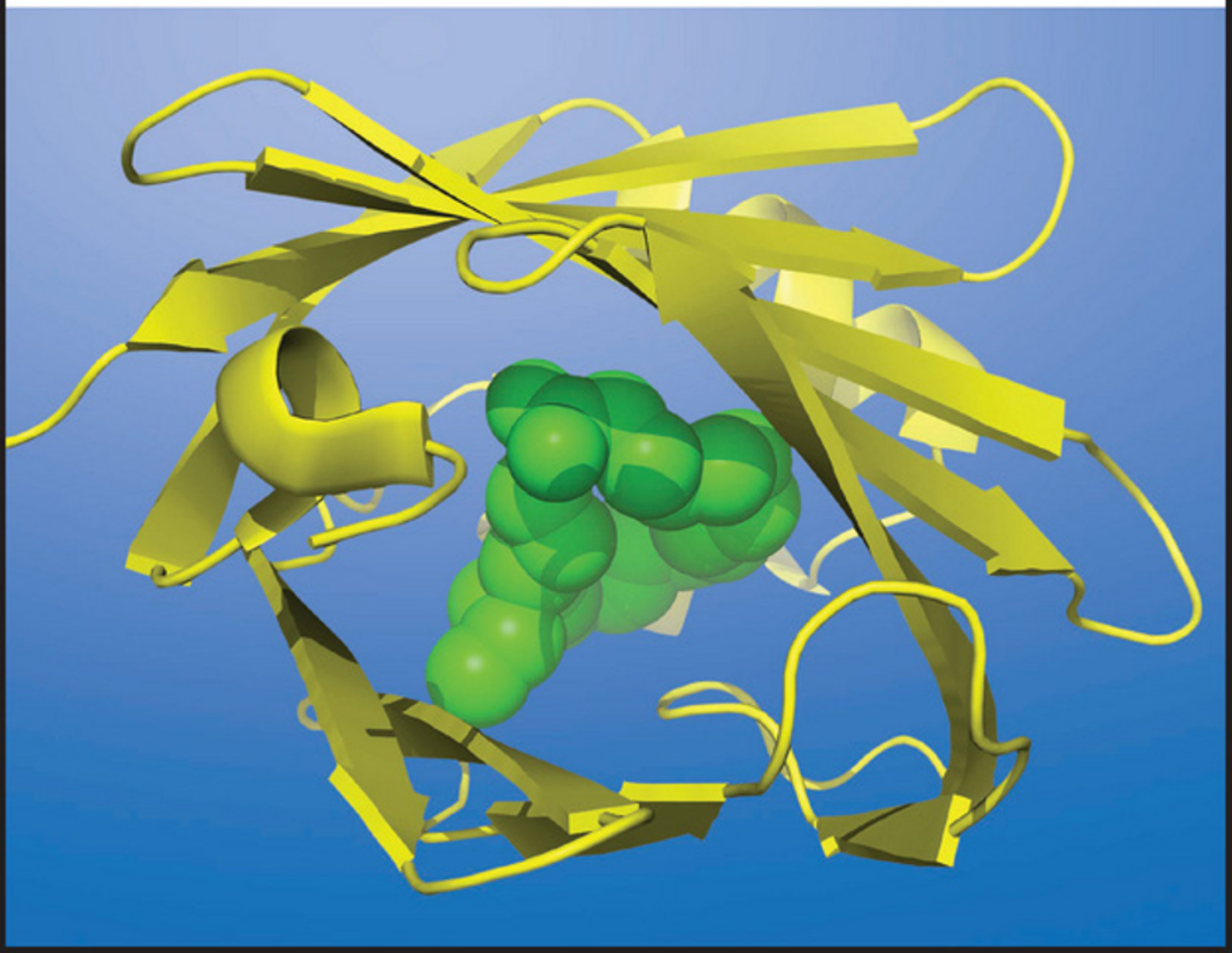
WILEY

Andrea Bellelli and Jannette Carey

Reversible Ligand Binding

Theory and Experiment



Reversible Ligand Binding

Reversible Ligand Binding

Theory and Experiment

Andrea Bellelli
Department of Biochemical Sciences,
Sapienza University of Rome, Italy

Jannette Carey
Department of Chemistry, Princeton University, USA

WILEY

This edition first published 2018

Registered Office(s)
John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Office
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Library of Congress Cataloging-in-Publication Data

Names: Bellelli, Andrea, 1958– author. | Carey, Jannette, author.
Title: Reversible ligand binding : theory and experiment / by Andrea Bellelli, Jannette Carey.
Description: First edition. | Hoboken, NJ : John Wiley & Sons, 2018. |
Includes bibliographical references and index. |
Identifiers: LCCN 2017026931 (print) | LCCN 2017040990 (ebook) | ISBN 9781119238478 (pdf) |
ISBN 9781119238492 (epub) | ISBN 9781119238485 (cloth)
Subjects: LCSH: Ligand binding (Biochemistry)
Classification: LCC QP517.L54 (ebook) | LCC QP517.L54 B45 2017 (print) | DDC 572/.33–dc23
LC record available at https://lccn.loc.gov/2017026931

Cover design by Wiley
Cover Image: Courtesy of RCSB Protein Data Bank (entry 1vyf, deposited by Angelucci et al. 2004)

Set in 10/12pt Warnock by SPi Global, Pondicherry, India

10 9 8 7 6 5 4 3 2 1

A.B. dedicates this book to Maurizio Brunori, teacher, mentor and friend for thirty-eight years until now and with more to come.

Contents

Preface

Ligand binding is a crucial event in virtually every biological phenomenon. Detailed understanding of many biologically relevant events including enzymatic catalysis, transport, and molecular recognition requires quantitative description of ligand binding. Such description may prove exquisitely complex because biological macromolecules may bind multiple ligands at once or alternatively and their reactions may present several types of thermodynamic linkage. The scope of this book is to provide a comprehensive view of the various biochemical considerations that govern reversible as well as irreversible ligand binding. Special attention is devoted to enzymology, a field usually treated separately from ligand binding, but actually governed by identical thermodynamic relationships.

This book is intended for PhD students and researchers, and aims at providing the understanding necessary to interpret ligand-binding experiments, formulate plausible reaction schemes, and analyze the data according to the chosen model(s). Attention is given to the design of the experiment because a properly designed experiment may provide clear evidence of biochemical features that can otherwise escape notice. Classical experiments are reviewed in order to further highlight the importance of the design of the experiment.

The book includes treatment of thermodynamic relationships that are most often left to the specialized literature, for example, ligand-linked dissociation. To make the book accessible to a general audience, we simplified the analysis of these relationships to the maximum possible extent, for example, whenever possible we adopted homodimeric proteins as model systems instead of more complex macromolecular assemblies. This choice allowed us to explore a large range of effects with minimally complex equations. Our scope throughout the book has been to present all the essential and distinguishing aspects of the phenomena we describe rigorously, but at the same time in the simplest possible form. Indeed, we are confident that every reader having competence in elementary algebra may take advantage of our work.

Acknowledgments

The authors of scientific books always contract debts with many people who, directly or indirectly, contributed ideas, suggestions, and comments. This book contains considerations and ideas that we elaborated over many years, long before our decision to write, and the list of people we would like to thank is very long. Moreover, the background of the authors being different, our two lists sum without overlapping.

We acknowledge the innumerable hours of education on the subject matters discussed in this book by our teachers and mentors. A.B. expresses his thanks to Maurizio Brunori, Eraldo Antonini, Bruno Giardina, Gino Amiconi, and Quentin H. Gibson.

We enjoyed many long and fruitful discussions with our colleagues, that often shaped our views as expressed in this book. Thanks are due to W.A. Eaton, S.J. Gill, T. Yonetani, M. Coletta, P. Ascenzi, G. Antonini, P. Sarti, F. Malatesta, E. Henry, P. Brzezinski, J.S. Olson, W.F. Xue, S. Linse, G.K. Ackers, and many others.

We express our thanks to our marvelous students, whose questions made us grow. Many of them are now well reputed researchers and close friends: R. Ippoliti, E. Lendaro, G. Boumis, A. Brancaccio, S. Santanché, A. Arcovito, F. Angelucci, L. Jin, B. Harish and R. Grandori. Special thanks are due to F. Saccoccia and A. Di Matteo who read the manuscript more than once and pointed out errors and unclear passages: their contribution has been invaluable.

Part I

Ligand Binding to Single Binding Site Targets

1

Theory of Ligand Binding to Monomeric Proteins

1.1 Importance of Ligand-Binding Phenomena in Biology

Reversible interactions between or among molecules underlie nearly every aspect of biology. To understand these interactions in a chemical way means to describe them quantitatively. To do so we must be able to determine their affinity, stoichiometry, and cooperativity by carrying out ligand-binding experiments. We use the term "ligand" in a way distinct from its use to discuss coordination complexes within inorganic chemistry. In ligand-binding theory we use this term to mean any interacting partner. Although many people consider a ligand to be a small molecule that interacts with a macromolecule, in fact, either partner can be considered to be the ligand of the other. In a typical experiment the concentration of one partner is held fixed while the concentration of the other partner is incremented. In ligand-binding theory and practice we define the ligand operationally as the partner whose concentration is incremented during the experiment. Such experiments resemble pH titrations both practically and theoretically, and thus are referred to as titrations. The partner whose concentration is held fixed is referred to as the "target," and again the definition is strictly operational, that is, either a small molecule or a macromolecule can be the target, depending on how the titration is set up.

As we will show later, there are good reasons to carry out experiments with first one partner, then the other, treated as the ligand; however, depending on the chemical properties of the two partners in the reaction, there may be practical limitations to, or conceptual difficulties in, the possibility of interchanging their respective roles as the ligand and the target. *In particular, if either of the reaction partners has multiple binding sites for the other, one may not obtain superimposable binding isotherms when exchanging the target and the ligand.*

Affinity refers to the strength of interaction between partners. Affinity is quantitatively expressed by an equilibrium constant that we measure in our experiment or, equivalently, a free energy difference between the bound and free states of the system that we calculate from the equilibrium constant.

Stoichiometry refers to the number of molecules of each partner that participate in the binding process, and it must also be determined by our measurements. In practice, what we really mean by stoichiometry is more often *molar ratio*. For example, in a process involving four molecules of one kind with two of another kind, the stoichiometry

Reversible Ligand Binding: Theory and Experiment, First Edition. Andrea Bellelli and Jannette Carey.

is 4:2 but the molar ratio is 2:1. The determination of true stoichiometries usually requires additional information from sources other than a binding experiment (e.g., molecular weights, and state of aggregation of the target and ligand in solution).

Thermodynamic linkage is a general term that applies to ligand-binding experiments in which the same target binds two or more molecules of the same or different ligands, and each ligand modulates the affinity of the target for the other. There are at least three different types of linkage, called identical, homotropic, and heterotropic (Wyman and Gill, 1990). *Identical linkage* occurs when two different ligands compete for the same binding site on the target, and their binding is mutually exclusive. This type of linkage is discussed in Section 1.8 for single binding site targets, and in Section 4.8 for multiple binding site targets. Competitive enzyme inhibition is a very important case of identical linkage and is described in Section 8.6. *Homotropic linkage* occurs when the target can bind more than a single molecule of the same ligand, with different affinity. Homotropic linkage can occur only in targets with multiple binding sites, thus its analysis is deferred to Chapter 4. Finally, *heterotropic linkage* occurs when the target can bind two different ligands in a non-exclusive manner and the binding of one ligand alters the affinity of the other. It is described in Section 1.9, and for targets with multiple binding sites in Section 4.9. Important examples of heterotropic linkage are uncompetitive enzyme-inhibition (Section 8.7), and regulation of the oxygen affinity of hemoglobin by effectors, including protons (Bohr effect), diphosphoglycerate, or inositol hexaphosphate, dealt with in Chapter 7.

Homotropic and heterotropic linkage are typically regarded as an emergent property unique to proteins, but some non-protein molecules of ~500–1000 Da have been shown to exhibit cooperative binding of their ligands (Rebek, 1985). An interesting and biologically relevant example is provided by the axial ligands of iron-porphyrins (Traylor and Sharma, 1992).

Homotropic and heterotropic linkages may be either positive (if each ligand increases the affinity of the other) or negative (if each ligand decreases the affinity of the other). *Cooperativity* has often been used as a synonym of linkage, but unfortunately not always with the necessary precision. Often, cooperativity, or positive cooperativity, is used to indicate positive homotropic linkage, but the terms of negative cooperativity or anti-cooperativity may be used to indicate a negative homotropic or heterotropic linkage. The definition of cooperativity is sufficiently general to encompass cases in which even monomeric proteins can respond cooperatively to two different ligands, for example, the physiological ligand and an ionic component of the solution (Weber, 1992). Note that in such cases the ion must be also considered as a physiological effector. Because the general applicability of a definition is inversely related to its precision, in this book, we shall prefer the terms positive or negative homotropic or heterotropic linkage whenever precision is required. Positive cooperativity occurs in several proteins and has special relevance in physiology. For example, the binding of oxygen to hemoglobin is cooperative in that oxygen affinity becomes stronger as binding progresses, as described in detail in Chapter 7.

In this chapter we describe the theoretical bases of ligand binding under equilibrium conditions for protein:ligand complexes with 1:1 stoichiometry; in the next chapter we discuss the kinetics of the same system, and in Chapter 3 we consider some practical aspects of experimental design, and some common sources of errors.

1.2 Preliminary Requirements for Ligand-Binding Study

Whether or not a given ligand binds to a given macromolecular target may be known from prior experiment or inferred from physiological or chemical data. When no such information is available, or when the equilibrium constant for the reaction is required, then the interaction must be evaluated from titration, using a method of analysis that is suitable for the solution conditions of interest and the concentrations of partners that can be practically achieved. Although specific methods will not be detailed here, all useful methods have in common that they offer some observable that changes during, and thus reports on, the binding process. Any observable is referred to as the *signal*, and its change relates to a shift between the bound and free states of the system.

Whether or not a ligand binds reversibly and without any chemical transformation must also be established. Ligand-binding theory and practice generally refers only to such cases; but we shall cover in this book also some common non-conformant cases, for example, thiol reagents and enzyme-substrate reactions, that of course do begin with a ligand-binding process followed by a chemical transformation. The most common approach to determining reversibility is to simply assume it; this is not appropriate for rigorous scientific work. Reversibility is generally established by showing that the signal change is reversed when concentrations are reduced, for example, by dilution or dialysis. A more stringent criterion is to show that the separated ligand and target are recovered unchanged after their interaction, although it can be difficult to rule out a small extent or minor degree of change. One of the best ways to do so is to repeat the binding measurement itself with the recovered materials to evaluate whether the affinity is the same. However, this method can fail with labile partners.

Finally, the molar ratio and stoichiometry must be determined from the same kinds of binding experiments that are used to determine affinity and cooperativity, using strategies that will be outlined later. Without knowing the correct molar ratio the interpretation of the affinity may be plagued by high uncertainty. Another common, unfortunate, and sometimes untested, assumption is that an interaction has a 1:1 molar ratio. As we will show in Chapter 3, the experimental design required to establish molar ratio is not difficult conceptually, and usually is not practically difficult either. Thus there is no reason whatsoever to leave this very important feature to untested assumption. The analysis of binding data is simpler for the 1:1 stoichiometry case, thus we will treat this case first before expanding the treatment to cases of any molar ratio.

1.3 Chemical Equilibrium and the Law of Mass Action

Every chemical reaction, *if allowed enough time*, reaches an equilibrium condition in which the rate of product formation from reactants equals the rate of product degradation to reactants. When this condition is reached the concentrations of reactants and products undergo no further net change. The specific chemical composition of the reaction mixture at equilibrium depends on experimental conditions such as temperature, pH, solution composition, and so on, but at any given set of conditions it obeys the law of mass action. This law states that, under the equilibrium condition, the ratio between the product of product concentrations and the product of reactant concentrations,

each raised to a power corresponding to its stoichiometric coefficient, equals a constant, the *equilibrium constant*. Technically, *molar activities should be used*, but under most experimental conditions the concentrations of ligand and protein are low enough to allow the researcher to neglect this distinction.

The equilibrium constant of a chemical reaction is independent of the initial concentrations of reagents and products, but varies with the solution conditions, as discussed further below. The equilibrium constant of a chemical reaction has the units of molarity elevated to a positive or negative, usually integer, power factor that corresponds to the difference between the stoichiometric coefficients of the products and reagents. If the sum of the stoichiometric coefficients of the products equals that of reagents, the equilibrium constant is a pure number.

Ligand binding to proteins usually conforms to the above description, and can be described as:

$$P + X \Leftrightarrow PX \qquad \text{(eqn. 1.1)}$$

where P represents the unliganded protein, X the molecule that binds to it and PX the protein-molecule complex. For practical reasons, it is often convenient to keep constant the protein concentration and to vary in the course of the experiment the concentration of the molecule X, in which case P is the target and X the ligand. Thus, unless differently specified, we shall assume that X is the ligand.

When the above reaction reaches its equilibrium condition, the law of mass action dictates:

$$K_a = [PX]/[P][X] \qquad \text{(eqn. 1.2)}$$

where the square brackets indicate the molar concentrations of the chemical species involved, and the subscript "a" indicates the association direction of the reaction.

A typical example of the above reaction is that of respiratory proteins that reversibly bind oxygen, for example, myoglobin (Mb):

$$Mb + O_2 \Leftrightarrow MbO_2$$

Reversible chemical reactions may be written, and observed experimentally, in either forward or reverse direction. Thus one may write reaction 1.1 in the form of the dissociation of the protein-ligand complex:

$$PX \Leftrightarrow P + X$$

with

$$K_d = [P][X]/[PX] \qquad \text{(eqn. 1.3)}$$

where the subscript "d" refers to the dissociation direction. From equations 1.2 and 1.3 we observe that the equilibrium constant of the dissociation reaction is the reciprocal of that of the combination reaction. The dissociation equilibrium constant has units of molar concentration, whereas K_a has units of reciprocal concentration. Thus K_d can be directly compared to the concentration of the free ligand. Exceptions to this rule do exist (see Section 1.11) and can be a cause of confusion.

A consistent and unique formulation for ligand-binding experiments, using either the association or the dissociation reactions and equilibrium constants, would probably be

desirable, but is not likely to be universally adopted. Indeed, we usually design an experiment considering the association reaction, because adding the ligand to the unliganded protein is more obvious and easier to do than dissociating an already-formed complex (although dissociation is possible, e.g., by dilution, chemical transformation of the unbound ligand, or phase extraction). Once the association experiment has been carried out, however, we often switch our reasoning to the dissociation reaction because the dissociation constant, having the units of a molar concentration, corresponds to a point on the ligand concentration axis of the binding plot, and actually the ligand concentration itself can be expressed as a multiple or sub-multiple of K_d. Thus, in the literature one finds experiments, analyses, and models developed in both ways and must be familiar with both.

Unless one directly measures [P], [X], and [PX] after the equilibrium condition has been reached, one only knows the total concentrations of the ligand and the protein, that is, $[X]_{tot} = [X] + [PX]$ and $[P]_{tot} = [P] + [PX]$ (for a 1:1 reaction). If this is the case, eqn. 1.3 should be rewritten as:

$$K_d = \left([P]_{tot} - [PX]\right)\left([X]_{tot} - [PX]\right)/[PX] \qquad \text{(eqn. 1.4)}$$

Although eqn. 1.4 can be easily solved for [PX], yielding a second-degree equation, the procedure is not completely straightforward. Indeed, *the more [PX] approaches $[X]_{tot}$, the greater the uncertainty in [X] and, consequently, in K_d.*

A great simplification can be achieved under conditions in which either (i) both the free and bound ligand (i.e., [X] and [PX]) can be measured directly (e.g., by using equilibrium dialysis); or (ii) $[X]_{tot} >> [PX]$. As we shall demonstrate in Section 1.4, the latter condition implies $[P]_{tot} << K_d$. If the experiment can be run under this condition, only a small fraction of the total ligand will be bound and $[X] \approx [X]_{tot}$, making the use of eqn. 1.4 unnecessary, and allowing direct use of eqns. 1.2 or 1.3. It may happen that, depending on K_d and the experimental method chosen, this condition cannot be met, as it would require protein concentrations too low to be detected. We shall discuss experimental approaches that may overcome this limitation in Sections 1.8 and 1.9.

1.4 The Hyperbolic and Sigmoidal Representations of the Ligand-Binding Isotherms

The graphical representation of binding measurements is important because it is usually difficult to visualize equations like those in the above paragraph or the more complex ones we shall encounter in the following chapters. Some graphical representations may offer clear indications of some property of the system, but may distort or alter other properties. Thus some caution in their use is in order, especially when we want not only to look at them, but to use them for quantitative analysis, that is, to determine the values of the parameters describing the binding reaction. In the present section we shall describe the simplest graphical representations of the ligand-binding isotherm, that is, the hyperbolic plot of [PX] versus [X] and its variants. These representations do not usually distort the experimental error, and should be preferred for quantitative analysis. More complex, and distorting, plots will be considered in a following section.

The soundest and statistically least biased way to represent ligand-binding data is to plot the signal recorded in the experiment, whatever it may be, as a function of the free

ligand concentration. The signal will be discussed further in Chapter 3; it is a detectable physical property of the system that depends on the concentrations of P, PX, or both. Thus the graph of the signal is essentially equivalent to a plot of [P] or [PX] versus [X]. However, for the sake of clarity, the researcher may decide to calculate and report the concentration of the bound protein ([PX]) or its fraction ($[PX]/[P]_{tot}$) rather than the experimentally recorded signal. The graph of the fraction of bound protein, that is, the ratio between bound and total binding sites, versus the concentration of the free ligand, or total ligand if $[X]_{tot} >> [PX]$, is probably the most commonly adopted representation in the biological context.

In order to define the relationship between [PX] and [X], we need to define the *binding polynomial* of the reaction. The binding polynomial expresses the sum of all species of the target as a function of one of them that is adopted as a reference species. For example, if the experiment conforms to eqn. 1.3 and we adopt as a reference the concentration of the unliganded protein P, we can write:

$$\begin{aligned} [PX] &= [P][X]/K_d \\ [P]_{tot} &= [P]+[PX] = [P]\left(1+[X]/K_d\right) \end{aligned} \qquad \text{(eqn. 1.5a)}$$

Eqn. 1.5 represents the binding polynomial of reaction 1.1. Notice that the direction in which the reaction is written, either combination or dissociation, is irrelevant because the binding polynomial does not distinguish reactants and products. The definition of the binding polynomial for such a simple chemical mechanism is obvious, but we write it explicitly in view of its pedagogical value for more complex reaction schemes, to be described in the following chapters.

From eqn. 1.5 we obtain:

$$[P] = [P]_{tot} K_d/\left(K_d + [X]\right) \qquad \text{(eqn. 1.5b)}$$

or

$$[PX] = [P]_{tot} - [P] = [P]_{tot}[X]/\left(K_d + [X]\right) \qquad \text{(eqn. 1.5c)}$$

Both equations correspond to rectangular hyperbolas, which asymptotically tend either to zero (1.5b) or to $[P]_{tot}$ (eqn. 1.5c).

Eqn. 1.5c is most often employed and can be rearranged to represent the fractional saturation, defined as the fraction of bound over total ligand binding sites:

$$\bar{X} = [PX]/[P]_{tot} = [X]/\left(K_d + [X]\right) \qquad \text{(eqn. 1.6)}$$

The fractional saturation is called in the ligand-binding literature $\bar{Y}$, $\bar{X}$, θ, or ν. We prefer $\bar{X}$ since in reactions involving two or more ligands this allows us to call $\bar{X}$ the fractional saturation of the target for ligand X, $\bar{Y}$ that for ligand Y and so on (see below).

Eqn. 1.6 describes a rectangular hyperbola with unitary asymptote (Figure 1.1A). For many biological systems this is the most meaningful representation of experimental data and it has been preferred traditionally by physiologists studying systems as different as oxygen carriers, hormone receptors, and enzymes. This representation has also the advantage of introducing minimal distortions in the experimental data and their errors. It is important to stress that the ligand concentration that appears in the law of

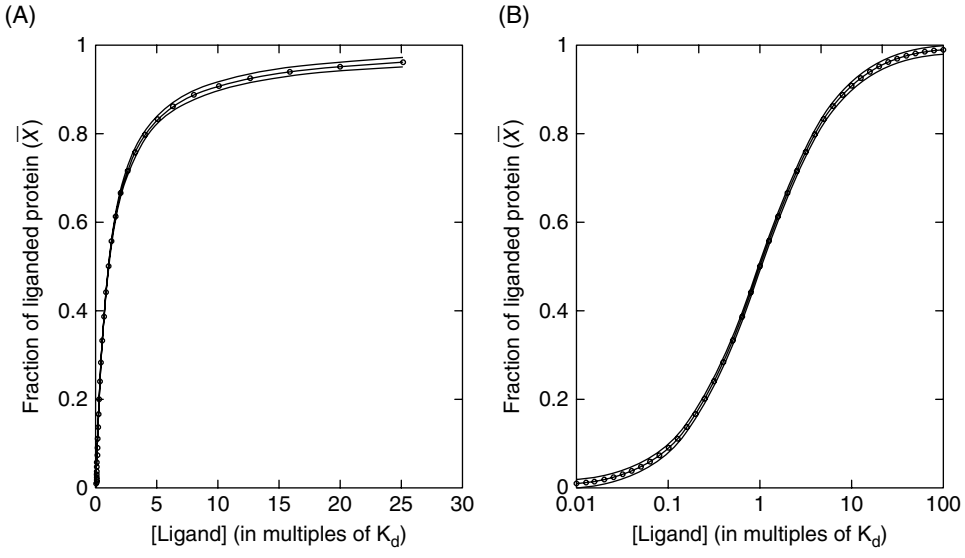


Figure 1.1 Two common representations of ligand-binding isotherms. Panel A: the $\bar{X}$ versus [X] plot; Panel B: same as A but [X] is logarithmically scaled. In both panels the points are calculated from eqn. 1.6. The two lines above and below the points on each panel are obtained by adding and subtracting a constant error of 0.01 $\bar{X}$ units, and demonstrate that in these representations there is no distortion of the error.

mass action, and hence in all the equations we developed thus far, except eqn. 1.4, refers to the free ligand, that is, $[X]_{tot}$-[PX].

A variant of the hyperbolic plot $\bar{X}$ versus [X] is obtained if one reports $\bar{X}$ as a function of the logarithm of the free ligand concentration. Chemists may prefer this representation over the hyperbolic one because the chemical potential of the ligand is proportional to the logarithm of its concentration (or activity), and thus the plot emphasizes some properties of the system that do not appear clearly in the hyperbolic plot. However, it is unusual for the physiological ligand concentrations to vary over several orders of magnitude; thus, this plot compresses the physiological range of ligand concentrations in a small region of the curve and physiologists seldom use this representation. The $\bar{X}$ versus log([X]) plot has the same overall accuracy as $\bar{X}$ versus [X] plot, and does not distort the experimental error. Actually, the error on ligand concentration, which is usually assumed to be negligible, may in some cases be better represented on a logarithmic than a linear scale (see Chapter 3).

From eqn. 1.6 one may calculate the free ligand concentration (expressed as a multiple or sub-multiple of K_d, as shown on Figure 1.1) required to achieve any value of $\bar{X}$. As an example, we may calculate that at [X]=1/3 K_d we have $\bar{X} = 0.25$ and at [X]=3 K_d, $\bar{X} = 0.75$; or that to raise $\bar{X}$ from 0.1 to 0.9 an increase of the logarithm of [X] of 1.91 units is required. Generalizing, if we consider ligand concentrations that are multiples or sub-multiples of the K_d by the same factor i, we have that $\bar{X}_{[X]=Kd/i} = 1 - \bar{X}_{[X]=iKd}$, which demonstrates that *the $\bar{X}$ versus log [X] plot is symmetric* (at least in the cases of 1:1 and 1:2 stoichiometries).

An important property of the $\bar{X}$ versus log([X]) plot is that its shape is perfectly invariant with respect to K_d: that is, if one explores a set of experimental conditions that cause K_d to vary, one obtains a series of symmetric sigmoidal curves of identical shape, shifted right or left according to their different K_ds. The shape of the plot can be quantified

as the approximate slope of the straight line passing through two points symmetric with respect to the center of the curve. For example, the points $\bar{X}_{[X]=Kd/3} = 0.25$ and $\bar{X}_{[X]=3Kd} = 0.75$ are symmetric in the sense defined above, and are joined by a straight line with slope ~0.52 (for a single-site macromolecule). The slope of the $\bar{X}$ versus log ([X]) plot was called the *binding capacity* of the macromolecule by Di Cera and Gill (1988), in analogy with the physical concept of heat capacity. This terminology is potentially confusing because the term capacity is also used to indicate the total amount of some substance that the sample may contain (e.g., the maximum amount of oxygen that a given volume of blood can transport is commonly referred to as its oxygen capacity), which is not what these authors meant. Nevertheless, the concept conveyed by the slope is important because changes in slope from the canonical value given above provide important clues on the binding properties of the protein, and/or the composition of the experimental system. In particular, as we shall demonstrate further on, positive cooperativity increases the slope, whereas chemical heterogeneity of the protein or negative cooperativity decreases it (see Chapters 3 and 4).

In the present chapter we deal only with single-site targets, whose changes in ligand affinity shift the position of the $\bar{X}$ versus log([X]) curve along the X axis but do not change its shape, and in particular do not change its slope; this will be true for all types of linkage considered in this chapter (Sections 1.8–1.10) and for the effect of temperature (Section 1.7). However we call the attention of the reader to this point, because in later chapters we shall describe systems having steeper or shallower $\bar{X}$ versus log[X] plots, and we shall develop an interpretation of the increased or decreased slope of these plots, which may constitute the first indication that the protein:ligand stoichiometry differs from 1:1.

The derivative of the $\bar{X}$ versus log([X]) plot, $\Delta\bar{X}$ versus log([X]), has been used by S.J. Gill to describe the ligand-binding isotherms of hemoglobin, as recorded using the thin layer dilution method (Figure 1.2) (Gill *et al.*, 1987). This is a very specialized

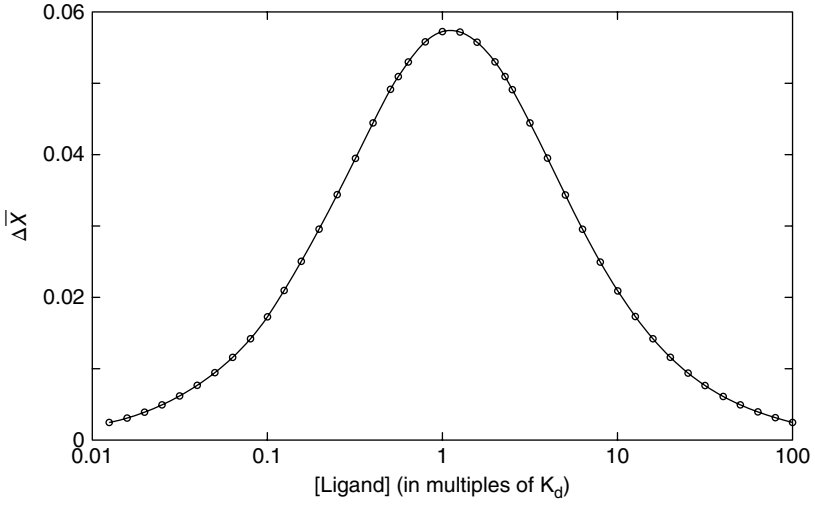


Figure 1.2 Plot of $\bar{X}$ increments versus [X] (on a log scale). Same data set as in Figure 1.1; because the errors added to calculated $\bar{X}$ values are constant and independent of $\bar{X}$, in this plot the points and the lines exactly overlap. Notice that the values on the ordinate axis depend on the dimension of the intervals of log [X] explored and that to obtain a readable plot, these should be kept constant.

plot justified by the experimental method used, and true to the experimental data in that it aims to minimize the error distortion implicit in other types of plots. However, this graphical representation is strongly linked to the experimental method, and greatly benefits from constant logarithmic steps of increase or decrease of the ligand concentration, which is why it never gained widespread use. A more general use of the derivative plot of the sigmoid-binding curve, applicable independently of the specific experimental setup, is to reveal the inflection point (i.e., the midpoint, $\mathbf{X_{1/2}}$, whose significance is discussed in the next Section) even when the sigmoid is incomplete.

1.5 The Important Concept of $X_{1/2}$

An important point of the $\bar{X}$ versus [X] or the $\bar{X}$ versus log [X] plots is their midpoint. We may calculate that $\bar{X}$ equals 0.5 (saturation of half the available binding sites) when [X]=K_d, consistent with the symmetry of the $\bar{X}$ versus log [X] plot for monomeric proteins.

The free ligand concentration required to achieve $\bar{X} = 0.5$ is called $\mathbf{X_{1/2}}$ (or $\mathbf{X_{50}}$) and is always a function of K_d. However, the simple equation $\mathbf{X_{1/2}}$=K_d applies only to the simple cases of monomeric, single-site proteins or multiple, identical, non-interacting sites; we shall consider below some cases in which this relationship is more complex and we shall devote Chapters 4, 5 and 7 to proteins with multiple interacting ligand binding sites.

In view of the widespread use and misuse of the term $\mathbf{X_{1/2}}$, and its equivalence with $\mathbf{IC_{50}}$ in enzymology (Section 8.6), it is important to provide a clear and unequivocal definition. *$\boldsymbol{X_{1/2}}$ is the free ligand concentration required to saturate half the available binding sites under the chosen experimental conditions, provided that chemical equilibrium has been reached.* This definition is consistent with that of $\mathbf{IC_{50}}$ (Cheng and Prusoff, 1973) and stresses the fact that $\mathbf{X_{1/2}}$ is a thermodynamic parameter that may depend on the experimental conditions, but must be independent of the direction in which the equilibrium condition is approached (e.g., whether the experiment was carried out by successive additions of the ligand to the target, as in a titration, or by removing the ligand by dilution or dialysis). Moreover, $\mathbf{X_{1/2}}$ cannot depend on the time required by the mixture to reach its equilibrium state.

A related, but perhaps less intuitive, concept is that of **Xm**, the ligand concentration required to express half the binding free energy change. If the $\bar{X}$ versus log [X] plot is symmetric, as always occurs in monomeric, single-site proteins, $\mathbf{Xm} = \mathbf{X_{1/2}}$ (Wyman 1963; Ackers *et al.,* 1983).

Due to its necessary relationship with K_d, $\mathbf{X_{1/2}}$ is commonly used as an overall empirical parameter to define the apparent affinity of the protein for its ligand. However, $\mathbf{X_{1/2}}$ may depend not only on K_d, but also on the concentration of components of the mixture other than X. If this happens, the type of dependence, whether linear, hyperbolic, or other, is a clue to the type of thermodynamic linkage between X and these components (Sections 1.8 and 1.9).

1.6 Other Representations of the Ligand-Binding Isotherm

Before the widespread use of electronic computers, the analysis of binding isotherms was carried out using graphical methods, and representations aimed at linearizing the [PX] versus [X] hyperbola were widely employed. Unfortunately, the transformations required to linearize the hyperbola entail significant distortion of the magnitude of the

experimental error, and are statistically unsound, thus they should not be used to obtain a quantitative estimate of the thermodynamic parameters (i.e., K_d or $\mathbf{X_{1/2}}$). They are nevertheless discussed here because a thorough understanding of these representations (and their weaknesses) is required to read and understand many classical papers on ligand binding; moreover, these representations may still have some limited usefulness because they may provide visual evidence of some property of the system that must be confirmed using statistically sounder methods.

A classical and widely employed plot was proposed by the Nobel laureate Archibald Vivian Hill to represent his hypothesis on hemoglobin cooperativity. The Hill plot is based on a simple re-elaboration of eqn. 1.6, which yields:

$$[\mathrm{PX}]/[\mathrm{P}] = \bar{X}/(1-\bar{X}) = [X]/\mathrm{K_d}$$

and in logarithmic form:

$$\log\left[\bar{X}/(1-\bar{X})\right] = \log([\mathrm{X}]) - \log(\mathrm{K_d})$$

The Hill plot for a single site macromolecule is a straight line with unitary slope and intercept -log(K_d). Although this representation of experimental data lacks a sound thermodynamic basis and was based on a hypothesis for the function and structure of hemoglobin that was later proven wrong (Bellelli, 2010), it had considerable success due to its simplicity and apparent information content. Indeed, it allows the researcher to measure log (K_d) using a ruler, by drawing a straight line through the experimental points in the range $0.1 < Y < 0.9$ (outside this range the experimental error affects the data very significantly, as shown in Figure 1.3), an important

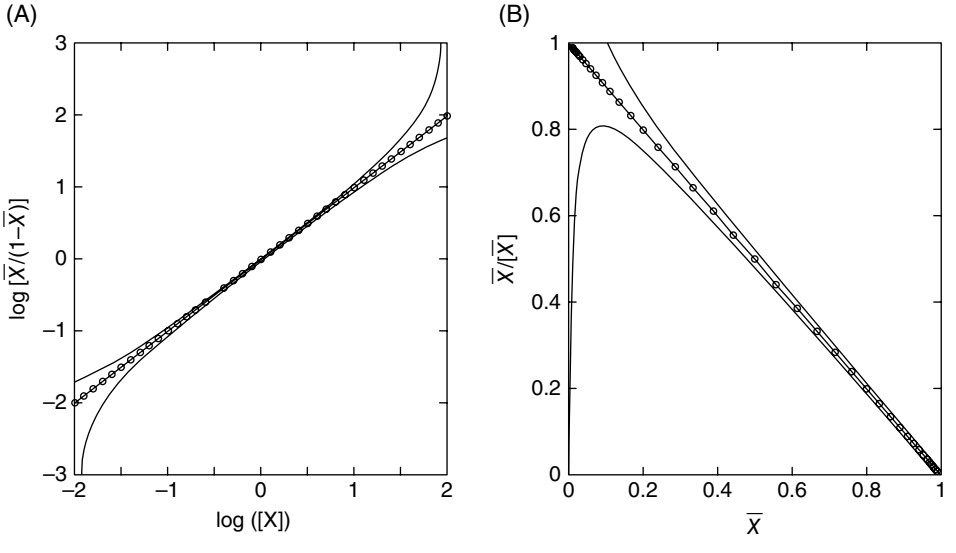


Figure 1.3 The Hill (A) and Scatchard (B) plots.

Ligand concentration is expressed as multiples of K_d (i.e., $[X]/K_d$).

These representations severely distort the experimental errors (continuous lines are calculated to represent an error of 0.01 on $\bar{X}$, i.e., $\bar{X} + \mathbf{0.01}$ and $\bar{X} - \mathbf{0.01}$), and should never be used for quantitative analyses. However, they may be helpful to visualize some qualitative details of the ligand-binding curve (e.g., the presence of heterogeneity or cooperativity), and may maintain some value for figurative purposes.

advantage at a time when personal computers able to implement sounder statistical methods were not available. Moreover, deviations from the unitary slope may indicate either heterogeneity or cooperativity, consistent with Hill's original objective (see Chapter 7).

Like several other graphical representations, the Hill plot severely distorts the experimental errors of the original data set. Thus it can be used, at most, as an instrument to depict a property of the system, but should not be used for quantitative analysis (Forsén and Linse, 1995). Some authors advocated a quantitative use of the Hill plot for the special cases of instruments with demonstrated greater precision (and smaller errors) at very low and very high values of ligand saturations (e.g., Imai *et al.,* 1970), but this suggestion cannot be generalized.

Figure 1.3A reports the Hill plot of the data set of Figure 1.1, together with the lines representing the error range of +/−0.01 independent of $\bar{X}$. The actual amplitude of the error of the experimental points cannot be defined *a priori*: it is a function of the experimental method used, of the value of the variable one measures, and often of $\bar{X}$ itself (e.g., the error on absorbance measurements is usually a small fraction of the actual value recorded by the instrument, which is in turn a function of $\bar{X}$, the protein concentration, and the extinction coefficients of the liganded and unliganded derivatives). Thus, the +/−0.01 confidence interval in Figure 1.3A is only representative of how severely a constant error is distorted in the Hill plot. The figure shows that the central region of the Hill plot, in the range $-1 < \log(\bar{X}/(1-\bar{X})) < 1$ (approximately equivalent to $0.1 < \bar{X} < 0.9$) is relatively free from error distortion, whereas, outside this range, the distortion becomes very pronounced.

The Hill plot has been widely used in the past to empirically evaluate the cooperativity of multimeric proteins (as described in Chapters 4 and 7); however, the advent of personal computers has made it obsolete. Moreover, its slope fails to provide a thermodynamically interpretable quantitative measure of the magnitude of cooperativity (Forsén and Linse, 1995) and remains an empirical parameter essentially devoid of any direct physical meaning. However, since the Hill plot has been widely employed in the past, interpretation of a host of literature data entails an understanding of its parameters (log ($\mathbf{X_{1/2}}$) and the slope, often called the Hill coefficient n).

The Scatchard plot is another type of graphical representation of great historical value that is unsuitable for quantitative data analysis, despite continuing to be in use. It aims to highlight different properties of a binding system than the Hill plot, but shares the same concerns for error distortion. This graphical representation was proposed by G. Scatchard (Scatchard, 1949) and is based on the following linear rearrangement of eqn. 1.6:

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

or:

$$\bar{X}/[X] = 1/K_d - \bar{X}/K_d$$

Thus, a plot of [PX]/[X] versus [PX] (or $\bar{X}/[X]$ versus $\bar{X}$) yields a straight line with slope $-1/K_d$ and two intercepts: with the abscissa at $[P]_{tot}/K_d$ (or $1/K_d$) and with the ordinate at $[P]_{tot}$ (for the case of 1:1 binding) (Figure 1.3B).

The Scatchard plot may seem to have little to offer: (i) it severely distorts the experimental errors; (ii) it uses [PX], which is error-prone, in both the dependent and independent variables; and (iii) it is strongly counterintuitive, as can be seen from a comparison of Figure 1.3B with the same data plotted in Figure 1.1.

The Scatchard plot was used to graphically detect heterogeneity of the protein preparation (i.e., the presence of isoforms or non-equivalent binding sites; see Section 3.10); however, it is obsolete, because the same information can be obtained by sounder statistical methods at the expense of minimal computing power.

1.7 Effect of Temperature: Thermodynamic Relationships

The reaction scheme of eqn. 1.1 neglects any energy contribution to the formation or dissociation of the protein-ligand complex; a more reasonable scheme would be:

$$P + X + Q \Leftrightarrow PX \quad \text{or} \quad P + X \Leftrightarrow PX + Q \qquad \text{(eqn. 1.7)}$$

In which energy (Q, which may indicate the free energy, the enthalpy or the entropy change) appears as a reactant or as a product of the reaction. Many ligand-binding reactions are exothermic, that is, they release heat as a consequence of the formation of the protein ligand bond(s); however, there are also examples of endothermic reactions, in which the amount of heat released upon bond formation is exceeded by the heat absorbed because of changes within the protein or the ligand (eqn. 1.7).

The fundamental relationship between the reaction free energy and the equilibrium constant is:

$$\Delta G^0 = -RT \ln K \qquad \text{(eqn. 1.8)}$$

The $\Delta G^0 \left(= G^0_{\text{products}} - G^0_{\text{reagents}}\right)$ is defined for standard thermodynamic conditions (1 M for every species in solution, 1 atm for gases), and measures the free energy that is absorbed or released during the process in which the system, starting from the standard conditions, reaches its chemical equilibrium state (at constant temperature). Biological systems rarely or never can be equilibrated under standard conditions, thus their free energy is corrected for their actual conditions and is called their $\Delta G'$; the correction is such that a system under equilibrium conditions has $\Delta G' = 0$. $\Delta G'$, if different from zero, defines the direction in which the system will evolve: if $\Delta G' < 0$ then evolution is in the direction of a net increase of products at the expense of the reactants; if $\Delta G' > 0$ then evolution is in the opposite direction. Progression of the reaction changes the concentrations of reactants and products, hence the $\Delta G'$ that becomes zero when equilibrium is reached.

Eqn. 1.8 allows the researcher to calculate the binding energy of the protein-ligand pair, or, to be more precise, the free-energy difference between the PX complex and the dissociated P and X pair. Thus, one might completely replace the equilibrium constant with the binding free energy. Although we do not suggest the reader to take this step, it is useful to keep in mind the order of magnitude of the binding energies commonly encountered in biological systems; thus a K_d of 1 mM corresponds to a free-energy difference of approximately 4 kcal/mole of complex at 25 °C, a K_d of 1 uM to 8 kcal/mole,

and a K_d of 1 nM to 12 kcal/mole. These values have a positive sign in the direction of dissociation, that is, free energy is released when the complex forms and is absorbed when the complex dissociates.

One should resist the simplistic idea of equating ΔG^0 to the protein-ligand bond energy, since this parameter also includes the energy of any ligand-linked structural rearrangement of the protein and the ligand, and any solvation-desolvation processes of the ligand and the binding site.

The equivalence of equilibrium constants and energies clarifies that no chemical reaction can be truly irreversible, as this would imply an infinite ΔG^0 and would violate the first principle of thermodynamics. What we usually call irreversible is a chemical reaction whose equilibrium constant is so large that the concentration of reagents at equilibrium escapes detection by any practical means. For example, the solubility product of mercuric sulfide is so low that the precipitate is in equilibrium with less than one of the constitutive ions in essentially any volume of water one may practically use, the equilibrium concentrations of mercuric and sulfide ions being of the order of 10^{-27} M. In cases like this one, thermodynamic relationships are used instead of direct measurements to infer the equilibrium constants, and we should imagine "concentrations" as the probability of at least one molecule of the reagent to be present in solution at any given time in the reaction mixture.

Protein-ligand interactions rarely approach these affinities, but may still reach levels that may make it difficult or impossible to detect the free ligand. For example, the K_d of the avidin-biotin complex is in the order of 10^{-15} M and that of heme-hemopexin is $<10^{-12}$ M, as is that of the Mb-NO complex. In these and similar cases competition (see below and Chapter 3) or kinetic (Chapter 2) experiments may be required to determine equilibrium constants.

The study of the dependence of the equilibrium constant on the temperature allows the researchers to better quantify the components of its ΔG^0. Indeed, if we combine eqn. 1.8 with the Gibbs' free energy equation:

$$\Delta G = \Delta H - T\,\Delta S$$

we obtain:

$$\ln K = -\left(\Delta H^0/RT\right) + \left(\Delta S^0/R\right) \qquad \text{(eqn. 1.9)}$$

In many simple chemical equilibria ΔH^0 is independent of temperature (at least on small temperature intervals) and the plot ln K versus 1/T yields a straight line, with slope $-\Delta H^0/R$ and intercept $\Delta S^0/R$ (the van t'Hoff plot; Figure 1.4). In such cases one may derive all three fundamental thermodynamic parameters for the reaction, ΔG, ΔH, and ΔS from determinations of the equilibrium constant at different temperatures. When ΔH is not temperature independent, the van t'Hoff plot is curved, and the dependence of ΔH on temperature conveys information about the heat capacity change for the system.

We stress once more that when we say that a function is a straight line whose slope and intercept correspond to certain parameters, we do not imply that the statistically soundest method to determine those parameters is by linear regression: our aim is to offer to the reader the description of a function or a concept in a way that is easy to visualize and remember. Actually, the soundest method to determine the thermodynamic