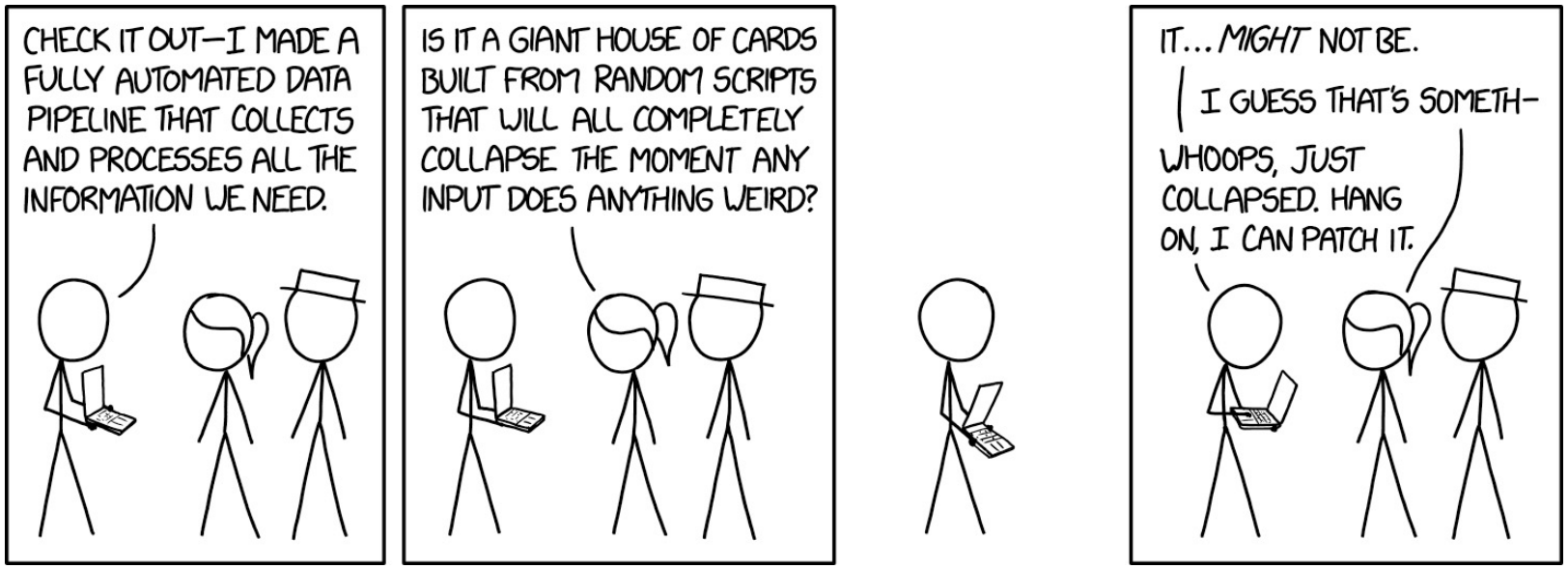
M2D1: Review small molecule microarray (SMM) experiment and results

1. Prelab
2. Walk through SMM
3. Examine chemical structure of hits



Due dates are approaching!

Mod2

- Journal Club presentation (15%)
  - Individual
  - Presentations on 10/26 & 10/28
- Research article (20%)
  - Individual
  - due 11/22
- Laboratory quizzes (collectively 5%)
  - M2D4 and M2D7
- Notebook (collectively 5%)
  - one entry will be graded by Thomas 24 hr after M2D7
- Blog (part of 5% Participation)
  - due 10/30 & 11/23 via Slack channel

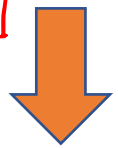
Wrap-up Mod 1

- Data summary due Wed. 10/13 @10pm
  - via Stellar
- Research talk due Sat. 10/16 @10pm
  - via GMAIL
- Data summary revisions due Sat. 10/23 @10pm
  - via Stellar
- Blog due 10/18 @ 10pm
  - via Slack

Module 2 Roadmap

→ potential SM drugs

Determine putative PF3D7_20109-F21 binders via high throughput screening (SMM)



Create plasmid of PF3D7_20109-F21 to use in validation assays

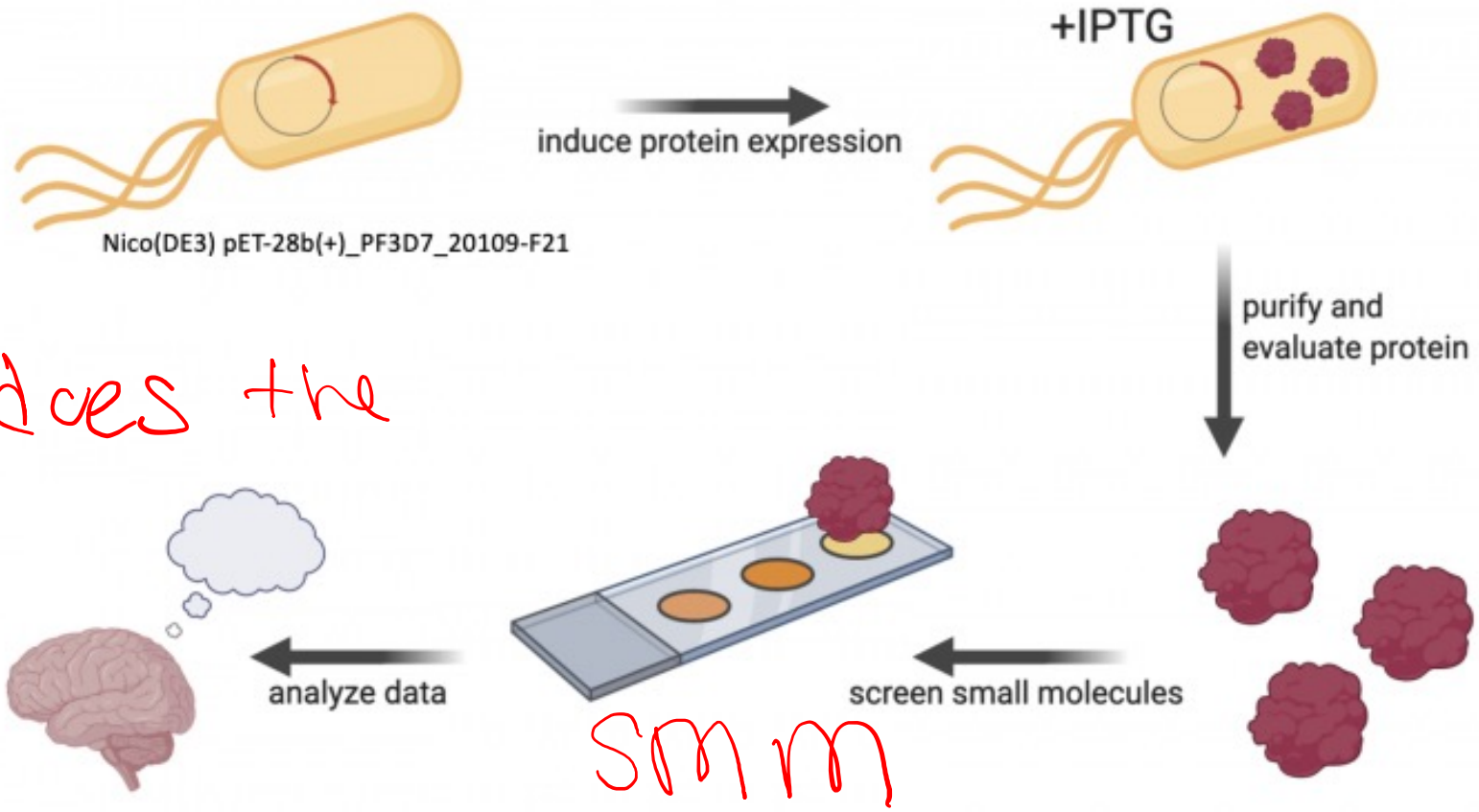
Express PF3D7_20109-F21 (from plasmid) in bacteria and purify protein

Assess purity and concentration of purified protein

Use purified protein to validate binding of small molecules identified in SMM

SMM Overview

Research goal: Identify small molecules that bind to the PF3D7_20109-F21 protein in *Plasmodium falciparum* using small-molecule microarray



where does the protein stick?

SMM

← protein of interest

Why are we taking this approach?

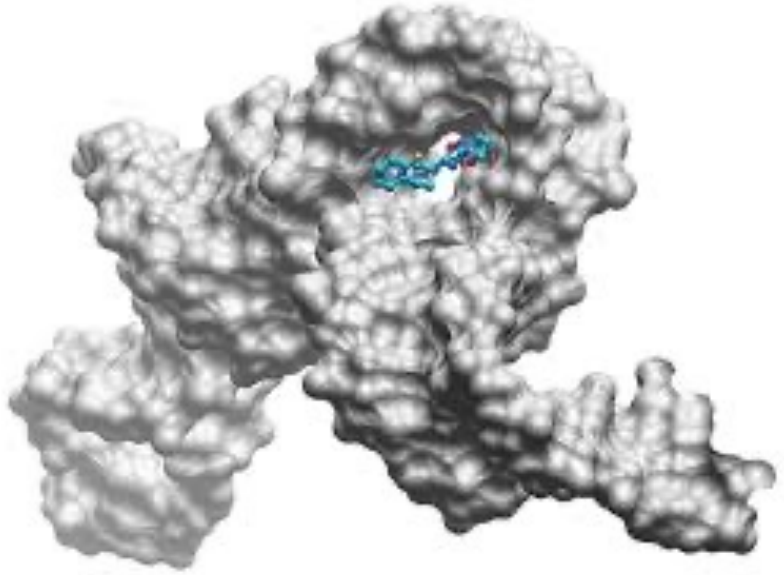
limited info

unbiased

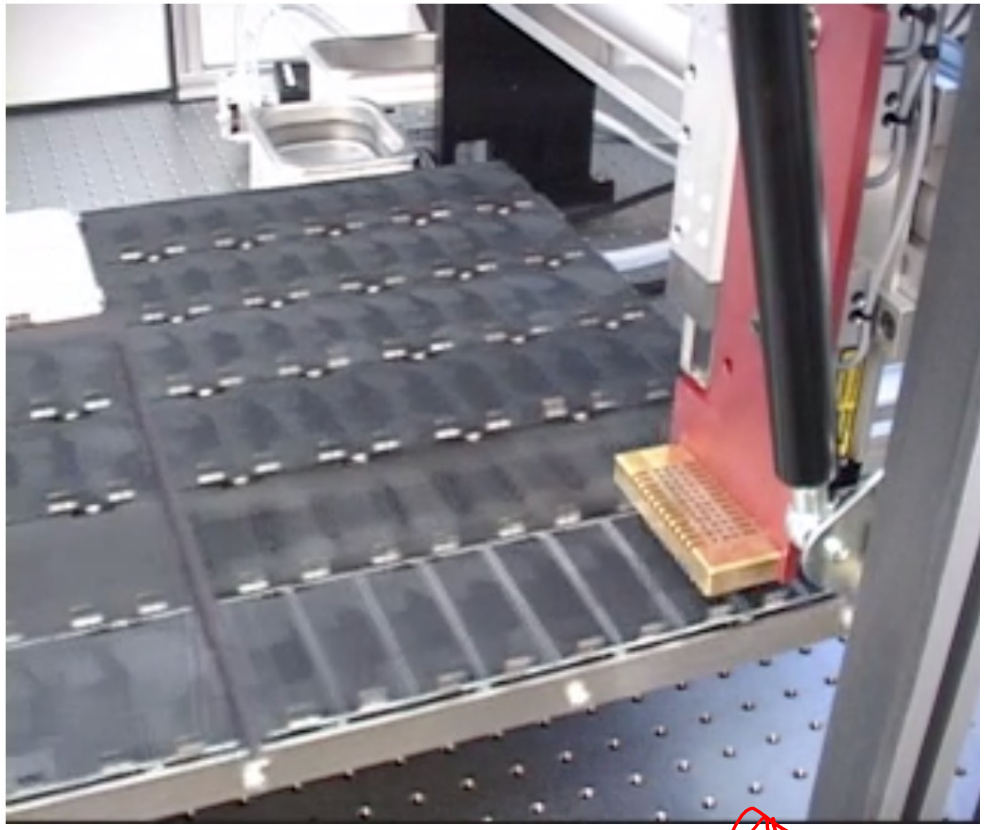
- High throughput assays are useful in screening potential therapeutic targets
  - Allows unbiased exploration of potential therapeutics
  - Allows examination of targets with limited information

- Small molecules
  - Mw < 500 Da
  - Natural or synthetic
  - Frequently comprised of Carbon/Nitrogen/Oxygen

into cells/across barriers

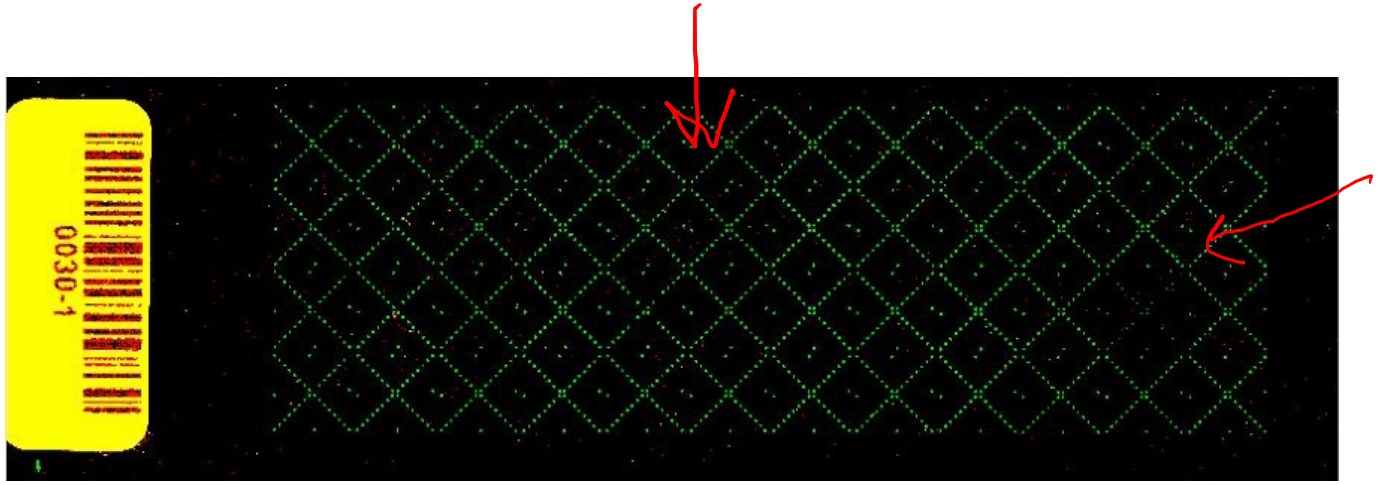


Small Molecule Microarray (SMM)



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

robot

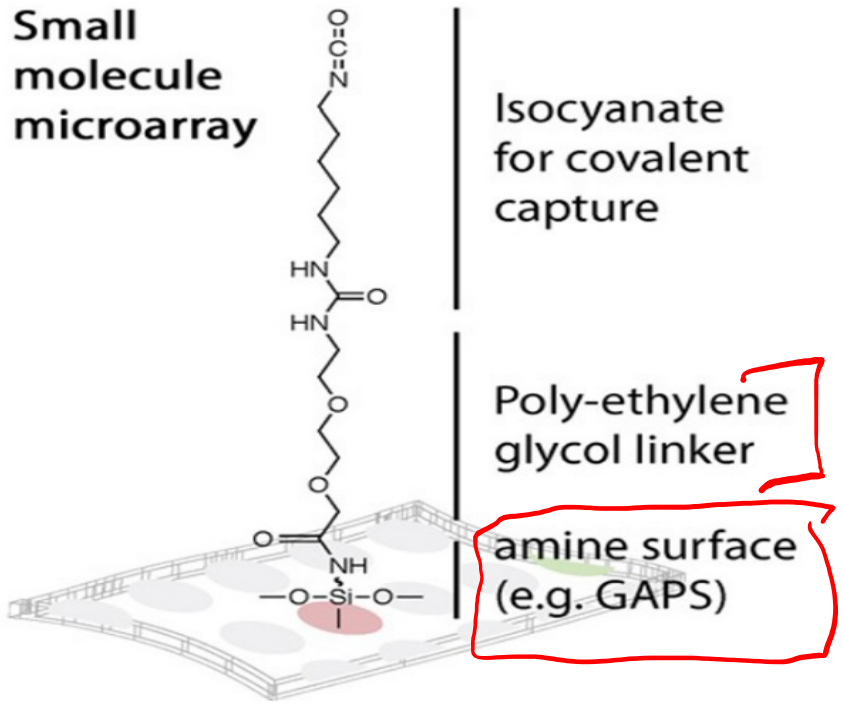


green dots map

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

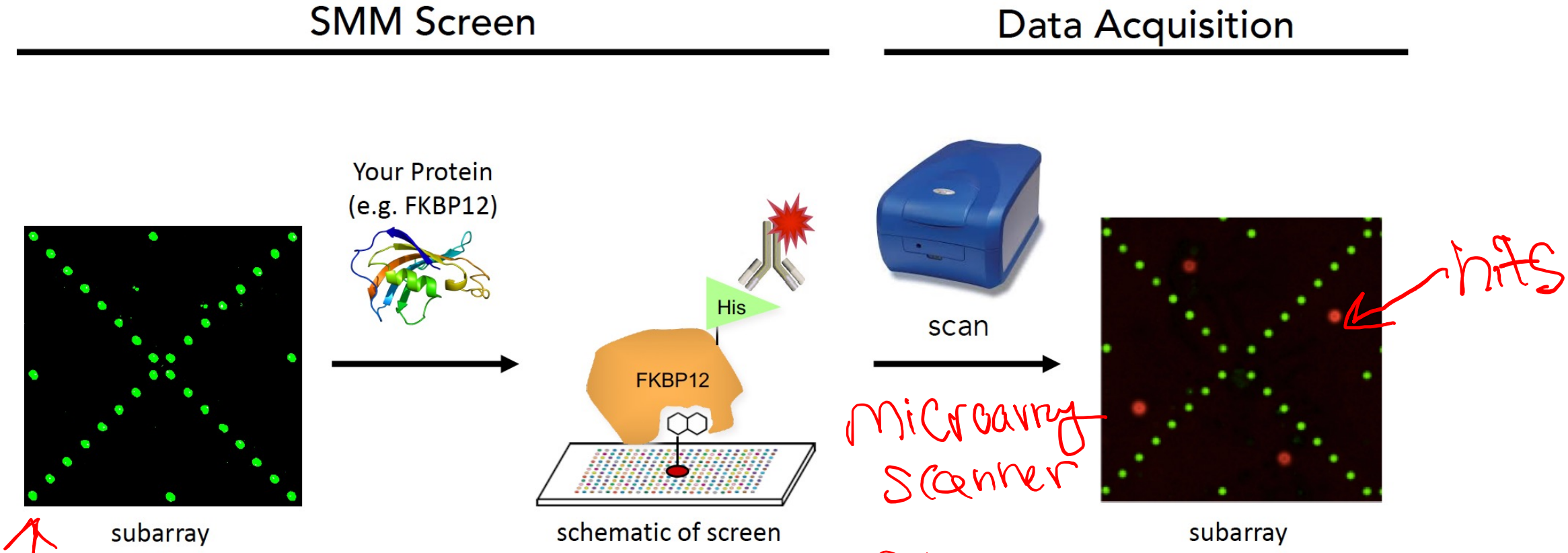
SMM slide preparation

- Gamma-aminopropylsilane (GAPS) slide coated with polyethylene glycol (PEG) spacer
- 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



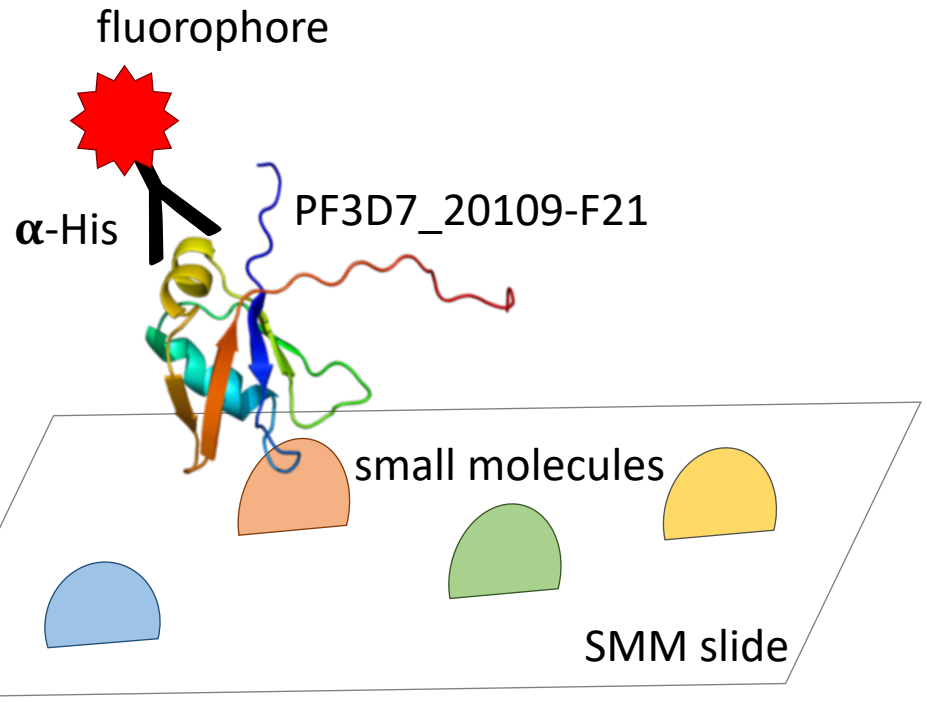
B. Leifer, S. Doyle

How do we screen for ligands that bind PF3D7_20109-F21?

- Incubate the SMM slide with 3ml of our purified PF3D7_20109-F21
- Wash away unbound protein
- Incubate SMM slide with AlexaFlour 647 anti-His antibody

→ where is protein binding?

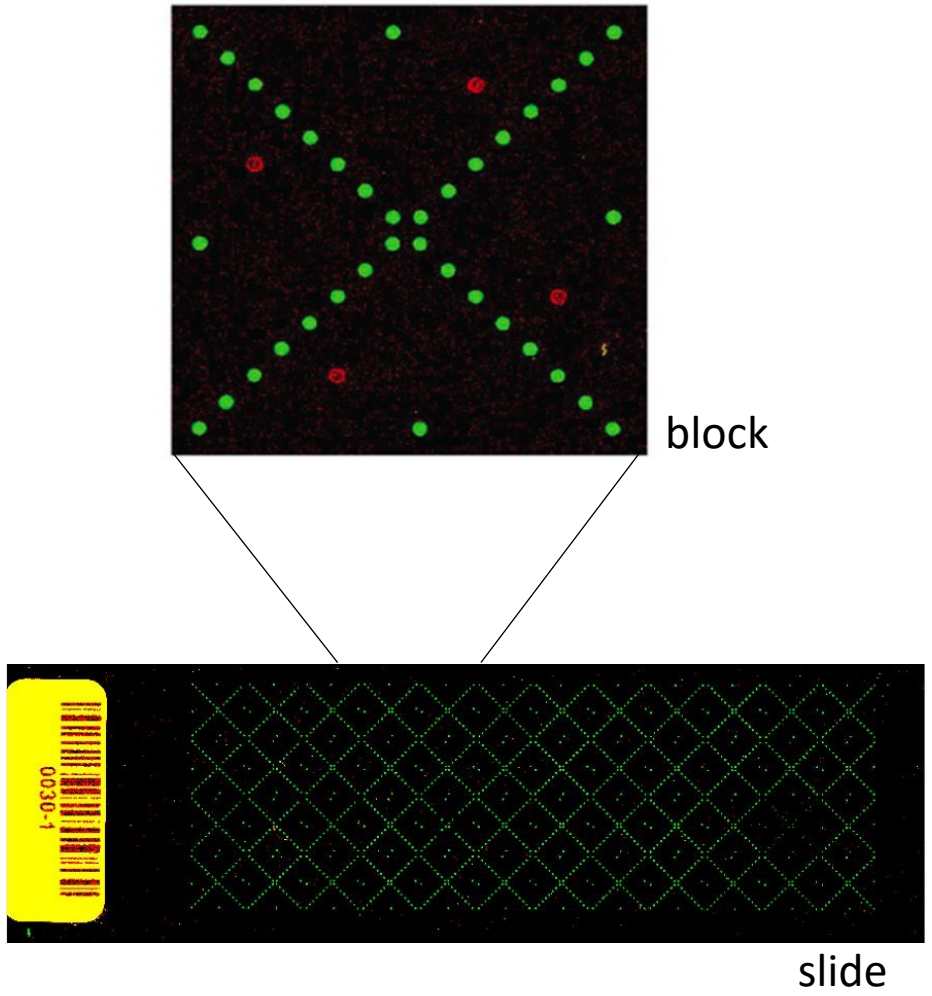
- Wash away excess antibody
- Store for scanning



Workflow for SMM data analysis

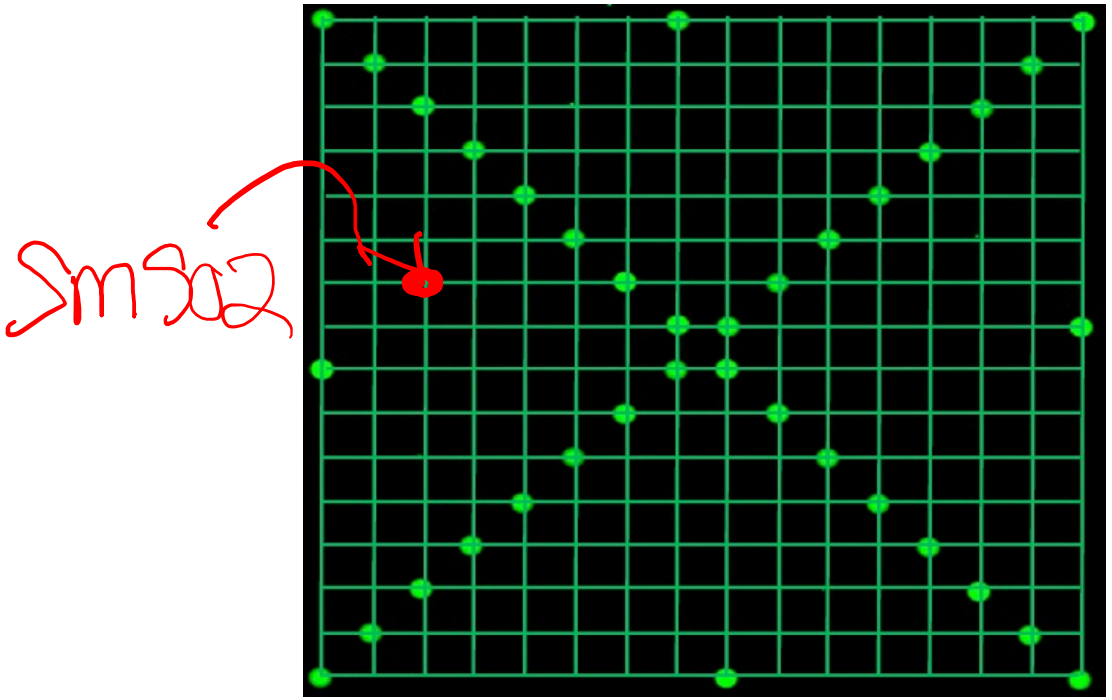
1. Align spots using fluorescence on 532 nm channel (sentinel spots)
2. Quantify fluorescence on 635 nm channel
3. Identify ‘hits’ with improbably high fluorescence
4. Complete ‘by eye’ analysis of putative hits to manually remove false positives

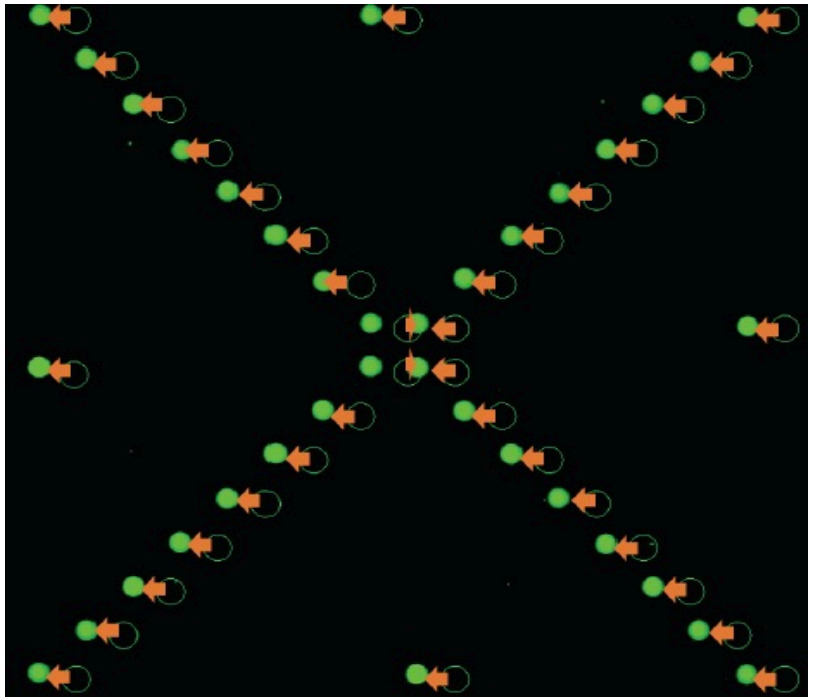
anti-His Ab



Align SMM using sentinel spots

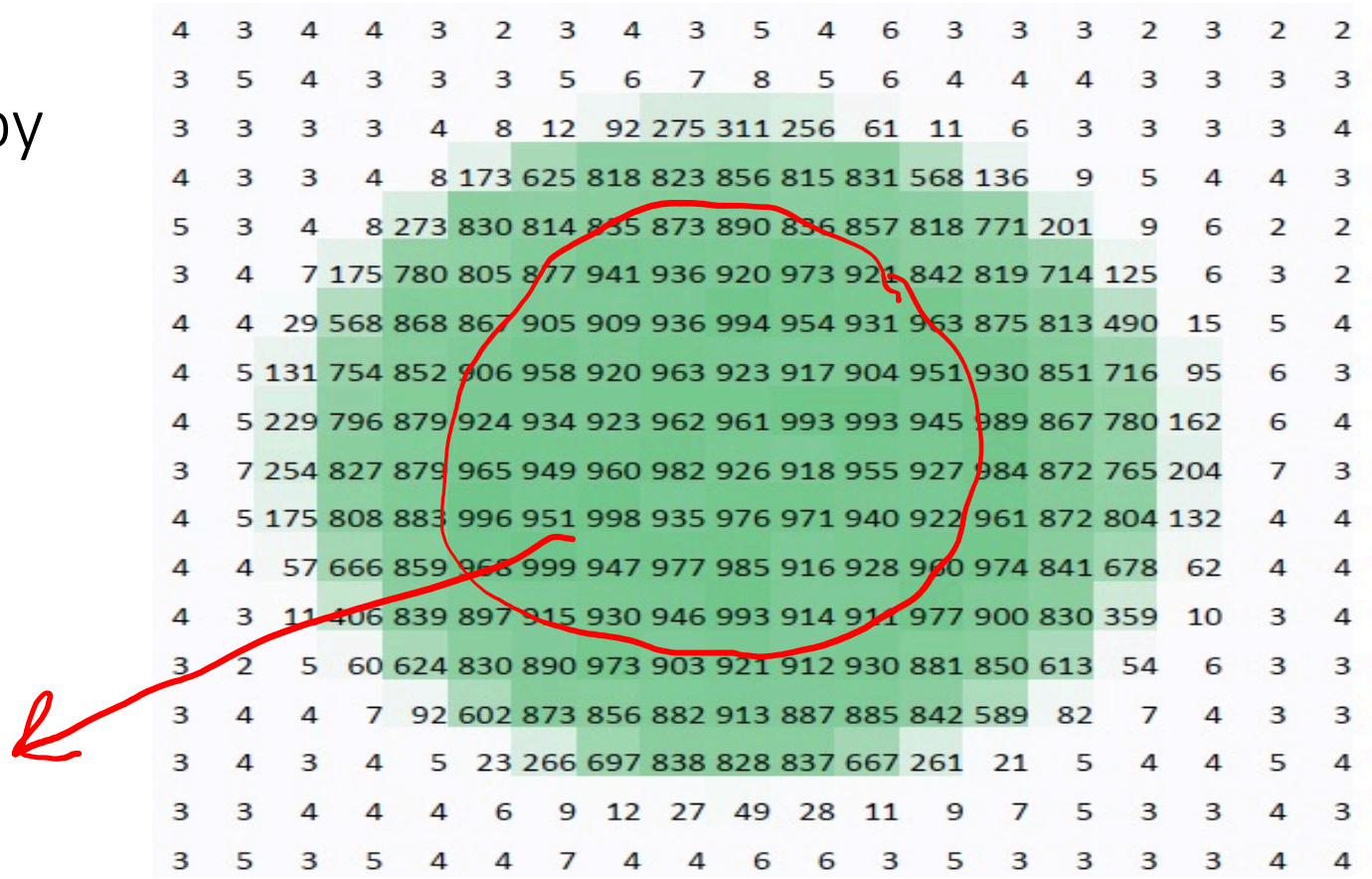
- Slides are printed in block patterns (16 rows x 16 columns)
- Each ligand spot is identifiable via intersecting lines between sentinels





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’



potential hit

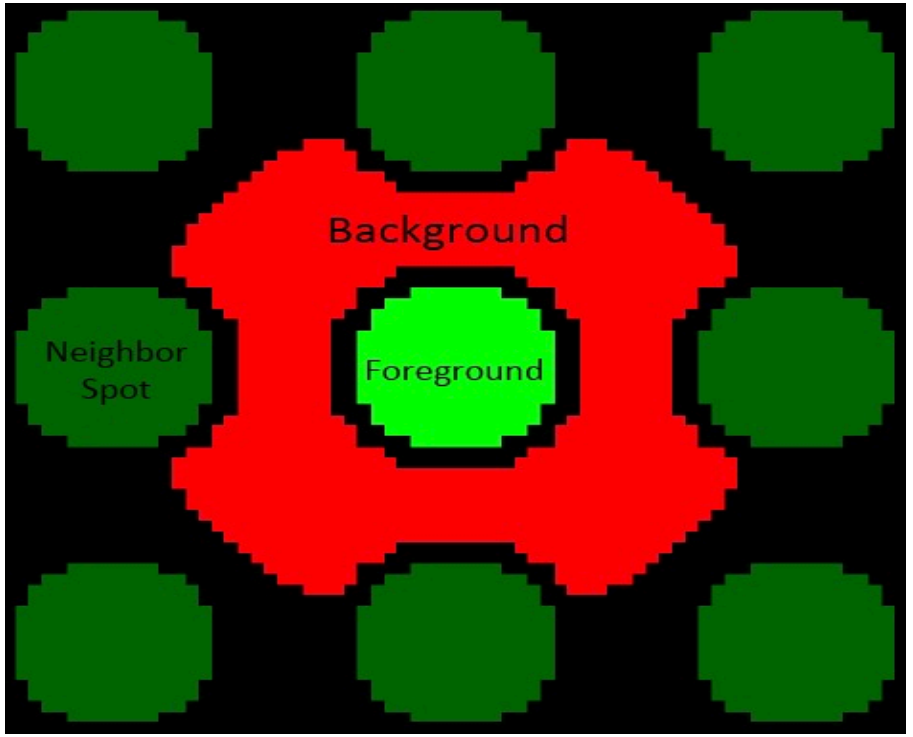
Fluorescence is quantified to identify hits

- Foreground:

SM spot

- Background:

area immediately around SM spot



→ non-specific binding?

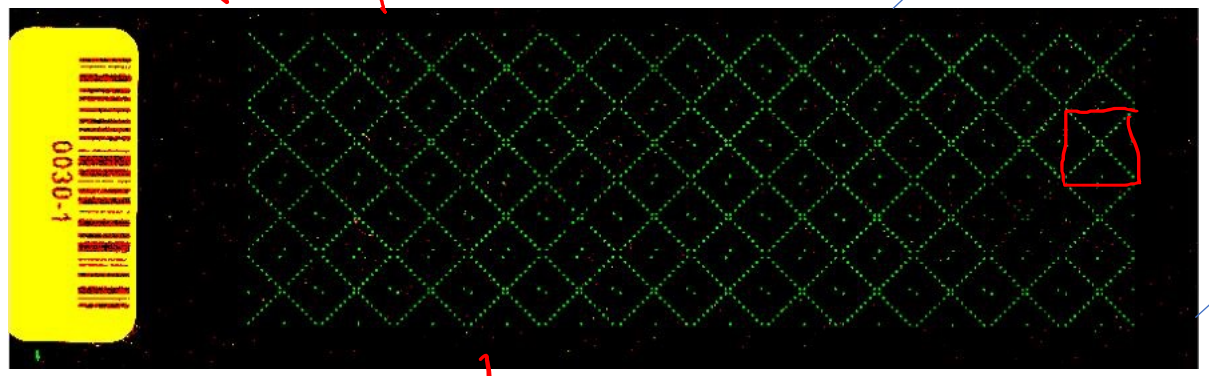
$$\text{Signal-to-noise ratio (SNR)} = \frac{\mu_{foreground} - \mu_{background}}{\sigma_{background}}$$

→ calculate Z score

How do you identify hits from the SMM data?

First, consider bias that exists in the data set

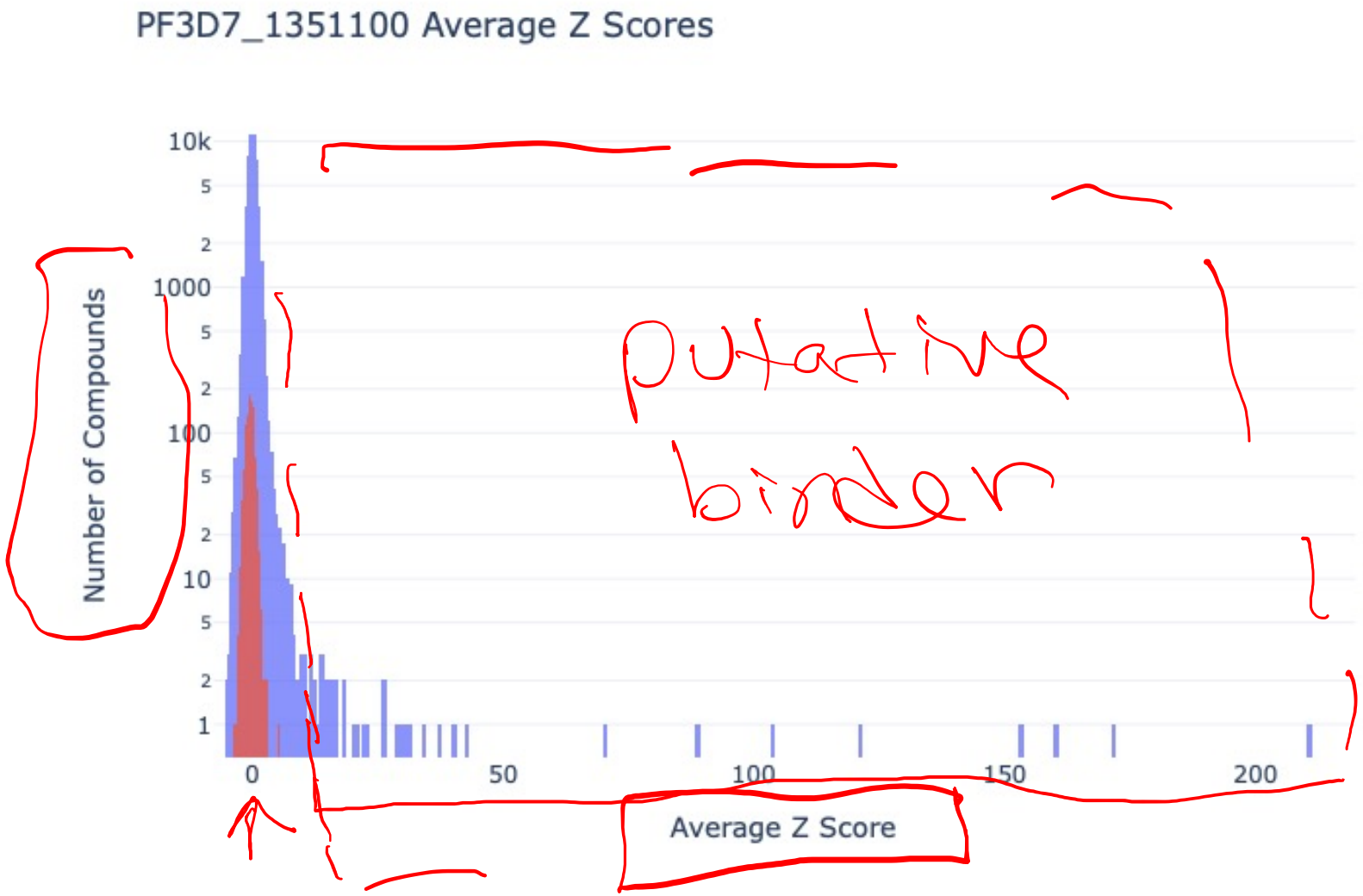
- Across all slides
- Within each block
- Within each slide



Then, identify hits with significantly higher fluorescence over background

Lastly, manually confirm hits to eliminate false positives

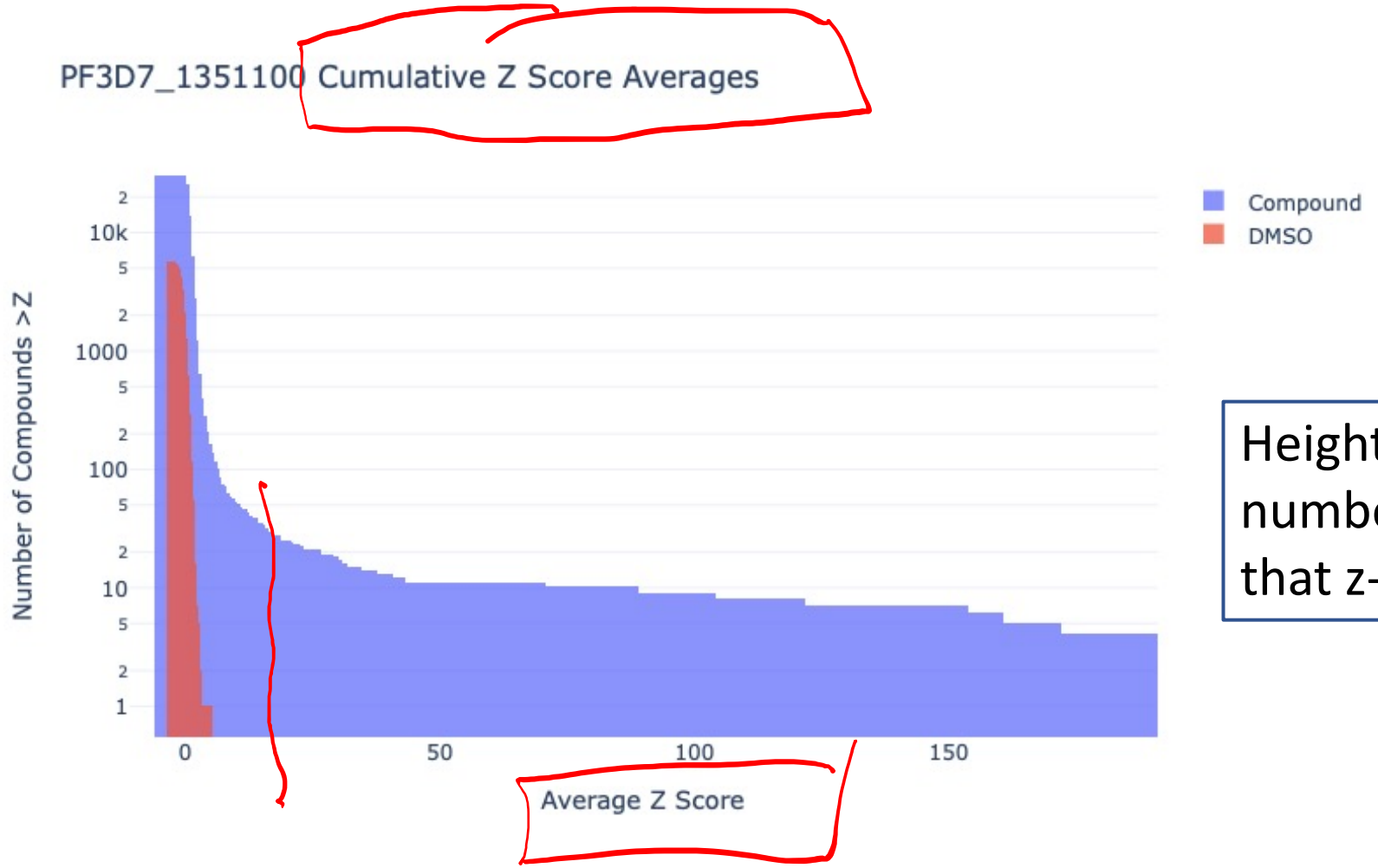
Average Z-score calculated for all compounds



Replicate spots averaged

Each count = unique compound

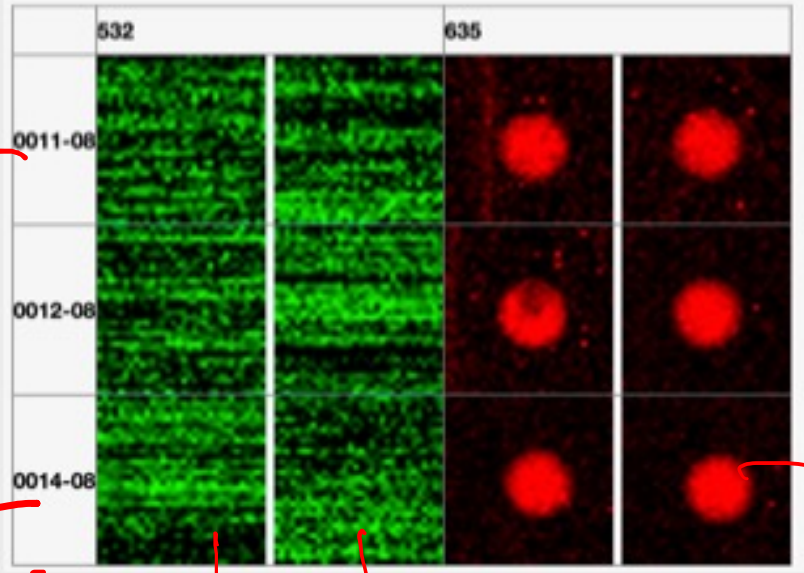
How do you determine a threshold Z-score?



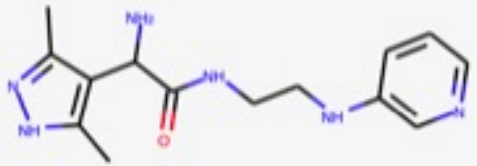
Height at x (average z score) = number of compounds (y) with that z-score or higher

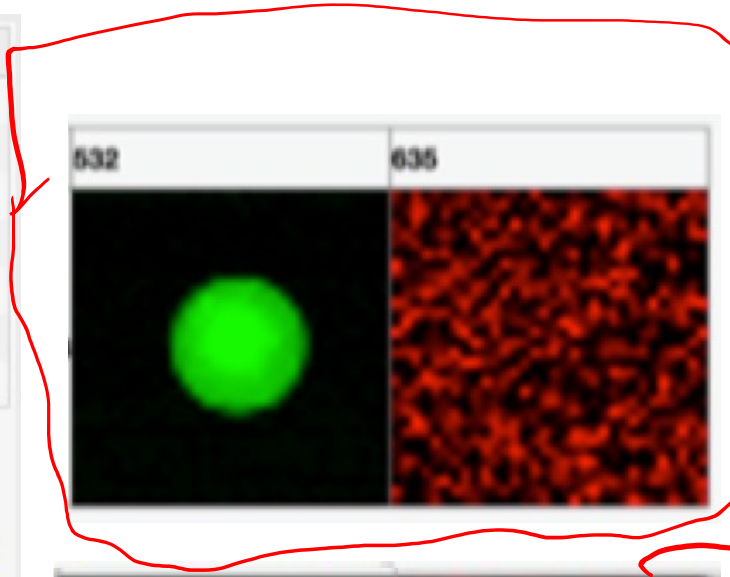
How do you validate hits manually?

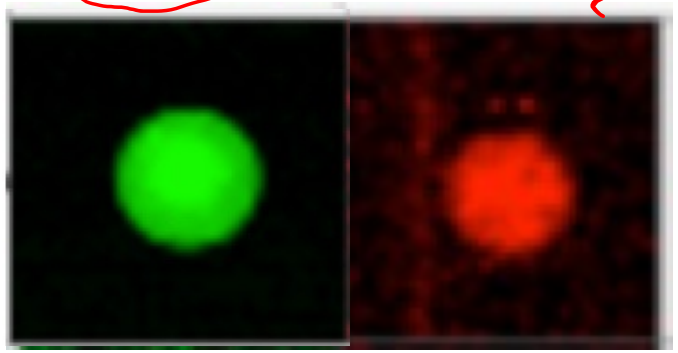
	ID	Robust Z	SMILES	Validated
49592	13:KI0001...	51.03151	C[C@H](C...	-1
42089	11:KI0001...	45.09263	CC1=C(C(...	example
6782	02:KI0001...	39.91118	CCNC(=O...	-1
29108	08:KI0001...	39.59436	C1C(C2=...	-1
44736	12:KI0001...	33.03555	C1CN(C2...	-1
29660	08:KI0001...	31.94118	CC1=NC2...	-1
11360	03:KI0001...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



duplicate spots







→ red spot = hit

Spike replicates → no signal

For Today

- Work through SMM procedure
- Evaluate chemical structures of identified hits
- Discuss close reading of scientific papers with Noreen

For M2D2

- Choose a journal article and sign up on the wiki
  - An article can be presented by only 1 person in a section (first come first served)

- Write and submit a short summary based on wiki guidelines