bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Rigorous equations for isothermal titration calorimetry: theoretical and practical consequences

Philippe Dumas
IGBMC, Department of integrated structural biology, Strasbourg university
F67404 Illkirch CEDEX, France
dumasp@igbmc.fr

Abstract

A preliminary version of this work, [doi.org/10.1101/512780], was deposited in bioRχiv on January 6, 2019, and in a revised form on May 15, 2019

New mathematical methods have been developed for processing titration curves (TC) obtained from Isothermal Titration Calorimetry (ITC). Exact TC equations for the usual multi-injection method (MIM), or for the single-injection method (SIM) with continuous injection, were derived by taking into account rigorously the effect of dilution resulting from the titration process. Several practical consequences of these results are discussed. An exact fit of a TC can thus be obtained, even with large injected volumes leading to important dilution. All available programs show systematic differences with the exact results, NanoAnalyze from TA being significantly more accurate. A part of the results have been incorporated into AFFINImeter from S4SD. It is also examined how certain multi-step mechanisms are in fact thermodynamically equivalent to the one-step association/dissociation mechanism. They will thus never explain any "atypical" TC not showing the classical sigmoid shape. Although only a single pair of parameters (K_d and ΔH) can explain an experimental TC compatible with the one-step mechanism, an infinite number of parameters explain equally well the same data with such an equivalent multi-step mechanism. An explicit link between the two sets of parameters is given. A parallel with the concept of gauge invariance in physics is proposed.

Statement of significance This work reconsiders in full the processing of ITC data by taking into account exactly the dilution due to the titration. New equations are obtained both for the single and multiple injection methods, which suppresses systematic problems in commonly used programs. The method is made practically available to everyone.

Index terms— ITC, Single Injection Method, Multiple Injection Method, Titration curve equations, Gauge Invariance.

1 INTRODUCTION

Isothermal Titration Calorimetry (ITC) is now a widely spread technique in chemistry and biology. This is due to the development of commercially available instruments with small cell volumes and high sensitivity (down to 200 μl and 0.1 μW, respectively). The range of application of ITC extends from classical chemical reactions to lipid-membrane studies, from macromolecule-ligand to macromolecule-macromolecule interactions; ITC also allows studying bacteria in solution and in biofilms by monitoring the evolved heat attached to their metabolism [1]. Interestingly, the list is not closed and new developments are being made like those on the determination of surface tension of liquids (A. Piñeiro, personal communication). Here, we will focus on the most common application of ITC consisting in titrating one molecule (the *titrant*) against another one initially present in the measurement cell (the *titrate*). This covers the vast area of all possible molecular interactions of chemical or biological interest. The aim of this work is of presenting the development and the consequences of a new mathematical method for processing such experimental data. The sole hypothesis being made is that the cell content is always well mixed, which allows to derive

© 2013 The Authors
0006-3495/08/09/2624/12 $2.00

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

an exact equation of a titration curve (TC) for the simple one-step association/dissociation mechanism obeying A + B ⇌ C and being described by one pair of thermodynamic parameters: the dissociation constant K_d and the enthalpy variation ΔH. It is also shown how to extend exactly the method to any complex mechanism without invoking any simplification. Several practical and theoretical consequences are presented. In a second part, these results are used with n-step mechanisms that one may be tempted to consider to account for atypical titration curves that cannot be explained by the usual one-step mechanism. Two such examples of composite mechanisms involving n pairs of parameters (K_{d_i}, ΔH_i, $i = 1, n$), will be shown to be in fact thermodynamically equivalent to the one-step association/dissociation mechanism fully described by a single pair of parameters $K_d(global)$ and $\Delta H(global)$. The exact, and by far non trivial, correspondence between the two sets of thermodynamic parameters is then obtained from the TC equations, which is a strong mark of theoretical consistency.

2 MATERIALS AND METHODS

2.1 Initial considerations

Two methods can be used in an ITC experiment: the Single Injection Method (SIM) wherein a single continuous injection at a steady rate is performed and the more common Multiple Injection Method (MIM) wherein successive injections lasting a few seconds are performed. Importantly, the time left between successive injections with MIM has to be sufficient to allow the reaction to reach equilibrium (that is to allow the power curve to reach the baseline), which can last much longer than the injection time. In this work, this hypothesis is supposed to hold true. Each heat power curve is integrated to yield the heat Q produced or absorbed in the syringe after each injection. The experimental data, the so-called *titration curves* (TC), or *binding isotherms*, thus correspond with MIM to discrete values $Q(V_i)$ of the function $Q(V)$ where V_i is the sum of all injected volumes from the first to the ith injection. We are interested in an equation yielding a rigorous evaluation of dQ/dV. It will be shown how this differential quantity can be used to take into account injections of finite size.

We first consider the simple situation corresponding to a single step:

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{C} \qquad (K_a,\, \Delta H) \tag{1}$$

where compound A is initially alone in the measurement cell at a concentration A_0 and compound B in the syringe at a concentration B_0 (the concentration $[X]$ of a compound X is noted simply X). Note that these neutral notations A and B are used to emphasize that one does not presuppose that a macromolecule (protein, nucleic acid,...) is initially in the cell. This reaction is characterized by an association constant K_a (or a dissociation constant $K_d = 1/K_a$) and an enthalpy variation ΔH per mole of C produced during the reaction. If A_0 and B_0 are known, the $Q(V_i)$ can be translated into $Q(s_i)$ where the s_i are the successive stoichiometric (or molar) ratios B_{tot}/A_{tot} with $A_{tot} = A + C$ and $B_{tot} = B + C$. Although the experiments with MIM only yield a discrete sampling of the function $Q(V)$, it is legitimate to consider the function itself with V varying continuously as in SIM.

2.2 The 'overfill mode'

Most often, any injected volume δV of compound B implies that an equivalent volume δV of the reaction mixture has to leave the measurement cell. This corresponds to the *overfill mode* and this implies that the measurement cell is already filled at the beginning of injections. Note that, with the instruments from TA, the user may or may not choose the overfill mode. In the following, we will only consider the overfill mode.

2.3 The dilution problem with the 'overfill mode'

A simplified method for obtaining dQ/dV is of considering that B in the syringe is much more concentrated than A ($B_0 \gg A_0$) and, therefore, that there is negligible dilution of the reaction mixture during the titration (as this was supposed in the seminal paper by Wiseman *et al.* [2]). As a consequence, any infinitesimal variation dC of the concentration of C is due to the reaction, which implies:

$$\frac{dQ}{dV} = \Delta H\, V_{cell}\, \frac{dC}{dV} = \Delta H\, \frac{dC}{dv} \tag{2}$$

where the reduced volume $v = V/V_{cell}$ has been introduced and will be used throughout to obtain equations valid for all instruments. However, this is most often an oversimplified hypothesis, particularly in the overfill mode enhancing the effect of dilution. Assuming that **the cell content is always well mixed**, the infinitesimal variations of the total concentrations of A and B are thus $dA_{tot} = -A_{tot}\, dV/V_{cell} = -A_{tot}\, dv$ and $dB_{tot} = (B_0 - B_{tot})\, dV/V_{cell} = (B_0 - B_{tot})\, dv$. The equation for compound B takes into account both the injection of new material and the dilution effect. After integration, it is obtained:

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$A_{tot}(v) = A_0\, e^{-v}; \qquad B_{tot}(v) = B_0\, \left(1 - e^{-v}\right) \tag{3}$$

The current stoichiometric ratio is thus $s(v) = B_{tot}(v)/A_{tot}(v) = r\,(e^v - 1)$ where $r = B_0/A_0$, which implies:

$$e^v = 1 + \frac{s(v)}{r}; \qquad \frac{ds}{dv} = r + s \tag{4}$$

The total concentrations can thus be expressed as functions of s considered as the independent variable:

$$A_{tot}(s) = A_0\, \frac{r}{r+s}; \qquad B_{tot}(s) = B_0\, \frac{s}{r+s} \tag{5}$$

2.4 Accounting for the dilution by the usual processing programs

Equations (3, 4) have already been derived [3, 4]. They are both very simple and "exact" (provided that complete mixing is always achieved). Strangely, various approximations of them are considered in all established programs. For example, the programs PEAQ from Microcal/Malvern and AFFINImeter from S4SD (http://software4science.com/), rely on the rational function $(1 - v/2)/(1 + v/2)$ in lieu of the term e^{-v} in equations 3 (see [5]). This rational function obtained after approximate considerations appears to be a first-order *Padé* approximant of the exact term e^{-v} (see equation (24) in [6]). The approximation is excellent when $v = V/V_{cell} \ll 1$, but deteriorates rapidly for larger values, which has serious consequences when the experimental conditions impose to inject large volumes due to insufficient concentration in the syringe. Furthermore, as already noticed in [7], an additional approximation of the approximation is made by using $B_{tot}(v) \approx B_0\, v\,(1 - v/2)$ instead of $B_{tot}(v) \approx B_0\, v/(1 + v/2)$. A better approximation than the *Padé* approximant for large v values, but not an optimal one for small values, is based on $e^{-v} \sim (1 - v/n)^n$ where v/n is the (reduced) injected volume at each injection and n the number of injections to reach v (see [8, 9]). This approximation is used in NanoAnalyze from TA. Finally, the program SEDPHAT [10] makes use of still another approximation.

2.5 Limits of the complete mixing hypothesis

The unique assumption made in this work is that the cell content is always well mixed, which may not be the case in short time intervals following quick injections. Compound B being injected at a distance from the output tubing, incomplete cell mixing tends to push out of the cell material less concentrated in free B than average. One may thus foresee further improvements in this area by introducing kinetic methods as in the full kinITC procedure [11]. To achieve realistic modelling (i) of the mixing of a stirred liquid, (ii) of the chemical reaction and of the resulting thermogenesis (variable both in time and space) and (iii) of heat transfer, it is likely that empirical methods used in engineering will be useful.

2.6 Using of a symbolic mathematical tool

Several results of this work could only be obtained, or checked, by using a mathematical software with symbolic capabilities. Here, *Mathematica 11.3* from Wolfram Research (Champaign, IL) was used. All figures were made with *Mathematica* too.

3 RESULTS

3.1 Accounting rigorously for the influence of dilution

Here, the function $Q(v)$ is evaluated and, for that, we need to consider the effect of dilution on compound C. The infinitesimal concentration variation dC is the sum $dC_{chem} + dC_{dil}$ of a term dC_{chem} due to the chemical reaction consuming A and B to produce C, and of dC_{dil} due to the dilution resulting from the injection of $dV = V_{cell}\, dv$. As for dA_{tot}, one has $dC_{dil} = -C(v)\, dv$. The reason why there was no distinction between chemical and dilution effects for A_{tot} and B_{tot} is that $d(A + C)_{chem} = d(B + C)_{chem} = 0$ and, therefore, $A_{tot} = A + C$ and $B_{tot} = B + C$ are only affected by the dilution.

Obviously, the heat evolved or absorbed during the titration is only linked to the variation of the concentration of C due to the chemical reaction, not to C leaving the cell. Therefore, taking into account v instead of V, equation 2 has to be replaced by:

$$\frac{dQ}{dv} = \Delta H\, V_{cell} \frac{dC_{chem}}{dv} \tag{6}$$

From the preceding, $dC_{chem} = dC - dC_{dil} = dC + C(v)\, dv$, which gives an exact expression for the **heat per injected mole of B**:

$$\mathcal{Q}(v) = \frac{1}{B_0\, V_{cell}} \frac{dQ}{dv} = \frac{\Delta H}{B_0} \left[\frac{dC}{dv} + C(v)\right] \tag{7}$$

where the new notation $\mathcal{Q}(v)$ for $dQ/(B_0\, V_{cell}\, dv)$ has been introduced. Considering that $ds/dv = r + s$ (equation 4), an alternative form with s instead of v as the

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

variable is obtained as:

$$\mathcal{Q}(s) = \frac{\Delta H}{B_0}\left[(r+s)\,\frac{dC}{ds} + C(s)\right] \tag{8}$$

Equations 7,8, which do not seem to have been mentioned yet, open the way for a rigorous treatment of ITC experiments. The difference with all approximate treatments is the term $C(v)$ added to dC/dv. Its justification is not immediately intuitive, which explains why it remained overlooked. Its importance is highlighted in the next section. Equation 7 may be considered from two different view points. First, one may see it as a differential equation yielding $C(v)$ by integration if the thermogenesis term $\mathcal{Q}(v)$ is known:

$$C(v) = \frac{B_0}{\Delta H}\int_0^v \mathcal{Q}(u)\,e^{-(v-u)}\,du \tag{9}$$

This tells that $C(v)$ can be obtained from the convolution of $\mathcal{Q}(v)$ with the function $\mathsf{H}(v)\,exp(-v)$ describing the progressive dilution in the measurement cell (H is the Heaviside step function: $\mathsf{H}(v<0)=0$, $\mathsf{H}(v\geq 0)=1$). Alternatively, one may consider that equation 7 yields $\mathcal{Q}(v)$ if $C(v)$ is known from the equations of equilibrium. Clearly, we are usually interested in the latter view point as we seek to compute $\mathcal{Q}(v)$ (or equivalently $\mathcal{Q}(s)$) at any step of a titration to determine K_a and ΔH by fitting the observed TC. Interestingly, if $C(v)$ is known analytically, it derives that $\mathcal{Q}(v)$ is also known analytically from equation 7 (see 3.3). Also, this result can readily be used with any complex mechanism involving several reactions and thus several functions $C_k(v)$ by considering the sum:

$$\mathcal{Q}(v) = \sum_{k=1}^{n}\frac{\Delta H_k}{B_0}\left[C_k(v) + \frac{dC_k}{dv}\right] \tag{10}$$

where n is the number of independent products participating in the thermogenesis and ΔH_k the enthalpic term specific of the kth product. It is of utmost importance to note that C_k **is different from the concentration of the kth product itself** if this compound participates in different reactions. This requires to be examined with great care and will be illustrated in the following with three independent products among five compounds in total being engaged in three reactions (see section 3.8.2 and section A.2 in Supplementary information).

3.2 *Influence of the term $C(v)$*

The importance of the influence of the term $C(v)$ on $\mathcal{Q}(v)$ (equation 7) is illustrated by considering the single association/dissociation mechanism described by equation 1. From mass action law ($A\times B/C = K_d$) and the conservation equations ($A+C=A_{tot}$ and $B+C=B_{tot}$), the concentration of C is obtained as a function of A_{tot} and B_{tot} at any stage of a titration:

$$\begin{aligned} C &= \frac{1}{2}\left[S-\sqrt{S^2-4\,P}\right] \\ S &= A_{tot}+B_{tot}+K_d \qquad P = A_{tot}\,B_{tot} \end{aligned} \tag{11}$$

The latter result on C is readily transformed into a function $C(v)$ by replacing A_{tot} and B_{tot} with equations 3. The variation of $C(v)$ is shown with figure 1. Importantly, for increasing values of v, dC/dv becomes asymptotically equal to $-C(v)$ (verified with *Mathematica*), which means that $\mathcal{Q}(v)\longrightarrow 0$ from equation 7, as it should. This important feature derives directly from the new exact equations, and not from the previous approximate treatments. As a matter of fact, we will see that three commonly used programs, out of four, do not fulfill this requirement of a null asymptotic value of $\mathcal{Q}(v)$.

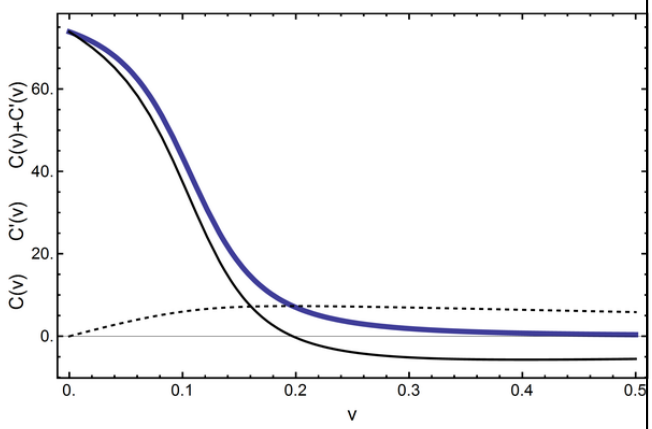

Figure 1: Influence of $C(v)$. Dashed curve: $C(v)$, solid thin curve: $C'(v)$, solid thick curve: $C(v)+C'(v)$.

3.3 *An exact analytical expression for $\mathcal{Q}(v)$ (and $\mathcal{Q}(s)$)*

We continue with the classical situation of a single association/dissociation mechanism described by equation 1. Here, only the logic is exposed as it is not useful to show the details. The latter result on $C(v)$ allows to obtain the heat per injected mole $\mathcal{Q}(v)$ by using equation 7. Then, the result is transformed by using equation 4 in order to express the heat per injected mole as a function of the stoichiometric ratio $s = B_{tot}/A_{tot}$, as usually in practice. The final result is:

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$\mathcal{Q}(s) = \frac{\left(\gamma^{-1}+1\right)\left[Y(s)-r\right]-\left(\gamma^{-1}+\gamma+2\right)s+\gamma-1}{2\,Y(s)}\Delta H$$
$$Y(s) = \sqrt{\left[\gamma\,(s-1)+r+s\right]^2+4\,\gamma\,(r+s)}$$
$$r = B_0/A_0;\;\; c = A_0/K_d\,(\text{Wiseman parameter});\;\; \gamma = r\,c \tag{12}$$

This corresponds to an exact explicit solution replacing the numerical approach described in [9]. It may be verified that for $r \longrightarrow \infty$, corresponding to negligible dilution, this is identical to the Wiseman isotherm (equation 3 in [2] where X_r stands for s and r for $1/c$).

One might argue that this rigorous result is of limited practical use because, being based upon differential calculus, it would only be significant for a continuous process with an infinite number of infinitely small injections. However, precisely for that reason, this provides us with an exact description of the continuous variation of s, or v, with SIM (if equilibrium is always reached), which is examined in the next section. Moreover, as we will see, one can take into account rigorously finite injections in MIM.

3.4 Equation for the measured heat power during a continuous injection in SIM

The continuous injection (cITC) used with SIM was presented by Markova and Hallen in [12]. The method has been analyzed in depth for being used with enzyme kinetics [13]. With cITC, the chemical reaction is perturbed at all times by the continuous injection of new material. In theory, therefore, equilibrium is never reached during the titration. However, in situations where the continuous injection is sufficiently slow that (i) the reaction is always very close to equilibrium and (ii) that the instrument response time may be neglected, $\mathcal{Q}(s)$ given by equation 12 can be used directly to represent the evolution of the heat per injected mole during the titration. In general, however, the instrument response time τ_r has to be taken into account. This is done by estimating the measured heat per mole $\mathcal{Q}_m(s)$ by:

$$\mathcal{Q}(s) = \mathcal{Q}_m(s) + \tau_r\,\frac{ds}{dt}\,\frac{d\mathcal{Q}_m}{ds}, \tag{13}$$

which is known in calorimetry as the Tian equation (for accessible references, see [14, 15]). From equation 4 it is obtained:

$$\frac{ds}{dt} = r\,\varphi\,e^{v} = \varphi\,(r+s)$$
$$r = B_0/A_0; \quad \varphi = dv/dt \tag{14}$$

with φ being the reduced injection rate (in s^{-1}). $\mathcal{Q}(s)$ being known (equation 12), $\mathcal{Q}_m(s)$ can be obtained by numerical integration of equation 13. Finally, $\mathcal{Q}_m(s)$ is transformed into $\varphi\,V_{cell}\,B_0\,\mathcal{Q}_m(s)$ to translate a heat per injected mole into a thermal power. Therefore,

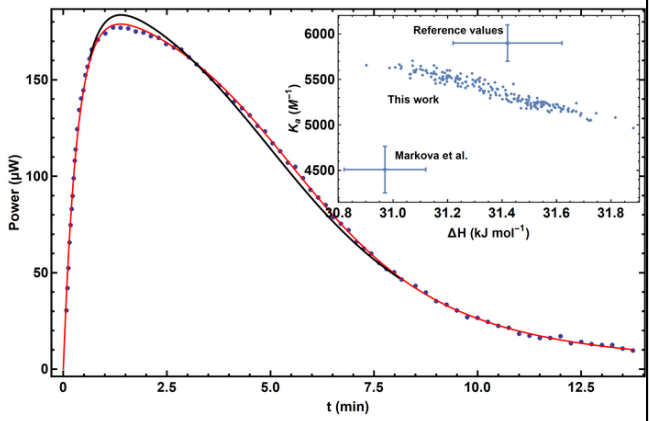

Figure 2: Fit of a continuous SIM curve (red) with equation 13. The experimental points were obtained from figure 2B in [12]. A correction of dilution was made as mentioned in [12]. The origin of times was displaced by 0.11 min to account for a clear lag before the abrupt rise of power in the original figure. A normally distributed random noise of zero mean and 0.5 μW standard deviation was added to the points to account for the small errors in the determination of their coordinates from the figure. This yielded reasonable error estimates on K_a and ΔH through multiple repeats of the fitting procedure each time with a new set of noise (inset: cloud of points corresponding to 200 repeats). The black curve corresponds to a fit with the nominal value of $V_{cell} = 1.36$ ml, whereas the red curve was obtained with $V_{cell} = 1.069 \times 1.36 = 1.45$ ml (see text).

$\varphi\,V_{cell}\,B_0\,\mathcal{Q}_m[s(t)]$ can be used to fit the measured experimental power curve $P_m(t)$. However, in [12], the curve to be fit was not the measured power curve, but an ideal power curve $P_i(t)$ that would be obtained with an ideal instrument having a null response time. Such an ideal power curve is represented by $\varphi\,V_{cell}\,B_0\,\mathcal{Q}[s(t)]$. It is argued here that the two methods are not equivalent in practice because the information content in $P_m(t)$ is higher than in $P_i(t)$ due to the necessary convolution operation to derive the ideal power curve from the measured power curve (see section A.4 in Supplementary information).

The fit of the measured power curve has been tested with one experimental dataset reported in [12]. These data obtained with a VP-ITC-like instrument are from the continuous titration of Ba^{++} with 18-crown-6. (Another experiment was reported about the titration of cytidine-2' monophosphate by bovine RNase A, but could not be used here because the raw power curve

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

was not shown). It appeared that the fit of these data with equation 13 was unacceptably poor when the expected values $r = B_0/A_0 = 29.95$ and $V_{cell} = 1.36\,\mathrm{ml}$ were used (figure 2, black curve), but that it improved considerably either after decreasing r by *ca.* $5-6\%$, or after increasing V_{cell} by the same relative amount. *A priori* neither r nor V_{cell} are free parameters; however, accurate measurements that were made in view of calibrating one VP-ITC instrument have shown that the effective cell volume could be significantly different from its nominal value, which led to an effective volume smaller than expected by $5-7\,\%$ [7]. Here, it was thus attempted to include the cell volume among the free parameters, which effectively led to a perfect fit with K_a and ΔH values in better agreement with the reference values in [16] than those obtained in [12] (figure 2, inset). In addition, the obtained value $\tau_r = (19.6 \pm 0.2)\,s$ for the instrument response time agrees perfectly with the expectation for a VP-ITC-like instrument. At variance with the decrease observed in [7], the effective cell volume had to be significantly increased by $(6.9 \pm 0.4)\,\%$ to fit the data. (Note that this is somewhat a fudge parameter that also aggregates any errors on A_0 and B_0, and thus on r). Although not proved firmly, this result looks reasonable.

In conclusion, equations 13 and 14 provide us with an accurate description of the measured heat power in SIM since this led to introducing an unexpected, but likely justified, correction on the cell volume. This could incite to use more widely this method introduced in [12]. It should be recalled, however, that this accuracy only holds if one is always very close to equilibrium. Indeed, if the continuous injection is too fast in comparison of the kinetics of equilibration of the reaction, then the latter equations are no more valid. In such a case, a true kinetic analysis (kinITC), as described in [11, 17], would ne necessary. This will only require a slight adaption to SIM of the description of the kinetic process that was performed for MIM.

3.5 Equation for the measured heat per mole for finite injections in MIM

Here, we consider how to account for finite-size injections. In practice, the acquisition program of an ITC instrument reports the average value of the heat power during short time intervals (all of the same duration). The heat evolved or absorbed following any injection comprised between the stoichiometry limits s_1 and s_2 is then the sum of these elementary power measurements multiplied by the short time interval. This is equivalent to the average value of the elementary power measurements mutiplied by the overall time interval between the stoichiometry limits s_1 and s_2. We can therefore obtain this exactly as the average value $\overline{\mathcal{Q}}(s_1, s_2)$ of $\mathcal{Q}(s)$ between s_1 and s_2. From equation 8 leading to a fortunate cancellation upon integration, it is obtained:

$$\overline{\mathcal{Q}}(s_1, s_2) = \frac{\int_{s_1}^{s_2} \mathcal{Q}(s)\,ds}{s_2 - s_1} = \frac{(r+s_2)\,C(s_2) - (r+s_1)\,C(s_1)}{(s_2 - s_1)}\,\frac{\Delta H}{B_0} \tag{15}$$

where $C(s)$ is the equilibrium concentration of the complex C as a function of s. This corresponds to the exact counterpart of the common approximation:

$$\overline{\mathcal{Q}}(s_1, s_2) \simeq \frac{C(s_2) - C(s_1)}{\delta V/V_{cell}}\,\frac{\Delta H}{B_0} = \frac{C(s_2) - C(s_1)}{v_2 - v_1}\,\frac{\Delta H}{B_0} \tag{16}$$

Equation 15 is not sufficient by itself since $C(s_1)$ and $C(s_2)$ have to be evaluated. By direct integration of $\mathcal{Q}(s)$ from equation 12 a stand-alone equation may be obtained as:

$$\overline{\mathcal{Q}}(s_1, s_2) = \frac{1 + \gamma - [Z(s_2) - Z(s_1)]\,/(s_2 - s_1)}{2\,\gamma}\,\Delta H$$
$$Z(s) = \sqrt{(r+\gamma)^2 + 2\,[r + \gamma + (r-\gamma)\,\gamma]\;s + (1+\gamma)^2\,s^2} \tag{17}$$

The integral was calculated with *Mathematica*. The latter form is quite different from the former because it implicitly includes the evaluation of $C(s_1)$ and $C(s_2)$. By definition of a derivative, $\overline{\mathcal{Q}}(s_1, s_2) \longrightarrow \mathcal{Q}(s_2)$ when $\delta s = s_2 - s_1 \longrightarrow 0$. For sufficiently close injection limits s_1 and s_2, the difference between the two estimates is very small if $\mathcal{Q}$ is evaluated at $s = \overline{s} = (s_1 + s_2)/2$, but it becomes quite significant when δs increases (Fig. 3).

Practical consequences: Equation 17 involves elementary mathematical functions and is easily programmed in widely used spreadsheet tools. This allows anyone to process TCs for the simple association/dissociation mechanism with rigorous equations by using the 'solver' functionality of such software (a Microsoft® Excel file with explanations is available on line. See section A.1 in Supplementary information). Another consequence is that, as with a continuous injection, one can represent the evolution of the heat per mole *vs.* s as a smooth curve being sampled at discrete points s_i depending on the successive injected volumes. This is illustrated with figures 3-6, 8.

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

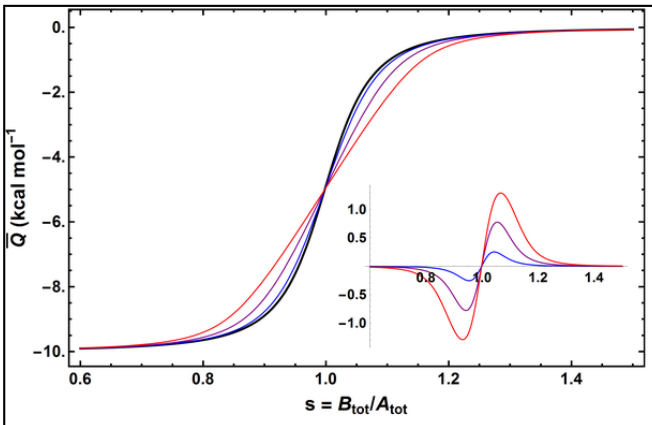

Figure 3: Influence of the injection width $\delta s = s_2 - s_1$ on the resulting TC calculated with $\overline{\mathcal{Q}}(s_1, s_2)$. The curves were calculated with $r = 10$, $c = 700$ and with δs varying from 0, which corresponds to $\mathcal{Q}(s)$ (black), to 0.3 (red), in steps of 0.1. The inset shows the difference curves corresponding to $\overline{\mathcal{Q}}(s_1, s_2) - \mathcal{Q}(s)$ with the same color code. Note that the large values reached by δs in this numerical test are unrealistic in practice and were used to show clearly the evolution of the resulting TCs. However, $\delta s = 0.1$ is a realistic value, which already shows well the difference between $\overline{\mathcal{Q}}(s_1, s_2)$ and $\mathcal{Q}(s)$ with the value $c = 700$ used. See below (equation 20) for a discussion on the influence of c.

3.6 Comparison of the exact and approximate equations

Here, the results obtained in this work are compared with those from different programs (PEAQ from Microcal/Malvern, NanoAnalyze from TA instruments, AFFINImeter from S4SD, and SEDPHAT [10]). This is done by comparing the theoretical TCs that are obtained with the same parameters (figures 4, 5, 6). Two K_d values, leading to $c = 10$ and $c = 1000$, were considered to illustrate smooth and sharp TCs. A low value $r = B_0/A_0 = 3$, implying important dilution of the cell content, was chosen to better visualize the influence of the approximations used for the stoichiometric ratio estimates in the different programs. The results from PEAQ and AFFINImeter were virtually identical: only the results from PEAQ are shown. There are significant errors in the ordinate values $\overline{\mathcal{Q}}$ obtained from PEAQ, AFFINImeter and SEDPHAT, and much less errors from NanoAnalyze (insets in figures 4, 5, 6). For sharp TCs (high c value), the errors are systematically high around the unit stoichiometry ($s \approx 1$). Also, all programs, except NanoAnalyze, show TCs going wrongly from negative to positive values at high stoichiometric ratios. This is obviously due to the lack of the corrective term $C(v)$ in equation (7) whose influence is visualized in figure 1. The negative consequences of this are discussed in section 3.7. According to the *model editor* in NanoAnalyze, the heat at the nth injection is evaluated as $(C_n - C_{n-1})\, V_{cell}\, \Delta H$ after a correction of the concentration of the bound species C_{n-1} reached at injection $n-1$ for the dilution resulting from the nth injection. This is only a partial correction since it does not correct C_n itself for the dilution effect. It is thus not clear how NanoAnalyze circumvents the lack of the term $C(v)$ in equation (7).

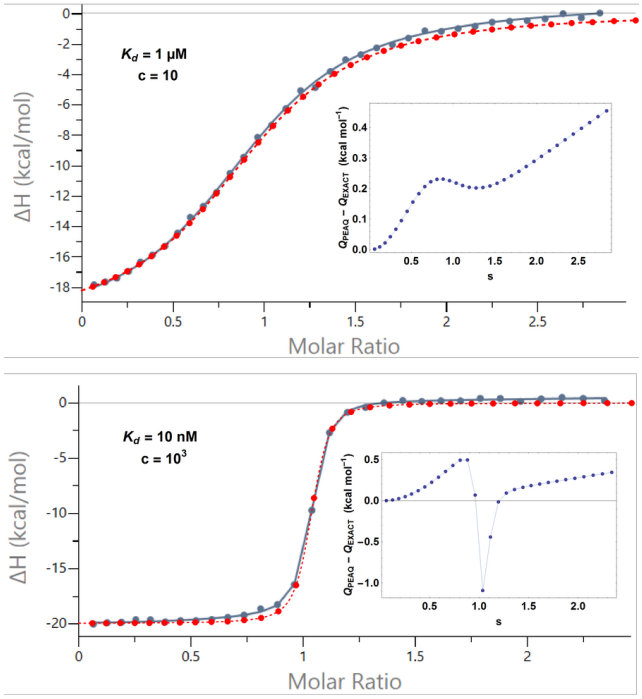

Figure 4: Comparison of the titration curves obtained from equation 17 (dashed red) and from PEAQ software from Malvern (blue). The results are virtually identical for the comparison with AFFINImeter. The calculations were done with the values $V_{cell} = 200\,\mu l$, $\delta V = 4\,\mu l$, $A_0 = 10\,\mu M$, $B_0 = 30\,\mu M$ ($r = 3$), $\Delta H = -20\,kcal\,mol^{-1}$ and either $K_d = 1\,\mu M$ ($c = 10$), or $K_d = 10\,nM$ ($c = 10^3$). The exact curve is superimposed onto the figure produced by the PEAQ software. The points for the same injection may differ signicantly in their s values (molar ratios) due to the approximation used by PEAQ (section 2.1). The inset shows the differences between the ordinate values of points for the same injection. Note that the TC from PEAQ reaches the horizontal axis at the end point for $K_d = 1\,\mu M$ and crosses the horizontal axis at $s = 1.4$ for $K_d = 10\,nM$, whereas the exact curve remains negative and reaches the horizontal axis asymptotically.

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

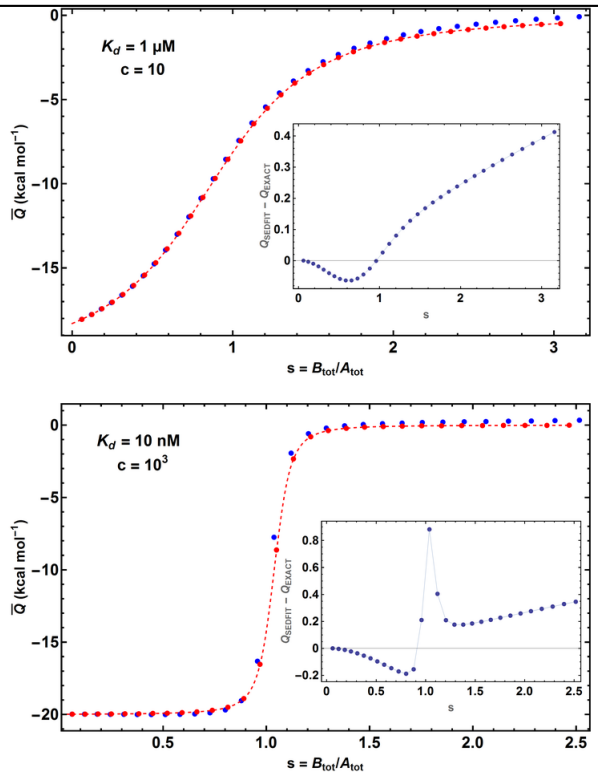

Figure 5: Comparison of the titration curves obtained from equation 17 (dashed red) and from SEDPHAT [10] (blue). Same parameters as in figure 4. Here also the TCs cross wrongly the axis $\overline{\mathcal{Q}} = 0$

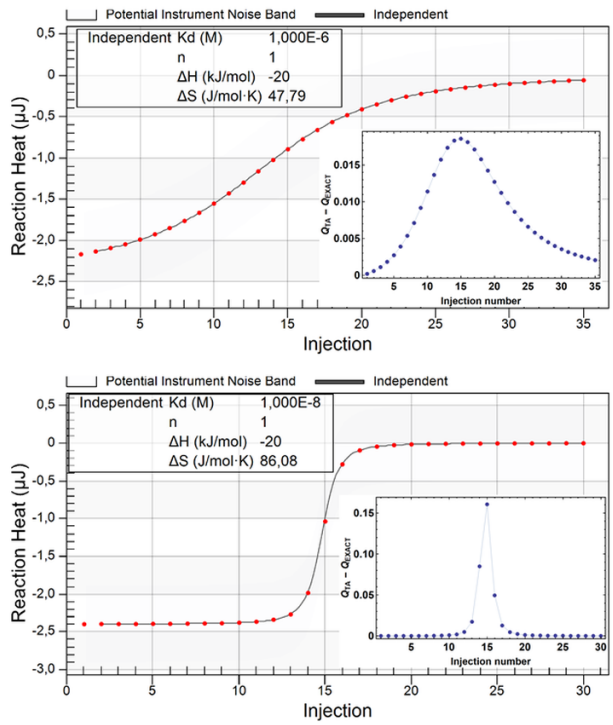

Figure 6: Comparison of the titration curves obtained from equation 17 (red dots) and from NanoAnalyze from TA (thin curves). Same parameters as in figure 4. Here, NanoAnalyze uses a particular axis labelling. The horizontal axis corresponds to the injection numbers. One cannot therefore visualize the (small) departure of the stoichiometric ratios from the exact values. The ordinate axis gives the absolute heat evolved during the injection, not the heat per injected mole. The errors on the ordinate values are marginal for $c = 10$ and become significant for $c = 10^3$, but only around $s = 1$ (injection # 15). At variance with PEAQ, AFFINImeter and SEDPHAT the theoretical TCs do not cross wrongly the axis $\mathcal{Q} = 0$.

3.7 Consequences of the analytical expression for $\mathcal{Q}(s)$ and $\overline{\mathcal{Q}}(s_1, s_2)$

Several quantitative results potentially with interesting practical consequences derive from equations 12 and 17. Some of these results were obtained by Velázquez-Campoy within the approximation of negligible dilution effects [18].

3.7.1 Stoichiometry of the inflection point (if any)

An exact expression is obtained along with a simple (and excellent) approximation for $r \gg 1$:

$$s_{infl} = 1 - \frac{1 + r + \gamma\,(r+3)}{(\gamma+1)^2}; \quad r = B_0/A_0; \quad \gamma = B_0/K_d$$
$$s_{infl} \simeq 1 - \frac{1}{c} \quad \text{for } r \gg 1; \quad c = A_0/K_d = \gamma/r \tag{18}$$

The inflection point is thus always located at $s < 1$ (this obviously does not take into account the complications due to concentration errors and/or active fractions different from 1). The simple form $s_{infl} \simeq 1 - 1/c$ was already obtained in [18].

An inflection point exists if $s_{infl} > 0$, which requires from the preceding exact result that $\gamma = B_0/K_d > (c+1)/(c-1)$, or equivalently $A_0/K_d > (B_0 + K_d)/(B_0 - K_d)$. In usual situations with $B_0 \gg K_d$, this is very close to the simple condition $A_0 > K_d$, as expected from the approximation.

Practical consequence: when a TC has no visible inflection point, one can conclude that $K_d \gtrsim A_0$ (which is equivalent to $c \lesssim 1$). However, when one is in the 'twilight zone', one cannot judge by eye whether an

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

inflection point is still present or not. Therefore, when this applies, only a rough estimate of K_d is obtained.

3.7.2 Maximum slope at the inflection point

We consider here the maximum slope derived either from $\mathcal{Q}(s)$ for an infinite number of infinitely small injections (equation 12), or from $\overline{\mathcal{Q}}(s_1, s_2)$ for finite-size injections in MIM (equation 17). In the first case, it is obtained:

$$\begin{aligned} \left(\frac{d\mathcal{Q}}{ds}\right)_{max} &= -\frac{(1+\gamma)^3}{4\,\gamma^2\sqrt{r+\gamma(1+r)}}\,\Delta H \\ \left(\frac{d\mathcal{Q}}{ds}\right)_{max} &\simeq -\frac{c^{1/2}}{4}\,\Delta H \quad \text{for } r \gg 1 \end{aligned} \tag{19}$$

The approximate result was mentioned in [18]. For the second case, one has to consider $\overline{\mathcal{Q}}(s_1, s_2)$ with $s_1 = s - \delta s/2$ and $s_2 = s + \delta s/2$ ($\delta s = s_2 - s_1$) to be comparable with $\mathcal{Q}(s)$. The exact expression of the derivative is complicate and of no interest; we thus only consider the approximate form that can be derived from it with *Mathematica*:

$$\left(\frac{d\overline{\mathcal{Q}}(s_1, s_2)}{ds}\right)_{max} \simeq -\frac{c^{1/2}}{4\,\sqrt{1+c\,(\delta s/4)^2}}\,\Delta H \quad \text{for } r \gg 1 \tag{20}$$

Since usually $r = B_0/A_0 \sim 10$, the approximation is often valuable. Comparison of equations 19 and 20 shows that the maximum slope derived from $\overline{\mathcal{Q}}(s_1, s_2)$ is less than that from $\mathcal{Q}(s)$, which is clearly apparent in figure 3. This was expected since $\overline{\mathcal{Q}}(s_1, s_2)$ results from the averaging of $\mathcal{Q}(s)$ between s_1 and s_2.

Practical consequence 1: When the transition zone of the sigmoid-shaped curve is sharp (large c value), the stoichiometry increment δs between successive injections roughly has to be less than $4\,c^{-1/2}$ to sample the transition zone at least once for sure (the estimate is safe since it is based on equation 19 giving the largest maximum slope). Therefore, this is the maximum acceptable δs value to obtain significant information on K_d. On the contrary, if $\delta s \gtrsim 4\,c^{-1/2}$, the sampling of the transition zone is a random event since its exact position depends on any errors on the concentrations and on the active fractions of A and B being potentially less than 1. In ideal conditions (no concentration errors, active fractions = 1), one can impose that the experimental point of any particular injection (say injection #i) fall very close to the inflection point. For that, the total injected volume V_i reached at that injection has to be equal to $V_{cell}\,ln[1+1/r]$ (that is $v_i = ln[1+1/r]$) since the corresponding stoichiometric ratio s_i is then equal to 1, which is very close to the exact position ($\sim 1 - 1/c$ from equation 18) of the inflection point when c is large. The practical interest of this is that one can compute in advance $V_{cell}\,ln[1+1/r]$ since $r = B_0/A_0$ is known.

The previous considerations point to the fact that it would be desirable to use variable injected volumes δV for the different injections. Indeed, with large c values, it would be better to decrease δV around the inflection point and increase it in the almost flat regions of the TC. A practical consequence is that the data acquisition programs should be able to change δV automatically during the acquisition itself. Ultimately, in all simple situations described by equation 1, the acquisition programs should be improved to perform the first steps of data processing and, hence, should tune in real time δV and the time interval between injections, particularly close to the unit stoichiometry $s = 1$ where the shape of the injection curves can change significantly. The latter point is of great importance for the ability of obtaining kinetic information by *kinITC* [11, 17, 19].

Practical consequence 2: The previous considerations give insight into the error resulting from using the ideal equation 12 (which is equivalent to the Wiseman isotherm when the dilution may be neglected) instead of equation 17. The comparison of equations 19 and 20 shows that the error remains small as far as $c\,(\delta s/4)^2 \ll 1$ and becomes significant when $c\,(\delta s/4)^2 > 10^{-1}$. When $c\,(\delta s/4)^2$ increases significantly above 1, the error resulting from using equation 12 to fit an experimental TC is severe since an erroneous c value that tends towards the limiting value $(4/\delta s)^2$ would be obtained. Therefore, in such situations, equation 12 would lead to K_d approaching the limiting value $A_0\,(\delta s/4)^2$, which only depends on "freely" adjustable experimental parameters, and not on the actual chemical interaction. More or less erratic results in the literature might originate from this. Definitely, equation 17 taking in consideration the injection width has to be used.

3.7.3 Value of $\mathcal{Q}_0$ for $s \to 0$

It is usually considered that, for large c values, $\mathcal{Q}_0 = \mathcal{Q}(s \to 0) = \Delta H$. However, from equations 12 and 17 the exact value (also mentioned in [18]) is:

$$\mathcal{Q}_0 = \frac{c}{1+c}\,\Delta H \tag{21}$$

Practical consequence: If c is known, ΔH is obtained as $(1+1/c)\,\mathcal{Q}_0$. If c is not yet known and no inflection point is visible (*i.e.* $c \lesssim 1$ from the preceding), one may

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

conclude that $|\Delta H| \gtrsim 2\,|\mathcal{Q}_0|$ (the sign being obviously known).

Theoretical consequence: Equation 21 joined to equation 10 will lead to a stringent test of theoretical consistency (section 3.8.2). Therefore, equation 21 is rigorous and any processing method not respecting it is incorrect. An example of this is in [20] where their equations 4.153, 4.154 imply $\Delta H = (1 + 1/(r\,c))\,\mathcal{Q}_0$, instead of the r-independent exact value $(1 + 1/c)\,\mathcal{Q}_0$. This results in underestimating $|\Delta H|$ with a relative error $|\delta\Delta H/\Delta H_{exact}| = (1 - 1/r)/(1 + c)$, which is as high as 45% with $c = 1$ and $r = 10$.

3.7.4 Asymptotic behavior of $\mathcal{Q}(s)$ for high stoichiometric ratios

Base-line correction is a recurrent problem due to more or less small-amplitude injections often being still present at high stoichiometric ratios. Normally, the so-called "blank experiments" performed by injecting the buffer without ligand, or by injecting the ligand into the buffer, are used to know whether or not the contribution of such injections may be legitimately subtracted. However, these blank experiments may be missing and, when present, may lead to bad corrections, particularly when the signal-to-noise ratio is high. It is thus of interest to know the asymptotic behavior of $\mathcal{Q}(s)$ when the stoichiometric ratio s becomes large to decide whether the experimental TC should be base-line corrected prior to the fitting procedure, or whether an additional free parameter should be used in the fitting procedure to determine the base-line. Since we consider flat regions of the TC, equation 12 may be used instead of equation 17, which yields the following exact and approximate asymptotic dependences for large s values:

$$\mathcal{Q}(s) \sim \frac{c\,r^2(1+c+c\,r)}{(1+c\,r)^3}\frac{\Delta H}{s^2}; \quad r = B_0/A_0;\ c = A_0/K_d$$
$$\mathcal{Q}(s) \sim \frac{1}{c}\frac{\Delta H}{s^2} \quad \text{for } r \gg 1 \tag{22}$$

Theoretical and practical consequence 1: This shows first that $\mathcal{Q}(s) \longrightarrow 0$ when $s \longrightarrow \infty$ and, second, that the sign of $\mathcal{Q}(s)$ cannot change. The same conclusion also applies to $\overline{\mathcal{Q}}(s_1, s_2)$ for s_1 and $s_2 \longrightarrow \infty$ (see figure 3 for a visual "proof"). Therefore, with all simple reactions A + B $\leftrightarrows$ C, an ideal TC (*i.e.* not affected by a baseline shift) cannot cross the horizontal axis $\mathcal{Q}(s) = 0$. It was seen that three programs out of four do not fulfill this requirement. As a consequence, no baseline correction can be made safely on an experimental basis by using "blank experiments" because the processing method introduces a systematic error on the theoretical baseline. Only by using a baseline shift as a fudge parameter in the fitting procedure can one "correct" the problem with these approximate methods. However, this will introduce errors on ΔH and, in turn, on K_d.

Practical consequence 2: Equations 22 also provide us with an objective criterion to locate a practical titration-curve end point by determining s_{max} such that $\mathcal{Q}(s > s_{max})$ may be considered null in practice. The criterion is that $\mathcal{Q}(s > s_{max})$ should become less than some fraction (say 1/2) of the experimental error $\sigma(\mathcal{Q})$ (*i.e* the standard deviation of the heat-per-mole values in the flat base line). In usual situations for which $r \gg 1$, this gives:

$$s_{max} \simeq \sqrt{\frac{2}{c}\frac{|\Delta H|}{\sigma(\mathcal{Q})}} \tag{23}$$

This result is interesting for its simplicity but two remarks should be made. First, it depends on the arbitrary choice for the fraction of the experimental error. Second, the obtained value for s_{max} is valid only if it is large enough to justify using the previous asymptotic expressions. Numerical calculations have shown that it should be greater than 1.5 to be significant, which will not be the case with too high c values and/or with too low signal-to-noise ratio $|\Delta H|/\sigma(\mathcal{Q})$. In any case, equation 12 yields an exact estimate of s_{max} in all situations by solving numerically the equation $\mathcal{Q}(s_{max}) = \sigma(\mathcal{Q})/2$. This shows that s_{max} is well above usual experimental stoichiometry limits when c is not large (Figure 7). It resuts that when c is roughly around, or less

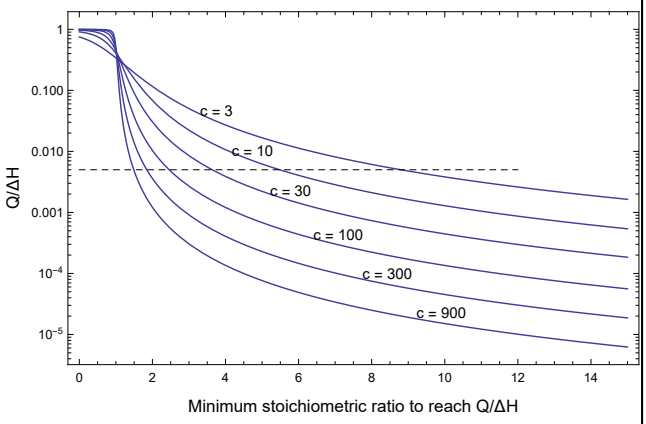

Figure 7: Minimum value of the stoichiometric ratio to be reached for various c values so that the residual heat signal becomes less than some fraction of the ΔH (valid for the usual one-step mechanism). The dashed line is for a residual signal equal to $5 \times 10^{-3}\,\Delta H$. As a rule of thumb, this 0.5 % residual signal is reached for a stoichiometric ratio close to $12\,c^{-1/3}$.

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

than 100, the base-line is most often not reached in practice and it is then incorrect to perform a base-line correction by subtracting the observed Q values at the highest stoichiometric ratios. Instead, it is safer to consider the modified theoretical function $\overline{\mathcal{Q}}(s_1, s_2) + \delta\mathcal{Q}$ from equation 17, $\delta\mathcal{Q}$ being an additional free parameter in the fitting procedure. An even more accurate procedure (used in AFFINImeter) is to consider that $\delta\mathcal{Q}$ is in fact a function $\delta\mathcal{Q}_n$ of the injection number n to take into account the difference of concentrations between the newly diluted ligand and the free ligand during the injection.

3.8 Consequences of the previous results on complex mechanisms

It is not uncommon to observe atypical TCs that cannot be explained by the usual simple mechanism involving one binding site characterized by one dissociation constant K_d and one enthalpic term ΔH. Examples of such curves are in Fig. 8.

These curves were obtained in [21] during the titration of thiamine pyrophosphate (TPP) with its target, the TPP riboswitch found in the plant *Arabidopsis thaliana* [22, 23]. A riboswitch is a non-coding RNA sequence present in the 5'- or 3'-untranslated region (UTR) of a messenger RNA (here, in the 3'-UTR of the gene THIC). The biological role of this interaction is of sensing the presence of the cognate ligand and of triggering a specific response depending on whether or not the ligand is present in sufficient concentration [24]. For this plant riboswitch, the interaction with TPP triggers a conformational change of the RNA provoking an alternative splicing mechanism that directs the mRNA to rapid degradation [25, 22]. Because it was needed to explain the non-classical shape of the TCs and because of this dynamic process, it was natural to suppose that the riboswitch could be present under two conformations R_1 and R_2, and that each conformation would recognize the TPP with a specific K_d and a specific enthalpic term ΔH. It was also natural to suppose that the two riboswitch conformations R_1 and R_2 were in equilibrium with an equilibrium constant $K = R_2/R_1$. Such alternative RNA conformations are common and many such examples having functional consequences have been described [26, 27, 28, 29, 30, 31]. The following three-step mechanism was thus considered in order to explain the curves shown in Fig.8.

$$\mathrm{R}_1 \leftrightarrows \mathrm{R}_2 \qquad (K,\, \Delta H) \tag{24}$$

$$\begin{aligned} \mathrm{R}_1 + \mathrm{TPP} &\leftrightarrows \mathrm{C}_1 \qquad (K_{a1} = 1/K_{d1},\, \Delta H_1) \\ \mathrm{R}_2 + \mathrm{TPP} &\leftrightarrows \mathrm{C}_2 \qquad (K_{a2} = 1/K_{d2},\, \Delta H_2) \end{aligned} \tag{25}$$

It was thought that with as many as six free parameters, instead of two, one would easily explain the non classical features of these TCs. Rather strangely at first sight, all attempts at obtaining such non-monotonous curves failed and purely classical sigmoid shapes were invariably obtained. On this practical basis, it was thus concluded that such attempts were vain. The fits shown in figure 8, therefore, **do not** correspond to this three-step mechanism. In fact, suppressing the interconversion step (equation 24) between the two species R_1 and R_2 is necessary and sufficient to fit the TCs in figure 8. This was rather unexpected since this corresponds to a decrease of the number of free parameters. It is explained in the following why this is so.

3.8.1 The three-step mechanism is equivalent to a classical one-step mechanism

Mass action law applied to equilibria corresponding to equations 25 indicates that the ratios R_1/C_1 and R_2/C_2 are equal respectively to K_{d1}/L and K_{d2}/L with L (for ligand) the concentration of free TPP. Taking now in consideration the preequilibrium between the two species R_1 and R_2 (equation 24) yields $R_2 = KR_1$ and, considering the former ratios, the following proportionality relationship $C_2 = \beta\, C_1$ with $\beta = KK_{d1}/K_{d2}$ is obtained. (It is important to note for the following that such a proportionality obviously does not exist if R_1 and R_2 are not engaged in this preequilibrium). It can thus be derived:

$$\begin{aligned} R_1 + R_2 &= (1+K)\, R_1 = (1+1/K)\, R_2 \\ C_1 + C_2 &= (1+\beta)\, C_1 = (1+1/\beta)\, C_2 \end{aligned} \tag{26}$$

If one considers the first equalities in equations 26, one obtains :

$$\frac{R_1 + R_2}{C_1 + C_2} = \frac{1+K}{1+\beta}\frac{R_1}{C_1} = \frac{1+K}{1+\beta}\frac{K_{d1}}{L} \tag{27}$$

It thus appears that the ratio of the two concentration sums $R_1 + R_2$ and $C_1 + C_2$ is of the form $K_d(global)/L$, exactly as the ratios R_1/C_1 and R_2/C_2 are of the form K_{d1}/L and K_{d2}/L, respectively. Therefore, everything happens as if there were binding of the ligand to a single species of RNA of concentration $R_1 + R_2$ to form a single complex of concentration $C_1 + C_2$. The global dissociation constant can be expressed under the two following alternative forms depending on whether K_{d1} or K_{d2} is taken as the reference:

$$K_d(global) = \frac{1+K}{1+\beta}K_{d1} = \frac{1+1/K}{1+1/\beta}K_{d2} \tag{28}$$

Taking into account $\beta = K\,K_{d1}/K_{d2}$ this can be

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

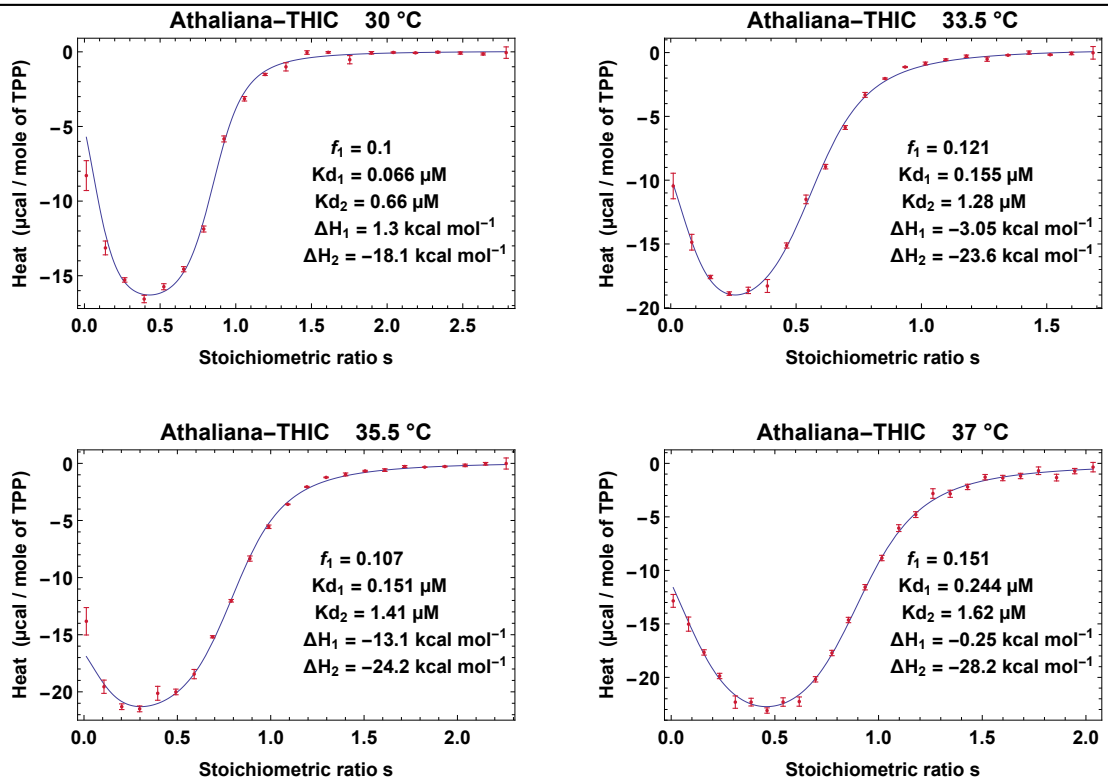

Figure 8: "Atypical" titration curves obtained at different temperatures for the binding of thiamine pyrophosphate (TPP) to the TPP riboswitch from *Arabidobsis thaliana* (the conditions are those described in [21]). The continuous lines fit the experimental points with a model involving two RNA species (section 3.8). The parameter values are indicated in each figure. The fitting was performed with a "global thermodynamic treatment" to constrain K_{d2}, ΔH_2 of the major species to respect the van't Hoff equation (as described in [11]), whereas K_{d1}, ΔH_1 of the minor species could not be constrained. f_1 and $1-f_1$ are the fractions of the minor and major species, respectively. It was observed that imposing $f_1 = K_{d1}/K_{d2}$ yielded a very good fit of the data, which likely is the mark of some unexplained feature of the TPP/riboswitch interaction. Therefore, the model in use allows to illustrate the methodological aspects of this work but, despite the quality of the fits, may not be the best model on a mechanistic point of view.

reduced to the following form (symmetric in K_{a1} and K_{a2} through the transformation $K \longrightarrow 1/K$ as it should):

$$K_a(global) = \frac{K_{a1} + K\,K_{a2}}{1+K} \tag{29}$$

As far as the concentrations are concerned, the complex mechanism involving three equilibria is thus equivalent to the simple mechanism characterized by a single association/dissociation equilibrium (equation 1). This result has already been obtained with a completely different method based on the binding polynomials approach (equation 113 in [32]). Since the shape of a TC for such a simple mechanism is uniquely determined by the initial concentrations and the equilibrium constant (its amplitude being proportional to the enthalpy variation), one understands why all attempts at obtaining non classical TCs failed. It is now clear that the interconversion step implying the proportionnality $R_2 = K\,R_1$ is responsible for this situation and that its suppression may lead to a TC departing from a sigmoid shape. It remains (i) to determine what is the value of the global enthalpic term $\Delta H(global)$ attached to $K_a(global)$ and (ii), to check that these two parameters alone account rigorously for the complex system involving three equilibria with three independent pairs of parameters (K, ΔH), (K_{a1}, ΔH_1) and (K_{a2}, ΔH_2).

3.8.2 *Determination of* $\Delta H(global)$

The determination of $\Delta H(global)$ can be obtained in two different ways. The first method makes use of equation 10 to evaluate $Q_0 = Q(s \to 0)$ and then of equation 21 to derive ΔH from Q_0. This allows us illustrating how equation 10 should be used when a product in one step is also a reactant in (an)other step(s). Being based on a specific technique, this method may not seem of general validity. However, the

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

universally valid van't Hoff equation also allows us to obtain a result independent of any technique. That the two results obtained by completely independent routes prove to be identical may be viewed as a proof that equations 10, 21 are exact.

$\Delta H(global)$ determination by the ITC-based method. Using of equation 10 requires great care since the thermogenesis resulting from the interconversion step (equation 24) is not simply related to $R_2(v)$ because the concentration variation of R_2 is also influenced by the third step (second equation 25). Instead, one can show that $R_2(v)$ should be replaced with $1/2\,[(R_2+C_2)-(R_1+C_1)]$ (see section A.2 in Supplementary information). Taking this in consideration, along with $R_2 = KR_1$, $C_2 = \beta\, C_1$ and $C_1 = C_2 = 0$ at $v = 0$, it may be shown:

$$Q_0 = Q(s \to 0) = \frac{1}{L_0}\left[\Delta H_1\, C_1'(0) + \Delta H_2\, C_2'(0) + \frac{\Delta H}{2}\left(\frac{K-1}{K+1}\, R_0 + (K-1)\, R_1'(0) + (\beta-1)\, C_1'(0)\right)\right] \tag{30}$$

with L_0 the concentration of the TPP ligand in the syringe and R_0 the initial RNA concentration in the cell. The terms $C_1'(0)$, $C_2'(0)$ and $R_1'(0)$ are the derivatives dC_1/dv, dC_2/dv and dR_1/dv at $v = 0$ and can be calculated without explicit knowledge of the functions themselves (see section A.3 in Supplementary information). Finally, with $\Delta H(global) = Q_0\,(1 + 1/c)$ from equation 21, it may be obtained after simplification:

$$\Delta H(global) = \frac{\Delta H_1 + \beta\, \Delta H_2 + \frac{\beta - K}{1+K}\, \Delta H}{1+\beta} \tag{31}$$

As expected, this is a symmetric form in ΔH_1, ΔH_2 since these terms appear in a weighted sum with coefficients $1/(1+\beta)$ and $\beta/(1+\beta)$ that are exchanged through the transformation $\beta \longrightarrow 1/\beta$ (recall that $\beta = KK_{d1}/K_{d2}$). Also, this symmetric form is analogous to that obtained for $K_a(global)$ with K replacing β in equation 29.

$\Delta H(global)$ determination by the van't Hoff equation. The van't Hoff equation applied to $K_a(global)$ yields:

$$\frac{\Delta H(global)}{RT^2} = \frac{\partial_T K_a(global)}{K_a(global)} \tag{32}$$

By replacing $K_a(global)$ from equation 29 and calculating the derivative, it is obtained:

$$\frac{\Delta H(global)}{RT^2} = \frac{\partial_T Ln\, K_{a1} + \beta\, \partial_T Ln\, K_{a2} + \frac{\beta - K}{1+K}\, \partial_T Ln\, K}{1+\beta} \tag{33}$$

which is effectively equivalent to equation 31 when considering the van't Hoff equation applied to the three equilibrium constants K_{a1}, K_{a2} and K. The method used here in a particular case has already been expressed in general terms (equations 106 and 107 in [32]).

We have thus obtained the sought after global ΔH term taking into account the three independent pairs of parameters (K, ΔH), (K_{a1}, ΔH_1) and (K_{a2}, ΔH_2). Repeated numerical calculations of TCs by considering either the three-step mechanism with any values for these six terms, or the single-step mechanism with the resulting terms $K_a(global)$ and $\Delta H(global)$, confirmed that they coincide exactly.

3.9 *Other example: competition of different modes of binding on a unique site*

The same kind of reasoning can be made with a n-step mechanism corresponding to the competition of n modes of binding on a unique site. This may happen with a large (flexible) ligand that can make more than one contact with its target and that sterically prevents the binding of another ligand as soon as one is bound in any possible way. Such a mechanism excludes any negative or positive cooperativity between the different binding modes. Although this situation has already been examined [33, 34], it is mentioned here as another illustration of the present considerations. We thus consider the n equilibria:

$$\mathrm{A} + \mathrm{B} \leftrightarrows \mathrm{C}_i \quad (i = 1, n) \tag{34}$$

with the obvious notations $K_{a\,\boldsymbol{i}}, K_{d\,\boldsymbol{i}}, \Delta H_{\boldsymbol{i}}$ for the respective thermodynamic parameters. The results are given without proof. The term $K_a(global)$ is here just the sum of the affinities of all modes of binding :

$$K_a(global) = \sum_{i=1}^{n} K_{a\,\boldsymbol{i}} \tag{35}$$

With the same van't Hoff-based method that led to equation 32, one obtains for $\Delta H(global)$ the following weighted sum of the ΔH_i s of all modes of binding:

$$\Delta H(global) = \frac{\sum_i K_{a\,\boldsymbol{i}}\, \Delta H_{\boldsymbol{i}}}{\sum_i K_{a\,\boldsymbol{i}}} = \frac{\sum_i \Delta H_{\boldsymbol{i}}/K_{d\,\boldsymbol{i}}}{\sum_i 1/K_{d\,\boldsymbol{i}}} \tag{36}$$

(Therefore, not only K_a, but $K_a\,\Delta H$ too is additive with this simple competition mechanism, which means that $K_a(global)\,\Delta H(global) = \sum_i K_{a\,\boldsymbol{i}}\,\Delta H_{\boldsymbol{i}}$). Here also, the ITC-based method using equation 10 gives the same result (which was verified for $n = 2$). It was also verified

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that the TCs obtained by considering either a two-step mechanism with any values for $K_{a1}, K_{a2}, \Delta H_1$ and ΔH_2, or the single-step mechanism with the resulting terms $K_a(global)$ and $\Delta H(global)$, coincide exactly.

3.9.1 Consequences of the equivalence of the composite and simple mechanisms

It has thus been rigorously proved that, thermodynamically, the three-step mechanism (equations 24, 25), or the n-step mechanism (equations 34) are strictly equivalent to the classical single-step mechanism (equation 1) and that the parameters describing each composite mechanism allow to compute the exact global parameters $K_a(global)$ and $\Delta H(global)$. One can thus definitely exclude the possibility of invoking such composite mechanisms to account for any atypical TCs as shown in Fig. 8. Of course, the invoked equivalence only holds as far as thermodynamics is concerned since, kinetically, the composite and simple mechanisms cannot be equivalent. This will be highlighted in a rather unexpected way (see section A.5 in supplementary information).

Reciprocally, there is another consequence of this strict thermodynamic equivalence: observing a typical monotonous TC and explaining it perfectly with the usual single-step mechanism will, alone, never exclude the possibility that it could correspond to one of the two composite mechanisms (or to others to be identified). However, if one is willing to explain a classical monotonous TC by invoking, for example the three-step mechanism (equations 24,25), one would need to prove (i) the existence of the alternative conformations R_1 and R_2 and (ii) that these can interconvert following equation 24. The latter requirement is fundamental because the existence of two **non-interconvertible forms** R_1 and R_2 allows to explain atypical TCs as shown with their fit with such a model in Fig. 8, whereas the same TCs cannot be explained if the interconversion step in equation 24 is active.

3.9.2 Number of degrees of freedom for the composite mechanisms

It is clear that there is no one-to-one correspondence between a given pair of parameters ($K_a(global)$, $\Delta H(global)$) and a single set of values for the thermodynamic parameters of each equivalent composite mechanism. For example, given ($K_a(global)$, $\Delta H(global)$), there are infinitely many solutions to equations 29 and 31 giving (K,ΔH), (K_{a1}, ΔH_1) and (K_{a2}, ΔH_2). First, equation 29 shows that two equilibrium constants among K, K_{a1} and K_{a2} can be chosen freely (within the limits of positive values) to obtain a given value of $K_a(global)$. Second, K and β being known, equation 31 shows that two enthalpic terms among ΔH, ΔH_1 and ΔH_2 can be chosen freely to obtain a given value of $\Delta H(global)$. In such a case, therefore, the composite mechanism has four degrees of freedom, whereas the equivalent simple mechanism has zero degree of freedom since only one pair ($K_d(global)$, $\Delta H(global)$) can explain the data.

Analogously, it is easy to verify that equations 35 and 36 give rise to $2\,(n-1)$ degrees of freedom since $(n-1)$ association constants $K_{a\,\boldsymbol{i}}$ (within the limits of positive values) and $(n-1)$ enthalpic terms $\Delta H_{\boldsymbol{i}}$ can be chosen freely. Altogether, one may express these observations with the following Gibbs-like phase rule: the number of degrees of freedom, or the *variance* in Gibbs' terminology, is equal to $2\,(n-1)$ where n is the number of equilibria in the simple ($n = 1$) or composite mechanisms ($n > 1$) being considered. These results lead to a tentative link with the concept of *gauge invariance* in physics (see section A.5 in Supplementary information)

4 CONCLUSION

The results exposed in this work represent a significant improvement on commonly used methods in ITC. At a theoretical level, self-consistent equations were derived, particularly concerning the asymptotic behavior of the measured heat for large s values. At a practical level, the results should allow obtaining more accurate results, particularly with important dilution. The work necessary for their introduction in existing programs is limited for the usual simple mechanism $A + B \rightleftharpoons C$. For more complex mechanisms it will be necessary to take great care of using correctly equation 10. At the time of writing the new equations for the simple mechanism have been implemented in a special version of AFFINImeter for Windows [35] (https://www.affinimeter.com/site/download/affinimeter-itc-windows/). They are also available with the Excel file accessible on line as supplementary material.

4.1 Acknowledgements

I am particularly indebted to A. Velázquez-Campoy who drew my attention to the work by Tellinghuisen [7] and made numerous constructive comments. I also thank A. Piñeiro and D. Burnouf for their critical reading, and M. Bastos, C. Birck and A. Schon for providing me with specific data.

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

[1] Chaires, J. B., L. D. Hansen, S. Keller, C. A. Brautigam, H. Zhao, and P. Schuck, 2015. Biocalorimetry. Methods (San Diego, Calif.) 76:1.

[2] Wiseman, T., S. Williston, J. F. Brandts, and L. N. Lin, 1989. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Analytical biochemistry 179:131–137.

[3] Sigurskjold, B. W., 2000. Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. Analytical biochemistry 277:260–266.

[4] Herrera, I., and M. A. Winnik, 2013. Differential binding models for isothermal titration calorimetry: moving beyond the Wiseman isotherm. The Journal of Physical Chemistry B 117:8659–8672.

[5] 2015. User manual. Microcal PEAQ-ITC analysis software (Copyright © 2015). Documentation, Malvern Instruments Limited.

[6] Weisstein, E. Padé approximant. http://mathworld.wolfram.com/PadeApproximant.html.

[7] Tellinghuisen, J., 2007. Calibration in isothermal titration calorimetry: Heat and cell volume from heat of dilution of NaCl(aq). Analytical Biochemistry 360:47–55.

[8] Freire, E., A. Schön, and A. Velazquez-Campoy, 2009. Isothermal titration calorimetry: general formalism using binding polynomials. Methods in enzymology 455:127–155.

[9] Poon, G. M. K., 2010. Explicit formulation of titration models for isothermal titration calorimetry. Analytical biochemistry 400:229–236.

[10] Zhao, H., G. Piszczek, and P. Schuck, 2015. SEDPHAT – A platform for global ITC analysis and global multimethod analysis of molecular interactions. Methods 76:137–148.

[11] Burnouf, D., E. Ennifar, S. Guedich, B. Puffer, G. Hoffmann, G. Bec, F. Disdier, M. Baltzinger, and P. Dumas, 2012. kinITC: a new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. Journal of the American Chemical Society 134:559–565.

[12] Markova, N., and D. Hallen, 2004. The development of a continuous isothermal titration calorimetric method for equilibrium studies. Analytical Biochemistry 331:77–88.

[13] Transtrum, M. K., L. D. Hansen, and C. Quinn, 2015. Enzyme kinetics determined by single-injection isothermal titration calorimetry. Methods 76:194–200.

[14] Calvet, E., 1962. Experimental Thermochemistry, Interscience: New York, volume 2, 385–410.

[15] Hansen, C. W., L. D. Hansen, A. D. Nicholson, M. C. Chilton, N. Thomas, J. Clark, and J. C. Hansen, 2010. Correction for instrument time constant and baseline in determination of reaction kinetics. International Journal of Chemical Kinetics 43:53–61.

[16] Briggner, L.-E., and I. Wadsö, 1991. Test and calibration processes for microcalorimeters, with special reference to heat conduction instruments used with aqueous systems. Journal of Biochemical and Biophysical Methods 22:101–118.

[17] Dumas, P., E. Ennifar, C. Da Veiga, G. Bec, W. Palau, C. Di Primo, A. Piñeiro, J. Sabin, E. Muñoz, and J. Rial, 2016. Extending ITC to Kinetics with kinITC. Methods in enzymology 567:157–180.

[18] Velázquez-Campoy, A., 2015. Geometric features of the Wiseman isotherm in isothermal titration calorimetry. Journal of Thermal Analysis and Calorimetry 122:1477–1483.

[19] Dumas, P., 2016. Joining thermodynamics and kinetics by kinITC, CRC Press, chapter 15, 281–300.

[20] Del Rio, J. M., and J. P. E. Grolier, 2017. Enthalpy and Internal Energy, Royal Society of Chemistry, chapter Chapter 4. Isothermal Titration Calorimetry, 96–132.

[21] Guedich, S., B. Puffer-Enders, M. Baltzinger, G. Hoffmann, C. Da Veiga, F. Jossinet, S. Thore, G. Bec, E. Ennifar, D. Burnouf, and P. Dumas, 2016. Quantitative and predictive model of kinetic regulation by E. coli TPP riboswitches. RNA biology 13:373–390.

[22] Cheah, M. T., A. Wachter, N. Sudarsan, and R. R. Breaker, 2007. Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. Nature 447:497–500.

[23] Bocobza, S., A. Adato, T. Mandel, M. Shapira, E. Nudler, and A. Aharoni, 2007. Riboswitch-dependent gene regulation and its evolution in the plant kingdom. Genes & development 21:2874–2879.

[24] Barrick, J. E., and R. R. Breaker, 2007. The distributions, mechanisms, and structures of metabolite-binding riboswitches. Genome Biology 8:1.

[25] Li, S., and R. R. Breaker, 2013. Eukaryotic TPP riboswitch regulation of alternative splicing involving long-distance base pairing. Nucleic acids research 41:3022–3031.

[26] Solomatin, S. V., M. Greenfeld, S. Chu, and D. Herschlag, 2010. Multiple native states reveal persistent ruggedness of an RNA folding landscape. Nature 463:681–684.

[27] Micura, R., and C. Höbartner, 2003. On secondary structure rearrangements and equilibria of small RNAs. Chembiochem : a European journal of chemical biology 4:984–990.

[28] Höbartner, C., and R. Micura, 2003. Bistable secondary structures of small RNAs and their structural probing by comparative imino proton NMR spectroscopy. Journal of molecular biology 325:421–431.

[29] Mahen, E. M., J. W. Harger, E. M. Calderon, and M. J. Fedor, 2005. Kinetics and thermodynamics make different contributions to RNA folding in vitro and in yeast. Molecular cell 19:27–37.

[30] Nagel, J. H. A., C. Flamm, I. L. Hofacker, K. Franke, M. H. de Smit, P. Schuster, and C. W. A. Pleij, 2006. Structural parameters affecting the kinetics of RNA hairpin formation. Nucleic acids research 34:3568–3576.

[31] Reining, A., S. Nozinovic, K. Schlepckow, F. Buhr, B. Fürtig, and H. Schwalbe, 2013. Three-state mechanism couples ligand and temperature sensing in riboswitches. Nature 499:355–359.

[32] Vega, S., O. Abian, and A. Velázquez-Campoy, 2015.

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A unified framework based on the binding polynomial for characterizing biological systems by isothermal titration calorimetry. Methods 76:99–115.

[33] Eftink, M., and R. Biltonen, 1980. Biological Calorimetry, London: Academic Press, chapter Thermodynamics of interacting biological systems, 343–412.

[34] Wyman, J., and S. J. Gill, 1990. Binding and Linkage: Functional Chemistry of Biological Macromolecules. Univ Science Books.

[35] Piñeiro, Á., E. Muñoz, J. Sabín, M. Costas, M. Bastos, A. Velázquez-Campoy, P. F. Garrido, P. Dumas, E. Ennifar, L. García-Río, J. Rial, D. Pérez, P. Fraga, A. Rodríguez, and C. Cotelo, 2019. AFFINImeter: A software to analyze molecular recognition processes from experimental data. Analytical Biochemistry .

[36] Landau, L., and E. Lifshitz, 1987. The classical theory of fields. Butterworth and Heinemann.

[37] Aharonov, Y., and D. Bohm, 1959. Significance of Electromagnetic Potentials in the Quantum Theory. Phys. Rev. 115:485–491.

A Supplementary information

A.1 Using the Excel solver with the usual one-step mechanism

The fitting of a titration curve with equation 17 is made available to everyone by using the solver functionality in Microsoft® Excel. An example is furnished (file ITC-Processing.xslx). All necessary explanations are given in the Excel file (sheet "Explanations").

A.2 Evaluation of $\mathcal{Q}$ with several reactions

When several reactions are simultaneously active one has to take into account several sources of thermogenesis. It was obtained (equation 10 in the main text):

$$\mathcal{Q}(v) = \sum_{k=1}^{N_c} \frac{\Delta H_k}{B_0} \left[C_k(v) + \frac{dC_k}{dv} \right] \tag{37}$$

It is then very important to evaluate correctly the terms $C_k(v)$ which may differ from the concentration of the kth product. The problem is illustrated with the following examples. Consider first the set of reactions (where L, for ligand, corresponds to compound B in the syringe):

$$\mathrm{R_1 + L \leftrightarrows C_1} \qquad \mathrm{R_2 + L \leftrightarrows C_2} \tag{38}$$

characterized by ΔH_1 and ΔH_2 for the production of one mole of each product $\mathrm{C_1}$ and $\mathrm{C_2}$, respectively. In such a case, there is no difficulty since neither product appears in the other reaction and the thermogenesis is given by:

$$\mathcal{Q}(v) = \frac{\Delta H_1}{L_0} \left[C_1(v) + \frac{dC_1}{dv} \right] + \frac{\Delta H_2}{L_0} \left[C_2(v) + \frac{dC_2}{dv} \right] \tag{39}$$

where $C_1(v)$ and $C_2(v)$ are indeed the concentrations of the two products $\mathrm{C_1}$ and $\mathrm{C_2}$, respectively. Consider now again the three-step reaction that was invoked about the TPP riboswitch:

$$\begin{aligned} &\mathrm{R_1 \leftrightarrows R_2} \\ &\mathrm{R_1 + L \leftrightarrows C_1} \\ &\mathrm{R_2 + L \leftrightarrows C_2} \end{aligned} \tag{40}$$

each step being characterized by ΔH, ΔH_1 and ΔH_2 for the production of one mole of each product $\mathrm{R_2}$, $\mathrm{C_1}$ and $\mathrm{C_2}$, respectively. The variation δR_2 of R_2 due to the first reaction, being accompanied by an opposite variation of R_1, is equal to $1/2\,\delta(R_2 - R_1)$. Now, the situation is less simple since $\mathrm{R_1}$ in the first reaction also appears as a reactant in the second reaction and $\mathrm{R_2}$ as a reactant in the third reaction. Thus, one cannot ascribe the variation of $1/2\,(R_2 - R_1)$ to the first reaction only. Therefore, to evaluate the thermogenesis due specifically to the

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

first reaction, one has to find a linear combination of the concentrations representing specifically the variation of $1/2\,(R_2 - R_1)$ in the first reaction. The solution is $S = 1/2\,[(R_2 + C_2) - (R_1 + C_1)]$ for the following reasons. First, $(R_2 + C_2)$ cannot be affected by the second reaction where neither R_2 nor C_2 appear and, second, $(R_2 + C_2)$ does not vary in the third reaction where R_2 and C_2 have opposite variations. Therefore the variation of $(R_2 + C_2)$ can only be ascribed to the first reaction. The same kind of arguments applies for $(R_1 + C_1)$, which implies that the variation of S can only result from the first reaction and the correct application of equation 37 leads to:

$$\mathcal{Q}(v) = \frac{\Delta H}{L_0}\left[S(v) + \frac{dS}{dv}\right] + \frac{\Delta H_1}{L_0}\left[C_1(v) + \frac{dC_1}{dv}\right] + \frac{\Delta H_2}{L_0}\left[C_2(v) + \frac{dC_2}{dv}\right] \tag{41}$$

In conclusion, if the algebraic equations of mass action law and of conservation lead for the different concentrations to explicit functions of v, this exact formalism also allows to obtain the titration-curve function in closed form by use of equation 37. Note that this is the case for both the 2-step and 3-step mechanisms just considered. If the different concentrations can only be obtained numerically, the same equation allows to obtain an 'exact' numerical solution.

A.3 Determination of the derivatives of concentrations at $v = 0$

One seeks to determine the derivatives dC_1/dv, dC_2/dv and dR_1/dv at $v = 0$ appearing in equation 30 for the determination of ΔH_{global}. The concentrations C_1, C_2 and R_1 are those of the species in equations 40. From the conservation of species (joined to equation 3) and mass action law, one derives:

$$\begin{aligned} &R_1 + R_2 + C_1 + C_2 = R_{tot} = R_0\, e^{-v} \\ &C_1 + C_2 + L = L_{tot} = L_0\,(e^{v} - 1) \\ &L = K_{d1}\,\frac{C_1}{R_1} \end{aligned} \tag{42}$$

from which one obtains:

$$\begin{aligned} &R_1'(0) + R_2'(0) + C_1'(0) + C_2'(0) = -R_0 \\ &C_1'(0) + C_2'(0) + L'(0) = L_0 \\ &L'(0) = \frac{K_{d1}}{R_1(0)}\,C_1'(0) = \frac{K_{d1}\,(1+K)}{R_0}\,C_1'(0) \end{aligned} \tag{43}$$

By inserting the proportionality relationships $R_2 = K R_1$ and $C_2 = \beta\, C_1$ with $\beta = K\, K_{d1}/K_{d2}$ (see section 3.8.1), a linear system of three equations with three unknowns is obtained, which leads to:

$$\begin{aligned} R_1'(0) &= -\frac{(1+\beta)\,(L_0 + R_0) + (1+K)\,K_{d1}}{(1+K)\,[1 + \beta + (1+K)\,K_{d1}/R_0]} \\ C_1'(0) &= \frac{L_0}{1 + \beta + (1+K)\,K_{d1}/R_0} \\ C_2'(0) &= \beta\, C_1'(0) \end{aligned} \tag{44}$$

A.4 Continuous titration with SIM

Equations 13 and 14 have been tested with experimental data reported in [12]. In this work, Markova & Hallen sought to develop the method of continuous titration (cITC) with SIM and they compared two types of calorimeter. Type 1 calorimeter was a heat-flow twin device with a long response time and type 2 calorimeter was an instrument close to the VP-ITC from Microcal-Malvern. Type 1 instrument was not considered here since it is not a total-fill overflow instrument, which does not fit with the theoretical analysis to be tested. The experimental data considered here are from the continuous titration of Ba^{++} with 18-crown-6. The measured experimental power curves $P_m(t)$ are like that shown in figure 9: they start at 0 at $t = 0$, raise abruptly to a maximum value and decrease smoothly to 0 for $t \longrightarrow \infty$. However, in [12], the curve to be fit was an ideal power curve $P_i(t) = \varphi\, V_{cell}\, B_0\, \mathcal{Q}[s(t)]$ that would be measured with an ideal instrument with a null response time τ_r. According to equation 13 in the main text $P_i(t)$ is derived from $P_m(t)$ as follows:

$$P_i(t) = P_m(t) + \tau_r\,\frac{P_m(t)}{dt} \tag{45}$$

Comparison of the two curves shows that the abrupt rise of $P_m(t)$ at $t = 0$ has disappeared in favor of a slow variation (figure 9). This results from the previous transformation $P_m(t) \longrightarrow P_i(t)$ being equivalent to the convolution:

$$\begin{aligned} &P_i(t) = [P_m * E_{\tau_r}]\,(t) \\ &\text{with } E_{\tau_r}(t < 0) = 0\,,\; E_{\tau_r}(t \geq 0) = \frac{1}{\tau_r}\,e^{-t/\tau_r} \end{aligned} \tag{46}$$

A convolution is essentially a local averaging (here of $P_m(t)$), hence a *contractant* transformation, which tends to smooth any sharp amplitude variations. (This could be expressed quantitatively with Laplace transform considerations, but figure 9 is sufficient to illustrate the problem). Any experimental information conveyed by such sharp variations is thus affected by a convolution. It is therefore not equivalent to fit the measured power curve corresponding to $P_m(t)$, or the ideal power curve

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

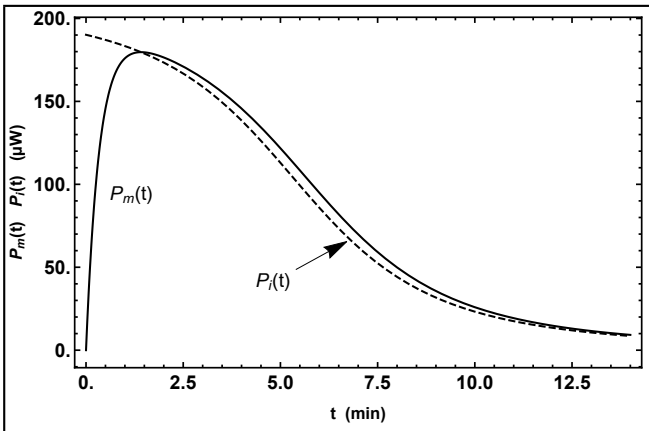

Figure 9: Comparison of the measured and ideal power curves $P_m(t)$ and $P_i(t)$ to illustrate the disappearance in $P_i(t)$ of the abrupt rise of power in $P_m(t)$. The curves were calculated with the parameters of the titration of Ba^{++} with 18-crown-6 (section 3.4).

corresponding to $P_i(t)$, because a part of the information content in $P_m(t)$ has disappeared in $P_i(t)$. In conclusion, fitting the measured power curve with $P_m(t)$ highlights, instead of masking, any disagreement. This appeared well with the black curve in figure 2 of the main text that cannot be accepted as a good fit. In comparison, the fit of the ideal power curve in [12] did not reveal any such disagreement (figure 3 in [12]) and yet, led to less accurate values of the parameters (inset of figure 2 in the main text).

A.5 Tentative link with the concept of gauge invariance in physics

Here, unusual considerations are developed. It was shown that certain composite mechanisms are strictly equivalent to a simple association/dissociation mechanism. It appeared that, in such situations, the thermodynamic parameters characterizing the composite mechanism can be chosen freely among an infinite number of possible values to represent the simple equivalent mechanism characterized by a single set of values $K_a(global)$ and $\Delta H(global)$. Such a situation is reminiscent of 'gauge invariance' in physics which, historically, appeared first with the fact that a mechanical or electric potential is defined to within an arbitrary constant. Later, it was recognized that an electromagnetic field, characterized by the measurable (hence, uniquely determined) electric and magnetic vector fields **E** and **H**, can be indifferently described by anyone of infinitely many scalar and vector potentials V and **A** through the Maxwell equations (p. 51 in [36]):

$$\begin{aligned} \mathbf{E} &= -\mathbf{grad}\, V - \frac{1}{c}\frac{\partial\ \mathbf{A}}{\partial t} \\ \mathbf{H} &= \mathbf{curl}\, \mathbf{A} \end{aligned} \tag{47}$$

All possible potentials V and **A** yielding the same vector fields **E** and **H** derive from the so-called *gauge transformations* $V \to V - \partial f/\partial t$ and $\mathbf{A} \to \mathbf{A} + \mathrm{grad} f$, where f is any arbitrary derivable function of space and time (p. 52 in [36]). In the frame of our comparison, the pair of experimentally measurable terms $K_a(global)$ and $\Delta H(global)$ stands for the pair of measurable electric and magnetic fields **E** and **H**, and the infinitely many possible values of the thermodynamic parameters given by equations 29, 31 or 35, 36 stand for the infinitely many possible functions f.

It is interesting to pursue the comparison by noting that, in classical electromagnetism, the potentials V and **A** may be seen as purely auxiliary quantities without real physical significance since applying any arbitrary gauge transformation to them will not change the really observable quantities **E** and **H**. The same is true in our case since, considering the first examined composite mechanism (equations 24, 25), any legitimate set of values of the six quantities K,ΔH, K_{a1}, ΔH_1 and K_{a2}, ΔH_2 will yield the same values of the parameters ($K_a(global)$, $\Delta H(global)$) allowing to interpret an observed TC. Therefore, in the frame of our problem, the six former parameters may be seen as purely auxiliary and devoid of physical significance, whereas only $K_a(global)$ and $\Delta H(global)$ would be physically meaningful. The same kind of conclusions would be reached with the other composite mechanism (equation 34). This interpretation is correct, **but only as far as thermodynamics is concerned** since, as already mentioned, a composite multi-step mechanism could have a specific kinetic signature (visible in the shapes of the injection curves) different from that of the thermodynamically equivalent one-step mechanism.

One could therefore think that the invoked comparison is imperfect since the 'gauge invariance' seen in electromagnetism would not apply strictly in our problem as a distinction should be made between thermodynamic and kinetic considerations. It turns out, however, that the same duality exists in physics when considering quantum and classical electromagnetism. The typical example of the need of considering a quantum description in electromagnetism is the so-called Ahranov-Bohm effect that was predited theoretically [37]. This effect consists in the interference phenomenon that can be observed with a split electron beam that followed two paths on opposite sides of a solenoid, exactly as light interference can be observed

bioRxiv preprint doi: https://doi.org/10.1101/512780; this version posted May 15, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with a split light beam that followed two paths of different lengths. Remarkably, the interference pattern depends on the current in the solenoid and, in turn, on the magnetic field **inside** the solenoid. The crux of the interpretation for the Ahranov-Bohm effect is that **E** and **H**, the only real quantities in classical physics, are null along the electron-beam path **outside** the solenoid and cannot be invoked to explain the observed dependence on the current intensity. Therefore, only the potentials V and **A**, which are non null **outside** the solenoid, can be considered, which imposes to consider them as physically meaningful in the frame of quantum electromagnetism [37].

The proposed comparison is thus fully consistent with the kinetic/thermodynamic opposition in our problem being equivalent to the quantum/classical opposition in electromagnetism. In other words, the potentials V and **A**, which may be seen as dummy in classical electromagnetism, become physically meaningful and necessary to account for quantum effects and, analogously, the parameters describing the composite mechanism (*e.g.* K,ΔH, K_{a1}, ΔH_1 and K_{a2}, ΔH_2), which may be seen as dummy in the strict frame of thermodynamics, may also become physically meaningful and necessary to account for kinetic aspects. This suggests to consider quantum electromagnetism as a more informative "kinetic facet" of a "thermodynamic-like" classical electromagnetism. Whether these considerations are of real heuristic interest, or merely suggestive, requires to be examined in more details.