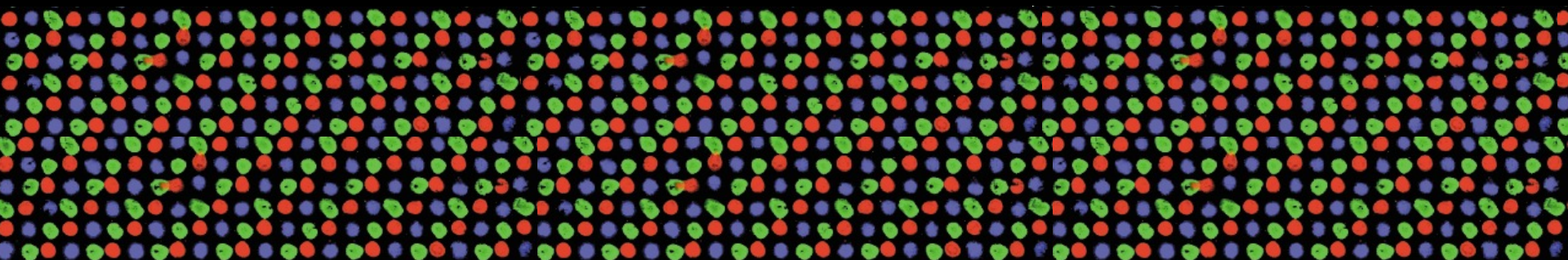


Lecture 2

Small Molecule Microarrays



a low-tech ligand discovery platform

February 15, 2024

Start of the millennium - the view from 2000

Diabetes (type 2)



< 100 Mendelian disease genes
(e.g. CFTR in cystic fibrosis, HEXA in Tay-Sachs)

12 common disease genetic variants
(e.g. CTLA4^{Thr17Ala} in Type 1 Diabetes, PRNP^{Met129Val} in Creutzfeld-Jacob)

PPAR γ

2000

2 decades+ on from the Human Genome Project

- Asthma
- Atrial fibrillation
- Breast cancer
- Crohn's disease
- Diabetes (type 1)
- Diabetes (type 2)
- Hypercholesterolemia
- Lupus
- Macular regeneration
- Myocardial infarction
- Obesity
- Prostate Cancer
- Others...

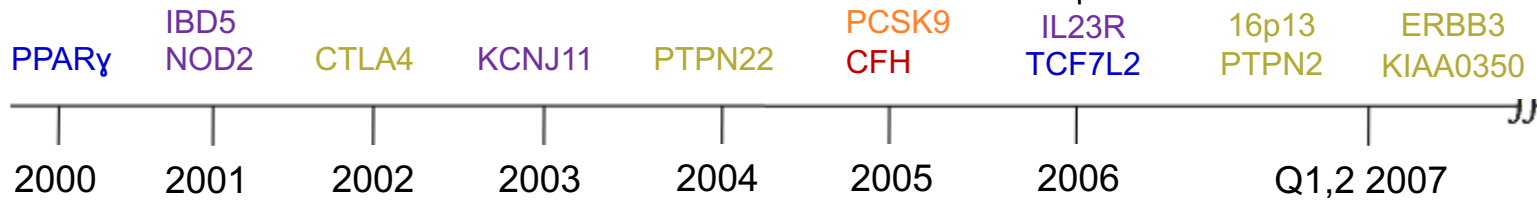


Thousands of loci affecting hundreds of common diseases

- LSP1
- HEX
- CDKAL1
- ORMDL3
- 4q25
- TCF2
- TCF2
- GCKR
- FTO
- CDKN2B/A
- FGFR2
- TNRC9
- MAP3K1
- PTPN2
- CDKN2B/A
- 8q24
- ATG16L1
- 5p13
- 10q21
- IRGM
- NKX2-3
- IL12B
- 3p21
- 1q24
- PTPN2
- IGF2BP2
- 8q24
- C12orf30
- ERBB3
- KIAA0350


















- IFIH1
- PCSK9
- CBF/C2
- LOC387715
- 8q24
- IL23R
- TCF7L2

- CD25
- IRF5
- PCSK9
- CFH



2024 – Gene-Disease Catalog (GDC)

Show SNPs for

 Digestive system disease	369
 Cardiovascular disease	1071
 Metabolic disease	489
 Immune system disease	1948
 Nervous system disease	1425
 Liver enzyme measurement	160
 Lipid or lipoprotein measurement	1026
 Inflammatory marker measurement	458
 Hematological measurement	5008
 Body weights and measures	2432
 Cardiovascular measurement	1001
 Other measurement	14018
 Response to drug	334
 Biological process	2342
 Cancer	1540
 Other disease	1921
 Other trait	1391



Drugging the genome

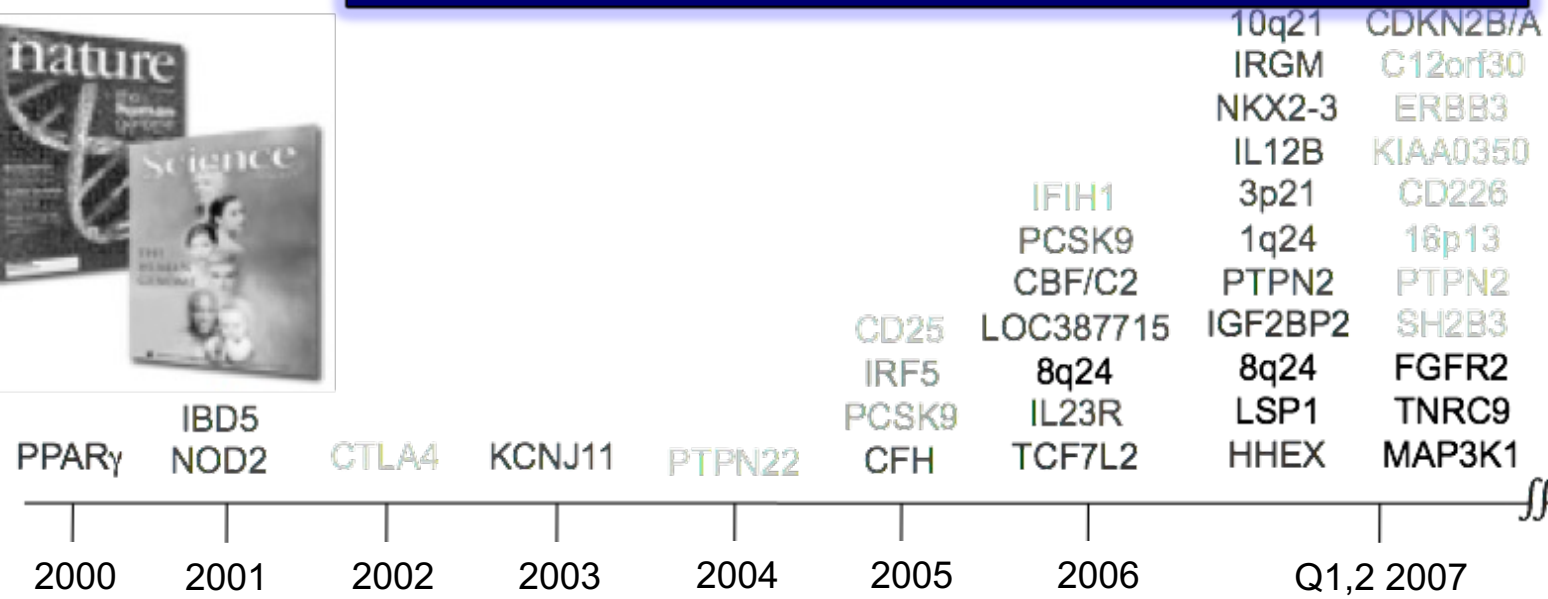
- Asthma
- Atrial fibrillation
- Breast cancer
- Crohn's disease
- Diabetes (type 1)
- Diabetes (type 2)
- Hypercholesterolemia
- Lupus
- Macular degeneration
- Myocardial infarction
- Obesity
- Prostate cancer
- Others...

of proteins targeted by the full armamentarium of drugs on the market <735

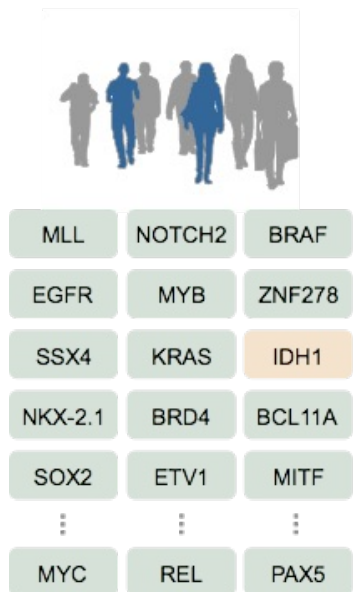
John P. Overington, EMBL-European Bioinformatics Institute



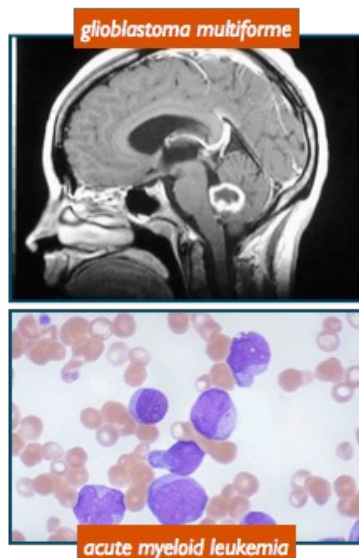
Thousands of loci affecting hundreds of common diseases



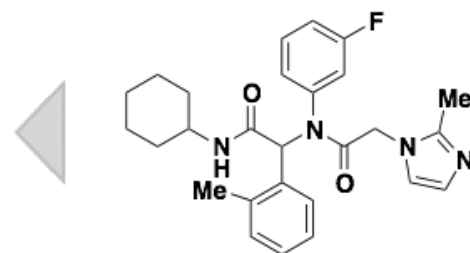
From Lecture 1 - Chemical probes of disease biology



patient samples reveal
list of disease genes



physiologic settings to test
the impact of disease genes



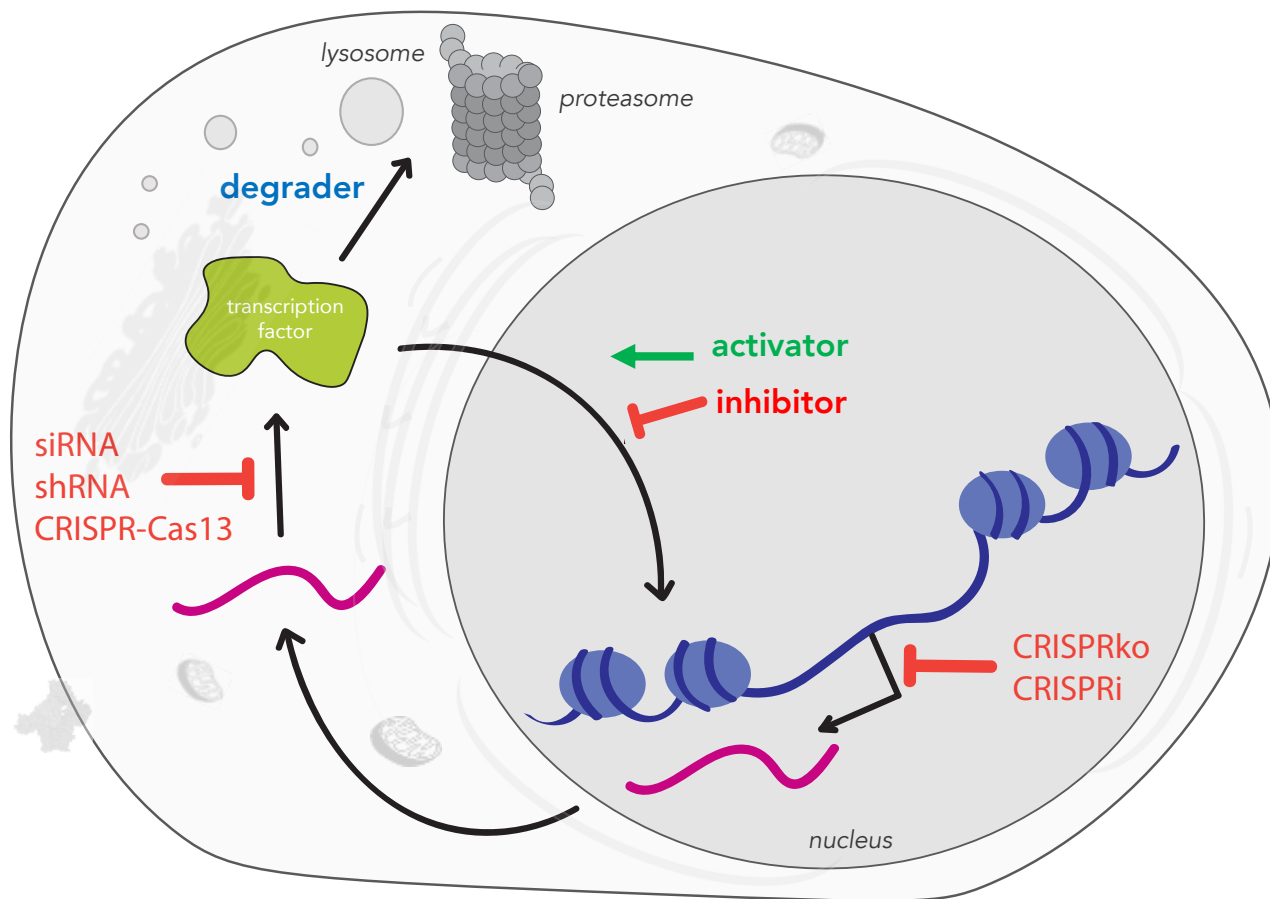
discover or develop small
molecules that reverse the
impact of disease genes

Approach: use small molecules to test emerging concepts in human disease
in physiologically relevant settings

Output: validated small-molecule probe to facilitate human clinical development
or diagnostic applications

An engineer's perspective on perturbation of proteins

intervention can take place at various parts of the system



genetic perturbants

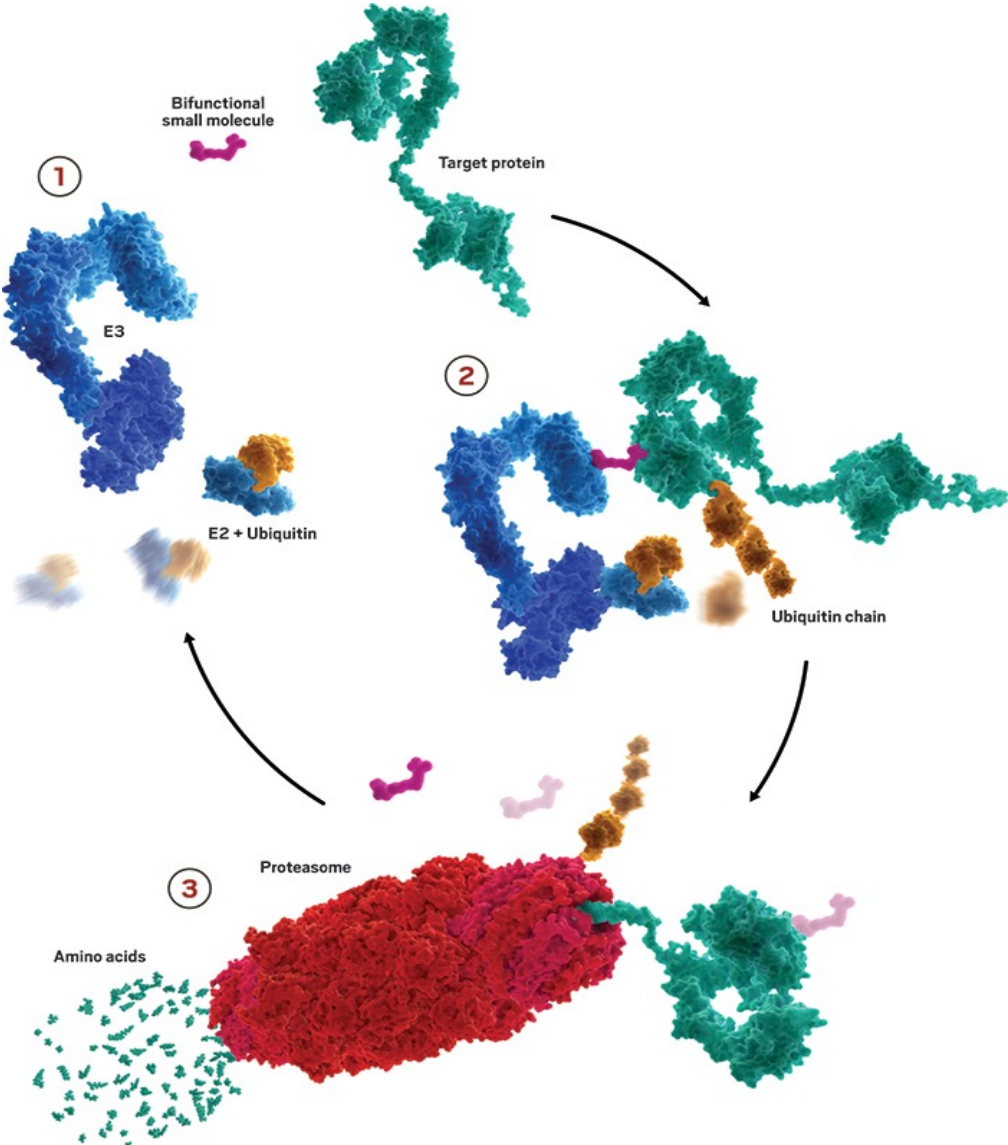
- ✓ shRNA
- ✓ CRISPR

chemical perturbants

- inhibitor
- activator
- degrader

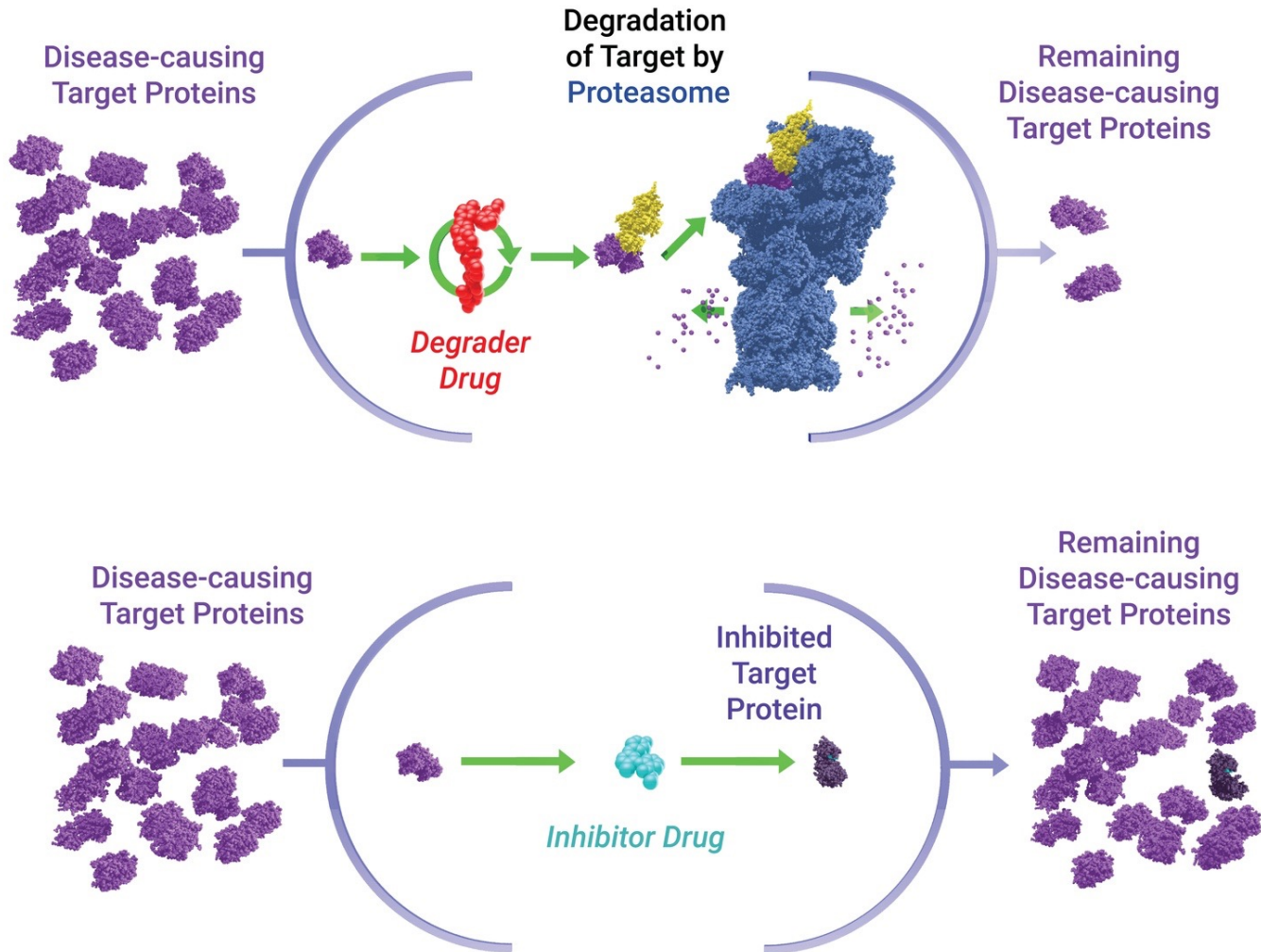
Your **MAX** chemical probes may utilize any of these mechanisms

A new approach - targeted protein degradation

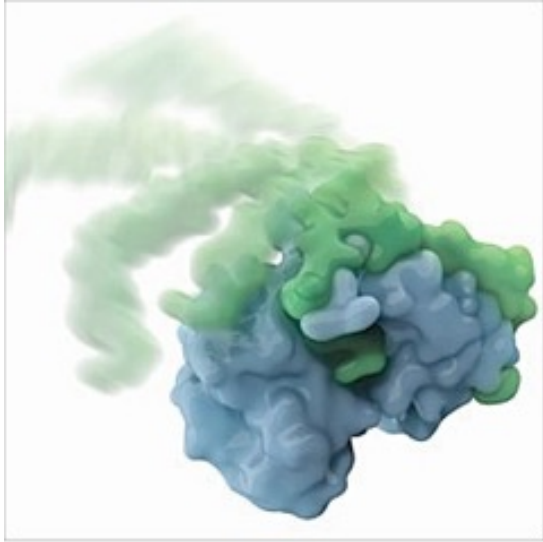


ADD text for 1, 2, 3

Targeted protein degradation



'Undruggable' targets are aplenty



disordered proteins

e.g., amyloids, transcription factors, enzymes



*DNA binding proteins
protein-protein interactors*

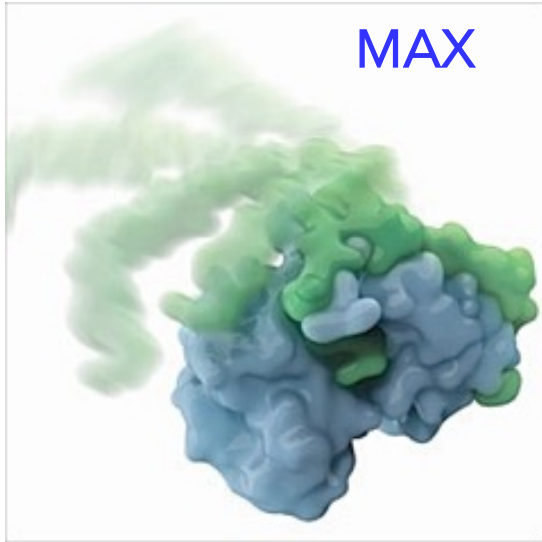
e.g., transcription factors, extracellular growth factors, scaffold proteins



integral membrane proteins

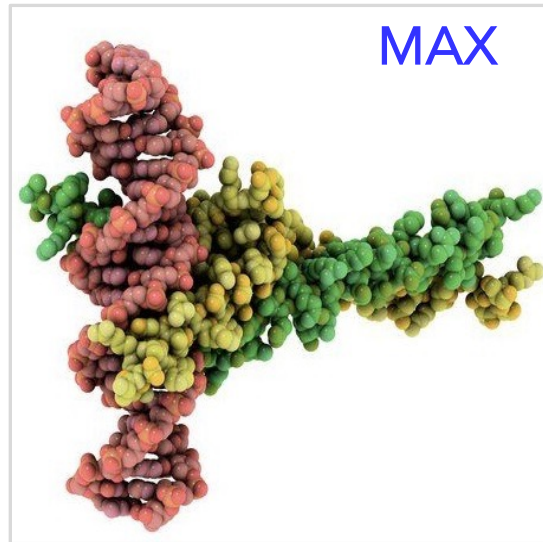
e.g., cell adhesion proteins, enzymes, receptors

'Undruggable' targets are aplenty



disordered proteins

e.g., amyloids, transcription factors, enzymes



*DNA binding proteins
protein-protein interactors*

e.g., transcription factors, extracellular growth factors, scaffold proteins

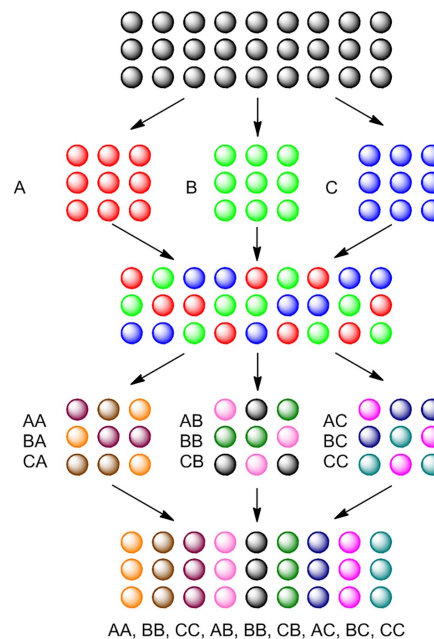
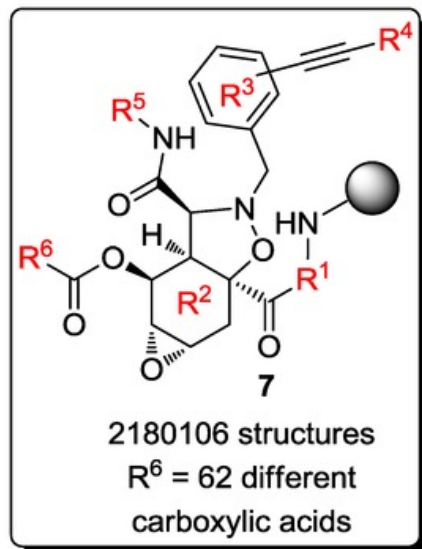


integral membrane proteins

e.g., cell adhesion proteins, enzymes, receptors

1998 – 'on-bead' binding assays

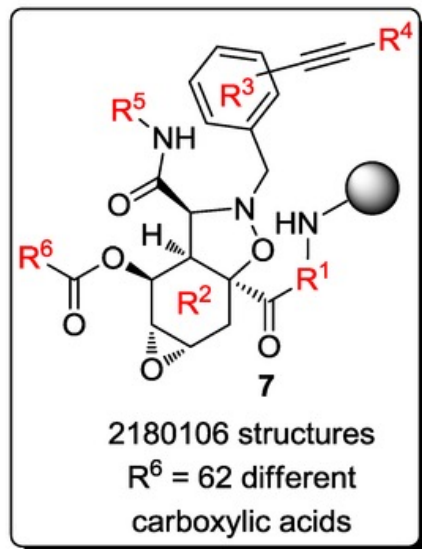
Chemical Library =
2.18M compounds on
90 μm Tentagel beads



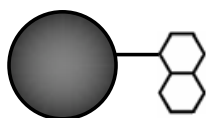
Split-Pool
Combinatorial
Synthesis

1998 – 'on-bead' binding assays

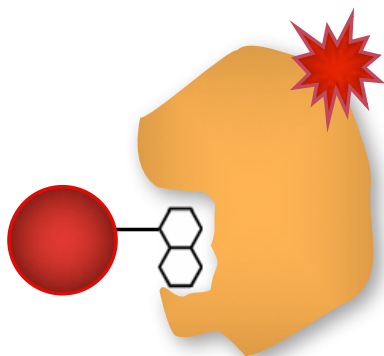
Chemical Library =
2.18M compounds on
90 μm Tentagel beads



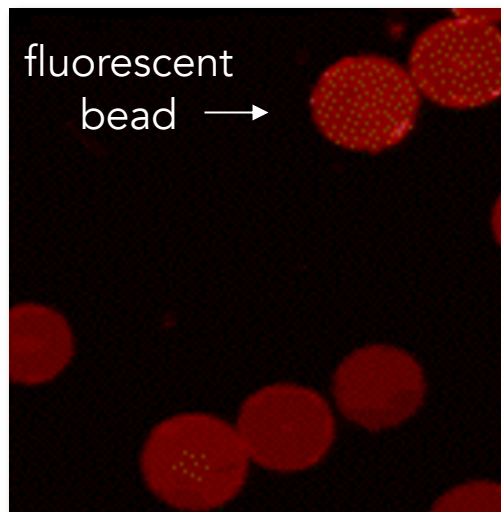
'Gradbot'
Angela
@ Harvard



no
binding

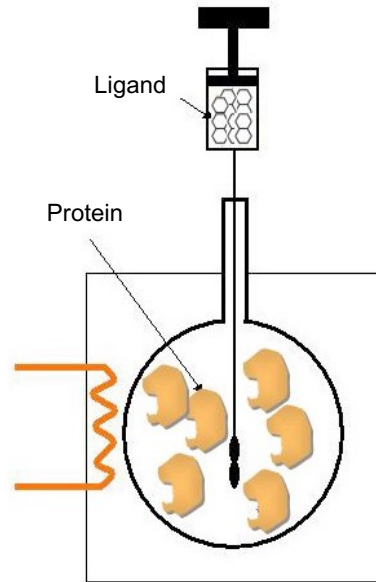
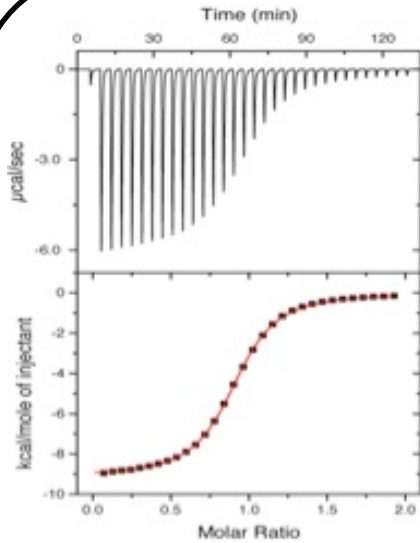


assay
positive



rhodamine dye
540/625 nm

1998 - other binding assay formats

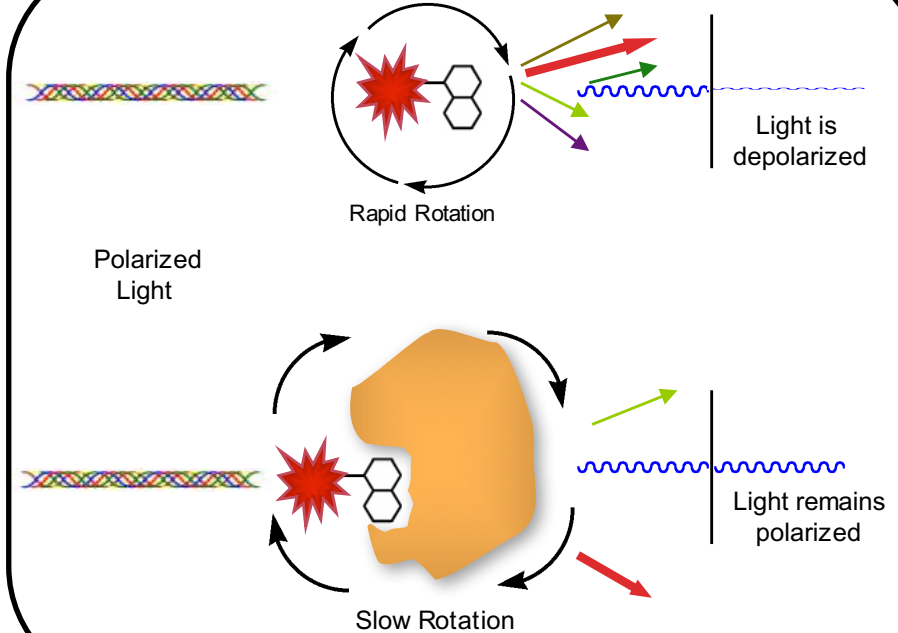


K_D equil, ΔG , ΔH

From 20.110 $\Rightarrow \Delta G = -RT \ln K_a = \Delta H - T\Delta S$

isothermal titration calorimetry

*measure changes in temperature upon binding,
plotted as power needed to maintain a constant T*



fluorescence polarization

*measure changes in rate of rotation
upon binding*

Late 1990s - 'Spatially addressable systems'

Dr. Patrick O. Brown

Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray

Mark Schena,* Dari Shalon,*† Ronald W. Davis,
Patrick O. Brown‡

A high-capacity system was developed to monitor the expression of many genes in parallel. Microarrays prepared by high-speed robotic printing of complementary DNAs on glass were used for quantitative expression measurements of the corresponding genes. Because of the small format and high density of the arrays, hybridization volumes of 2 microliters could be used that enabled detection of rare transcripts in probe mixtures derived from 2 micrograms of total cellular messenger RNA. Differential expression measurements of 45 *Arabidopsis* genes were made by means of simultaneous, two-color fluorescence hybridization.

SCIENCE • VOL. 270 • 20 OCTOBER 1995

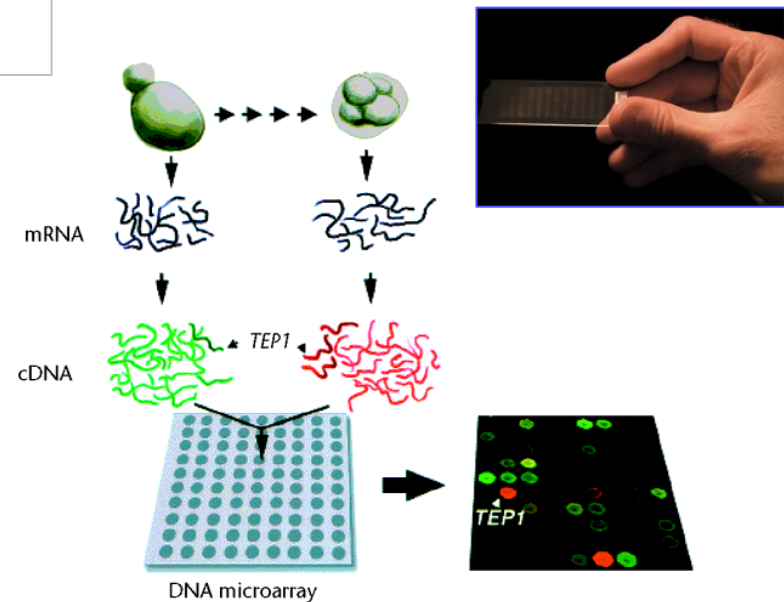


Exploring the new world of the genome with DNA microarrays

Patrick O. Brown^{1,3} & David Botstein²

Departments of ¹Biochemistry and ²Genetics, and the ³Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305, USA. e-mail: pbrown@cmgm.stanford.edu

Thousands of genes are being discovered for the first time by sequencing the genomes of model organisms, an exhilarating reminder that much of the natural world remains to be explored at the molecular level. DNA microarrays provide a natural vehicle for this exploration. The model organisms are the first for which comprehensive genome-wide surveys of gene expression patterns or function are possible. The results can be viewed as maps that reflect the order and logic of the genetic program, rather than the physical order of genes on chromosomes. Exploration of the genome using DNA microarrays and other genome-scale technologies should narrow the gap in our knowledge of gene function and molecular biology between the currently-favoured model organisms and other species.



follow changes in gene expression during yeast sporulation

Late 1990s - 'Spatially addressable systems'

Dr. Patrick O. Brown

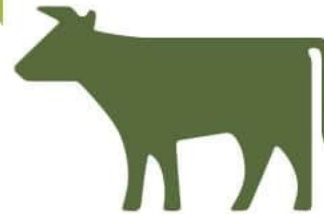
Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray

Mark Schena,* Dari Shalon,*† Ronald W. Davis,
Patrick O. Brown

A high-capacity system was developed in parallel. Microarrays prepared by high-density spotting on glass were used for quantitative expression analysis. Because of the small format and high throughput, microliters could be used that enabled the derivation of 2 micrograms of total RNA from 45 *Arabidopsis* genes for fluorescence hybridization.

SCIENCE

IMPOSSIBLE

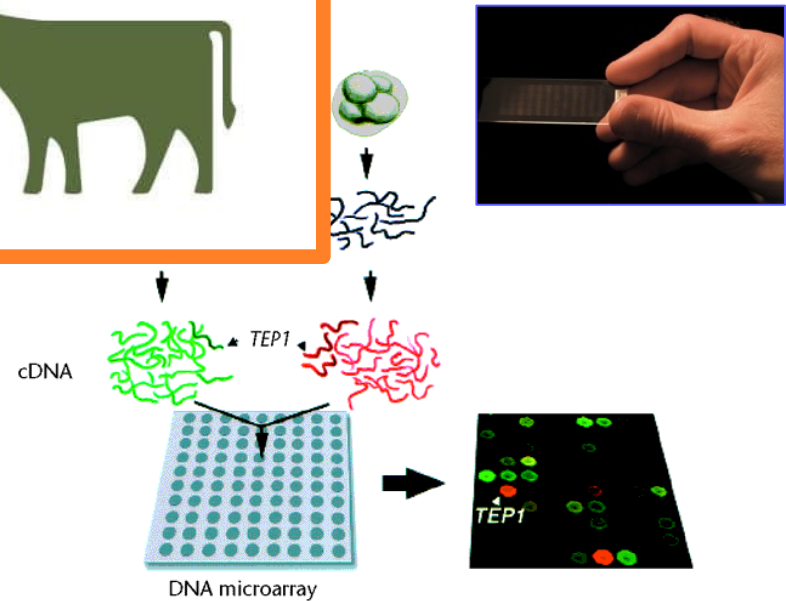


Exploring the new world with DNA

Patrick O. Brown^{1,3} & David Botstein²

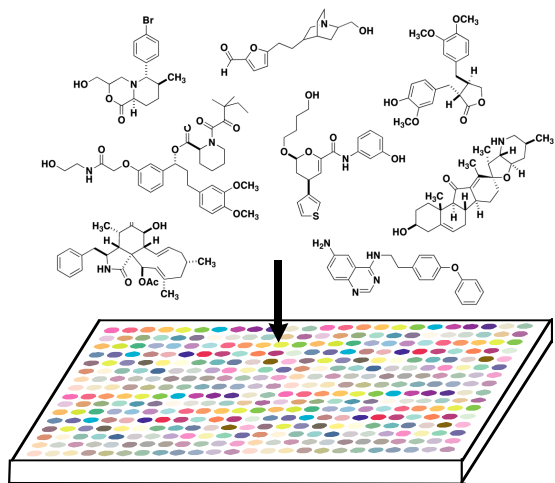
Departments of ¹Biochemistry and ²Genetics, and the ³Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305, USA. e-mail: pbrown@cmgm.stanford.edu

Thousands of genes are being discovered for the first time by sequencing the genomes of model organisms, an exhilarating reminder that much of the natural world remains to be explored at the molecular level. DNA microarrays provide a natural vehicle for this exploration. The model organisms are the first for which comprehensive genome-wide surveys of gene expression patterns or function are possible. The results can be viewed as maps that reflect the order and logic of the genetic program, rather than the physical order of genes on chromosomes. Exploration of the genome using DNA microarrays and other genome-scale technologies should narrow the gap in our knowledge of gene function and molecular biology between the currently-favoured model organisms and other species.



follow changes in gene expression during yeast sporulation

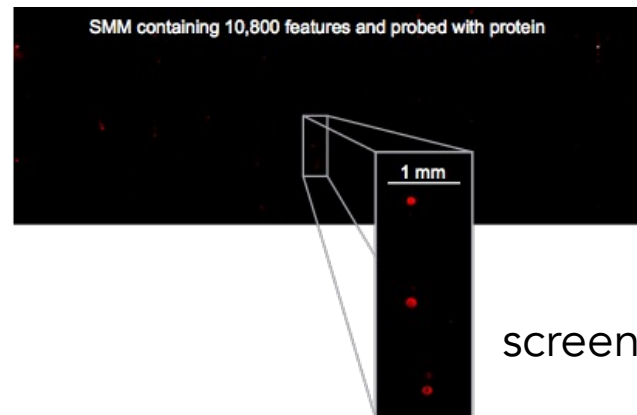
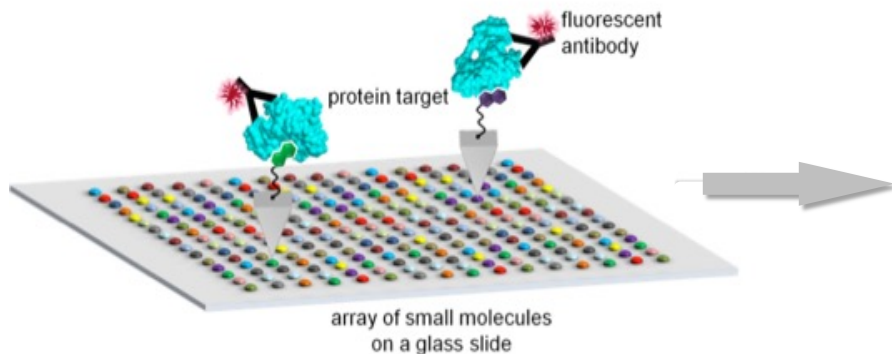
Small Molecule Microarrays (SMMs)



compound stock solutions



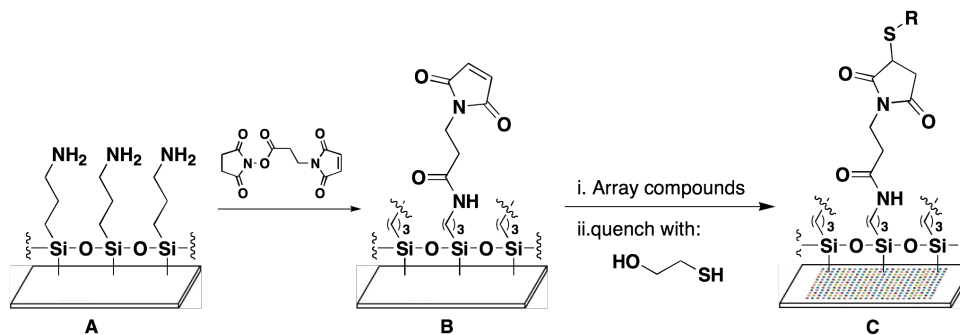
SMM manufacture and screening



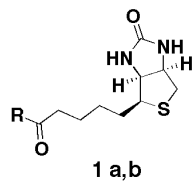
screened SMM

Proof-of-concept experiments for SMMs

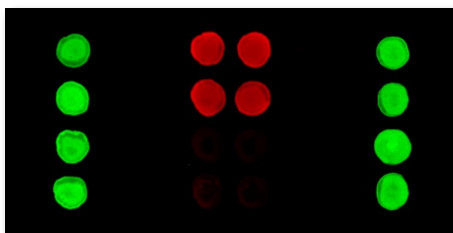
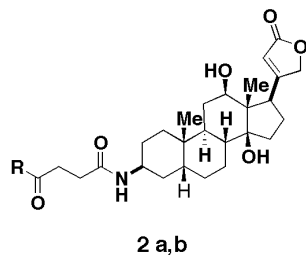
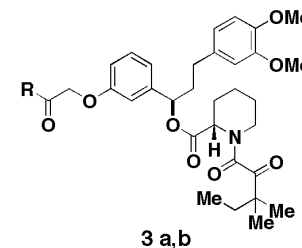
detecting known protein-ligand interactions



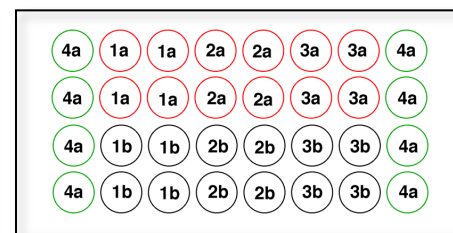
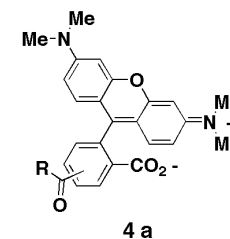
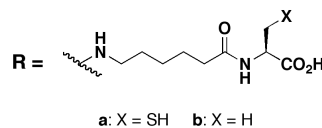
Streptavidin



FKBP12

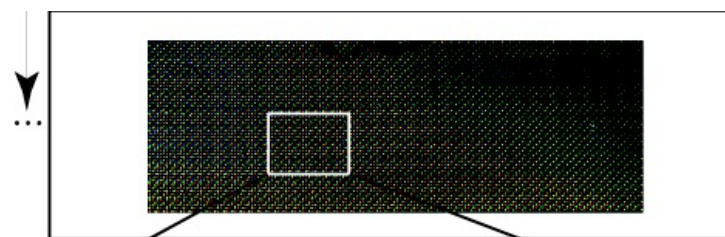
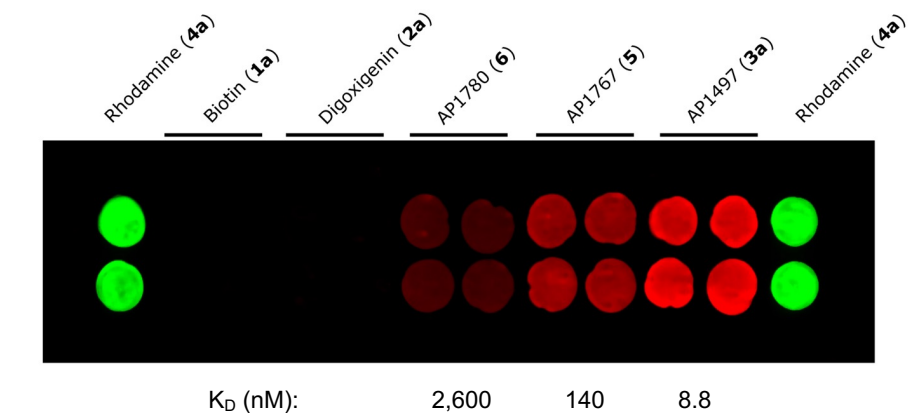


Anti-Digoxin mAb

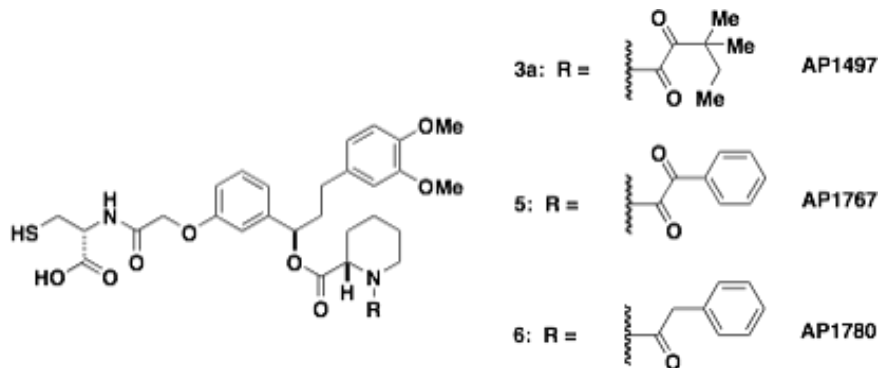
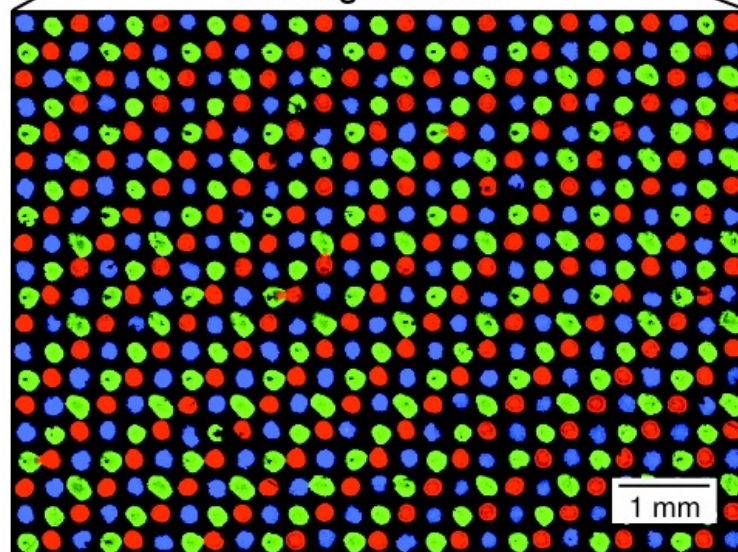


Proof-of-concept experiments for SMMs

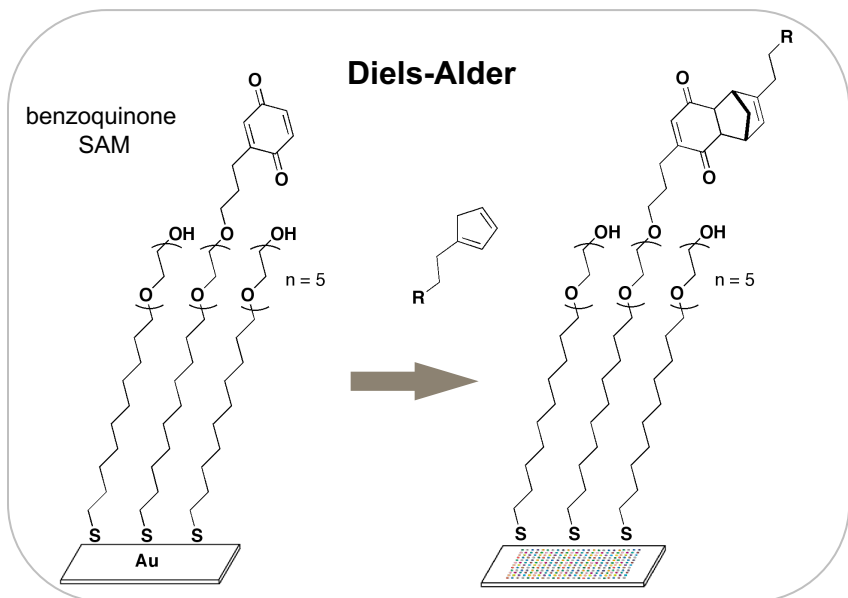
evaluating affinities and multiplexed formats



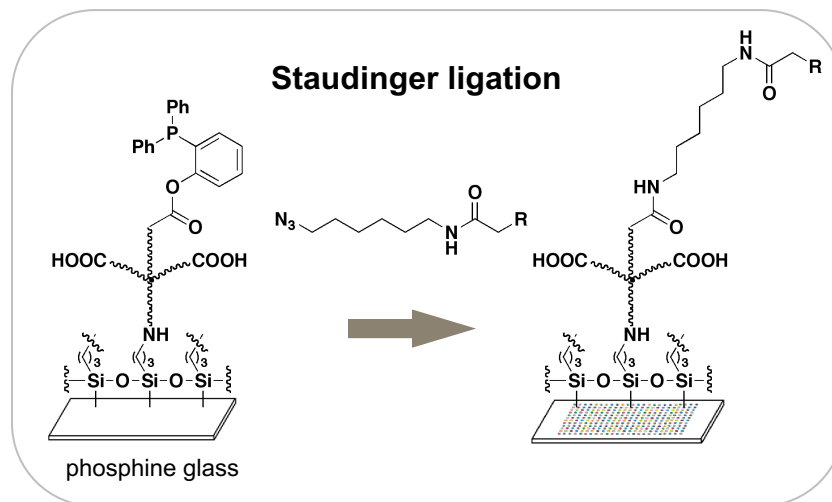
10x magnification



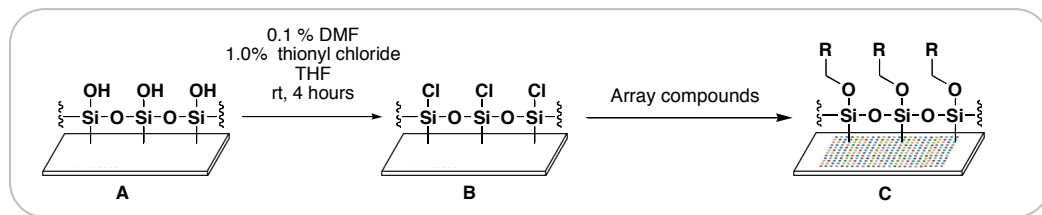
Capture chemistries for making SMMs



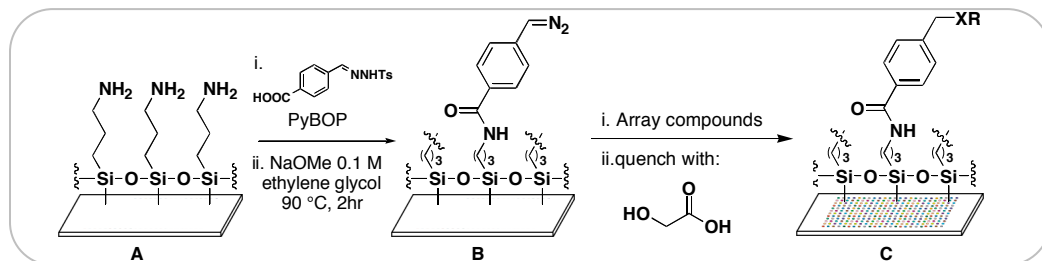
Houseman, B.T., Mrksich, M. *Chem. Biol.* 9, 443-454, 2002



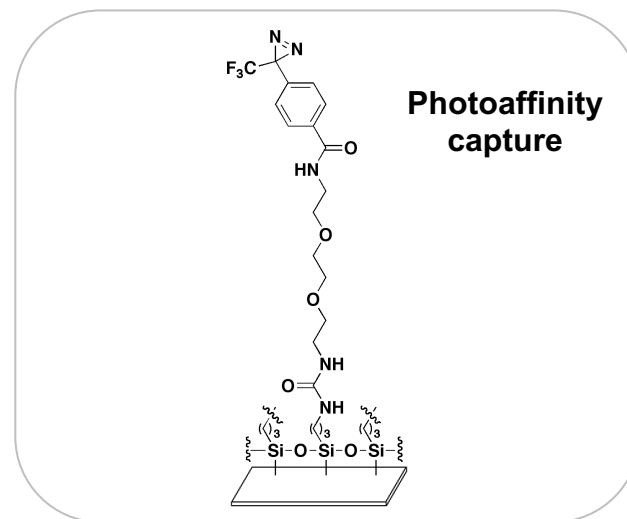
Köhn et al., *Angew. Chem. Int. Ed.* 42, 5830-5834, 2003



Hergenrother et al., *J. Am. Chem. Soc.* 122, 7849-7850, 1999

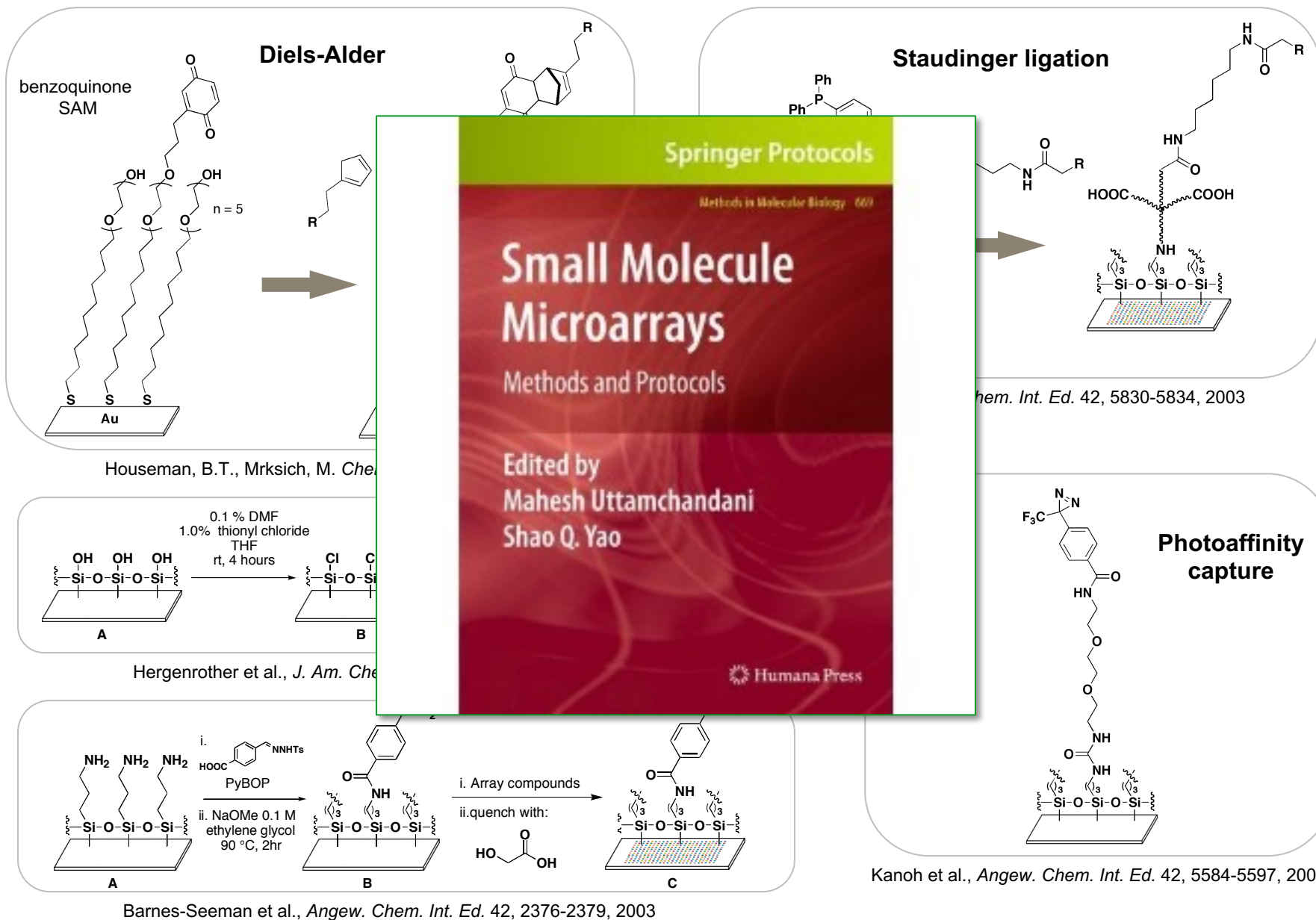


Barnes-Seeman et al., *Angew. Chem. Int. Ed.* 42, 2376-2379, 2003



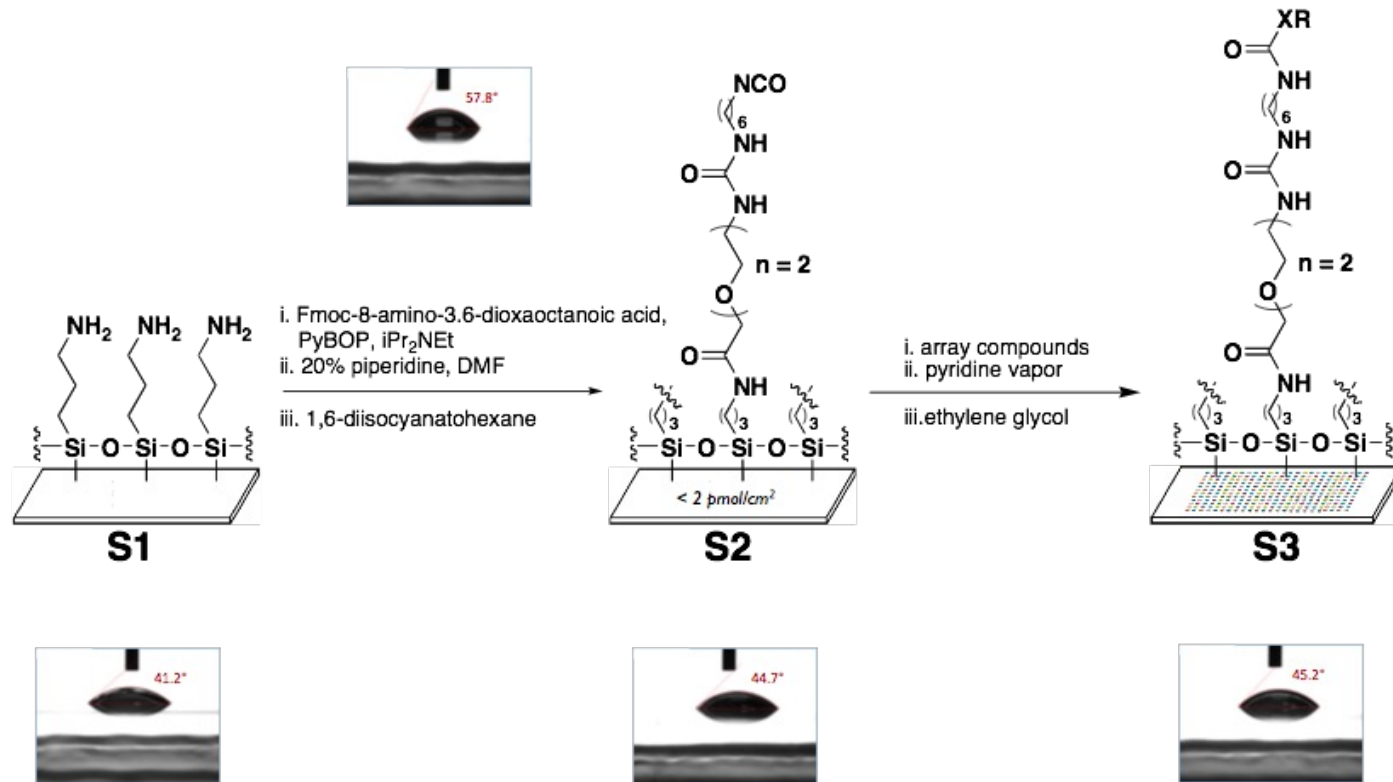
Kanoh et al., *Angew. Chem. Int. Ed.* 42, 5584-5597, 2003

Capture chemistries for making SMMs



Primary capture chemistry for making SMMs

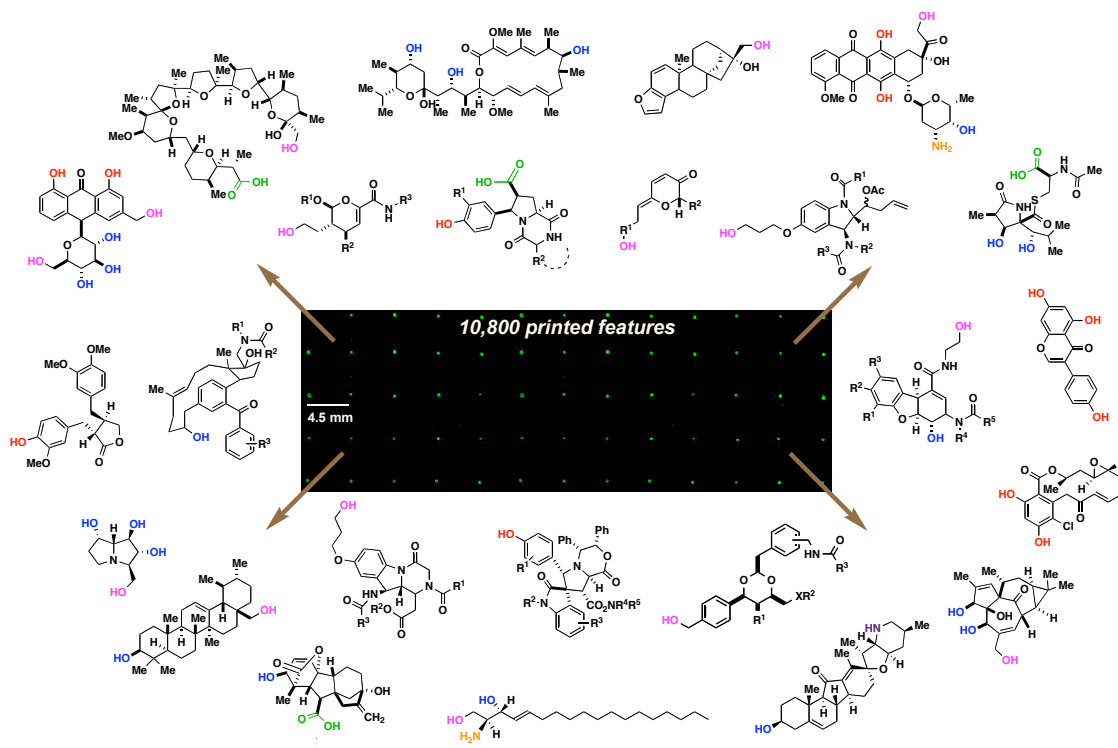
isocyanate coating reacts with nucleophilic functional groups



Bradner, J. E., McPherson, O. M., Mazitschek, R. M., Barnes-Seeman, D., Shen, J. P., Dhaliwal, J., Stevenson, K., Duffner, J. L., Park, S. B., Nghiem, P. T., Schreiber, S. L., Koehler, A. N., Chem Biol, 13, 493-504 (2006)

Bradner, J. E., McPherson, O. M., Koehler, A. N., Nature Protocols, 1, 2344-2352 (2006)

SMMs contain compounds from a variety of sources



 **KOCHINSTITUTE**
for Integrative Cancer Research at MIT

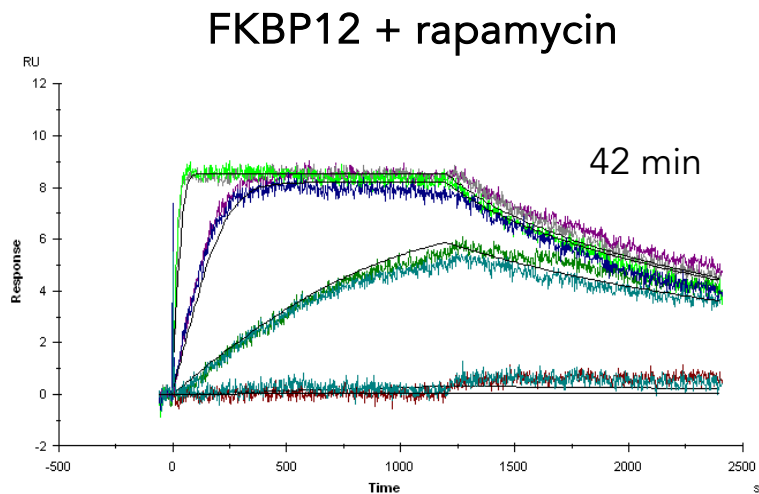
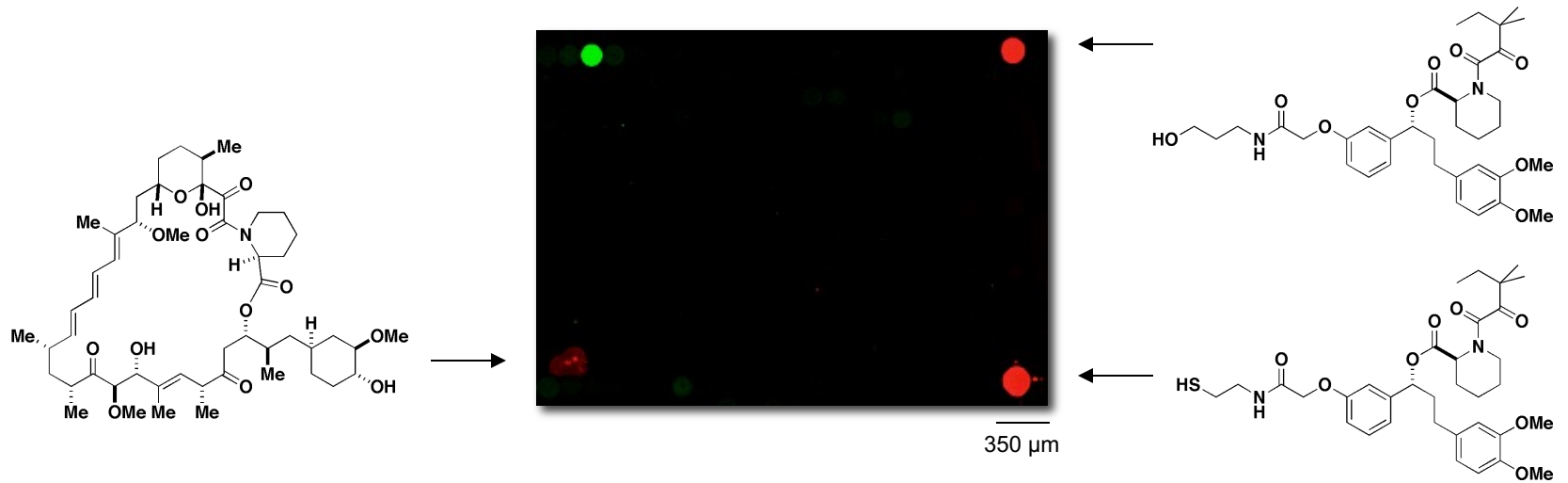


MIT CENTER FOR
PRECISION
CANCER MEDICINE

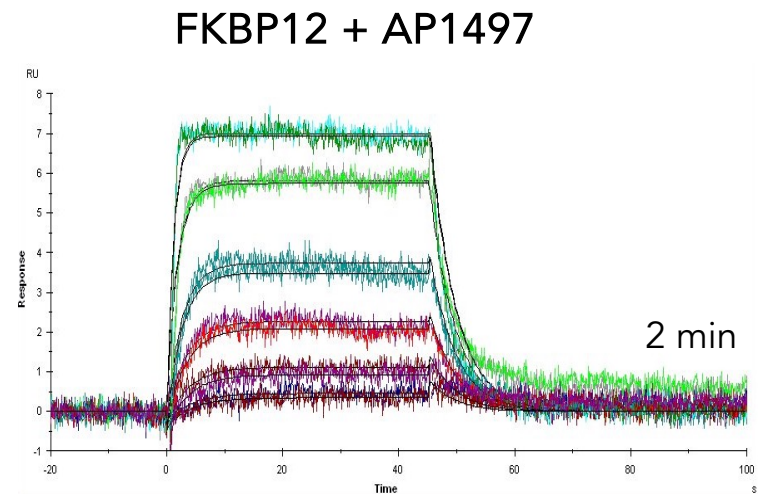
>135,000 commercials
~4,000 macrocycles
~4,000 bioactives, drugs
~4,000 Boston University
<1,000 MIT synthetics
100,000 Broad

In silico analysis of 400,000 'National Library' for screens:
>75% isocyanate-reactive

Interactions with varying kinetics can be visualized

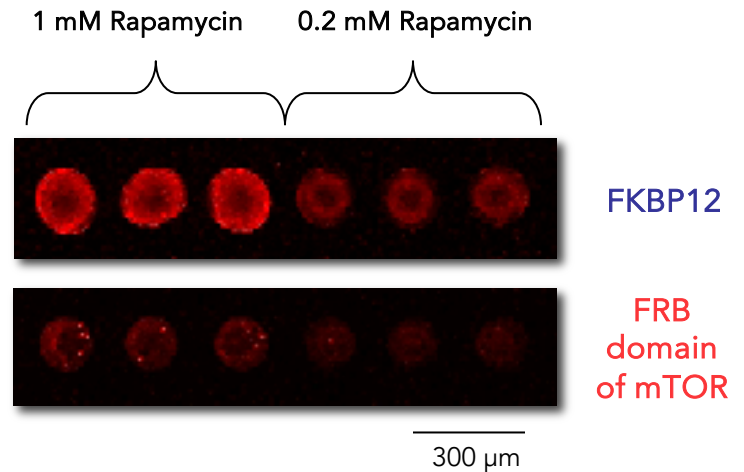
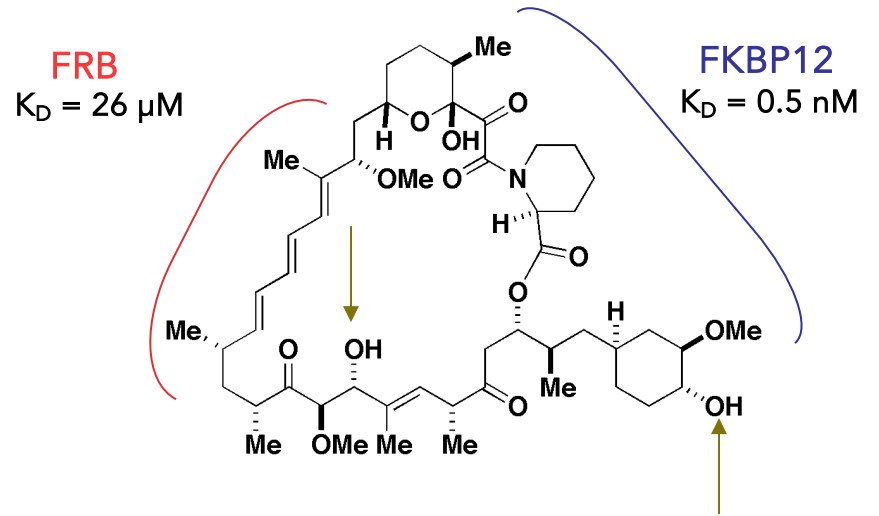
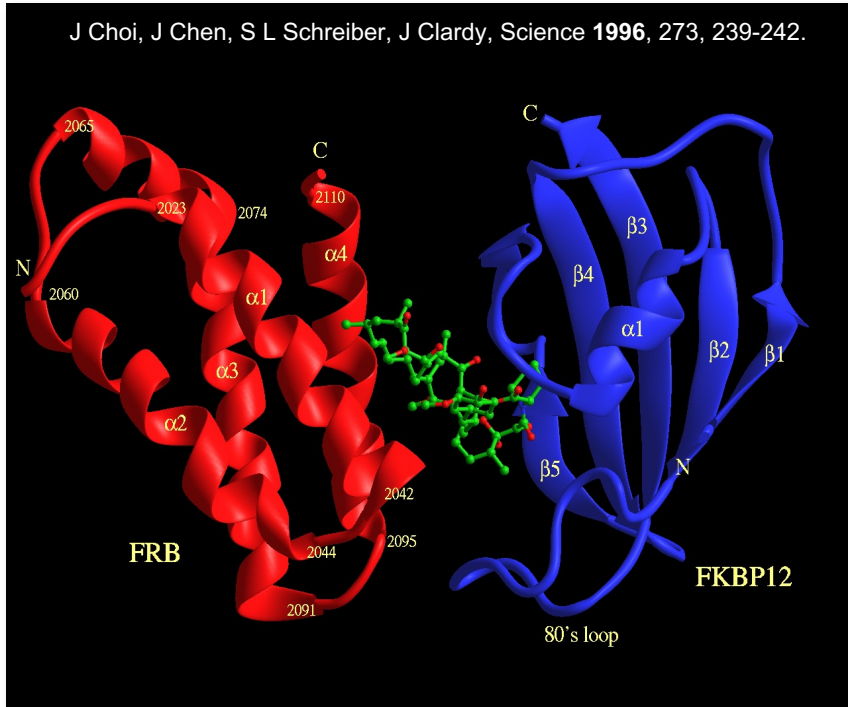


$$K_D = 0.5 \text{ nM}$$
$$K_d = 0.000965 \text{ sec}^{-1}$$



$$K_D = 18 \text{ nM}$$
$$K_d = 0.226 \text{ sec}^{-1}$$

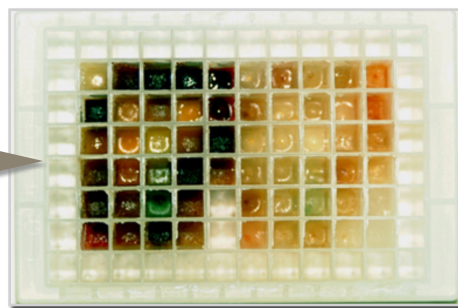
Detecting multiple interactions with Rapamycin



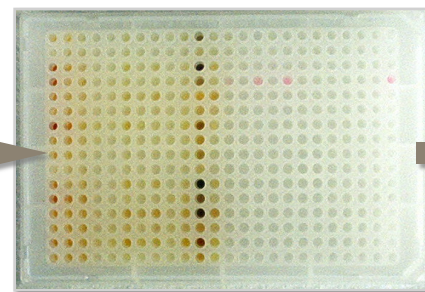
SMMs containing natural product extracts



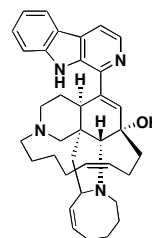
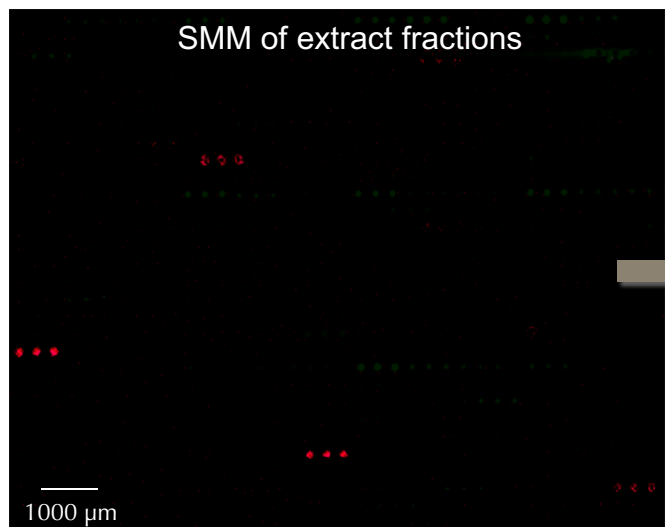
Didemnum roberti



crude extracts

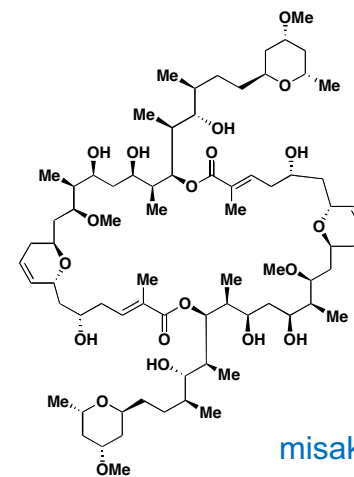


fractions of varying purity



manzamine A

eIF4a binder
stimulates IRES- and
cap-dependent translation

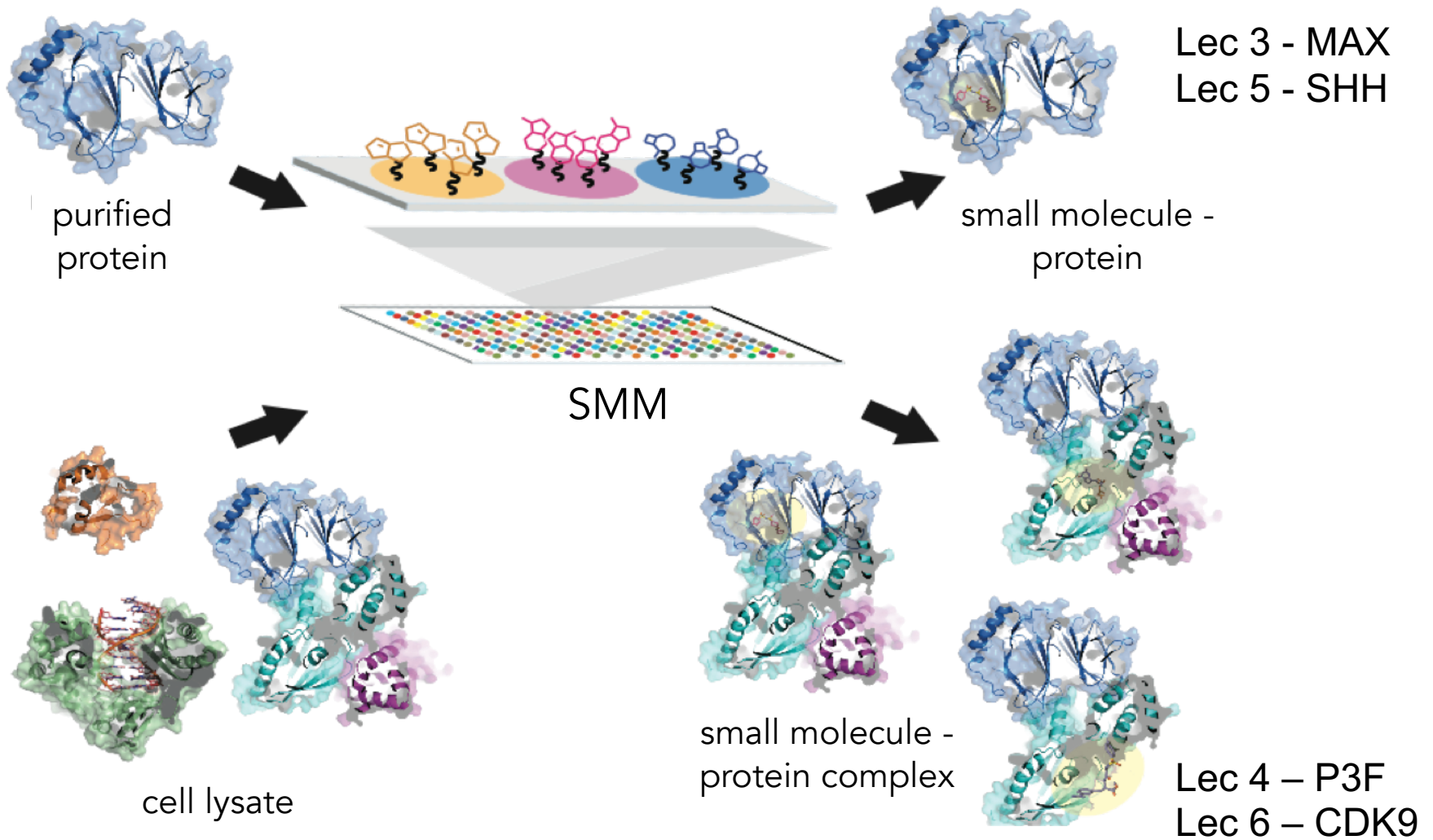


misakinolide A

tubulin binder

SMMs enable a new type of screen

target-directed assays in a native environment



SMMs enable surveys across panels of proteins

assess 'targetability' – typically with pure proteins or domains

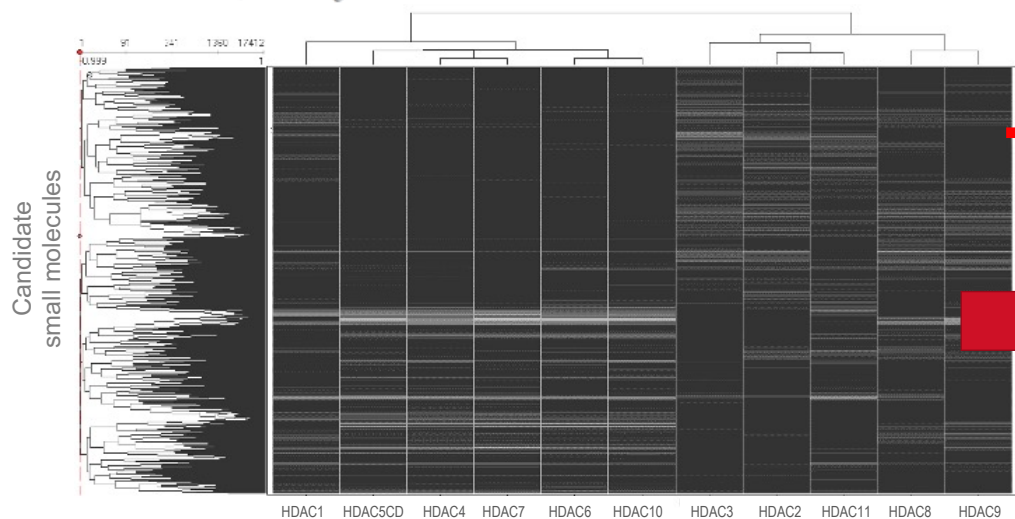


Articles
pubs.acs.org/acschemicalbiology

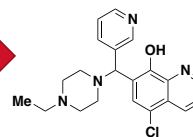
Inhibition of Zinc-Dependent Histone Deacetylases with a Chemically Triggered Electrophile

Zarko V. Boskovic,^{†,‡,¶} Melissa M. Kemp,^{†,¶} Allyson M. Freedy,^{†,‡} Vasanthi S. Viswanathan,[†] Marius S. Pop,^{§,||} Jason H. Fuller,[†] Nicole M. Martinez,[†] Samuel O. Figueroa Lazú,[†] Jiyoung A. Hong,^{†,§,‡} Timothy A. Lewis,[†] Daniel Calarese,[¶] James D. Love,[¶] Amedeo Vetere,[†] Steven C. Almo,[¶] Stuart L. Schreiber,^{†,‡} and Angela N. Koehler^{¶,†,§,||}

100 transcription factors (Clemons et al., PNAS, 2010)
11 HDACs (Boskovic et al., ACS Chem Biol, 2016)
20 RNA-binding proteins (recently completed)
34 cytokines (recently completed)



Array of HDAC proteins



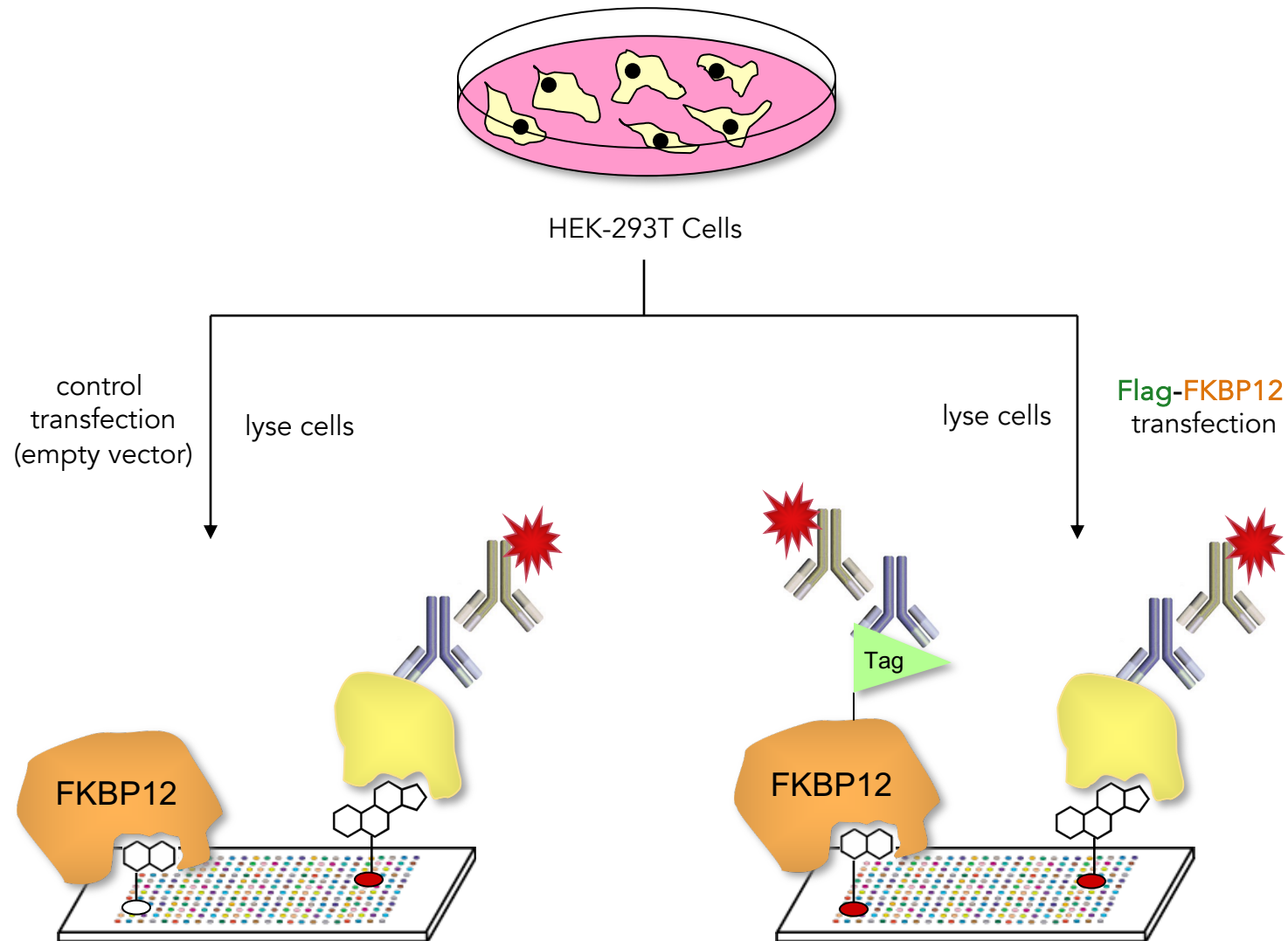
BRD-4354

selectively inhibits HDACs 5 and 9

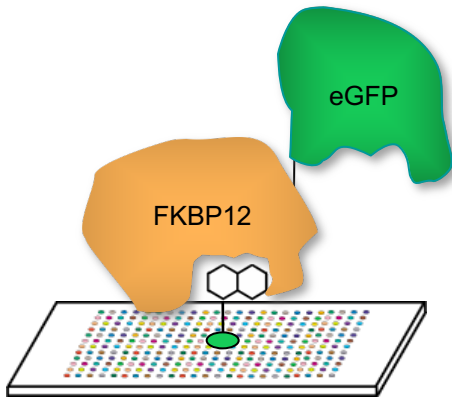
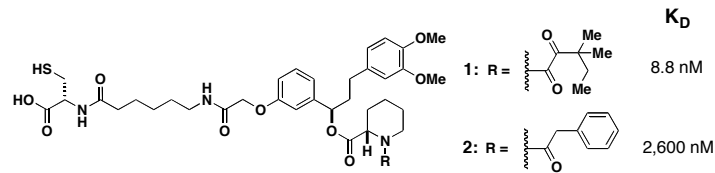
on-array selectivity
on-array SBR patterns
fragmentation (e.g., BRICS, RECAP, Bemis-Murcko, etc.)

signatures as training sets for machine learning

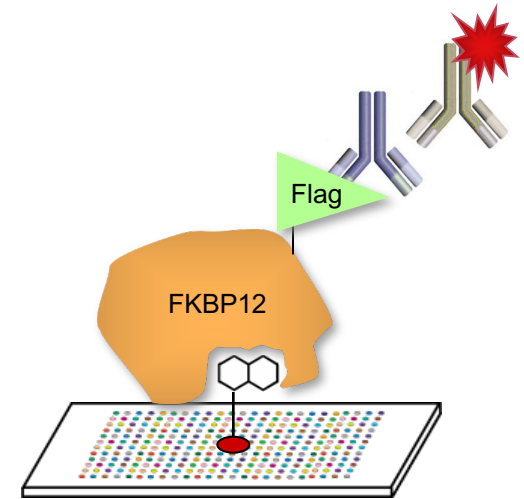
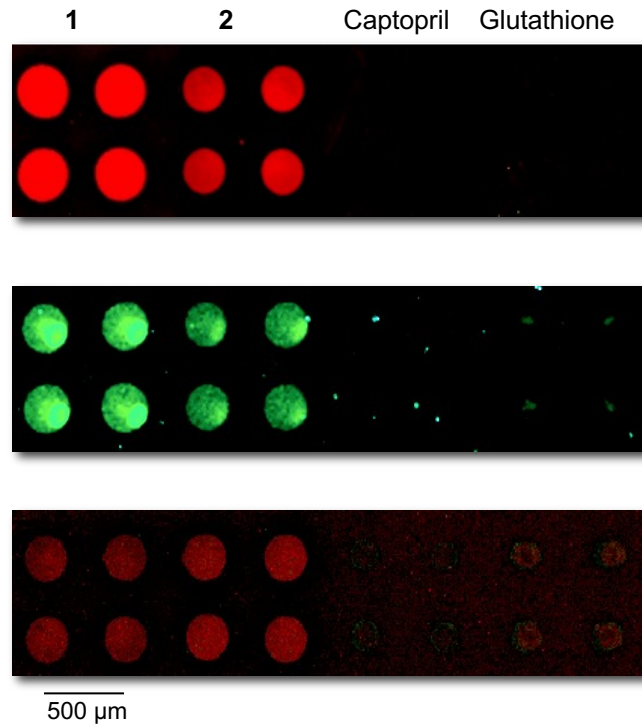
Binding screens involving cell lysates



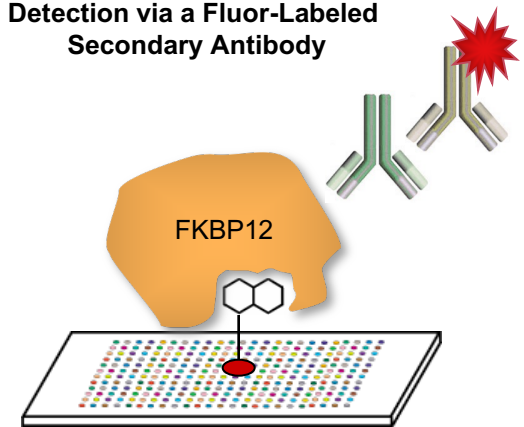
Comparing detection methods using lysates



Detection via Green Fluorescent Protein

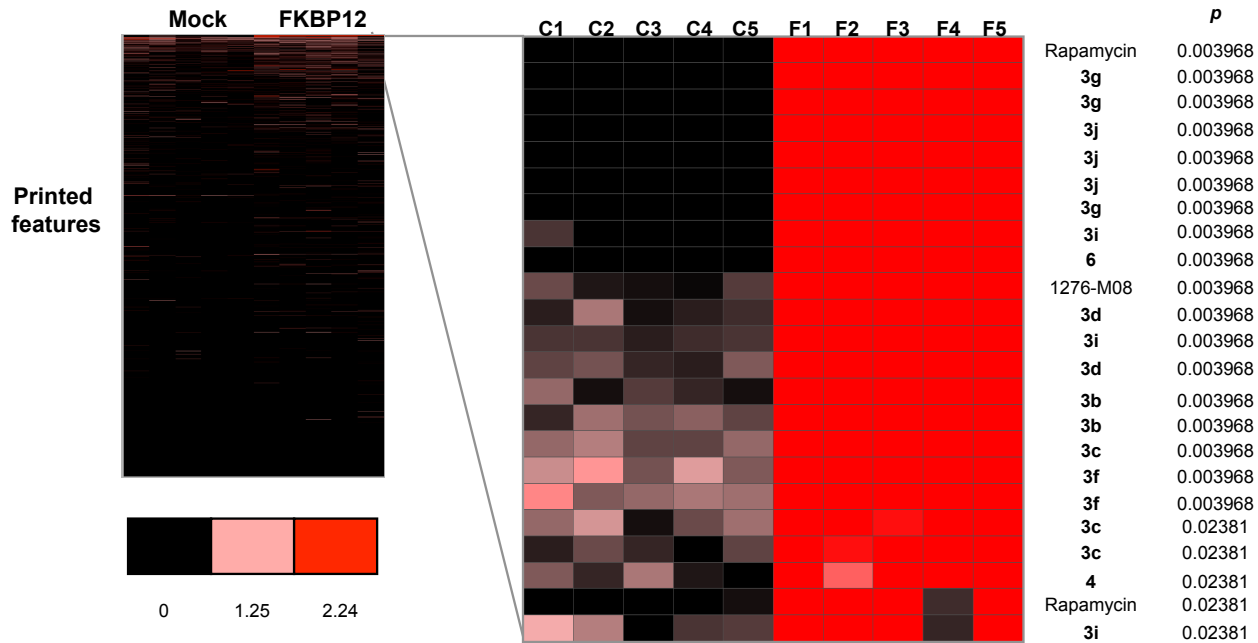
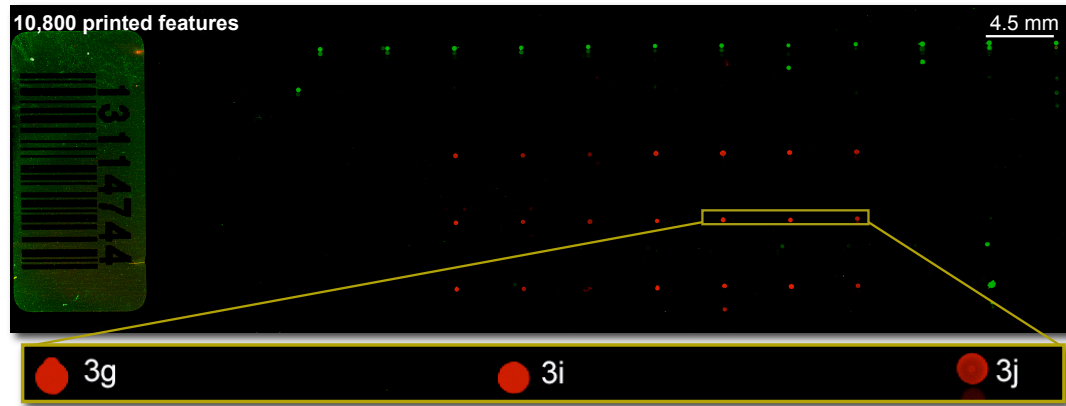


Detection via a Fluor-Labeled Secondary Antibody

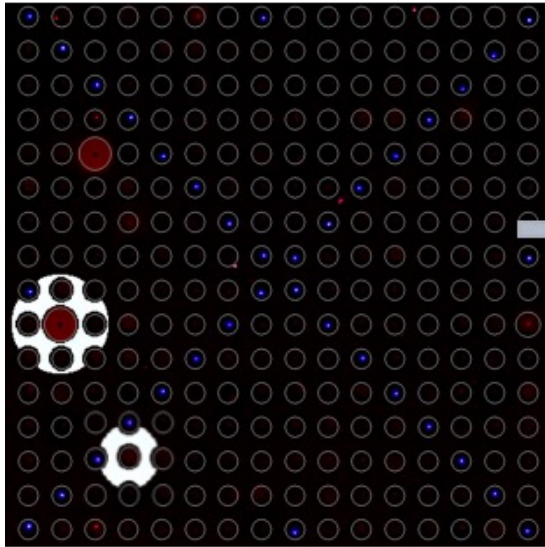


Detection via anti-FKBP12 Antibody and Labeled Secondary

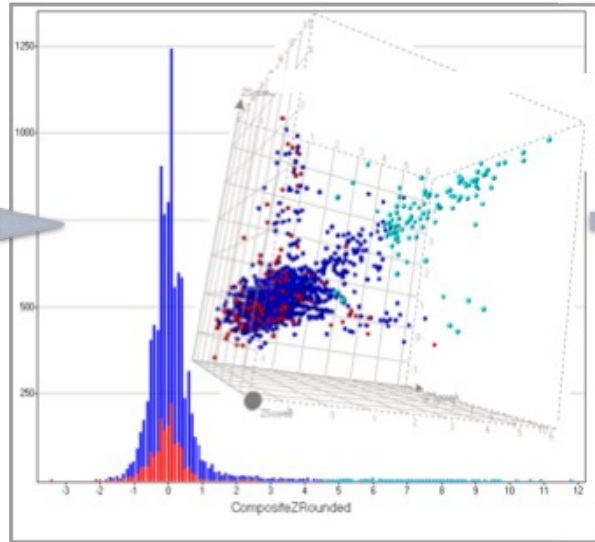
Binding screen using in cell lysates



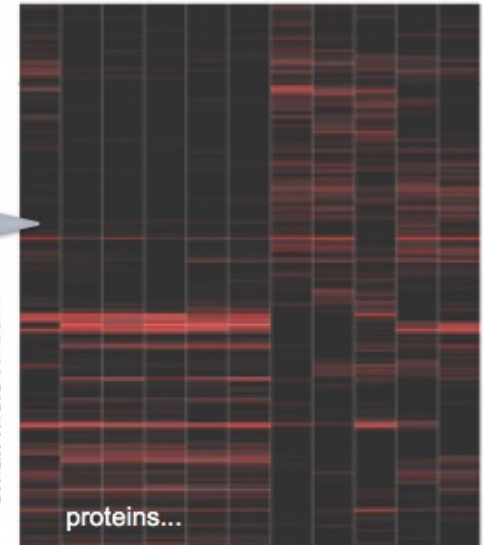
Analysis pipeline – the simple version



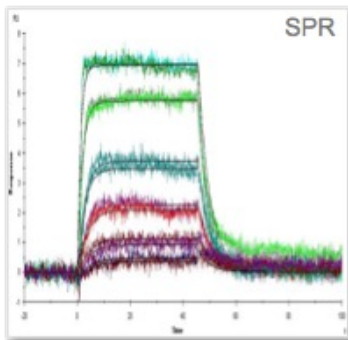
fluorescent features reveal putative interactions



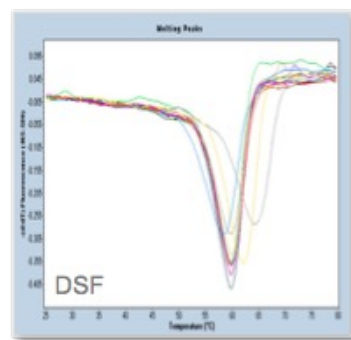
compute composite Z-scores (hit calls)



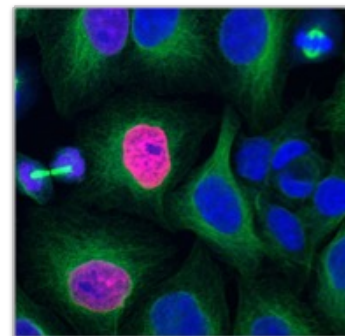
specificity analysis



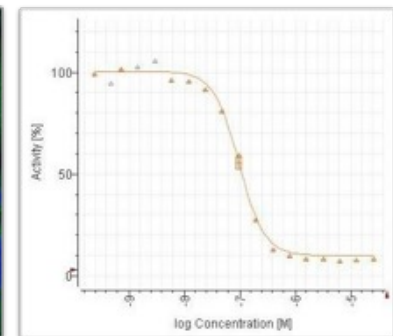
secondary binding assays



+



functional assays




Public access for SMM data sets

<http://chembank.broad.mit.edu>

DSA-ChemBank: 796,063 curated compounds, 1,963 assays, 149 projects, 16,942,065 well measurements

ChemBank: 528,062 curated compounds, 529 assays, 45 projects, 5,764,724 well measurements

43,651 users at 8,309 organizations in 154 countries



<http://pubchem.ncbi.nlm.nih.gov>

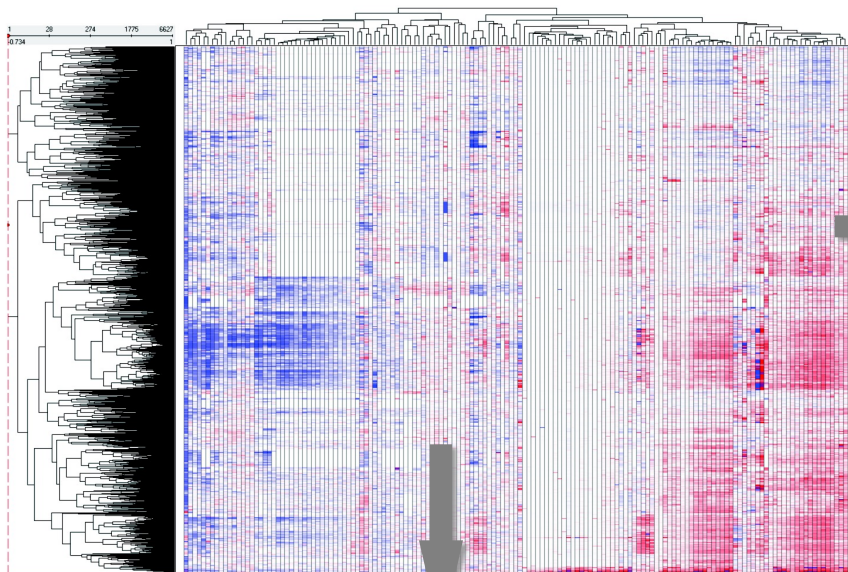


<http://bard.nih.gov/drupal>

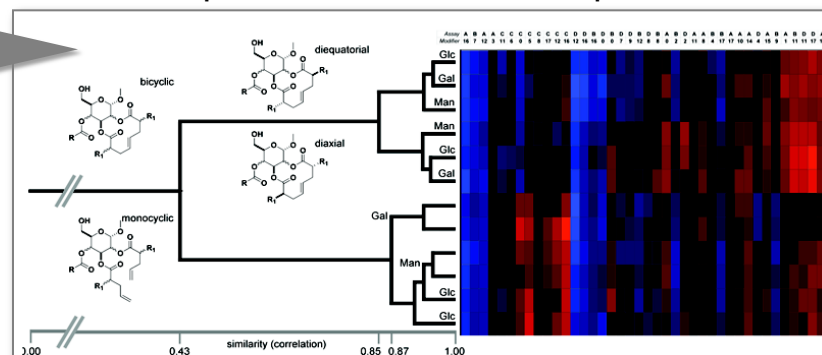
ChemBank: an analytical tool for the community

assays (cell-based, biochemical, binding)

small molecules



relationship of structure to screen performance



hit in a PRMT1 SMM binding assay

(structural analog)

DOS compound

inhibitor of arginine HMT (PRMT1) enzymatic activity

CHEMBANK

Find Small Molecules

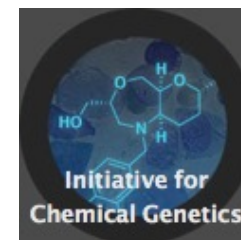
- By substructure**
Search compound collection by substructure using the SMARTS string, or stream with AME Molecular Editor.
- By similarity**
Search compound collection by similarity to a structure using the Simlab or Stream with AME Molecular Editor.
- By descriptor**
Filter compound collection using calculated molecular descriptor values.
- By assay**
Find compounds scoring as 'hit' in biological assays.
- By function**
Find compounds with known biochemical interactions, therapeutic uses, or molecular functions.
- By vendor**
Find compounds made by a particular chemist, or sold by a particular vendor.
- By molecule name**
Find compounds with a particular name, or containing a part of a name.

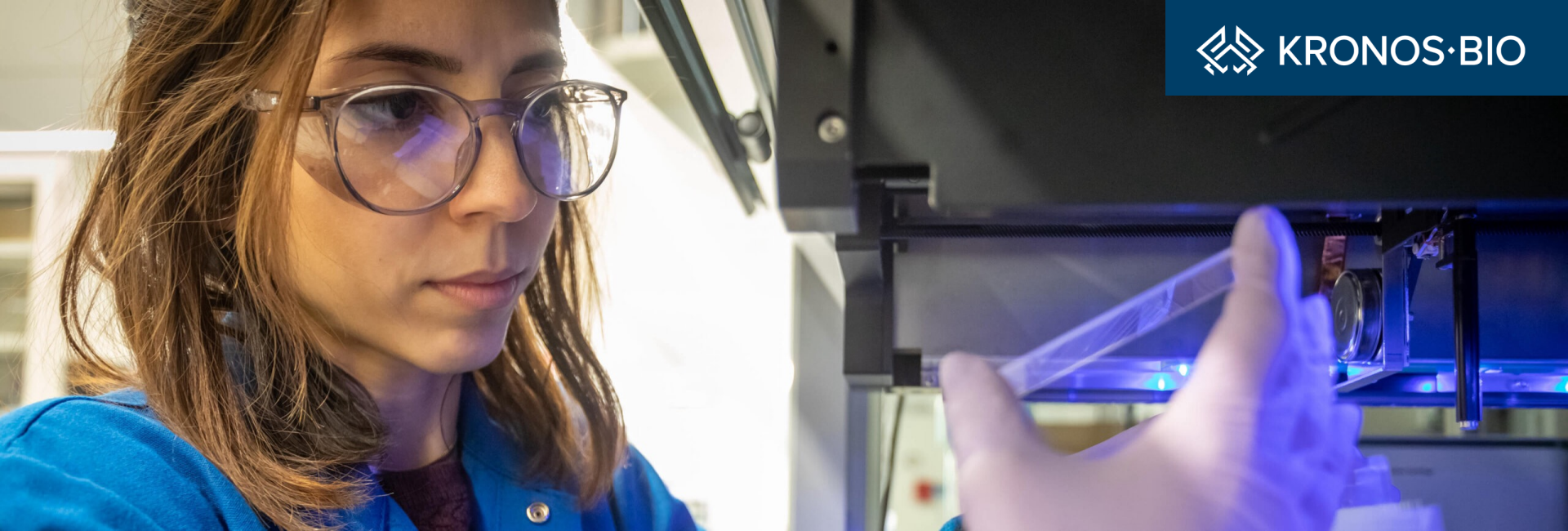
Find Assays

- High throughput screen**
Find all high-throughput screening (HTS) assays.
- Find small molecule microarray (SMM) assays**
- By screener**
Find assays performed by a particular screener.
- By protocol assay search**
Find small molecule assays by screener, assay type, and project description.
- Find Proteins**
- By name or ID**
Find proteins by entering a single criterion that queries name, official symbol, GeneID, UniProt ID, Swiss-Prot ID, RefSeq, Accession #, Entrez Gene ID, or PDB ID criteria.
- By protocol search**
Find proteins using a combination of name, official symbol, GeneID, UniProt ID, Swiss-Prot ID, RefSeq, Accession #, Entrez Gene ID, or PDB ID criteria.

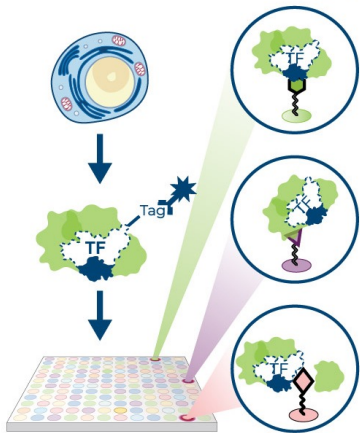
high-signal outlier in a Wnt assay

relationships between assays (protein and phenotype)





Small Molecule Microarray (SMM)

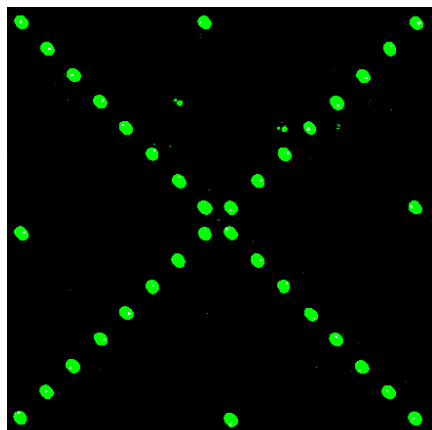


○ SMM hits

TRN	Candidate & Indication		Discovery	IND-enabling	Phase 1	Phase 2	Phase 3
MYC/TF Fusions	KB-0742 (CDK9 Inhibitor)	MYC-amplified solid tumors and other transcriptionally addicted tumors	▶				
IRF4	KB-9558 (p300 KAT inhibitor)	R/R Multiple Myeloma	▶				
MYC	Undisclosed		▶				
β-Catenin	Undisclosed		▶				
Undisclosed	Discovery Collaboration	Genentech <i>A Member of the Roche Group</i>	▶				
Multiple	Undisclosed		▶				

Additional proprietary programs from mapping and screening the MYC, AR, MYB, IRF4 and other TRNs

20.109 MAX screens (Spring 2023)

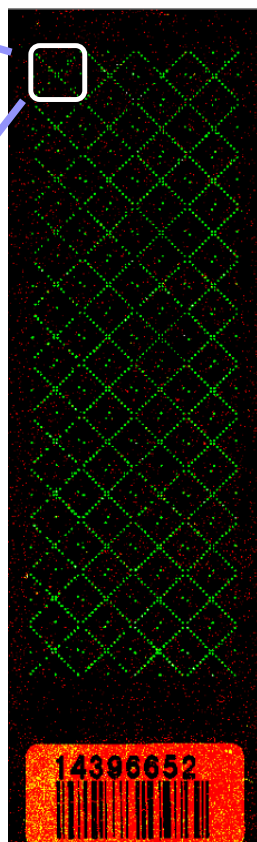


subarray with sentinel pattern for alignment

each team screens
~10,000 unique
compounds

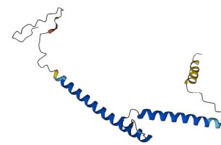
16x16x48 = **12,288**
2 replicate slides

4 replicates for each compound

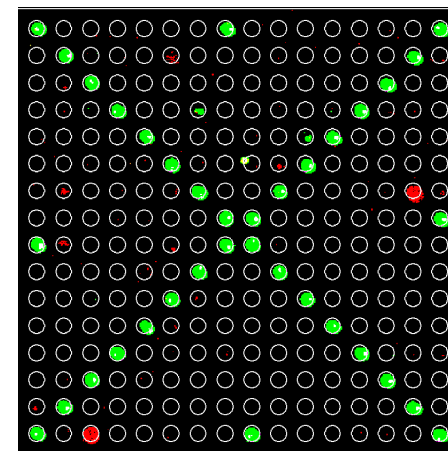


full array with 48
subarrays (4 x 12)

student-
purified
MAX



scan



subarray with 'gal file'
(genepix alignment) file
superimposed



small molecule 'hits'
for your 2024 studies

Upcoming Lectures

2/8/24	Lecture 1	Intro to chemical biology: small molecules, probes, and screens
2/13/24	No Lecture	Snow Day
2/15/24	Lecture 2	Small Molecule Microarrays
2/20/24	No Lecture	
2/22/24	Lecture 3	Our protein target – MAX
2/27/24	Lecture 4	Quantitative evaluation of protein-ligand interactions
2/29/24	Lecture 5	KB-0742: A Phase 2 clinical candidate discovered by SMMs
3/5/23	Lecture 6	Wrap up discussion for Mod 1 experiments and report