AFFINImeter: A software to analyze molecular recognition processes from experimental data

Ángel Piñeiro[1*], Eva Muñoz[2], Juan Sabín[2], Miguel Costas[3], Margarida Bastos[4], Adrián Velázquez-Campoy[5-8], Pablo F. Garrido[1], Philippe Dumas[9], Eric Ennifar[10], Luis García-Río[11], Javier Rial[2], Daniel Pérez[2], Patricia Fraga[2], Aurelio Rodríguez[12], Carmen Cotelo[12]

[1]Departamento de Física de Aplicada, Facultade de Física, Universidade de Santiago de Compostela, E-15782 Santiago de Compostela, Spain. [2]AFFINImeter Scientific & Development team, Software 4 Science Developments, S. L. Ed. Emprendia, Santiago de Compostela, A Coruña 15782, Spain. [3]Laboratorio de Biofisicoquímica, Departamento de Fisicoquímica, Facultad de Química, Universidad Nacional Autónoma de México, México City 04510, Mexico. [4]CIQ-UP, Departamento de Química e Bioquímica, Faculdade de Ciências da Universidade do Porto, R. Campo Alegre 687, P-4169-007 Porto, Portugal. [5]Institute of Biocomputation and Physics of Complex Systems (BIFI), Joint Units IQFR-CSIC-BIFI, and GBsC-CSIC-BIFI, Universidad de Zaragoza, Zaragoza 50018, Spain; and Department of Biochemistry and Molecular and Cell Biology, Universidad de Zaragoza, 50009 Zaragoza, Spain. [6]Aragon Institute for Health Research (IIS Aragon), 50009 Zaragoza, Spain. [7]Biomedical Research Networking Centre for Liver and Digestive Diseases (CIBERehd), 28029 Madrid, Spain. [8]Fundacion ARAID, Government of Aragon, 50018 Zaragoza, Spain. [9]IGBMC, Dept of Integrative Biology. Strasbourg University. F67404 Illkirch CEDEX, France. [10]CNRS, Architecture et Réactivité de l'ARN, Institut de Biologie Moléculaire et Cellulaire, Université de Strasbourg, UPR 9002, F-67000 Strasbourg, France. [11]Centro de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS), Departamento de Química Física, Universidade de Santiago, 15782 Santiago, Spain. [12]Fundación Pública Galega Centro Tecnolóxico de Supercomputación de Galicia (CESGA) Avda. de Vigo s/n 15705, Santiago de Compostela

*E-mail: Angel.Pineiro@usc.es

Abstract

The comprehension of molecular recognition phenomena demands the understanding of the energetic and kinetic processes involved. General equations valid for the thermodynamic analysis of any observable that is assessed as a function of the concentration of the involved compounds are described, together with their implementation in the AFFINImeter software. Here, a maximum of three different molecular species that can interact with each other to form an enormous variety of supramolecular complexes are considered. The corrections currently employed to take into account the effects of dilution, volume displacement, concentration errors and those due to external factors, especially in the case of ITC measurements, are included. The methods used to fit the model parameters to the experimental data, and to generate the uncertainties are described in detail. A simulation tool and the so called kinITC analysis to get kinetic information from calorimetric experiments are also presented. An example of how to take advantage of the AFFINImeter software for the global multi-temperature analysis of a system exhibiting cooperative 1:2 interactions is presented and the results are compared with data previously published. Some useful recommendations for the analysis of experiments aimed at studying molecular interactions are provided.

Key Words: Molecular Recognition; Affinity; Isothermal Titration Calorimetry; Thermodynamics; Kinetics; Analysis Software

Table of contents:

1.- Introduction

The understanding of the underlying mechanisms involved in molecular recognition processes requires the knowledge of the structural, energetic and kinetic features associated with these events. This in turn allows the rational design or optimization of chemical and biochemical compounds for specific applications. Examples of processes that can be studied in solution are protein-ligand association, the dissociation/association of protein oligomers, the encapsulation of small molecules in hydrophobic cavities or in molecular aggregates formed by amphiphilic compounds, step-wise binding of small molecules to a large polymer, spontaneous aggregation of molecules into larger self-organized structures, micelle formation/dissociation, and adsorption of molecules onto solid interfaces [1-2].

For the quantitative characterization of these processes we need to apply a thermodynamic model to a set of experimentally obtained data of an observable property that is sensitive to the amount of free and bound species in the system [3-5]. For recognition or binding processes between any two molecules (M and A) at a given temperature and pressure, the representation of such an observable as a function of the total concentration of M or A (or, alternatively, as a function of the concentration ratio [M]/[A] or [A]/[M]) is typically referred to as "equilibrium isotherm" or "saturation binding curve". From this curve, quantitative information on the thermodynamic properties of the studied process (equilibrium constant, Gibbs energy, enthalpy, entropy, heat capacity) can be obtained. This equilibrium thermodynamic information can in turn be used to interpret the data at molecular level, together with information from other experiments and/or models (X-ray diffraction, NMR, molecular dynamics simulations, statistical thermodynamics, *etc*). Additionally, when the observable is also monitored as a function of time for different M or A concentrations, a kinetic characterization of the binding event is also possible [6].

Different experimental techniques are available to study molecular recognition. The most commonly used methods are Isothermal Titration Calorimetry (ITC), Nuclear Magnetic Resonance (NMR), Surface Plasmon Resonance (SPR), Microscale Thermophoresis (MST) and a variety of spectroscopic methods including fluorescence, circular dichroism, IR and UV-vis. The thermodynamic information obtained from any of these techniques for a given process must *in principle* be the same. However, differences are expected when the experimental setup is not identical, *e.g*. the use of different sample volume and concentrations in order to maximize the signal-to-noise ratio, the immobilization of the target compound to a substrate or the use of an external marker (*e.g*. a fluorescent tag). Particularly in the cases of immobilized or labelled reactants the system is expected to behave differently from experiments where free, unlabelled molecules are present in solution. Different techniques may also differ in the way the system is perturbed and the observed signal is measured. For instance, ITC is based on sequential perturbations of equilibrium states followed by return to equilibrium. The measured signal is the heat released or absorbed upon small concentration changes due to several factors: dilution, aggregation/dissociation or the transformation between supramolecular complexes of different stoichiometry. Therefore, the kinetics of the binding process may affect the results obtained for the thermodynamic parameters obtained from ITC, mainly for slow-binding ligands. In contrast, most of the techniques designed for the study of intermolecular interactions register an observable signal of well equilibrated samples, *i.e.*, the

samples are not perturbed during the experiment. Nevertheless, even under the same conditions and techniques, different values for the thermodynamic parameters are found in the literature. There are a large number of reasons for these discrepancies to occur, ranging from uncertainties in sample concentration, incorrect instrument calibration, technical difficulties to obtain a good signal-to-noise ratio, and finally differences arising from the data fitting. A detailed discussion of the relevance of each of the factors leading to the inconsistencies between the thermodynamic parameters obtained from different techniques is indeed outside the scope of the present work. However, it should be noted that efforts are being made towards the standardization of results obtained from different biophysical techniques [7].

In order to improve the characterization of a large variety of molecular processes that are often involved in fundamental or applied research, several companies are continuously making efforts to optimize the sensitivity and performance of the instruments, as well as improving the hardware and software aimed at collecting the primary data (raw signal). Concurrently, complementary software is needed to properly analyze experimental information of different degrees of complexity and retrieve the thermodynamic constants that characterize the processes under study. For this, the selection of a proper binding model, the correct estimation of the involved parameters, avoiding local minima and over-parameterization, as well as getting a good estimation of uncertainties, are of paramount importance. AFFINImeter was designed in order to fulfil these aims. This software offers the traditional binding models, but more importantly, it makes possible to develop user-tailored models in a very simple way, a possibility not publicly available in any other analysis software. Furthermore, during the processing of the raw signal useful information is sometimes discarded, as it has been the case for ITC measurements –a technique that has been traditionally employed mainly to get thermodynamic data– where the kinetic information has remained mostly unaddressed. AFFINImeter allows users to obtain valuable kinetic information from the raw ITC data, with no additional effort nor sample consumption, using a particular development called kinITC [8]. At present, AFFINImeter can analyse binding curves from ITC, from NMR, and from spectroscopy data. These different AFFINImeter applications are referred to as AFF-ITC, AFF-NMR and AFF-Spectroscopy. In the present work, we describe the main features and technical details of this software, with special emphasis on the implementation of thermodynamic binding models that represent a large variety of binding and molecular recognition events and mechanisms.

2.- Software description

The software currently available for the analysis of binding events from experimental measurements has different origins: (i) software attached to a specific instrument, typically developed and provided by the equipment manufacturers; (ii) third-party software specialized in a single technique for several instrument brands; and (iii) third-party software for a range of techniques and instrument brands. The last two are usually developed by private companies or academic groups.

AFFINImeter is currently a shareware software belonging to the last group, thus aiming at having special applications to different techniques. The AFFINImeter branch for ITC (AFF-ITC) is

currently more complete than those for NMR and other spectroscopies (AFF-NMR and AFF-Spectroscopy) since the former includes the analysis of the raw data for ITC instruments commercialized by the main manufacturers. The raw data analysis for NMR measurements is externally performed by MNova (http://mestrelab.com/software/mnova/binding/), which works jointly with AFF-NMR. The analysis of raw data for other spectroscopies will be implemented in AFFINImeter in the future. Belonging to group (iii), the models, fitting engines and algorithms, as well as the treatment of uncertainties and local minima –the core of AFFINImeter– are common for AFF-ITC, AFF-NMR and AFF-Spectroscopy. In what follows, the main features of AFFINImeter are illustrated through their application to ITC data.

The main features of the AFFINImeter software are briefly described here. They are explained in detail in the following sections.

- Fully automated integration of the ITC raw data, including noise removal, baseline correction and calculation of the uncertainty for each data point of the equilibrium isotherm;
- Complete flexibility in the design of models for the thermodynamic analysis, defined by the user in a comprehensive and simple form. Examples of possible events/models are self-association, multi-site binding, competitive interactions, formation of ternary complexes (consisting of any stoichiometric combination of 3 different molecules) and coupling between several processes (*e.g.* self-association + 1:1 binding);
- Kinetic analysis using the kinITC method [9] applied to the ITC raw data corresponding to 1:1 and 1:n interactions. This method allows obtaining the rate constants (k_{on} and k_{off}) with no extra effort, using the same titration experiment data employed for the thermodynamic analysis;
- Fitting parameters that allow correction for the concentration of active (i.e. competent to form complexes) titrant (syringe) and titrate (calorimetric cell);
- No need to provide numerical starting values for the fitted parameters;
- Dynamic mathematical/logical relationships between parameters and their maximum/minimum boundaries. This is useful mainly when including hybrid models and also to prevent over-parameterization;
- Powerful analysis of local minima and final parameter uncertainties;
- Global analysis of several measurements, including direct + reverse experiments, repeats of the same experiment or multi-temperature experiments, by sharing fitting parameters when convenient;
- Powerful simulation tool using an unlimited number of models;
- Results supplied in detail, including species distribution plot. The automated report of the obtained results is generated in html and pdf formats. The values represented in each plot can also be downloaded in csv format for further treatment by the users;
- In addition to standard user accounts, administrator accounts are also available to manage user accounts. Administrator accounts have access to the data of their associated users;
- Accessibility from any device with internet connection.

2.1.- Treatment of raw data to obtain the equilibrium isotherm

The processing of the raw data to generate the equilibrium isotherm is a very important step in the analysis of the experimental data. AFF-NMR does not directly analyze the NMR raw data, but takes the equilibrium isotherm from the MNova Software [http://mestrelab.com/]. For AFF-Spectroscopy it is currently assumed that the raw data has been externally processed. In contrast, AFF-ITC uses the raw data from the most common ITC instruments (Malvern-MicroCal and TA Instruments) generating the equilibrium isotherm and the equilibration time curve (ETC) used to yield kinetic information from the ITC experiment (see section 2.8 for details). The basis of the raw data treatment and data management in AFF-ITC and kinITC has been reported in [10-11]. Here we describe only the relevant parts in data treatment that allow comprehension of the thermodynamic parameters' retrieval.

The raw data of a power-compensation ITC instrument consists of a two column file containing the power supplied to the sample cell to keep the temperature constant and the corresponding time. During a typical experiment, the power supplied signal changes due to small perturbations during baseline periods (stirring, small environmental temperature changes, etc.) and also due to the larger perturbation produced by the injection of small aliquots of the solution contained in the syringe into a solution located in the sample cell (injection-recovering period). After the system in the cell has returned to equilibrium, a stable baseline is recovered. Thus, the signal corresponding to the injection-recovering period is recorded as a power *vs* time peak. To perform the thermodynamic analysis, the peaks of a complete titration have to be integrated. For this, the software automatically generates a baseline for each injection, which is subtracted from the raw signal. These individual baselines are obtained by fitting a smooth function joining the first point of the signal injection –just before the perturbation begins– to the trend of the last part of the signal, i.e. excluding the perturbed region. The employed function for this fitting is a linear combination of orthogonal functions (Legendre polynomials of first, second and third order). Finally, the noise is reduced producing a corrected titration curve that can be directly visualized. The noise removal is based on analysis of local deviations to the signal trend. Several controls are available to the user in order to manipulate the level of noise to be removed as well as to choose the time period between peaks to be considered as baseline. Thereafter the peaks are integrated, generating the equilibrium isotherm, and the associated uncertainties per point, σ, are calculated. The process is fully automated by default, but the user can easily modify the integration limits. The integral of each peak represents the heat released or absorbed due to the events (binding for example) that take place in the sample cell as a result of each injection. This energy is normalized by the number of moles of the injected compound (or number of moles of reactant in the cell), leading to the molar enthalpy change. The corresponding uncertainty for each point in the isotherm is the sum of three contributions [10]: (i) that due to the signal's noise, determined as the product of its standard deviation in the baseline region times the duration of the signal perturbation upon injection; (ii) that emerging from possible outliers in the perturbed region, that is evaluated as the difference between the area obtained by numerical integration of the experimental peak and the area corresponding to the integration of a mathematical function fitted to the same data (the employed mathematical function is a truncated series expansion of peak functions, with the number of terms used being chosen by a F-test for statistical significance); and (iii) that arising from the dispersion of

the points constituting the isotherm, estimated from the difference between the interpolation or extrapolation of first and second order polynomials to groups of sequential three and four points in the equilibrium isotherm. The software performs several rounds of interpolations and extrapolations taking the points from lower to higher titrant concentrations and in the opposite direction. The minimum difference between the experimental value and the extrapolated/interpolated point found among all these rounds is taken as a contribution to the final uncertainty. The calculation of this contribution is clearly described by Keller *et al* [12]. Our implementation is an adaptation of the method described by these authors. The total uncertainty is determined as the square root of the sum of squares of all contributions. This allows assigning an uncertainty to every experimental point in the equilibrium isotherm. This is possible even for points where the signal abruptly changes, provided that the dispersion between the two regions (*i.e.* before and after the abrupt change) is small, as typically occurs for very strong interactions.

An example for the automatic treatment of ITC raw data is shown in Figure 1. By default, the binding isotherm and the so called Equilibration Time Curve (the curve obtained by plotting the time required for each injection to recover the baseline vs the corresponding concentration ratio, see Figure 1 and section 2.8) are analysed assuming a simple 1:1 interaction. An F-test is employed to decide if corrections due to the dilution of the titrant and/or to the concentration of active reactants are required. The analysis for a single experiment typically takes a few seconds and it is fully automated for the simplest interactions. A maximum of 10 experiments can be simultaneously uploaded and processed. This is useful to simultaneously process multi-temperature experiments where the software employs the van't Hoff and Eyring equations. If the interaction model is more complex, then the user can specifically design a reaction scheme using the model builder tool to analyse the experimental data (see Figure 2 and this video: https://www.youtube.com/watch?v=dhfW-PQ-qVA).

Figure 1. Graphs automatically displayed in AFFINImeter-ITC upon raw data upload and processing. A) ITC data preview, which shows the raw thermogram; B) processed thermogram after baseline correction and noise removal; C) binding isotherm, resulting from the peak-by-peak integration of the processed thermogram as a function of At/Mt (titrant to titrate molar ratio) with the error bars determined as explained in the text; D) Equilibration time curve (ETC) resulting from the equilibration time for each injection as a function of the stoichiometric ratio At/Mt (the error bars of the ETC are determined using only the third contribution explained in the text and the time resolution of the experiment, typically 1 second). Solid red lines in C and D are the results of the fittings to a 1:1 binding model and to equation 8 in [11], respectively.

2.2.- The use of generalized equations to build interaction models

There is a wide variety of software available for the analysis of intermolecular interactions from data obtained with specific biophysical techniques/applications (i.e. Origin, Nanoanalyze and HypCa for ITC, Scrubber for SPR, Titan for NMR, and Sedphat for a number of techniques including AUC, ITC and DLS) as well as generic analysis software of binding curves (i.e. GraphPad, SigmaPlot and DynaFit). Generally, these computational programs can be classified as software that only offers a list of predefined models that are independently implemented, or as software that permits the implementation of new models. In the latter case, generation and validation of new models is not straightforward and requires a good knowledge of equilibrium thermodynamics, mathematics, and algorithms, and possibly also computational programming. Further, for some instruments it is necessary to be familiar with technical details that could affect the equations to be programmed. It is worth to mention that the validation of new models is not trivial mainly because typically no reference results are available. For these reasons, only experienced users are able to use these advanced tools to create models.

AFFINImeter has solved these problems by adopting a different strategy –the use of generalized equations for the two main families of binding models, namely “stoichiometric equilibria” (SE) and “independent sites” (IS) models (see below). Since the equations are generalized, the amount of available models for each family is virtually unlimited. The validation is also much more simple and robust, since any error present in relatively simple models (easy to detect) should also be present in more complex models. Thus, the mathematical induction method can be employed to demonstrate the validity of non-trivial models (typically difficult and in many cases impossible mainly due to the lack of appropriate data) from the results obtained with simple ones. Predictions based on simulation of experiments upon variation of the different parameters can also be used to check for the consistency of the results. Note that AFFINImeter allows the simulation of experiments corresponding to any of these two model families.

The SE models are based on stoichiometric (or macroscopic) equilibrium constants and they are typically applied to low stoichiometry interactions where the different binding sites can be coupled to each other. The IS models are based on site-specific (or microscopic) equilibrium constants where the different binding sites do not interact with each other, although they can be grouped in sets where the sites have identical affinity and enthalpy of interaction with the involved species. The main features of the SE and IS models are summarized in Table 1, and both models are described in detail in the following sections. Standard available software for

the study of intermolecular interactions includes just a few predefined SE models, as well as non-competitive IS models for one-set and two-sets of independent sites. This translates into a number of limitations, namely (i) competitive IS interactions are not implemented; (ii) no SE competitive models, other than for 1:1 interactions, are available; (iii) hybrid binding models such as coupled dissociation/binding and those that represent non-standard experiments (e.g. those where there is an "extra" compound or cosolute simultaneously present in both the syringe and cell in the ITC instrument) are missing; (iv) fitting parameters that account for potential deviations between nominal and true active concentration of the compounds participating in the binding event can only be considered for 1:1 interactions; and (v) the enthalpic contribution of the dilution of the titrant to the total signal in ITC instruments is not properly modelled (or not modelled at all), thus requiring complementary experiments in many cases. These five limitations have been surpassed in AFFINImeter, as described and discussed in what follows.

Table 1.- Main features of Stoichiometric Equilibria and Independent Sites models as implemented in AFFINImeter. A detailed description of both types of models together with some examples are provided in sections 2.2.3 and 2.2.4.

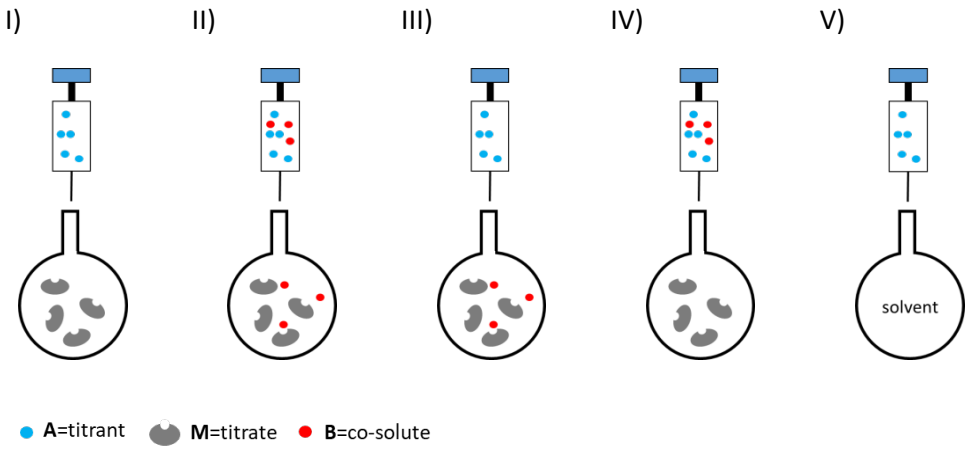
Stoichiometric Equilibria models	Independent Sites models
Consist of a set of connected equilibria (reaction scheme) between stoichiometric species, written in chemical language. The stoichiometric equilibria can be established between free species (uncomplexed initial compounds) and a complex of defined stoichiometry or between two different complexes.	Consist of one or more sets of sites. Each set is characterized by a number of independent sites that are thermodynamically equivalent. Sites of the same set will bind the ligand with the same binding affinity and enthalpy change. Sites of different sets are not equivalent. All sites are independent to each other.
Allow the thermodynamic characterization of equilibria between stoichiometric species.	Allow the thermodynamic characterization of binding equilibria between the ligand and the receptor site(s) and defines the stoichiometry of the ligand-receptor complex.
Considers the possibility that binding of one ligand may influence the binding of other ligand(s), *i.e.* cooperative binding in multi-site binding events.	Assume that binding sites are independent, and therefore it is not applicable to the characterization of cooperative interactions.
Each equilibrium considered is characterized by the corresponding stoichiometric (or macroscopic) equilibrium constant and enthalpy change.	The binding sites are grouped in sets with identical site (or microscopic) equilibrium constant and enthalpy, thus allowing the interaction of a macromolecule with a large number of ligand molecules, without increasing the number of fitting parameters
Competitive ligands can be considered	Competitive ligands can be considered

2.2.1.- Nomenclature and experimental setups in ITC experiments

General binding models that involve the participation of 1, 2 or 3 different molecules have been implemented in AFFINImeter for both model families (SE and IS). In a standard ITC experiment there is a target molecule –typically a macromolecule located in the sample cell–

to which one ligand candidate –initially located in the syringe at a higher concentration– can bind. However, several factors like the possible low solubility of the ligand or limitations in the amounts of the compound available, lead to exchange the reactants in the syringe and cell, performing the often called reverse experiment. From the thermodynamic point of view direct and reverse experiments are symmetrical, *i.e.* the thermodynamic equations and parameters describing the interaction between two molecules are the same, irrespective of the experiment's design. AFFINImeter denotes the main solute in the sample cell as M (as a reference to "macromolecule", the most typical setup) and the main solute in the syringe as A regardless the position of the reactants. A third compound participating in the interaction, B, may be present in both the sample cell and the syringe, only in the sample cell, or only in the syringe. Additionally, dilution experiments where only the compound A is initially located in the syringe and the sample cell is filled with the solvent can be convenient to test for any potential self-association/aggregation of compound A. Dilution experiments are also used to study the dissociation of protein oligomers, polymer aggregates or de-micellization. The ITC experimental setups that can be considered in AFF-ITC are depicted in Figure 2.

A)



B)

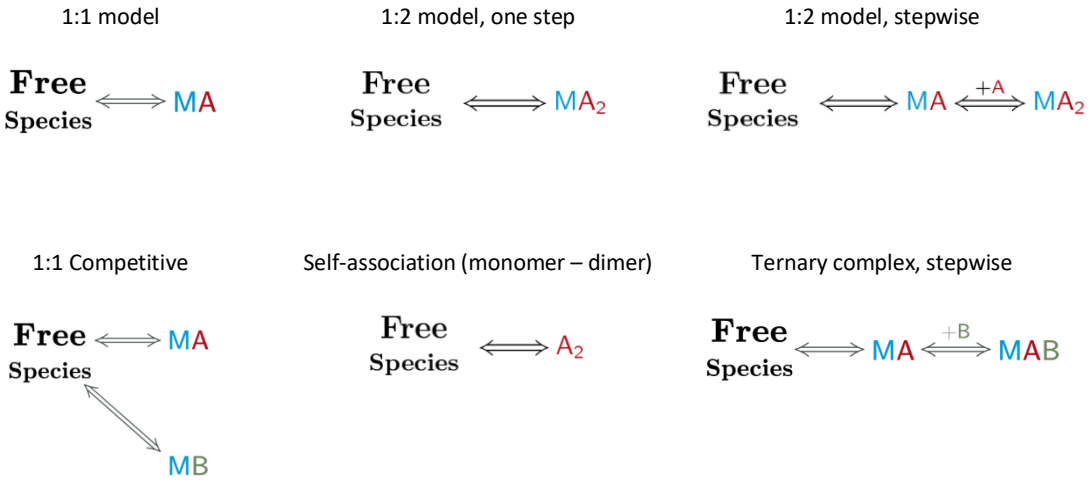


Figure 2. A) Schematic representation of ITC experimental setups and nomenclature used in AFFINImeter showing: I) standard titration of A into M; titration of A into M in the presence of II) B in the cell and in the syringe; III) B mixed with M in the cell; IV) B mixed with A in the syringe; V) dilution experiment of A, in syringe, into solvent in the cell. B) A few illustrative examples of AFFINImeter binding models. Note that the possible presence of a co-solute (B) in the syringe, in the cell or in both locations significantly expands the variety of models that can be applied, including displacement experiments or the formation of relatively complex chemical species consisting of more than two molecules.

2.2.2.- Equilibrium constants – stoichiometric *vs* site models

SE models consider the equilibria between stoichiometric species regardless the location of the ligand in a specific site of the receptor, while IS models explicitly assumes binding of the ligand to specific sites. Therefore, the equilibrium constants in each model family have a different meaning, although they are in fact related.

Let's consider the interaction of a ligand (A) with a divalent macromolecule (M) as depicted in Figure 3. Using the SE approach two stoichiometric species can be formed, M_1A_1 and M_1A_2, and the following stepwise chemical reactions can be written:

$$M + A \leftrightarrow M_1A_1 \qquad \text{and} \qquad M_1A_1 + A \leftrightarrow M_1A_2 \tag{1}$$

with the corresponding stepwise stoichiometric equilibrium constants being:

$$K_{11} = [M_1A_1]/([M]\cdot[A]) \qquad \text{and} \qquad K_{12} = [M_1A_2]/([M_1A_1]\cdot[A]) \tag{2}$$

where [M], [A], $[M_1A_1]$ and $[M_1A_2]$ denote the molar concentrations of free M and A species and of complexes of stoichiometry 1:1 and 1:2, respectively. K_{11} is the binding constant characterizing the formation of the complex M_1A_1 from the free species, regardless of the binding site, while K_{12} characterizes the formation of the complex M_1A_2 from M_1A_1 and a second A molecule.

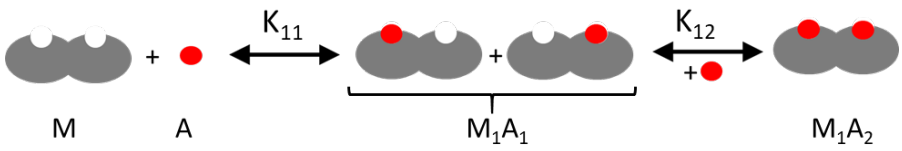


Figure 3. Schematic diagram of binding equilibria based on stoichiometric constants, for the interaction between a bivalent receptor, M, with a monovalent ligand, A, is represented.

The formation of the M_1A_2 complex can also be written as a single step process:

$$M + 2\cdot A \leftrightarrow M_1A_2 \tag{3}$$

leading to the **global** stoichiometric equilibrium constant:

$$\beta_{11} = [M_1A_2]/([M]\cdot[A]^2) \quad (4)$$

The global and stepwise constants are related as $\beta_{11} = K_{11}\cdot K_{12}$ and thus the units of the global and stepwise equilibrium constants are different. Note that the change in free energy for the formation of a given complex is determined as the natural logarithm of the product of the corresponding equilibrium constant by the ratio of standard state concentrations of reactants and products, raised to the appropriate powers. Thus, the argument of the logarithm is dimensionless, as required. The values of global and stepwise constants cannot be directly compared and differ, typically, by several orders of magnitude. AFFINImeter uses stepwise equilibrium constants and so they are always denoted by *K*, even though the equilibria consisting any model can skip specific stoichiometric species, i.e. the difference of the sum of the stoichiometric coefficients of a given chemical species upon a reaction can be higher than 1. For instance, for the reaction scheme $M + A \leftrightarrow M_1A_1 + 2\cdot A \leftrightarrow M_1A_3$ the complex M_1A_2 is not present and the difference between the stoichiometric coefficient of M_1A_1 and M_1A_3 is 2. The values of *K* for different orders of association cannot be directly compared. The best way to compare the weight of two reactions is to observe the species distribution plots (see below).

On the other hand, for the same system, i.e. for a monovalent ligand A binding to a divalent macromolecule M, the two involved sites can be considered to be distinguishable (see Figure 4). The explicit consideration of specific binding sites involves the incorporation of four complexes and four equilibria into the model:

$M + A \leftrightarrow M_{s1}A_1$ $M + A \leftrightarrow M_{s2}A_1$ $M_{s1}A_1 + A \leftrightarrow M_1A_2$ $M_{s2}A_1 + A \leftrightarrow M_1A_2$

where $M_{s1}A_1$ is the 1:1 complex with site 1 occupied by A, while $M_{s2}A_1$ is the 1:1 complex with the site 2 occupied. This reaction scheme implies a larger number of parameters, as a different equilibrium constant is needed for each site:

$$K_{s1} = [M_{s1}A_1] / ([M]\cdot[A]) \quad \text{and} \quad K_{s2} = [M_{s2}A_1] / ([M]\cdot[A]) \quad (5)$$

and two additional constants for the saturation of M, depending on which site was first occupied:

$$K_{s1s2} = [M_1A_2] / ([M_{s1}A_1]\cdot[A]) \quad \text{and} \quad K_{s2s1} = [M_1A_2] / ([M_{s2}A_1]\cdot[A]) \quad (6)$$

Thus in this case we are dealing with the so called microscopic, or ‘site specific’ binding constants.

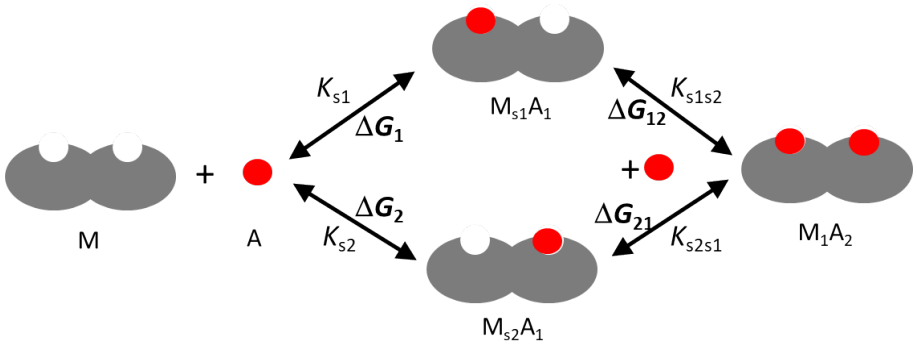


Figure 4. Schematic diagram of binding equilibria based on site constants, for the interaction between a bivalent receptor, M, with a monovalent ligand, A. The site equilibrium constant and the change in Gibbs energy associated to each reaction is defined next to each arrow.

Since the stoichiometric species M_1A_1 in the SE model represents the sum of possible complexes of 1:1 stoichiometry formed between M and A, then $M_1A_1 = M_{s1}A_1 + M_{s2}A_1$. Therefore, the relationship between stoichiometric and site-specific equilibrium constants is:

$$K_{11} = K_{s1} + K_{s2} \qquad \text{and} \qquad K_{12} = (K_{s1s2} \cdot K_{s2s1}) / (K_{s1s2} + K_{s2s1}) \tag{7}$$

From the above descriptions and definitions it is clear that site equilibrium constants are more informative than stoichiometric constants, but the number of site equilibrium constants is, in this particular case, twice that of stoichiometric constants. The problem is simplified by the fact that, in this case, the 1:2 complex is the same regardless the pathway employed to reach it, so the reactions represented in Figure 4 are not independent: $\Delta G_1+\Delta G_{12}=\Delta G_2+\Delta G_{21}$. This introduces a connection between the constants given by equations (5) and (6): $K_{s1}/K_{s2} = K_{s2s1}/K_{s1s2}$. Anyway, for more complex interactions, of higher order stoichiometry, the number of required thermodynamic parameters would make the problem using site constants unapproachable. The situation can be simplified if:

$$K_{s2s1} = K_{s1} \quad \text{and} \quad K_{s1s2} = K_{s2} \tag{8}$$

This simplification means that the sites are considered independent from each other, *i.e.* binding of A to site 2 does not influence the binding of a second molecule of A to site 1, and vice-versa. The relationship between stoichiometric (macroscopic) and site-specific (microscopic) constants then becomes:

$$K_{11} = K_{s1} + K_{s2} \qquad \text{and} \qquad K_{12} = (K_{s1} \cdot K_{s2}) / (K_{s1} + K_{s2}) \tag{9}$$

Furthermore, if sites are independent **and** equivalent, all site constants will be equal:

$$K_{s1} = K_{s2} = K_{s1s2} = K_{s2s1} \tag{10}$$

and then,

$$K_{11} = 2 \cdot K_{s1} \qquad \text{and} \qquad K_{12} = K_{s1} / 2 \tag{11}$$

Thus, finally we can write for the system under consideration

$$K_{12} = K_{11} / 4 \quad (12)$$

An equivalent derivation of this equation can be found in [3]. It is worth to mention that equation (12) fulfils for a divalent interaction with two independent and equivalent sites. However, fulfilment of the previous equations does not necessarily imply that the interactions are divalent, since when dealing with an experimental data set arising from a different type of mechanism –perhaps even more complex than a divalent interaction– the system might also be well described by those same equations. For example, apparently complex three-steps mechanisms can be described exactly by a trivial one-step mechanism. This is the case for a macromolecule exhibiting an equilibrium between two conformational states with different affinities and enthalpies for a given ligand. Such a composite mechanism requires one set of equilibrium constant and enthalpic terms for the interconversion step and two additional sets for the two association/dissociation steps. However, as far as thermodynamics is concerned, the same system can be rigorously described by a simple one-step mechanism with a single set of apparent equilibrium constant and enthalpy terms that can be expressed as functions of the original three sets of equilibrium constants and enthalpies [4].

2.2.3.- Stoichiometric Equilibrium (SE) models

Equations 1-4 can be generalized to describe the interaction between three molecules M, A and B, forming complexes of any stoichiometry. Since the reaction scheme leading to a complex of certain stoichiometry in terms of stepwise reactions is not unique, it is convenient to write the equations in terms of global equilibrium constants (it is worth to remind again that AFFINImeter uses stepwise constants and that the global and the stepwise equilibrium constants can be easily related):

$$m\cdot M + a\cdot A + b\cdot B \leftrightarrow M_mA_aB_b \qquad \text{and} \qquad \beta_{mab} = [M_mA_aB_b] / ([M]^m\cdot[A]^a\cdot[B]^b) \quad (13)$$

Thus, if a particular observable X (e.g. molar enthalpy of complex formation) is sensitive to the presence of $M_mA_aB_b$, the observed signal is the sum of population-weighted contributions of free and bound species, and the averaged signal contained in a volume *V* would be given by eq. (14). It is worth to note that some experimental observables require signal average based on different concentration units such as mole fraction (e.g. chemical shift) or mole fraction of total intensity (e.g. fluorescence anisotropy.

$$X = \sum_{s,j,k=0}^{m,a,b} X_{sjk} \cdot [M_sA_jB_k] \cdot V \quad (14)$$

where X_{sjk} is the molar contribution of $M_sA_jB_k$ to the observable X and

$$[M_sA_jB_k] = \beta_{sjk}\cdot[M]^s\cdot[A]^j\cdot[B]^k \quad (15)$$

where β_{sjk} is the global stoichiometric equilibrium constant corresponding to the formation of $M_sA_jB_k$ and the total concentration of each compound ($[M]_T$, $[A]_T$ and $[B]_T$) is distributed between the existing free and complexed forms:

$$[M]_T = \sum_{s,j,k=0}^{m,a,b} s \cdot [M_sA_jB_k] \quad (16)$$

$$[A]_T = \sum_{s,j,k=0}^{m,a,b} j \cdot [M_sA_jB_k] \quad (17)$$

$$[B]_T = \sum_{s,j,k=0}^{m,a,b} k \cdot [M_s A_j B_k] \quad (18)$$

where the factors *s*, *j* and *k* in equations 16-18 account for the number of M, A and B molecules within each $M_sA_jB_k$ complex, respectively. Note that Equations 16-18 include in the sums the possibility of self-association of a single compound (M_s, A_j, or B_k), as well as the formation of various combinations of hetero-complexes formed by two compounds (M_sA_j, A_jB_k or M_sB_k).

Equations (14-18) include two parameters for each chemical species: X_{sjk} and β_{sjk}. The values for those parameters can be numerically obtained by fitting the equation (14) to the experimental data, i.e. to the observable *X* as a function of the total concentration of the involved compounds. For each set of {X_{sjk}, β_{sjk}} values used during the iterative fitting procedure, the concentration of the different chemical species can also be numerically determined using equations (16-18), provided that the total concentration of the three compounds is known. For the final values of {X_{sjk}, β_{sjk}}, the concentration of each complex, as well as its contribution to the averaged signal, can be determined for every concentration of the three compounds. These equations can be directly used to obtain stoichiometric equilibrium constants from the analysis of binding curves obtained by NMR titrations and general spectrometric measurements (available in AFF-NMR and AFF-Spectroscopy). For ITC their use is not straightforward when the experiments are performed in instruments of the "filled type" (i.e., Malvern MicroCal or TA instruments) as (i) due to instrument's design, upon each injection an equivalent volume of solution is expelled from the effective cell volume, implying that the concentration of the compounds in the cell is changed after each injection in a way in which calculation might depends on various assumptions; (ii) the measured heat does not correspond to a specific concentration, but rather to the heat absorbed or delivered during the concentration change, *i.e.* the shape of the equilibrium isotherm may depend on the volume and speed of the injections, as well as on stirring speed; and (iii) the total signal may contain the contributions of the heat of dilution, mainly those of the injected compounds. Thus, for filled type ITCs it is necessary to consider the calculation of the concentration change in the cell after each injection, as well as the corresponding corrections to the detected heat. Additionally, the contribution of the dilution effect to the signal should be subtracted. This last effect can be done after the experiment if separate dilution experiments are performed, or can be included in the equations if modelled. Both options can be applied with AFFINImeter, although the last one is usually more practical and no less accurate. Including all the above mentioned terms, the final equation used in AFFINImeter for the analysis of ITC data under the stoichiometric equilibrium scheme is:

$$\Delta Q_i = \{-Q_{syr,i} + (Q_{i+1} - Q_i) + Q_{corr,i} + \Delta H_{dil} \cdot ([A_{syr}] - [A]_{T,i}) + Q_{db}\} / ([A_{syr}] \cdot v_{inj,i}) \quad (19)$$

where ΔQ_i is the heat absorbed or delivered per mole of A added during the injection i. Note that the presence of M and B in any experiment is optional (see Figure 2), but it is compulsory that compound A be present at some concentration in the syringe. $Q_{syr,i}$ is the total heat required to form all chemical species present in the injected volume ($v_{inj,i}$). It can be calculated as:

$$Q_{syr,i} = \sum_{s,j,k=0}^{m,a,b} \Delta H_{sjk} \cdot [M_s A_j B_k] \cdot v_{inj,i} \quad (20)$$

with ΔH_{sjk} being the molar enthalpy change corresponding to the formation of the complex $M_s A_j B_k$ from the free species (equivalent to X_{sjk} in eq. 14). Similarly, Q_i is the total heat required to form all chemical species present in the cell, and is given by:

$$Q_i = \sum_{s,j,k=0}^{m,a,b} \Delta H_{sjk} \cdot [M_s A_j B_k]_i \cdot V_i \quad (21)$$

One should note that whereas for $Q_{syr,i}$ the concentration of each chemical species does not depend on the injection since the concentration of the syringe is constant, in the case of Q_i the concentration of the complexes depends on the total concentration of each species in the cell, so it is expected to be different for each injection. In eq. 21 the index i was included for the total volume of the solution in the cell (V_i), although for Malvern MicroCal or TA ITC instruments it is a constant. In equation (19), $Q_{corr,i}$ is the heat contribution associated with the sample that is displaced (removed from the section of cell where the sensors are located) during the injection. Since the concentration changes after each injection, and the contributions Q_{syr} and Q_i are calculated using the concentrations in the cell after i and i+1 injections, $Q_{corr,i}$ is taken as the average between the heat delivered or absorbed corresponding to a volume $v_{inj,i}$ that leaves the cell upon injection:

$$Q_{corr,i} = [(Q_i + Q_{i+1}) / 2] \cdot v_{inj.i} / V_i \quad (22)$$

In contrast to other software for the analysis of ITC measurements, in AFFINImeter the dilution contribution is calculated as the product of the dilution molar enthalpy (ΔH_{dil}) and the difference between the concentration of A in the syringe [A_{syr}] and that in the sample cell after each injection $[A]_{T,i}$ (see eq. 19). As expected, this term is maximum for the first injection and decreases as the concentration of A in the cell approaches that in the syringe. Finally, in Equation (19) there is an additional term (Q_{db}) that represents a constant correction for a potential baseline shift, generally due to external (not specific for the studied system) physical phenomena such as mechanical mixing of two fluid samples, temperature equilibration due to temperature gradient along the injection syringe, composition mismatches between syringe and cell solutions, electrical noise or temperature fluctuations, both in the room where the instrument is located, etc.

In summary, the generalization of the stoichiometric equilibrium equations as implemented in AFFINImeter allows an easy application of many models to the experimental ITC data, involving a maximum of three different compounds. Although in AFFINImeter all the stoichiometric equilibrium constants are considered as stepwise (and hence denoted as *K*), global stoichiometric binding constants can also be used provided all the reactions introduced by the user in the model builder (https://www.youtube.com/watch?v=dhfW-PQ-qVA) use the free species as unique reactants to form all the involved complexes. As stated at the beginning of this section, we only indicated here the equations corresponding to the global equilibrium constants because the equations based on stepwise constants depend on the specific reaction scheme or reaction pathway chosen for the studied system. However, these equations are easy to derive from the above given expressions (see previous section). The sequential equilibrium models implemented in AFFINImeter have been employed in a number of studies including proteins, DNA and other heterogeneous systems [13-17].

2.2.4.- Implementation of Independent Sites (IS) models

General site equilibrium constants were introduced in section 2.2.2. It was shown that when the sites are independent and equivalent, all the site equilibrium constants are identical. This is a useful model for the study of many multivalent interactions where these conditions are fulfilled. In this model, the equilibrium constant and the enthalpy change for the interaction can be obtained for each site, together with the number of binding sites, considered a fitting parameter. One-set-of-independent-sites and two-sets-of-independent-sites models are available in several programs for the analysis of ITC data. In these models, all binding sites within each set are independent and equivalent to each other. The sites of different sets are also independent, but they have different affinity and enthalpy of interaction. To our knowledge, the equations for competitive IS models, *i.e.* considering two competitive ligands for each site, are not publicly available yet. AFFINImeter implemented the generalization of IS models for multiple sets of independent and equivalent sites for each set, with two possible ligands (A and B) competing with different affinity and enthalpy of interaction for all the sites. Some publications taking advantage of these functionalities have been recently reviewed in [18]. The equations of this general model are described next. From a kinetic point of view:

$$v_a^{A_s} = k_a^{A_s} (n_s - \langle A \rangle_s - \langle B \rangle_s) \cdot [A] \qquad \text{and} \qquad v_d^{A_s} = k_d^{A_s} \langle A \rangle_s \tag{23}$$

where $v_a^{A_s}$ and $v_d^{A_s}$ are the rate of association and dissociation of the molecule A to the set s of a given macromolecule M, $k_a^{A_s}$ and $k_d^{A_s}$ are the kinetic constants for the same process, n_s is the total number of sites for the considered set, $\langle A \rangle_s$ and $\langle B \rangle_s$ are the average number of A and B molecules per mol of M in the set s, and [A] is the concentration of free A in the solution. Under equilibrium conditions the rates of adsorption and desorption match:

$$k_a^{A_s} (n_s - \langle A \rangle_s - \langle B \rangle_s) \cdot [A] = k_d^{A_s} \langle A \rangle_s \tag{24}$$

thus,

$$K_{A_s} \cdot [A] = \theta_{A_s} / (1 - \theta_{A_s} - \theta_{B_s}) \tag{25}$$

where the microscopic equilibrium constant for the binding of A to any site of the set s in M is defined as $K_{A_s} = k_a^{A_s} / k_d^{A_s}$ and the fraction of occupied sites are defined by $\theta_{A_s} = \langle A \rangle_s / n_s$ and $\theta_{B_s} = \langle B \rangle_s / n_s$. The same equations can be written for B:

$$k_a^{B_s} (n_s - \langle A \rangle_s - \langle B \rangle_s) \cdot [B] = k_d^{B_s} \langle B \rangle_s \tag{26}$$

$$K_{B_s} \cdot [B] = \theta_{B_s} / (1 - \theta_{A_s} - \theta_{B_s}) \tag{27}$$

Combining eqs (23) and (25) and solving for θ_{A_s} and θ_{B_s}:

$$\theta_{A_s} = K_{A_s} \cdot [A] \; / \; (1 + K_{A_s} \cdot [A] + K_{B_s} \cdot [B]) \tag{28}$$

$$\theta_{B_s} = K_{B_s} \cdot [B] \; / \; (1 + K_{A_s} \cdot [A] + K_{B_s} \cdot [B]) \tag{29}$$

Considering that the sites are completely independent regardless of whether they belong to the same or to different sets. They do compete with each other to capture the molecules of

the ligands A and B. The total concentration of each compound is distributed between the different sets as:

$$[M]_T = [M] + [M]_T \cdot \sum_s n_s \cdot \left(\theta_{A_s} + \theta_{B_s}\right) \quad (30)$$

$$[A]_T = [A] + [M]_T \cdot \sum_s n_s \cdot \theta_{A_s} \quad (31)$$

$$[B]_T = [B] + [M]_T \cdot \sum_s n_s \cdot \theta_{B_s} \quad (32)$$

Note that, in contrast to the stoichiometric equilibrium scheme, the equations for M are different from those for A and B. This is because the sites are assumed to be located in M and the model is not symmetrical in this sense. Thus, the application of this model using AFFINImeter requires the presence of at least a macromolecule M in the sample cell and a ligand A in the syringe.

Again, the model can be applied to any property that is sensitive to the fraction of occupied sites for the different sets. For the specific case of ITC, the equation equivalent to eq. 19 in this interaction scheme is:

$$\Delta Q_i = \{(Q_{i+1} - Q_i) + \Delta Q_{corr,i} + \Delta H_{dil} \cdot ([A_{syr}] - [A]_{T,i}) + Q_{db}\} / ([A_{syr}] \cdot v_{inj,i}) \quad (33)$$

The term $Q_{syr,i}$ was removed since in its current version for this model the macromolecule M cannot be in the syringe and the model does not consider the possibility of binding between A and B or their self-association. All the terms in equation 33 are identical to those in eq. 19 except Q_i and Q_{i+1}, which are given by:

$$Q_i = [M]_{T,i} \cdot V_i \cdot \left\{\sum_s n_s \cdot \left[\left(\theta_{A_s}\right)_i \cdot \Delta H_{A_s} + \left(\theta_{B_s}\right)_i \cdot \Delta H_{B_s}\right]\right\} \quad (34)$$

This model is well suited mainly for the study of molecular interactions with macromolecules with multiple binding sites under the assumption that they are independent to each other. Notably, with our implementation, it is possible to study competitive binding in this kind of systems, and to consider also multiple sets of binding sites. Actually, an IS model is defined by the number of sets available for binding. For each new set there are, in principle, three new parameters (the number of sites, the site equilibrium constant and the associated enthalpy). As in the case of stoichiometric equilibrium models, the dilution enthalpy and Q_{db} could also be globally fitted. Some examples of the application of the IS family of models using AFFINImeter can be found in references [19-21].

2.2.5.- Concentration correction

The determination of equilibrium constants is based on the concentrations of the different chemical species present in the sample. Such concentrations are calculated from the nominal concentrations of the participating compounds. Mismatch between nominal and true active concentrations (due to *i.e.* chemical degradation, aggregation or conformational changes that prevent the binding of a fraction of the reactant present) have a serious impact in the results. In order to quantitatively consider these events, we have introduced a correction factor for each compound: r_M, r_A and r_B; that are equivalent to the ratio between active concentration and nominal concentration of M, A or B, respectively. The introduction in the equations is

simple since the concentration of each compound is replaced by its product with the corresponding factor. The only additional consideration is the determination of ΔQ_i when r_A is different from unity. By convention, ΔQ_i is the heat absorbed or delivered per mole of A added during the injection i. The values of the experimental binding isotherm are calculated using the concentration of nominal A, without any correction. Accordingly, the value of A_{syr} in the denominator of equations 19 and 33 is not corrected, allowing the comparison between the fitted and the experimental values.

Thus the equations 19 and 33 corresponding to the SE and IS models, respectively, became:

$$\Delta Q_i = \{-Q_{syr,i} + (Q_{i+1} - Q_i) + Q_{corr,i} + \Delta H_{dil}\cdot([A_{syr}] - [A]_{T,i}) + Q_{db}\} / ([A_{syr}]^{nominal} \cdot v_{inj,i}) \quad (35)$$

$$\Delta Q_i = \{(Q_{i+1} - Q_i) + \Delta Q_{corr,i} + \Delta H_{dil}\cdot([A_{syr}] - [A_T]_i) + Q_{db}\} / ([A_{syr}]^{nominal} \cdot v_{inj,i}) \quad (36)$$

were $[A_{syr}]^{nominal}$ is the nominal, non-corrected concentration of A.

2.2.6.- Non-fitted parameters and hybrid SE/IS models

In the previous sections, we have described in detail how the SE and the IS models are implemented in AFFINImeter. For SE models we have 2 fitting parameters for each chemical reaction while for IS models there are 3 additional fitting parameters *per* set of sites. Additionally the ΔH_{dil}, Q_{db} and the r_M, r_A and r_B values can be fitted for each equilibrium isotherm. Often, it is convenient to keep constant some of these parameters, to avoid overparameterization while enough degrees of freedom in the fitting should be taken to make it valuable. For instance, if the concentrations are well known, the three correction factors introduced for this aim must be set to 1 and not allowed to vary. It is also experimentally found that in most cases, the "external" heat contributions considered by Q_{db} can be dismissed, and sometimes the ΔH_{dil} is not significant and can be neglected. This can be easily tested by performing the fitting allowing their variation or keeping them fixed. Additionally, the value of any other parameter might be known from other experimental or theoretical source, and thus be introduced as a fixed parameter. On the other hand, AFFINImeter allows mathematical and logical relationships between parameters. This is a powerful functionality for parameters that we wish to restraint as a combination of values of other parameters, *i.e*. they are not fitted nor kept constants but dynamically changed during the iterative minimization. This kind of dynamical restrictions can also be set for the upper and lower bounds of parameters that are fitted. By using such relationships it is possible to introduce hybrid models in AFFINImeter, i.e. models that combine independent and stoichiometric binding sites. For instance, we could have a macromolecule with four binding sites where only two of them are independent and equivalent. If the non-independent sites cannot be occupied before the saturation of the independent ones, we could use a SE model with only two parameters for the first two sites, by using eq. 12 to connect the stepwise stoichiometric binding constants and matching the enthalpies for the first two reactions. The dynamical relationships between fitting parameters can be employed to introduce a large variety of restrictions that minimize the number of fitting parameters. Such functionality is usually employed when specific experimental information is available to connect the parameters on a rational basis. Another application would be the fitting of data corresponding to the same

system at different temperatures, by using the van't Hoff equation to connect the equilibrium constants (see section 3).

2.3.- Minimization procedure

For ITC data, the fitting parameters for each model are numerically determined by minimizing the objective function:

$$\chi^2 = \frac{1}{p} \sum_1^p \frac{\left(\Delta Q_i^{exp} - \Delta Q_i\right)^2}{\sigma_i^2} \quad (37)$$

where p is the total number of points, (ΔQ_i^{exp}) the values of the experimental equilibrium isotherm, ΔQ_i is given by eqs 19 or 33 (depending on the model), and σ_i is the uncertainty of each point (see [10] and section 2.1.). In order to optimize the quality and the speed of convergence, the Simulated Annealing and Levenberg-Marquardt algorithms are sequentially used [22]. In addition, for each combination of proposed fitting parameters during the iterative process, a damped Newton-Raphson algorithm using the Armijo rule [23] was employed to obtain the concentration of free M, A and B compounds from equations 16-18 or 30-32. This latter calculation is individually performed for every point in the equilibrium isotherm, although global controls, such as backpropagation fittings based on the homogeneous increasing/decreasing of the free concentrations as the total concentration of M, A and B change, were introduced in the code to guarantee the consistency of the results.

From eq. 37 it is clear that the lower the χ^2 value the better the fitting. Ideally, for a combination of reasonably good experimental measurements, fair estimation of uncertainties, and right selection of the interaction model, the χ^2 value should approach the unity, which means that the differences between the ΔQ_i^{exp} and the ΔQ_i are essentially accounted for by the uncertainties σ_i. Actually, the estimation of the uncertainties is not a trivial task. Thus, good fittings with χ^2 significantly larger than 1 are expected if the σ values are unrealistically low and vice versa (equation 37). Although in general the uncertainties in AFFINImeter seem to be fairly determined, we have detected some trend towards underestimated uncertainty values for each individual titration point (used as weighting factors in equation 37).

In theory, if the uncertainties are well estimated, the χ^2 distribution gives us the probability of obtaining a value of the χ^2 above, or below, some threshold. However, in practice, the value of χ^2 may be difficult to translate into an informative measure of the quality of the fitting that guides the user in the selection of the right model to analyze the data. In general, χ^2 values around unity correspond to good fittings, but the absolute value of this parameter depends too much on the uncertainties of the points and its value is not bounded, so it is difficult to say whether or not a value of χ^2 around 5, 10 or 20 is acceptable. So we have implemented a Goodness of Fit (GoF) estimation, defined as the probability of finding the fitted value of each experimental point within a normal distribution with half-width equal to the corresponding uncertainty. This parameter is basically a normalization of χ^2 expressed in percentage units. The GoF of a given curve is determined as the mean GoF of all its points. Ideally, the GoF value of a perfect fit would be 100%. The advantage of this descriptor is that it is bounded between 0 and 100, so it is easy to interpret. Nevertheless, one should not forget that it still depends

too much on the uncertainties of the experimental points. Improvements of this parameter to solve this problem are expected to be included in AFFINImeter in the future.

2.4.- Seeds and local minima

Non-trivial models may have a large number of fitting parameters, increasing the probability of reaching and being stuck at local minima. In order to identify the presence of such results, we have implemented the possibility of repeating the fittings starting from different seeds/initial values for the parameters. By default, 20 different independent minimizations are performed using 20 different sets of random seeds. The number of independent sets can be changed by the user. Unless the user modifies the default options, the seeds are always randomly chosen between the lower and upper bounds allowed for each fitting parameter. By default, constant values are provided for such bounds ([1, 10^8] M^{-1} for the equilibrium constants and [-10^5, 10^5] J/mol for the enthalpies). The results for each independent fitting are shown at the bottom of the results' window, and they can be downloaded in csv format for further treatment. Several minima with low χ^2 values different from each other likely indicate that the model is overparameterized. Ideally, all the minimizations should converge to the same result regardless of the employed initial value, but in some cases local minima can be identified. This can be used as a consistency test. A fitting where the lowest χ^2 only appears once is an evidence of convergence problems. When this happens, the number of independent minimizations should be increased. A maximum of 1000 repeats is allowed in AFFINImeter. Obviously, the larger the number of repeats, the longer the time taken to achieve the results.

2.5.- Uncertainties

Uncertainties are part of the results in any quantitative scientific experiment. We have made an effort to provide reliable estimation of uncertainties for the fitting parameters by determining them in two different ways, named in AFFINImeter as standard and statistical errors. The standard errors are determined as the square root of the diagonal elements of the variance-covariance matrix, obtained as the inverse of the Hessian matrix (the matrix of the second partial derivatives $\partial^2\chi^2/\partial p_i \partial p_j$ of the objective function with respect to the fitting parameters). These uncertainties are determined on the flight, since the Hessian matrix is already required for the Levenberg-Marquardt method used for the minimization; therefore no expensive extra calculations are required. Since our implementation of the models is fully general, as described in the previous sections, we do not use explicit analytical derivatives and the elements of the Hessian matrix are numerically obtained. The uncertainties determined in this way are less reliable than those obtained from analytical derivatives. Thus, we decided to implement a non-parametric bootstrap method to obtain what we call in AFFINImeter statistical errors: the Jackknife resampling method [24,25]. These uncertainties are more reliable, since they are determined from the standard deviations of the parameters obtained by repeating the fitting as many times as experimental points are available in the target experiment, each time ignoring one point in these extra fittings. For instance, in an equilibrium isotherm consisting of 30 points, all of them are employed in the main minimization process. The resulting parameters are used as seeds for 30 independent fittings (of 29 points each) where the data points are sequentially removed, one at a time. In this case, AFFINImeter gets 30 independent solutions for each parameter (in addition to the one provided by the main

fitting). Then, the standard deviation for each parameter, using the values obtained in these 30 independent fittings, is obtained and taken as “statistical uncertainty”. The removal of the points with a high weight in the fittings induces a significant change in the parameters and so increases the standard deviations. By default, the statistical uncertainties are activated in AFFINImeter, but the user can also access to the standard errors by clicking on the corresponding option.

2.6.- Global analysis

In the previous sections we have described how AFFINImeter works to fit the parameters of a model to the molar enthalpies obtained from an ITC experiment. Note that a single experiment in any technique is rarely conclusive. One always needs to perform several repeats of the same assay and to fit all the results either together or individually using the same or complementary models. AFFINImeter allows global fitting of multiple experiments for this aim.

Further, in cases of competitive binding studies of two different ligands A and B to a given macromolecule M, several experiments can be designed to study the interactions between these molecules: the titration of A or B separately into a solution of M, the titration of both A and B together into the same solution, the titration of A into a solution of M+B, and the titration of B into a solution of M+A. Additionally this can be done at different concentrations of each compound. AFFINImeter allows analyzing all these experiments together, by optionally sharing the fitting parameters. Other example is the case where the same system is studied at several temperatures. In that case, the equilibrium constants can be connected to each other by using the van’t Hoff equation and the enthalpies can also be restrained by assuming a linear or quadratic dependence involving the heat capacity change upon binding (see section 3). Similar situations can be proposed for many other interactions. The results obtained in this way are expected to be much more robust than fitting individual experiments separately.

2.7.- Fitting results, simulation tool and report

In the previous sections we have explained how the SE and IS models have been implemented in AFFINImeter, including parameters to consider the ratio or percentage of active compounds in the solution, corrections for volume displacement, external contributions to the heat signal, and dilution of the injected compound. Those equations can be used to fit experiments as well as to perform simulations. The simulations are useful in order to design experiments –if the user has an estimation of the expected parameters–, to understand the sensitivity of the isotherm to the different parameters, and also to understand better the interaction mechanisms. As explained above, the evaluation of the equations 19 and 33 require the determination of the concentration of free species, as well as the concentration of every complex present in the solution (or fraction of occupied sites in the case of IS models). This allows determining the contribution of each chemical species to the total molar heat. The combination of the three plots (total molar heat, species distribution, and different contributions to the total signal, together with the value of the model parameters) indicate which stoichiometry or binding site dominates at specific concentration ranges, and it is useful both for optimizing new experiments and to understand complex interaction mechanisms. An example is shown in Figure 5. All this information is provided in csv (numerical data for further

external treatment), html or pdf formats, and can be freely shared by e-mail from the AFFINImeter software.

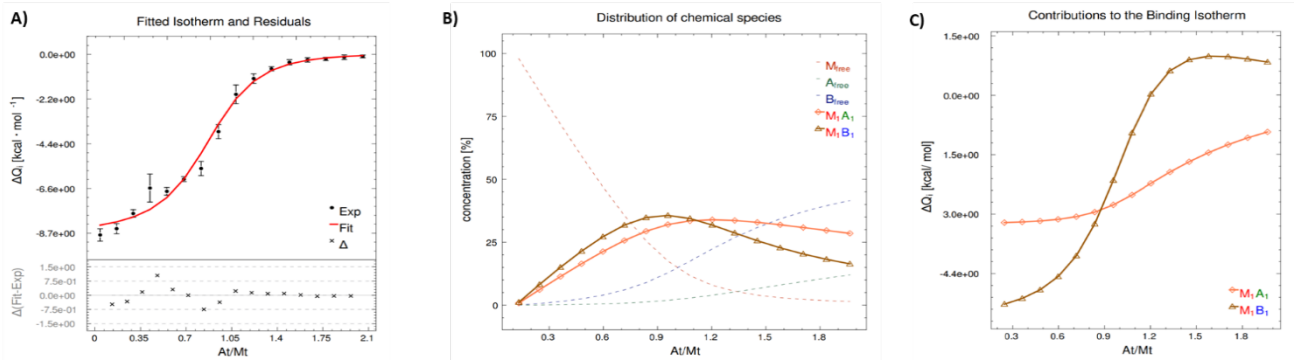


Figure 5. Example of plots generated in AFFINImeter after data analysis. A) Binding isotherm and fitting result; B) Species distribution plot; C) Contribution of complexes formation to the overall signal. In this example the data analyzed were for a titration where a receptor M is titrated with a mixture of ligands in syringe (A + B). The competitive model MA ↔ FS ↔ MB was used.

2.8.- The evaluation of kinetic constants using kinITC

ITC raw data have intrinsically kinetic information, as they are heat power as a function of time. Traditionally, the perturbations in this signal caused by the injection of small aliquots of one titrant solution into a titrate solution are integrated thus leading to an energy (heat) profile as a function of the stoichiometric ratio. Such processed data can be employed for the application of thermodynamic models, thus contributing to elucidate interaction mechanisms as well as the entropic and enthalpic contributions to the Gibbs energy change corresponding to the formation of the different chemical species. This classical treatment, however, results in the loss of useful kinetic information. Although the possibility of obtaining kinetic information from ITC experiments has long been known, strangely it was not at all used in biology and even considered as impossible by many users. A few years ago, the ITC community was re-awakened to this possibility [8] the method was named kinITC. A simplification of this method that considers the variation of the time of return to equilibrium at each injection [26] has been implemented into AFFINImeter [9]. It is worth mentioning that ordinary titration data with reasonable values of the Wiseman parameter are often amenable to this analysis. Thus, even old ITC data can be directly reanalysed to extract potential kinetic information. In a recent review, a compilation of kinITC studies using AFFINImeter has been summarized [18]. Also a paper was recently published by Zihlmann et al. [27], who retrieved the kinetic constants of 29 mannosides binding to the bacterial adhesion FimH using the kinITC method implemented in AFFINImeter and performed a thoughtful validation of results with SPR data. The kinITC method provides results that are in excellent agreement with those obtained by SPR for a wide range of Kd values.

2.9.- Accessibility and management of lab projects

The default version of AFFINImeter is a cloud software with user accounts protected by password. This means that it can be accessed from the web browser of any device with internet connection. The calculations required to fit experimental data often require a significant computational cost that is performed using the AFFINImeter servers. Local installations of AFFINImeter are also available upon request. Once the user access to his/her account, all the past projects are available and they can be read or edited. They can also be organized in folders and the results of complete fitting or simulation projects can be shared with other people by introducing the corresponding e-mail address.

3.- Example of how to use AFFINImeter to analyze complex interactions with multiple ITC measurements

Although AFFINImeter has been made public recently, many research groups from academia, research institutes and pharmaceutical companies have already benefited from the advanced tools of analysis that the software offers. Indeed, a number of research papers have been published that manifest the value of AFFINImeter for the robust thermodynamic and kinetic analysis of ITC data. Referencing all these works is out of the scope of this paper but a rather comprehensive review is provided in a recent literature review article [18]. A book chapter has also been recently published with a detailed analysis on a number of heparin derivatives [29]. As illustrative examples we would like to highlight the work performed by Shiu-Hin Chan et al. [17] and the work performed by Marques-Carvalho et al [28] where advanced stoichiometric models were used in the thermodynamic and mechanistic characterization of complex interactions. Besides, the excellent work performed by Zihlmann et al. focused on testing the reliability of KinITC-ETC to retrieve kinetic information [27].

In this section we will show that AFFINImeter is well suited to analyze complex intermolecular mechanisms, without the need of developing and validating new code, thus saving time and effort. To this end, we show now a global multitemperature analysis –using the van't Hoff equation to connect the equilibrium constants and the change in heat capacity to linearly constrain the corresponding enthalpies at different temperatures– for a set of measurements of a system exhibiting cooperative 1:2 interactions (a similar analysis has been published in [8]). First we provide a description of the system, which is required to understand the interaction mechanism, and then the analysis method using AFFINImeter is explained in detail. Finally, the obtained results following such methodology are shown.

3.1.- Description of the system and analysis performed in the literature

NQO1 NAD(P)H (NADH, nicotinamide adenine dinucleotide) quinone dehydrogenase 1 (EC 1.6.5.2) is a cytosolic FAD-binding (FAD, flavin adenine dinucleotide) homodimeric protein responsible for the two-electron reduction of quinones to hydroquinones, with preference for short-chain acceptor quinones, thus preventing the one-electron reduction of radical species, intervening in detoxification processes and activation of antitumor quinone-related prodrugs [30]. NQO1 stabilizes and protects tumor suppressor p53 from degradation. Thus, decreased expression of NQO1 is associated with decreased p53 stability, as well as chemoresistance stemming from diminished prodrug bioactivation [30]. P187S is a widespread polymorphism in NQO1 associated with different types of cancer and low efficacy of anticancer drugs. P187S causes lowered structural stability and activity. NQO1-FAD is an excellent biological system to

study 1:2 protein:ligand interaction: 1) NQO1 is able to bind two FAD molecules per dimer; 2) the binding of the two FAD molecules is cooperative (homotropic interaction); 3) FAD binding and FAD homotropy are highly temperature dependent; and 4) mutations or polymorphisms in NQO1, in particular P187S associated with cancer susceptibility, result in considerable distortion of the cooperativity of binding. Therefore, NQO1-FAD interaction is an appropriate system to perform a global analysis of calorimetric titrations conducted at different temperatures applying a 1:2 model with homotropic binding cooperativity and variable temperature (considering the extended van't Hoff equation with non-zero binding heat capacity) in order to estimate the intrinsic FAD binding parameters and the cooperative FAD binding parameters. At particular temperatures each variant (wild-type and P187S NQO1) shows an apparent non-cooperative binding behaviour (because of the temperature-dependency of the cooperativity binding parameters); therefore, in biological systems with stoichiometry higher than 1 it is always advisable to perform titrations at different temperatures in order to confirm whether or not the binding is cooperative. It was hypothesized that the apparent binding cooperativity features observed in NQO1-FAD titrations were in fact the manifestation of a significant percentage of aggregated protein in the calorimetric cell, and, consequently, reverse titrations (FAD in cell and NQO1 in syringe) were more appropriate. However, it was later demonstrated that there is true binding cooperativity in NQO1-FAD interaction, that NQO1 potential aggregation has no influence on the titration features (different from reducing the fraction of active binding-competent protein), and that direct and reverse titrations provided a common set of estimated intrinsic and cooperative parameters [30] applying the same 1:2 model with binding cooperativity.

3.2.- Analysis with AFFINImeter

The calorimetric titrations for FAD binding to wild-type and P187S NQO1 were reanalysed using AFFINImeter. We used the model builder tool (https://www.youtube.com/watch?v=dhfW-PQ-qVA) to apply the model described in the original reference [30] (Fig. 6).

Figure 6. Equilibrium equation describing the interaction model applied for FAD binding to wild-type and P187S NQO1. Dummy reactions were also introduced in the model builder tool in order to account for the global parameters employed in the model (using the "enthalpy" as a global parameter and setting the corresponding equilibrium constant to zero).

The binding constants corresponding to the formation of the complexes with stoichiometries 1:1 and 1:2 were connected by a cooperativity constant. The model assumes that there are two identical binding sites with cooperativity. There are two sets of binding parameters: K_a, ΔH, and ΔC_P are the association constant, the binding enthalpy, and the binding heat capacity for each the two identical binding sites; α, Δh, and Δc_P are the cooperativity interaction constant, the cooperativity binding enthalpy, and the cooperativity binding heat capacity for the binding of the second ligand. Thus, the first ligand binds with binding parameters: K_a, ΔH, and ΔC_P; and the second ligand binds with binding parameters: $K_a\alpha$, $\Delta H+\Delta h$, and $\Delta C_P+\Delta c_P$.

Because experiments were performed at different temperatures, the van't Hoff relationship was employed to establish the temperature dependency of the equilibrium constants (K_a and α) as a function of the enthalpies and heat capacities (ΔH, ΔC_P, Δh, and Δc_P), and the enthalpies were assumed to vary linearly with temperature. Thus, the estimated parameters are those at 25°C.

The equations describing the temperature dependence of the enthalpy, overall equilibrium constants and cooperativity constant can be found in the original reference [30]. The introduction of those equations into AFFINImeter was performed by using the following procedure: i) using the model builder, a dummy reaction was added to the model for each global parameter (α, ΔC_p, Δh and Δc_p); ii) for each dummy reaction (let's say the formation of A_2 from free species) we set an equilibrium constant equal to zero and we use the corresponding "enthalpy" as a global parameter that can be introduced in the equations to connect the equilibrium constants or the enthalpies. Using this procedure, the variety of models that can be applied with AFFINImeter is significantly expanded since any kind of global parameters and restrictions between the actual chemical reactions describing the system can be introduced, including heat capacity changes and cooperativity constants. In particular, the equations describing the model applied in [30] are easily implemented in the model builder. The results of the analysis using AFFINImeter are shown in Figures 7-8 and in Table 1. It can be seen that the global fitting of the two sets of data is good and that the parameters compare well with those published in the main paper (Table 1). Differences can be ascribed mainly to the fact that the analysis performed with AFFINImeter is done by weighting the uncertainties of the molar heat values (see Figures 7 and 8). It is worth to comment that the analysis applied here is not directly implemented in any public software. For its application in the original paper [30] it was necessary to develop specific computational code. In AFFINImeter, the model is easily constructed using the model builder tool. The introduction of the global fitting parameters connecting equilibrium constants and enthalpies is done by adding dummy reactions in the model builder, with the corresponding mathematical relationships in the "fit setting" form.

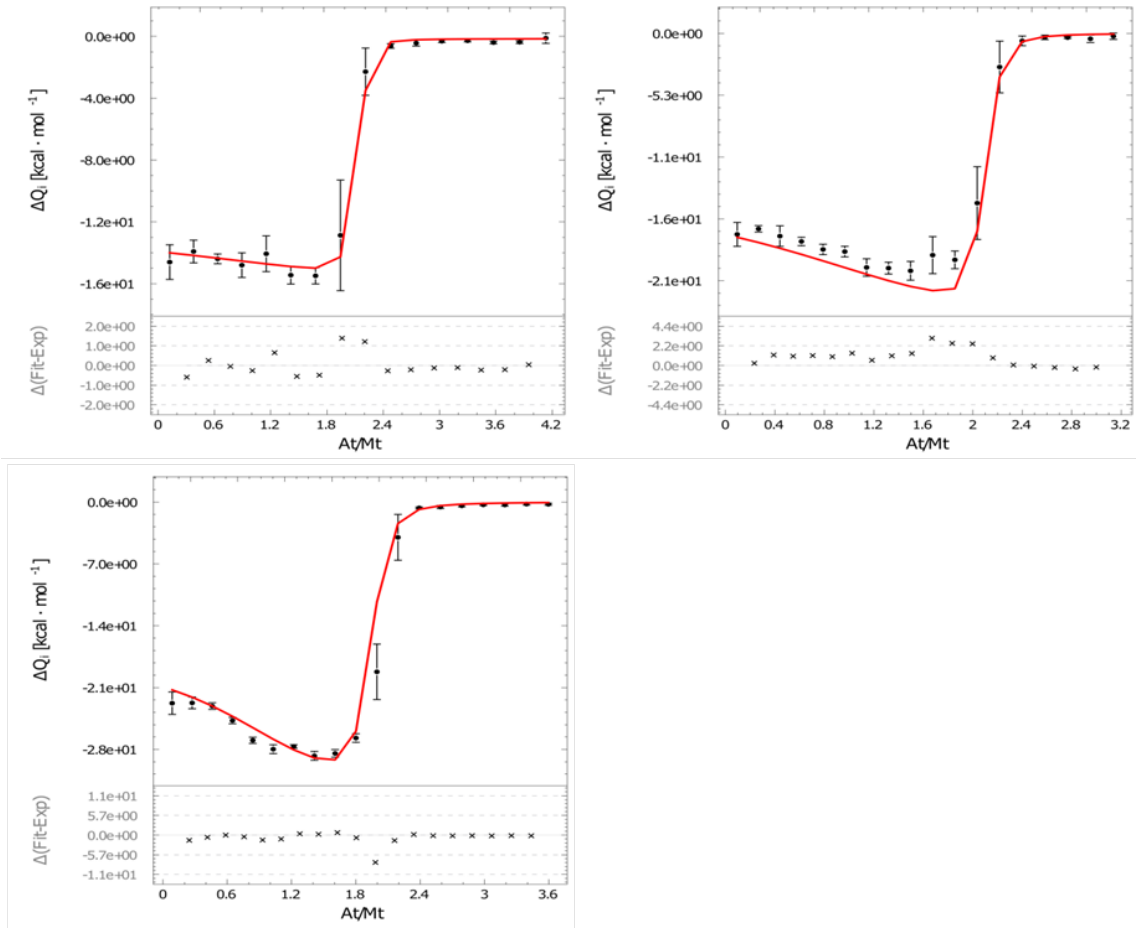


Figure 7. Experimental binding isotherms obtained for the interaction between FAD and the wild type NQO1 at 288.15, 293.15, and 298.15 K, together with the curves obtained by a global fitting of the three experiments together to the model described in [30] using AFFINImeter, as explained in the text.

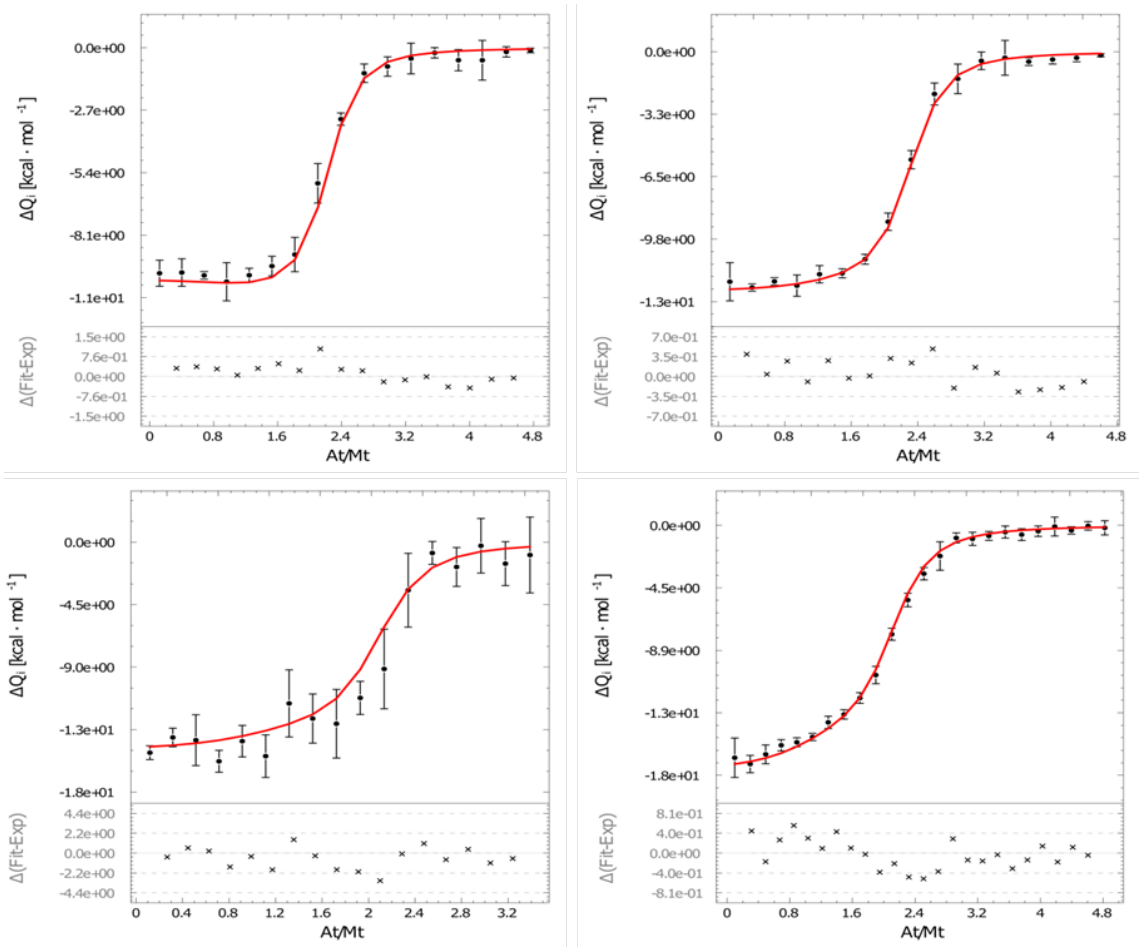


Figure 8. Experimental binding isotherms obtained for the interaction between FAD and the P187S NQO1 mutant at 288.15, 292.15, 295.15 and 298.15 K, together with the curves obtained by a global fitting of the four experiments together to the model described in [30] using AFFINImeter, as explained in the text.

Table 2.- Values of the parameters obtained from the fitting of the data shown in Figures 7 and 8 to the model described in [30] using AFFINImeter, as described in the text.

		K_a (10^7) (M^{-1})	ΔH (kcal mol^{-1})	ΔC_p (kcal mol^{-1} K^{-1})	α	Δh (kcal mol^{-1})	Δc_p (kcal mol^{-1} K^{-1})
WT	AFFINImeter	5.8 ± 3.7	-20.69 ± 0.12	-0.699 ± 0.033	0.585 ± 0.092	-11.38 ± 0.32	-0.98 ± 0.23
	Ref. [30]	8.8 ± 0.9	-22.2 ± 0.5	-0. 8 ± 0.1	0.53 ± 0.04	-12.0 ± 0.6	-1.1 ± 0.2
P187S	AFFINImeter	1.13 ± 0.48	-17.22 ± 0.35	-0.798 ± 0.047	0.35 ± 0.17	2.40 ± 0.87	0.34 ± 0.14
	Ref. [30]	0.81 ± 0.09	-17.6 ± 0.5	-0.9 ± 0.1	0.62 ± 0.05	3.1 ± 0.4	0.6 ± 0.2

These binding parameters are related to those introduced in previous sections:

K_a is equal to K_{s1} and K_{s2} (two identical binding sites) and equal to $K_{11}/2$.

$K_a\alpha$ is equal to K_{s1s2} and K_{s2s1} and equal to $2K_{12}$

4.- Best-practice recommendations for analysis of molecular interactions

A number of recommendations, of global use are appropriate at this stage, to improve the quality of the retrieved results.

- **Raw data processing.**

  In a first step, AFFINImeter performs the automatic processing of the raw data and automatic thermodynamic and kinetic analysis using a 1:1 simple model. Although this automated process is already optimized, in some cases manual refinement could be convenient, especially when the data are of poor quality. For this, it is important to verify peak by peak the processed thermogram to assess that the processing (baseline correction, noise removal, peak integration and molar heat uncertainty) is adequate.

- **Generating and using tailored binding models.**

  When creating a new binding model, use the minimum amount of complex species that are expected to be present in the binding event, although including all the species that are assumed to be formed. The application of a wrong model is an important source of error in the analysis of intermolecular interactions, so the model to perform the final analysis should be reasonably and carefully validated.

- **Global analysis.**

  Performing global analysis is always recommended, as well as its comparison to individual fitting. Even for the characterization of a 1:1 interaction, where it is advisable to perform at least 2-3 repeats. For the characterization of complex binding systems with models involving more than one equilibrium, global analysis is typically a requirement in order to reduce the number of degrees of freedom in the fitting process. In these cases, it is recommended to perform several titration experiments using different experimental setups *i.e.* direct and reverse experiments, and/or modification of titrant and/or titrate concentrations. Global analysis of this dataset will give much more robust results and will prevent over-parameterized fittings. The determination of uncertainties for global analysis of several experiments is exactly the same as for the single experiments (see below).

- **Use or error bars.**

  AFFINImeter automatically includes the uncertainty associated to the integral calculation of each peak in the binding isotherm (error bars). By default, the software takes into account these uncertainties throughout the fitting to yield a more reliable result. However, the unweighted fitting (disregarding the error bars) can also be performed.

- **Use of the parameter Q_{dil} to account for the heat of dilution.**

  Blank experiments are useful in order to investigate self-association or aggregation in the titrant solution. Self-association and some kind of aggregation can be considered in AFFINImeter by stoichiometric equilibrium models. Non-specific dilution effects can be

modeled using the Q_{dil} parameter. In many cases, the use of this parameter is a very convenient option that can replace the blank subtraction operation (although it is always convenient to perform the blank experiment in order to test for possible extraneous heat effects and to have an experimental estimate of the dilution heat).

- **Validation tools.**
  - Visual inspection: use it to determine if differences between the experimental data and theoretical curve are random or systematic. In the second case, it can be an indication that the model employed is not appropriate.
  - GoF: use it to compare the goodness of fit between different analyses of the same isotherm with the same number of fitting parameters.
  - Uncertainty associated to a parameter value: very large uncertainties (larger than the nominal value itself) are a sign of over-parameterized fitting.
  - List of local minima: use it to check for potential over-parameterization.
  - Parameter saturation: check that the result obtained for all the fitting parameters are not equal to the maximum or minimum allowed value. If that is the case that parameter should be kept as a constant or, alternatively, modify the fitting range the parameter can access during the fitting process. By default, such range is fixed to several orders of magnitude for each fitting parameter but it can be manually modified by the user.
- **Statistical & standard errors**

  As explained in the main text, AFFINImeter provides both statistical and standard errors. By default, the statistical errors are activated, since we consider they are a more reliable estimation of the parameter uncertainties. However the user has also access to the standard errors. For highly involved models it could be convenient to ask for many independent fitting repeats (a maximum of 1000 repeats can be performed in AFFINImeter) to then perform a statistical study of the results (all the independent minima can be downloaded in csv format from the results window).

- **Visualization and fitting of ETC.**

  The ETC curve represents the duration of the different peaks throughout the titrations as a function of the concentration ratio ([A]/[M]). Typically the high concentration region of standard sigmoidal isotherms where saturation is reached ([A]/[M] significantly larger than 1) is noisy (low heat effects) and does not provide much kinetic information. Thus, by default, AFFINImeter sets only uses the concentration region where the ETC curve contains useful information. However all the data range is accessible by just activating the corresponding option below the ETC plot. The duration of the experiments in the ETC curve is automatically detected, but it can also be modified by the user.

5.- Perspectives

The current version of AFFINImeter for the analysis of ITC measurements offers an unlimited number of thermodynamic models that can be freely designed by the user. Although this software is relatively new and its potential is still far from being fully exploited, a number of groups have already took advantage of its original functionalities and analysis methods (see [27-30] and references therein). New experiments can now be designed, performed and properly analysed, leading to characterization of non-trivial interactions (see for instance those shown in the Appendix). In the present manuscript we aimed at describing in detail the theory behind the software in order to facilitate its correct use by the scientific community interested in the characterization of intermolecular interactions using experimental techniques, in the present case mainly ITC. Optimization of available instruments regarding automation, reduction of sample consumption, sensitivity, implementation of simultaneous channels for measurements under the same conditions (for reproducibility or statistics) or different one (such as concentrations or temperature, to be used in global analysis), would allow the community to even take more advantage of AFFINImeter. We foresee that there is also room for improvement and development in the software along the following lines: (i) the removal of unnecessary assumptions in the employed equations; (ii) the improvement of algorithms for better or/and faster convergence; (iii) the expansion of functionalities; and (iv) the inclusion of new applications to other experimental techniques. As an example of developments in progress, more rigorous equations for the analysis of ITC data [31] are being incorporated –these include exact concentration corrections due to the displacement of volume and exact equations for the effect of dilution in the molar heat resulting from the titration. One major problem that is solved by these exact equations is the erroneous asymptotic heat value obtained at large molar ratios [Mt]/[At]. The implementation of these results will allow obtaining equilibrium constant and enthalpic terms not affected by this up-to-now erroneous basic estimate. AFFINImeter is also continuously evolving by taking advantage of new algorithms and more powerful computational resources. The expansion of functionalities will be reflected mainly in a redesign of the software to facilitate the multi-temperature and global analysis, amongst a number of specific interaction models. Finally, the recent release of AFF-NMR and AFF-Spectroscopy shows that the strategy of the AFFINImeter´s Scientific and Development team is to develop a multi-technique software for the orthogonal analysis of intermolecular interactions within the same platform.

Acknowledgments

The authors thank the financial support of the Spanish Ministry of Economy and Competitiveness (projects MAT2015-71826-P to Á. P., BFU2016-78232-P to A.V.C. and CTQ2017-84354-P to L. G-R), Xunta de Galicia (projects AGRUP2015/11, GPC ED431B 2017/21, GR 2007/085, IN607C 2016/03 and Centro singular de investigación de Galicia accreditation 2016-2019, ED431G/09), and Fundação para a Ciência e Tecnologia (FCT-Portugal) (project UID/QUI/0081/2013 to M.B.). These research projects were partially supported by FEDER. Financial support from PAIP-FQ-UNAM (grant 5000-9018 to M.C.) and FEDER (COMPETE 2020) (project Norte-01-0145-FEDER-000028 to M. B.) is acknowledged. P. F. G. thanks the Spanish

Ministry of Economy and Competitiveness and the European Social Fund for his predoctoral research grant, reference BES-2016-076761. Facilities provided by the Galician Supercomputing Centre (CESGA) are also acknowledged.

References

[1] Surfactants and Interfacial Phenomena; Rosen MJ and Kunjappu JT - 2012, Wiley & Sons. ISBN: 978-0-470-54194-4

[2] Proteins in solution and at interfaces: methods and applications in biotechnology and materials science; Edited by Ruso JM and Piñeiro Á - 2013. Wiley & Sons. ISBN: 978-0-470-95251-1

[3] Binding constants. The measurement of Molecular Complex Stability. Connors KA - 1987, John Wiley & Sons. ISBN: 978-0471830832.

[4] Vega S, Abian O, Velazquez-Campoy A, A unified framework based on the binding polynomial for characterizing biological systems by isothermal titration calorimetry. Methods 2015, 76:99-115. DOI: 10.1016/j.ymeth.2014.09.010

[5] Velazquez-Campoy A, Allostery and cooperative interactions in proteins assessed by Isothermal Titration Calorimetry, Chapter 12, pp 223-246, in Biocalorimetry: foundations and contemporary approaches; Edited by Bastos M, Boca Raton, FL : CRC Press, Taylor & Francis Group, [2016] ©20 ©2015. ISBN: 9781482246667 148224666X. DOI: 10.1201/b20161-16

[6] Pang X and Zhou H-X. Rate Constants and Mechanisms of Protein–Ligand Binding. Annual Review of Biophysics, 2017, 46:105-130. DOI: 10.1146/annurev-biophys-070816-033639

[7] COST action CA15126, 'Between Atom and Cell; Integrating Molecular Biophysics Approaches for Biology and Health Care, Working Group 4, 'Optimization of data quality', https://arbre-mobieu.eu/

[8] Burnouf D, Ennifar E, Guedich S, Puffer B, Hoffmann G, Bec G, Disdier F, Baltzinger M, and Dumas P. kinITC: A New Method for Obtaining Joint Thermodynamic and Kinetic Data by Isothermal Titration Calorimetry. J. Am. Chem. Soc., 2012, 134 (1):559–565. DOI: 10.1021/ja209057d

[9] Dumas P, Ennifar E, Bec G, Piñeiro Á, Sabín J, Muñoz E, Rial J. Implementation of kinITC into AFFINImeter. Malvern Panalytical Application Note. 2015. https://www.malvernpanalytical.com/en/learn/knowledge-center/application-notes/AN150427kinItcAffiniMeter.html

[10] Muñoz E, Sabín J, Rial J, Pérez D, Ennifar E, Dumas P and Piñeiro Á, "Thermodynamic and kinetic analysis of isothermal titration calorimetry experiments by using kinITC in AFFINImeter" in "Microcalorimetry of Biological Molecules: Methods and Protocols". Edited by Eric Ennifar, Springer, series title: Methods Molecular Biology, vol 1964, ISBN:978-1-4939-9178-5.

[11] Dumas P, Ennifar E, Da Veiga C, Bec G, Palau W, Di Primo C, Piñeiro Á, Sabin J, Muñoz E, Rial J. Chapter Seven - Extending ITC to Kinetics with kinITC. Methods in Enzymology. 2016, 567: 157-180. DOI: 10.1016/bs.mie.2015.08.026.

[12] Keller S, Vargas C, Zhao H, Piszczek G, Brautigam CA, Shuck P, High-Precision Isothermal Titration Calorimetry with Automated Peak-Shape Analysis. Analytical Chemistry 2012, 84: 5066-5073. DOI: 10.1021/ac3007522

[13] Vigonsky E, Fish I, Livnat-Levanon N, Ovcharenko E, Ben-Tal N and Lewinson O, Metal binding spectrum and model structure of the Bacillus anthracis virulence determinant MntA. Metallomics, 2015,7(10):1407-1419. DOI: 10.1039/c5mt00100e

[14] Marques-Carvalho MJ and Morais-Cabral JH, Interaction of Calmodulin with the EAG1 Potassium Channel. Biophysical Journal. 2015, 108(2): 23A. DOI: 10.1016/j.bpj.2014.11.150

[15] Gómez-González B, Francisco V, Montecinos R, and García-Río L. Investigation of the binding modes of a positively charged pillar[5]arene: internal and external guest complexation. Org. Biomol. Chem., 2017,15: 911-919. DOI: 10.1039/C6OB02573K

[16] Rivera-Najera LY, Saab-Rincón G, Battaglia M, Amero C, Pulido NO, García-Hernández E, Solórzano RM, Reyes JL and Covarrubias AA. A Group 6 Late Embryogenesis Abundant Protein from Common Bean Is a Disordered Protein with Extended Helical Structure and Oligomer-forming Properties. J. Biol. Chem. 2014, 289, 31995-32009. DOI: 10.1074/jbc.M114.583369

[17] Chan DS-H, Seetoh W-G, McConnell BN, Matak-Vinković D, Thomas SE, Mendes V, Blaszczyk M, Coyne AG, Blundell TL and Abell C, Structural insights into the EthR–DNA interaction using native mass spectrometry. Chem. Commun., 2017,53: 3527-3530. DOI: 10.1039/C7CC00804J

[18] Muñoz E, and Piñeiro Á, AFFINImeter Software: from its Beginnings to Future Trends A Literature review. J. Appl. Bioanal., 2018, 4(4):123-138. DOI: http://dx.doi.org/10.17145/jab.18.017

[19] Silva T, Claro B, Silva BFB, Vale N, Gomes P, Gomes MS, Funari S, Teixeira J, Uhríková D, and Bastos M, Unravelling a Mechanism of Action for a Cecropin A-Melittin Hybrid Antimicrobial Peptide: The Induced Formation of Multilamellar Lipid Stacks. Langmuir, 2018, 34(5): 2158–2170. DOI: 10.1021/acs.langmuir.7b03639

[20] Dias R, Brás NF, Fernandes I, Pérez-Gregorio M, Mateus N, Freitas V, Molecular insights on the interaction and preventive potential of epigallocatechin-3-gallate in Celiac Disease. Int J Biol Macromol. 2018,112:1029-1037. DOI: 10.1016/j.ijbiomac.2018.02.055

[21] Soares S, García-Estévez I, Ferrer-Galego R, Brás NF, Brandão E, Silva M, Teixeira N, Fonseca F, Sousa SF, Ferreira-da-Silva F, Mateus N, de Freitas V. Study of human salivary proline-rich proteins interaction with food tannins. Food Chem. 2018, 243:175-185. DOI: 10.1016/j.foodchem.2017.09.063

[22] Numerical Recipes in C: The Art of Scientific Computing, Second Edition 2nd Edition. Press WH, Flannery BP, Teukolsky SA, Vetterling WT – 1988, Cambridge University Press. ISBN-10: 0521431085.

[23] Armijo L, Minimization of Functions having Lipschitz continuous first partial derivatives. Pacific. J. Mathematics. 1966, 16(1): 1-3. https://projecteuclid.org/euclid.pjm/1102995080.

[24] Tukey, JW, Bias and confidence in not quite large samples. The Annals of Mathematical Statistics. 1958, 29(2): 614. DOI: 10.1214/aoms/1177706647.

[25] Efrom B, and Stein C, The Jackknife estimate of variance. The Annals of Statistics 1981, 9(3): 586-596. https://projecteuclid.org/euclid.aos/1176345462

[26] Dumas P, Joining Thermodynamics and Kinetics by kinITC, Chapter 15, pp 281-300, in Biocalorimetry: foundations and contemporary approaches; Edited by Bastos M, Boca Raton, FL : CRC Press, Taylor & Francis Group, [2016] ©20 ©2015. ISBN: 9781482246667 148224666X. DOI: 10.1201/b20161-19

[27] Zihlmann P, Slbermann M, Sharpe T, Jiang X, Mühlethaler T, Jakob R, Rabbani S, Sager C, Frei P, Maier T, Ernst B, KinITC - One method supports both thermodynamic and kinetic SARs as exemplified on FimH antagonists. Chem-Eur J. 2018, 24(49):13049-13057. DOI: 10.1002/chem.201802599

[28] Marques-Carvalho M, Oppermann J, Muñoz E, Fernandes A, Gabant G, Cadene M, Heinemann S, Schönherr R, and Morais-Cabral J, Molecular Insights into the Mechanism of Calmodulin Inhibition of the EAG1 Potassium Channel. Structure 2016, 24(10), 1742-1754. DOI: 10.1016/j.str.2016.07.020

[29] Muñoz E, Sabín J, Piñeiro Á, Franco Rodríguez G, The Use of ITC and the Software AFFINImeter for the Quantification of the Anticoagulant Pentasaccharide in Low Molecular Weight Heparin" in "Microcalorimetry of Biological Molecules, methods and protocols". Edited by Eric Ennifar, Springer, series title: Methods Molecular Biology, vol 1964, ISBN:978-1-4939-9178-5.

[30] Clavería-Gimeno R, Velazquez-Campoy A, Pey AL, Thermodynamics of cooperative binding of FAD to human NQO1: Implications to understanding cofactor-dependent function and stability of the flavoproteome. Archives of Biochemistry and Biophysics. 2017, 636:17-27. DOI: 10.1016/j.abb.2017.10.020

[31] Dumas P, Rigorous equations for Isothermal Titration Calorimetry: theoretical and practical consequences. https://doi.org/10.1101/512780

Appendix : The use of the AFFINImeter simulation tool.

Using two examples, we illustrate the employment of the AFFINImeter simulation tool to determine the binding isotherms of the possible ITC experiments and the corresponding species distributions.

First we considered a system consisting of a ligand **L** with a protein **P** that is already bound to another ligand **L'**. The first ligand **L** is assumed to self-associate forming dimers, but only the monomeric species can interact with the protein. In the notation employed in AFFINImeter, the main compounds in the syringe and in the cell are A and M, respectively. An additional compound B can be in the sample cell, in the syringe or in both. The chosen system can be studied by using the protein as titrant or as a titrate. This system will be simulated using an equilibrium stoichiometry (SE) model. Thus, using the model builder tool (https://www.youtube.com/watch?v=dhfW-PQ-qVA), the reaction schemes required to simulate both the direct and the reverse experiments are shown in Figure I.1. In one case the protein and **L'** are in the syringe at the same concentration forming 1:1 complexes, while **L** is in the sample cell with both monomers and dimers in equilibrium. When the protein is titrated to the cell, the **PL'** complex and the $\mathbf{L_2}$ dimers dissociate to form the **PL** complex and free **L'** molecules. This occurs because the equilibrium constant corresponding to the formation of **PL** has been chosen to be much higher than those corresponding to the formation of **PL'** and $\mathbf{L_2}$ (see Table I.1.). Once all the **L** molecules reacted with **P**, the **PL'** complexes remain formed upon injection because there are no free **L** molecules to compete with **L'** for the interaction with **P** (Figure I.1, top). In the second case, i.e. when the dimer $\mathbf{L_2}$ is injected from the syringe, it dissociates forming monomers that then interact with **P**. Again, **L** competes with **L'** for the interaction with **P**. Once the protein is saturated with **L**, all the **L'** molecules are free in the cell. Beyond that injection, **L** remains in dimeric form when injected from the syringe (Figure I.1, bottom).

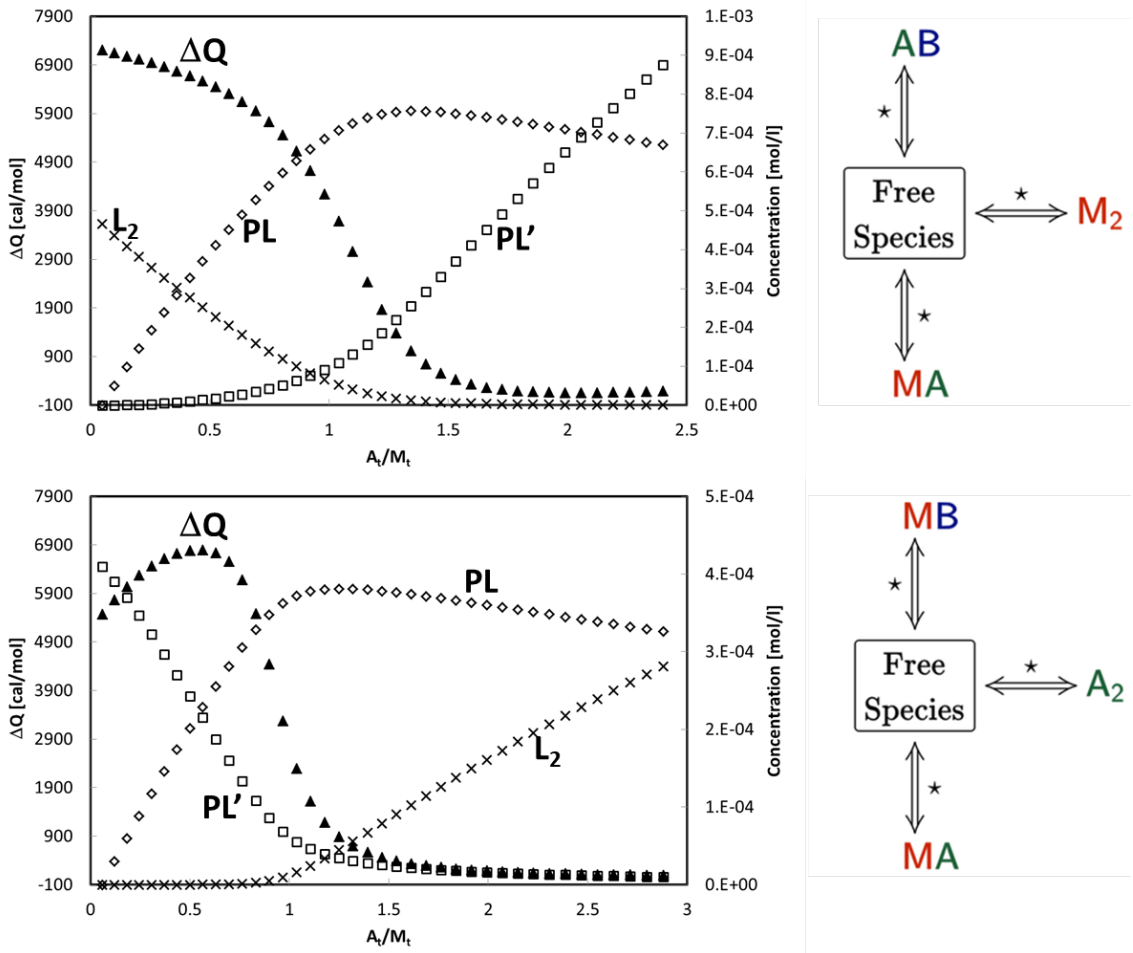


Figure I.1. Simulated binding isotherms (▲; left axes) for the reaction schemes showed on the right of each plot, together with the concentration in the sample cell of the different chemical species (see labels in the plots; right axes). The two reaction schemes describe exactly the same molecular interactions for the direct and reverse experiments (see main text). The concentrations employed for these simulation are: {[M] = [**L**] = 1 mM; [A_{syr}] = [**P**] = 5 mM; [B_{syr}] = [**L'**] = 5 mM} for the top plot and {[M] = [**P**] = 0.5 mM; [A_{syr}] = [**L**] = 3 mM; [B_{cell}] = [**L'**] = 0.5 mM} for the bottom plot. The employed thermodynamic parameters are in Table I.1. In all cases 40 injections of 2 μl each in a sample cell of 0.2 ml were performed.

Table I.1. Thermodynamic parameters employed for the simulation of the curves shown in Figure I.1.

	K	ΔH
FS↔ L_2	10^5	-10^3
FS↔PL	10^7	$3 \cdot 10^3$
FS↔PL'	$5 \cdot 10^4$	$-4 \cdot 10^3$

Clearly, the simulation tool is useful to understand the behaviour of a molecular system for a given set of thermodynamic parameters. Additionally, it is also useful to design the ITC experiments. In this example, the fitting of the model to the experiment corresponding to the first simulation (Figure I.1, top) will probably be overparameterized since it requires at least 6

parameters to describe a smooth curve. However, the binding isotherm corresponding to the second simulation (Figure I.1, bottom) present stronger curvature changes and more parameters would be statistically justified. Of course, if the two experiments are performed the global fitting of both binding isotherms would provide more reliable values for the parameters.

The second simulation corresponds to the interaction of a hypothetical nanoparticle (**NP**) initially located in the sample cell at 0.5 mM with two competitor ligands initially at 40 mM in the syringe (**A**) and at 5 mM or 10 mM in the sample cell (**B**). **NP** is assumed to have two different interaction regions with 20 and 5 sites, respectively. It is also assumed that all the interaction sites in both regions are independent from each other. This system is well suited for an IS model. The employed site affinity constants and enthalpies for each ligand and site are given in Table I.2.

Table I.2. Thermodynamic parameters employed for the simulation of the curves shown in Figure I.2.

	Region 1; n = 20		Region 2; n = 5	
	A	B	A	B
K	10^8	10^6	10^7	10^4
ΔH	$-5\cdot10^3$	-10^3	10^3	$5\cdot10^3$

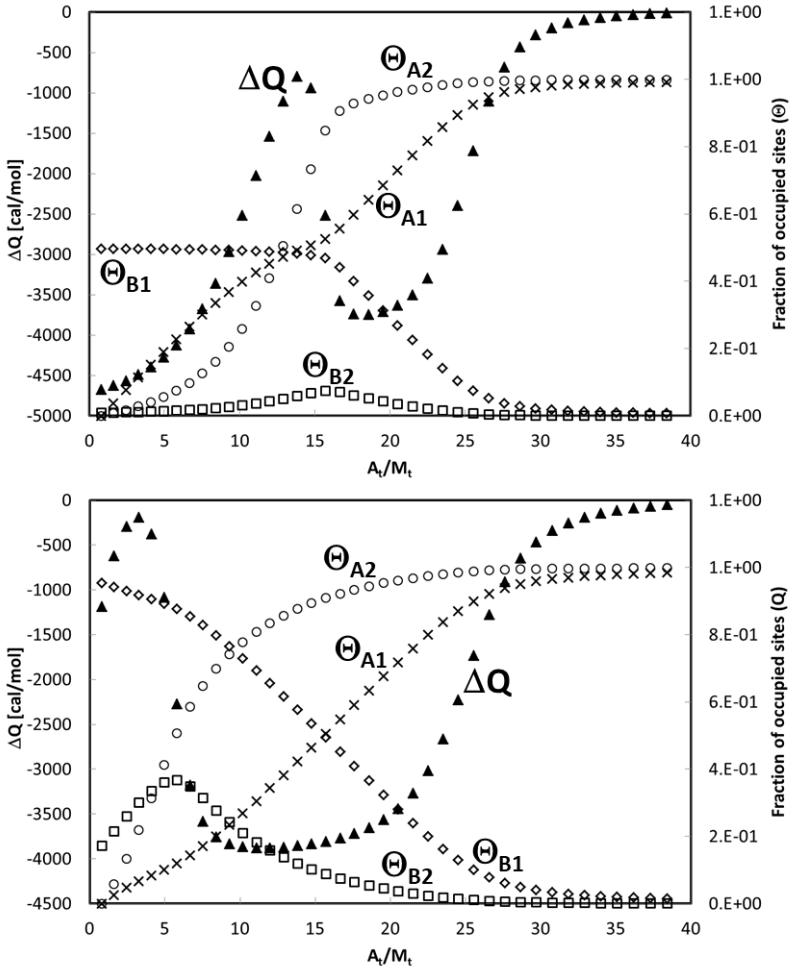


Figure I.2. Simulated binding isotherms (▲; left axes) for the interaction of a hypothetical nanoparticle with two competitor ligands (see main text), together with the fraction of occupied sites (see labels in the plots; right axes) for the two interaction regions (subscripts 1 and 2). The two binding isotherms reflect the same molecular interactions at slightly different concentrations of B in the sample cell (5 mM for the top plot and 10 mM for the bottom plot).

At the beginning of the experiment the concentration of A in the cell is null and the concentration of B (5 mM) is not enough to saturate both regions of NP. In the first experiment the first region (with n=20) is half saturated of B while the second region is empty, as expected from the corresponding equilibrium constants. The addition of A displaces B from the first region and it starts to populate the second region (see the small maximum at At/Mt~16). Then A also displaces B from the second region. At the end of the experiment, the whole NP is saturated by molecules of A ($\Theta_{A1} = \Theta_{A2} = 1$) while B remains free in solution ($\Theta_{B1} = \Theta_{B2} = 0$). The second experiments (B concentration 10 mM) behaves similarly but it begins at a larger concentration of B, enough to saturate the first region and to partially populate the second one. Note that the binding isotherms are quite different for each of these two hypothetical experiments and hence they would be well suited to perform a global fitting and obtain a good thermodynamic characterization.