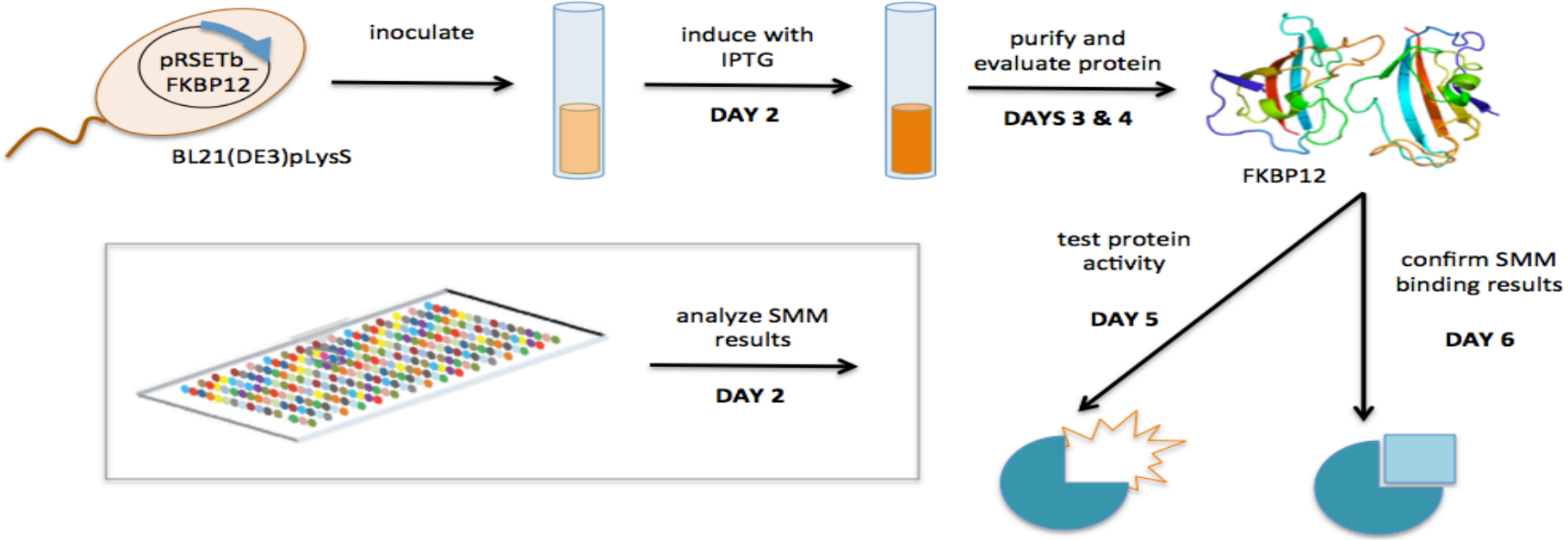
M1D7:
Complete data analysis

1. Pre-lab discussion
2. Practice statistics exercise
3. Analyze PPIase & DSF data

Important due dates!

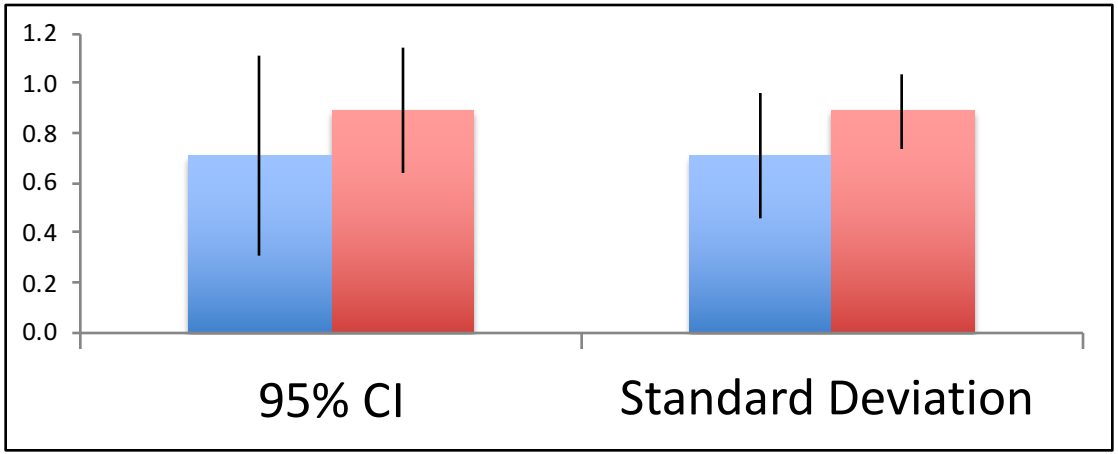
- Data summary draft due Mon, Mar 12 at 10 pm
  - Extra office hours Sat, Mar 10 from 10 am – 5p in 56-302
  - Standing office hours as scheduled
  - By appointment!
- Mini-presentation due Sat, Mar 17 at 10 pm
  - Stay tuned for additional office hours next week
- Blog post due Sun, Mar 18 at 10 pm
  - Watch for invite email next week

Overview of Mod1 experiments



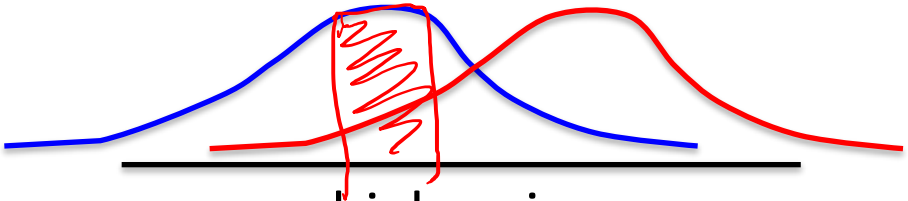
Confidence intervals show the variance in the data set

- At 95% confidence interval, there is a 95% chance that the true mean is within the defined range

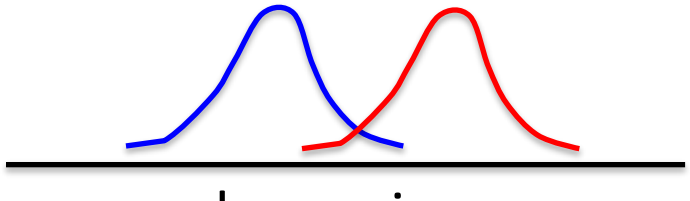


Student's *t*-test used to determine if populations are significantly different

- Follows *t*-distribution under null hypothesis
- At $p < 0.05$, there is less than a 5% chance that populations are the same (or there is a 95% chance that populations are different)
- Examines signal (means):noise (variance) ratio



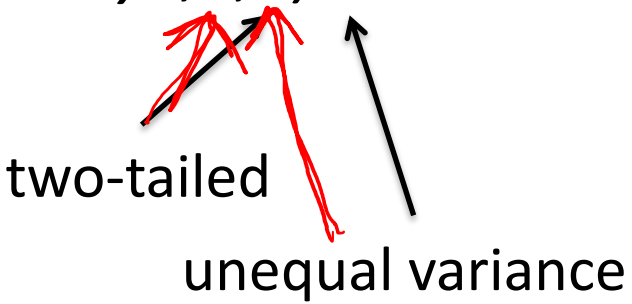
high noise



low noise

Calculating Student's *t* in excel

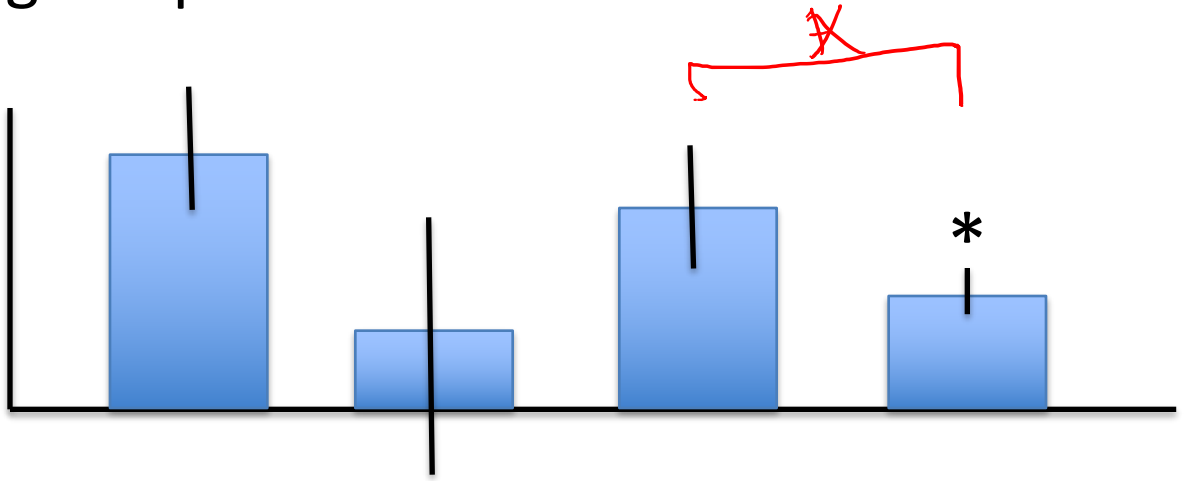
p = TTEST (array1, array2,2,3)



Can only compare two data sets at a time!

How will you use statistics in your data analysis?

- Specific activity values calculated from PPIase
- Melting temperatures determined from DSF



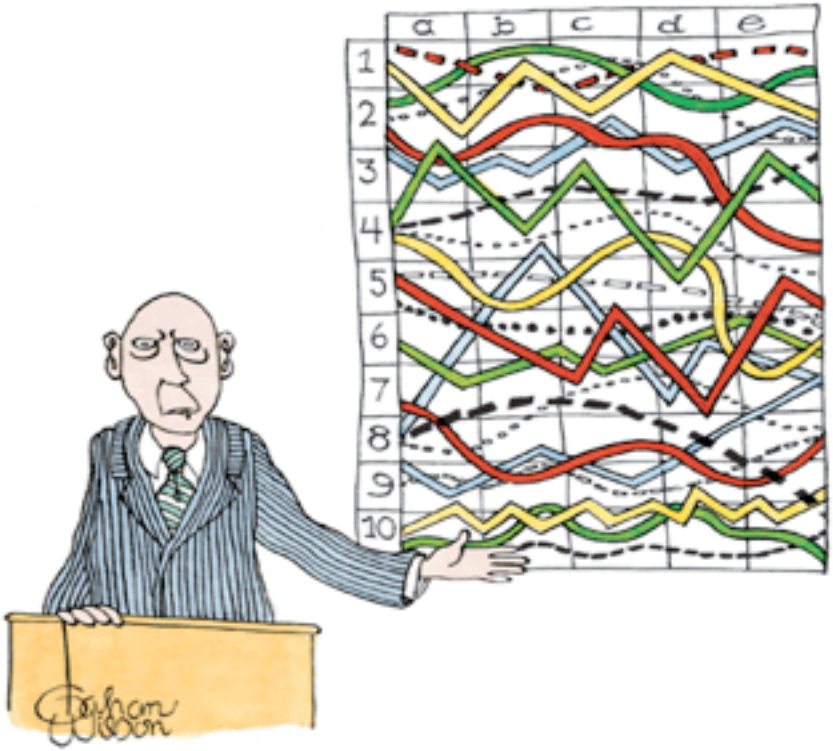
What if the data are not statistically significant?

What were your *expected* results?

- For the PPIase assay:

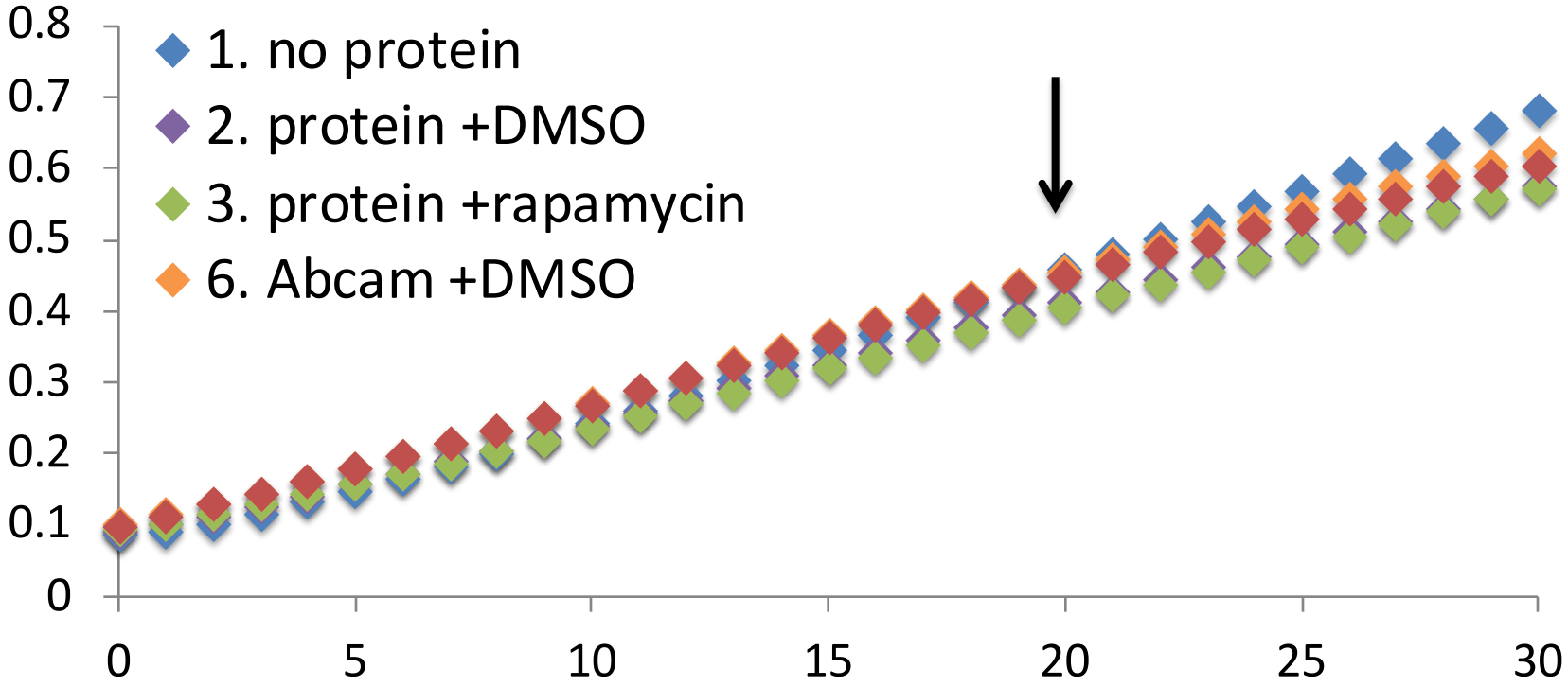
- For the DSF assay:

What were your *actual* results?

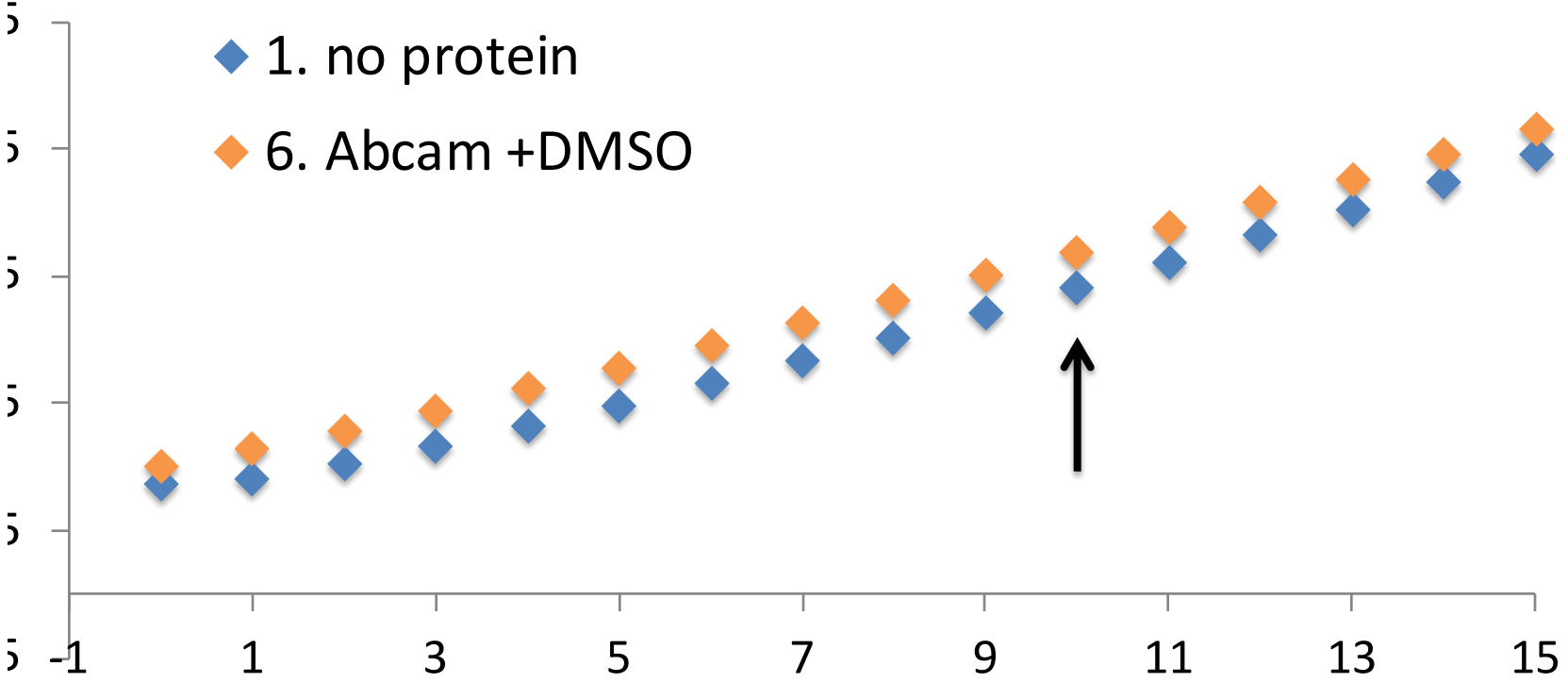


“I’ll pause for a moment so you can let this information sink in.”

Pooled class data show ‘crossover’ for test (Abcam) and control samples



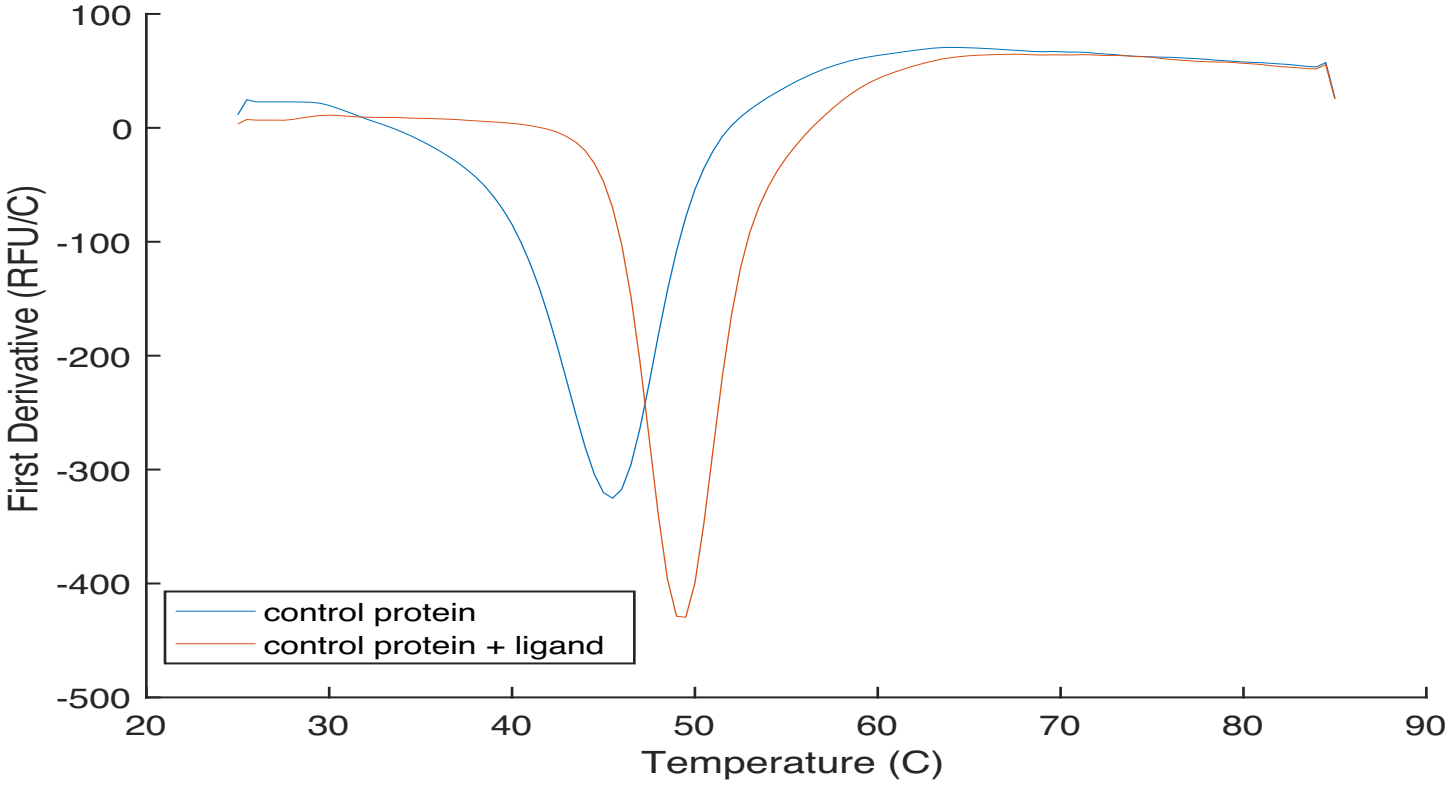
Let’s take a closer look



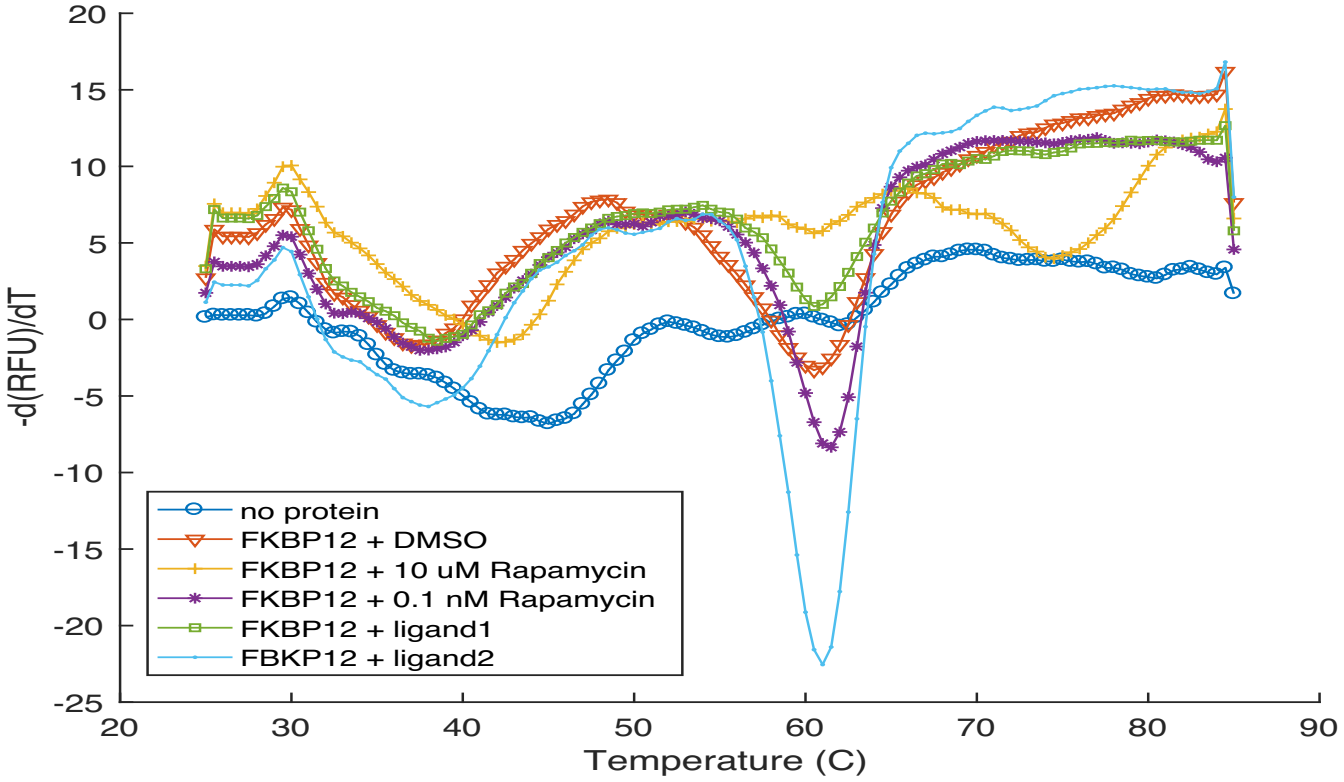
How will you analyze your PPIase data?

- You will be provided pooled class data
  - Plots for 30 min and 15 min timecourse
- Use t = 10 minutes for final timepoint in specific activity calculations
  - Obtain values from pooled data for Conditions #1, #2, #3, #6, and #7
  - Should still report your individual data!
- Compare your +ligand data to pooled data

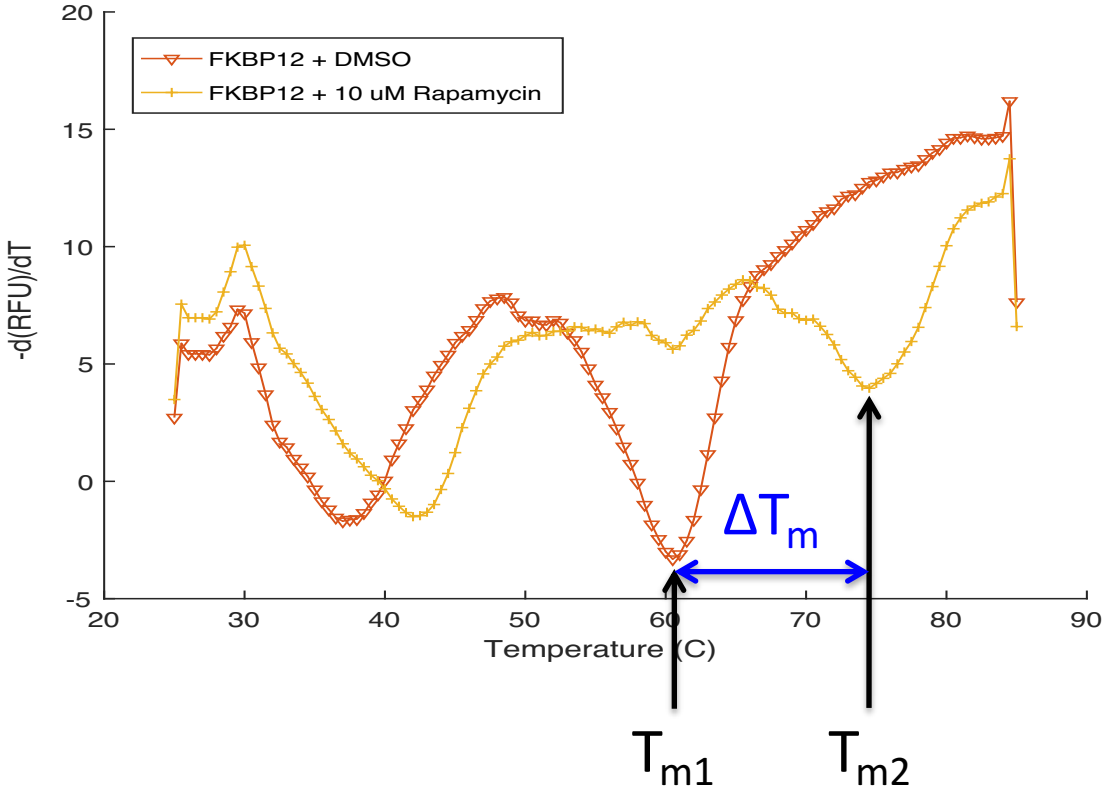
DSF assay controls look great!



Now, let's look at your data

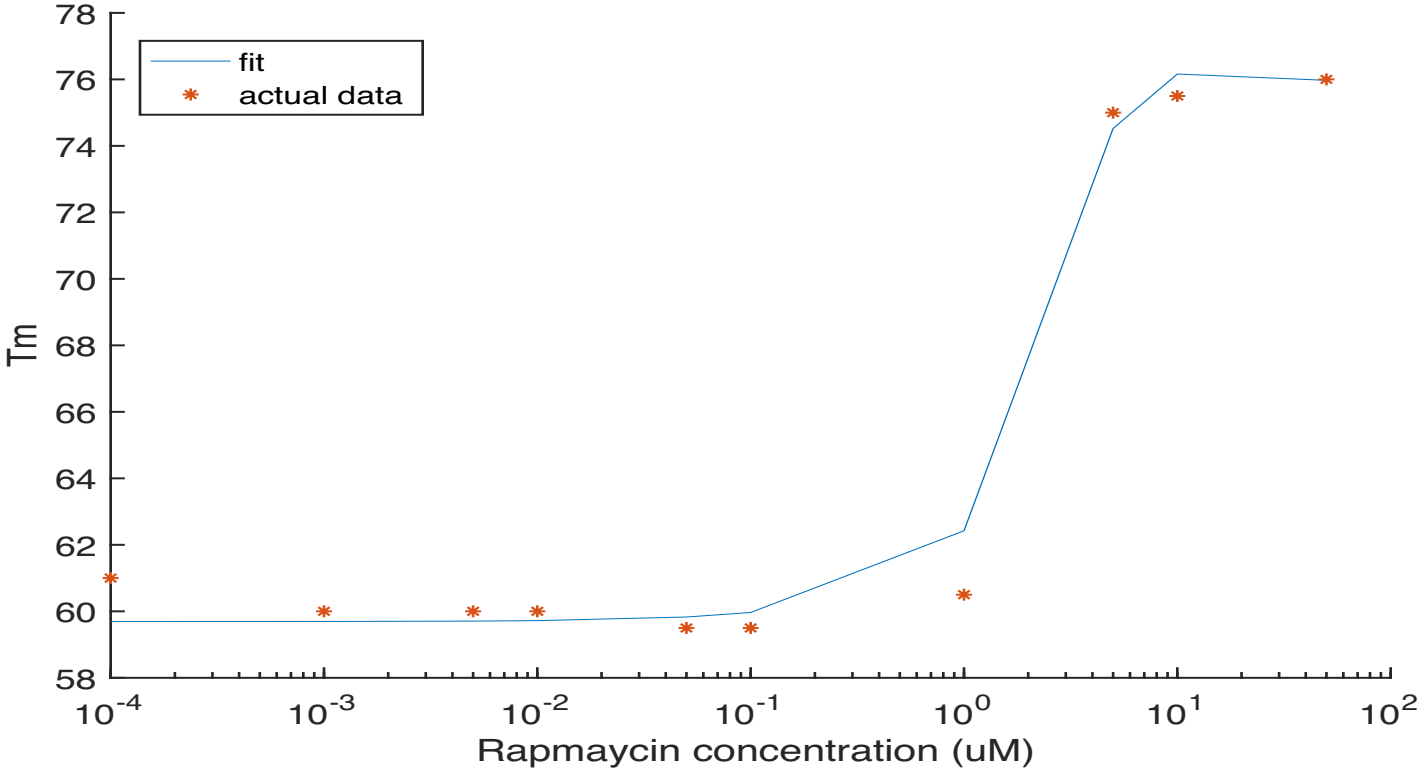


Define T_m using peak at highest temperature



- Find minimum of first derivative in relevant temperature range

Additionally, calculate the apparent Kd for DSF data analysis



How will you analyze your DSF data?

- Complete by-eye determination of T_m from plot of first derivatives
- Additional rapamycin concentrations used to calculate apparent K_d
  - You will be provided pooled class data
  - Use fit in MATLAB script to back-out value
  - If that fails, use your eyeballs

Be sure to post your data to the wiki!

- For the PPIase assay:
  - Single plot with all curves
  - Specific activity calculations
- For the DSF assay:
  - Single plot with all first derivative curves
  - Tm values
- Should be uploaded by 10 pm tonight!

Be sure all information is clearly labeled in excel spreadsheet

Notes on your Data summary

- Required to use class data
  - PPIase: pooled data AND comparison(s)
  - DSF: pooled K_d data AND comparison(s)
- Completed with your partner
  - Use individual assignments to generate a ‘polished’ draft
- Follow the format guidelines on the wiki
  - Review the example ‘data’ slide
- Redundancy serves a purpose!

Today in lab…

- Data analysis

Notebook due 10pm!

MIDI

For next time…

- Read Mod 2 overview and M2D1 introduction
- Prepare for in-class journal article discussion
  - Everyone expected to participate!

Lastly, some notes on previous homework

- Results slide draft (figure, title, caption, text)
  - Use specific nouns: protein vs FKBP12
  - Describe all data represented in the figure and specifically reference in the text

– Concision! In order to... vs to...

- Mini-presentation outline
  - Include your hypothesis
  - Be mindful of time limit and focus on key experiments