Analysis of Macromolecular Interactions by Isothermal Titration Calorimetry

A primer on the experimental thermodynamic characterisation of binding of biomolecules

Gonzalo Obal
Protein Biophysics Unit
Structural Biology Platform
Institut Pasteur de Montevideo
Biomolecular interactions: why are they important?

Corpora non agunt nisi ligata

Paul Ehrlich

All (or almost all) biological phenomena depends on interactions between molecules

Antibody-antigen / Enzyme-substrate / Receptor-hormone / Signaling cascades
Protein-lipid / Protein-carbohydrate / Protein-peptide / Protein-DNA
Protein-RNA / ADN-ARN / ...etc, etc, etc.....

Binding of macromolecules is the basis of molecular specificity (ie discrimination between partners and non-partners)

So, without binding there is no biology
Protein-protein interaction network map in yeast.
A remarkable example: metal sensing *E. coli*

a. The volume of an *E. coli* cell is about $1.8 \times 10^{-15}$ L...

...thus, the lowest intracellular Zn$^{2+}$ concentration is about $1 \times 10^{-9}$ M (i.e. 1 ion Zn$^{2+}$ per bacteria!).

b. Zn$^{2+}$ sensor in *E. coli* manage uptake/expulsion of ions, being sensible to concentration below $10 \times 10^{-15}$ M ...

...thus, in typical conditions, intracellular concentration of Zn$^{2+}$ should be of less than 1 átomo Zn$^{2+}$/*E. coli*


Thus...a complete interpretation of any interaction under a particular biological scenario requires the knowledge of both the strength of binding and concentration of the molecules involved

So, consider it when studying binding:

- The biological relevance of Kd depends on the actual concentrations of interacting partners

- Free protein and ligand concentrations will dictate the extent of binding, via Kd

- Thermodynamic and kinetic control of reactions exists.....´even´ in biology
The goals of binding studies are to answer the questions:

**How many?**  **How tightly?**  **How fast?**

**Where?**  **Why?**  **How?**

Binding studies should ultimately elucidate each question in order to provide a complete understanding of a biomolecular interaction......

(this is a long, complex, interdisciplinary task)
What information do we need to get a full characterisation of an interaction?

(from a biophysical point of view)

**THERMODYNAMICS**
- Stoichimetry
- Affinity (strength)
- other thermodynamic properties ($\Delta H$, $\Delta S$, $\Delta C_p$)

**KINETICS**
- rate constants for association ($k_{on}$) and disociation ($k_{off}$)
- Reaction mechanism

**STRUCTURE**
- Three-dimensional structure of the individual interacting partners and their complex(es)

**DYNAMICS**
- Molecular dynamics of individual interacting partners and their complex(es)
How to analyse protein-ligand interactions?

There’s a wide toolkit of methods (with both advantages and pitfalls)...

... qualitative methodologies (yes/no information)...

- Double-hybrid
- Pull-down method (TAP, co-precipitation) *in vivo* and *in vitro*
- Protein arrays

... and quantitatives

- Equilibrium dialysis
- EMSA / native PAGE / Blue Native Page
- ELISA / RIA
- Fluorescence (FRET, quenching, anisotropy)
- Light scattering
- Surface plasmon resonance
- Surface wave resonance
- Asymmetric flow-filled fractionation
- Stopped-flow (coupled to a wide option of detectors)
- ITC
- NMR
- AUC
- QCM (quartz crystal microbalance/piezoelectric acoustic sensor)
The basics of binding interactions

Definition of binding affinity for macromolecular recognition

The binding of two (any number) of proteins can be viewed as a reversible process, in an equilibrium governed by the law of mass action.
The (very) basics of binding interactions

Definition of binding affinity for macromolecular recognition

**Association and Dissociation Constants:**

For a binding reaction at equilibrium:

\[ P + A \rightleftharpoons PA \]

\[ k_1 \quad \text{and} \quad k_{-1} \]

- The rate of formation of \([PA]\) is \(k_1[P][A]\), where \(k_1\) is a second order rate constant with units of \(1 \text{ mol}^{-1}\text{s}^{-1}\).
- The rate of breakdown of \([PA]\) is \(k_{-1}[PA]\); \(k_{-1}\) is a first order rate constant with units of \(s^{-1}\).
- At equilibrium, the rate of formation of \([PA]\) equals the rate of its breakdown, so \(k_1[P][A]=k_{-1}[PA]\).

Also recall that:

\[ K_d (M) = k_{-1}/k_1 = [P][A]/[PA] = 1/K_a \]

Or, in terms of fraction of protein binding sites occupied (\(y\)), which is often convenient to measure:

\[ y = [PA]/[P] + [PA] = K_a[A]/1 + K_a[A] \]

- Divide through by \([P]\)
- Recall \([PA]/[P] = K_a[A]\)

What happens when \(K_a[A] \ll 1\); when \(K_a[A] \gg 1\); when \(K_a[A] = 1\)?
General properties of binding isotherms: Fractional Saturation

\[ Y = \frac{[X]}{([X]+K_d)} \]

- For \( [X] = 0 \) \( \quad Y = 0 \) nothing bound
- For \( [X] \to \infty \) \( \quad Y = 1 \) full occupancy
- For \( [X] = K_D \) \( \quad Y = 0.5 \) half occupancy
Thermodynamic properties of a binding reaction

Binding constants provide an entry into thermodynamics...and vice versa
Thermodynamic properties of a binding reaction

Binding constants provide an entry into thermodynamics...and vice versa

\[ \Delta G = -RT \ln K_{eq} \]

\[ \Delta G = \Delta H - T \Delta S \]

\[ \Delta G = \Delta H_r - T \Delta S_r + \Delta C_p [(T - T_r) - T \ln (T/T_r)] \]

van’t Hoff Expression: \[ \ln K_a = \frac{-\Delta H^\circ + \Delta S^\circ}{RT} \]
-To analyse a binding reaction we need somehow to "see" how substrates are being consumed and/or products are being formed

-So, we'll follow the extent of the reaction

\[ K_A = \frac{[MX]}{[M] \times [X]} = \frac{1}{K_D} \]

Basically, the idea of every binding experiment involves fixing the concentration of one of the interactors (typically M) and varying the one of the other, having found some "signal" that changes proportionally to the amount of complex formed.
...so we need some signal for monitoring a binding reaction (i.e. follow/determine Y) ...

- direct measurement of the concentration of one interactor
  - fluorescence (intrinsic/extrinsic
    - anisotropy
    - FRET
    - heat
    - .........

... ¿which one and how do we use it?

**Two general methods for determining binding constants:**

1) Measure bound vs. Free ligand/protein at equilibrium as a function of concentration

2) Measure association and dissociation rate constants and use these to calculate binding constants
How do we calculate affinity from a titration experiment?

First: \[ K_d = \frac{[X][M]}{[MX]} \]

where: \([MX],[M],[X]\) are the concentrations for free complex, macromolecule, and ligand

Total macromolecule and ligand concentration are:

\[
\begin{align*}
[X]_t &= [X] + [MX] \\
[M]_t &= [M] + [MX]
\end{align*}
\]

So we write \(K_d\) as:

\[ K_d = \frac{[X][([M]_t - [X]_t + [X])]}{[X]_t - [X]} \]

Rearranging terms:

\[
[X]^2 + ([M]_t - [X]_t + K_d)[X] - K_d[X]_t = 0
\]

As:

\[
[MX] = [X]_t - [X],
\]

We re-write:

\[
[MX] = \frac{[X]_t - [M]_t - K_d + \sqrt{([M]_t - [X]_t + K_d)^2 + 4[X]_t K_d}}{2}
\]

\[
[MX] = \frac{[X]_t - [M]_t - K_d + \sqrt{([M]_t - [X]_t + K_d)^2 + 4[X]_t K_d}}{2}
\]

O en su forma más común:

so...we ``follow'' a signal proportional to the amount of complex formed as a function of x.
WORKFLOW OF QUANTITATIVE CHARACTERISATION OF BIOMOLECULAR BINDING INTERACTIONS

- PRODUCTION OF REACTANTS
- CHARACTERISATION OF REACTANTS
- SELECTION OF METHOD(S)
- IMPROVED EXPERIMENTAL DESIGN
- SIMULATION
- BINDING EXPERIMENT
- Equilibrium MODEL SELECTION and/or CONSTRUCTION
- RESULTS
- DATA ANALYSIS (FITTING & STATISTICS)
- BINDING EQUATIONS
- INTERPRETATION
  (BIOLOGICAL SIGNIFICANCE)
Isothermal Titration Calorimetry

Overview and Applications

Theory and Instrumentation

Experimental setup

Concepts of ITC data analysis

Examples: interpreting ITC results in context of structural thermodynamics
Isothermal Titration Calorimetry

• ITC measures the heat uptake or release during a biomolecular reaction
  – Heat is taken up (absorbed, endothermic)
  – Heat is evolved (released, exothermic)

• Calorimetry is the only method that can directly measure the binding energetics of biological processes
# Methods to determine Binding Constants

<table>
<thead>
<tr>
<th><strong>Spectroscopy (Fluorescence, UV/Vis, CD)</strong></th>
<th><strong>Signal</strong>&lt;br&gt;change of absorption or emission of light</th>
<th><strong>Information</strong>&lt;br&gt;$K_D$ ($10^{-4}$-$10^{-11}$M)&lt;br&gt;ΔH, ΔS, n</th>
<th><strong>Advantage</strong>&lt;br&gt;in solution</th>
<th><strong>Disadvantage</strong>&lt;br&gt;probe needed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Microcalorimetry</strong></td>
<td>heat of binding</td>
<td>$K_D$ ($10^{-3}$-$10^{-11}$M)</td>
<td>no labels, in solution</td>
<td>large sample</td>
</tr>
<tr>
<td><strong>Surface Plasmon Resonance</strong></td>
<td>change of refractive index due to mass</td>
<td>$K_D$ ($10^{-3}$-$10^{-13}$M) $k_i, k_s$</td>
<td>small sample, automated</td>
<td>surface coupled, ligand must have large mass</td>
</tr>
<tr>
<td><strong>Stopped-Flow</strong></td>
<td>coupled to spectroscopy</td>
<td>$K_D$ ($10^{-3}$-$10^{-12}$M) $k_i, k_s$</td>
<td>fast</td>
<td>probe needed</td>
</tr>
<tr>
<td><strong>Analytical Ultracentrifugation</strong></td>
<td>absorption at different radii for different times</td>
<td>$K_D$ ($10^{-3}$-$10^{-6}$M)</td>
<td>good for homomeric interactions</td>
<td>slow</td>
</tr>
<tr>
<td><strong>Nuclear Magnetic Resonance</strong></td>
<td>shift of magnetic resonance frequency</td>
<td>$K_D$ ($10^{-3}$-$10^{-6}$M)</td>
<td>in solution, structural information</td>
<td>slow, large sample, expensive</td>
</tr>
<tr>
<td><strong>Binding Assays</strong></td>
<td>various, e.g. SDS-PAGE, densitometry, radioactivity</td>
<td>$K_D$ ($10^{-3}$-$10^{-15}$M)</td>
<td>can be most sensitive</td>
<td>sometimes inaccurate</td>
</tr>
</tbody>
</table>
Microcalorimetry provides binding stoichiometry

• Number of ligand binding sites per macromolecule – on a molar basis, model-independent
  (by convention a ´´Ligand´´ has one binding site, and a ´´Macromolecule´´ can have more than one)
Microcalorimetry is a gold-standard for binding analysis

- Label-free
- In-solution
- No MW limitations
- Optical transparency/clarity unimportant
- Minimal assay development
Microcalorimetry is (almost) universally applicable to study interactions between:

- Protein-small molecule
- Enzyme-inhibitor
- Protein-protein
- Protein-DNA
- Protein-RNA
- Protein-lipid/liposomes
- Protein-carbohydrate
- Other non-biological binding reactions
- Oligomerisation
Microcalorimetry provides a total picture of binding energetics (however, it is not easy to relate physicochemical properties ($\Delta G$, $\Delta H$, $\Delta S$) specifically and directly to binding mechanisms)

**Very broadly:**

i) $\Delta H$ reflects energy changes associated with making (-$\Delta H$) and breaking (+ $\Delta H$) of bonds (hydrogen, van der Walls bonding, solvent)

Entalphy-driven reactions (high affinity)

ii) $\Delta S$ reflects changes associated with increasing (+ $\Delta S$) or decreasing (- $\Delta S$) the number of microscopic configurations (hydrophobic interactions, flexibility, solvent)

Entropy-driven reactions (high specificity)
Importantly, the relative contribution of $\Delta H$ and $\Delta S$ to $\Delta G$ govern/distinguish the functional characteristics of R-L interactions.

This is a major concern in drug discovery and agonist/antagonist mechanism.
The idea is to titrate a macromolecule (in the cell) by serially injecting a ligand (in the syringe), while recording the evolved (absorbed/release) heat.
VP-ITC (GE, Microcal) at UBP-BS/IPMON
ITC – Before titration

- Ligand – in syringe
- Macromolecule in ITC cell
Titration begins: First injection

- Ligand in syringe
- Macromolecule in cell
- Macromolecule-ligand complex

As the first injection is made, all injected ligand is bound to target macromolecule.
Return to baseline

The signal returns to baseline before the next injection.
Second injection

As a second injection is made, again all injected ligand becomes bound to the target.
Second return to baseline

Signal again returns to baseline before next injection.
Injections continue

As the injections continue, the target becomes saturated with ligand, so less binding occurs and the heat change starts to decrease.
Injections continue

As the injections continue, the target becomes saturated with ligand so less binding occurs and the heat change starts to decrease.
End of titration

When the macromolecule is saturated with ligand, no more binding occurs, and only heat of dilution is observed.
Step by step...
...how do we get information from an ITC experiment?

1. titration:

First, we perform `data collection`:

- power (µcal o µJ/seg) as a function of time (s)

Every binding event (injection) is associated to a given amount of heat released (exothermic) or absorbed (endothermic)

\[ P = \frac{dQ}{dt} \]
...recapitulating, the concept is:

1) In each injection we add \( n \) moles of \( X \) to a fixed amount of \( M ([M]_T) \), releasing or absorbing some heat \( q_i \).

2) Thus... at every \([X]_T/[M]_T\), we directly measure its associated \( q_i \)...
Finally...

...given $q_i$ for every injection we have:

$$\sum_{i=1}^{n} q_i = Q_{\text{reacción}}$$

...experiment is done a constant T and P, thus:

$$Q_{\text{reacción}} = \Delta H$$

(Wiseman Isotherm)

(Ajuste a modelo $n$ sitios de unión)

$$\Delta G_{\text{obs}} = -RT \ln K_{B,\text{obs}}$$

$$\Delta S_{\text{obs}} = \frac{(\Delta H_{\text{obs}} - \Delta G_{\text{obs}})}{T}$$

$$\Delta C_{p,\text{obs}} = \frac{d\Delta H_{\text{obs}}}{dT}$$
The shape of binding isotherms depends on the value of $K_D$ (this is important from a practical consideration, as this limits the conditions of the experiment).

### C Values in ITC

\[
C = \left( \frac{[M]_{\text{tot}}}{K_D} \right) \times N
\]

- $C = 10-100$ very good
- $C = 5-500$ good
- $C = 1-5$ and 500-1000 OK
- $C = < 1$ and $> 1000$ not wanted
Displacement ITC – HIV-1 Protease-Inhibitor Binding

Unable to determine KB

KB of $3.1 \times 10^{10}$ M$^{-1}$
Stoichiometry

“N” is the average number of binding sites per mole of protein in your solution, assuming:

- that all binding sites are identical and independent
- that you have pure protein (and ligand)
- that you have given the correct protein and ligand concentrations
- that all your protein is correctly folded and active
Data analysis

i) *Global ITC data analysis*

Global fitting routines to different binding models
Situation could complicate for n>2
Cooperativity is sometimes difficult to analyse (specially negative cooperative)

ii) *Optimisation and Statistical error analysis*

Monte Carlo simulations of binding data
Variance-Covariance Matricial analysis of statistical significance

iii) *Global Multimethod Analysis (GMMA)*

Sometimes ITC (or any other technique) data are not sufficient for obtaining a plausible explanation
Combination of other methodologies are required (orthogonal information)
Other complication: proteins (biomolecules) are dynamics (native ensemble)
Other complication: proteins (biomolecules) are dynamics (native ensemble)

- Induced fit
- Conformational selection
- Conformational selection + Induced fit
A: $\text{SOS}^{\text{DH-PH-cat}}$ vs $\text{SOS}^{\text{Histone}}$R153A

- Time (min)
- $\mu$cal/sec
- kcal/mole of injectant

$N = 0.91 \pm 0.01$
$K_d = 2.52 \pm 0.35 \mu M$

B: $\text{SOS}^{\text{DH-PH-cat}}$ vs $\text{SOS}^{\text{Histone}}$D140A

- Time (min)
- $\mu$cal/sec
- kcal/mole of injectant

no binding

C: $\text{SOS}^{\text{DH-PH-cat}}$R552A vs $\text{SOS}^{\text{Histone}}$R153A

- Time (min)
- $\mu$cal/sec
- kcal/mole of injectant

no binding
cAMP binding to catabolite inhibitory protein CAP

CAP is an homodimeric protein

what does the ITC shows?

Suggests two sites (independent? sequential?)
...where binding of a second cAMP molecules is entropically disfavourable
How can we explain negative cooperativity from the thermodynamic properties of?

Check for entropy change (conformational) of the protein during cAMP binding...
Thus, a thermodynamic model can be propose:

**No cAMP:**
Both monomers are flexible
Affinity for cAMP is identical at both sites

**Binding of 1 cAMP:**
Bound monomer rigidifies

**Binding of 2 AMP:**
All protein rigidifies
Thanks!

Questions?