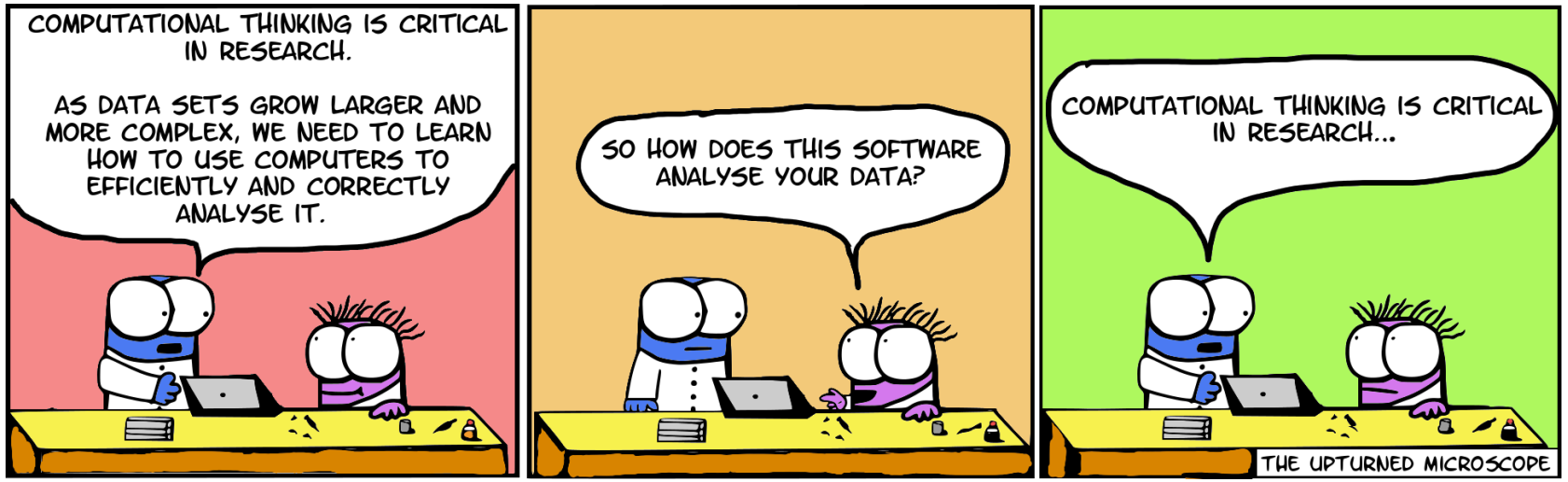
M2D4: Review small molecule microarray (SMM) technology and data analysis

1. Comm Lab
2. Quiz
3. Prelab
4. Walk through SMM



A note on JC Presentations

1) You will be recorded
2) This recording is for review between you and Noreen and will be deleted immediately after review
3) Presentations are due on the day YOU present, first come first serve with regards to choosing your order on the day of
4) Food will be provided

Journal Article Presentation Days

Wed (11/2)

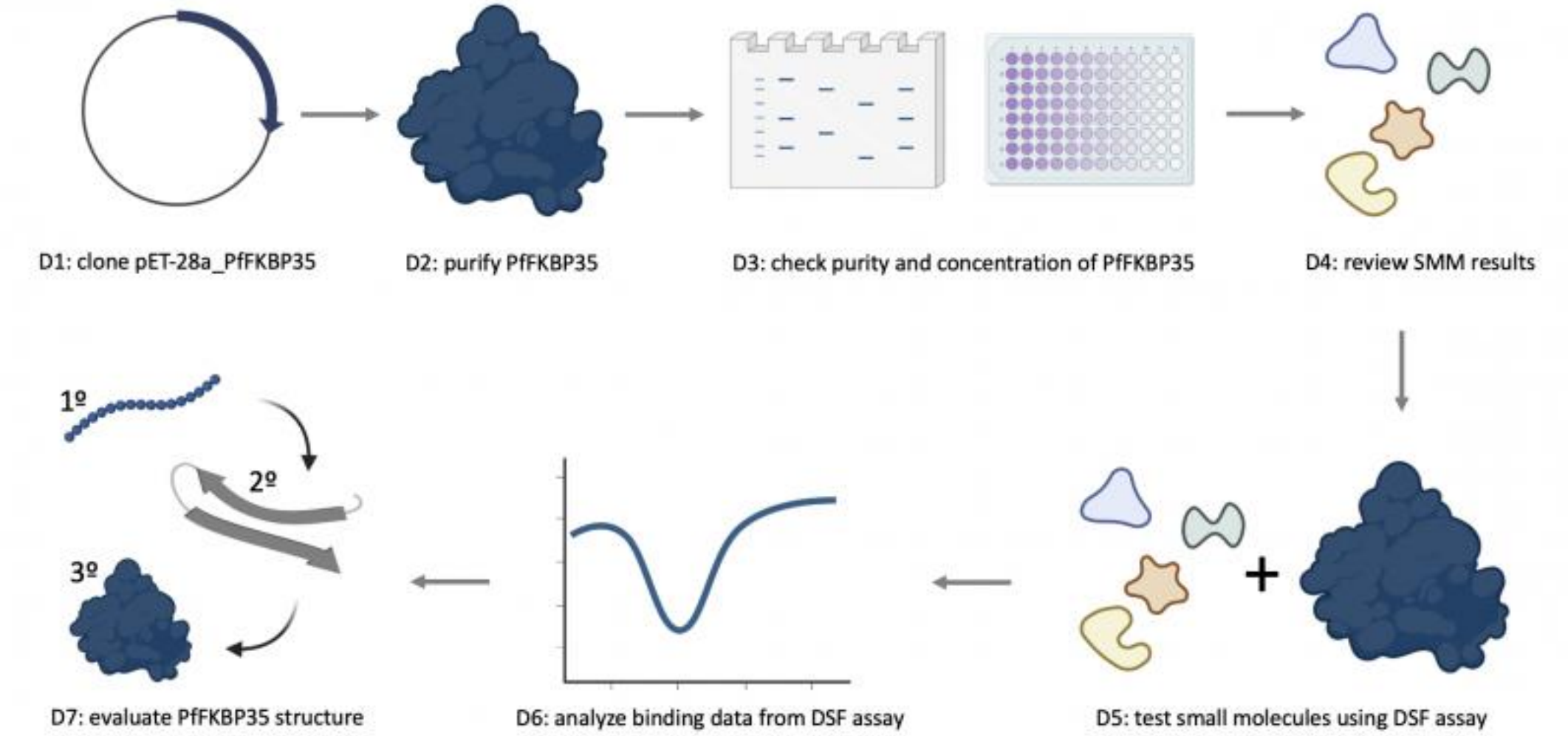
- Andrew
- Jessica
- Natalie
- Sarah

Fri (11/4)

- Calli
- Demi
- Franco
- Maddy

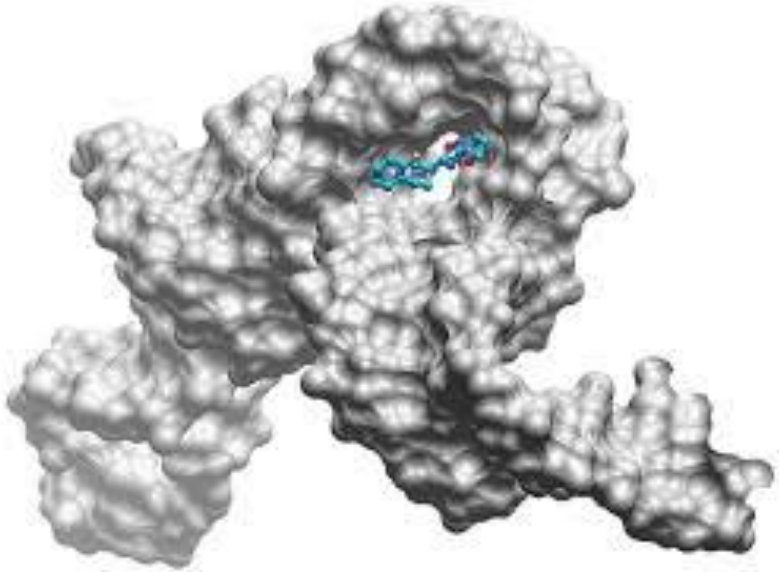
Overview of M2: drug discovery

Research goal: Test small molecules for binding to the *Plasmodium falciparum* FKBP35 protein using a functional assay.

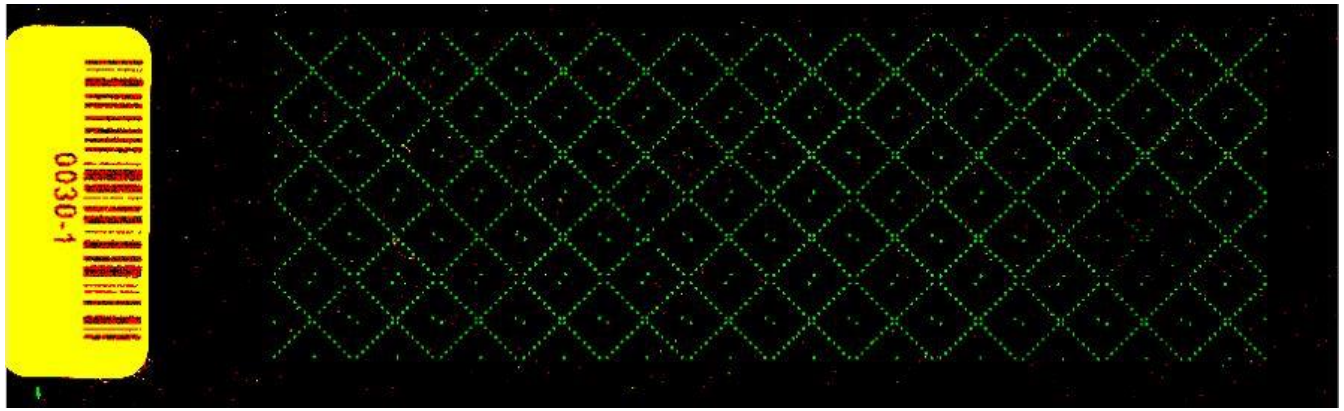


Why are we discussing this approach?

- High throughput assays are useful in screening potential therapeutic targets
  - Allows unbiased exploration of potential therapeutics
  - Allows examination of targets with limited information
- Small molecules
  - Mw < 500 Da
  - Natural or synthetic
  - Frequently comprised of Carbon/Nitrogen/Oxygen
- Our small molecule library is based on FK506 (known binder)
  - **Why limit ourselves to FK506?**



Small Molecule Microarray (SMM)



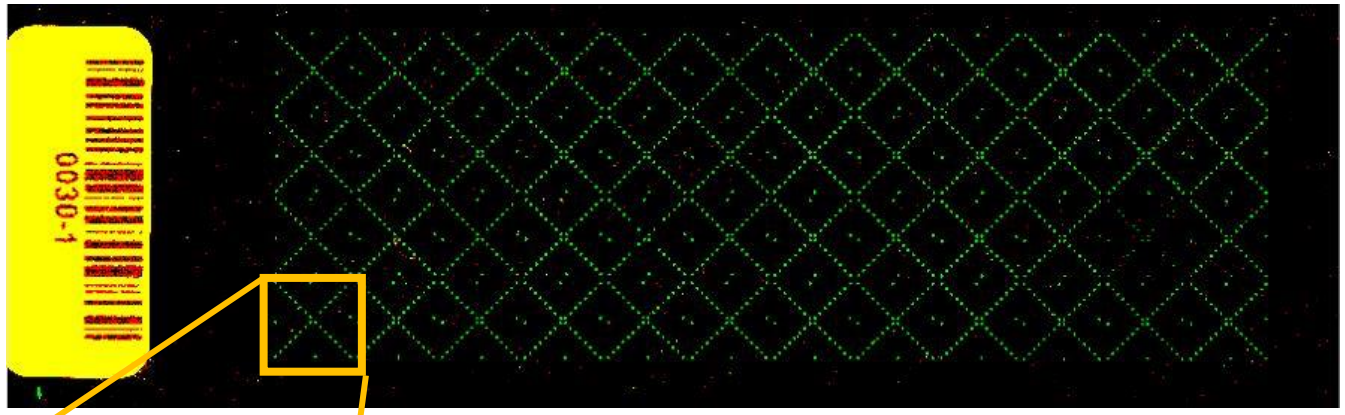
- Each slide contains ~12,000 spots
  - ~4,200 small molecules / ligands (in duplicate = ~8,400)
  - Fluorescein sentinel spots
  - DMSO negative control spots

Koehler Lab 2014 - Small-molecule Microarrays from Koehler Lab on Vimeo.

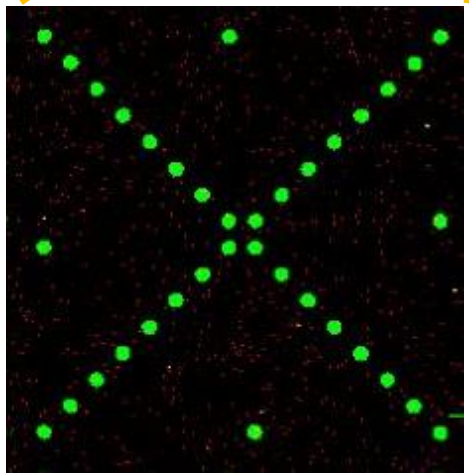
Guide to the SMM slide

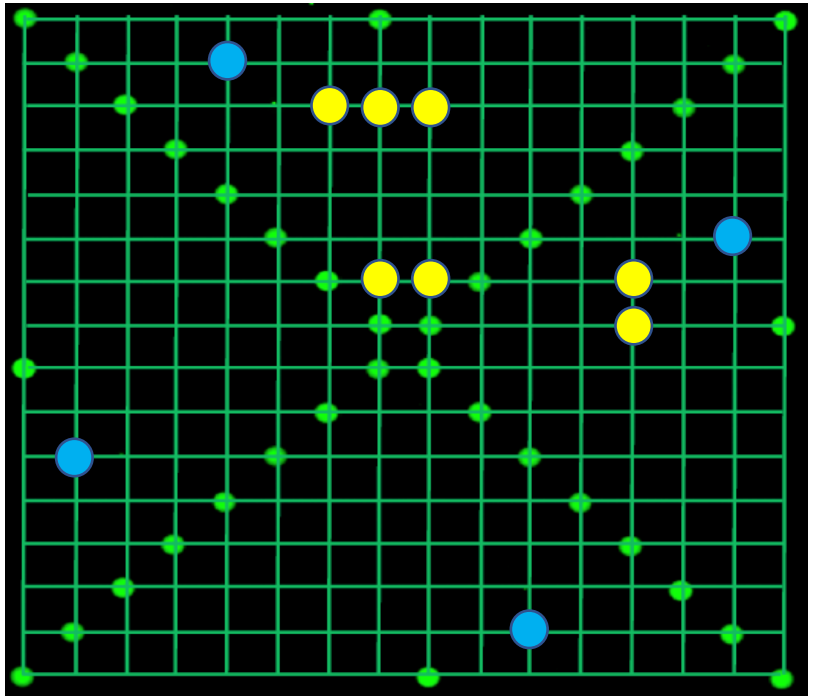
- Each slide has several blocks
- Each block has sentinel spots which are landmarks
- Rest of dots are small molecules and controls
- Can overlay a computational map to identify the location of each small molecule

Slide



Block





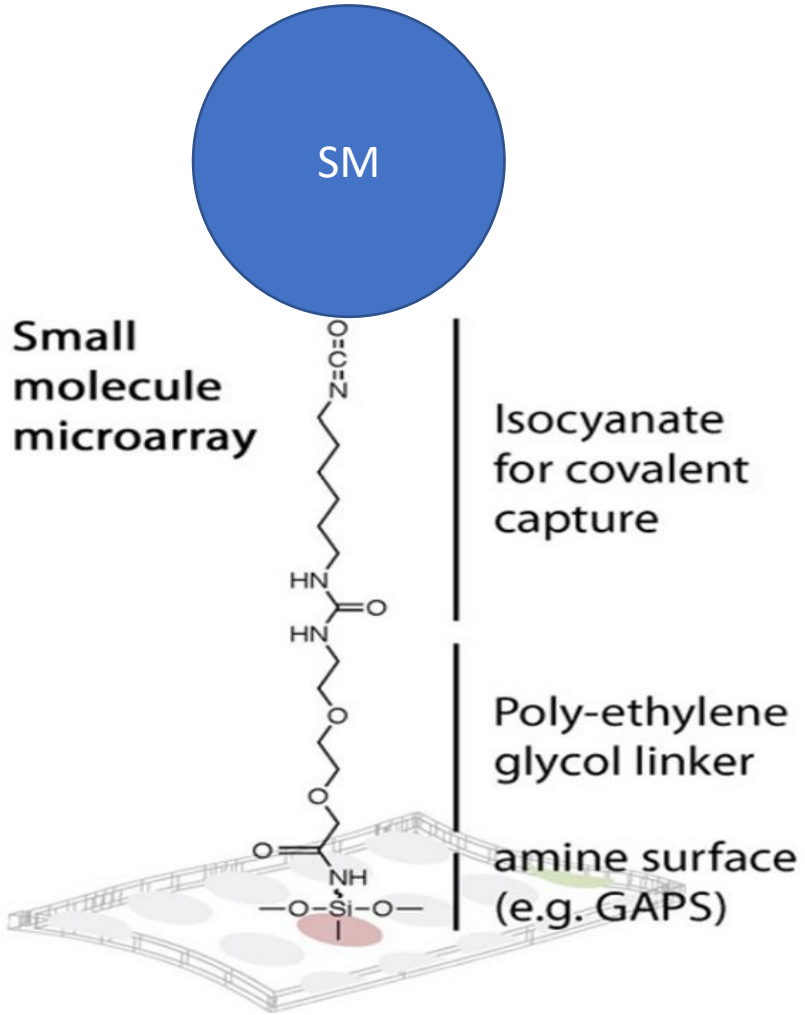
Green= sentinel spots (fluorescein dye)

Blue= DMSO

Yellow= SM

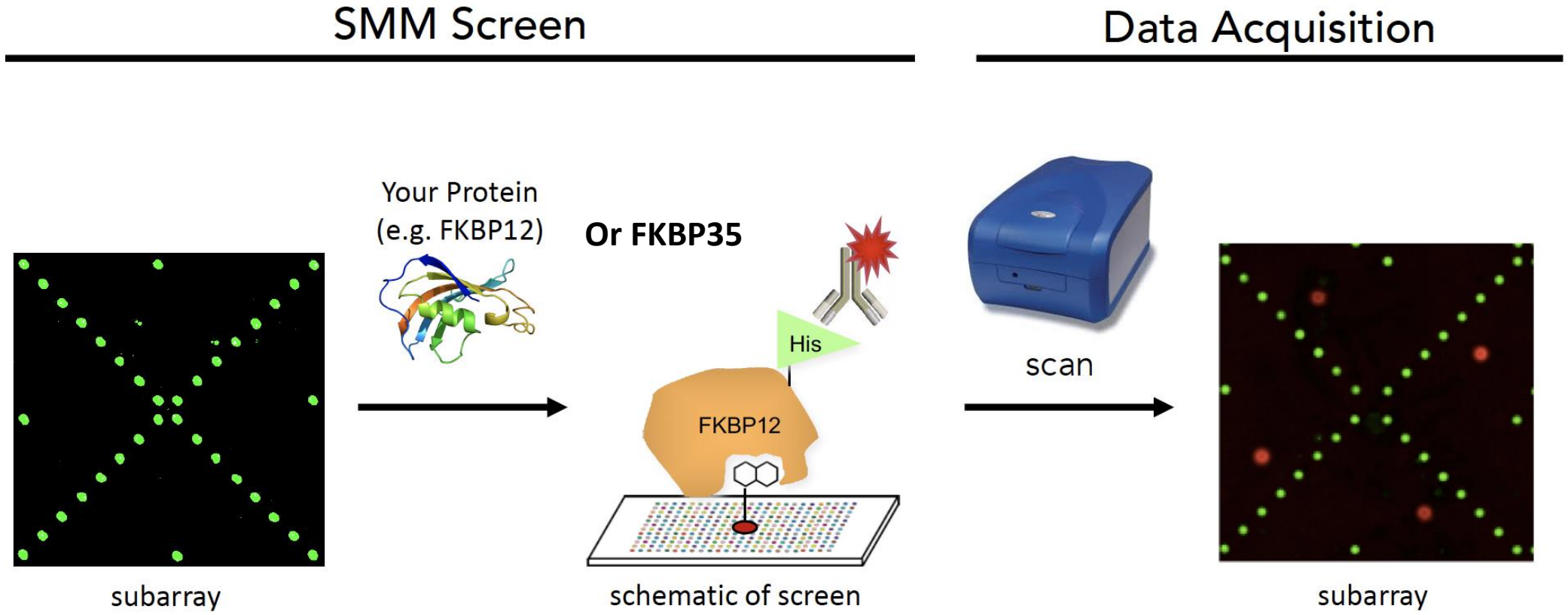
SMM slide preparation

- Gamma-aminopropylsilane (GAPS) coated slide with polyethylene glycol (PEG) spacer
  - Why might it be a good idea to include a spacer here?
- PEG coupled to 1,6-diisocyanatohexane to generate isocyanate-functionalized slide
- Isocyanate able to react with nucleophilic functional groups



Leifer et al. Methods in Enzymology. 2018.

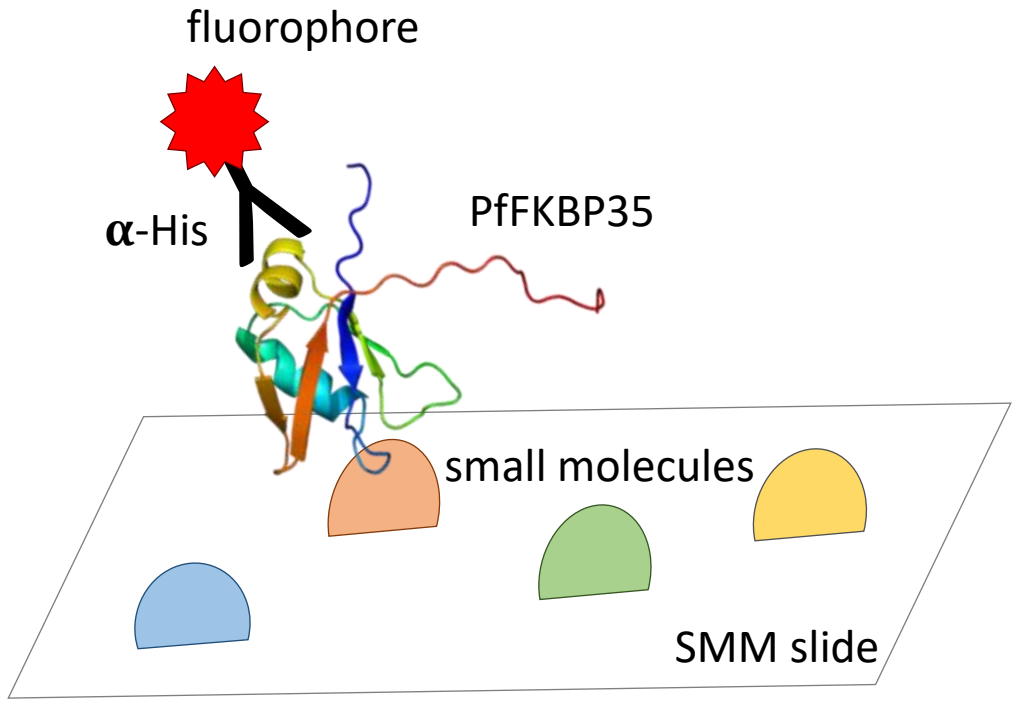
SMM workflow



B. Leifer, S. Doyle

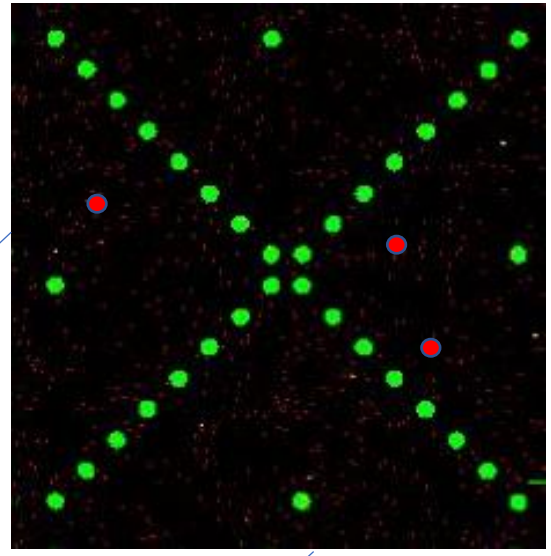
How would we screen for ligands that bind PfFKBP35?

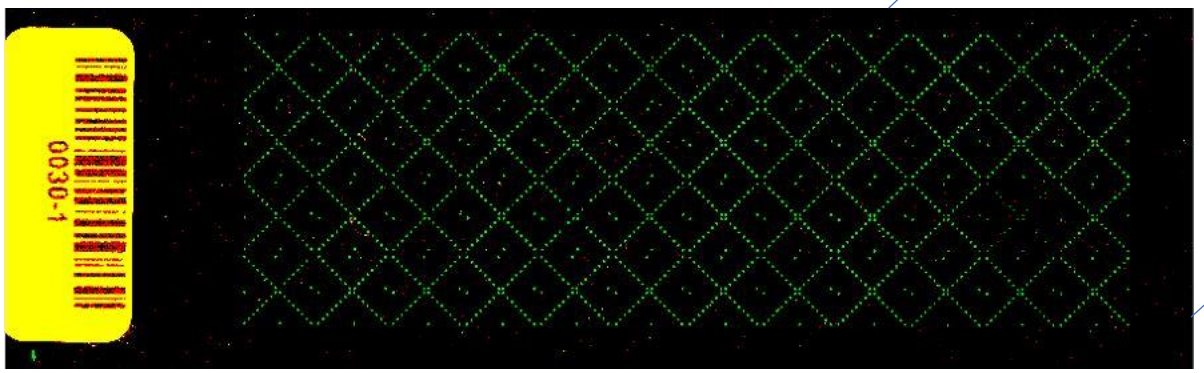
- Incubate the SMM slide with 3ml of our purified PfFKBP35
- Wash away unbound protein
- Incubate SMM slide with AlexaFlour 647 anti-His antibody
- Wash away excess antibody
- Store for scanning



What do putative binders look like on the SMM slide?

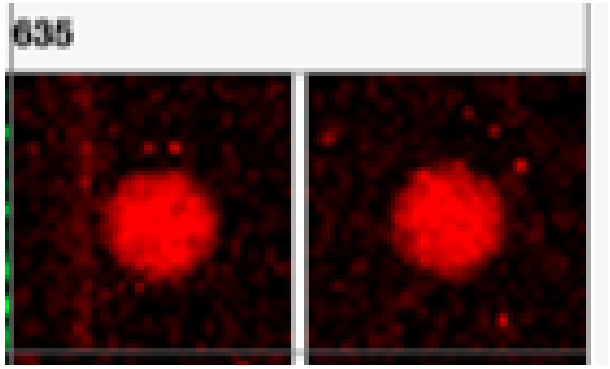
Block





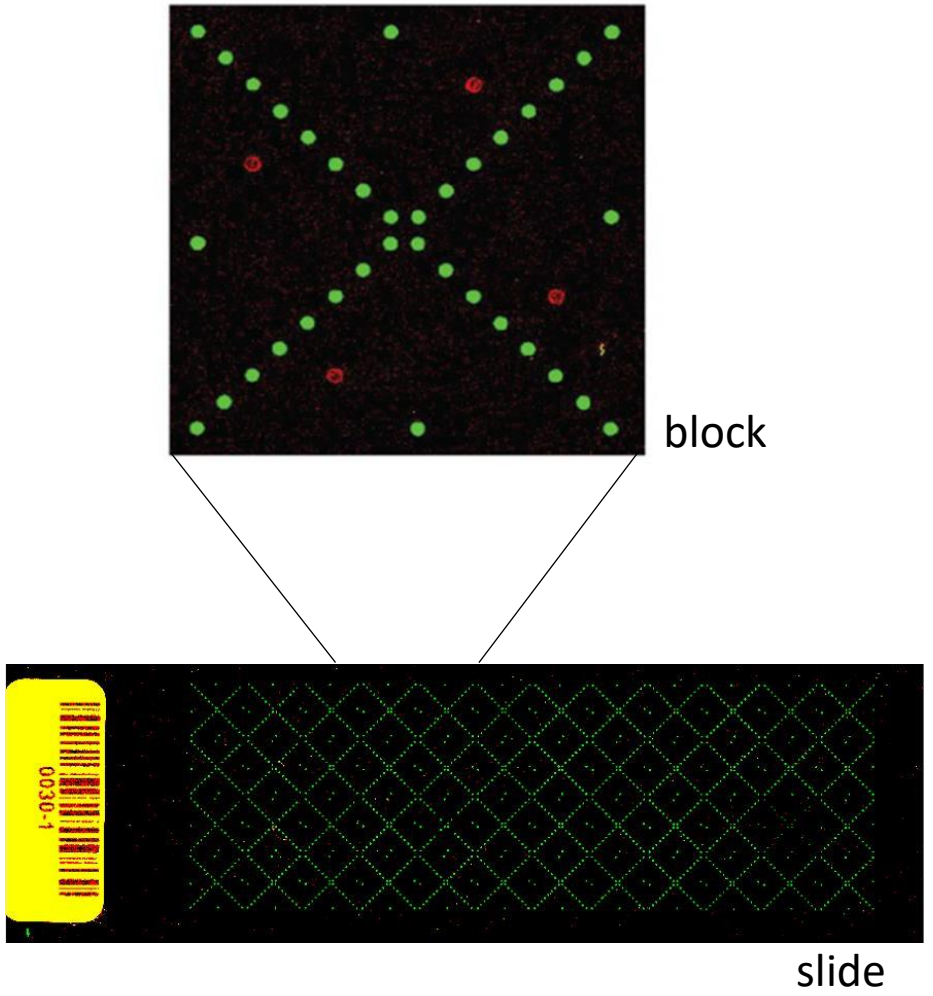
Slide

(RED)



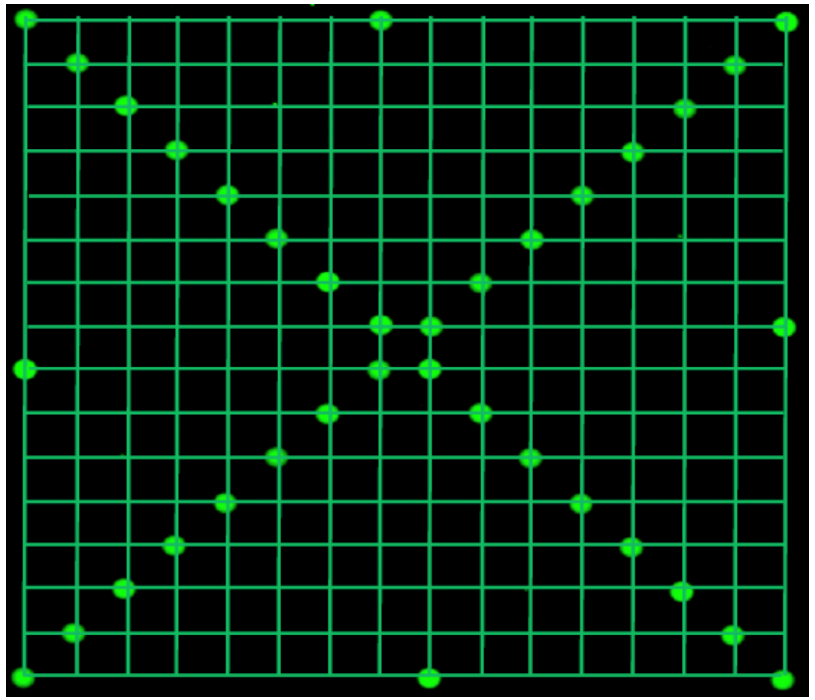
Workflow for SMM data analysis

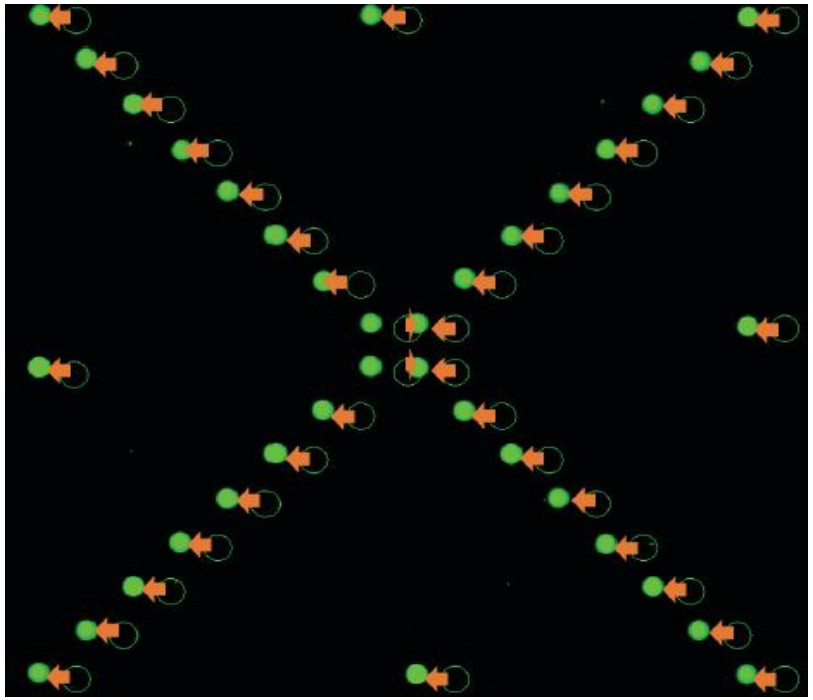
1. Align spots using fluorescence on 532 nm channel (sentinel spots)
2. Quantify fluorescence on 635 nm channel
3. Identify ‘hits’ with improbably high fluorescence
4. Complete ‘by eye’ analysis of putative hits to manually remove false positives



Align SMM using sentinel spots

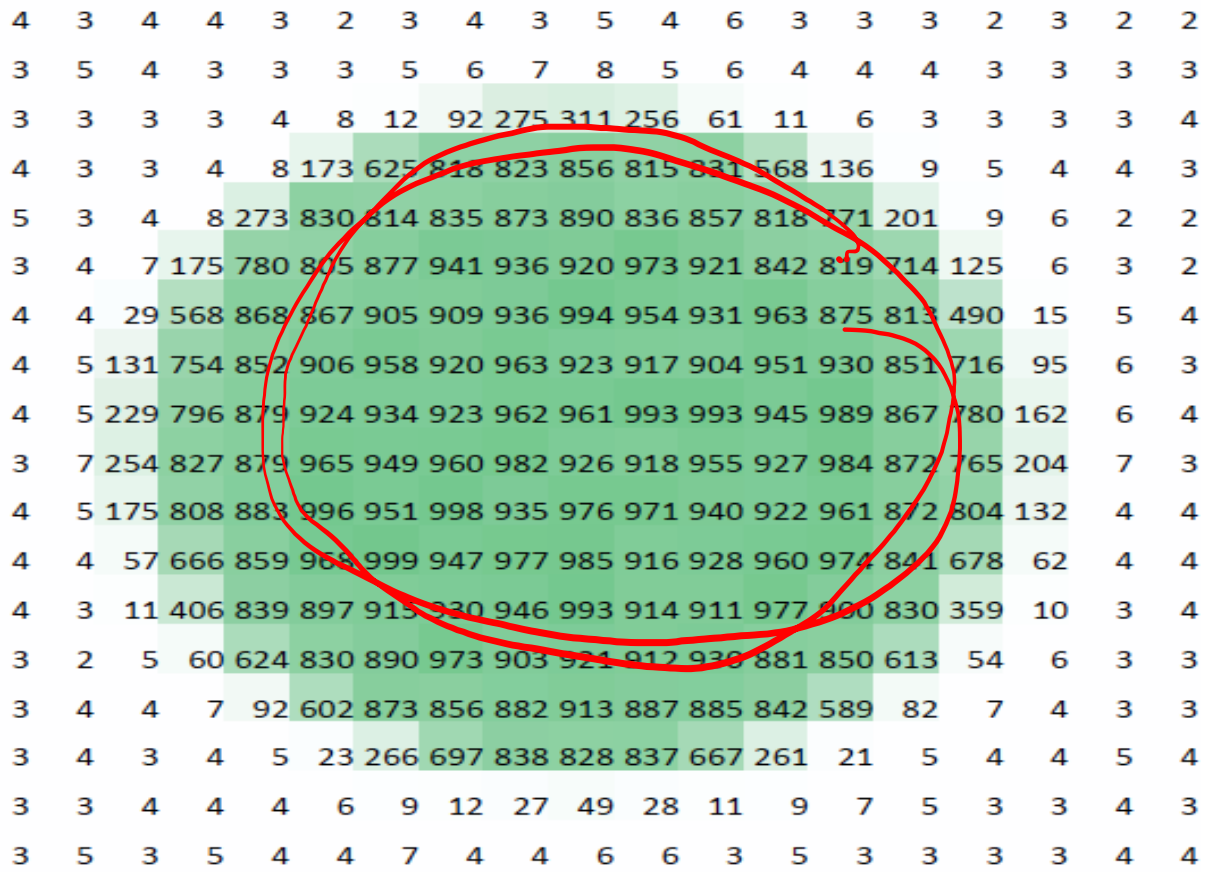
- Slides are printed in block patterns (16 rows x 16 columns)
- Each ligand spot is identifiable via intersecting lines between sentinels





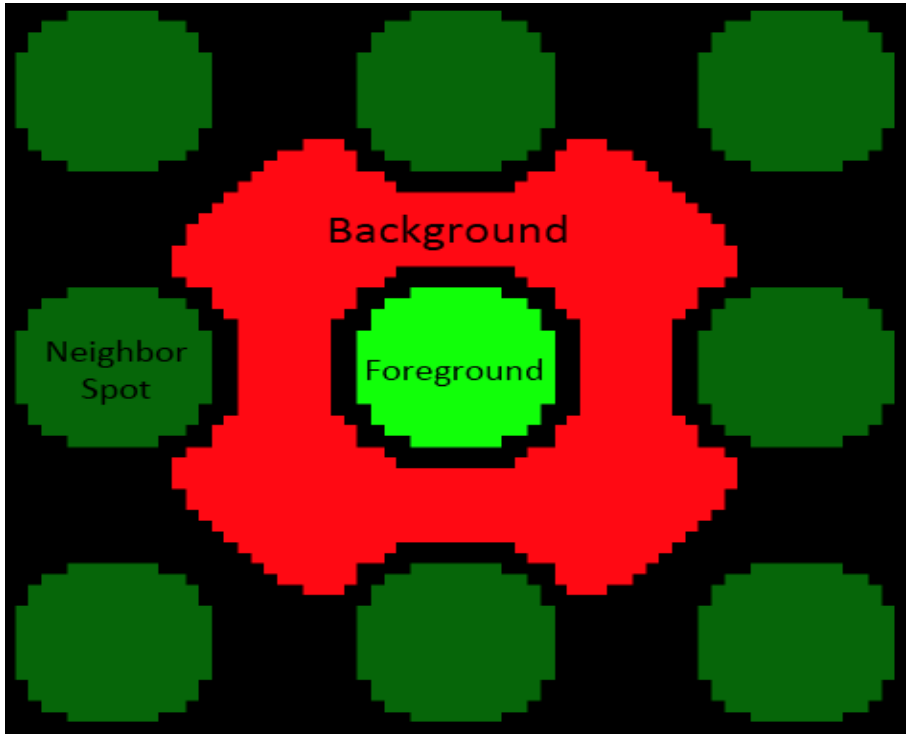
Spots are represented by an array of numerical values

- 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 Attached
- Background:
  - Area around the SM

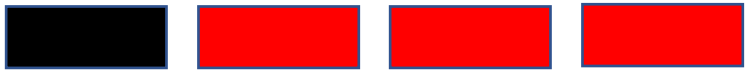


$$\text{Signal-to-noise ratio (SNR)} = \frac{\mu_{foreground} - \mu_{background}}{\sigma_{background}}$$

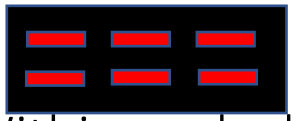
How do you identify hits from the SMM data?

First, consider **bias** that exists in the data set

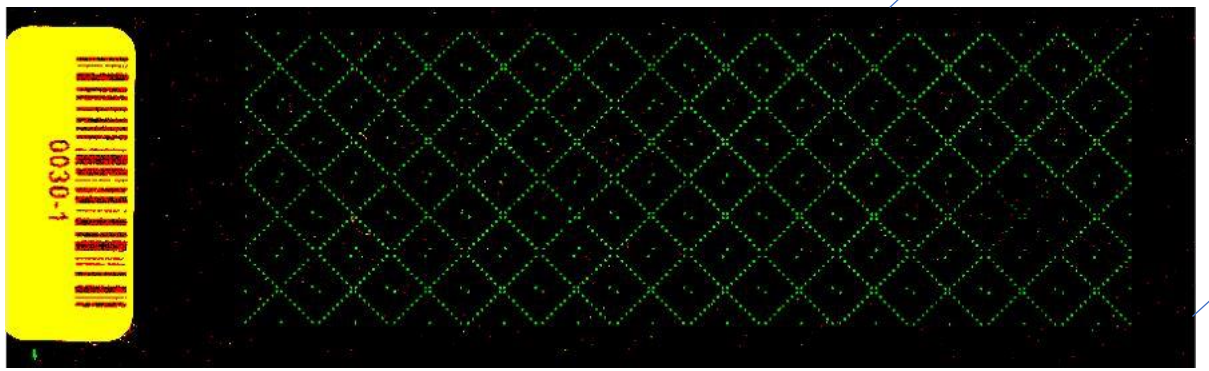
- Across all slides



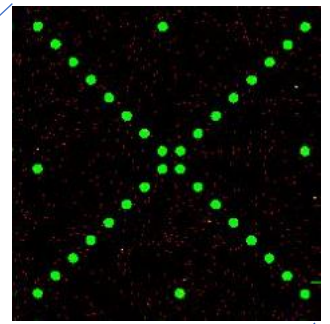
- Within each block



- Within each slide



Block

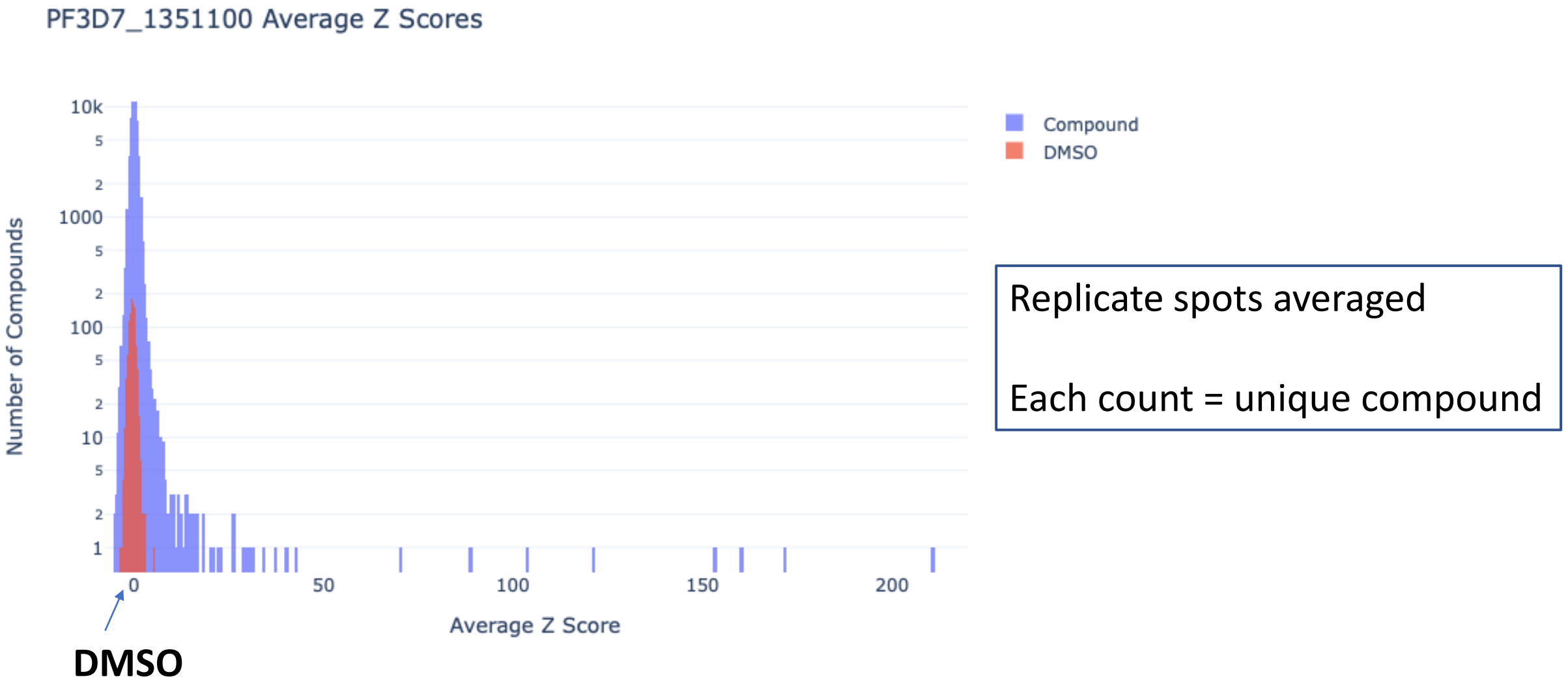


Slide

Then, identify hits with significantly higher fluorescence over background

Lastly, manually confirm hits to eliminate false positives

Average Z-score calculated for all compounds

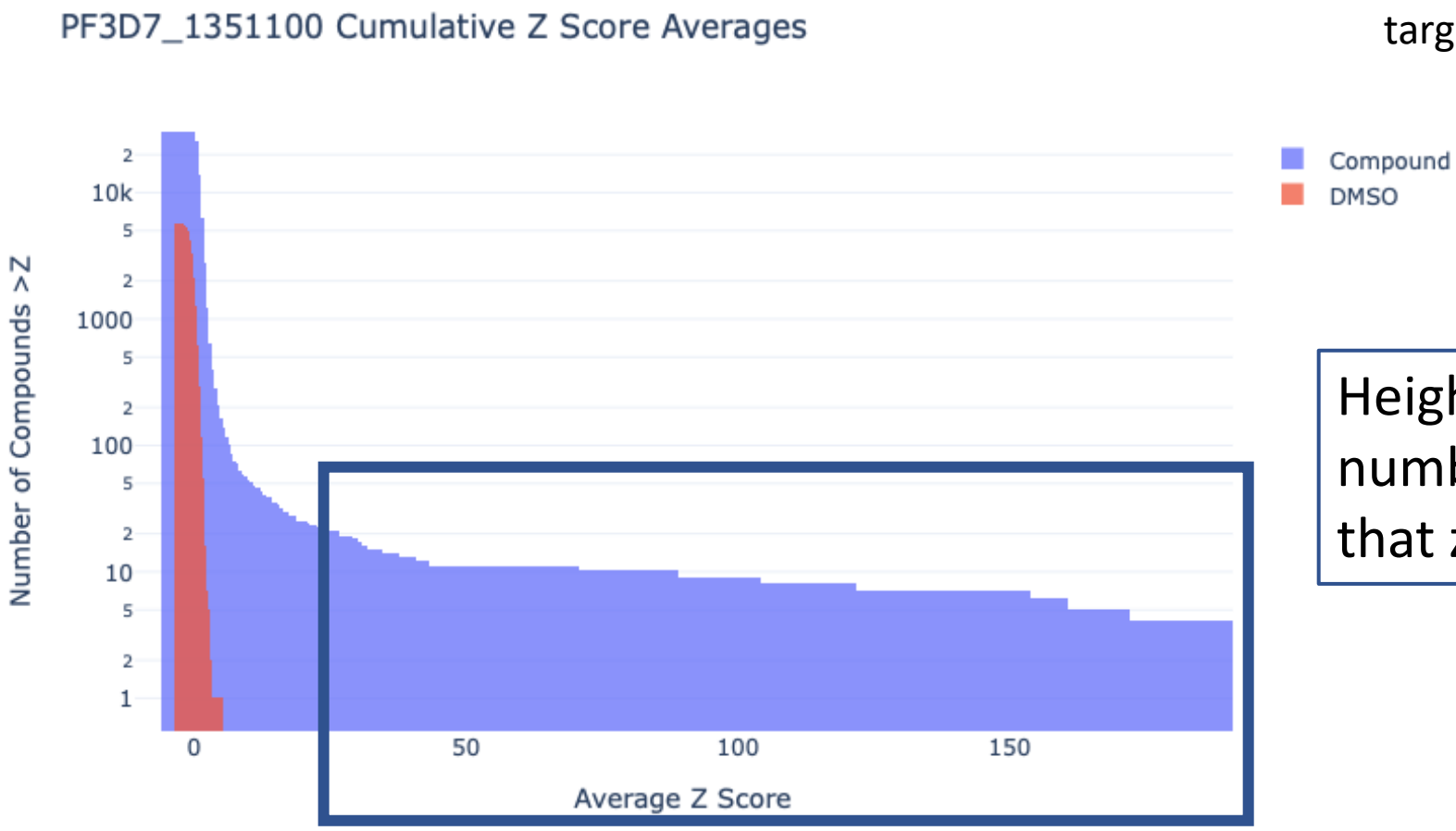


Replicate spots averaged

Each count = unique compound

How do you determine a threshold Z-score?

These screens are **TOOLS** – they alone cannot tell you whether a target IS a candidate drug or not.



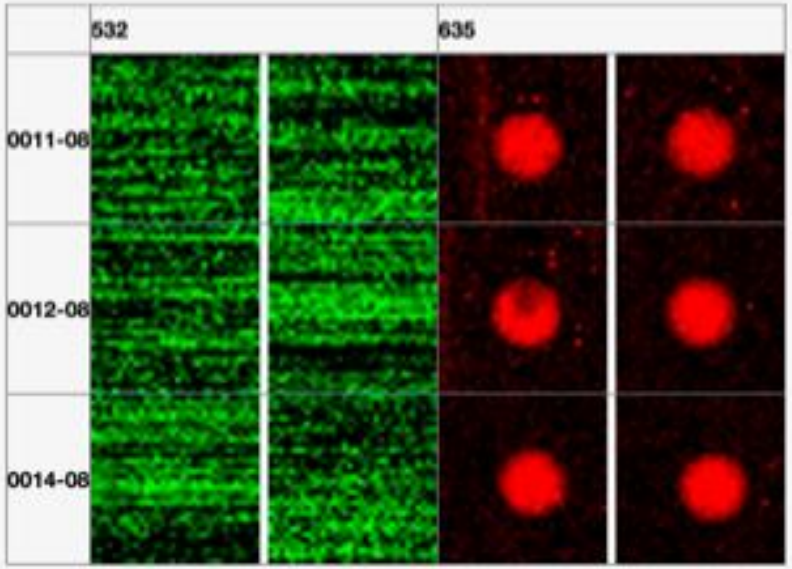
Height at x (average z score) = number of compounds (y) with that z-score or higher

How do you validate hits manually?

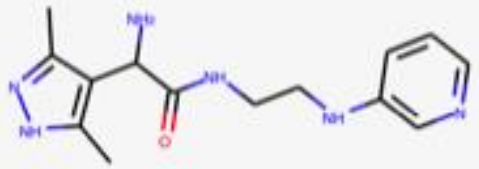
1) True Hit
2) Sentinel Spot
3) ???

532 – Fluoroscein
635 – Anti-His
Left – Green / Right - Red

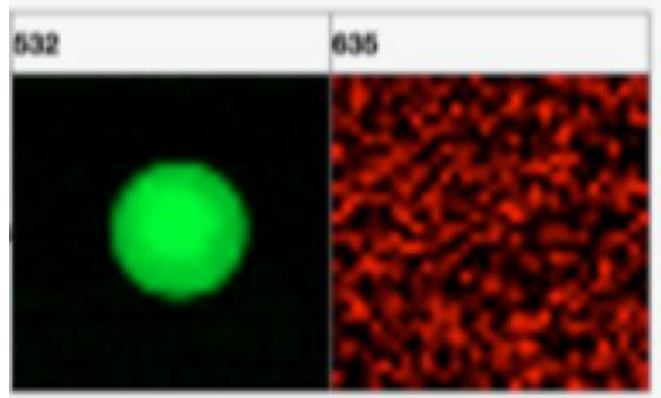
	ID	Robust Z	SMILES	Validated
49592	13:KI0001...	51.03151	C[C@H](C...	-1
42089	11:KI0001...	45.09263	CC1=C(C(...	example
6782	02:KI0001...	39.91118	CCNC(=O...	-1
29108	08:KI0001...	39.59436	C1C(C2=...	-1
44736	12:KI0001...	33.03555	C1CN(C2...	-1
29660	08:KI0001...	31.94118	CC1=NC2...	-1
11360	03:KI0001...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

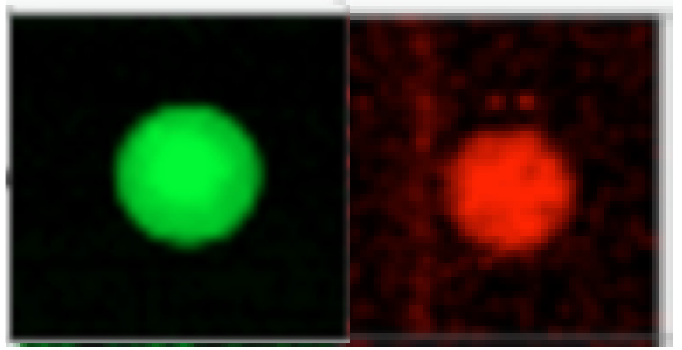


B True Hit



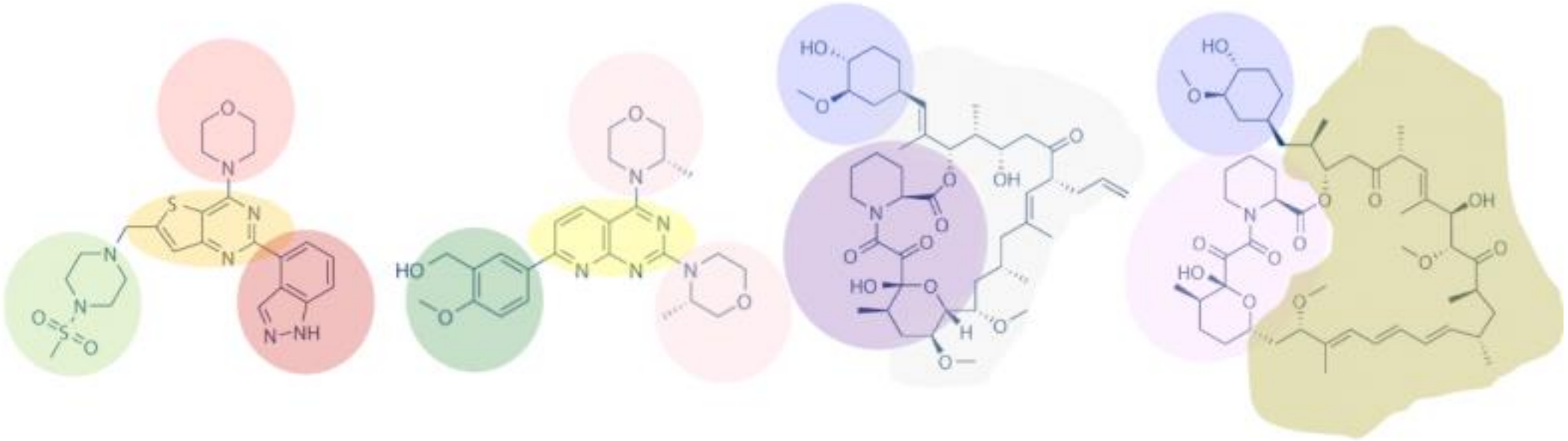
A Sentinel Spot





C Autofluorescer

Compare chemical structures of identified compounds



For Today

- Work through SMM procedure
- Evaluate chemical structures of small molecules that will be used in next lab's assay

** If you would like to revise your JP slide and replace your homework, you can do so. Turn it in by 10pm tonight **

For M2D5

- Draft an outline of the introduction for your Research Article
  - Use guidance on the Wiki section for Homework and the Research Article assignment