

# 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

## Office Hours:

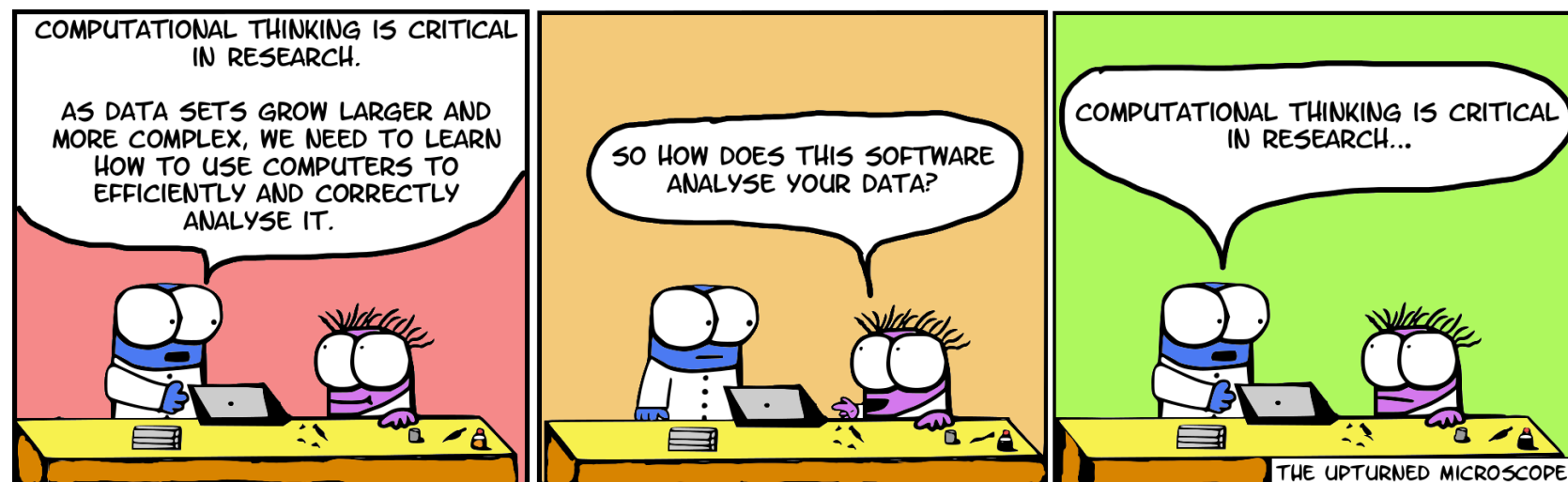
Monday: 1-2pm @ 16-319 and Zoom [Becky](#)  
3-5pm @ 16-317 and Zoom [Noreen](#)

Tuesday: 10-11am @ 1-390 [Becky & Jamie](#)

Thursday: 10-11am @ 1-390 [Noreen & Jamie](#)

\*After lecture by request

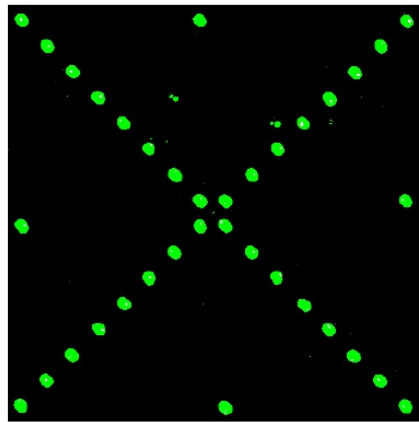
\*Also available by appointment



# SMM workflow

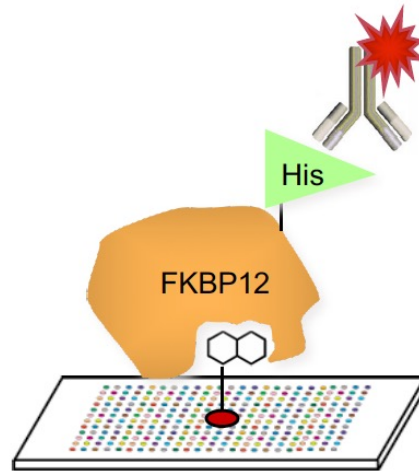
## SMM Screen

## Data Acquisition



subarray

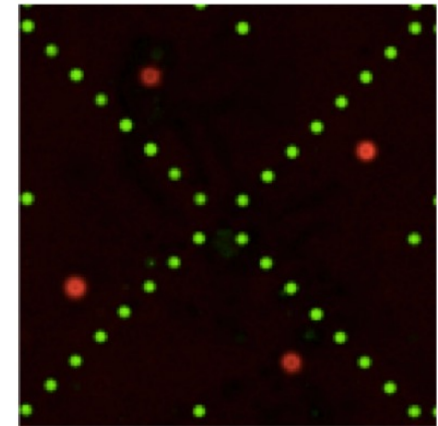
Your Protein  
(e.g. FKBP12)



schematic of screen



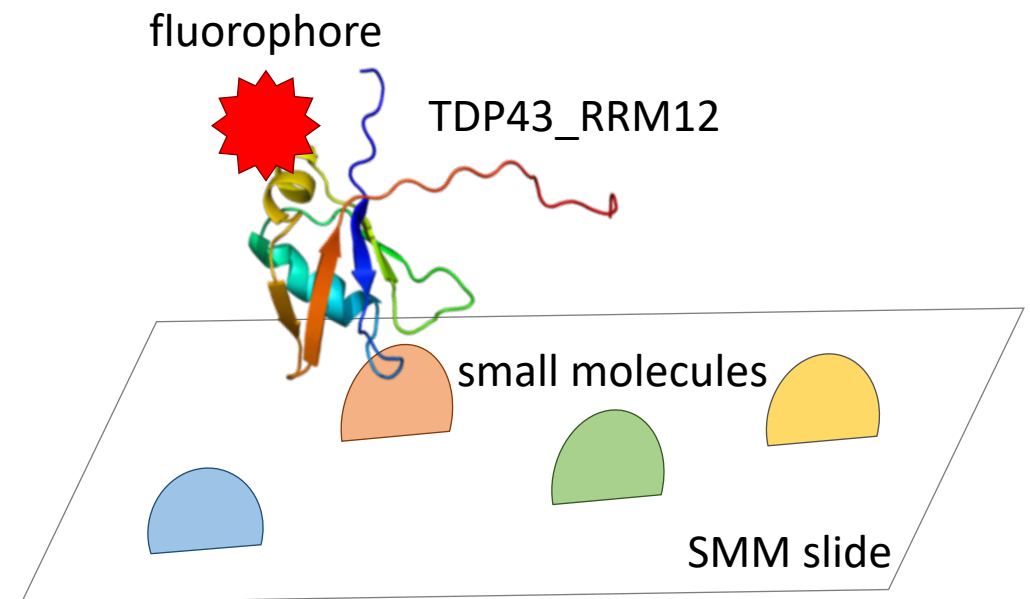
scan



subarray

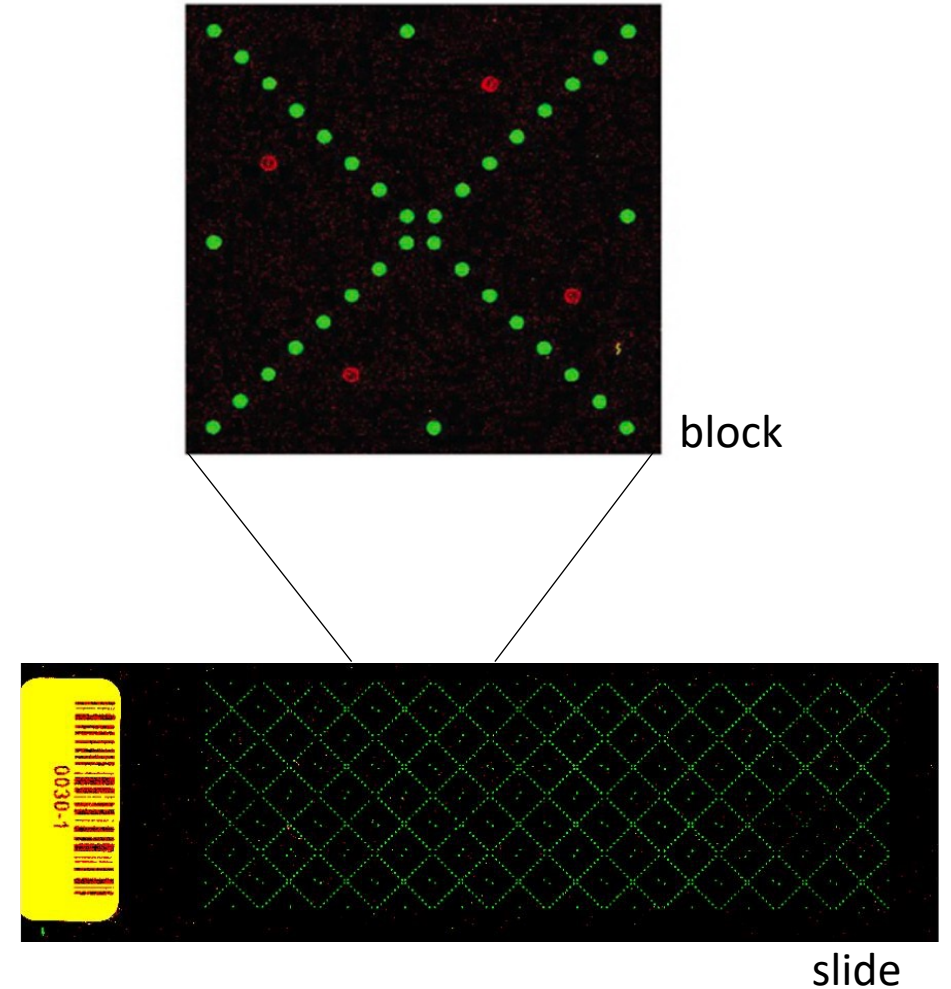
# How do we use the SMM to screen for ligands that bind our protein of interest?

- Create a recombinant protein of the TDP43 RNA binding domains (TDP43\_RRM12)
  - Label this protein with a Alexa647 fluorophore
- Incubate the SMM slide with our purified and labeled TDP-43\_RRM12
- Wash away unbound protein
- Store for scanning



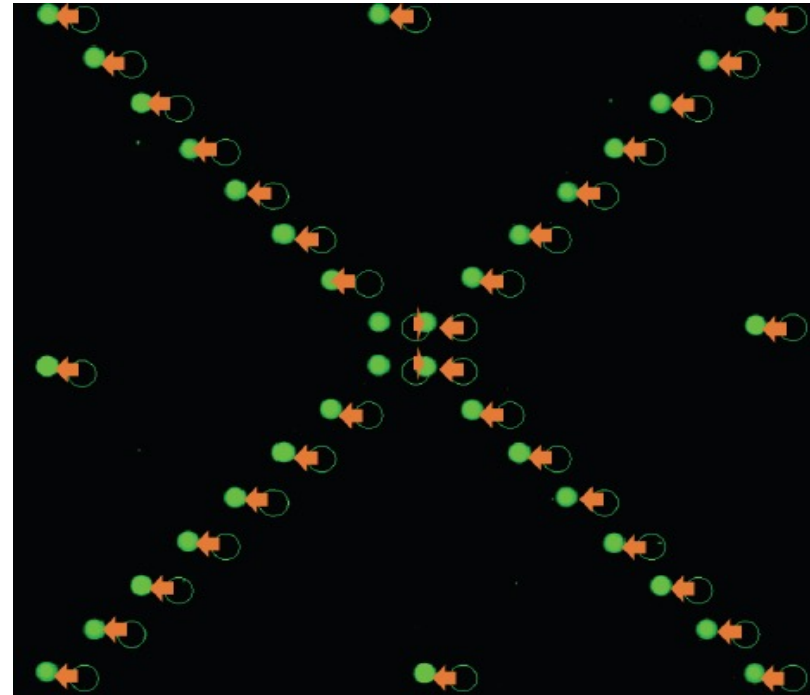
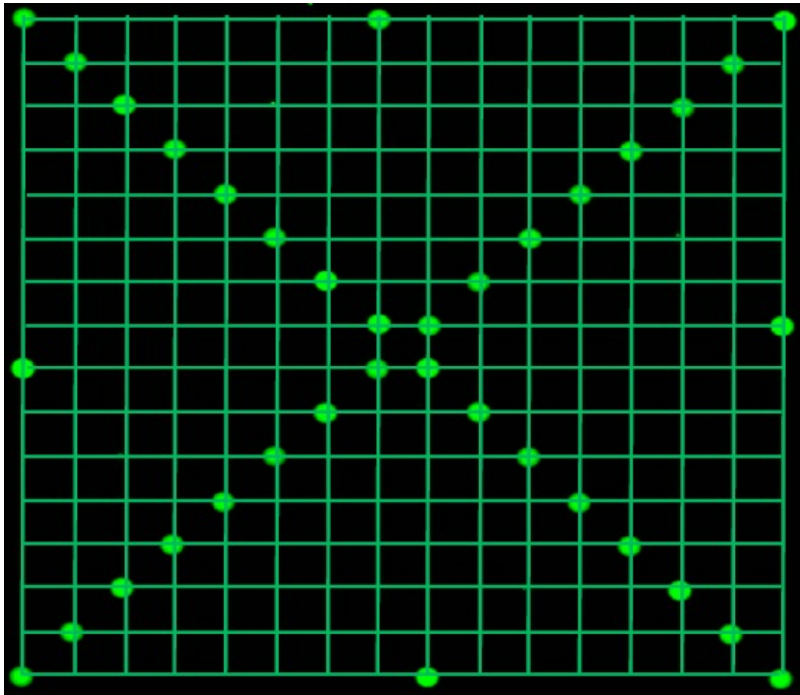
# Workflow for SMM data analysis

1. Align spots using fluorescence on 532 nm channel (sentinel spots)  
↓
2. Quantify fluorescence on 635 nm channel  
↓
3. Identify 'hits' with improbably high fluorescence  
↓
4. Complete 'by eye' analysis of putative hits to manually remove false positives



# Align SMM using sentinel spots

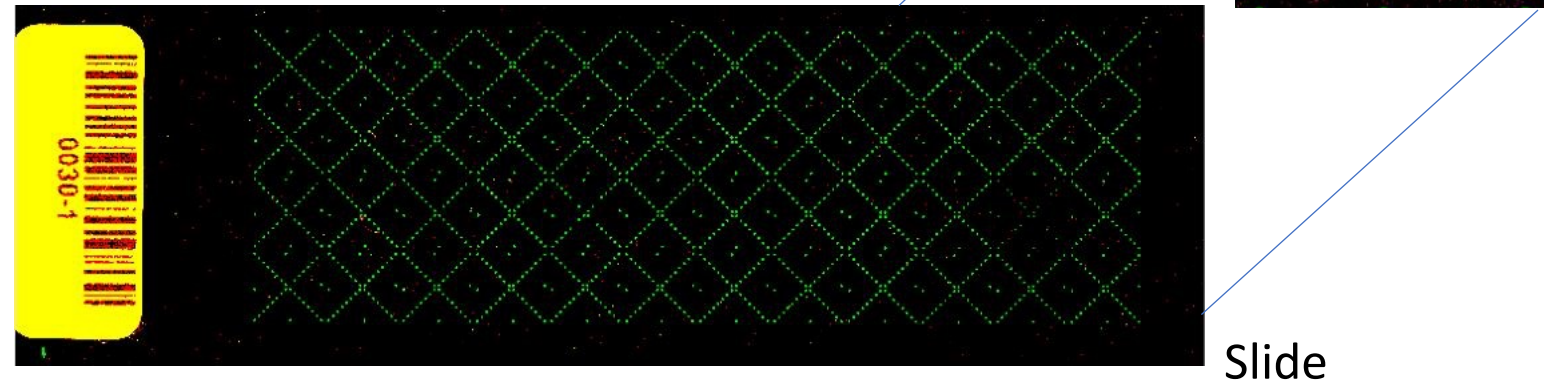
- Slides are printed in block patterns (16 rows x 16 columns)
- Each ligand spot is identifiable via intersecting lines between sentinels



# 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



Then, identify hits with significantly higher fluorescence over background

Lastly, manually confirm hits to eliminate false positives

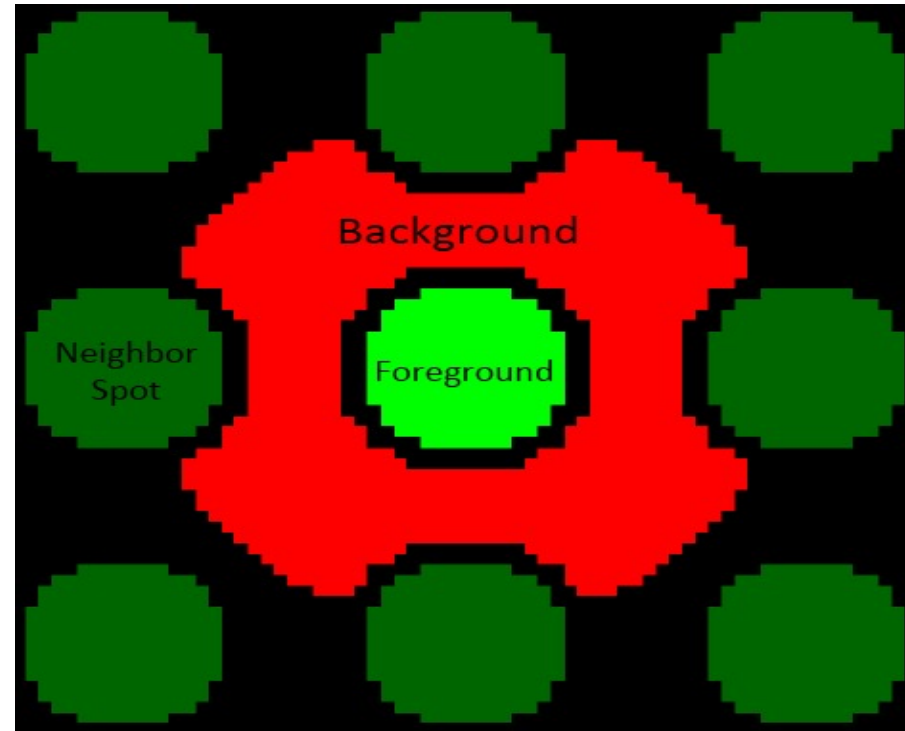
# Spots are represented by an array of numerical values

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

4	3	4	4	3	2	3	4	3	5	4	6	3	3	3	2	3	2	2
3	5	4	3	3	3	5	6	7	8	5	6	4	4	4	3	3	3	3
3	3	3	3	4	8	12	92	275	311	256	61	11	6	3	3	3	3	4
4	3	3	4	8	173	625	818	823	856	815	831	568	136	9	5	4	4	3
5	3	4	8	273	830	814	835	873	890	836	857	818	771	201	9	6	2	2
3	4	7	175	780	805	877	941	936	920	973	921	842	819	714	125	6	3	2
4	4	29	568	868	867	905	909	936	994	954	931	963	875	813	490	15	5	4
4	5	131	754	852	906	958	920	963	923	917	904	951	930	851	716	95	6	3
4	5	229	796	879	924	934	923	962	961	993	993	945	989	867	780	162	6	4
3	7	254	827	879	965	949	960	982	926	918	955	927	984	872	765	204	7	3
4	5	175	808	883	996	951	998	935	976	971	940	922	961	872	804	132	4	4
4	4	57	666	859	968	999	947	977	985	916	928	960	974	841	678	62	4	4
4	3	11	406	839	897	915	930	946	993	914	911	977	900	830	359	10	3	4
3	2	5	60	624	830	890	973	903	921	912	930	881	850	613	54	6	3	3
3	4	4	7	92	602	873	856	882	913	887	885	842	589	82	7	4	3	3
3	4	3	4	5	23	266	697	838	828	837	667	261	21	5	4	4	5	4
3	3	4	4	4	6	9	12	27	49	28	11	9	7	5	3	3	4	3
3	5	3	5	4	4	7	4	4	6	6	3	5	3	3	3	3	4	4

# Fluorescence is quantified to identify hits

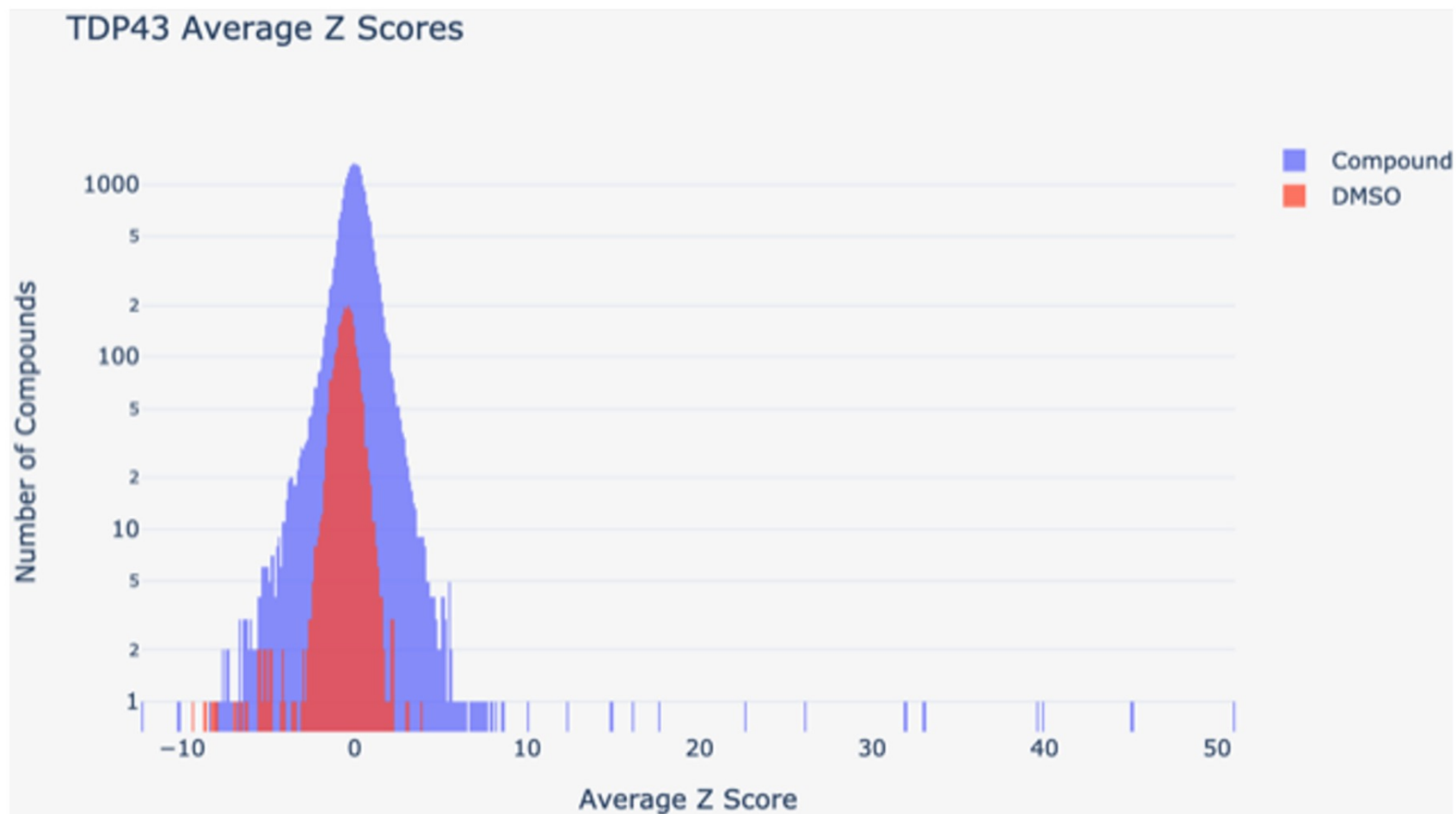
- Foreground:
- Background:



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



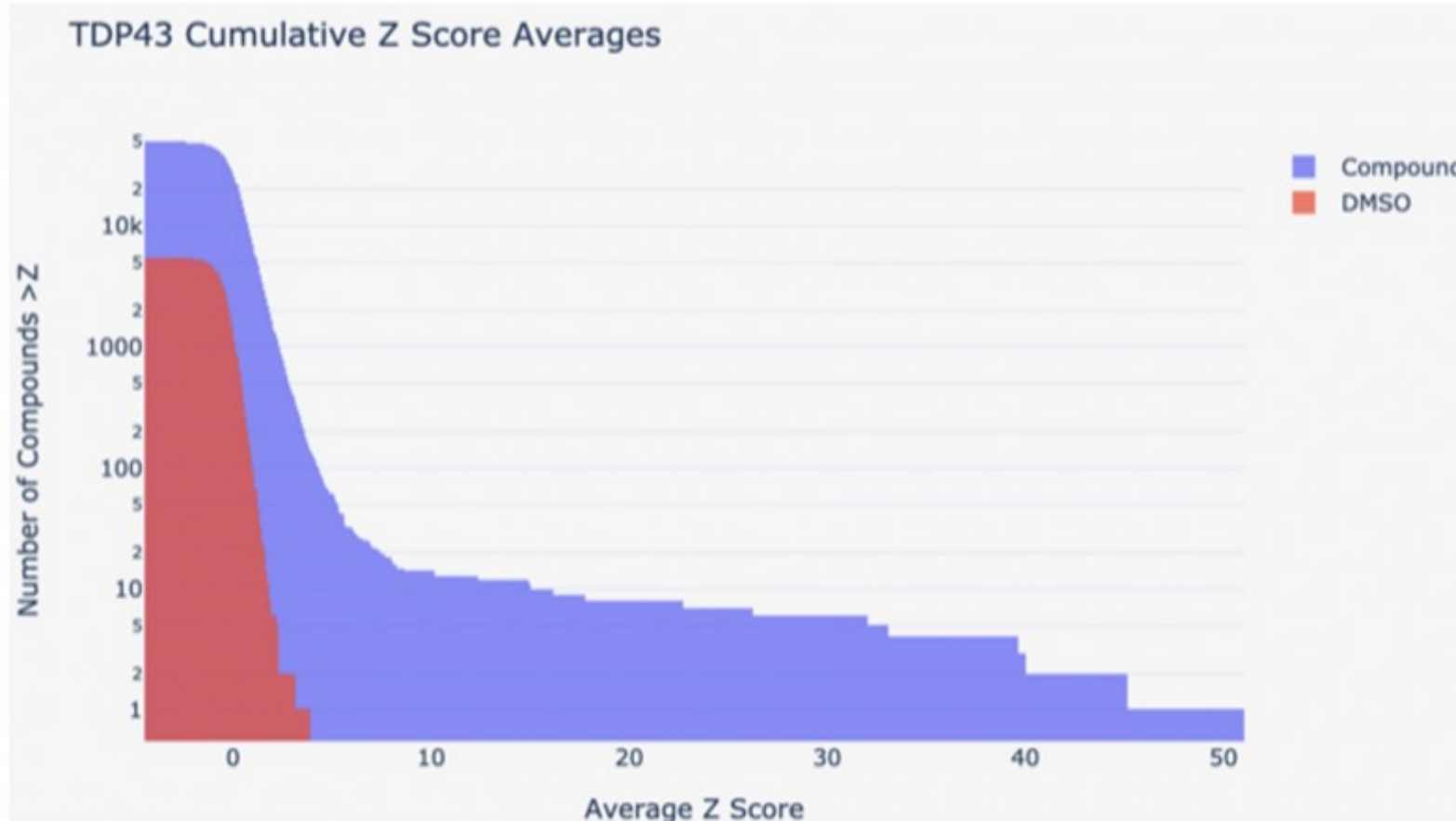
# Average Z-score calculated for all compounds



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?

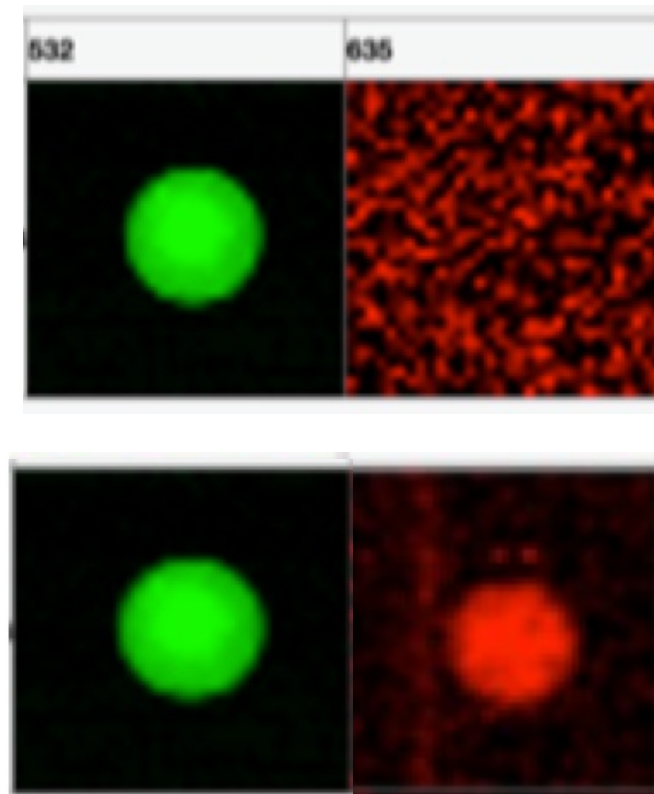
ID	Robust Z	SMILES	Validated
49592	51.03151	C[C@H](C...	-1
42089	45.09263	CC1=C(C...	example
6782	39.91118	CCNC(=O...	-1
29108	39.59436	C1C(C2=...	-1
44736	33.03555	C1CN(C2...	-1
29660	31.94118	CC1=NC2...	-1
11360	26.13059	C1CN(CC...	-1

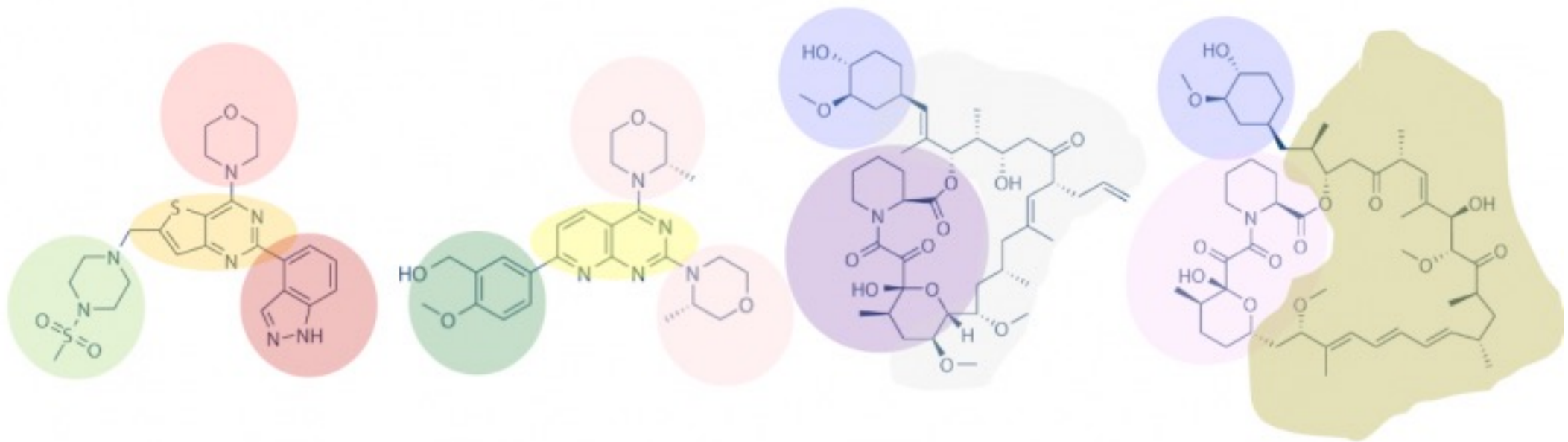
	532	635
0011-08		
0012-08		
0014-08		

  
Cc1c(C)nc(C)nc1C(=O)NCCCNc1cccnc1

Chemical structure of a hit molecule, featuring a thiazole ring substituted with methyl groups and a side chain containing a secondary amide, a primary amine, and a pyridine ring.

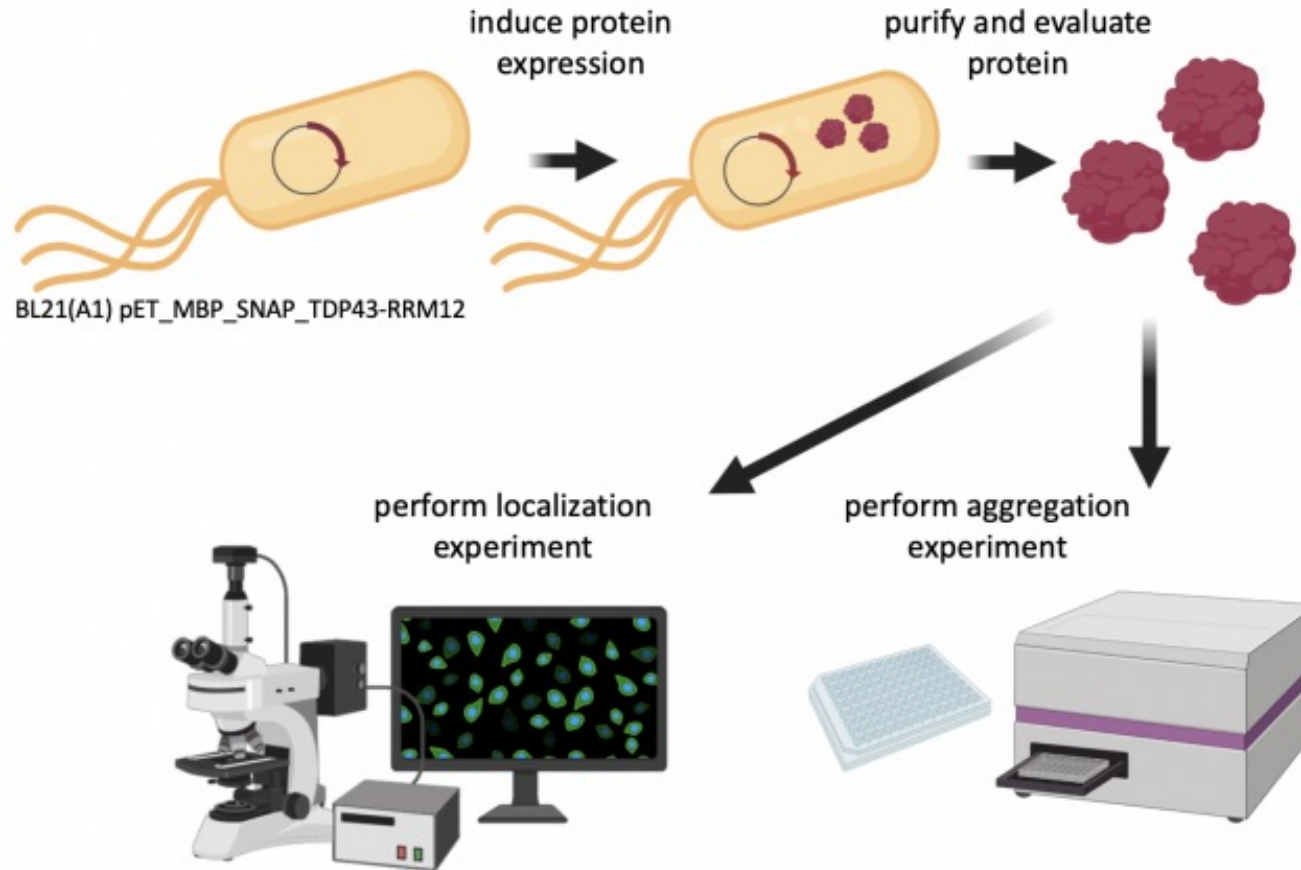


How will you identify common structures?



# 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...