

Pragmatic approaches to discover chemistry for biology

John J Irwin

UCSF Pharmaceutical Chemistry
jji@cgl.ucsf.edu

Acknowledgments



University of California
San Francisco

- **Brian Shoichet**
UC San Francisco
 - Michael Mysinger
 - Kelan Thomas
 - Christian Laggner
- **Andrej Sali**
UC San Francisco
 - Eswar Narayanan



UNC
SCHOOL OF MEDICINE

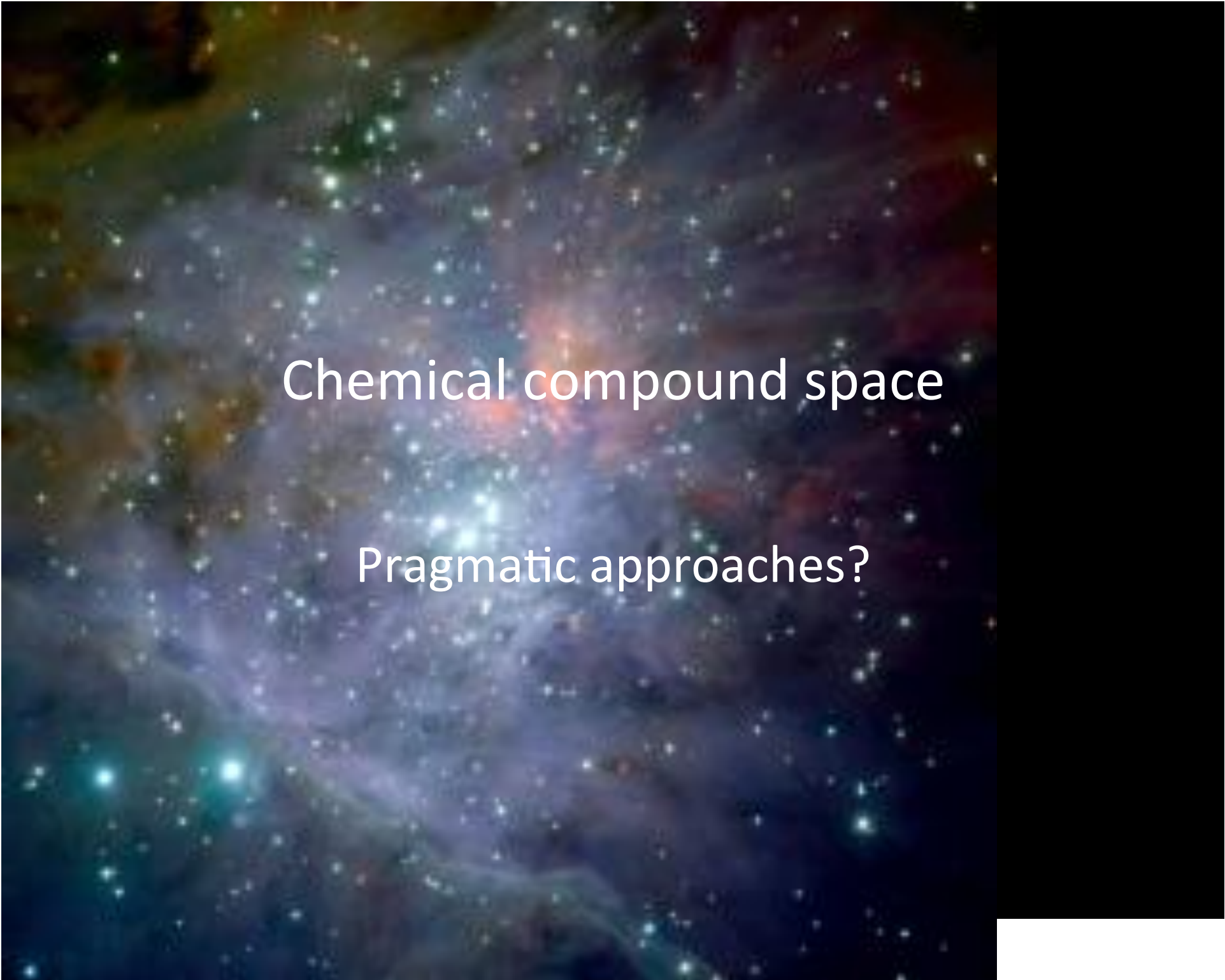
- **Bryan Roth**
University of North
Carolina School of
Medicine
 - Vincent Setola



NOVARTIS

- **Jeremy Jenkins**
Novartis
 - Eugen Lounkine

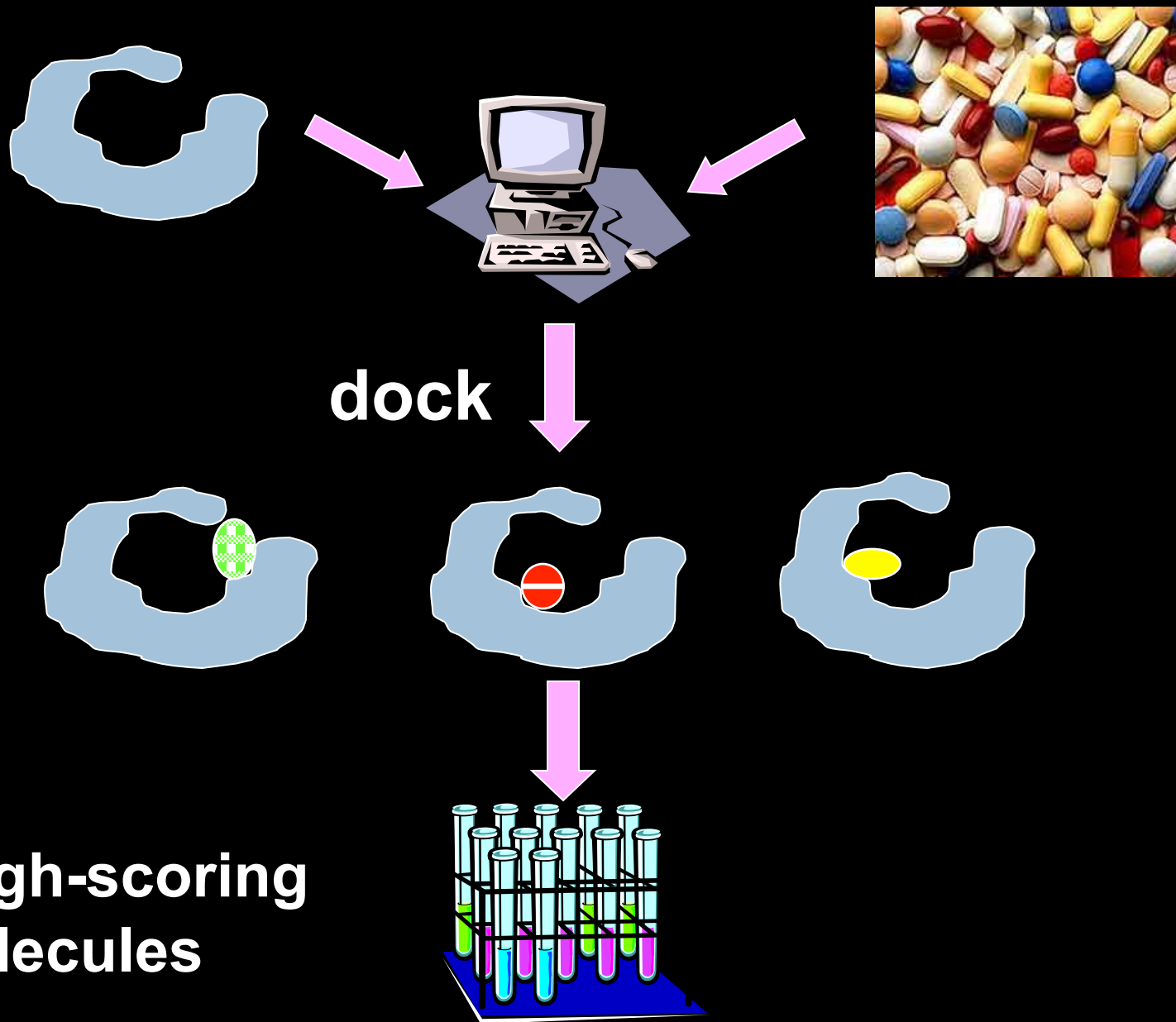
- \$\$ = NIH



Chemical compound space

Pragmatic approaches?

Molecular Docking for Ligand Discovery



**Test high-scoring
molecules**

Web interface for docking screens

DOCK Blaster - A Free V... x

Google

http://blaster.docking.org/start.shtml

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center docking.org

Prepare input

DOCK Blaster

Home | **DOCK!** | Documentation | Mailing Lists

Preparer | Scrutinizer | Target Prep | Calibration | Docking | Results

Q'd Run Done Q'd Run Done Qd Run

Welcome to the DOCK Blaster Job Preparation Form! To start docking, please complete the form below. If you need help, please read the [documentation](#). We urge you to attempt at least one [tutorial](#) using prepared data before attempting to use DOCK Blaster with your own data.

[Leave A Message](#)

[Live Help Offline](#)

Target: rec.pdb

Indicate binding site: xtal-lig.mol2

Actives: actives.smi

Inactives: inactives.smi

Email for reports: (optional)

By clicking "DOCK!" you agree to the [Terms and Conditions](#).

Optional PIN:

The aim of this calculation is...?

Angiotensin II converting enzyme (ACE) from the DUD benchmark. This run uses a ligand to both indicate the binding site and as a control, and includes annotated actives and computed property matched decoys from the DUD benchmark.

Cofactor.par: No file chosen Expert.tar: No file chosen

[Expert Help](#)

A product of [BCIRC](#), the Bioinformatics and Chemical Informatics Research Center @ UCSF. Last updated Feb 11, 2009. questions and discussion to blaster-fans at docking.org; bug reports to support at docking.org; any other correspondence to comments at docking.org. [Terms of use](#). [Privacy policy](#).

Review calibration data

Watching Job #39669

Preparer Scrutinizer Target Prep Calibration Docking Results

Files supplied: Receptor, Ligand

Status: Calibration docking complete

Summary Report of Calibration Docking (What is this?)

Sampling	Scoring	
	Polarized	AMBER
Coarser	0.703/0	0.742/0
Finer	0.704/0	0.876/0

Detailed reports: [plots](#) and [poses and scores](#).

Congratulations! You are now eligible for database docking (How to use this?)

Sampling	Scoring	Database type	Database (#) number of molecules
Coarser	Polarized	By Property	clean-fragments (#12) 120727

If you have any doubts, please consider screening the "clean fragments" (subset 12) first. Three reasons: 1. it will be done faster. 2. fragments are less complex, and therefore more efficient than leads for sampling chemical space. 3. you can still dock the leads afterwards. Still not convinced?

DOCK Blaster - A Free Virtu... ZINC - A free database for... DOCK Blaster - A Free Virtu...

http://blaster.docking.org/cgi-bin/resultsbrowser.pl?job_id=41412&i=decoys&j=1&k=A&pin=

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center

Results Browser Job #41412

Home DOCK! Documentation Mailing Lists

Preparer Scrutinizer Target Prep Calibration Docking Results

Files supplied: [Receptor](#), [Ligand](#)

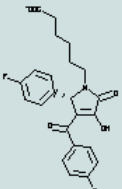
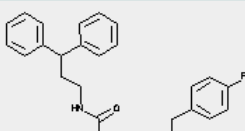
Visualize using: [CHIMERA](#) [PYMOL](#) [Important note about broken molecules.](#)

Download hits: [PDB](#) [SDF](#) [CEL1](#) [mol2](#)

Leave A Message

Live Help Offline

Get Chimeral Get Pymol

Rank	ZINC ID	Compound	Availability	Properties	Annotations
1	9370567 -114.34		Asinex: BAS01029113 IBScreen: STOCK2S-66522 Pharmeks: PHAR041448 ChemBridge: 5951033 Aurora Fine Chemicals: kbsb-0079151	Mwt: 489.317 xLogP: 4.15 Charge: -1 RotBond: 9 # Protomers: 3 Contact: 159 ES: -106.35 VdW: -26.24 Desolv: p=19.61, ap=-1.36	PDB Go SEA!
2	23242305 -101.19		Life Chemicals (Virtual): F5485-0181	Mwt: 486.592 xLogP: 6.35 Charge: -1 RotBond: 10 # Protomers: 2	PDB Go SEA!

Review hits in Chimera...

UCSF Chimera

ViewDock - <http://blaster.docking.org/cgi-bin/getfile4.pl?job=39669,h=09,f=r...>

S	Number	Name	Energy score
V	1	ZINC19517626	-111.14
V	2	ZINC18130364	-109.14
V	3	ZINC31287867	-107.91
V	4	ZINC22178324	-107.67
V	5	ZINC19422208	-106.72
V	6	ZINC20149538	-106.6
V	7	ZINC15746425	-106.22
V	8	ZINC20258718	-106.19
V	9	ZINC22220643	-105.94
V	10	ZINC19480954	-105.73

Chimera Model #1.1

REMARK Number : 1
REMARK Source num : 1
REMARK Name : ZINC19517626
REMARK Description : (2S)-2-(1H-indole-3-carboxylamino)propanoic
REMARK Energy score : -111.14
REMARK van der Waals component : -19.99
REMARK electrostatic component : -113.15
REMARK polar solvation component : 22.97
REMARK apolar solvation component : -0.97

Change Compound State
☒ Viable ☐ Deleted ☐ Purged

Show all downloads...

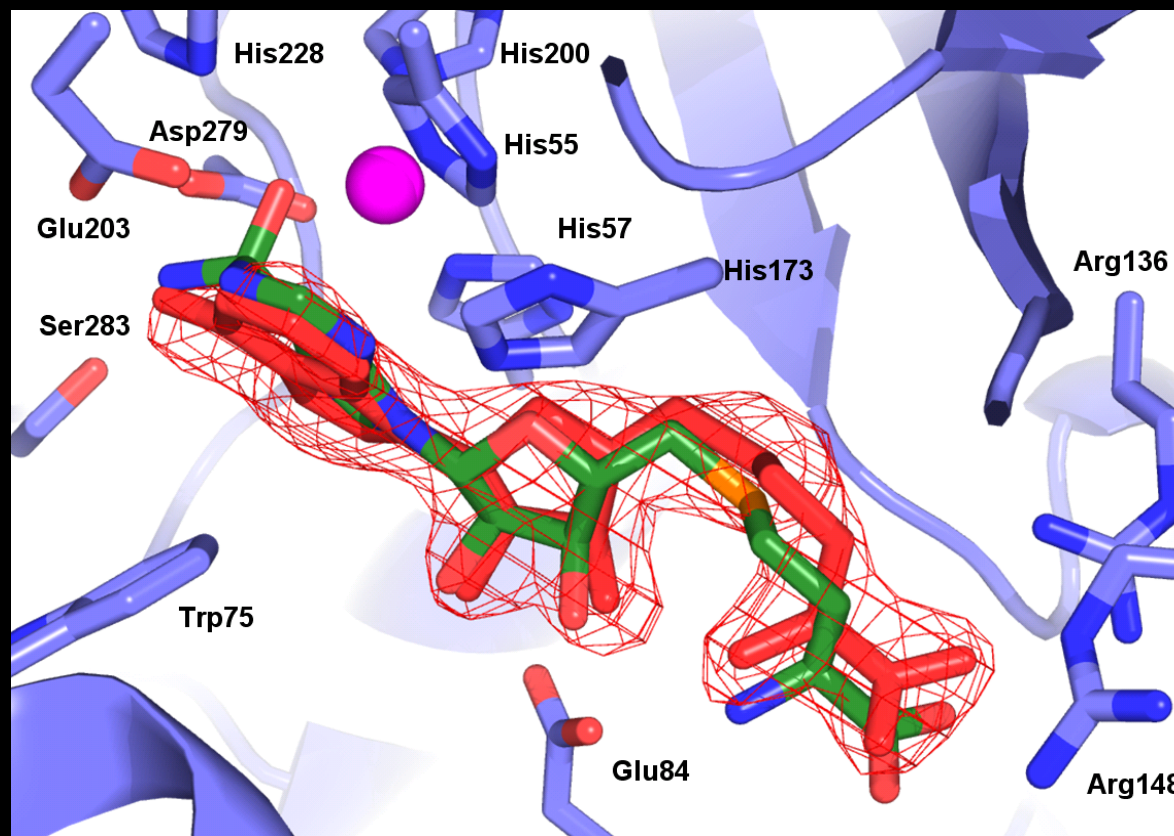
Start

Microsoft PowerP... DOCK Blaster - A... UCSF Chimera ViewDock - http://... Viewing

5:00 PM

The image is a collage of 30 molecular docking images. The top row shows six different ligands (cyan, yellow, and magenta) docked into a protein binding site (grey). The second row shows seven more ligands docked into the same site. The third row shows five ligands docked into the site, with labels for residues: Gln120, Asn152, Tyr150, Ser64, Ala318, Lys315, Water, Phosphate Ion, His228, His200, His55, His57, His173, Arg136, Arg148, Glu203, Asp279, Ser283, Trp75, Glu84, Y150, K315, S64, and Gln102. The bottom row shows five more ligands docked into the site. The images are arranged in a grid-like fashion, with some showing the ligand in a yellow mesh surface and others showing it in a blue mesh surface. The protein is shown in a grey ribbon representation.

Protein Function Prediction by docking high energy intermediates

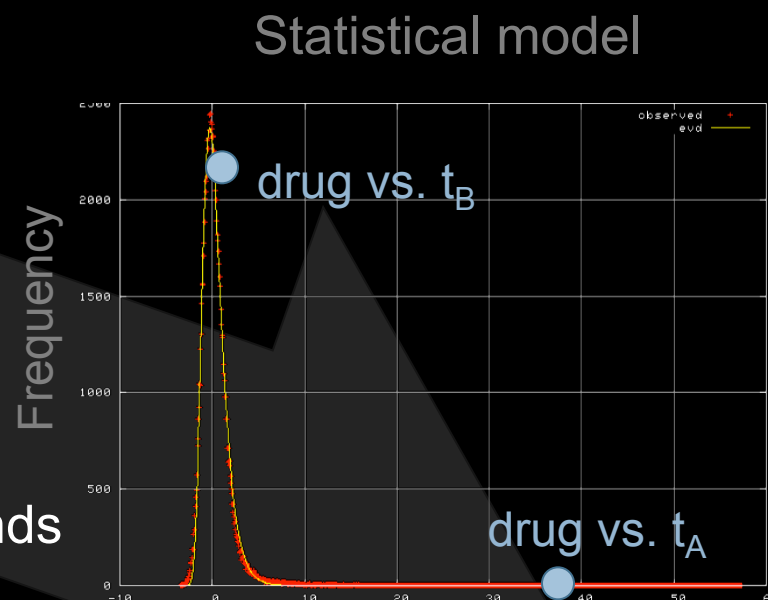
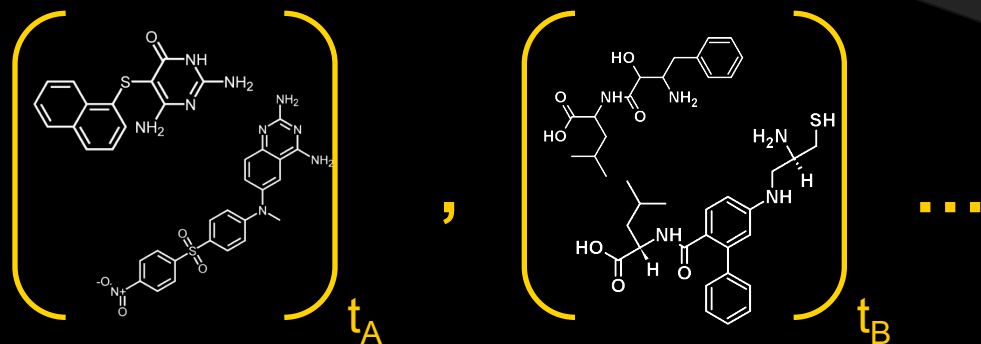


Hermann JC, Marti-Arbona R, Fedorov AA, Fedorov E, Almo SC, Shoichet BK, Raushel FM. Structure-based activity prediction for an enzyme of unknown function. *Nature* 448,775-9 (2007).

Use of DOCK to predict the substrate of an enzyme of previously unknown function.

Technology: Similarity Ensemble Approach (SEA)

- 1** For each drug,
- 2** Compare it against target t_x 's ligands



- Z-score
- 3** Rank by significance

Keiser et al., *Nature Biotech.* 2007; Keiser et al., *Nature*, 2009.



The ZINC Database

<http://zinc.docking.org>

20 million compounds
commercially available
structures calculated
multiple conformations
properties (charge, solv, etc...)
links to suppliers

Free to the community

Multiple subsets

8 M drug-like (Lipinski)

3 M lead-like (Oprea...)

400 K fragment-like (Astex, ...)

Available in popular formats

SMILES, SDF, mol2, flexibase

The screenshot shows the ZINC database homepage. At the top, there's a navigation bar with links to UCSF, About UCSF, Search UCSF, and UCSF Medical Center. The main header features the ZINC logo with the tagline 'A free database for virtual screening' and 'is not commercial'. Below the header, there's a welcome message stating that ZINC contains over 13 million purchasable compounds in ready-to-dock, 3D formats. It mentions that ZINC is provided by the Shoichet Laboratory at UCSF and provides a reference to Irwin and Shoichet (2005). A 'Caveat Emptor' section follows, stating that the database does not guarantee the quality of any molecule. A list of links is provided, including 'About ZINC', 'Search and Browse', 'About BCIRC', 'About Shoichet Lab', 'About UCSF Pharmaceutical Chemistry', 'Acknowledgements', 'User Upload', 'Fun ZINC Facts!', 'How much is ZINC used?', 'I'm Feeling Lucky', and 'Sightings'. The 'Molecule of the month' section displays the chemical structure of ZINC9922817. At the bottom, a footer note states that ZINC is a product of BCIRC, the Bioinformatics and Chemical Informatics Research Center at UCSF, and provides contact information for support and discussion.

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center

Shoichet Laboratory docking.org

Version 10

Home Subsets HELPI Mailing Lists

Welcome to ZINC, a free database of commercially-available compounds for virtual screening. ZINC contains over 13 million purchasable compounds in ready-to-dock, 3D formats. ZINC is provided by the [Shoichet Laboratory](#) in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF). To cite ZINC, please reference: Irwin and Shoichet, *J. Chem. Inf. Model.* 2005;45(1):177-82 [PDE](#), [DOI](#). We thank [NIGMS](#) for financial support (GM71896). There are [release notes](#) for ZINC 10. - We have a [survey](#) where you can give us feedback.

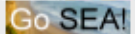
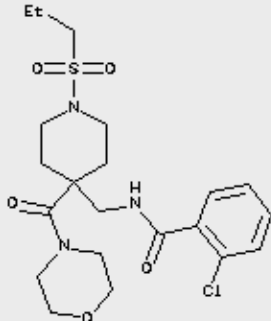
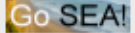
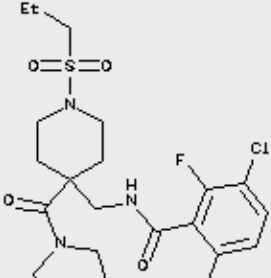
Caveat Emptor: We do not guarantee the quality of any molecule for any purpose and take no responsibility for errors arising from the use of this database. ZINC is provided in the hope that it will be useful, but you must use it at your own risk.

- [About ZINC](#)
- [Search and Browse](#)
- [About BCIRC](#)
- [About Shoichet Lab](#)
- [About UCSF Pharmaceutical Chemistry](#)
- [Acknowledgements](#)
- [User Upload](#)
- [Fun ZINC Facts!](#)
- [How much is ZINC used?](#)
- [I'm Feeling Lucky](#)
- [Sightings](#)

Molecule of the month

[ZINC9922817](#)

A product of [BCIRC](#), the Bioinformatics and Chemical Informatics Research Center @ UCSF. Last updated Aug 6, 2009, questions and discussion to [zinc-fans at docking.org](#); bug reports to support at [docking.org](#); any other correspondence to comments at [docking.org](#). [Terms of use](#), [Privacy policy](#).

ZINC ID	Supplier information; Representations Properties: xLogP, ap & p desolvation, HBD,HBA,Charge,Mwt,NRB Annotations; Similarity	Structure Click for quick 3D display
36143466 #1  FIND SIMILAR  Go SEA!	ChEMBL-08: CHEMBL219714 ChEMBL-08 10uM: CHEMBL219714 ChEMBL-08 1uM: CHEMBL219714 ChEMBL09: CHEMBL219714 PubChem: 11662843...5 total ref: mol2, SDF, SMILES, flexibase 1.71,4.64,-19.78,1,8,0, 472.007, 7 Glycine transporter 1: 9.7 (ChEMBL09)Glycine transporter 1: 18.3 (ChEMBL09)Sodium- and chloride-dependent glycine transporter 1: 95.3 (ChEMBL09) Similar to: 34608435. SEA predictions: ADAM10 pPvalue:40 maxTC:34 (annotated) Muscle glycogen phosphorylase pPvalue:10 maxTC:31 (annotated) Voltage-gated potassium channel subunit Kv1.2 pPvalue:24 maxTC:28 (annotated) Voltage-gated potassium channel subunit Kv1.3 pPvalue:12 maxTC:28 (annotated) Sodium- and chloride-dependent glycine transporter 1 pPvalue:155 maxTC:100 (annotated) ... 16 in total.	
36143438 #2  FIND SIMILAR  Go SEA!	ChEMBL-08: CHEMBL222720 ChEMBL-08 10uM: CHEMBL222720 ChEMBL-08 1uM: CHEMBL222720 ChEMBL09: CHEMBL222720 PubChem: 44420132...5 total ref: mol2, SDF, SMILES, flexibase 1.95,4.57,-20.79,1,8,0, 507.987, 7 Glycine transporter 1: 8.4 (ChEMBL09)Glycine transporter 1: 19.7 (ChEMBL09)Sodium- and chloride-dependent glycine transporter 1: 42.8 (ChEMBL09) SEA predictions: ADAM10 pPvalue:25 maxTC:32 (annotated)	

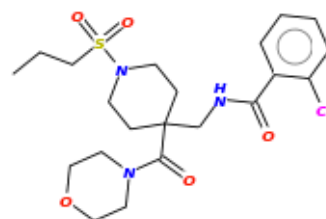
sea viewer

Query Info

Code 75.18.164.137

Name SEArch 75.18.164.137

Size 1 (of 1)



[Back to Index](#)

Page << < 1 of 3 > >>

Code	#Ligands	Reference Name	E-value	MaxTC	Molecules
1 SC6A9_MOUSE_100	6	Sodium- and chloride-dependent glycine transporter 1 [100 nM]	6.38e-130	1.00	view >>
2 SC6A9_RAT_100	6	Glycine transporter 1 [100 nM]	6.38e-130	1.00	view >>
3 SC6A9_HUMAN_1000	202	Glycine transporter 1 [1000 nM]	8.02e-112	1.00	view >>
4 SC6A9_HUMAN_10000	216	Glycine transporter 1 [10000 nM]	1.13e-109	1.00	view >>
5 SC6A9_HUMAN_100	140	Glycine transporter 1 [100 nM]	7.73e-98	1.00	view >>
6 SC6A9_HUMAN_10	62	Glycine transporter 1 [10 nM]	3.76e-83	1.00	view >>
7 ADA10_HUMAN_100	122	ADAM10 [100 nM]	1.98e-33	0.34	view >>
8 ADA10_HUMAN_1000	183	ADAM10 [1000 nM]	4.92e-33	0.34	view >>
9 ADA10_HUMAN_10000	190	ADAM10 [10000 nM]	6.07e-33	0.34	view >>
10 Q8DNV6_STRR6_1000	11	UDP-N-acetylmuramoylalanine-D-glutamyl-lysine-D-alanyl-D-alanine ligase [1000 nM]	3.20e-29	0.31	view >>
11 Q8DNV6_STRR6_10000	32	UDP-N-acetylmuramoylalanine-D-glutamyl-lysine-D-alanyl-D-alanine ligase [10000 nM]	4.26e-29	0.32	view >>
12 SC6A9_HUMAN_1	7	Glycine transporter 1 [1 nM]	7.74e-21	0.34	view >>

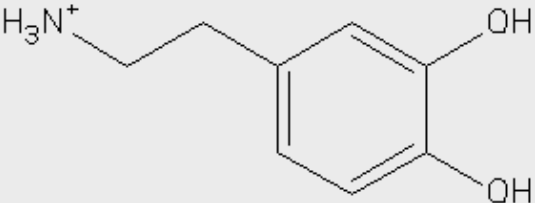
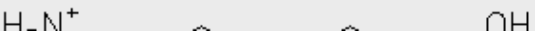
Initial search took 257 second(s).

Beginning : #1 Hits: 22 Display: per page= 100

[Next Page](#) [Download Table](#) [Purchasing Info](#)

Download: [SMILES](#) [MOL2](#) [SDF](#) [Flexibase](#) ☐ Show Detail?

Done

ZINC ID	Supplier information; Representations Properties: xLogP, ap & p desolvation, HBD, HBA, Charge, Mwt, NRB Annotations; Similarity	Structure Click for quick 3D display
33882 #1  Go SEA!	Acros Organics: 12200 Alfa-Aesar: A11136 APiChem: AC-12190 Bosche Scientific: D1738 Enamine: T5885989...95 total ref: mol2, SDF, SMILES, flexibase -0.05,-1.62,-50.24,5,3,1, 154.189, 2 SupplierCAS: 62-31-7 (SourceSDF)Dopamine D1 receptor: 192 (ChEMBL09)Dopamine D2 receptor: 1860 (ChEMBL09)Dopamine D1 receptor: 4300 (ChEMBL09)Dopamine D4 receptor: 15 (ChEMBL09)Delta opioid receptor: 3045 (ChEMBL09)Kappa opioid receptor: 3045 (ChEMBL09)Mu opioid receptor: 3045 (ChEMBL09)Sigma opioid receptor: 3045 (ChEMBL09)Dopamine D3 receptor: 13.2 (ChEMBL09)Dopamine transporter: 6400 (ChEMBL09)Dopamine D5 receptor: 460 (ChEMBL09) ... 32 in total. Similar to: 6091848, 36489899. SEA predictions: Beta-1 adrenergic receptor pPvalue:43 maxTC:29 (annotated) Insulin-like growth factor binding protein 3 pPvalue:20 maxTC:31 (annotated) Enoyl-[acyl-carrier-protein] reductase pPvalue:35 maxTC:33 (annotated) Dopamine D4 receptor pPvalue:10 maxTC:61 (annotated) Arachidonate 15-lipoxygenase pPvalue:50 maxTC:48 (annotated) ... 17 in total.	
	Acros Organics: 16605 Labotest: LT0333117 National Cancer Institute: 000001	

ZINC¹²

John Irwin – options sign out

(Default Cart – Empty)

[About](#) [Search](#) [Browse](#) [Help](#)

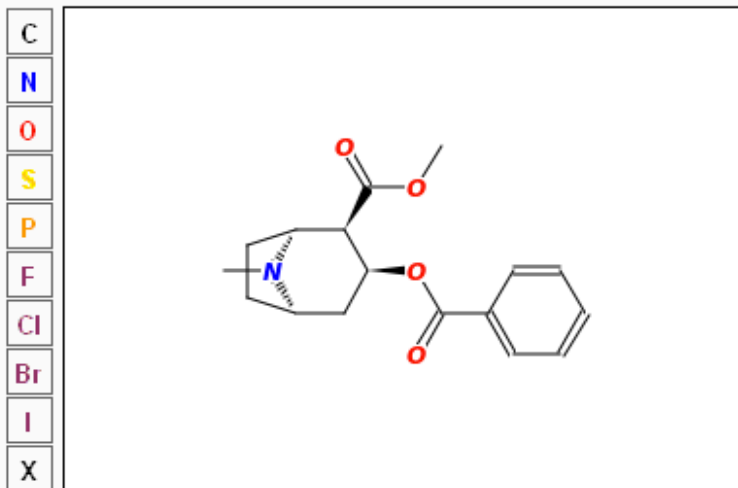
ZINC ID or SMILES

Go

[Structure](#) [Properties](#) [Vendors](#) [ZINC](#) [Advanced](#)

100 Results per page Result format: Summary List

Run Query



SMILES

Beautify

Clear

COC(=O)[C@H]2[C@@H](OC(=O)c1ccccc1)C[C@@H]3CC[C@H]2N3

90%

Seachange
PHARMACEUTICALS

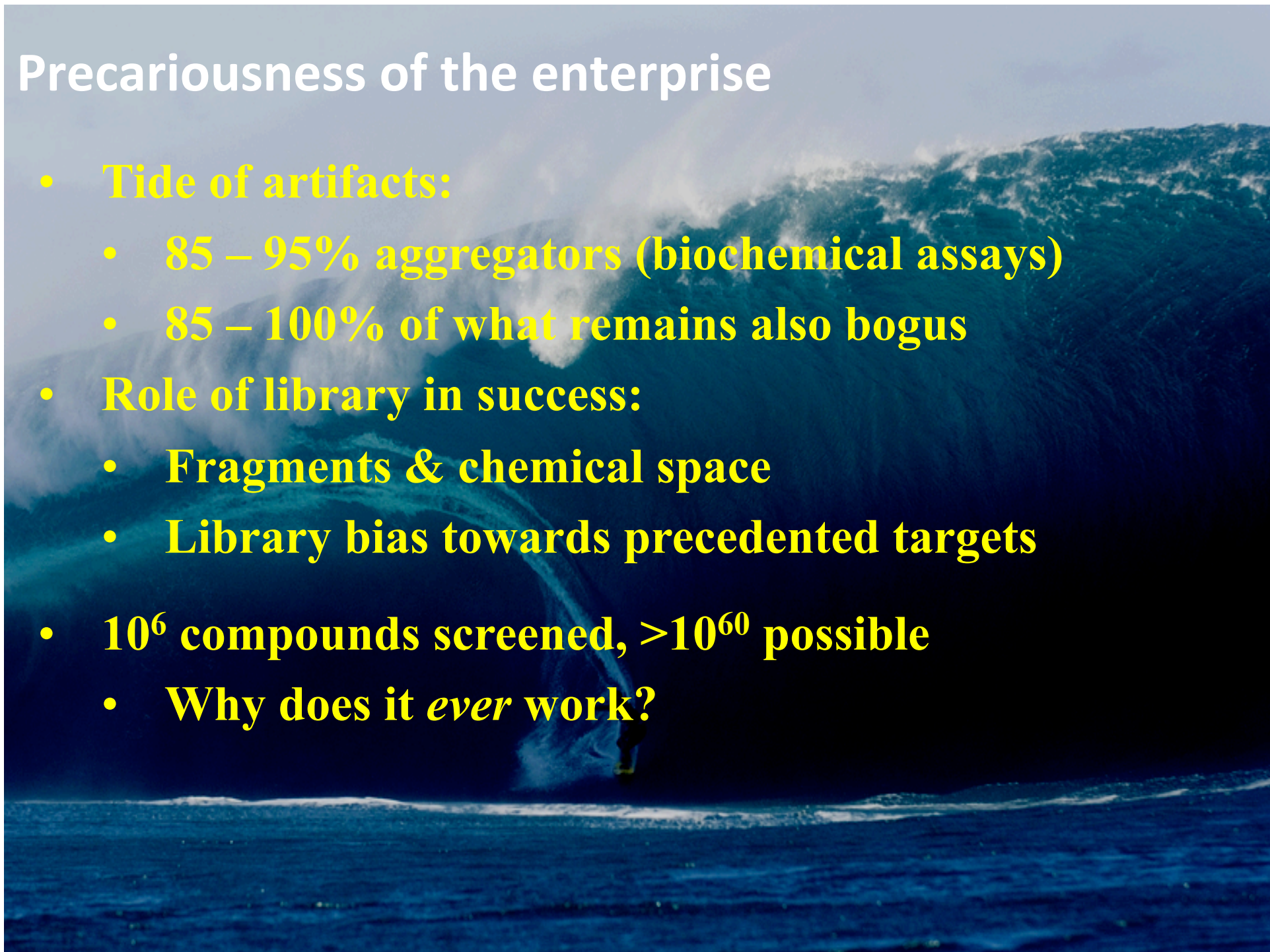
Library bias toward GPCRs

Target	ZINC molecules similar to known ligands	Docking screen	
		Hit rate	potency (nM)
Adenosine receptors	4240	35%	200
Adrenergic receptors	4146	24%	9
Adenylyl Cyclases	565	4%	50,000
AmpC β - lactamase	545	2–5%	26,000

Carlsson et al., JMC 2010

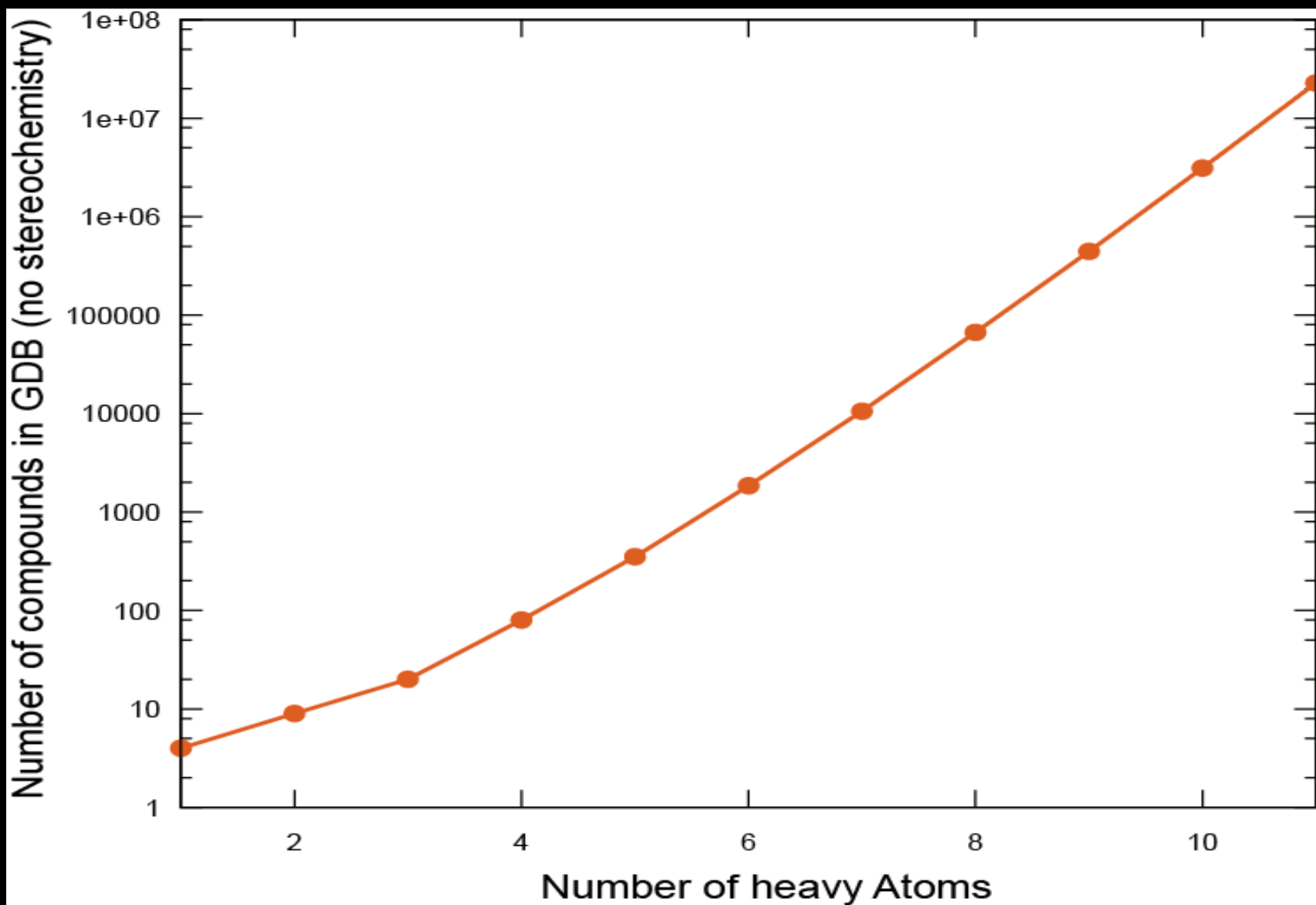
Precariousness of the enterprise

- Tide of artifacts:
 - 85 – 95% aggregators (biochemical assays)
 - 85 – 100% of what remains also bogus
- Role of library in success:
 - Fragments & chemical space
 - Library bias towards precededented targets
- 10^6 compounds screened, $>10^{60}$ possible
 - Why does it *ever* work?



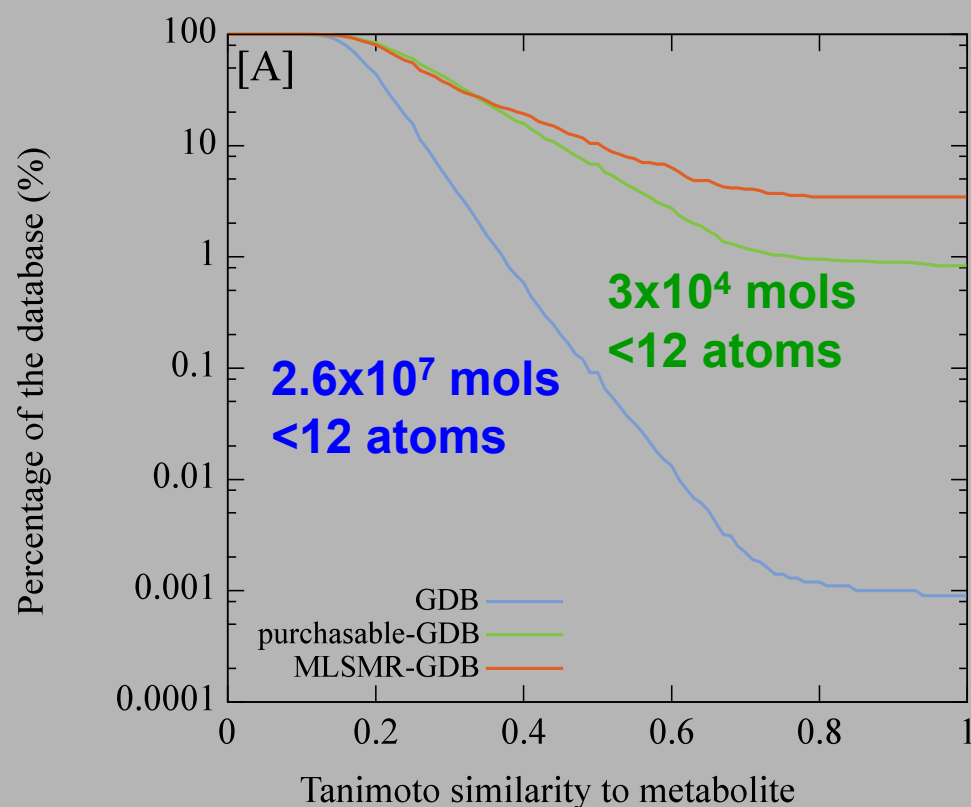
Enumerating *all* of chemical space: the GDB

Molecules composed of H, C, N, O, F; <12 atoms

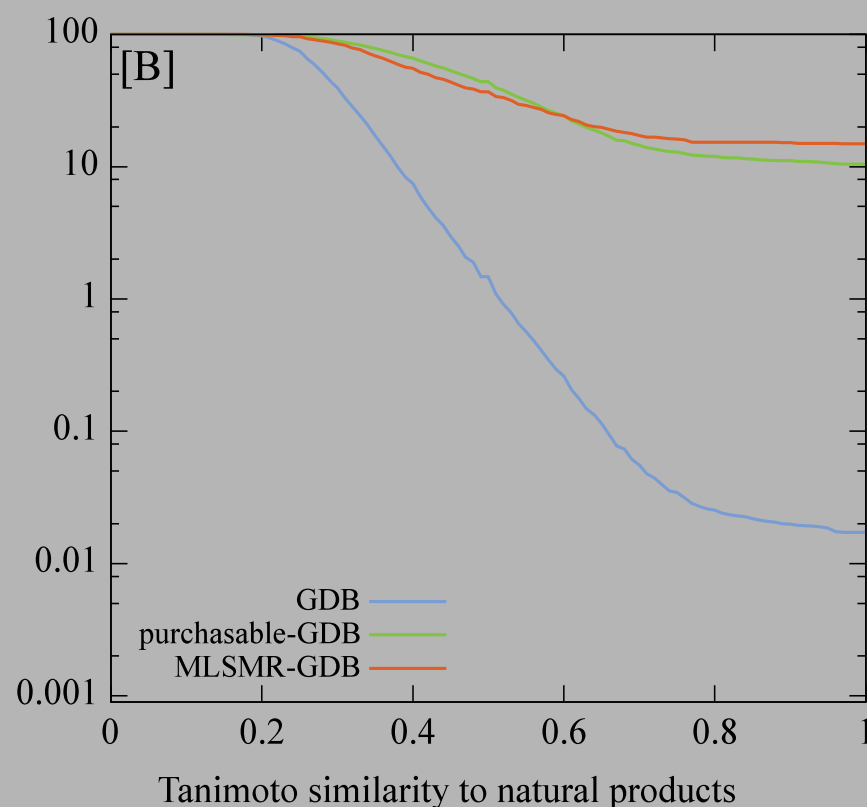


1000-fold *Bias* towards biogenic molecules: Similarity of **purchasable** to biogenic molecules vs. **GDB**

GDB & **p-GDB** vs. Metabolites

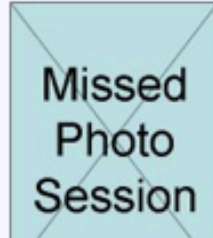


GDB & **p-GDB** vs. Natural Products



To further increase bias, fill the biogenic holes
(scaffolds at once novel *and* precedented)

Natural product	Purchasable Nearest Neighbor	Natural product	Purchasable Nearest Neighbor

ChEMBL 	ChEMBL09 no email p.no phone f.no fax	2011-02-22	84535 ▼ 45097 NYA	90%: 44925 80%: 25334 70%: 13949 60%: 6325
DrugBank Experimental	DrugBank-experimental not for sale p.phone f.not for sale	4854 2011-02-15	3438 ▼ 897 122 NYA	90%: 1685 80%: 1334 70%: 1022 60%: 717
DrugBank Nutriceuticals	DrugBank-nutriceuticals not for sale p.phone f.not for sale	66 2008.5	75 ▼ 4 4 NYA	90%: 48 80%: 38 70%: 32 60%: 26
DrugBank Drugs	DrugBank-small molecules not for sale p.phone f.not for sale	1382 2011-02-15	1488 ▼ 254 35 NYA	90%: 900 80%: 760 70%: 630 60%: 473
	FDA-approved drugs (via DSSTOX) no email p.N/A f.N/A	1217 2005-03-01	2041 ▼ 28 190 NYA	90%: 784 80%: 650 70%: 537 60%: 392
	Human Metabolome Database not for sale p.phone f.not for sale	7886 2008.5	5406 ▼ 4280 311 NYA	90%: 1340 80%: 749 70%: 443 60%: 270
	Street drugs no email p.no phone f.no fax	172 2009.4	167 ▼ 0 4 NYA	90%: 83 80%: 70 70%: 52 60%: 41
	TCM Database @ Taiwan info at tcm.cmu.edu.tw p.no phone f.no fax	3897 2011-03-07	754 ▼ 1411 34 NYA	90%: 447 80%: 298 70%: 191 60%: 124

Summary

A free database for virtual screening

ZINC
is not commercial

Home ▾

Subsets ▾

HELP! ▾

Mailing Lists ▾

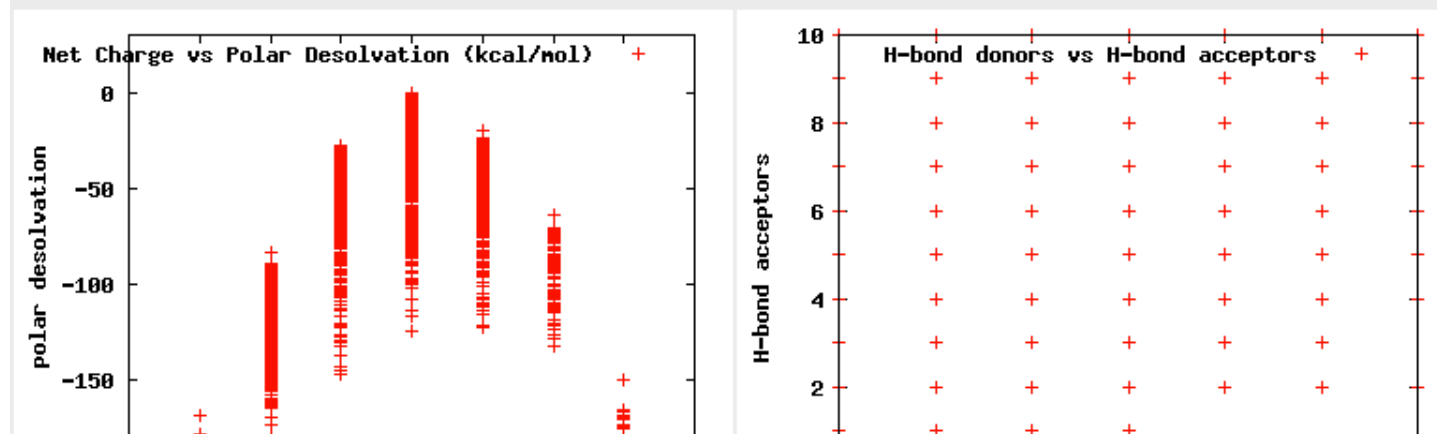
"DNP Natural Products" subset (# 23) - 19881 entries

Contents: [General Information](#) [Property Distributions](#) [Clustering and Diversity](#) [Downloads](#)

General Information

"Usual" Subset Downloads: [mol2](#) | [SDF](#) | [Flexibase](#) | [more...](#)

Property Distributions



How can we predict (off-target) binding?

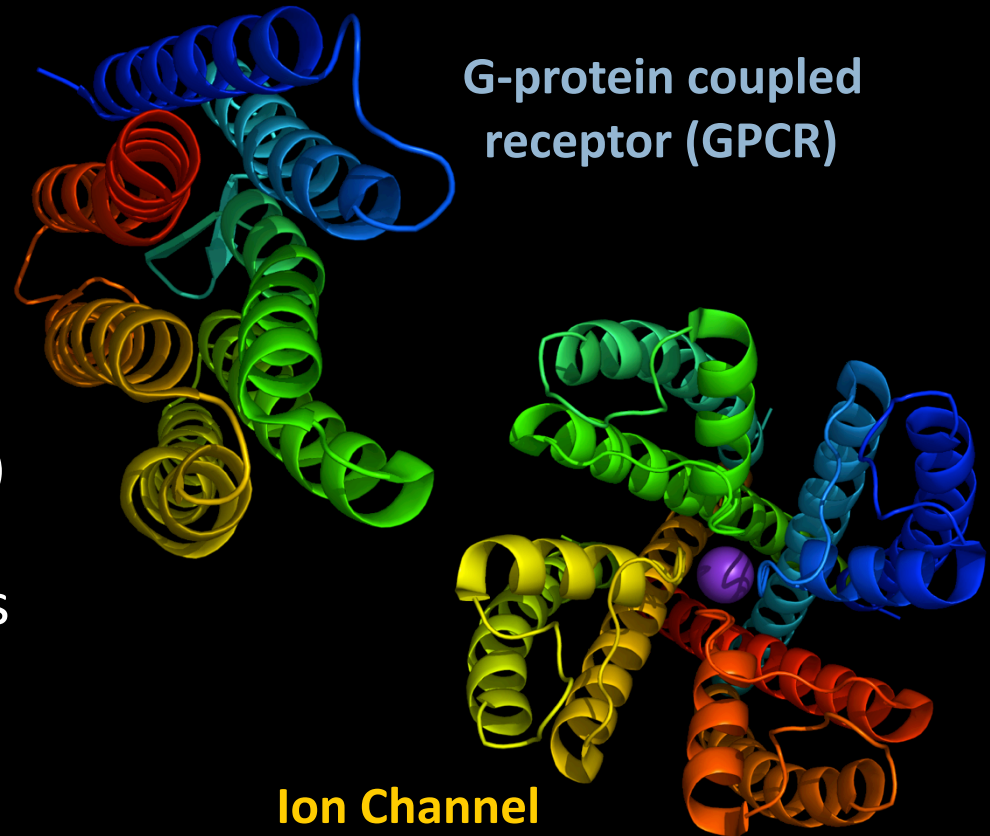
- We relate drugs to proteins **pharmacologically**.
 - We represent proteins solely by their ligands.
 - This works even when the protein targets are unrelated by biological criteria.
 - We can predict **new targets for known drugs**.

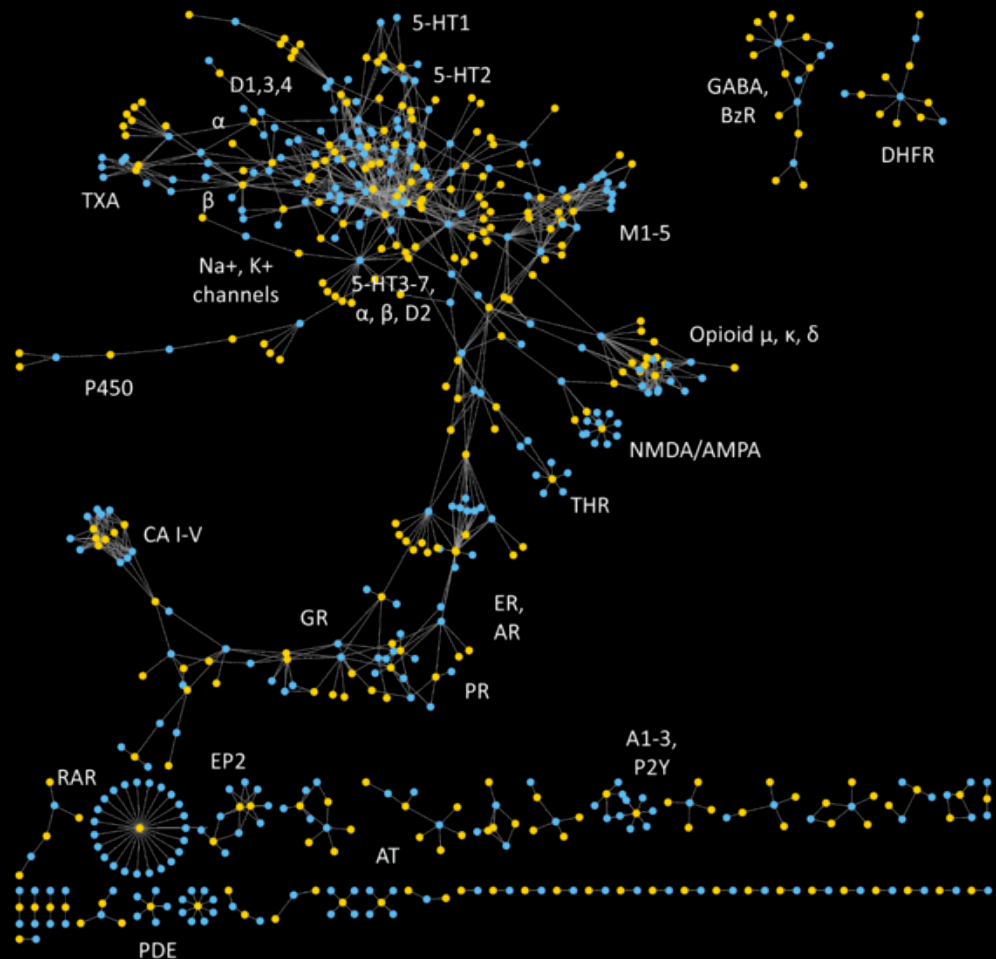
Keiser MJ, Setola V, et al. Predicting new molecular targets for known drugs. *Nature* 462 (7270), 175-81 (2009).

Keiser MJ, et al. Relating protein pharmacology by ligand chemistry. *Nat Biotech* 25 (2), 197-206 (2007).

Core idea: Relate proteins by their ligands, not by sequence or structure

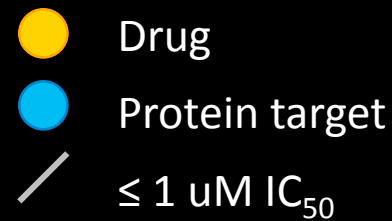
- Many drugs and ligands act through dissimilar proteins.
- Methadone
 - μ opioid receptor (GPCR)
 - NMDA receptor (Ion Channel)
- Serotonin and serotonergics
 - 5-HT_{1-2,4-7} receptors (GPCR)
 - 5-HT₃ receptor (Ion Channel)
- Ligands are the functional link between drug targets.





Target key

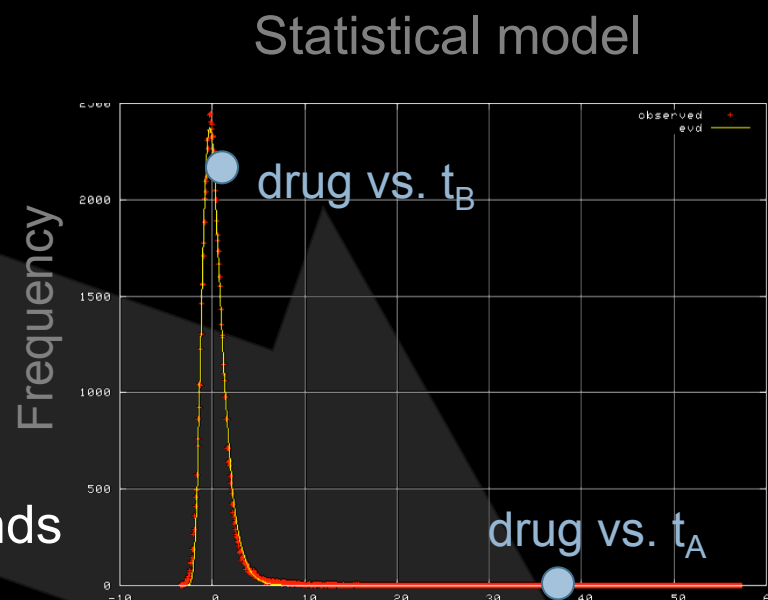
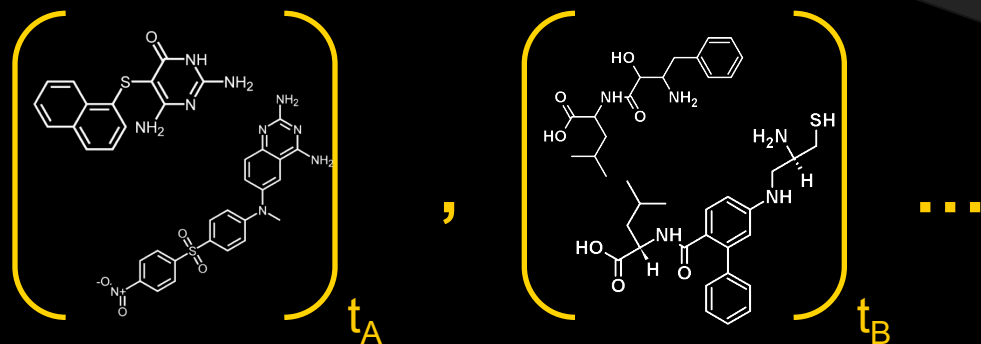
α	α adrenergic
β	β adrenergic
δ	δ-opioid
κ	κ-opioid
μ	μ-opioid
5-HT	Serotonin
A1-3	Adenosine A1-A3
AMPA	α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate
AR	Androgen
AT	Angiotensin II
BzR	Peripheral-type benzodiazepene
CA I-V	Carbonic anhydrase I-V
D1-4	Dopamine D1-D4
DHFR	Dihydrofolate reductase
EP2	Prostaglandin E2
ER	Estrogen
GABA	γ-aminobutyric acid
GR	Glucocorticoid
K+	Potassium channel
M1-5	Muscarinic M1-M5
Na+	Sodium channel
NMDA	N-methyl-D-aspartic acid
P2Y	P2Y purinoceptor
P450	Cytochrome P450
PDE	Phosphodiesterase
PR	Progesterone
RAR	Retinoic acid
THR	Thyroid hormone
TXA	Thromboxane A2 / Prostaglandin H2



Known polypharmacology.
Can it be expanded?

Technology: Similarity Ensemble Approach (SEA)

- 1** For each drug,
- 2** Compare it against target t_x 's ligands

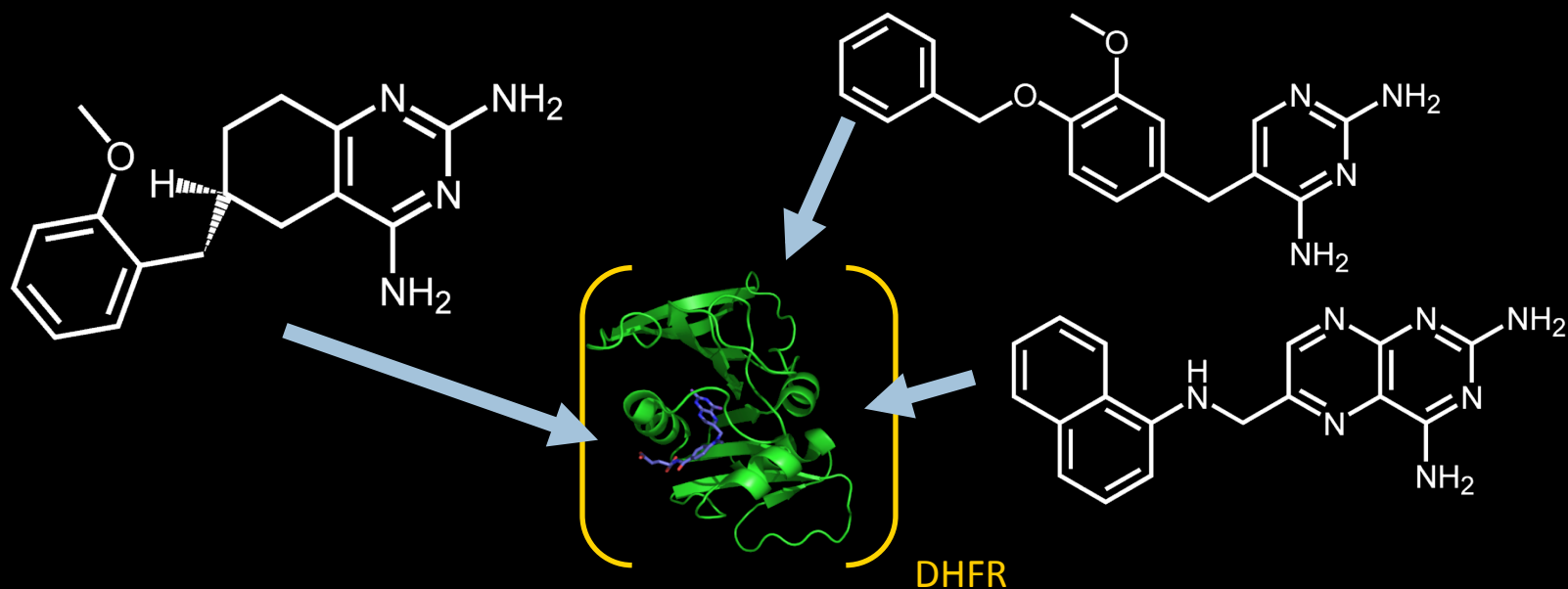


- Z-score
- 3** Rank by significance

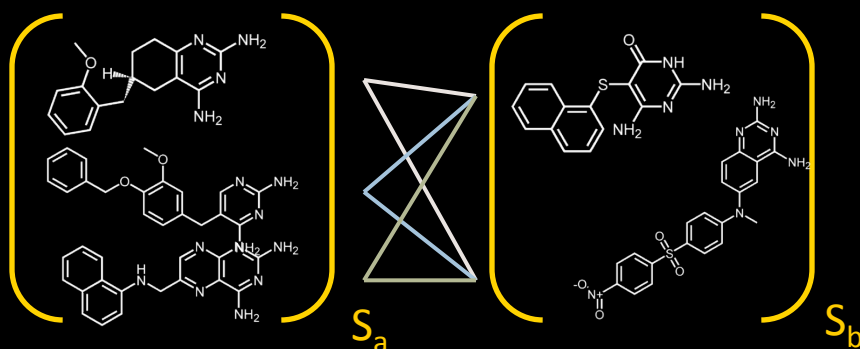
Where do we get the ligands?

- From compound activity databases

Database	Ligands	Targets	Data points
MDDR	201,761	631	391,406
WOMBAT	254,679	2,100	760,605
ChEMBL-protein	222,177	3,153	600,165
ChEMBL-complete	578,715	7,493	2,787,240



How can we compare two ligand sets?



$$TC(a_1, b_1) = 0.91$$

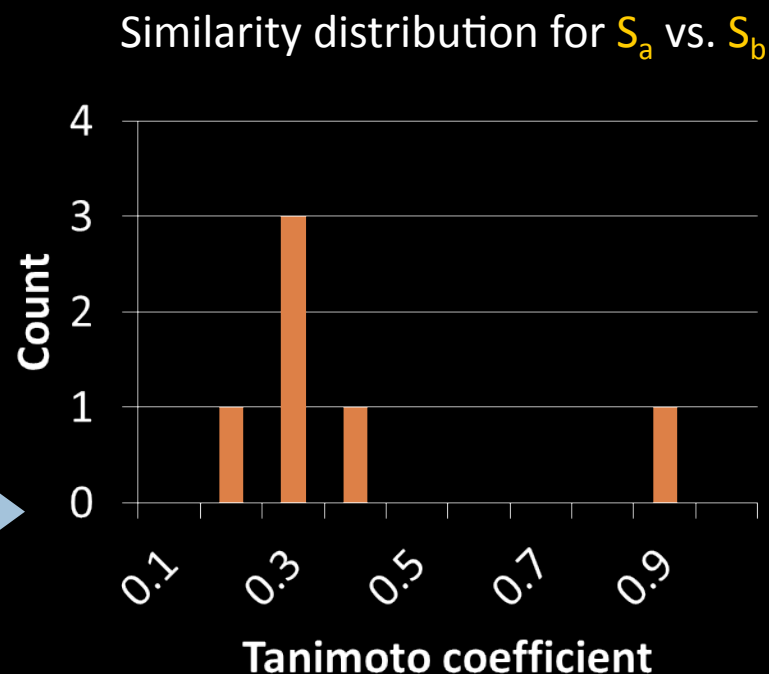
$$TC(a_1, b_2) = 0.34$$

$$TC(a_2, b_1) = 0.28$$

$$TC(a_2, b_2) = 0.31$$

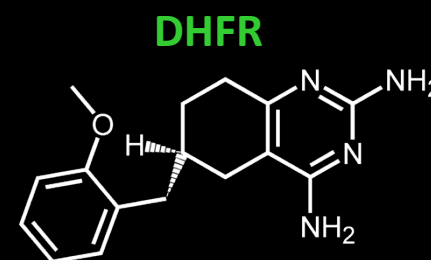
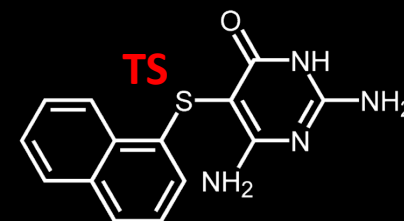
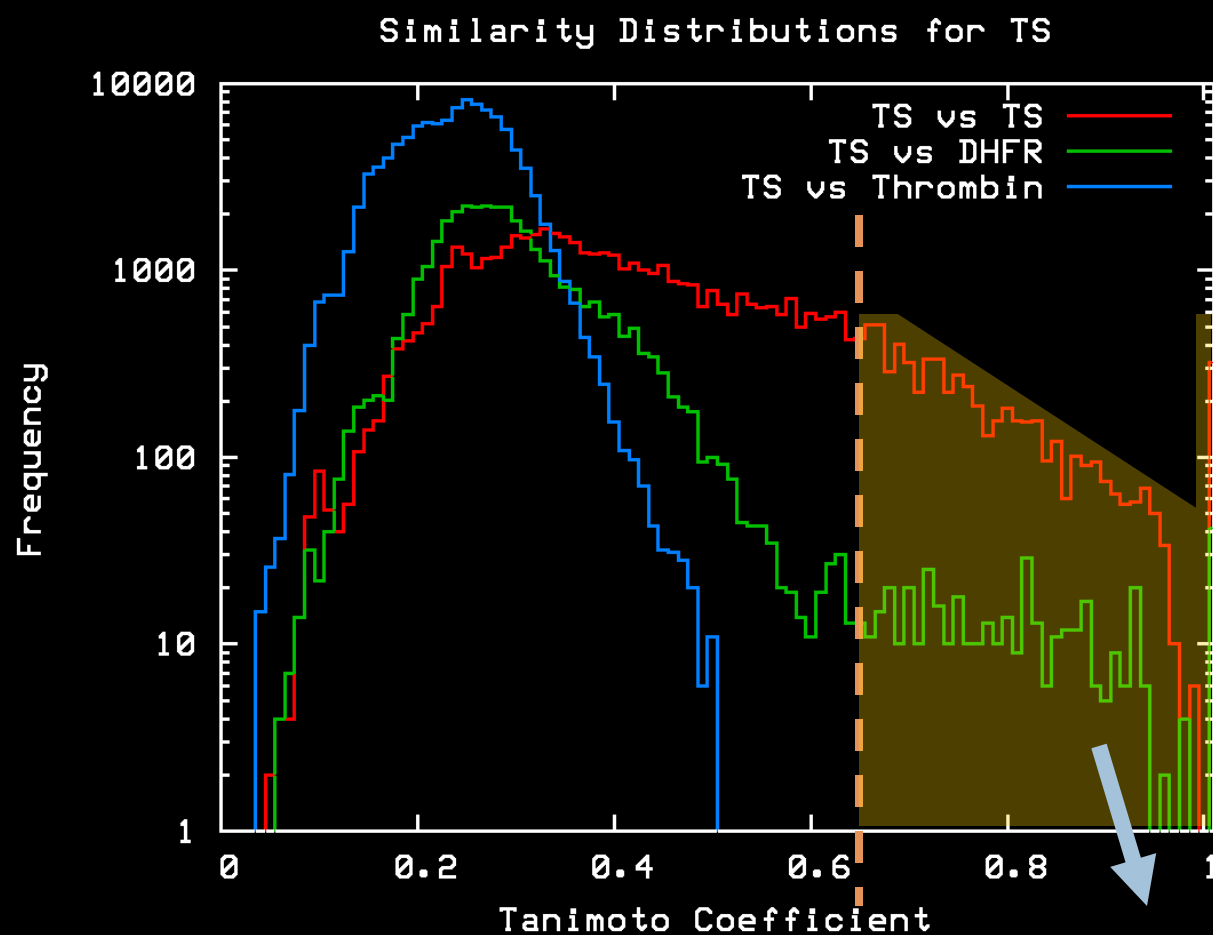
$$TC(a_3, b_1) = 0.43$$

$$TC(a_3, b_2) = 0.39$$



TC = Tanimoto coefficient, using ECFP4 fingerprints

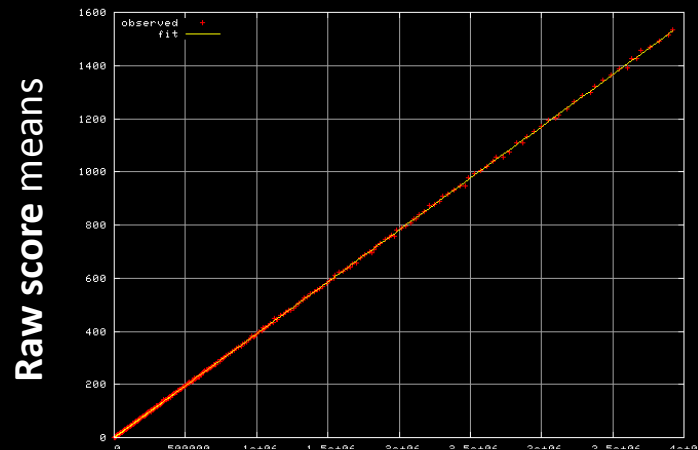
Distributions: Thymidylate synthetase



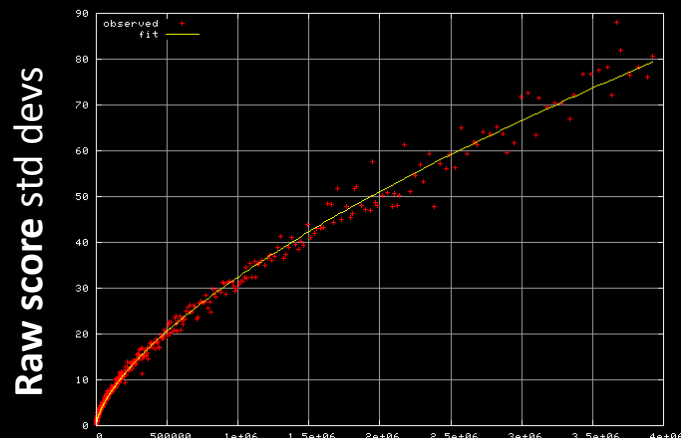
“Raw Score”

Use statistics to find unlikely similarity

Random raw scores

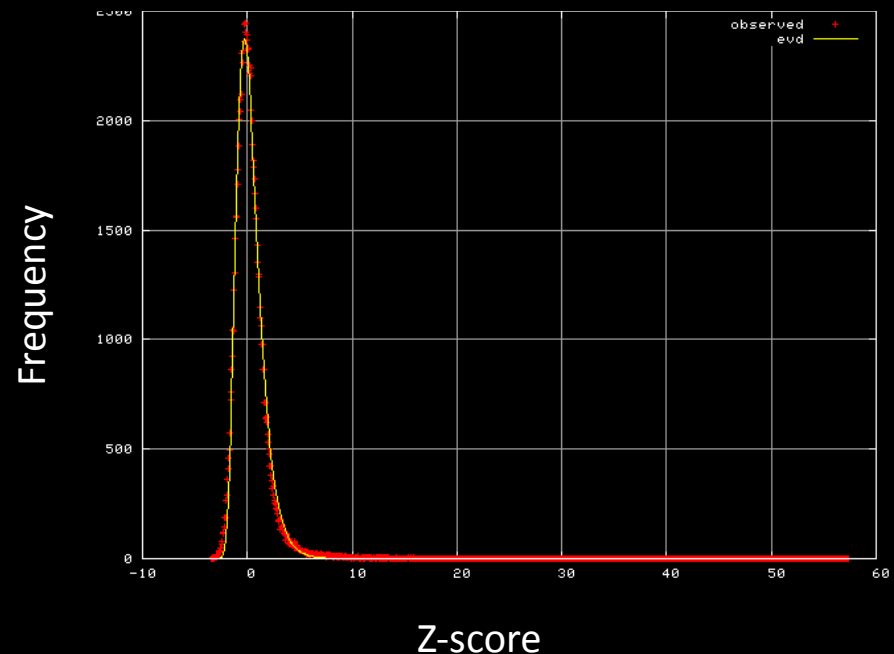


RS Standard deviations



Product of set sizes

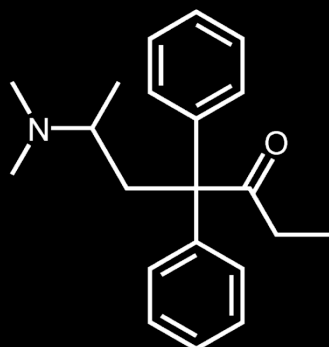
Extreme Value Distribution (EVD)



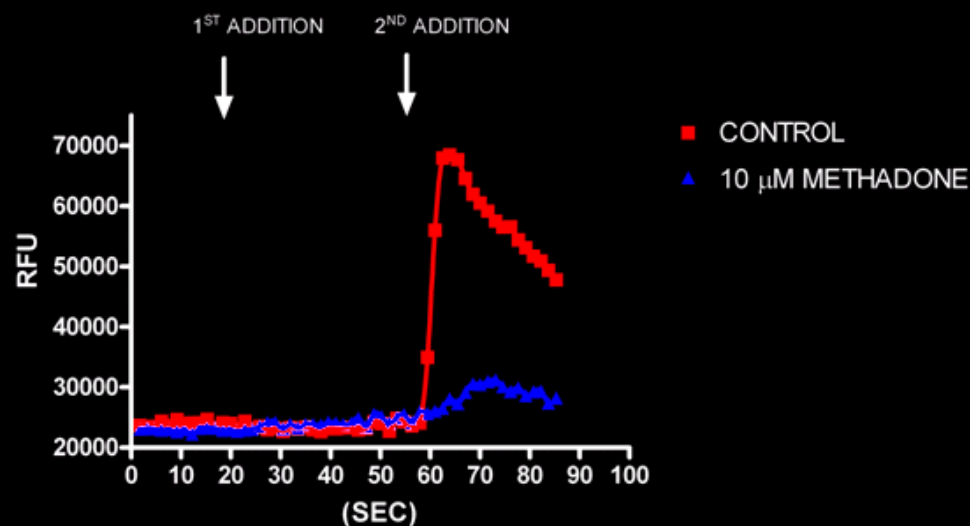
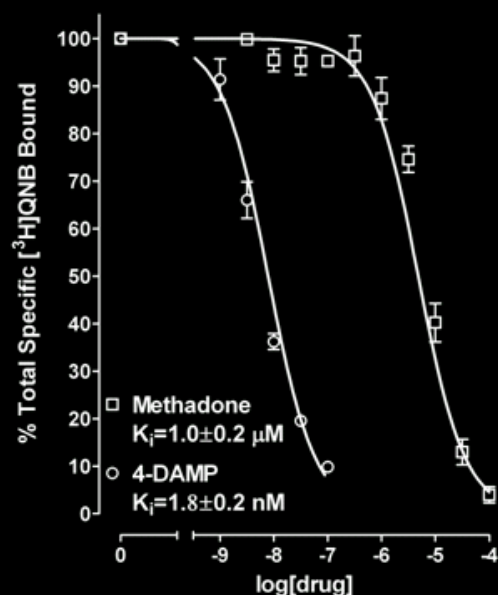
Probability of any given set comparison's score happening by random chance alone:

$$p(z) = 1 - e^{-e^{-z\pi/\sqrt{6}-\gamma}}$$

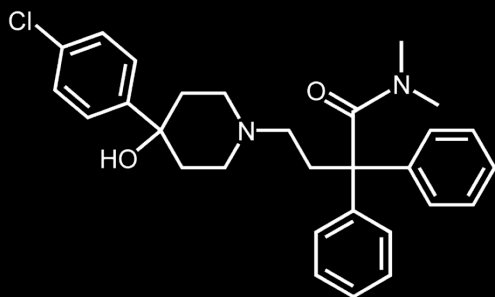
First experimental tests of predictions: Methadone as an antimuscarinic



	Activity	E-value	Max TC
1	Antimuscarinic	4.45e-50	0.77
2	Muscarinic M3 Antagonist	1.22e-11	0.67
3	Opioid Agonist	1.84	0.61
4	NMDA Receptor Antagonist	9.04	0.67

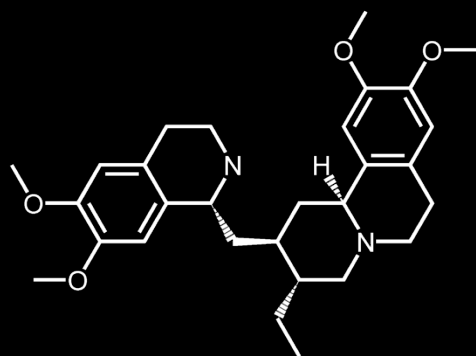


First experimental tests: Emetine & Imodium



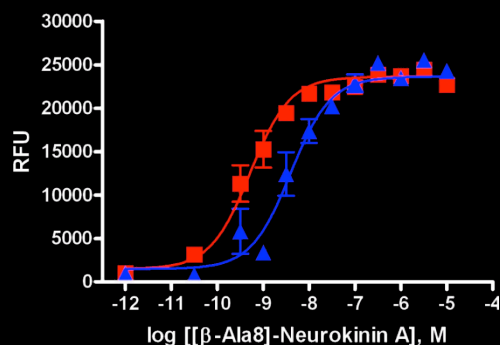
Imodium

	Activity Class	E-Value
1	Neurokinin NK2 Antagonist	1.55e-20
2	Substance P Antagonist	2.12e-15
3	Neurokinin NK3 Antagonist	2.63e-14

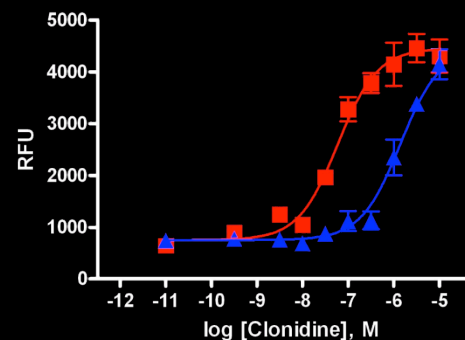


Emetine

	Activity Class	E-Value
1	Adrenergic (α2) Blocker	4.34e-118
2	Dipeptidyl Aminopeptidase IV Inhibitor	6.50e-17
3	Dopamine (D1) Antagonist	1.23e-10



Imodium
Ref. antag.

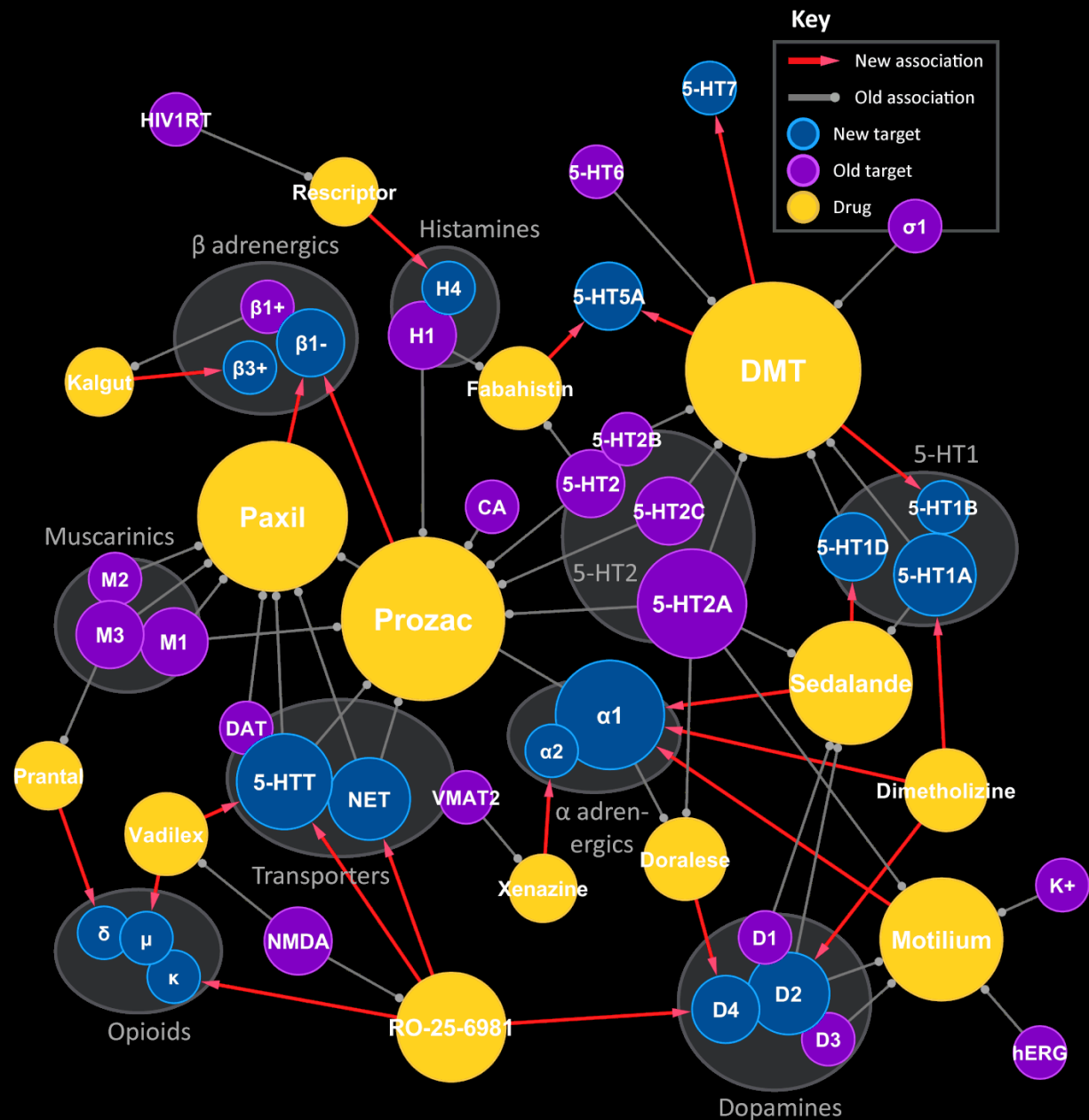


Emetine
Ref. antag.

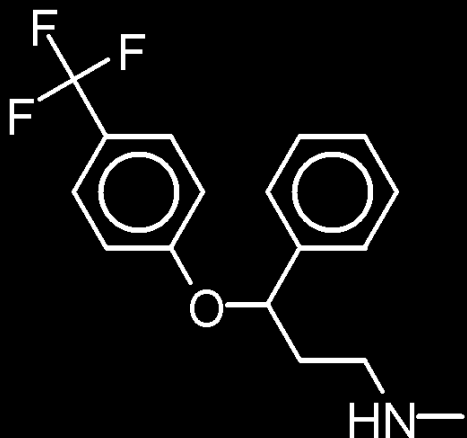
Keiser et al, *Nature Biotech*, 2007

We've
found 23+
novel
off-targets
for known
drugs

(...of 20 assays,
15 at < 1 μ M
5 at < 100 nM)



Screen for new off-targets: Side effects

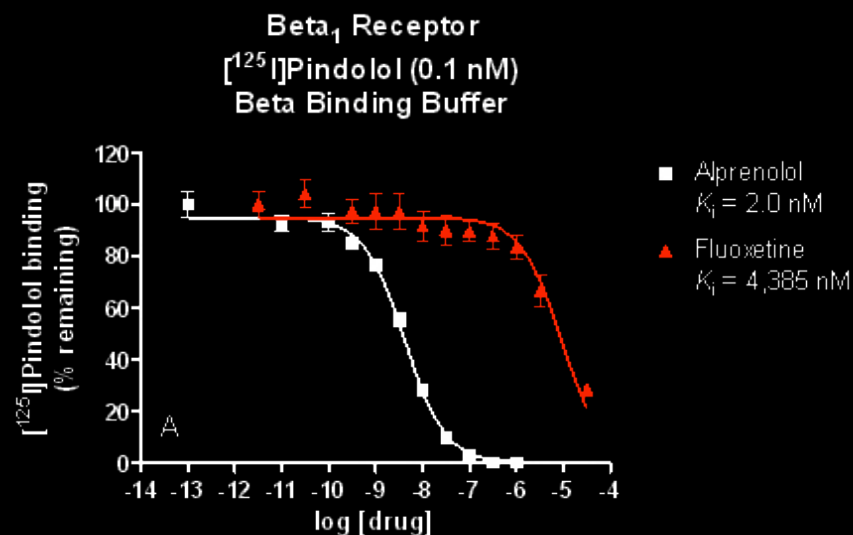


Prozac binds β_1 (4 μM)

Paxil binds β_1 (10 μM)

Prozac's serum plasma levels: 0.3-1.6 μM
Target consistent with SSRI side effects.

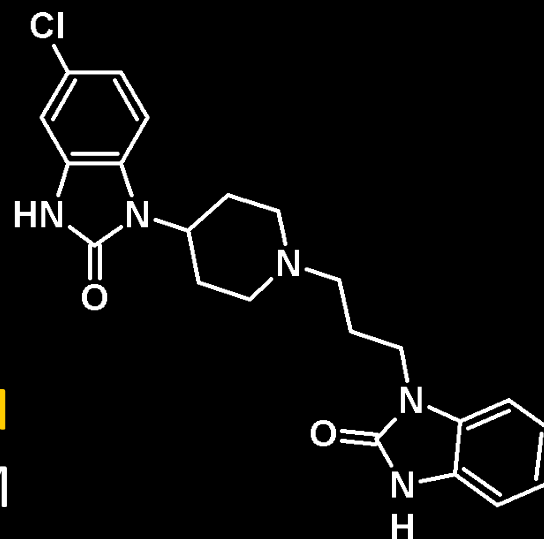
	Target	E-value	MaxTc
1	Adrenergic (β) blocker	3.92e-15	0.55
2	5 HT1A Antagonist	2.02e-9	0.55
3	Acetylcholinesterase Inhibitor	1.14e+1	0.41



Keiser, et al. *Nature*, 2009; Thomas, et al. *Psychopharmacol Bulletin*, 2010.

New targets & adverse drug reactions

- 2004 FDA warning letter against Motilium's import due to "adverse cardiac events"



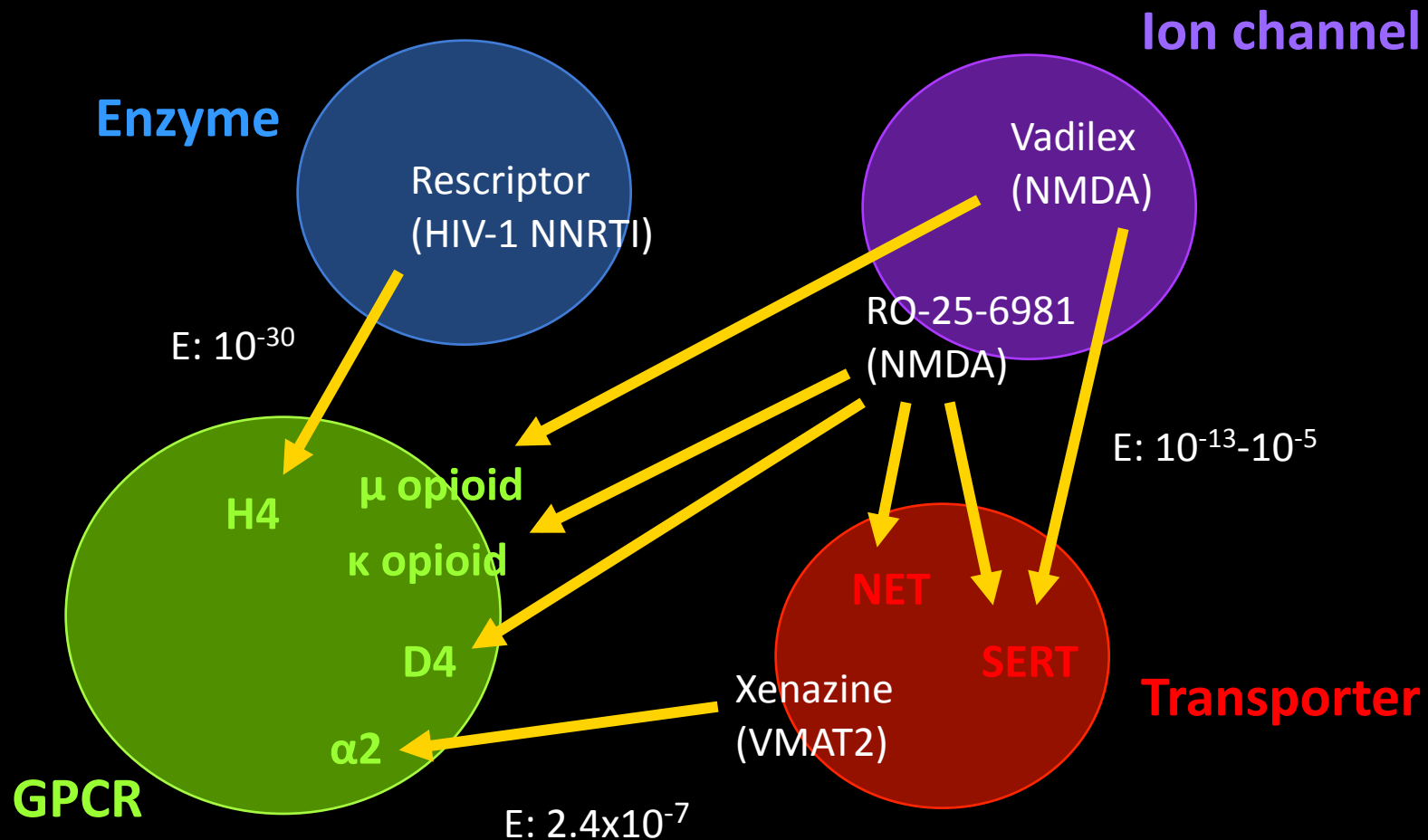
- Targets + SEA α_{1B} : 525 nM
– D1R, D2R hERG: 5,000 nM

- Uses

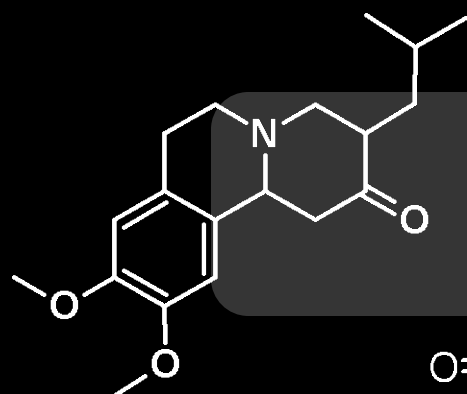
- Stimulates lactation
 - Off-label, and frequent
- Suppresses nausea and vomiting

Predicted Target		E-value
1	Adrenergic (α_1) Blocker	3.28e-13
2	Dopamine (D4) Antagonist	5.57e-12
3	μ Agonist	2.05e-08

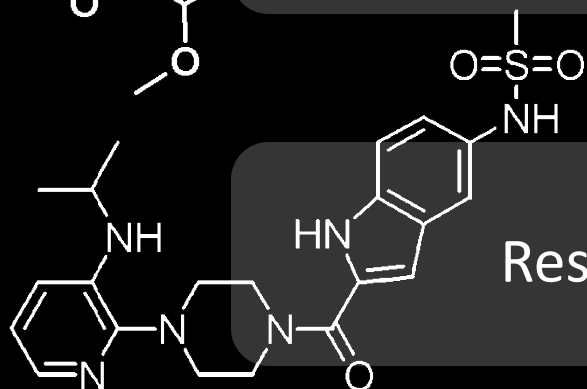
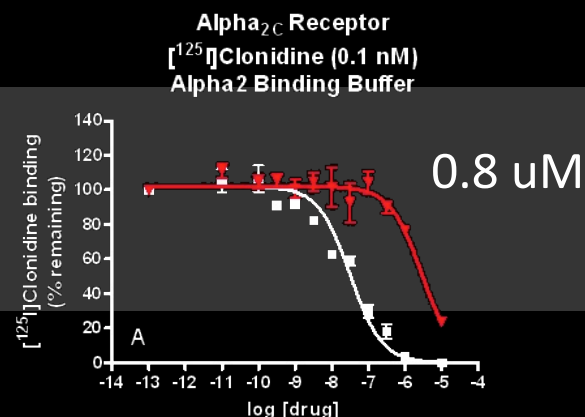
Systematic screen: Crossing target boundaries



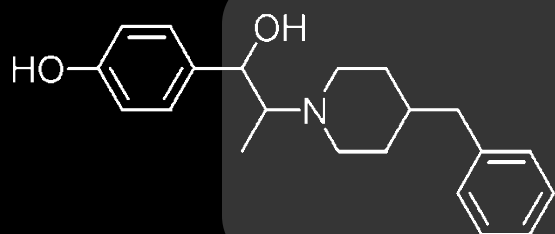
Systematic screen: Crossing target boundaries



Xenazine α_2 AR



Rescriptor H4 histamine (GPCR) 5.0 uM

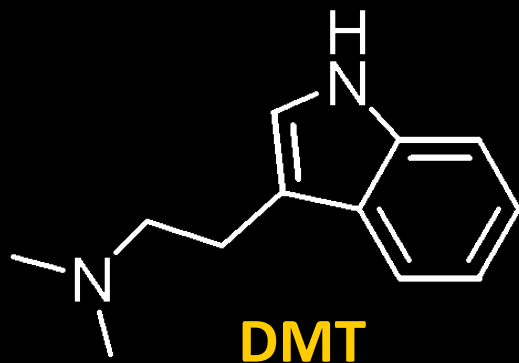


Vadilex μ opioid (GPCR) 0.4 uM
5HTT (transporter) 0.1 uM
DDR4 (GPCR) 0.1 uM
NET (transporter) 1.3 uM

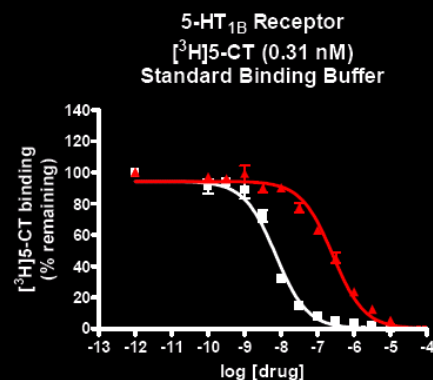
- I. Demystifying mode of action
- II. New drug uses
- III. Target ID in phenotypic screens
- IV. Broad-scale ADR predictions

SEA APPLIED

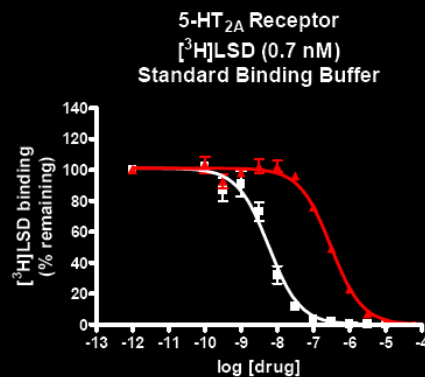
Systematic screen for new targets: MOA



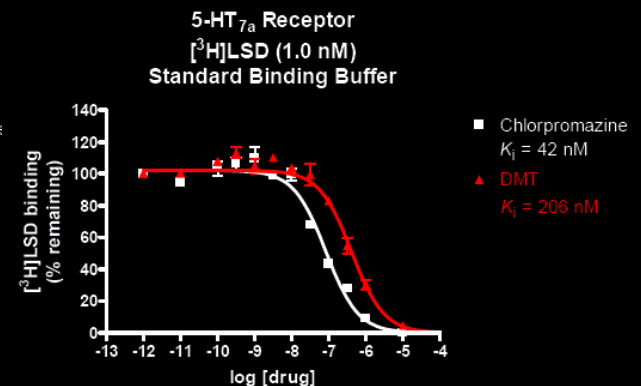
Target	E-value
5-HT _{1B}	3.1e-21
5-HT _{2A}	1.2e-13
5-HT _{5A}	1.1e-7
5-HT ₇	5.0e-6



5-HT_{1B}: 129 nM



5-HT_{2A}: 127 nM



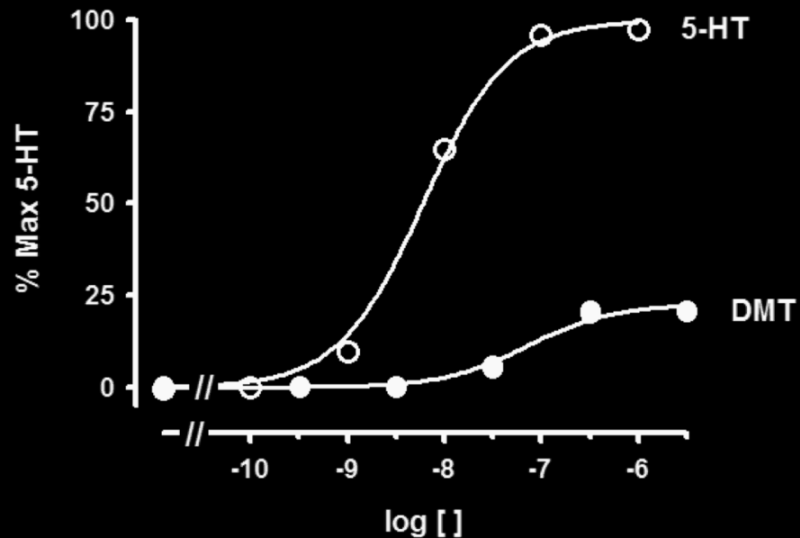
5-HT_{7A}: 206 nM

DMT-induced hallucination blocked in 5HT2A KO

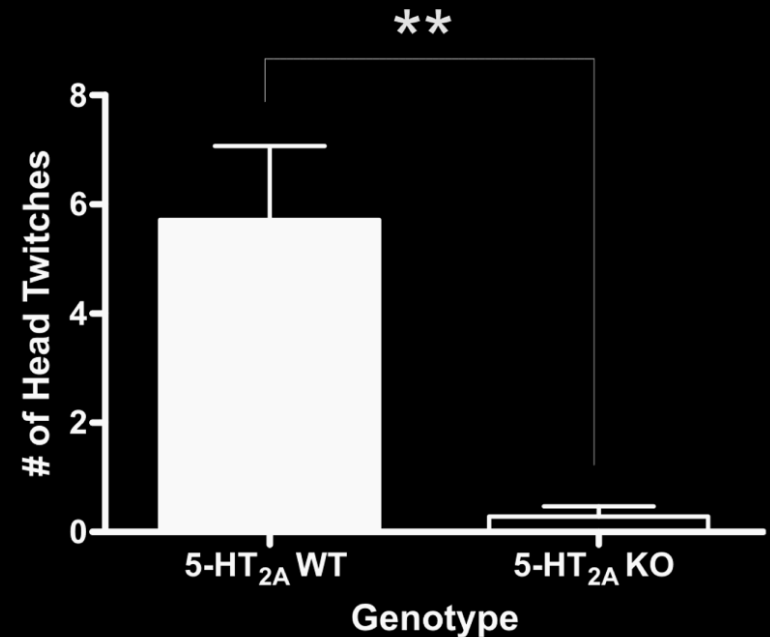
DMT **antagonizes** 5-HT_{2A}

Knock-out **abolishes** hallucination

5-HT_{2A}-mediated
Ca²⁺ response *in vitro*



DMT-induced head twitch in
5-HT_{2A} wild-type and knockout mice

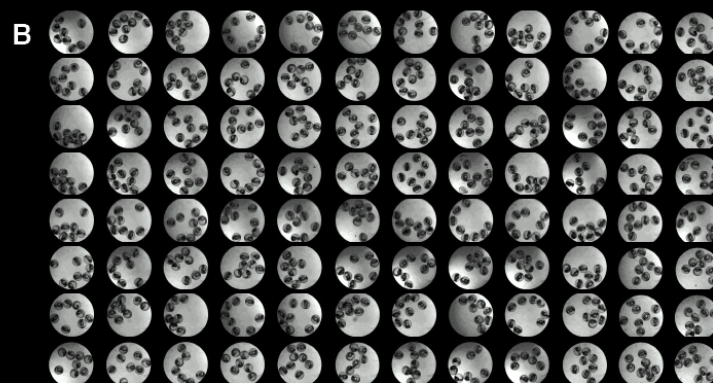
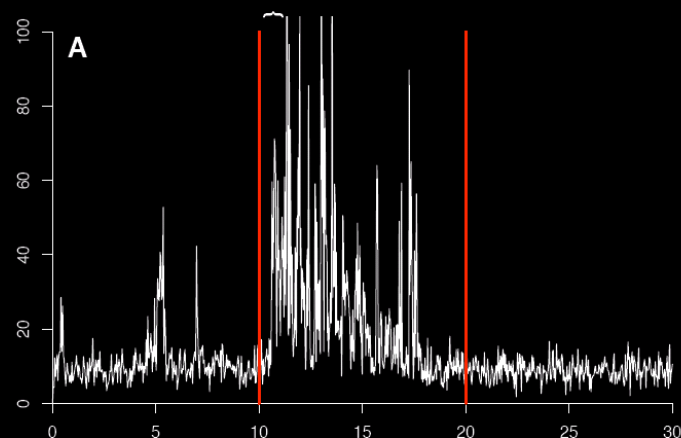


Supports 5-HT_{2A} as the **primary target** for DMT's
hallucinogenic effects.

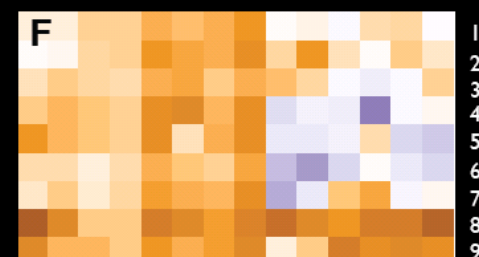
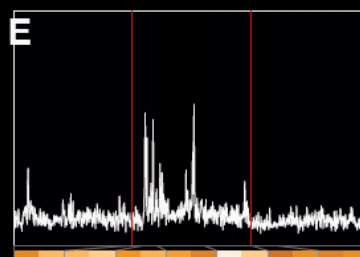
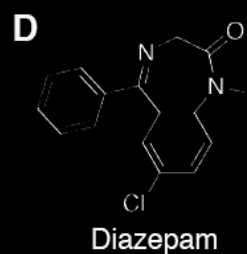
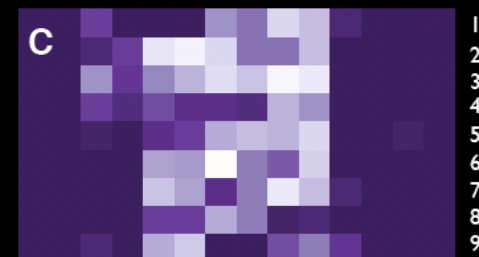
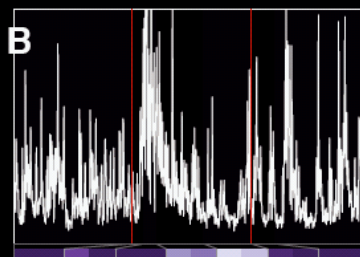
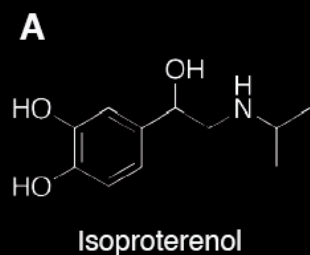


Target ID in phenotypic screens: Zebrafish response to light flash

1 Quantify response

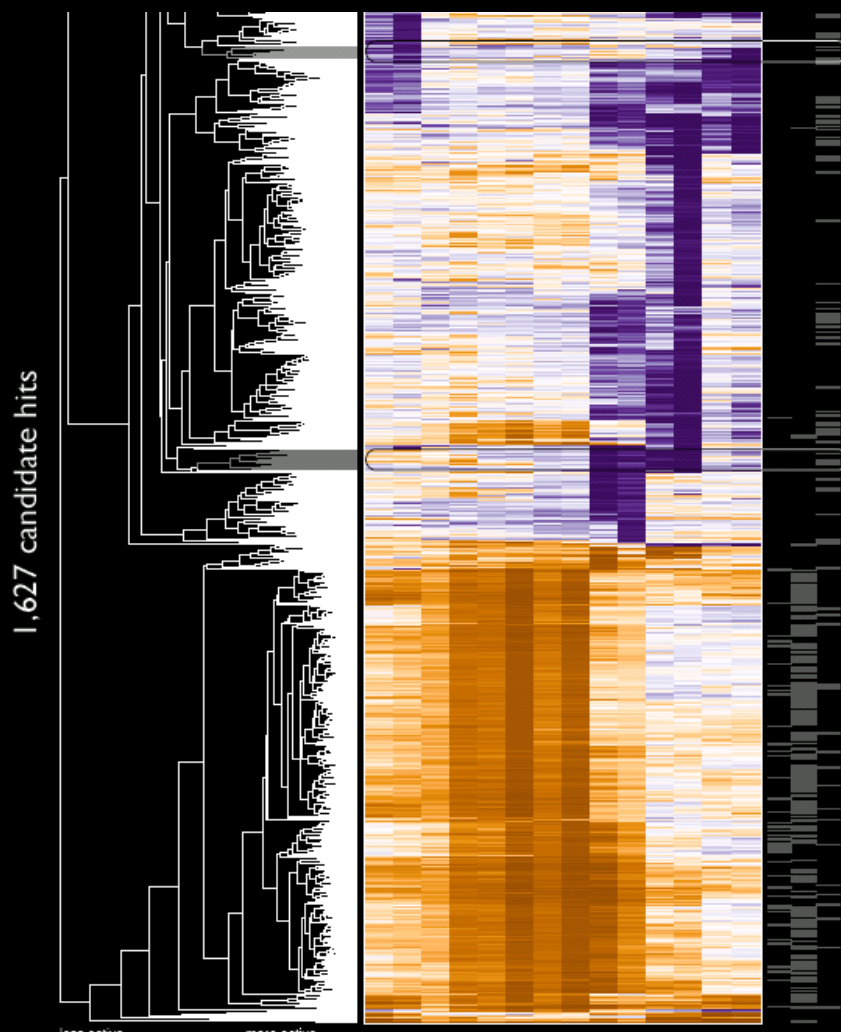


2 Express it as a fingerprint



Kokel et al, *Nature Chem. Biol.* 2010.

But what are the targets?

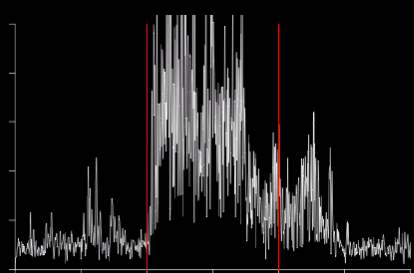


14,000 compounds
screened:

1627 hits

Clusters by phenotype

Fishing for targets with SEA

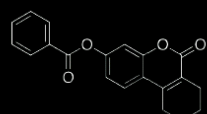
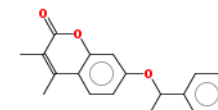
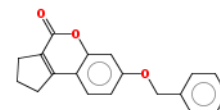
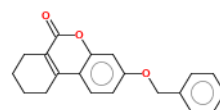


MAG-1

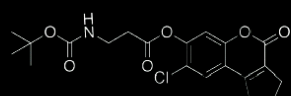
MAG-2

"amine oxidase [flavin-containing] B [-]" Structures (365 compounds)

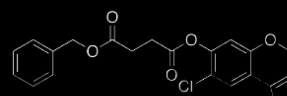
rank	1	rank	2	rank	3
code	00075613	code	00075612	code	00075611
ecfp4	0.6122	ecfp4	0.5800	ecfp4	0.4000
daylight	0.9468	daylight	0.6923	daylight	0.6598



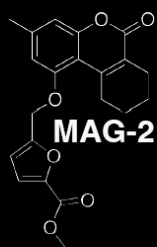
MAG-1



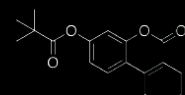
MAG-3



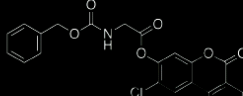
MAG-6



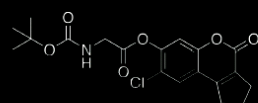
MAG-2



MAG-4

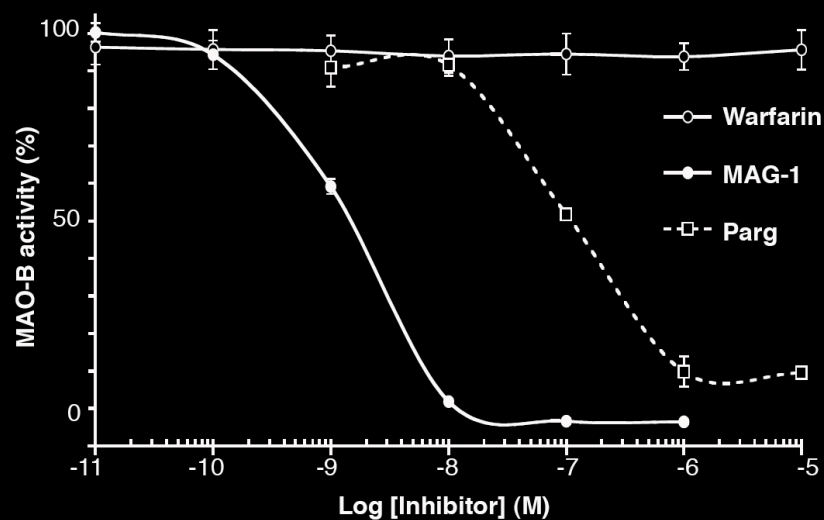


MAG-5



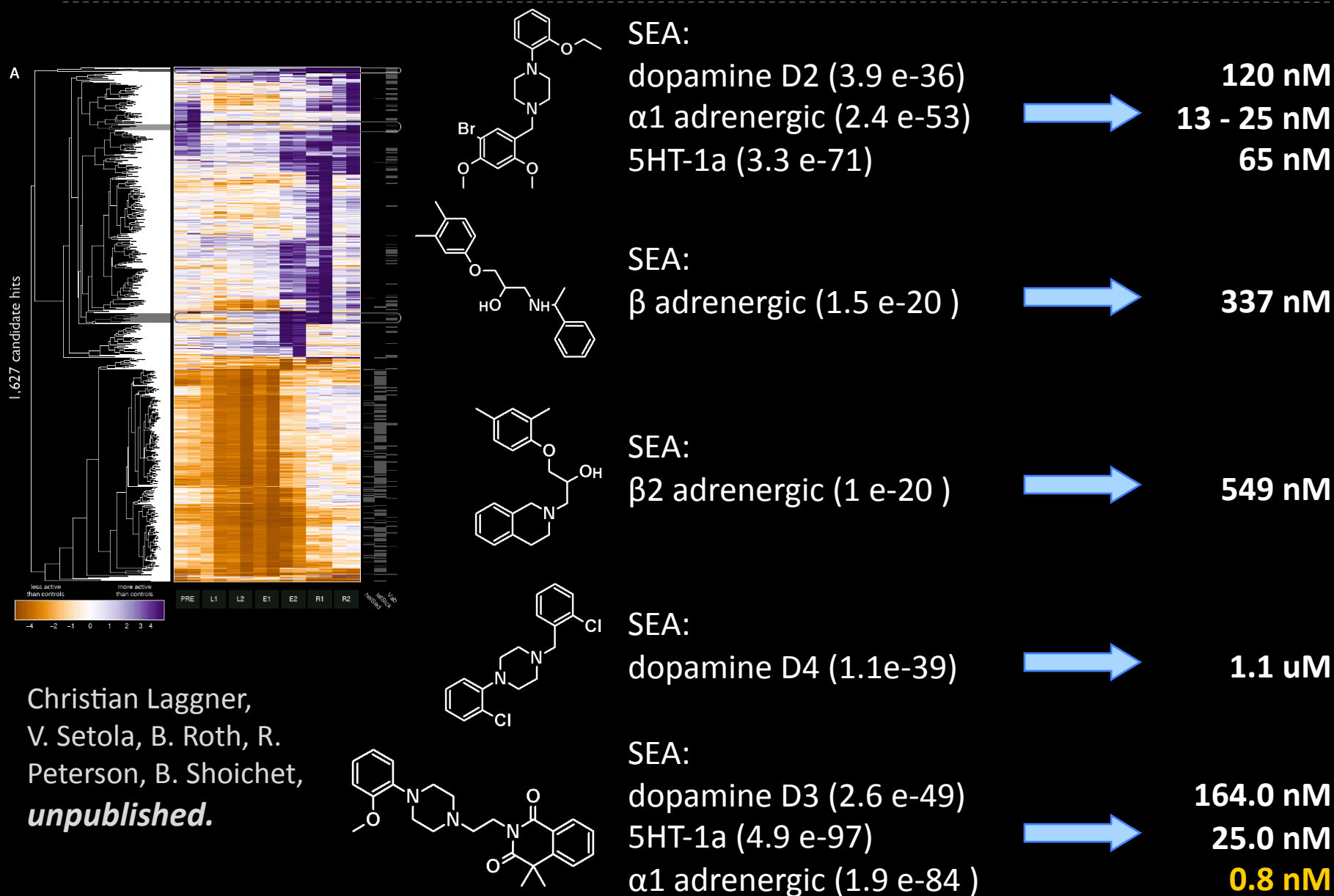
MAG-7

SEA Screen



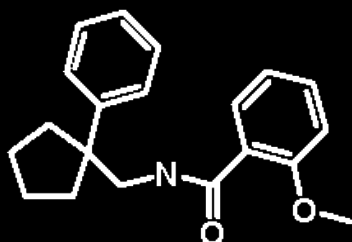
D. Kokel, et al. (C. Laggner, R. Peterson), *Nat. Chem. Biol.* 2010

Target ID by chemoinformatic similarity

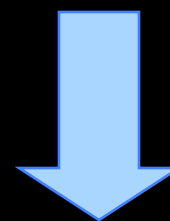
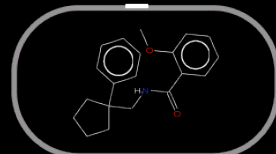


Christian Laggner,
V. Setola, B. Roth, R.
Peterson, B. Shoichet,
unpublished.

Targets of stimulatory hit cluster?

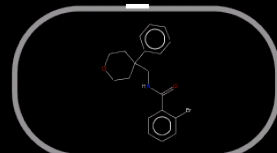


CdId_6519

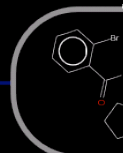


SEA

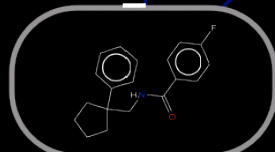
CdId_5901



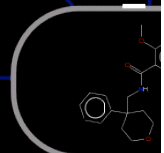
CdId_6519



CdId_2826



CdId_6519



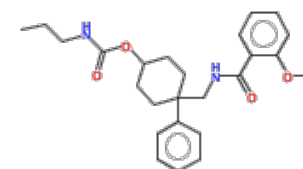
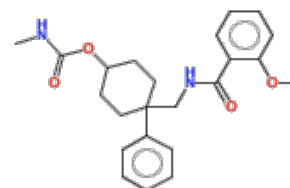
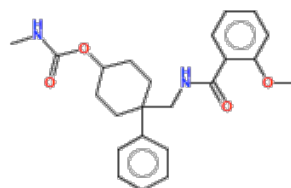
[Back to Index](#)

Page << 1 of 1 >>

Code	#Ligands	Reference Name	E-value	MaxTCMolecules
1 KCNA2_HUMAN	11	Potassium voltage-gated channel subfamily A member 2 (Voltage-gated potassium channel subunit Kv1.2) (HBK5) (NGK1) (HUKIV)	5.50e-77	0.53 view >>
2 KCNA1_HUMAN	14	Potassium voltage-gated channel subfamily A member 1 (Voltage-gated potassium channel subunit Kv1.1) (HUKI) (HBK1)	1.14e-64	0.53 view >>
3 KCNA3_HUMAN	98	Potassium voltage-gated channel subfamily A member 3 (Voltage-gated potassium channel subunit Kv1.3) (HPCN3) (HGK5) (HUKIII) (HLK3)	2.92e-13	0.53 view >>

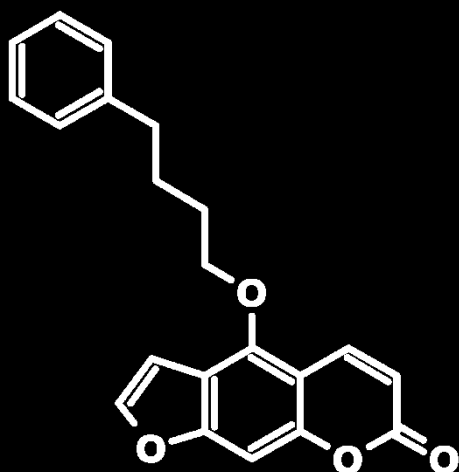
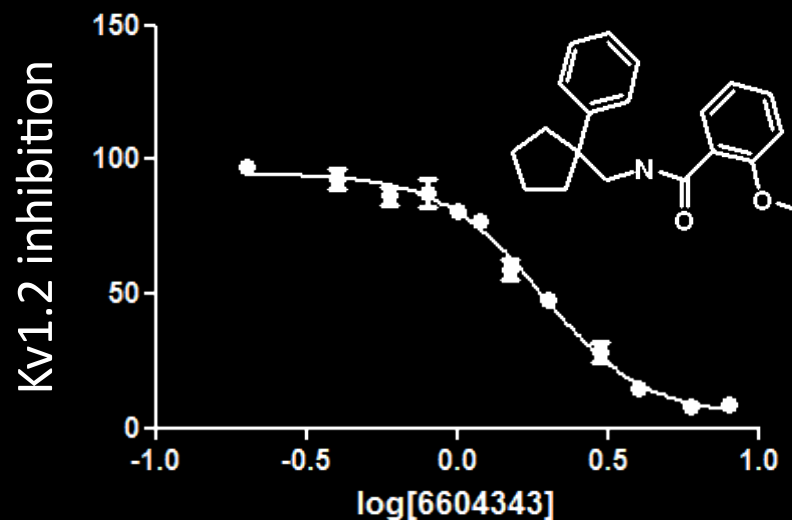
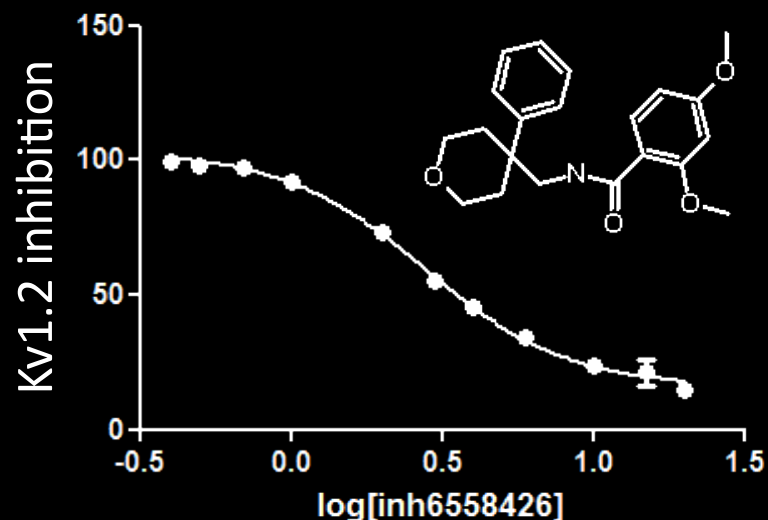
Potassium voltage-gated channel subfamily A member 2 (Voltage-gated potassium channel subunit Kv1.2) (HBK5) (NGK1) (HUKIV)" Structures (11 compounds)

rank	1	rank	2	rank	3
code	18007	code	17913	code	17967
ecfp4	0.5593	ecfp4	0.5593	ecfp4	0.5574
daylight	0.6131	daylight	0.6131	daylight	0.6093



C. Laggner, D. Kokel, D. Mi

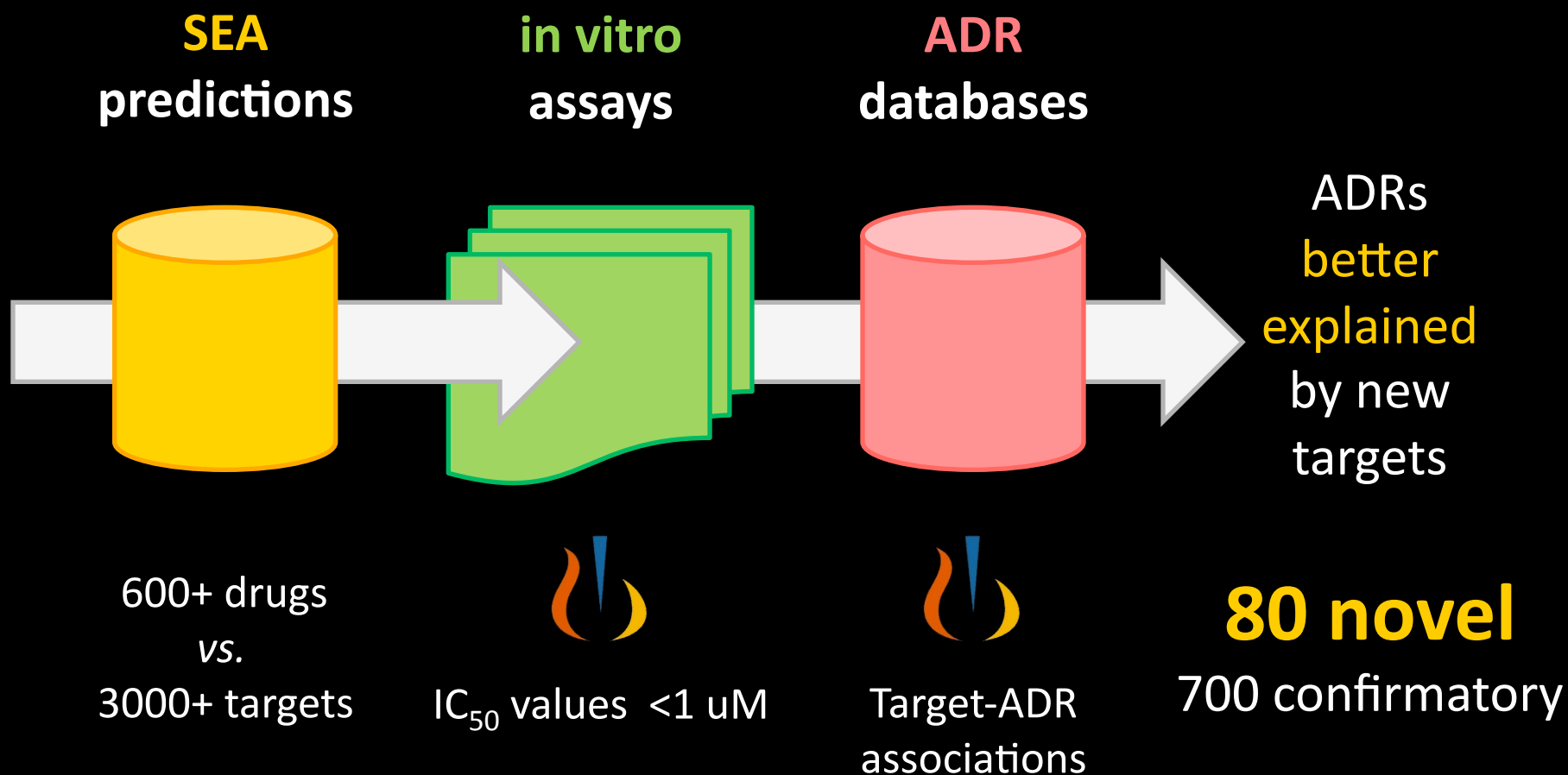
Hits block Kv1.2 channels at 2 to 4 μM



Psora-4, a 3 nM Kv1.n blocker, **phenocopies** the hits from the fish light assays.

D. Minor, Christian Laggner, D. Kokel, R. Peterson, B. Shoichet, *unpublished*.

IV Large-scale adverse drug reaction predictions



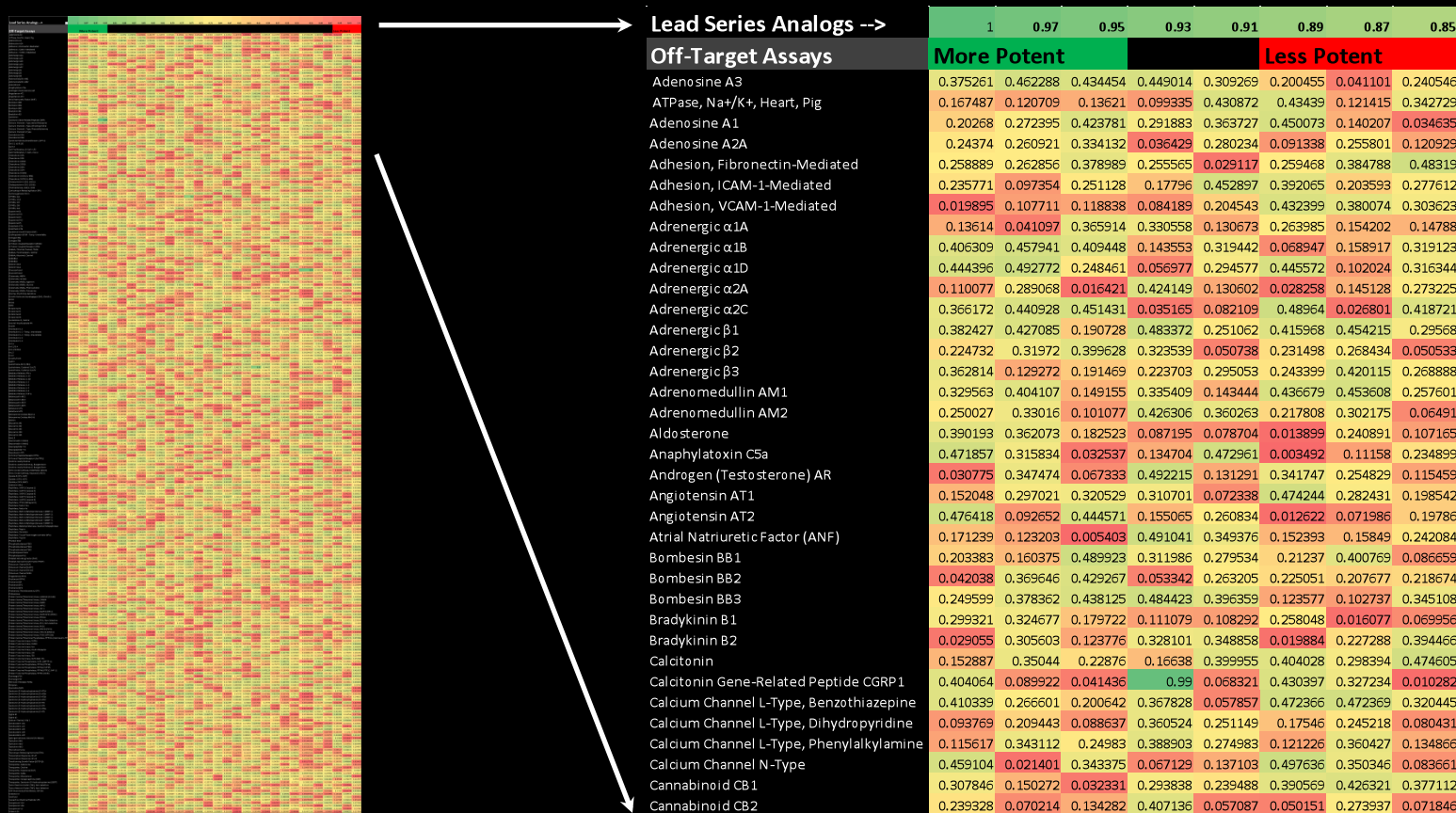
With Jeremy Jenkins & Eugen Lounkine



University of California
San Francisco

SEA broad-scale side-effect heat maps

Screen compound analog series for toxicity targets.



Note: The heat maps depicted here were built from *random data* only.

Standard H2L & LO profiling assays from Cerep, ChanTest, other CROs

Summary

- Pragmatic approaches to ligand discovery in chemical compound space
 - Purchasability
 - Bias towards biogenic compounds
 - Fragments

You can try SEA yourself

<http://sea.bkslab.org>



Similarity ensemble approach (SEA)

The SEA relates proteins based on the chemical similarity of their ligands to rapidly search large databases and build similarity maps among groups of compounds.

SEA is provided by the Shoichet Laboratory in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF).

Online search tool

The SEA search tool (beta) is now available.

Precalculated results

Database	Results
MDL Drug Data Report	Similarity map ¹
	Similarity map ² - in progress
	Significance scores

1 AT&T Graphviz
2 Cytoscape

michael.james.keiser@ucsf.edu