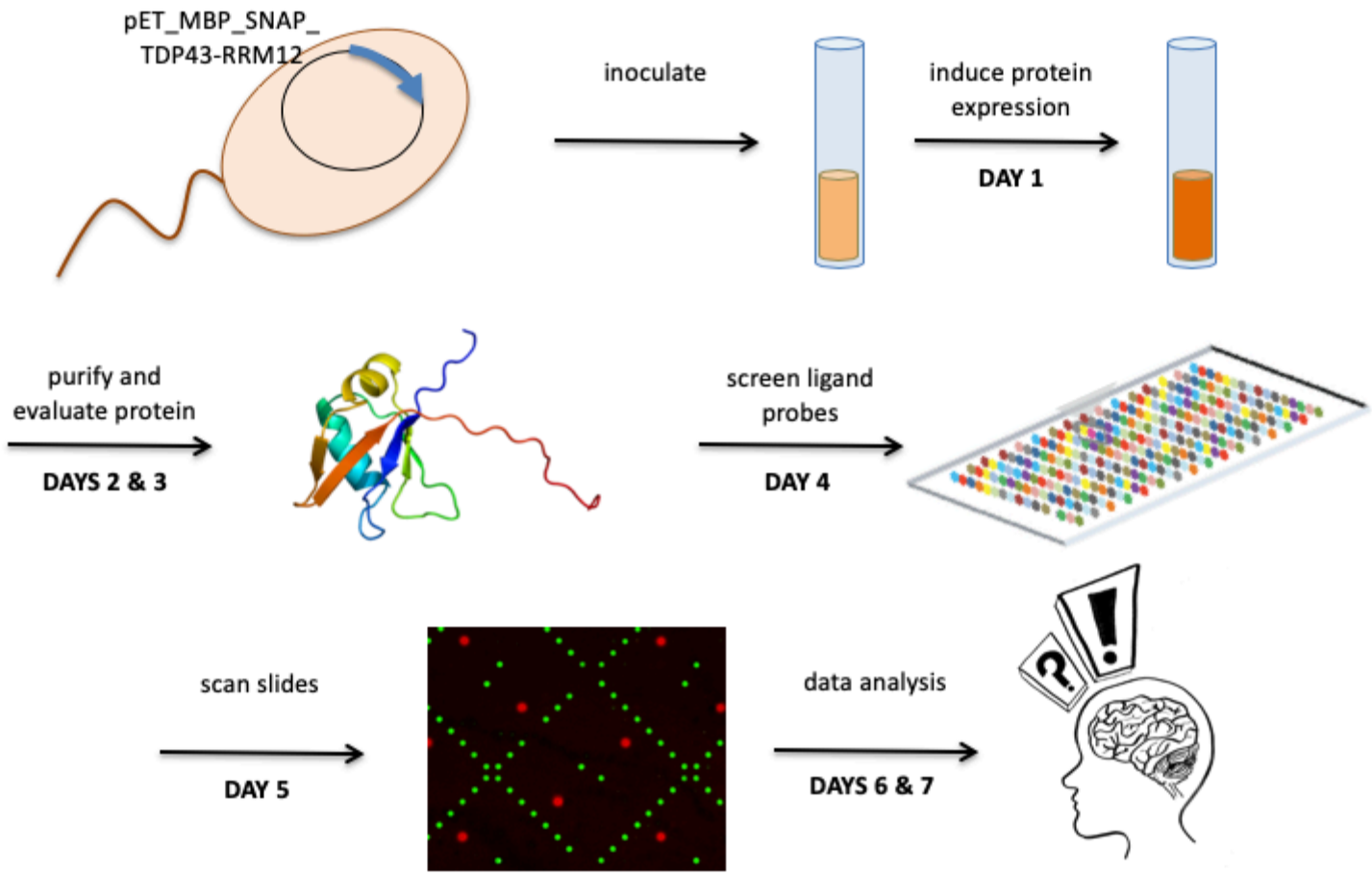
M1D6:
Analyze SMM data

1. BE Comm Lab workshop
2. Analyze SMM data
3. Work on Data summary!

Important due dates are approaching!

- Data summary (15%)
  - completed in teams and submitted via Stellar
  - draft due 3/8 at 10p, final revision due 3/22
  - format in bullet points
  - Extra Office hours Saturday, March 7 at TBD in 56-302
- Mini-presentation (5%)
  - completed individually and submitted via Gmail
  - due 3/15 at 10p
- Notebook (part of 10% Homework and Notebook)
  - due 3/4 at 10p via email to Kevin
- Blog (part of 5% Participation)
  - due 3/16 at 10p via Blogspot

Overview of Mod1 experiments

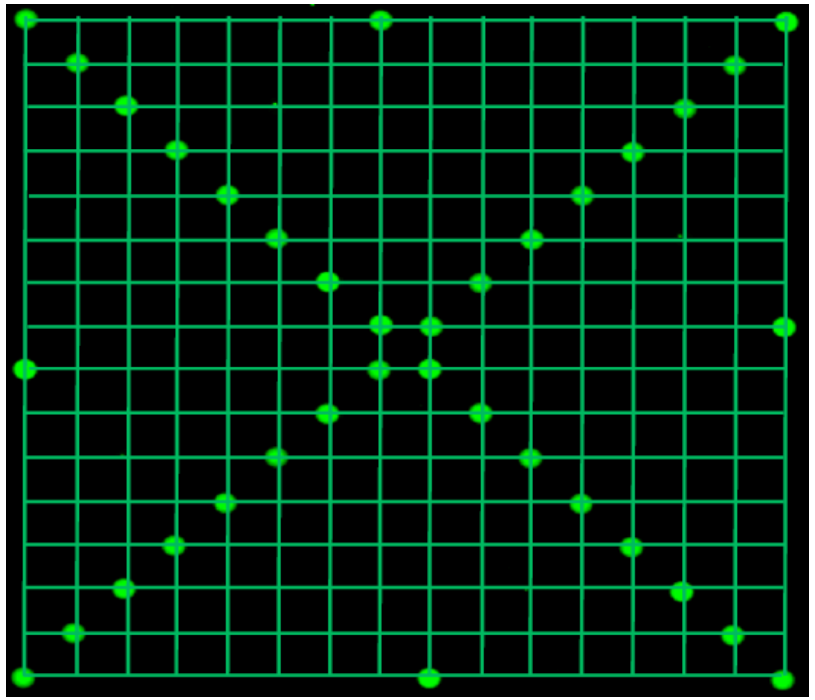


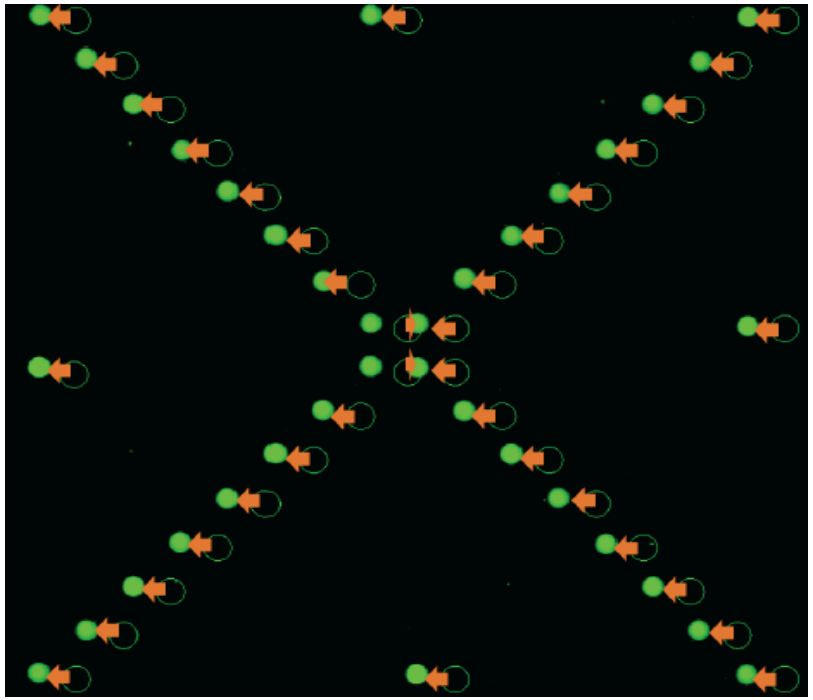
Workflow for SMM data analysis

1. Align spots using fluorescence on 532 nm channel (sentinel spots)
   - What is the source of the fluorescence?
2. Quantify fluorescence on 635 nm channel
   - What is the source of the fluorescence?
3. Identify ‘hits’ with improbably high fluorescence
4. Complete ‘by eye’ analysis of putative hits

Align SMM using sentinel spots

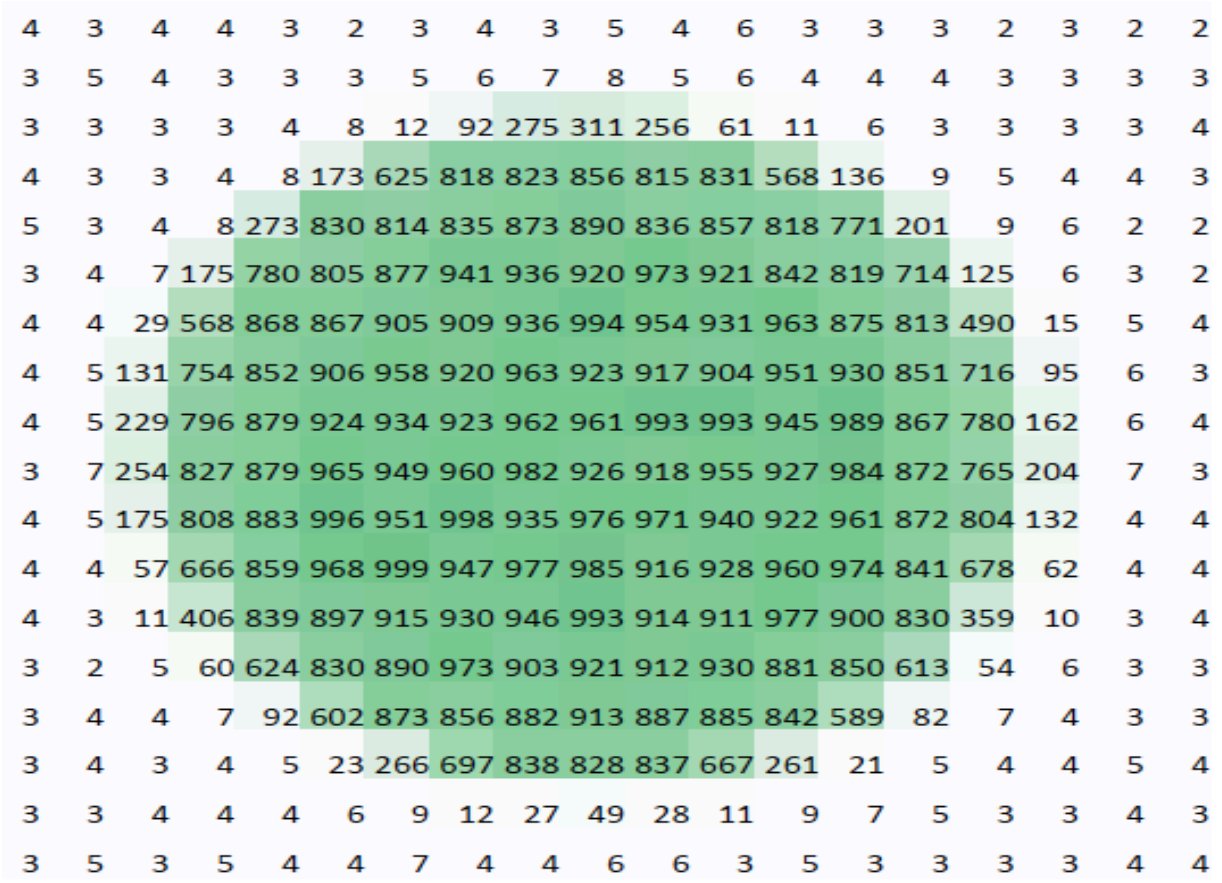
- Slides are printed in block patterns (16 rows x 16 columns)
- Each ligand spot is identifiable via intersecting lines between sentinels





Spots are represented by an array of numerical values

- Each pixel is represented by a number that indicates intensity of the signal
- Computational analysis used to define ‘hits’



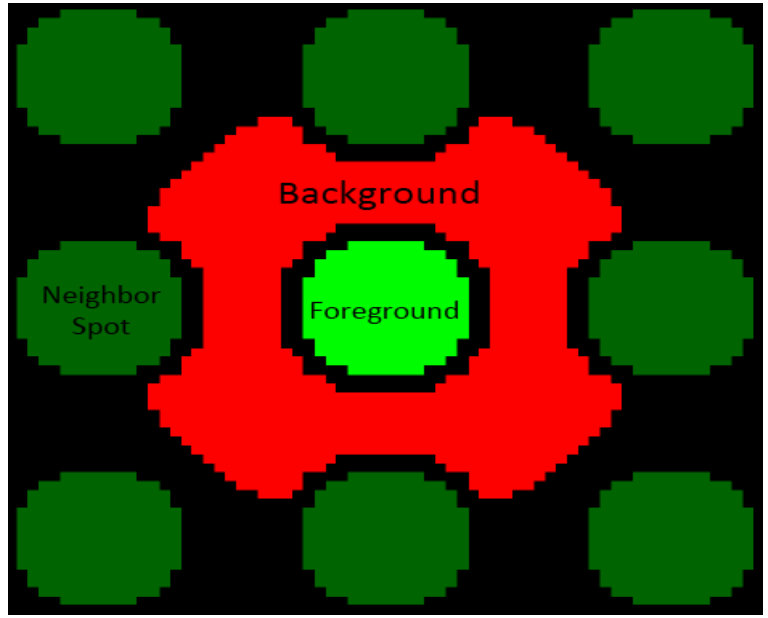
Fluorescence is quantified to identify hits

- Foreground:

ligand is printed

- Background:

carryover from ligand spot



$$\text{Signal-to-noise ratio (SNR)} = \frac{\mu_{foreground} - \mu_{background}}{\sigma_{background}}$$

How will you identify hits from the SMM data?

First, consider bias that exists in the data set

- Across all slides
- Within each block
- Within each slide

Then, identify hits with significantly higher fluorescence over background

Lastly, manually confirm hits to eliminate false positives

Identifying hits with significant fluorescence

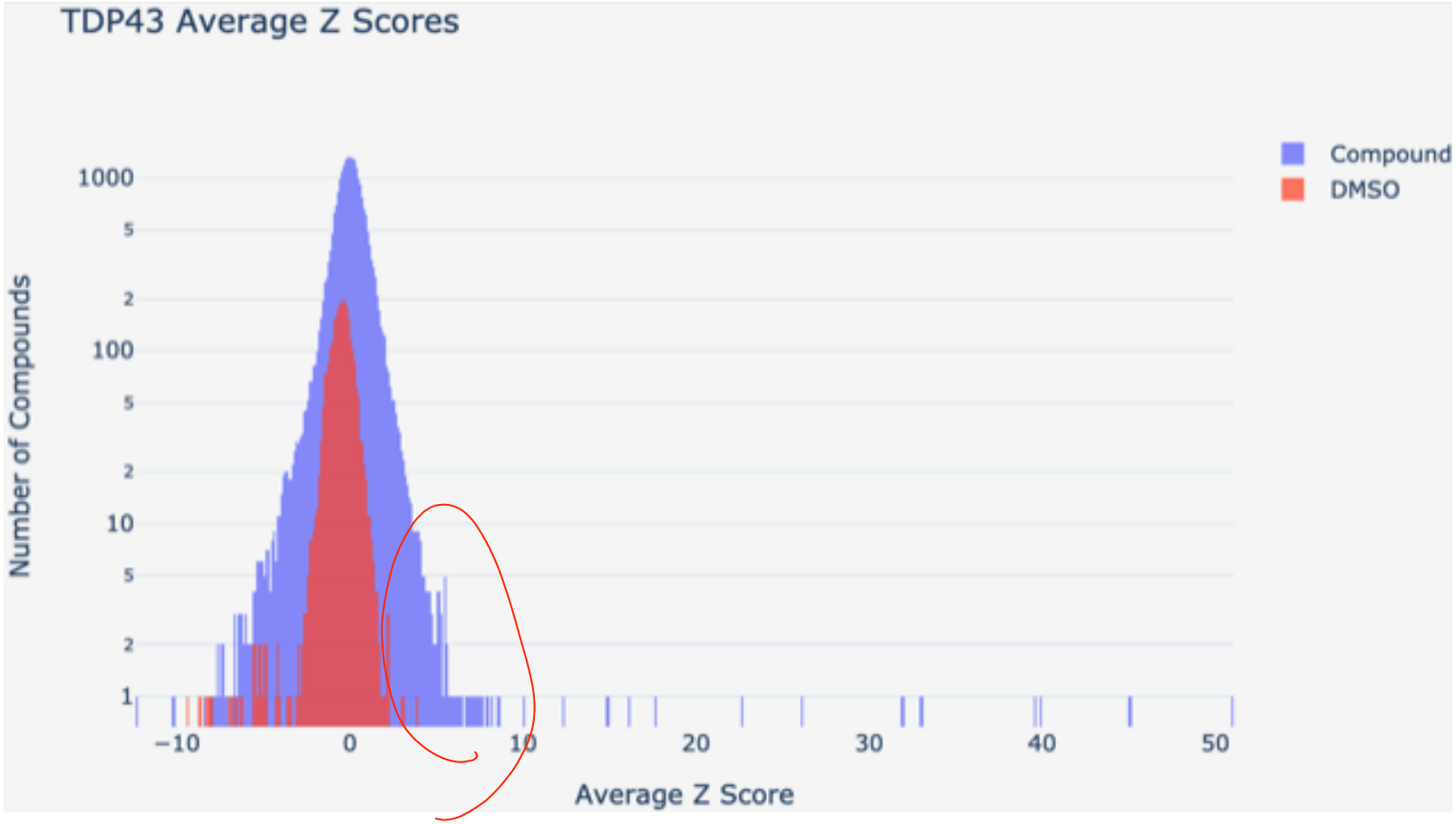
median absolute deviation (MAD)

$$\text{Robust Z-score} = \frac{SNR_i - \text{median}(SNR)}{\text{median}(|SNR_i - \text{median}(SNR)|) * 1.48}$$

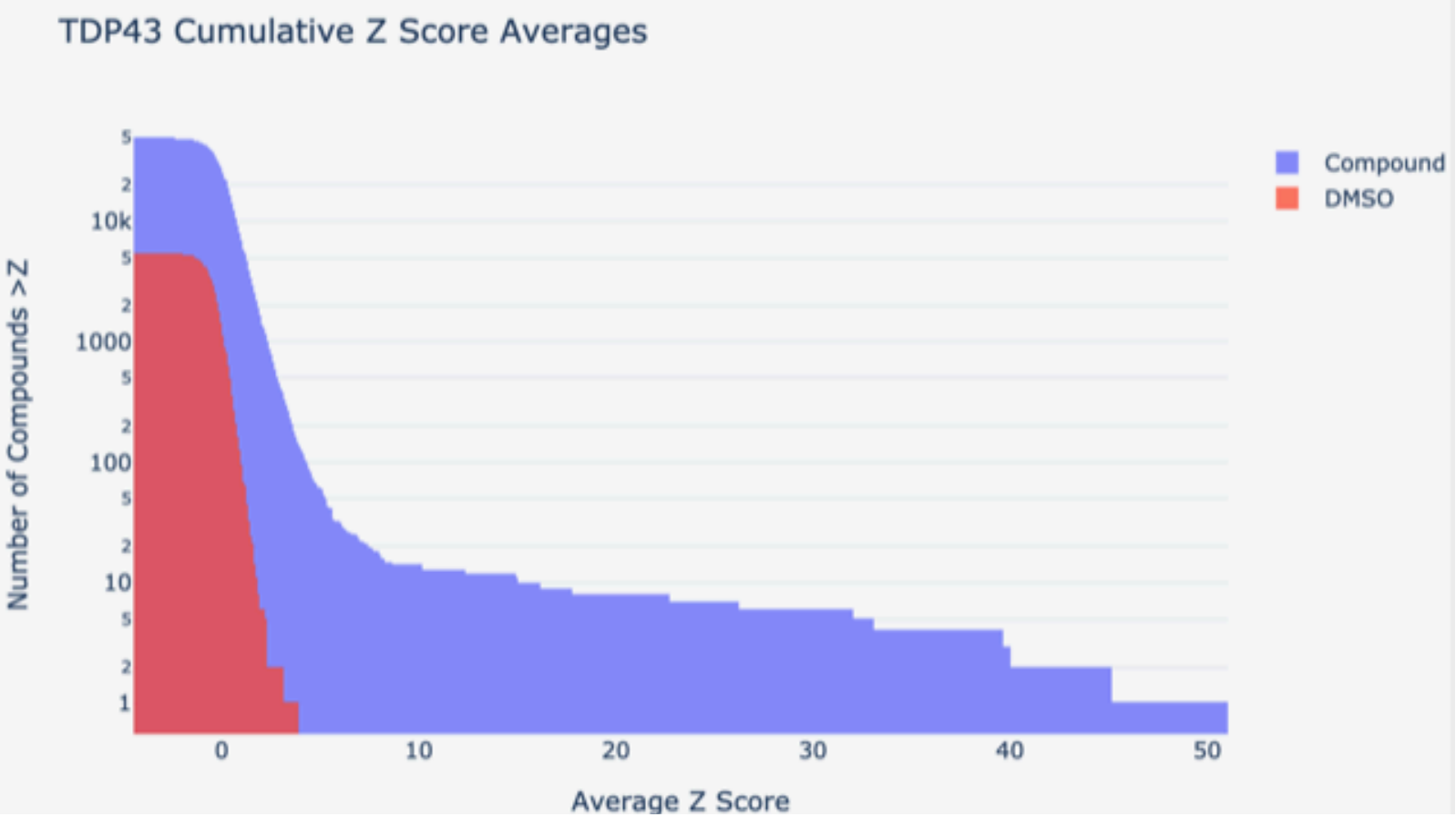
scale factor for the normal distribution

Robust Z-scores eliminate the influence of outliers

Average Z-score calculated for all compounds

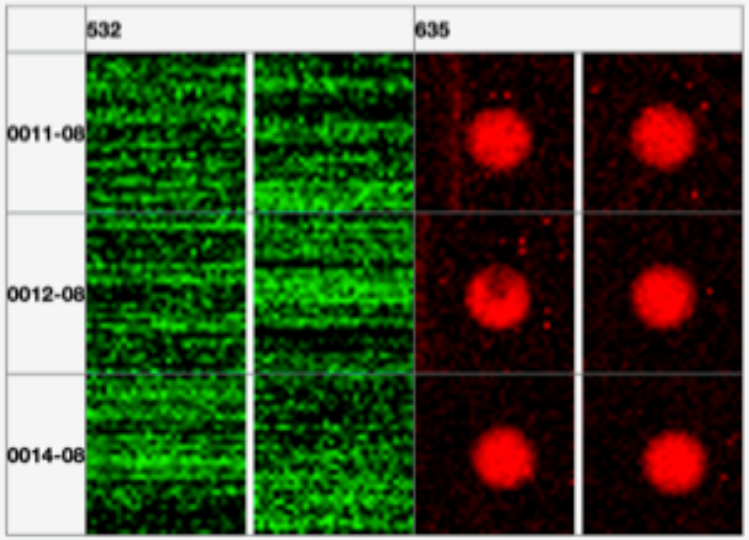


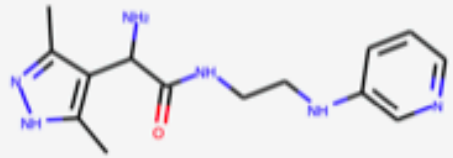
How will you determine a threshold Z-score?

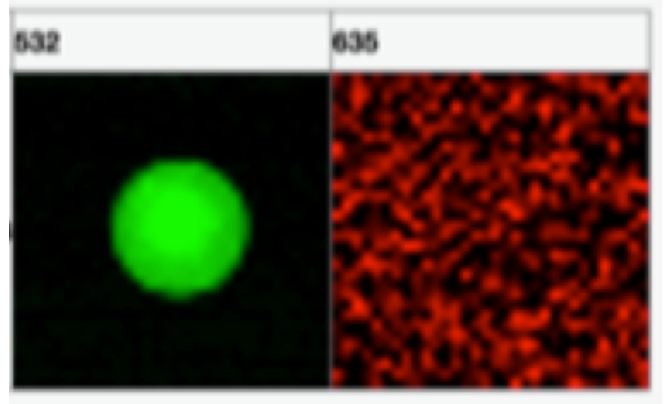


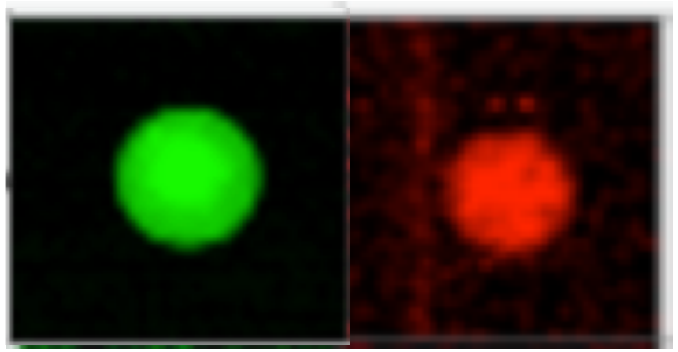
How will you validate hits?

	ID	Robust Z	SMILES	Validated
49592	13:KI0001...	51.03151	C[C@H](C...	-1
42089	11:KI0001...	45.09263	CC1=C(C(...	example
6782	02:KI0001...	39.91118	CCNC(=O...	-1
29108	08:KI0001...	39.59436	C1C(C2=...	-1
44736	12:KI0001...	33.03555	C1CN(C2...	-1
29660	08:KI0001...	31.94118	CC1=NC2...	-1
11360	03:KI0001...	26.13059	C1CN(CC...	-1









For today…

- You should have a list of confirmed hits when you leave!

For M1D7…

- Outline Future works ideas for Data summary
- Revise methods section for protocols completed on M1D1– M1D3
  - Include confirmation digest procedures

Notes on Implications & Future works…

- Implications should relate to the problems / goals in your introduction
- Be mindful of overreaching the data
- Future works should be the ‘next steps’ needed to further the research

Implications and Future Work: potential topics [edit]

- ***Topic***: What is the positive hit rate? Is this consistent with similar research?
- ***Topic***: Do your hits share any common chemical structures?
  - If no, provide a putative explanation. If yes, how can you further test if this structure is important in binding?
- ***Topic***: How can you use your TDP43-RRM12 binders to further research focused on this protein?
- ***Topic***: How might this method be improved?
- ***Topic***: How might this assay be used in the clinic? in industry?

Notes on Methods revision…

- Minor intermediate steps used in an experimental protocol should not be distinct sub-sections
- Do not include information about tubes or water
  - Water is the assumed diluent unless otherwise noted
- Include the final concentration (not amount of a stock)
  - Rather than “added to a final concentration of 1mM” just write “1 mM added”
- Provide the actual times for incubations
- Comments apply to all text
  - Only specifically addressed first example to be corrected in homework feedback

Making progress on the Data summary!

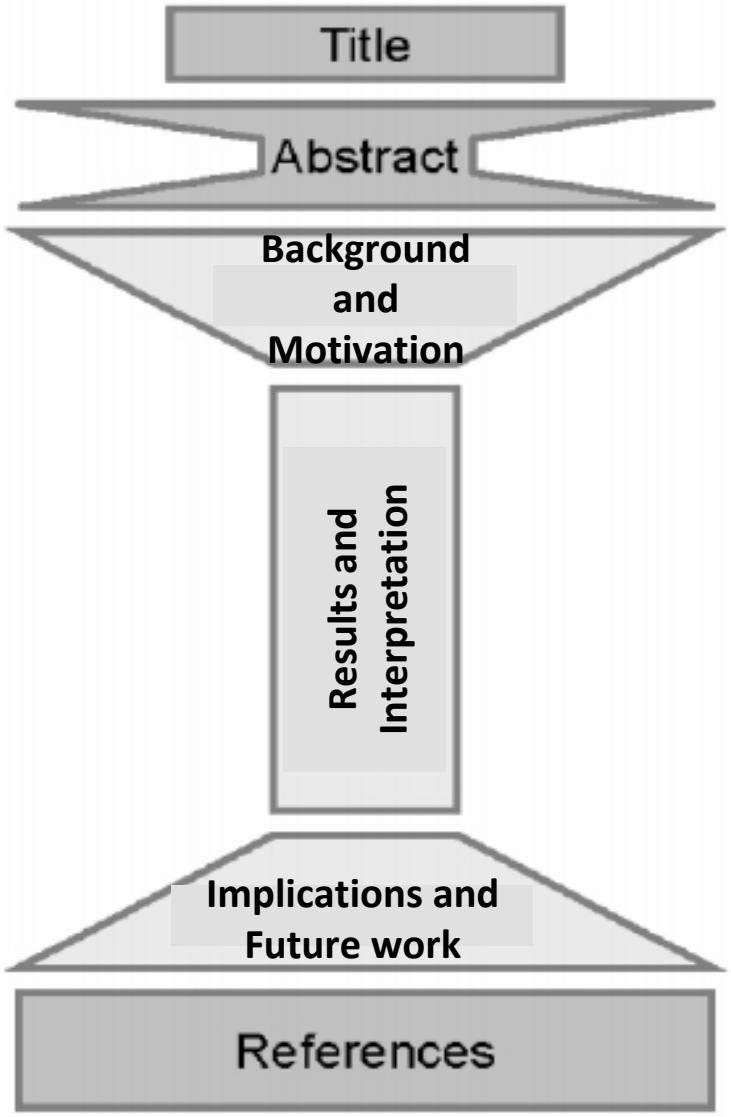
Title: take-home message
Abstract: Paragraph, NOT in bullet points!

In bullet points:
Background and Motivation (include citations)

Results and Interpretation

Implications and Future work (include citations)

References (see wiki for format suggestions)



Results and Interpretation: potential topics and figures [edit]

Figures and topics are listed below according to the three major phases of your experiment. Within each phase, you should look for sub-groupings of interest, rather than treat each piece of data in isolation. In other words, try to both interpret and communicate outcomes holistically.

Keep in mind that you described the detailed methods in a separate homework assignment and it does not need to be included in this report. Therefore, figure captions and/or supporting text should include only the most relevant aspects of the methods, such as the names of the important reagents, experimental techniques, or assays.

Protein purification [edit]

- **Schematic**: Experimental design.
  - Do not include minor technical details that are not necessary to understand your experimental conditions.
- ***Topic***: TDP43-RRM12 purification.
- **Figure**: Image of polyacrylamide gel.
- **Figure**: Graph or table displaying cell protein concentration.

Ligand screening [edit]

- **Schematic**: Experimental approach.
  - Do not include minor technical details that are not necessary to understand the goal of the experiment.
- ***Topic***: Identification of positive hits.
- **Figure**: Image from representative scan.
- **Figure**: Graph or table comparing z-scores and p-values.
- ***Topic***: Chemical structure comparison.
- **Figure**: Images of positive hits.

Data summary structure / logistics

- To be submitted as a powerpoint file!
  - Change page settings such that 'slides' are portrait and 8.5" x 11"
  - Upload to Stellar (draft due Mar 8 at 10pm, revision due Mar 22 at 10pm)

- Each figure will be included as a separate Data slide
  - Image should be at the top of the slide with title and caption
  - Results / Interpretation text should be included on same slide
  - Though figures are separated into Data slides, the story should be cohesive between figures!