M2D1: Examine SMM data collected using TDP43 protein

1. Prelab

2. Walk through SMM analysis

3. Examine chemical structure of hits

4. Discuss journal article

SMM workflow
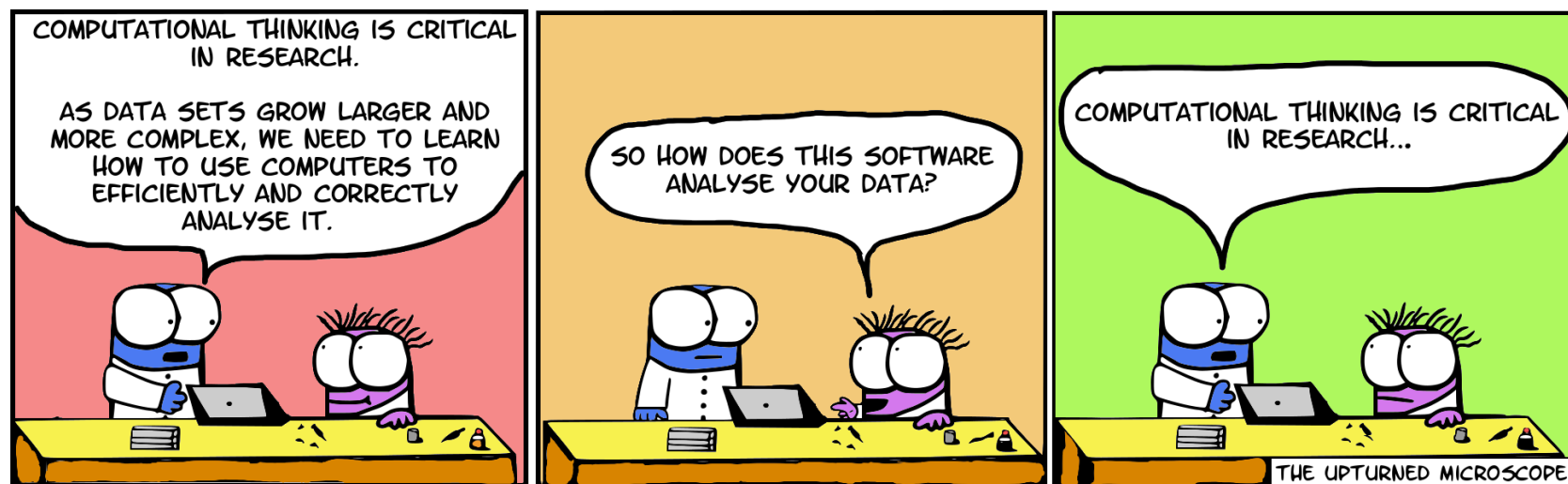

B. Leifer, S. Doyle
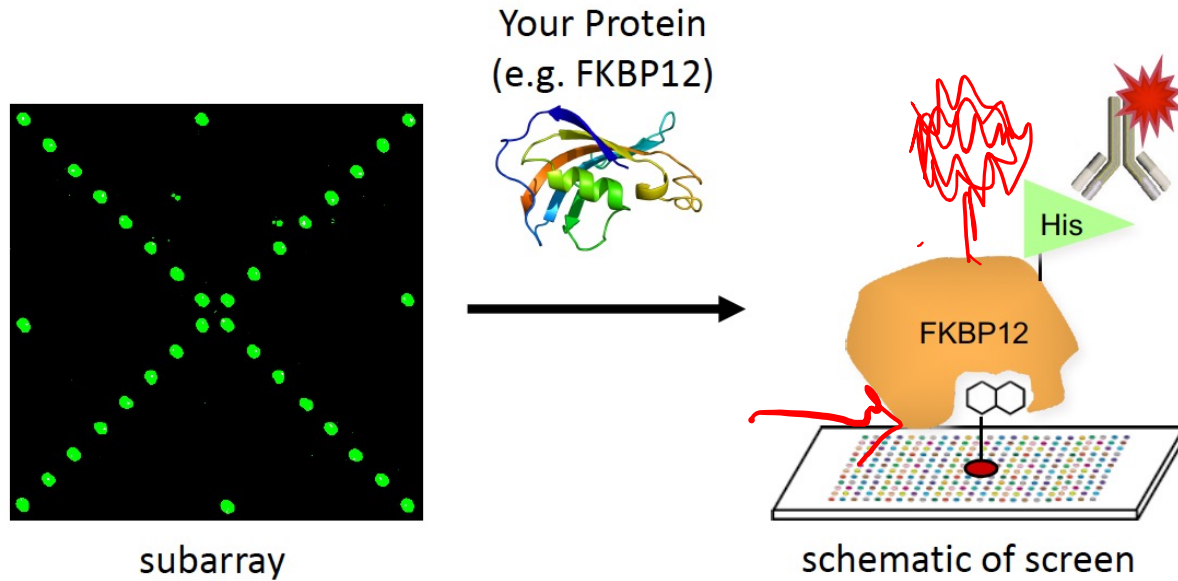
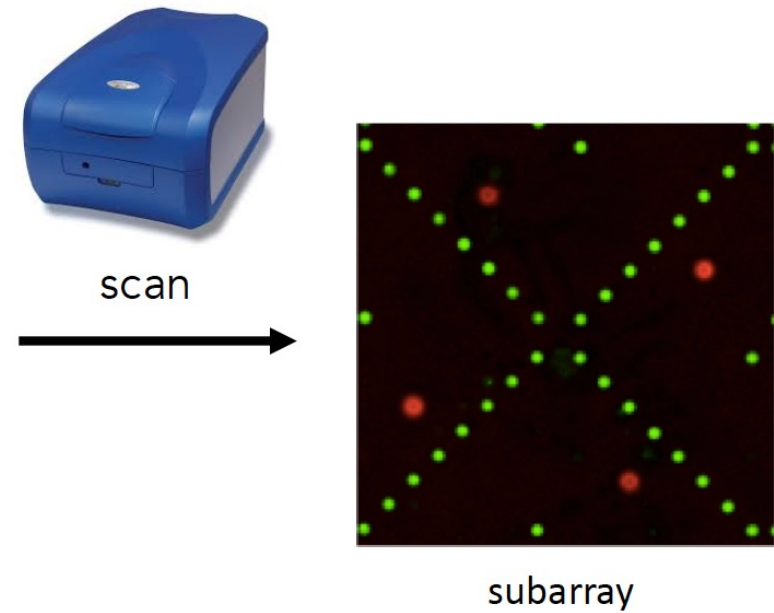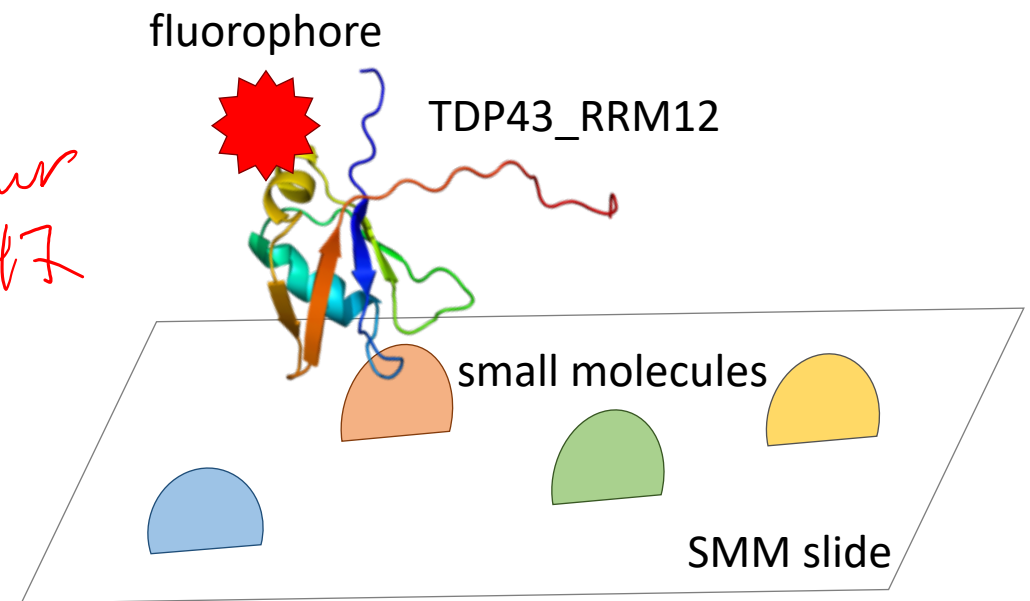# How do we use the SMM to screen for ligands that bind our protein of interest?

- Create a recombinant protein of the TDP43 RNA binding domains (TDP43_RM12)
    - Label this protein with a Alexa647 fluorophore ~~—SNAP tag— AlexaFlour 647~~

- Incubate the SMM slide with our purified and labeled TDP-43_RRM12
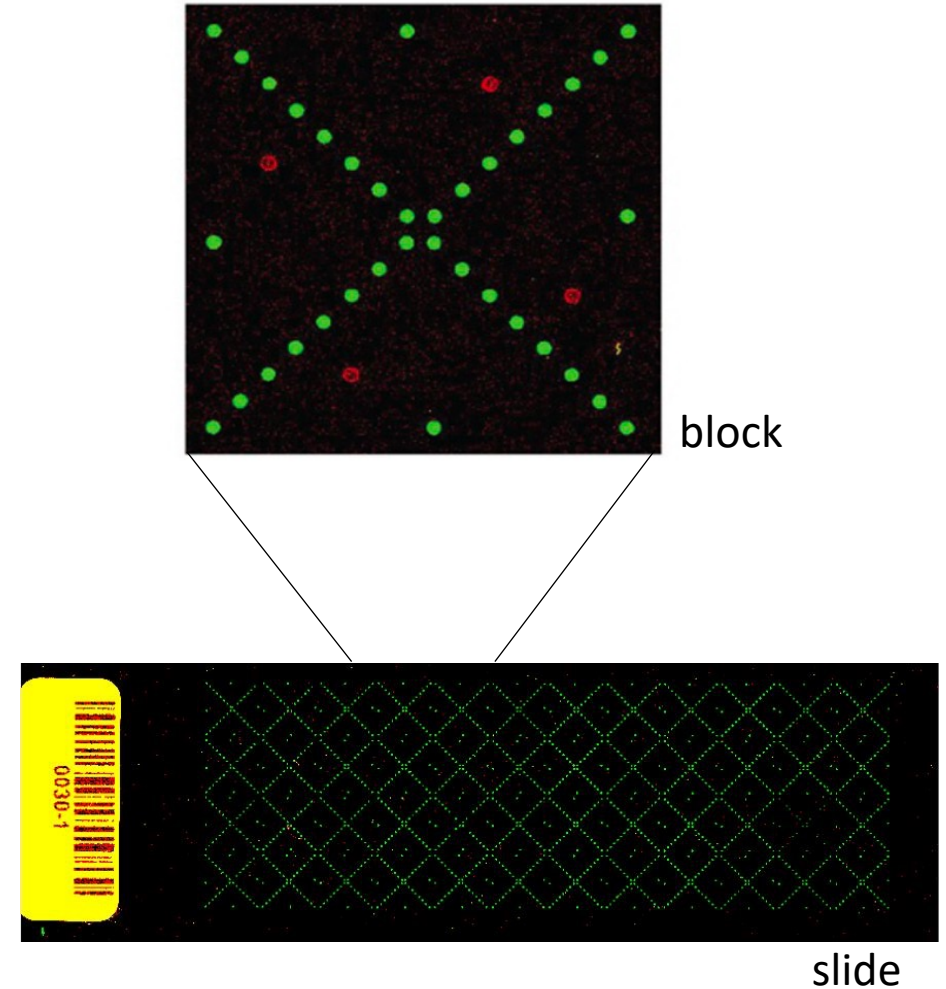
- Wash away unbound protein

- Store for scanning

Workflow for SMM data analysis

1. Align spots using fluorescence on 532 nm channel (sentinel spots)

   ⬇

2. Quantify fluorescence on 635 nm channel

   ⬇

   — Red
   — Protein

3. Identify 'hits' with improbably high fluorescence

   ⬇

4. Complete 'by eye' analysis of putative hits to manually remove false positives

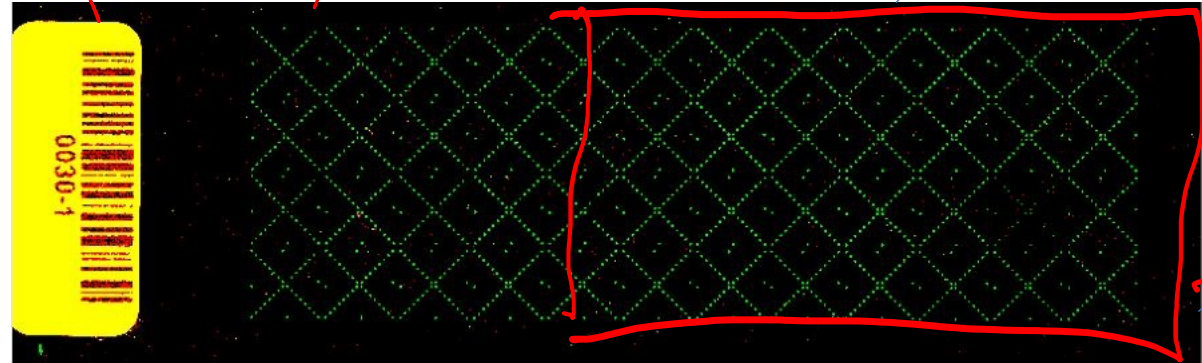block

slide

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
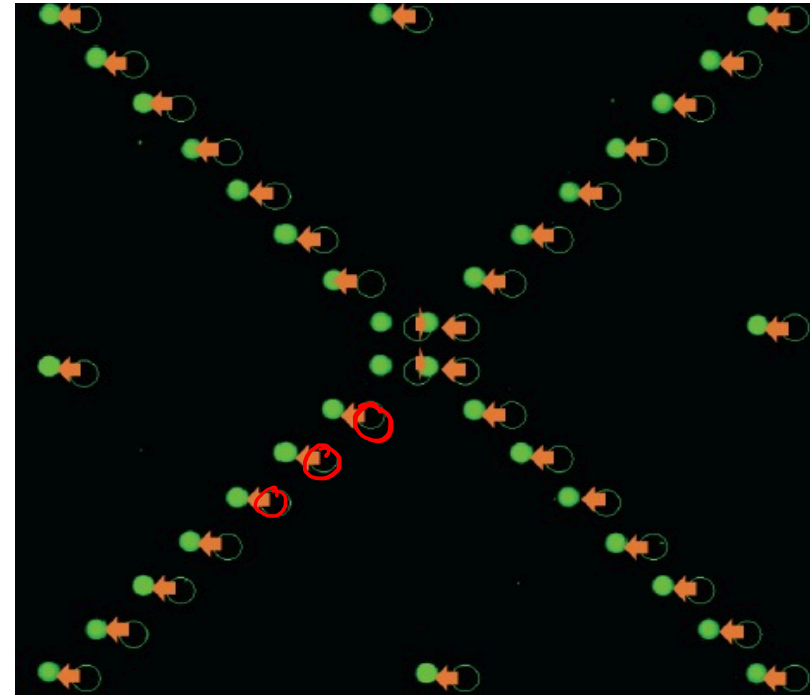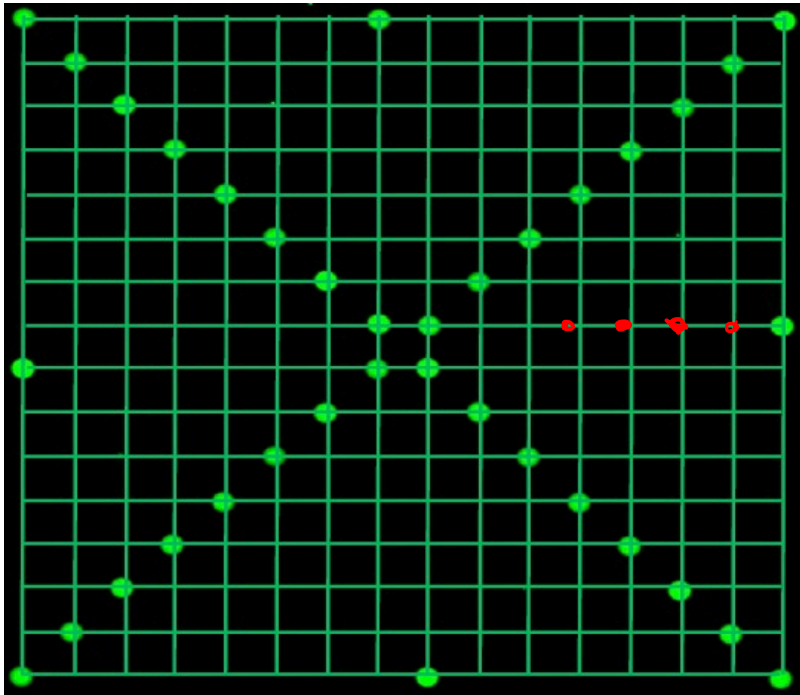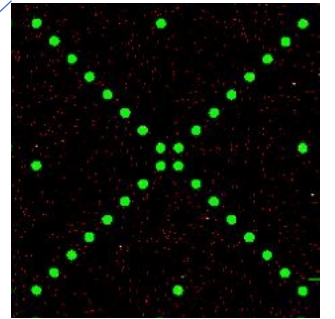
How do you identify hits from the SMM data?

First, consider bias that exists in the data set
- Across all slides
- Within each block
- Within each slide

printer/pins (handwritten)

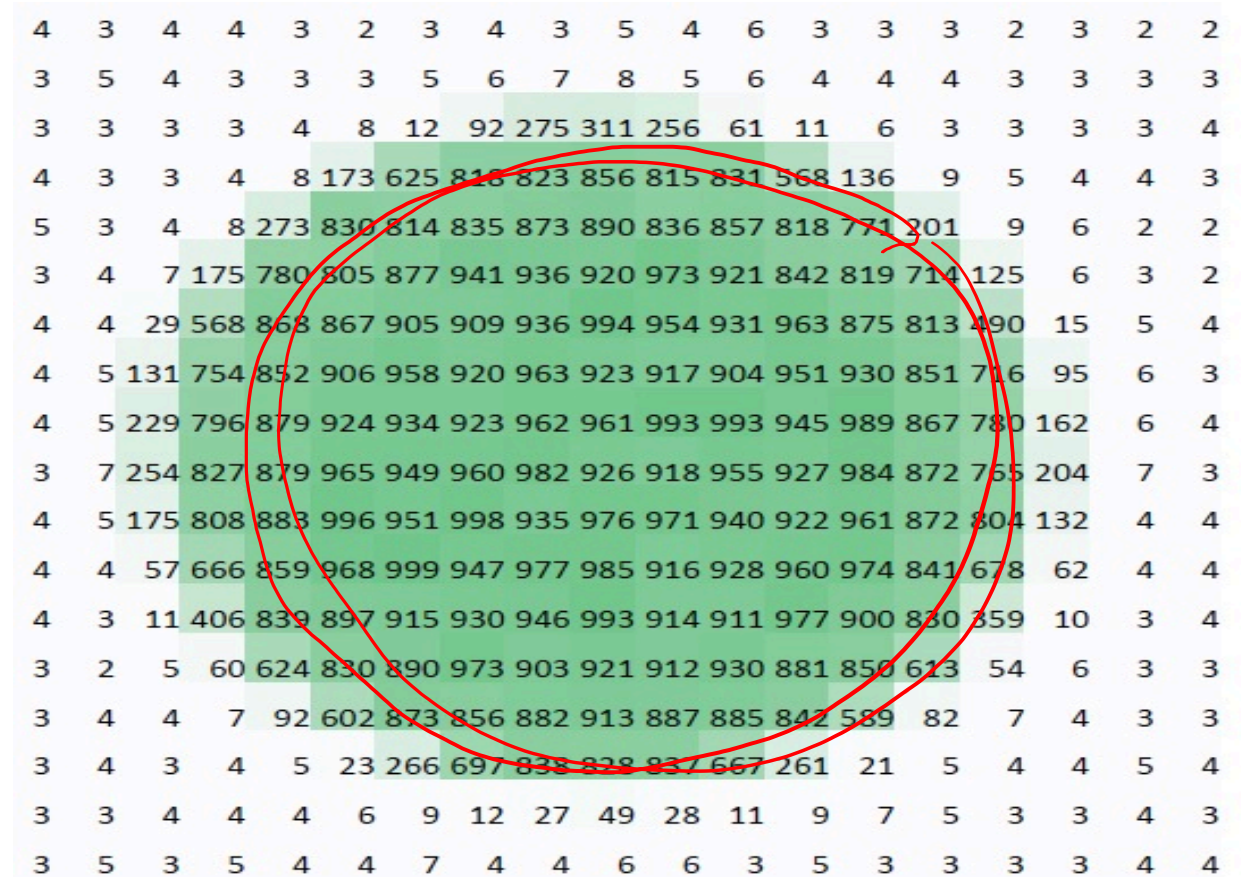*extra background* (handwritten)

Block

Slide

Then, identify hits with significantly higher fluorescence over background

Lastly, manually confirm hits to eliminate false positives

Spots are represented by an array of numerical values

- Each pixel is represented by a number that indicates intensity of the signal

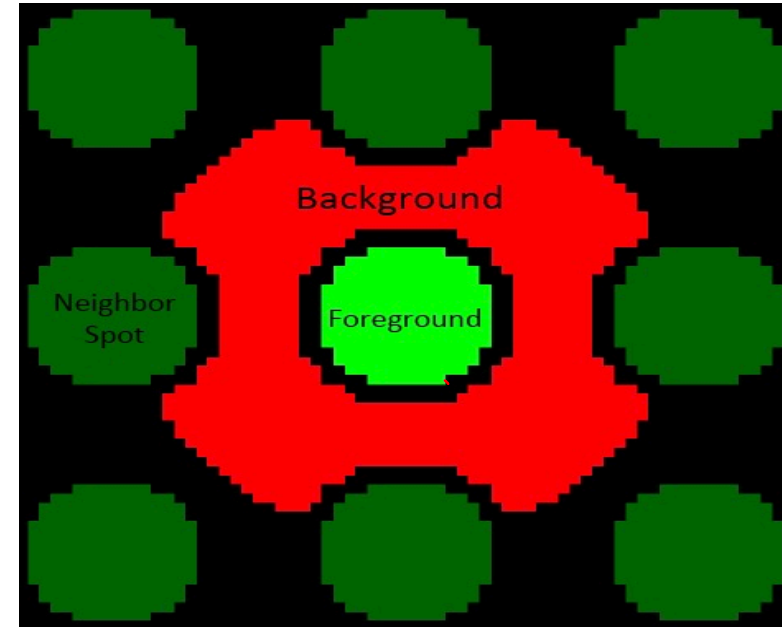- Computational analysis used to define 'hits'

Fluorescence is quantified to identify hits

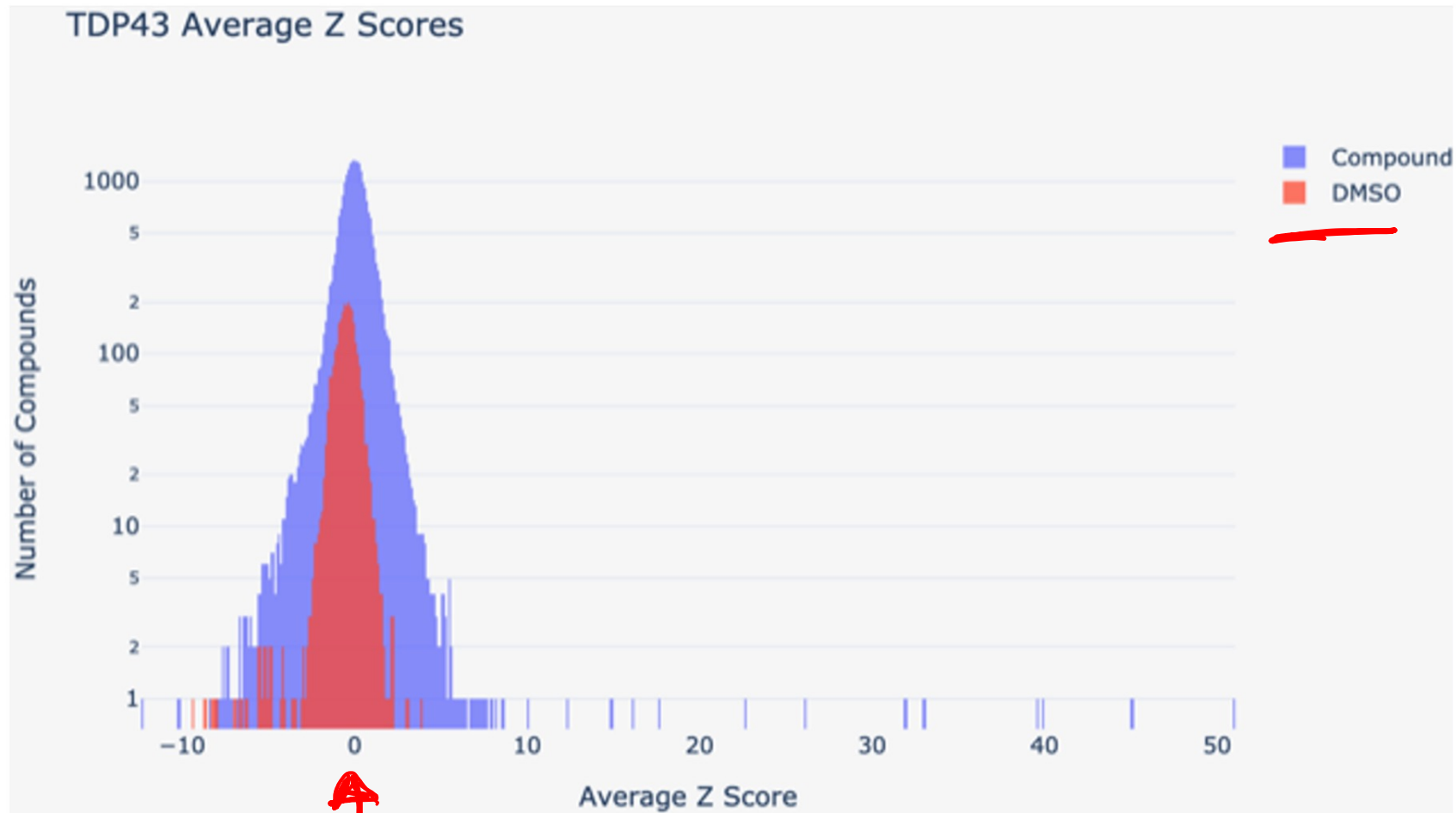- Foreground:

  SM printed

- Background:

  noise

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

SNR is then used to calculate the **robust z score**
- How different is the foreground signal from the background?
- Able to plot the distribution of the z scores to give an overview of whole data set

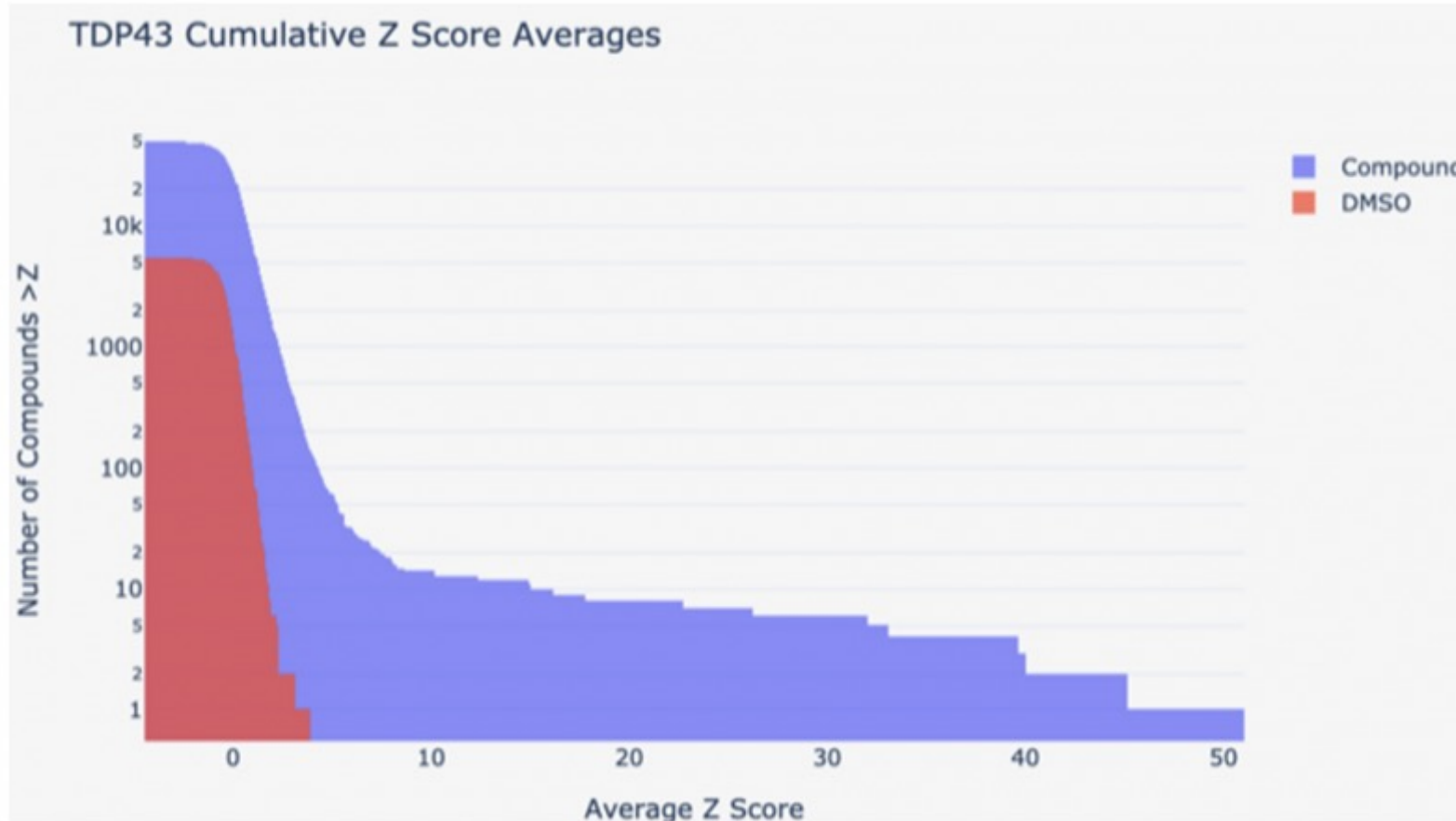# Average Z-score calculated for all compounds

TDP43 Average Z Scores

Replicate spots averaged

Each count = unique compound

How many compounds have a particular z score?

Height at x (average z score) = number of compounds (y) with that z-score or higher

- Useful for setting a threshold to exclude likely non-binders

How do you validate hits manually?

How will you identify common structures?

Overview of Mod1 experiments

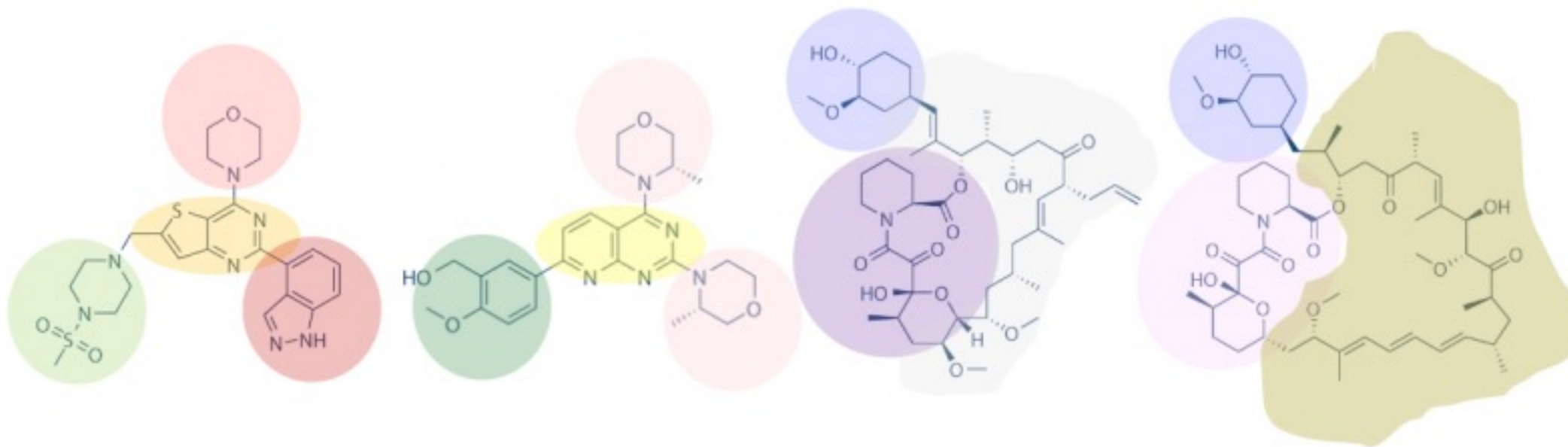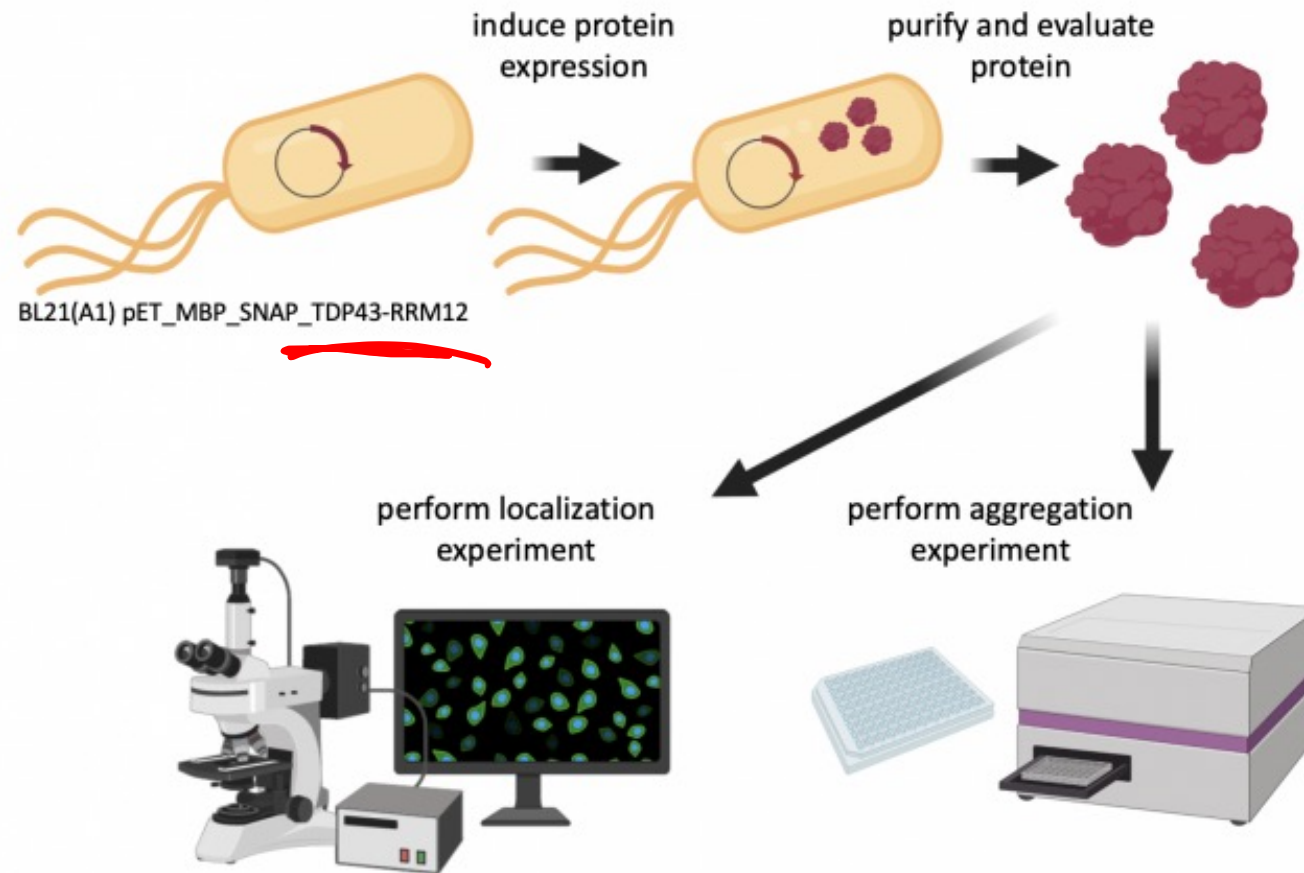**Research goal: Use functional assays to characterize ligands identified as binders to TDP43 from SMM technology**

For Today

- Work through SMM analysis procedure
- Evaluate chemical structures of identified hits
- Discuss reading of scientific papers with Noreen
  - Group 1: Teal, White, Gray, Purple, Pink
  - Group 2: Blue, Green, Yellow, Orange, Red

For M1D3

- Begin thinking about Background and Motivation for Data Summary
  - Submit document answering questions on the Homework section of wiki
  - Due Thursday, Feb. 10 at 1:05pm on Stellar

- Visit Comm Lab by M1D5
  - Can visit to discuss an assignment from any class, a personal statement for an internship application, etc…