



Review

# Analysis of the Equilibrium Distribution of Ligands in Heterogeneous Media—Approaches and Pitfalls

Maria João Moreno <sup>1,2,\*</sup> , Luís M. S. Loura <sup>1,3</sup> , Jorge Martins <sup>4</sup> , Armindo Salvador <sup>1,5,6</sup>   
and Adrian Velazquez-Campoy <sup>7,8,9,10</sup>

- <sup>1</sup> Coimbra Chemistry Center—Institute of Molecular Sciences (CQC-IMS), University of Coimbra, 3004-535 Coimbra, Portugal
  - <sup>2</sup> Department of Chemistry, Faculty of Sciences and Technology, University of Coimbra, 3004-535 Coimbra, Portugal
  - <sup>3</sup> Faculty of Pharmacy, University of Coimbra, 3000-548 Coimbra, Portugal
  - <sup>4</sup> Centro de Ciências do Mar (CCMAR/CIMAR, LA) and DCBB-FCT, Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal
  - <sup>5</sup> CNC—Center for Neuroscience and Cell Biology, 3004-504 Coimbra, Portugal
  - <sup>6</sup> Institute for Interdisciplinary Research, University of Coimbra, 3030-789 Coimbra, Portugal
  - <sup>7</sup> Institute for Biocomputation and Physics of Complex Systems (BIFI), Joint Unit GBsC-CSIC-BIFI, Universidad de Zaragoza, 50018 Zaragoza, Spain
  - <sup>8</sup> Departamento de Bioquímica y Biología Molecular y Celular, Universidad de Zaragoza, 50009 Zaragoza, Spain
  - <sup>9</sup> Instituto de Investigación Sanitaria de Aragón (IIS Aragón), 50009 Zaragoza, Spain
  - <sup>10</sup> Centro de Investigación Biomédica en Red en el Área Temática de Enfermedades Hepáticas Digestivas (CIBERehd), 28029 Madrid, Spain
- \* Correspondence: mmoreno@ci.uc.pt



**Citation:** Moreno, M.J.; Loura, L.M.S.; Martins, J.; Salvador, A.; Velazquez-Campoy, A. Analysis of the Equilibrium Distribution of Ligands in Heterogeneous Media—Approaches and Pitfalls. *Int. J. Mol. Sci.* **2022**, *23*, 9757. <https://doi.org/10.3390/ijms23179757>

Academic Editor: David M. Jameson

Received: 31 July 2022

Accepted: 24 August 2022

Published: 28 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** The equilibrium distribution of small molecules (ligands) between binding agents in heterogeneous media is an important property that determines their activity. Heterogeneous systems containing proteins and lipid membranes are particularly relevant due to their prevalence in biological systems, and their importance to ligand distribution, which, in turn, is crucial to ligand's availability and biological activity. In this work, we review several approaches and formalisms for the analysis of the equilibrium distribution of ligands in the presence of proteins, lipid membranes, or both. Special attention is given to common pitfalls in the analysis, with the establishment of the validity limits for the distinct approaches. Due to its widespread use, special attention is given to the characterization of ligand binding through the analysis of Stern–Volmer plots of protein fluorescence quenching. Systems of increasing complexity are considered, from proteins with single to multiple binding sites, from ligands interacting with proteins only to biomembranes containing lipid bilayers and membrane proteins. A new formalism is proposed, in which ligand binding is treated as a partition process, while considering the saturation of protein binding sites. This formalism is particularly useful for the characterization of interaction with membrane proteins.

**Keywords:** binding affinity; partition coefficient; proteins; lipid bilayers; biomembranes; Stern–Volmer plots; Scatchard plots

## 1. Introduction

Biological systems are very heterogeneous media, with macromolecules or supramolecular aggregates dispersed in an aqueous medium. The hydrophobic effect is the major force that drives the formation of these structures [1,2] where the dispersed structures are characterized by a polar or ionic shell and a nonpolar core. Small molecules distribute between the aqueous phase and the non-aqueous structures as a result of the established interactions. The hydrophobic effect is the major force for non-specific binding, while polar interactions (type, configuration and geometry) are responsible for interaction specificity.

The distribution of the small molecules (ligands) between the distinct structures determines their activity and availability.

The most common examples of heterogeneous media are proteins in aqueous solution [3]. Most of these macromolecules fold into very organized and well-defined structures, with distinct regions where the other molecules may bind with higher or lower affinity and specificity [4,5].

Biological membranes are another example of heterogeneous media, formed by amphiphilic lipids assembled in a bilayer and containing proteins embedded in the bilayer or associated with its surface [6,7]. A major difference between lipid membranes and proteins is the inexistence of well-defined binding sites in the former. Being formed from a large number of molecules associated by many weak forces, lipid bilayers of biomembranes are generally in the fluid phase and can dissolve or incorporate small molecules without changing their properties, namely their ability to dissolve additional molecules. Thus, although the properties of the membranes vary strongly across the bilayer, the distribution of small molecules between the aqueous medium and membranes is usually well-defined by a partition coefficient or a partition constant [8–10].

Other examples of heterogeneous media are micelle solutions, with properties differing from the previous two examples. Similar to lipid bilayers, micelles are also formed by the association of many amphiphilic molecules held together by weak forces. They are therefore dynamic structures, able to dissolve small molecules. However, due to their small size (comparable to that of proteins), micelles saturate at relatively small numbers of solute occupancy, with the average number of solutes per micelle  $\langle n \rangle$  being usually  $\leq 1$  [11]. At small solute concentrations, the distribution between the aqueous medium and the micelles is usually well described by a partition. However, similarly to lipid bilayers, the local concentration of solute is not homogeneous within the micelle [12,13]. For  $\langle n \rangle = 1$ , the number of solutes encountered per micelle usually follows a Poisson distribution, with about 40% of the micelles being without solute, while 26% of the micelles having 2 or more solutes. The distinct and small number of solutes in the micelles leads to stochastic effects, strongly influencing the system's properties and reactivity [14–17].

Another distinctive characteristic of heterogeneous media is their dynamics. Folded proteins are structures formed by a single macromolecule or a few polypeptide chains with relatively strong interactions between them. Conversely, lipid bilayers and micelles are supramolecular structures held together by weak forces. The number of amphiphiles in each micelle varies from a few molecules (e.g., bile salt micelles [18]) to a few hundred [19,20], while biomembranes are formed by millions of amphiphilic lipids. The amphiphiles that associate as micelles also have a higher hydrophilic/lipophilic balance and are therefore more soluble in the aqueous medium as monomers [21]. Both properties lead to faster dynamics of amphiphile insertion/desorption and assembly/disassembly of the supramolecular structure for the case of micelles, with important consequences for solute distribution [22–25]. Given the particularities of micellar heterogeneous media, and their relatively lower importance in biological systems, they will not be further addressed in this work.

When considering the distribution of small molecules in heterogeneous media, it is important to distinguish between affinity and specificity. The binding affinity of a ligand to a protein or biomembrane reflects to a high extent its poor affinity for the solvent (its hydrophobicity). In turn, binding specificity depends on the establishment of several interactions with the binding agent in a well-defined, three-dimensional arrangement of the groups involved in the interaction. Although the maximization of affinity usually leads also to a high specificity, ref. [26] the latter may vary widely for the case of ligands with low-to-moderate hydrophobicity [27]. Interactions with lipid bilayers are usually not specific, with the association affinities depending essentially on the fluidity and charge of the membrane [9,28–38]. Some exceptions may occur for the case of membranes with lateral heterogeneity and/or when glycolipids are present [39–42]. In contrast, proteins have well-defined binding sites with a characteristic spatial distribution of functional groups

that lead to a stronger interaction with some specific molecules [43–47]. However, for some proteins a wide variability is observed in the properties of ligands with similar affinity (with serum albumins being a striking example), showing that binding to proteins can also be non-specific. This review only addresses the methodologies and formalisms for characterization of binding affinity.

Because proteins have well-defined binding sites, the characterization of ligand affinity may be done following two approaches: (i) from the fraction of protein with ligand bound; and (ii) from the fraction of ligand bound to the protein. Methodologies based on both approaches will be addressed in this review, and in principle both lead to the same parameters. In contrast, due to the absence of well-defined binding sites in lipid membranes, very different results are obtained depending on whether the fraction of ligand bound (increasing the amount of membrane) or the local concentration of ligand in the membrane (increasing the amount of ligand) is followed. The former approach provides the ligand binding affinity for an unperturbed bilayer, while the later provides information on the saturation of the membrane with the ligand (which corresponds necessarily to a membrane with altered properties).

It is the objective of this review to compile different formalisms and experimental approaches for the characterization of the binding affinity of small molecules to proteins and lipid membranes. Focus is given to practical aspects, with the clear identification of the applicability of different formalisms and common pitfalls. The complexity of the considered heterogeneous systems increases throughout the review, from proteins with a single binding site to biomembranes containing both proteins and lipid bilayers.

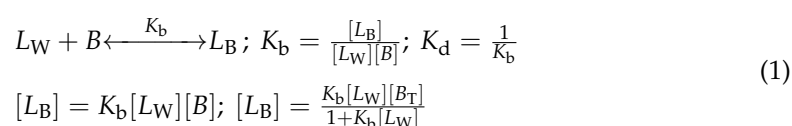
## 2. Proteins in Solution

### 2.1. Proteins with a Single Binding Site

#### 2.1.1. Analytical Solution of the Full Equation

Ligand association is more frequently followed through variations in a protein property that is proportional to the advance of the binding process (i.e., saturation fraction). Several properties may be followed, including the protein enzymatic activity, ref. [48] protein structure or intrinsic fluorescence, refs. [49–53] changes in the charge or mass, refs. [54–56] or properties that reflect the binding itself, such as the heat evolved or changes in the ionization of the ligand or protein [47,49,56].

If the protein contains only one type of binding sites (one or several equivalent and independent binding sites on each protein molecule), ligand association may be described by the following scheme and equilibrium relations:



where  $[L_B]$ ,  $[L_W]$ ,  $[B]$ , and  $[B_T]$  are, respectively, the concentrations of the ligand bound to the protein, the ligand in the aqueous phase, unoccupied binding sites in the proteins, and the total concentration of binding sites (all with respect to the total volume of the solution), whereas  $K_b$  and  $K_d$  are, respectively, the equilibrium binding and dissociation constants. Considering the mass conservation equations for this system ( $[L_T] = [L_W] + [L_B]$ ;  $[B_T] = [B] + [L_B]$ ) the equation from which the concentration of occupied binding sites may be calculated is:

$$(x)^2 - x(K_d + [B_T] + [L_T]) + [L_T][B_T] = 0; x = [L_B] \quad (2)$$

where the dissociation is used instead of the association constant due to the simplification of the resulting equation.

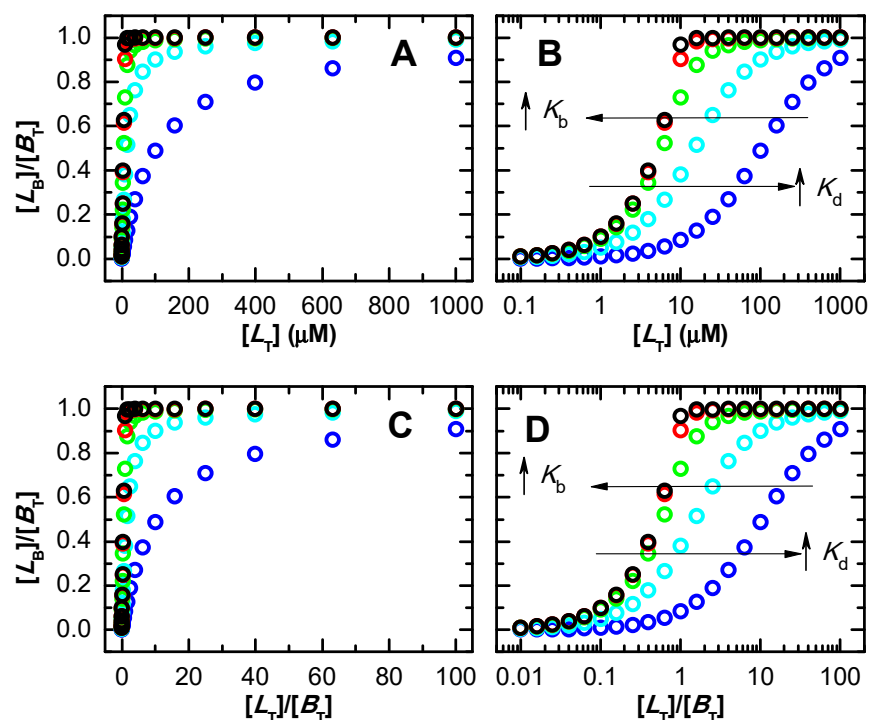
To use this quadratic equation for the quantitative analysis of experimental results, it is first necessary to identify which of the two mathematically possible solutions, Equation (3), has physical meaning.

$$x_+ = [L_B] = \frac{(K_d + [B_T] + [L_T]) + \sqrt{(K_d + [B_T] + [L_T])^2 - 4[L_T][B_T]}}{2} \quad (3)$$

$$x_- = [L_B] = \frac{(K_d + [B_T] + [L_T]) - \sqrt{(K_d + [B_T] + [L_T])^2 - 4[L_T][B_T]}}{2}$$

In the case of association between two entities, it may be easily identified that the physically meaningful solution is always  $x_-$ , because  $x_+$  is higher than  $[B_T]$  and/or  $[L_T]$ . Where not evident, the physically meaningful solution may be decided by calculating  $x_{\pm}$  over a large range of parameter and concentration values, with the constraints that it must be  $>0$ ,  $<[B_T]$  and  $<[L_T]$ . When the protein saturation is experimentally followed, the concentration of ligand is the independent variable, and the association affinity ( $K_b$  or  $K_d$ ) may be found through the best fit of the correct solution of the quadratic equation to the experimental results. The solution of Equation (3) is valid in all conditions, low- or high-affinity binding, comparable or dissimilar concentrations of ligand and protein.

Figure 1 shows the dependence of the protein saturation, with  $[L_T]$  (top plots) and with the  $[L_T]/[B_T]$  ratio (bottom plots) for several values of the association affinity. As the affinity of the ligand to the protein increases (higher  $K_b$  or lower  $K_d$ ), the protein becomes saturated at lower ligand concentrations and at lower ratios of total ligand to protein-binding sites. The representation of the x axis in the logarithmic scale (plot B and D), which is proportional to the natural thermodynamic variable (the chemical potential), facilitates the interpretation of the results. For high-affinity binding, the protein is fully saturated for  $[L_T]/[B_T] > 1$ , smoothly decreasing to zero at lower ligand concentrations. The decrease in the binding affinity leads to a lower fraction of protein saturation at the same  $[L_T]$  (or ratio  $[L_T]/[B_T]$ ), with a smoother transition between low and high protein saturation.



**Figure 1.** Saturation of binding sites predicted from the full description of the system, Equation (3), as a function of the total concentration of ligand (plots (A,B)), or of the ratio of total ligand to binding sites (plots (C,D)). Note the logarithmic scale on the x-axis of plots (B,D). The total concentration of binding sites was 10  $\mu\text{M}$ , and the association affinity ( $K_d$ ) was varied: 100  $\mu\text{M}$  ( $\circ$ ), 10  $\mu\text{M}$  ( $\square$ ), 1  $\mu\text{M}$  ( $\triangle$ ), 0.1  $\mu\text{M}$  ( $\diamond$ ), and 0.01  $\mu\text{M}$  ( $\circ$ ).

Note that the asymmetry of the isotherms obtained for high-affinity binding is due to the representation as a function of  $[L_T]$ . If the concentration of free ligand is used instead, symmetric isotherms are always obtained, as shown in Figure S1. However, the free ligand concentration is not a known parameter, as it depends on the unknown binding affinity. Thus, to evaluate the adequacy of the conditions regarding the characterization of the binding affinity, one has to consider the known total ligand concentration. From Figure 1 one notes that the binding isotherms for the two highest affinities considered (red and black) are mostly overlapped. This shows that for the total concentration of protein considered in those simulations (10  $\mu\text{M}$ ), the highest affinity that may be accurately characterized is  $K_b = 10^7 \text{ M}^{-1}$  ( $K_d = 0.1 \mu\text{M}$ ). For the characterization of higher affinity binding, a lower protein concentration must be used. Illustrating this point, simulations for 1  $\mu\text{M}$  total protein concentration (Figure S2) show that the isotherms for  $K_d = 0.1 \mu\text{M}$  and  $0.01 \mu\text{M}$  no longer fully overlap in these conditions, allowing their experimental discrimination. The upper limit for accurate estimation of the binding affinity is now  $K_d = 0.01 \mu\text{M}$  (note the overlap of black and grey symbols in Figure S2A).

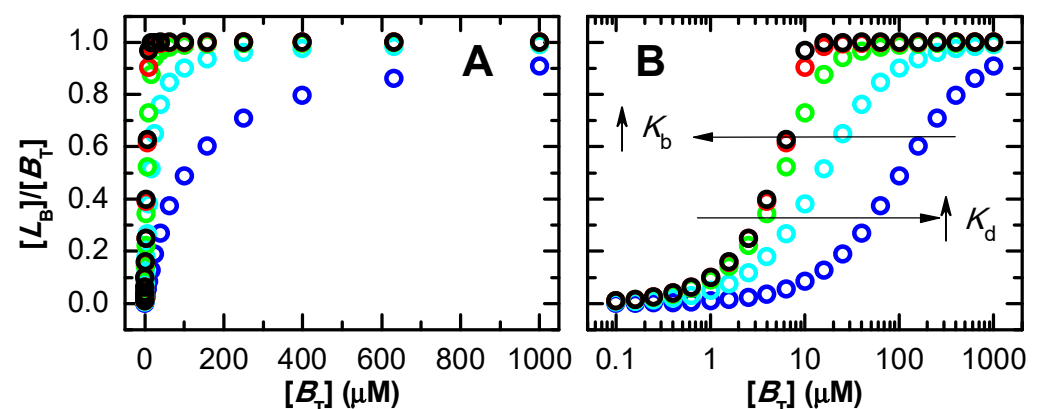
In the previous figures it was considered that protein saturation was being experimentally followed. If a property of the ligand is monitored instead (e.g., ligand fluorescence), refs. [29,57–60] the dependent variable is now the fraction of ligand bound, and the total concentration of protein is the independent variable. The concentration of ligand bound is again calculated from the  $x_-$  solution of Equation (3), from which its fraction may be calculated.

The fraction of bound ligand for several values of the association affinity is shown in Figure 2, as a function of the concentration of binding sites. As expected, the curves are equivalent to those shown in Figure 1 when saturation of the protein is being followed. In fact, the variations observed in the experimentally followed variable ( $\Delta S$ ) are proportional to  $[L_B]$ , which is the same for a given set of  $K_d$ ,  $[B_T]$  and  $[L_T]$ , Equation (4). The normalization of  $\Delta S$  by its limit value ( $\Delta S_{\text{max}}$  corresponding to full protein saturation) leads exactly to the same curve shape as when a property of the protein or ligand is being followed, Equation (5).

$$[L_B] = K_b[L_W][B] = \frac{[L_W][B]}{K_d} \quad (4)$$

$$\Delta S = f[L_B]; \Delta S_{\text{max}} = f[L_B]_{\text{max}} \quad (5)$$

$$\Delta S = \frac{\Delta S_{\text{max}}}{[L_B]_{\text{max}}} [L_B]$$



**Figure 2.** Fraction of bound ligand predicted from the full description of the system, Equation (3), as a function of the total concentration of binding sites, represented with linear (A) or logarithmic (B) x-axis scale. The total concentration of ligand was 10  $\mu\text{M}$ , and the association affinity was varied from  $K_d$  equal to 100  $\mu\text{M}$  ( $\circ$ ), 10  $\mu\text{M}$  ( $\square$ ), 1  $\mu\text{M}$  ( $\triangle$ ), 0.1  $\mu\text{M}$  ( $\diamond$ ), and 0.01  $\mu\text{M}$  ( $\circ$ ).

### 2.1.2. Approximations Commonly Used

Although the analysis methods currently available allow the easy use of the full equation, the use of simplified equations to quantitatively characterize the association between ligands and proteins is still frequent. The two most common approximations are presented below, with a critical analysis of their validity.

When protein saturation is followed, it is commonly assumed that the ligand is in large excess ( $[L_T] \gg [B_T]$ ). Under those conditions (corresponding to negligible ligand depletion due to binding), Equation (3) simplifies to the classical hyperbolic equation:

$$\frac{[L_B]}{[B_T]} = \frac{K_b[L_T]}{1 + K_b[L_T]} \cong \frac{[L_T]}{K_d + [L_T]} \text{ for } [L_T] \gg [B_T] \quad (6)$$

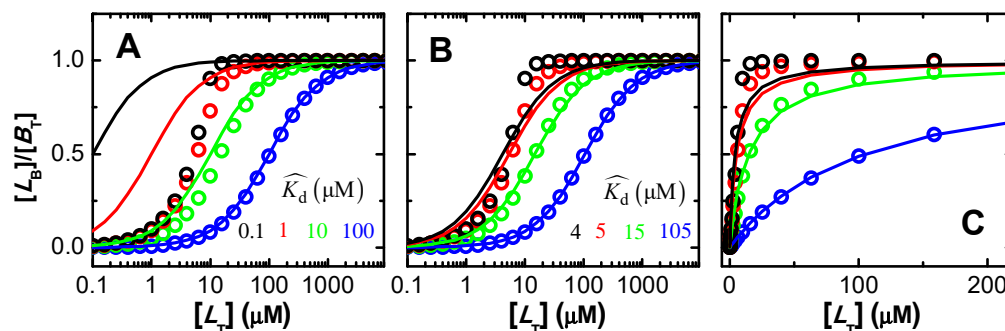
This approximation of excess ligand was originally considered in the context of enzyme kinetics [61], where for in vitro studies the approximation is usually valid because the enzyme is in very low concentrations. However, as shown below, when studying the association between ligands and proteins this approximation may easily become inapplicable and lead to incorrect estimations of the binding affinity.

When binding is followed through a ligand property, it is the fraction of bound ligand that is accessed. The total concentration of protein is usually the independent variable that is varied until nearly all ligand is bound. If it may be considered that the protein is in large excess ( $[B_T] \gg [L_T]$ ), the general equation simplifies to:

$$\frac{[L_B]}{[L_T]} \cong \frac{K_b[B_T]}{1 + K_b[B_T]} = \frac{[B_T]}{K_d + [B_T]} \text{ for } [B_T] \gg [L_T] \quad (7)$$

Figure 3 (plot A) represents the real saturation of the protein (calculated from Equation (3), circles), and that predicted when assuming excess ligand (continuous lines, Equation (6)). The concentration of binding sites was always 10  $\mu\text{M}$ , and ligand concentration was varied between  $10^{-1}$  and  $10^4$   $\mu\text{M}$ . For the lower affinity considered ( $K_d = 100$   $\mu\text{M}$ ), half saturation occurs at a ligand concentration close to 100  $\mu\text{M}$ , where the condition of excess ligand is a good approximation. The best fit of Equation (6) leads to  $K_d = 105$   $\mu\text{M}$ , differing by only 5% from the true value. However, as expected, higher affinities lead to significant binding at ligand concentrations comparable to or smaller than the concentration of binding sites. Under those conditions, the approximation of excess ligand is not valid, and the best fit of Equation (6) leads to incorrect parameter estimations. For  $K_d = 10$   $\mu\text{M}$  this parameter is overestimated by 50%, and the approximation is totally inappropriate for lower dissociation constants (higher affinities). This inadequacy results from the fact that the most relevant part of the saturation curve for parameter estimation occurs for ligand concentrations that are lower than the concentration of binding sites, violating the assumption of excess ligand. An equivalent behaviour is obtained when the approximation of excess protein is used (Figure S3). In this case, the concentration of binding sites is the independent variable and the fractional binding of ligand is the dependent one.

It should be noted that the general equations for ligand binding to proteins with a single type of binding sites, Equation (1), are valid at all ligand and protein concentrations, for low- and high-affinity binding. The limits of validity of Equations (6) and (7) arise from the substitution of the concentration of free species (ligand or binding sites) with their total concentration. The inadequacy of the approximations of excess ligand or protein is well-known and has been repeatedly highlighted in the literature [61,62]. However, although more careful researchers and specialized software take this into consideration, refs. [47,63–66] it is still frequently overlooked.



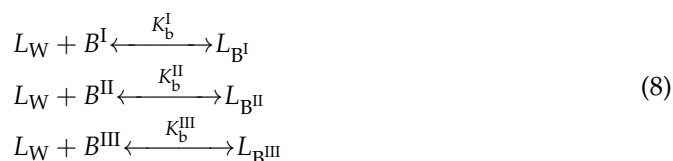
**Figure 3.** Saturation of binding sites predicted from the full description of the system (Equation (3), circles) for  $[B_T] = 10 \mu\text{M}$  and different dissociation constants ( $K_d$ ):  $100 \mu\text{M}$  ( $\circ$ ),  $10 \mu\text{M}$  ( $\odot$ ),  $1 \mu\text{M}$  ( $\circ$ ), and  $0.1 \mu\text{M}$  ( $\circ$ ). The lines were calculated from Equation (6) (excess ligand approximation). Plot (A)—predicted saturation of binding sites for the different binding affinities considered; Plots (B,C)—best fit of Equation (6) to the saturation of binding sites observed (note the logarithmic scale in plot (B) and linear scale in plot (C)). The dissociation constants obtained from the best fit of Equation (6) are indicated in plot (B).

2.2. Proteins with More than One Binding Site

2.2.1. Formalism—Microscopic and Macroscopic Association Constants

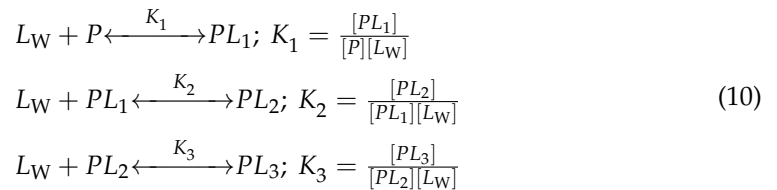
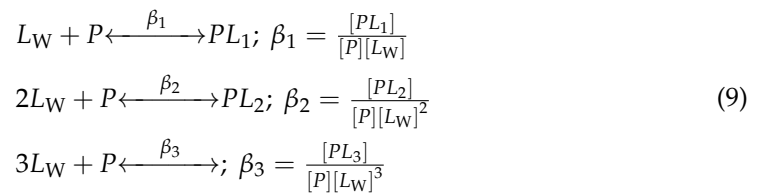
A commonly encountered difficulty when studying protein/ligand association occurs when the number of binding sites per protein is not well-characterized. If protein saturation may be achieved and followed quantitatively, the macroscopic binding associations may be obtained and from their dependence on protein saturation some conclusions may be drawn regarding the number of binding sites, their similarity and independence. In some conditions, one may also obtain information regarding the microscopic binding association to the various binding sites and the degree of cooperativity. Excellent reviews are available in the literature for the full and accurate analysis of protein saturation for the case of multiple binding sites (e.g., [67]). The cases of three independent binding sites (equal or different) will be briefly analysed below, calling attention to common problems when oversimplifying the analysis of the data.

There are three possibly distinct binding sites in the system (in this case all three in the same protein, but the same formalism applies if each is in a distinct protein because they are independent), so binding is described by the following scheme:



where  $B^i$  and  $L_{B^i}$  correspond respectively to the free and occupied binding sites  $i$ , and  $K_b^i$  is the respective microscopic association constant, with  $i$  representing the distinct binding sites (I, II, or III). The association of the ligand with each binding site cannot be characterized independently. Instead, a general equation considering ligand association with all binding sites must be defined and solved. If the binding sites are equivalent and independent, the solution is simply to consider the total concentration of binding sites ( $[B] = [B^I] + [B^{II}] + [B^{III}]$ ) in Equation (2) and use the solution of the quadratic equation as indicated above. However, if the binding sites are not equivalent, the system is described by a set of  $n + 2$  equations where  $n$  is the number of binding sites, and must be solved numerically. If the concentration of protein is known and each protein contains the three distinct binding sites, protein saturation with ligand is best described by the macroscopic

overall association constants  $\beta_i$ , or the stepwise macroscopic association constants  $K_i$ , defined in Equations (9) and (10), respectively.



Note that  $PL_i$  represents a sub-ensemble of protein–ligand complexes with  $i$  ligands bound per protein. It does not contain information regarding the binding sites where the ligands are bound. For example,  $PL_1$  represents the collection of protein molecules with only one ligand bound, which may be in site 1, 2 or 3. The relative fraction of ligand bound to each of the possible binding sites depends on the microscopic binding affinities ( $K_b^i$ ). The same holds for  $PL_2$  and  $PL_3$ , although in this case the relative fractions are not simply a result of the relative values of  $K_b^i$  but also of possible dependencies between the distinct binding sites.

The two sets of macroscopic association constants (overall and stepwise) are interrelated by:

$$\beta_1 = K_1; \beta_2 = K_1K_2; \beta_3 = K_1K_2K_3 \quad (11)$$

The set of mass balance equations that needs to be solved to calculate the concentration of free ligand in solution is:

$$\begin{aligned} [P_T] &= [P] + [PL_1] + [PL_2] + [PL_3] = [P] + \sum_{i=1}^n \beta_i [P][L_W]^i \\ [L_T] &= [L_W] + [L_B] = [L_W] + [L_{B^I}] + [L_{B^{II}}] + [L_{B^{III}}] \\ &= [L_W] + [PL_1] + 2[PL_2] + 3[PL_3] = [L_W] + \sum_{i=1}^n i\beta_i [P][L_W]^i \end{aligned} \quad (12)$$

where  $L_{B^x}$  represents ligand bound to site  $x$ .

Protein saturation as a function of the concentration of free ligand is given by:

$$\bar{n}_P = \langle i \rangle = \frac{[L_B]}{[P_T]} = \frac{\sum_{i=1}^n i\beta_i [P][L_W]^i}{[P_T]} \quad (13)$$

where  $\langle i \rangle$  is the average number of ligands bound per protein molecule.

The ligand saturation is given by:

$$\bar{n}_L = \frac{[L_B]}{[L_T]} = \frac{\sum_{i=1}^n i\beta_i [P][L_W]^i}{[L_T]} \quad (14)$$

which corresponds to the fraction of ligand molecules that is bound.



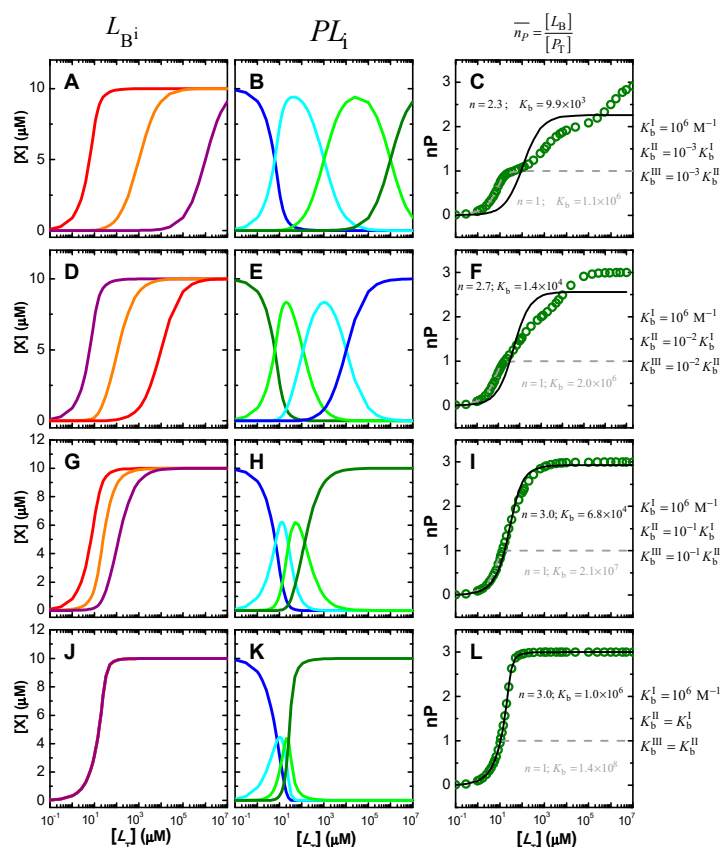
If the binding sites are independent, the stepwise macroscopic and the microscopic association constants are related by:

$$\begin{aligned}
 K_1 &= K_b^I + K_b^{II} + K_b^{III} \\
 K_2 &= \frac{K_b^I K_b^{II} + K_b^I K_b^{III} + K_b^{II} K_b^{III}}{K_b^I + K_b^{II} + K_b^{III}} \\
 K_3 &= \frac{K_b^I K_b^{II} K_b^{III}}{K_b^I K_b^{II} + K_b^I K_b^{III} + K_b^{II} K_b^{III}}
 \end{aligned}
 \tag{15}$$

The derivation of these equations, the relations between the stepwise macroscopic association constants for the case of dependent binding sites, and details regarding the numerical methods used to calculate the free ligand concentration may be found in the literature [4,46,62,67,68].

### 2.2.2. Characterization of the Binding Parameters from Protein Saturation Curves

The saturation of a protein with 3 independent binding sites is shown in Figure 4 with a high affinity for binding site I ( $K_b^I = 10^6 \text{ M}^{-1}$ ), and distinct relations to the other binding sites to which the ligand binds with equal or lower affinity. The total concentration considered for the protein is  $10 \text{ }\mu\text{M}$  and the total ligand concentration is varied to cover the full range of protein saturation.



**Figure 4.** Variation of the concentration of occupied binding sites ( $LB^I$  —,  $LB^{II}$  —, and  $LB^{III}$  —) (A,D,G,J), protein with  $i$  ligands bound ( $P$  —,  $PL_1$  —,  $PL_2$  —, and  $PL_3$  —) (B,E,H,K), and average number of ligands per protein ( $\bar{n}_p$  ○) (C,F,I,L), as a function of  $[L_T]$  for  $[P_T] = 10 \text{ }\mu\text{M}$ . The microscopic association constant of highest affinity was equal to  $10^6 \text{ M}^{-1}$ , while the relative values of the intermediate and lower affinity binding sites were varied from top to bottom. The lines in the right plots (C,F,I,L) are the best fits considering only one type of binding sites, Equation (3), with the number of binding sites fixed at 1 (---) or as an adjustable variable (—), and the estimated values shown in the respective plot.

When the affinities for the distinct binding sites differ by 3 orders of magnitude (upper plots), saturation of the distinct sites occurs at very different ligand concentrations (plot A), and proteins associate with 1, 2 or 3 ligands at different ligand concentrations (plot B), leading to a saturation curve with well-defined steps (plot C). A similar behaviour is observed for binding affinities that differ by 2 orders of magnitude (plots D to F), although the stepwise saturation of protein with ligand becomes less pronounced, with extensive overlap of the ligand concentration range for formation of  $PL_1$ ,  $PL_2$ , and  $PL_3$ . This overlap is even more pronounced when the affinities differ by only one order of magnitude (plots G to I), in which case the saturation steps can no longer be observed in the protein saturation curve. The lower plots show the case where all 3 binding sites in the protein have equal affinity for the ligand. In this case, proteins with only 1 or 2 ligands have a relatively low abundance, and over a very limited ligand concentration range. In this case, the narrowness of the ligand concentration range where the proteins have intermediate numbers of ligands bound arises from the following two reasons. First, taking each binding site from 10 to 90% saturation involves nearly two orders of magnitude variation in  $[L_W]$ , from  $0.11 K_d$  to  $9 K_d$ . When the binding affinities are very different, significant occupation of each binding site occurs at very distinct ligand concentrations, leading to a wide concentration range for protein saturation. In turn, for similar binding affinities, the concentration ranges of significant occupation with ligand overlap. Second, for similar binding affinities and intermediate values of the average protein saturation, it is equally probable that the ligand is associated with different binding sites in the same protein, or to the same or a different type of binding site in distinct proteins, thus decreasing the relative population of protein molecules with intermediate values of ligands bound.

The best fit of the saturation curves with a model considering an adjustable number of equal and independent binding sites per protein (solution  $x_-$  of Equation (3)) is shown in the plots on the right. As expected, a perfect fit and the correct binding parameters ( $n = 3$  and  $K_b = 10^6 \text{ M}^{-1}$ ) are obtained when the binding affinity for the distinct binding sites is the same (plot L). On the other extreme, if the distinct binding sites have well separated ligand affinities, this simplified model leads to clear misfits to the saturation curve, with no risk of accidental inadequate use (plots C and F). However, the full description of the protein saturation curve requires varying the ligand concentration over 8 orders of magnitude. Such a wide range of ligand concentrations is not usually evaluated experimentally. If only 2 or 3 orders of magnitude are considered, only the binding of the first ligand would be observed, leading to a good fit. For binding affinities differing by only an order of magnitude (plot I), the stepwise association is not evident in the shape of the saturation curve, which becomes reasonably well described by the simplified model considering equal and independent binding sites. The best fit leads to a similar value for the number of binding sites, but to an intermediate value of the association constant. The simulations shown in Figure 4 assume that the protein concentration is well known and that protein saturation may be quantitatively and directly followed. Obviously, the model that assumes one binding site per protein is not adequate to describe the saturation curve in this case (dashed grey lines).

The analysis of the simulations depicted in Figure 4 shows that when the number of ligands per protein may be directly obtained, the analysis of the results with a model considering a variable number of equal and independent binding sites is a good first option. If the model adequately describes the shape of the saturation curve (plots I and L in this case study), the association constant estimated corresponds to the microscopic association constant if the binding sites are equal and independent. If the binding affinity for the binding sites is different, the value obtained is close to the average value (within the same order of magnitude) of all association constants. If the model does not adequately describe the shape of the saturation curve, the values obtained are therefore devoid of physical meaning, and more complicated models must be used [56,62–65,67–70]. The accurate description of the saturation curve permits obtainment of the global association constants, Equation (9), and the stepwise macroscopic association constants, Equation (10). If the binding sites are

independent, the number of microscopic and macroscopic binding constants is the same and it may in principle be possible to obtain one set of constants from the other. However, if the binding sites are not independent, the number of microscopic binding constants exceeds the number of parameters that may be obtained from the protein saturation curves, and therefore cannot be quantitatively characterized. In this case, only some relations between them may be obtained, as well as possible indication of cooperativity for binding to the distinct binding sites.

### 2.2.3. Scatchard Plots and Evaluation of Cooperativity

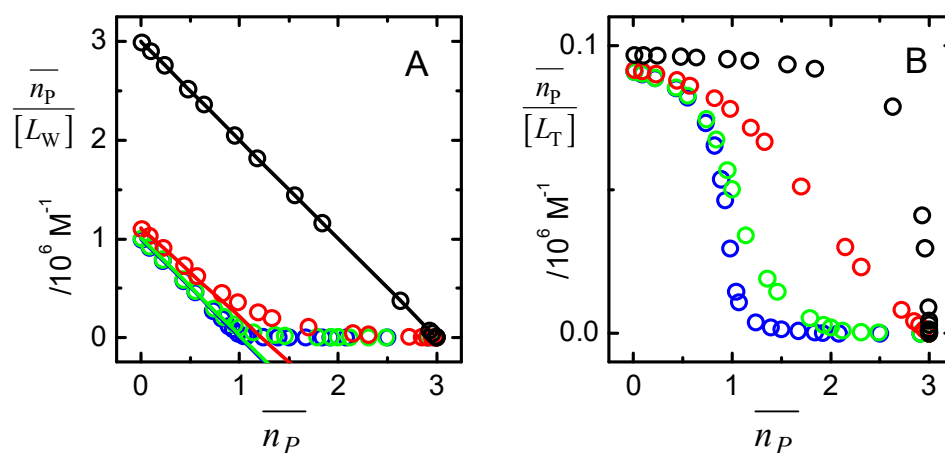
The analysis of the protein saturation curve with complex models requires the use of complicated fitting procedures, including numerical methods to obtain the concentration of free ligand and protein, the equivalent to Equation (12) but adjusted to the number of binding sites. To avoid the laborious procedure of data analysis (or, in an attempt to complement the analysis), specific graphical representations of the data are usually performed to evaluate for the presence of cooperativity between the binding sites. Scatchard plots are a prominent example, where  $\bar{n}_P/[L_W]$  is represented as a function of  $\bar{n}_P$ . A linear dependence between both variables is expected for the case of a protein with  $n$  independent binding sites, and deviations from linearity are interpreted as indicative of cooperativity.

$$\frac{\bar{n}_P}{[L_W]} = nK_b - \bar{n}_PK_b \quad (16)$$

Before analysing the results obtained with Scatchard plots, some remarks must be made regarding this and other linearization procedures. A first big limitation is that, because the binding constants are still not known (the objective is to obtain them from this analysis), if the concentrations of bound or free ligand are not directly measured,  $[L_W]$  cannot be calculated. In these situations,  $[L_T]$  is commonly used, leading to incorrect interpretations. Another important concern is that these linearization procedures do not comply with the assumptions imposed by non-linear least-square regression analysis, e.g., that the  $x$ -axis must be an independent variable with negligible error. Also, while the uncertainty associated with concentrations may be assumed to be normally distributed and to be constant within the range analysed (allowing the use of regressions not weighted by the uncertainty in the variables represented in the  $x$  and  $y$  axis), this is not the case for the inverse of concentration (which shows a much higher uncertainty for large values of the variable, corresponding to low concentrations). The confidence intervals are also not symmetric for inverse concentrations, with both properties leading to best fits that do not reflect the average behaviour of the system [71]. Finally, those linearization procedures were introduced at a time when computers were not available for data analysis. While the risk of misinterpreting the data through the use of linearization procedures was acceptable then, it is certainly not justified now.

To further evaluate the adequacy of data linearization and representation in Scatchard plots, this procedure was applied to the case of 3 independent binding sites in the conditions considered previously for Figure 4. The results are shown in Figure 5. Plot A represents the protein saturation divided by the concentration of free ligand, as a function of protein saturation. As expected, when the ligand affinity to all binding sites is the same (black), a linear dependence is observed with the intercept in the  $y$ -axis being equal to  $nK_b$  ( $3 \times 10^6 \text{ M}^{-1}$ ), the  $x$ -axis intercept being  $n$ , and the slope equal to  $K_b$ . In the other extreme, when the binding affinities differ by 3 orders of magnitude (blue), only the population of the highest affinity site is captured (although in the concentration range analysed they all become occupied, as indicated in the  $x$ -axis that goes up to  $\bar{n}_P = 3$ ), with a linear variation leading to  $nK_b(1.0 \times 10^6 \text{ M}^{-1})$  and  $n = 1.0$ . A similar behaviour is observed when the affinities differ by two orders of magnitude (green), the best linear fit for low values of  $\bar{n}_P$  leading to  $nK_b(9.9 \times 10^5 \text{ M}^{-1})$  and  $n = 1.0$ . When the binding affinities differ by only one order of magnitude (red), non-linear Scatchard plots are obtained, with the best fit for small  $\bar{n}_P$  leading to  $nK_b(9.0 \times 10^5 \text{ M}^{-1})$  and  $n = 1.2$ . The non-linear dependence may be

incorrectly interpreted as negative cooperativity, despite resulting from binding to different, but independent, binding sites. Therefore, unless the protein being studied is known to have several equivalent binding sites (as is usually the case for multimeric proteins) and the sample is of high purity, negative cooperativity cannot be inferred from non-linear Scatchard plots.



**Figure 5.** Scatchard plots for the simulation results considering ligand association to proteins with 3 independent binding sites with the microscopic association constant of highest affinity  $K_b^I = 10^6 \text{ M}^{-1}$ , and different relative values of the intermediate and lower affinity binding sites:  $K_b^{\text{III}} = K_b^{\text{II}} = K_b^I$  ( $\circ$ ),  $K_b^{\text{III}} = 10^{-1} K_b^{\text{II}}$ ;  $K_b^{\text{II}} = 10^{-1} K_b^I$  ( $\odot$ ),  $K_b^{\text{III}} = 10^{-2} K_b^{\text{II}}$ ;  $K_b^{\text{II}} = 10^{-2} K_b^I$  ( $\ominus$ ),  $K_b^{\text{III}} = 10^{-3} K_b^{\text{II}}$ ;  $K_b^{\text{II}} = 10^{-3} K_b^I$  ( $\odot$ ). The lines in plot A are best fits of straight lines to the region of low protein saturation. Note that in plot (A) the protein saturation in the  $y$ -axis is divided by the concentration of free ligand, while in plot (B) the total ligand concentration is used.

As expected, if the total concentration of ligand is used instead of the free concentration (plot B), the information obtained is meaningless and misleading. In this case, non-linear dependences are obtained for all cases analysed, with the obtained downward curvatures suggesting positive cooperativity. Unfortunately, this incorrect data analysis may be found in the literature. Hopefully this review will alert inexperienced readers to possible pitfalls and will contribute to a proper data analysis.

#### 2.2.4. Indirect Characterization of the Binding Parameters Indirectly from Changes in a Protein Property

In the previous analysis, it was assumed that the number of ligands per protein could be directly characterized. However, if protein saturation is not directly accessed, but rather inferred from a property that varies with protein saturation ( $\Delta S$ ), the number of binding sites can only be evaluated from the analysis of the results, and inadequate procedures may lead to an incorrect description of the system. This will be addressed in the current section, considering proteins with 3 independent binding sites and several cases with distinct relative values of the binding affinities.

The signal from the protein at a given concentration of ligand,  $S$ , depends on the concentration of all different proteins species,  $[PL^i]$ , and on the signal characteristic of each species,  $s_i$ , Equation (17).

$$\begin{aligned} \Delta S &= S - S_0 = s_0[P] + \sum_{i=1}^N s_i [PL^i] - s_i [P_T] \\ &= \sum_{i=1}^N (s_i - s_0) [PL^i] \end{aligned} \quad (17)$$

Note that  $PL^i$  does not represent a protein with  $i$  ligands, which would be  $PL_i$ , in the notation followed in this work. A protein with a given number of ligands corresponds to

many distinct species, depending on the binding sites to which the ligands are bound. Thus, in Equation (17),  $N$  is not equal to the number of binding sites ( $n$ ), but rather to the number of possible combinations of bound ligands. Assuming that the same signal is obtained for all combinations with the same number of bound ligands, Equation (17) becomes

$$\Delta S = \sum_{i=1}^n (s_i - s_0) [PL_i] \quad (18)$$

and may be expressed in terms of the macroscopic binding associations, Equation (19).

$$\Delta S = \sum_{i=1}^n (s_i - s_0) \beta_i [P][L_W]^i \quad (19)$$

In this section, it is considered that the signal variation is the same when the ligand binds to any of the binding sites, and that the magnitude of the signal variation is proportional to the number of bound ligands,  $(s_i - s_0) = \Delta s_{\max}(i/n)$ . This is a good approximation when properties such as the overall mass or charge of the ligand–protein complex are being followed. However, the observed property may vary differently, and this aspect will be discussed in Section 2.3 in the context of quenching of protein fluorescence by bound ligand.

From the analysis of the signal variation as a function of the ligand concentration, it is in principle possible to obtain the intrinsic signal variation for each number of bound ligands ( $s_i - s_0$ ), and the macroscopic association constants  $\beta_i$ . In turn, from the relations between the macroscopic association constants, some information may be obtained regarding the similarity and/or dependency of the different binding sites, Equation (20).

$$\frac{\left(\frac{\beta_i}{\binom{n}{i}}\right)^{1/i}}{\left(\frac{\beta_{i-1}}{\binom{n}{i-1}}\right)^{1/(i-1)}} = 1 \text{ identical and independent binding sites}$$

$$< 1 \text{ non – identical binding sites}$$

$$\text{or identical but not independent (negative cooperativity)} \quad (20)$$

$$> 1 \text{ non – identical binding sites}$$

$$\text{or identical but not independent (positive cooperativity)}$$

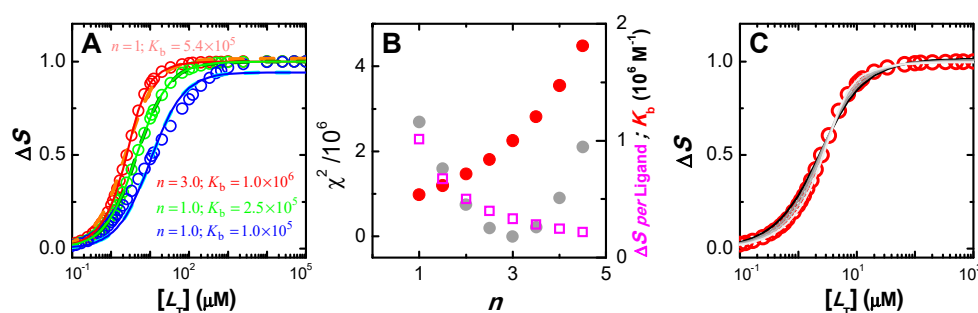
In the case of equal and independent binding sites, the macroscopic association constants depend on the microscopic association constant according to [4]:

$$\beta_i = \prod_{i=1}^n K_i \quad (21)$$

$$K_i = \frac{n-i+1}{i} K_b$$

The simulations shown in Figure 6A were performed for  $[P_T] = 1 \mu\text{M}$ ,  $K_b^I = 10^6 \text{ M}^{-1}$ , and at different relative values of the binding affinity to the other two binding sites in the protein ( $K_b^I/K_b^{II} = K_b^{II}/K_b^{III} = 10$  in blue;  $K_b^I/K_b^{II} = K_b^{II}/K_b^{III} = 3$  in green, and  $K_b^I/K_b^{II} = K_b^{II}/K_b^{III} = 1$  in red). The lines correspond to the best fit of Equation (3) assuming an adjustable number of equal and independent binding sites ( $n \geq 1$ , solid thin lines), or considering only one binding site per protein ( $n = 1$ , dashed thick lines). The two sets of fitting curves cannot be visually distinguished and lead to the same parameters for the cases considered of binding sites with different affinities (curves in blue and green). The quality of the best fit for different values of  $n$  is presented in Figure S4, showing a small improvement as  $n$  decreases. The fitting parameters are thus the same for both models, with the association constant obtained from the best fit being similar to that of

the binding site with intermediate affinity ( $K_b^{\text{II}} = 10^5 \text{ M}^{-1}$ ) and  $n = 1$ . However, different parameters are obtained for the best fit of the simulation results when all binding sites are equal (Figure 6, curve and parameters in red). In this case, the best fit with adjustable  $n$  shows a minimum in the squared deviation for  $n = 3$  (plot B), with reasonable sensitivity for both lower and higher values of  $n$ . Nevertheless, the representation of the best fit curves for values of  $n$  from 1 to 4 (plot C) shows that the fitting sensitivity is insufficient for the accurate characterization in the case of experimental data due to the unavoidable presence of noise (here represented as large symbols for the simulated results).

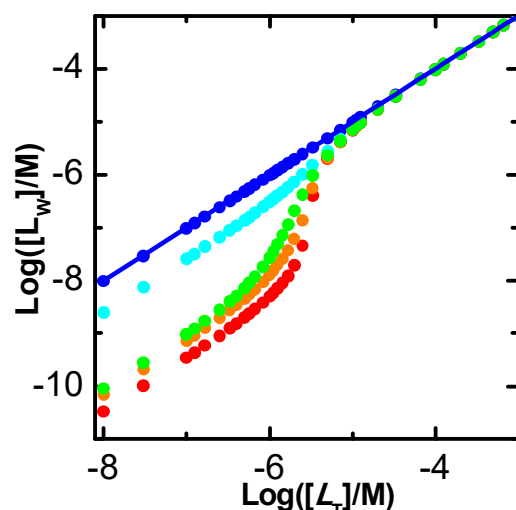


**Figure 6.** Plot (A): variation of a hypothetical signal proportional to the number of ligands associated with a protein containing 3 independent binding sites:  $K_b^{\text{I}} = 10^6 \text{ M}^{-1}$ ; and  $K_b^{\text{I}}/K_b^{\text{II}} = K_b^{\text{II}}/K_b^{\text{III}} = 10$  ( $\circ$ ),  $K_b^{\text{I}}/K_b^{\text{II}} = K_b^{\text{II}}/K_b^{\text{III}} = 3$  ( $\circ$ ), and  $K_b^{\text{I}}/K_b^{\text{II}} = K_b^{\text{II}}/K_b^{\text{III}} = 1$  ( $\circ$ ). The lines are the best fits considering only one type of binding sites, Equation (3), with the number of binding sites fixed at 1 (thick dashed lines) or as an adjustable variable (thin continuous lines). The parameters obtained from the best fit are also shown. Plot (B): effect of the number of binding sites per protein in terms of the quality of the best fit ( $\bullet$ ), association constant ( $\bullet$ ) and signal variation per ligand bound ( $\square$ ), for the case  $K_b^{\text{I}}/K_b^{\text{II}} = K_b^{\text{II}}/K_b^{\text{III}} = 1$ . Plot (C): Fitting curves obtained for  $n = 1$  ( $\bullet$ ),  $n = 2$  ( $\bullet$ ),  $n = 3$  ( $\bullet$ ), and  $n = 4$  ( $\bullet$ ).

In all cases considered, the magnitude of the signal variation per ligand bound decreases as the number of binding sites considered increases, such that the total variation in the signal is obtained only when the protein is fully saturated (Figures 6B and S4, plots A and C). Somewhat surprisingly, the association constant obtained from the best fit decreases with the decrease in the number of binding sites, leading to a large variation in the overall protein binding affinity obtained from the best fit:  $nK_b$  varies from  $5.4 \times 10^5 \text{ M}^{-1}$  for  $n = 1$  to  $5.9 \times 10^6 \text{ M}^{-1}$  for  $n = 4$  in the case of equal binding sites, with the correct description of the system for  $n = 3$  leading to  $nK_b = 3.0 \times 10^6 \text{ M}^{-1}$ . However, the quality of the distinct best fits to experimental results would not allow a discrimination between the distinct set of parameters, and the simpler model would thus be selected. This would underestimate the global binding affinity by nearly one order of magnitude. Attempting to understand this effect, distinct simulations have been performed at lower and higher affinities. The results are shown in Figure S5. For  $K_b^{\text{I}} = 10^4 \text{ M}^{-1}$  (plots A to C), the quality of the best fit with a variable number of equal and independent binding sites is even less sensitive to  $n$ , with  $K_b$  being always approximately  $1 \times 10^4 \text{ M}^{-1}$  for the case of equal binding sites (plot B and C), and close to the average  $K_b^{\text{I}}$  for different affinity binding sites. In the case of high-affinity binding,  $K_b^{\text{I}} = 10^8 \text{ M}^{-1}$  (plot D), a high sensitivity to  $n$  is observed for the case of equal binding sites (plots E and F) and for  $K_b^{\text{I}}/K_b^{\text{II}} = K_b^{\text{II}}/K_b^{\text{III}} = 3$  (plots G and H), but not for  $K_b^{\text{I}}/K_b^{\text{II}} = K_b^{\text{II}}/K_b^{\text{III}} = 10$  (plots I and J).

The analysis of the data presented in Figure 6 and S5 shows that the sensitivity of the best fit to  $n$  depends on the fraction of ligand associated with the protein, which is presented in Figure 7 for the various conditions considered. If the ligand is in large excess, the shape of the protein saturation curve is not dependent on the considered number of binding sites. In this case, it is not possible to obtain the number of binding sites from the best fit of the saturation curve. On the other hand, if the fraction of ligand bound to the protein is significant, increasing the number of binding sites leads to a decrease

in the fraction of ligand free in solution, and consequently to differences in the shape of the protein saturation curve. In this case, the protein saturation curve is a straight line until almost full saturation. When the data is tentatively described by a model with fewer binding sites, the predicted fraction of free ligand is higher and the saturation curve linearly reaches saturation at lower total ligand concentrations. The best fit then obtained underestimates saturation at high values and overestimates it at low saturation, leading to a lower association constant.



**Figure 7.** Variation of the concentration of free ligand in solution as a function of the total ligand concentration, for  $[P_T] = 1 \mu\text{M}$  and considering 3 equal and independent binding sites with  $K_b$  equal to  $10^4 \text{ M}^{-1}$  (●),  $10^6 \text{ M}^{-1}$  (●) or  $10^8 \text{ M}^{-1}$  (●), and for different binding sites with  $K_b^I = 10^8 \text{ M}^{-1}$ , and  $K_b^I/K_b^II = K_b^II/K_b^III = 3$  (●), or  $K_b^I/K_b^II = K_b^II/K_b^III = 10$  (●). The blue line is for  $[L_w] = [L_T]$ .

In conclusion, when protein saturation with ligand is followed indirectly for proteins with several similar binding sites, information regarding the number of binding sites and their respective association constants can only be obtained if binding is limited by ligand availability. Otherwise, the dependence of the property on the total ligand concentration follows a hyperbola and is insensitive to the number of binding sites in the protein. In this situation, there is no justification to use complex models, and the simplest one ( $n = 1$ ) is usually used. The association constant obtained is similar to or smaller than the true binding association, thus leading to a significant underestimation of the fraction of ligand bound to the protein. If binding is limited by ligand availability but the data is noisy, it will also be impossible to distinguish between models with different  $n$ . The description of the system with the simplest model may in this case lead to a severe underestimation of the amount of ligand associated with the protein. It is therefore essential to perform the experiment at different protein concentrations, to check for consistency and find the most adequate model and parameter values.

### 2.3. Quenching of Protein Fluorescence: Stern–Volmer Plots

The characterization of binding to proteins through ligand quenching of protein fluorescence is a method that is very commonly used and deserves special attention. The fluorescence decrease is usually described by the Stern–Volmer equation,

$$\frac{S_0}{S} = 1 + K_{SV}[L_w] \quad (22)$$

where  $S_0$  and  $S$  represent the fluorescence intensity in the absence and in the presence of ligand, respectively. The latter depends on the concentration of all different protein species,  $[PL^i]$ , and on the signal characteristics of each species,  $s_i$ , Equation (17) in Section 2.2.4.

The equation for  $S_0/S$  is thus:

$$\frac{S_0}{S} = \frac{s_0[P_T]}{s_0[P] + \sum_{i=1}^N s_i[PL^i]} \quad (23)$$

Remember that  $PL^i$  does not represent a protein with  $i$  ligands, but rather each different combination of  $i$  ligands bound to the distinct binding sites. Thus, in Equations (17) and (23)  $N$  is not equal to the number of binding sites ( $n$ ), but rather to the number of possible combinations of bound ligands. Assuming that the same quenching efficiency is obtained for all combinations with the same number of bound ligands, Equation (23) may be expressed in terms of the macroscopic binding associations, where  $N$  is now equal to  $n$ :

$$\frac{S_0}{S} = \frac{[P_T]}{[P] + \sum_{i=1}^n \frac{s_i}{s_0} [PL_i]} = \frac{[P] + \sum_{i=1}^n \beta_i [P][L_w]^i}{[P] + \sum_{i=1}^n \frac{s_i}{s_0} \beta_i [P][L_w]^i} = \frac{1 + \sum_{i=1}^n \beta_i [L_w]^i}{1 + \sum_{i=1}^n \frac{s_i}{s_0} \beta_i [L_w]^i} \quad (24)$$

It should be highlighted that the validity of this approximation is difficult to evaluate and will depend both on the protein and ligand. For example, for proteins with several fluorescent groups (e.g., bovine serum albumin, which has two tryptophan residues), the fluorescence quantum yield of each group may be different, and each of them will most likely be quenched with different efficiencies by ligands in the distinct binding sites. This approximation is also invalid if ligand binding at some of the binding sites in the protein causes conformational changes that affect the protein fluorescence quantum yield and quenching efficiency by subsequent ligands. In the following analysis the possible non-validity of this approximation will be ignored, but this should be kept in mind whenever the results obtained by protein fluorescence quenching do not agree with those from other approaches.

Equation (24) is further simplified if it may be considered that the residual protein fluorescence is zero when any of the binding sites is occupied with ligand ( $s_i = 0$  for  $i \neq 0$ , corresponding to 100% efficient fluorescence quenching by the ligand):

$$\frac{S_0}{S} = \frac{[P_T]}{[P]} = \frac{[P] + \sum_{i=1}^n \beta_i [P][L_w]^i}{[P]} = 1 + \sum_{i=1}^n \beta_i [L_w]^i \quad (25)$$

Although usually assumed, full quenching of the protein fluorescence by the ligand is frequently not supported experimentally, and in fact it is commonly not expected. Attention will therefore be given to the consequences of its non-validity later in this section.

The relation between the Stern–Volmer constant and the binding constants to the different binding sites may be obtained by comparing Equation (25) with Equation (22), leading to:

$$K_{SV} = \frac{\sum_{i=1}^n \beta_i [L_w]^i}{[L_w]} \quad (26)$$

If the protein has only one binding site, Equation (26) simplifies to Equation (27), and the Stern–Volmer constant corresponds to the association constant:

$$K_{SV} = \beta_1 = K_b \quad (27)$$

Thus, for proteins with a single binding site, when ligand binding leads to full quenching of protein fluorescence, and if the free (not the total) ligand concentration is used as the independent variable, the slope of a linear Stern–Volmer plot is equal to the association binding constant. If any of the above conditions is violated,  $K_{SV}$  will deviate from  $K_b$ .



If the proteins contain several binding sites, there is no simple correspondence between  $K_{SV}$  and  $K_b$ . In fact, in this case  $K_{SV}$  is not a constant, its value is dependent on the concentration of ligand, and thus non-linear Stern–Volmer plots are obtained. This is frequently observed when characterizing ligand association with proteins containing several binding sites, such as serum albumin. Although this is the expected behaviour, the authors usually analyse the results with the simplified Equation (22) which is unable to capture the observed behaviour and leads to misinterpretation. However, some relevant information may be obtained when the linear range of  $S_0/S$  variation is considered (that is, for low ligand concentrations). In this case, Equation (26) simplifies to Equation (28). Therefore, in the case of proteins with several binding sites with full quenching by the ligand in any of the binding sites, the value of  $K_{SV}$  obtained from the linear fit at low ligand concentrations provides an estimate of the overall binding affinity.

$$K_{SV}(\text{small } [L_W]) = \beta_1 = \sum_{i=1}^n K_b^i \quad (28)$$

In the simplest case, with  $n$  identical and independent binding sites,  $\beta_1 = n K_b$ .

If the quenching efficiency is not 100%, the limit at low ligand concentrations may still provide useful information. From the general Equation (23) one may obtain Equation (29). The derivation of Equations (28) and (29) is provided in Section S3 of the Supplementary Materials.

$$K_{SV}(\text{small } [L_W]) = \sum_{i=1}^n \frac{s_0 - s_i}{s_0} K_b^i \quad (29)$$

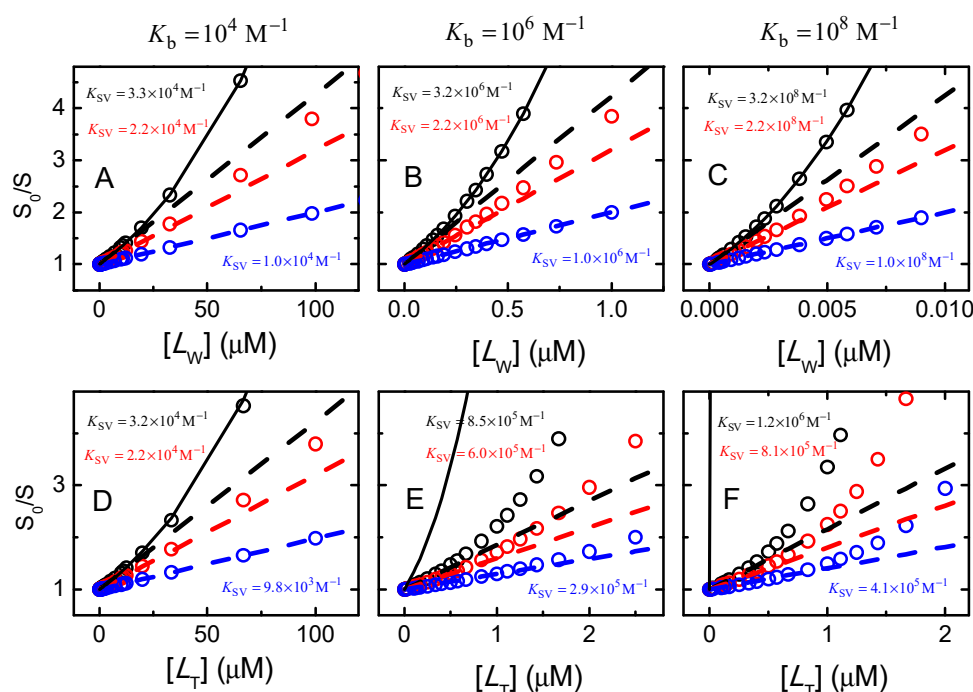
The quenching efficiency is usually an unknown parameter, which prevents researchers from using Equation (29). If Equation (28) is used but the quenching efficiency is not 100%, the value obtained for the parameter  $\beta_1$  will underestimate the true value. For example, if the quenching efficiency is only 10% ( $s_i = 0.9$ ), the estimated value for  $\beta_1$  will be an order of magnitude lower than its true value. Higher quenching efficiencies lead to a more accurate estimation of  $\beta_1$ .

To exemplify some expected Stern–Volmer plots for quenching of a protein with several binding sites, a set of conditions has been simulated for ligand binding to a protein with 3 binding sites, Figures 8 and 9. In Figure 8 the 3 binding sites are considered equal and independent, with the ligand affinity changing from low ( $K_b = 10^4 \text{ M}^{-1}$ ), to intermediate ( $K_b = 10^6 \text{ M}^{-1}$ ), and high ( $K_b = 10^8 \text{ M}^{-1}$ ). However, the possibility of different quenching efficiencies for the ligand bound to each binding site was accounted for in the examined situations: 100% quenching efficiency regardless of the binding site to which the ligand is associated ( $s_i = 0$  for  $i = I$  to  $III$ , circles in black),  $s_I = 0.6$ ,  $s_{II} = 0.3$ , and  $s_{III} = 0$  (circles in red), and only ligand in binding site  $III$  leading to fluorescence quenching  $s_I = s_{II} = 1$ , and  $s_{III} = 0$  (circles in blue). For the case of distinct quenching efficiencies, it was considered that when two binding sites are occupied the fluorescence quenching efficiency is given by the product of the residual fluorescence when the ligand is in each of the binding sites ( $s_{I\&II} = s_I s_{II}$ ). The concentration of protein with a given combination of ligands ( $[PL^i]$ ) was calculated from  $[PL_i]$  and the relative value of the binding affinities that lead to that ligand occupation. As an example, the concentration of protein with ligand only in binding site  $I$  is equal to  $[PL_1]K_b^I / (K_b^I + K_b^{II} + K_b^{III})$ , and with ligand in binding sites  $I$  and  $II$  is equal to  $[PL_2](K_b^I + K_b^{II}) / (2(K_b^I + K_b^{II} + K_b^{III}))$ .

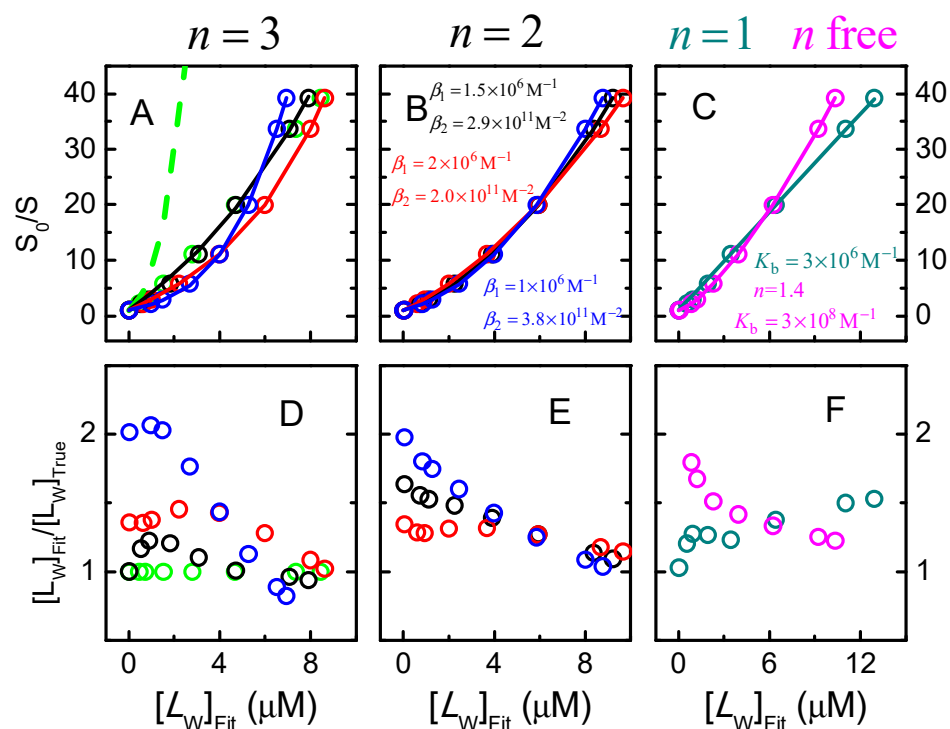
The first observation from the results shown in Figure 8 is that Stern–Volmer plots with an upward curvature are generally obtained. The sole exceptions occur when only the ligand bound to site  $III$  leads to fluorescence quenching (blue circles). This situation is formally equivalent to having only one binding site (i.e., the other binding sites are “silent”), and in this case a linear dependence of  $S_0/S$  with  $[L_W]$  is in fact expected, Equation (27). However, if (incorrectly) the total concentration of ligand is considered (bottom plots), an upward curvature is observed even in this case when the binding affinity is intermediate or high (plots E and F).

Due to the non-linear Stern–Volmer plots, the value of  $K_{SV}$  obtained from the best fit of the simplified Equation (22) depends on the considered range of ligand concentrations and the corresponding magnitude of fluorescence variation. The dashed lines shown in Figure 8 were obtained from the best fit for  $S_0/S$  between 1 and 1.1. When the concentration of free ligand is considered (upper plots), the values obtained for  $K_{SV}$  lead to reasonable estimates of  $K_b$  (within the same order of magnitude) for all considered conditions, in agreement with the predictions from Equations (28) and (29). As predicted from Equation (29), the value of  $K_{SV}$  depends on the extent of protein fluorescence quenching by the ligand. The larger the quenching efficiency, the higher the value obtained for  $K_{SV}$ . The upper limit of  $K_{SV}$  is obtained for  $s_i = 0, i \neq 0$  and should have been equal to  $K_b^I + K_b^{II} + K_b^{III}$ , Equation (28). However, somewhat higher values are obtained in some situations due to the upward curvature of  $S_0/S$ .

The above discussion shows that the order of magnitude of the binding affinity may be obtained from the analysis of the linear region of Stern–Volmer plots (Equations (22), (27), (28) or (29), depending on the specific conditions) if the concentration of free ligand is known. However, unless measured directly, the concentration of free ligand cannot be known before calculation of the binding affinity. If the total concentration of ligand is considered instead, deviations by several orders of magnitude may be obtained (e.g., for  $K_b = 10^8 \text{ M}^{-1}$ ;  $K_{SV}([L_T])$  is  $1.2 \times 10^6 \text{ M}^{-1}$  if  $s_i = 0, i \neq 0$ ,  $8.1 \times 10^5 \text{ M}^{-1}$  if  $s_I = 0.6, s_{II} = 0.3, s_{III} = 0$ , and  $4.1 \times 10^5 \text{ M}^{-1}$  if  $s_I = s_{II} = 1, s_{III} = 0$ ).



**Figure 8.** Stern–Volmer plots for the quenching of protein fluorescence due to ligand binding. Three equal and independent binding sites in the protein are considered, with low ( $K_b = 10^4 \text{ M}^{-1}$ , plot (A,D)), intermediate ( $K_b = 10^6 \text{ M}^{-1}$ , plot (B,E)) or high ( $K_b = 10^8 \text{ M}^{-1}$ , plot (C,F)) affinity. 100% quenching of protein fluorescence is considered for the ligand associated with any of the binding sites ( $s_i = 0, i \neq 0$ ,  $\circ$ ), 40%, 70% or 100% for ligand in binding site I, II, or III respectively ( $s_I = 0.6, s_{II} = 0.3, s_{III} = 0$ ,  $\square$ ), or no quenching by ligands in binding sites I or II, and 100% quenching by ligand in binding site III ( $s_I = s_{II} = 1, s_{III} = 0$ ,  $\triangle$ ). The values obtained for  $K_{SV}$  from the slope at  $S_0/S \leq 1.1$  (Equation (22), dashed lines) are given in each plot with the corresponding colours. The continuous lines are the predictions from Equation (25), considering the “true” binding parameters. In the upper plots the free ligand concentration is considered, while the bottom plots consider the total concentration of ligand.



**Figure 9.** Stern–Volmer plots obtained for the best fit of the results simulated for ligand binding to a protein with 3 equal and independent binding sites ( $K_b = 10^6 \text{ M}^{-1}$ ; corresponding to  $\beta_1 = 3 \times 10^6 \text{ M}^{-1}$ ,  $\beta_2 = 3 \times 10^{12} \text{ M}^{-2}$ , and  $\beta_3 = 1 \times 10^{18} \text{ M}^{-3}$ ) and partial fluorescence quenching by the ligand in sites I and II ( $s_I = 0.6$ ,  $s_{II} = 0.3$ ,  $s_{III} = 0$ ). The upper plots show  $S_0/S$  as a function of the concentration of free ligand obtained from the best fit  $[L_W]_{\text{FIT}}$ , and the ratio of  $[L_W]_{\text{FIT}}$  and the “true” concentration of free ligand is shown in the lower plots. Plot (A)  $S_0/S$  obtained for the binding and quenching parameters considered in the simulation ( $\circ$ ), and predicted by Equation (25) for the same values of  $\beta_i$  ( $-$ ). Some examples of best fit obtained for different values of  $\beta_i$  are also shown:  $\beta_1 = 3 \times 10^6 \text{ M}^{-1}$  (fixed), leading to  $\beta_2 = 1.2 \times 10^{11} \text{ M}^{-2}$ , and  $\beta_3 = 1.5 \times 10^{16} \text{ M}^{-3}$  ( $\circ$  and  $-$ );  $\beta_1 = 2 \times 10^6 \text{ M}^{-1}$  (fixed), leading to  $\beta_2 = 1.4 \times 10^8 \text{ M}^{-2}$ , and  $\beta_3 = 3.3 \times 10^{16} \text{ M}^{-3}$  ( $\circ$  and  $-$ ); and  $\beta_1 = 1 \times 10^6 \text{ M}^{-1}$  (fixed), leading to  $\beta_2 = 2.3 \times 10^8 \text{ M}^{-2}$ , and  $\beta_3 = 9.4 \times 10^{16} \text{ M}^{-3}$  ( $\circ$  and  $-$ ). Several combinations of association constants considering 2 binding sites are shown in Plot (B), and considering 1 or an adjustable number of binding sites are shown in Plot (C). The resulting association constants are given in the plot with the same colour as that of the respective curves.

To increase the accuracy of  $K_{SV}$  as an estimate of the protein-binding ability, its evaluation must include the calculation of the corresponding free ligand concentrations. This approach is not feasible for proteins with several binding sites with distinct ligand affinities and/or binding cooperativity, because in these cases  $[L_W]$  must be calculated numerically for each evaluation of  $K_{SV}$  (and corresponding  $K_b$ ). A reasonable compromise between accuracy and feasibility may be achieved by considering that all binding sites are equal and independent. In this case,  $[L_W]$  may be calculated from  $[L_T] - [L_B]$ , with  $[L_B]$  being the  $x_-$  solution of Equation (3). In this case (considering equal and independent binding sites), the binding affinity for the binding sites ( $K_b$ ) may be obtained from the dependence of  $S_1/S$  vs.  $[L_W]$  through several alternative approaches: (i)  $K_{SV}$  may be obtained from the best fit of a straight line at low quenching,  $S_1/S$  between 1 and 1.1 being usually a good balance between sensitivity and little deviation from linearity. The binding constant  $K_b$  is then obtained from  $K_{SV}$  using Equation (28), considering 100% quenching efficiency ( $s_i = 0$ ,  $i \neq 0$ ), or (29) if the quenching efficiency is known. In this approach it is assumed that each protein has  $n$  equal and independent binding sites, with the corresponding  $[B_T]$  in Equation (3) being given by  $n[P_T]$ . (ii)  $K_b$  may also be obtained from the best fit of the

whole range of  $S_1/S$ , assuming that the same quenching efficiency is attained for proteins with 1 or more ligands bound ( $s_i/s_0 = s_x/s_0$ ,  $i \neq 0$ ), and with  $S_1/S$  given by Equation (30).

$$\frac{S_0}{S} = \frac{1 + \sum_{i=1}^n \beta_i [LW]^i}{1 + \frac{s_x}{s_0} \sum_{i=1}^n \beta_i [LW]^i} \quad (30)$$

with  $\beta_i$  being related with  $K_b$  through Equation (21) in Section 2.2.4, for the case of  $n$  equal and independent binding sites.

A Microsoft Excel file has been developed for the analysis of Stern–Volmer-like plots for protein–ligand binding experiments and will be provided to interested readers upon request.

The fitting procedure indicated above was applied to the simulated data to evaluate its applicability limits. For low binding affinity ( $K_b = 10^4 \text{ M}^{-1}$ ) and 100% quenching efficiency, the best linear fits to  $S_1/S < 1.1$  lead to good estimates of  $n K_b$ , independently of the number of binding sites considered ( $n = 1$  to 4). When the whole curve was analysed, assuming  $n = 1$  leads to a poor fit and overestimation of  $n K_b$ . A better fit and more accurate parameter estimates were obtained for  $n = 3$ , suggesting that this analysis may provide some information on the number of binding sites *per* protein. As observed before (Figure 8), as the quenching efficiency is progressively decreased,  $n K_b$  tends to be underestimated, both from the linear best fit at low  $S_1/S$  or from the whole range of  $S_1/S$ . An equivalent situation is obtained for intermediate binding affinities ( $K_b = 10^6 \text{ M}^{-1}$ ), with the linear best fit leading to good estimates and some discrimination regarding the number of binding sites when the whole curve is analysed. For high binding affinities and 100% quenching efficiency, only  $n = 1$  could be used for the best linear fit to  $S_1/S < 1.1$ . Setting  $n > 1$  leads to poor sensitivity and estimates of  $n K_b$  diverging to infinity. This is due to the upward curvature that occurs even at very low quenching. Again, as the quenching efficiency is decreased, the estimates obtained for  $n K_b$  deviate significantly from the true values, underestimating this parameter by more than two orders of magnitude for  $s_I = s_{II} = 1$ ,  $s_{III} = 0$ . The values obtained for the parameters of all conditions considered are shown in the supplementary materials (Table S1). Importantly, the previous results show that the analysis of Stern–Volmer plots *per se* underestimates the values of high-affinity binding constants, and thus cannot be relied upon for their estimation. For high-affinity binding, the direct quantification of the free ligand concentration is required. This allows understanding of the observation that the binding affinities obtained from the analysis of Stern–Volmer plots are never much higher than  $10^6 \text{ M}^{-1}$ , leading to dependences of  $n K_b$  for homologous series (e.g., with increased hydrophobicity) that level-off close to  $10^6 \text{ M}^{-1}$  [72].

In the above analysis, the models used considered that the binding sites in the protein were equal and independent. This was justified when describing the simulated results as they were in fact obtained for equal and independent binding sites. This is usually not the case in real systems where the complexity of the system may be significantly higher, with the binding sites in the protein being different and/or displaying cooperativity. If the general Equation (25) is used without any assumption regarding the relation between the macroscopic binding constants, from their relations the dissimilarity of the binding sites or eventual cooperativity in ligand binding could in principle be evaluated, Equation (20) in Section 2.2.4. However, this procedure is strongly discouraged because it could lead to severe misinterpretation of the data. To elucidate this, the simulations obtained assuming that the protein contains 3 equal and independent binding sites with  $K_b = 10^6 \text{ M}^{-1}$  and a fluorescence quenching efficiency dependent on the binding site ( $s_1 = 0.6$ ,  $s_2 = 0.3$ ,  $s_3 = 0$ ) was analysed with several possible combinations of  $\beta_i$ . The results obtained are shown in Figure 9.

As expected, the variation of  $S_0/S$  predicted from Equation (25) for the set of binding parameters used to obtain the simulation,  $n = 3$  and  $K_b^i = 1.0 \times 10^6 \text{ M}^{-1}$  for all  $i$  (green line on plot A), does not accurately describe the simulation results. This is because only partial fluorescence quenching was considered,  $s_1 = 0.6$ ,  $s_2 = 0.3$ ,  $s_3 = 0$ , which is not in

agreement with the approximations required to obtain Equation (25). To obtain a good fit it is necessary to adjust the binding parameters which will thus deviate from the “true” ones. Many different combinations of parameter estimates yield equally good fits of the observed dependence of  $S_0/S$  with  $[L_W]$ . The examples shown in Figure 9 correspond to the best fit of equation (25), allowing  $\beta_i$  to vary while fixing the number of binding sites at  $n = 3$  (plot A),  $n = 2$  (plot B), or  $n = 1$  (plot C). The best fit with  $n$  as an adjustable parameter and considering infinite cooperativity, a procedure commonly followed to describe Stern–Volmer plots with upward curvature [53,72], is also shown in plot C. The ratio between the values of the concentration of free ligand obtained from each best fit,  $[L_W]_{\text{Fit}}$ , and its “true” value, is shown in plots D to F. For this particular situation this ratio is almost always higher than 1, to compensate for the apparent weaker binding affinity that results from the inefficient fluorescence quenching. The values obtained for  $\beta_i$  vary widely, suggesting positive cooperativity in some cases and negative cooperativity in others, with the binding affinity ranging from  $\cong 10^6 \text{ M}^{-1}$  (close to the true one) to over  $10^8 \text{ M}^{-1}$ .

This shows very clearly that detailed information regarding the binding properties of proteins with multiple binding sites cannot be obtained from Stern–Volmer plots of protein fluorescence quenching by the ligand. At most, the order of magnitude of the overall binding affinity may be obtained. It is therefore preferable to consider the simplest models, with only one binding site per protein, or several equal and independent binding sites at most.

The simulations shown in this section considered proteins with several binding sites, and it was shown that an upward curvature in the Stern–Volmer plots is expected, Figures 8 and 9. The remainder of this section will be dedicated to the simpler case of proteins with a single binding site, and it will be seen that in this case downward curvature is the general behaviour. A downward curvature in Stern–Volmer plots is commonly observed in protein-binding experiments, and it is usually interpreted as indication that there are several fluorophores in the protein and not all are being quenched by the ligand. For the case of proteins containing several tryptophan residues this is a particularly convenient interpretation. However, the reason behind this behaviour may be simply that the presence of the ligand in the binding site does not lead to full quenching ( $s_i \neq 0$ ).

Figure 10 shows simulations for a protein with one binding site, that binds the ligand with low, moderate or high affinity, leading to full or only partial quenching of the protein fluorescence. The simulation results for  $S_0/S$  were analysed to obtain estimates of the binding affinity and fluorescence quenching efficiency. The equations used for the linear best fit at small  $S_0/S$  and for the best fit of the whole variation in  $S_0/S$  are easily obtained from Equations (29) and (30), respectively, and are shown below for completeness.

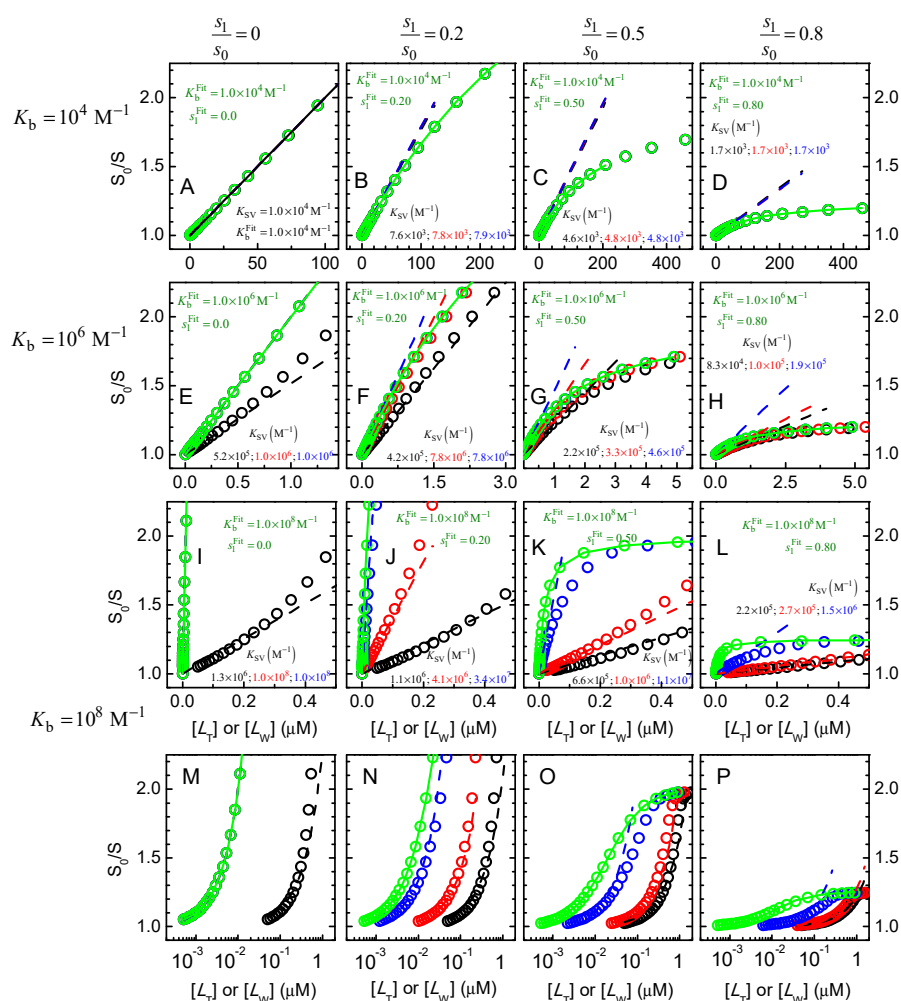
The slope at  $S_0/S$  close to unity is related to the binding affinity and quenching efficiency by:

$$K_{\text{SV}}(\text{small } [L_W]) = \left(1 - \frac{s_1}{s_0}\right) K_b \quad (31)$$

and the whole variation of  $S_0/S$  is related to the binding affinity and quenching efficiency by:

$$\frac{S_0}{S} = \frac{1 + K_b[L_W]}{1 + \frac{s_1}{s_0}K_b[L_W]} \quad (32)$$

Equation (32) is formally equivalent to the equation commonly used to account for downward curvature in Stern–Volmer plots, with  $s_1/s_0$  being interpreted as the fraction of fluorescent groups in the protein that are not accessible to the ligand ( $1 - f_a$ ). However, the interpretation of the results obtained from both approaches is fundamentally different. In Equation (32),  $s_1/s_0$  provides information regarding the proximity of the ligand to the protein fluorophores in a single binding site protein. In contrast, in the usual analysis,  $f_a$  is interpreted as the fraction of ligands in a given binding site, assuming *a priori* and without any support that quenching is 100% efficient and that the protein has multiple binding sites.



**Figure 10.** Stern–Volmer plots for the quenching of protein fluorescence (at  $[P_T] = 10^{-6}$  M) due to ligand binding. One binding site in the protein is considered, with low ( $K_b = 10^4$  M $^{-1}$ , plot (A–D)), intermediate ( $K_b = 10^6$  M $^{-1}$ , plot (E–H)) or high ( $K_b = 10^8$  M $^{-1}$ , plot (I–P), note the logarithmic scale in the y axis for the lower plots) affinity. Different efficiencies of quenching of protein fluorescence are considered ( $s_1/s_0 = 0$ , leftmost plots;  $s_1/s_0 = 0.2$ , 2nd from left;  $s_1/s_0 = 0.5$ , 3rd from left; and  $s_1/s_0 = 0.8$ , rightmost). The  $S_0/S$  ratio at low quenching was analysed with the simplified Stern–Volmer Equation (22) considering that  $[L_W] = [L_T]$  ( $\circ$ , and  $\ominus$ ), or calculating  $[L_W]$  from Equation (1) and  $x_-$  solution of (3) assuming that  $K_b = K_{SV}$  ( $\circ$ , and  $\ominus$ ), or using Equation (31) to calculate  $K_b$  from  $K_{SV}$  ( $\circ$ , and  $\ominus$ ). The full variation of  $S_0/S$  was also analysed with Equation (32), to obtain directly  $K_b$  and  $s_1/s_0$  ( $\circ$ , and  $\ominus$ ). The values of  $K_{SV}$  obtained from the best fit at low  $S_0/S$  are shown in the plots with the same colour as the respective best fit lines. The parameters obtained from the best fit of the whole variation of  $S_0/S$  are shown in green.

In Figure 10, the quenching efficiency was varied from 100% to 20% ( $s_1/s_0 = 0, 0.2, 0.5$ , and  $0.8$ , left to right plots), and the binding affinity was varied from low to high ( $K_b = 10^4, 10^6$ , and  $10^8$  M $^{-1}$ , top to bottom). The dashed lines in each plot correspond to the best fit of Equation (22) for  $S_0/S$  close to unity, with the relation of  $K_{SV}$  with  $K_b$  given by Equation (31), and the concentration of free ligand calculated from the  $x_-$  solution of Equation (3). The circles and dashed lines in red assume full quenching, while the value of  $s_1/s_0$  considered in the simulations is used for the circles and dashed lines in blue. To highlight the inadequacy of using  $[L_T]$  instead of  $[L_W]$ , this condition is also shown with black circles and dashed lines. Finally, the best fit of Equation (32) to the whole variation of  $S_0/S$  is shown as green circles and a continuous line. In this case, the quenching efficiency

was an adjustable parameter. The estimated parameter obtained from the distinct best fits is shown in the respective plots with the same colours as the corresponding curves.

When the data is analysed with the best fit of the linear region, only one parameter may be estimated,  $K_{SV}$ . If the quenching efficiency is known,  $K_b$  may be calculated accurately from  $K_{SV}$  using Equation (31), otherwise  $K_b$  will be underestimated. For low-affinity binding (plots A to D), the value obtained for  $K_{SV}$  is essentially independent on the fitting procedure because in this case binding of the ligand to the protein does not significantly deplete it in the aqueous phase. As the quenching efficiency decreases, the  $K_{SV}$  obtained decreases as well, underestimating  $K_b$  if  $s_1/s_0$  is assumed to be zero. For intermediate binding affinity (plots E to H), significant depletion of solute in the aqueous medium is observed and underestimates of  $K_b$  will be obtained if  $[L_T]$  is considered (black). For the other fitting procedures, the conclusions are the same as given above for low-affinity binding.

The case of high binding affinity is represented in plots I to L (and M to P with a logarithmic scale for the  $x$  axis). Neglect of ligand depletion leads to underestimation of  $K_b$  by several orders of magnitude (black). Considering  $[L_W]$  obtained from the best fit and the incorrect assumption that  $s_1/s_0 = 0$  (red) will also lead to underestimation of  $K_b$  (for  $s_1/s_0 = 0.8$ , the value obtained for  $K_{SV}$  is  $2.7 \times 10^5 \text{ M}^{-1}$ , leading to the same value for  $K_b$  if full quenching is assumed). If both  $[L_W]$  from the fit and the correct  $s_1/s_0$  are considered (blue),  $K_{SV}$  is  $1.5 \times 10^6 \text{ M}^{-1}$  leading to  $K_b = 7.5 \times 10^6 \text{ M}^{-1}$ , which although closer still underestimates the true value.

The correct parameters ( $K_b$  and  $s_1/s_0$ ) are obtained when the full variation of  $S_0/S$  is analysed with Equation (32) (green) for all binding affinities and quenching efficiencies. However, some care must be taken regarding the sensitivity of the best fit to the estimated parameters. This is illustrated in Figure S6 for low, intermediate and high affinity. At the concentration of protein considered in those simulations ( $1 \mu\text{M}$ ), the sensitivity is high for low binding affinity, with the best fit, considering ratios of 1.5 or 0.67 between the estimated and the true  $K_b$ , leading to clear misfits (dark blue and orange). A good sensitivity is still obtained for the case of intermediate affinity, although when the  $x$  axis is represented in the logarithmic scale it is seen that a clear misfit is only obtained for ratios of 2 and 0.5. This is due to the significant ligand depletion, which is dependent on the estimated  $K_b$  and shifts the curve of  $S_0/S$  in the  $x$  axis. This effect becomes more evident for high-affinity binding (right plots), where essentially the same quality is obtained for the best fit with the estimated parameter varying from 0.5 to 2 times the true  $K_b$ . A clear misfit of the experimental  $S_0/S$  is only observed for an underestimation of  $K_b$  by at least one order of magnitude, and essentially no sensitivity is observed for overestimation of  $K_b$ . A better sensitivity in the estimation of  $K_b$  is obtained if the total concentration of protein is decreased, with  $[P_T] = 10^{-8} \text{ M}$  leading to a good parameter estimation when  $K_b = 10^8 \text{ M}^{-1}$ .

In conclusion, quenching of protein fluorescence due to ligand binding is an appropriate methodology to quantitatively characterize the binding to single binding site proteins, providing both the binding affinity and the quenching efficiency. However, the range of ligand concentrations must be large enough to lead to non-linear Stern–Volmer plots, and the whole variation of  $S_0/S$  should be analysed with Equation (32). Additionally, to ensure a good sensitivity regarding the estimated binding affinity, the total concentration of protein should be lower than the estimated value for the dissociation constant. If the solubility of the ligand in the aqueous medium does not allow for reaching high enough concentrations to generate non-linear Stern–Volmer plots, the variation of  $S_0/S$  must be analysed with Equation (22), and  $K_b$  assumed equal to  $K_{SV}$ . This procedure will underestimate  $K_b$  if the quenching efficiency is not 100%. If  $s_1/s_0$  may be obtained independently, an accurate estimation of  $K_b$  may be obtained from  $K_{SV}$  using Equation (31).

When the proteins have several binding sites, the quenching of protein fluorescence may very easily lead to incorrect estimations of the binding parameters. To increase the accuracy, a wide range of ligand concentrations should be used to lead to non-linear Stern–Volmer plots (which in this case are expected to show upward curvature) and  $S_0/S$

should be analysed with models considering several equal and independent binding sites, Equations (21) and (30). This procedure is expected to yield good estimates of the total binding affinity. If moderate-to-high binding affinity is obtained, the experiments should be repeated at a lower total concentration of protein to increase the confidence in the parameter estimated for high binding affinity.

In any case, the use of the simple Stern–Volmer equation with  $[L_W]$  replaced by  $[L_T]$  is an incorrect procedure that may underestimate the binding affinity by several orders of magnitude. The concentration of free ligand is dependent on the binding affinity and must be calculated for each set of parameters. The assumption that all binding sites in the protein are equal and independent leads to a simple procedure for the calculation of  $[L_W]$ , being a good compromise between accuracy and convenience.

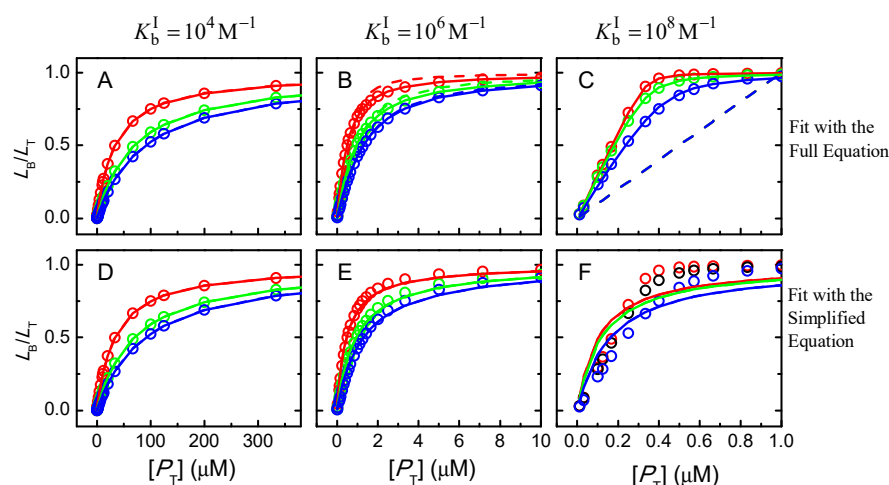
#### 2.4. Characterization of the Binding Parameters from Ligand Saturation Curves

It is sometimes easier or more convenient to follow the interaction of the ligand with the protein through a variation in the properties of the former. This is a convenient situation in the case of fluorescent ligands that change their fluorescence properties upon binding. It may also be a necessity if the ligand solubility does not allow attainment of protein saturation with ligand, limiting the maximum concentration achieved and/or leading to the formation of aggregates.

If the protein has only one binding site and the concentration of protein is known, the formalism is simple, and following protein saturation or ligand saturation leads to equivalent and accurate results. Equation (3) or Equation (7) can be used for this purpose, the latter equation only if the protein is in large excess. However, accuracy and equivalence are unwarranted when the protein has several binding sites, and/or when its concentration is not well-characterized. Some specific examples will be analysed below.

##### 2.4.1. Ligand Binding to Proteins with Multiple Binding Sites

The ligand saturation curves obtained for the case of proteins with three independent binding sites are shown in Figure 11 for different binding affinities and distinct relations between the affinities for the distinct binding sites in the protein.



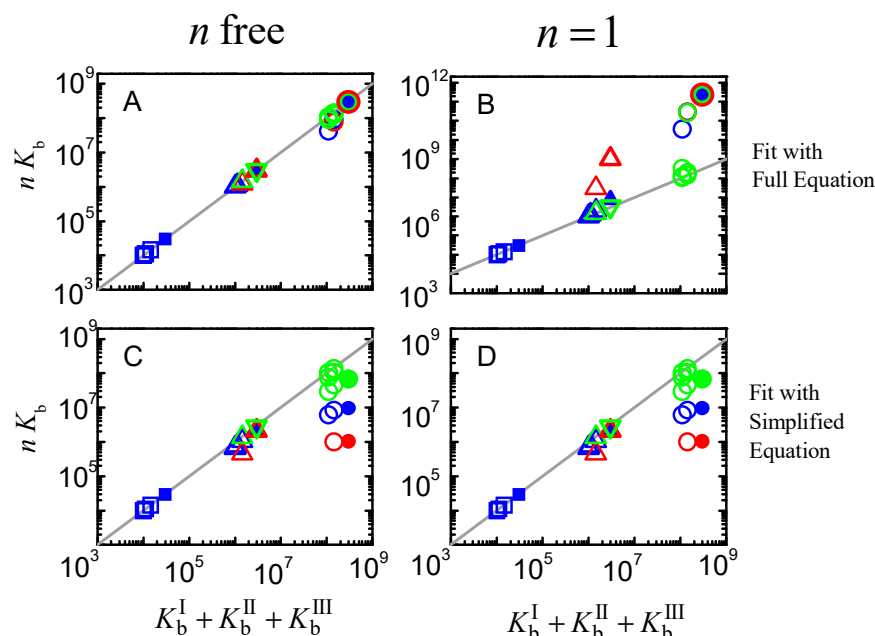
**Figure 11.** Fraction of ligand bound to a protein with 3 independent binding sites of equal affinity ( $\circ$ ), with  $K_b^I/K_b^{II} = K_b^{II}/K_b^{III} = 3$  ( $\circ$ ) or with  $K_b^I/K_b^{II} = K_b^{II}/K_b^{III} = 10$  ( $\circ$ ) for binding with low affinity ( $K_b^I = 10^4 \text{ M}^{-1}$ , plots (A,D)), intermediate affinity ( $K_b^I = 10^6 \text{ M}^{-1}$ , plots (B,E)), and high affinity ( $K_b^I = 10^8 \text{ M}^{-1}$ , plots (C,F)). In the upper plots the lines are the best fit of Equation (3) with an adjustable number of binding sites  $n$  with equal affinity  $K_b$  (continuous lines), or for a fixed number of binding sites  $n = 1$  (dashed lines). The equivalent is shown in the lower plots but showing the best fit of the simplified Equation (7), considering that the protein is in large excess. The total concentration of ligand is  $1 \mu\text{M}$  in all plots.



The first observation is that a good fit is always obtained when the fraction of bound ligand is described by a model that considers an adjustable number of equal and independent binding sites, as in Equation (3). This clearly shows that the shape of the ligand saturation curve does not allow to distinguish between equal or different binding sites in the protein. For intermediate and high-affinity binding, it is observed that the best fit with an adjustable number of binding sites leads to a better description of the ligand saturation curve than when  $n = 1$  is imposed (continuous vs. dashed lines, respectively). The misfit for  $n = 1$  is clear in the case of high-affinity binding, but for low or intermediate affinity it would not be possible to distinguish between the two situations due to the noise in the real data.

The bottom plots in Figure 11 show the best fit of the simplified Equation (7), which considers that the protein is in large excess. As expected, a good fit is obtained for the case of low affinity (binding occurs at protein concentrations of several hundred  $\mu\text{M}$  and the ligand concentration is  $1 \mu\text{M}$ ). A clear misfit is obtained for high-affinity binding, and a reasonable fit for intermediate affinity.

The next question is whether the good fits lead to accurate estimations of the ligand overall binding affinity of the protein. Because different binding sites cannot be distinguished, the relevant properties are the overall binding affinity ( $nK_b$  vs.  $\sum K_b^i$ ), the highest binding affinity ( $K_b$  vs.  $K_b^I$ ), and the average binding affinity ( $K_b$  vs.  $\bar{K}_b^i = \sqrt[n]{\left(\prod_{i=1}^n K_b^i\right)}$ ). Because the sites are all assumed equal in the models considered to describe the ligand saturation profiles, the estimates obtained for the highest affinity and for the average affinity deviate significantly from the true values when the binding sites are different. However, the overall binding affinity may be estimated with good accuracy for most conditions. The estimates obtained from the best fit of ligand saturation curves for a large set of conditions are presented in Figure 12.



**Figure 12.** Relation between the overall binding affinity obtained from the best fit considering equal and independent binding sites and the real overall binding affinity ( $K_b^I + K_b^II + K_b^III$ ). Plots (A,B) show the results from the best fitting of Equation (3), while plots (C,D) consider the simplified Equation (5). In the left plots the number of binding sites was an adjustable parameter, while it was fixed at  $n = 1$  in the right plots. The total concentration of ligand was  $1 \mu\text{M}$  in the blue symbols, lower in the green symbols (down to  $1 \text{ nM}$ ), and  $10 \mu\text{M}$  in the red symbols. The value of  $K_b^I$  was  $10^4 \text{ M}^{-1}$  (squares),  $10^6 \text{ M}^{-1}$  (triangles), or  $10^8 \text{ M}^{-1}$  (circles), with lower affinities for the other two binding sites (hollow symbols) or equal affinity for all binding sites (filled symbols).

The fitting approach that yields the best estimate of the overall protein binding affinity is the use of the full equation considering an adjustable number of equal binding sites (plot A). This procedure leads to an excellent prediction of this property for low and intermediate affinities (less than 15% deviation) and reasonable predictions for high-affinity binding. In this case, the quality of the estimate gets poorer for higher ligand concentrations and for binding sites with similar (but not equal) affinities. This is due to a significant population of binding sites with intermediate affinity, which is not captured by the model that considers only one type of binding sites. When the number of binding sites is fixed at  $n = 1$  (plot B) the predictions are usually good for low-, intermediate- and high-affinity binding at very low ligand concentrations (symbols in green), but overestimates the binding affinity by several orders of magnitude for high ligand concentrations (symbols in red). It should be noted that at high ligand concentrations the quality of the fit was very poor (Figure 11 plot C, dashed lines), which is related to the inadequate calculation of the free ligand concentration, as discussed when interpreting Figure 7.

The bottom plots in Figure 12 correspond to the estimates when the best fit of the fraction of ligand bound is obtained with the simplified equations that assume excess protein, Equation (5). As expected, the quality of the estimates is not dependent on the number of binding sites, as the results in plot C are equal to those in plot D, although the best fit was obtained for  $n \neq 1$  in plot C. This is because in the used equation  $n$  is always associated with  $K_b$ . The parameters obtained are only good in the case of low affinity binding, and lead to an underestimation of the overall binding affinity for intermediate and high-affinity binding. Surprisingly at first, the quality of the estimates improves as the concentration of ligand decreases, changing from  $nK_b / \sum K_b^i$  equal to 0.03 at  $[L_T] = 1 \mu\text{M}$  to 0.23 at  $[L_T] = 0.1 \mu\text{M}$  in the case of  $K_b^I = K_b^{II} = K_b^{III} = 10^8 \text{ M}^{-1}$ . This is because, at very low ligand concentrations, significant ligand binding occurs only at total protein concentrations that are in fact much higher than the ligand concentration, as assumed in the derivation of the simplified solution.

From this detailed analysis it may be concluded that, when analysing ligand saturation without knowledge of the detailed binding details (number of binding sites and whether they are equal or different), it is a reasonable approach to consider that all binding sites are equal, and perform the analysis in terms of binding sites using the full equation. If a good fit is obtained, this procedure is expected to yield an accurate prediction of the overall protein-binding affinity. The correct number of binding sites may also be obtained if they are all equal and if binding leads to a significant depletion of the total protein. However, if the binding sites are not equal, the value of  $n$  estimated from the best fit is an underestimation. The deviation between the estimated and the real number of binding sites in the protein depends on the values of the distinct binding affinities, the relation being  $nK_b = \sum K_b^i$  when a good estimate of the overall protein-binding affinity is obtained. Concomitantly, following the ligand saturation does not lead to good estimates of the highest binding affinity. Exceptions occur when the affinity is the same for all binding sites, or when the difference between the distinct binding sites is very large (more than one order of magnitude), in this case leading only to the characterization of the highest affinity binding site.

There are several examples in the literature where this procedure was followed to characterize the association of fluorescent amphiphiles with proteins [29,52,57–60,73–77]. In most cases the objective was only to evaluate the overall binding affinity. As explained above, this parameter is obtained accurately, even for the case of proteins with multiple binding sites when very low concentrations of ligands are used [29,58–60,73,74].

#### 2.4.2. Ligand Binding as a Partition Coefficient

In some situations, the protein under study is not well-characterized (e.g., in terms of protein molecular weight, and number of binding sites), and only the total mass of protein in solution is known. In these cases, it is useful to describe binding as a partition coefficient between the aqueous medium and the protein. It may also be convenient to

describe binding in terms of partition for the case of proteins with several binding sites such as serum albumin [72,78,79]. If the concentration of the ligand is much lower than that of the protein, the volume of protein available for binding ( $V_P$ ) is maintained throughout the titration, and the equilibrium distribution of ligand is given by Equation (33).

$$L_W \xrightleftharpoons{K_P^{W \rightarrow P} V_P} L_P; K_P^{W \rightarrow P} = \frac{nL_P/V_P}{nL_W/V_W} \quad (33)$$

The partition coefficient obtained is related with the overall equilibrium association by:

$$\begin{aligned} K_P^{W \rightarrow P} &= \frac{nL_P}{nL_W} \frac{V_W}{n_P \bar{V}_P} \cong \frac{nL_P}{nL_W} \frac{V_T}{n_P \bar{V}_P} = \frac{nL_P}{nL_W} \frac{1}{[P]} \frac{1}{\bar{V}_P} = K_b^{\text{obs}} \frac{1}{\bar{V}_P} \\ &\Leftrightarrow K_b^{\text{obs}} = K_P^{W \rightarrow P} \bar{V}_P \end{aligned} \quad (34)$$

where  $K_b^{\text{obs}}$  is the observed association constant when considering the whole protein, and  $\bar{V}_P$  is the molar volume of the whole protein. The assumption made ( $V_w \cong V_T$ ) is valid for dilute solutions, which is the case for studies performed using model systems. Its validity must be verified if this formalism is applied to biological systems or when very high concentrations of binding agents are used.

The concentrations of ligand associated with the protein and in the aqueous phase are given by:

$$\begin{aligned} [L_P] &= [L_T] \frac{K_P^{W \rightarrow P} (V_P/V_T)}{V_W/V_T + K_P^{W \rightarrow P} (V_P/V_T)} \\ [L_W] &= [L_T] \frac{V_W/V_T}{V_W/V_T + K_P^{W \rightarrow P} (V_P/V_T)} \end{aligned} \quad (35)$$

If the concentration of ligand is comparable to that of the protein, it is not correct to assume that the volume of the protein available for binding remains constant throughout the titration. In this case, it is necessary to consider the available volume of protein,  $V_P^*$ , which is given by:

$$V_P^* = V_P - nL_P \frac{\bar{V}_P}{\#} \quad (36)$$

where # is the number of binding sites per molar volume. The scheme describing the partition of the ligand between the aqueous medium and the protein is obtained replacing  $V_P$  by  $V_P^*$  in Equation (33).

The volume of protein available to interact with the ligand is calculated from the quadratic equation:

$$\left(\frac{V_P^*}{V_T}\right)^2 K_P^{W \rightarrow P} + \frac{V_P^*}{V_T} \left\{ \frac{V_W}{V_T} + K_P^{W \rightarrow P} \left( \frac{nS_T \bar{V}_P}{V_T \#} - \frac{V_P}{V_T} \right) \right\} - \frac{V_W}{V_T} \frac{V_P}{V_T} = 0 \quad (37)$$

where the physically meaningful solution is always  $(V_P^*)_+$ .

The concentration of ligand associated with the protein or in the aqueous phase may be calculated from the equivalent to Equation (35) but with the volume of the protein being replaced by  $V_P^*$ , which is given by the solution of Equation (37).

The application of this formalism to the analysis of ligand binding to proteins with multiple binding sites leads to results identical to those shown in Figures 11 and 12. A good fit is always obtained when considering a variable number of binding sites, and the volume of free protein from Equation (37). The partition coefficient is a good estimate for the overall binding affinity, with  $K_b^{\text{obs}}$  being close to  $\sum K_b^i$ . The simplified Equation (35) is only valid for very low ligand concentrations or very small  $K_P^{W \rightarrow P}$ , which in both cases leads to a low saturation of the protein.

The advantage of this formalism is that it allows to recover the overall binding affinity even when the concentration of protein is not known, only its total mass or volume. If the ligand concentration is high enough to saturate the protein, the best fit is dependent on  $n$ , and an estimate is obtained for the number of binding sites in the protein (considered

equal and independent). From this it is possible to estimate the volume per binding site. However, if the best fit is not dependent on  $n$ , only the partition coefficient is obtained, and it is not possible to calculate the observed binding constant. Nevertheless, from the partition coefficient one may directly calculate the fraction of ligand associated with the protein for a given amount of protein, which is usually the relevant property.

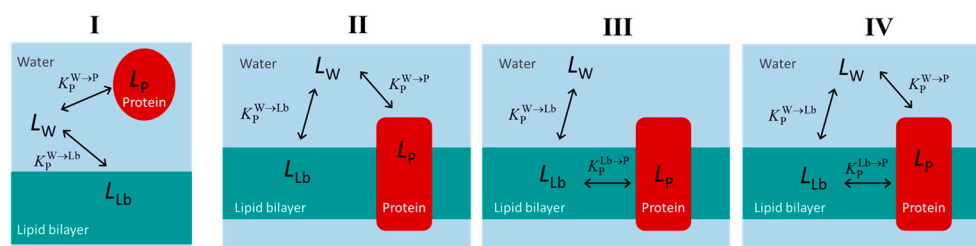
Another advantage of this formalism is the ease of its application to the description of ligand distribution in heterogeneous media with increasing complexity. An example is when membranes and proteins are present in the same aqueous solution, as observed in all biological systems. This situation will be considered in the next section.

### 3. Ligand Distribution in the Presence of Proteins and Membranes

Ligand association with membranes is usually treated as a partition because the membranes do not have well-defined binding sites and because they can accumulate relatively high amounts of ligand without changing their binding properties. For small uncharged or singly charged ligands, the membrane may accommodate at least up to 5 mol% of the former without changing their binding properties [33,34]; depending on the ligand and membrane composition, local ligand concentrations as high as 20 mol% may be achieved without significant changes in the membrane binding affinity [80–82]. Considering the typical molar volume of a membrane lipid (0.8 dm<sup>3</sup>/mol), the local ligand concentration in the membrane may thus exceed 0.1 M.

The use of the partition formalism allows a direct comparison of the binding affinity of lipid membranes and proteins. For a protein with  $\bar{V}_P = 50$  dm<sup>3</sup>/mol (corresponding to a molecular weight of ~70 kg/mol), a moderate binding affinity  $K_b^{\text{obs}} = 10^6$  M<sup>-1</sup> corresponds to  $K_P^{W \rightarrow P}$  equal to  $2 \times 10^4$ . At low ligand concentrations, for an equivalent volume of protein and lipid bilayer the ligand will be equally distributed for  $K_P^{W \rightarrow Lb}$  equal to  $2 \times 10^4$ . However, at high ligand concentrations the protein becomes saturated while the membrane is still able to accommodate the ligand. This will shift the distribution of ligand from the protein to the membrane. The maximum local concentration of ligand in the protein is equal to  $1/(\bar{V}_P/\#)$ . For a protein with  $\bar{V}_P = 50$  dm<sup>3</sup>/mol and a single binding site this concentration is 0.02 M, and is thus lower than the maximum ligand concentration achievable in a membrane without changing its properties as a binding agent.

A quantitative analysis of the ligand distribution between the aqueous medium, the protein and the lipid membrane will be presented in the following sections. Focus will be given to the effect of ligand concentrations regarding the fraction of ligand bound to the protein and membrane. Several distinct conditions will be considered: proteins soluble in the aqueous phase (case I), and membrane proteins (case II to IV), as shown in Figure 13.



**Figure 13.** Distribution of ligands in media containing proteins and lipid membranes. Case (I)—proteins soluble in the aqueous medium; Cases (II–IV)—membrane proteins, in the case of ligand equilibration with both binding systems from the aqueous medium (II), interaction with the protein from the membrane only (III) or from both the aqueous medium and the membrane (IV).

#### 3.1. Lipid Membranes and Proteins in Solution

When using the partition formalism, the equations that describe the distribution of ligand between the aqueous phase and the different binding agents are easily obtained.

This distribution is given by Equation (38) for the case of lipid membranes and aqueous soluble proteins (case I).

$$\begin{aligned} [L_W] &= [L_T] \frac{V_W/V_T}{V_W/V_T + K_P^{W \rightarrow P} (V_P^*/V_T) + K_P^{W \rightarrow Lb} (V_{Lb}/V_T)} \\ [L_P] &= [L_T] \frac{K_P^{W \rightarrow P} (V_P^*/V_T)}{V_W/V_T + K_P^{W \rightarrow P} (V_P^*/V_T) + K_P^{W \rightarrow Lb} (V_{Lb}/V_T)} \\ [L_{Lb}] &= [L_T] \frac{K_P^{W \rightarrow Lb} (V_{Lb}/V_T)}{V_W/V_T + K_P^{W \rightarrow P} (V_P^*/V_T) + K_P^{W \rightarrow Lb} (V_{Lb}/V_T)} \end{aligned} \quad (38)$$

The volume of lipid bilayer available to receive ligand is considered constant (non-saturable binding agent), while that of the protein depends on the amount of ligand bound. The volume of protein available to the ligand is obtained from the solution of the quadratic equation given in Equation (39), the physically meaningful solution being always  $x_+$ :

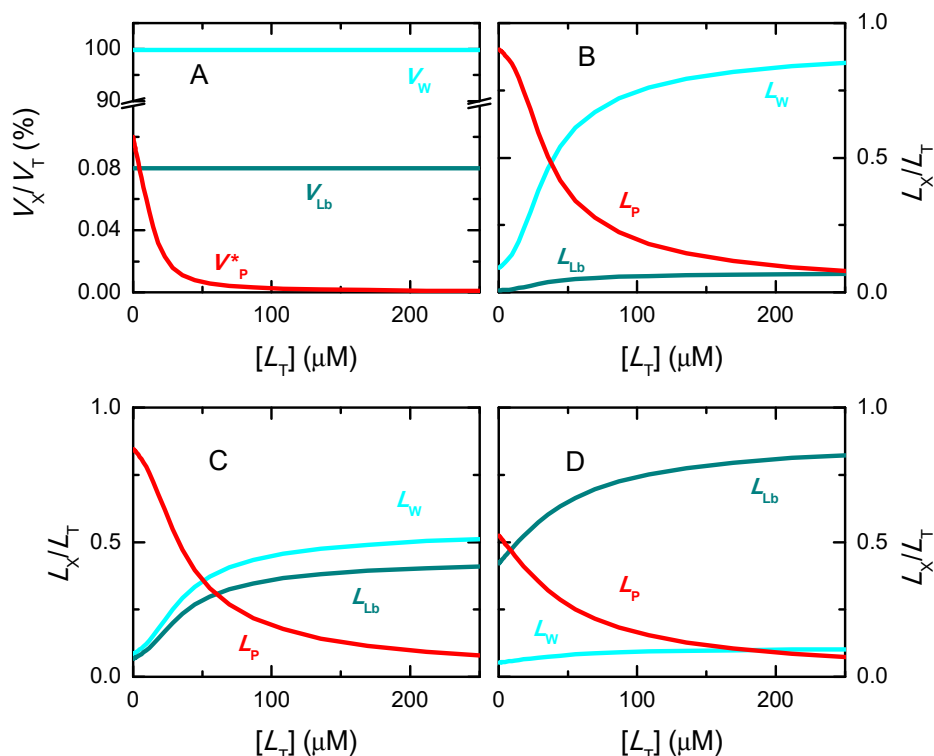
$$\left(\frac{V_P^*}{V_T}\right)^2 K_P^{W \rightarrow P} + \frac{V_P^*}{V_T} \left\{ \frac{V_W}{V_T} + K_P^{W \rightarrow Lb} \frac{V_{Lb}}{V_T} + K_P^{W \rightarrow P} \left( \frac{nS_T \bar{V}_P}{V_T} - \frac{V_P}{V_T} \right) \right\} - \frac{V_P}{V_T} \left( \frac{V_W}{V_T} + K_P^{W \rightarrow Lb} \frac{V_{Lb}}{V_T} \right) = 0 \quad (39)$$

The effect of ligand concentration on the membrane volume and that of protein available for ligand binding is represented in Figure 14A for the case of moderate affinity of the ligand for the protein ( $K_b^{obs} = 5 \times 10^5 \text{ M}^{-1}$ ,  $K_P^{W \rightarrow P} = 10^4$ ), and low affinity for the lipid bilayer ( $K_P^{W \rightarrow Lb} = 10^2$ ). In the absence of ligand, at the protein and lipid concentrations considered, the volume of protein available for ligand binding is slightly higher than that of the lipid membrane. However, as the concentration of ligand is increased, the protein binding sites are occupied and the protein volume available for binding decreases. In contrast, the membrane volume available for binding is maintained. The fraction of ligand in the distinct media is presented in plot B. At very low ligand concentrations, most ligand is associated with the protein. However, as the protein becomes saturated with the ligand, it accumulates in the aqueous medium and in the membrane. For the low membrane affinity considered in this simulation, only about 10% of the ligand is associated with the membrane. A higher membrane affinity is considered in plots C ( $K_P^{W \rightarrow Lb} = 10^3$ ) and D ( $K_P^{W \rightarrow Lb} = 10^4$ ). In these conditions the fraction of ligand associated with the membrane becomes significant and may represent up to 90%. It should be noted that in plot D, the overall affinities of the ligand for the protein and membrane are equal ( $K_P^{W \rightarrow P} = K_P^{W \rightarrow Lb}$ ), and that in the absence of ligand the volume of protein available is slightly higher than that of the membrane. However, the membrane becomes the dominant binding agent for ligand concentrations higher than 9  $\mu\text{M}$ .

### 3.2. Membranes Containing Lipids and Membrane Proteins

In this case one has to consider three possibilities for the equilibration of ligand with the membrane protein: from the aqueous phase only (Case II), from the membrane only (Case III), or from both (Case IV).

The equations that describe the distribution of ligand at equilibrium for case II are the same as those for soluble proteins, Equations (38) and (39). All the observations from the previous section are thus valid in this case, with the presence of the membrane decreasing the apparent affinity of the ligand for the protein.



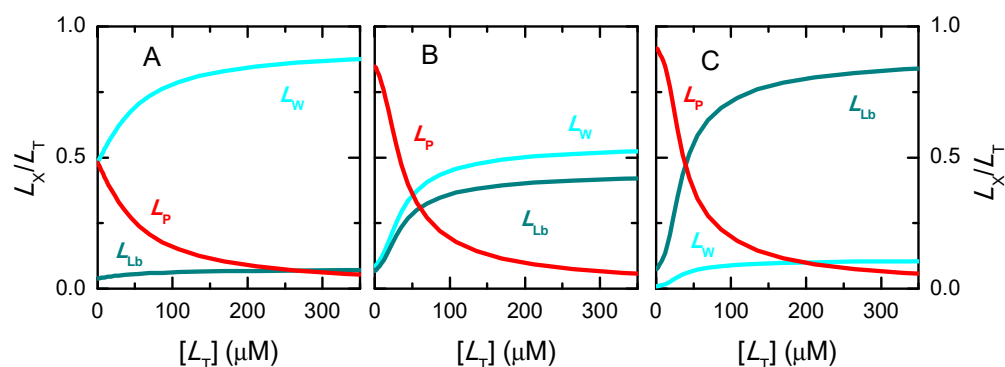
**Figure 14.** Distribution of a ligand between the aqueous medium, a lipid membrane and a protein soluble in the aqueous medium. The protein is at a concentration of 10  $\mu\text{M}$  and has a molar volume of 50  $\text{dm}^3/\text{mol}$  with a single binding site, while a lipid concentration of 1  $\text{mM}$  is considered with a lipid molar volume of 0.8  $\text{dm}^3/\text{mol}$ . The partition coefficient between the aqueous medium and the protein is the same in all plots ( $K_P^{W \rightarrow P} = 10^4$ ) while the affinity for the membrane is low in plots (A) and (B) ( $K_P^{W \rightarrow Lb} = 10^2$ ), and increases for plots (C) ( $K_P^{W \rightarrow Lb} = 10^3$ ) and (D) ( $K_P^{W \rightarrow Lb} = 10^4$ ).

If ligand binding to the protein occurs only after ligand association with the membrane (Case III), the equilibrium distribution of the ligand may be calculated from Equations (40) and (41).

$$\begin{aligned}
 [L_W] &= [L_T] \frac{V_W/V_T}{V_W/V_T + K_P^{W \rightarrow Lb} \{ (V_{Lb}/V_T) + (V_P^*/V_T) K_P^{Lb \rightarrow P} \}} \\
 [L_P] &= [L_T] \frac{K_P^{W \rightarrow Lb} K_P^{Lb \rightarrow P} (V_P^*/V_T)}{V_W/V_T + K_P^{W \rightarrow Lb} \{ (V_{Lb}/V_T) + (V_P^*/V_T) K_P^{Lb \rightarrow P} \}} \\
 [L_{Lb}] &= [L_T] \frac{K_P^{W \rightarrow Lb} (V_{Lb}/V_T)}{V_W/V_T + K_P^{W \rightarrow Lb} \{ (V_{Lb}/V_T) + (V_P^*/V_T) K_P^{Lb \rightarrow P} \}}
 \end{aligned} \tag{40}$$

$$\left( \frac{V_P^*}{V_T} \right)^2 K_P^{W \rightarrow Lb} K_P^{Lb \rightarrow P} + \frac{V_P^*}{V_T} \left\{ \frac{V_W}{V_T} + K_P^{W \rightarrow Lb} \frac{V_{Lb}}{V_T} + K_P^{W \rightarrow Lb} K_P^{Lb \rightarrow P} \left( \frac{nS_T \bar{V}_P}{V_T} - \frac{V_P}{V_T} \right) \right\} - \frac{V_P}{V_T} \left( \frac{V_W}{V_T} + K_P^{W \rightarrow Lb} \frac{V_{Lb}}{V_T} \right) = 0 \tag{41}$$

The results obtained for a moderate relative affinity between the membrane and the protein ( $K_P^{Lb \rightarrow P} = 10$ ) are shown in Figure 15, at affinities between the aqueous medium and the membrane,  $K_P^{W \rightarrow Lb}$ , increasing from  $10^2$  (plot A) to  $10^4$  (plot C). In contrast to what is observed for proteins in the aqueous medium (Figure 14), increasing the affinity of the ligand for the lipid bilayer leads to an apparent increase in the ligand affinity for the protein. This result was expected because it is the ligand in the membrane that is able to interact with the protein. However, a closer analysis of Figure 15 shows that this increased affinity for the protein is only significant for low ligand concentrations. As the protein-binding sites become saturated, most ligand accumulates in the membrane despite the lower intrinsic affinity.



**Figure 15.** Distribution of a ligand between the aqueous media, the lipid bilayer and the protein in a membrane composed of a lipid bilayer with embedded membrane proteins. The protein and lipid concentrations are the same as in Figure 14, leading to 56% of the membrane being occupied by protein and 44% by the lipid bilayer. The partition coefficient between the lipid bilayer and the protein is the same in all plots ( $K_P^{Lb \rightarrow P} = 10^2$ ) while the affinity for the membrane is low in plot (A) ( $K_P^{W \rightarrow Lb} = 10^2$ ), and increases for plots (B) ( $K_P^{W \rightarrow Lb} = 10^3$ ) and (C) ( $K_P^{W \rightarrow Lb} = 10^4$ ).

Even though this model assumes that the ligand cannot associate with the membrane protein directly from the aqueous medium, this step closes a thermodynamic circle and thus the apparent partition coefficient may be calculated from the other two equilibria, Equation (42).

$$K_P^{W \rightarrow P} = K_P^{W \rightarrow Lb} K_P^{Lb \rightarrow P} \quad (42)$$

Due to the thermodynamic cycle involving the ligand in the aqueous medium and associated with the lipid bilayer or with the protein, Case IV in Figure 13 cannot be formally distinguished from Case III in what concerns the equilibrium distribution of the ligand. In fact, Case II is also formally indistinguishable from Case III and IV, Equations (40) and (41) being obtained if  $K_P^{W \rightarrow P}$  in Equations (38) and (39) is replaced by its relation with  $K_P^{W \rightarrow Lb}$  and  $K_P^{Lb \rightarrow P}$  given by Equation (42). Another mechanism could additionally be considered, with the ligand interaction with the lipid bilayer being mediated by interaction with the membrane protein, which is also described by the same formalism. This situation was not explicitly considered because interaction with the lipid membrane is usually non-specific, being observed for the vast majority of ligands with moderate and high hydrophobicity.

The formalism developed in this section may be of particular relevance for the characterization of ligand association with membrane proteins. It provides a simple method to quantitatively compare the ligand affinity for the lipid bilayer and for the proteins, and to evaluate the effect of ligand concentration and relative concentration of the membrane proteins. Those aspects will be further explored in a dedicated manuscript. For now, we will just define another important experimental observable, the partition coefficient between the aqueous medium and the membrane as a whole, as shown in Equation (43).

$$K_P^{W \rightarrow M} = \frac{(nL_{Lb} + nL_P)/(V_{Lb} + V_P)}{nL_W/V_W} = \frac{K_P^{W \rightarrow Lb}V_{Lb} + K_P^{W \rightarrow P}V_P}{V_{Lb} + V_P} = \frac{K_P^{W \rightarrow Lb}(V_{Lb} + K_P^{Lb \rightarrow P}V_P)}{V_{Lb} + V_P} \quad (43)$$

Thus, the characterization of the partition coefficient at low ligand concentrations in the presence and absence of the membrane protein, allows obtaining the water to protein and the lipid bilayer to protein partition coefficients. This may be of high relevance when studying interactions with biological membranes.

The situations discussed above correspond to different mechanisms of signal transduction and ligand transport through biological membranes with impact in the areas of biochemistry, pharmacology, neurochemistry and endocrinology. Distinction between the mechanisms of interaction with the membrane protein has been attempted, namely on the basis of the dependence on ligand hydrophobicity [83,84]. However, an increase in ligand lipophilicity would lead to an increased apparent affinity for the membrane protein

in all cases, provided that the relative affinity for the lipid bilayer and the protein is maintained. The mechanisms may eventually be distinguished if the kinetics of the interaction is accessed. If the ligand in the aqueous medium can only interact with the membrane, a relative accumulation of ligand in the membrane is expected at short times followed by its redistribution to the membrane protein. The reverse is expected if the ligand in the aqueous medium can only interact with the membrane protein. If direct interaction is established with both binding agents, the time evolution in each medium would depend on the relative rate and equilibrium constants.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23179757/s1>.

**Author Contributions:** Conceptualization, M.J.M.; methodology, M.J.M.; formal analysis, M.J.M. and A.S.; validation, J.M., L.M.S.L. and A.V.-C.; writing—original draft preparation, M.J.M.; writing—review and editing, M.J.M., L.M.S.L., J.M., A.S. and A.V.-C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the European Regional Development Fund, through COMPE TE2020—Operational Program for Competitiveness and Internationalization, and Portuguese funds via FCT-Fundação para a Ciência e a Tecnologia, under projects UIDB/04539/2020, UIDP/04539/2020, UIDB/00313/2020 and UIDP/00313/2020 (Portugal).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Tanford, C. Hydrophobic Free-Energy, Micelle Formation and Association of Proteins with Amphiphiles. *J. Mol. Biol.* **1972**, *67*, 59. [[CrossRef](#)]
2. Southall, N.T.; Dill, K.A.; Haymet, A.D.J. A view of the hydrophobic effect. *J. Phys. Chem. B* **2002**, *106*, 521–533. [[CrossRef](#)]
3. Creighton, T.E. *Proteins. Structures and Molecular Properties*; W. H. Freeman and Company: New York, NY, USA, 1993.
4. Klotz, I.M.; Hunston, D.L. Protein Interactions with Small Molecules—Relationships Between Stoichiometric Binding Constants, Site Binding Constants, and Empirical Binding Parameters. *J. Biol. Chem.* **1975**, *250*, 3001–3009. [[CrossRef](#)]
5. Cera, E.D. Mechanisms of ligand binding. *Biophys. Rev.* **2020**, *1*, 011303. [[CrossRef](#)] [[PubMed](#)]
6. Gennis, R.B. *Biomembranes. Molecular Structure and Function*; Springer: New York, NY, USA, 1989.
7. Singer, S.J.; Nicolson, G.L. Fluid Mosaic Model of Structure of Cell-Membranes. *Science* **1972**, *175*, 720. [[CrossRef](#)]
8. Deyoung, L.R.; Dill, K.A. Solute Partitioning Into Lipid Bilayer-Membranes. *Biochemistry* **1988**, *27*, 5281–5289. [[CrossRef](#)]
9. Seelig, A.; Allegrini, P.R.; Seelig, J. Partitioning of Local-Anesthetics Into Membranes—Surface-Charge Effects Monitored by the Phospholipid Headgroup. *Biochim. Biophys. Acta* **1988**, *939*, 267–276. [[CrossRef](#)]
10. Santos, N.C.; Prieto, M.; Castanho, M.A.R.B. Quantifying molecular partition into model systems of biomembranes: An emphasis on optical spectroscopic methods. *Biochim. Biophys. Acta-Biomembr.* **2003**, *1612*, 123–135. [[CrossRef](#)]
11. Melo, E.; Freitas, A.A.; Chang, Y.W.; Quina, F.H. On the significance of the solubilization power of detergents. *Langmuir* **2001**, *17*, 7980–7981. [[CrossRef](#)]
12. Melo, E.C.C.; Costa, S.M.B. A Critical Evaluation of Partition Constants in Nonionic Micelles. *J. Phys. Chem.* **1987**, *91*, 5635–5640. [[CrossRef](#)]
13. Lopes, A.; Demelo, J.S.; Martins, A.J.; Macanita, A.L.; Pina, F.S.; Wamhoff, H.; Melo, E. Partition of Pesticides of the Coumarin Family between Water and Amphiphilic Aggregates. *Environ. Sci. Technol.* **1995**, *29*, 562–570. [[CrossRef](#)] [[PubMed](#)]
14. Moreno, M.J. *Kinetics of Bimolecular Reactions with Stochastic Effects and Dimensional Constraints*; Universidade Nova de Lisboa: Oeiras, Portugal, 2000.
15. Moreno, M.J.; Herrmann, C.; Ivo, R.; Lourtie, I.M.G.; Wamhoff, H.; Melo, E. Consequences of the Partition between Water and Molecular Aggregates on the Photodegradation Pattern and Kinetics of a Pesticide of the Coumarin Family. *Environ. Sci. Technol.* **1995**, *29*, 136–141. [[CrossRef](#)] [[PubMed](#)]
16. Moreno, M.J.; Lourtie, I.M.G.; Melo, E. Stochastic effects on the time-dependent rate constant of photodimerization of 12-(9-anthroyloxy)stearic acid in micelles. *J. Phys. Chem.* **1996**, *100*, 18192–18200. [[CrossRef](#)]
17. Berberansantos, M.N.; Prieto, M.J.E. Energy-Transfer in Spherical Geometry—Application to Micelles. *J. Chem. Soc.-Faraday Trans. II* **1987**, *83*, 1391–1410. [[CrossRef](#)]



18. Matsuoka, K.; Moroi, Y. Micelle formation of sodium deoxycholate and sodium ursodeoxycholate (Part 1). *Biochim. Biophys. Acta* **2002**, *1580*, 189–199. [[CrossRef](#)]
19. Tanford, C. Thermodynamics of Micelle Formation—Prediction of Micelle Size and Size Distribution. *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 1811–1815. [[CrossRef](#)]
20. Oliver, R.C.; Lipfert, J.; Fox, D.A.; Lo, R.H.; Doniach, S.; Columbus, L. Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group. *PLoS ONE* **2013**, *8*, 2488. [[CrossRef](#)]
21. Davis, H.T. Factors Determining Emulsion-Type—Hydrophile-Lipophile Balance and Beyond. *Colloids Surf. A* **1994**, *91*, 9–24. [[CrossRef](#)]
22. Menger, F.M.; McCreery, M.J. Kinetic Characterization of Bile-Salt Micelles. *J. Am. Chem. Soc.* **1974**, *96*, 121–126. [[CrossRef](#)]
23. Bolt, J.D.; Turro, N.J. Measurement of the Rates of Detergent Exchange Between Micelles and the Aqueous Phase Using Phosphorescent Labeled Detergents. *J. Phys. Chem.* **1981**, *85*, 4029–4033. [[CrossRef](#)]
24. Rharbi, Y.; Winnik, M.A. Solute exchange between surfactant micelles by micelle fragmentation and fusion. *Adv. Colloid Interface Sci.* **2001**, *89*, 25–46. [[CrossRef](#)]
25. Coreta-Gomes, F.M.; Vaz, W.L.C.; Moreno, M.J. Effect of Acyl Chain Length on the Rate of Phospholipid Flip-Flop and Intermembrane Transfer. *J. Membr. Biol.* **2018**, *251*, 431–442. [[CrossRef](#)] [[PubMed](#)]
26. Eaton, B.E.; Gold, L.; Zichi, D.A. Lets Get Specific—The Relationship Between Specificity and Affinity. *Chem. Biol.* **1995**, *2*, 633–638. [[CrossRef](#)]
27. Gupta, A.B.; Mukherjee, S.; Pan, C.Q.; Velazquez-Campoy, A.; Sivaratanan, J.; Low, B.C. Spatial arrangement of LD motif-interacting residues on focal adhesion targeting domain of Focal Adhesion Kinase determine domain-motif interaction affinity and specificity. *Biochim. Biophys. Acta-Gen. Subj.* **2020**, *1864*, 129450. [[CrossRef](#)] [[PubMed](#)]
28. Cardoso, R.M.S.; Martins, P.A.T.; Ramos, C.V.; Cordeiro, M.M.; Leote, R.J.B.; Naqvi, K.R.; Vaz, W.L.C.; Moreno, M.J. Effect of dipole moment on amphiphile solubility and partition into liquid ordered and liquid disordered phases in lipid bilayers. *Biochim. Biophys. Acta (BBA)-Biomembr.* **2020**, *1862*, 183157. [[CrossRef](#)] [[PubMed](#)]
29. Filipe, H.A.L.; Coreta-Gomes, F.M.; Velazquez-Campoy, A.; Almeida, A.R.; Peixoto, A.F.; Pereira, M.M.; Vaz, W.L.C.; Moreno, M.J. Synthesis and Characterization of a Lipidic Alpha Amino Acid: Solubility and Interaction with Serum Albumin and Lipid Bilayers. *J. Phys. Chem. B* **2013**, *117*, 3439–3448. [[CrossRef](#)]
30. Kramer, S.D.; Braun, A.; Jakits-Deiser, C.; Wunderli-Allenspach, H. Towards the predictability of drug-lipid membrane interactions: The pH-dependent affinity of propranolol to phosphatidylinositol containing liposomes. *Pharm. Res.* **1998**, *15*, 739–744. [[CrossRef](#)]
31. Liu, X.Y.; Yang, Q.; Kamo, N.; Miyake, J. Effect of liposome type and membrane fluidity on drug-membrane partitioning analyzed by immobilized liposome chromatography. *J. Chromatogr. A* **2001**, *913*, 123–131. [[CrossRef](#)]
32. Manuel, M.; Martins, J. Partitioning of 1-pyrenesulfonate into zwitterionic and mixed zwitterionic/anionic fluid phospholipid bilayers. *Chem. Phys. Lipids* **2008**, *154*, 79–86. [[CrossRef](#)]
33. Martins, P.T.; Velazquez-Campoy, A.; Vaz, W.L.C.; Cardoso, R.M.S.; Valerio, J.; Moreno, M.J. Kinetics and Thermodynamics of Chlorpromazine Interaction with Lipid Bilayers: Effect of Charge and Cholesterol. *J. Am. Chem. Soc.* **2012**, *134*, 4184–4195. [[CrossRef](#)]
34. Moreno, M.J.; Bastos, M.; Velazquez-Campoy, A. Partition of amphiphilic molecules to lipid bilayers by isothermal titration calorimetry. *Anal. Biochem.* **2010**, *399*, 44–47. [[CrossRef](#)] [[PubMed](#)]
35. Moreno, M.J.; Prieto, M. Interaction of the Peptide-Hormone Adrenocorticotropin, ACTH(1- 24), with a Membrane Model System—A Fluorescence Study. *Photochem. Photobiol.* **1993**, *57*, 431–437. [[CrossRef](#)] [[PubMed](#)]
36. Saveyn, P.; Cocquyt, E.; Sinnaeve, D.; Martins, J.C.; Topgaard, D.; Van der Meeren, P. NMR study of the sorption behavior of benzyl alcohol derivatives into sonicated and extruded dioctadecyldimethylammonium chloride (DODAC) dispersions: The relevance of membrane fluidity. *Langmuir* **2008**, *24*, 3082–3089. [[CrossRef](#)]
37. Thomae, A.V.; Koch, T.; Panse, C.; Wunderli-Allenspach, H.; Kramer, S.D. Comparing the lipid membrane affinity and permeation of drug-like acids: The intriguing effects of cholesterol and charged lipids. *Pharm. Res.* **2007**, *24*, 1457–1472. [[CrossRef](#)] [[PubMed](#)]
38. Thomae, A.V.; Wunderli-Allenspach, H.; Kramer, S.D. Permeation of aromatic carboxylic acids across lipid bilayers: The pH-partition hypothesis revisited. *Biophys. J.* **2005**, *89*, 1802–1811. [[CrossRef](#)]
39. Bender, F.C.; Whitbeck, J.C.; de Leon, M.P.; Lou, H.; Eisenberg, R.J.; Cohen, G.H. Specific association of glycoprotein B with lipid rafts during herpes simplex virus entry. *J. Virol.* **2003**, *77*, 9542–9552. [[CrossRef](#)]
40. Kamlekar, R.K.; Simanshu, D.K.; Gao, Y.G.; Kenoth, R.; Pike, H.M.; Prendergast, F.G.; Malinina, L.; Molotkovsky, J.G.; Venyaminov, S.Y.; Patel, D.J.; et al. The glycolipid transfer protein (GLTP) domain of phosphoinositol 4-phosphate adaptor protein-2 (FAPP2): Structure drives preference for simple neutral glycosphingolipids. *Biochim. Biophys. Acta* **2013**, *1831*, 417–427. [[CrossRef](#)]
41. Pina, D.G.; Johannes, L.; Castanho, M. Shiga toxin B-subunit sequential binding to its natural receptor in lipid membranes. *Biochim. Biophys. Acta-Biomembr.* **2007**, *1768*, 628–636. [[CrossRef](#)]
42. Ruyschaert, J.M.; Lonz, C. Role of lipid microdomains in TLR-mediated signalling. *Biochim. Biophys. Acta-Biomembr.* **2015**, *1848*, 1860–1867. [[CrossRef](#)]
43. Ladbury, J.E. Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chem. Biol.* **1996**, *3*, 973–980. [[CrossRef](#)]
44. Schreiber, G.; Keating, A.E. Protein binding specificity versus promiscuity. *Curr. Opin. Struct. Biol.* **2011**, *21*, 50–61. [[CrossRef](#)] [[PubMed](#)]

45. Claveria-Gimeno, R.; Vega, S.; Abian, O.; Velazquez-Campoy, A. A look at ligand binding thermodynamics in drug discovery. *Expert Opin. Drug Discov.* **2017**, *12*, 363–377. [[CrossRef](#)] [[PubMed](#)]
46. Vega, S.; Abian, O.; Velazquez-Campoy, A. Handling complexity in biological interactions: Allostery and cooperativity in proteins. *J. Therm. Anal. Calorim.* **2019**, *138*, 3229–3248. [[CrossRef](#)]
47. Velazquez-Campoy, A.; Freire, E. ITC in the Post-Genomic era...? Priceless. *Biophys. Chem.* **2005**, *115*, 115–124. [[CrossRef](#)]
48. Litman, T.; Zeuthen, T.; Skovsgaard, T.; Stein, W.D. Competitive, non-competitive and cooperative interactions between substrates of P-glycoprotein as measured by its ATPase activity. *Biochim. Biophys. Acta* **1997**, *1361*, 169–176. [[CrossRef](#)]
49. Rabbani, G.; Baig, M.H.; Lee, E.J.; Cho, W.K.; Ma, J.Y.; Choi, I. Biophysical Study on the Interaction between Eperisone Hydrochloride and Human Serum Albumin Using Spectroscopic, Calorimetric, and Molecular Docking Analyses. *Mol. Pharm.* **2017**, *14*, 1656–1665. [[CrossRef](#)]
50. Gelamo, E.L.; Silva, C.; Imasato, H.; Tabak, M. Interaction of bovine (BSA) and human (HSA) serum albumins with ionic surfactants: Spectroscopy and modelling. *Biochim. Biophys. Acta* **2002**, *1594*, 84–99. [[CrossRef](#)]
51. Li, G.X.; Liu, B.S.; Zhang, Q.J.; Han, R. Investigation on the effect of fluorescence quenching of bovine serum albumin by cefoxitin sodium using fluorescence spectroscopy and synchronous fluorescence spectroscopy. *Luminescence* **2016**, *31*, 1054–1062. [[CrossRef](#)]
52. Sengupta, B.; Sengupta, P.K. Binding of quercetin with human serum albumin: A critical spectroscopic study. *Biopolymers* **2003**, *72*, 427–434. [[CrossRef](#)]
53. Hebert, P.C.; MacManus-Spencer, L.A. Development of a Fluorescence Model for the Binding of Medium- to Long-Chain Perfluoroalkyl Acids to Human Serum Albumin Through a Mechanistic Evaluation of Spectroscopic Evidence. *Anal. Chem.* **2010**, *82*, 6463–6471. [[CrossRef](#)]
54. Vuignier, K.; Schappler, J.; Veuthey, J.L.; Carrupt, P.A.; Martel, S. Drug-protein binding: A critical review of analytical tools. *Anal. Bioanal. Chem.* **2010**, *398*, 53–66. [[CrossRef](#)] [[PubMed](#)]
55. Cooper, M.A. Label-free screening of bio-molecular interactions. *Anal. Bioanal. Chem.* **2003**, *377*, 834–842. [[CrossRef](#)] [[PubMed](#)]
56. Du, X.; Li, Y.; Xia, Y.L.; Ai, S.M.; Liang, J.; Sang, P.; Ji, X.L.; Liu, S.Q. Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int. J. Mol. Sci.* **2016**, *17*, 144. [[CrossRef](#)]
57. Mackay, D.; Panjehshahin, M.R.; Bowmer, C.J. Analysis of the Binding of Fluorescent Ligands to Soluble- Proteins—Use of Simultaneous Nonlinear Least-Squares Regression to Obtain Estimates of Binding Parameters. *Biochem. Pharmacol.* **1991**, *41*, 2011–2018. [[CrossRef](#)]
58. Abreu, M.S.C.; Estronca, L.M.B.B.; Moreno, M.J.; Vaz, W.L.C. Binding of a Fluorescent Lipid Amphiphile to Albumin and its Transfer to Lipid Bilayer Membranes. *Biophys. J.* **2003**, *84*, 386–399. [[CrossRef](#)]
59. Cardoso, R.M.S.; Filipe, H.A.L.; Gomes, F.; Moreira, N.D.; Vaz, W.L.C.; Moreno, M.J. Chain Length Effect on the Binding of Amphiphiles to Serum Albumin and to POPC Bilayers. *J. Phys. Chem. B* **2010**, *114*, 16337–16346. [[CrossRef](#)]
60. Haque, M.E.; Ray, S.; Chakrabarti, A. Polarity estimate of the hydrophobic binding sites in erythroid spectrin: A study by pyrene fluorescence. *J. Fluoresc.* **2000**, *10*, 1–6. [[CrossRef](#)]
61. Srinivasan, B. A guide to the Michaelis-Menten equation: Steady state and beyond. *FEBS J.* **2021**, 16124. [[CrossRef](#)]
62. Velazquez-Campoy, A.; Goni, G.; Peregrina, J.R.; Medina, M. Exact analysis of heterotropic interactions in proteins: Characterization of cooperative ligand binding by isothermal titration calorimetry. *Biophys. J.* **2006**, *91*, 1887–1904. [[CrossRef](#)]
63. Freire, E.; Kawasaki, Y.; Velazquez-Campoy, A.; Schon, A. Characterisation of Ligand Binding by Calorimetry. In *Biophysical Approaches Determining Ligand Binding to Biomolecular Targets: Detection, Measurement and Modelling*; Podjarny, A., Dejaegere, A.P., Kieffer, B., Eds.; Royal Society of Chemistry: Paris, France, 2011; pp. 275–299.
64. Bastos, M.; Velazquez-Campoy, A. Isothermal titration calorimetry (ITC): A standard operating procedure (SOP). *Eur. Biophys. J. Biophys. Lett.* **2021**, *50*, 363–371. [[CrossRef](#)]
65. Claveria-Gimeno, R.; Velazquez-Campoy, A.; Pey, A.L. Thermodynamics of cooperative binding of FAD to human NQO1: Implications to understanding cofactor-dependent function and stability of the flavoproteome. *Arch. Biochem. Biophys.* **2017**, *636*, 17–27. [[CrossRef](#)] [[PubMed](#)]
66. Claveria-Gimeno, R.; Vega, S.; Abian, O.; Velazquez-Campoy, A. Tinkering with Binding Polynomials in Isothermal Titration Calorimetry. In *Microcalorimetry of Biological Molecules: Methods and Protocols*; Ennifar, E., Ed.; Springer: Berlin/Heidelberg, Germany, 2019; Volume 1964, pp. 185–213.
67. Freire, E.; Schon, A.; Velazquez-Campoy, A. Isothermal Titration Calorimetry: General Formalism using Binding Polynomials. *Method. Enzymol.* **2009**, *455*, 127–155. [[CrossRef](#)]
68. Vega, S.; Abian, O.; Velazquez-Campoy, A. A unified framework based on the binding polynomial for characterizing biological systems by isothermal titration calorimetry. *Methods* **2015**, *76*, 99–115. [[CrossRef](#)]
69. Taneva, S.G.; Banuelos, S.; Falces, J.; Arregi, I.; Muga, A.; Konarev, P.V.; Svergun, D.I.; Velazquez-Campoy, A.; Urbaneja, M.A. A Mechanism for Histone Chaperone Activity of Nucleoplasmin: Thermodynamic and Structural Models. *J. Mol. Biol.* **2009**, *393*, 448–463. [[CrossRef](#)] [[PubMed](#)]
70. Ward, L.D.; Timasheff, S.N. Cooperative Multiple Binding of Bisans and Daunomycin to Tubulin. *Biochemistry* **1994**, *33*, 11891–11899. [[CrossRef](#)] [[PubMed](#)]
71. Kutner, M.; Nachtsheim, C.; Neter, J.; Li, W. *Applied Linear Statistical Models*, 5th ed.; McGraw-Hill/Irwin: Boston, MA, USA, 2004.
72. Alesio, J.L.; Slitt, A.; Bothun, G.D. Critical new insights into the binding of poly- and perfluoroalkyl substances (PFAS) to albumin protein. *Chemosphere* **2022**, *287*, 131979. [[CrossRef](#)]

73. Estronca, L.M.B.B.; Moreno, M.J.; Vaz, W.L.C. Kinetics and thermodynamics of the association of dehydroergosterol with lipid bilayer membranes. *Biophys. J.* **2007**, *93*, 4244–4253. [[CrossRef](#)]
74. Martins, P.A.T.; Domingues, N.; Pires, C.; Alves, A.M.; Palmeira, T.; Samelo, J.; Cardoso, R.; Velazquez-Campoy, A.; Moreno, M.J. Molecular crowding effects on the distribution of amphiphiles in biological media. *Colloids Surf. B* **2019**, *180*, 319–325. [[CrossRef](#)]
75. Martins, P.A.; Gomes, F.; Vaz, W.L.C.; Moreno, M.J. Binding of Phospholipids to  $\alpha$ -Lactoglobulin and their Transfer to Lipid Bilayers. *Biochim. Biophys. Acta* **2008**, *1778*, 1308–1315. [[CrossRef](#)]
76. Abe, Y.; Fukui, S.; Koshiji, Y.; Kobayashi, M.; Shoji, T.; Sugata, S.; Nishizawa, H.; Suzuki, H.; Iwata, I. Enantioselective binding sites on bovine serum albumin to dansyl amino acids. *Biochim. Biophys. Acta* **1999**, *1433*, 188–197. [[CrossRef](#)]
77. Epps, D.E.; Raub, T.J.; Kezdy, F.J. A General, Wide-Range Spectrofluorometric Method for Measuring the Site-Specific Affinities of Drugs Towards Human Serum- Albumin. *Anal. Biochem.* **1995**, *227*, 342–350. [[CrossRef](#)] [[PubMed](#)]
78. Bischel, H.N.; MacManus-Spencer, L.A.; Zhang, C.J.; Luthy, R.G. Strong Associations of Short-Chain Perfluoroalkyl Acids with Serum Albumin and Investigation of Binding Mechanisms. *Environ. Toxicol. Chem.* **2011**, *30*, 2423–2430. [[CrossRef](#)] [[PubMed](#)]
79. Allendorf, F.; Berger, U.; Goss, K.U.; Ulrich, N. Partition coefficients of four perfluoroalkyl acid alternatives between bovine serum albumin (BSA) and water in comparison to ten classical perfluoroalkyl acids. *Environ. Sci.-Processes Impacts* **2019**, *21*, 1852–1863. [[CrossRef](#)] [[PubMed](#)]
80. Videira, R.A.; Antunesmadeira, M.D.; Custodio, J.B.A.; Madeira, V.M.C. Partition of Dde in Synthetic and Native Membranes Determined by Ultraviolet Derivative Spectroscopy. *Biochim. Biophys. Acta-Biomembr.* **1995**, *1238*, 22–28. [[CrossRef](#)]
81. Andrieux, K.; Forte, L.; Lesieur, S.; Paternostre, M.; Ollivon, M.; Grabielle-Madelmont, C. Insertion and partition of sodium taurocholate into egg phosphatidylcholine vesicles. *Pharm. Res.* **2004**, *21*, 1505–1516. [[CrossRef](#)] [[PubMed](#)]
82. Heerklotz, H. Interactions of surfactants with lipid membranes. *Q. Rev. Biophys.* **2008**, *41*, 205–264. [[CrossRef](#)]
83. Lolicato, F.; Juhola, H.; Zak, A.; Postila, P.A.; Saukko, A.; Rissanen, S.; Enkavi, G.; Vattulainen, I.; Kepczynski, M.; Rog, T. Membrane-Dependent Binding and Entry Mechanism of Dopamine into Its Receptor. *ACS Chem. Neurosci.* **2020**, *11*, 1914–1924. [[CrossRef](#)]
84. Shahoei, R.; Tajkhorshid, E. Menthol Binding to the Human  $\alpha$ 4  $\beta$ 2 Nicotinic Acetylcholine Receptor Facilitated by Its Strong Partitioning in the Membrane. *J. Phys. Chem. B* **2020**, *124*, 1866–1880. [[CrossRef](#)]