20.320, notes for Sep 11

Tuesday, September 11, 2012
9:35 AM
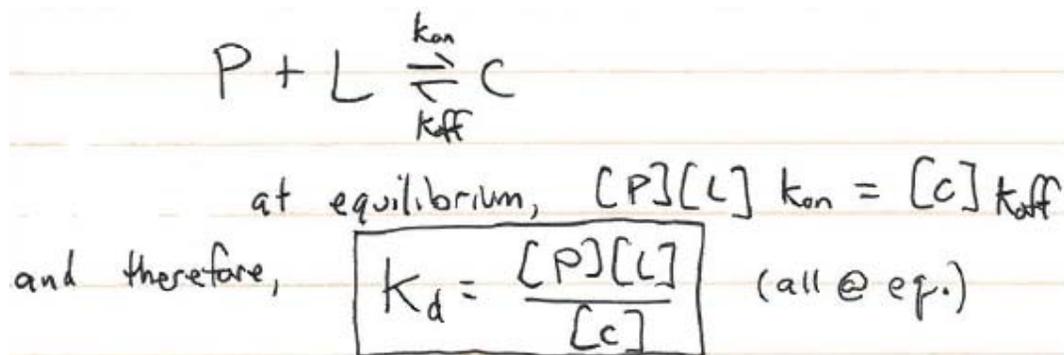
Today we'll talk about some experimental techniques that can give us the useful parameters for our models.

Titration analysis

We start simple.

1. Attach protein to plate
2. Add fluorescently labeled ligand
3. Rinse off unbound ligand
4. Read fluorescence on plate
5. This test provides $[C]_{eq}$, the complex concentration at equilibrium.

It's called a titration because we'll do it first with a given amount of ligand, then repeat with a larger amount. As we increase the amount of ligand in solution, we get different equilibrium complex concentrations. This doesn't yet tell us, however, how this relates to our useful parameters. When we go back to our original equation,

$$P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} C$$

at equilibrium, $[P][L]\, k_{on} = [C]\, k_{off}$

and therefore,

$$K_d = \frac{[P][L]}{[C]} \quad \text{(all @ eq.)}$$

This is not perfect, though, because we don't always know [P] or [L]. A more useful term, then, is called fractional saturation.
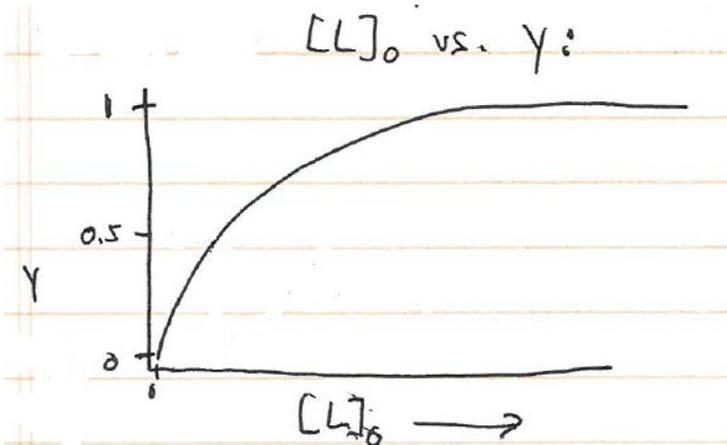
Fractional saturation

$$Y = \frac{[C]_{eq}}{[P]_0} = \frac{\text{amount of complex @ eq.}}{\text{total initial protein}}$$

It ranges from 0 to 1, and is composed only of terms that we can measure with some accuracy. How does Y, then, relate to K_d?

Given that $[P]_0 = [C] + [P]$,

$$Y = \frac{[L]}{[L] + K_d}$$

Now, the graph of Y versus $[L]_o$ looks like this:

$[L]_o$ vs. Y:

HOWEVER, our equation for Y doesn't contain $[L]_o$, it contains only $[L]$.

$$Y = \frac{[L]}{[L] + K_d}$$

See?

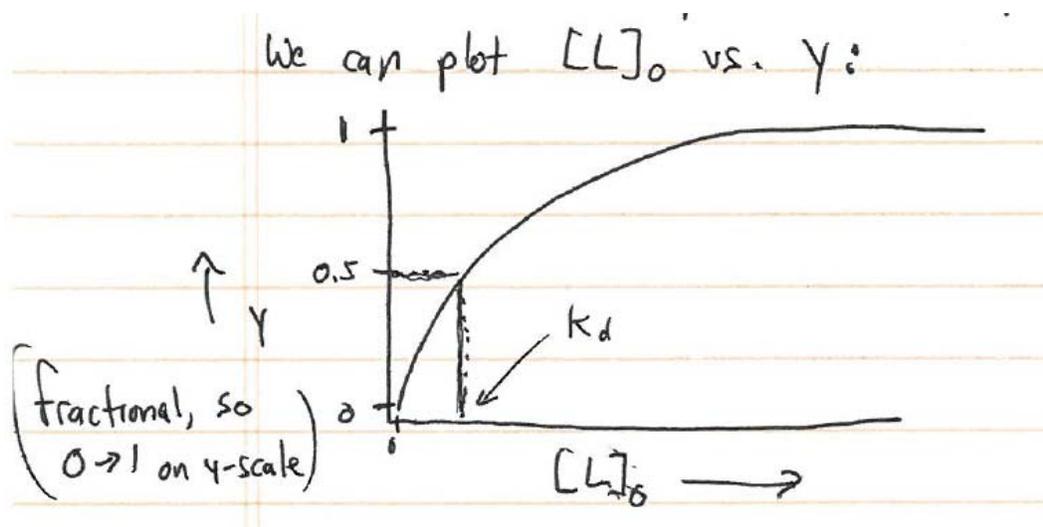## The Pseudo-First Order Approximation, PFOA

$[L]_o$ is not usually the same as $[L]$, but there are some conditions where they're more or less equivalent. If we have an enormous initial excess of ligand (excess $[L]_o$, then reaction with all the protein will barely affect the ligand concentration.

In this case, $[L]_o \gg [P]_o$ and therefore $[L] \cong [L]_o$.

And this limiting case of excess ligand is called the Pseudo-First Order Approximation (PFOA). And if we make our experiments within the acceptable range for the PFOA, then, we can make our assumption and draw this graph again.

$$Y = \frac{[L]_o}{[L]_o + K_d} \quad \left(\begin{array}{c}\text{pseudo-first}\\\text{order approx.}\end{array}\right)$$

Notice how, by simplifying the equation, we can figure out that y will equal 0.5 when $[L]_o$ equals K_d.

We can plot $[L]_o$ vs. y:

$\uparrow$ Y

(fractional, so $0 \rightarrow 1$ on y-scale)

$[L]_o \longrightarrow$

K_d

Where is K_d?

$$K_d = [L]_o \text{ for } Y = 0.5$$

And what values of K_d correspond to a weak or strong interaction?

- Strong Interaction : $K_D = 10^{-8} \rightarrow 10^{-12}$ M
  weak Interaction : $K_d = 10^{-3} \rightarrow 10^{-5}$ M

Notice the pattern. **A smaller K_d corresponds to a stronger interaction.** Think of K_d as the ligand concentration necessary to complex half the available proteins at equilibrium.

Outside PFOA

Now, what happens when we can't assume PFOA because the ligand in solution is simply not in sufficient excess? We must do some math. It's not trivial, but it's solvable.

$$[L]_{eq} = [L]_0 - [c]_{eq} \quad (\text{plug in for } [c]_{eq}?)$$

$$\nearrow \quad y = \frac{[c]}{[L] + k_d}$$

$$[c]_{eq} = [c]_0 - y[P]_0 \quad (\text{plug into } \text{ equation for } y)$$

$$y = \frac{[L]_0 - y[P]_0}{[c]_0 - y[P_0] + k_d} \qquad \text{solve for } y$$

$$y = \frac{\left(k_d + [L]_0 + [P]_0\right) - \sqrt{\left(k_d + [L]_0 + [P]_0\right)^2 - 4[P]_0[L]_0}}{2[P]_0}$$

Isothermal Titration Calorimetry, ITC

And so, titration can give us a decent approximation of K_d. However, there are cases where we want more information or we also want some other thermodynamic characteristics. That's where ITC can come in. Like all calorimetry, it relies on measuring heat flow from a reaction. The calorimeter is made from two cells, one reference and one sample, both containing about 1ml of the same solution but only one containing protein.
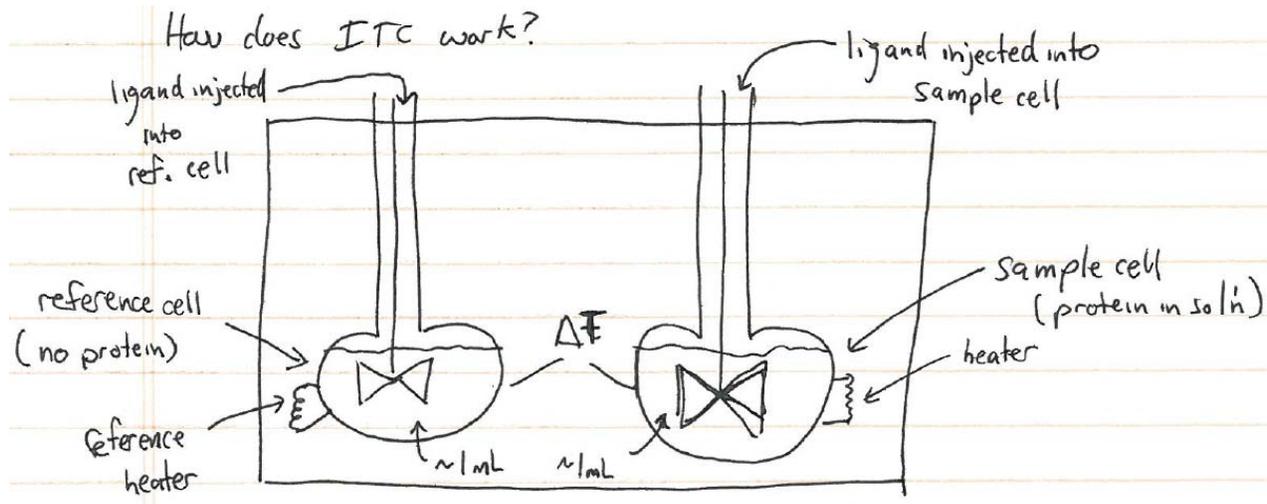

Both chambers have heaters, which are used to equalize chamber temperatures and for which the precise heat output can be measured. The idea is to heat up both chambers slightly, at the same time that you add ligand. If the protein-ligand coupling reaction releases or absorbs heat, a different amount of heat will be needed for the sample chamber to reach a given temperature. Since you're keeping them at the same temperature, what you do is measure the heat flow required for this. You make very precise and continuous heat flow measurements.

A plot of heat flow versus time shows you how the flow is affected by the addition of ligand (which starts the exo- or endothermic reaction).
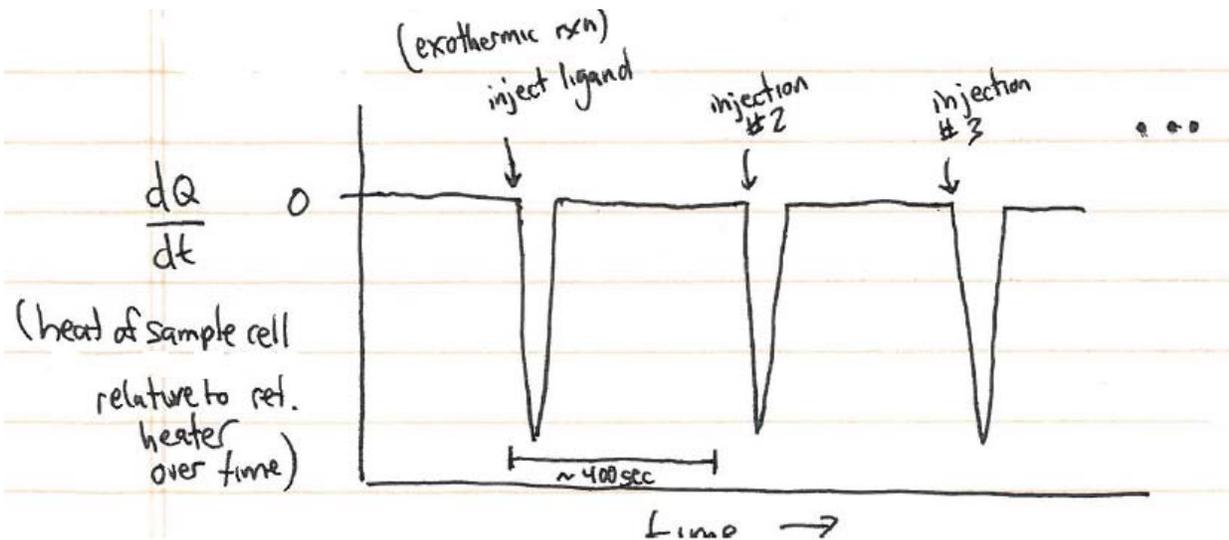
As the experiment continues, each new addition of ligand has a smaller effect. There is less and less free protein to form complexes, so less heat is released for the same amount of ligand.
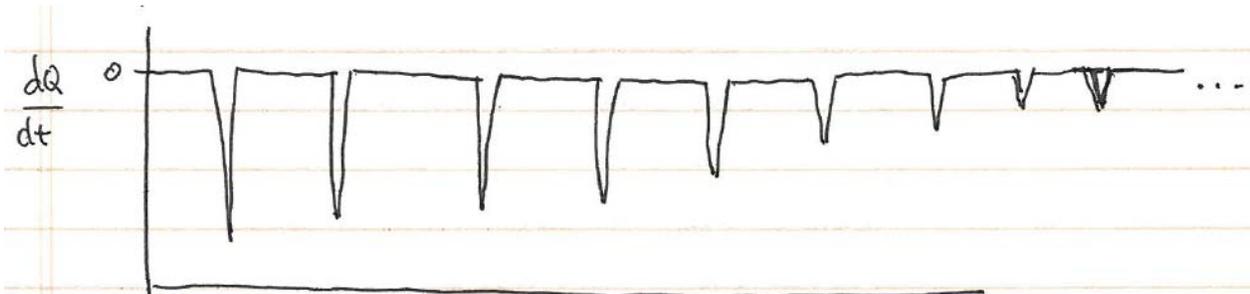
And we can integrate the peaks over time. If we curve-fit, we'll be similarly able to very roughly estimate K_d.
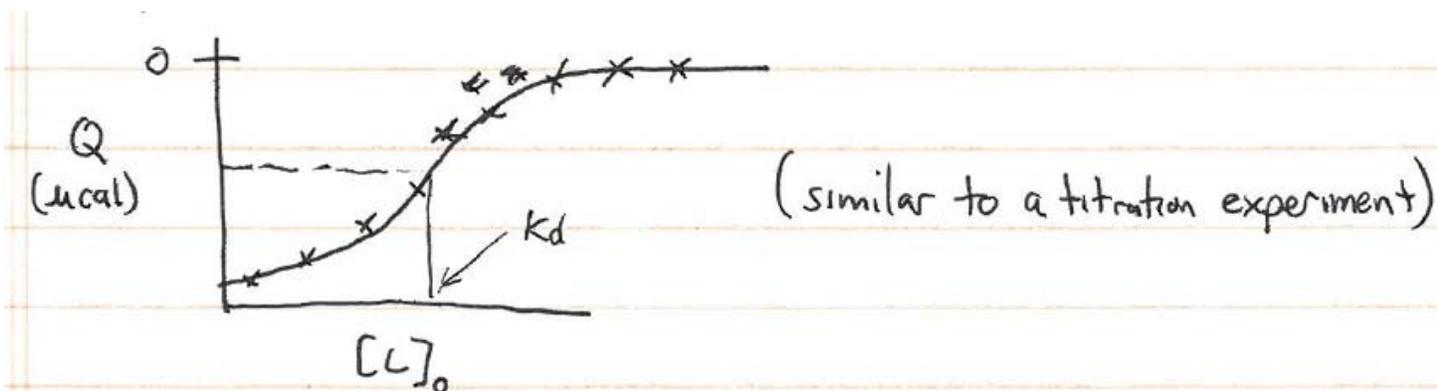
The reason why the estimate is rough is that we're not actually within PFOA. We are depleting the ligand in each run, so it's clearly not in great excess. It's an easy start, though, so it's useful. Calorimetry might not be the best at calculating K_d, but it does yield interesting thermodynamic parameters.

Enthalpy from ITC
Enthalpy is the easiest, since it is closely related to the amount of heat transferred (plus some other parameters that we can figure out).

$$Q_i = V \, \Delta H_{app} \left([c]_i - [c]_{i-1} \right)$$

complex formed from the ith injection
– complex formed from the i-1 injection

heat from ith injection

V = Volume of cell, in liters (typically $10^{-3 \text{ to}} L$)

ΔH_{app} = Apparent enthalpy, in cal/mol

The least simple thing to calculate here is $[C]_i$. We can't just assume that all the ligand going in will react. At this point, we might want to ask what we *do* know.

We know: K_d, $[L]_o$, $[P]_o$

And with these parameters we can calculate Y:

$$Y = \frac{\left(K_d + [L]_o + [P]_o \right) - \sqrt{\left(K_d + [L]_o + [P]_o \right)^2 - 4[P]_o[L]_o}}{2[P]_o}$$

And since we know Y, we can also solve for [C]:

$$Y = \frac{[C]}{[P]_o}$$

And so, we can use all of the above to solve for **ΔH**.

Entropy from ITC

From K_d we can calculate ΔG, and with that and ΔH we can get ΔS.

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

And so, with a single 2-hour experiment we can calculate K_d, ΔG, ΔH, and ΔS. These are extremely useful parameters for our models, and a great way to get started.

Limitations of ITC

- **You need a lot of both protein and ligand** for this to work. This is the rule of thumb, the relationship between $[P]_o$ and K_d necessary for ITC to work.

$$10 < \frac{[P]_0}{k_d} < 100 \quad \$ \quad [L]_0 \sim 10[P]_0.$$

So if $k_d \approx 10^{-7}$ M
$$[P]_0 \approx 10^{-6} \text{ or } 10^{-5} M$$
$$\$ \quad [L] \approx 10^{-5} \text{ or } 10^{-4} M$$

This might not seem like a problem, but each ligand injection is diluted about 1:100. Therefore, you need a really pure and concentrated solution. Protein is similar. Given the above, for a 50k Da protein you'd need about 50 to 500µg of protein at hand. This is **a lot**. That's why getting enough protein and ligand is the hardest part about ITC. If you're screening multiple ligands , for example, the cost and effort quickly becomes ridiculous. People do this only when they really need to calculate ΔS and ΔH.

- **The K_d range is limited.** A K_d between 10^{-5} and 10^{-5} is ideal. If the K_d is $< 10^{-8}$ (that's a very strong interaction) then at first almost all the ligand is bound with each injection. At some point, you suddenly deplete the protein and get no more binding. The curve then becomes a square wave.
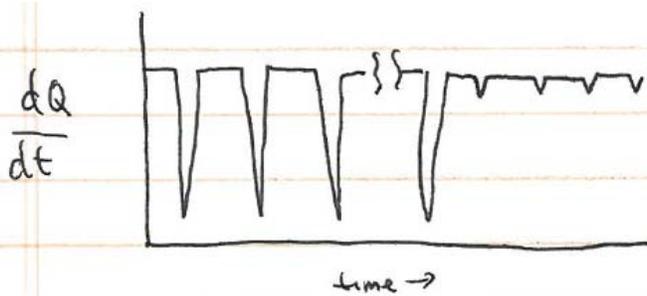
effectively a square-wave rather than a curve. Tough to estimate k_d.

If you have the opposite problem, too weak of an interaction (high K_d), then the signal is very low. There's almost no heat generated, and the protein/ligand is not significantly depleted. You don't want to be the grad student stuck with this experiment.
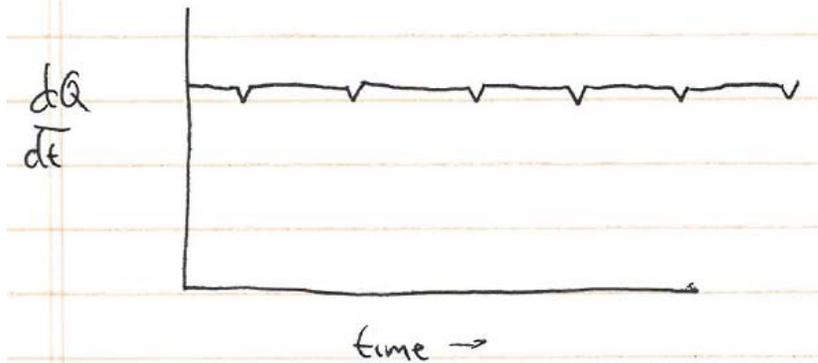
- With 2 hours per experiment, it is **slow**.
- **No kinetic data**, but good equilibrium.

Benefits of ITC

With all these limitations, why would anyone use ITC?

① One of the few technique to provide ΔG, ΔH, ΔS & potentially other thermo parameters.

② Label free, in solution, no modification, immobilization, etc...

③ Applicable to most systems: large proteins, DNA, small molecules

How do you detect interaction?

So say you're not exactly calculating the strength of an interaction ,but rather you wonder whether there's any interaction at all. How do you know whether two proteins interact?

An interesting point to note is that any technique you use must be able to tell the difference between protein, ligand, and their complex.

- FRET (depends on fluorescence).
- Co-IP (followed by mass spectrometry).
- Microscale Thermophoresis (MST).
  - From Wikipedia,
    Microscale thermophoresis is the directed movement of particles in a microscopic temperature gradient. Any change of the hydration shell of biomolecules due to changes in their structure/conformation results in a relative change of movement along the temperature gradient and is used to determine binding affinities, binding kinetics and activity kinetics. Events such as thephosphorylation of a protein or the binding of small molecules to a target can be monitored. MST allows measurement of interactions directly in solution without the need of immobilization to a surface (immobilization-free technology).

    Pasted from <http://en.wikipedia.org/wiki/Microscale_thermophoresis>

Let's focus a little bit on Mass Spectrometry

Mass Spectrometry

Mass Spec provides a really convenient way of measuring the ratio of C to P & L. If the two proteins have a different mass-to-charge ratio (m/z), we will see a different peak for each of the proteins and for the complex.

Change in m/z ratio:

Intensity

ligand

protein

m/z →

1:1 P:L complex

2nd 1:2 P:L complex
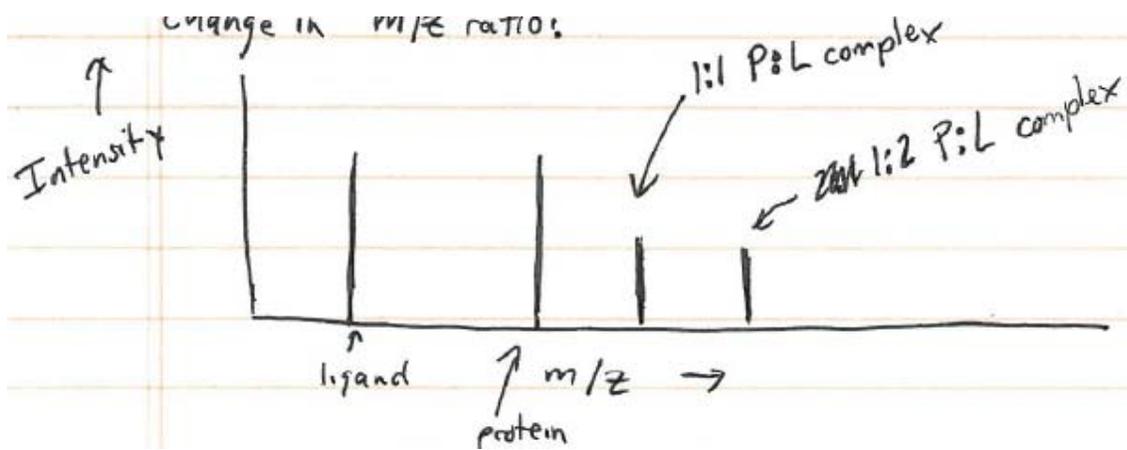
The nice thing about this experiment is that it gives you the stoichiometry of the reaction. Just look at the m/z values of the different signals we get. Note that for all Mass Spec experiments we're assuming that the complex can survive the process of ionization and analysis without breaking into protein and complex.

So why is mass spec so popular? Because it's really simple . We could use fluorescence, but we'd have to develop tags for each different compound. Mass spec, on the other hand, allows one to screen large compound libraries quickly and cheaply. That's why it's such an important tool of the pharmaceutical industry.

20.320 Analysis of Biomolecular and Cellular Systems
Fall 2012