

Data Fitting, Errors and Binding Curves

Chris Johnson and Stephen McLaughlin

Biophysics

MRC Laboratory of Molecular Biology

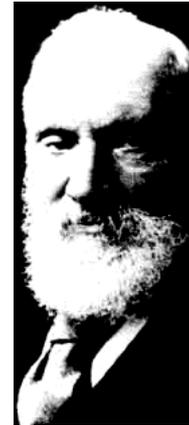


Overview

- Properties of Measurement and data
- General data fitting
- Evaluating 'goodness of fit'
- Fitting binding curve

Measurement according to William Thomson aka Lord Kelvin

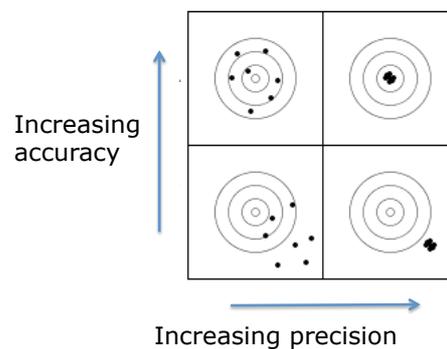
"when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of Science, whatever the matter may be."

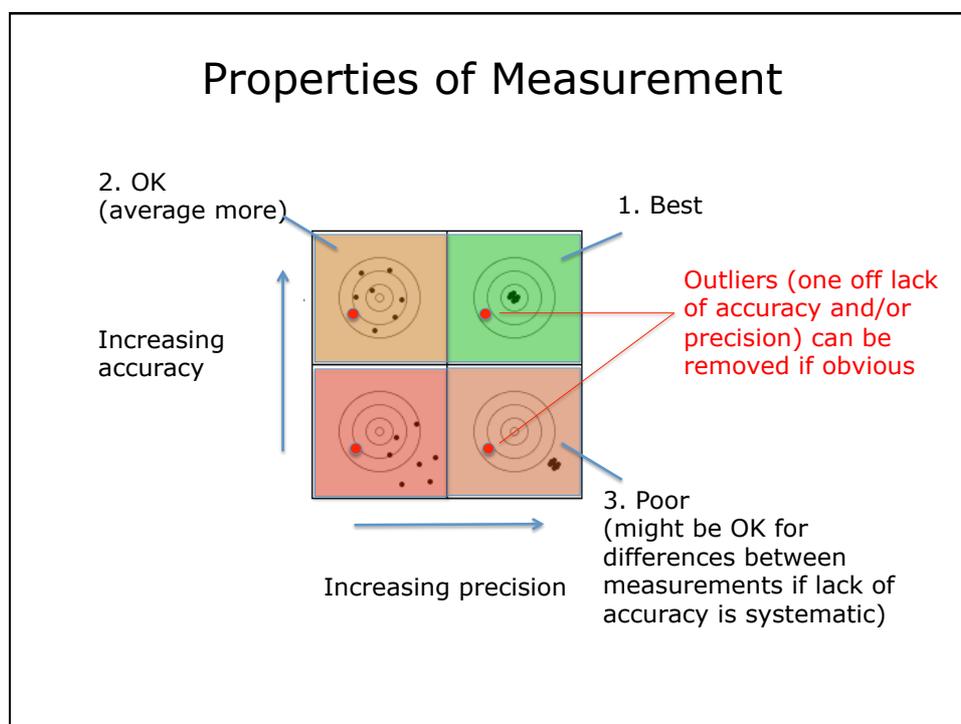


Art is the "I" while science is the "We"

Properties of Measurement

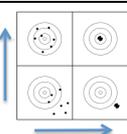
- **Accuracy**; how close a measurement(s) reproduces the value of the parameter being measured
- Reliability or **Precision**; the differences between multiple 'identical' measurements (scatter)
- **Validity**; how well the measured variable represents or reports on the attribute being examined
- **Units**



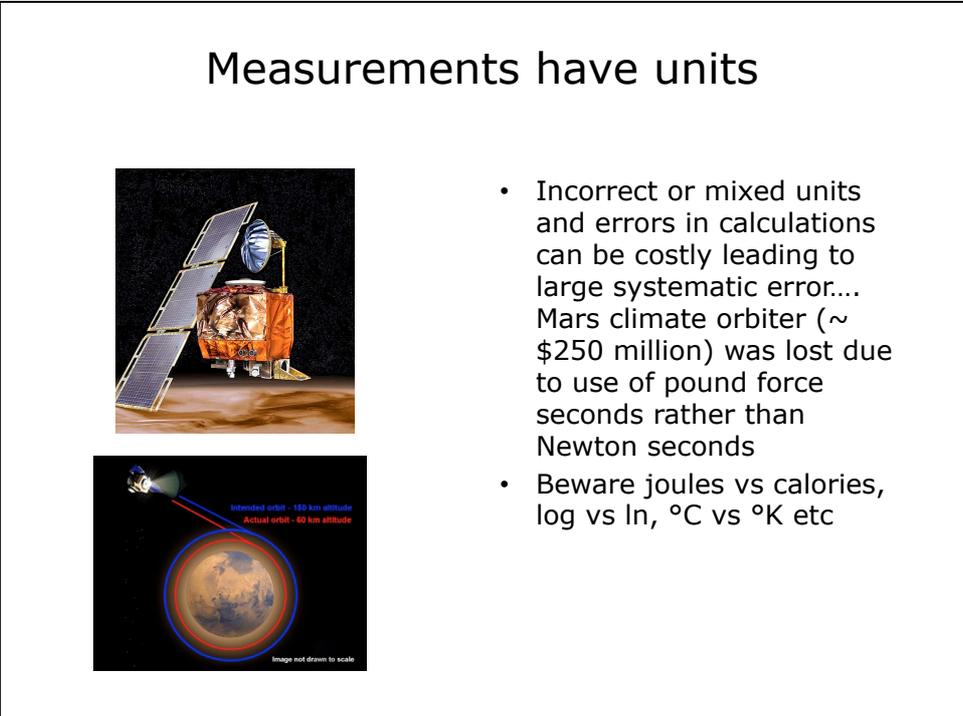
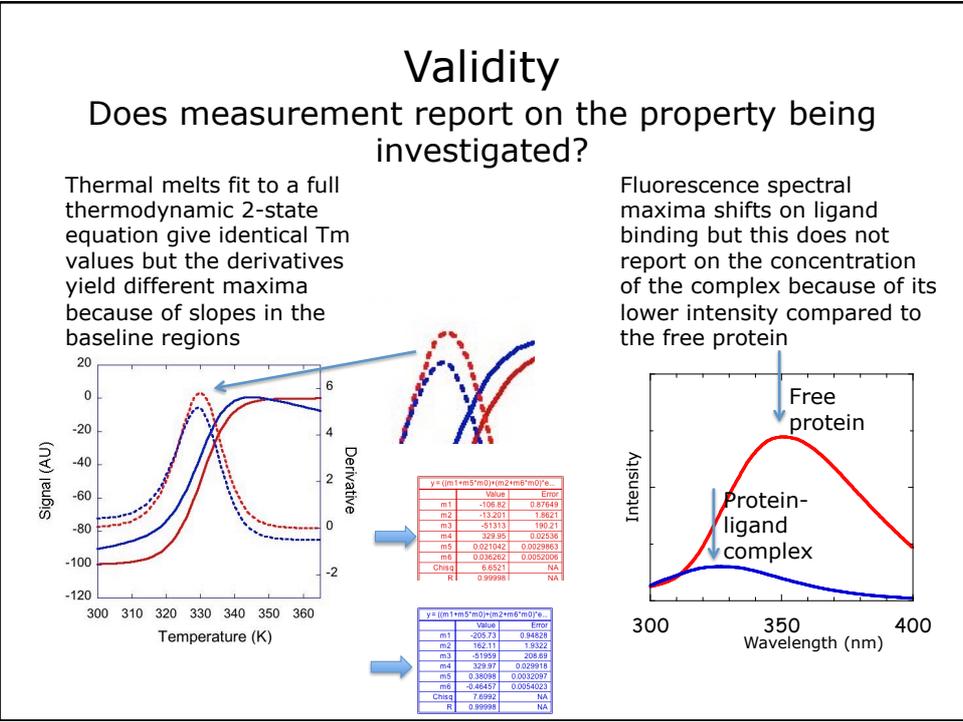


Uncertainty in Measurement

Precision and Accuracy are subject to random and systematic error



- Random errors (noise)
 - Measurement range of instrument, signal magnitude relative to instrument noise (S/N)
 - S/N of tools used in preparation of samples, dilutions etc
 - **Random errors (lack of precision) will average out**
- Systematic errors ('mistakes')
 - Instrument calibration, instrument measurement range / response (offscale)
 - Calibration of tools used in preparation of samples (pipettes, concentration of stock solutions, etc)
 - Operator error / mistakes (calculations, missing component of assay, etc)
 - Temperature equilibration
 - Kinetic equilibration (time for endpoint to be reached before measurement, etc)
 - **Systematic errors (lack of accuracy) are not removed by averaging**



General data fitting Non-linear least squares (NLLS)

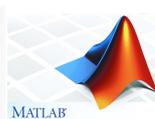
NLLS curve fitting software



Spreadsheets
Good for simulation

Common packages
Intuitive user interface
Formula entry in text
Data nesting within plots
Global fitting

Advanced mathematical and programming options
Steep learning curve





Get yourself a function, $Y=f(x)$ and fit using NLLS!

- “With four parameters I can fit an elephant, and with five I can make him wiggle his trunk.” John von Neumann
- With enough terms anything (including elephants) can be fit.
- But is there then a ‘model’?

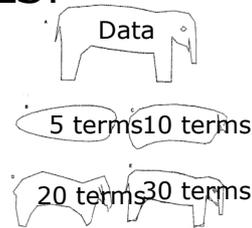
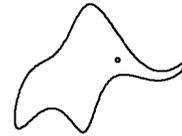


FIGURE 1.2. “How many parameters does it take to fit an elephant?” was answered by Wei (1975). He started with an standard elephant (A) defined by 20 points and used four square Fourier sine series fits of the form $f(x) = a_0 + \sum_{n=1}^N a_n \sin(n\pi x/2)$ and $g(y) = b_0 + \sum_{m=1}^M b_m \sin(m\pi y/2)$. He examined the best $N = 5, 10, 20$ and M between 0 to 21 and stopped with the fit of a 30 terms model. He concluded that the 30 terms model “not only fit best, but also, he would care your classical engineers into preliminary design.”

Wei J. 1975. Least square fitting of an elephant. Chemtech 5: 128-129



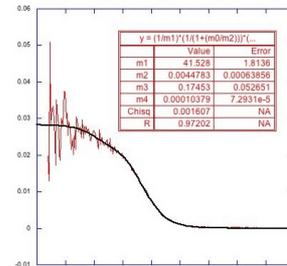
“Drawing an elephant with four complex parameters” by Jurgen Mayer, Khaled Khairy, and Jonathon Howard, Am. J. Phys. 78, 648 (2010)

NLLS goals and assumptions

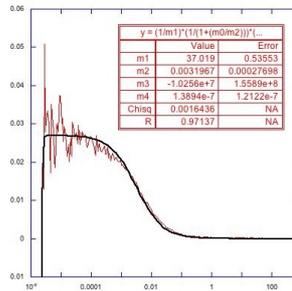
- NLLS tests if a given mathematical function is **consistent** with the data and determines the values of parameters in the function that have the highest probability of being correct.
- You know x accurately; uncertainty is in the dependent y variable
- Measurements (points) are independent
- Variability in y is random and gaussian
- Variability in y is equal as a percentage error for all points (unless points have associated errors; weighting)

NLLS; Getting started

- The program only looks for least squares fit making small iterations in the variables
- It may not find the logical and global least squares minimum unless you start fitting with reasonable values
- The more variables the higher dimensionality of the least squares fitting surface and the more local minima
- Some programs allow constraint (boundaries) for variables

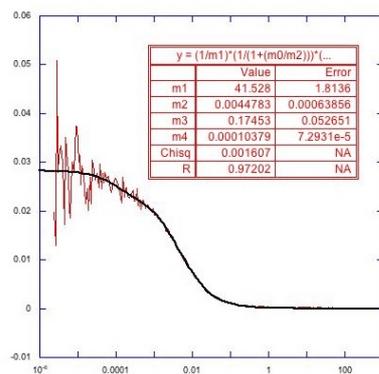


Start fit m4 = 1e5



Start fit m4 = 1e7

Cure fitting results: Goodness of fitw



- How precise or how well constrained are the best fit parameters (**standard errors**)?
- How good are the global indices of fit (**R and chi squared**)?
- **Visually examine the residuals** look for systematic (non random) distribution...

Error of individual parameters

- Compute the dependence of Y as fit variable is changed from the best fit value at each value of X keeping other variables fixed. Repeat for all variables and combine with complex matrix algebra. (this is not dependent on the number of data points and data scatter)
- Quantify the sum of squares of the data from the fit (how scattered are the data)
- Include the degrees of freedom in the data set (number of points - number of variables)
- Combine => Error estimate

y = m1*exp(-m2*m0)+m3*exp(-m...		
	Value	Error
m1	0.032669	6.092e-5
m2	32515	96.8
m3	0.024653	0.00019861
m4	6.3453e+5	8493.6
m6	-0.050038	1.4588e-5
Chisq	0.1338	NA
R	0.98006	NA

- 95% confidence interval is ~ double the SE
- Actually the SE * t , where t comes from the t distribution appropriate to the degrees of freedom in the data set
- For 95% and large degrees freedom (many more points than variables) t= 2.2 to 1.96 (n= 10 to ∞)

Error as fraction of value indicates how good the fit is and how well constrained (necessary) is the fit variable

Quality of the overall fit Chi squared and R

- Chi Square Value
- Sum of the squared error between the data and the fit divided by predicted standard deviation at that point, or an average across the data set

$$\chi^2 = \sum_i \left(\frac{y_i - f(x_i)}{\sigma_i} \right)^2$$

y = m1*exp(-m2*m0)+m3*exp(-m...		
	Value	Error
m1	0.032669	6.092e-5
m2	32515	96.8
m3	0.024653	0.00019861
m4	6.3453e+5	8493.6
m6	-0.050038	1.4588e-5
Chisq	0.1338	NA
R	0.98006	NA

- R and R²
- Coefficient of determination... how much of the total variation in y is predicted by the fit.

$$R = \sqrt{1 - \frac{\chi^2}{\sum_i \sigma_i (y_i - \bar{y})^2}}$$

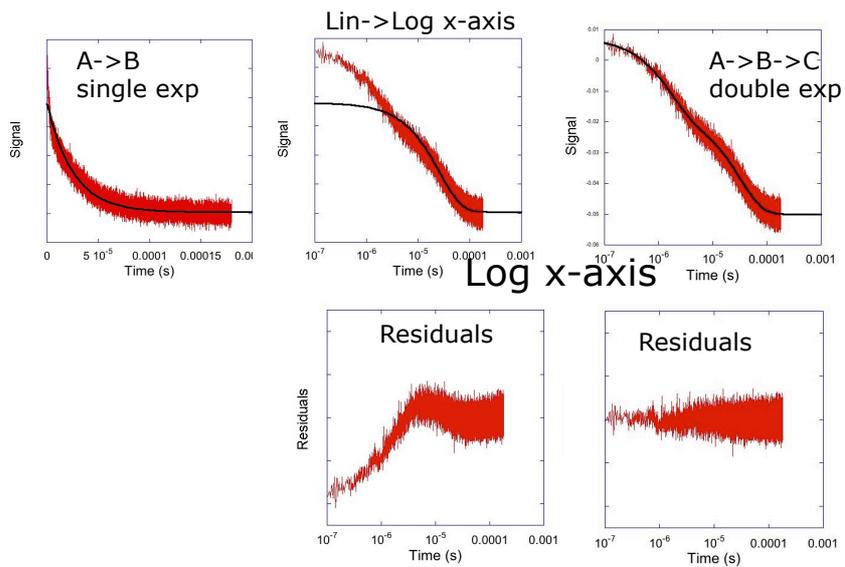
Qualitative ('common sense') assessment of the fit

- Are the best fit parameters physically reasonable? (negative rates/concentrations)
- Do some parameters have insignificant values / amplitude and might be omitted?
- Parameters and errors take no account of the realistic precision of the measurement so forget all the decimals.
- **Does the fit look 'good' for all values of x ? This may require a non linear (log) view of x in biological processes such as binding or exponential kinetics**
- **Visually inspect the residuals (difference between the data and the fit). These should be randomly / uniformly distributed.**

$$y = m1 \cdot \exp(-m2 \cdot m0) + m3 \cdot \exp(-m4 \cdot m0) + m6$$

	Value	Error
m1	0.032669	6.092e-5
m2	32515	96.8
m3	0.024653	0.00019861
m4	6.3453e+5	8493.6
m6	-0.050038	1.4588e-5
Chisq	0.1338	NA
R	0.98006	NA

Random distribution of Residuals?

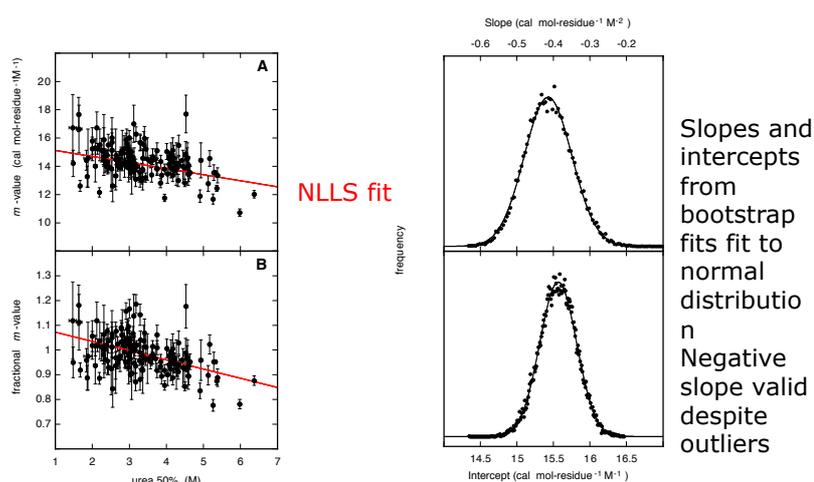


Constraint or bias in data sets

- Some points are more important in constraining the fit especially where the x axis distribution is not uniform.
- Many transforms which linearize data sets for convenient plotting are subject to this bias (Lineweaver Burke, Hill plot, etc)
- Significance of these points to the reported error can be determined in a variety of statistical resampling methods (bootstrap, jackknifing, etc)
- Regress+ free Mac software allows simple bootstrap analysis
- Create new data set by random picking n times from **the full data set**
- Fit to function
- Repeat process 1000+ times
- Collate fitted parameters and analyse as a normal distribution to determine standard deviation of fitted values including bias

http://www.causascientia.org/software/Regress_plus.html

Example of bootstrap analysis



Real Errors and Error Propagation

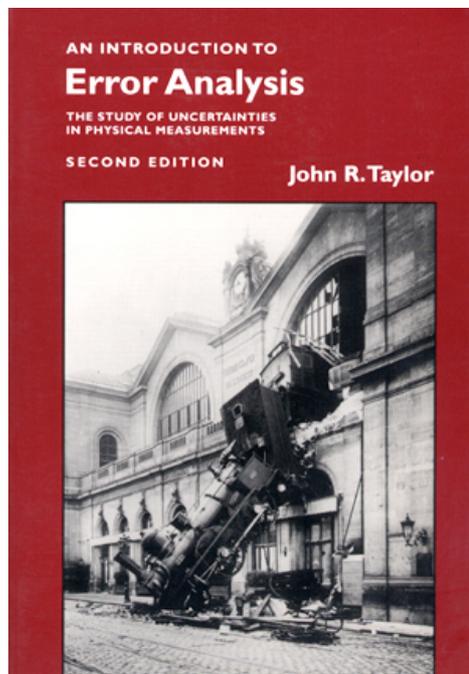
- Curve fitting only reports on the least squares fit to a particular data set
- Random errors are reduced by repeating measurements
- Systematic errors may be reduced providing their source is removed from the measurement (calibration, care, adjusted protocol)
- Ultimately all the materials used for measurement may have to be obtained independently to eliminate all sources of error.
- Common sense dictates how far errors need to be minimised.
- Fitted values may have to be combined or extrapolated both of which compound the associated error
- Simple formula allow the propagation of error to be accounted for;

$$A = B + C + \dots$$

$$\partial A = \sqrt{(\partial B)^2 + (\partial C)^2} \dots$$

$$A = B \times C \times \dots$$

$$\frac{\partial A}{A} = \sqrt{\left(\frac{\partial B}{B}\right)^2 + \left(\frac{\partial C}{C}\right)^2} \dots$$



Data Fitting in Practice

“..one must not underestimate the ingenuity of humans to invent new ways to deceive themselves.”

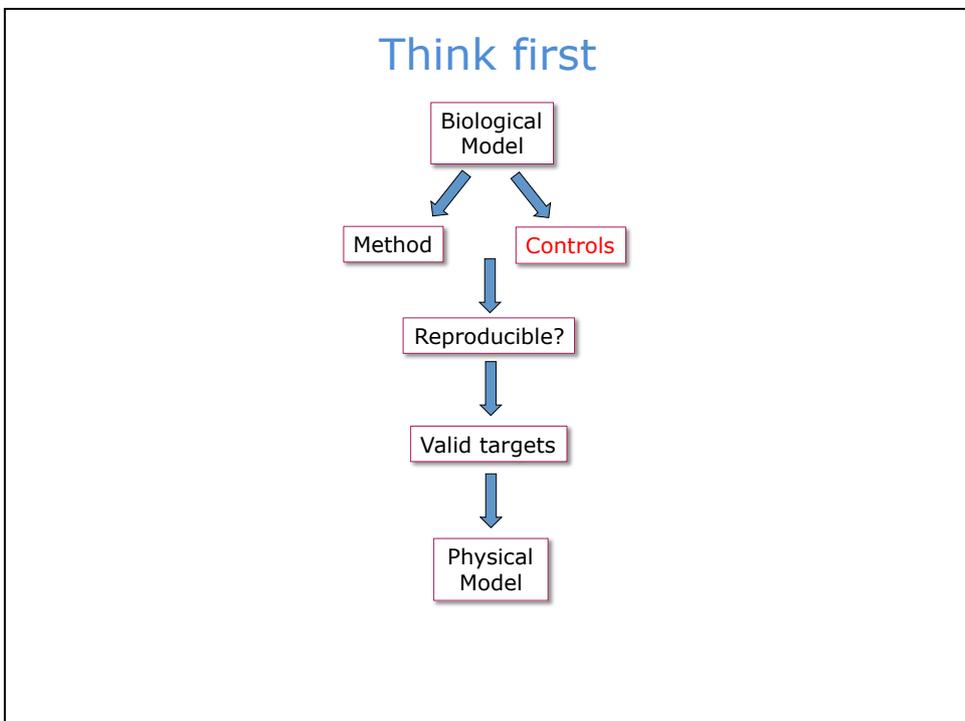
Richard Henderson (2013) PNAS

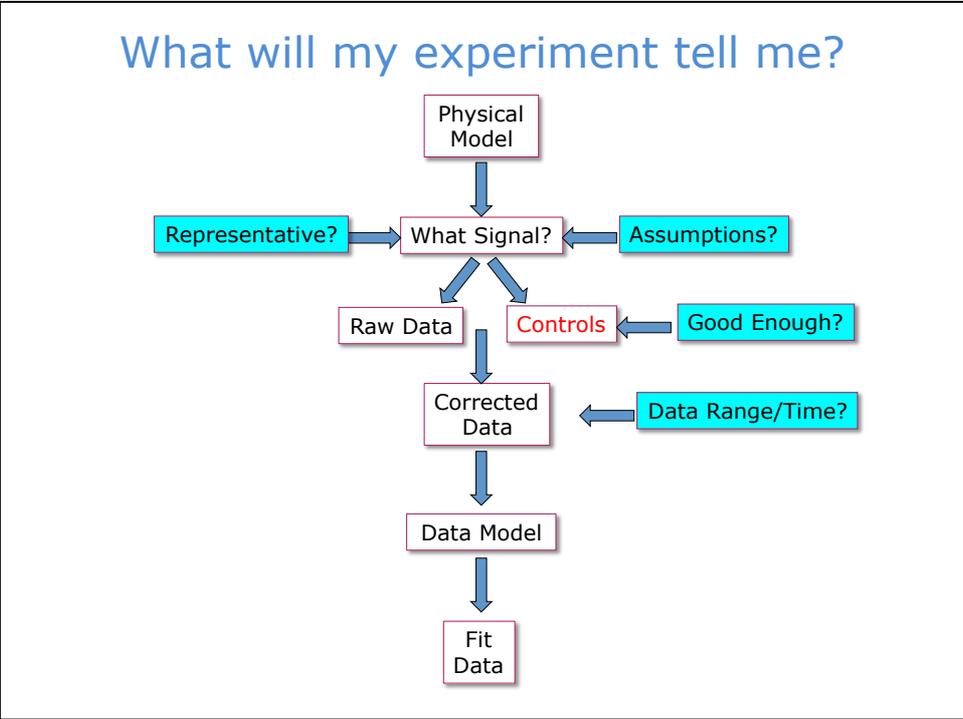
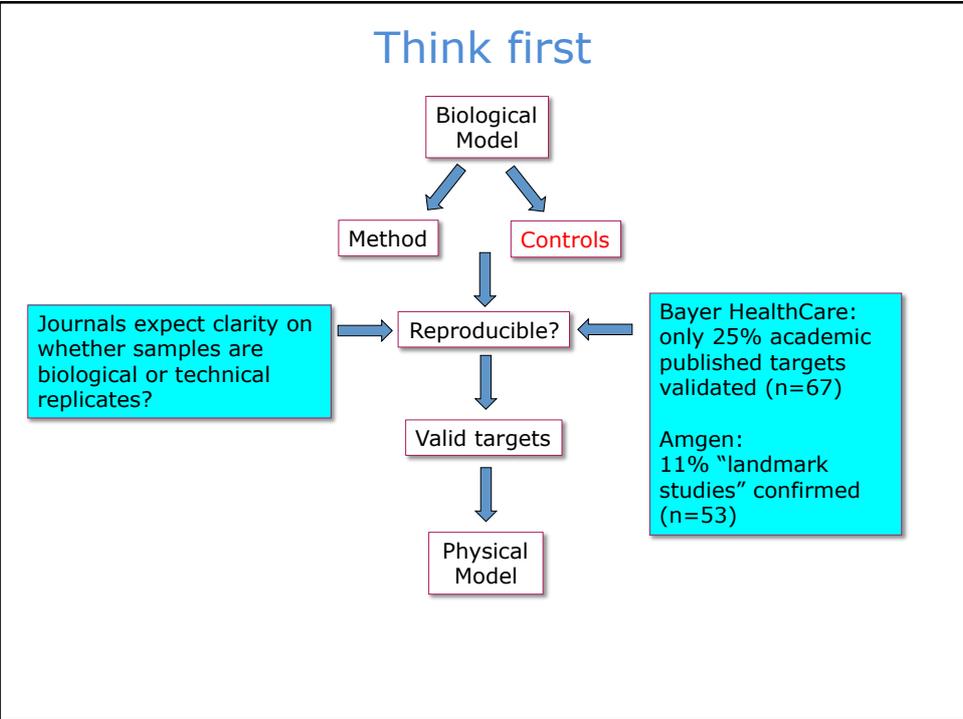
Einstein from the noise

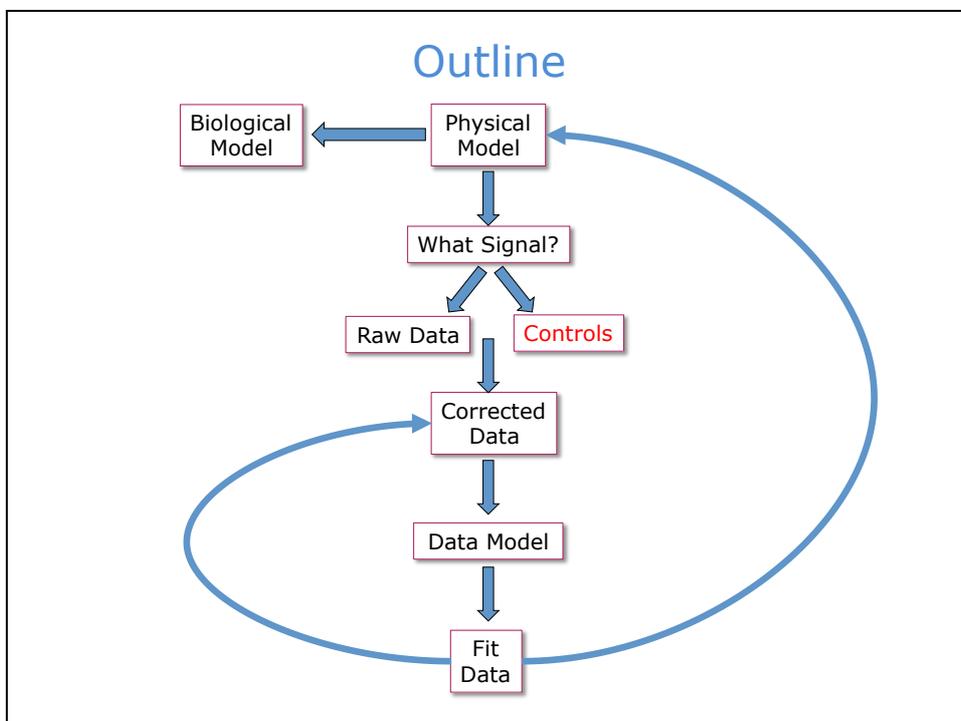
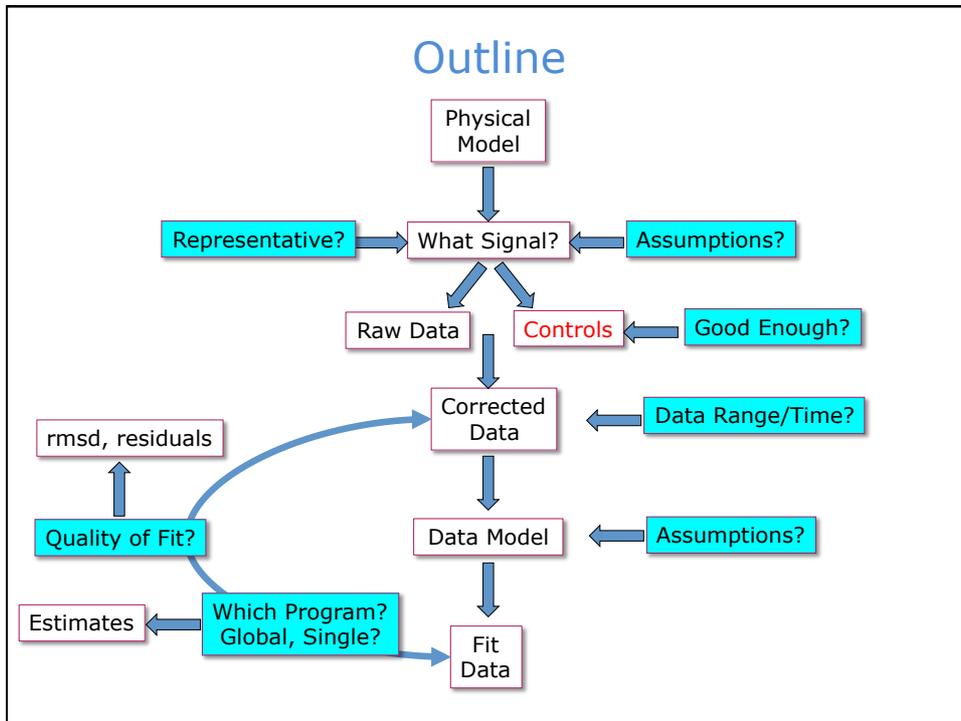


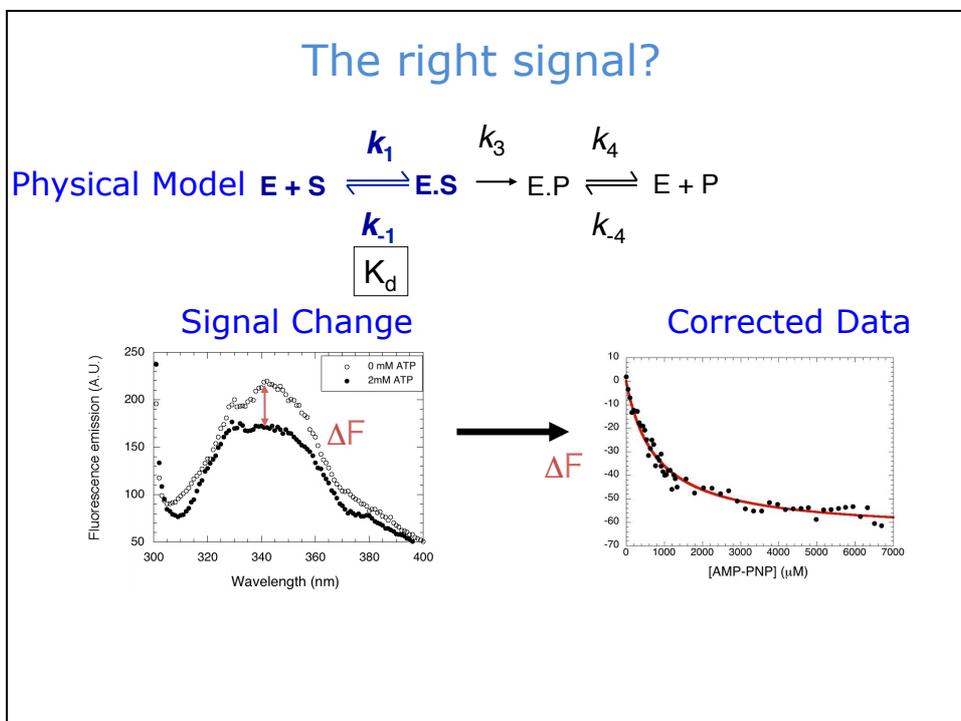
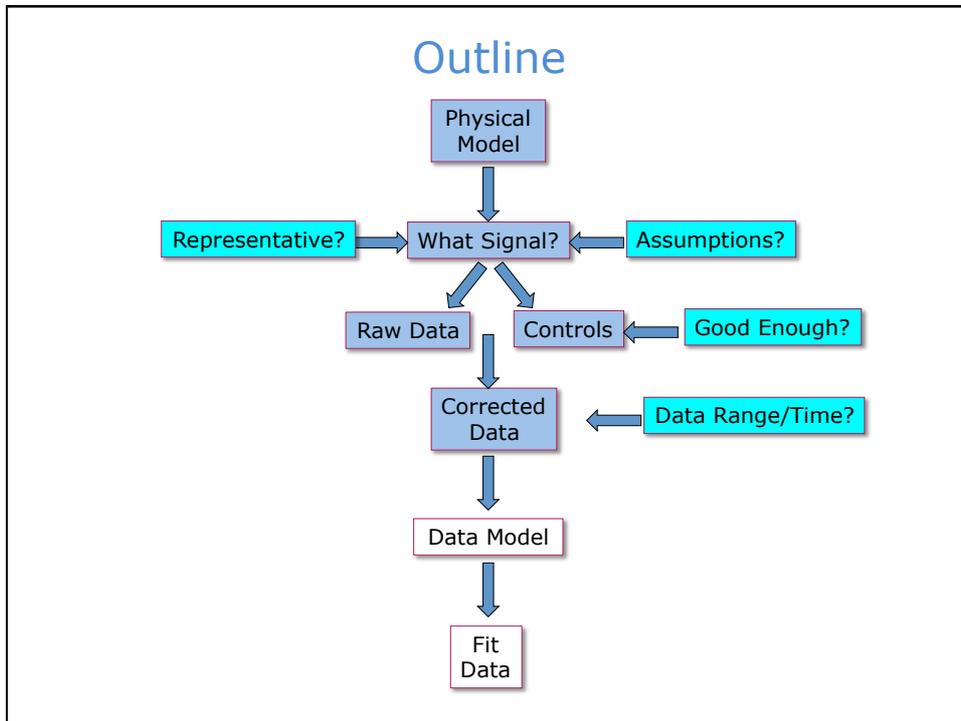
Model dependent bias

Maxim Shatsky *et al.* *J Struct Biol.* (2009)









The right signal?

$$\text{Physical Model } E + S \xrightleftharpoons[k_{-1}]{k_1} E.S \xrightarrow{k_3} E.P \xrightleftharpoons[k_{-4}]{k_4} E + P$$

Corrected data

Data model

$$\Delta F = \frac{\Delta F_{\max} \cdot [L]}{K_d + [L]}$$

Assumptions

1:1 model
 $[E.S] \propto \Delta F, k_3=0$

The right signal?

$$E + S \xrightleftharpoons[k_{-1}]{k_1} E.S \xrightleftharpoons[k_{-2}]{k_2} E^*.S \xrightarrow{k_3} E.P \xrightleftharpoons[k_{-4}]{k_4} E + P$$

K_c

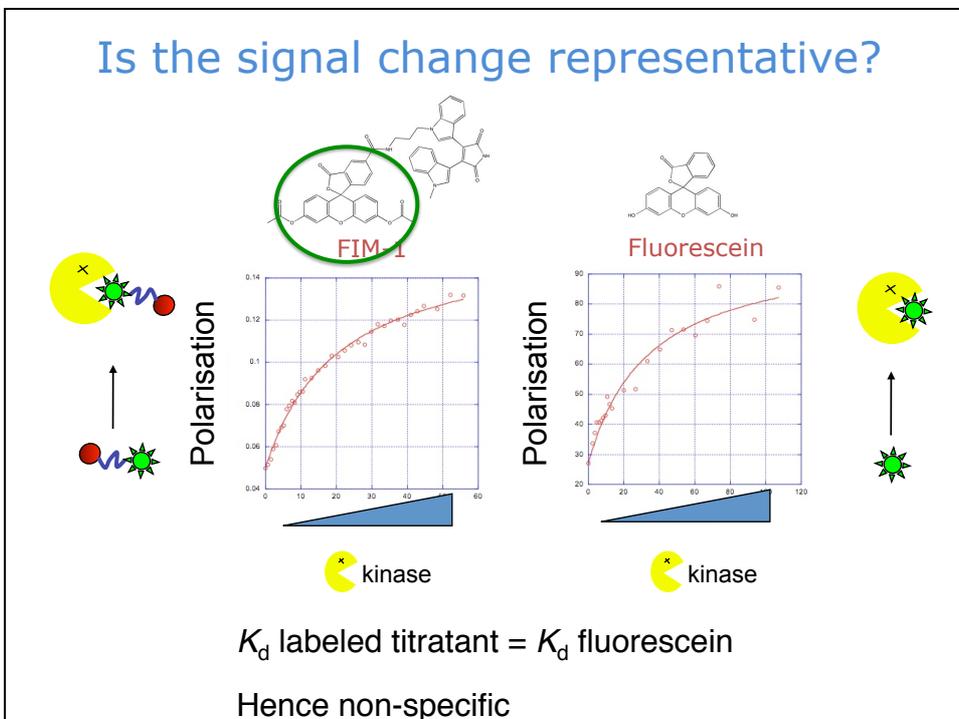
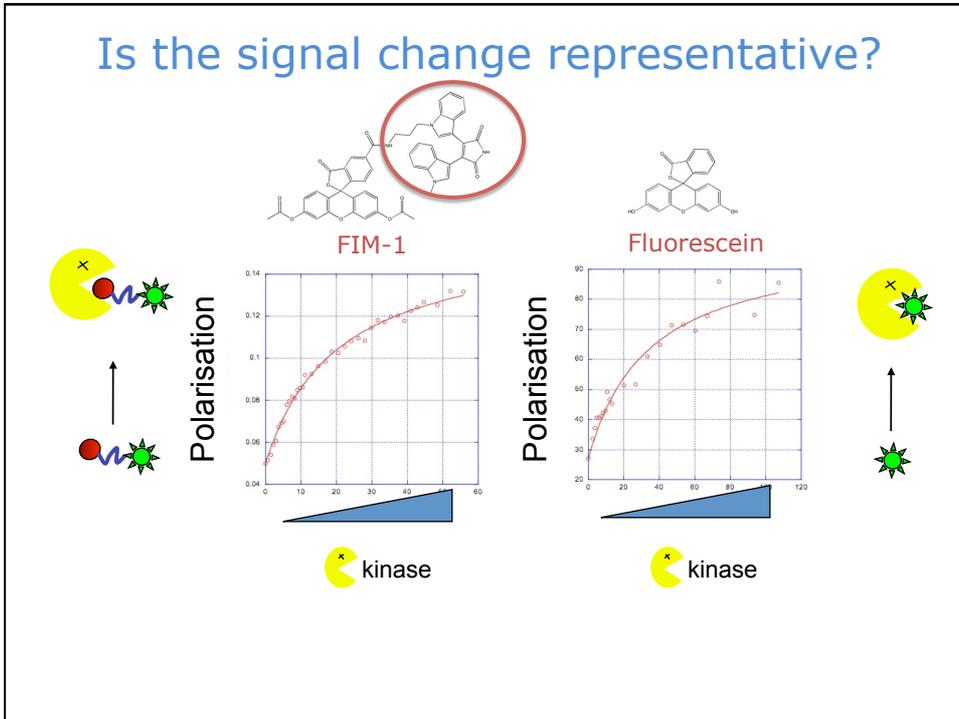
$$\begin{matrix} \updownarrow \\ E' \end{matrix}$$

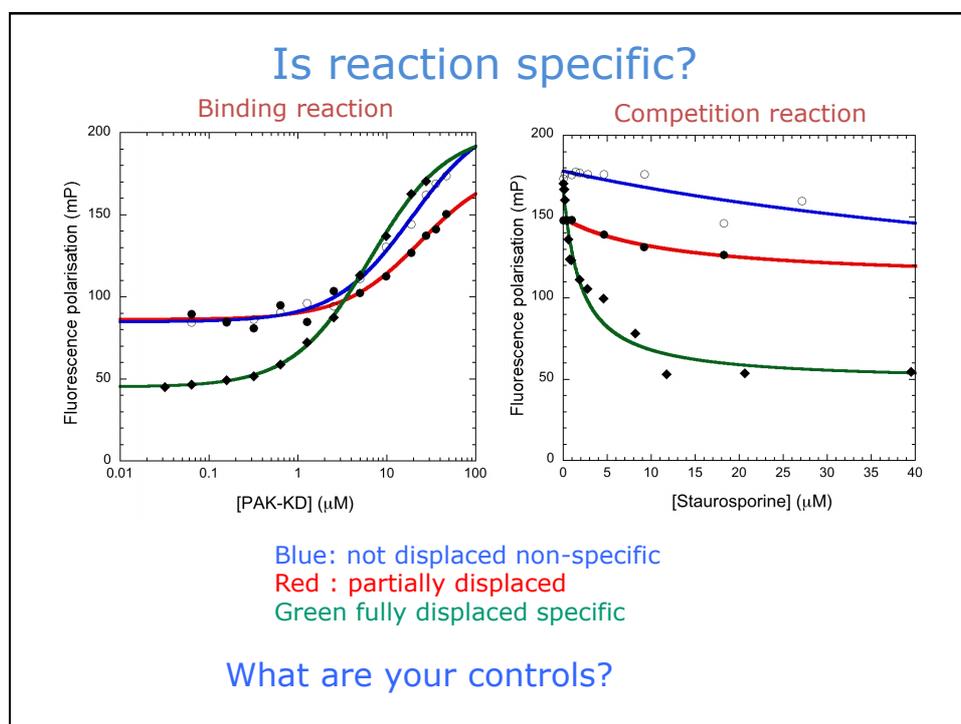
K_d

Physical model could be incomplete leading to incorrect assumptions

$\Delta F \propto ? [E.S] + [E^*.S] + \dots ?$

- Can the physical measurement be ascribed to a specific step in the reaction?
- Are there any missing steps in the analysis?





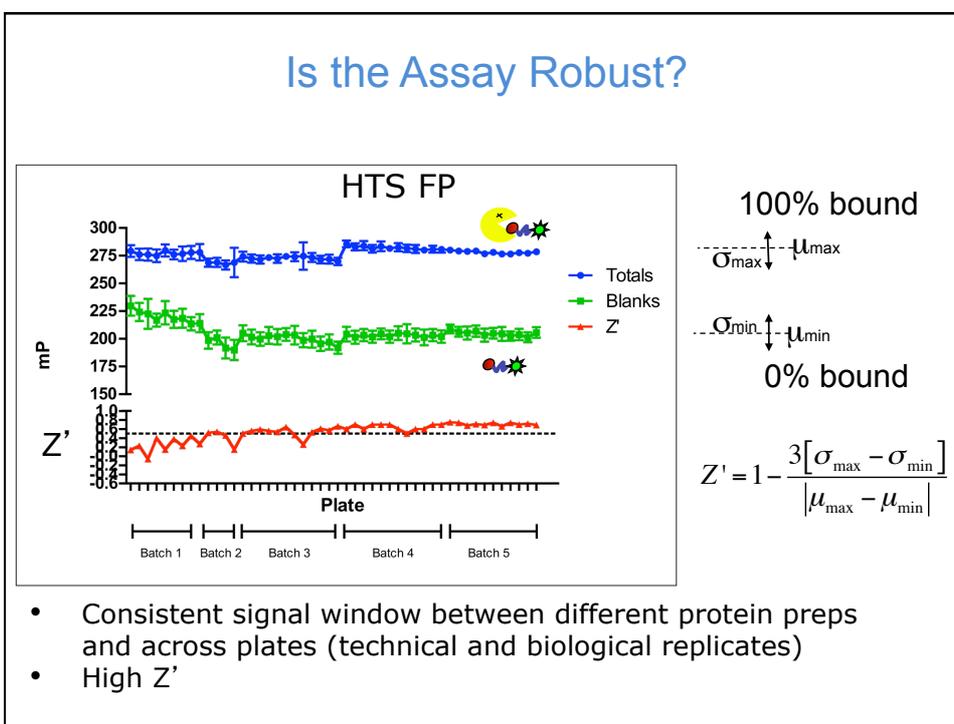
Practical Tips for Good Signals

Before Experiment

- Buffers and samples filtered
- Take care of aggregation (check at end as well)
- Concentrations accurate? How do you measure?
- Sample stability? Temp., pH, and Time.

<p>Methodology</p> <ul style="list-style-type: none"> Good signal change? Background? Buffer Effects? Accurate titrations? Temperature control? Time to reach equilibrium? 	<p>Controls</p> <ul style="list-style-type: none"> Positive control? Negative control? Reagents defined and stable?
---	---

Is the Assay Robust?



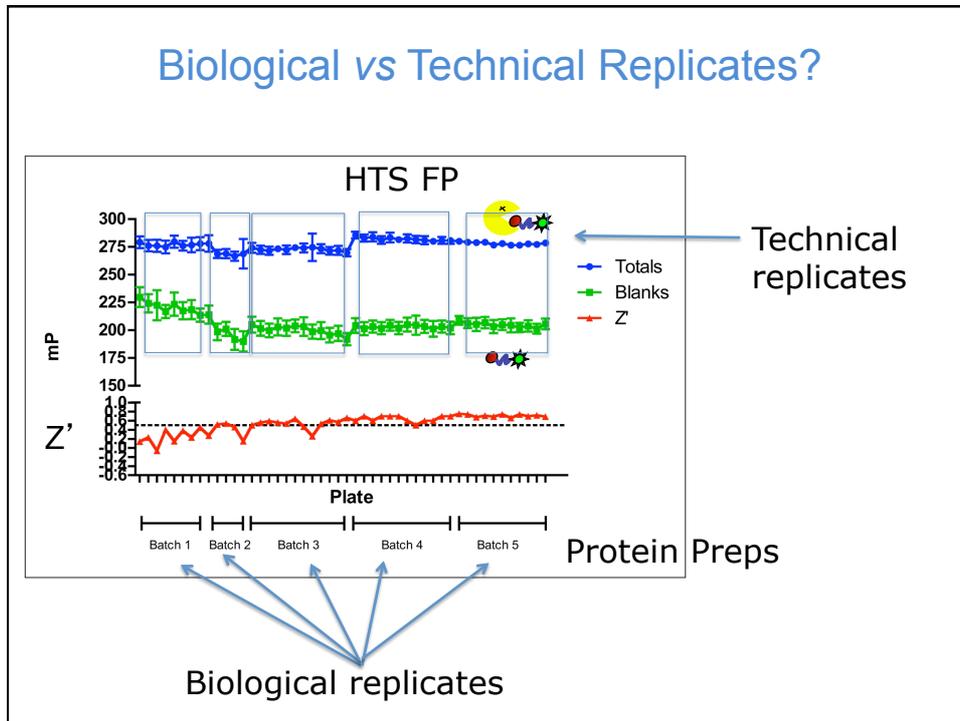
Equations for determining assay performance or sensitivity

Parameter	Equation	Comment
Coefficient of variation	$\%CV = \frac{\sigma}{\mu} \times 100$	A measure of the precision relative to the mean value, calculated for the maximum and minimum signals. An acceptable limit is <15%.
Signal to noise	$S:N = \frac{\mu_{\max} - \mu_{\min}}{\sigma_{\min}}$	A measure of the signal strength. This equation is sometimes given as ^a : $\frac{\mu_{\max} - \mu_{\min}}{\sqrt{\sigma_{\max}^2 + \sigma_{\min}^2}}$
Signal to background	$S:B = \frac{\mu_{\max}}{\mu_{\min}}$	Usually calculated using control compounds, acceptable value >2-fold.
Signal window	$SW = \frac{\mu_{\max} - \mu_{\min} - 3(\sigma_{\max} + \sigma_{\min})}{\sigma_{\max}}$	The significant signal between max and min controls, acceptable value >2-fold ^b .
Z' factor	$Z' = 1 - \frac{(3\sigma_{\max} + 3\sigma_{\min})}{ \mu_{\max} - \mu_{\min} }$	Representation of the SW using a score where a value >0.5 represents an acceptable assay. Z' is measured in the absence of library compounds; Z is in the presence.
Minimum significance ratio	$MSR = 10^{2\alpha_4}$	Smallest potency ratio between two measurements that is statistically significant (with 95% confidence).
Cheng-Prusoff	$K_i = \frac{IC_{50}}{(1 + [S]/K_m)}$	Used to derive the K_i from the IC_{50} under conditions of competitive inhibition.
Cheng-Prusoff (ligand-binding)	$K_i = \frac{IC_{50}}{(1 + [L]/K_d)}$	Used to derive the K_i from the IC_{50} under conditions of competitive binding.

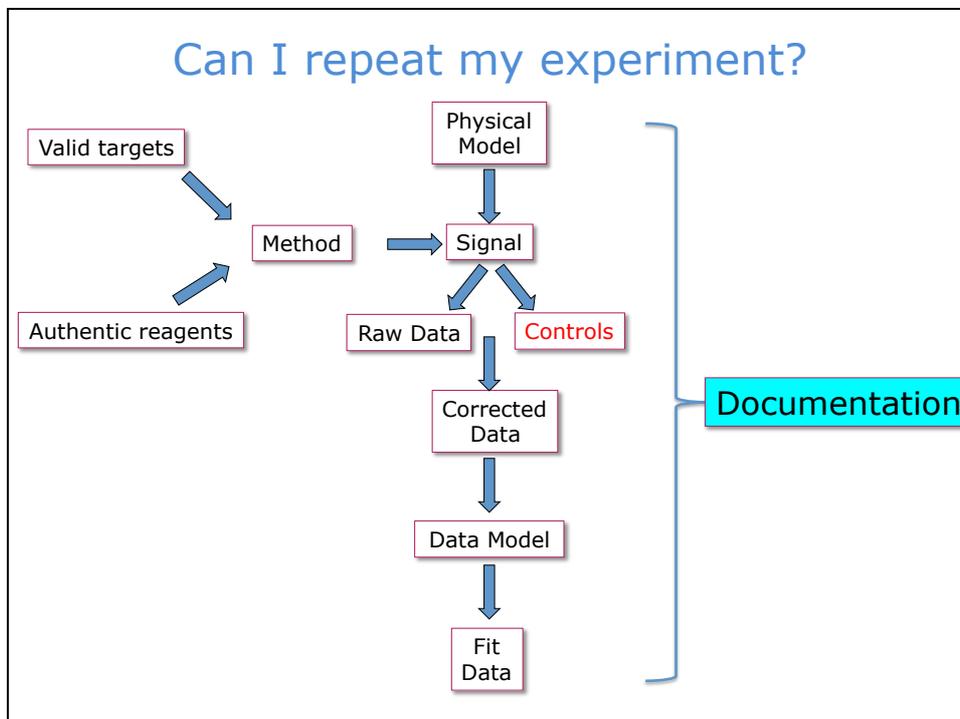
σ , s.d. of the assay signal; μ , mean of the assay signal; σ_d , s.d. of the difference in log potency; max, maximum signal; min, minimum signal. ^aThis is a redundant description of the signal window equation. ^bSome assays, such as those with ratiometric data, may perform adequately with as little as a 20% change in signal (for example, see ref. 111) or with $0 < Z' < 0.5$.

Inglese *et al.* (2007) Nature Chemical Biology 3, 466 - 479 (2007)

Biological vs Technical Replicates?

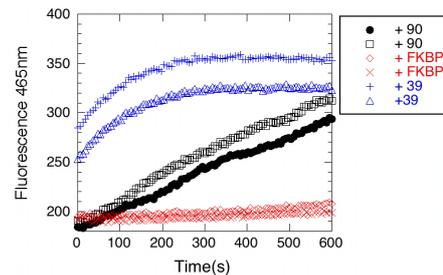
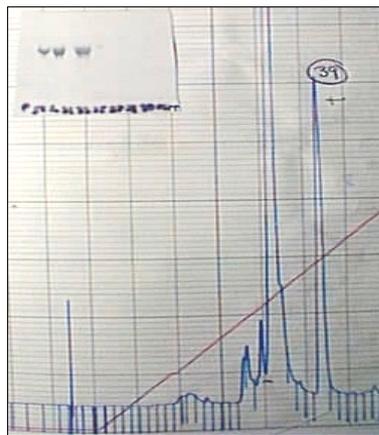


Can I repeat my experiment?



"Do not waste clean thinking on dirty enzymes" Efraim Racker and Arthur Kornberg.

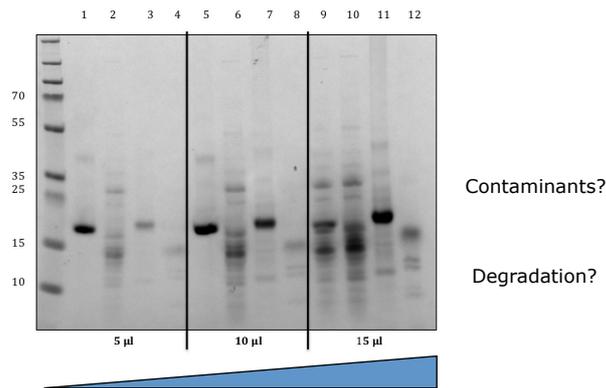
Case study: Does co-chaperone FKBP59 stimulate Hsp90 ATPase activity? Initial tests were exciting!



FKBP59 appeared homogeneous but further purification resolved contaminating ATPase activity

"Wasted time on dirty proteins"

Case study: How does mutation effect dimerisation?



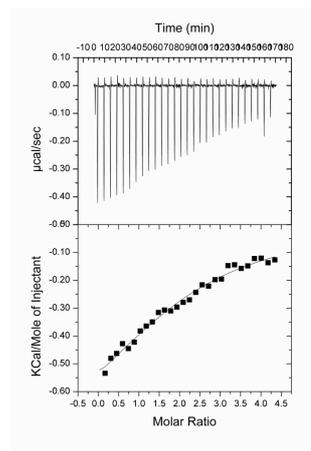
Gel loading

What do you do to ensure the quality of your sample?
Have you checked the validity of the targets/reagents?

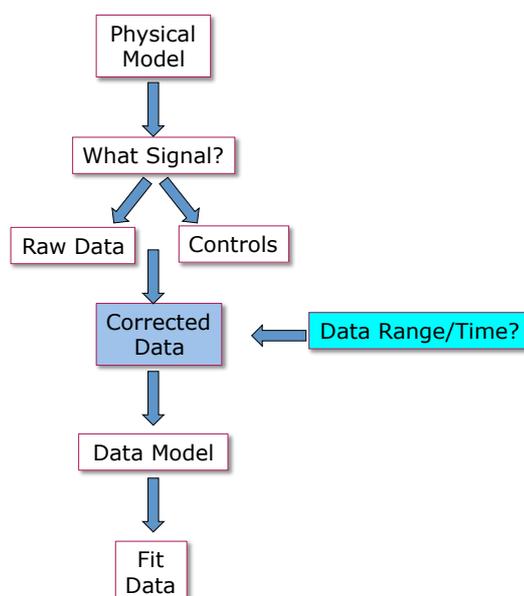
Background Signal?

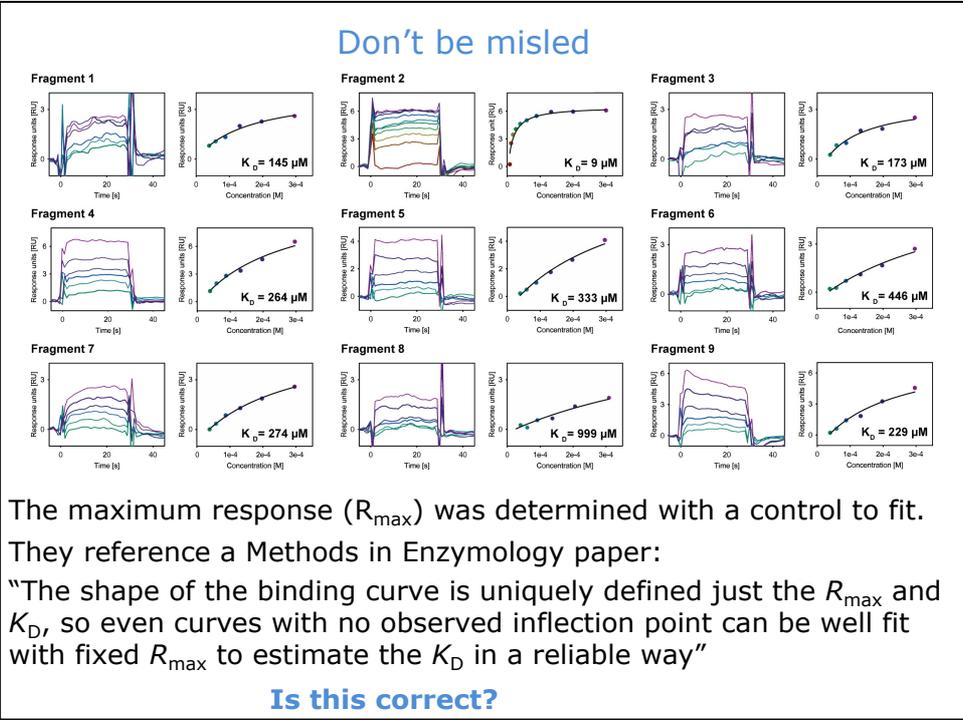
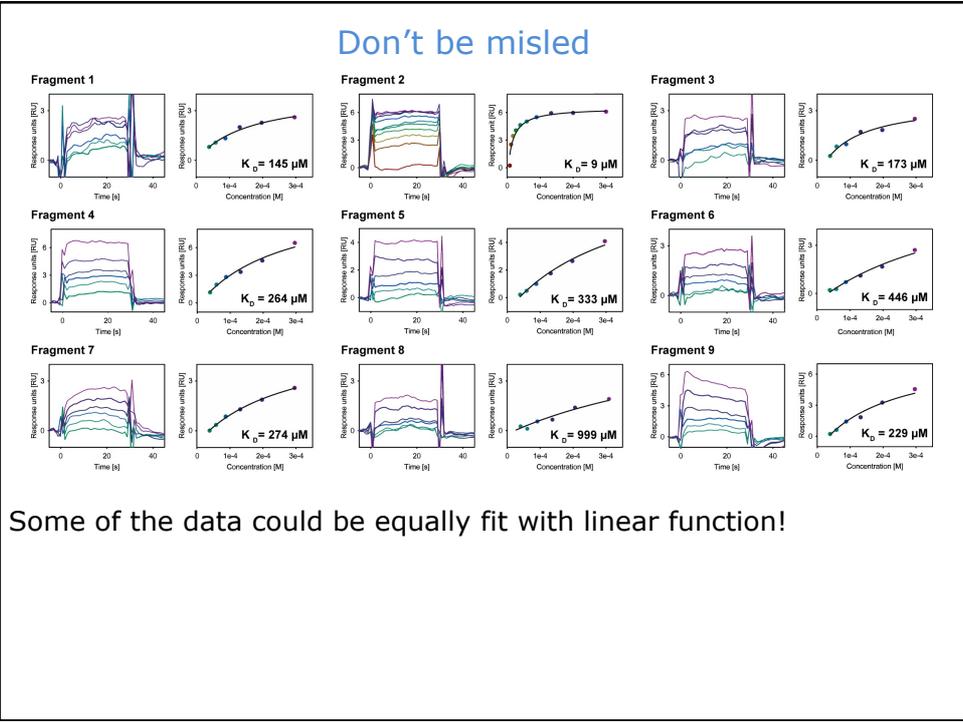
Does titrant effect signal or measurement?

- Titrant may contribute to signal e.g. light-scattering, heat of dilution in ITC.
- Buffer mismatch e.g. in AUC buffer sediments, DMSO in SPR and ITC.
- Titrant effects measurement e.g. the incident light (inner-filter effect)

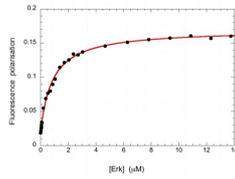


Outline

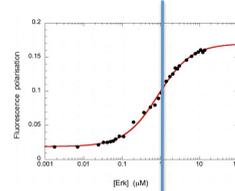




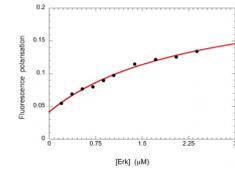
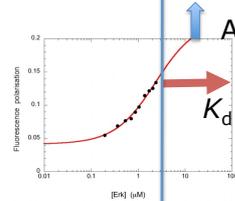
Right Range? Taking the Log View



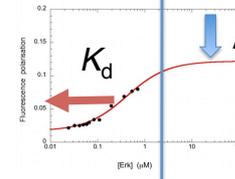
Full range



Limited no of points



Too short range



Rule of thumb $[Final] \geq 5-20 \times K_d$

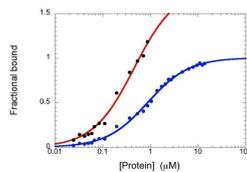
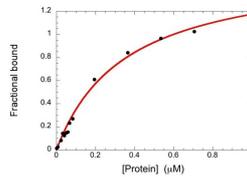
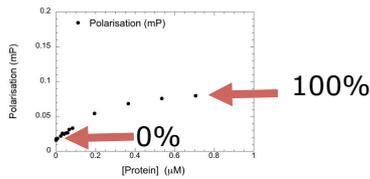
Don't Normalise to Fit

I think I've reached saturation

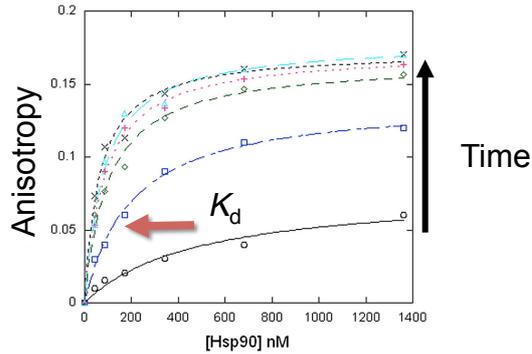
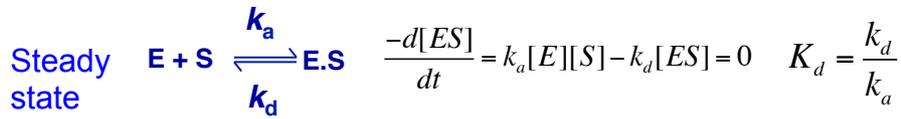
$$f_b = \left(\frac{F - F_0}{F_1 - F_0} \right)$$

Fit of normalised data to a simple curve looks ok

Comparison with real full data set (blue): K_d underestimated



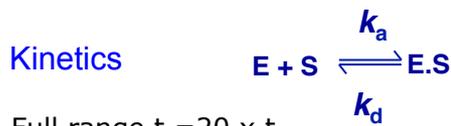
Right Range? Taking the Long View



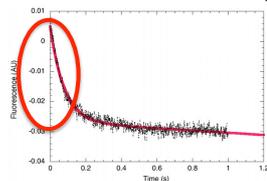
Has the reaction reached equilibrium?

Compare data set at different times or increase time between additions

Right Range? Taking the Long View



Full range $t = 20 \times t_{1/2}$

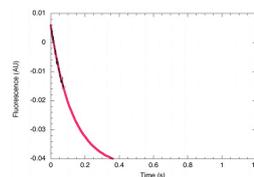
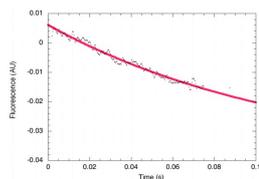


$$k_{obs} = k_a[S] + k_d$$

$$y = A \cdot \exp(k_{obs} \cdot t) + B \cdot t + C$$

$t = 2 \times t_{1/2}$

Has the endpoint been reached?

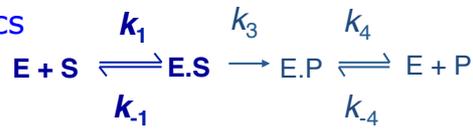


Wrong amplitude
Wrong rate

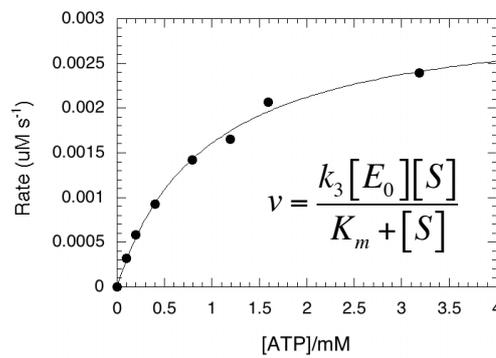
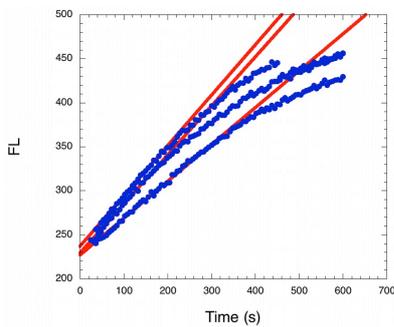
Rule of thumb $t = 10-20 \times t_{1/2}$

Right Range? Taking the Short View

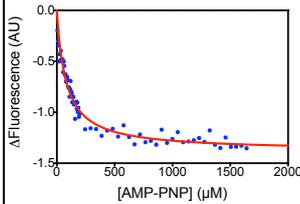
Enzyme kinetics



Assumption: $[P] = 0$, therefore look at initial rates
 $[E] \ll [S]$, so $[S] \approx$ [free substrate]

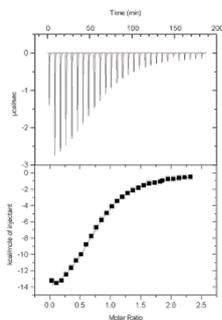


Appropriate Concentrations



Fluorescence, FP, MST
 Labeled $[L]$ (total) $\ll K_d$
 Assume titrating $[P]$ doesn't change as $[PL]$ small

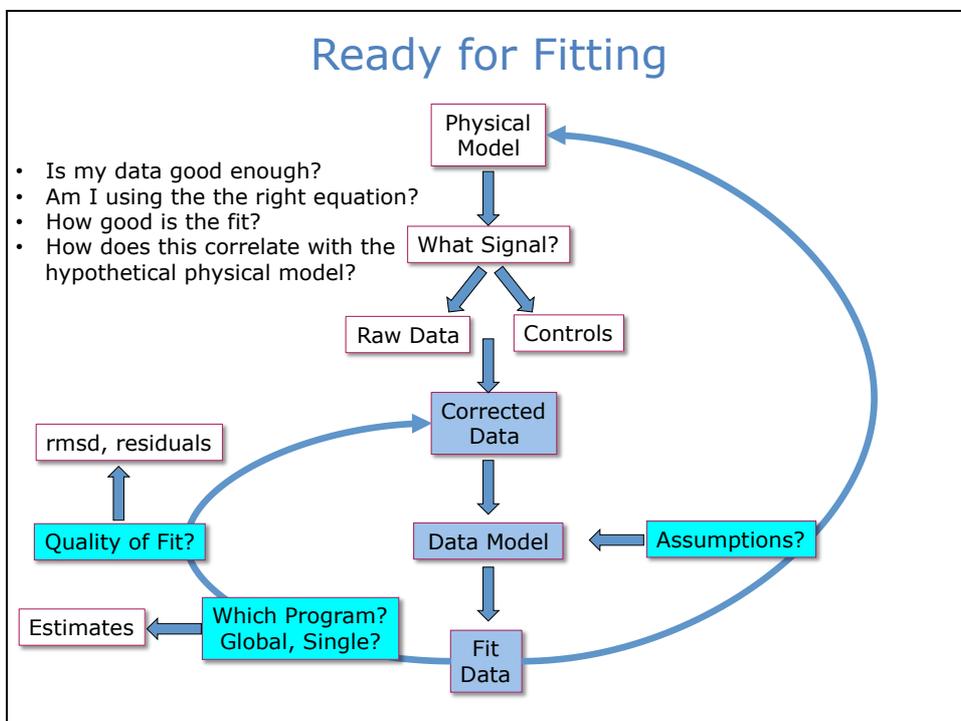
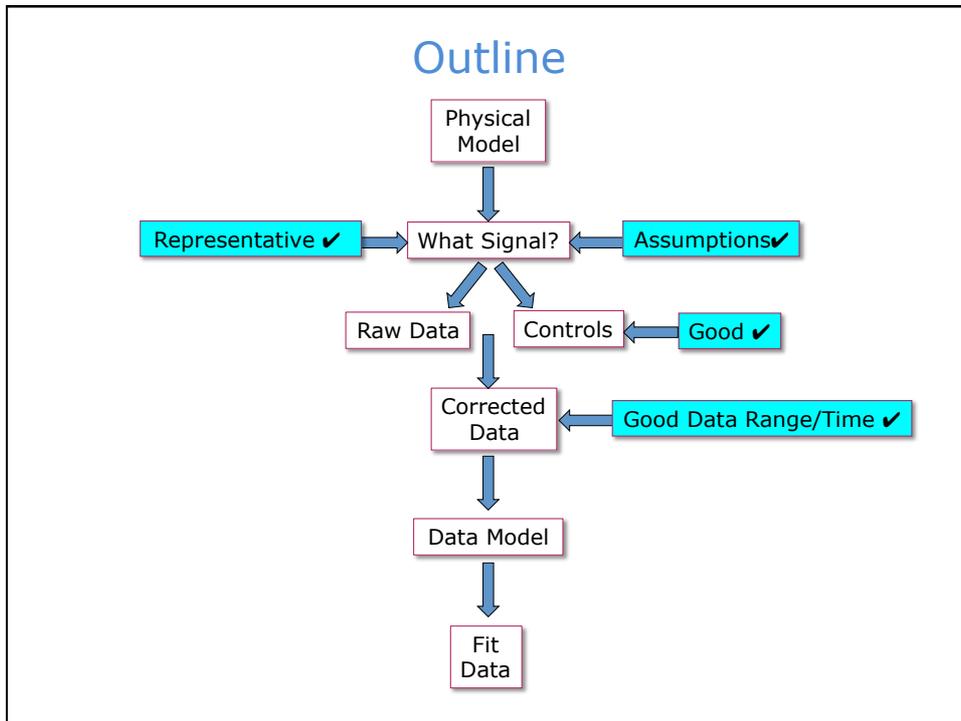
If $K_d = [L]/2$, then affinity much tighter than expected



ITC
 $[L]$ (in cell) $> K_d$
 P titrates with reaction

Need to have rough idea and then repeat

Have you accurately measured concentration?



Bad Fitting: Ugly Feet in Bigger Shoe

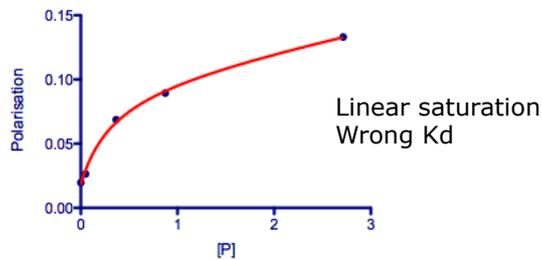
What assumptions have been made in the data model?
How do all the parameters relate to the signal?

Easy to get a good fit to bad data with more complicated equation
For example:

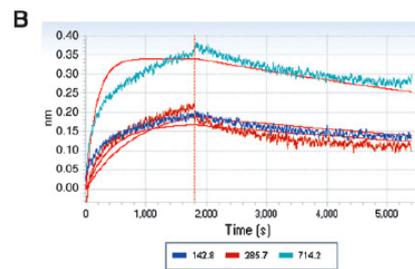
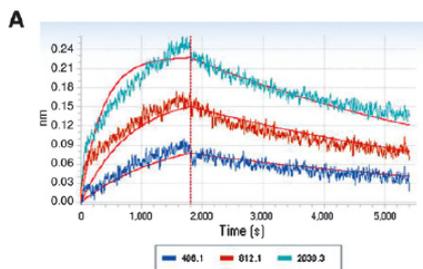
In Prism one-site binding model:

$$Y = B_{max} * X / (K_d + X) + NS * X + Background$$

NS*X is linear non-specific binding? Is this real for my system?



Bad Fitting: Wrong Shoe, Wrong Feet; Pinches and Gaps

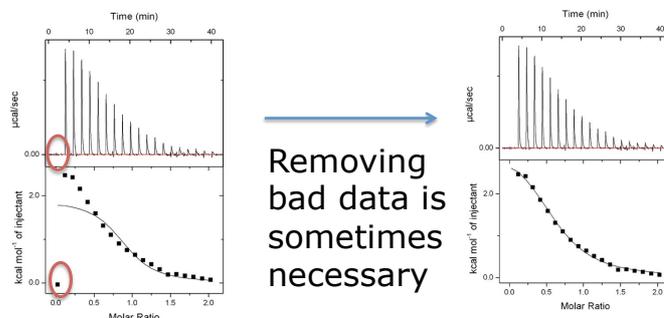


Assumption that data should fit to supplied software model:
BLI (Octet) kinetic data fitted to 1:1 interaction
Data does not fit model

Globally fitted to 2 decimal places!



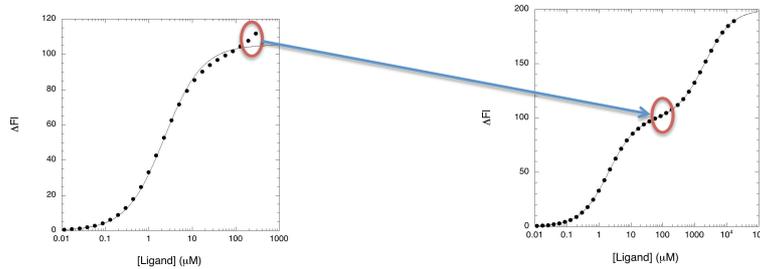
Outliers and Outright Lying!



Outliers and Outright Lying!



Outliers and Outright Lying!



Is there a reproducible deviation of the data from the model? Or is the outlier seen in only 1% of data?

Is it random chance (Gaussian) or experimental error?

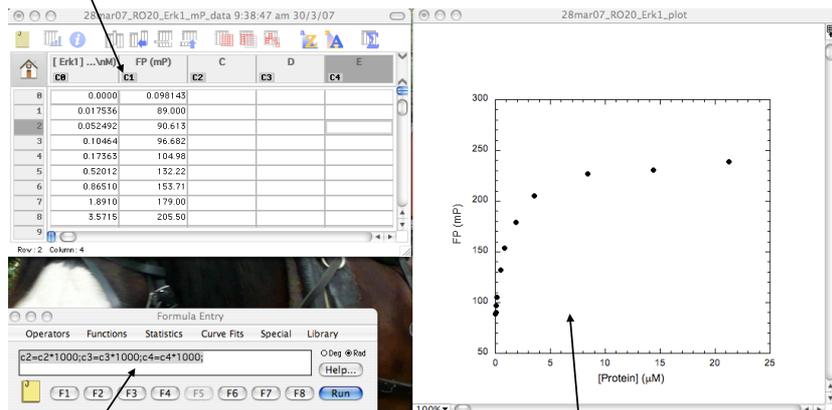
Is attribution of experiment error self-deception?

Better assay conditions?

Do you have the wrong data range and/or wrong model?

Fitting in Kaleidagraph

Data columns

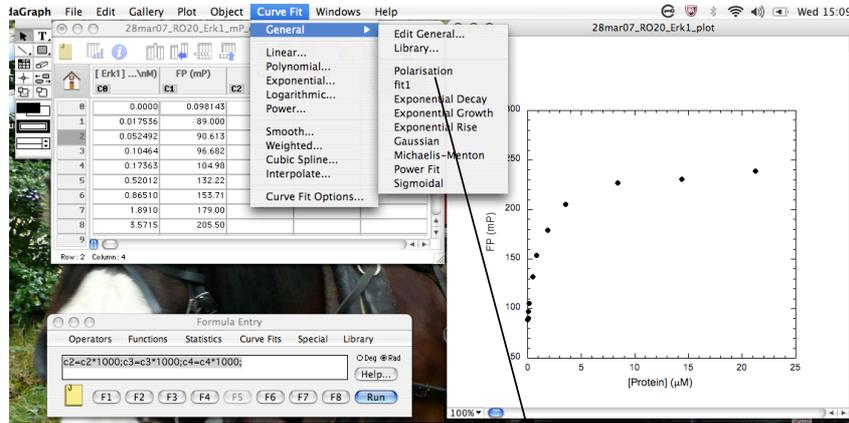


Transformation tool

Plot

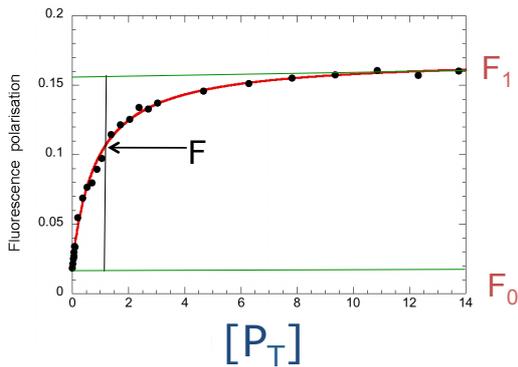
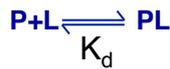
Always fit raw not normalised data

Fitting in Kaleidagraph: Data Model



Saved models

Data Model



$$K_d = \frac{[P] \cdot [L]}{[PL]}$$

$$F = F_0 + (F_1 - F_0) \frac{[PL]}{[L_T]}$$

$$[PL] = \frac{([P_T] - [PL]) \cdot ([L_T] - [PL])}{K_d}$$

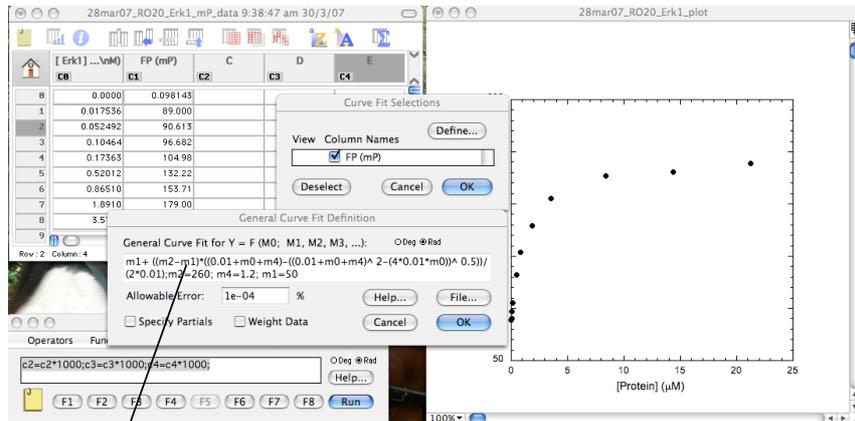
$$[L_T] = [L] + [PL]$$

$$[P_T] = [P] + [PL]$$

[P_T]

$$F = F_0 + \frac{(F_1 - F_0) \left\{ ([P_T] + [L_T] + K_d) - \sqrt{([P_T] + [L_T] + K_d)^2 - 4[P_T][L_T]} \right\}}{2[L_T]}$$

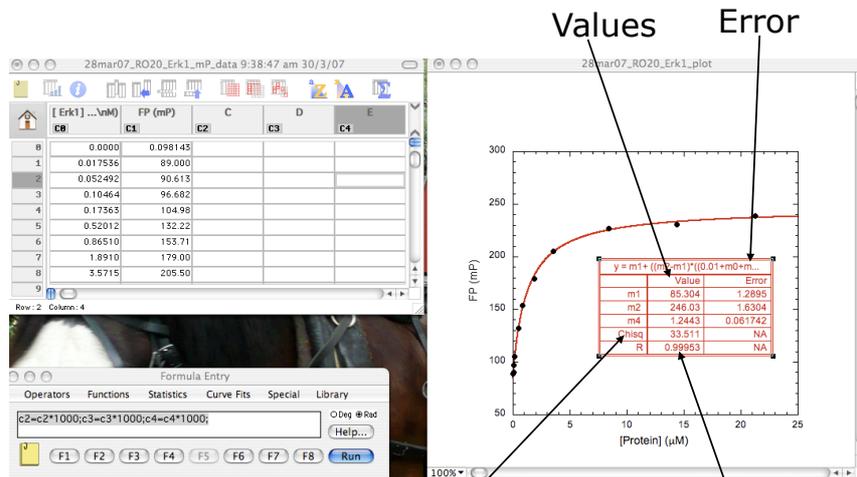
Fitting in Kaleidagraph: Parameters



Parameters

$$F = F_0 + \frac{(F_1 - F_0) \left\{ ([P_T] + [L_T] + K_d) - \sqrt{([P_T] + [L_T] + K_d)^2 - 4[P_T][L_T]} \right\}}{2[L_T]}$$

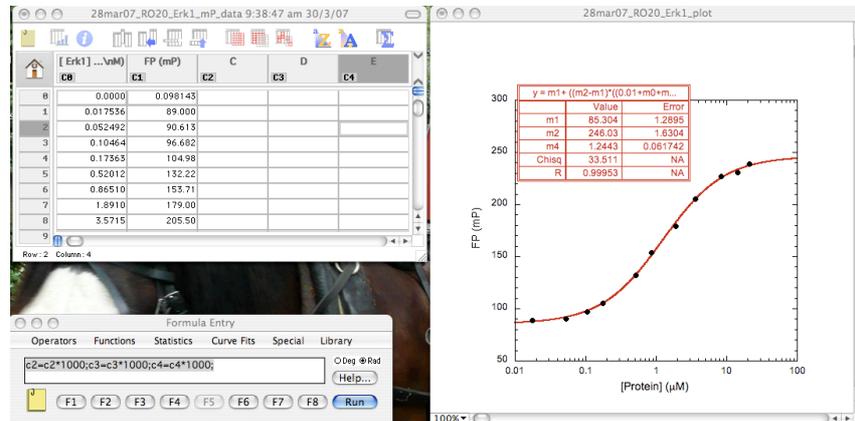
Fitting in Kaleidagraph: Result



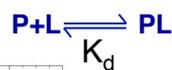
Chi²

Correlation coefficient

Fitting: Log View



Data Model



Assume $[PL] \propto \Delta F$

$$K_d = \frac{[P] \cdot [L]}{[PL]}$$

$$K_d = \frac{([Pt] - [PL]) \cdot [L]}{[PL]}$$

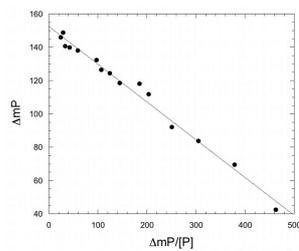
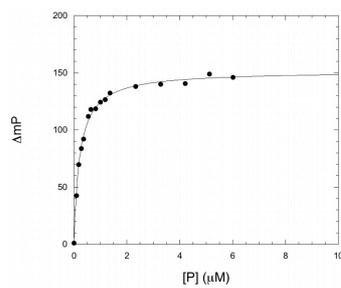
$$[PL] \equiv \Delta F; [Pt] \equiv \Delta F_{\max}$$

$$K_d = \frac{(\Delta F_{\max} - \Delta F) \cdot [L]}{\Delta F}$$

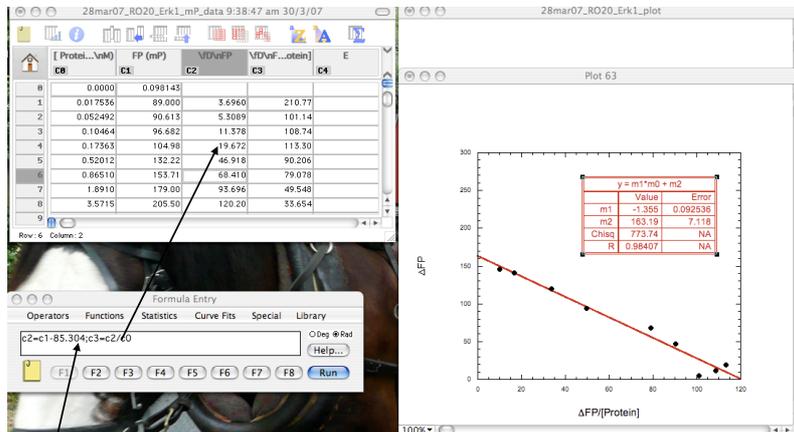
$$\Delta F = \frac{\Delta F_{\max} \cdot [L]}{K_d + [L]}$$

Transform data

$$\Delta F = \Delta F_{\max} - K_d \cdot \frac{\Delta F}{[L]}$$



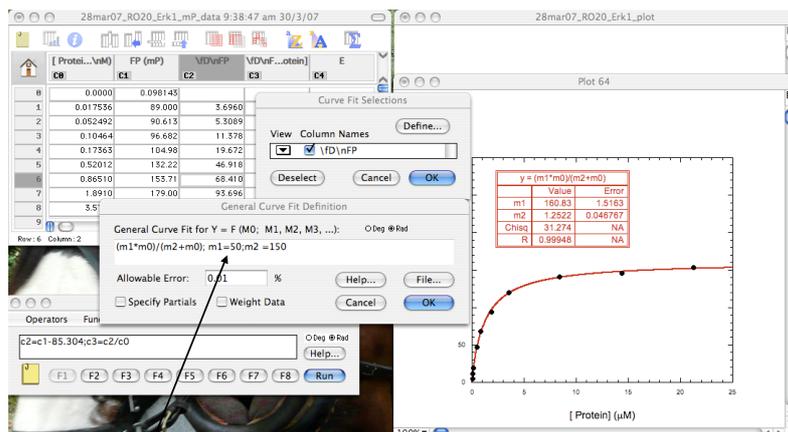
Fitting: Linear Transformation



Transform data
(if appropriate)

$$\Delta F = \Delta F_{\max} - K_d \cdot \frac{\Delta F}{[L]}$$

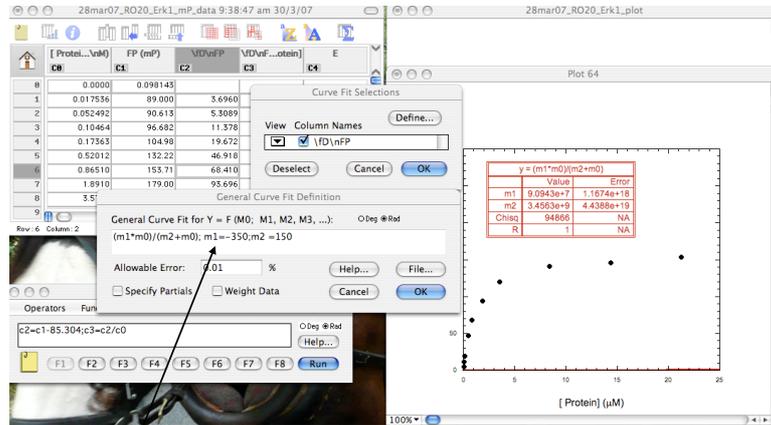
Fitting in Kaleidagraph: Estimates



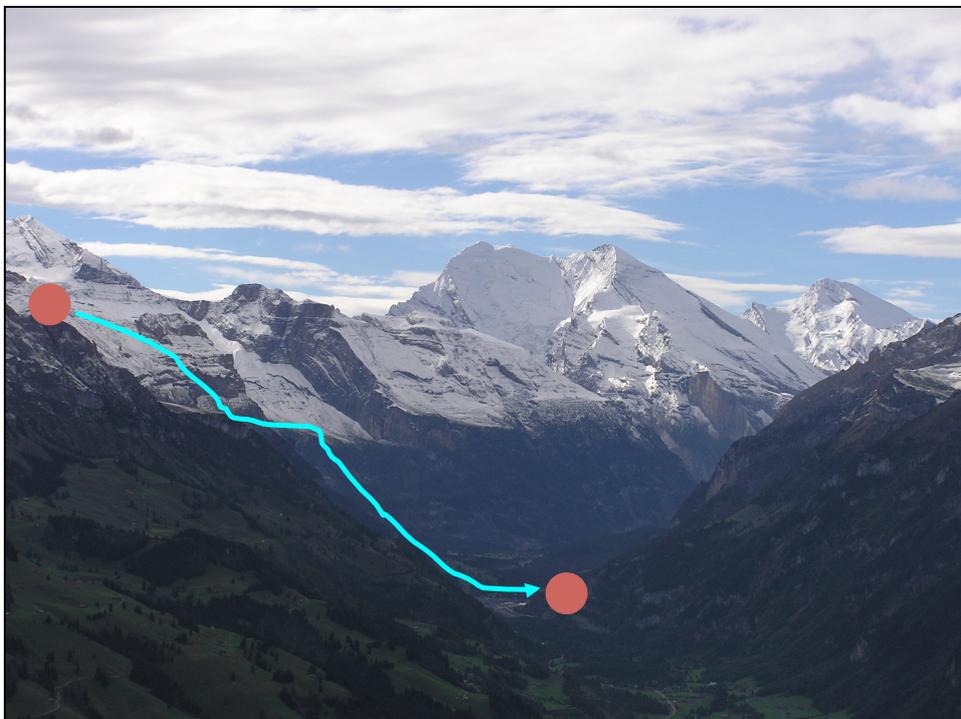
Estimates

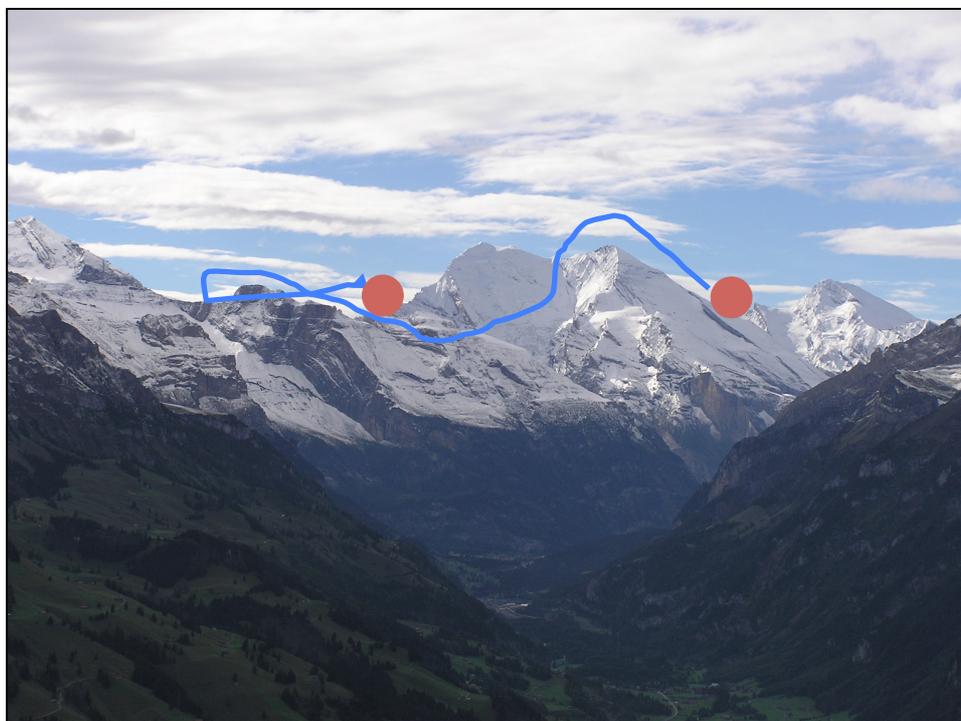
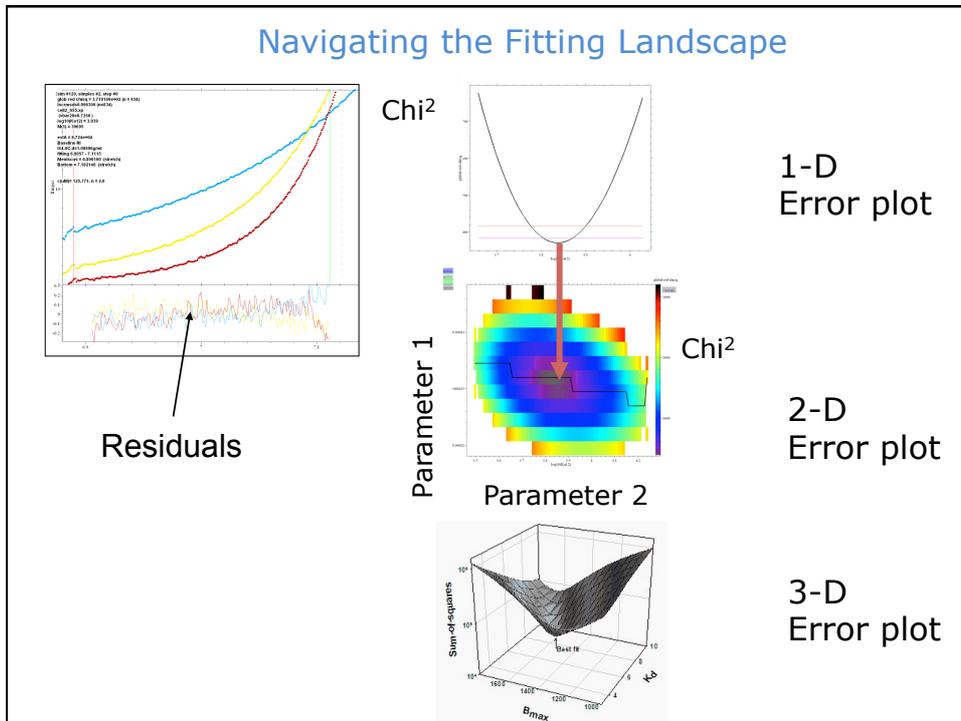
$$\Delta F = \frac{\Delta F_{\max} \cdot [L]}{K_d + [L]}$$

Fitting in Kaleidagraph: Estimates

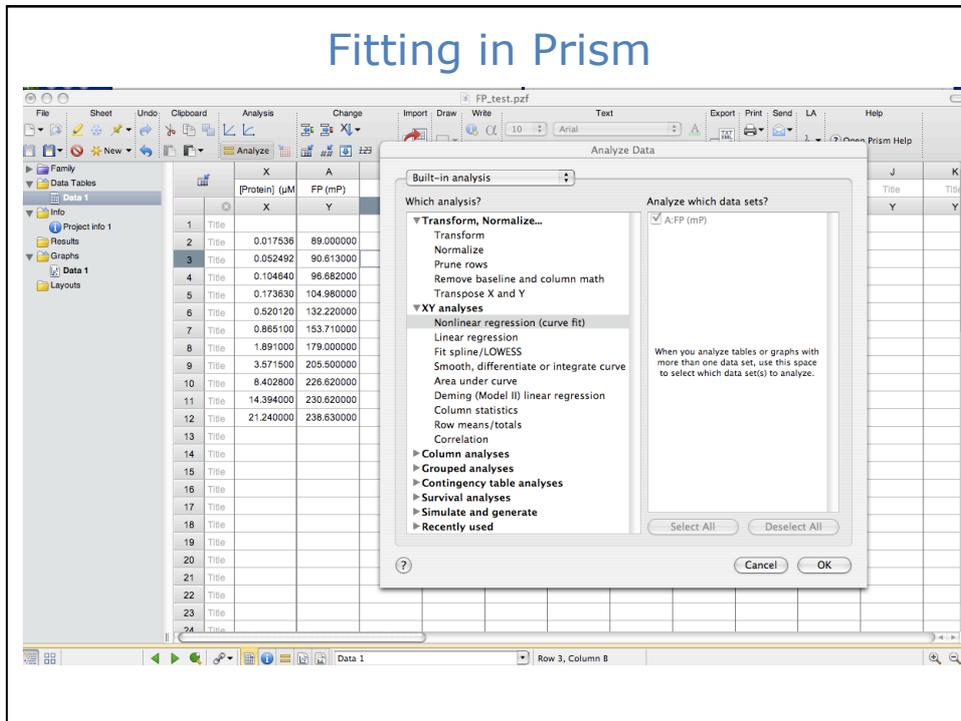


Bad estimates

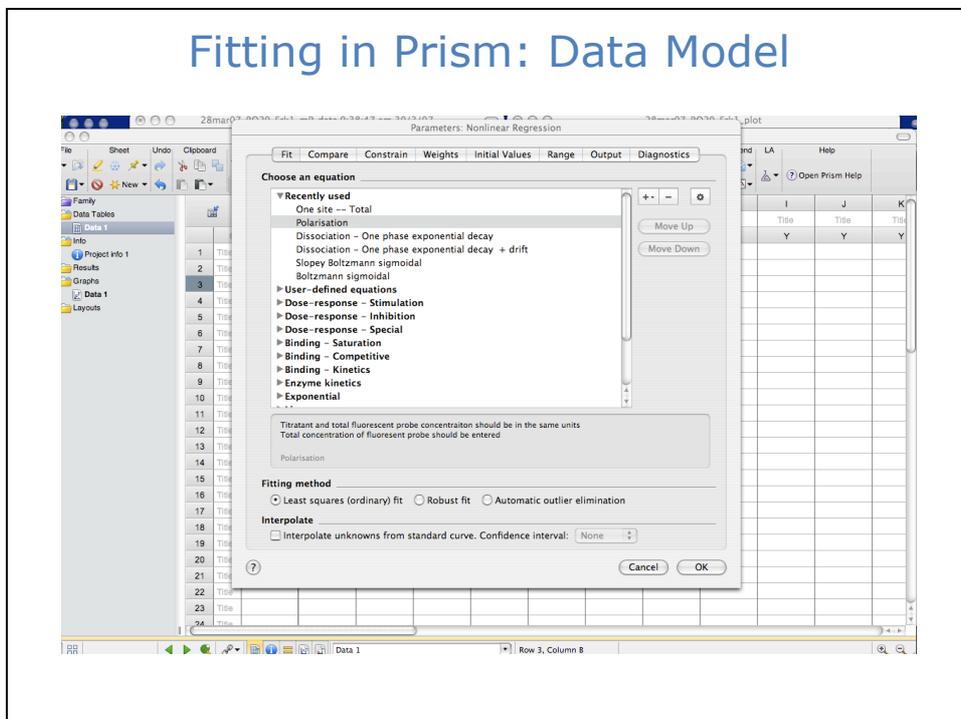




Fitting in Prism



Fitting in Prism: Data Model



Fitting in Prism: Equation

Equation Rules for Initial Values Default Constraints Transforms to Report

Polarisation equation
 Tip: Titrant and total fluorescent probe concentration should be in the same units
 Total concentration of fluorescent probe should be entered

Bottom: polarisation in absence of titrant
 Top: polarisation at saturation
 C: is the constant fluorescent probe concentration
 Kd: dissociation constant

Slopey Boltzmann sigmoidal
 Boltzmann sigmoidal
 User-defined equations
 Slopey Boltzmann sigmoidal
 Slopey Boltzmann sigmoidal [2]
 Dissociation - One phase exponential decay + drift
 Polarisation

$$Y=Bottom + ((Top-Bottom)/((C+X+Kd)-(((C+X+Kd)^2-(4*C*X))^0.5)))/(2*C)$$

Fitting method
 Least squares (ordinary) fit Robust fit Interpolate



Fitting in Prism: Parameter Estimates

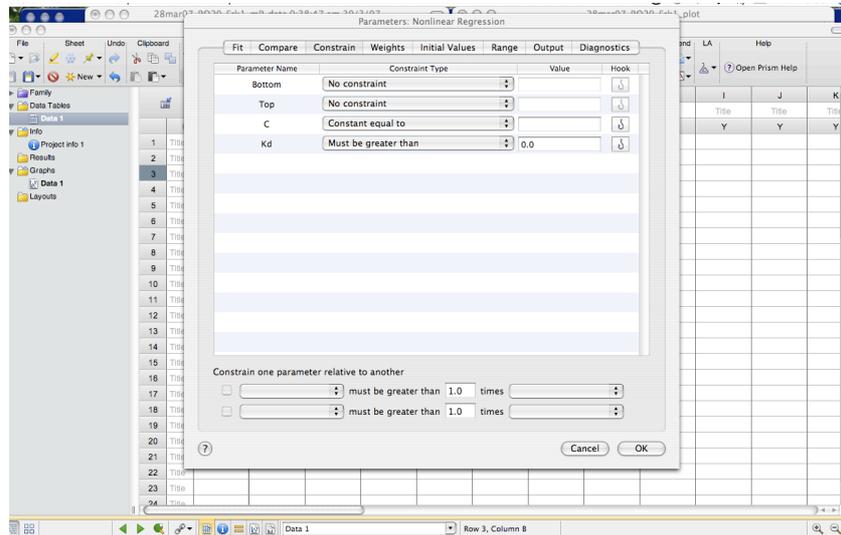
Parameters: Nonlinear Regression

Define a set of rules to compute initial values to use as a default every time you curve fit this equation.

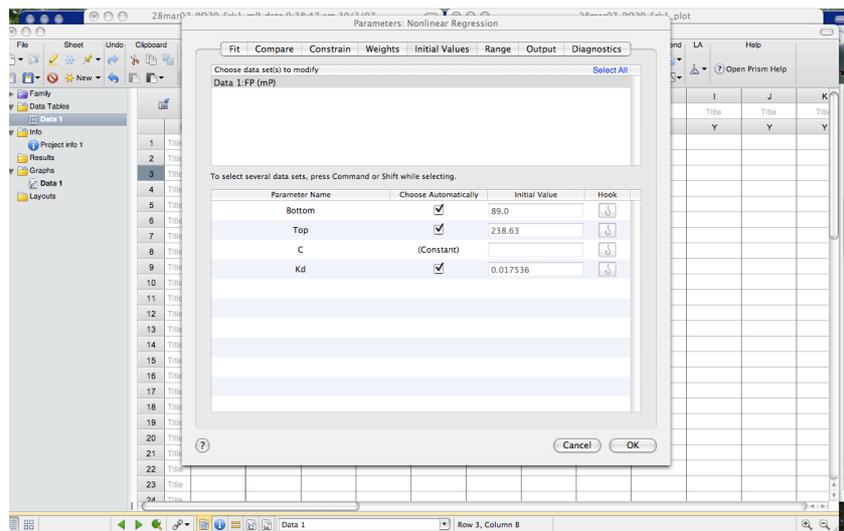
Parameter	Value	Rule
Bottom	1.0	*YMIN
Top	1.0	*YMAX
C	1.0	(Initial value, to be fit)
Kd	1.0	*(Value of X at YMIN)

Default range
 Start graphing the curve at: The smallest X value X = 0

Fitting in Prism: Constraints



Fitting in Prism: Estimate Check



Fitting in Prism: Result

Values →

Parameter	Best-fit value
Bottom	85.24
Top	246.0
C	= 0.0100
Kd	1.244
Std. Error	
Bottom	1.289
Top	1.630
Kd	0.0517
95% Confidence Intervals	
Bottom	82.33 to 88.36
Top	242.3 to 249.8
Kd	1.102 to 1.387
Goodness of Fit	
Degrees of Freedom	8
R square	0.9991
Absolute Sum of Squares	33.49
Sy.x	2.046
Constraints	
C	C = 0.0100
Kd	Kd > 0.0
Number of points	11

Goodness of Fit $R^2 = \frac{SS_{reg}}{SS_{tot}}$

Error →

How Good is My Fit? Residuals

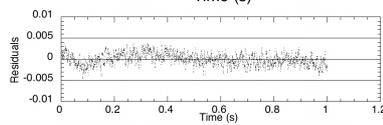
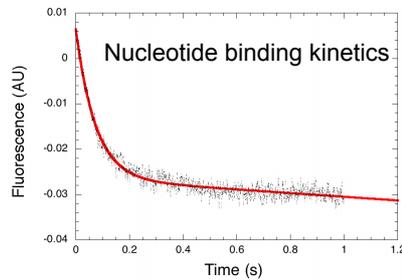
In Kaleidagraph:

In Prism:

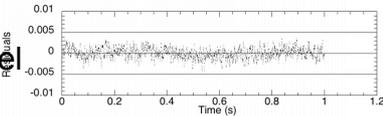
Single

Double

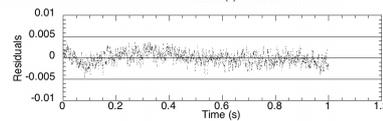
Does The Shoe Fit the Data ? Residuals



Single



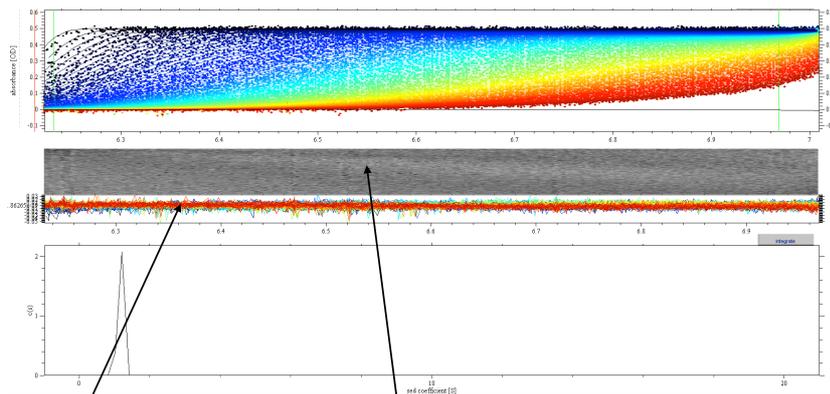
Single + drift



Double

Systematic deviation?
Decide which data model
Change physical model

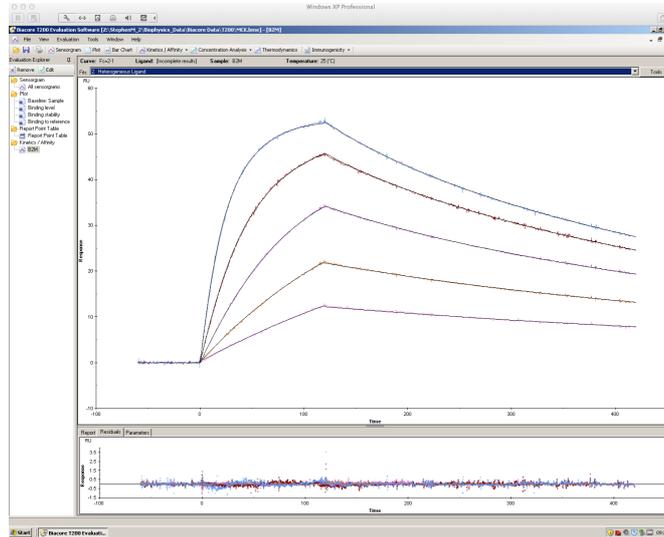
Automatic Residuals: AUC



Residuals overlay

Residual bitmap

Automatic Residuals: SPR



Global vs Individual Fitting

Why do you need to fit globally? Constraining all data to fit to one set of parameters. Experimental variation?
 Is the data model multi-component and difficult to fit e.g. AUC velocity or is it bad data?
 How does global compare to individual fits

Parameters: Nonlinear Regression

Parameter Name	Constraint Type	Value	lock
HillNM	Data set constant (=column title)		
Koff	Shared value for all data sets		
Bmax	No constraint		
Time0	Constant equal to	150.0	
NS	Constant equal to	0.0	

Constrain one parameter relative to another

- must be greater than 1.0 times
- must be greater than 1.0 times

Table of results		Global (shared)
2	Best-fit values	
3	HillNM	
4	Kon	1920
5	Koff	4.4e-005
6	Bmax	
7	Time0	
8	NS	
9	KD	7.774e-009
10	Std. Error	
11	Kon	5.907
12	Koff	4.849e-007
13	Bmax	
14	KD	2.895e-10
15	95% Confidence Intervals	
16	Kon	1910 to 1930
17	Koff	1.406e-005 to 1.594e-005
18	Bmax	
19	KD	7.289e-001 to 8.287e-009
20	Goodness of Fit	
21	Degrees of Freedom	6358
22	R square	0.9993
23	Absolute Sum of Squares	128780
24	Syx	3.710
25	Constraints	
26	HillNM	

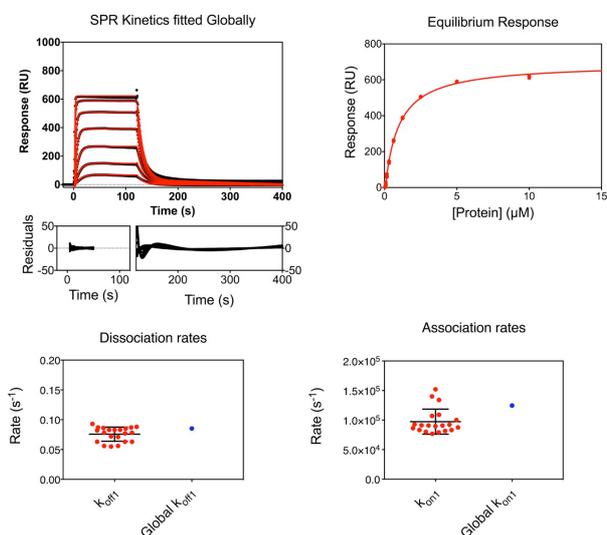
RU vs Time (s)

Residuals vs Time (s)

Shared values

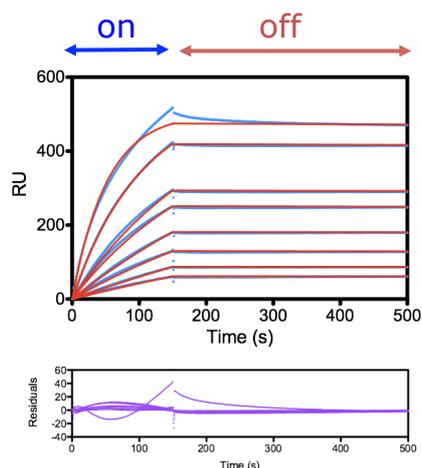
Global parameters

Global vs Individual Fitting



Fitting individually see distribution of constants
Fitting globally see only error of parameter fit

Poor Global Fitting Can Lead to Error

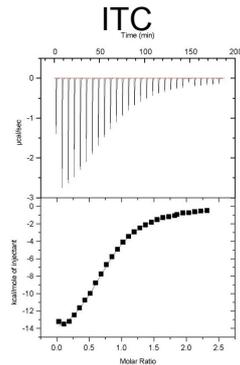


Global fit gives poor fitting for k_{off} very low value

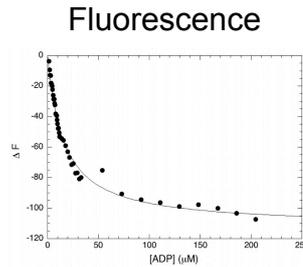
Global fit: estimated K_d approximately 30 nM
Individual fits to association: $K_d = 1.8 \mu\text{M}$
From ITC: $K_d = 2.2 \mu\text{M}$

Compare Techniques

Do fits to different physical signals give the same answer?
 Do the signals (after correction) represent the process?
 Is the physical model correct?



$K_d = 7.3 \pm 0.3 \mu\text{M}$
 $n = 0.8 \pm 0.1$



$K_d = 11 \pm 2 \mu\text{M}$

And of course repeat and repeat...

Statistics (Kaleidagraph)

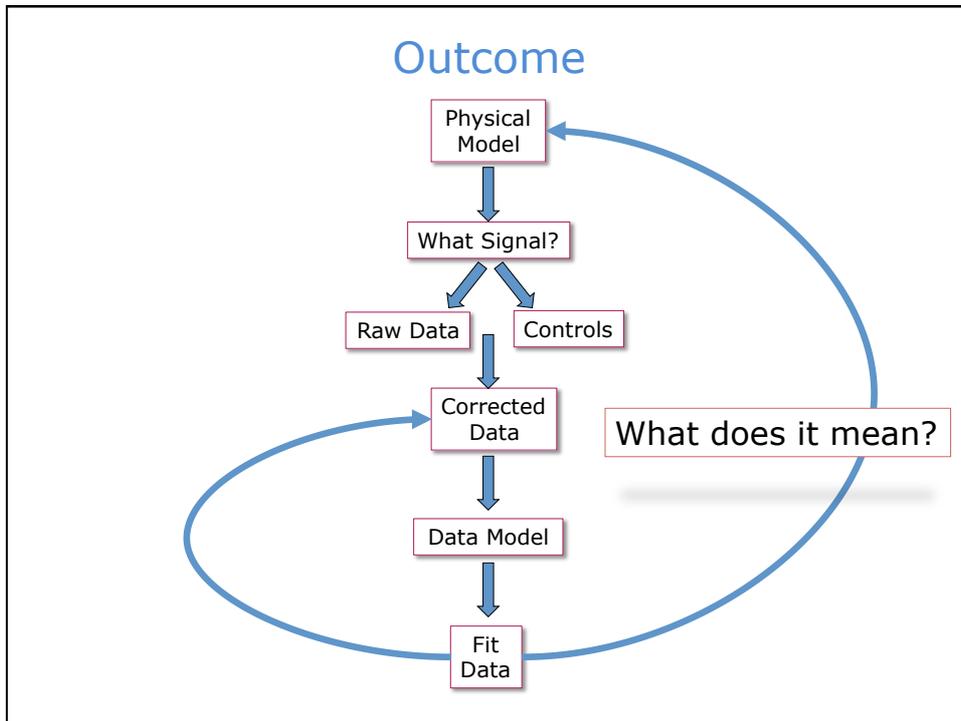
Parameter	Minimum	Maximum	Sum	Points	Mean	Median	Std Dev	Std Error	Skewness	Kurtosis	
Lm	79.0000	245.000	1950.0	12.000	162.50	143.00	14.415	42.221	20.480	15.094	0.75421
Tb	81.0000	891.000	5507.0	15.000	367.13	456.00	471.14	214.64	46071	59.551	0.54427
Tc	170.000	398.000	3020.0	15.000	201.33	195.00	247.85	69.931	4995.4	19.395	1.2397
dNb	0.20000	1.4000	7.8000	10.000	0.78000	0.56500	0.37016	0.34042	0.11958	6.10700	0.80057

Statistics (Prism)

The screenshot shows the Prism software interface. On the left is a data table with columns A through H and rows 1 through 18. The 'Analyze' menu is open, showing various statistical options. On the right, the 'Parameters: Column Statistics' dialog box is visible, with 'Descriptive Statistics' checked, including 'Minimum and maximum', 'Quartiles (Median, 25th and 75th percentile)', 'Percentile (50.0)', 'Mean, SD, SEM', 'Coefficient of variation', 'Geometric mean', 'Skewness and kurtosis', and 'Column sum'. Below this, 'Confidence intervals' and 'Test if the values come from a Gaussian distribution' are also checked. The 'Calculations' section shows 'Subcolumns' set to 'Compute the mean of the subcolumns for each row, and then calculate column statistics of those means'. At the bottom right, a box plot shows 'Kd' on the y-axis (log scale from 0.01 to 100000) and 'Protein' on the x-axis. The plot displays individual data points with error bars for each protein.

Statistics (Prism)

This screenshot is identical to the one above, but with the equation $\Delta G = -RT \ln K$ displayed above the box plot. The box plot shows 'Kd' on the y-axis (log scale from 0.01 to 100000) and 'Protein' on the x-axis. The plot displays individual data points with error bars for each protein.



And finally.....

A medieval-style illustration of a man in a black robe, likely William of Occam, holding a quill and looking at a book. The background features a circular diagram and some text.

Occam's Razor

A cartoon illustration of a woman in an orange dress sweeping a blue rug with a broom. A cloud of dust is shown being swept away.

Occam's Broom