

A unified framework based on the binding polynomial for characterizing biological systems by isothermal titration calorimetry

Sonia Vega¹, Olga Abian^{1,2,3,4*} and Adrian Velazquez-Campoy^{1,5,6*}

¹ Institute of Biocomputation and Physics of Complex Systems (BIFI), Joint Unit IQFR-CSIC-BIFI, Universidad de Zaragoza, Zaragoza, Spain

² Instituto Aragonés de Ciencias de la Salud (IACS), Zaragoza, Spain

³ IIS Aragón, Zaragoza, Spain

⁴ Centro de Investigación Biomédica en Red en el Área Temática de Enfermedades Hepáticas y Digestivas (CIBERehd), Spain

⁵ Department of Biochemistry and Molecular and Cell Biology, University of Zaragoza, Zaragoza, Spain

⁶ Fundacion ARAID, Government of Aragon, Spain

***Co-corresponding authors:**

Olga Abian (oabifra@unizar.es)

Adrian Velazquez-Campoy (adrianvc@unizar.es)

Short title: Unified framework for ITC in biological systems

Abstract

Isothermal titration calorimetry (ITC) has become the gold-standard technique for studying binding processes due to its high precision and sensitivity, as well as its capability for the simultaneous determination of the association equilibrium constant, the binding enthalpy and the binding stoichiometry. The current widespread use of ITC for biological systems has been facilitated by technical advances and the availability of commercial calorimeters. However, the complexity of data analysis for non-standard models is one of the most significant drawbacks in ITC. Many models for studying macromolecular interactions can be found in the literature, but it looks like each biological system requires specific modeling and data analysis approaches. The aim of this article is to solve this lack of unity and provide a unified methodological framework for studying binding interactions by ITC that can be applied to any experimental system. The apparent complexity of this methodology, based on the binding polynomial, is overcome by its easy generalization to complex systems.

Keywords: isothermal titration calorimetry, binding polynomial, macromolecular interactions, ligand binding, homotropic and heterotropic chemical linkage, allosteric and polymeric linkage

Introduction

Isothermal titration calorimetry (ITC) has become a standard laboratory technique worldwide. ITC allows the simultaneous determination of the association equilibrium constant, the binding enthalpy and the binding stoichiometry in a binding process where non-covalent interactions are established in solution. Regarding its applicability, researchers employ ITC for characterizing biomacromolecular interactions involving proteins, nucleic acids, carbohydrates, lipids, low-molecular-weight compounds, and even cellular organelles and whole cells. The practical use of ITC ranges from the study of the interaction between two natural binding partners or the elucidation of binding cooperativity phenomena underlying allosteric regulation, to the optimization of lead compounds in drug discovery and development.

Just a few decades ago calorimetry was considered a marginal technique in the applied chemistry field. Nowadays, the widespread use of ITC for biological systems, facilitated by technical developments (e.g. increase in sensitivity, miniaturization and automation) and the availability of user-friendly commercial calorimeters, has led to its incorporation in many biophysical/biochemical labs, as well as its consideration as the gold-standard for molecular interaction studies.

Among the advantages in ITC, these can be highlighted:

- Complete thermodynamic characterization in a single experiment
- Broad practical range for binding affinity determination
- No need for molecular labeling
- No need for molecular immobilization

At the same time, among the disadvantages in ITC, these can be highlighted:

- Time and sample consumption
- Inability to reliably determine binding affinities lower than nanomolar (and higher than millimolar) applying a simple direct set-up
- Complexity of data analysis for non-standard models

The last point is even applicable to standard models (e.g. protein with a single binding site), because of many ITC users lacking the appropriate experimental training and knowledge about fundamental concepts associated with binding and, in particular, calorimetry. Thus, it is not uncommon to observe misinterpretations and systematic errors in published work. In addition, commercial software packages for ITC analysis contain only a few binding models and, consequently, users must either develop their own fitting routines for specific binding models or rely on the capability to apply standard models to more complex systems with the appropriate assumptions (under certain circumstances a complex model can be reasonably substituted by a simple, approximate model).

Many models for studying macromolecular interactions can be found in the literature: protein with a single binding site, protein with several independent binding sites, protein with several binding sites exhibiting cooperative interactions (chemical homotropic and heterotropic linkage), oligomeric protein dissociation, protein self-association/dissociation coupled to ligand binding (polysteric linkage). Reviewing a representative set of these publications it is easy to conclude that each experimental scenario requires its own specific hypotheses, strategies, and approximations, leading to a rather heterogeneous set of approaches lacking unity.

The aim of this article is to provide a unified mathematical formalism for studying binding interactions by ITC that can be applied to any experimental system. This methodology is based on the binding polynomial, the partition function of the system containing all relevant information [1-3]. As it occurs with many general methodologies, it may seem artificially complex for simple systems (i.e. like "using a sledgehammer to crack a nut"), but the easy and straightforward generalization to complex systems represents a remarkable benefit. Providing a unified framework does not preclude the user for acquiring proficient training and knowledge about experimental and theoretical aspects of binding interactions, as well as data analysis. The methodology and tools described here are directed to the appropriate description of the binding model, its implementation and its connection with the real system regarding the estimation of the thermodynamic binding parameters. The strategies for

obtaining additional information (e.g. performing experiments at different temperatures, ionic strengths, pH's, co-solutes, etc. in order to estimate the binding heat capacity, the net number of protons/ions/molecules exchanged upon binding, etc., or adopting different experimental approaches), as well as technical descriptions and protocols for ITC experiments, can be readily found elsewhere. Specific details about the data pre-analysis (e.g. baseline correction, integration of peaks, etc.) are expected to be already known to the ITC user.

Before advancing to the following sections two important observations must be made. First, the set of binding equations does not imply any restrictive assumption (at least, other than dilute solutions are considered and, therefore, activities can be replaced by concentrations; alternatively, the activity coefficients can be considered included in the equilibrium constants) and it is valid for any technique suitable for binding studies (e.g. circular dichroism, fluorescence spectroscopy, nuclear magnetic resonance). The basic difference consists of the final observable employed in each technique (e.g. heat, ellipticity, fluorescence intensity, chemical shift). Second, ITC is not a differential technique, but a finite-difference technique. Thus, finite changes in the excess molar enthalpy of the system (heat evolved at constant pressure) triggered by finite changes in the composition of the system are monitored. Therefore, cumulative data must not be employed (minimizing systematic errors) and differential equations should not be employed for modeling the system. Differential equations-based approaches are fine for academic purposes, but they really do not have added value compared to finite-differences-based approaches and, furthermore, do not resemble the actual scenario.

The binding polynomial

For a biological macromolecule, P , with several binding sites for a given ligand, L , different liganded states can be considered: $\{PL_i, i = 0, \dots, n\}$. This ensemble of states is in equilibrium and overall association constants, β_i , or step-wise association constants, K_i , can be defined:

$$\begin{aligned}\beta_i &= \frac{[PL_i]}{[P][L]^i} \\ K_i &= \frac{[PL_i]}{[PL_{i-1}][L]}\end{aligned}\tag{1}$$

The meaning of these equilibrium constants is simple and rather ambiguous at the same time, since they are not associated with specific binding sites, but with subsets of macromolecule-ligand complexes. The constant β_i reflects the binding of i ligand molecules to any i binding sites of the n possible binding sites (no matter their binding affinity and the specific location of the binding sites) to form the PL_i complex (that is, it is associated to a maximum number of $\binom{n}{i}$ liganded states), whereas the constant K_i reflects the binding of the i th ligand to the PL_{i-1} complex to form the PL_i complex (no matter the binding affinity and the specific location of the binding site).

These two sets of equilibrium constants are equivalent and are interrelated:

$$\begin{aligned}\beta_i &= \prod_{j=1}^{j=i} K_j \\ K_i &= \frac{\beta_i}{\beta_{i-1}}\end{aligned}\tag{2}$$

From now on, and for the sake of simplicity, we will only employ the overall association constants. These association constants constitute a set of macroscopic phenomenological model-free parameters governing the binding equilibrium (i.e. statistical distribution of the macromolecule into the different liganded states), and no mechanistic interpretation should be inferred from them. However, based on these association constants, appropriate binding models based on site-specific microscopic association constants and, if necessary, cooperativity constants, can be discriminated.

The partition function of the system, Z , is defined as:

$$Z = \sum_{i=0}^n \exp(-\Delta\hat{G}_i/RT)\tag{3}$$

where R is the ideal gas constant, T is the absolute temperature and $\Delta\hat{G}_i$ is the Gibbs energy difference between the liganded state with i ligands bound and a reference

state. The ligand-free macromolecule, the state with higher energy (since ligand binding always lowers the Gibbs energy of the system), is taken as the reference state, and each exponential term is equal to the relative population ratio of the corresponding liganded state:

$$\exp(-\Delta\hat{G}_i/RT) = \frac{[PL_i]}{[P]} = \beta_i[L]^i \quad (4)$$

The Gibbs energy $\Delta\hat{G}_i$ does not coincide with the usual definition for the Gibbs energy $\Delta G_i = -RT\ln\beta_i$, but it contains an additional term $-iRT\ln[L]$ accounting for the additional stabilization effect (i.e. reduction in Gibbs energy) of entropic nature induced by the presence of the ligand.

Therefore, partition function can be expressed in terms of the relative population ratios, or the association constants and the concentration of free ligand:

$$Z = \sum_{i=0}^n \frac{[PL_i]}{[P]} = \sum_{i=0}^n \beta_i[L]^i = \mathcal{P}_n([L], \{\beta_i\}) \quad (5)$$

Thus, the partition function is an n -degree polynomial in $[L]$, with positive coefficients equal to the overall association constants. Although it has not been indicated in the previous expression, the dependence on the environmental variables (pressure, pH, ionic strength, co-solute concentrations,...) is implicitly assumed. The previous equation provides an operational method for defining the binding polynomial for any biological system: sum the concentrations of all possible liganded states for the macromolecule and divide by the concentration of ligand-free macromolecule. The molar fraction of each liganded state, χ_i , is given by:

$$\chi_i = \frac{[PL_i]}{[P]_T} = \frac{\beta_i[L]^i}{Z} \quad (6)$$

where $[P]_T$ is the total concentration of macromolecule. The molar fraction of ligand-free macromolecule is equal to $1/Z$.

The partition function is also called the binding polynomial because it contains all information related to the binding in the system; in particular, it defines the average number of ligand molecules bound per macromolecule, $n_{LB} = [L]_{bound}/[P]_T$ (which is the main parameter in a binding experiment), the average excess molar enthalpy of the system, $\langle\Delta H\rangle$, and the average excess molar Gibbs energy of the system, $\langle\Delta G\rangle$ [2-4]:

$$\begin{aligned}
n_{LB} &= \langle i \rangle = \left(\frac{\partial \ln Z}{\partial \ln [L]} \right)_{T,p,\dots} = \frac{\sum_{i=0}^n i \beta_i [L]^i}{\sum_{i=0}^n \beta_i [L]^i} = \sum_{i=0}^n i \chi_i \\
\langle \Delta H \rangle &= RT^2 \left(\frac{\partial \ln Z}{\partial T} \right)_{[L],p,\dots} = -T^2 \left(\frac{\partial \left(\frac{\Delta G}{T} \right)}{\partial T} \right)_{[L],p,\dots} = \frac{\sum_{i=0}^n \Delta H_i \beta_i [L]^i}{\sum_{i=0}^n \beta_i [L]^i} = \sum_{i=0}^n \Delta H_i \chi_i \quad (7) \\
\langle \Delta G \rangle &= -RT \ln Z = -RT \ln \left(\sum_{i=0}^n \beta_i [L]^i \right)
\end{aligned}$$

where ΔH_i is the overall enthalpy associated with the formation of complex PL_i . This set of equations constitutes the foundation for the binding equations to be employed in any binding model.

From the previous equations the following relationships can be derived:

$$\begin{aligned}
\left(\frac{\partial \ln \chi_i}{\partial \ln [L]} \right)_{T,p,\dots} &= i - n_{LB} \\
\left(\frac{\partial \ln \chi_i}{\partial T} \right)_{[L],p,\dots} &= \frac{1}{RT^2} (\Delta H_i - \langle \Delta H \rangle) \quad (8)
\end{aligned}$$

indicating that the molar fraction of complex PL_i reaches a maximum value, regarding the concentration of free ligand and the temperature, when the average ligation number and the average excess molar enthalpy of the system coincide with the particular ligation number and excess enthalpy of the complex PL_i , respectively. In addition, the homotropic derivative of n_{LB} defines the binding capacity [3,5,6]:

$$\left(\frac{\partial n_{LB}}{\partial \ln [L]} \right)_{T,p,\dots} = \langle i^2 \rangle - \langle i \rangle^2 \quad (9)$$

which is a measure of the extent of the binding cooperativity and the capability of the macromolecule for storing and delivering chemical energy (ligand molecules) with the environment upon changes in the chemical potential of the ligand (similar to the heat capacity, the capability for storing and delivering thermal energy with the environment). The binding capacity is related to the Hill coefficient n_H (slope of the Hill plot), another index for quantifying binding cooperativity [3,6]:

$$n_H = \left(\frac{\partial \ln \left(\frac{n_{LB}}{n - n_{LB}} \right)}{\partial \ln [L]} \right)_{T,p,\dots} = \frac{\left(\frac{\partial n_{LB}}{\partial \ln [L]} \right)_{T,p,\dots}}{n_{LB} \left(1 - \frac{n_{LB}}{n} \right)} \quad (10)$$

where the denominator is the binding capacity for a macromolecule with n identical and independent binding sites.

The factorability is an important property of the binding polynomial. If the binding polynomial can be factorized into N terms:

$$Z = \sum_{i=0}^n \beta_i [L]^i = \prod_{g=1}^N (\sum_{i=0}^{n_g} \beta_{g,i} [L]^i) = \prod_{g=1}^N Z_g \quad (11)$$

where the n binding sites are distributed into N groups of n_g binding sites ($\sum_{g=1}^N n_g = n$), then, the system can be considered as a set of N independent subsystems: a set of N independent subsets of binding sites (i.e. binding sites belonging to different subsets behave independently, but binding sites belonging to the same subset behave with (positive) cooperativity). In addition, if the binding sites within a given group can be considered independent, Z_g can be further factorized:

$$Z = \prod_{g=1}^N (\prod_{i=0}^{n_g} (1 + k_{g,i} [L])) \quad (12)$$

and, if all binding sites within a group are identical:

$$Z = \prod_{g=1}^N (1 + k_g [L])^{n_g} \quad (13)$$

Factorability of the binding polynomial is possible if the following relationship is fulfilled for every i :

$$\rho_i = \frac{\left(\frac{\beta_i}{\binom{n}{i}}\right)^{1/i}}{\left(\frac{\beta_{i-1}}{\binom{n}{i-1}}\right)^{1/i-1}} \leq 1 \quad (14)$$

If all $\rho_i = 1$ the binding sites are identical and independent; if all $\rho_i \leq 1$ the binding sites are either non-identical and independent or exhibit negative cooperativity (negative homotropy); and if some $\rho_i > 1$ some of the binding sites exhibit positive cooperativity (positive homotropy) [3,7]. Figure 1 illustrates different scenarios for a macromolecule with three binding sites: identical and independent binding sites, and binding sites with negative and positive homotropic interactions.

Extensions of the binding polynomial

The binding polynomial can be generalized for a macromolecule capable of binding two ligands, A with n_A sites and B with n_B sites [3]:

$$Z = \sum_{t=0}^{n_B} \sum_{s=0}^{n_A} \frac{[PA_s B_t]}{[P]} = \sum_{t=0}^{n_B} \sum_{s=0}^{n_A} \beta_{st} [A]^s [B]^t \quad (15)$$

where β_{st} is the overall association constant for the complex $PA_s B_t$. The constant β_{st} can be split into two factors: $\beta_{st} = \beta_{s0} \beta_{t/s}$, where β_{s0} is the overall association constant for $P + A_s \leftrightarrow PA_s$, and $\beta_{t/s}$ is the overall association constant for $PA_s + B_t \leftrightarrow PA_s B_t$. Thus, from the point of view of ligand A , the binding polynomial can be written as:

$$Z = \sum_{s=0}^{n_A} \beta_{s0} [A]^s \sum_{t=0}^{n_B} \beta_{t/s} [B]^t = \sum_{s=0}^{n_A} \beta_{s0} Z_{s,B} [A]^s = \sum_{s=0}^{n_A} \beta_s^{app} [A]^s \quad (16)$$

where $Z_{s,B}$ is the binding polynomial for B restricted to the subset of PA_s complexes.

From the binding polynomial average quantities can be calculated:

$$\begin{aligned} n_{AB} &= \left(\frac{\partial \ln Z}{\partial \ln [A]} \right)_{T,p,[B],\dots} \\ n_{BB} &= \left(\frac{\partial \ln Z}{\partial \ln [B]} \right)_{T,p,[A],\dots} \\ \langle \Delta H \rangle &= RT^2 \left(\frac{\partial \ln Z}{\partial T} \right)_{p,[A],[B],\dots} \\ \langle \Delta G \rangle &= -RT \ln Z \end{aligned} \quad (17)$$

And, by virtue of Maxwell relations and Legendre transformations, chemical heterotropic linkage relationships between ligands A and B can be obtained [3].

The binding polynomial can also be generalized for a macromolecule exhibiting different conformations with different ligand binding affinities, giving rise to the general allosteric model [3]. In a broad sense, allostery can be considered as the modulation of conformational equilibrium by ligand binding. Different conformations of a macromolecule interact with a given ligand with different binding affinities, and, thus, in the presence of that ligand the conformational equilibrium is redistributed towards those conformational states able to bind that ligand. In addition, the different conformational states possess different biological activities (i.e. different ability to interact with other biological partners). Therefore, the interaction of a macromolecule with a ligand will influence the ability of that macromolecule to interact with other ligands. According to the general allosteric model the binding polynomial for a protein with $m+1$ conformations and n ligand binding sites is given by:

$$Z = \sum_{s=0}^m \sum_{i=0}^n \frac{[P_s L_i]}{[P]} = \sum_{s=0}^m \sum_{i=0}^n \frac{\beta_{si} [P_s] [L]^i}{[P]} = \sum_{s=0}^m \sum_{i=0}^n \beta_{si} \gamma_s [L]^i \quad (18)$$

where β_{si} is the overall association constant for the complex $P_s L_i$ and γ_s is the conformational equilibrium constant for the P_s conformation ($\gamma_s = \frac{[P_s]}{[P]}$). Regrouping

terms the binding polynomial can be written as:

$$\begin{aligned} Z &= \sum_{s=0}^m \gamma_s \sum_{i=0}^n \beta_{si} [L]^i = \sum_{s=0}^m \gamma_s Z_s \\ Z &= \sum_{i=0}^n \left(\sum_{s=0}^m \beta_{si} \gamma_s \right) [L]^i = \sum_{i=0}^n \beta_i^{app} [L]^i \end{aligned} \quad (19)$$

where Z_s is the binding polynomial restricted to the P_s conformation.

Special cases can be readily derived from the general allosteric model by imposing certain constraints [8]. For example, for a homo-oligomeric macromolecule with only two possible conformations, R and T , and n independent ligand binding sites, the binding polynomial is given by:

$$Z = (1 + k_R[L])^n + \gamma(1 + k_T[L])^n = \sum_{i=0}^n \binom{n}{i} (k_R^i + \gamma k_T^i) [L]^i = \sum_{i=0}^n \beta_i^{app} [L]^i \quad (20)$$

where $\gamma = [T]/[R]$, k_R and k_T are the site-specific microscopic association constants for the subunits in R and T conformation, respectively ($k_R > k_T$). This is the Monod-Wyman-Changeux (MWC) model, also known as “concerted model” for positive cooperativity, because the macromolecule consists of n subunits undergoing a simultaneous conformational change, with no mixed conformational species, modulated by ligand binding [9]. Interestingly, conformations R and T are not intrinsically cooperative regarding ligand binding, but ligand binding drives a cooperative transition between the two conformations. In particular, if there is just one ligand binding site (i.e. no cooperativity):

$$Z = 1 + k_R[L] + \gamma(1 + k_T[L]) = (1 + \gamma) + (k_R + \gamma k_T)[L] \quad (21)$$

the macromolecule exhibits a conformational equilibrium between two states with different ligand binding affinities.

If the macromolecule consists of n subunits undergoing individual, sequential conformational change between two conformations, R and T (with $k_T = 0$), upon ligand binding (with mixed conformational species, and $s = i$), the binding polynomial is given by:

$$Z = \sum_{i=0}^n \beta_{ii} \gamma_i [L]^i = \sum_{i=0}^n \beta_i^{app} [L]^i \quad (22)$$

where β_{ii} is the overall association constant for i ligand molecules to the macromolecule with i subunits in R conformation. Each term $\beta_{ii} \gamma_i$ contains the intrinsic association constant and several factors accounting for binding cooperative interactions and conformational changes. This is the Pauling-Koshland-Némethy-Filmer (PKNF) model, also known as “sequential model”, which, contrary to the MWC model, can describe negative cooperativity [10]. A fundamental difference between the PKNF and MWC models is that in the former the binding of ligand triggers the subunit

conformational change (induced fit) and the macromolecule shows intrinsic cooperativity regarding ligand binding, whereas in the latter the ligand binding shifts the pre-existing conformational equilibrium (conformational selection) and each conformation shows no intrinsic ligand binding cooperativity.

Finally, the binding polynomial can also be generalized for a self-associating macromolecule [3]. The binding polynomial is also a polynomial in $[L]$, but the positive coefficients are explicitly dependent on the macromolecule concentration $[P]$ and the self-association constants β_{sm} :

$$Z = \sum_{m=1}^{m_s} \sum_{i=0}^{n_m} \frac{m[P_m L_i]}{[P]} = \sum_{m=1}^{m_s} \sum_{i=0}^{n_m} m\beta_{sm}\beta_{mi}[P]^{m-1}[L]^i \quad (23)$$

where m is the order of the self-association state of the macromolecule (it is included as a normalizing factor to quantify the macromolecule concentration in a monomer basis), m_s is its maximal self-association order, β_{sm} is the m -order self-association constant, and β_{mi} and n_m are the overall association constant for the binding of i ligands to and the number of binding sites in the m -order self-association state of the macromolecule, respectively. Then, the binding polynomial can be expressed in terms of the partial binding polynomials Z_m for each m -order self-association state of the macromolecule:

$$Z = \sum_{m=1}^{m_s} m\beta_{sm}[P]^{m-1}(\sum_{i=0}^{n_m} \beta_{mi}[L]^i) = \sum_{m=1}^{m_s} m\beta_{sm}[P]^{m-1}Z_m \quad (24)$$

and from that binding polynomial average quantities can be calculated (Eq. 7).

For an m -order self-associating macromolecule not all low-order self-association states are significantly populated under certain conditions. For example, a 4-subunit macromolecule may exhibit a monomer-dimer-tetramer equilibrium, or a monomer-tetramer equilibrium. Then, the binding polynomial can be considerably simplified.

Binding equations

The set of binding equations is obtained by combining the chemical equilibrium equations and the mass conservation equations for both the macromolecule and the ligand:

$$\begin{aligned} [P]_T &= [P]Z \\ [L]_T &= [L] + [P]_T \frac{\partial \ln Z}{\partial \ln [L]} \end{aligned} \quad (25)$$

where, for the sake of clarity the notation of the partial derivative has been simplified, or alternatively:

$$\begin{aligned} [P]_T &= [P]Z \\ [L]_T &= [L] + [P][L] \frac{\partial Z}{\partial [L]} \end{aligned} \quad (26)$$

In the first form (method 1) an n+1-degree polynomial equation in [L] is obtained:

$$[L] + [P]_T \frac{\partial \ln Z}{\partial \ln [L]} - [L]_T = 0 \quad (27)$$

whereas in the second form (method 2) a set of non-linear equations in [P] and [L] is obtained:

$$\begin{aligned} [P]Z - [P]_T &= 0 \\ [L] + [P][L] \frac{\partial Z}{\partial [L]} - [L]_T &= 0 \end{aligned} \quad (28)$$

These sets of equations can be solved analytically for simple systems (e.g. one or two classes of independent binding sites) or numerically for any system (e.g. using the Newton-Raphson or the bisection method) [11]. Thus, solving for [P] and [L] for given values of $[P]_T$, $[L]_T$ and $\{\beta_i\}$, the concentration of each complex can be calculated (see Eq. 6):

$$[PL_i] = [P]_T \chi_i = [P]_T \frac{\beta_i [L]^i}{Z} = \beta_i [P][L]^i \quad (29)$$

Although both approaches are equivalent, method 2 is more easily generalized (in particular, for polymeric systems, where ligand binding effects are linked to association/dissociation of a macromolecule, and systems with more than one ligand).

Solving equilibrium equations

Because analytical methods have limited use for solving polynomial equations, numerical methods are preferred. Newton-Raphson method provides an easily programmable and generalizable algorithm with reasonably fast convergence for solving the binding equations. For method 1, a recursive relationship is defined:

$$\begin{aligned} [L] + [P]_T \frac{\partial \ln Z}{\partial \ln [L]} - [L]_T = 0 &\Rightarrow f([L]) = 0 \\ [L]_{k+1} &= [L]_k - \frac{f([L]_k)}{f'([L]_k)}, \text{ until } |[L]_{k+1} - [L]_k| < \varepsilon \end{aligned} \quad (30)$$

Selecting an appropriate initial value and a convergence limit, [L] can be calculated for given values of $[P]_T$, $[L]_T$ and $\{\beta_i\}$. For method 2, a similar recursive relationship is defined:

$$\left. \begin{aligned} [P]Z - [P]_T &= 0 \\ [L] + [P][L] \frac{\partial Z}{\partial [L]} - [L]_T &= 0 \end{aligned} \right\} \Rightarrow \begin{cases} f_1([P], [L]) = 0 \\ f_2([P], [L]) = 0 \end{cases} \Rightarrow \vec{f}(\vec{x}) = 0 \quad (31)$$

$$\vec{x}_{k+1} = \vec{x}_k - J^{-1}(\vec{x}_k) \cdot \vec{f}(\vec{x}_k), \text{ until } |\vec{x}_{k+1} - \vec{x}_k| < \varepsilon$$

where J is the Jacobian matrix ($J_{pq} = \frac{\partial f_p(\vec{x})}{\partial x_q}$) associated with the set of equations.

Similarly, selecting appropriate initial values and a convergence limit, $[P]$ and $[L]$ can be calculated for given values of $[P]_T$, $[L]_T$ and $\{\beta_i\}$. And, once $[P]$ and $[L]$ are known, the concentration of complexes PL_i can be calculated.

Application of binding equations to ITC

Up to now, we have dealt with solving a binding equilibrium for given total concentrations of macromolecule and ligand. In an ITC experiment the composition of the calorimetric cell changes as the solution from the syringe is injected sequentially and the heat associated with each injection is measured. As a result, the system is going through a discrete set of equilibrium states of different chemical composition. The direct observable in an ITC experiment is the heat per injection, which reflects the formation/dissociation of macromolecule-ligand complexes elicited by the addition of titrant solution. Therefore, the change in the concentration of each complex PL_i upon injection j must be calculated, in order to estimate the expected heat effect associated. This involves solving the binding equations for each injection (experimental point) along the titration. And, for this calculation, the total concentration of macromolecule and ligand, $[P]_T$ and $[L]_T$, in the calorimetric cell must be calculated. For any chemical species located in the syringe at concentration $[S]_0$ the concentration in the calorimetric cell after injection j is given by:

$$[S]_{T,j} = [S]_0 \left(1 - \prod_{r=1}^j \left(1 - \frac{v_r}{V_0} \right) \right) \quad (32)$$

where v_r is the volume for each injection and V_0 is the cell volume. Similarly, for any chemical species located initially in the cell at concentration $[C]_0$ the concentration in the calorimetric cell after injection j is given by:

$$[C]_{T,j} = [C]_0 \prod_{r=1}^j \left(1 - \frac{v_r}{V_0} \right) \quad (33)$$

These expressions are valid for a displacement-type calorimeter operating at constant cell volume, that is, when injecting a given volume from the syringe, an equivalent volume is expelled from the cell with no simultaneous mixing. There are other

expressions for calculating these concentrations upon other assumptions (e.g. simultaneous mixing of ligand solution while injecting) [12,13], but all of them are equivalent provided that $v_j/V_0 \ll 1$.

Once the concentrations of macromolecule-ligand complexes after each injection, $[PL_i]_j$, are calculated, the heat effect associated with each injection, q_j , can be calculated as a finite-difference:

$$q_j = V_0 \left([P]_{T,j} \langle \Delta H \rangle_j - [P]_{T,j-1} \langle \Delta H \rangle_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) = V_0 \sum_{i=1}^n \Delta H_i \left([PL_i]_j - [PL_i]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) \quad (34)$$

Usually, an additional adjustable parameter q_d accounting for non-zero background injection heat due to titrant dilution/injection is included in the right-hand side of the previous equation. Also, the heat effect is normalized by the amount of titrant injected, Δn_{LT} :

$$Q_j = \frac{q_j}{v_j[S]_0} \quad (35)$$

For very small injection volumes (compared to V_0) Q_j could be considered identical to the derivative dq_j/dn_{LT} . However, Q_j is actually $q_j/\Delta n_{LT}$, a quotient of finite-differences (i.e. an incremental or difference quotient).

Finally, the experimental data (experimental heats per injection) are compared to the calculated values and the thermodynamic binding parameters ($\{\beta_i\}$, $\{\Delta H_i\}$) can be estimated through an iterative non-linear least-squares regression analysis. For this purpose the Levenberg-Marquardt algorithm has been employed (as implemented in Origin 7.0). Other appropriate algorithms (e.g. steepest descent or conjugate gradient search) can be adopted instead. In the following sections different macromolecular systems will be described by applying this methodology.

Summarizing, the sequence of steps for constructing a model for ITC:

1. Enumerate all the macromolecular states and construct the binding polynomial.
2. Calculate the total concentration of macromolecule and ligand(s) in the calorimetric cell, according to the experimental set-up.

3. Solve binding equations and calculate the concentration of all complexes after each injection.
4. Calculate the normalized heat associated with each injection.
5. Estimate the thermodynamic binding parameters through non-linear regression analysis.

Macromolecule with one ligand binding site

This is the simplest case and it will be employed as an example for comprehensively describing all calculation steps. As indicated above, this case can be handled using much simpler and direct approaches, but mastering the general methodology at this level guarantees a successful application to more complex systems.

Evidently, in this system the association constant, β , and the overall binding enthalpy, ΔH , coincide with the step-wise association constant K_i and step-wise association enthalpy ΔH , as well as with the microscopic site-specific association constant k and the binding enthalpy Δh .

A macromolecule with a single ligand binding site is represented by an overall association constant β and an overall binding enthalpy ΔH . The binding polynomial is given by:

$$Z = 1 + \beta[L] \quad (36)$$

from which:

$$n_{LB} = \frac{\beta[L]}{1+\beta[L]} \quad (37)$$

$$\langle \Delta H \rangle = \frac{\beta[L]}{1+\beta[L]} \Delta H$$

and the molar fractions for the ligand-free and ligand-bound macromolecule are:

$$\chi_0 = \frac{1}{1+\beta[L]} \quad (38)$$

$$\chi_1 = \frac{\beta[L]}{1+\beta[L]}$$

According to method 1 the binding equations are expressed as follows:

$$\begin{aligned} [P]_T &= [P](1 + \beta[L]) \\ [L]_T &= [L] + [P]_T \frac{\beta[L]}{1 + \beta[L]} \end{aligned} \quad (39)$$

from which a quadratic equation in $[L]$ is obtained:

$$[L]^2\beta + [L](1 + \beta([P]_T - [L]_T)) - [L]_T = 0 \quad (40)$$

that can be solved numerically by the Newton-Raphson iterative algorithm:

$$[L]_{k+1} = [L]_k - \frac{\beta[L]_k^2 + [L]_k(1 + \beta([P]_T - [L]_T)) - [L]_T}{2\beta[L]_k + 1 + \beta([P]_T - [L]_T)} \quad (41)$$

According to method 2 the binding equations are expressed as follows:

$$\begin{aligned} [P]_T &= [P](1 + \beta[L]) \\ [L]_T &= [L] + \beta[P][L] \end{aligned} \quad (42)$$

from which a set of non-linear equations in $[P]$ and $[L]$ is obtained:

$$\begin{aligned} [P] + \beta[P][L] - [P]_T &= 0 \\ [L] + \beta[P][L] - [L]_T &= 0 \end{aligned} \quad (43)$$

that can be solved numerically by the Newton-Raphson iterative algorithm:

$$\begin{aligned} \begin{pmatrix} [P]_{k+1} \\ [L]_{k+1} \end{pmatrix} &= \\ \begin{pmatrix} [P]_k \\ [L]_k \end{pmatrix} - \frac{1}{1 + \beta([P]_k + [L]_k)} \begin{pmatrix} 1 + \beta[P]_k & -\beta[P]_k \\ -\beta[L]_k & 1 + \beta[L]_k \end{pmatrix} \cdot \begin{pmatrix} [P]_k + \beta[P]_k[L]_k - [P]_T \\ [L]_k + \beta[P]_k[L]_k - [L]_T \end{pmatrix} \end{aligned} \quad (44)$$

The concentration of the macromolecule-ligand complex can be readily calculated from the free macromolecule and ligand concentrations, $[P]$ and $[L]$:

$$[PL] = [P]_T \frac{\beta[L]}{1 + \beta[L]} = \beta[P][L] \quad (45)$$

If these calculations are performed for each injection j , then, the normalized heat associated with injection j along the calorimetric titration is given by:

$$Q_j = \frac{1}{v_j[L]_0} \left(V_0 \Delta H \left([PL]_j - [PL]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) + q_d \right) \quad (46)$$

and the non-linear least squares regression analysis allows estimating the thermodynamic binding parameters. A macromolecule with a single binding site is the most common situation found in the literature. Figure 2 shows two examples of a single ligand binding: EDTA interacting with calcium and NS3 protease from the hepatitis C virus interacting with its inhibitor danoprevir.

These equations can be applied to any experimental set-up: ligand A solution titrated into a macromolecule solution (direct titration), macromolecule solution

titrated into a ligand A solution (reverse titration), and ligand A/macromolecule solution diluted into buffer solution.

Macromolecule with two ligand binding sites

The next level in complexity corresponds to a macromolecule with two binding sites for a ligand. The two binding sites may be identical or non-identical, and independent or cooperative. Although the data analysis to be performed can be based on the microscopic site-specific association constants, k_1 and k_2 , the site-specific binding enthalpies, Δh_1 and Δh_2 , and the cooperativity parameters, α and Δh , it is best to employ the general phenomenological model based on the overall parameters and later select the appropriate specific model depending on their values. A detailed description and different examples can be found elsewhere [14,15].

A macromolecule with two ligand binding sites is represented by two overall association constants, β_1 and β_2 , and two overall binding enthalpies, ΔH_1 and ΔH_2 . The binding polynomial is given by:

$$Z = 1 + \beta_1[L] + \beta_2[L]^2 \quad (47)$$

from which:

$$n_{LB} = \frac{\beta_1[L] + 2\beta_2[L]^2}{1 + \beta_1[L] + \beta_2[L]^2} \quad (48)$$

$$\langle \Delta H \rangle = \frac{\beta_1[L]\Delta H_1 + \beta_2[L]^2\Delta H_2}{1 + \beta_1[L] + \beta_2[L]^2}$$

and the molar fractions for the ligand-free macromolecule and the different macromolecule-ligand complexes are:

$$\chi_0 = \frac{1}{1 + \beta_1[L] + \beta_2[L]^2}$$

$$\chi_1 = \frac{\beta_1[L]}{1 + \beta_1[L] + \beta_2[L]^2} \quad (49)$$

$$\chi_2 = \frac{\beta_2[L]^2}{1 + \beta_1[L] + \beta_2[L]^2}$$

According to method 1 the binding equations are expressed as follows:

$$[P]_T = [P](1 + \beta_1[L] + \beta_2[L]^2)$$

$$[L]_T = [L] + [P]_T \frac{\beta_1[L] + 2\beta_2[L]^2}{1 + \beta_1[L] + \beta_2[L]^2} \quad (50)$$

from which a cubic equation in $[L]$ is obtained:

$$[L]^3\beta_2 + [L]^2(\beta_1 + \beta_2(2[P]_T - [L]_T)) + [L](1 + \beta_1([P]_T - [L]_T)) - [L]_T = 0 \quad (51)$$

that can be solved numerically by the Newton-Raphson iterative algorithm:

$$[L]_{k+1} = [L]_k - \frac{f([L]_k)}{f'([L]_k)} \quad (52)$$

According to method 2 the binding equations are expressed as follows:

$$\begin{aligned} [P]_T &= [P](1 + \beta_1[L] + \beta_2[L]^2) \\ [L]_T &= [L] + \beta_1[P][L] + 2\beta_2[P][L]^2 \end{aligned} \quad (53)$$

from which a set of non-linear equations in $[P]$ and $[L]$ is obtained:

$$\begin{aligned} [P] + \beta_1[P][L] + \beta_2[P][L]^2 - [P]_T &= 0 \\ [L] + \beta_1[P][L] + 2\beta_2[P][L]^2 - [L]_T &= 0 \end{aligned} \quad (54)$$

that can be solved numerically by the Newton-Raphson iterative algorithm:

$$\begin{pmatrix} [P]_{k+1} \\ [L]_{k+1} \end{pmatrix} = \begin{pmatrix} [P]_k \\ [L]_k \end{pmatrix} - J^{-1}([P]_k, [L]_k) \cdot \begin{pmatrix} f_1([P]_k, [L]_k) \\ f_2([P]_k, [L]_k) \end{pmatrix} \quad (55)$$

The concentration of the macromolecule-ligand complexes can be readily calculated from the free macromolecule and ligand concentrations, $[P]$ and $[L]$:

$$\begin{aligned} [PL] &= [P]_T \frac{\beta_1[L]}{1 + \beta_1[L] + \beta_2[L]^2} = \beta_1[P][L] \\ [PL_2] &= [P]_T \frac{\beta_2[L]^2}{1 + \beta_1[L] + \beta_2[L]^2} = \beta_2[P][L]^2 \end{aligned} \quad (56)$$

If these calculations are performed for each injection j , then, the normalized heat associated with injection j along the calorimetric titration is given by:

$$Q_j = \frac{1}{v_j[L]_0} \left(V_0 \left(\Delta H_1 \left([PL]_j - [PL]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) + \Delta H_2 \left([PL_2]_j - [PL_2]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) \right) + q_d \right) \quad (57)$$

and the non-linear least squares regression analysis allows estimating the thermodynamic binding parameters. Recent examples of the application of this model are: inhibitor binding to choline kinase $\alpha 1$ [16], inhibitor from horseshoe crab binding to subtilisin [17], and lipid binding to GrIR [18].

Once the overall thermodynamic parameters are determined, the appropriate specific model can be selected according to the value of the following parameter (see Eq. 14):

$$\rho_2 = \frac{4\beta_2}{\beta_1^2} \quad (58)$$

If $\rho_2 \leq 1$ the binding polynomial is factorable. If $\rho_2 = 1$ the binding sites are identical and independent; if $\rho_2 < 1$ the binding sites are either non-identical and independent or exhibit negative cooperativity; and if $\rho_2 > 1$ the binding sites exhibit positive cooperativity. Then, the binding polynomial can be expressed in terms of the site-specific association constants:

$$\begin{aligned} Z &= 1 + 2k[L] + k^2[L]^2 = (1 + k[L])^2 \\ Z &= 1 + (k_1 + k_2)[L] + k_1k_2[L]^2 = (1 + k_1[L])(1 + k_2[L]) \\ Z &= 1 + 2k[L] + \alpha k^2[L]^2 \end{aligned} \quad (59)$$

for identical and independent binding sites, non-identical and independent binding sites, and identical and cooperative binding sites (PKNF model), respectively, where k is the microscopic association constant for each identical site (associated with a microscopic binding enthalpy Δh), k_1 and k_2 are the microscopic association constants for the non-identical sites (associated with a microscopic binding enthalpies Δh_1 and Δh_2), and α ($= \frac{4\beta_2}{\beta_1^2} = \rho_2$, see Eq. 14) is the cooperativity interaction constant associated with a cooperativity binding enthalpy Δh ($= \Delta H_2 - 2\Delta H_1$). The described methodology can now be applied starting from any of these binding polynomials. Alternatively, the transformation relationships for calculating the site-specific binding parameters from the overall binding parameters can be used (see the Discussion section) [15].

Unfortunately, not all cases are well-defined, because, for example, non-identical independent binding sites and identical binding sites with negative cooperativity are mathematically equivalent models and cannot be distinguished without employing extra-thermodynamic information (e.g. kinetic or structural information). In addition, mixed ill-defined cases can be found in the literature (e.g. non-identical binding sites ($\rho_2 < 1$) exhibit positive cooperativity ($\rho_2 > 1$), resulting in an overall $\rho_2 > 1$). Therefore, the description in terms of the overall binding parameters is always valid and well justified, whereas the description in terms of the microscopic site-specific parameters might be sometimes questionable.

The binding polynomial expressed in terms of the step-wise equilibrium association constants deserves special attention (see Eq. 2-5):

$$Z = 1 + K_1[L] + K_1K_2[L]^2 \quad (60)$$

Because of its nested expression, it is usually called “sequential binding model”, very often implying some kind of binding site occupation order. This mechanistic misinterpretation erroneously considers that: 1) K_1 and K_2 are the association constants for the first and second binding sites, respectively; and 2) if K_1 and K_2 are different, the macromolecule presents two non-identical binding sites or the binding sites are cooperative. However, K_1 represents the binding of the first ligand to any of the two binding sites, and K_2 represents the binding of the second ligand to the remaining unoccupied binding site (thus, for stating that “the first ligand binds first”, no special name for the model is needed). Another mistake adding to the incorrect interpretation regarding the occupation order is that the ligand always binds first to the high affinity site. If the binding sites are non-identical and independent, there is always a certain fraction of macromolecule with a single ligand bound to the low affinity binding site; unless the affinity difference is very large, all possible liganded states have finite non-zero population. If the binding sites show negative cooperativity, the ligand always binds first to the high affinity site; on the contrary, if the binding sites show positive cooperativity, the ligand always binds first to the low affinity site. Binding cooperativity occurs if $\rho_2 = 4K_2/K_1$ is larger or smaller than 1 (that is, if $K_2 > K_1/4$ or $K_2 < K_1/4$, considering appropriately the statistical configuration factors involved) (see Eq. 2 and 14). In fact, if $K_1 = K_2$, there is positive cooperativity. We propose to abolish the name “sequential binding model”, since this is just the general binding model expressed in terms of the step-wise association constants and that name is misleading.

Macromolecule with several classes of independent ligand binding sites

A macromolecule with n binding sites is represented by a set of association constants, β_i , and a set of overall binding enthalpies, ΔH_i . The general expression for the binding polynomial leads to consider $2n$ binding parameters (association constants and binding enthalpies). It is difficult to reliably independently determine such a large set of parameters from the titration curve. In this case, error analysis and parameter

correlation analysis are crucial to properly establish uncertainties and dependencies among parameters. In addition, global analysis of sets of experiments at different concentrations or with different set-ups (e.g. direct and reverse titrations) may help in reducing parameter uncertainties and dependencies.

In certain cases, the number of parameters can be reduced. As explained before, if the n binding sites are distributed into N groups of n_g binding sites ($\sum_{g=1}^N n_g = n$) and if the g groups of binding sites are considered independent, the binding polynomial can be factorized into g factors:

$$Z = \prod_{g=1}^N Z_g = \prod_{g=1}^N \left(\sum_{i=0}^{n_g} \beta_{g,i} [L]^i \right) \quad (61)$$

In addition, if the binding sites within a given group can be considered independent, the g factors can be further factorized:

$$Z = \prod_{g=1}^N \left(\prod_{i=0}^{n_g} (1 + k_{g,i} [L]) \right) \quad (62)$$

and, if all binding sites within a group are identical:

$$Z = \prod_{g=1}^N (1 + k_g [L])^{n_g} \quad (63)$$

where the set of binding parameters has been reduced to $2g$ (g site-specific microscopic association constants, k_g , and g site-specific microscopic binding enthalpies, Δh_g). The described methodology can now be easily applied starting from any of the previous binding polynomial expressions. Examples of a macromolecule with different classes of binding sites or binding sites exhibiting homotropic interactions can be found in the literature [19-23].

Macromolecule with two different cooperative ligands

In many cases the binding of ligand A is regulated by the binding of a second ligand B in an allosteric cooperative fashion. Thus, the macromolecule can bind two different ligands A and B , a special case of Eq. 15-16 with $n_A = 1$ and $n_B = 1$. The binding of ligand B will affect the thermodynamic parameters for ligand A binding. Conversely, the binding of ligand A will affect the thermodynamic parameters for ligand B binding. Because of energy conservation (irrespective of molecular sizes and structural steric and accessibility considerations), the effect of ligand B on ligand A binding affinity (α -fold) is the same as the effect of ligand A on ligand B binding affinity. The binding

cooperativity may be due to short range direct ligand-ligand interaction or long-range indirect interaction through a macromolecule conformational change. A detailed description and analysis of this model can be found in the literature [14,24,25].

If the ligand *A* is titrated into a macromolecule solution with the ligand *B*, then, the binding affinity and the binding enthalpy for ligand *A* are modulated by the binding parameters and concentration of ligand *B* according to a simplified model [24,26]:

$$\beta_A^{app} = \beta_A \frac{1 + \alpha \beta_B [B]}{1 + \beta_B [B]}$$

$$\Delta H_A^{app} = RT^2 \left(\frac{\partial \ln \beta_A^{app}}{\partial T} \right)_{[A],[B],p,\dots} = \Delta H_A - \frac{\beta_B [B]}{1 + \beta_B [B]} \Delta H_B + \frac{\alpha \beta_B [B]}{1 + \alpha \beta_B [B]} (\Delta H_B + \Delta h) \quad (64)$$

Under certain circumstances, the ternary system (*P*, *A*, *B*) can be substituted by a simple approximated model from the point of view of ligand *A* defined by the binding apparent binding parameters and the cooperativity binding parameters (α and Δh) can be calculated from the intrinsic binding parameters for ligands *A* and *B* [27]. Nevertheless, there is no need for using these equations derived from a simplified model, since the exact ternary model can be employed.

A macromolecule capable of binding two different interacting ligands is represented by two association constants, β_A and β_B , and two binding enthalpies, ΔH_A and ΔH_B , plus a cooperativity interaction constant, α , and a cooperativity binding enthalpy, Δh . Positive heterotropic cooperativity corresponds to $\alpha > 1$, whereas negative heterotropic cooperativity corresponds to $\alpha < 1$. If $\alpha = 1$ the ligands behave independently and the data analysis is trivial. Since $n_A = n_B = 1$, there is no practical distinction between overall, step-wise and microscopic association constants and binding enthalpies. The binding polynomial is given by:

$$Z = 1 + \beta_A [A] + \beta_B [B] + \alpha \beta_A \beta_B [A][B] \quad (65)$$

from which:

$$n_{AB} = \frac{\beta_A [A] + \alpha \beta_A \beta_B [A][B]}{1 + \beta_A [A] + \beta_B [B] + \alpha \beta_A \beta_B [A][B]}$$

$$n_{BB} = \frac{\beta_B [B] + \alpha \beta_A \beta_B [A][B]}{1 + \beta_A [A] + \beta_B [B] + \alpha \beta_A \beta_B [A][B]} \quad (66)$$

$$\langle \Delta H \rangle = \frac{\beta_A [A] \Delta H_A + \beta_B [B] \Delta H_B + \alpha \beta_A \beta_B [A][B] (\Delta H_A + \Delta H_B + \Delta h)}{1 + \beta_A [A] + \beta_B [B] + \alpha \beta_A \beta_B [A][B]}$$

and the molar fractions for the ligand-free macromolecule and the different macromolecule-ligand complexes are:

$$\begin{aligned}
\chi_0 &= \frac{1}{1+\beta_A[A]+\beta_B[B]+\alpha\beta_A\beta_B[A][B]} \\
\chi_A &= \frac{\beta_A[A]}{1+\beta_A[A]+\beta_B[B]+\alpha\beta_A\beta_B[A][B]} \\
\chi_B &= \frac{\beta_B[B]}{1+\beta_A[A]+\beta_B[B]+\alpha\beta_A\beta_B[A][B]} \\
\chi_{AB} &= \frac{\alpha\beta_A\beta_B[A][B]}{1+\beta_A[A]+\beta_B[B]+\alpha\beta_A\beta_B[A][B]}
\end{aligned} \tag{67}$$

According to method 1 the binding equations are expressed as follows:

$$\begin{aligned}
[P]_T &= [P](1 + \beta_A[A] + \beta_B[B] + \alpha\beta_A\beta_B[A][B]) \\
[A]_T &= [A] + [P]_T \frac{\beta_A[A]+\alpha\beta_A\beta_B[A][B]}{1+\beta_A[A]+\beta_B[B]+\alpha\beta_A\beta_B[A][B]} \\
[B]_T &= [B] + [P]_T \frac{\beta_B[B]+\alpha\beta_A\beta_B[A][B]}{1+\beta_A[A]+\beta_B[B]+\alpha\beta_A\beta_B[A][B]}
\end{aligned} \tag{68}$$

from which two quadratic equations in $[A]$ and $[B]$, and a quintic equation in $[P]$ are obtained, and they can be solved numerically by the Newton-Raphson iterative algorithm.

According to method 2 the binding equations are expressed as follows:

$$\begin{aligned}
[P]_T &= [P](1 + \beta_A[A] + \beta_B[B] + \alpha\beta_A\beta_B[A][B]) \\
[A]_T &= [A] + \beta_A[P][A] + \alpha\beta_A\beta_B[P][A][B] \\
[B]_T &= [B] + \beta_B[P][B] + \alpha\beta_A\beta_B[P][A][B]
\end{aligned} \tag{69}$$

from which a set of non-linear equations in $[P]$, $[A]$ and $[B]$ is obtained:

$$\begin{aligned}
[P] + \beta_A[P][A] + \beta_B[P][B] + \alpha\beta_A\beta_B[P][A][B] - [P]_T &= 0 \\
[A] + \beta_A[P][A] + \alpha\beta_A\beta_B[P][A][B] - [A]_T &= 0 \\
[B] + \beta_B[P][B] + \alpha\beta_A\beta_B[P][A][B] - [B]_T &= 0
\end{aligned} \tag{70}$$

that can be solved numerically by the Newton-Raphson iterative algorithm:

$$\begin{pmatrix} [P]_{k+1} \\ [A]_{k+1} \\ [B]_{k+1} \end{pmatrix} = \begin{pmatrix} [P]_k \\ [A]_k \\ [B]_k \end{pmatrix} - J^{-1}([P]_k, [A]_k, [B]_k) \cdot \begin{pmatrix} f_1([P]_k, [A]_k, [B]_k) \\ f_2([P]_k, [A]_k, [B]_k) \\ f_3([P]_k, [A]_k, [B]_k) \end{pmatrix} \tag{71}$$

The concentration of macromolecule-ligand complexes can be readily calculated from the free macromolecule and ligands concentrations, $[P]$, $[A]$ and $[B]$:

$$\begin{aligned}
[PA] &= [P]_T \frac{\beta_A[A]}{1+\beta_A[A]+\beta_B[B]+\alpha\beta_A\beta_B[A][B]} = \beta_A[P][A] \\
[PB] &= [P]_T \frac{\beta_B[B]}{1+\beta_A[A]+\beta_B[B]+\alpha\beta_A\beta_B[A][B]} = \beta_B[P][B] \\
[PAB] &= [P]_T \frac{\alpha\beta_A\beta_B[A][B]}{1+\beta_A[A]+\beta_B[B]+\alpha\beta_A\beta_B[A][B]} = \alpha\beta_A\beta_B[P][A][B]
\end{aligned} \tag{72}$$

If these calculations are performed for each injection j , then, the normalized heat associated with injection j along the calorimetric titration is given by:

$$Q_j = \frac{1}{v_j[L]_0} \left(V_0 \left(\Delta H_A \left([PA]_j - [PA]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) + \Delta H_B \left([PB]_j - [PB]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) \right) + (\Delta H_A + \Delta H_B + \Delta h) \left([PAB]_j - [PAB]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) \right) + q_d \quad (73)$$

and the non-linear least squares regression analysis allows estimating the thermodynamic binding parameters. Figure 3 shows an example corresponding to the NS3 protease from the hepatitis C virus, capable of interacting with its substrate and the activating accessory viral NS4A protein.

These equations can be applied to any experimental set-up: ligand *A* solution titrated into a macromolecule/ligand *B* solution, macromolecule solution titrated into a ligand *A*/ligand *B* solution, and ligand *A*/ligand *B* solution titrated into a macromolecule solution. Moreover, complete saturation of macromolecule with ligand *B* when both placed together in the cell is not required. On the other hand, there are no particular requirements for the relative binding affinities or relative binding enthalpies of ligand *A* and ligand *B* (except when $\alpha \ll 1$ and placed in syringe and cell, respectively, as seen in the next section).

Macromolecule with two different competitive ligands

In many cases the binding affinity for a given ligand *A* is too high (too low) to be reliably determined in a standard titration. Although there are several ways to overcome this difficulty (e.g. change temperature or pH), the most appropriate approach consists of performing a titration in the presence of a competing ligand *B* with lower (higher) binding affinity. Thus, competition or displacement titration can be employed to extend the practical window for experimental determination of ligand *A* binding affinity, from very high (higher than nanomolar) to very low values (lower than millimolar) [28,29]. In the first case, the potent ligand *A* is usually titrated into a macromolecule solution with the weaker ligand *B*; then, the binding affinity and the binding enthalpy for ligand *A* become modulated by the binding parameters and concentration of ligand *B* according to a simplified model (taking $\alpha = 0$ and $\Delta h = 0$ in the previous simplified model):

$$\beta_A^{app} = \frac{\beta_A}{1+\beta_B[B]} \quad (74)$$

$$\Delta H_A^{app} = \Delta H_A - \frac{\beta_B[B]}{1+\beta_B[B]} \Delta H_B$$

and, selecting appropriately the ligand B and its concentration, the apparent binding affinity is lowered to experimentally accessible values. The intrinsic binding parameters for ligand A can be calculated from its apparent binding parameters and the intrinsic binding parameters for ligand B . In the second case, ligand B plays the role of potent ligand, while ligand A is the weak ligand. Similar equations hold, except for reversing the roles of both ligands. The intrinsic binding parameters for ligand A can be calculated from the apparent and the intrinsic binding parameters for ligand B . Nevertheless, there is no need for using these equations derived from a simplified model, since the exact ternary model can be employed and applied to any experimental set-up [12,30-32].

Now, a macromolecule can bind two different ligands A ($n_A = 1$) and B ($n_B = 1$). The binding of ligand A is completely dependent on ligand B binding, because they are competitive ligands (i.e. their binding is mutually exclusive). The reciprocal excluding binding effect may be due to direct ligand-ligand steric hindrance or long-range indirect interaction through a macromolecule conformational change between two conformational states, each one capable of binding ligand A or ligand B only. This case is a special limiting case of the previous one for maximal negative cooperativity ($\alpha = 0$ and $\Delta h = 0$).

Therefore, a macromolecule capable of binding two different competitive ligands is represented by two association constants, β_A and β_B , and two binding enthalpies, ΔH_A and ΔH_B . The binding polynomial is given by:

$$Z = 1 + \beta_A[A] + \beta_B[B] \quad (75)$$

from which:

$$n_{AB} = \frac{\beta_A[A]}{1+\beta_A[A]+\beta_B[B]}$$

$$n_{BB} = \frac{\beta_B[B]}{1+\beta_A[A]+\beta_B[B]} \quad (76)$$

$$\langle \Delta H \rangle = \frac{\beta_A[A]\Delta H_A + \beta_B[B]\Delta H_B}{1+\beta_A[A]+\beta_B[B]}$$

and the molar fractions for the ligand-free macromolecule and the different macromolecule-ligand complexes are:

$$\begin{aligned}\chi_0 &= \frac{1}{1+\beta_A[A]+\beta_B[B]} \\ \chi_A &= \frac{\beta_A[A]}{1+\beta_A[A]+\beta_B[B]} \\ \chi_B &= \frac{\beta_B[B]}{1+\beta_A[A]+\beta_B[B]}\end{aligned}\quad (77)$$

According to method 1 the binding equations are expressed as follows:

$$\begin{aligned}[P]_T &= [P](1 + \beta_A[A] + \beta_B[B]) \\ [A]_T &= [A] + [P]_T \frac{\beta_A[A]}{1+\beta_A[A]+\beta_B[B]} \\ [B]_T &= [B] + [P]_T \frac{\beta_B[B]}{1+\beta_A[A]+\beta_B[B]}\end{aligned}\quad (78)$$

from which a cubic equation in $[P]$ is obtained:

$$\begin{aligned}[P]^3\beta_A\beta_B + [P]^2(\beta_A + \beta_B + \beta_A\beta_B([A]_T + [B]_T - [P]_T)) + [P](1 + \beta_A([A]_T - [P]_T) + \beta_B([B]_T - [P]_T)) - [P]_T &= 0\end{aligned}\quad (79)$$

and it can be solved numerically by the Newton-Raphson iterative algorithm:

$$[P]_{k+1} = [P]_k - \frac{f([P]_k)}{f'([P]_k)}\quad (80)$$

and $[A]$ and $[B]$ are calculated from:

$$\begin{aligned}[A] &= [A]_T \frac{1}{1+\beta_A[P]} \\ [B] &= [B]_T \frac{1}{1+\beta_B[P]}\end{aligned}\quad (81)$$

According to method 2 the binding equations are expressed as follows:

$$\begin{aligned}[P]_T &= [P](1 + \beta_A[A] + \beta_B[B]) \\ [A]_T &= [A] + \beta_A[P][A] \\ [B]_T &= [B] + \beta_B[P][B]\end{aligned}\quad (82)$$

from which a set of non-linear equations in $[P]$, $[A]$ and $[B]$ is obtained:

$$\begin{aligned}[P] + \beta_A[P][A] + \beta_B[P][B] - [P]_T &= 0 \\ [A] + \beta_A[P][A] - [A]_T &= 0 \\ [B] + \beta_B[P][B] - [B]_T &= 0\end{aligned}\quad (83)$$

that can be solved numerically by the Newton-Raphson iterative algorithm:

$$\begin{pmatrix} [P]_{k+1} \\ [A]_{k+1} \\ [B]_{k+1} \end{pmatrix} = \begin{pmatrix} [P]_k \\ [A]_k \\ [B]_k \end{pmatrix} - J^{-1}([P]_k, [A]_k, [B]_k) \cdot \begin{pmatrix} f_1([P]_k, [A]_k, [B]_k) \\ f_2([P]_k, [A]_k, [B]_k) \\ f_3([P]_k, [A]_k, [B]_k) \end{pmatrix}\quad (84)$$

The concentration of macromolecule-ligand complexes can be readily calculated from the free macromolecule and ligands concentrations, $[P]$, $[A]$ and $[B]$:

$$\begin{aligned}
[PA] &= [P]_T \frac{\beta_A[A]}{1+\beta_A[A]+\beta_B[B]} = \beta_A[P][A] \\
[PB] &= [P]_T \frac{\beta_B[B]}{1+\beta_A[A]+\beta_B[B]} = \beta_B[P][B]
\end{aligned}
\tag{85}$$

If these calculations are performed for each injection j , then, the normalized heat associated with injection j along the calorimetric titration is given by:

$$Q_j = \frac{1}{v_j[L]_0} \left(V_0 \left(\Delta H_A \left([PA]_j - [PA]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) + \Delta H_B \left([PB]_j - [PB]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) \right) + q_d \right)
\tag{86}$$

and the non-linear least squares regression analysis allows estimating the thermodynamic binding parameters. Figure 4 shows an example corresponding to two competitive ligands (calcium and magnesium ions) binding to the same molecule (EDTA). This approach has been successfully applied to determining the thermodynamic binding parameters of first- and second-generation inhibitors of the HIV-1 protease [33-37].

These equations can be applied to any experimental set-up: ligand A solution titrated into a macromolecule/ligand B solution, macromolecule solution titrated into a ligand A /ligand B solution, and ligand A /ligand B solution titrated into a macromolecule solution [38,39]. Moreover, complete saturation of macromolecule with ligand B when both placed together is not required. If ligand A solution is titrated into a macromolecule/ligand B solution, the binding affinity of ligand A must be higher than that of ligand B ; otherwise, binding of ligand A by displacing ligand B will not occur.

Self-associating macromolecule

The energetics of macromolecular heterodimerization is easily monitored by ITC, just by locating one subunit in the syringe and the other in the cell. However, the energetics of homodimerization cannot be determined using that experimental scheme. A macromolecule solution will always contain dimers, P_2 , and monomers, P , and monomers cannot be isolated (except by lowering the concentration). This problem is worked out by performing the titration of a concentrated macromolecule

solution into a buffer solution. The observed heat effect reflects the dissociation of homodimers into monomers triggered by the dilution upon injection [40,41].

A macromolecule exhibiting monomer-dimer equilibrium is represented by a dimerization constant, $\beta (= [P_2]/[P]^2)$, and a dimerization enthalpy, ΔH . Expressing the concentration of macromolecule on a monomer basis, the binding polynomial is given by:

$$Z = 1 + 2\beta[P] \quad (87)$$

from which:

$$\begin{aligned} n_{PB} &= \frac{2\beta[P]}{1+2\beta[P]} \\ \langle \Delta H \rangle &= \frac{2\beta[P]}{1+2\beta[P]} \Delta H \end{aligned} \quad (88)$$

and the molar fractions for the monomeric and dimeric macromolecule are:

$$\begin{aligned} \chi_1 &= \frac{1}{1+2\beta[P]} \\ \chi_2 &= \frac{2\beta[P]}{1+2\beta[P]} \end{aligned} \quad (89)$$

Because there is only one component there is no distinction between method 1 and 2. The binding equation is expressed as follows:

$$[P]_T = [P](1 + 2\beta[P]) \quad (90)$$

from which a quadratic equation in $[P]$ is obtained:

$$[P]^2 2\beta + [P] - [P]_T = 0 \quad (91)$$

and it can be solved analytically or numerically by the Newton-Raphson iterative algorithm:

$$[P]_{k+1} = [P]_k - \frac{2\beta[P]_k^2 + [P]_k - [P]_T}{4\beta[P]_k + 1} \quad (92)$$

The concentration of dimers can be readily calculated from the concentration of monomers $[P]$:

$$[P_2] = [P]_T \frac{\beta[P]}{1+2\beta[P]} = \beta[P]^2 \quad (93)$$

If these calculations are performed for each injection j , then, the normalized heat associated with injection j along the calorimetric titration is given by:

$$Q_j = \frac{1}{v_j[P]_0} \left(V_0 \left((-\Delta H) \left([P]_j - [P]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) - \chi_{1,0}[P]_0 \frac{v_j}{V_0} \right) + q_d \right) \quad (94)$$

where $[P]_0$ and $\chi_{1,0}$ are the concentration of macromolecule (in a monomer basis) and the molar fraction of macromolecule monomers in the syringe, respectively, and the

dimerization enthalpy is considered per monomer. The quantity $\chi_{1,0}$ can be calculated as follows:

$$\chi_{1,0} = \frac{\sqrt{1+8\beta[P]_0}-1}{4\beta[P]_0} \quad (95)$$

The expression for the heat effect associated with each injection reflects the change (increase) in monomer concentration in the cell elicited by that injection. Because the macromolecule in the syringe is not completely in dimeric form, a certain amount of monomers, not contributing to the heat effect, are injected upon each injection. That is the reason for including the correcting term in the previous equation. Non-linear least squares regression analysis allows estimating the thermodynamic binding parameters.

These equations are also valid for characterizing the energetics of heterodimer formation by performing heterodimer dissociation experiments. And this is not limited to dimers, but higher-order homo-oligomers (P_n) can also be characterized by this procedure [42].

Self-associating macromolecule with a single ligand binding site

As a final example of the versatility of the described methodology, a macromolecule exhibiting a monomer-dimer equilibrium modulated by ligand binding (an example of polysteric linkage) will be considered [3,43]. Because of the conservation energy principle, the coupling effects between self-association and ligand binding are reciprocal, as it will be shown below: if the ligand binds preferentially to the dimer (i.e. higher binding affinity for the dimeric macromolecule), ligand binding promotes macromolecule dimerization (i.e. higher monomer-monomer affinity when bound to ligands); and if the ligand binds preferentially to the monomer (i.e. lower binding affinity for the dimeric macromolecule), ligand binding promotes macromolecule monomerization (i.e. lower monomer-monomer affinity when bound to ligands).

A macromolecule exhibiting monomer-dimer equilibrium with a ligand binding site in each monomer is represented by three association constants (β_{11} , β_{21} and β_{22}) and

three binding enthalpies (ΔH_{11} , ΔH_{21} and ΔH_{22}) corresponding to the complexes PL , P_2L and P_2L_2 , as well as three dimerization constants (β_0 , β_1 and β_2) and three dimerization enthalpies (ΔH_0 , ΔH_1 and ΔH_2) corresponding to the complexes P_2 , P_2L and P_2L_2 :

$$\begin{aligned} [P_2] &= \beta_0 [P]^2 \\ [PL] &= \beta_{11} [P][L] \\ [P_2L] &= \beta_{21} \beta_0 [P]^2 [L] = \beta_{11} \beta_1 [P]^2 [L] \\ [P_2L_2] &= \beta_{22} \beta_0 [P]^2 [L]^2 = \beta_{11}^2 \beta_2 [P]^2 [L]^2 \end{aligned} \quad (96)$$

As it can be seen, all these equilibrium constants (as well as their associated binding enthalpies) are not independent, but certain linkage constraints exist among them. In particular:

$$\begin{aligned} \beta_{21} \beta_0 &= \beta_{11} \beta_1 \\ \beta_{22} \beta_0 &= \beta_{11}^2 \beta_2 \end{aligned} \quad (97)$$

from which the following relationships can be derived:

$$\begin{aligned} \frac{\beta_{21}}{\beta_{11}} &= \frac{\beta_1}{\beta_0} \\ \frac{\beta_{22}}{\beta_{21} \beta_{11}} &= \frac{\beta_2}{\beta_1} \end{aligned} \quad (98)$$

reflecting the reciprocal coupling between the self-association and the ligand binding. For example, the first one states that if the ligand binding affinity is higher for the dimer compared to the monomer ($\beta_{2/1} > \beta_{1/1}$), the strength of the dimer is higher when a ligand is bound ($\beta_1 > \beta_0$), that is, ligand binding promotes dimerization of the macromolecule. Besides the linkage between self-association and ligand binding, the binding of ligand to the dimer may exhibit homotropic (positive or negative) cooperativity if $\alpha = \frac{\beta_{22}}{\beta_{21}^2} \neq 1$ (with $\Delta h = \Delta H_{22} - 2\Delta H_{21}$).

Expressing the concentration of macromolecule on a monomer basis, the binding polynomial is given by:

$$Z = 1 + 2\beta_0 [P] + (\beta_{11} + 2\beta_{21} \beta_0 [P])[L] + 2\beta_{22} \beta_0 [P][L]^2 \quad (99)$$

or expressed in terms of the partial binding polynomials Z_m for each m -order self-association state of the macromolecule ($m_s = 2$, $n_1 = 1$, $n_2 = 2$):

$$Z = (1 + \beta_{11} [L]) + 2\beta_0 [P] (1 + \beta_{21} [L] + \beta_{22} [L]^2) \quad (100)$$

from which:

$$n_{LB} = \frac{(\beta_{11} + 2\beta_{21}\beta_0[P])[L] + 4\beta_{22}\beta_0[P][L]^2}{1 + 2\beta_0[P] + (\beta_{11} + 2\beta_{21}\beta_0[P])[L] + 2\beta_{22}\beta_0[P][L]^2} \quad (101)$$

$$\langle \Delta H \rangle = \frac{2\beta_0[P]\Delta H_0 + \beta_{11}[L]\Delta H_{11} + 2\beta_{21}\beta_0[P][L](\Delta H_{21} + \Delta H_0) + 2\beta_{22}\beta_0[P][L]^2(\Delta H_{22} + \Delta H_0)}{1 + 2\beta_0[P] + (\beta_{11} + 2\beta_{21}\beta_0[P])[L] + 2\beta_{22}\beta_0[P][L]^2}$$

and the molar fractions for the ligand-free macromolecule and the different macromolecule-ligand complexes are:

$$\begin{aligned} \chi_{1/0} &= \frac{1}{1 + 2\beta_0[P] + (\beta_{11} + 2\beta_{21}\beta_0[P])[L] + 2\beta_{22}\beta_0[P][L]^2} \\ \chi_{2/0} &= \frac{2\beta_0[P]}{1 + 2\beta_0[P] + (\beta_{11} + 2\beta_{21}\beta_0[P])[L] + 2\beta_{22}\beta_0[P][L]^2} \\ \chi_{1/1} &= \frac{\beta_{11}[L]}{1 + 2\beta_0[P] + (\beta_{11} + 2\beta_{21}\beta_0[P])[L] + 2\beta_{22}\beta_0[P][L]^2} \\ \chi_{2/1} &= \frac{2\beta_{21}\beta_0[P][L]}{1 + 2\beta_0[P] + (\beta_{11} + 2\beta_{21}\beta_0[P])[L] + 2\beta_{22}\beta_0[P][L]^2} \\ \chi_{2/2} &= \frac{2\beta_{22}\beta_0[P][L]^2}{1 + 2\beta_0[P] + (\beta_{11} + 2\beta_{21}\beta_0[P])[L] + 2\beta_{22}\beta_0[P][L]^2} \end{aligned} \quad (102)$$

As mentioned before, method 1 is not appropriate for models with more than two components or involving self-associating macromolecules. According to method 2 the binding equations are expressed as follows:

$$\begin{aligned} [P]_T &= [P](1 + 2\beta_0[P] + (\beta_{11} + 2\beta_{21}\beta_0[P])[L] + 2\beta_{22}\beta_0[P][L]^2) \\ [L]_T &= [L] + \beta_{11}[P][L] + 2\beta_{21}\beta_0[P]^2[L] + 4\beta_{22}\beta_0[P]^2[L]^2 \end{aligned} \quad (103)$$

from which a set of non-linear equations in $[P]$ and $[L]$ is obtained:

$$\begin{aligned} [P] + 2\beta_0[P]^2 + \beta_{11}[P][L] + 2\beta_{21}\beta_0[P]^2[L] + 2\beta_{22}\beta_0[P]^2[L]^2 - [P]_T &= 0 \\ [L] + \beta_{11}[P][L] + 2\beta_{21}\beta_0[P]^2[L] + 4\beta_{22}\beta_0[P]^2[L]^2 - [L]_T &= 0 \end{aligned} \quad (104)$$

that can be solved numerically by the Newton-Raphson iterative algorithm (see Eq. 55). The concentration of macromolecule-ligand complexes can be readily calculated from the free macromolecule and ligands concentrations, $[P]$ and $[L]$.

If these calculations are performed for each injection j , then, the normalized heat associated with injection j along the calorimetric titration is given by:

$$\begin{aligned} Q_j &= \frac{1}{v_j[L]_0} \left(V_0 \left(\Delta H_0 \left([P_2]_j - [P_2]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) + \Delta H_{11} \left([PL]_j - [PL]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) \right) \right. \\ &\quad \left. + (\Delta H_{21} + \Delta H_0) \left([P_2L]_j - [P_2L]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) + (\Delta H_{22} + \Delta H_0) \left([P_2L_2]_j - \right. \right. \\ &\quad \left. \left. [P_2L_2]_{j-1} \left(1 - \frac{v_j}{V_0} \right) \right) \right) + q_d \end{aligned} \quad (105)$$

where the dimerization enthalpies are considered per dimer, and the non-linear least squares regression analysis allows estimating the thermodynamic binding parameters. This model has been recently employed to characterize the coupling between ligand binding and monomer-dimer equilibrium in β -lactoglobulin [44].

Although a homodimeric macromolecule with a binding site for a ligand on each monomer has been considered, the complexity of the system can be increased: 1) heterodimeric molecule; 2) higher self-association orders (e.g. trimer, tetramer, etc.); and 3) higher number of binding sites per macromolecule subunit.

Discussion

The binding polynomial was introduced several decades ago for describing binding interactions in biological systems. This concept allows handling systems of different complexity levels in a simple and general manner. However, its implementation for ITC data analysis (and other binding techniques) has been considered somewhat difficult and the modeling and the data analysis strategies for different biological systems found in the literature lack uniformity, with many system-specific approaches. The purpose of this article is to provide a unified framework, based on the binding polynomial, revealing the versatility of this concept for the mathematical description of biological systems and the data analysis in ITC.

Being ITC a finite-differences technique (finite increments or changes in concentrations, average enthalpy and saturation fraction), the formalism is defined in terms of finite-differences and incremental quotients. Other approaches defined in terms of differential equations or derivatives do not present any real advantage and, furthermore, do not resemble the real experimental scenario (there are no continuous changes in the variables) [45-47]. Just consider, for example, a given experimental set-up where the number of injections is 4-fold larger and the injection volume is 4-fold smaller. In that case, not only the titration has a different appearance (there are more points along the saturation progress), but the relative errors, as well as the uncertainties in the estimated thermodynamic parameters, are larger. On the contrary, spectroscopic techniques provide a different experimental situation, where an

arbitrary number of titrant additions does not change the appearance of the titration curve (other than increasing the number of experimental points).

The binding equations, arising from the binding polynomial (chemical equilibrium) and the mass conservation principle, lead to a single polynomial equation with a free concentration as unknown (e.g. $[P]$ or $[L]$) or a set of non-linear equations with all free concentrations as unknowns ($[P]$, $[L]$...). These equations can only be solved analytically for simple systems or quasi-simple approximations of complex systems by simple systems under certain conditions. However, they can always be solved using numerical methods. Among the usual numerical methods, Newton-Raphson method provides a general, easily programmable recursive algorithm, no matter the degree of the polynomial equation or the set of non-linear equations. To guarantee fast convergence to the sought solution, appropriate seeding values in the recursive algorithm are required. Because the binding equations must be solved at any injection for calculating the free concentration of reactants after any injection, the free concentrations of reactants after the previous injection are a good choice as seeding values (that is, the free concentrations of reactants after two consecutive injections are fairly close).

A normalization parameter N is usually introduced as a factor for the nominal concentration of macromolecule. This parameter accounts for the percentage of active (binding competent) fraction of macromolecule; due to partial unfolding or misfolding, or even impurities, macromolecule concentration determined by spectrophotometric methods may be affected by a significant uncertainty. If the concentration of ligand (usually a low molecular weight chemical) can be precisely determined, it can be employed as a calibrating ("titrating") solution for the macromolecule.

As it has been shown with many examples, the mathematical framework can be applied to any biological system (e.g. homotropic and heterotropic interactions, allosteric and polysteric systems). Thus, the level of complexity in the biological system is reflected in the number of binding parameters and the strategies required for reliably estimating those binding parameters (e.g. need for binary titrations when studying ternary interactions, or need for a self-dissociation dilution titration when

studying self-association coupled to ligand binding), but it is not reflected in the complexity of the mathematical treatment or in a system-specific approach. For example, homotropic cooperativity is mathematically easier than heterotropic cooperativity (homotropy with $n = 2$ leads to a cubic equation, but heterotropy with $n = 1$ leads to a quintic equation); however, the mathematical descriptions based on the binding polynomial and solving the binding equations applying the Newton-Raphson method are similar in complexity.

In addition, this mathematical framework is independent of the experimental set-up of the titration (e.g. direct or reverse titration, competitive binding with competing ligand in cell or syringe), because the binding equations are the same; the differences are only concerned with the calculation of the total concentrations of both reactants (whether a reactant is initially placed in the syringe or the cell) and the heat normalization.

The data analysis should be first performed based on the phenomenological overall parameters (association constants and binding enthalpies). Then, the appropriate model based on the microscopic site-specific binding parameters can be discriminated and defined. Transformation equations between parameters corresponding to the phenomenological model and the site-specific model can be employed or, alternatively, the data analysis can be performed using the site-specific binding polynomial. The transformation equations relating equilibrium constants from different models can be used together with the van't Hoff equation to obtain the transformation equations for the corresponding binding enthalpies. Thus, given a certain relationship between association constants:

$$\beta = F(\{\kappa_q\}) \quad (106)$$

the relationship between the associated binding enthalpies is given by:

$$\Delta H = RT^2 \left(\frac{\partial \ln \beta}{\partial T} \right)_{p,[P],[L],\dots} = RT^2 \left(\frac{\partial \ln F(\{\kappa_q\})}{\partial T} \right)_{p,[P],[L],\dots} = G(\{\Delta h_q\}, \{\kappa_q\}) \quad (107)$$

For example, the step-wise association constant K_i is equal to β_i / β_{i-1} , the ratio of two consecutive overall association constants; then, the step-wise binding enthalpy associated with K_i is equal to $\Delta H_i - \Delta H_{i-1}$, the difference between two consecutive

overall binding enthalpies. Or, the second-order overall association constant β_2 is equal to αk^2 in the PKNF model for two cooperative binding sites; then, the overall binding enthalpy ΔH_2 is equal to $2\Delta H + \Delta h$.

As in many other techniques, the combined use of a set of titrations (e.g. direct and reverse titrations, titrations performed at different macromolecule concentrations), or even adding more experimental techniques (e.g. spectroscopy, ultracentrifugation, X-ray diffraction, small-angle X-ray scattering) can be very helpful in properly defining the binding model for the biological system and reliably determining the binding parameters [14,22,48].

In certain circumstances there may be two types of binding macromolecules in the cell. In that case, the same formalism can be employed with a binding polynomial for each type of macromolecule, Z_1 and Z_2 , from which:

from which:

$$\begin{aligned} [P_1]_T &= [P_1]Z_1 \\ [P_2]_T &= [P_2]Z_2 \\ [L]_T &= [L] + [P_1]_T \frac{\partial \ln Z_1}{\partial \ln L} + [P_2]_T \frac{\partial \ln Z_2}{\partial \ln L} = [L] + [P_1][L] \frac{\partial Z_1}{\partial L} + [P_2][L] \frac{\partial Z_2}{\partial L} \end{aligned} \quad (108)$$

and the same procedure can be followed finalizing in the expression for the normalized heat associated with injection j along the calorimetric titration:

$$Q_j = \frac{1}{v_j[L]_0} \left(V_0 \left(\Delta H_1 \left([P_1L]_j - [P_1L]_{j-1} \left(1 - \frac{v_j}{v_0} \right) \right) + \Delta H_2 \left([P_2L]_j - [P_2L]_{j-1} \left(1 - \frac{v_j}{v_0} \right) \right) \right) \right) + q_d \quad (109)$$

and the non-linear least squares regression analysis allows estimating the thermodynamic binding parameters. This scheme has been employed for estimating the binding parameters in protein-metal interaction eliminating the need for previous ion removal from the protein, either injecting metal solution into a protein solution with chelator excess, or injecting the chelator into a protein solution with metal excess [49-51].

Complex biological models can be, under certain circumstances, approximated by a quasi-simple model based on apparent binding parameters (as seen above for cooperative and competitive ligands). For example, a macromolecule exhibiting a conformational equilibrium between two conformational states with different affinities (β_1 and β_2) for a single ligand can be described by the following binding polynomial (see Eq. 21):

$$Z = 1 + \beta_1[L] + \gamma(1 + \beta_2[L]) \quad (110)$$

from which:

$$n_{LB} = \frac{\beta_1[L] + \gamma\beta_2[L]}{1 + \beta_1[L] + \gamma(1 + \beta_2[L])} \quad (111)$$

$$\langle \Delta H \rangle = \frac{\beta_1[L]\Delta H_1 + \gamma\Delta H_\gamma + \gamma\beta_2[L](\Delta H_2 + \Delta H_\gamma)}{1 + \beta_1[L] + \gamma(1 + \beta_2[L])}$$

and the description of the system explicitly includes the conformational equilibrium parameters (γ , ΔH_γ). However, apparent binding parameters for the ligand can be obtained from the previous expressions [3,25,52]:

$$\beta^{app} = \frac{\beta_1 + \gamma\beta_2}{1 + \gamma} \quad (112)$$

$$\Delta H^{app} = \frac{\beta_1\Delta H_1 + \gamma\beta_2(\Delta H_2 + \Delta H_\gamma)}{\beta_1 + \gamma\beta_2} - \frac{\gamma}{1 + \gamma} \Delta H_\gamma$$

where the conformational equilibrium parameters are implicitly included in the apparent ligand binding parameters. The binding polynomial can now be written as:

$$Z = 1 + \frac{(\beta_1 + \gamma\beta_2)}{1 + \gamma} [L] = 1 + \beta^{app}[L] \quad (113)$$

This expression of the binding polynomial for the quasi-simple system has been obtained applying a normalization factor equal to the sum of the statistical weights of ligand-free macromolecular species (which is equivalent to taking the ensemble of L -free macromolecular states as a reference). From this:

$$n_{LB} = \frac{\beta^{app}[L]}{1 + \beta^{app}[L]} \quad (114)$$

$$\langle \Delta H \rangle = \frac{\beta^{app}[L]}{1 + \beta^{app}[L]} \Delta H^{app}$$

Thus, if Z can be transformed to the form $Z = 1 + \beta^{app}[L]$ through an appropriate normalization, the system can be substituted by a quasi-simple system for a macromolecule with just a single ligand binding site [25]. The system is simplified at the expense of using more complex binding parameters (in fact, β^{app} and ΔH^{app} are not intrinsic binding parameters, because in general they will be dependent on other

intrinsic binding parameters, conformational parameters, and ligand and macromolecule concentrations). However, even if the model can be simplified and substituted by a quasi-simple model, sometimes that model cannot be employed (e.g. a cooperative ligand B is placed at a subsaturating concentration in the calorimetric cell, or two competitive ligands are injected simultaneously into a macromolecule solution [39]).

As it has been mentioned before, when Z contains more than one L -free macromolecular states, it is usual to normalize Z according to that subset of non-ligated states and express Z in terms of apparent overall association constants. For example, in the case of a macromolecule binding two different ligands A and B , from the point of view of ligand A we can normalize using the subset of A -free macromolecular states (see Eq. 15-16):

$$Z = \frac{\sum_{s=0}^{n_A} \beta_{s0} [A]^s \sum_{t=0}^{n_B} \beta_{t/s} [B]^t}{\sum_{t=0}^{n_B} \beta_{t/0} [B]^t} = \sum_{s=0}^{n_A} \beta_{s0} \frac{Z_{s,B}}{Z_{0,B}} [A]^s = \sum_{s=0}^{n_A} \beta_s^{app} [A]^s \quad (115)$$

where $Z_{s,B}$ is the binding sub-polynomial for ligand B restricted to the sub-ensemble of sA -ligated species. From this:

$$\begin{aligned} \beta_s^{app} &= \beta_{s0} \frac{Z_{s,B}}{Z_{0,B}} \\ \left(\frac{\partial \ln \beta_s^{app}}{\partial \ln [B]} \right)_{T,p,\dots} &= n_{s,BB} - n_{0,BB} = \Delta n_{s,BB} \\ RT^2 \left(\frac{\partial \ln \beta_s^{app}}{\partial T} \right)_{[B],p,\dots} &= \Delta H_s^{app} = \Delta H_{s0} + \langle \Delta H_{s,B} \rangle - \langle \Delta H_{0,B} \rangle \end{aligned} \quad (116)$$

where $n_{s,BB}$ is the average number of ligand B molecules bound in the sub-ensemble of sA -ligated species, $\Delta H_{s,0}$ is the binding enthalpy for the complex PA_s , and $\langle \Delta H_{s,B} \rangle$ is the average binding enthalpy for ligand B within the sub-ensemble of sA -ligated species. These are well-known linkage relationships (e.g. pH dependency of ligand binding parameters due to coupling of proton exchange at ionizable groups) [3,25]. In the case of a macromolecule exhibiting several conformational states we may normalize Z using the L -free macromolecular states (see Eq. 18-19):

$$Z = \frac{\sum_{s=0}^m \sum_{i=0}^n \beta_{si} \gamma_s [L]^i}{\sum_{s=0}^m \gamma_s} = \sum_{i=0}^n \frac{\sum_{s=0}^m \beta_{si} \gamma_s}{\sum_{s=0}^m \gamma_s} [L]^i = \sum_{i=0}^n \beta_i^{app} [L]^i \quad (117)$$

where β_i^{app} is the ensemble-averaged overall association constant. Or, alternatively, we may normalize Z using the subset of macromolecular states with $s = 0$:

$$Z = \frac{\sum_{s=0}^m \sum_{i=0}^n \beta_{si} \gamma_s [L]^i}{\sum_{i=0}^n \beta_{0i} [L]^i} = \sum_{s=0}^m \gamma_s \frac{Z_s}{Z_0} \quad (118)$$

where Z_s is the binding sub-polynomial restricted to the s macromolecular conformation. This arrangement of the partition function is the one employed for studying the conformational equilibrium of a macromolecule coupled to ligand binding (e.g. thermal or chemical stability as a function of ligand concentration) [3,53,54].

Even if the system is not approximated by a quasi-simple system, some assumptions can be made in order to simplify the model and reduce the number of possible macromolecular states (regarding ligand binding equilibrium, as well as macromolecular conformational and self-association equilibria), reducing also the number of parameters to be estimated. For instance, in the previous example we might consider the ligand does not bind to the low binding affinity conformational state of the macromolecule ($\beta_2 = 0$ in Eq. 108-111) [55,56]. In the macromolecule exhibiting a self-association equilibrium coupled to ligand binding we might consider the ligand does not bind to the monomer or to the dimer, by eliminating the corresponding terms in the binding polynomial (Eq. 99). Another example is the model for obligate binding order or conditional binding in which ligand B cannot bind unless ligand A is bound; this is a special case of cooperative ligands (eliminating the third term in the right hand side in Eq. 65) and has been applied to determining the energetics of phosphate ion binding to a protease-inhibitor complex [57].

A typical example is the macromolecule with n identical binding sites exhibiting maximal positive cooperativity ($\beta_i = 0$, for $i = 1, \dots, n-1$). In this case the binding polynomial is given by:

$$Z = 1 + \beta_n [L]^n \quad (119)$$

where all intermediate ligation states are neglected (maximal positive cooperativity), from which:

$$n_{LB} = \frac{n\beta_n [L]^n}{1 + \beta_n [L]^n} \quad (120)$$

$$\langle \Delta H \rangle = \frac{\beta_n [L]^n}{1 + \beta_n [L]^n} \Delta H_n$$

Then, the macromolecule can populate two ligation states, either free or bound to n ligands, and the Hill coefficient is equal to n . Obviously, this kind of approximation is

more legitimate the larger the cooperativity effect. This approach has been applied to the interaction of ecarpholin S, a phospholipase from a snake venom, with two binding sites exhibiting a Hill coefficient $n_H = 1.6$ for its inhibitor suramin [58].

Models even more complex than those described here can be considered under the same formalism. For example, a binding polynomial can be defined for cooperative and non-cooperative binding to linear lattice-like macromolecules with overlapping ligand binding sites, leading to a high degree polynomial that under certain conditions can be truncated to a fifth-degree polynomial [59,60]. An approach based on equivalent binding equations (chemical equilibrium and mass conservation) has been employed for cooperative and non-cooperative ligand binding to DNA or other linear polymers [61-65].

Finally, some words of caution about the selection of the most suitable model. The estimated overall (or step-wise) association constants, together with extra-thermodynamic information (e.g. structural or kinetic data) are instrumental in discriminating the best binding model. Sometimes two models can be equivalent, that is, they give rise to mathematically identical equations, but they differ in the mechanistic description of the binding process and the set (number) of thermodynamic parameters, which is a decisive factor for the capability of independently determining all the associated parameters. For example, the MWC model for a macromolecule with two binding sites is equivalent to the KNF model (except for the fact that the MWC model does not explain negative cooperativity), but the MWC model has three equilibrium constants (two association constants and one conformational constant), while the KNF model has two equilibrium constants (one association constant and one conformational/cooperativity constant). From one titration all parameters in the KNF model can be determined, but not in the MWC model. Thus, in general, sophisticated models can be developed accounting for subtle mechanistic or configurational interaction details, but it may be difficult to estimate many parameters experimentally, in particular if the number of experimental data along the titration does not provide sufficient resolution level. This potential problem can be minimized if sets of titrations performed under different conditions (e.g.

different concentrations of reactants) are globally analyzed. In addition, there are statistical parametric and non-parametric tests for comparing different models and the selection of optimal models from experimental data (e.g. parametric F-test for comparing nested models, non-parametric Akaike's information criterion).

Appendix

In order to show how to implement the described formalism, a comparison with the traditional method is provided for the simplest case: a macromolecule with a single ligand binding site. This system is represented by an association constant β and a binding enthalpy ΔH .

From simple chemical equilibrium considerations, the concentration of complex PL can be calculated from the total concentrations of macromolecule and ligand by solving a quadratic equation:

$$[PL] = \frac{1 + \beta([P]_T + [L]_T) - \sqrt{(1 + \beta([P]_T + [L]_T))^2 - 4\beta^2[P]_T[L]_T}}{2\beta} \quad (A1)$$

Then, the code in Origin for the fitting routine can be written as follows (code 1):

```
a=1+beta*(Ptb+Ltb);
b=4*beta^2*Ptb*Ltb;
PLb=(a-sqrt(a^2-b))/(2*beta);
a=1+beta*(Pta+Lta);
b=4*beta^2*Pta*Lta;
PLa=(a-sqrt(a^2-b))/(2*beta);
NDH=(V0*(dH*(PLa-PLb*Di)))/inju/Lsyr;
```

where P_{tb} and P_{ta} are the total concentrations of macromolecule before and after a given injection, L_{tb} and L_{ta} are the total concentrations of ligand before and after the injection, PL_{b} and PL_{a} are the concentrations of macromolecule-ligand complex before and after a given injection, V_0 is the cell volume, D_i is the dilution factor, in_j is the injection volume, L_{syr} is the concentration of ligand in the syringe and NDH is the normalized heat effect associated with that injection. We may include the parameter n as a factor multiplying $[P]_T$ in Eq. A1 or when previously calculating $[P]_T$.

Applying the general formalism (Eqs. 42-46), the code in Origin for the fitting routine based on the Newton-Raphson algorithm for solving the set of non-linear equations obtained from the binding polynomial can be written as follows (code 2):

```

Lold=Lseed;
Pold=Pseed;
Lnew=Lold+0.00000001;
Pnew=Pold+0.00000001;
for(ii=1;(abs(Lnew-Lold)>tolerance || abs(Pnew-Pold)>tolerance) &&
ii<max_ iterations;ii++) {
    Lold=Lnew;
    Pold=Pnew;
    F1=-1*(Lold+beta*Pold*Lold-Ltb);
    F2=-1*(Pold+beta*Pold*Lold-Ptb);
    J11=1+beta*Pold;
    J12=beta*Lold;
    J21=beta*Pold;
    J22=1+beta*Lold;
    detJ=J11*J22-J12*J21;
    detL=F1*J22-F2*J12;
    detP=-F1*J21+F2*J11;
    Lnew=Lold+detL/detJ;
    Pnew=Pold+detP/detJ;
};
PLb=beta*Pnew*Lnew;
Lold=Lseed;
Pold=Pseed;
Lnew=Lold+0.00000001;
Pnew=Pold+0.00000001;
for(ii=1;(abs(Lnew-Lold)>tolerance || abs(Pnew-Pold)>tolerance) &&
ii<max_ iterations;ii++) {
    Lold=Lnew;
    Pold=Pnew;

```

```

F1=-1*(Lold+beta*Pold*Lold-Lta);
F2=-1*(Pold+beta*Pold*Lold-Pta);
J11=1+beta*Pold;
J12=beta*Lold;
J21=beta*Pold;
J22=1+beta*Lold;
detJ=J11*J22-J12*J21;
detL=F1*J22-F2*J12;
detP=-F1*J21+F2*J11;
Lnew=Lold+detL/detJ;
Pnew=Pold+detP/detJ;
};

```

```
PLa=beta*Pnew*Lnew;
```

```
NDH=(V0*(dH*(PLa-PLb*Di)))/inju/Lsyr;
```

where J_{ij} are the elements of the Jacobian matrix defined from the functions F_i , L_{seed} and P_{seed} are initial values for the free concentration of macromolecule and ligand in the Newton-Raphson algorithm, and the final values for the free concentration of macromolecule and ligand are stored as P_{new} and L_{new} after convergence.

It is obvious that code 1 is much simpler than code 2, and, apparently, the general formalism complicates the mathematics unnecessarily. However, the advantage of the general formalism resides in its easy generalization, that can be fully recognized when handling more complex systems which become intractable by analytical methods, but very easy managed through numerical methods like the one implemented here. In particular, the user must only modify the definitions of F_i and J_{ij} according to the system under study, as well as the calculation of macromolecule-ligand complexes from the free, in the previous code.

Acknowledgements

This work was supported by Spanish Ministerio de Ciencia e Innovación (BFU2010-19451 to AVC, PTA2009-2341-I to SV), Miguel Servet Program from Instituto de Salud Carlos III (CP07/00289 to OA), Fondo de Investigaciones Sanitarias (PI10/00186 to OA),

Diputación General de Aragón (Protein Targets Group B89 to AVC, Digestive Pathology Group B01 to OA).

Figure captions

Figure 1. Molar fractions of the different macromolecular species (P , PL , PL_2 and PL_3) as a function of n_{LB} (Left) and n_{LB} as a function of $\ln[L]$ (Right) for a macromolecule with: three identical and independent binding sites (continuous lines, $\beta_1 = 3 \cdot 10^6 \text{ M}^{-1}$, $\beta_2 = 3 \cdot 10^{12} \text{ M}^{-2}$, $\beta_3 = 10^{18} \text{ M}^{-3}$, $n_H = 1$), three binding sites with negative homotropy (dashed lines, $\beta_1 = 3 \cdot 10^6 \text{ M}^{-1}$, $\beta_2 = 3 \cdot 10^{11} \text{ M}^{-2}$, $\beta_3 = 10^{16} \text{ M}^{-3}$, $n_H = 0.65$), and three binding sites with positive homotropy (grey lines, $\beta_1 = 3 \cdot 10^6 \text{ M}^{-1}$, $\beta_2 = 3 \cdot 10^{13} \text{ M}^{-2}$, $\beta_3 = 10^{20} \text{ M}^{-3}$, $n_H = 1.65$). As it can be seen on the left plot, the maximal value of the molar fraction χ_i is reached when $n_{LB} = i$. Binding cooperativity is reflected in the departure of the molar fractions of the intermediate ligation states from those corresponding to independent ligand binding: negative cooperativity increases the population of intermediate states, while positive cooperativity deplete those intermediate states. According to the values of the overall association constants, in the negative homotropy case $\rho_2 < \rho_3 < 1$, and in the positive homotropy case $\rho_2 > \rho_3 > 1$; therefore, the largest deviations from the independent binding occurs in the formation of the complex PL_2 , which is easily observed in the large differences observed in χ_2 for the three scenarios, compared to the smaller differences observed in χ_3 . The inset shows the binding capacity for the three cases, and the higher and lower binding capacity correspond to the positive and negative homotropy case, respectively.

Figure 2. (Left) Calorimetric titration for EDTA interacting with calcium. Calcium (5 mM) was titrated into EDTA (0.4 mM). The experiment was performed in MES 10 mM, pH 6, at 25°C. Data analysis was performed as explained in the text, providing the following binding parameters: $\beta_A = 2.1 \pm 0.1 \cdot 10^5 \text{ M}^{-1}$, $\Delta H_A = 4.0 \pm 0.1 \text{ kcal/mol}$. (Right) Calorimetric titration for NS3 protease interacting with its inhibitor danoprevir. Danoprevir (300 μM) was titrated into NS3 S139A (inactive) protease (20 μM). The experiment was performed in Tris 10 mM pH 7, CHAPS 1%, DTT 2 mM, DMSO 1.5%, glycerol 25%, at 25°C. Data analysis was performed as explained in the text, providing the following binding parameters: $\beta_A = 7.4 \pm 0.5 \cdot 10^7 \text{ M}^{-1}$, $\Delta H_A = -12.5 \pm 0.2 \text{ kcal/mol}$.

Figure 3. Calorimetric titrations for the NS3 protease-substrate-pNS4A system. Substrate (150 μM) was titrated into NS3 S139A (inactive) protease (10 μM) or into an NS3 protease solution (10 μM) with NS4A peptide (200 μM). Experiments were performed in Tris 10 mM pH 7, CHAPS 1%, DTT 2 mM, glycerol 25%, at 25°C. Data analysis was performed as explained in the text, providing the following parameters: substrate interacting with inactive NS3 protease (closed circles; $\beta_A = 8.1 \pm 0.6 \cdot 10^5 \text{ M}^{-1}$, $\Delta H_A = 3.9 \pm 0.2 \text{ kcal/mol}$), and substrate interacting with the NS4A-bound inactive protease (open circles, cooperative interaction constant $\alpha = 5.9 \pm 0.7$ and cooperative binding enthalpy $\Delta h = -7.9 \pm 0.3 \text{ kcal/mol}$). (Inset) The binding parameters for pNS4A interacting with NS3 protease were previously independently determined under the same conditions: $\beta_B = 8.9 \pm 0.8 \cdot 10^4 \text{ M}^{-1}$, $\Delta H_B = -1.5 \pm 0.2 \text{ kcal/mol}$. If the approximate model with a single binding site is used for the analysis of the substrate interacting with NS4A-bound NS3 protease, the following apparent parameters are obtained for the substrate interacting with the NS3 protease in the presence of pNS4A: $\beta_A^{app} = 4.2 \pm 0.3 \cdot 10^6 \text{ M}^{-1}$, $\Delta H_A^{app} = -3.9 \pm 0.2 \text{ kcal/mol}$. Substrate and activator peptide pNS4A show positive heterotropic cooperativity, with a cooperative Gibbs energy ($\Delta g = -1.1 \text{ kcal/mol}$) partitioned into a favorable enthalpic contribution ($\Delta h = -7.9 \text{ kcal/mol}$) and an unfavorable entropic contribution ($-T\Delta s = 6.8 \text{ kcal/mol}$). With the simplified quasi-simple model (Eqs. 64 and 114) the apparent binding parameters for the substrate are dependent on the concentration of pNS4A employed in the titration, whereas with the exact ternary model the intrinsic cooperative parameters are true constants. In this case, both sets of parameters are in agreement because pNS4A was employed at a saturating concentration.

Figure 4. Calorimetric titrations for the EDTA-calcium-magnesium system. (Left) calcium (480 μM) was titrated into EDTA (30 μM), (Middle) magnesium (500 μM) was titrated into EDTA (30 μM), and (Right) calcium (480 μM) was titrated into EDTA (30 μM) in the presence of magnesium (220 μM). Experiments were performed in Mops 10 mM pH 7, NaCl 150 mM, at 25°C. Data analysis was performed as explained in the text, providing the following parameters: calcium interacting with EDTA, $\beta_A = 1.1 \pm 0.1 \cdot 10^7 \text{ M}^{-1}$, $\Delta H_A = -5.2 \pm 0.2 \text{ kcal/mol}$; and magnesium interacting with EDTA, $\beta_B = 1.5 \pm 0.1 \cdot 10^5 \text{ M}^{-1}$,

$\Delta H_B = 3.1 \pm 0.2$ kcal/mol. The apparent parameters for calcium in the presence of magnesium (220 μM) can be obtained using a simplified quasi-simple model: $\beta_A^{app} = 3.1 \pm 0.2 \cdot 10^5 \text{ M}^{-1}$, $\Delta H_A^{app} = -8.4 \pm 0.2$ kcal/mol (in good agreement with Eqs. 74 and 114).

References

- [1] J. Wyman, Heme proteins, *Adv. Protein Chem.* 4 (1948) 407-531.
- [2] J. Wyman, Linked functions and reciprocal effects in hemoglobin: A second look, *Adv. Protein Chem.* 19 (1964) 223-286.
- [3] J. Wyman, S.J. Gill. *Binding and linkage: Functional chemistry of biological macromolecules*, University Science Books, Mill Valley, 1990.
- [4] J. Schellman, Macromolecular binding. *Biopolymers* 14 (1975) 999-1018.
- [5] E. Di Cera, S.J. Gill, J. Wyman, Binding capacity: Cooperativity and buffering in biopolymers. *Proc. Natl. Acad. Sci. USA* 85 (1988) 449-452.
- [6] J. Schellman, Fluctuation and linkage relations in macromolecular solution, *Biopolymers* 29 (1990) 215-224.
- [7] J. Wyman, P. Philipson, A probabilistic approach to cooperativity of ligand binding by a polyvalent molecule, *Proc. Natl. Acad. Sci. USA* 71 (1974) 3431-3434.
- [8] J. Wyman, On allosteric models, *Curr. Topics Cell. Reg.* 6 (1972) 207-223.
- [9] J. Monod, J. Wyman, J.P. Changeux, On the nature of allosteric transitions: A plausible model, *J. Mol. Biol.* 12 (1965) 88-118.
- [10] D.E. Koshland Jr., G. Nemethy, D. Filmer, Comparison of experimental binding data and theoretical models in proteins containing subunits, *Biochemistry* 5 (1966) 365-385.
- [11] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery. *Numerical Recipes: The Art of Scientific Computing*. Third Edition. Cambridge University Press (2007).
- [12] B.W. Sigurskjold, Exact analysis of competition ligand binding by displacement isothermal titration calorimetry, *Anal. Biochem.* 277 (2000) 260-266.
- [13] ITC data analysis in Origin®, Tutorial Guide (2004), www.microcalorimetry.com
- [14] J.C. Houtman, P.H. Brown, B. Bowden, H. Yamaguchi, E. Apella, L.E. Samelson, P. Schuck, Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: Application to adaptor protein complexes in cell signaling, *Protein Sci.* 16 (2007) 30-42.
- [15] E. Freire, A. Schön, A. Velazquez-Campoy, Isothermal titration calorimetry: General formalism using binding polynomials, *Methods Enzymol.* 455 (2009) 127-155.
- [16] M. Sahun-Roncero, B. Rubio-Ruiz, G. Saladino, A. Conejo-Garcia, A. Espinosa, A. Velazquez-Campoy, F.L. Gervasio, A. Entrena, R. Hurtado-Guerrero, The mechanism of

allosteric coupling in choline kinase $\alpha 1$ revealed by a rationally designed inhibitor, *Angew. Chem. Int. Ed.* 52 (2013) 4582-4586.

[17] R.T. Shenoy, S. Thangamani, A. Velazquez-Campoy, H. Bow, J. Ling-Ding, J. Sivaraman, Structural basis for dual-inhibition mechanism of a non-classical Kazal-type serine protease inhibitor from horseshoe crab in complex with subtilisin, *PLoS ONE* 6 (2011) e18838.

[18] C. Jobichen, A.Z. Fernandis, A. Velazquez-Campoy, K.Y. Leung, Y.K. Mok, M.R. Wenk, J. Sivaraman J, Identification and characterization of the lipid binding property of GrIR, a locus of enterocyte effacement regulator, *Biochem. J.* 420 (2009) 191-199.

[19] B. Gopal, C.P. Swaminathan, S. Bhattacharya, A. Bhattacharya, M.R. Murthy, A. Surolia, Thermodynamics of metal ion binding and denaturation of a calcium binding protein from *Entamoeba histolytica*, *Biochemistry* 36 (1997) 10910-10916.

[20] F.J. Bruzzese, P.R. Connelly, Allosteric properties of inosine monophosphate dehydrogenase revealed through the thermodynamics of binding of inosine 5'-monophosphate and mycophenolic acid. Temperature dependent heat capacity of binding as a signature of ligand-coupled conformational equilibria, *Biochemistry* 36 (1997) 10428-10438.

[21] F. Bou-Abdallah, M.R. Woodhall, A. Velazquez-Campoy, S.C. Andrews, N.D. Chasteen, Thermodynamic analysis of ferrous ion binding to *E. coli* ferritin EcFtnA, *Biochemistry* 44 (2005) 13837-13846.

[22] S.G. Taneva, S. Bañuelos, J. Falces, I. Arregi, A. Muga, P.V. Konarev, D.I. Svergun, A. Velazquez-Campoy, M.A. Urbaneja, A mechanism for histone chaperoning activity of nucleoplasm: thermodynamic and structural models, *J. Mol. Biol.* 393 (2009) 448-463.

[23] V.H. Le, R. Buscaglia, J.B. Chaires, E.A. Lewis, Modeling complex equilibria in isothermal titration calorimetry experiments: thermodynamic parameters estimation for a three-binding-site model, *Anal. Biochem.* 434 (2013) 233-241.

[24] A. Velazquez-Campoy, G. Goñi, J.R. Peregrina, M. Medina, Exact analysis of heterotropic interactions in proteins: Characterization of cooperative ligand binding by isothermal titration calorimetry, *Biophys. J.* 91 (2006) 1887:1904.

[25] M. Eftink, R.L. Biltonen, Thermodynamics of interacting biological systems, in: A.E. Beezer (Ed.), *Biological Microcalorimetry*, Academic Press, London, 1980.

- [26] W. Du, W.S. Liu, D.J. Payne, M.L. Doyle, Synergistic inhibitor binding to *Streptococcus pneumoniae* 5-enolpyruvylshikimate-3-phosphate synthase with both monovalent cations and substrate, *Biochemistry*, 39 (2000) 10140-10146.
- [27] S.E. Johnson, M.X. Ilagan, R. Kopan, D. Barrick, Thermodynamic analysis of the CSL x Notch interaction: distribution of binding energy of the Notch RAM region to the CSL beta-trefoil domain and the mode of competition with the viral transactivator EBNA2, *J. Biol. Chem.* 285 (2010) 6681-6692.
- [28] Y.L. Zhang, Z.Y. Zhang, Low-affinity binding determined by titration calorimetry using a high-affinity coupling ligand: a thermodynamic study of ligand binding to protein tyrosine phosphatase 1B, *Anal. Biochem.* 261 (1998) 139-148.
- [29] J.M. Bradshaw, V. Mitaxov, G. Waksman, Investigation of phosphotyrosine recognition by the SH2 domain of the Src kinase, *J. Mol. Biol.* 293 (1999) 971-985.
- [30] A. Velazquez-Campoy, E. Freire, ITC in the postgenomic era...? Priceless, *Biophys. Chem.* 115 (2005) 115-124.
- [31] A. Velazquez-Campoy, E. Freire, Isothermal titration calorimetry to determine association constants for high-affinity ligands, *Nat. Protoc.* 1 (2006) 186-191.
- [32] M.T. Henzl, M.E. Davis, A. Tan, Divalent ion binding properties of the timothy grass allergen, *Phl p 7*, *Biochemistry* 47 (2008) 7846-7856.
- [33] A. Velazquez-Campoy, Y. Kiso, E. Freire, The binding energetics of first and second generation HIV-1 protease inhibitors: Implications for drug design, *Arch. Biochem. Biophys.* 390 (2001) 169-175.
- [34] A. Velazquez-Campoy, E. Freire, Incorporating target heterogeneity in drug design, *Journal of Cellular Biochemistry* 84(S37) (2001) 82-88.
- [35] A. Velazquez-Campoy, S. Vega, E. Freire, Amplification of the effects of drug-resistance mutations by background polymorphisms in HIV-1 protease from African subtypes, *Biochemistry* 41 (2002) 8613-8619.
- [36] H. Ohtaka, A. Velazquez-Campoy, D. Xie, E. Freire, Overcoming drug resistance in HIV-1 chemotherapy: The binding thermodynamics of Amprenavir and TMC-126 to wild-type and drug-resistant mutants of the HIV-1 protease, *Protein Sci.* 11 (2002) 1908-1916.

- [37] S. Vega, L.W. Kang, A. Velazquez-Campoy, Y. Kiso, L.M. Amzel, E. Freire, A structural and thermodynamic escape mechanism from a drug resistant mutation of the HIV-1 protease, *Proteins* 55 (2004) 594-602.
- [38] G. Krainer, J. Broecker, C. Vargas, J. Fanghänel, S. Keller S, Quantifying high-affinity binding of hydrophobic ligands by isothermal titration calorimetry, *Anal. Chem.* 84 (2012) 10715-10722.
- [39] M.T. Henzl, L.A. Markus, M.E. Davis, A.T. McMillan, Simultaneous addition of two ligands: A potential strategy for estimating divalent ion affinities in EF-hand proteins by isothermal titration calorimetry, *Methods* 59 (2013) 336-348.
- [40] S.D. Burrows, M.L. Doyle, K.P. Murphy, S.G. Franklin, J.R. White, I. Brooks, D.E. McNulty, M.O. Scott, J.R. Knutson, D. Porter, P.R. Young, P. Hensley, Determination of the monomer-dimer equilibrium of interleukin-8 reveals it is a monomer at physiological concentrations, *Biochemistry* 33 (1994) 12741-12745.
- [41] A. Velazquez-Campoy, S.A. Leavit, E. Freire, Characterization of protein-protein interactions by isothermal titration calorimetry, *Methods Mol. Biol.* 261 (2004) 35-54.
- [42] K. Luke, D. Apiyo, P. Wittung-Stafshede, Dissecting homo-heptamer thermodynamics by isothermal titration calorimetry: entropy-driven assembly of co-chaperonin protein 10, *Biophys. J.* 89 (2005) 3332-3336.
- [43] M.L. Doyle, P.C. Weber, S.J. Gill, Carbon monoxide binding to *Rhodospirillum rubrum* ferrocyanochrome *c'*, *Biochemistry* 24 (1985) 1987-1991.
- [44] G. Gutierrez-Magdaleno, M. Bello, M.C. Portillo-Tellez, A. Rodriguez-Romero, E. Garcia-Hernandez, Ligand binding and self-association cooperativity of β -lactoglobulin, *J. Mol. Recognit.* 26 (2013) 67-75.
- [45] T. Wiseman, S. Williston, J.F. Brandts, L.N. Lin, Rapid measurement of binding constants and heat of binding using a new titration calorimeter, *Anal. Biochem.* 179 (1989) 131-137.
- [46] G.M. Poon, Explicit formulation of titration models for isothermal titration calorimetry, *Anal. Biochem.* 400 (2010) 229-236.
- [47] I. Herrera, M.A. Winnik, Differential binding models for isothermal titration calorimetry: Moving beyond the Wiseman isotherm, *J. Phys. Chem. B* 117 (2013) 8659-8672.

- [48] E. Freire, Y. Kawasaki, A. Velazquez-Campoy, A. Schön, Characterisation of ligand binding by calorimetry, in: A. Podjarny, A.P. Dejaegere, B. Kieffer (Eds.), *Biophysical Approaches Determining Ligand Binding to Biomolecular Targets: Detection, Measurement and Modelling*, RSC Biomolecular Sciences No.22, RSC Publishing (Royal Society of Chemistry), London, 2011.
- [49] A.D. Nielsen, C.C. Fuglsang, P. Westh P, Isothermal titration calorimetric procedure to determine protein-metal ion binding parameters in the presence of excess metal ion or chelator, *Anal. Biochem.* 314 (2003) 227-234.
- [50] X. Arias-Moreno, A. Velazquez-Campoy, J.C. Rodriguez, M. Pocovi, J. Sancho, Mechanism of low density lipoprotein (LDL) release in the endosome: implications of the stability and Ca²⁺ affinity of the fifth binding module of the LDL receptor, *J. Biol. Chem.* 283 (2008) 22670-22679.
- [51] O. Abian, J.L. Neira, A. Velazquez-Campoy, Thermodynamics of zinc binding to hepatitis C virus NS3 protease: a folding by binding event, *Proteins* 77 (2009) 624-636.
- [52] S.J. Gill, B. Richey, G. Bishop, J. Wyman, Generalized binding phenomena in an allosteric macromolecule, *Biophys. Chem.* 21 (1985) 1-14.
- [53] M. Straume, E. Freire, Two-dimensional differential scanning calorimetry: simultaneous resolution of intrinsic protein structural energetics and ligand binding interactions by global linkage analysis, *Anal. Biochem.* 203 (1992) 259-268.
- [54] M.S. Celej, S.A. Dassie, M. Gonzalez, M.L. Bianconi, G.D. Fidelio, Differential scanning calorimetry as a tool to estimate binding parameters in multiligand binding proteins, *Anal. Biochem.* 350 (2006) 277-284.
- [55] V. Dincbas-Renqvist, C. Lendel, J. Dogan, E. Wahlberg, T. Härd, Thermodynamics of folding, stabilization, and binding in an engineered protein-protein complex, *J. Am. Chem. Soc.* 126 (2004) 11220-11230.
- [56] M.J. Cliff, M.A. Williams, J. Brooke-Smith, D. Barford, J.E. Ladbury, Molecular recognition via coupled folding and binding in a TPR domain, *J. Mol. Biol.* 346 (2005) 717-32.
- [57] S.P. Edgcomb, B.M. Baker, K.P. Murphy, The energetics of phosphate binding to a protein complex, *Protein Sci.* 9 (2000) 927-933.

- [58] X. Zhou, T.C. Tan, S. Valiyaveetil, M.L. Go, R.M. Kini, A. Velazquez-Campoy, J. Sivaraman, Structural characterization of myotoxic ecarpholin S from *Echis carinatus* venom, *Biophys. J.* 95 (2008) 3366-3380.
- [59] G.M. Clore, A.M. Gronenborn, R.W. Davies, Theoretical aspects of specific and non-specific equilibrium binding of proteins to DNA as studied by the nitrocellulose filter binding assay. Co-operative and non-co-operative binding to a one-dimensional lattice, *J. Mol. Biol.* 155 (1982) 447-466.
- [60] G.M. Clore, A.M. Gronenborn, R.W. Davies, Cooperative non-specific DNA binding of the N-terminal core of the cyclic AMP receptor protein of *Escherichia coli* and its modulation by cyclic AMP, *FEBS Lett.* 164 (1983) 57-62.
- [61] J.D. McGhee, P.H. von Hippel, Theoretical aspects of DNA-protein interactions: Co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice, *J. Mol. Biol.* 86 (1974) 469-489.
- [62] W.B. Peters, S.P. Edmondson, J.W. Shriver, Thermodynamics of DNA binding and distortion by the hyperthermophile chromatin protein Sac7d, *J. Mol. Biol.* 343 (2004) 339-360.
- [63] W.B. Peters, S.P. Edmondson, J.W. Shriver, Effect of mutation of the Sac7d intercalating residues on the temperature dependence of DNA distortion and binding thermodynamics. *Biochemistry* 44 (2005) 4794-804.
- [64] A. Velazquez-Campoy, Ligand binding to one-dimensional lattice-like macromolecules: Analysis of the McGhee-von Hippel theory implemented in isothermal titration calorimetry. *Anal. Biochem.* 348 (2006) 94-104.
- [65] M.R. Kasimova, A. Velazquez-Campoy, H.M. Nielsen, On the temperature dependence of complex formation between chitosan and proteins, *Biomacromolecules*, 12 (2011) 2534-2543.

Figure 1

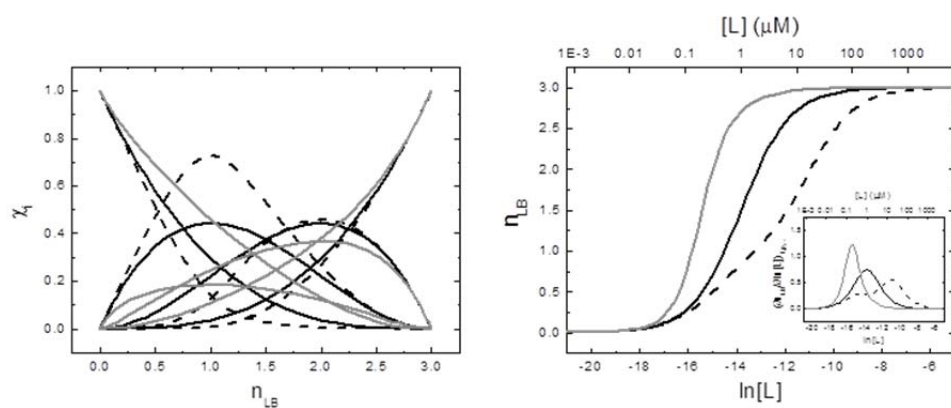


Figure 2

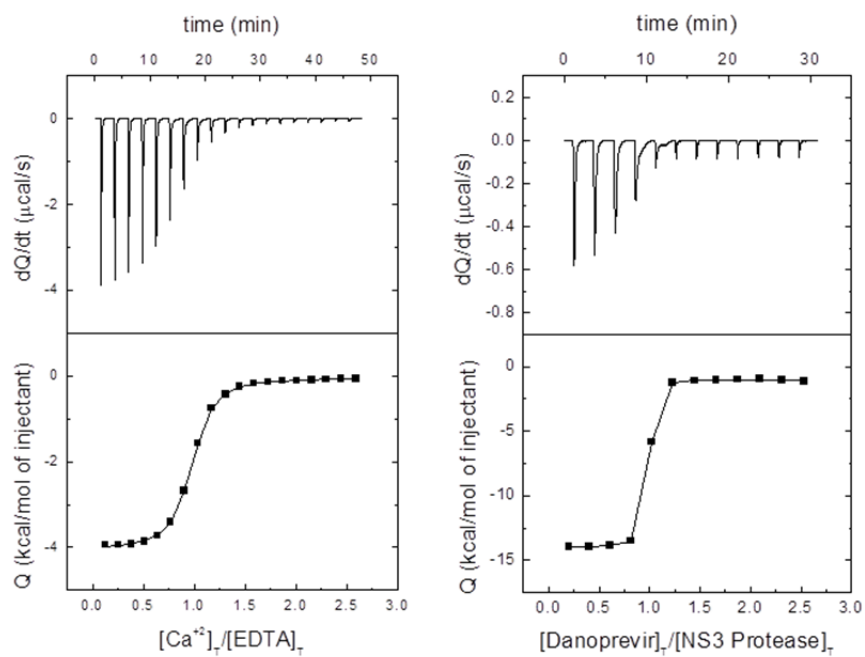


Figure 3

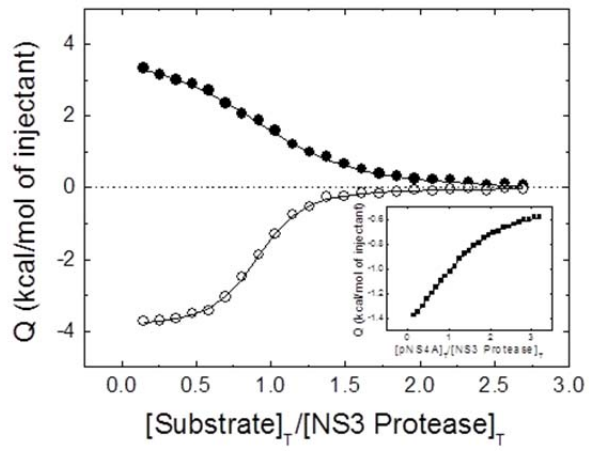


Figure 4

