  and  $K$ , which are obtained directly, rather than from their inverses. A sufficient number of points are required, however, to cover a range of ligand concentrations wide enough to include the value of  $K$ .



*Fig. 2.2 SCATCHARD plot*

Equally, we can define a function for the saturation of a protein by the ligand  $Y_L$ :

$$Y_L = \frac{\bar{v}}{n} = \frac{K(L)}{1 + K(L)}$$

During saturation,  $\bar{v}$  varies from 0 to  $n$  while  $Y_L$  varies from 0 to 1;  $Y_L$  represents the degree of saturation with respect to one site.

### 2.3. PROTEINS POSSESSING $n$ INDEPENDENT AND NON-EQUIVALENT SITES

A protein may contain several categories of binding site having different affinities for the ligand, but independent of each other. Let us suppose that  $m$  categories of binding site exist; in each category  $i$  the number of sites  $n_i$ , defined by their intrinsic association constants  $K_i$ , is equivalent and independent. The total number of sites,  $n$ , capable of binding the ligand is:

$$n = \sum_{i=1}^m n_i$$

The average number of sites occupied by the ligand is given by:

$$\bar{v} = \sum_{i=1}^m \frac{n_i K_i(L)}{1 + K_i(L)}$$

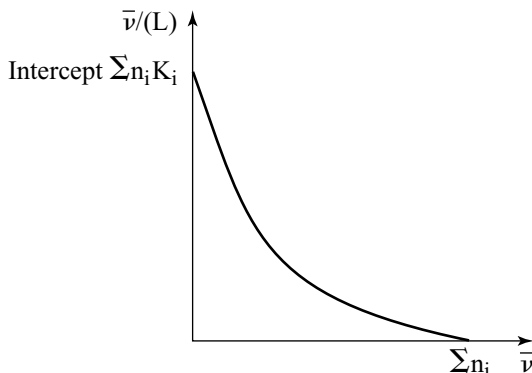
When several categories of site exist, the SCATCHARD plot deviates from linearity (Fig. 2.3). The value at the horizontal-axis intercept is equal to:

$$\sum_{i=1}^m n_i$$

in other words ( $n_1 + n_2$ ), if  $m = 2$ . The value of the vertical-axis intercept is:

$$\sum_{i=1}^m n_i K_i$$

i.e. ( $n_1 K_1 + n_2 K_2$ ) for  $m = 2$ .



**Fig. 2.3 SCATCHARD plot for a protein possessing  $n$  independent but non-equivalent sites**

In the case of a protein that contains two categories of independent site, such that  $n_1$  sites are defined by their microscopic constant  $K_1$  and  $n_2$  sites by their microscopic constant  $K_2$ , the relationship becomes:

$$\bar{v} = \frac{n_1 K_1(L)}{1 + K_1(L)} + \frac{n_2 K_2(L)}{1 + K_2(L)}$$

It is possible, on the condition that the intrinsic constants are sufficiently different, to deconstruct the SCATCHARD plot and estimate the binding parameters. If the constant  $K_1 \gg K_2$  by at least a factor of 50, we may estimate  $n_1$  by the intercept of the line of steepest slope on the horizontal-axis (values of  $\bar{v}$ ) and  $n_1 K_1$  by its vertical-axis intercept. After taking the difference, we can obtain  $n_2$ . By subtracting point by point the first saturation polynomial, we can estimate the parameters  $K_2$  and  $n_2$  using a new plot:

$$\bar{v} - \frac{n_1 K_1(L)}{1 + K_1(L)} = \frac{n_2 K_2(L)}{1 + K_2(L)}$$

The values of  $n_2$  and  $K_2$  can be reintroduced into the equation and after progressive refinement we may obtain the four system parameters with satisfactory precision.

When there are more than two categories of site, or even with only two but where the constants  $K_1$  and  $K_2$  are not sufficiently different, it becomes difficult, indeed impossible, to determine the binding parameters. For a complex system, an approximate solution may be found by curve-smoothing based on an initial hypothesis, starting with the simplest possible. Initial estimates for each parameter are used, which are subsequently refined in an iterative manner eventually revealing the solution.

## **2.4. PROTEINS POSSESSING *N* EQUIVALENT BUT DEPENDENT SITES**

There are proteins that possess  $n$  equivalent sites for which interactions exist between the sites. This dependence may occur for various reasons, for instance as a result of electrostatic interactions or steric effects, where the binding of one molecule interferes with the binding of a second, and so on. Additionally, it may be related to the existence of several conformational states of the protein, either induced or pre-existing, where ligand binding induces a conformational change or shifts the equilibrium between several forms of the protein.

### **2.4.1. EQUIVALENT SITES PRESENTING AN ELECTROSTATIC DEPENDENCE**

When a ligand molecule is electrically charged and its binding to the protein involves electrostatic interactions, the net charge of the protein varies according to

the extent of saturation. Consequently, there is an increase in work necessary to bind a second charged molecule and the law of saturation includes an additional term to represent the electrostatic contribution. The average number of occupied sites per protein molecule is described by the relationship:

$$\bar{v} = \frac{nK(L)e^{-2wZ}}{1 + K(L)e^{-2wZ}}$$

where  $w$  is the electrostatic interaction coefficient, defined by the formula:

$$w = \frac{\epsilon^2}{2DkT} \left( \frac{1}{b} - \frac{\chi}{1 + \chi a} \right)$$

$\epsilon$  is the charge carried by one electron ( $4.8 \times 10^{-10}$  electrostatic units);  $k$ , the BOLTZMANN constant;  $D$ , the dielectric constant of the medium;  $T$ , the absolute temperature;  $b$  is the radius of the protein approximated to a sphere whose charge is distributed uniformly over the surface;  $a$  is the exclusion radius (the combined radius of the protein plus ligand);  $\chi$  is a factor that depends on the dielectric constant of the medium, in water it is equal to  $0.38 \times 10^{-8} \sqrt{\mu}$ ,  $\mu$  being the ionic strength of the aqueous solution. The factor  $2wZ$  represents the electrical work that would be needed to discharge the sphere. In fact, everything proceeds as though the equilibrium constant were related to a macroscopic equilibrium constant extrapolated to zero charge:

$$K = K_0 e^{-2wZ}$$

Due to the existence of electrostatic interactions the SCATCHARD plot deviates from linearity, but using the formula:

$$\frac{\bar{v}}{(L)e^{-2wZ}} = K(n - \bar{v})$$

it is possible to obtain a linear graph by plotting  $\bar{v} / (L)e^{-2wZ}$  as a function of  $\bar{v}$ . In order to do this, both the electrostatic interaction factor and the protein charge must be known.

A classic example of multiple binding equilibria with electrostatic interactions between the sites is the coupling of protons to a category of basic groups on a protein, for example carboxylates or tyrosinates, during acid titration. We may divide all titratable groups on a protein into  $m$  categories of  $n_i$  titratable groups that contribute to electrostatic interactions. If the protein approximates an evenly charged sphere, the expression corresponding to proton binding becomes:

$$\bar{v} = \frac{\sum_{i=1}^m n_i K_i (L) e^{-2wZ}}{\sum_{i=1}^m 1 + K_i (L) e^{-2wZ}}$$

Here, the ligand is the proton;  $K_i$  is the inverse of the ionisation constant of those groups in category  $i$ , where:

$$K_{i,diss} = 1/K_i$$

and:

$$K_{i,diss} = \frac{\alpha_i(H^+)}{(1-\alpha_i)}$$

$\alpha_i$  being the average degree of dissociation:

$$K_{i,diss} = K_{i0,diss} e^{2wz}$$

giving the expression for titration curves, known as the LINDERSTRØM-LANG equation:

$$\log \frac{\alpha_i}{1-\alpha_i} = \text{pH} - \text{p}K_{i0,diss} + 0.865wz$$

## 2.4.2. EQUIVALENT SITES PRESENTING

### STERIC OR CONFORMATIONAL INTERACTIONS

#### 2.4.2.1. PHENOMENOLOGICAL ASPECT

Interactions between sites can lead to an association that is either cooperative or anti-cooperative. For enzymes comprising several subunits and therefore several substrate-binding sites, cooperative effects are generally observed, which result essentially from variations in the conformational state of the protein. Purely steric effects, which would be mainly anti-cooperative, are rare in enzymatic systems and for proteins in general. Anti-cooperativity, however, can also arise from conformational effects.

The interactions between ligand-binding sites on a protein have been discussed by many authors: HILL (1910), ADAIR (1925 and 1949), WYMAN (1948 and 1964), SCATCHARD (1949), (NOZAKI et al., 1957), EDSALL and WYMAN (1958), in terms of their phenomenological aspect; the allosteric models suggested later will be discussed in Part V.

In the case where interactions exist between sites, the events take place as if the microscopic constant varies as a function of the degree of saturation of the protein. We can define an apparent microscopic constant that is actually variable:

$$K' = \frac{\bar{v}}{(n-\bar{v})(L)}$$

$K'$  varies as a function of  $\bar{v}$ . As for electrostatic interactions, we can write:

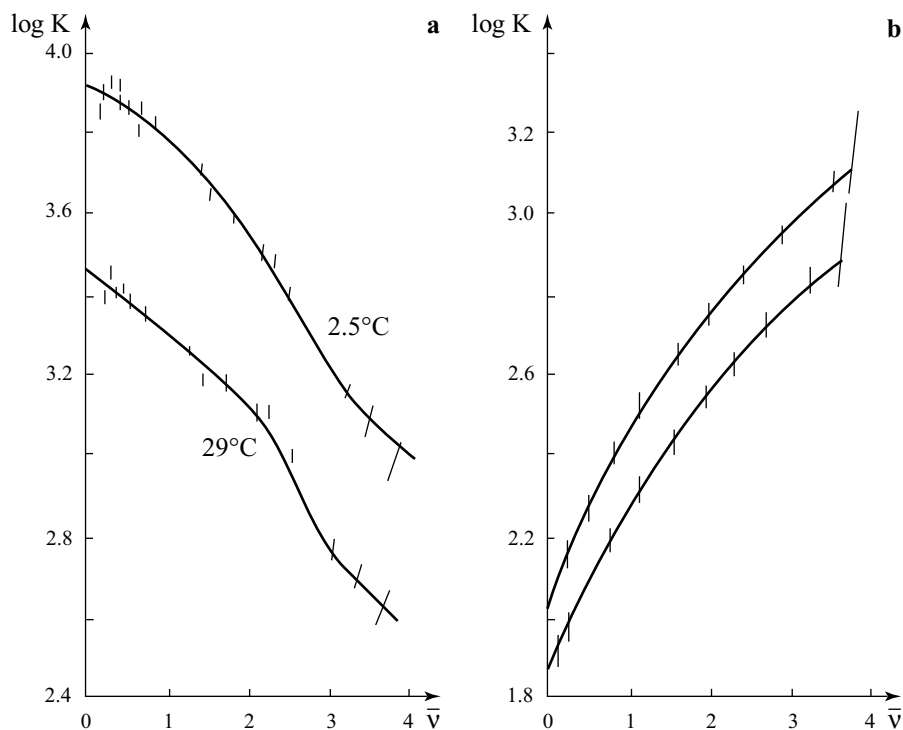
$$K' = K_0 e^{-f(\bar{v})}$$

If  $f(\bar{v})$  is an increasing function of  $\bar{v}$ ,  $K'$  increases gradually with saturation and there is cooperativity. Conversely, if  $f(\bar{v})$  is a decreasing function, there is anti-cooperativity.  $K_0$  is the intrinsic constant extrapolated at zero saturation.



By way of example, let us recall the system studied by NOZAKI et al. in 1957, which was not a biological system but nonetheless a good model for the interactions between ligand-binding sites. Their system concerned the interactions of copper and zinc metal ions with 4-methyl imidazole.

In the case of copper ions,  $K'$  increased as a function of  $\bar{v}$  (cooperative interaction); with zinc ions, this “constant” was reduced (anti-cooperative interaction). Figure 2.4 illustrates these results.

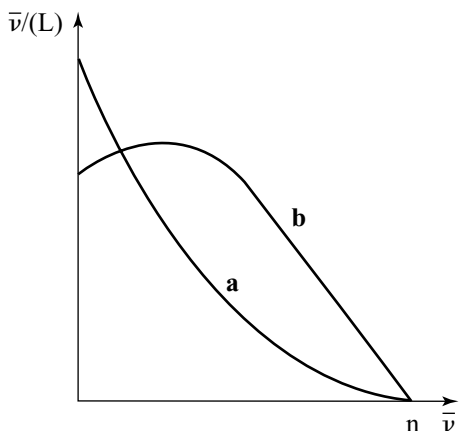


**Fig. 2.4 Interaction of 4-methyl imidazole**  
 (a) with  $\text{Cu}^{++}$ : an example of a negative interaction  
 (b) with  $\text{Zn}^{++}$ : an example of a positive interaction

When the sites are equivalent but show cooperative effects, the SCATCHARD plot deviates from linearity (Fig. 2.5 below). With cooperative binding, the plot is concave; with anti-cooperative binding, it is convex. However, the latter cannot be distinguished from the situation described previously when several categories of site exist. Where cooperativity is observed it is possible, by extrapolating the linear part of the graph, to obtain the number of binding sites and the microscopic constant corresponding to the binding of the first molecule.

If we consider, for example, the case of a tetrameric enzyme that has four binding sites for its substrate (or another specific ligand), the successive equilibrium constants are  $K_1$ ,  $K_2$ ,  $K_3$  and  $K_4$ , and the corresponding microscopic association

constants:  $K_1/4$ ,  $2K_2/3$ ,  $3K_3/2$  and  $4K_4$ . If  $K_1/4 < 2K_2/3 < 3K_3/2 < 4K_4$ , there is cooperativity between the sites. On the contrary, when  $K_1/4 > 2K_2/3 > 3K_3/2 > 4K_4$ , there is anti-cooperativity. In effect, it is as though the microscopic constant gradually increases (cooperativity) or, conversely, decreases (anti-cooperativity) during the course of saturation.



**Fig. 2.5 SCATCHARD plot of a protein possessing equivalent but non-independent sites**  
 (a) anti-cooperative effects  
 (b) cooperative effects

#### 2.4.2.2. INTERACTION ENERGY BETWEEN SITES

The interaction energy between two binding sites can be determined. The apparent free-energy change for the  $i$ th ligand molecule is:

$$\Delta G_0' = -RT \ln K_i$$

The constant  $K_i$  contains a statistical factor, indeed:

$$K_i = \frac{[n - (i - 1)]K}{i}$$

In order to determine the intrinsic free-energy change corresponding to the binding of the  $i$ th ligand the contribution of this statistical factor must be taken into account.

$$\Delta G'_{i,0} = -RT \ln K_i + RT \ln \frac{[n - (i - 1)]}{i}$$

The interaction energy  $\Delta G_{1,i,j}$  between two sites is the difference in association energy of the  $i$ th and the  $j$ th ligand. Thus:

$$\Delta G_{1,i,j} = \Delta G_{j,0} - \Delta G_{i,0}$$

or:

$$\Delta G_{1,i,j} = -RT \ln \frac{K_j}{K_i} + RT \ln \frac{[n - (j - 1)]i}{[n - (i - 1)]j}$$

If the  $j$ th ligand binds more strongly than the  $i$ th ligand, and  $\Delta G_{1,i,j} < 0$ , there is cooperativity. If  $\Delta G_{1,i,j} > 0$ , there is anti-cooperativity. If  $\Delta G_{1,i,j} = 0$ , the sites are equivalent and independent.

### 2.4.2.3. EMPIRICAL EQUATIONS

When interactions exist between sites the saturation function can no longer be represented linearly. The classic representations (see Sect. 2.6.9) have curvilinear profiles. Diverse empirical equations have been proposed to describe this phenomenon; the HILL equation (1910), introduced to account for the cooperative binding of oxygen to haemoglobin, is shown below:

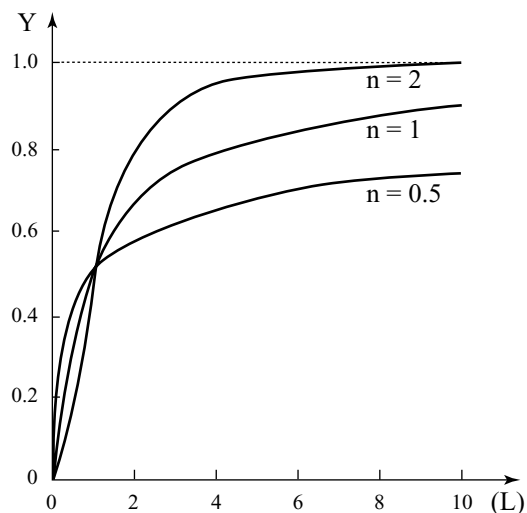
$$\frac{Y_L}{1 - Y_L} = K'(L)^{\alpha_H}$$

where  $\alpha_H$  is the HILL number and  $K'$ , the apparent association constant. The coefficient  $\alpha_H$  reflects the interactions between the sites. If  $\alpha_H > 1$ , there is cooperativity. If  $\alpha_H < 1$ , there is anti-cooperativity. When  $\alpha_H = 1$ , all sites are equivalent and independent.  $\alpha_H$  cannot be greater than the number of sites and is, at most, equal when the cooperativity is optimal. Figure 2.6 shows the form of the saturation curve for different values of  $\alpha_H$ . The HILL number is defined by the equation:

$$\alpha_H = \frac{d \ln[Y_L / (1 - Y_L)]}{d \ln(L)}$$

The HILL equation can be written in logarithmic form:

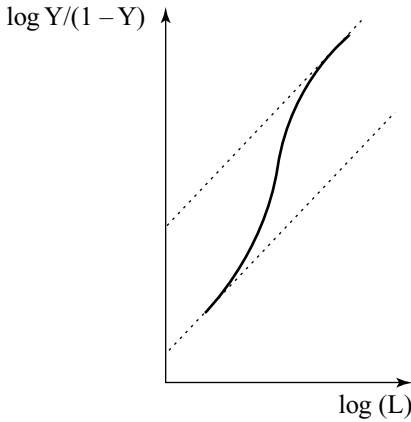
$$\log \frac{Y}{1 - Y} = \alpha_H \log(L) + \log K'$$



*Fig. 2.6 Saturation curves for different values (0.5, 1 and 2) of the HILL coefficient*

By plotting  $\log Y/(1 - Y)$  as a function of  $\log(L)$ , we obtain a curve better known as the **HILL plot**, which is often used to describe phenomena displaying cooperativity or anti-cooperativity. If there is an interaction between the sites, we obtain a curve that has an inflexion point where  $\alpha_H$  is a maximum (Fig. 2.7 below). The slope at the origin is 1; the slope at saturation is also 1. Thus, the HILL plot is a tangent to two straight lines each with a gradient of 1. In Part V (Chap. 13), we will analyse

the HILL plot, in particular the way in which the plot is used to determine the interaction energy between sites.



**Fig. 2.7 HILL plot**

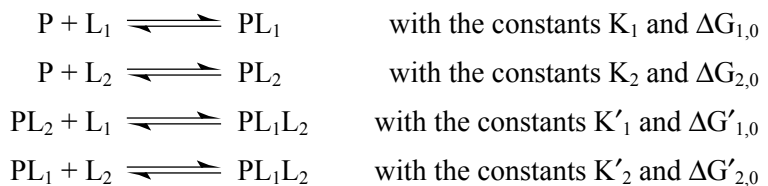
In 1925, ADAIR proposed a general equation to describe cooperative binding; the purpose of this study was also to explain the “abnormal” binding of oxygen to haemoglobin. The ADAIR equation assumes that all sites are equivalent at the beginning, but that the binding of one molecule modifies the intrinsic affinity of a protein for the following molecule, and so on. This amounts to attributing different microscopic constants to the saturation of successive sites on the protein. The ADAIR equation is expressed as follows:

$$\bar{v} = \frac{\sum_{i=1}^n i(L)^i \prod_{j=1}^i K_j}{1 + \sum_{i=1}^n (L)^i \prod_{j=1}^i K_j}$$

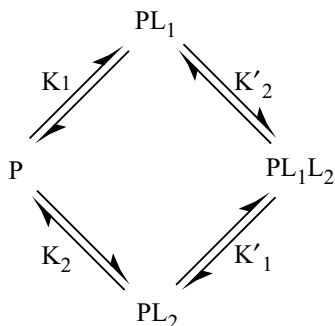
Thus, the binding of oxygen to haemoglobin is described by four microscopic constants, namely  $K_1/4 < 3K_2/2 < 2K_3/3 < 4K_4$ .

## 2.5. LINKED FUNCTIONS

Before embarking on the formalism of a general case, let us consider the simple binding to a protein of two ligand molecules  $L_1$  and  $L_2$ . We have the following equilibria:



which may be written in the following form:



These four equilibria are not independent. **They are linked** by the free energy relationship:

$$\Delta G_{0,1,2} = \Delta G_{0,1} + \Delta G'_{0,2} = \Delta G'_{0,1} + \Delta G_{0,2}$$

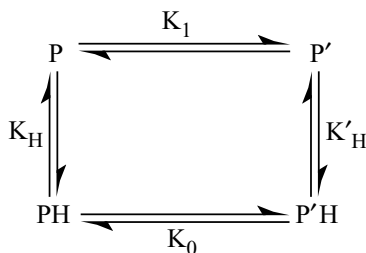
with the constants related as follows:

$$K_1 K'_2 = K_2 K'_1$$

We will have the opportunity to study equilibria of this type in the case of an enzyme having two substrates that bind randomly to their specific sites (see Chap. 5).

An example of linked functions is provided by proteins that change conformation when certain groups are ionised, in particular, groups that possess unusual ionisation pKs.

Let P and P' be two conformations of a protein in an unprotonated form, and PH and P'H the corresponding states when the protein is protonated. We have the following equilibria:



A theoretical treatment of such a schema has been given by TANFORD and co-workers (1960). The apparent constant of the protein's conformational change is:

$$K_{\text{app}} = \frac{(P') + (P'H)}{(P) + (PH)} = \frac{K_0(1 + K_1)}{K_1(1 + K_0)} K'_H = \frac{1 + K_1}{1 + K_0} K_H$$

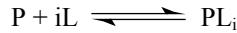
This shows that the conformational change is linked to the protonation of the ionisable group. Reciprocally, the pK of the ionisable group is an apparent pK that

comprises a contribution from the true ionisation constant and from the conformational aspect:

$$K_{H,app} = K_H \frac{1 + K_1}{1 + K_0}$$

These two examples have been developed to give a simple and practical illustration of the idea of linked equilibria.

- Let us now consider the more general case as developed by WYMAN in 1964 when several different ligands bind to a protein. It is useful, first of all, to introduce a few definitions given by WYMAN with respect to the binding of a single ligand type. Let P be a protein capable of binding a ligand L. For the binding of  $i$  ligand molecules, the global equilibrium is:



with the microscopic association constant:

$$K_{Li} = \frac{(PL_i)}{(P_0)(L)^i}$$

and  $P_0$  being the concentration of free protein,  $(L)$ , the concentration of free ligand. The following expression relates the total protein concentration to  $P_0$ :

$$P_{total} = \sum_{i=0}^n (PL_i) = P_0 \sum_{i=0}^n K_{Li} (L)^i = P_0 X_L$$

where  $X_L$  is what WYMAN defined as the **binding polynomial**. A similar type of expression can be applied to the total ligand:

$$L_{total} = \sum_{i=0}^n i(PL_i) = P_0 \sum_{i=0}^n iK_{Li} (L)^i = P_0 \frac{LdX_L}{dL}$$

The saturation function may be written:

$$Y_L = \frac{1}{n} \frac{L_{total}}{P_{total}} = \frac{1}{n} \frac{LdX_L}{nX_L dL}$$

In the presence of two ligands, L and L', we have the equilibrium:



with the apparent microscopic association constant:

$$K_{LiL'_j} = \frac{(PL_iL'_j)}{(P_0)(L)^i (L')^j}$$

The saturation functions for L and L' are, respectively:

$$Y_L = \frac{1}{n} \left( \frac{d \ln X_L X_{L'}}{d \ln L} \right)_L$$

$$Y_{L'} = \frac{1}{m} \left( \frac{d \ln X_L X_{L'}}{d \ln L'} \right)_{L'}$$

and the fundamental relationship for linked functions is written:

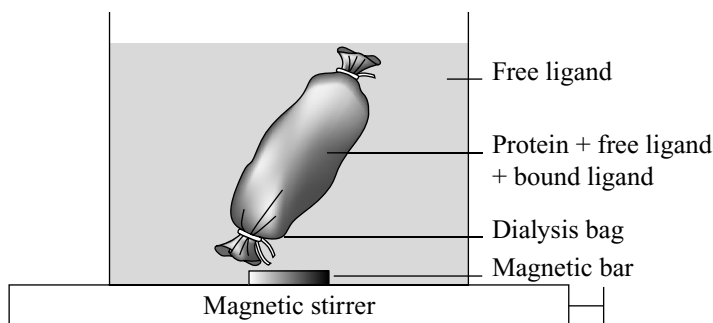
$$n \left( \frac{Y_L}{\ln L'} \right)_L = m \left( \frac{Y_{L'}}{\ln L} \right)_{L'}$$

## 2.6. METHODS TO STUDY LIGAND BINDING

A diverse range of methods is available to measure ligand-binding to proteins. The choice depends essentially on the nature of the ligand, its properties and the change in its properties upon protein binding. In certain cases, the spectral properties of the ligand or the protein (absorption or fluorescence) can be used; alternatively, binding studies may be made possible by employing radioactive ligands. In order to facilitate such experiments, many rapid methods have been developed. However, it is important to bear in mind that some techniques are prone to giving erroneous results, due to the fact that they disturb the thermodynamic equilibrium of ligand association.

### 2.6.1. EQUILIBRIUM DIALYSIS

Dialysis takes place across a semi-permeable membrane whose pores are relatively small and hence permeable to the small ligand molecules, but impermeable to the protein. In general, membranes are produced from cellulose, cellophane or colloidion. These membranes are sold commercially in the form of cylinders, which are closed at each end to create a **dialysis bag** (Fig. 2.8).



**Fig. 2.8** *Equilibrium dialysis experiment*

The protein solution is placed inside the bag, which also contains free ligand and any ligand bound to the protein. Free ligand alone is found in the external solution. When equilibrium is reached, the free-ligand concentration inside and outside the bag is practically the same. In reality, this is not always possible due to the existence of the **DONNAN effect**. Indeed, if we consider a dilute salt solution of NaCl, the

concentrations of  $\text{Cl}^-$  and  $\text{Na}^+$  ions inside and outside of the dialysis bag are not identical because of the existence of charges on the protein. In the interior of the bag (i), the following equality holds true:

$$(\text{Cl}^-)_i + (\text{OH}^-)_i = (\text{P})Z + (\text{Na}^+)_i + (\text{H}^+)_i$$

where (P)Z is the charge on the protein. In the external solution (e), we have:

$$(\text{Cl}^-)_e + (\text{OH}^-)_e = (\text{Na}^+)_e + (\text{H}^+)_e$$

The DONNAN equilibrium law leads to the following formula:

$$\frac{(\text{Na}^+)_i}{(\text{Na}^+)_e} = \frac{(\text{Cl}^-)_i}{(\text{Cl}^-)_e}$$

This effect can be overcome by working at higher ionic strength (for example, 0.1 M NaCl) so that the protein charge becomes negligible relative to the charges present in the solvent.

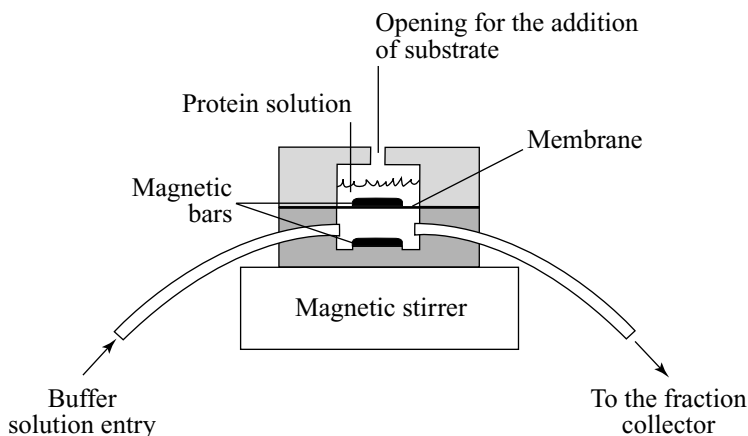
In order to study a ligand-binding equilibrium it is important to choose carefully the protein concentration and to vary the ligand concentration in such a way that there is significant variation in the concentration of free ligand. Clearly, this depends on the equilibrium constant. The use of dialysis bags requires quite large quantities of protein. When there is only a limited quantity of protein available (50–100  $\mu\text{L}$ ), it is beneficial to use small dialysis cells comprising two compartments separated by a semi-permeable membrane.

## 2.6.2. DYNAMIC DIALYSIS

► Equilibrium dialysis is quite a slow process; in order to reach equilibrium completely it may be necessary to wait several hours or, indeed, several days. This experimental requirement is a major limitation when using a particularly unstable protein. COLOWICK and WOMACK (1969) developed a continuous-flow dialysis method or dynamic dialysis, which enables a complete dialysis experiment to be carried out in 20 min. The principle of the method is outlined as follows. A dialysis cell is used whose dimensions have been precisely calculated. The cell contains two compartments separated by a semi-permeable membrane. The lower compartment receives a constant flow of buffer solution, which is directed towards a fraction collector (Fig. 2.9 opposite).

The upper compartment contains the protein in equilibrium with a radioactively labelled ligand. The protein-ligand solution rapidly equilibrates; free radioactive ligand passes into the lower compartment and is carried along by the flow towards the fraction collector. The fractions are then recovered and their radioactivity content measured. Very quickly a steady state is reached wherein the rates of entry and exit of the ligand to and from the lower compartment become equal. Under these conditions, the radioactive ligand concentration in the flow-through is exactly proportional to the concentration of free ligand in the lower compartment.

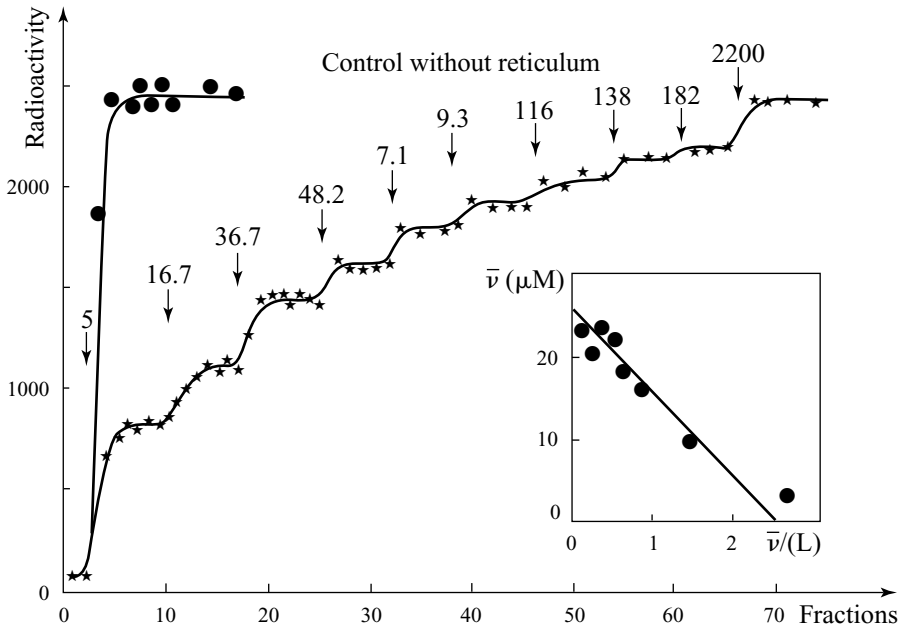




**Fig. 2.9** *Diagram of the apparatus for measuring ligand binding by dynamic dialysis*

The volume of buffer needed to reach this steady state is approximately four times the volume of the lower compartment. The diffusion rate is therefore constant and it remains so for a long time, as the quantity of substrate diffusing is low compared to its concentration in the upper compartment. In the device described by COLWICK and WOMACK the lower compartment has a volume of 2 mL and the flow rate is 8 mL/min; the steady state is therefore reached after one and a half minutes. The radioactive ligand is forced out incrementally with cold ligand, which is introduced in the upper compartment. First of all, ligand alone is placed in the upper compartment in order to determine the proportionality constant between the quantity of radioactive ligand diffusing out and the concentration of free ligand left in this compartment. More precisely, the experiment is performed by placing the enzyme in the presence of the ligand in the same conditions. A first plateau is reached from which the concentration of free ligand in these conditions may be deduced. A small volume of cold substrate solution is then added; a new steady state is reached as indicated by a new plateau and so on until the final plateau is reached.

This should be identical to the plateau measured in a control experiment lacking enzyme in the upper compartment. The corresponding diagram (Fig. 2.10 below) may be used to estimate the concentration of free ligand for each concentration of total ligand, represented by each plateau in the radioactivity level. This simplification results from the fact that the ligand concentration is exactly compensated by the isotopic dilution of the ligand. One consequence of this phenomenon is that, in the absence of enzyme, the addition of a non-radioactive ligand does not affect the plateau. The diagram allows us to obtain directly the free-ligand concentration by dividing each value corresponding to a plateau by the maximum value. From the difference, we can calculate the concentration of bound ligand. In this manner, from the experimental curve we can construct a SCATCHARD plot, from which the intrinsic affinity constant and the number of binding sites may be determined.



**Fig. 2.10 Binding of radioactive ADP [<sup>14</sup>C] to sarcoplasmic reticulum vesicles (shown as stars)**

The control curve without sarcoplasmic reticulum is also shown (black circles). Inset: the binding parameters are determined from the SCATCHARD plot,  $\bar{v} / (L)$  the concentration of ADP bound (in  $\mu M$ ) as a function of the concentration of free ADP. (Reprinted from *Biochimie*, 58, TENU J.P. *et al.*, Binding of nucleotides ATP and ADP to sarcoplasmic reticulum: study by rate of dialysis, 513. © (1976) with permission from Elsevier)

The quantitative aspect of the method when the steady state is reached can be easily demonstrated. If N is the number of radioactive ligand molecules in the lower compartment at time t, after addition of substrate to the upper compartment, we have:

$$\frac{dN}{dt} = S_1 D - N \frac{f}{V} \tag{1}$$

$S_1$  is the concentration of radioactive free ligand in the upper compartment, f the buffer flow rate (the buffer volume that crosses the lower compartment per unit time), V, the volume in the lower compartment, D is a constant that depends on the diffusing molecules and the characteristics of the apparatus. At steady state,  $dN/dt = 0$  and:

$$\frac{N}{V} = S_1 \frac{D}{f} \tag{2}$$

$N/V$  is the concentration of radioactive ligand in the lower compartment.

The time needed to approach the steady state is given by integrating the expression [1]:

$$\frac{N(t)}{V} = \frac{S_1 D}{f} (1 - e^{-(t/V)}) + \frac{N_{(t=0)}}{V} e^{-(t/V)}$$

Thus, when the volume crossing the lower compartment is less than four (or more) times the compartment volume, the exponential term  $e^{-(t/V)}$  is less than 0.018 and the concentration is greater than 98% of the value at steady state.

In order for this experiment to be well controlled quantitatively it is important that the flow rate, the mixing rate and the dimensions of the chamber are accurately calculated and that these rates are maintained perfectly constant.

### **2.6.3. MEASURING PROTEIN-LIGAND INTERACTIONS IN A BIPHASIC WATER-POLYMER SYSTEM**

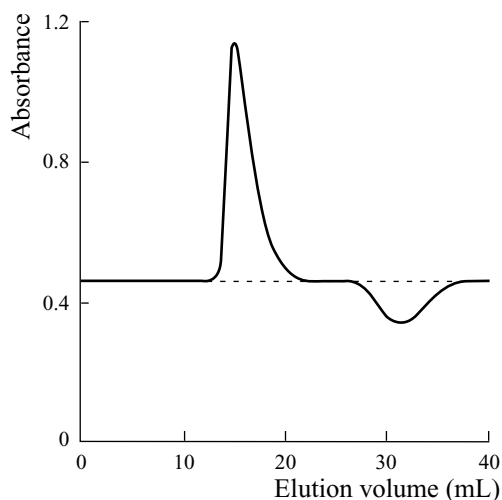
GRAY and CHAMBERLIN (1971) presented an alternative to the method of COLOWICK, with two advantages: speed of execution and the need for only small quantities of protein. The method involves partitioning in a two-phase system in which dextran (MW 500 000) and polyethylene glycol are mixed with the protein and ligand. The two phases are then separated; samples of each phase are removed and the concentrations of protein and ligand are measured in each. Proteins tend to partition preferentially in the dextran phase at weak ionic strength. Small ligands are found in both phases. The equilibrium of ligand and protein partitioning in the two phases is reached after a few minutes or even less. The technique requires only low quantities (microlitre volumes) of reactant; the validity of the method rests on the fact that the ligand molecules bound to protein partition into the same phase as the protein. In order to analyse the results, the fractions of protein and ligand molecules in both phases must be known. To achieve this, precise determinations are carried out prior to ligand-binding experiments. For this study, it is essential that the protein partitions preferentially into a phase in such a way that the concentration of bound ligand is always negligible compared to the free-ligand concentration in the other phase.

The dextran/polyethylene glycol system was chosen because the two phases can be quickly separated and protein molecules partition preferentially in the dextran phase if certain conditions are respected. However, numerous controls must be carried out as this type of technique has two important limitations. The first is that certain proteins, particularly small proteins, do not have a huge tendency to partition into a single phase; theoretical considerations show that this phenomenon depends on the size and on the surface properties of the molecule. The second limitation results from the fact that the environment of a protein in a biphasic system is not the same as in dilute aqueous solutions, in which ligand-binding studies are generally performed. Thus, this environment may affect the association properties of the ligand and protein. ▲

### **2.6.4. SIZE-EXCLUSION CHROMATOGRAPHY**

Molecular sieves, for example Sephadex, which retard the elution of smaller molecules and let larger molecules pass more quickly down a chromatography column, can be used to separate free ligand from bound ligand that is carried along during

the elution of the protein. The column used must be equilibrated with a buffer solution containing the ligand at a given concentration. The protein, in the presence of the same ligand concentration, is then loaded onto the column and eluted by the buffer solution including ligand. The ligand must have a property (radioactivity, absorption band) that allows its total concentration to be measured in the eluate (Fig. 2.11). This process is repeated for each ligand concentration (HUMMEL & DREYER, 1962). The introduction of techniques of high performance liquid chromatography (HPLC) has improved the precision and speed of obtaining results.



**Fig. 2.11** *Elution profile monitored by absorbance at 285 nm following the passage of ribonuclease down a Sephadex G25 column, previously equilibrated with 2'-cytidilic acid*

(Reprinted from *Biochimica Biophysica Acta*, 63, HUMMEL J.P. & DREYER W.J., Measurement of protein-binding phenomena by gel filtration, 530. © (1962) with permission from Elsevier)

### 2.6.5. ULTRAFILTRATION

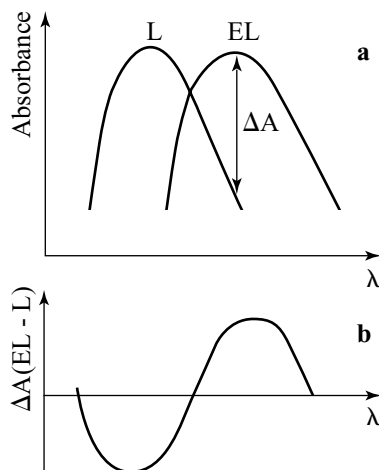
- Ultrafiltration uses the property of certain filters to absorb proteins while free ligand is not retained on them. In principle, it suffices to filter an homogenous, mixed solution of protein and ligand through these filters (e.g. nitrocellulose, semi-permeable membranes etc.). Subsequently, the proportions of ligand retained on the filter and free ligand in the filtrate are measured. Importantly, it must be stressed that this method may lead to errors because it disturbs the association equilibrium. Equally, it should be noted that the filters used might become saturated by relatively weak protein concentrations (YARUS & BERG, 1970).

### 2.6.6. ULTRACENTRIFUGATION

In this method the homogenous protein-ligand solution is subjected to a gravitational field. After a suitable period of time, at the top of the centrifugation tubes the solution contains only free ligand whose concentration can be determined (STEINBERG & SCHACHMAN, 1966). Originally, the method required analytical or preparative ultracentrifuges. It has since been adapted for use with ultra-rapid tabletop microcentrifuges (HOWLETT et al., 1978).

### 2.6.7. DIRECT SPECTROPHOTOMETRIC METHODS

If a substrate, or more generally, a ligand possesses a chromophore that absorbs in a region not interfering with the absorption spectrum of the protein, then spectroscopic methods offer a rapid and easy means to study its association to the protein. The absorption spectrum of the protein-ligand complex needs to shift with respect to the spectrum of the free ligand and is often the case. Furthermore, a wavelength must be found where the difference in absorbance is large enough for any variation to be measurable in a relatively large zone of ligand concentrations (Fig. 2.12).



**Fig. 2.12 Spectrophotometric method**  
 (a) absorbance spectrum of a ligand  $L$   
 and of the enzyme-ligand complex,  $EL$   
 (b) difference absorbance spectrum.

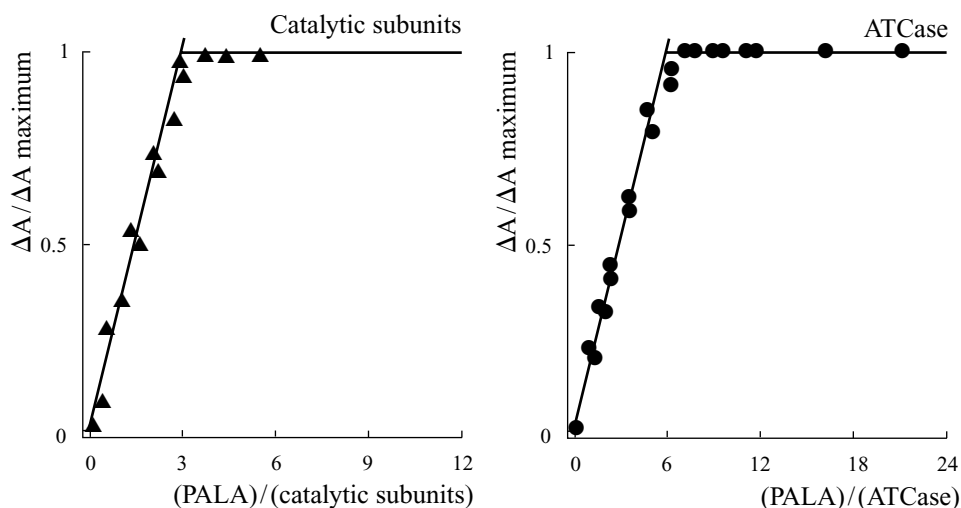
*The difference is significant,  
 so the conditions are favourable  
 for the application of this method*

It is often preferable to employ much more sensitive fluorescence methods where possible. Such methods are applicable when, upon binding to an enzyme, the fluorescence signal of the substrate changes sufficiently for detection. Since fluorescence methods are highly sensitive it is crucial to ensure that buffer solutions do not contain fluorescent impurities.

In certain cases, the fluorescence of the enzyme itself may be monitored if ligand binding significantly affects its signal.

### 2.6.8. DIRECT TITRATION OF A NUMBER OF ACTIVE SITES

When a ligand possesses a very strong affinity for a protein, the protein concentration can be much higher than the dissociation constant. Under these conditions, virtually all the added ligand binds and the number of binding sites may be directly titrated. In this way, for example, it was possible to titrate directly the catalytic sites of aspartate transcarbamylase with N-phosphonacetyl-L-aspartate (PALA), an analogue of a substrate transition state (KERBIRIOU et al., 1977). Figure 2.13 below shows below the results of the titration.



**Fig. 2.13 Spectrophotometric titration of the catalytic sites of aspartate transcarbamylase (ATCase) by *N*-phosphonacetyl-*L*-aspartate (PALA) (a) using isolated catalytic subunits (catalytic trimer) – (b) using the full-length enzyme, which possesses six catalytic subunits.** (From *J. Biol. Chem.*, **252**, KERBIRIOU D. *et al.*, 2881. © (1977) with permission from The American Society for Biochemistry and Molecular Biology)

### 2.6.9. INTERPRETATION OF EXPERIMENTAL DATA

The analysis of experimental data from ligand-binding studies is often very sensitive and may lead to errors in interpretation, notably when a protein possesses several categories of binding site, or when it displays cooperative or anti-cooperative interactions. If only one category of equivalent and independent binding site exists the situation is simpler, although certain precautions must still be taken. Firstly, the experimental study must span a wide ligand concentration range. If possible this should vary by two orders of magnitude and around the intrinsic affinity constant (or constants). Indeed, if the concentration range covered in the experiment is too narrow the conclusions may sometimes lead to the perception of only a single category of binding site, when in fact several exist.

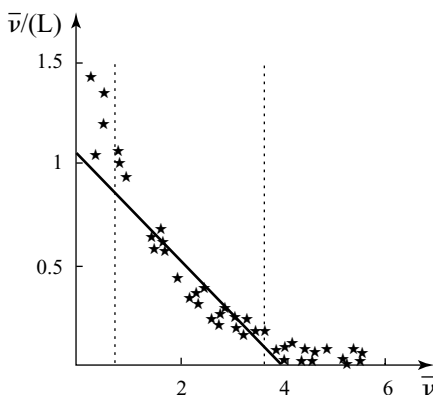
In order to illustrate this last point, we will consider one of the most discussed examples in articles regarding the analysis of ligand binding: that of carbamyl phosphate binding to aspartate transcarbamylase. SUTER and ROSENBUSH (1976) carried out several experiments in a concentration range of carbamyl phosphate varying between 0.5 and 200  $\mu\text{M}$ . The corresponding SCATCHARD plot is clearly not linear (Fig. 2.14 opposite). The results were interpreted on the basis of two classes of site, that is, three ( $2.9 \pm 0.2$ ) high-affinity sites ( $K_{1,\text{diss}} = 2.3 \mu\text{M}$ ) and three ( $2.9 \pm 0.2$ ) sites with lower affinity ( $K_{1,\text{diss}} = 62.5 \mu\text{M}$ ), the enzyme concentration being 1.7  $\mu\text{M}$ . If the authors had explored a narrower concentration range, they might have concluded that four equivalent and independent sites existed with an

affinity corresponding to  $K'_{\text{diss}} = 3.2 \mu\text{M}$ . This type of error arising through experiment flaws is still frequently encountered in the literature. Furthermore, it is not always easy to reach sufficiently high ligand concentrations (due to low solubility) or, conversely, sufficiently low concentrations (due to inadequate sensitivity). It is therefore essential to bear in mind that in some experiments the presence of certain sites may not be detectable and this must be taken into account during their interpretation.

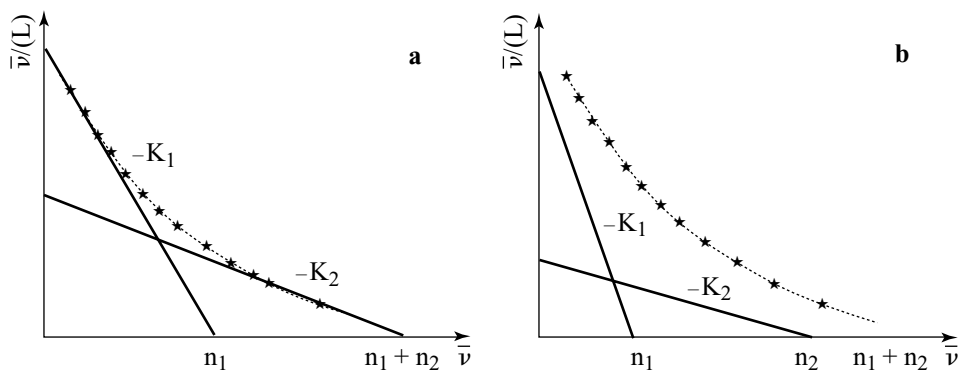
**Fig. 2.14 Binding of carbamyl-phosphate to aspartate transcarbamylase**

Dashed lines mark a narrower ligand concentration range.

(From *J. Biol. Chem.*, **251**, SUTER P. & ROSENBUSH J.P., 5986. © (1976) with permission from The American Society for Biochemistry and Molecular Biology)



Amongst the errors found in the literature, it is useful to point out the erroneous interpretations of the non-linear SCATCHARD plots mentioned by NØRBY *et al.* (1980). **Figure 2.15 illustrates this type of error.** Figure 2.15a represents the false interpretation, which involves comparing the slope of the linear parts of the diagram to the dissociation constant and simply extrapolating in order to have the number of sites of each category. Figure 2.15b represents the rigorous deconvolution of this diagram according to the procedure previously indicated (see Sect. 2.3.3).



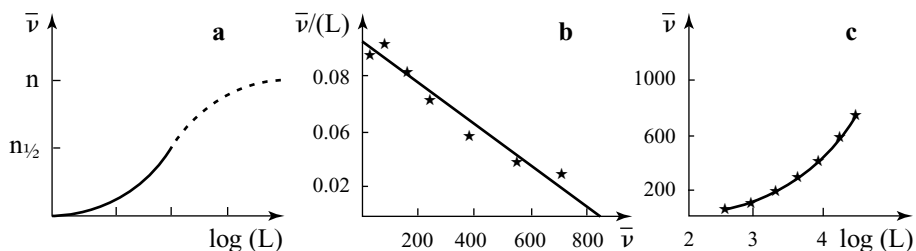
**Fig. 2.15 Theoretical SCATCHARD diagrams for the binding of a ligand to two categories of different sites**

(a) incorrect resolution of the data – (b) correct resolution of the data

(Reprinted from *Anal. Biochem.*, 102, NØRBY J.G. *et al.*, Scatchard plot: common misinterpretation of binding experiments, 318. © (1980) with permission from Elsevier)

- When the experiment is conducted in a correct manner and there is no ambiguity in the graphical plot, it is still important to select the best procedure for data analysis. This has been the object of numerous discussions in the literature. It has been previously underlined that among the linear representations, the SCATCHARD plot enables a better determination of the experimental parameters than the KLOTZ plot. This will be demonstrated in Part II (Chap. 5, Sect. 5.2.3) for the EADIE and LINEWEAVER-BURK plots, which, in the case of enzyme reactions, are equivalent to the SCATCHARD and KLOTZ plots, respectively.

KLOTZ (1982) criticised the use of the SCATCHARD plot, in particular when extrapolating to deduce the number of sites in biological systems, where several categories of receptor site with different affinities often exist; so too for non-specific binding, since often these experiments do not allow values close to saturation to be reached. KLOTZ proposed the use of another type of representation in which the concentration of bound molecules is plotted as a function of the logarithm of the free ligand concentration. If all sites are equivalent and independent, this representation has the following properties: the inflexion point corresponds to half saturation, the sigmoidal curve is symmetrical about the inflexion point, the plateau corresponding to  $n$ , the number of sites, is reached asymptotically for very large values of free-ligand concentration. KLOTZ emphasised that in many cases extrapolation of the SCATCHARD plot is carried out when actually the inflexion point is not yet reached. Figure 2.16 illustrates KLOTZ's argument. In fact, the same arguments can be made for the semi-logarithmic plot, because it is often very difficult to locate the inflexion point on such a diagram and the symmetry that enables deduction of the saturation is only valid when the sites are both equivalent and independent.



**Fig. 2.16 (a) semi-logarithmic representation of bound ligand as a function of the logarithm of the free-ligand concentration, for a receptor with  $n$  identical sites (b) SCATCHARD plot for the binding of diazoprane to benzodiazepine receptors from rat cerebral cortex membranes (c) changes in bound ligand as a function of the free-ligand concentration (same representation as in (a) with the data from (b))**

(From *Science*, 217, KLOTZ, Numbers of receptor sites from Scatchard graphs: facts and fantasies, 1248. © (1982) reprinted with permission from American Association for the Advancement of Science)

Incidentally, MUNSON and ROBARD (1983) raised a “constructive criticism” of both the SCATCHARD and KLOTZ representations and underlined that the semi-logarithmic plot is not superior to the SCATCHARD plot; the important point being that they should be used and interpreted correctly. They recommended using a statistical method to analyse the data. These authors (1980) suggested a mathematical procedure for the analysis of ligand-binding data. They employed a weighted least-squares



method for the determination of binding parameters by considering different multiple-binding models likely to be encountered in several areas of biology such as endocrinology, neurology, immunology, enzymology and the physico-chemistry of proteins. The SCATCHARD plot, despite its advantages, does not lend itself to statistical analysis, since the variables plotted on the horizontal and vertical axes are not independent. However, it is frequently used as an initial means to obtain provisional estimates for the constants. A detailed treatment of the statistical analysis methods is given in Part II (Chap. 5, § 5.2.3). Other calculation procedures have been suggested such as the simulation method based on the binding polynomial of WYMAN (see § 2.5) proposed by JOSÉ (1985). **Several programs for the statistical data analysis of protein-ligand binding have been developed and are available.** The program CMFITT, developed by M. DESMADRIL, is currently commercialised by the CNRS.

The fact remains that ligand-binding equilibria involving proteins that comprise different classes of binding site or interactions between the sites are always difficult to interpret and may lead to false conclusions even if the experiments have been conducted correctly and a statistical analysis of the data carried out. This notion has been emphasised by several authors, in particular LIGHT (1984) (see also the response of PAUL et al., 1984). Therefore, it is worth independently obtaining complementary information such as structural data, for example. Thus, in the case of the binding of carbamyl phosphate to aspartate transcarbamylase, the statistical treatment of FELDMAN (1983) and the comments by KLOTZ and HUNSTON (1984) clearly indicate that structural data were able to remove the ambiguity. ▲

## ***BIBLIOGRAPHY***

### ***BOOKS***

- ADAIR G.S. –1949– in *Haemoglobin*, F.J.W. ROUGHTON & J.C. KENDREW eds, Butterworths, Cambridge.
- EDSALL J.T. & WYMAN J. –1958– *Biophysical chemistry*, Acad. Press, New York.
- YON J. –1959– *Structure et dynamique conformationnelle des protéines*, Hermann, Paris.

### ***REVIEWS***

- KLOTZ I.M. –1953– in *The proteins*, H. NEURATH & K. BAILEY eds, 1st ed., Acad. Press, New York.
- WYMAN J. –1964– Linked functions and reciprocal effects in hemoglobin, a second look. *Adv. Prot. Chem.* **19**, 223–286.

### ***SPECIALISED ARTICLES***

- ADAIR G.S. –1925– *J. Biol. Chem.* **63**, 493.
- COLOWICK S.P. & WOMACK F.C. –1969– *J. Biol. Chem.* **244**, 774.

- FELDMAN H.A. –1983– *J. Biol. Chem.* **258**, 12865.
- GRAY C.W. & CHAMBERLIN M.J. –1971– *Anal. Biochem.* **41**, 83.
- HILL A.V. –1910– *J. Physiol. Lond.* **40**, 4.
- HOWLETT G.J., YEH E. & SCHACHMAN H.K. –1978– *Arch. Biochim. Biophys.* **190**, 809.
- HUMMEL J.P. & DREYER W.J. –1962– *Biochim. Biophys. Acta* **63**, 530.
- JOSÉ M. –1985– *Anal. Biochem.* **144**, 494.
- KERBIRIOU D., HERVÉ G. & GRIFFIN J.H. –1977– *J. Biol. Chem.* **252**, 2881.
- KLOTZ I.M. –1982– *Science* **217**, 1247.
- KLOTZ I.M. –1983– *Science* **220**, 981.
- KLOTZ I.M. & HUNSTON D.L. –1984– *J. Biol. Chem.* **259**, 10060.
- LIGHT K.E. –1984– *Science* **223**, 76.
- MUNSON P.J. & ROBBARD D. –1980– *Anal. Biochem.* **170**, 220.
- MUNSON P.J. & ROBBARD D. –1983– *Science* **220**, 979.
- MURALT A. (VON) –1930– *J. Am. Chem. Soc.* **52**, 3518.
- NØRBY J.G., OTTOLENGHI P. & JENSEN J. –1980– *Anal. Biochem.* **102**, 318.
- NOZAKI Y., GURD F.R.N., CHEN R. & EDSALL J.T. –1957– *J. Am. Chem. Soc.* **79**, 2123.
- PAUL S.M., HANGER R.L., HULIHAN-GIBLIN B.A. & SKOLNICK P. –1984– *Science* **223**, 77.
- SCATCHARD G. –1949– *Ann. N.Y. Acad. Sci.* **51**, 660.
- STEINBERG I.Z. & SCHACHMAN H.K. –1966– *Biochemistry* **5**, 3728.
- SUTER P. & ROSENBUSH J.P. –1976– *J. Biol. Chem.* **251**, 5986.
- TANFORD C., DE P.K. & TAGGART V.G. –1960– *J. Am. Chem. Soc.* **82**, 6028.
- TENU J.P., GHELIS C., YON J. & CHEVALLIER J. –1976– *Biochimie* **58**, 513.
- WYMAN J. –1948– *Adv. Prot. Chem.* **4**, 407–531.
- YARUS M. & BERG P. –1970– *Anal. Biochem.* **35**, 450.