Data Fitting, Errors and Binding Curves

- Properties of Measurement and Data
- General data fitting
- Evaluating 'goodness of fit'
- Fitting binding data (Stephen)
Measurement according to William Thomson (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 for “I” while science is about “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 property being examined
- **Units**
Properties of Measurement

1. Best
   - Increasing accuracy
   - Increasing precision

2. OK
   - Increasing accuracy (average more)

3. Poor
   - Decreasing accuracy (except OK for differences between measurements if lack of accuracy is systematic)

Outliers (one off lack of accuracy and/or precision) can be removed if obvious.
Uncertainty in Measurement

Precision and Accuracy are subject to random and systematic error

- Random errors (noise)
  - Measurement magnitude relative to instrument noise (S/N)
  - Precision of tools used in preparation of samples (pipettes, concentration of stock solutions, etc)
  - Data may have stochastic as well as deterministic components

- Random errors (lack of precision) will average out

- Systematic errors (‘mistakes’)
  - Instrument calibration, instrument measurement range (offscales)
  - 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 (unless they are a one off)
Validity

Does measurement report on the property being investigated?

Thermal melts fit to a full thermodynamic 2-state equation give identical Tm values but the derivatives yield different maxima because of slopes in the baseline regions.

Fluorescence spectral maxima shifts on ligand binding but this does not report on the concentration of the complex because of its lower intensity compared to the free protein.
Measurements have units

- Incorrect or mixed units and errors in calculations can be costly leading to large systematic error. Mars climate orbiter (~ $250 million) was lost due to use of pound force seconds rather than Newton seconds.
- Beware joules vs calories, log vs ln, °C vs °K etc.
General data fitting
Non-linear least squares (NLLS)
NLLS curve fitting software

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

Spreadsheets
Good for simulation

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 idealized drawing (A) defined by 36 points and used least-squares Fourier sine series fits of the form $y(x) = a_0 + \sum_{n=1}^{N} \frac{b_n}{\sqrt{2}} \sin(n\pi x)$ to fit the data shown in (B). The fit is shown for $N = 5, 10, 20, 30$ and 50. The conclusion is that the 50-term model may not satisfy the third-grade art teacher, but would make every most chemical engineer into a preliminary design.
NLLS goals and ‘ideal’ data assumptions

- NLLS tests if a given $Y=f(x)$ is consistent with the data (goes through the points) 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 and uniformly distributed.

- 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).

• You know $x$ accurately; uncertainty is in the dependent $y$ variable.

• Measurements (points) are independent and uniformly distributed.

• 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.

<table>
<thead>
<tr>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>41.528</td>
</tr>
<tr>
<td>m2</td>
<td>0.0044783</td>
</tr>
<tr>
<td>m3</td>
<td>0.17453</td>
</tr>
<tr>
<td>m4</td>
<td>0.00010379</td>
</tr>
<tr>
<td>Chiq</td>
<td>0.0018607</td>
</tr>
<tr>
<td>R</td>
<td>0.97202</td>
</tr>
</tbody>
</table>

Start fit $m_4 = 1e5$

<table>
<thead>
<tr>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>37.019</td>
</tr>
<tr>
<td>m2</td>
<td>0.0031957</td>
</tr>
<tr>
<td>m3</td>
<td>-1.02666e7</td>
</tr>
<tr>
<td>m4</td>
<td>1.3994e-7</td>
</tr>
<tr>
<td>Chiq</td>
<td>0.0016436</td>
</tr>
<tr>
<td>R</td>
<td>0.97137</td>
</tr>
</tbody>
</table>

Start fit $m_4 = 1e7$
Cure fitting results: Goodness of fit

- How good are the global indices of fit (R and chi squared)?
- How precise or how well constrained are the best fit parameters (standard errors)?
- Visually examine the residuals look for systematic (non random) distribution....
Quality (‘goodness’) 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 (weighted errors) or an average across the data set
- This is what is being minimised during fitting, but as a number it is not informative

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

- R and \( R^2 \)
- Coefficient of determination... how much of the total variation in y is predicted by the fit as compared to deviation from a mean

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

95% confidence interval is ~ double the standard error
- 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
Qualitative (‘common sense’) assessment of the fit

- Are the best fit parameters physically reasonable? (negative rates/concentrations)
- Do some parameters have insignificant values / amplitude, or in the sum of terms do they have similar values, and so 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.
Random distribution of Residuals?

**Lin->Log x-axis**

- **A->B**
  - single exp

- **Signal**
  - Time (s): $0, 5 \times 10^{-5}, 0.0001, 0.00015, 0.001$

**Log x-axis**

- **A->B->C**
  - double exp

- **Signal**
  - Time (s): $10^7, 10^6, 10^5, 0.0001, 0.001$

- **Residuals**
  - Time (s): $10^7, 10^6, 10^5, 0.0001, 0.001$
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 introduce 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

Slopes and intercepts from bootstrap fits fit to normal distribution. Negative slope valid despite outliers.
Fitting errors versus real errors.

**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 and required precision 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:

\[
A = B + C + \ldots \\
\partial A = \sqrt{\left(\partial B\right)^2 + \left(\partial C\right)^2} \ldots
\]

\[
A = B \times C \times \ldots \\
\frac{\partial A}{A} = \sqrt{\left(\frac{\partial B}{B}\right)^2 + \left(\frac{\partial C}{C}\right)^2} \ldots
\]
Data Fitting in Practice

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

Richard Henderson (2013) PNAS
fMRI of an Atlantic salmon

fMRI of a salmon showing brain regions responding to pictures of human emotion
Am I sure my data is statistically above the background or am I just looking at a dead salmon?

fMRI of a dead salmon demonstrating the importance of appropriate correction to avoid false positive results

Journals expect clarity on whether samples are biological or technical replicates?

Bayer HealthCare: only 25% academic published targets validated (n=67)

Amgen: 11% “landmark studies” confirmed (n=53)
Think first

Physical Model

What Signal?

Raw Data

Controls

Corrected Data

Data Model

Fit Data
What will my experiment tell me?

- Physical Model
  - Representative?
    - Raw Data
    - Controls
  - What Signal?
    - Corrected Data
  - Assumptions?
    - Good Enough?
    - Data Range/Time?
- Data Model
  - Fit Data
Outline

Physical Model

What Signal? Controls

Corrected Data

Data Model

Fit Data

Representative? Assumptions?

Raw Data

Data Range/Time?

Good Enough?
The right signal?

Physical Model

\[ E + S \rightleftharpoons E.S \]

\[ k_1 \]

\[ K_d \]

\[ k_{-1} \]

\[ k_3 \]

\[ k_4 \]

\[ E + P \rightleftharpoons E.P \]

\[ k_{-4} \]

Signal Change

\[ \Delta F \]

Corrected Data

\[ \Delta F \]
The right signal?

Physical Model: $E + S \rightleftharpoons E.S \rightarrow E.P \rightleftharpoons E + P$

Corrected data: $\Delta F$

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

Assumptions:

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

Physical model could be incomplete leading to incorrect assumptions

\[ \Delta F \propto \text{[E.S] + [E*.S] + ...?} \]

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

FIM-1

Polarisation

kinase
Is the signal change representative?

FIM-1

Fluorescein

Polarisation

Polarisation

kinase

kinase
Is the signal change representative?

$K_d$ labeled titratant = $K_d$ fluorescein

Hence non-specific
Is reaction specific?

All three probes bind to the kinase
Can they be displaced by a known active site inhibitor?

What are your controls?
Is reaction specific?

Blue: not displaced non-specific
Red: partially displaced
Green fully displaced specific

What are your controls?
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?

Methodology

Good signal change?
Background?
Buffer Effects?
Accurate titrations?
Temperature control?
Time to reach equilibrium?

Controls

Positive control?
Negative control?
Reagents defined and stable?
Is the Assay Robust?

- Consistent signal window between different protein preps and across plates (technical and biological replicates)
- High $Z'$

$$Z' = 1 - \frac{3(\sigma_{\text{max}} - \sigma_{\text{min}})}{\left|\mu_{\text{max}} - \mu_{\text{min}}\right|}$$
Equations for determining assay performance or sensitivity

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Equation</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient of variation</td>
<td>( %CV = \frac{\sigma}{\mu} \times 100 )</td>
<td>A measure of the precision relative to the mean value, calculated for the maximum and minimum signals. An acceptable limit is &lt;15%.</td>
</tr>
<tr>
<td>Signal to noise</td>
<td>( S:N = \frac{\mu_{\text{max}} - \mu_{\text{min}}}{\sigma_{\text{min}}} )</td>
<td>A measure of the signal strength. This equation is sometimes given as: ( \frac{\mu_{\text{max}} - \mu_{\text{min}}}{\sqrt{\sigma_{\text{max}}^2 + \sigma_{\text{min}}^2}} )</td>
</tr>
<tr>
<td>Signal to background</td>
<td>( S:B = \frac{\mu_{\text{max}}}{\mu_{\text{min}}} )</td>
<td>Usually calculated using control compounds, acceptable value &gt;2-fold.</td>
</tr>
<tr>
<td>Signal window</td>
<td>( SW = \frac{\mu_{\text{max}} - \mu_{\text{min}} - 3(\sigma_{\text{max}} + \sigma_{\text{min}})}{\sigma_{\text{max}}} )</td>
<td>The significant signal between max and min controls, acceptable value &gt;2-fold(^b).</td>
</tr>
<tr>
<td>( Z' ) factor</td>
<td>( Z' = 1 - \frac{(3\sigma_{\text{max}} + 3\sigma_{\text{min}})}{\mu_{\text{max}} - \mu_{\text{min}}} )</td>
<td>Representation of the SW using a score where a value &gt;0.5 represents an acceptable assay. ( Z' ) is measured in the absence of library compounds; ( Z ) is in the presence.</td>
</tr>
<tr>
<td>Minimum significance ratio</td>
<td>( MSR = 10^{2z_d} )</td>
<td>Smallest potency ratio between two measurements that is statistically significant (with 95% confidence).</td>
</tr>
</tbody>
</table>

\( \sigma \), s.d. of the assay signal; \( \mu \), mean of the assay signal; \( \sigma_d \), s.d. of the difference in log potency; \( \text{max} \), maximum signal; \( \text{min} \), minimum signal. \(^a\)This is a redundant description of the signal window equation. \(^b\)Some 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 \( 0 < Z' < 0.5 \).

Biological vs Technical Replicates?

**Biological replicates**

**Technical replicates**

**Protein Preps**

**Biological replicates**
Can I repeat my experiment?

- Valid targets
- Authentic reagents
- Method
- Physical Model
- Signal
- Raw Data
- Controls
- Corrected Data
- Data Model
- Fit Data
- Documentation
"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!
"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!

Further purification 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?
Does titratant effect signal or measurement?

- Titratant may contributes to signal e.g. light-scattering, heat of dilution in ITC.

- Buffer mismatch e.g. in AUC buffer sediments, DMSO in SPR and ITC.

- Titratant effects measurement e.g. the incident light (inner-filter effect)
Don’t be misled

Some of the data could be equally fit with linear function!
The shape of the binding curve is uniquely defined just the $R_{\text{max}}$ and $K_D$, so even curves with no observed inflection point can be well fit with fixed $R_{\text{max}}$ to estimate the $K_D$ in a reliable way.

Don’t be misled

They reference a Methods in Enzymology paper:

“The shape of the binding curve is uniquely defined just the $R_{\text{max}}$ and $K_D$, so even curves with no observed inflection point can be well fit with fixed $R_{\text{max}}$ to estimate the $K_D$ in a reliable way”

Is this correct?
Right Range? Taking the Log View

Full range

Limited no of points

Too short range

Amplitude

$K_d \geq \text{[Final]}$

Amplitude

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

Steady state

\[ E + S \xleftrightarrow{k_a}{k_d} ES \]

\[ \frac{-d[ES]}{dt} = k_a [E][S] - k_d [ES] = 0 \]

\[ K_d = \frac{k_d}{k_a} \]

Has the reaction reached equilibrium?
Compare data set at different times or increase time between additions
Right Range? Taking the Long View

Kinetics

\[ \text{E} + \text{S} \xrightleftharpoons[k_d]{k_a} \text{E.S} \]

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

\[ k_{\text{obs}} = k_a[S] + k_d \]

\[ y = A \times \exp(k_{\text{obs}} t) + B \times t + C \]

Has the endpoint been reached?

Rule of thumb \( t = 10-20 \times t_{1/2} \)
Right Range? Taking the Short View

Enzyme kinetics

\[ E + S \xleftrightarrow{k_{-1}} E.S \xrightarrow{k_1} E.P \xleftrightarrow{k_4} E + P \]

\[ E + S \xrightarrow{k_3} E.P \xleftrightarrow{k_4} E + P \]

Assumption: \([P] = 0\), therefore look at initial rates
\([E] \ll [S]\), so \([S] \approx [\text{free substrate}]\)

\[ v = \frac{k_3 [E_0][S]}{K_m + [S]} \]
Appropriate Concentrations

\[
P + L \rightleftharpoons PL \quad K_d = \frac{[P][L]}{[PL]}
\]

Fluorescence, FP, MST
Labeled \([L]\) (total) \(<< K_d\)
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?
Physical Model

What Signal?

Corrected Data

Raw Data

Controls

rmsd, residuals

Quality of Fit?

Estimates

Which Program? Global, Single?

Fit Data

Data Model

Assumptions?

• Is my data good enough?
• Am I using the right equation?
• How good is the fit?
• How does this correlate with the hypothetical physical model?
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:

\[ Y = \frac{B_{\text{max}} \cdot X}{K_d + X} + NS \cdot X + \text{Background} \]

\( NS \cdot X \) is linear non-specific binding? Is this real for my system?

Linear saturation
Wrong Kd
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!

Removing bad data is sometimes necessary.

Removing outliers can be cheating.
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

\[ F = F_0 + \left( F_1 - F_0 \right) \left( \frac{[P][L]}{[PL]} \right) \]

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

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

\[ [L_T] = [L] + [PL] \]

\[ [P_T] = [P] + [PL] \]

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

Values

Error

Chi²

Correlation coefficient
Fitting: Log View
\[ \Delta F = \frac{\Delta F_{\text{max}} [L]}{K_d + [L]} \]
Fitting in Kaleidagraph: Estimates

Bad estimates
Navigating the Fitting Landscape

- **Residuals**
- **1-D Error plot**
- **Chi²**
- **Parameter 1**
- **Parameter 2**
- **2-D Error plot**
- **Chi²**
- **3-D Error plot**
Fitting in Prism
Fitting in Prism: Data Model
Fitting in Prism: Equation
Fitting in Prism: Parameter Estimates
Fitting in Prism: Constraints
Fitting in Prism: Estimate Check
Values

Fitting in Prism: Result

Goodness of Fit

Error

\[ R^2 = \frac{SS_{req}}{SS_{tot}} \]
How Good is My Fit? Residuals

In Kaleidagraph:

In Prism:
Does The Shoe Fit the Data? Residuals

Nucleotide binding kinetics

Systematic deviation?
Decide which data model
Change physical model

Single
Single + drift
Double
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

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 \( k_{\text{off}} \) 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?

Hsp90 + ADP ⇌ Hsp90.ADP

(ITC)

Fluorescence

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

$K_d = 7.3 \pm 0.3 \, \mu \text{M}$

$n = 0.8 \pm 0.1$

And of course repeat and repeat...
Statistics (Kaleidagraph)
Beware P Values!!!
The equation for the change in free energy, $\Delta G = -RT \ln K$, is shown in the image. This equation relates the change in free energy ($\Delta G$) to the reaction constant ($K$) and the temperature ($T$) in the thermodynamics of a reaction. The figure also includes a graph with data points representing the relationship between different variables, likely related to protein binding or other biochemical processes.
Outcome

Physical Model

What Signal?

Raw Data

Controls

Corrected Data

Data Model

Fit Data

What does it mean?
And finally.....

Occam’s Razor

Occam’s Broom