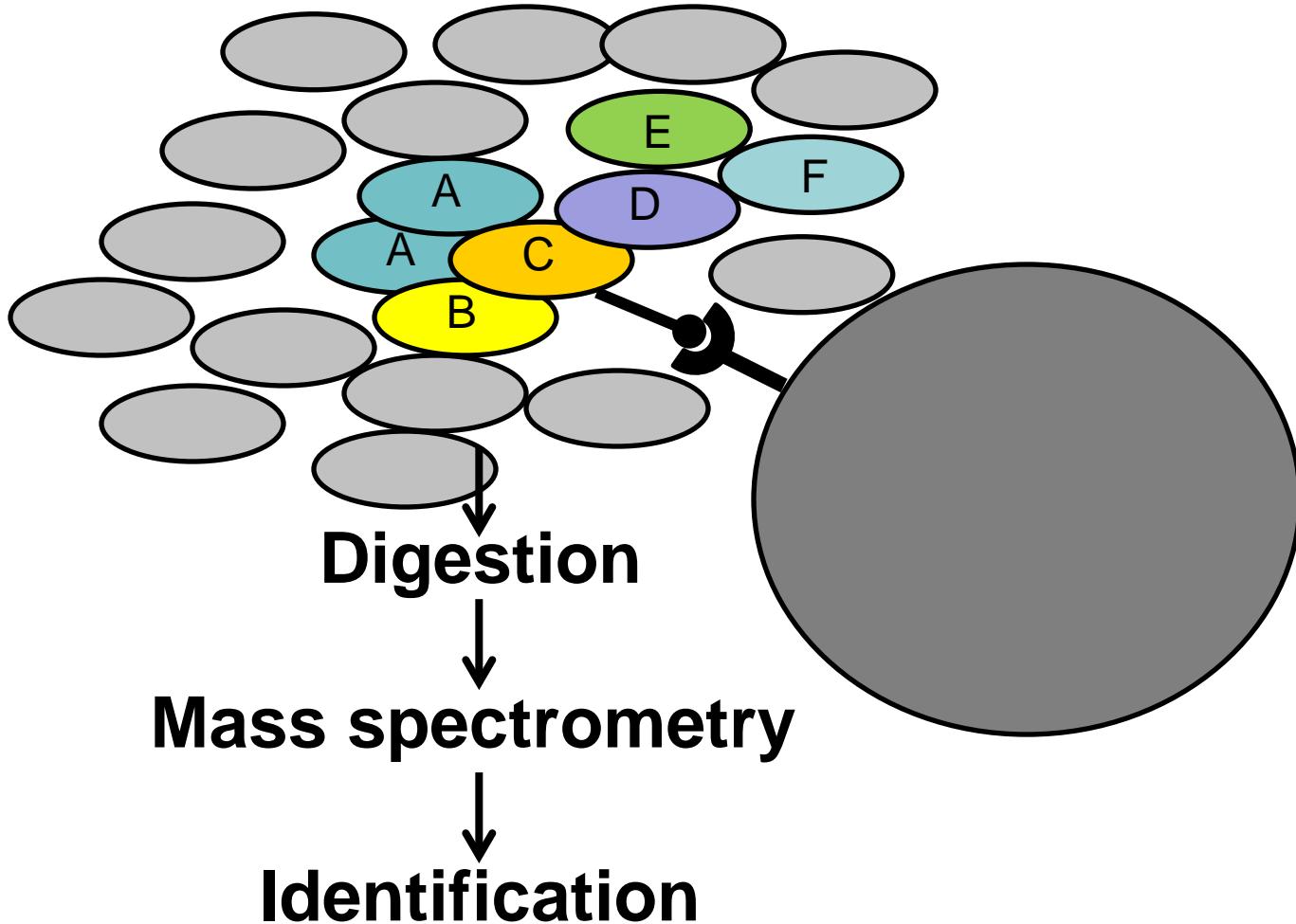
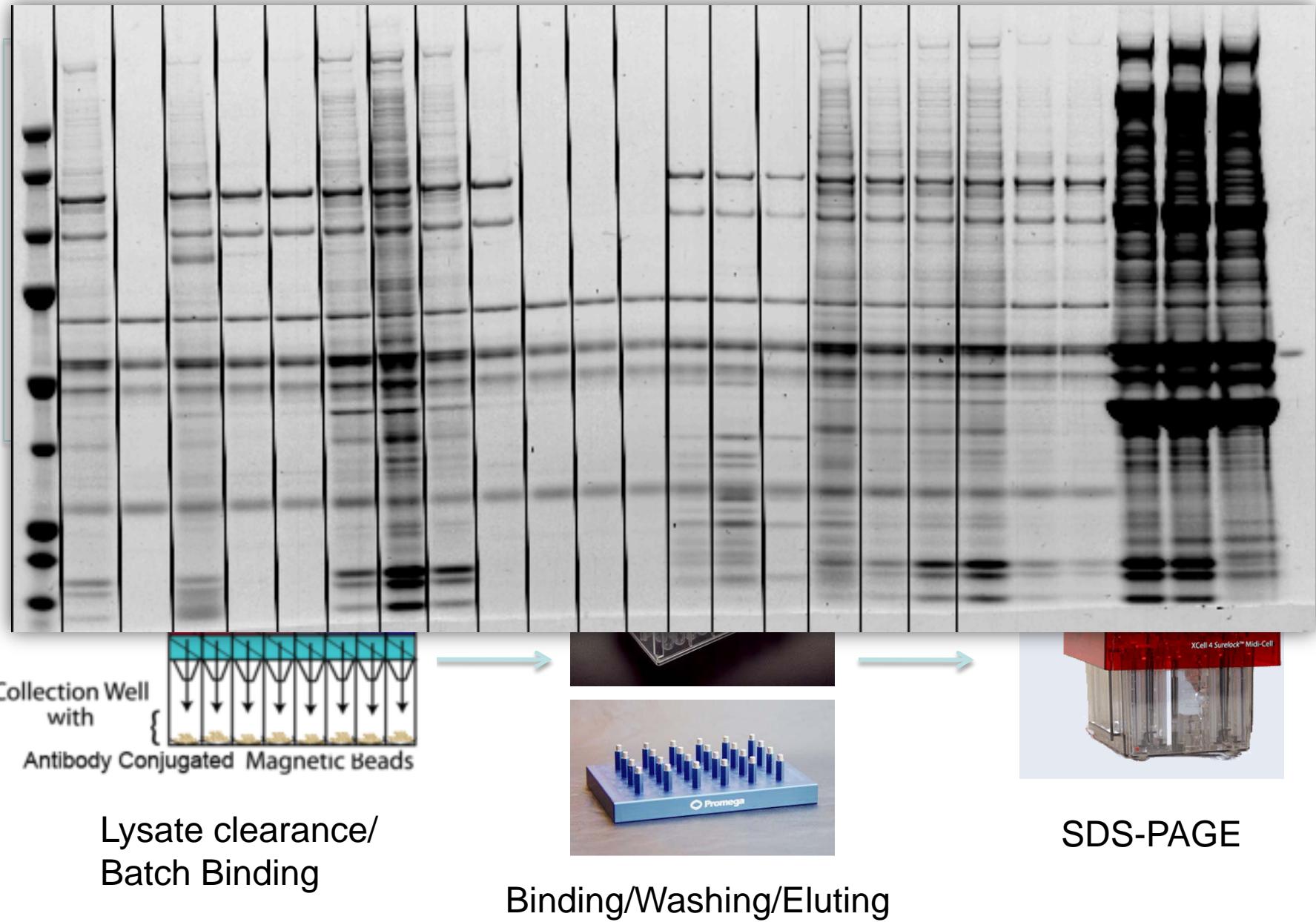


Proteomics Informatics - Protein Characterization II: Protein Interactions (Week 11)

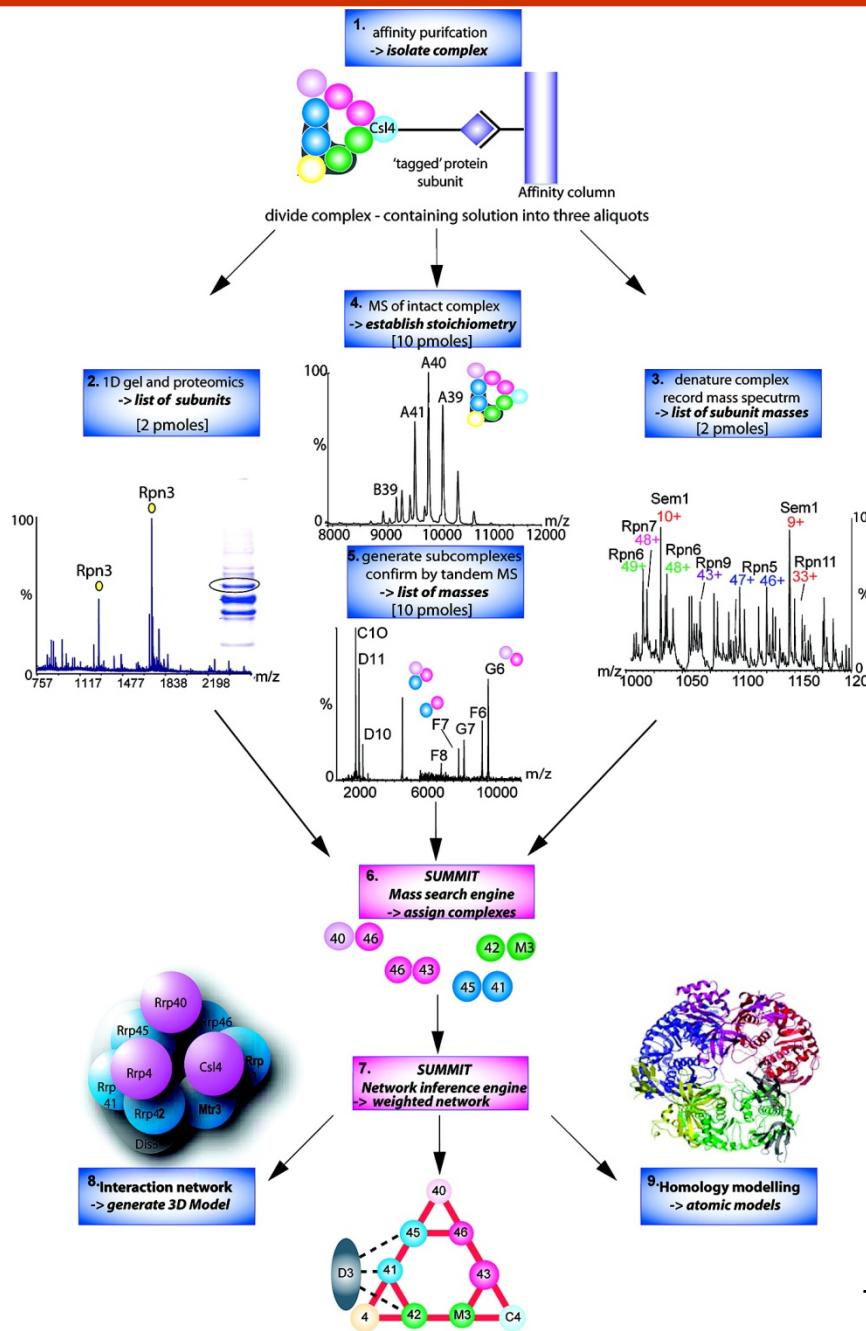
Discovering New Protein Interactions with Affinity Capture Mass Spectrometry



Affinity Capture Optimization Screen

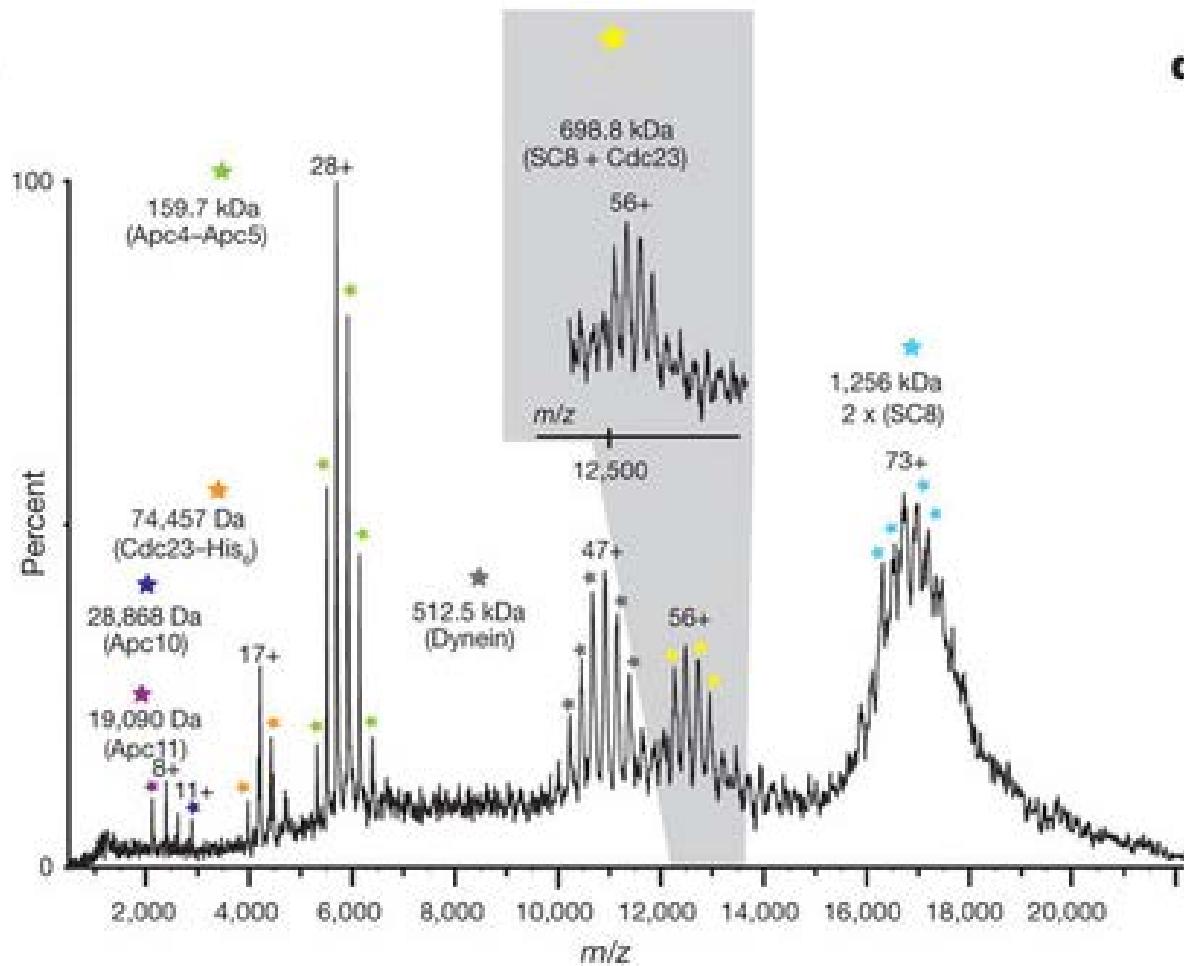


Analysis of Non-Covalent Protein Complexes



Non-Covalent Protein Complexes

c



d

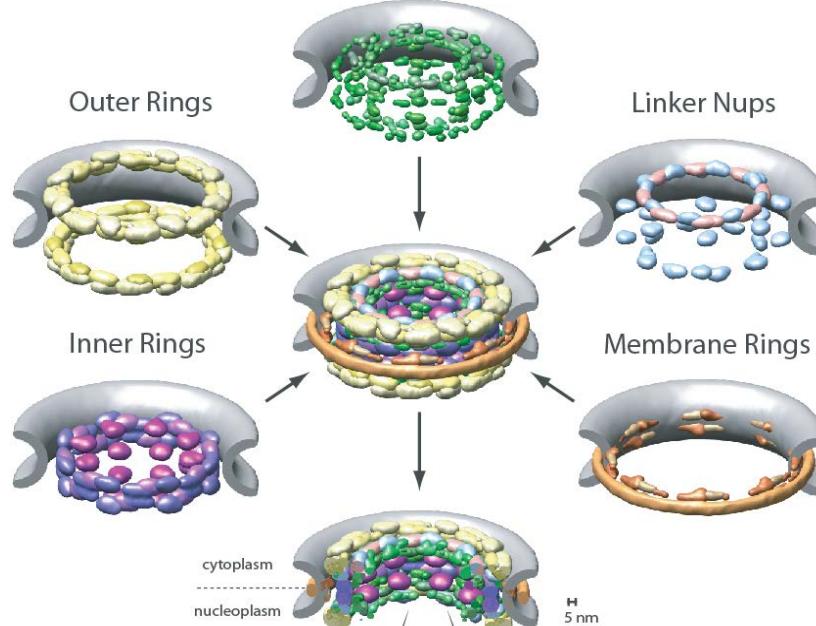
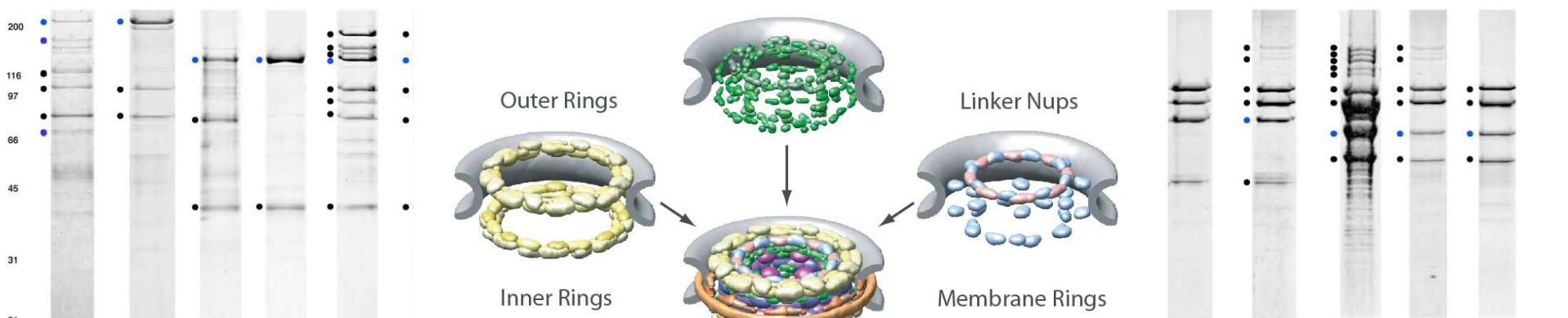
<i>S. cerevisiae</i> APC/C subunit	Stoichiometry
Apc1	1
Apc2	1
Apc4	1
Apc5	1
Apc10	1
Apc11	1
Cdc23	2
Cdc16	2
Cdc27	2
Apc9	ND
Mnd2	1
Apc13	1
Cdc26	2
Total MW (kDa)	1,127–1,158

Molecular Architecture of the NPC

Ultracentrifugation	Quantitative immunoblotting	Affinity purification	Overlay assay	Electron microscopy	Immuno-electron microscopy	Bioinformatics and membrane fractionation
30 S-values	1 S-value	30 relative abundances	75 composites 7 contacts	Electron microscopy 10,615 particles	Immuno-electron microscopy 10,615 particles	30 protein sequences

Over 20 different extraction and washing conditions ~ 10 years of art.
(41 pullouts are shown)

# 32	# 19	# 26	# 17	# 67	# 36	# 34	# 23	# 56	# 31	# 61	# 58	# 50	# 28	# 38	# 27	# 48	# 65	# 47	# 29
Nup159	Nup159	Nup116	Nup116	Nup116	Nup116	Nup82	Nup82	Nup82	Nup82	Nup82	Nsp1	Nsp1	Nsp6	Nsp6	Nup57	Nup57	Nup49	Nup49	Nup49

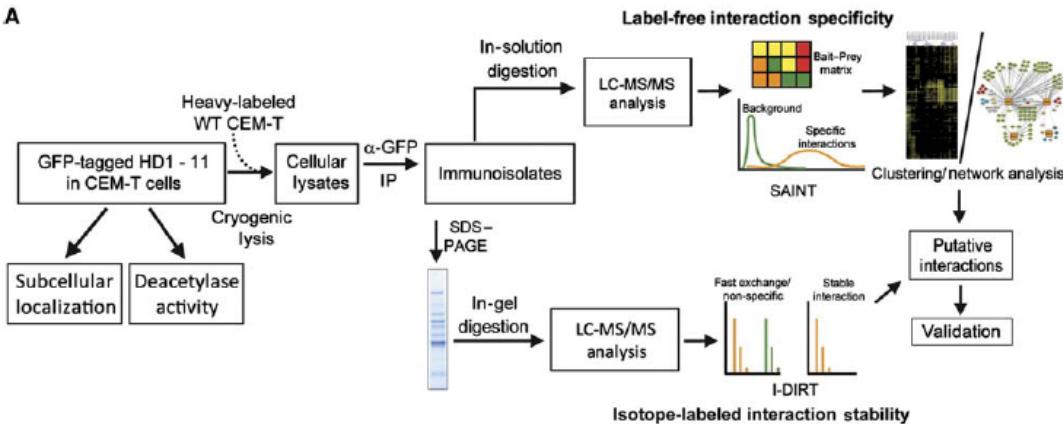


Actual model

Alber F. et al. Nature (450) 683-694. 2007
Alber F. et al. Nature (450) 695-700. 2007

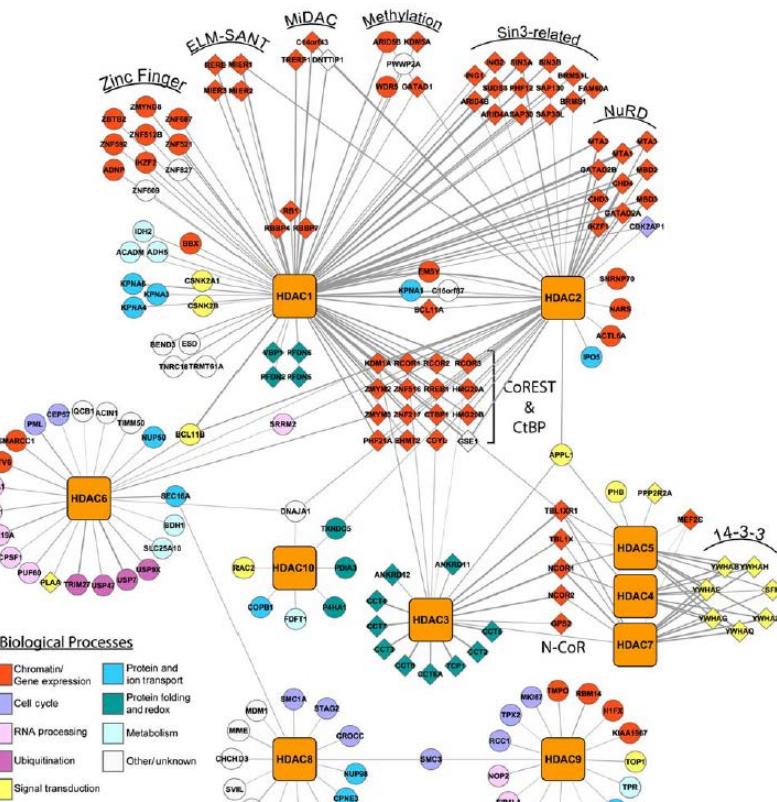
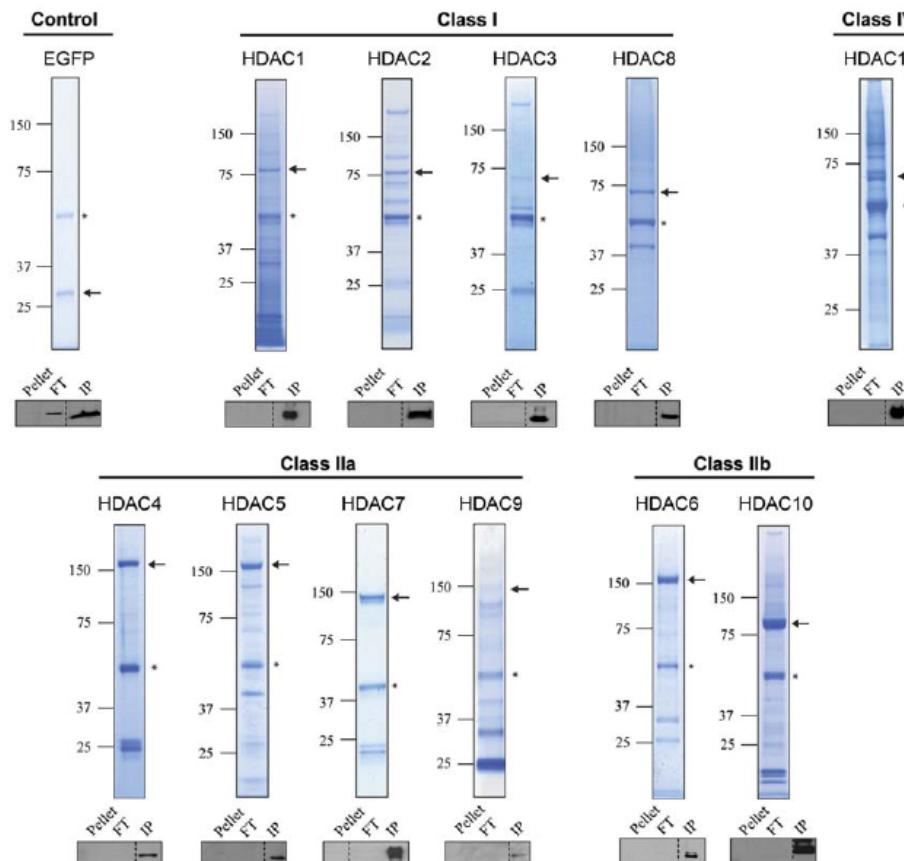
Interaction Map of Histone Deacetylases

A



Joshi et al. Molecular Systems Biology 9:672

B



Protein Complexes - specific/non-specific binding

E

Stats Table

	Bait 1	Bait 2	Bait 3	Bait 4	Bait k
Interactor 1	$X_{1,1}$	$X_{2,1}$	$X_{3,1}$	$X_{4,1}$	$X_{k,1}$
Interactor 2	$X_{1,2}$	$X_{2,2}$	$X_{3,2}$	$X_{4,2}$	$X_{k,2}$
Interactor 3	$X_{1,3}$	$X_{2,3}$	$X_{3,3}$	$X_{4,3}$	$X_{k,3}$
Interactor 4	$X_{1,4}$	$X_{2,4}$	$X_{3,4}$	$X_{4,4}$	$X_{k,4}$
Interactor m	$X_{1,m}$	$X_{2,m}$	$X_{3,m}$	$X_{4,m}$	$X_{k,m}$

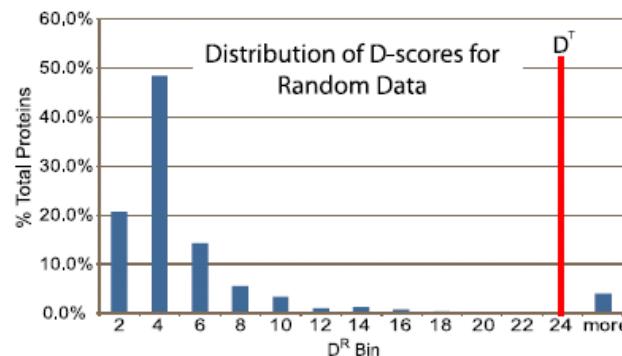
$X_{i,j}$ = total spectral counts for interactor j from bait i

$$\bar{X}_j = \frac{\sum_{i=1, j=n}^{i=k} X_{i,j}}{k} ; n = 1, 2, \dots, m \quad (\text{Eq. 1})$$

$$Z_{i,j} = \frac{X_{i,j} - \bar{X}_j}{\sigma_j} \quad (\text{Eq. 2})$$

$$f_{i,j} = \begin{cases} 1 & ; X_{i,j} > 0 \\ X_{i,j} & \end{cases} \quad p = \begin{cases} \text{number of replicates} \\ \text{runs in which} \\ \text{the interactor is present} \end{cases}$$

$$D^R_{i,j} = \sqrt{\left(\frac{k}{\sum_{i=1}^{i=k} f_{i,j}} \right)^p} X_{i,j} \quad (\text{Eq. 3})$$

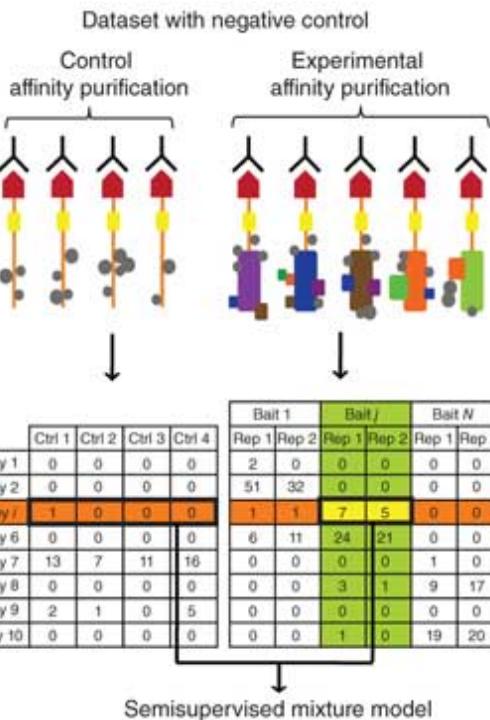


D^T = D-score threshold below which 95% of simulated data falls

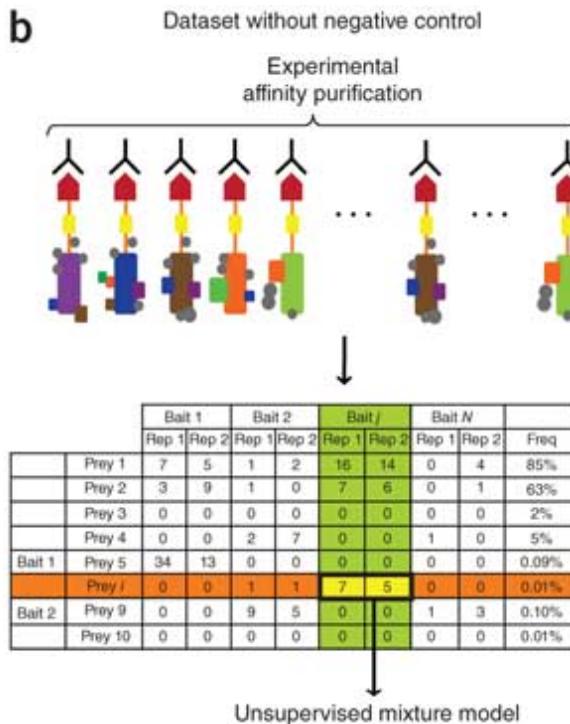
$$D^N_{i,j} = D^R_{i,j} / D^T \quad (\text{Eq. 4})$$

Protein Complexes - specific/non-specific binding

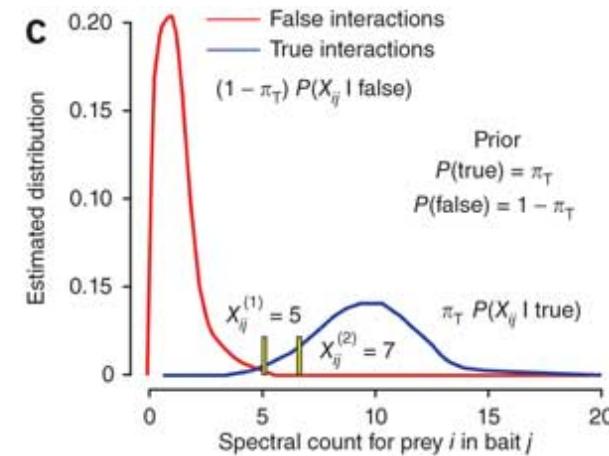
a



b



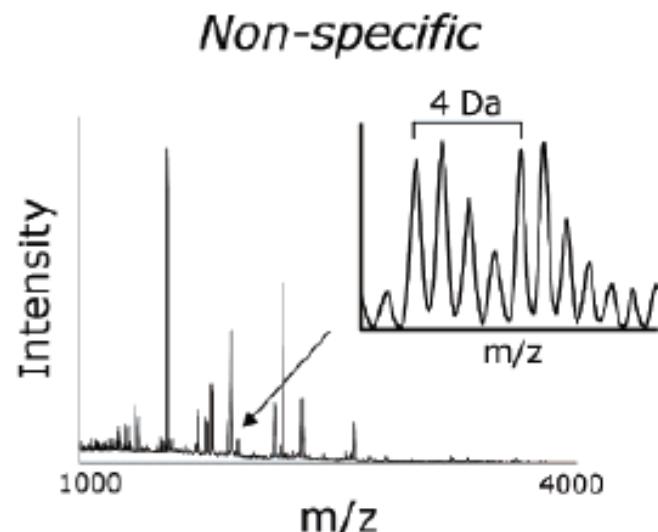
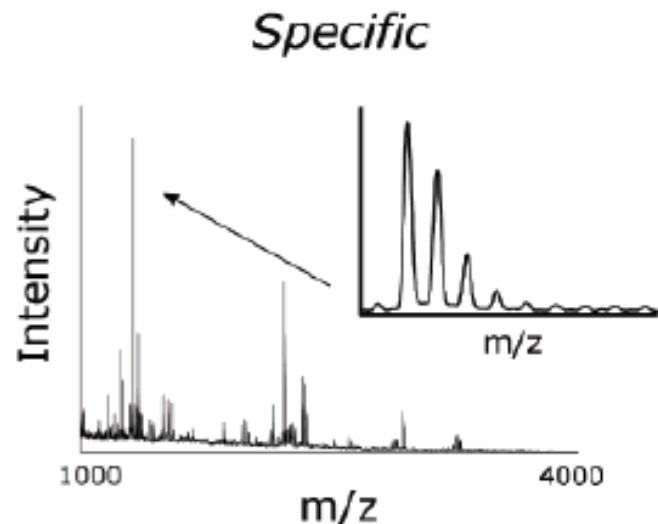
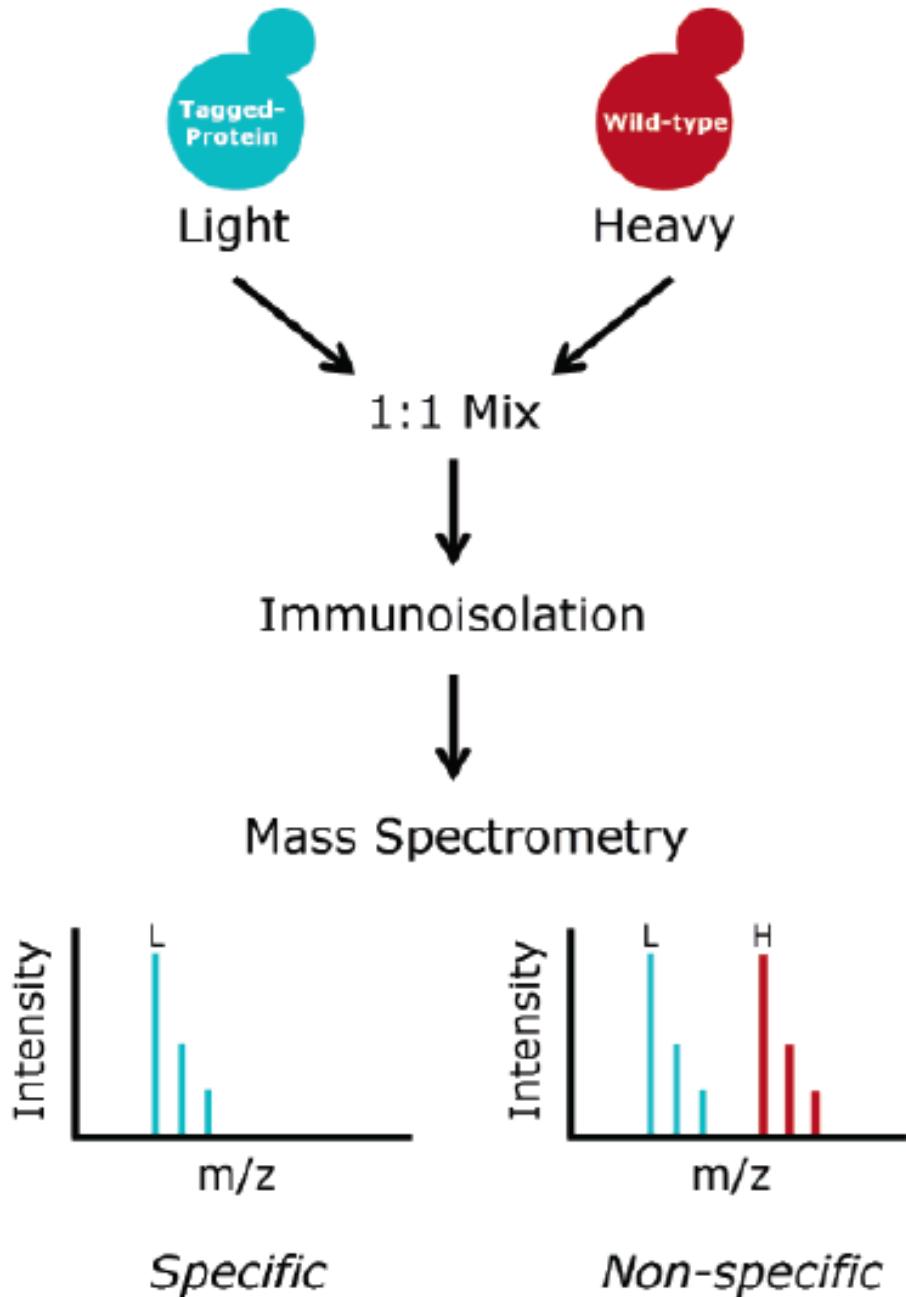
c



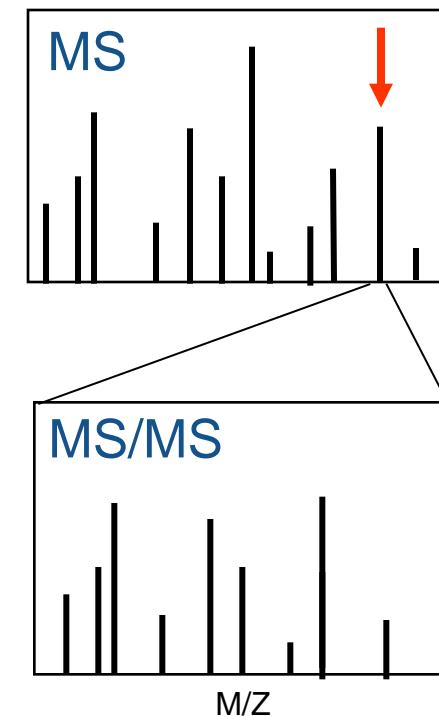
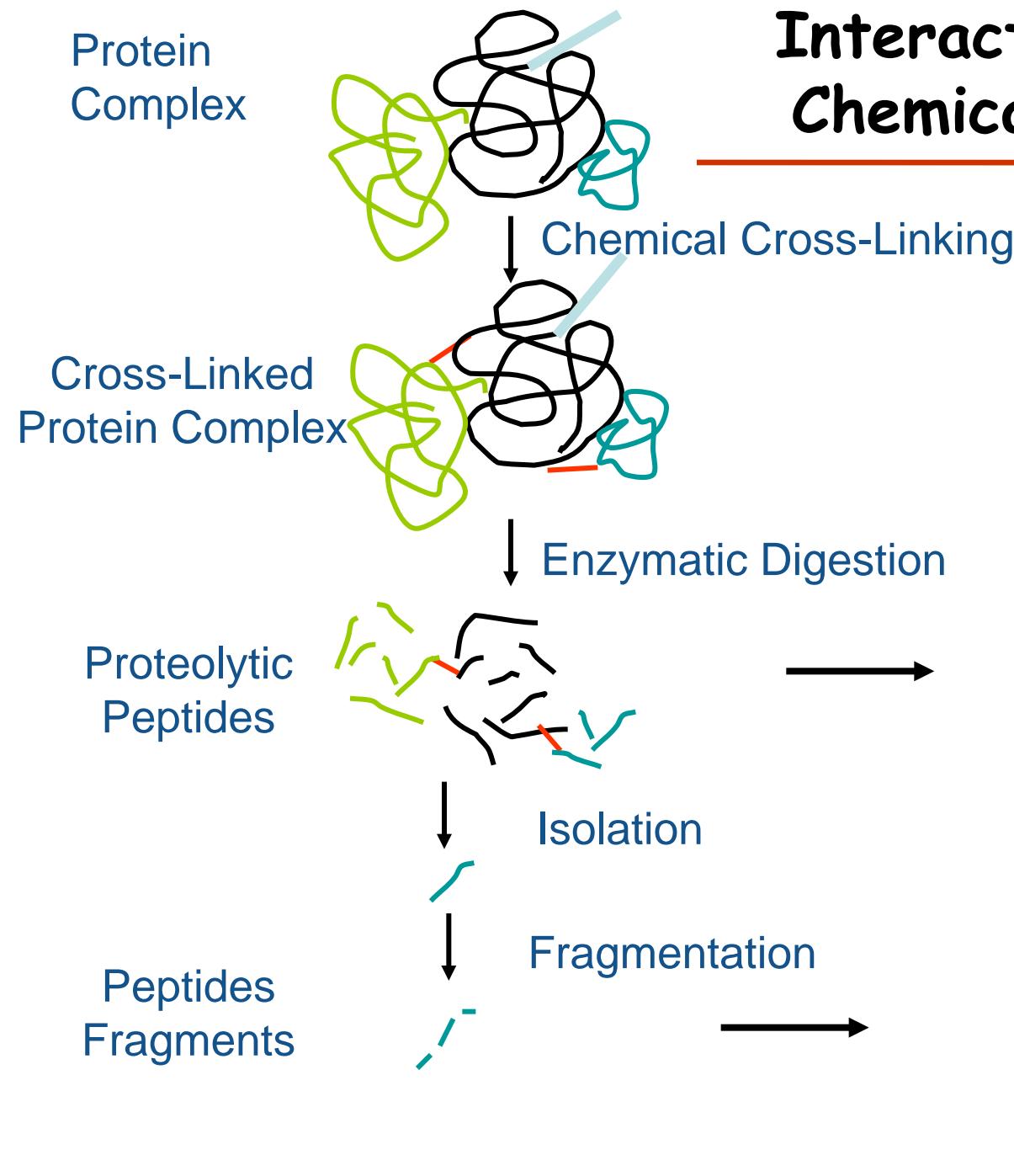
d

- Probability for interactions in individual replicates by Bayes rule
- ↓
- Summary probability for unique interaction pair
- ↓
- Select probability threshold
- Estimate false discovery rate

Protein Complexes - specific/non-specific binding

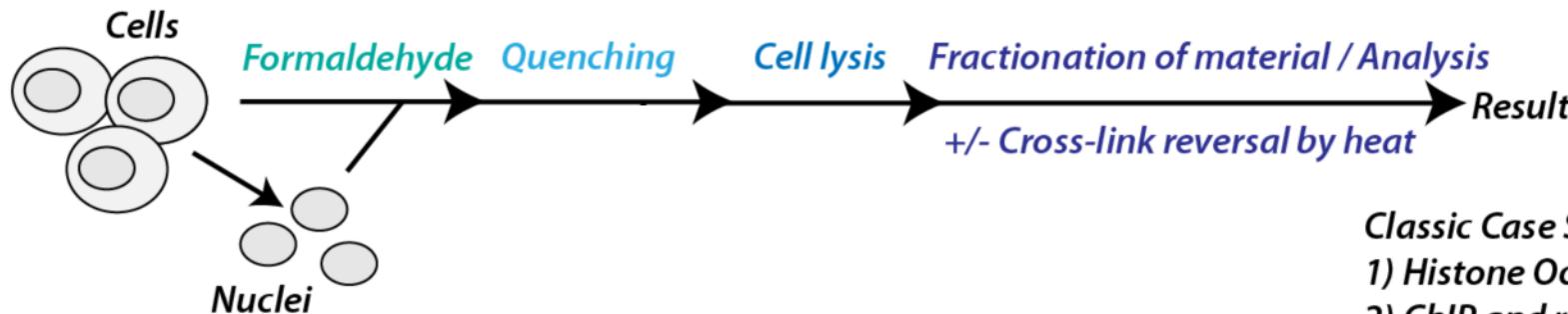


Interaction Partners by Chemical Cross-Linking



Protein Crosslinking by Formaldehyde

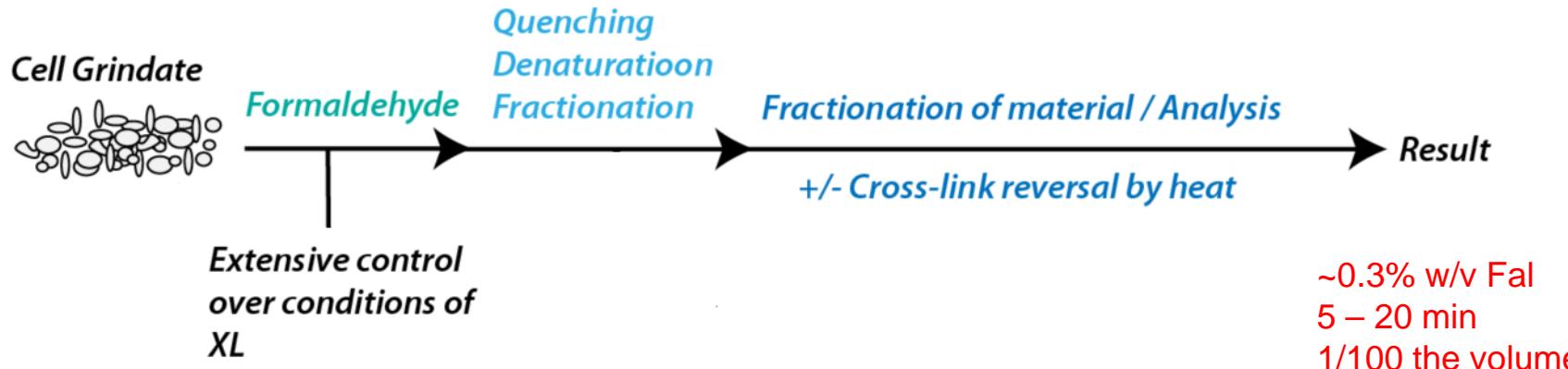
Example Traditional Formaldehyde Cross-linking:



Classic Case Studies:
1) Histone Octamer
2) ChIP and related methods

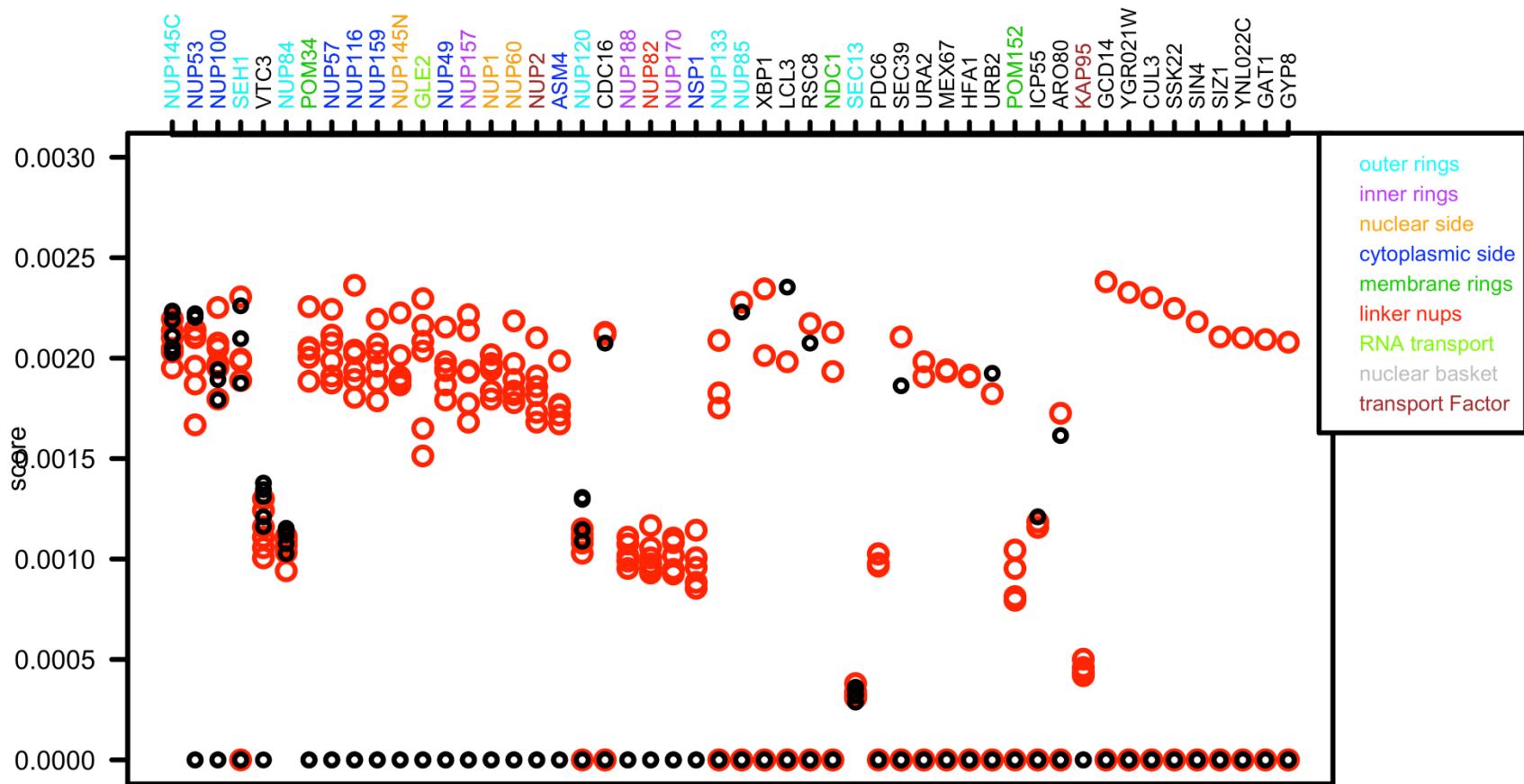
~1% w/v Fal
20 – 60 min

Example Modified Approach:



~0.3% w/v Fal
5 – 20 min
1/100 the volume

Protein Crosslinking by Formaldehyde

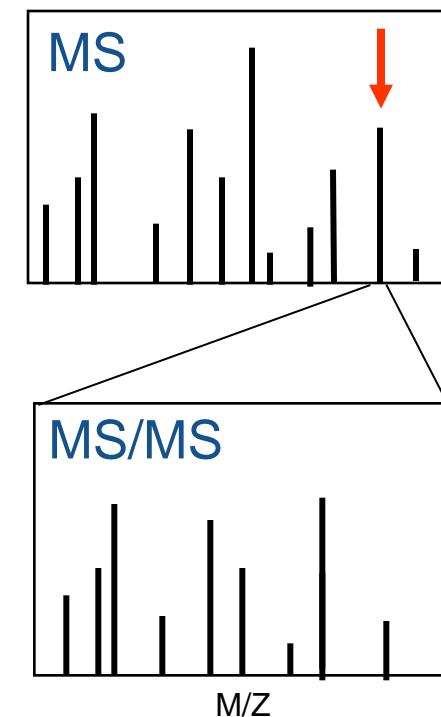
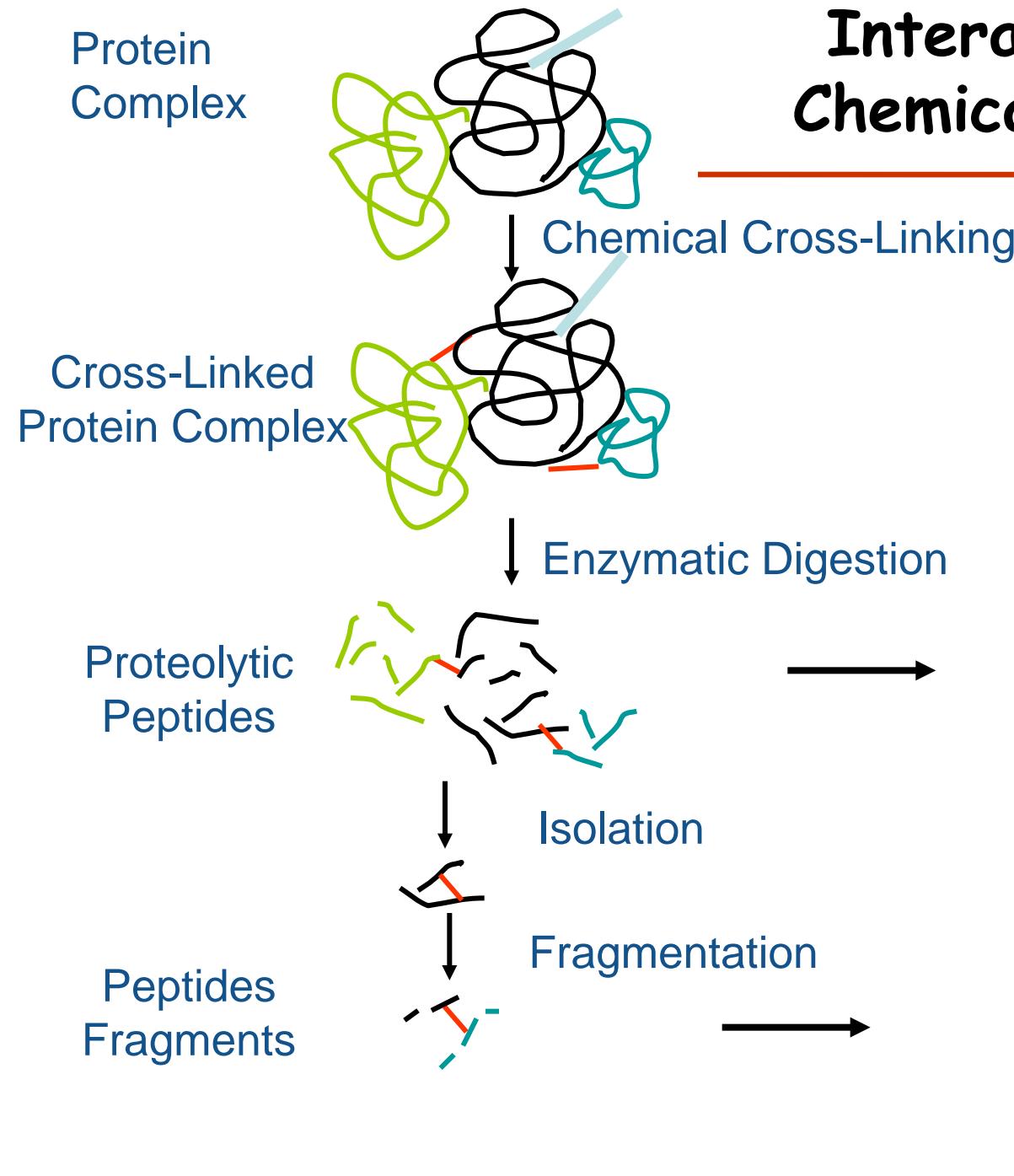


RED: Formaldehyde crosslinking

BLACK: No crosslinking

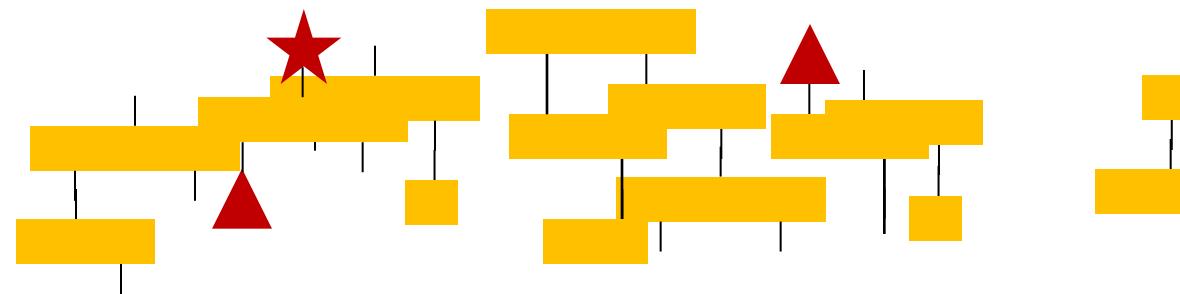
SCORE: Log Ion Current / Log protein abundance

Interaction Sites by Chemical Cross-Linking



Cross-linking

protein



n peptides with reactive groups



$(n-1)n/2$ potential ways to cross-link peptides pairwise
+ many additional uninformative forms

Cross-linking

Mass spectrometers have a limited dynamic range and it is therefore important to limit the number of possible reactions not to dilute the cross-linked peptides.

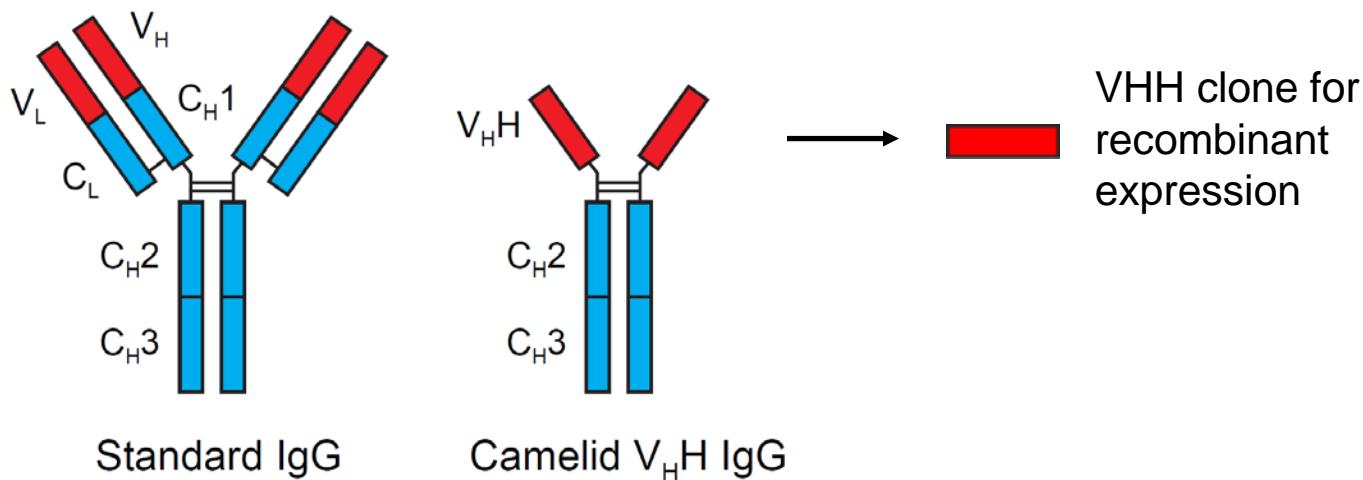
For identification of a cross-linked peptide pair, both peptides have to be sufficiently long and required to give informative fragmentation.

High mass accuracy MS/MS is recommended because the spectrum will be a mixture of fragment ions from two peptides.

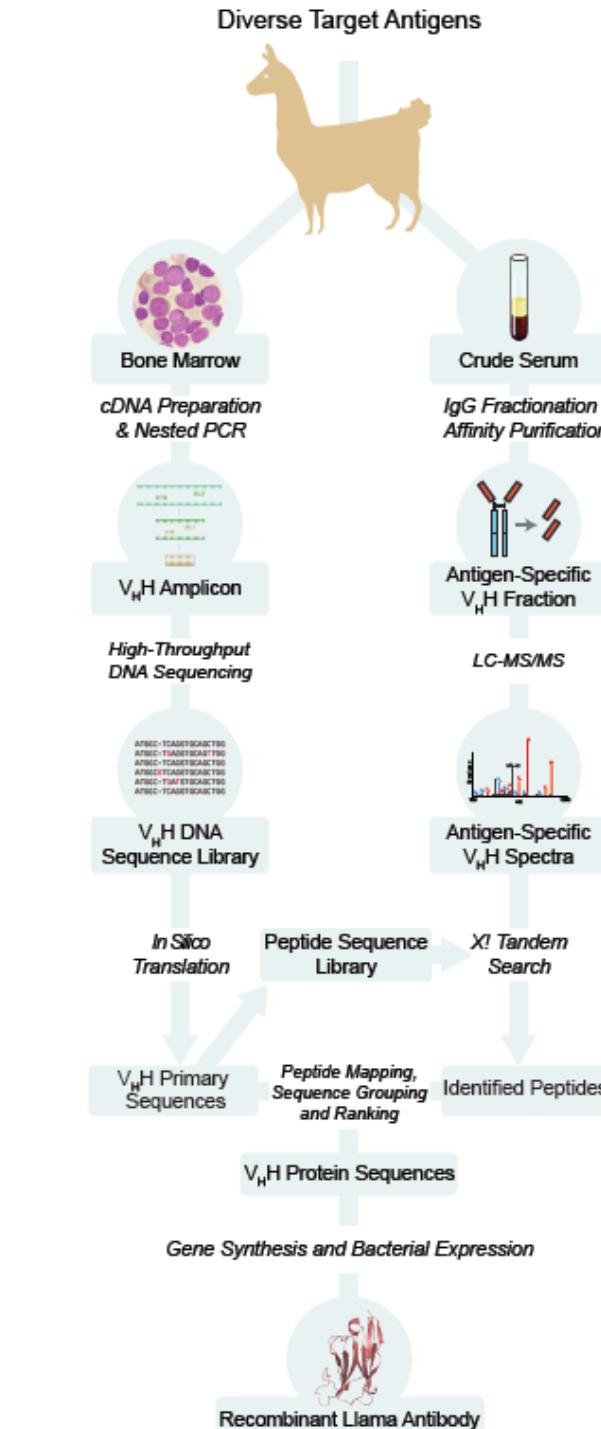
Because the cross-linked peptides are often large, CAD is not ideal, but instead ETD is recommended.

Cloning nanobodies for GFP pullouts

- Atypical heavy chain-only IgG antibody produced in camelid family – retain high affinity for antigen without light chain
- Aimed to clone individual single-domain VHH antibodies against GFP – only ~15 kDa, can be recombinantly expressed, used as bait for pullouts, etc.
- To identify full repertoire, will identify GFP binders through combination of high-throughput DNA sequencing and mass spectrometry



Cloning Llamabodies for GFP pullouts

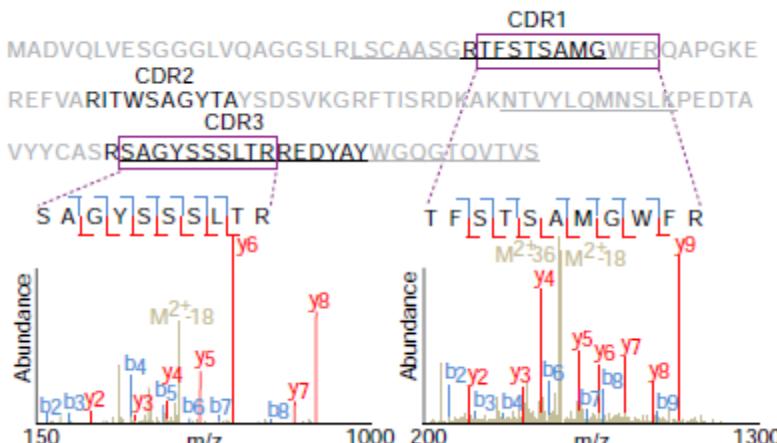


Identifying full-length sequences from peptides

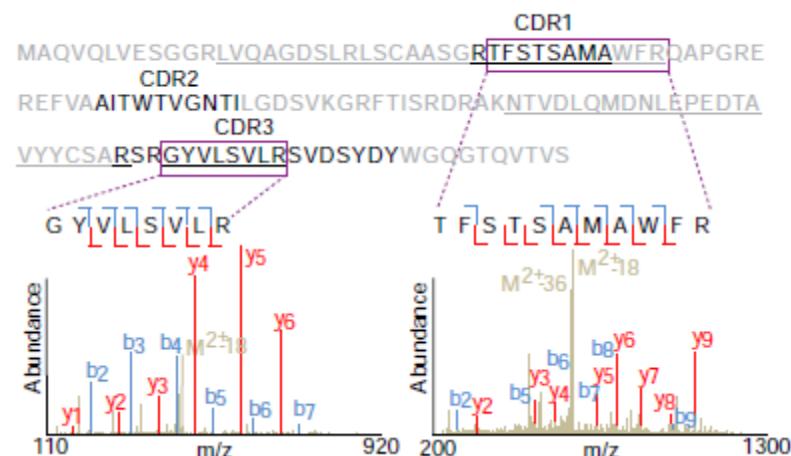
Rank of the group
Click to expand the group for viewing more sequences
This group contains 43 sequences
Click to expand the sequence for peptide mapping information
Sequence name
2 ▶ (43) ▶ HOKO31K01DSD3D_rev_fro CDR1: 100.0% (9/9); CDR2: 100.0% (10/10); CDR3: 100.0% (14/14); combined CDR: 100.0% (33/33); overall: 75.8% (94/124); HT-seq count: 10
MAQVQLVESGGGLVQAGGSLRLSCAASGRRTFSGYAMGWFRQTPGREREAVAAITWAHSTYYSDSVKDRFTISIDNTRNTGYLQMNSLKPEDTAVYYCTVRHGTWFTTSRYWTDWGQQTQVIVS
-1.5
-5.0
-4.2
-1.6
-8.7
-2.2
-3.6
The best score (log(e)) of this peptide is -4.2
The sequence of the candidate with CDR regions highlighted in red and areas mapped by MS identified peptides underlined.
Ranking metrics: CDR coverage percentages, overall coverage percent and "HT-seq count" (a count of the number of HT DNA sequencing reads that produced this sequence)
Click to expand for all matched sequences
Click to expand for MS spectra information

The metrics following the peptide are (a) 'uniqueness' score, which is 1.0 in this case; (b) spectra count from X!Tandem, which is 6 in this case; (c) number of sequences matched across the whole sequence database, which is 43 in this case.

LaG-9



LaG-16

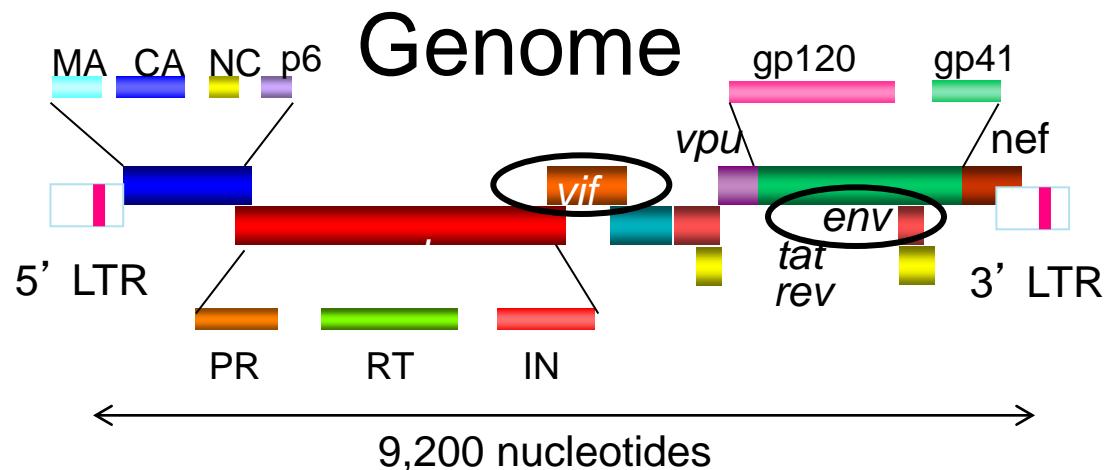
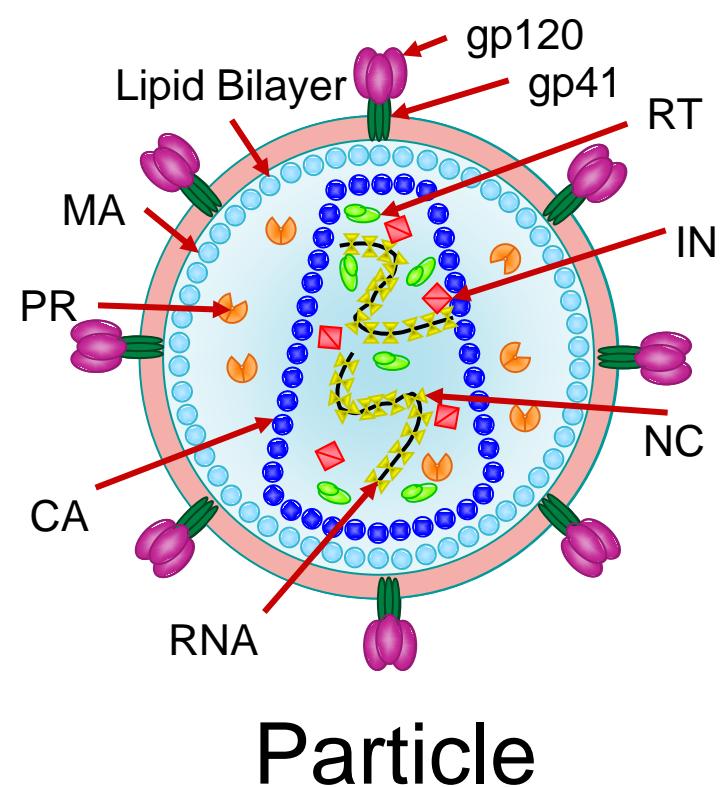


Sequence diversity of 26 verified anti-GFP nanobodies

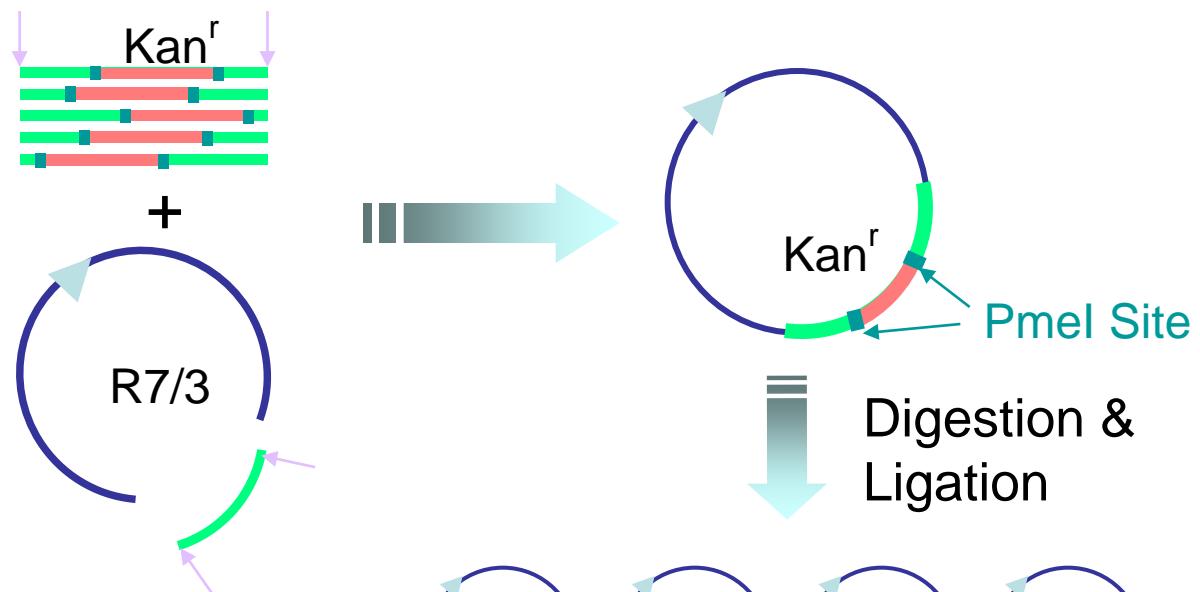
- Of ~200 positive sequence hits, 44 high confidence clones were synthesized and tested for expression and GFP binding: 26 were confirmed GFP binders.
- Sequences have characteristic conserved VHH residues, but significant diversity in CDR regions.

	FR1	CDR1	FR2	CDR2	FR3	CDR3	FR4
1	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY						
02p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
03p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
05p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
06p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
07p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
08p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
09p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
10p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
11p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
12p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
14p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
16p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
17p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
18p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGRTFSNYAMGWFRQAPGKEREFVAAISWTVGVSTY					
19p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					
21p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					
24p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					
26p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					
27p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					
29p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					
30p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					
35p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					
37p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					
41p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					
42p	MAQVQLVE	SGGGLVQAGGSGLRLSCAAASGPTG..AMAWFROAPGKEREFVAAISWTVGVSTY					

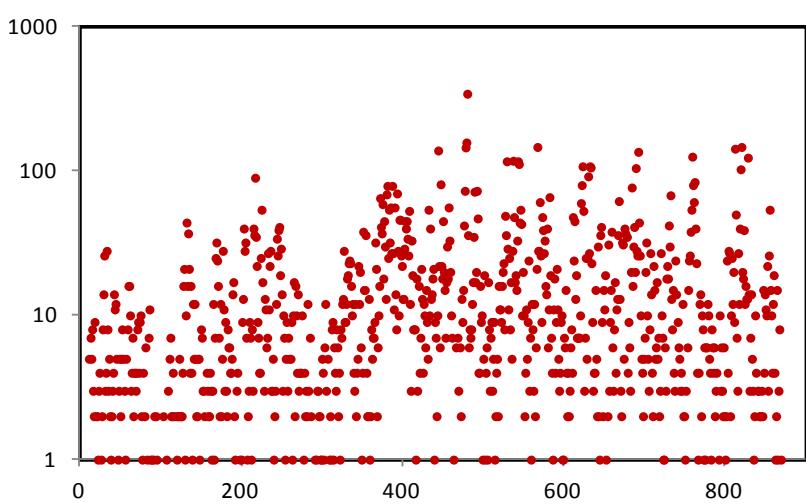
HIV-1



Random Insertion of 5 Amino Acids in Proviral DNA Clone

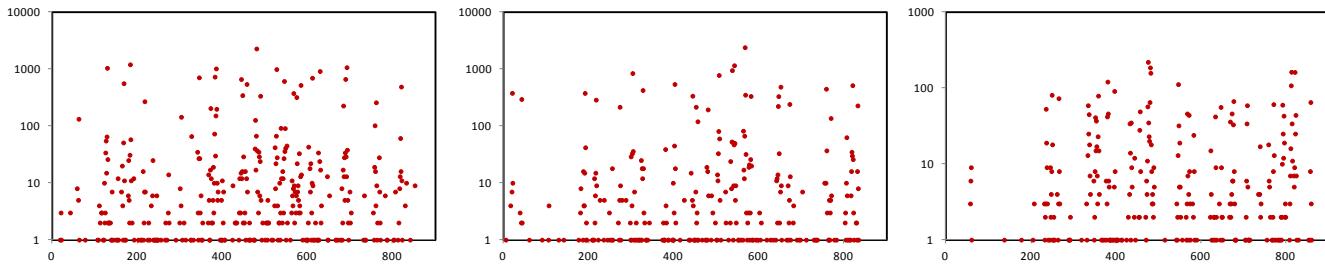


Random insertion of 5 amino acids (PmeI)
within specific viral coding region

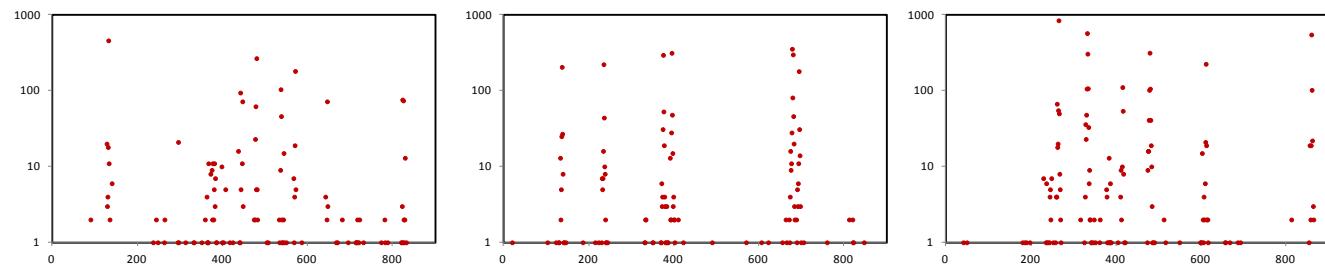


Fitness Landscape of Targeted Viral Segment

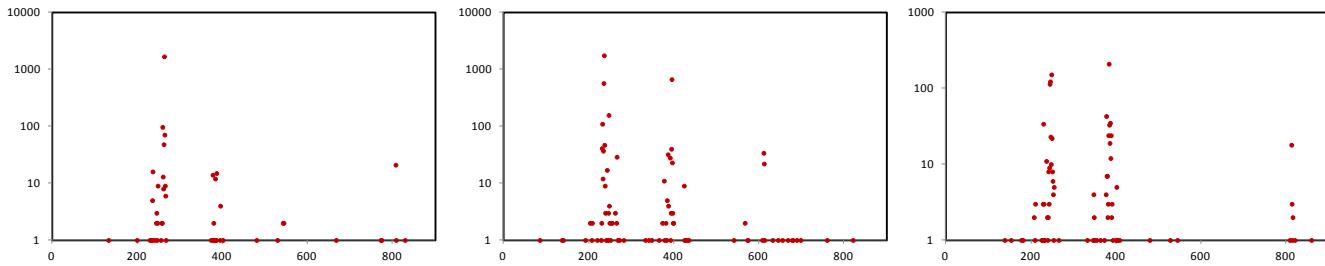
Day 1



Day 3

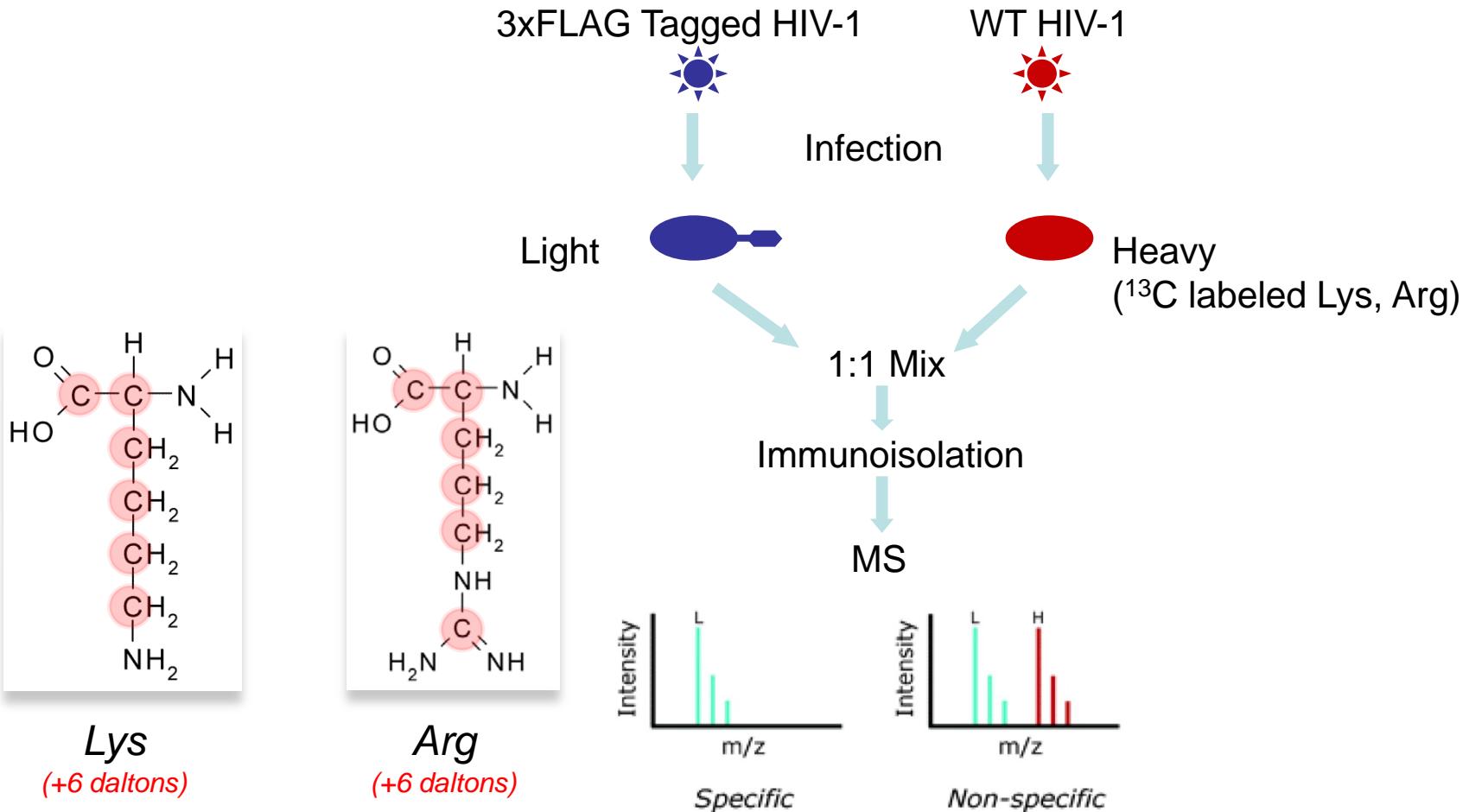


Day 6



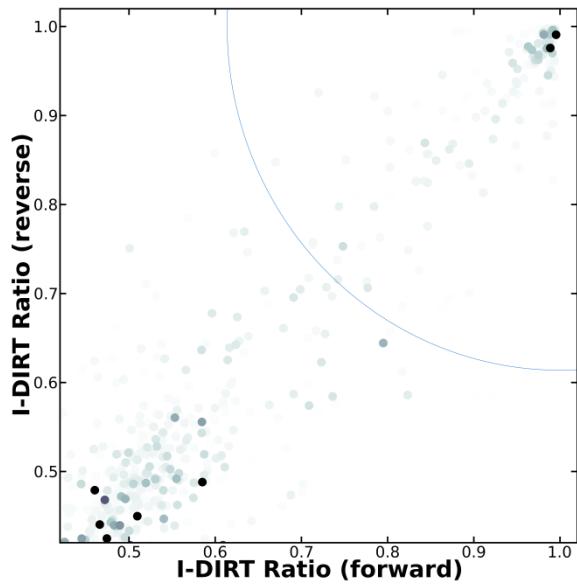
Specific and Non-Specific Interactors

I-DIRT = Isotopic Differentiation of Interactions as Random or Targeted

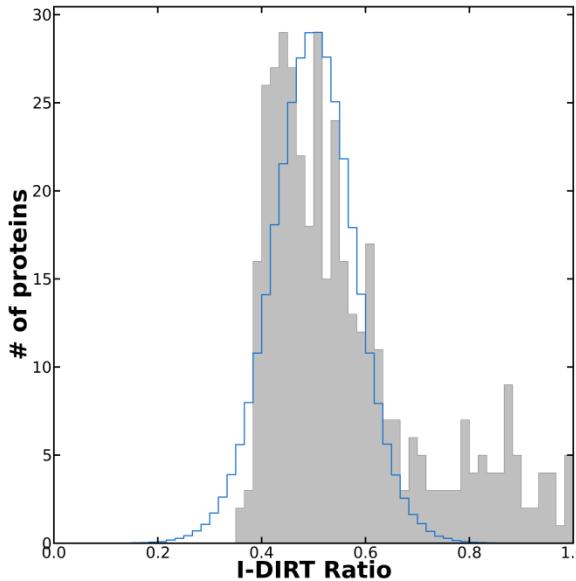
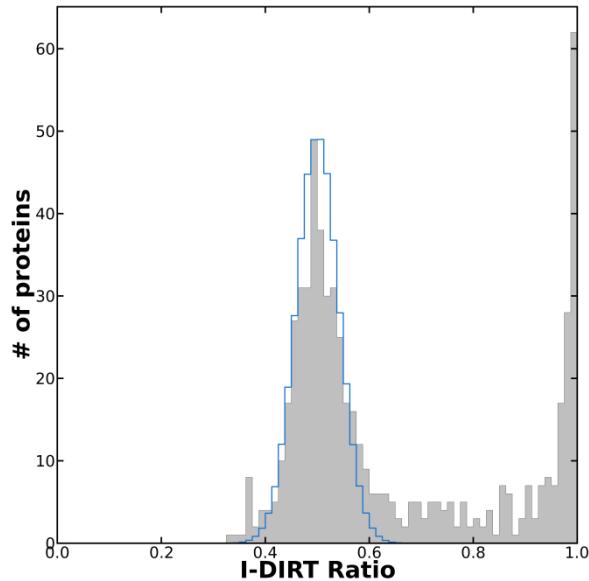
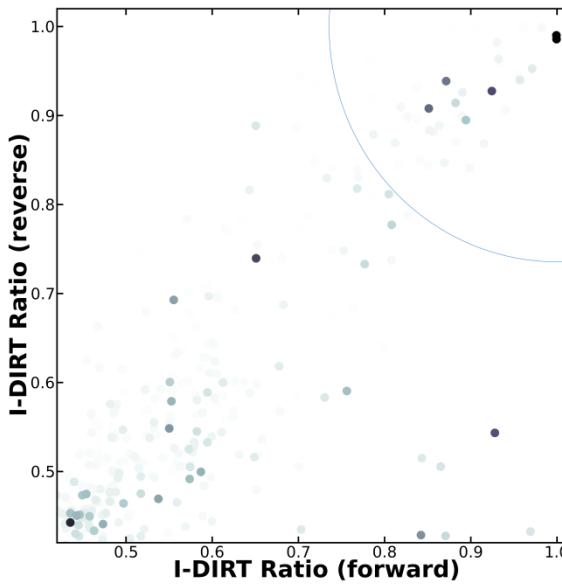


Specific and Non-Specific Interactors

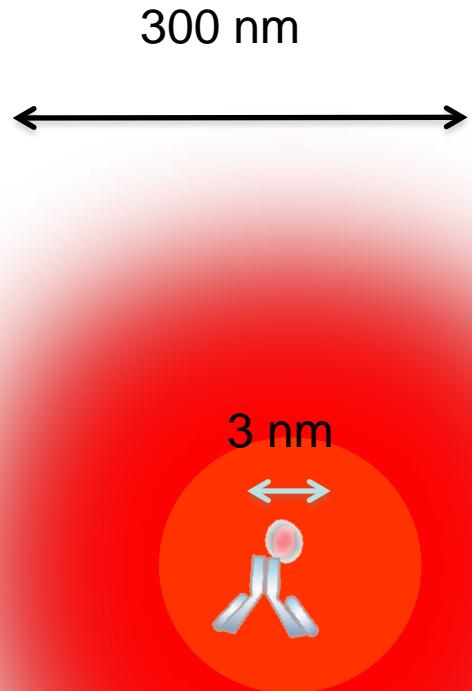
Env-3xFLAG



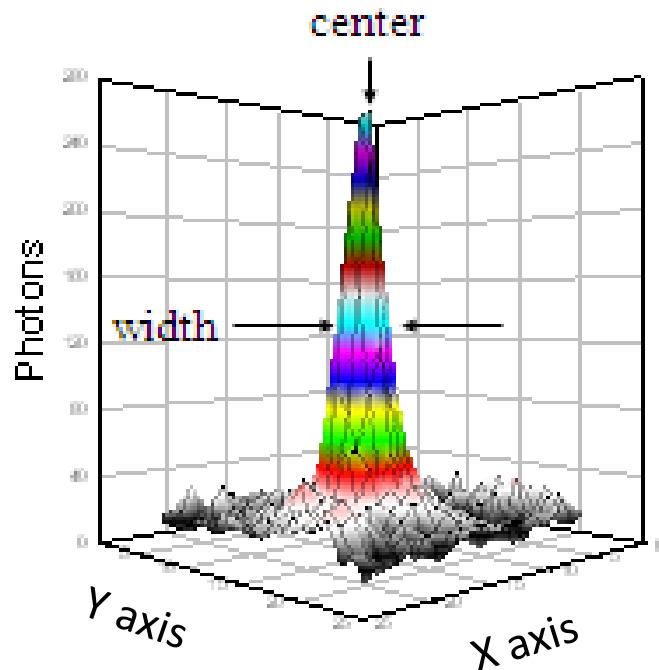
Vif-3xFLAG



Limitation of Light Microscopy



Fluorescent Imaging with One Nanometer Accuracy (FIONA)



CCD image of a single Cy3 molecule:

Width \sim 250nm

Center is localized within width/(S/N)

$$(S/N)^2 \sim N$$

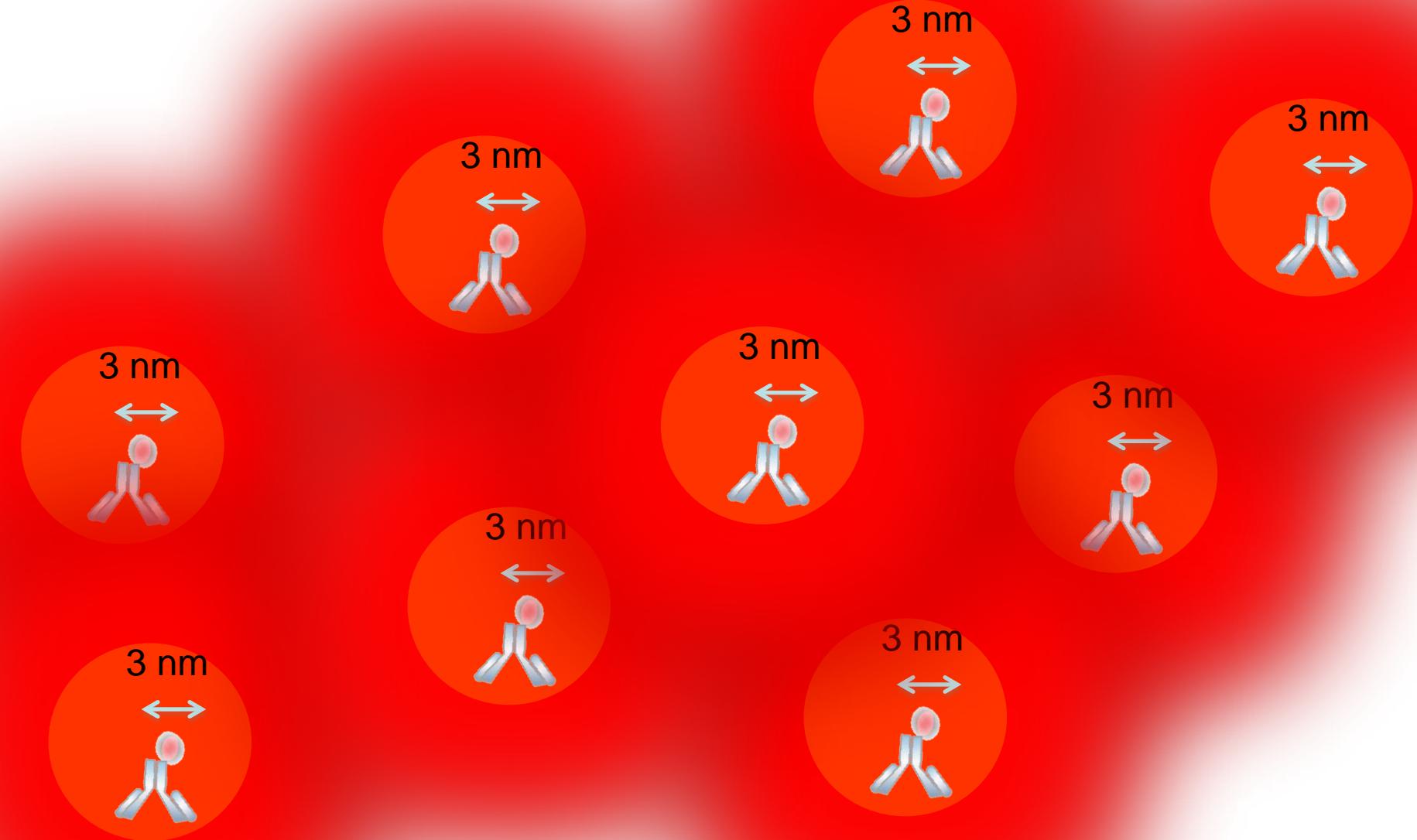
N = total # photon

(for $N \sim 10^4$ center within ~ 1.3 nm)

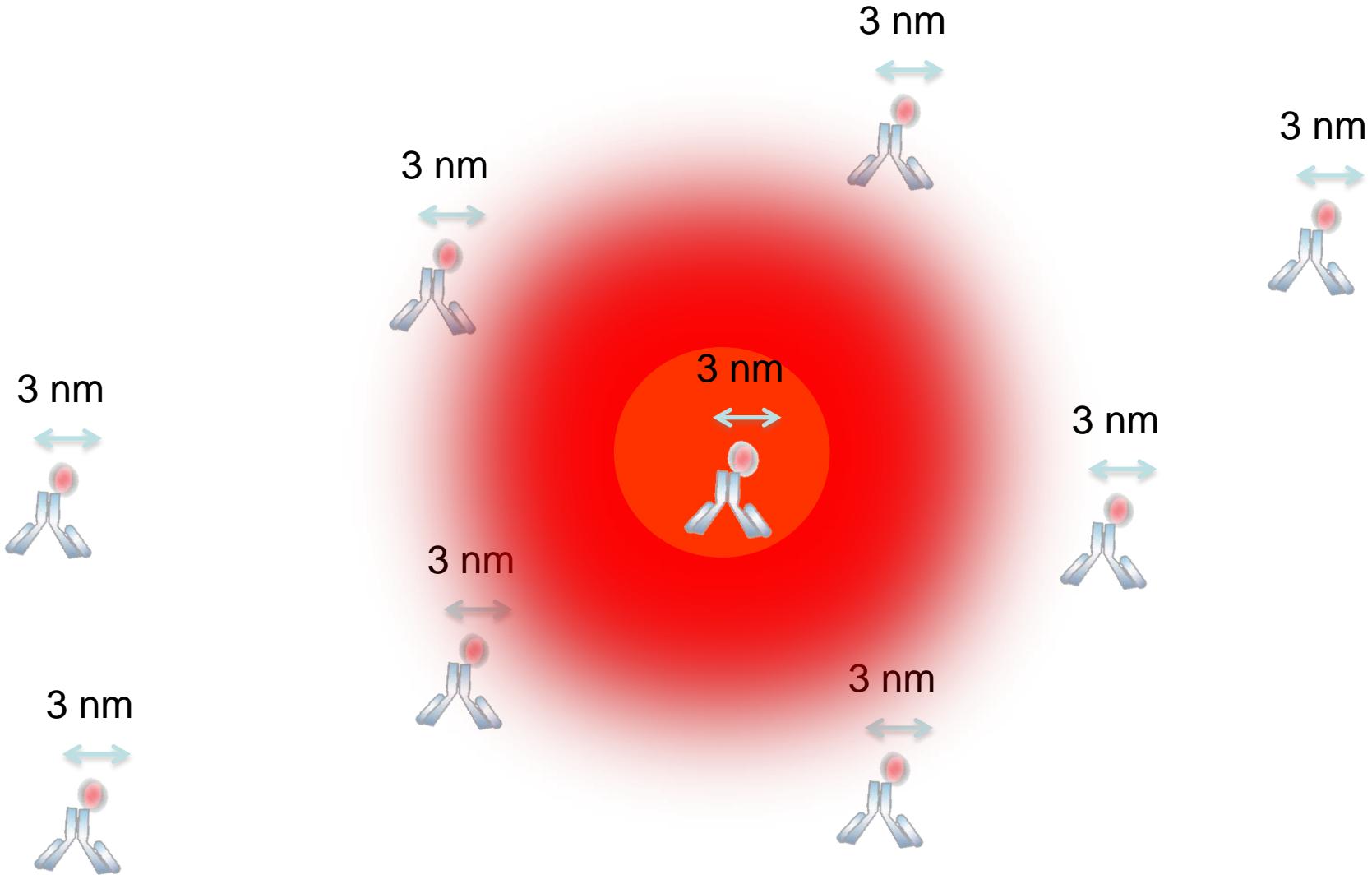
Yildiz et al, Science 2003.

Paul Selvin

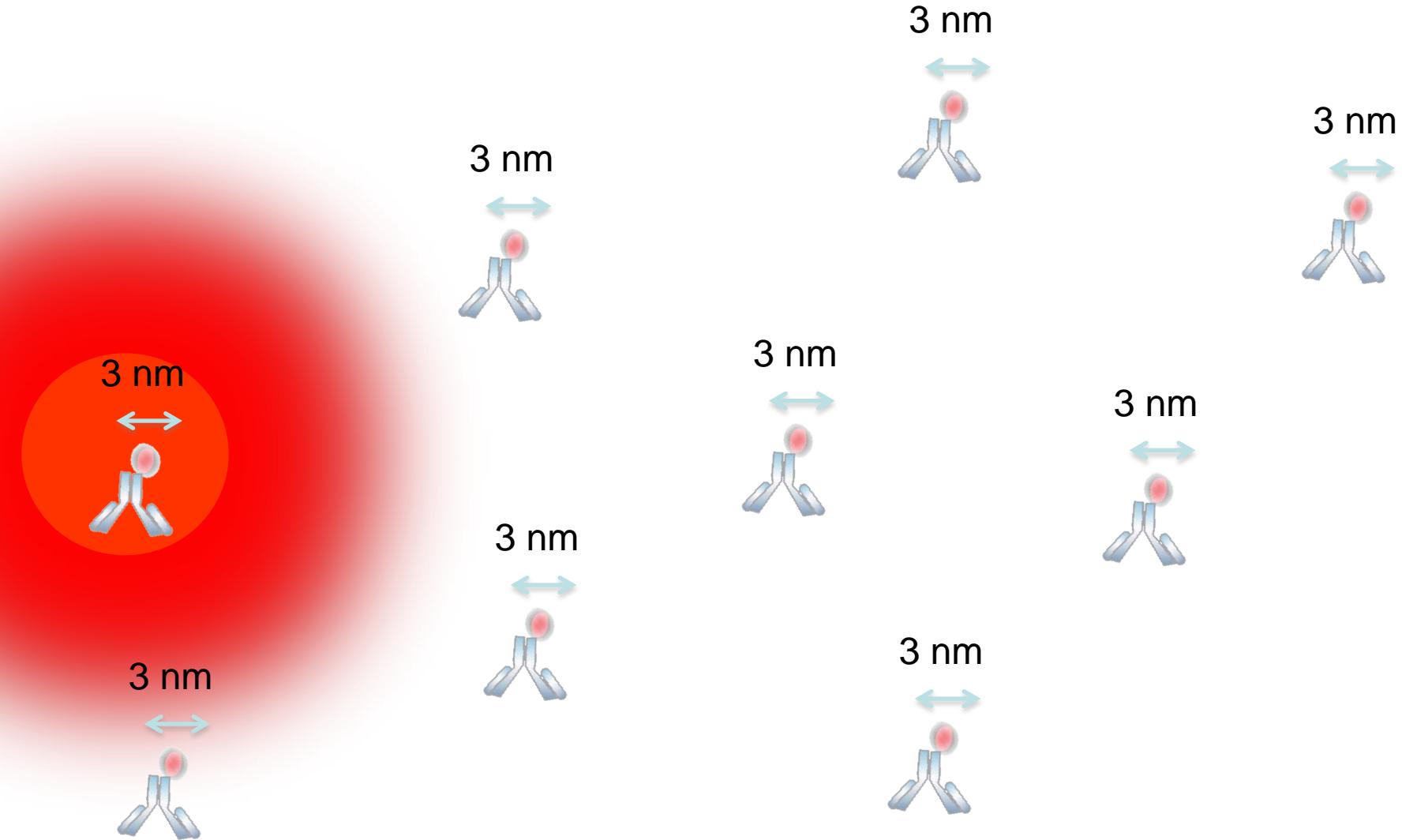
Limitation of Light Microscopy



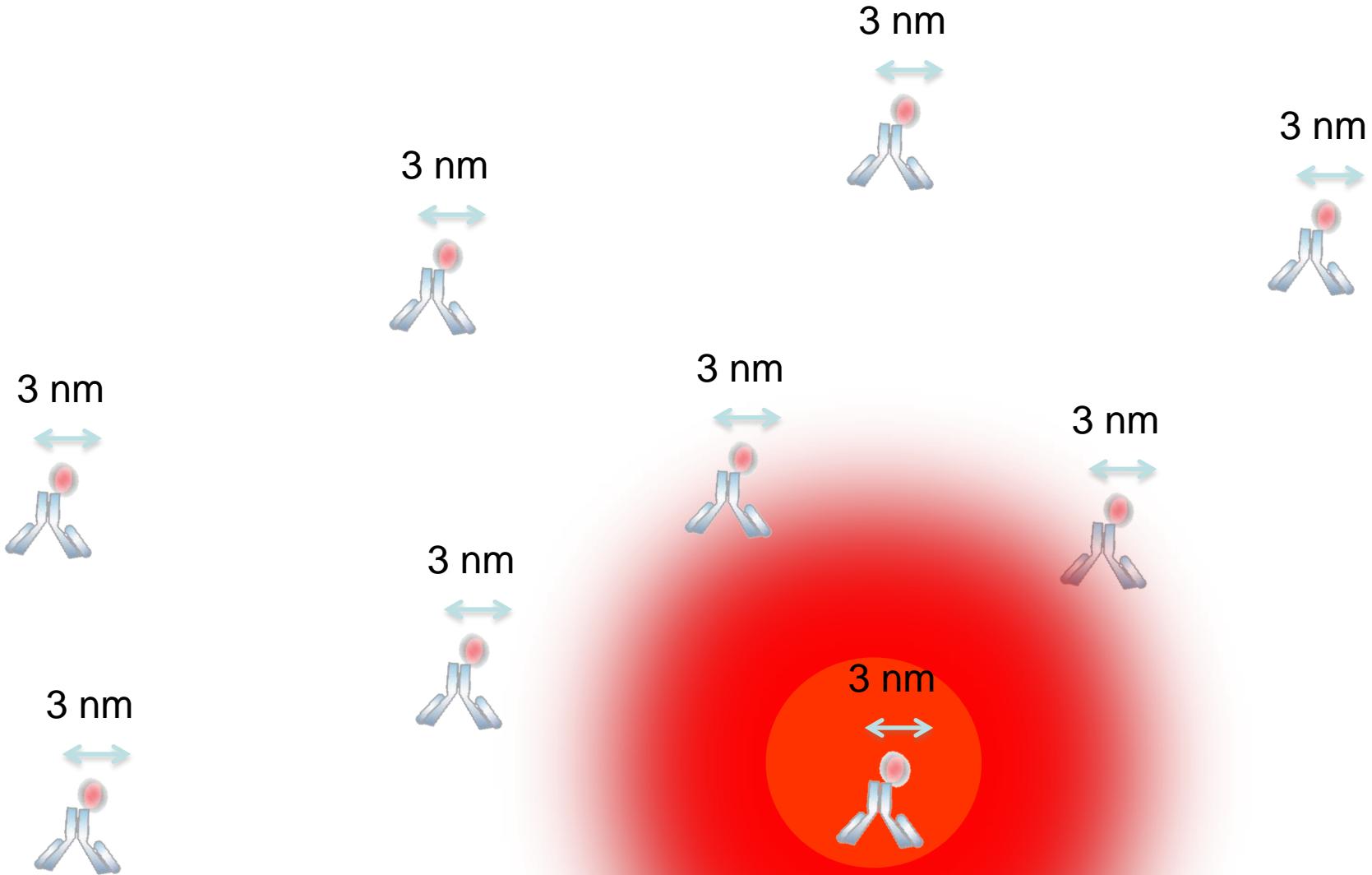
Limitation of Light Microscopy



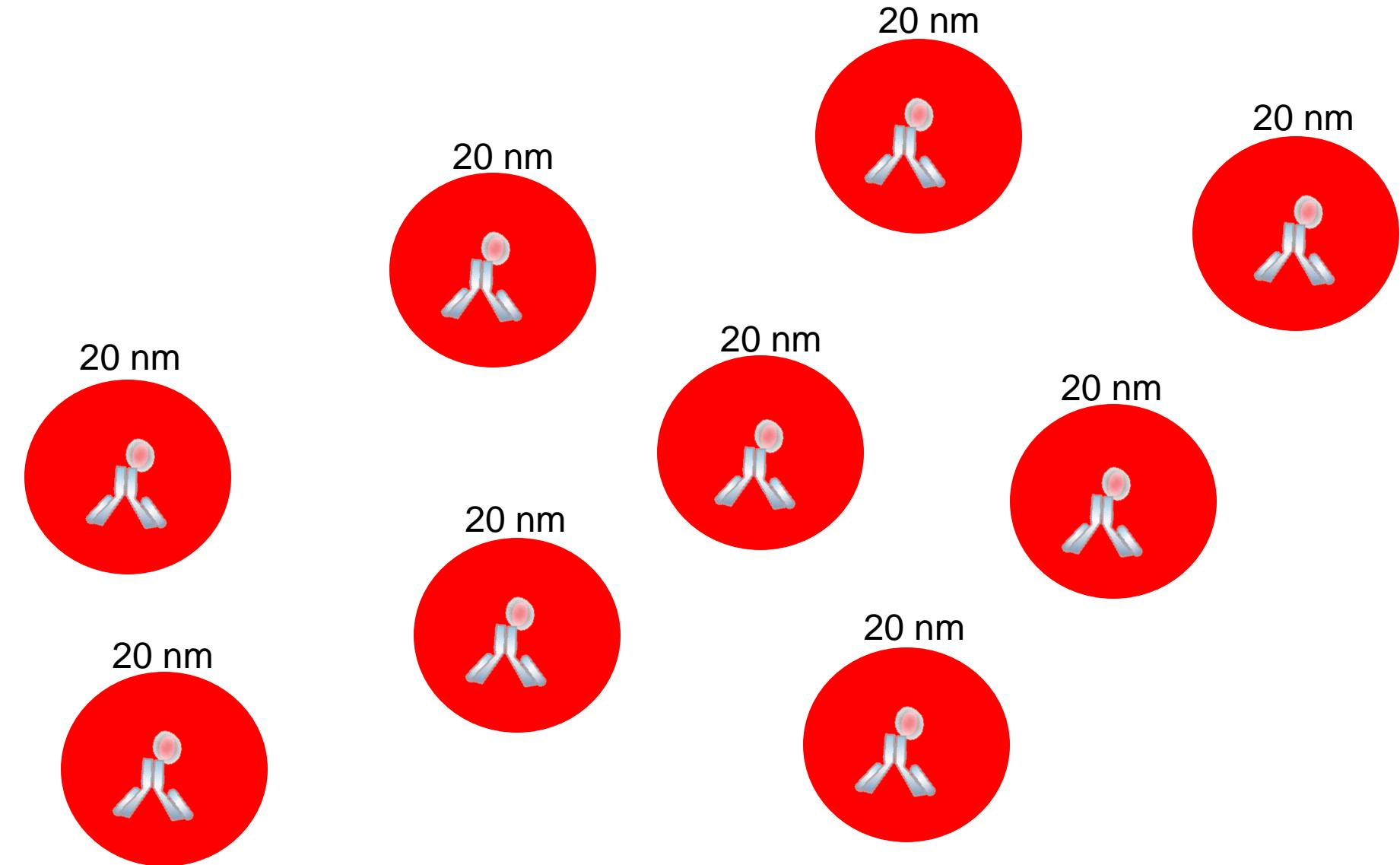
Limitation of Light Microscopy



Limitation of Light Microscopy



Limitation of Light Microscopy



Super-Resolution Localization Microscopy

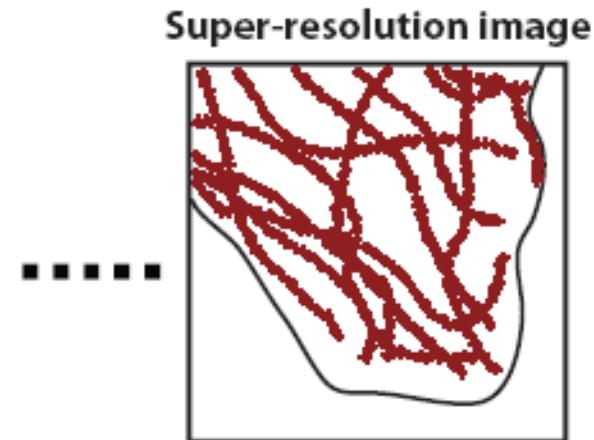
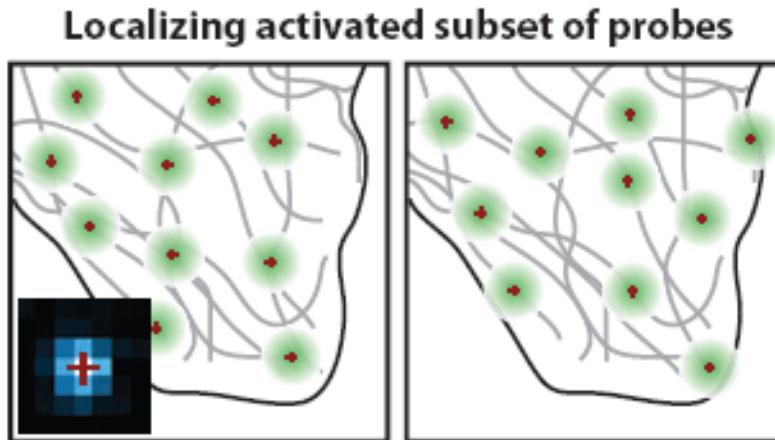
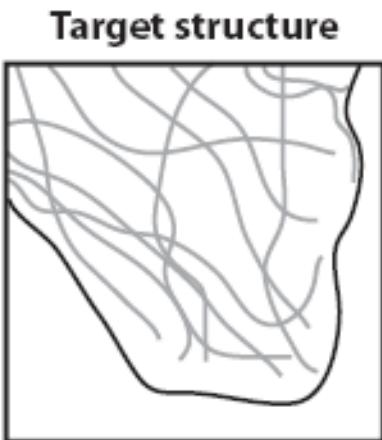
Using two lasers for interchangeable activation and excitation of probes

PALM: PhotoActivation Localization Microscopy
Using fluorescence proteins (mEOS, etc)

Betzig, 2006
Science

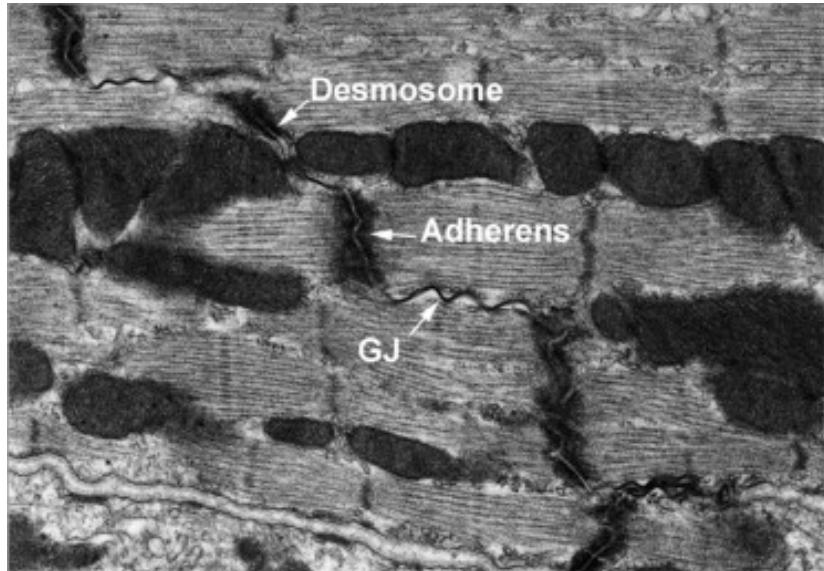
STORM: STochastic Optical Reconstruction Microscopy
Using doubly labeled (Cy3-Cy5) Ab

Bates, 2007 Science

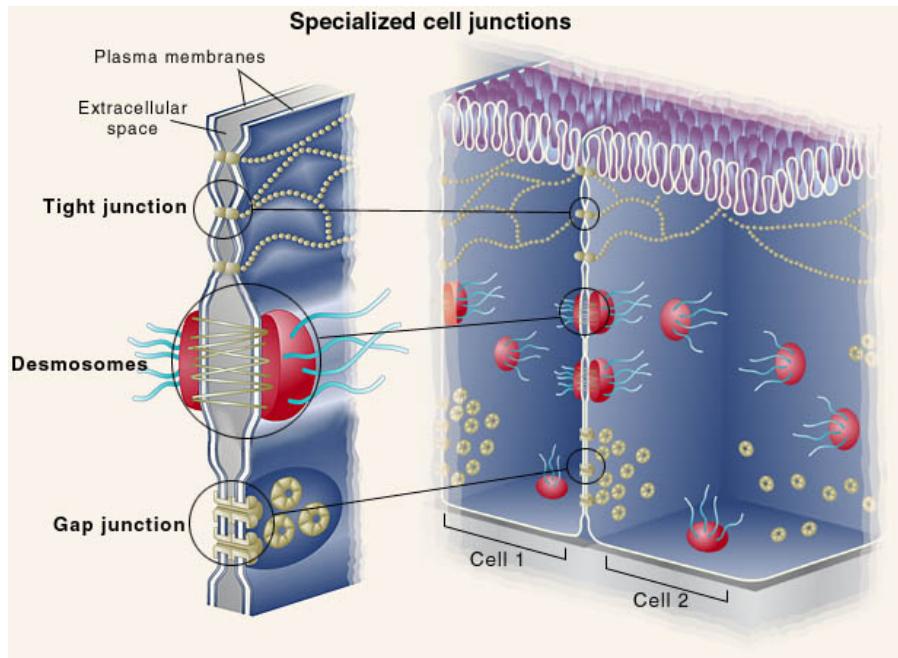


Huang, Annu. Rev. Biochem, 2009

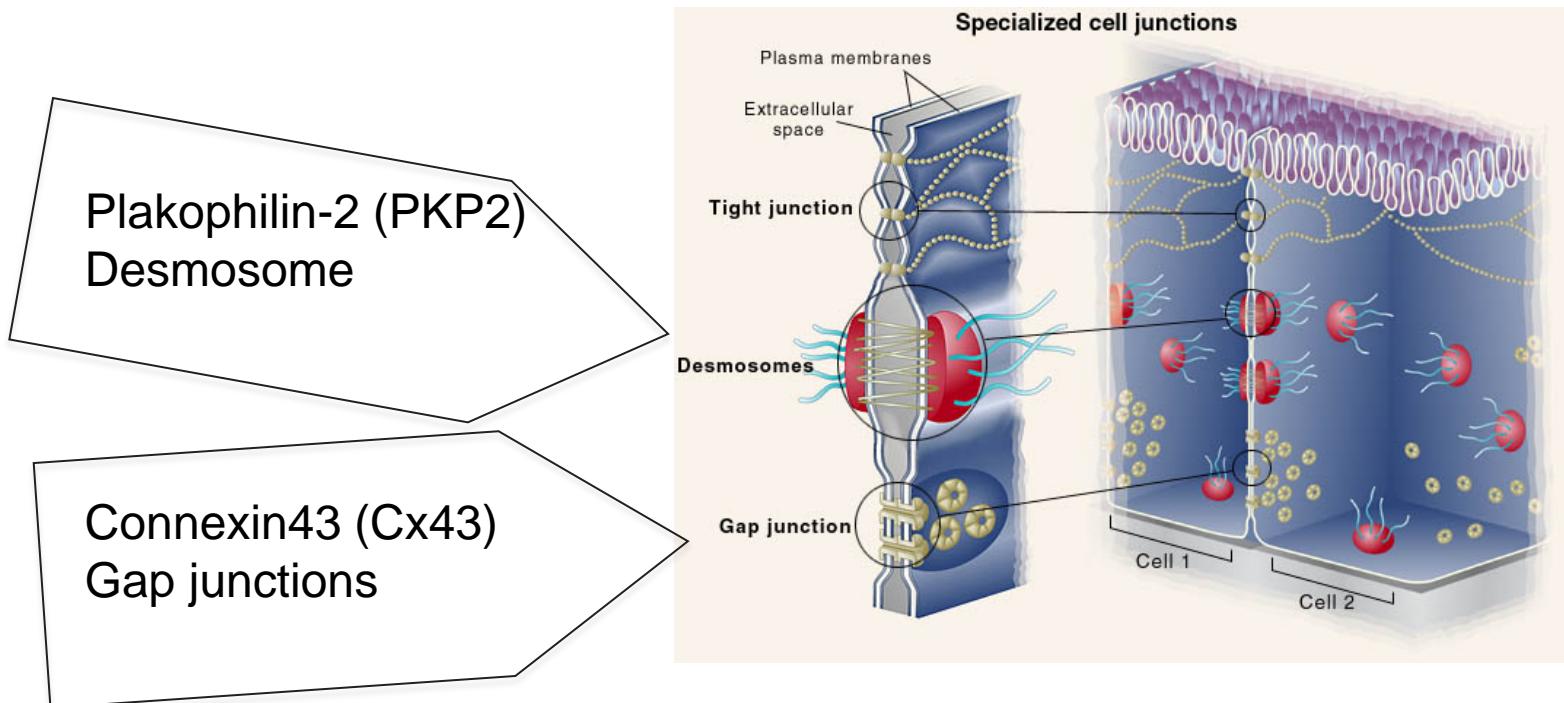
Molecular Organization of the Intercalated Disc



Saffitz, Heart Rhythm (2009)



Molecular Organization of the Intercalated Disc

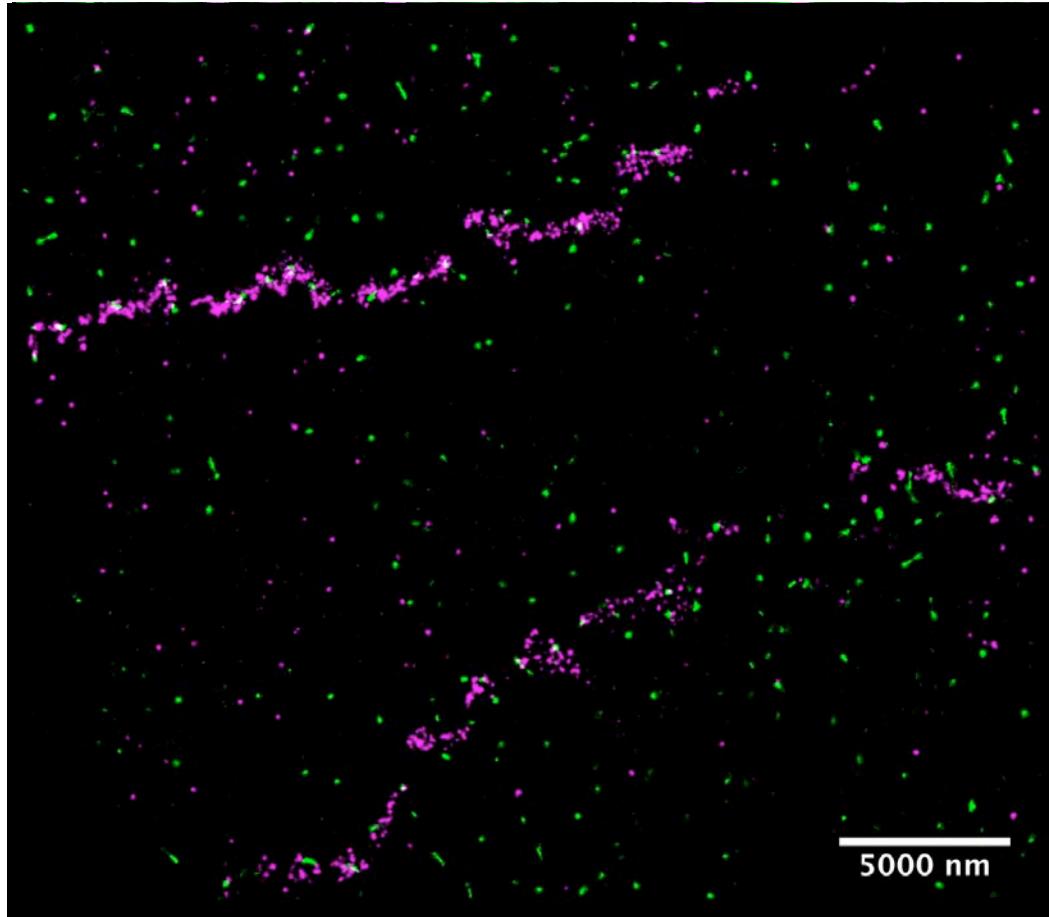


What is the interaction map of ID proteins?

Regular Microscopy v. Super-Resolution

Cx43

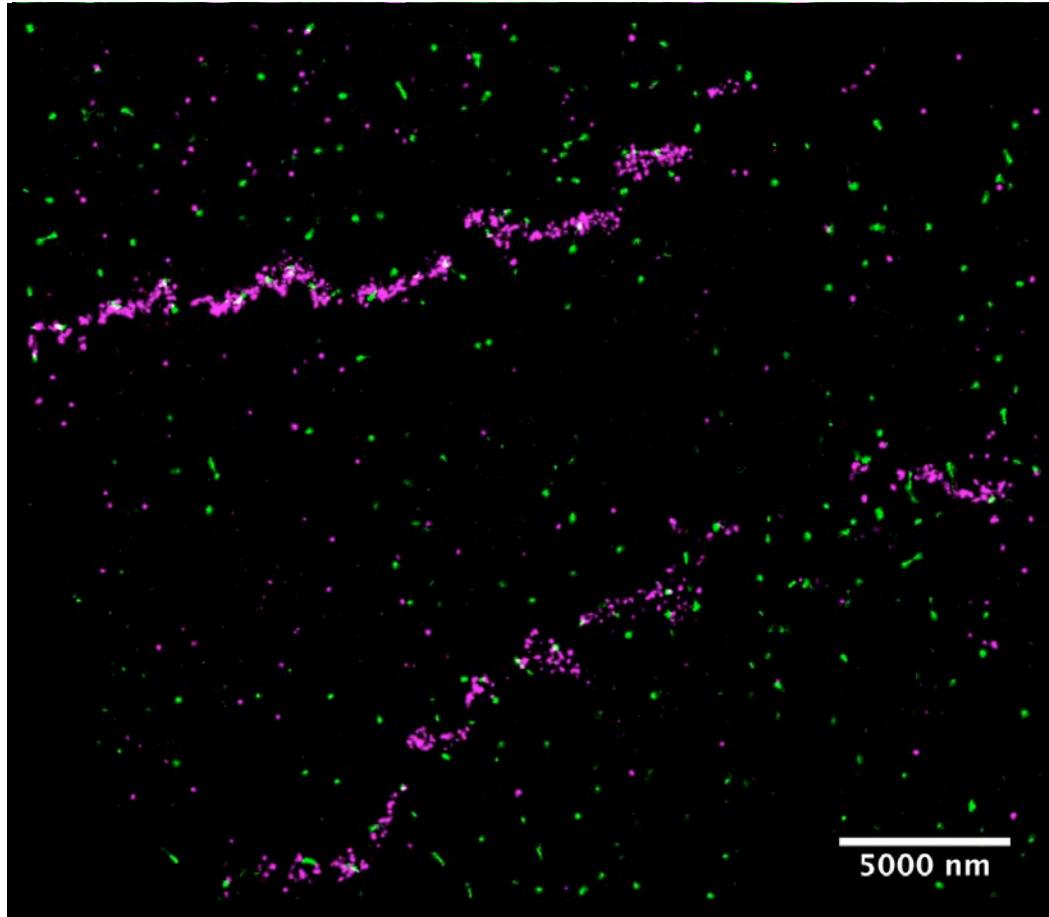
PKP2



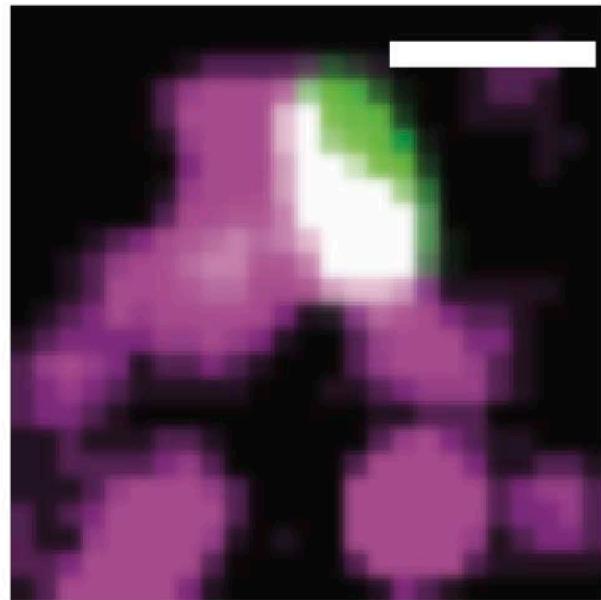
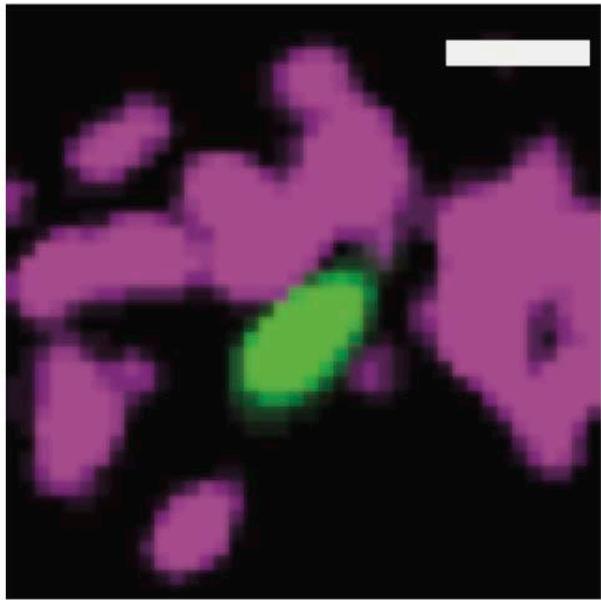
Regular Microscopy v. Super-Resolution

Cx43

PKP2



Regular Microscopy v. Super-Resolution



Cx43

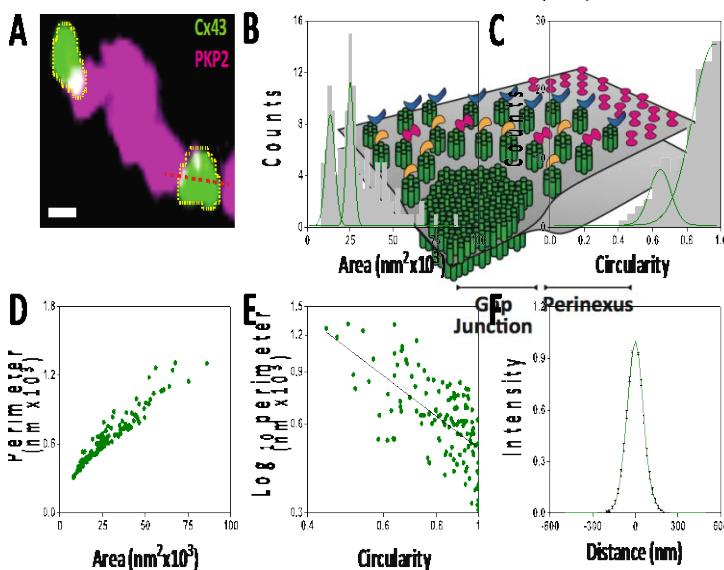
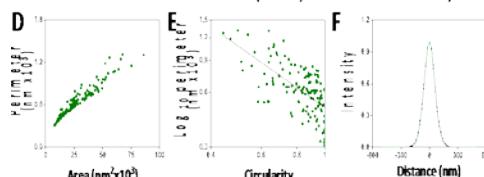
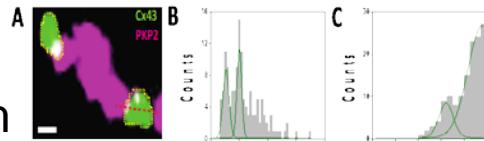
PKP2

What Do We Mean by Colocalization?

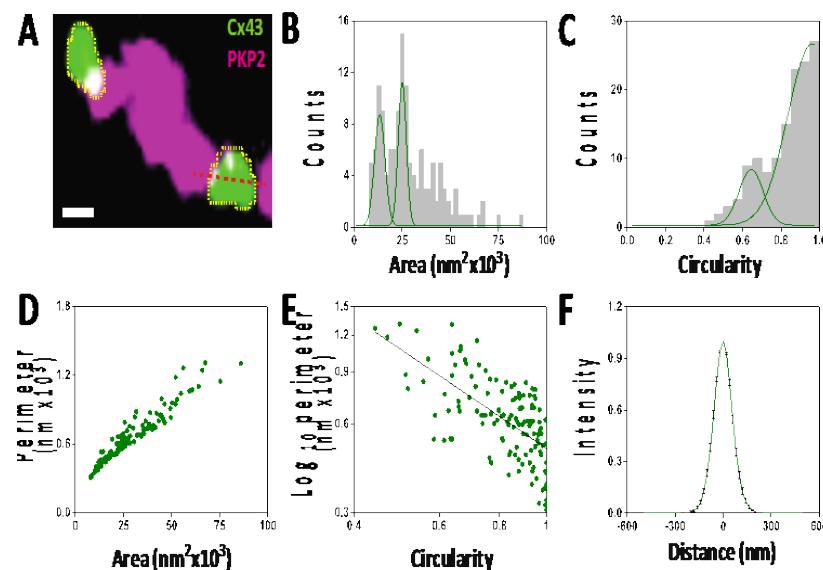


Characterization of Cx43 Clusters

Scale =200 nm



Two distinct size populations corresponding to hemi-channels and full channels.



Predominantly circular

Cx43-PKP2 Overlap Analysis

shRNA

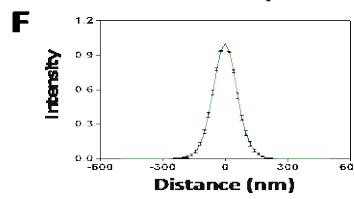
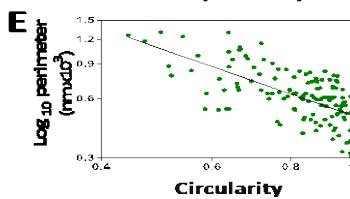
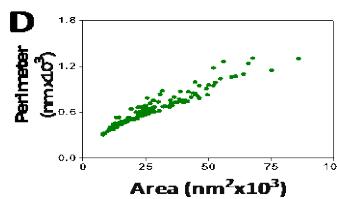
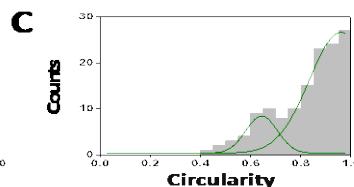
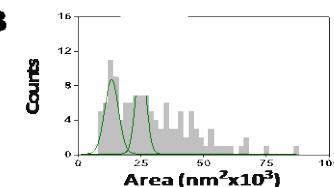
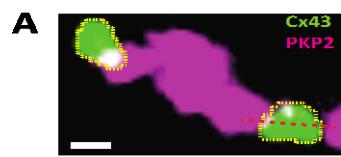
100%
overlap

50% overlap

Cx43

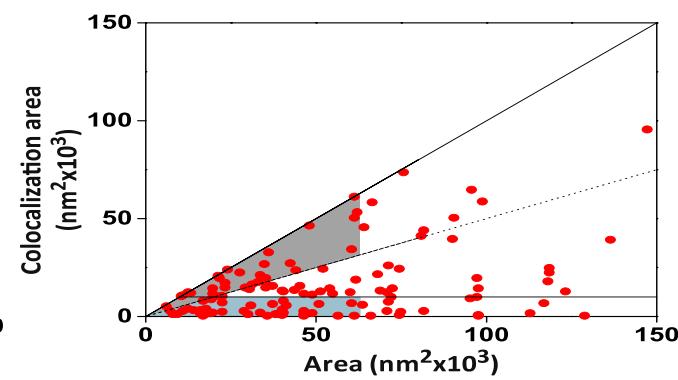
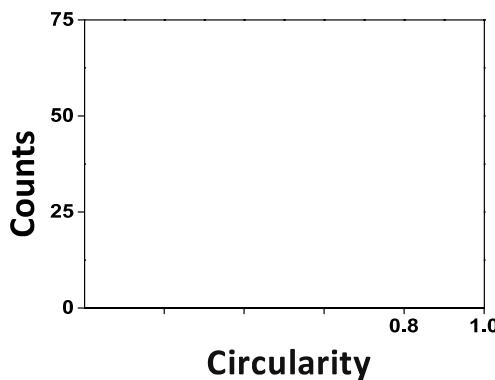
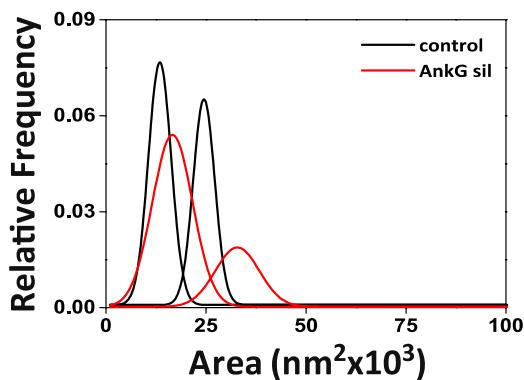
A correlation between overlap and Cx43 cluster area

Effect AnkG Silencing on Cx43



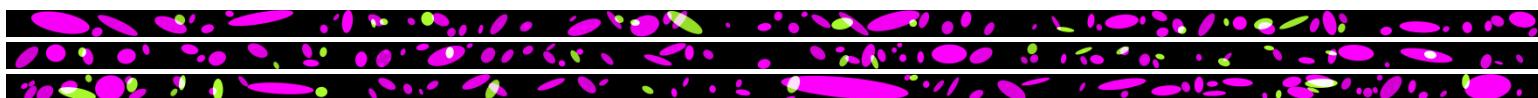
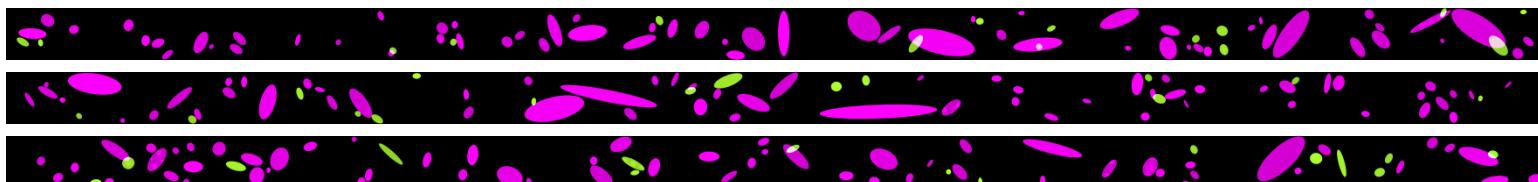
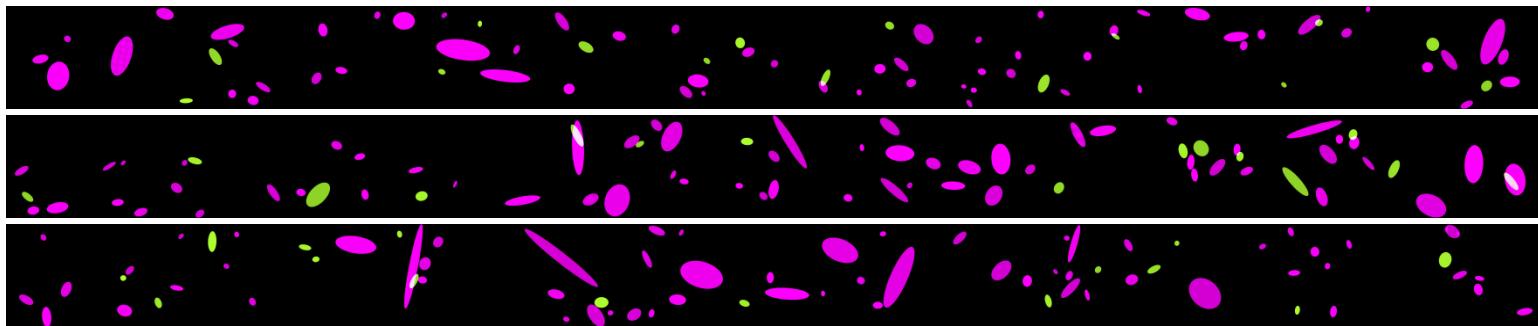
shRNA
100% overlap
50% overlap

AnkG Sil



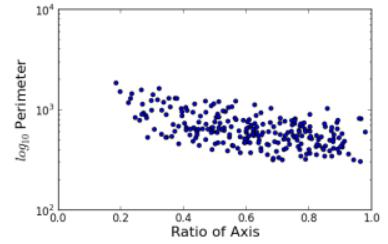
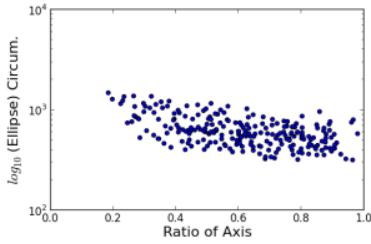
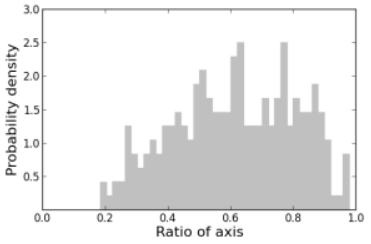
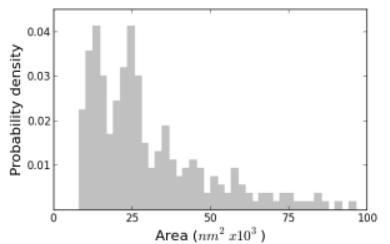
AnkG silencing results in increase of Cx43 cluster size and loss of circularity.

Monte-Carlo Simulations

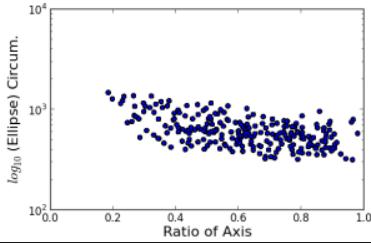
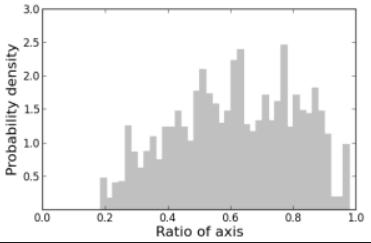
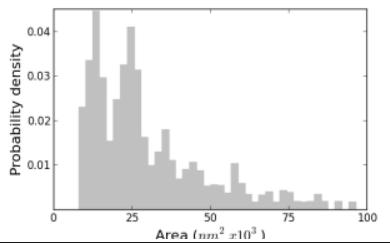


Monte-Carlo Simulations

Experiment

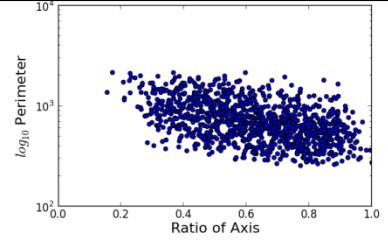
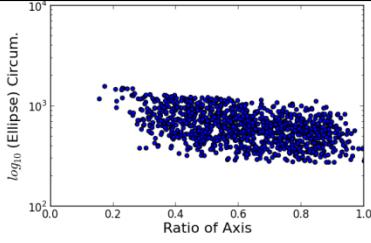
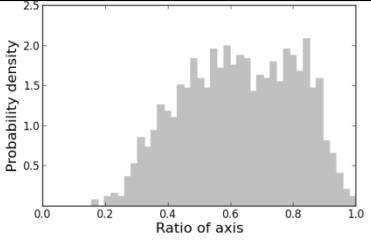
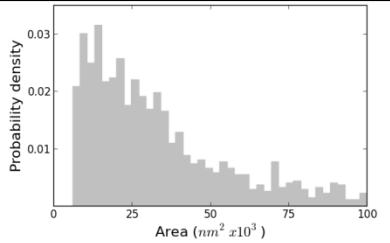


Simulation

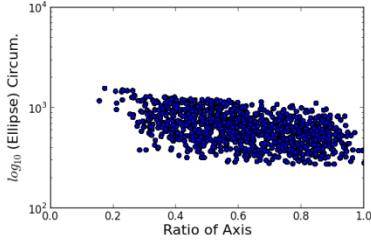
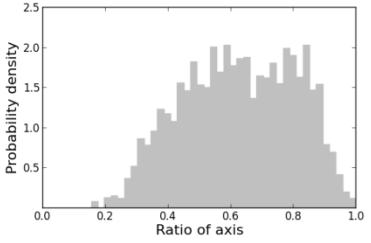
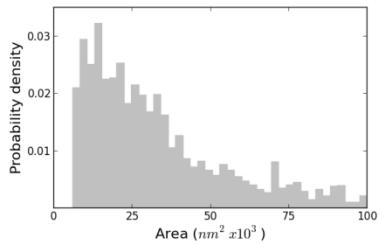


Cx43

Experiment



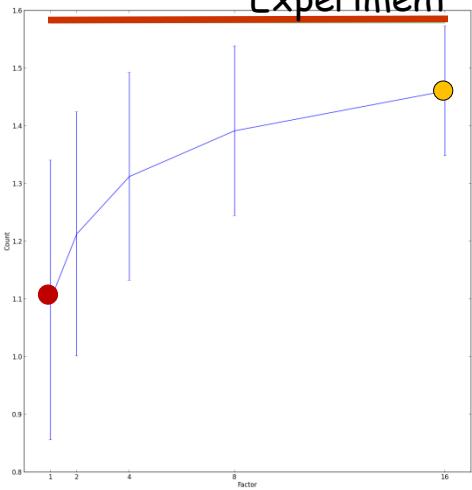
Simulation



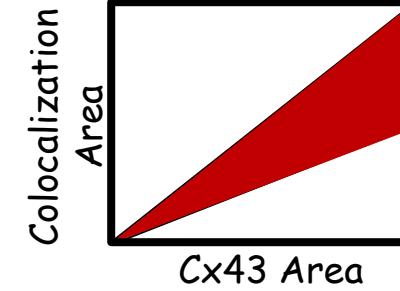
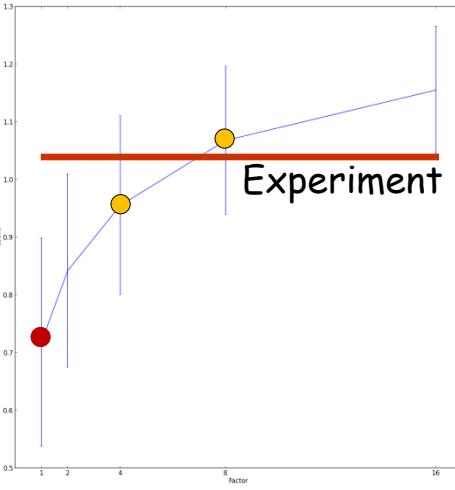
PKP2

Is the Observed Overlap Random?

Untreated
Experiment

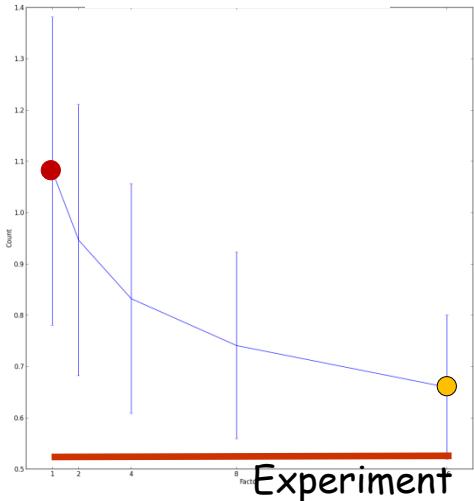


AnkG Silencing

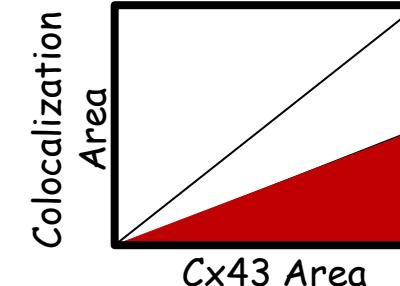
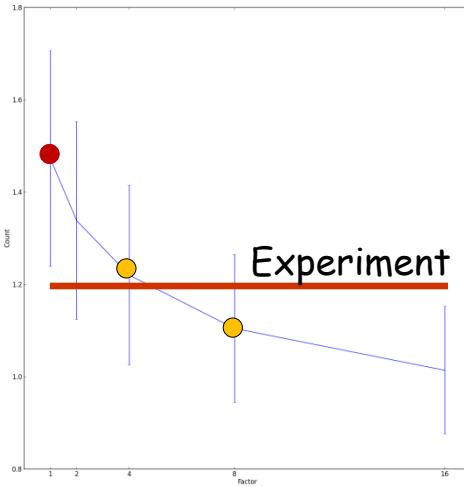


● Uniform
○ Non-uniform

Untreated



AnkG Silencing



Proteomics Informatics - Protein Characterization II: Protein Interactions (Week 11)
