

Model Fitting using Excel (and Gnuplot)

Biochemistry Boot Camp 2023

Session #8

Nicholas Fitzkee

nfitzkee@chemistry.msstate.edu

Think and Discuss

What is a scientific model?

Properties of Models

- Explain an observable quantity (e.g. measured heat, growth rate, etc.)
- Express quantity in terms of understandable parameters and fundamental constants (equilibrium constant, rate constants, etc.)
- Should be predictive (so we can test the model)

Examples: Biochemical Models

- Single site binding: $P + L \rightleftharpoons PL$

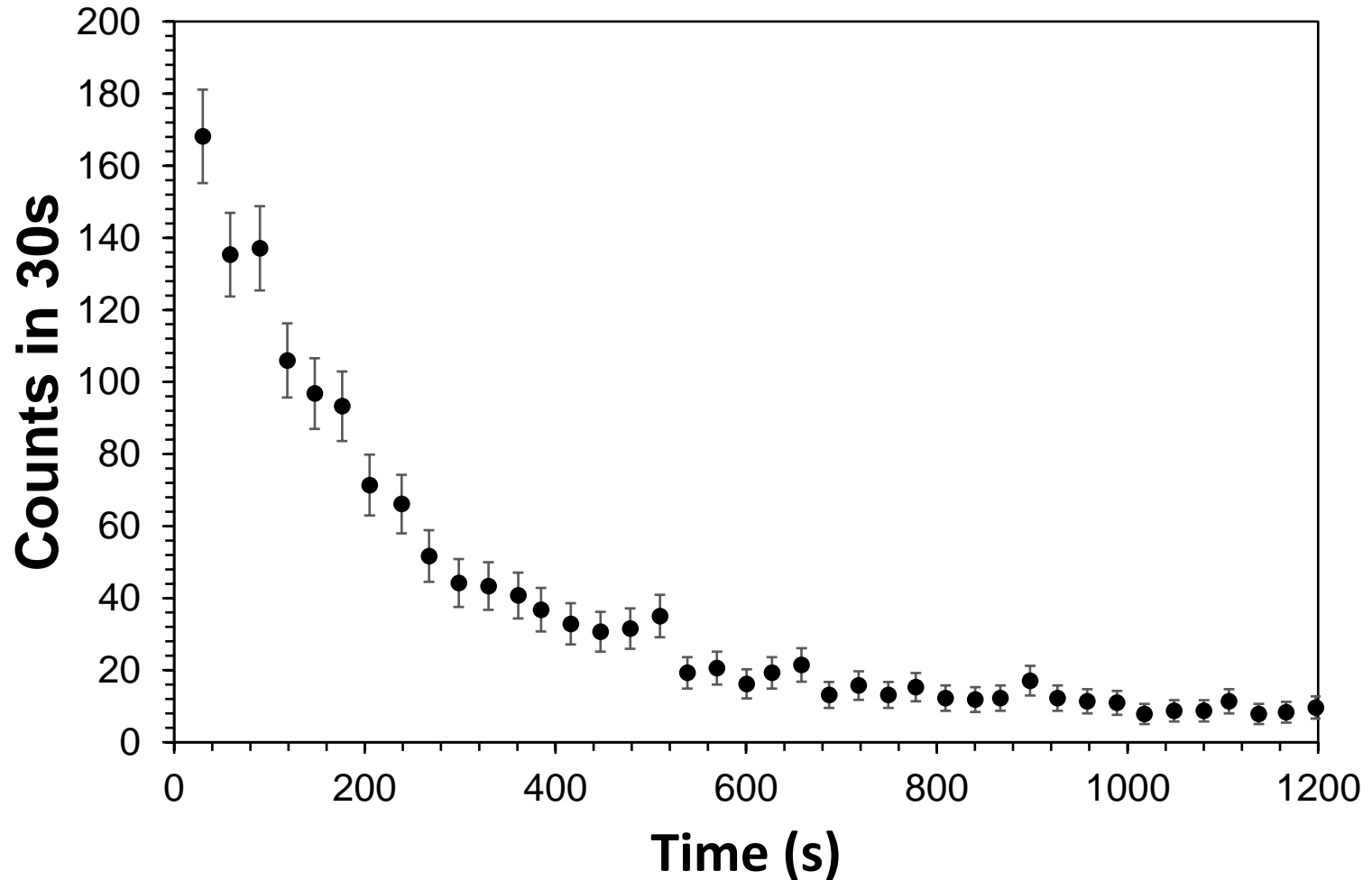
$$\bar{v} = \frac{[PL]}{P_0} = \frac{(P_0 + L_0 + K) - \sqrt{(P_0 + L_0 + K)^2 - 4P_0L_0}}{2P_0}$$

- Folding vs. Temperature: $N \rightleftharpoons U$

$$f = \frac{1}{1 + K} = \frac{1}{1 + e^{-(\Delta\bar{H}^0 - T\Delta\bar{S}^0)/RT}}$$

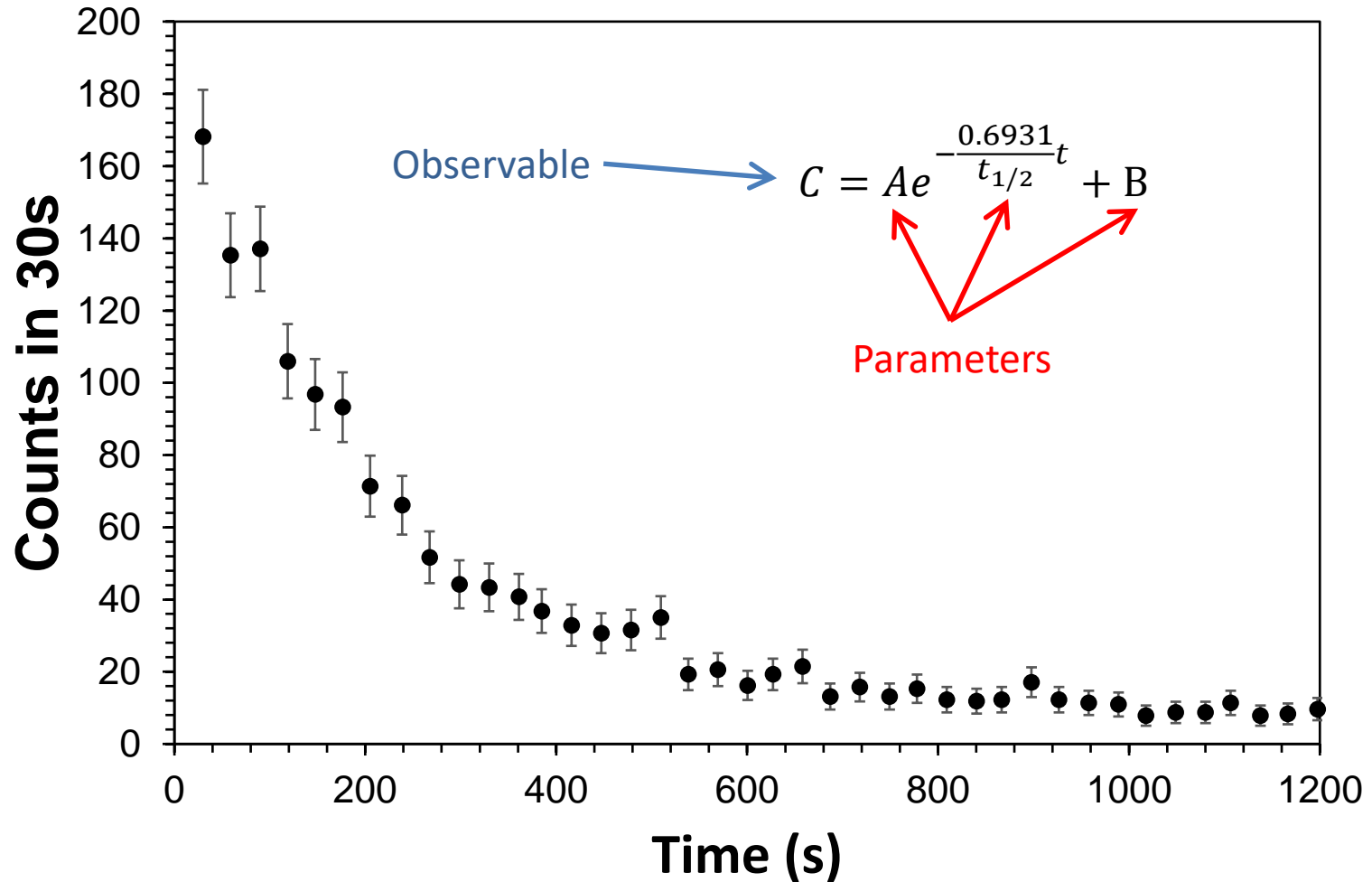
Example Radioactive Decay

Radioactive Decay of ^{137m}Ba



Example Radioactive Decay

Radioactive Decay of ^{137m}Ba



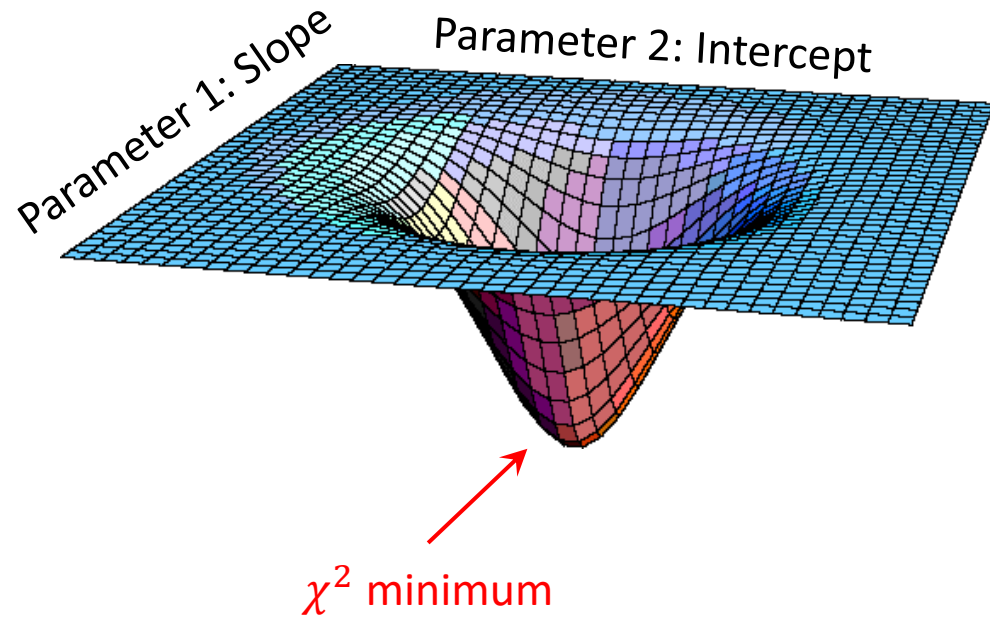
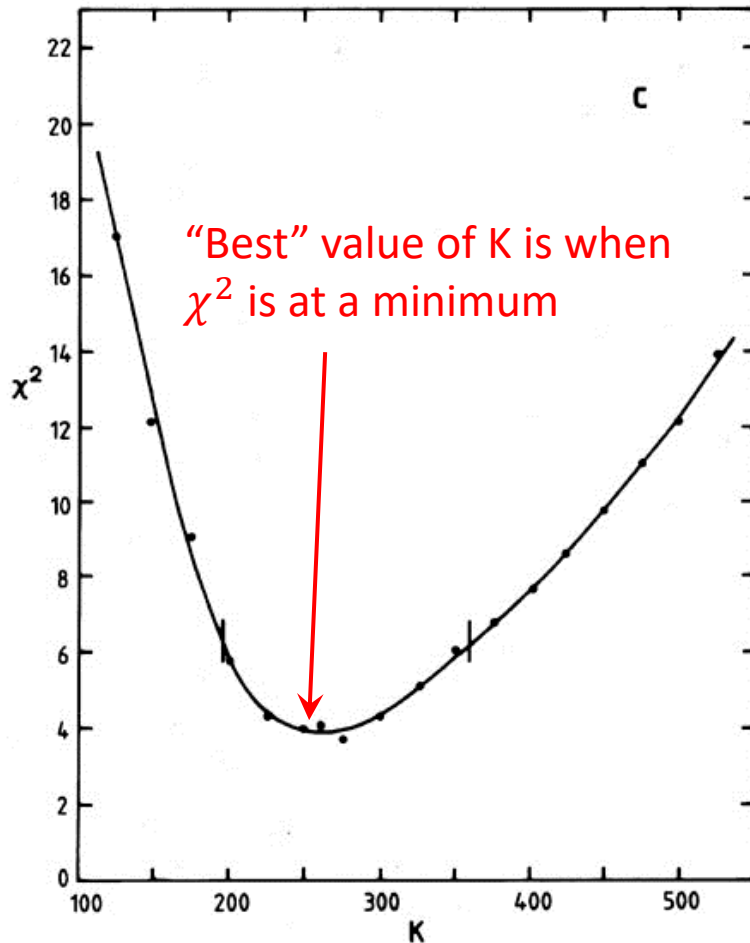
How to Find Parameters?

- Minimize the difference between the observed and model-calculated values:

$$\chi_n^2 = \frac{1}{N_d} \sum_i \left\{ \frac{1}{\sigma_i^2} [y_i - f(x_i)]^2 \right\}$$

- Definitions:
 - y_i = observed data point i
 - $f(x_i)$ = model calculated point x_i (will change when parameters are changed)
 - σ_i = uncertainty for point i
 - N_d = # of observations - # of parameters

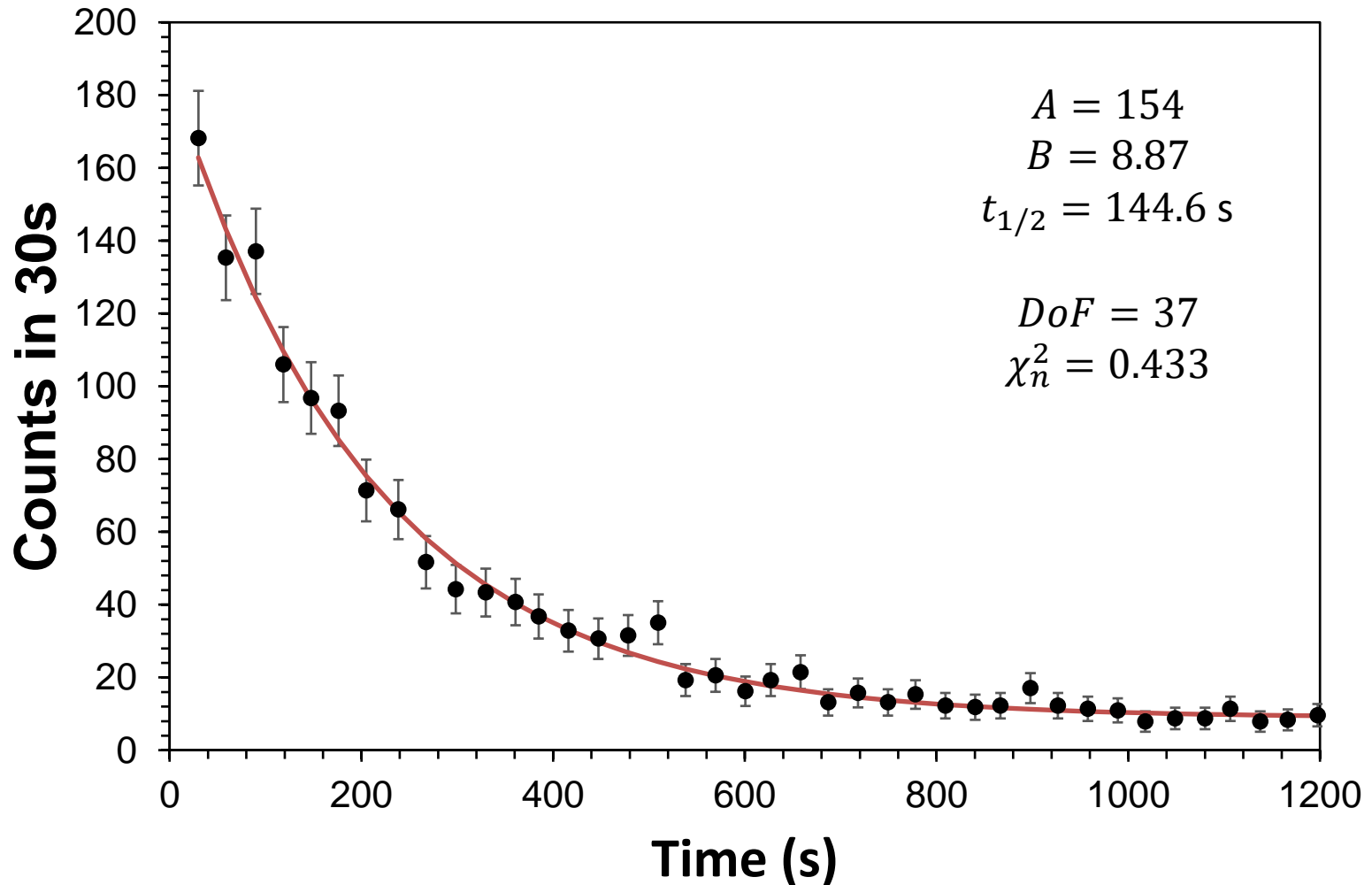
Minimizing Chi-Square



- Minimizing χ^2 works for multiple parameters, too
- Finding the global minimum can be hard

Example Radioactive Decay

Radioactive Decay of ^{137m}Ba

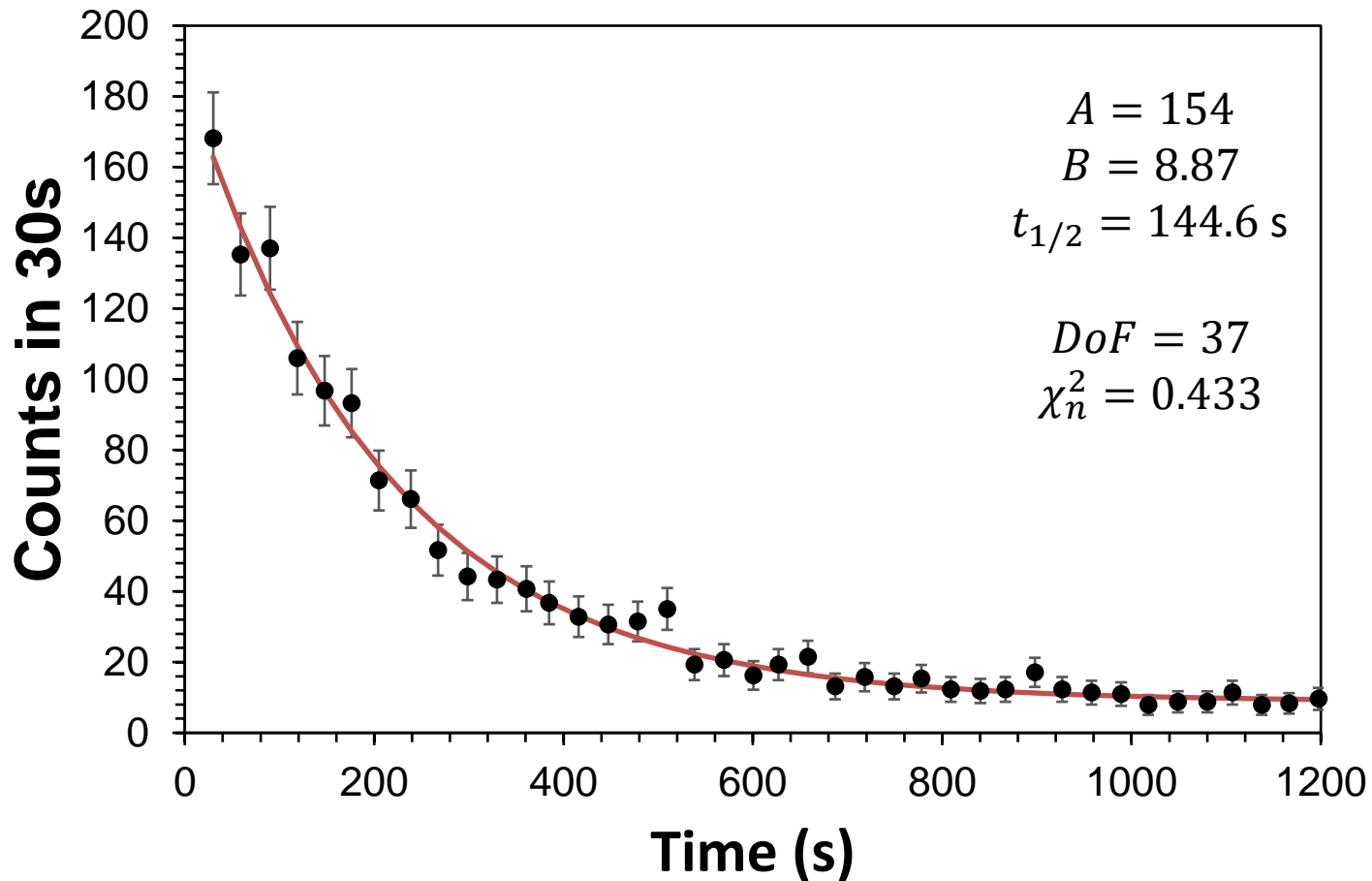


Think and Discuss

The accepted half-life for $^{137\text{m}}\text{Ba}$ is 153 s. Is our value of 144.6 good?

How to Assess a Fit:
#1: Does the fit look good?

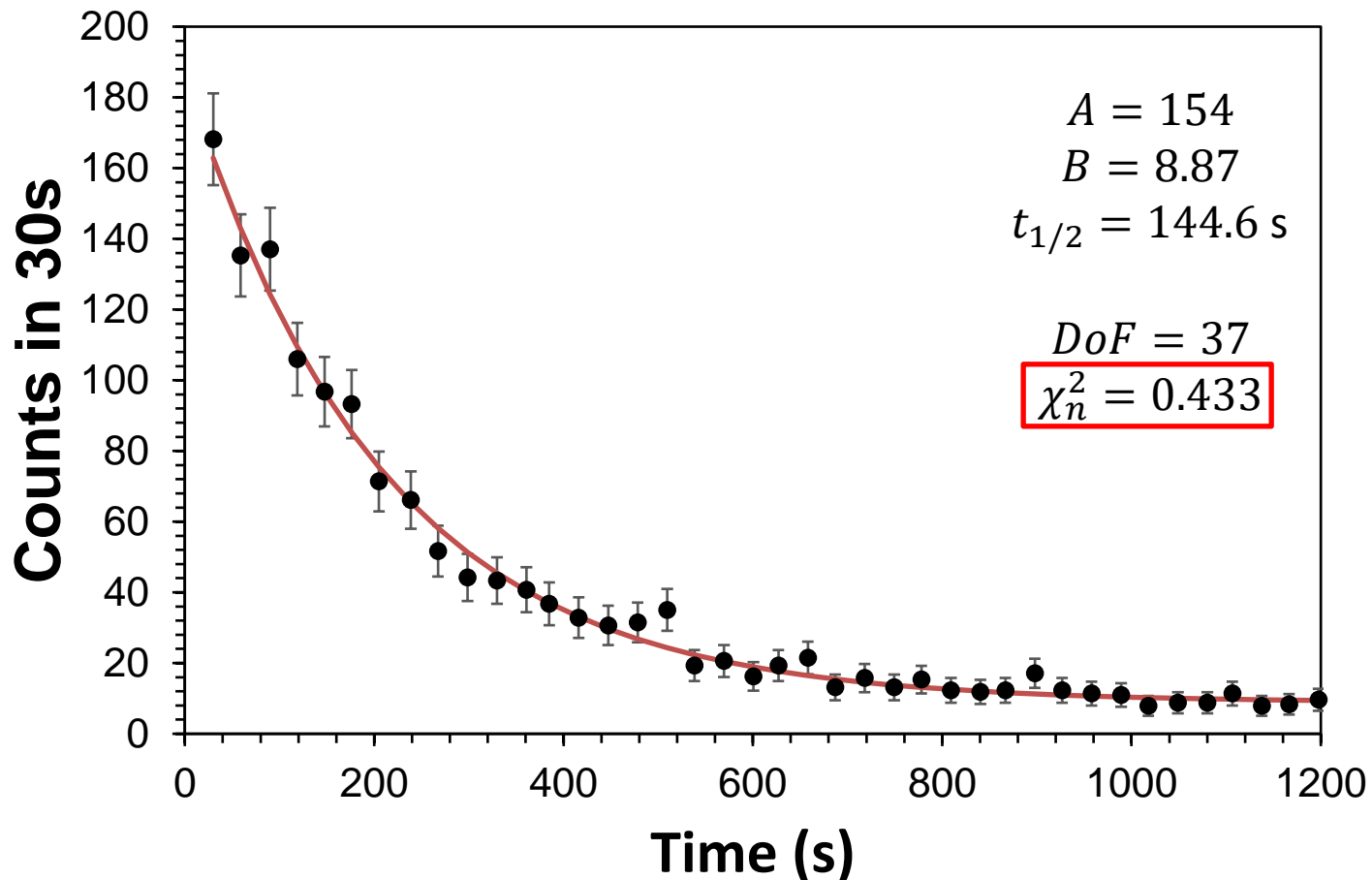
Radioactive Decay of ^{137m}Ba



How to Assess a Fit:

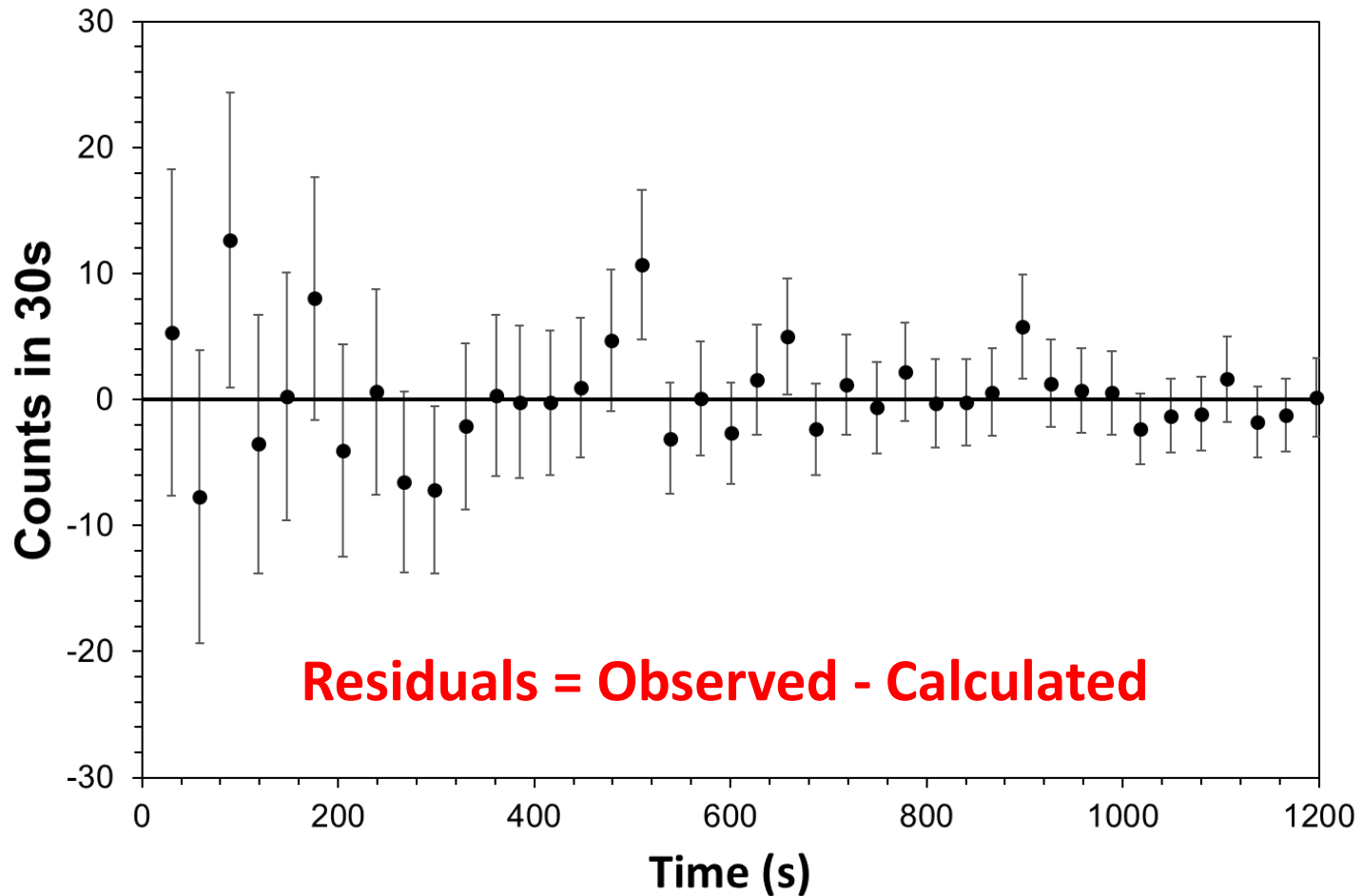
#2: Is Normalized Chi-Square (χ_n^2) ≤ 1.0 ?

Radioactive Decay of ^{137m}Ba



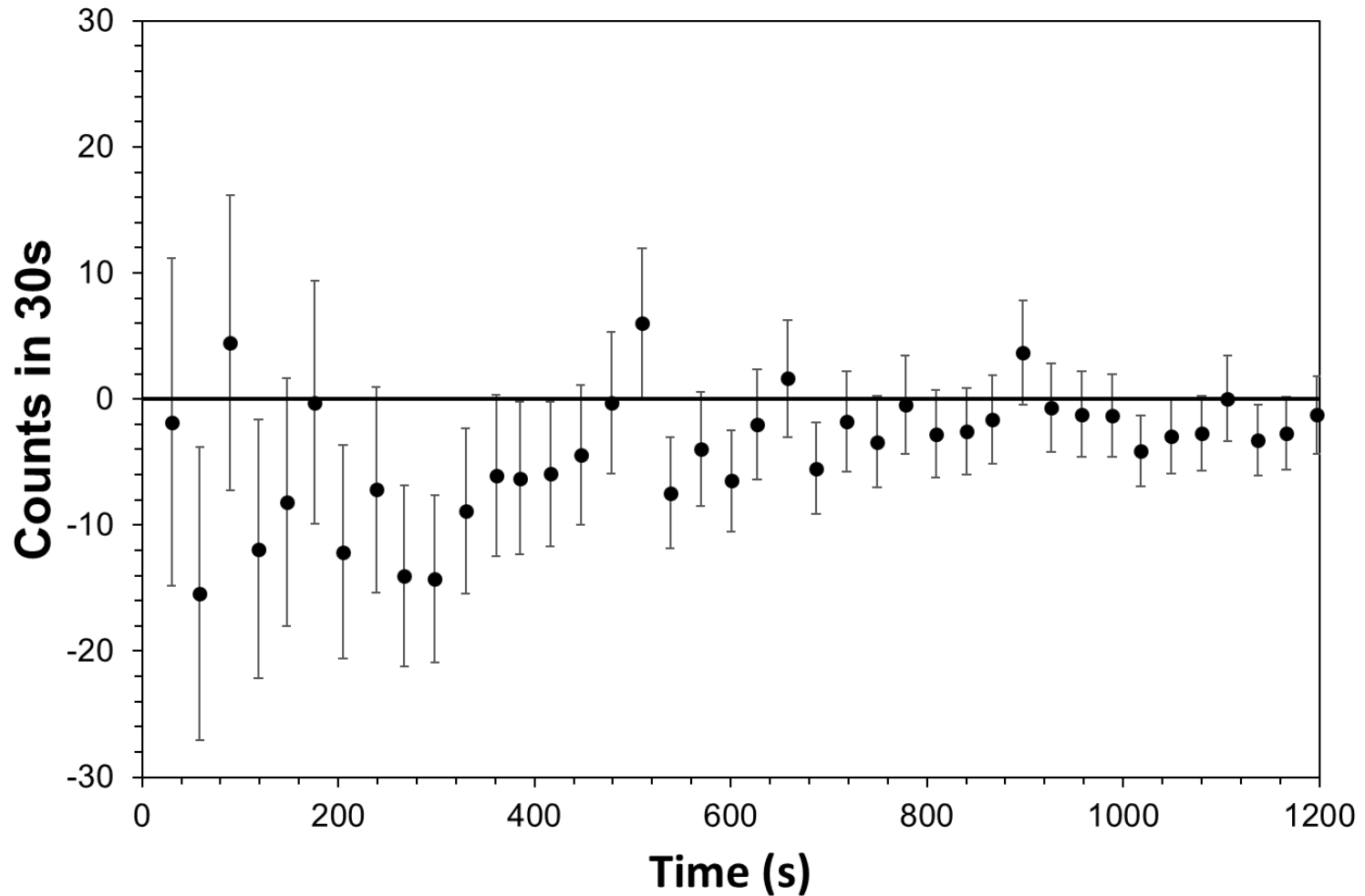
How to Assess a Fit: #3: Check the Residuals Plot

Radioactive Decay of $^{137\text{m}}\text{Ba}$



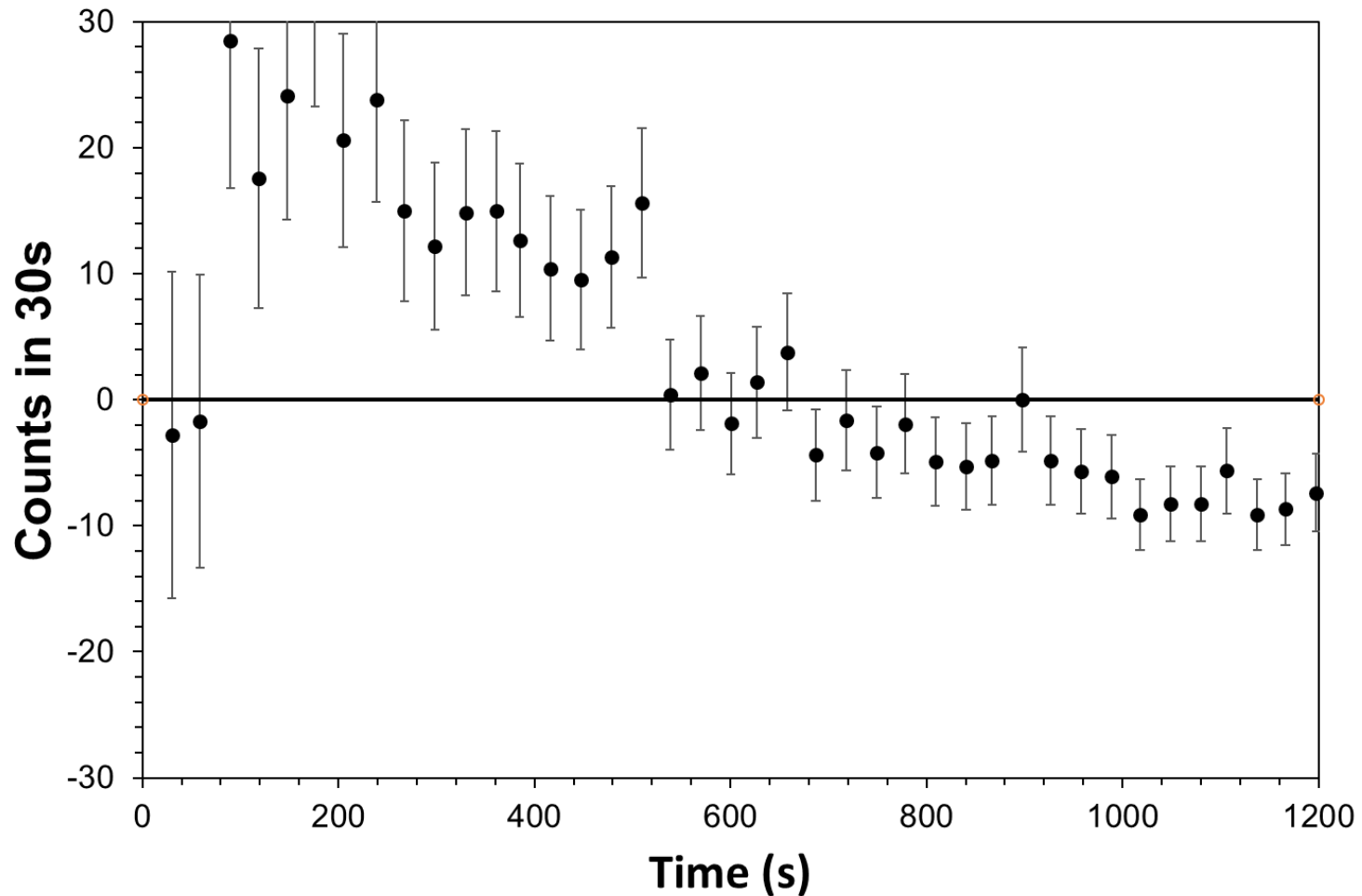
Bad Fit Example: Model Over-Predicts Data

Radioactive Decay of ^{137m}Ba



Bad Fit Example: Systematic Trends in Residuals

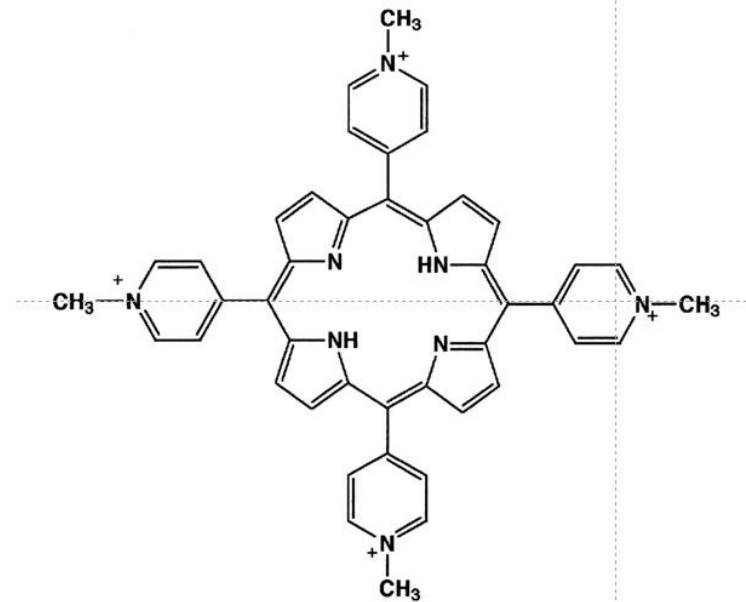
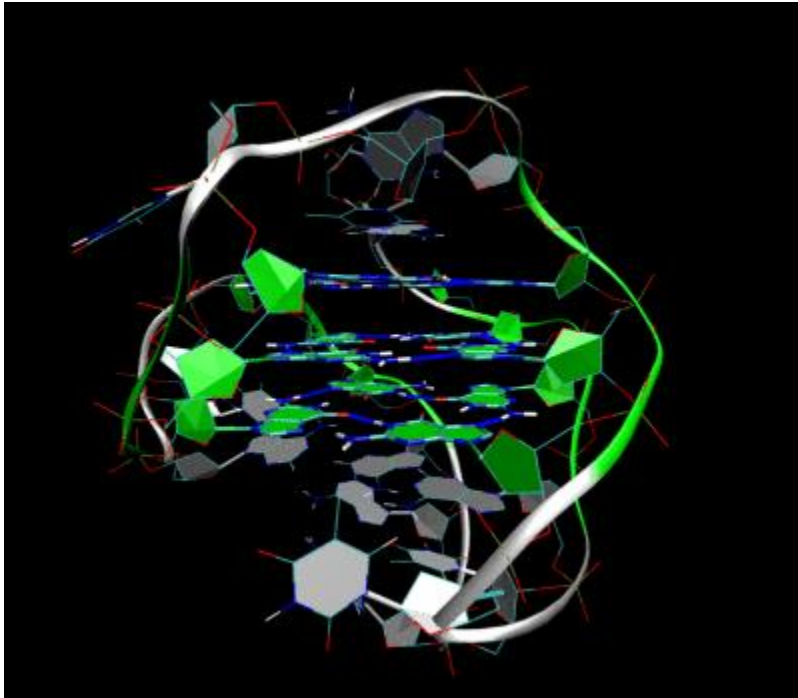
Radioactive Decay of ^{137m}Ba



Biochemical Model Building

- **Step 1:** Come up with a hypothesis about how a system works
 - How many binding sites?
 - Is there cooperativity?
- **Step 2:** Translate the qualitative hypotheses into an observable mathematical form with *parameters*
 - Example parameters: K , τ , N
 - Parameters may not be known
- **Step 3:** Design an experiment that that can produce observables from step 2; perform the experiment
 - *Optimize* the parameters to make the fit look as good as possible
- **Step 4:** Assess the fit – Is the agreement convincing?

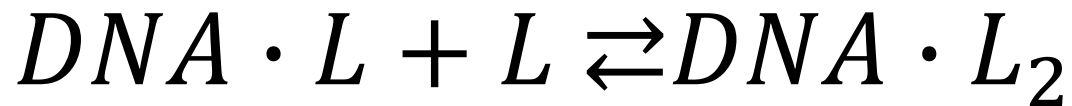
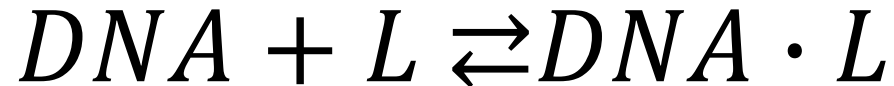
Practical Example: DNA Binding



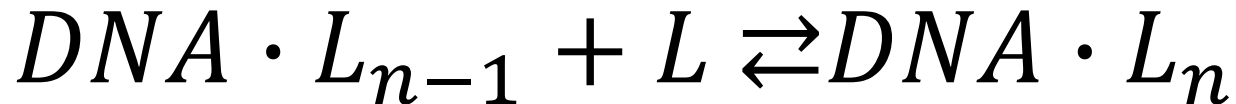
TMPyP4

- What is association constant?
- How many binding sites?

Practical Example: DNA Binding



...



- Assumption in our model: All equilibrium constants are identical (equivalent binding)

Practical Example: DNA Binding

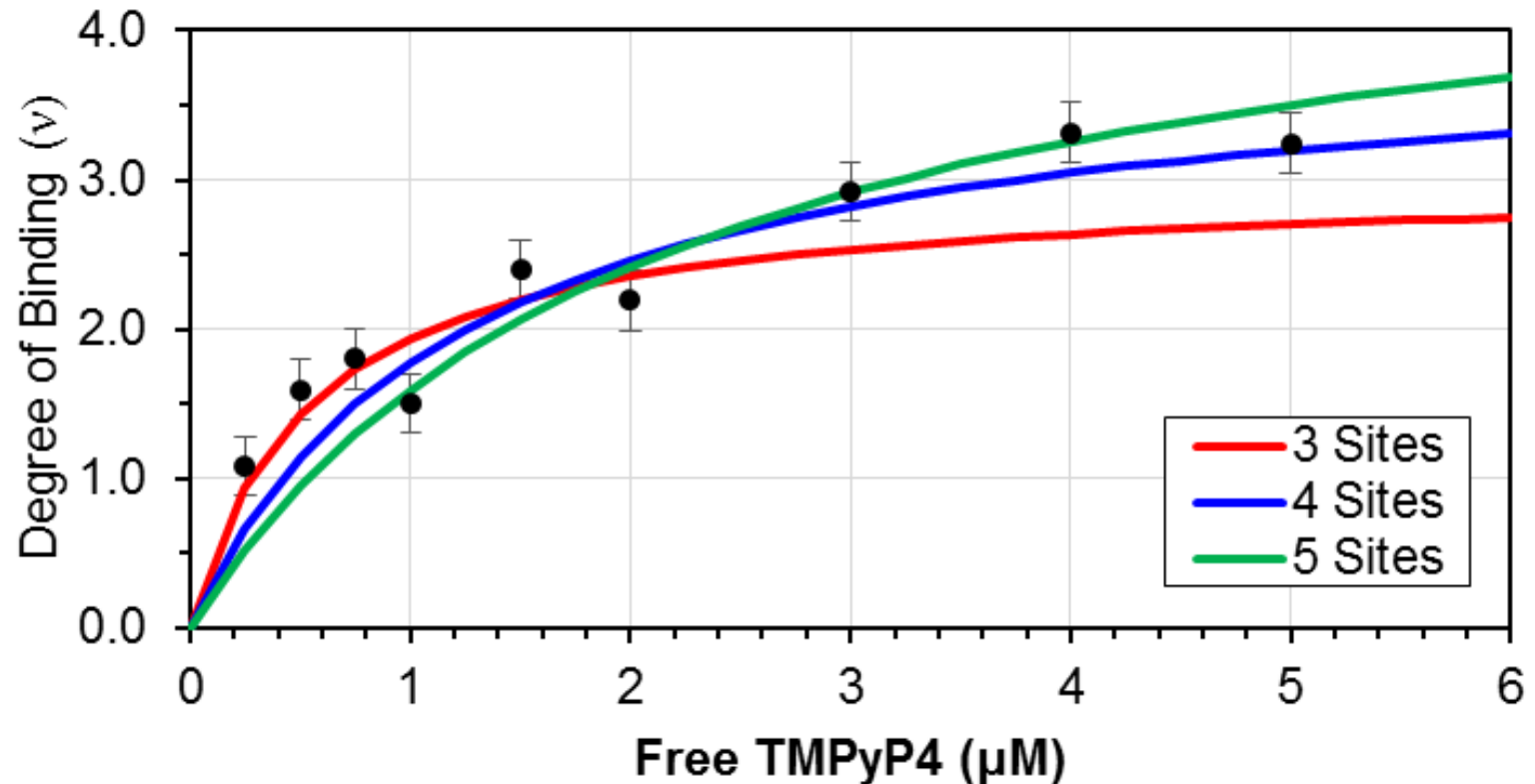
- Experimentally, it is possible to measure the degree of binding (\bar{v}) as a function of free drug concentration (L):

$$\bar{v} = \frac{nK_{eq}L}{1 + K_{eq}L}$$

- We observe \bar{v} and L ; K_{eq} and n are the parameters.

Practical Example: DNA Binding

Binding of TMPyP4 to Quadruplex DNA



- Download this data as `binding.txt` from the website.

The Power of Gnuplot

- Minimizing chi-square can be very tricky; Excel isn't always up to the task
- What about estimating uncertainties of the parameters?
- Gnuplot is a tool that can handle some of these challenges (see tutorial)

Summary

- Biochemical models allow us to quantify and predict the behavior of biological systems
- Fitting parameters allow us to optimize agreement between model and observations
- Programs like Gnuplot enable generalized model fitting with statistical analysis