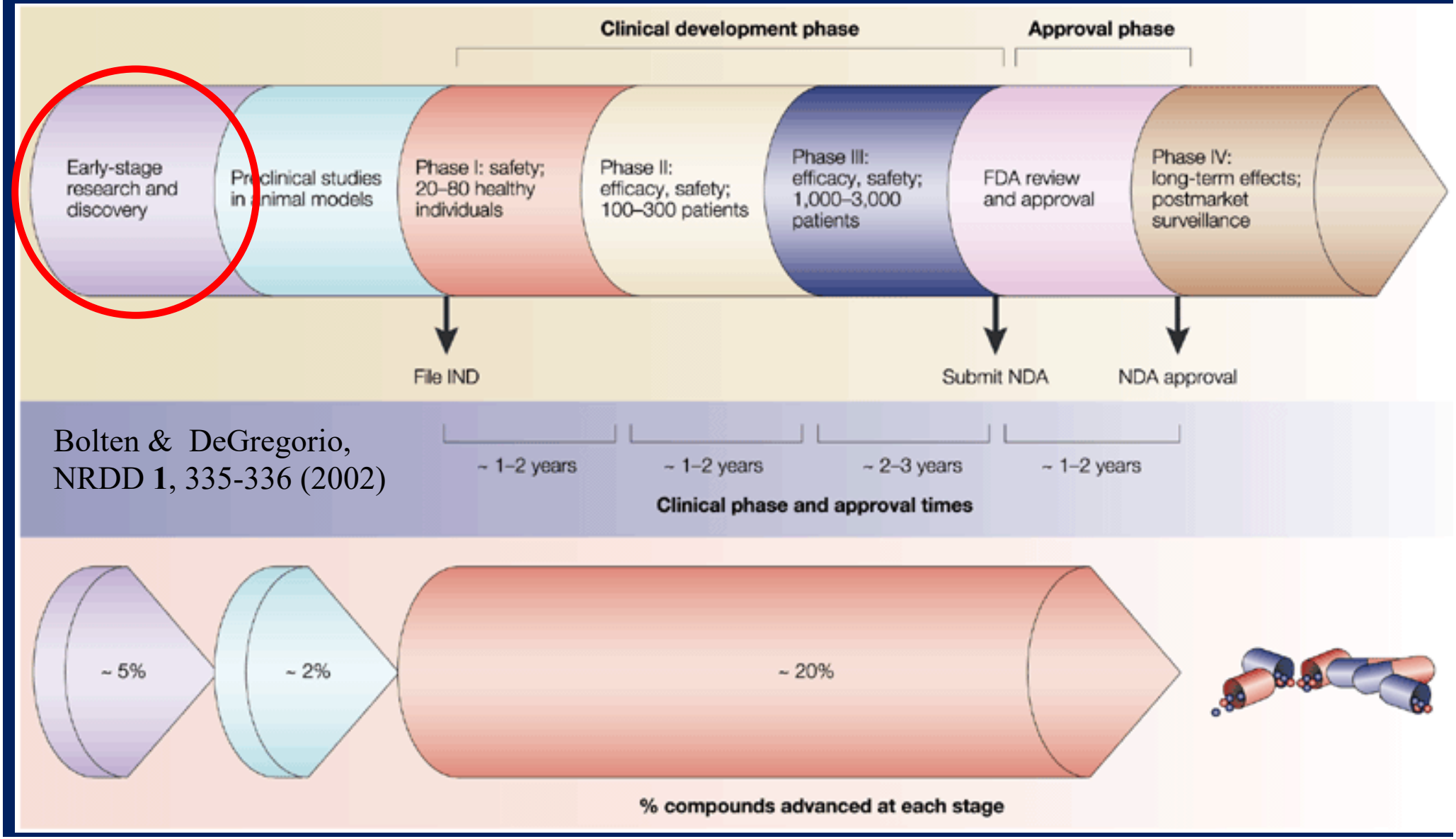
Using
Molecular Docking
For
Ligand Discovery

Promise: in the next hour you will learn about molecular docking

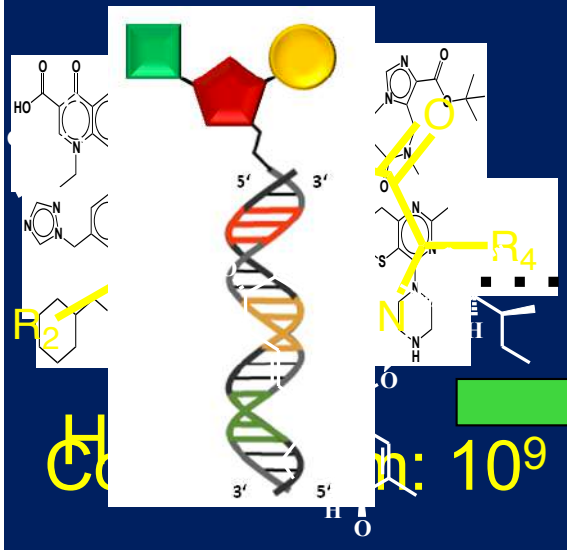
Context

The Drug Discovery Pipeline



Bolten & DeGregorio, NRDD **1**, 335-336 (2002)

Screening in early drug discovery



Co[illegible]n: 10^9

DELs: $10^{7.5}$

natural products

restricted diversity

10^{63} molecules with 30 heavy atoms containing N,C,S,O,H
(Bohacek, 1996)

Techniques in pre-clinical drug discovery

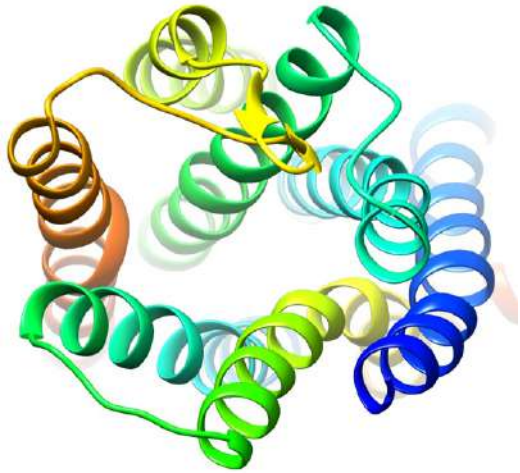
- **Ligand-based pharmacology/med. chem.**
- **Natural Product-based discovery**
- **High Throughput Screening**
- **Structure-Based**
  - Structure-based design
  - Virtual screens
- **Protein therapeutics**
- **Cell-based therapeutics**

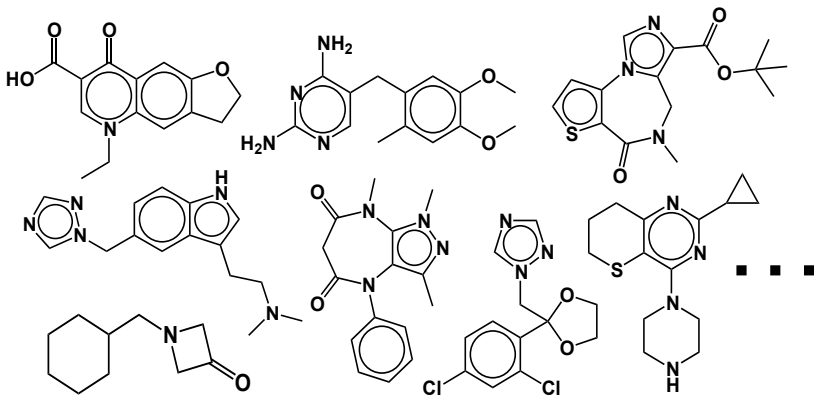
Molecular Docking: Lecture Goals

- What docking is
- How docking works
- Why docking is hard
- Success stories
- Opportunities / frontiers
- What you can expect from a docking campaign
- More stories
- Suggestions on how to get started
- Failures/limitations/caveats
- Summary
- Suggestions on how to get started

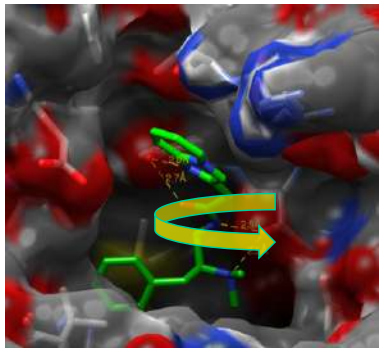
What docking is

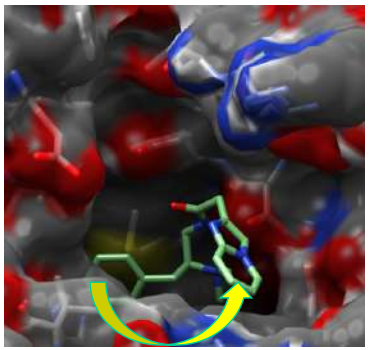
Docking large libraries for new chemotypes

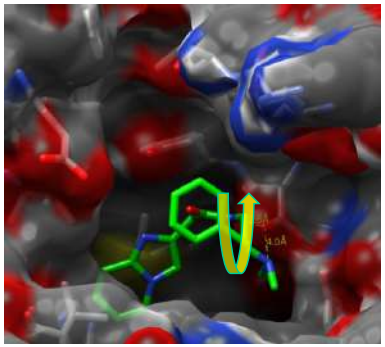


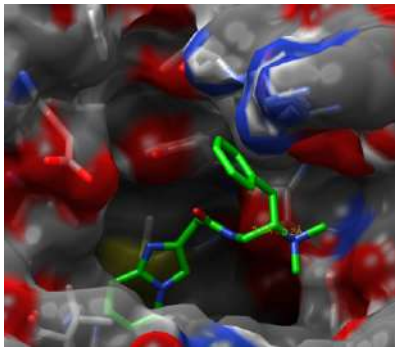


library of ***available*** molecules









each in ~10^6 orientations & conformations

Score for fit

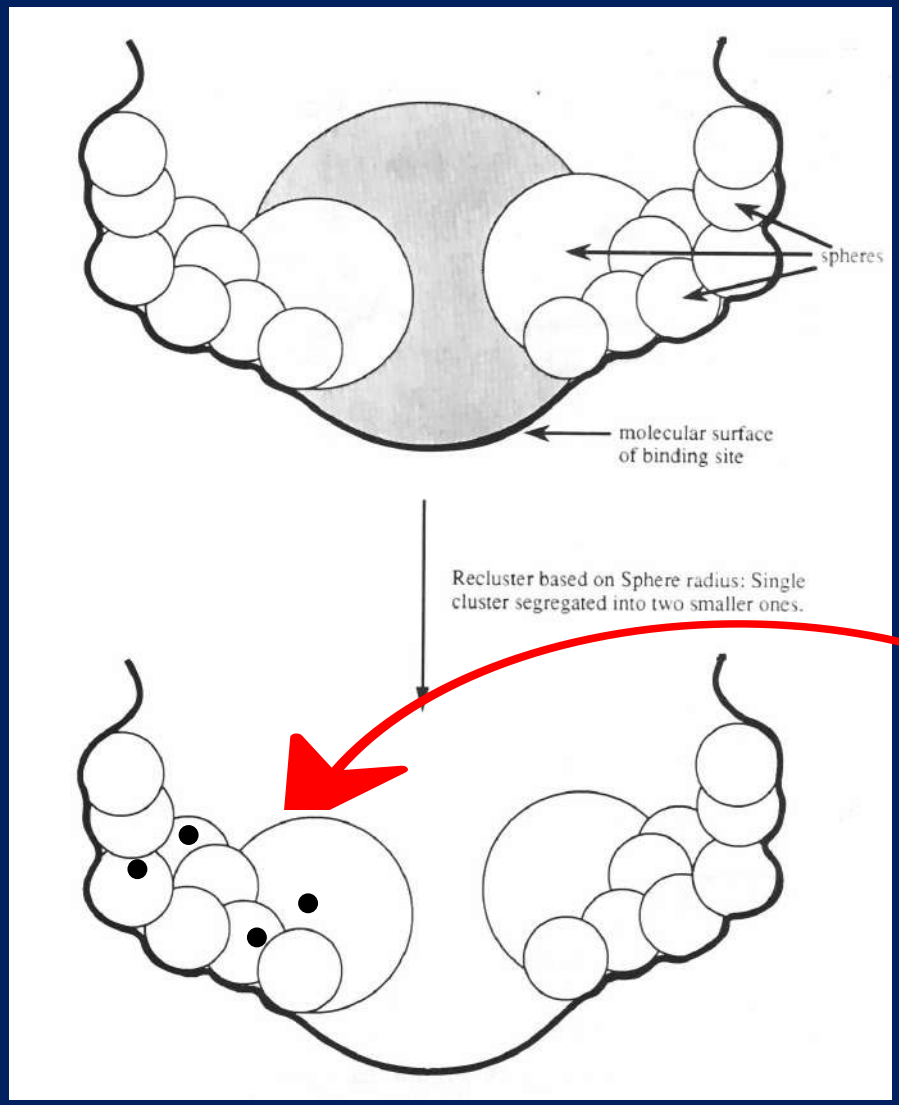
Test high-scoring molecules

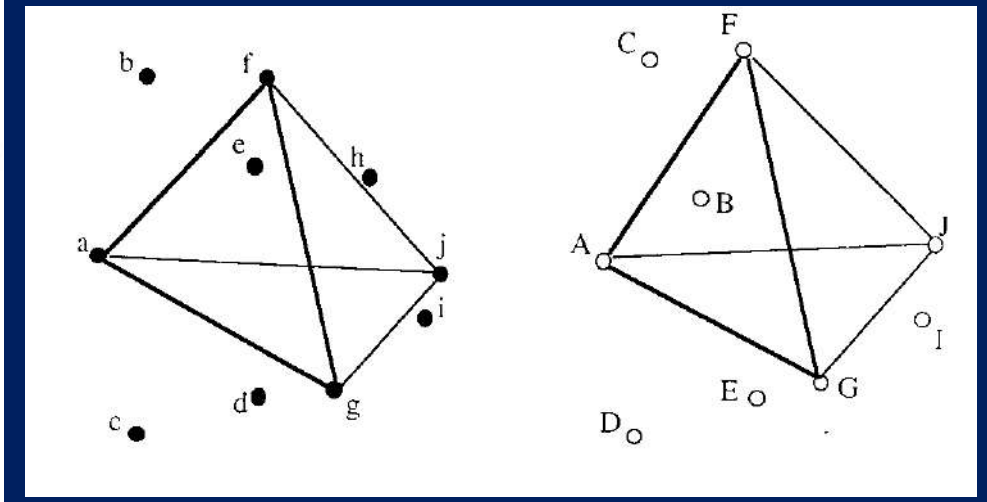
How docking works

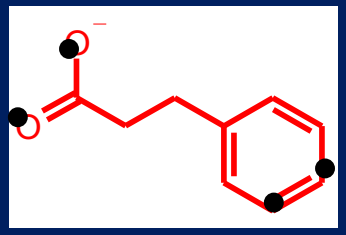
Calculating orientations in DOCK

hot spots on protein surface

Match ligand atoms onto
Hot spots using internal distances







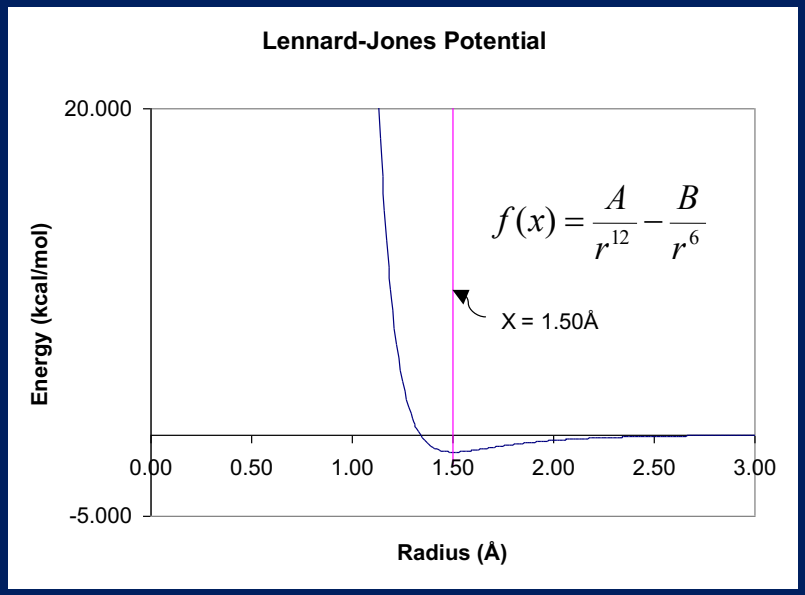
$^{n}C_{m}$: thousands of orientations per molecule

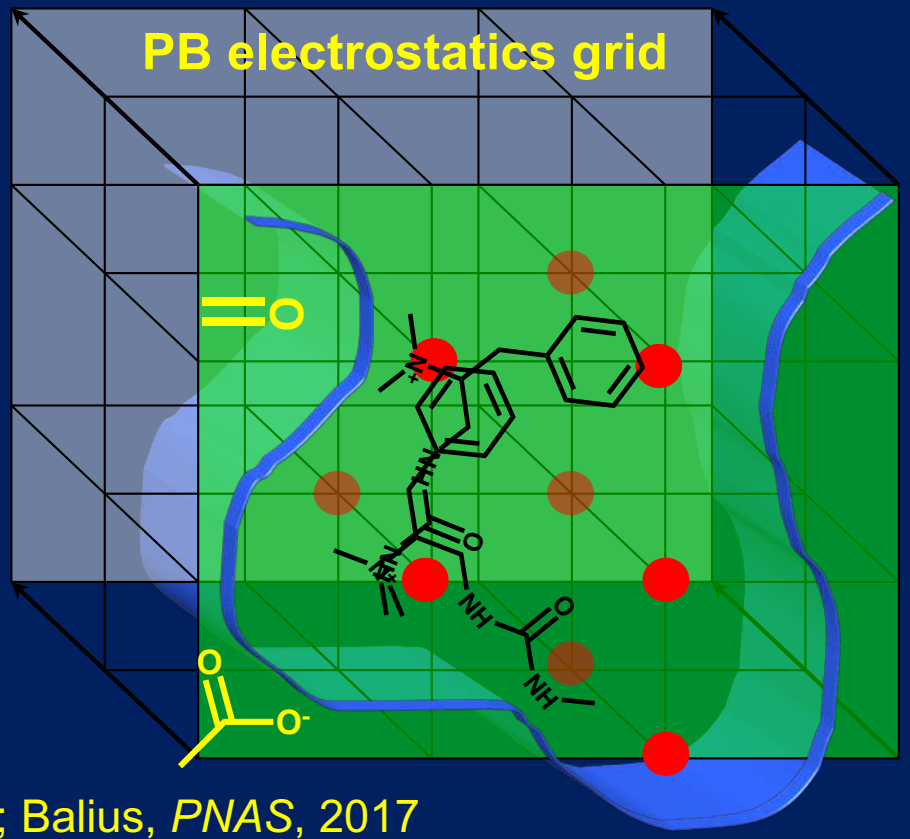
It's a miracle it ever works: subtracting large, inaccurate energies to get a small one

$$\Delta G_{interact} - \Delta G_{solv,\,L} - \Delta G_{solv,\,R} = \Delta G_{bind}$$

$$\Sigma(q_i P_i + v_i P_v) = \Delta G_{inter}$$

$$((1/D_0 - 1/D_w)/2r)*\Sigma Q^2{}_i*Born*(a_i \delta V \Sigma 1/r^4{}_{ik}) + \Delta H_{np} + IST = \Delta G_{solv}$$

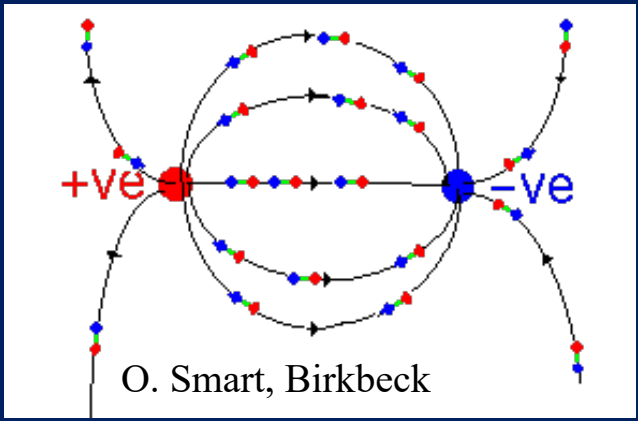




Mysinger, *JCIM* 2010; Fischer, *Nature Chem.* 2014; Balius, *PNAS*, 2017

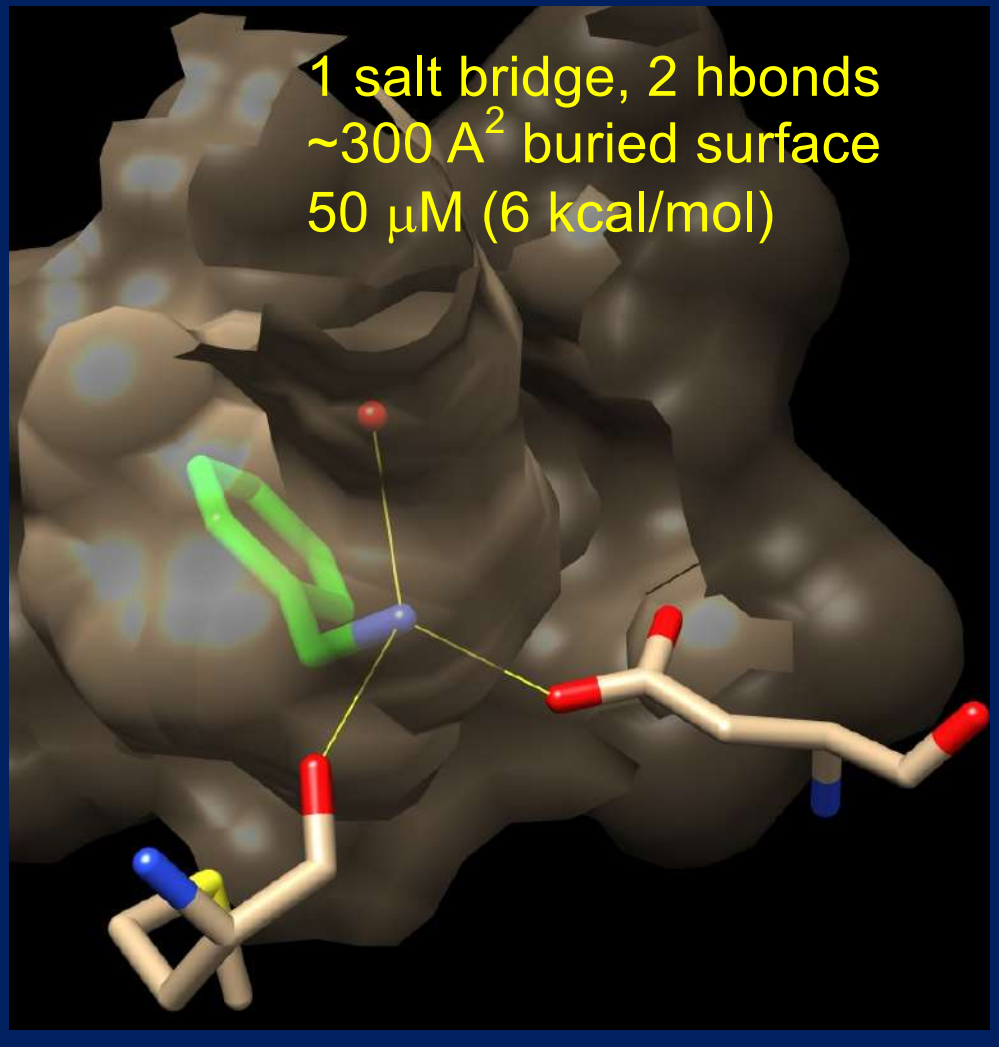
Ionic interactions ("salt bridges")

Ionic: $E = q_1q_2/4\pi\varepsilon r = 332 \times q_1q_2/\varepsilon r$ kcal/mol



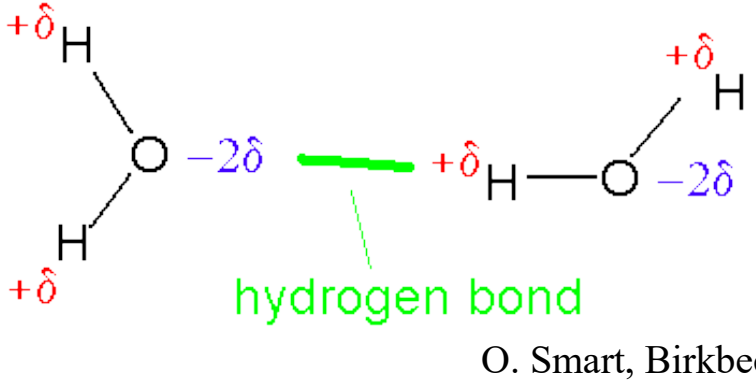
Dielectric Constants of some media

Water (20°C)	80.3
Water (0°C)	87.7
Methanol	33.6
Liquid H_2S (-85.5 °C)	9.3
Beeswax	2.9
Liquid Argon (-191°C)	1.5
Vacuum	1.0



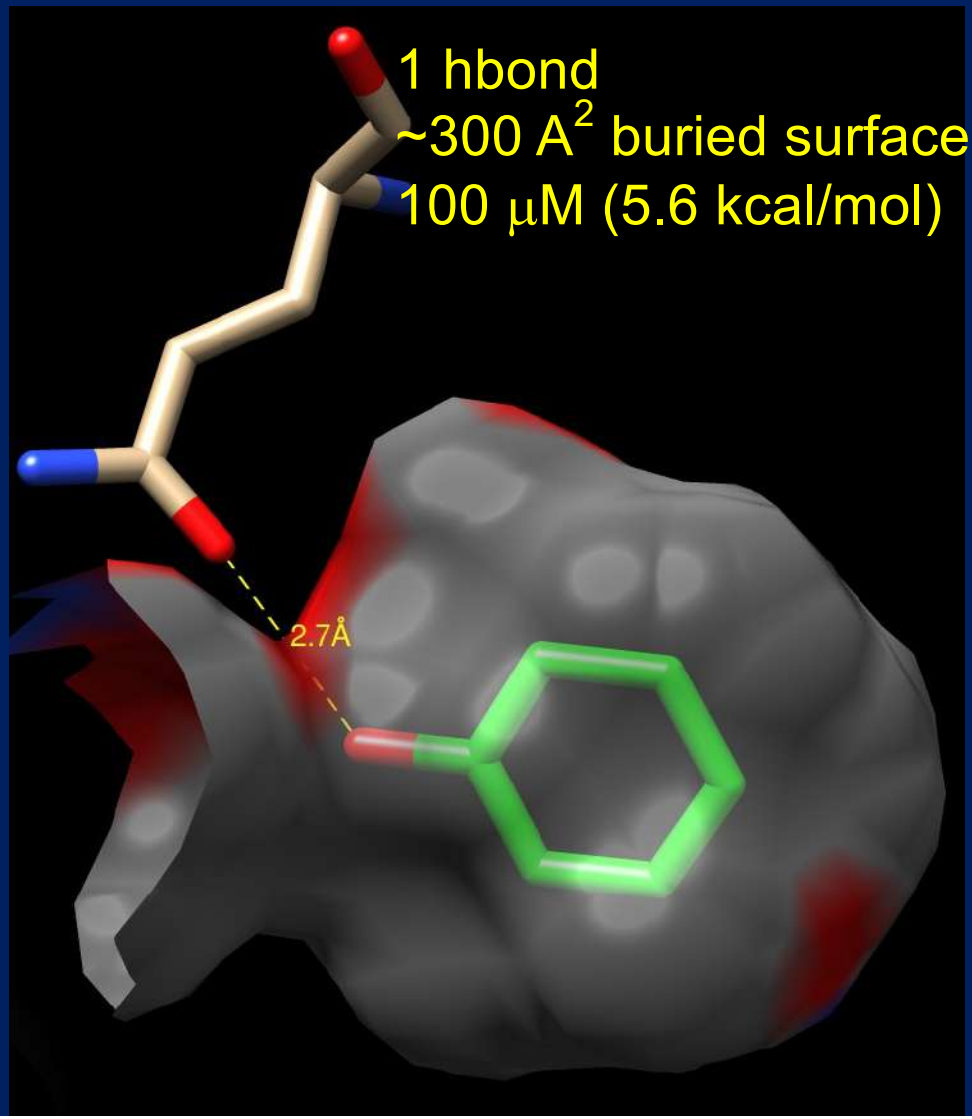
Hydrogen bonds

Dipole-Dipole ($1/r^3$, often modeled through partial charges)

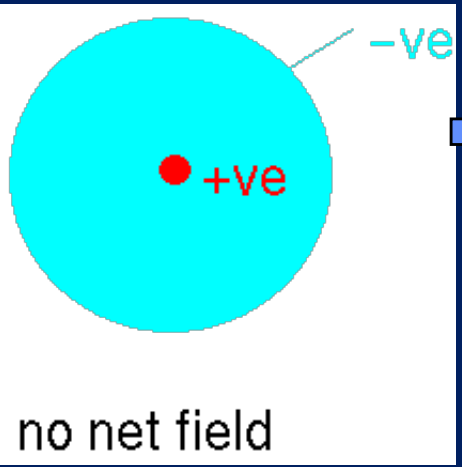


O. Smart, Birkbeck

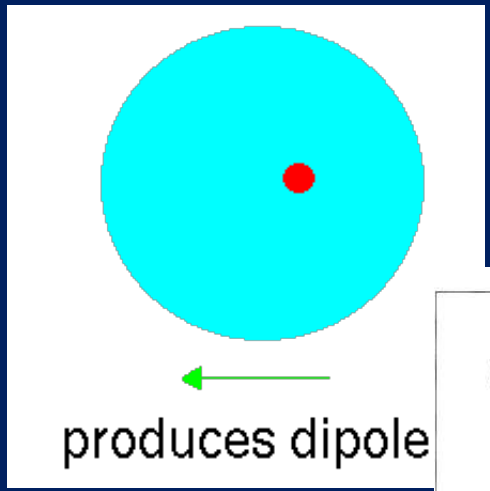
directionality important;
hydrogen bond in vacuum
~ 7 kcal/mol



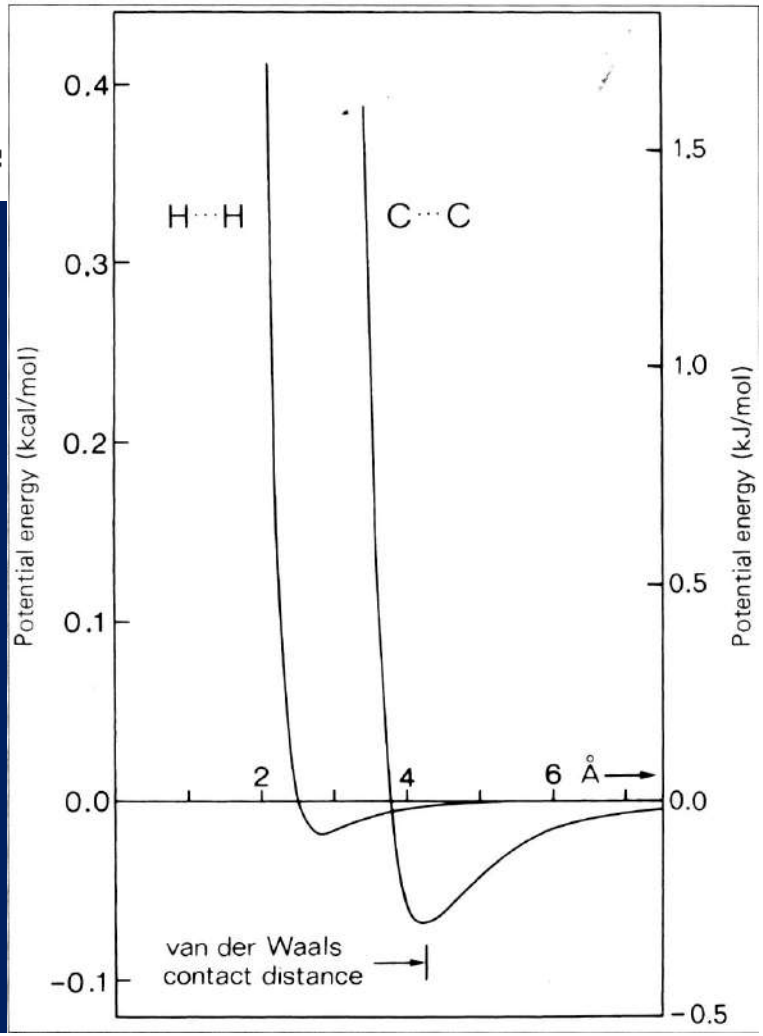
non-bonded interactions: dispersion-attraction (van der Waals)



Electronic fluctuation

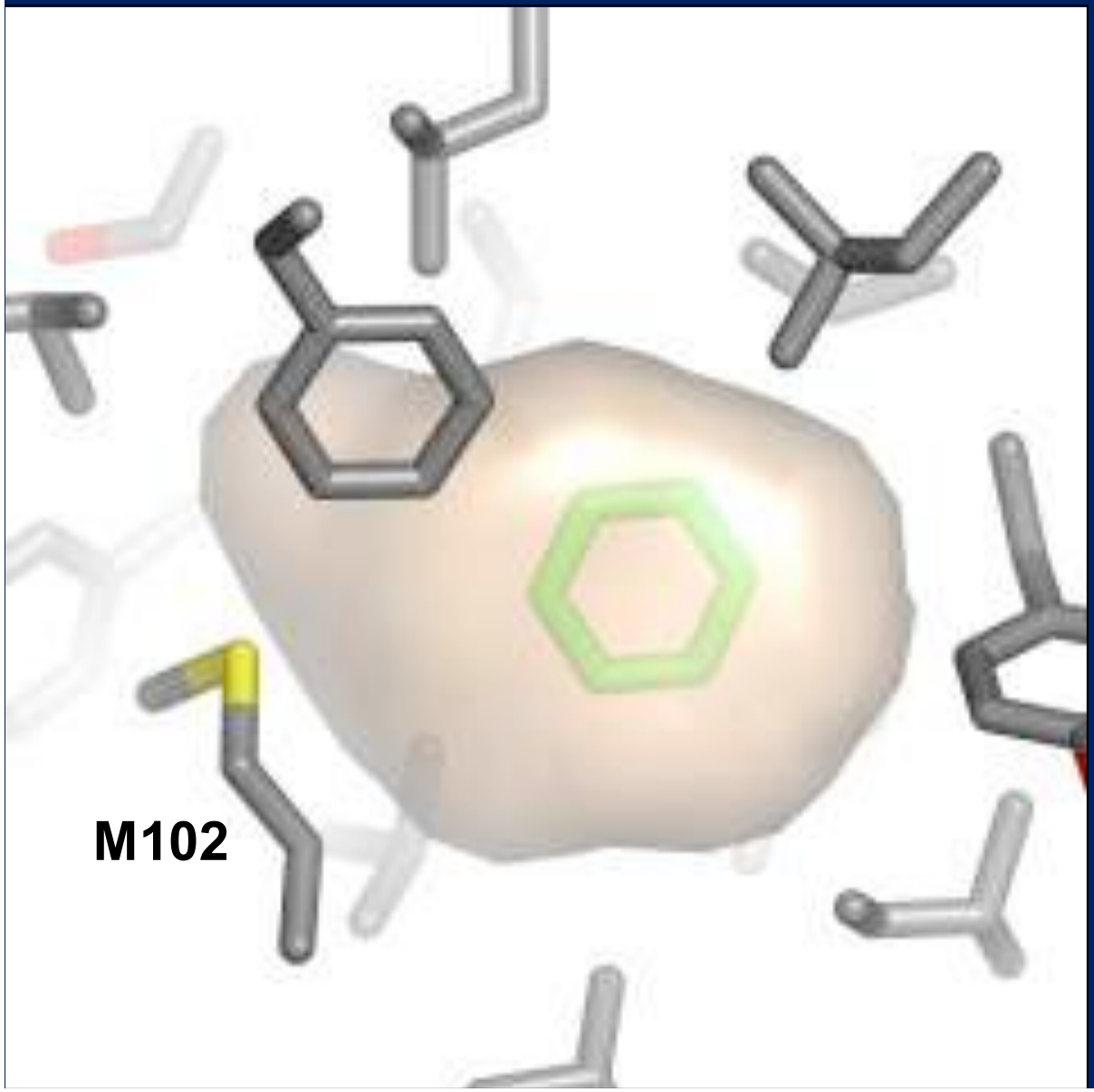


$$E = A_{ij}/r^{12} - B_{ij}/r^{6}$$



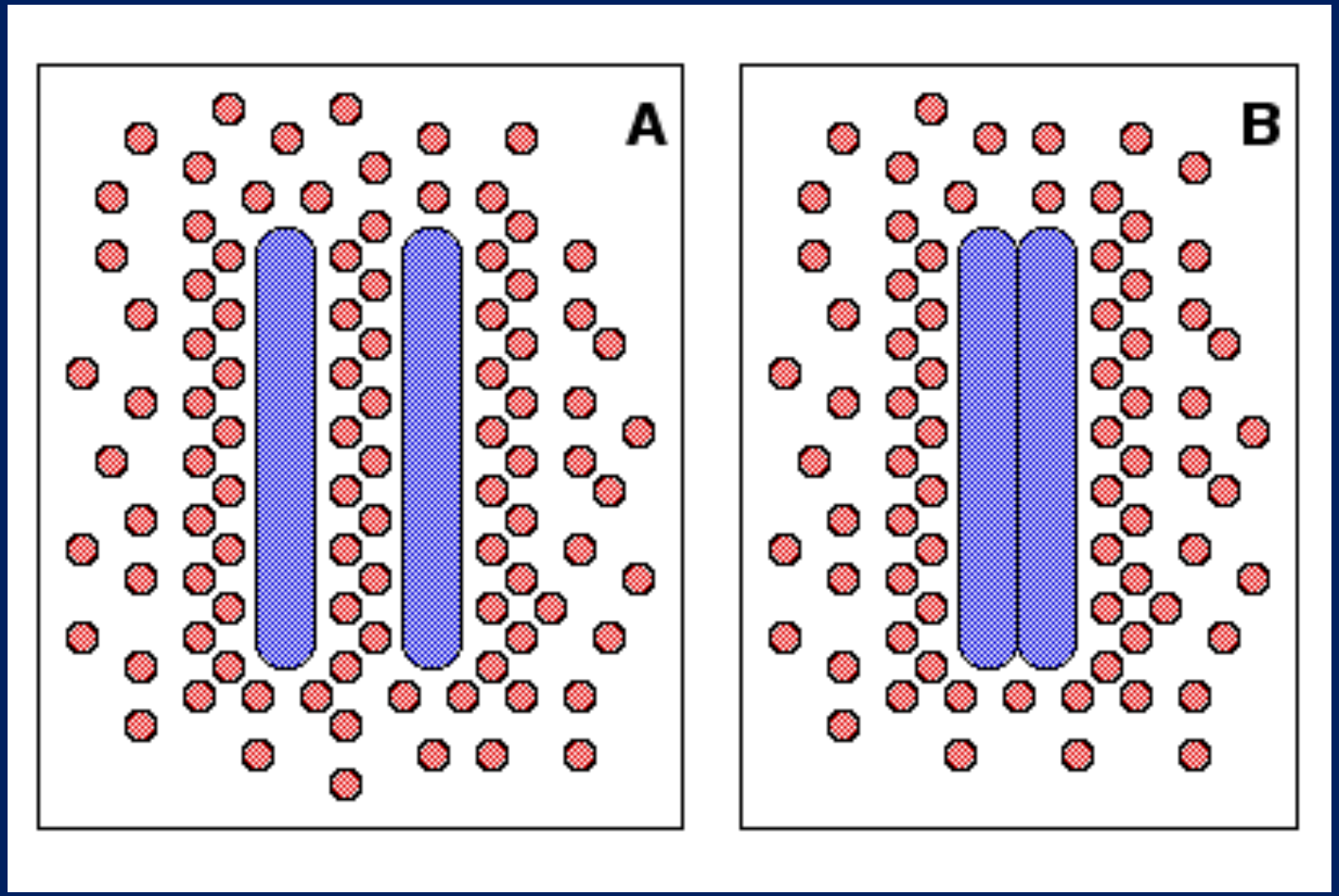
Binding dominated by Non-polar Interactions

T4L Core Mutant with Benzene (Kd 400 μM)



Interaction inventory:
~300 A^2 buried surface
4.8 kcal/mol

Non-bonded ‘interactions’: Hydrophobicity



Steve Mack

Other interactions: stacking, quadrupolar, chelation.

What do we approximate ✓ what do we miss? x

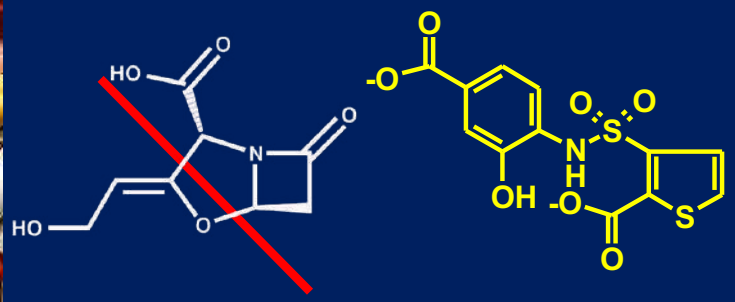
Types of Interaction	Strength	Effect of Distance	
Ionic	V. Strong	$1/r$, long range	✓
Ion-Dipole	Strong	$1/r^2$, short range	✓
Dipole-Dipole	Moderate	$1/r^3$, short range	✓
Hydrogen Bond	Moderate	$1/r^3$ (?) short range	x
Ion-Induced Dipole	Weak	$1/r^4$, v. short range	x
Dispersion	V. Weak	$1/r^6$, ext. short range	✓
Repulsion	Ext. Strong	$1/r^{12}$, ext. short range	x
"Hydrophobic"	Weak	??	

Conformationals states
Desolvation: Ligand, Protein
Entropy: conformational, configurational, vibrational

$$\sum(q_i P_i + v_i P_v) - \Delta H_{solv\_charg} + \Delta H_{solv\_np} = \text{dock score}$$

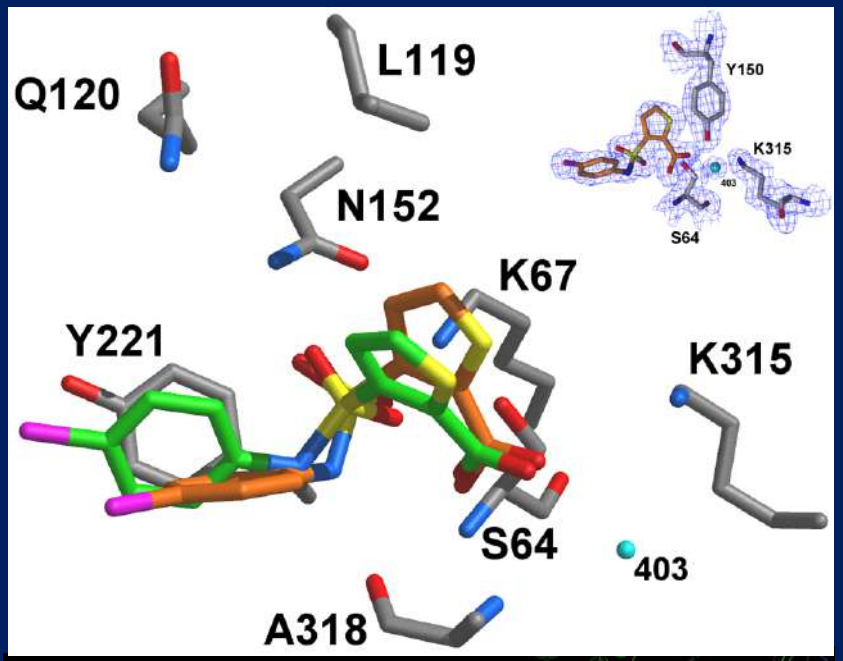
Why docking is hard

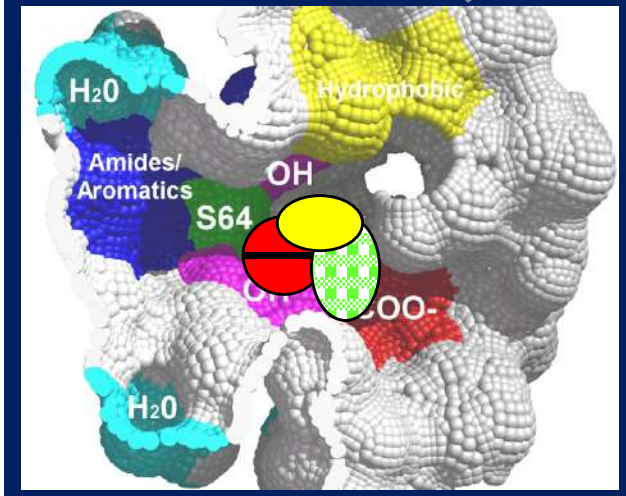
Challenges illustrated in two β-lactamase campaigns: 2-6% hit rates, ~30 uM K_is

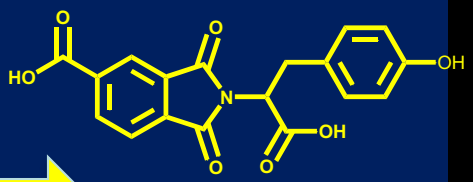


26 uM

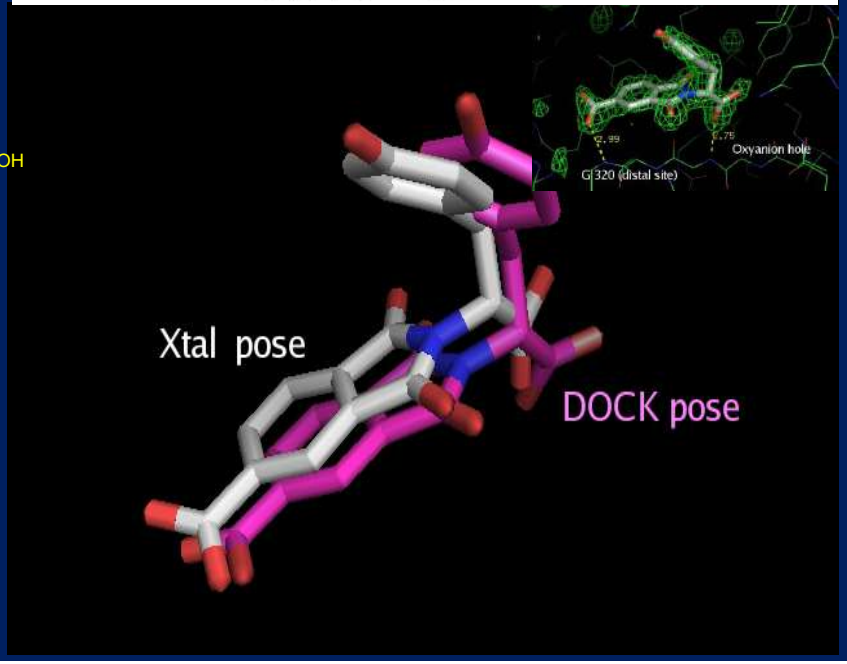
~10^5 cmpds







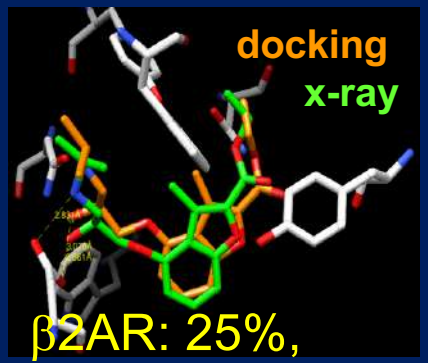
37 uM



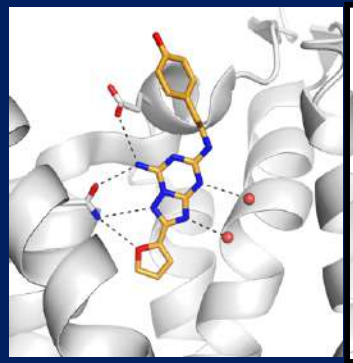
Powers, ***Structure***, 2002; Babaoglu, ***JMC***, 2008

Success stories

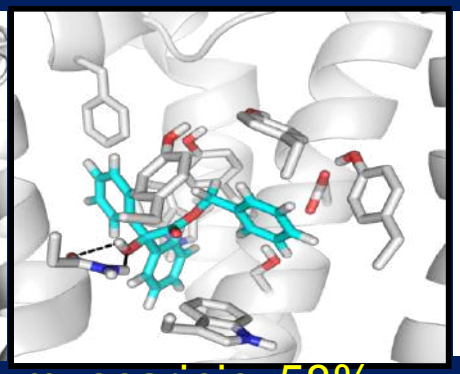
docking vs GPCRs: 17 to 58% hit rates, nM activities



β2AR: 25%,
0.01 to 3 uM
Kolb, *PNAS* 2009



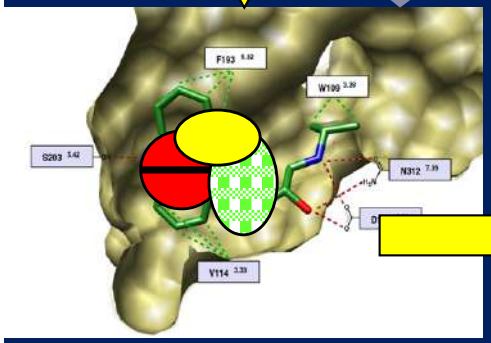
A2a: 35%,
0.2 to 3 uM
Carlsson, *JMC* 2010



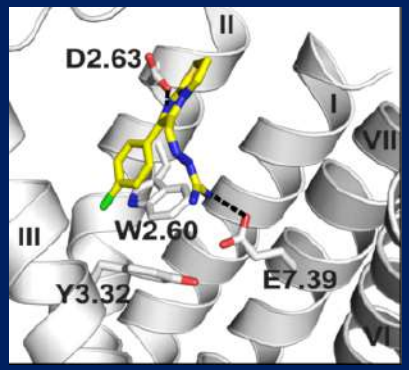
muscarinic: 58%
0.4 to 40uM
Kruze, *Mol Pharm* 2013

D3 model: 23%
0.2 to 3uM; Carlsson, *Nature CB* 2011

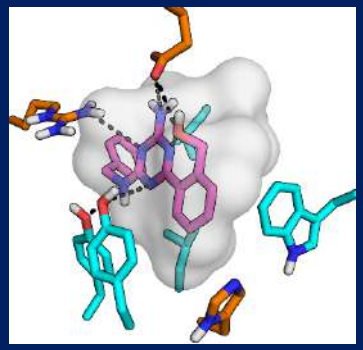
3x10^6 cmpds



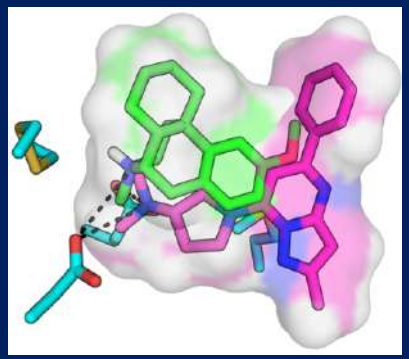
Test



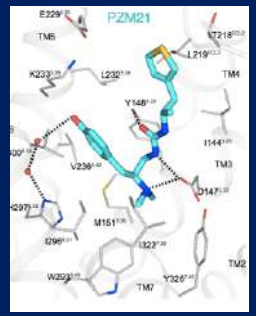
CXCR4:17%,
0.3 to 30uM
Mysinger, *PNAS* 2012



GPR68 probe
Huang, *Nature* 2015



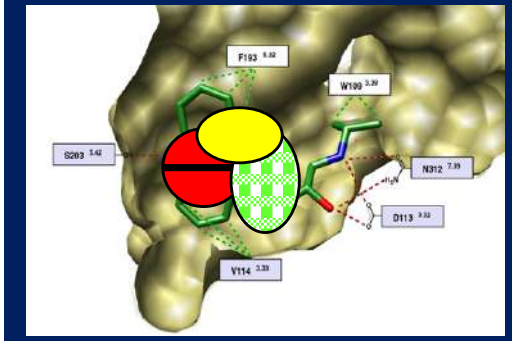
MRGPRX2 probe
Lansu, *Nature CB* 2017



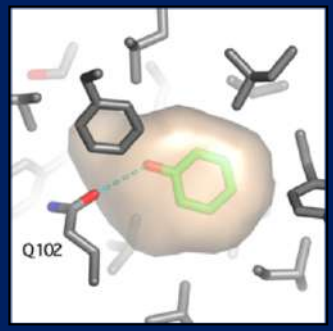
Mu Opioid: 30%, 4 nM
Manglik, *Nature* 2016

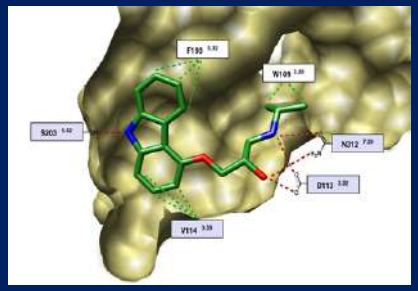
And yet (sometimes) it can prioritize likely ligands

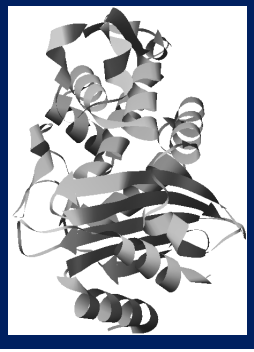
3x10^6 cmpds



Test







buried cavities:

docking	fragmt docking	HTS
	75%	-----

GPCRs:

docking	fragmt docking	HTS
17-60%	40-75%	~0.1%

AmpC:

docking	fragmt docking	HTS
2-11%	48%	0%

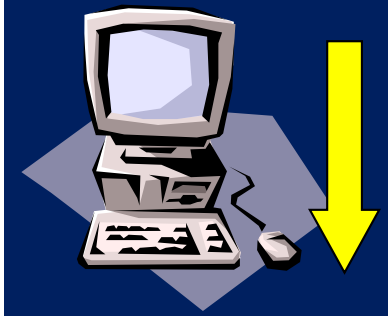
CTX-M:

docking	fragmt docking	HTS
0%	55%	---

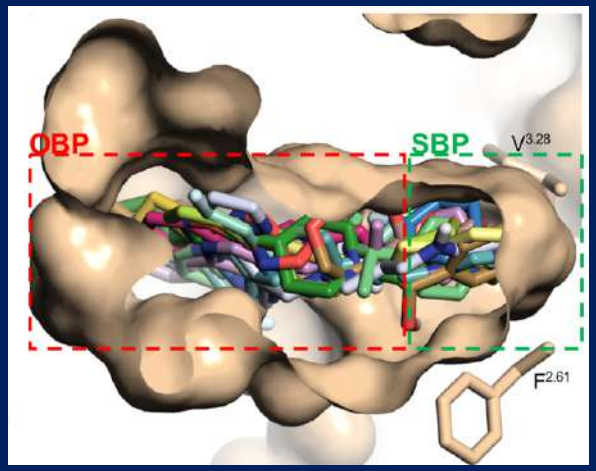
Opportunities / frontiers

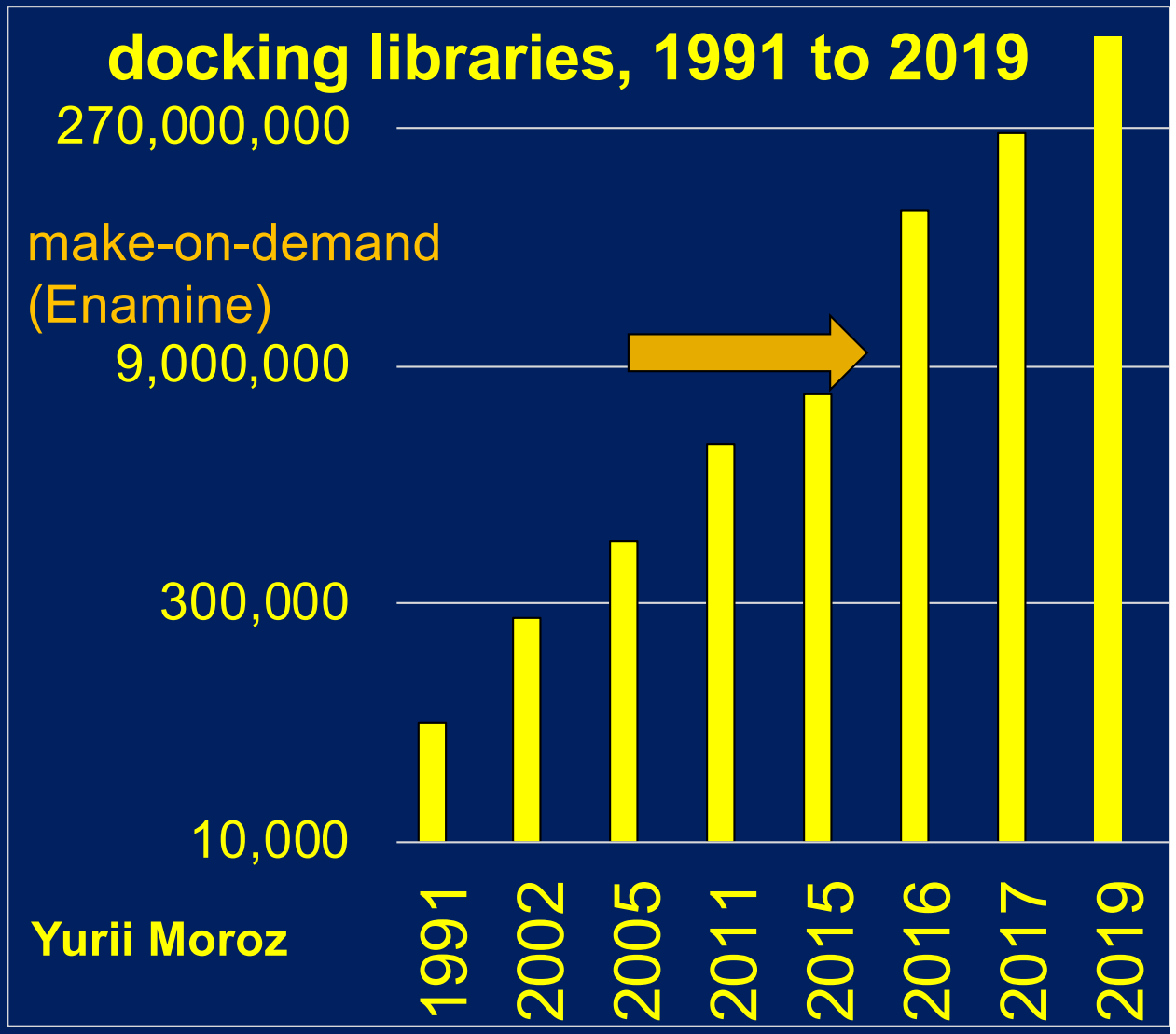
docking libraries have exploded

3→500 million compounds

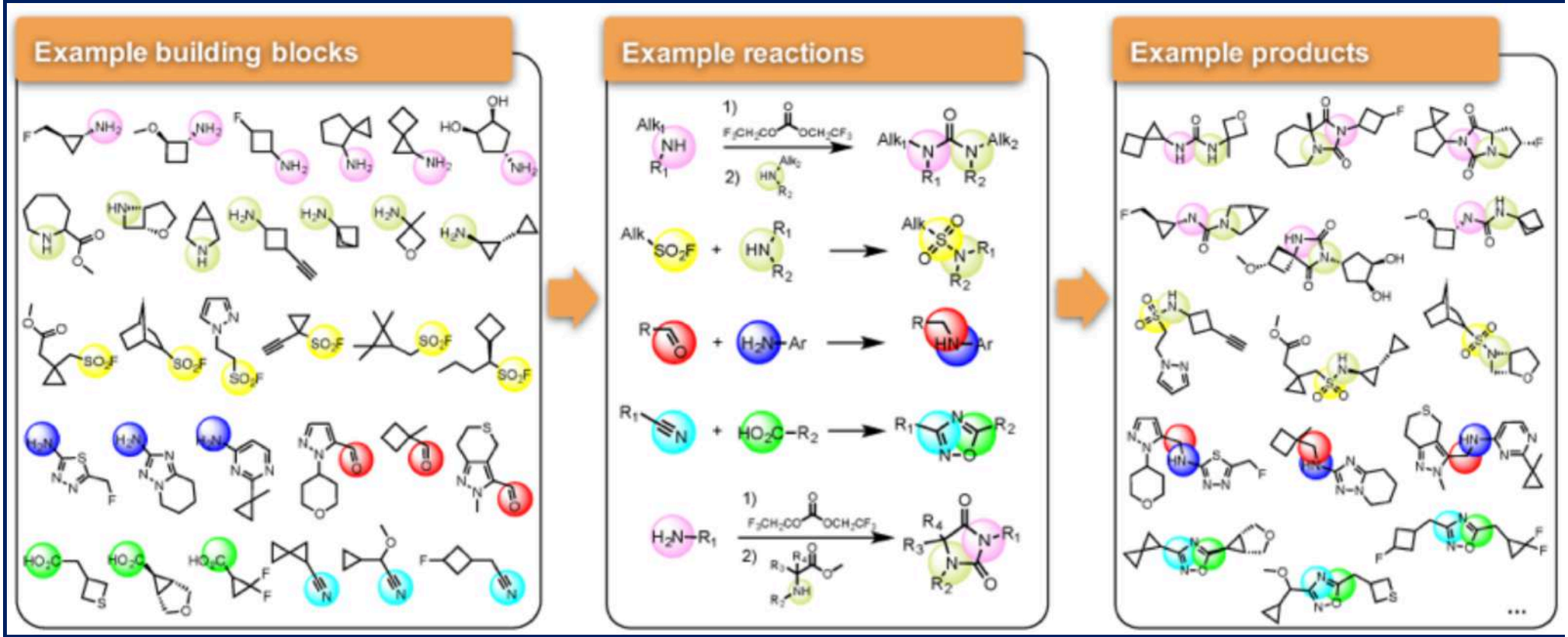


450 trillion complexes

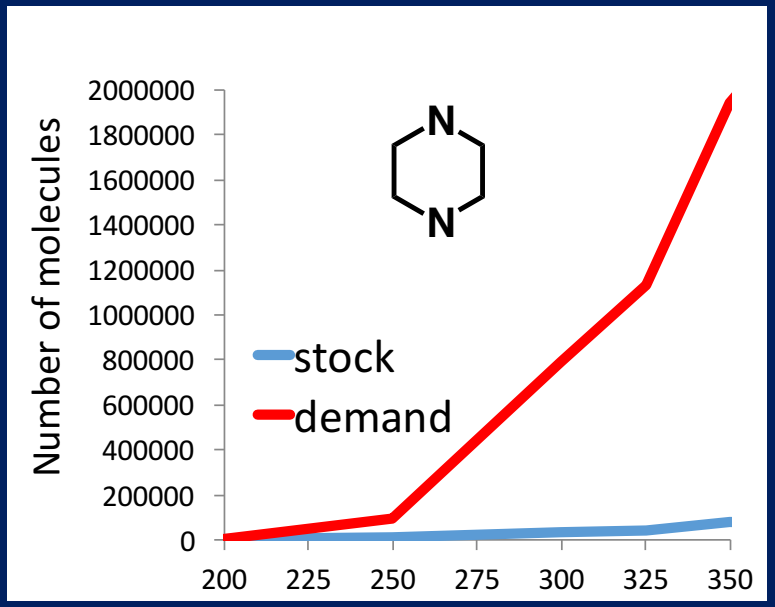




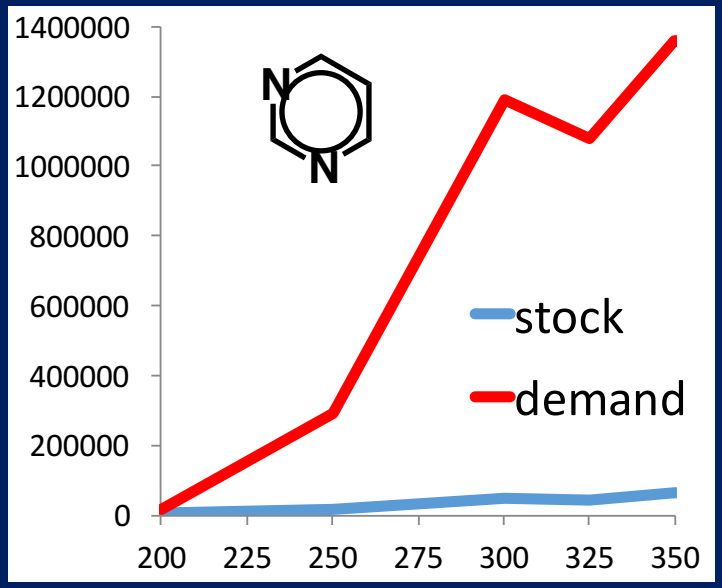
make-on-demand library:
109 2-component reactions
72,000 building blocks



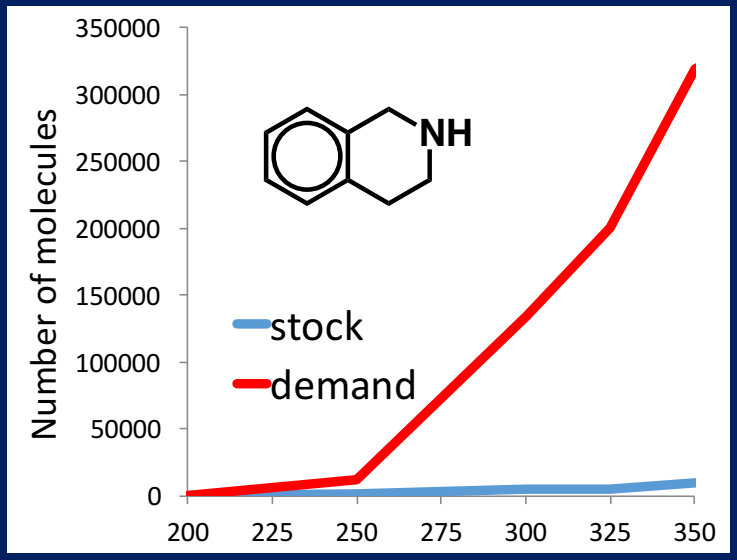
functional group coverage improves in “make-on-demand” vs. “in-stock”



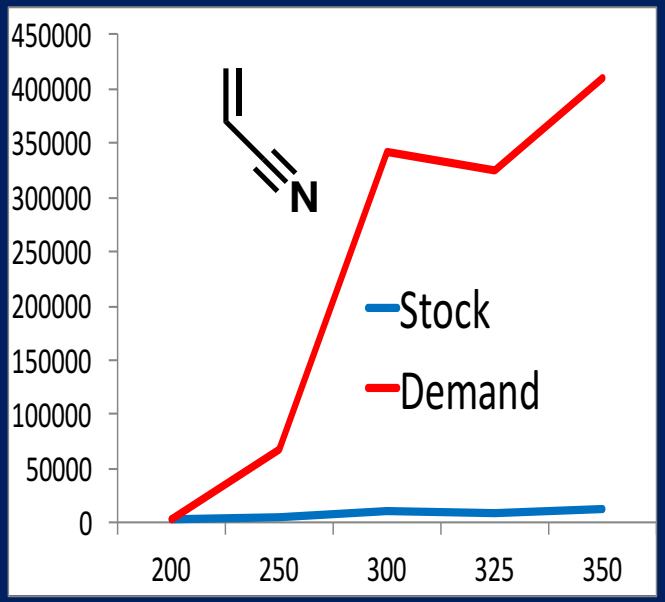
Mol. Wt



Mol. Wt



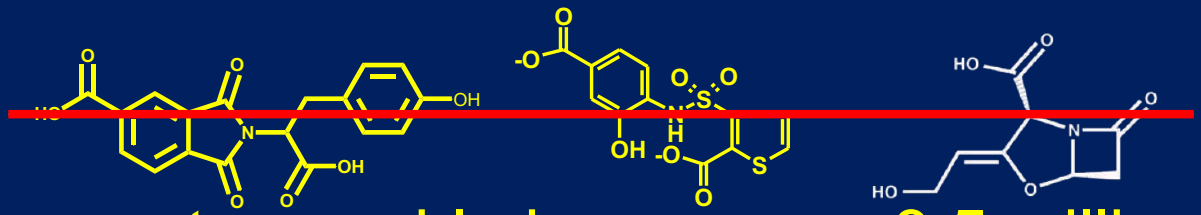
Mol. Wt



Mol. Wt

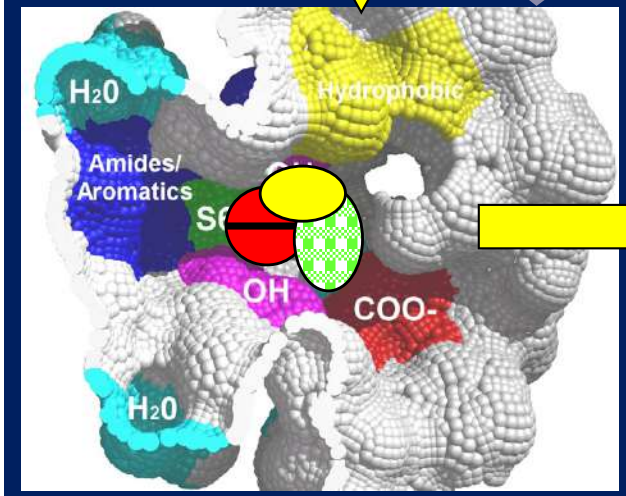
What you can expect from a docking campaign

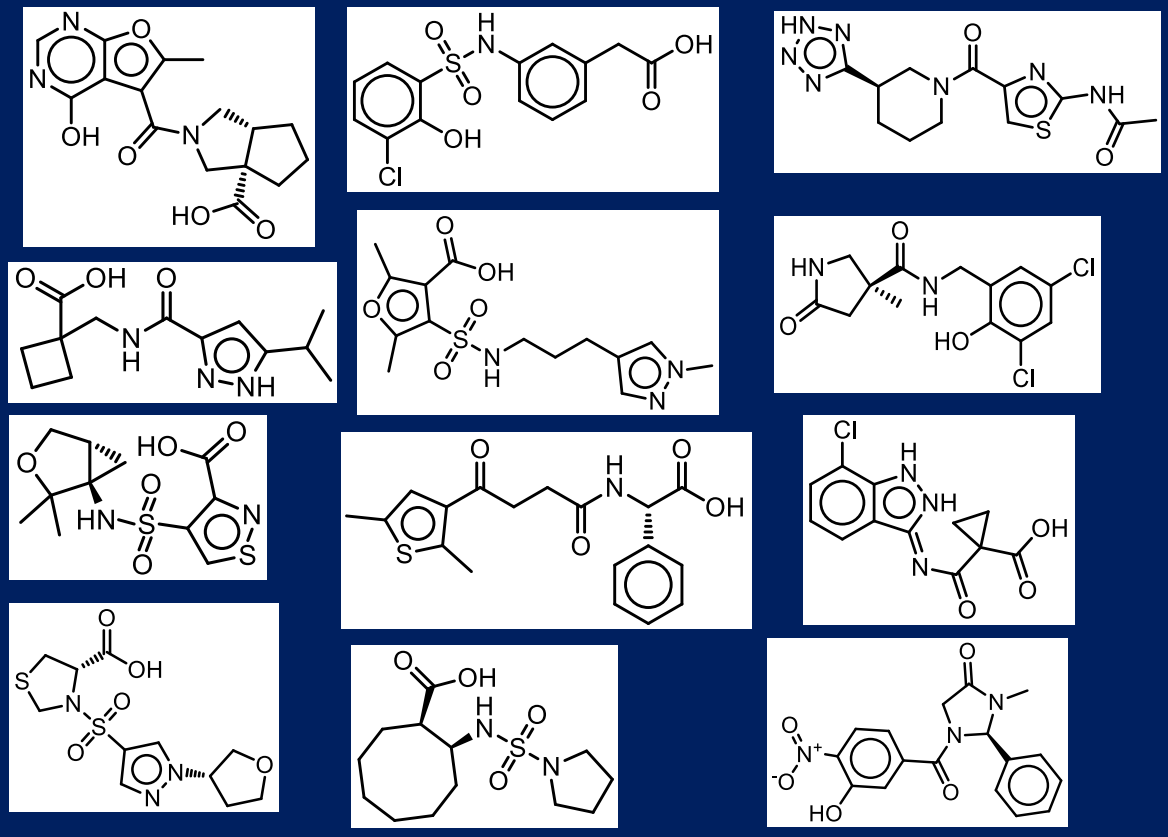
docking 99 million compounds vs β-lactamase predicts new chemotypes



cannot resemble knowns nor 3.5 million library

99 million cmpds



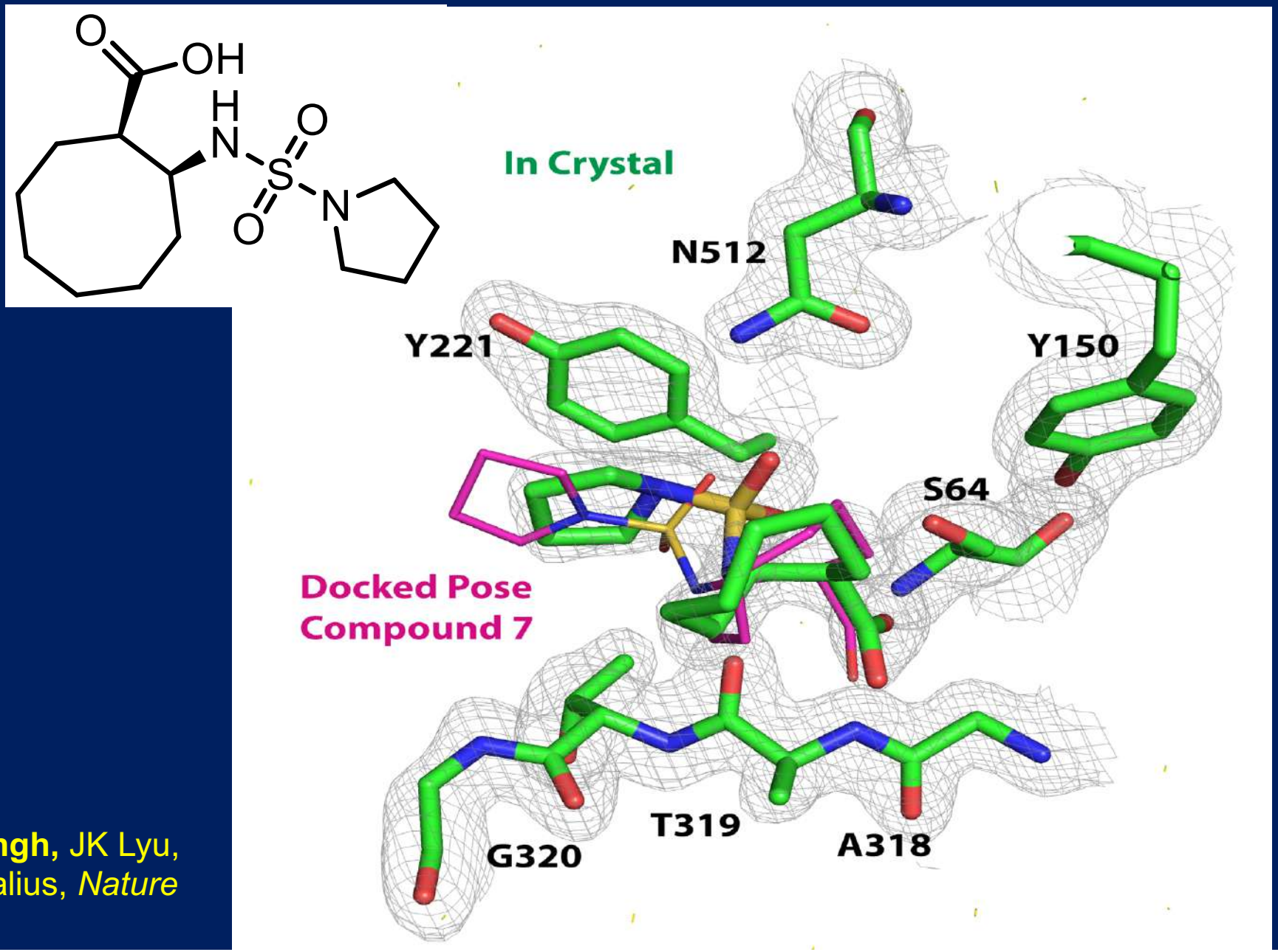


44 WACs picked from top 3000 clusters

28,310 core hours

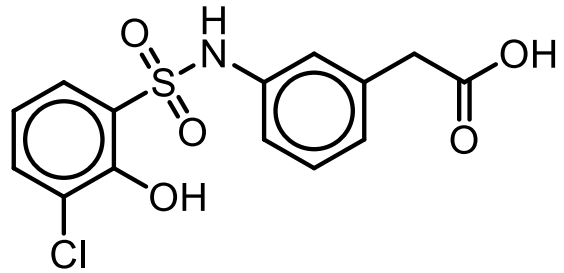
JK Lyu, Trent Balius, Isha Singh, unpublished

99 million vs β-lactamase: 11% hit rate, x-ray structure confirms a docking prediction



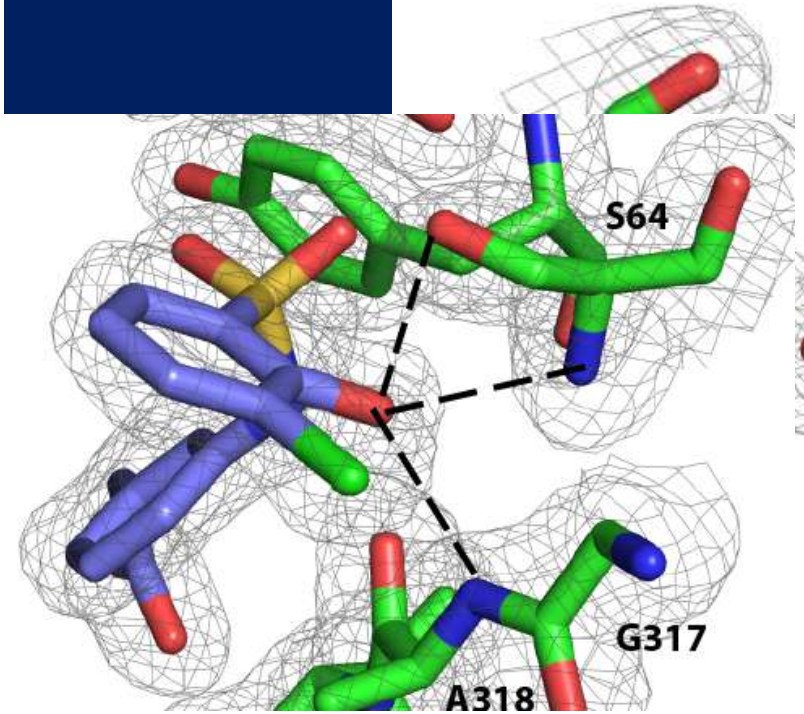
Isha Singh, JK Lyu, Trent Balius, *Nature* 2019

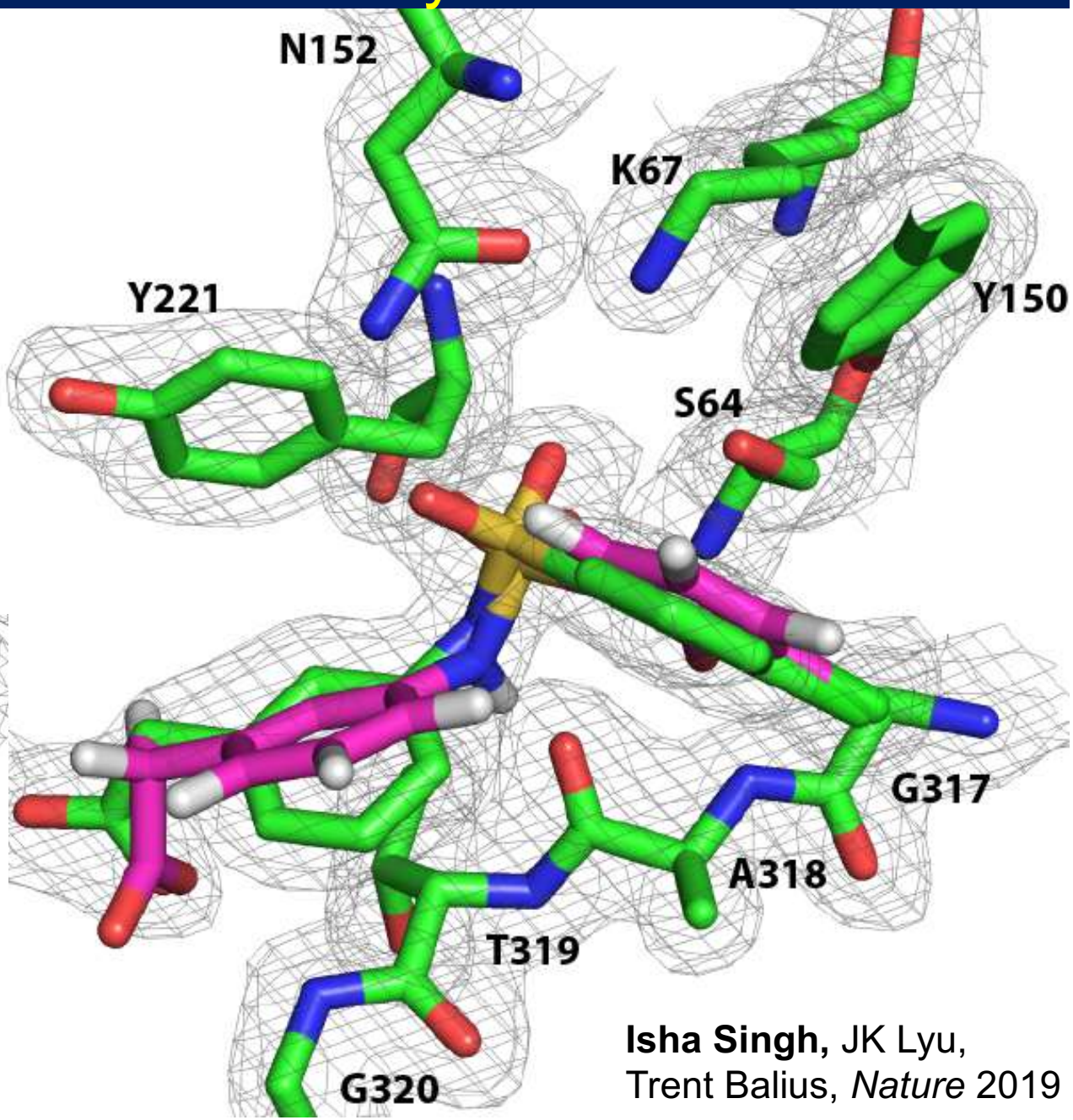
99 million vs β-lactamase: 11% hit rate, unprecedented 1.3 uM phenolate in the oxyanion hole



(Docked Pose)

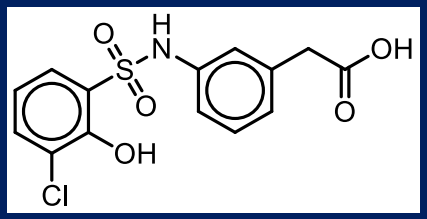
(RMSD 0.98 Å)

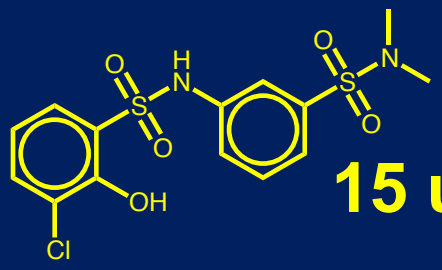




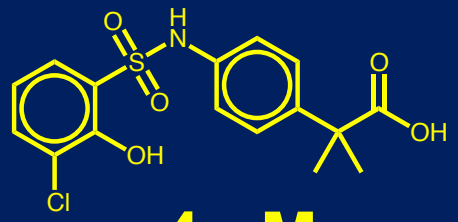
Isha Singh, JK Lyu, Trent Balius, *Nature* 2019

analog-by-catalog to 77 nM





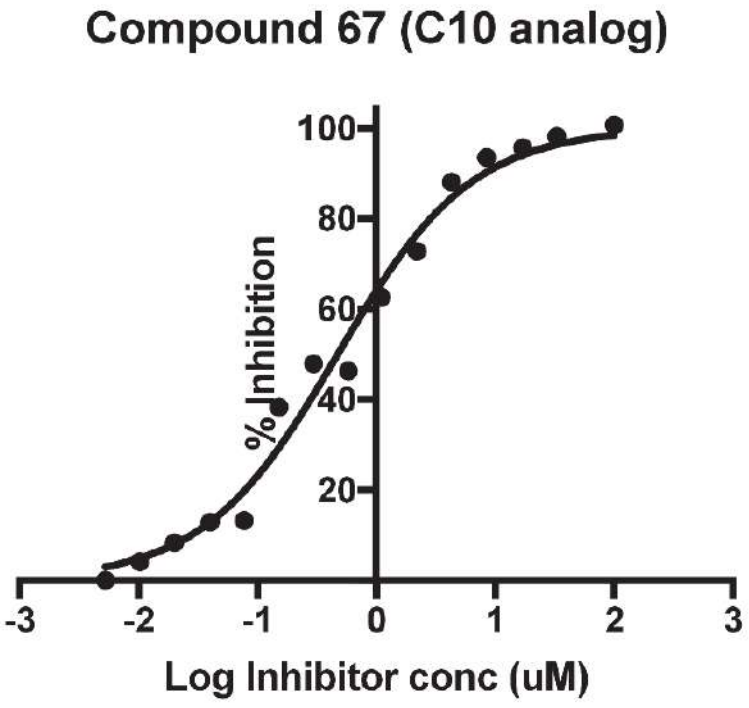
15 uM

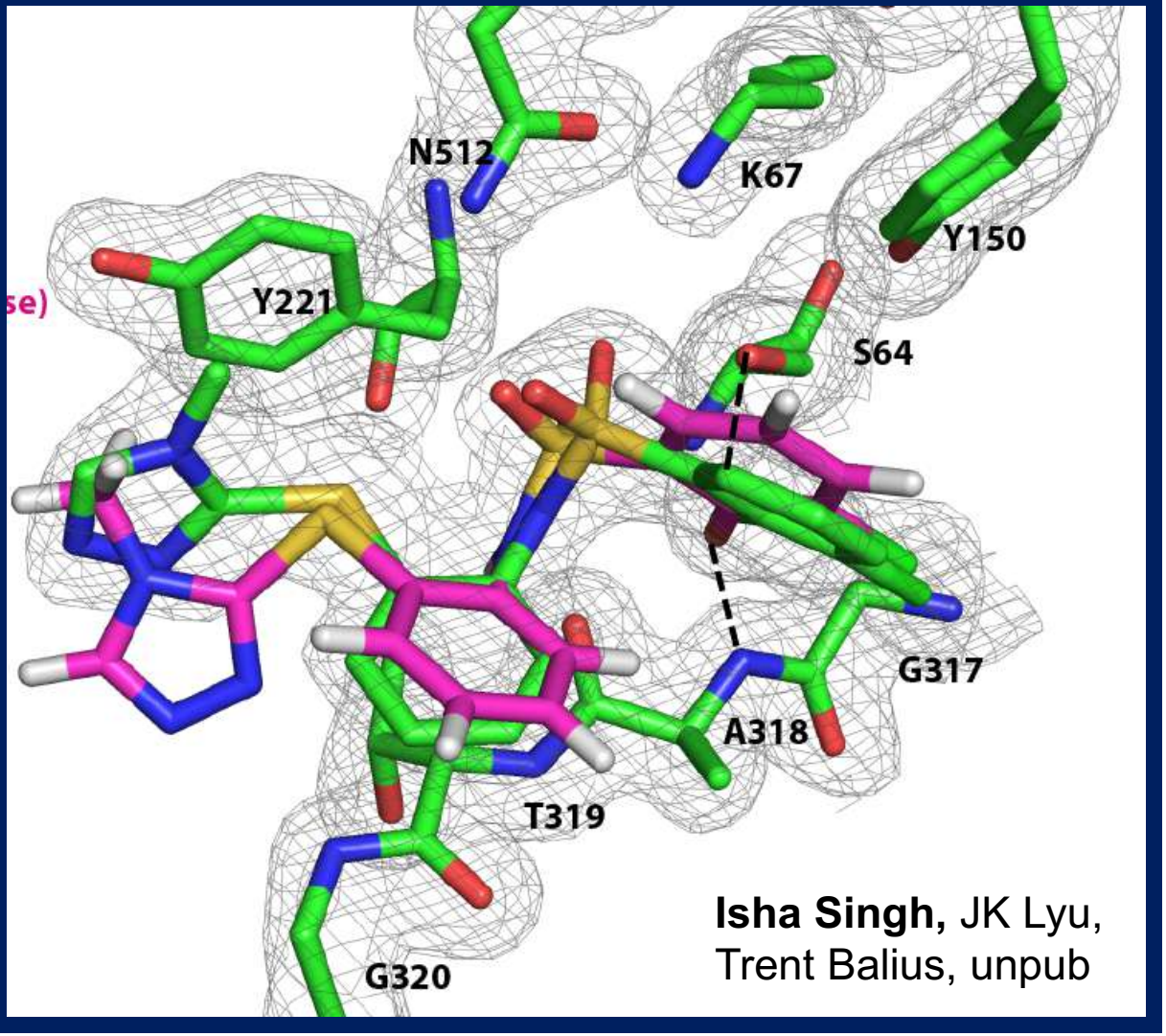


4 uM

77 nM

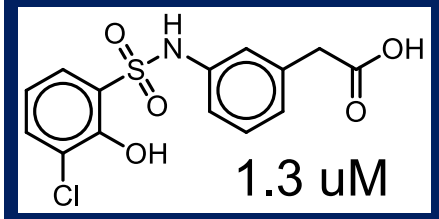
>400 analogs docked, test 10



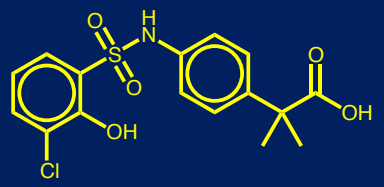


Isha Singh, JK Lyu, Trent Balius, unpub

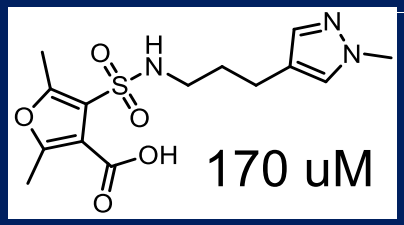
current problems: analog-by-catalog—
>2000 possible, which are best?



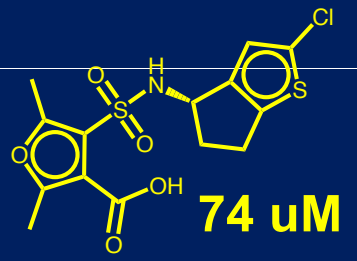
test 11

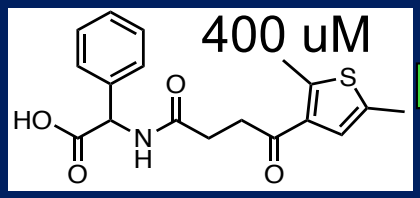


▪ ▪ ▪

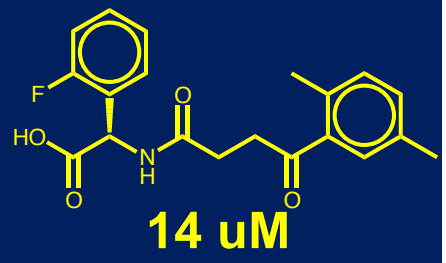


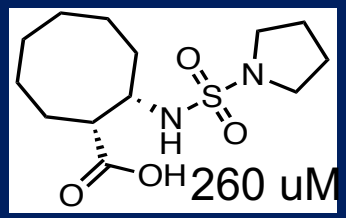
test 15



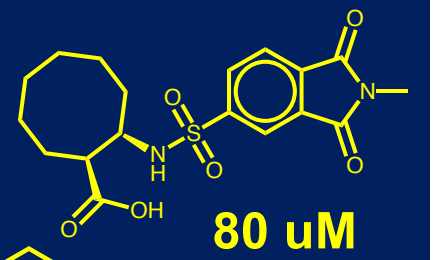


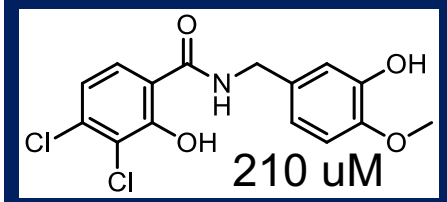
test 31



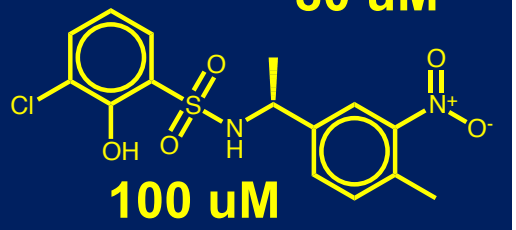


test 19





test 14

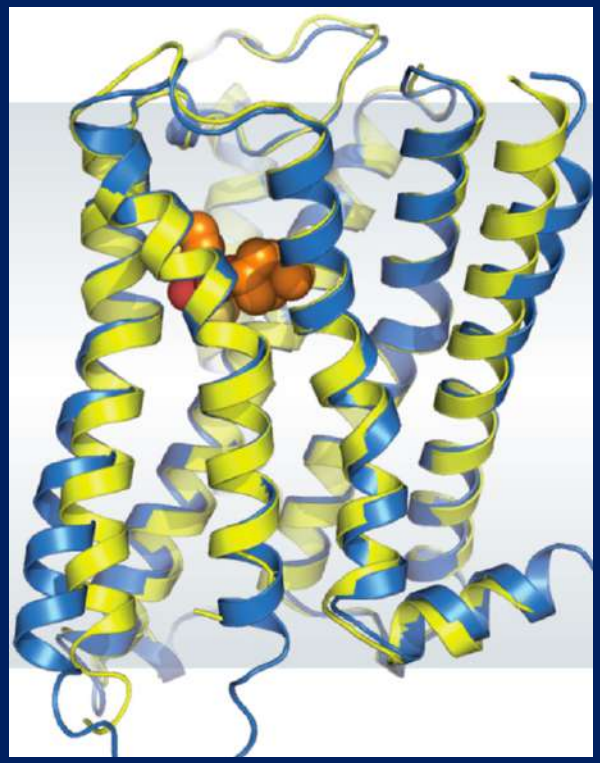


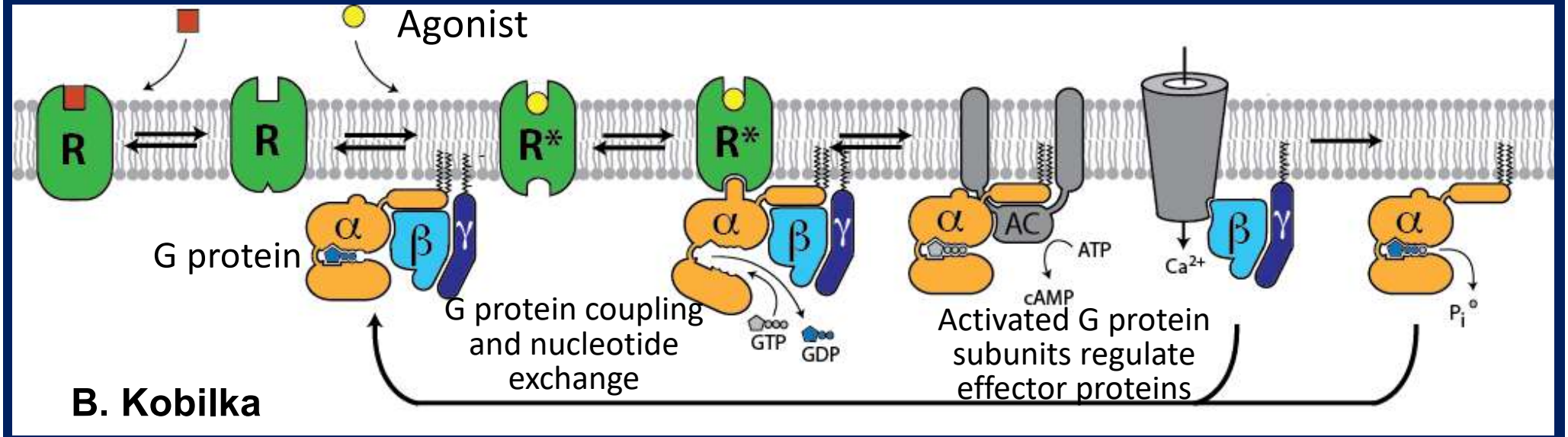
Isha Singh, JK Lyu, Trent Balius, *Nature* in press

Success stories:
GPCRs
DRD4
melatonin

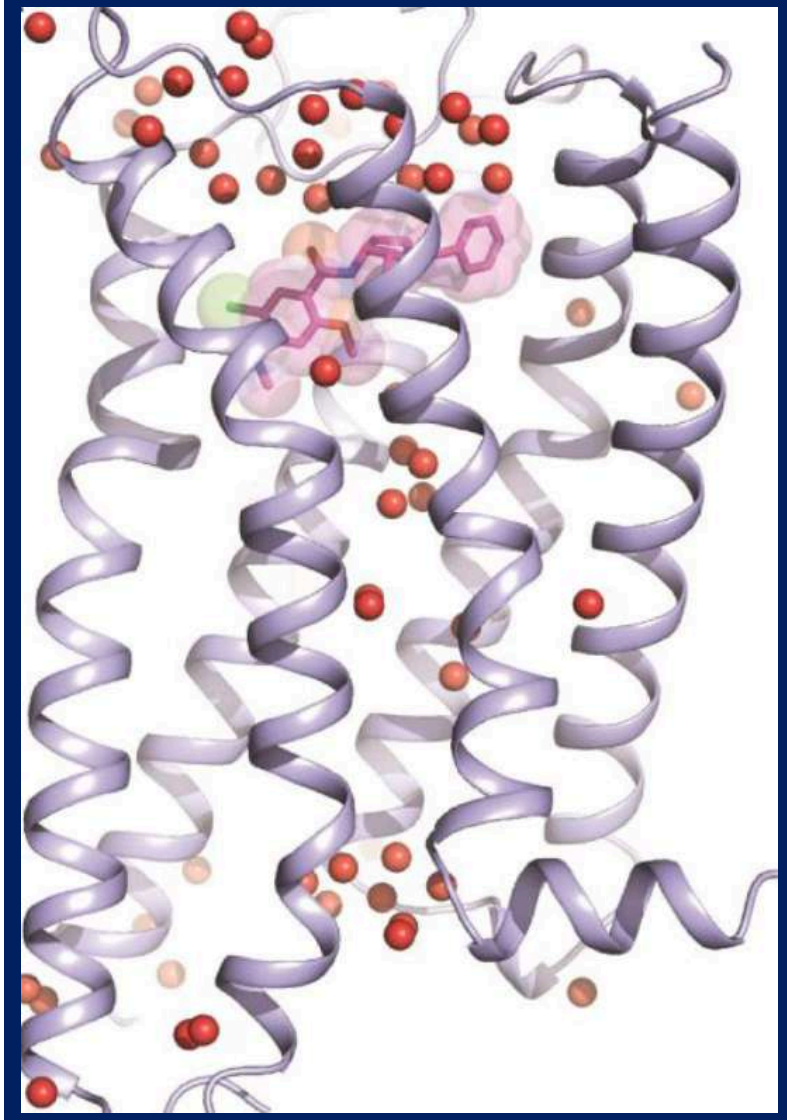
G Protein Coupled Receptors as targets

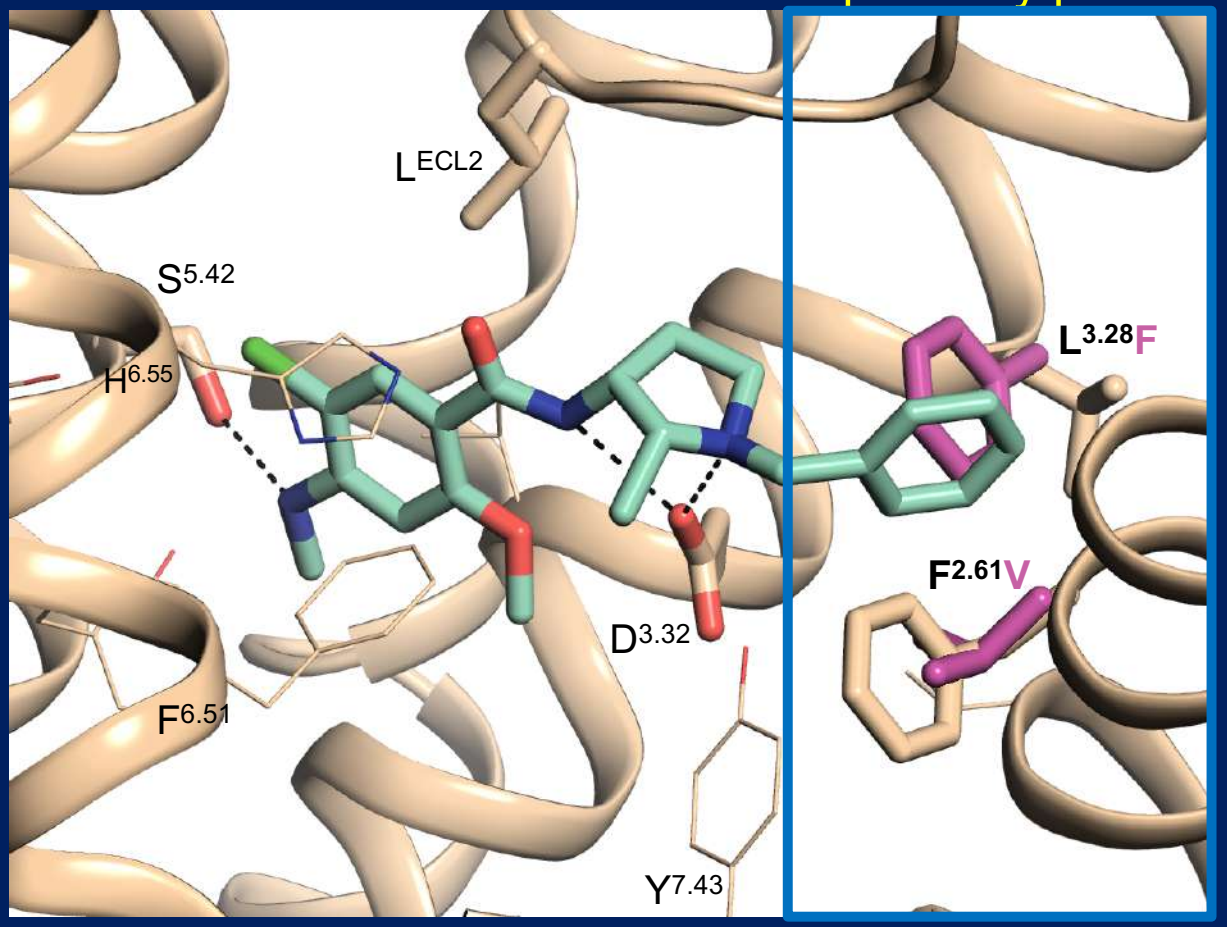
- 7-Transmembrane Spanners
- Signal across membrane
- Largest class of drug targets (>26% of drugs)
- Over >360 pharmacologically relevant GPCRs in the genome





x-ray structure of DRD4/nemonapride to 1.95 Å reveals a specificity pocket

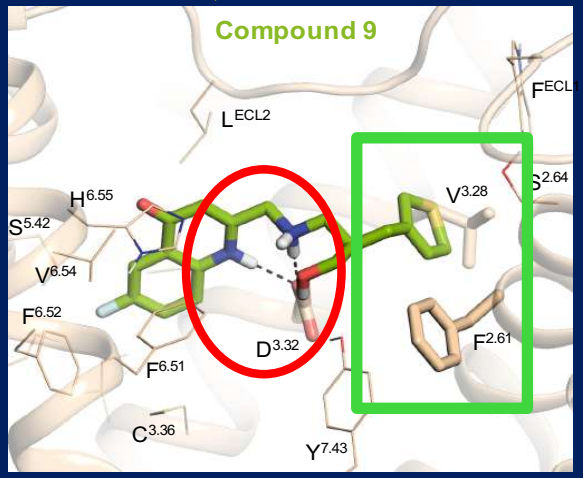


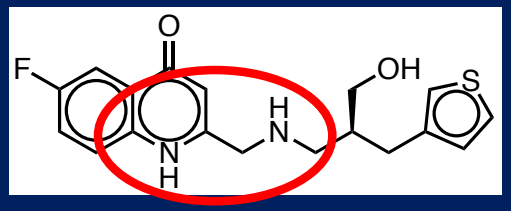


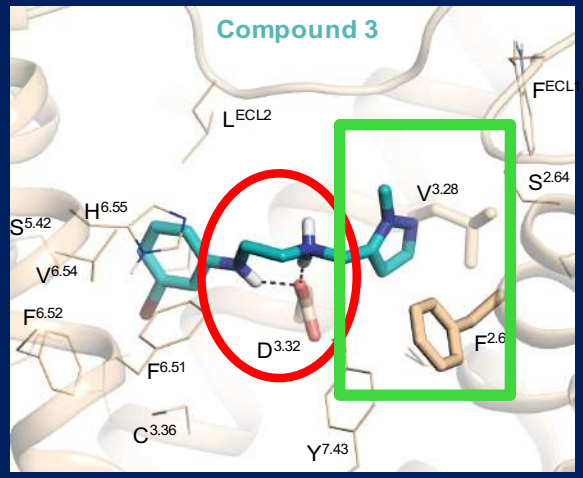
Wang, Wacker, Levit, ***Science*** in press

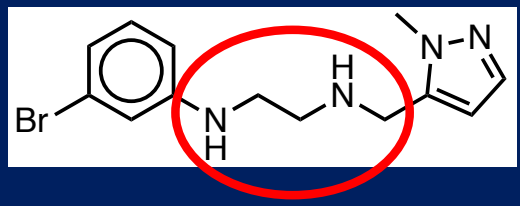
As structure refined, docking begins for novel, biased DRD4 agonists

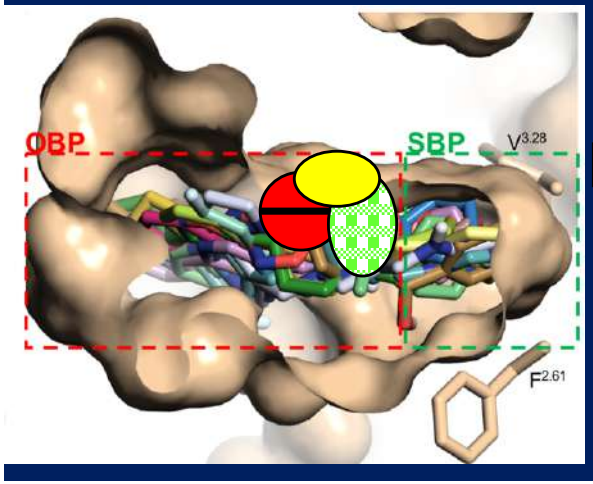
1.7 million molecules



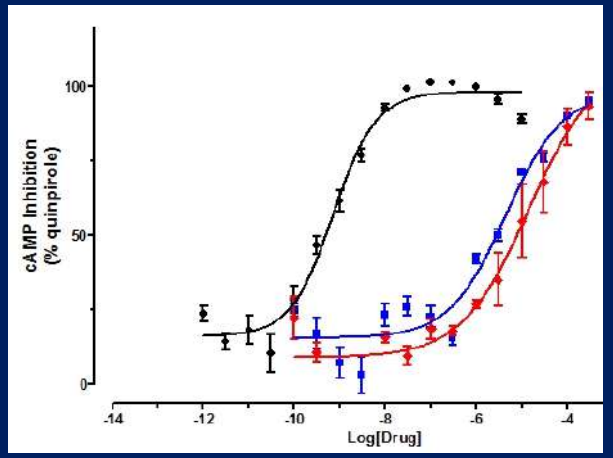








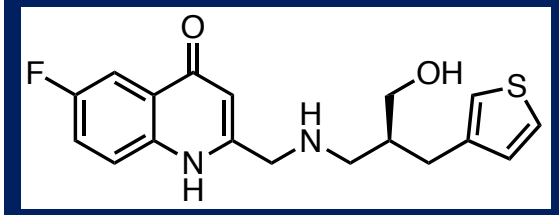
Test



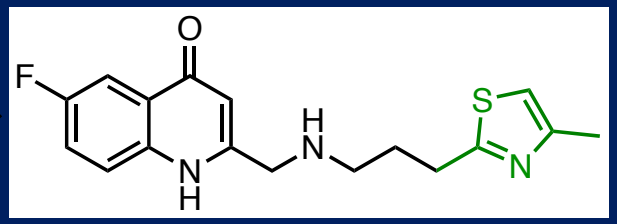
2 hits (10 tested):
DRD4: 56 & 213 nM (EC_{50}), 10x vs D2/D3

Wang, Wacker, Levit, ***Science*** in press

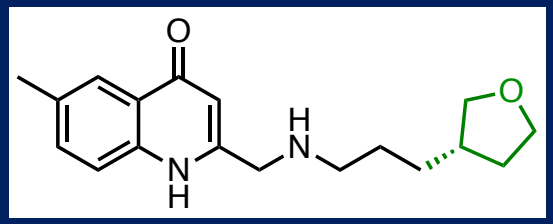
Analog-By-Catalog (ABC) improves affinity & specificity 10-fold (to 100x D2/D3)



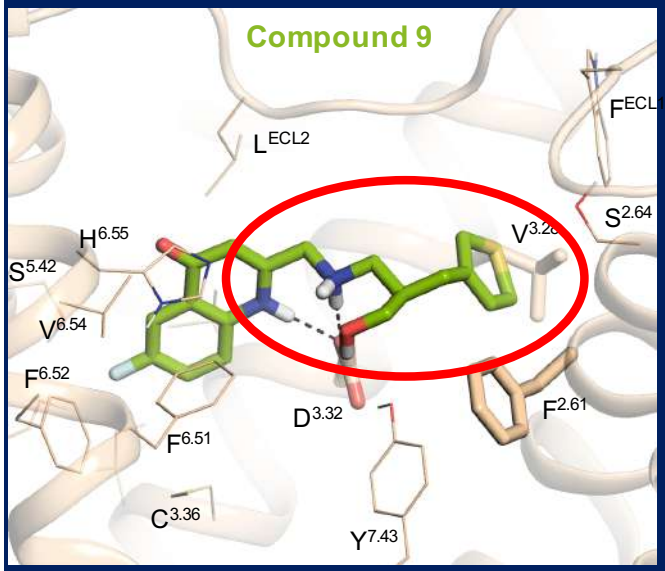
EC_{50} 213 nM

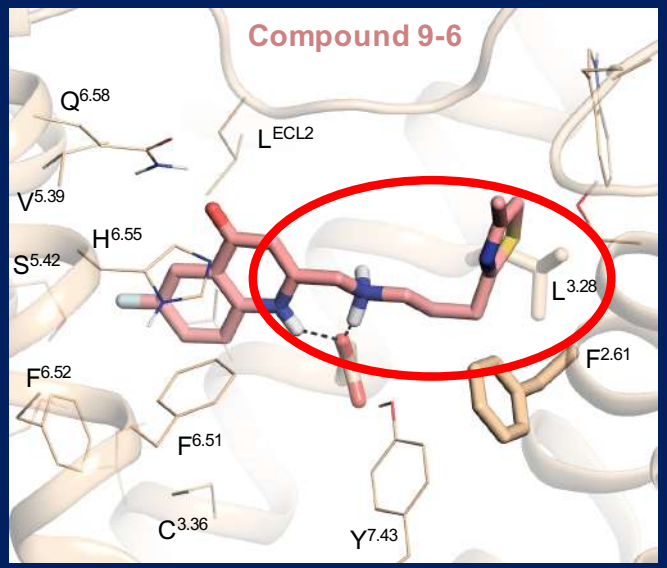


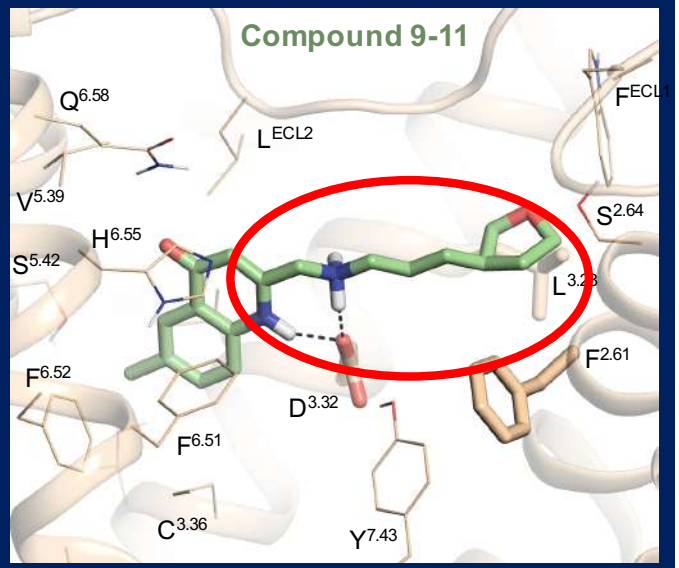
EC_{50} 33 nM



68 nM

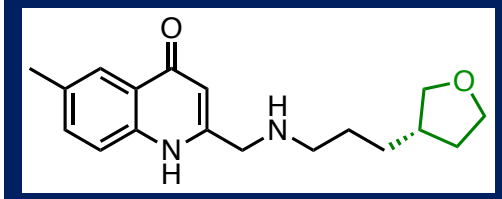


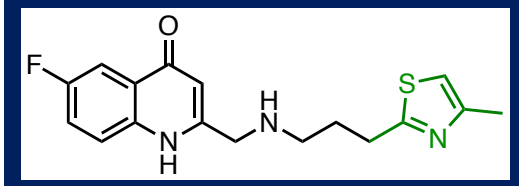


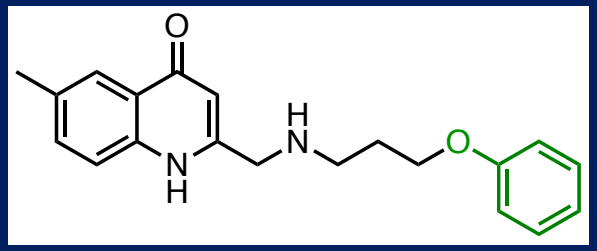


Wang, Wacker, Levit, ***Science*** in press

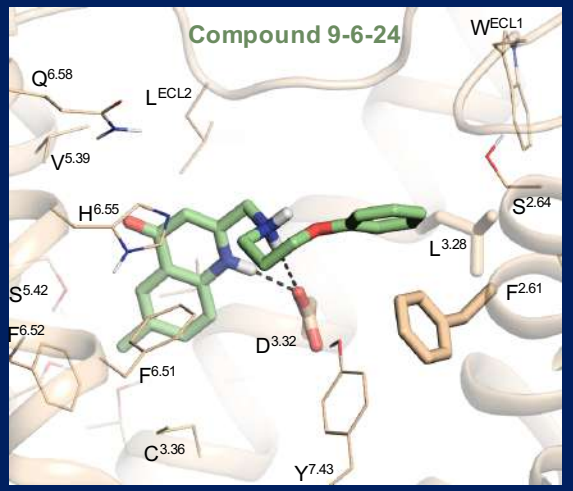
optimization to a 4nM Gi agonist (arrestin antagonist?) with 10,000-fold specificity vs D2/D3

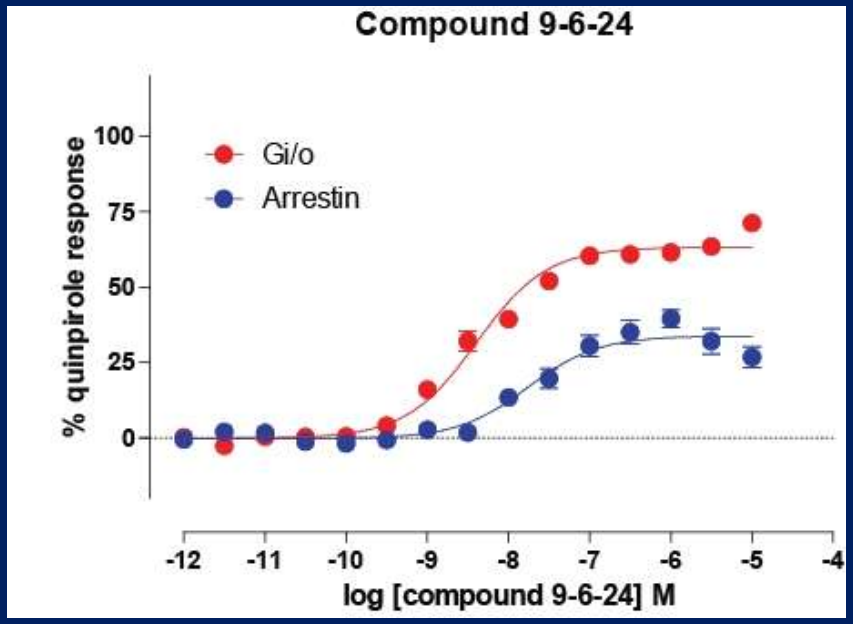


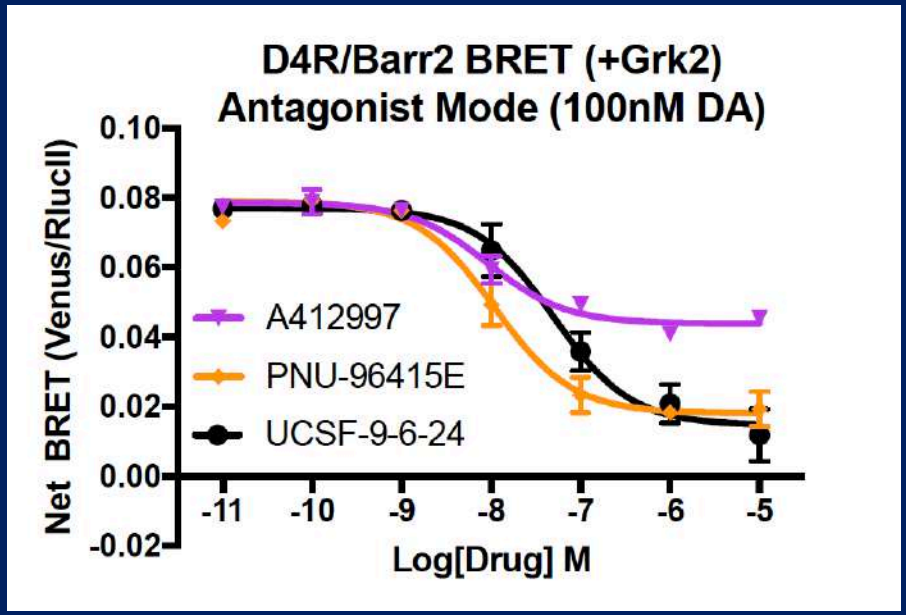




EC_{50} 4.2 nM (UCSF924)

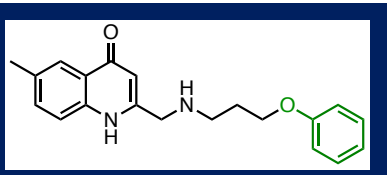


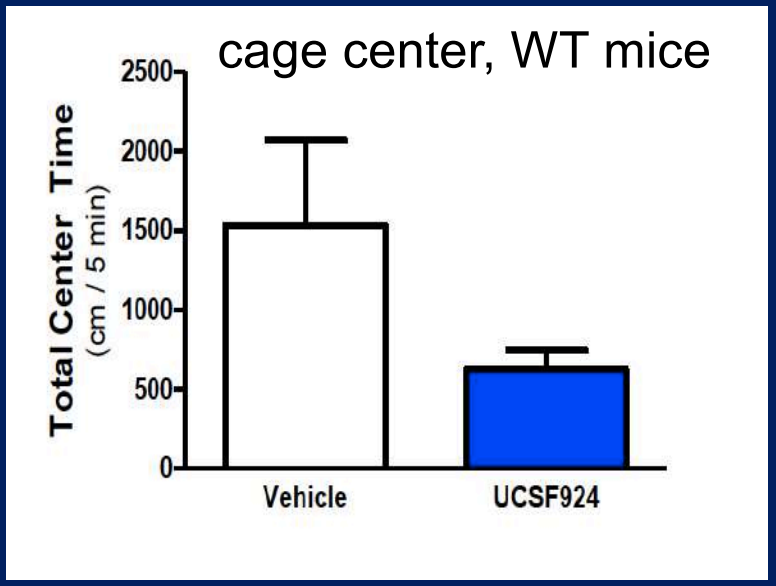


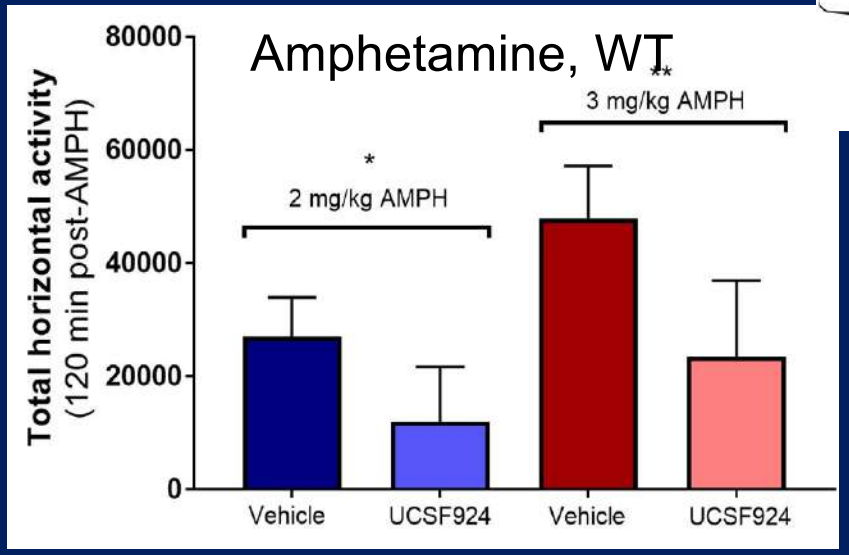


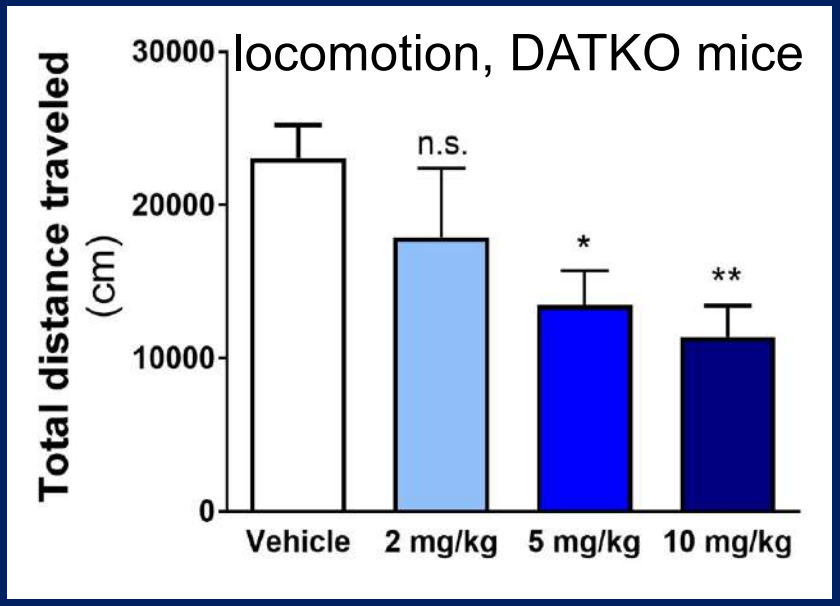
Wang, Wacker, Levit, ***Science*** in press; M. Caron & R. Chandrasekhar, unpub

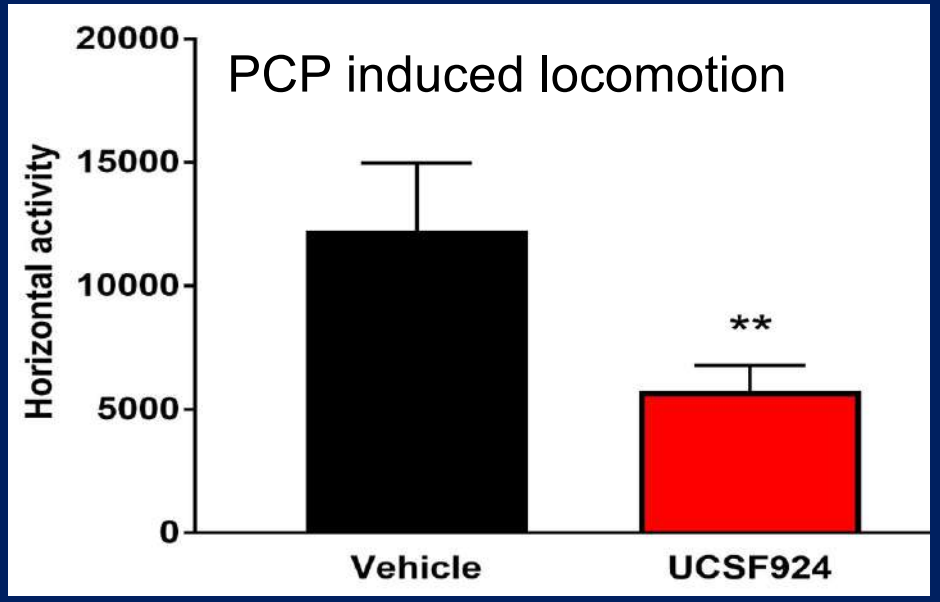
‘924 decreases mouse movement & exploration











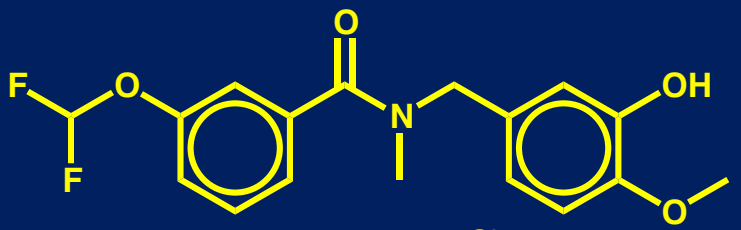
M. Caron & R. Chandrasekhar, unpub

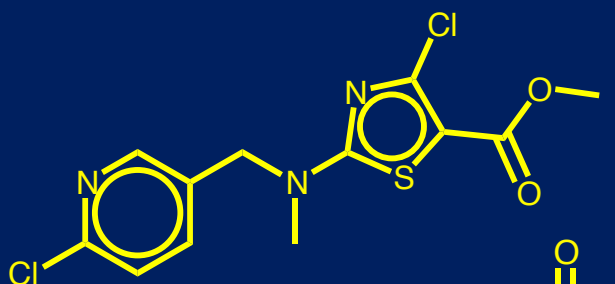
Success stories: melatonin

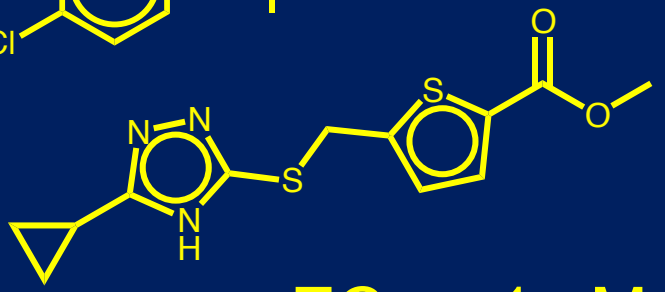
LSD vs melatonin GPCR crystal structures (2.9 A): novel ligands with new pharmacology

$1.5x10^8$ lead-like molecules
$1.7x10^6$ configurations
$2.5x10^{14}$ complexes
2 cluster days
BUT multiple control calculations

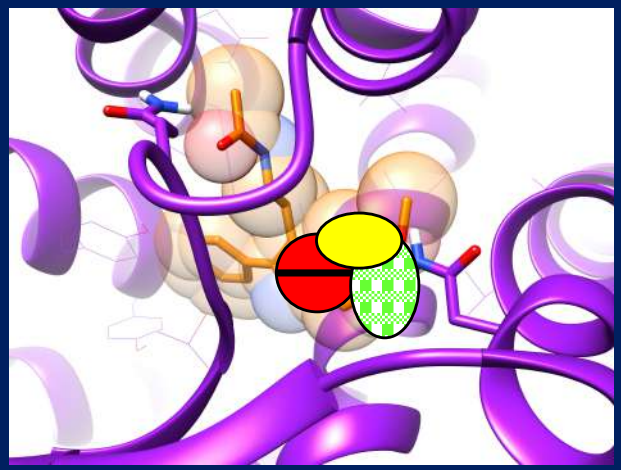
15 hits, 6 μM to 250 pM



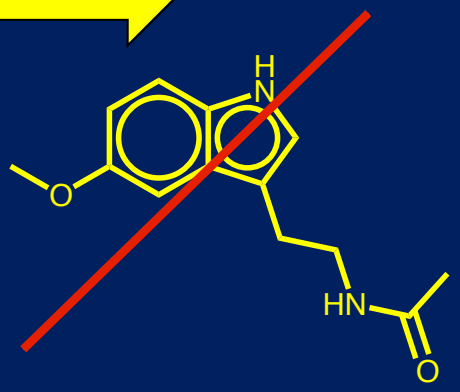




EC_{50} 1 nM

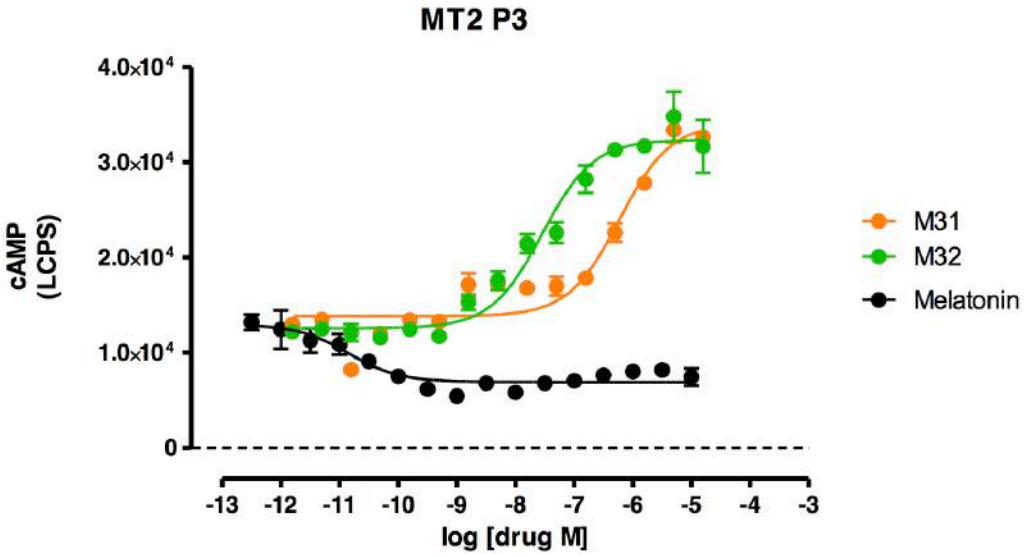


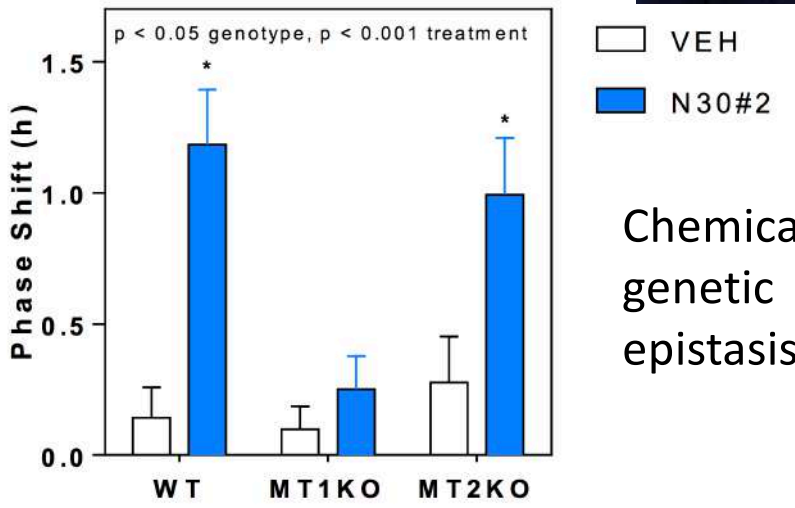
synthesize & test 38



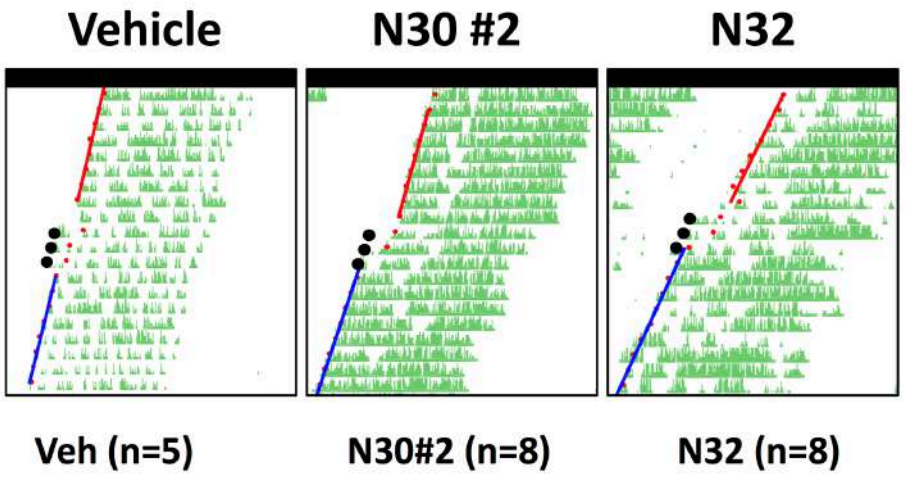
Reed Stein, Hye Jin Kang, Bryan Roth, **Enamine**

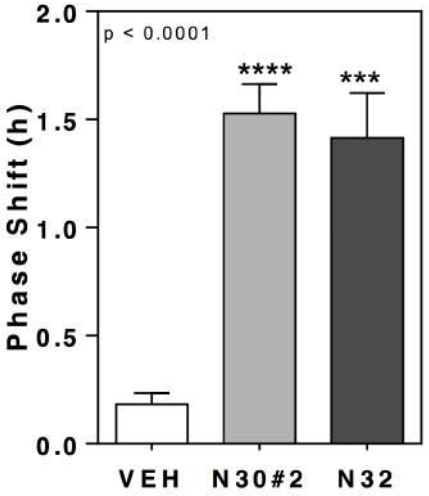
unexpected in vivo behavior of MT_1-selective inverse agonists in circadian rhythm





Chemical-genetic epistasis





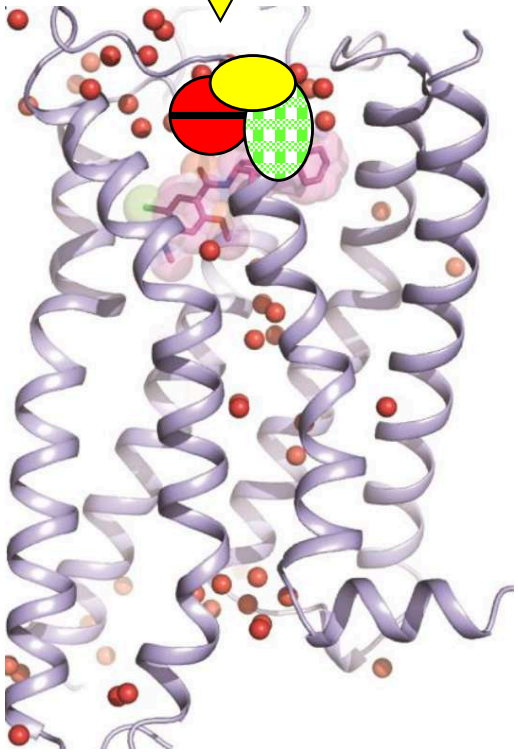
mouse circadian rhythm (1 mg/kg) shifted more than by melatonin

Margarita Dubocovich, in Stein et al, under review

Success stories: DRD4

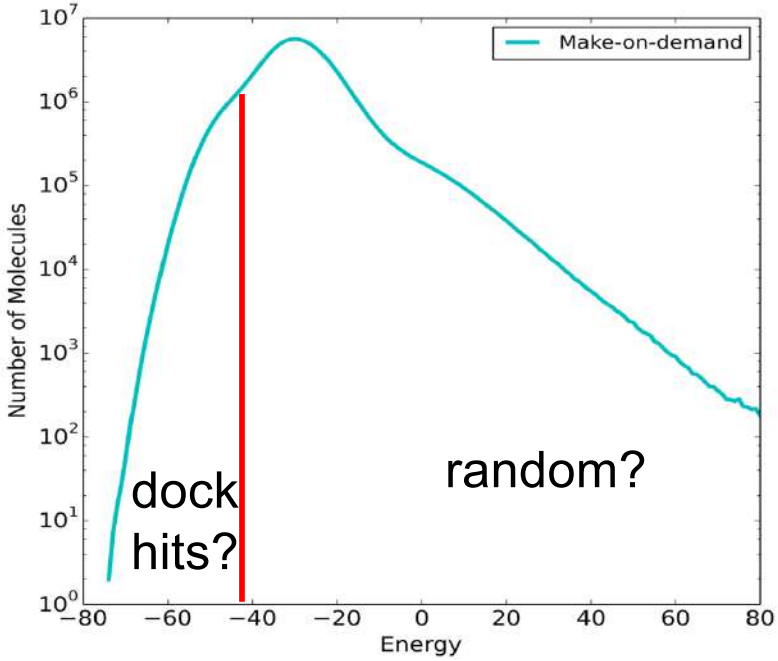
How many ligands *findable* in the library?
138 million vs. dopamine D4

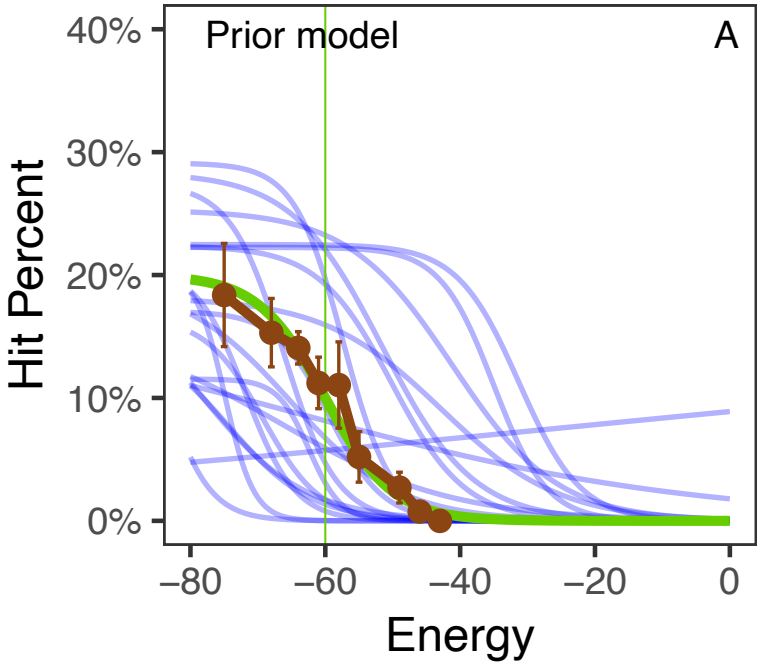
138 million molecules
10^{14} complexes
1 mol/sec;10^3 cores
1.8 cluster days



~~Test 30~~

Test 549



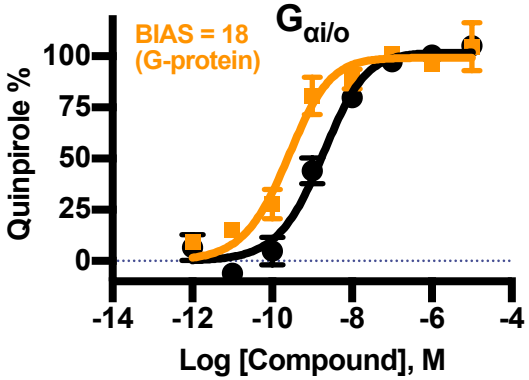


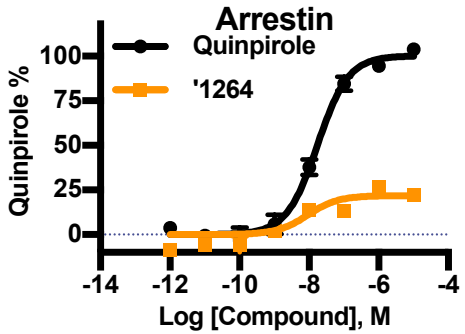
JK Liu, A Levit, T Balius, M. O'Meara

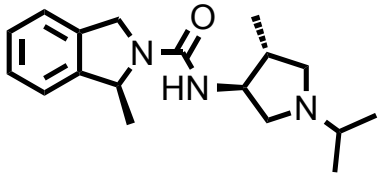
from top ranked: novel, selective, potent D4 ligands

ZINC361131264, 18x bias

Rank: 1067

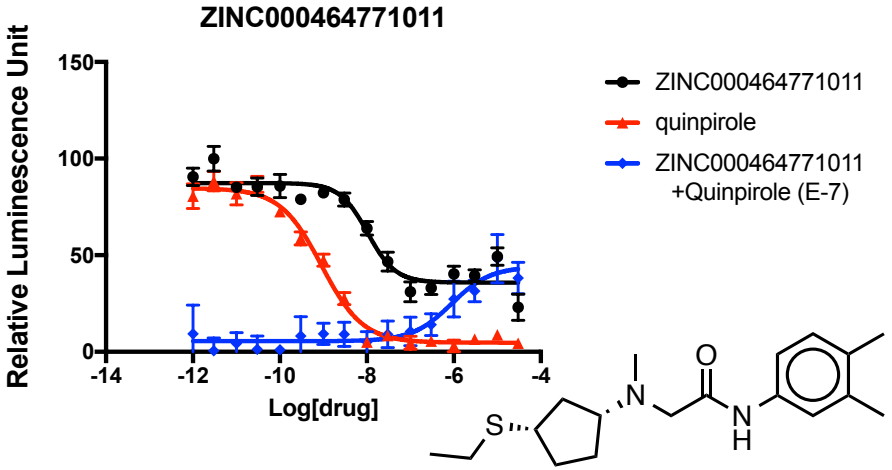






D4 = 0.18 nM
D2 > 10,000 nM
D3 > 10,000 nM

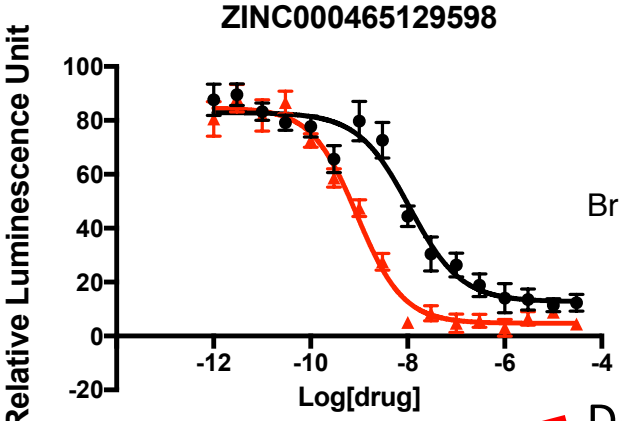
Rank: 936

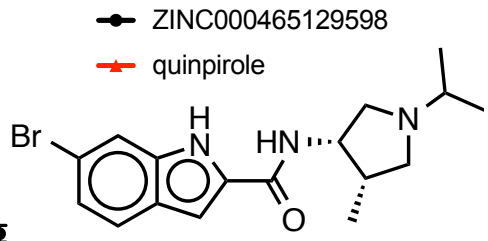


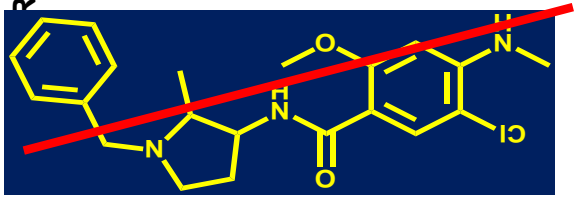
D4 = 10 nM
D2 > 10,000 nM
D3 > 10,000 nM

ZINC000465129598

Rank: 1565



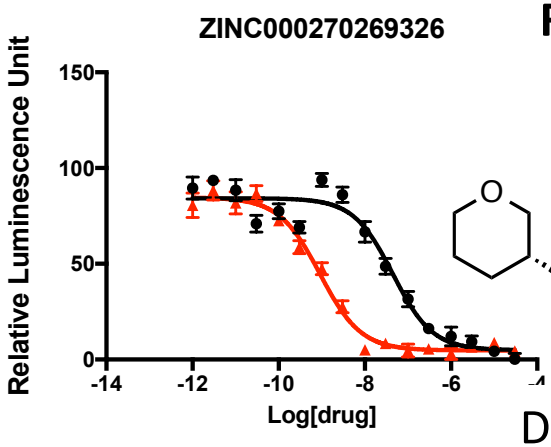


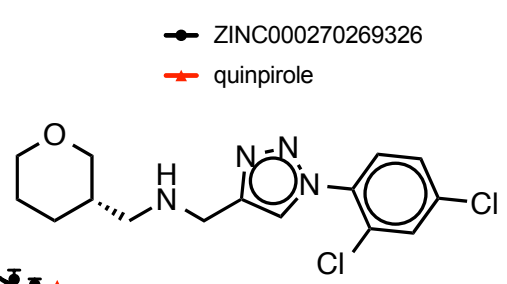


D4 = 11 nM
D2 > 10,000 nM
D3 > 10,000 nM

ZINC000270269326

Rank: 1510

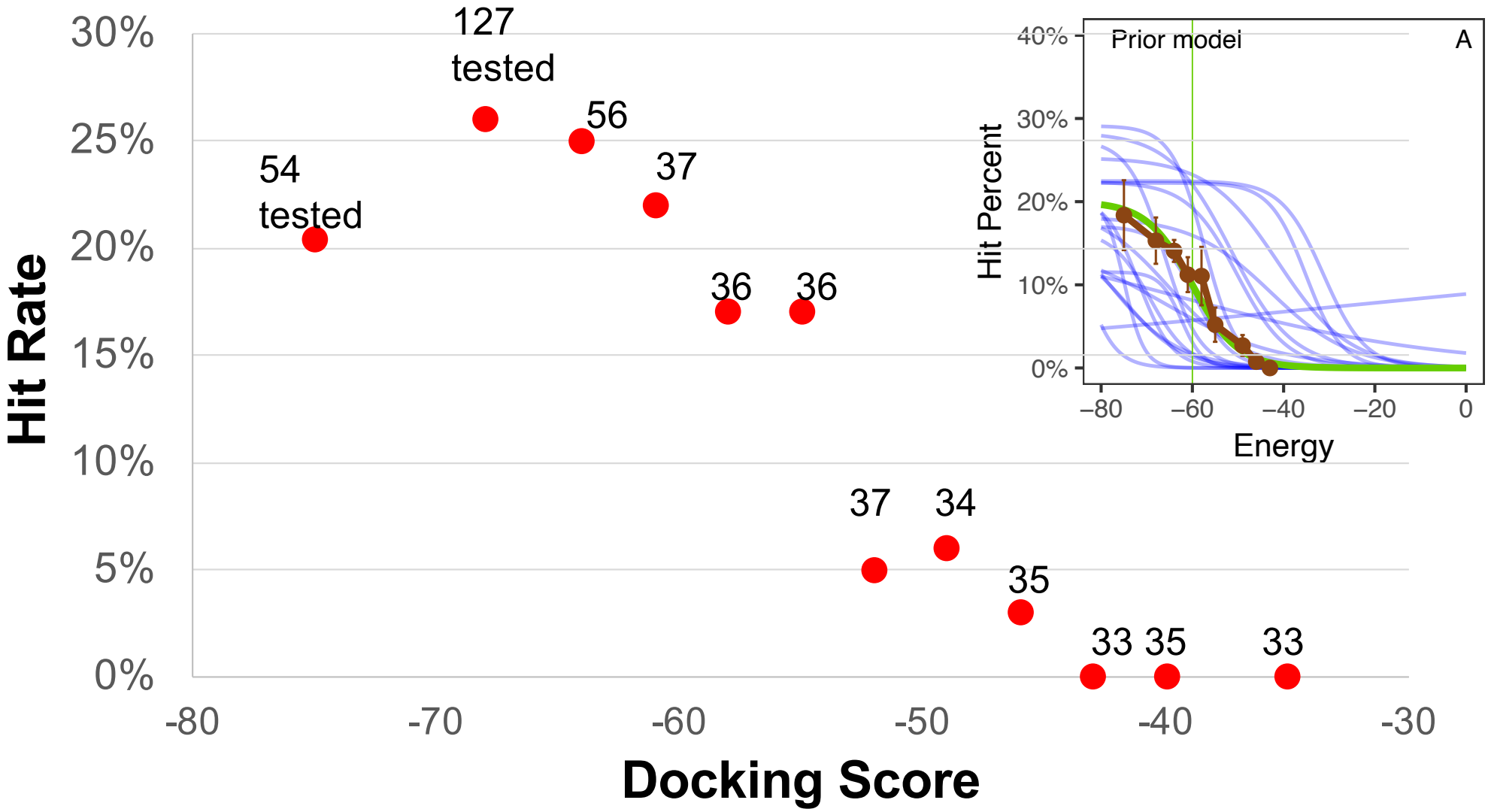




	ZINC000270269326
IC50	4.496e-008

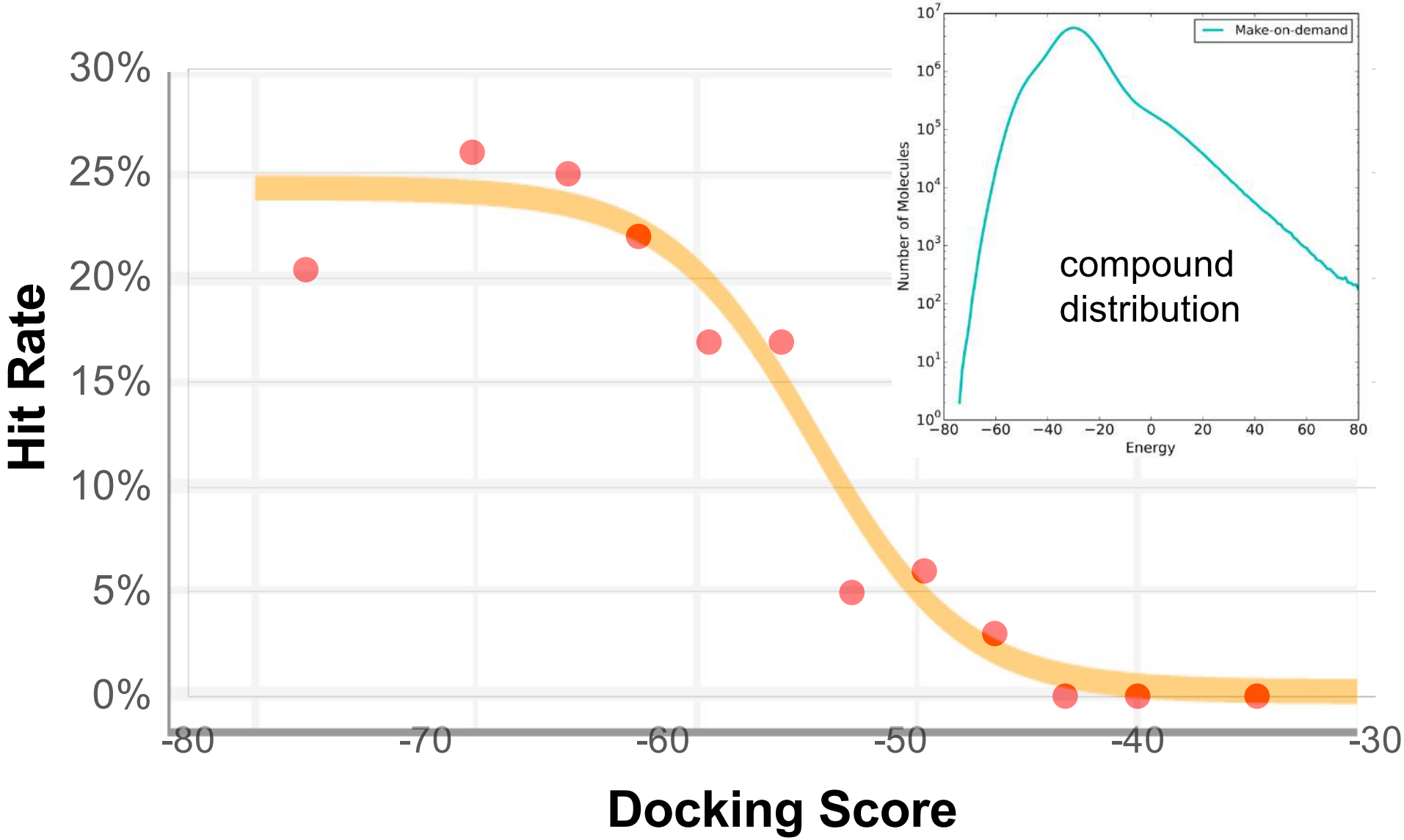
D4 = 45 nM
D2 > 10,000 nM
D3 > 10,000 nM

hit rates correlate with docking rank



Sheng Wang, Bryan Roth

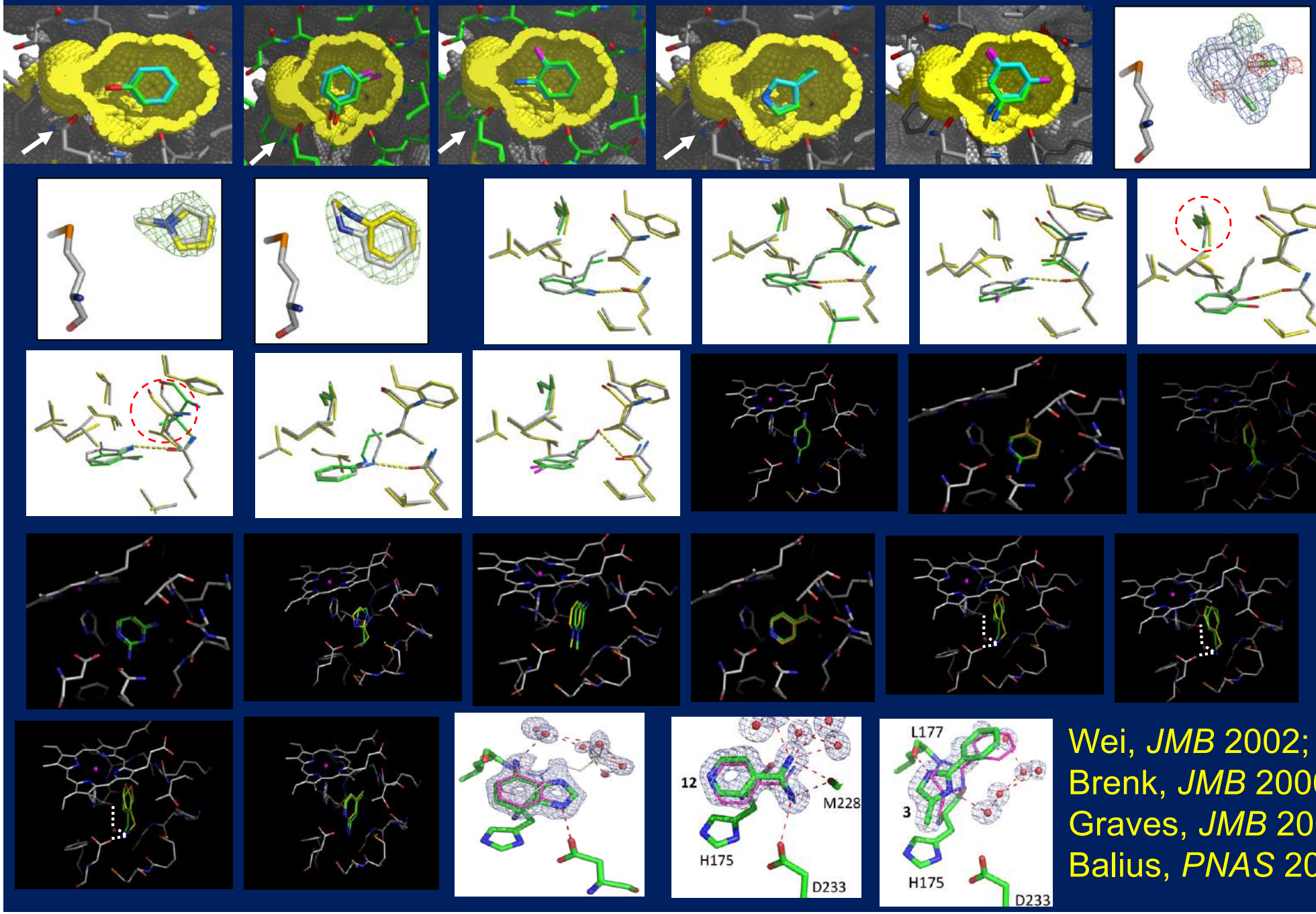
AUC predicts 479,000 (±40%) DRD4 actives (10 uM) in 72,000 (±30%) scaffold families



Matt O'Meara

Success stories: fragments

docking vs. xtal in cavity sites (if only they cured cancer)



Wei, *JMB* 2002;
Brenk, *JMB* 2006;
Graves, *JMB* 2008;
Balius, *PNAS* 2017

Failures/limitations/caveats

Current research questions

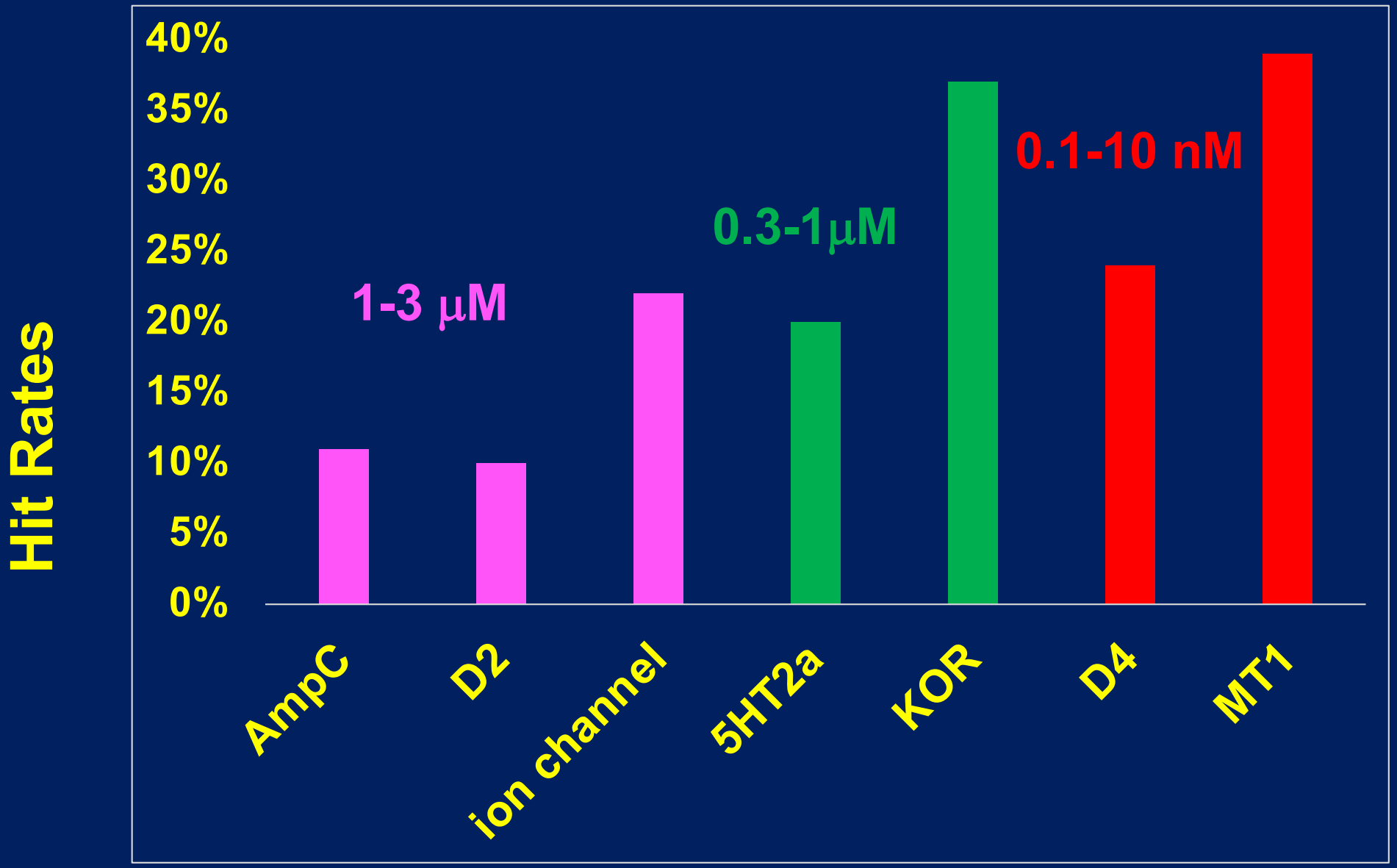
- When do we reach saturation?
- Reproducing D4R in two-to-three more targets
- Expanding the testing limits (to 1500, 5000?)
- Sampling the 10^{11} accessible molecules for docking
- Impact on polypharmacology?
- On selectivity?
- Bespoke, academic libraries?
  - Ellman, Hartwig
- Chemical Space (JK).
- Improving Analog-by-catalog with better theory (FEP)
- Tilting the plateau for better hit rates, rankings, affinities

Suggestions
on how to get started with docking

- Join a docking lab
- Attend a docking course (September, Byers Hall)
- Download the software and read the manual and the literature
- DOCK Blaster
- Make a friend in a docking lab
- Become a venture capitalist, hire people from a docking lab

Summary

results may vary: μM to pM in 9 LSD campaigns



Reed Stein, Magdalena Korczynska, Chase Webb, Anat Levit, unpublished