

Designing focused libraries

Computational Aspects of High-Throughput Screening planning and analysis

EMBL Heidelberg, Germany

Steffen Renner, Novartis

George Papadatos, EMBL-EBI

NOVARTIS

Overview

- Focused screening
- Where to focus
- Focus on biology
- Conclusions

Focused screening

Consciously descide what to screen

instead of

Screen everything at hand

Where to select from

Internal compound archive	Vendor catalogues / External resources	Chemical space
- Fast - Limited size? - Maybe propriatory - Maybe internal bio-data available	- Less fast - Larger than internal library - Nonpropriatory - Maybe external biodata available	- Slow - Designed libraries - Very large chemical space - propriatary

How many compounds to screen?

Higher chance not to miss hits / to find best compound

More compounds screened

Assays required with lower costs per cpd

Pilot screens for larger screens

Iterative screening

Avoid testing redundant chemical space

Less compounds screened

More complex assays possible with higher costs per cpd

NOVARTIS

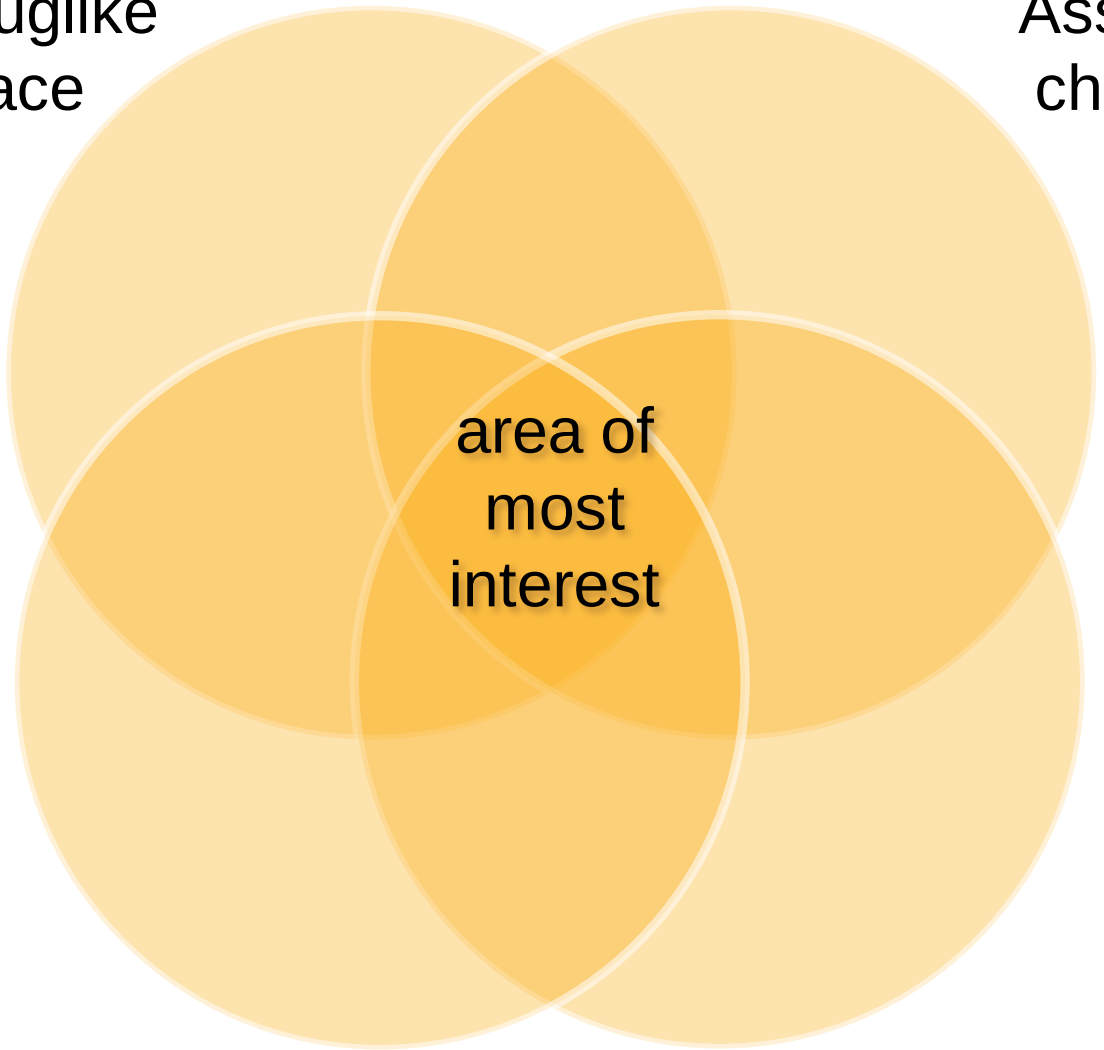
Overview

- Focused screening
- Where to focus
- Focus on biology
- Conclusions

What is interesting chemical space

Leadlike / druglike chemical space

Assay compatible chemical space



Target / biological relevant chemical space

Tox / adverse effect free chemical space

NOVARTIS

What are high quality hits: Interviews with chemists from Novartis project teams

30 Novartis project teams interviewed about high quality hits

Most important properties:
- MW <450 / leadlikeness
- water soluble
- membrane permeable

Hitlists:
- Scaffold diversity
-5-20 actives per series desired

=> Series size of 100-500 cpds

Stereocenters:
1-4 defined stereocenters are attractive

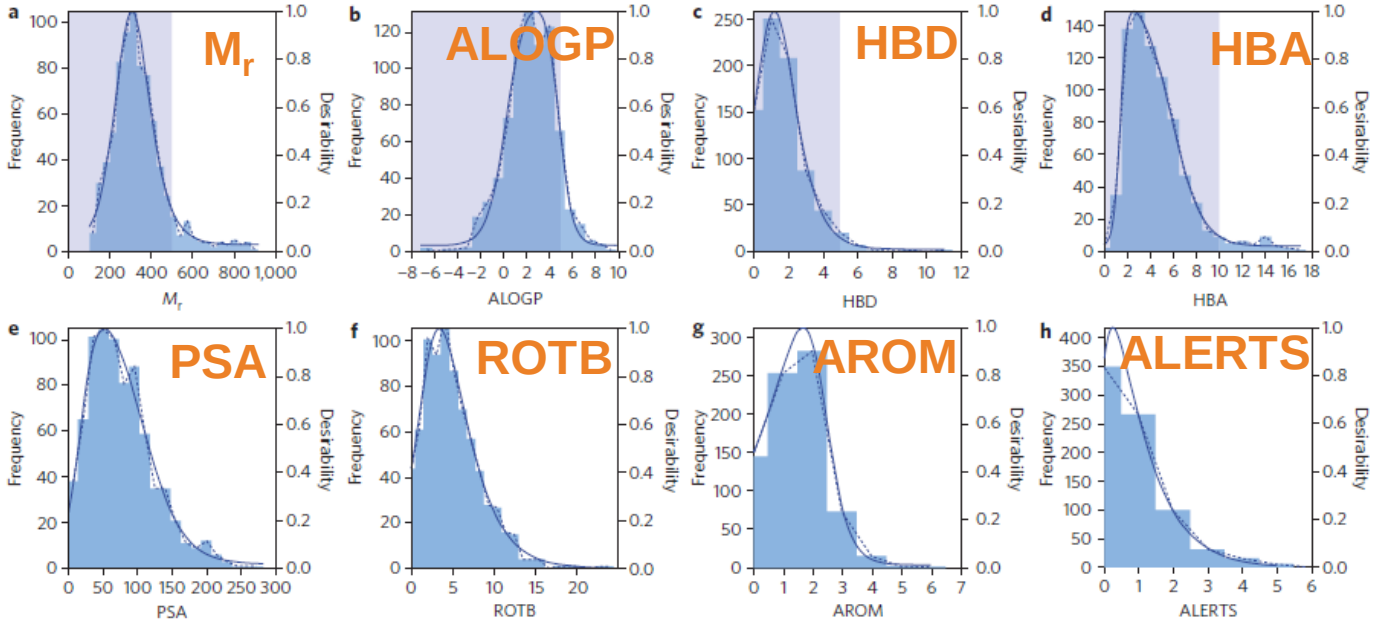
Non-/ Exclusive cpds:
- *Exclusive* high value
- *Nonexclusive* also valuable starting points for morphing – best without target class annotation

Renner S. et al, Future Med. Chem. 2011, 751-766

NOVARTIS

Quantifying the chemical beauty of drugs.

Quantitative estimate of druglikeness QED



$$QED_w = \exp\left(\frac{\sum_{i=1}^{n} w_i \ln d_i}{\sum_{i=1}^{n} w_i}\right)$$

d = individual desirability function
w = weight for property function
n = number of properties

Figure 1 | Histograms of eight selected molecular properties for a set of 771 orally absorbed small molecule drugs. a–h, Molecular properties M_r (**a**), lipophilicity estimated by atom-based prediction of ALOGP (**b**), number of HBDs (**c**), number of HBAs (**d**), PSA (**e**), number of ROTBs (**f**), number of AROMs (**g**) and number of ALERTS (**h**). The Lipinski-compliant areas are shown in pale blue in (**a**), (**b**), (**c**) and (**d**). The solid blue lines describe the ADS functions (equation (2)) used to model the histograms. The parameters for each function are given in Supplementary Table S1.

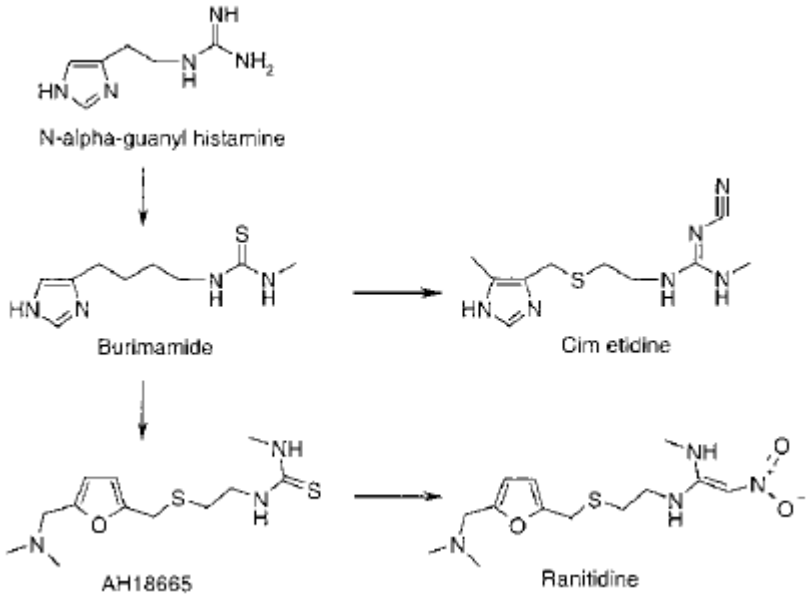
Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins AL, Nat. Chem. 2012, 4, 90-98

Quantitative measure more predictive than binary Ro5,
… but not leadlikeness, limited usefulness for screening library assessment..

NOVARTIS

Leadlikeness

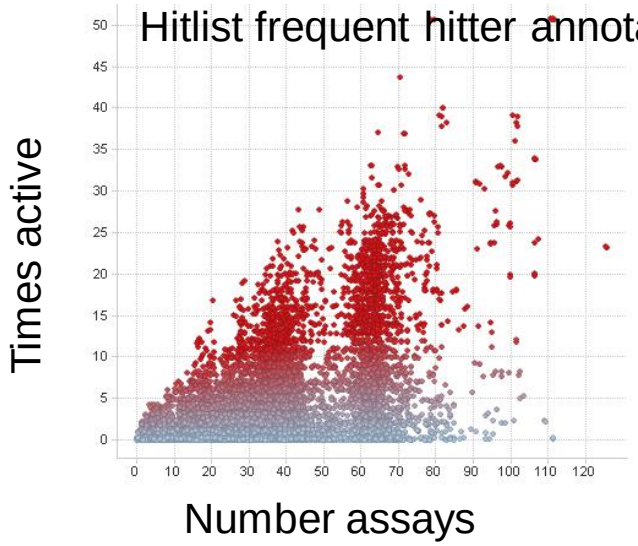
- Start with compounds that allow growing / optimization towards druglike molecules
- Start with compounds that have good LE and LipE if hits



Is there a difference between leads and drugs? A historical perspective.
Oprea TI, Davis AM, **Teague** SJ, Leeson PD.
J Chem Inf Comput Sci. 2001, 41, 1308-15.

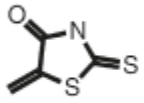
NOVARTIS

Assay compatible chemical space



Only possible for compounds from a screening deck that have been screened multiple times

New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays

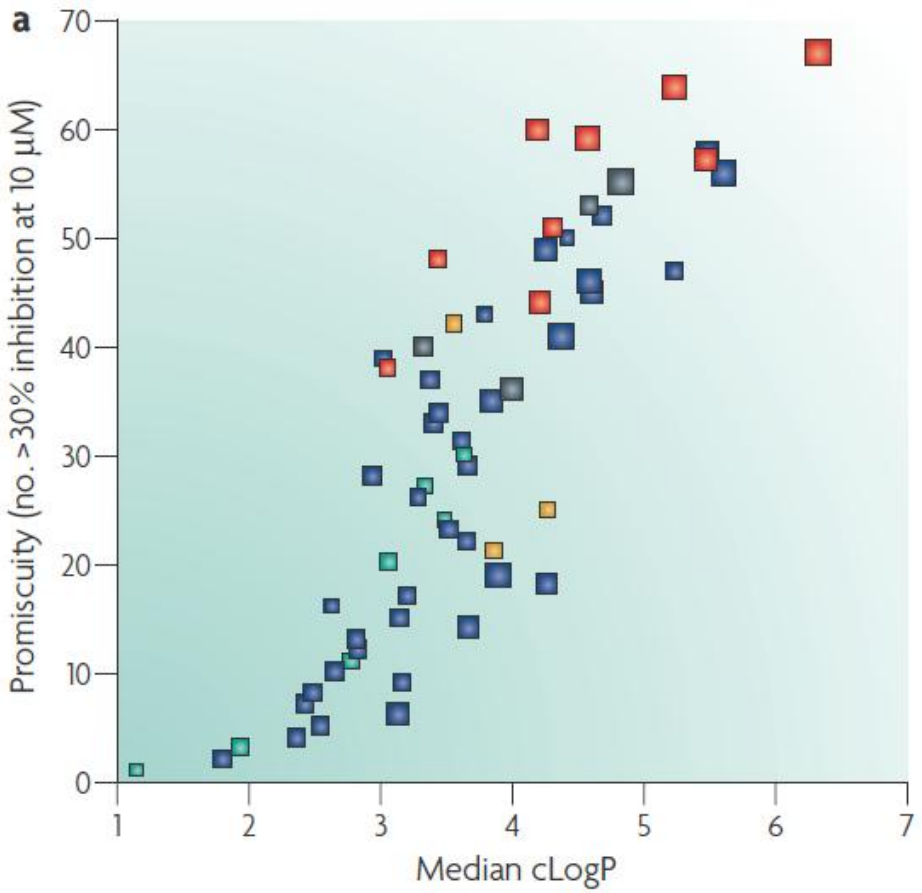
Substructure[a]	Assays Hit	# Cpds	Total Cpds	Enrichment[b]	Refs for hits	Refs for MOA[c]
Rhodanines  ene_rhod_A	0	60	235	227%	21-55	Reactivity.[6,7,36-40] Chelation.[22]
	1	39				
	2	32				
	3	26				
	4	21				
	5	41				
	6	16				

Reported MOAs eg ‘reactive’, ‘chelation’, ‘redox and protein reactive’ and ‘dyes’.

Baell JB, Holloway GA., J Med Chem. 2010, 53, 2719.

NOVARTIS

Avoid grease to get rid of sideactivities and toxicities



Leeson PD, Springthorpe B. Nat Rev Drug Discov. 2007, 6, 881-90

„Molecular obesity“

Hann, MM, Med. Chem. Commun. 2011, 2, 349

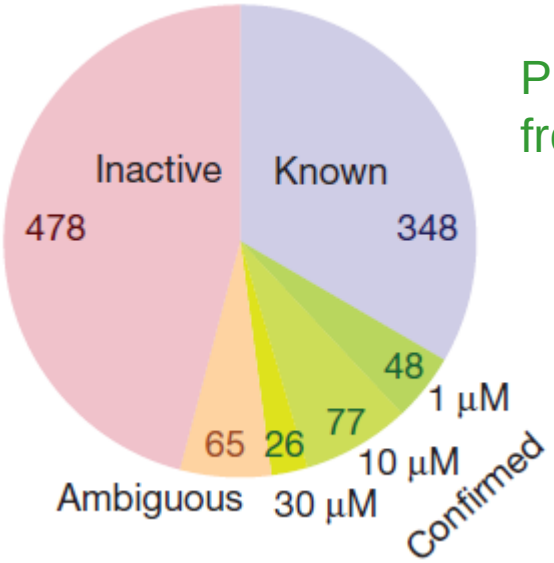
NOVARTIS

Avoid side effects directly

Large-scale prediction and testing of drug activity on side-effect targets

Eugen Lounkine[1]*, Michael J. Keiser[2,3]*, Steven Whitebread[1], Dmitri Mikhailov[1], Jacques Hamon[4], Jeremy L. Jenkins[1], Paul Lavan[4], Eckhard Weber[4], Allison K. Doak[3], Serge Côté[4], Brian K. Shoichet[3] & Laszlo Urban[1]

21 JUNE 2012 | VOL 486 | NATURE | 361



Prediction not confirmed

Prediction confirmed from literature

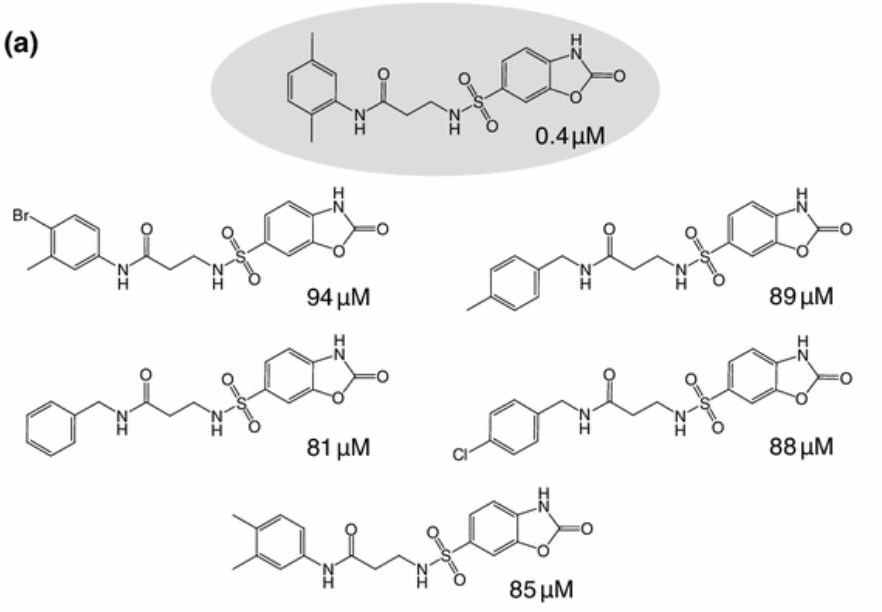
Prediction confirmed by experiment

- So far not sytematically applied for focused set design
- Critical for application is the accuracy of predictions

NOVARTIS

Chemical Diversity

- Similar molecules are more likely to have similar activities than non-similar molecules
- Testing many representative „good“ chemotypes preferrable over testing many analogues



... But it’s all about probabilities, Sometimes there is a single active cpd in a series

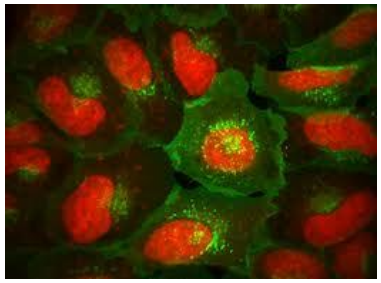
J Comput Aided Mol Des. 2013 Feb;27(2):115-24.
Activity cliffs in PubChem confirmatory bioassays taking inactive compounds into account.
Hu Y, Maggiora GM, Bajorath J.

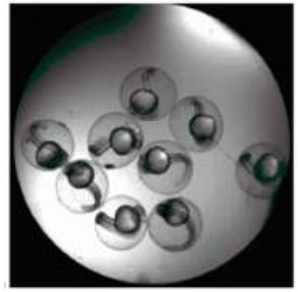
NOVARTIS

Overview

- Focused screening
- Where to focus
- Focus on biology
- Conclusions

Focused sets for phenotypic screening

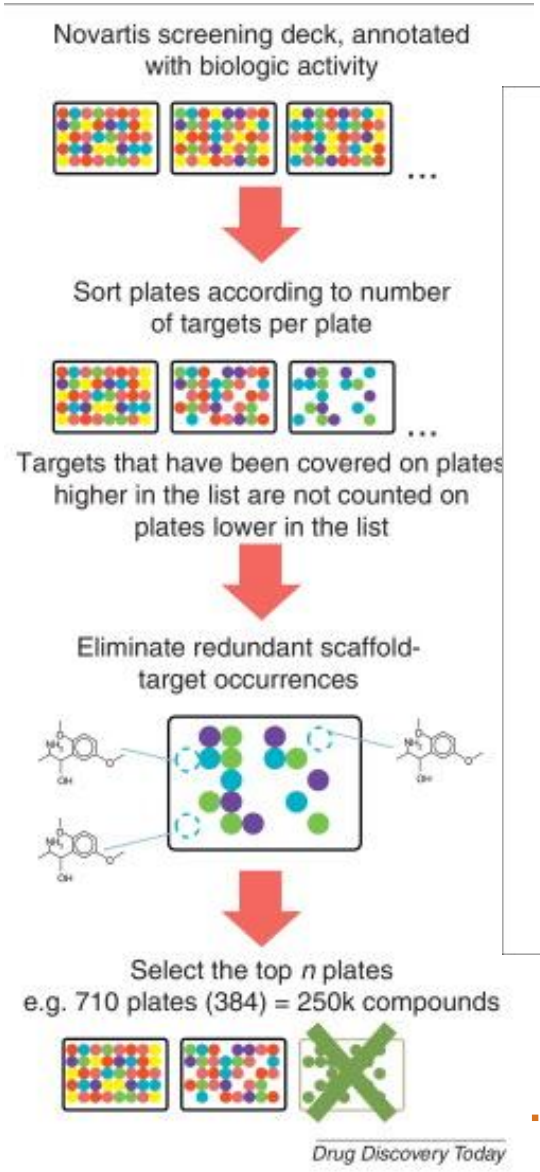




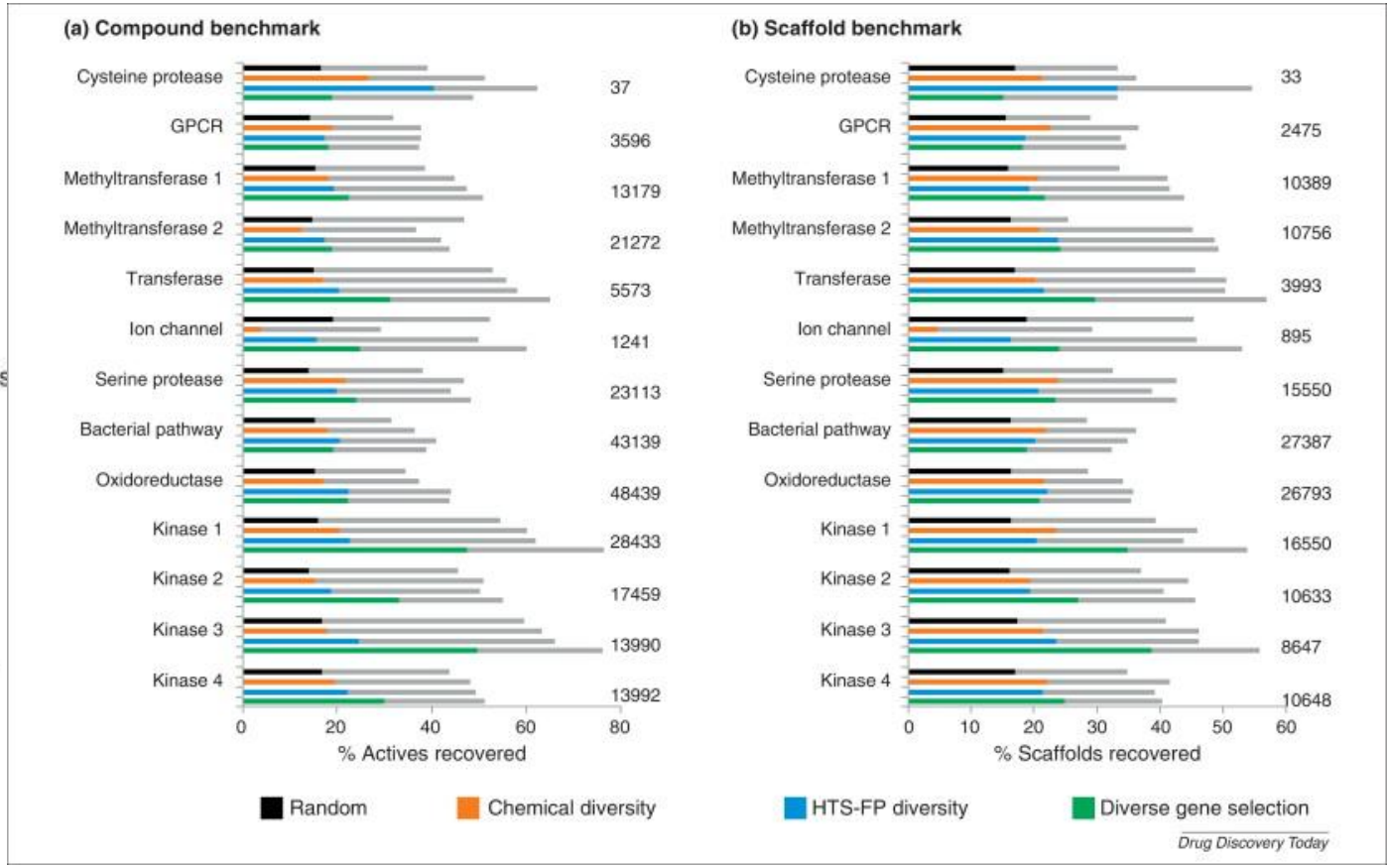
- Like for other focused screens, and in addition ...

- Cell permeability highly important
- Explore many biologies/ focused biologies with chemical probes
- Biological diversity – target many different targets
- Explore polypharmacologies
- Good annotation and understanding of compounds helpful
- «Chemical probe»-like compounds

Biodiversity for more diverse selections



Simulated 250k screens – recall of known hits



Petrone PM, et al
Drug Discov Today. 2013, 18, 674-80

… also better for targeted screens

NOVARTIS

Biology-focused sets

- Select compounds for relevant targets and pathways

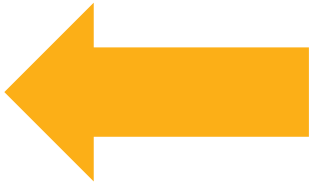
OMIM
Gene Ontology
Pathway databases
Gene expression DBs
...



Potential disease relevant genes



ChEMBL
PubChem
OpenPHACTS
Internal Screening DB
… DrugBank, Lopac

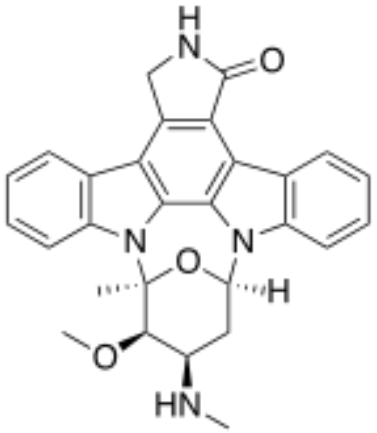


Compounds with biology of interest

NOVARTIS

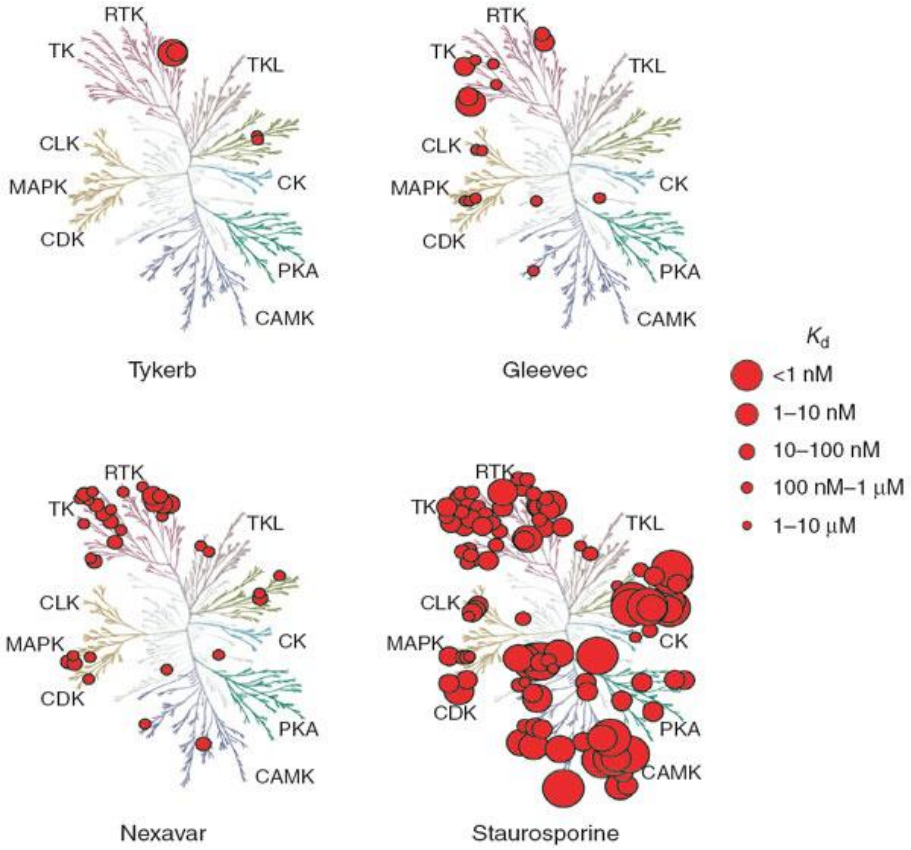
Criteria for selecting chemical probes

What we don't want!



„Staurosporine is the archetype of the 'probe gone bad',
with more than 8,000 publications"

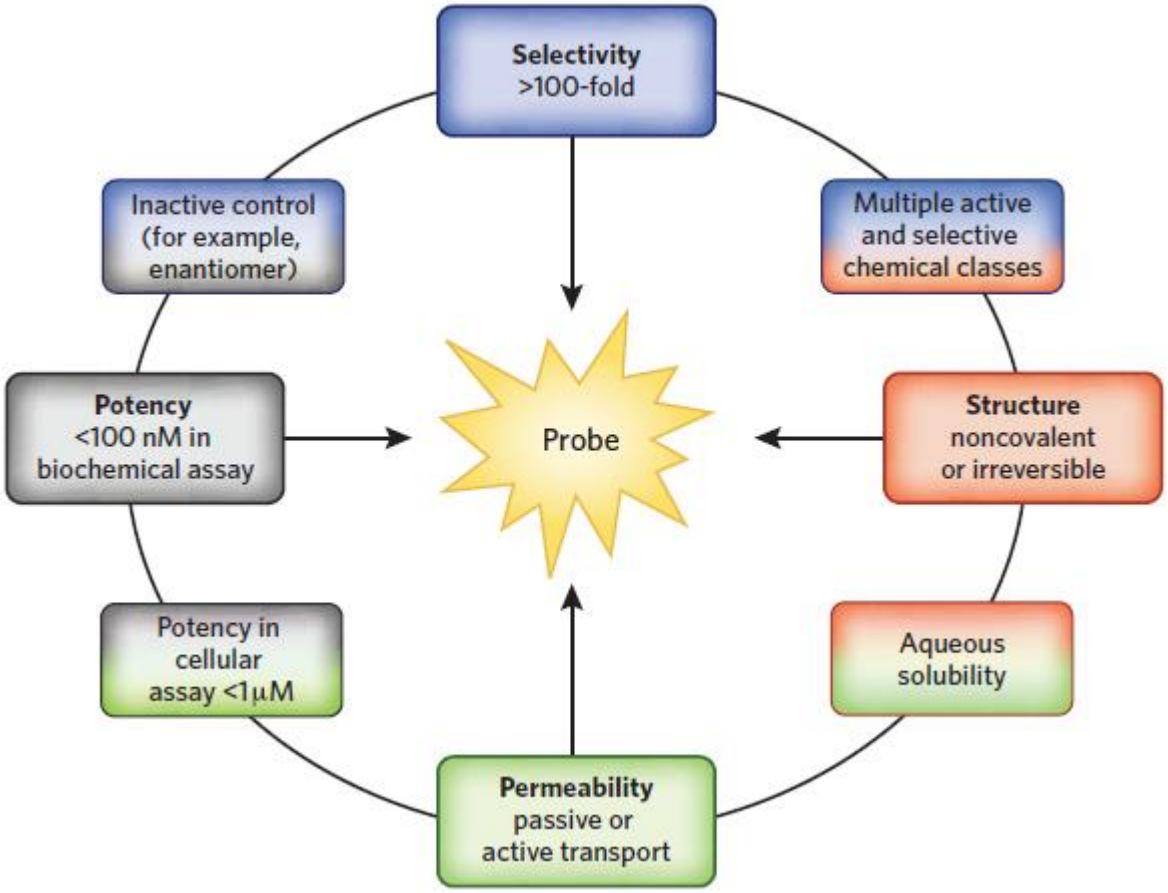
Nat Chem Biol. 2010 Mar;6(3):159-161.
The art of the chemical probe.
Frye SV.



Nat Chem Biol. 2006 Dec;2(12):689-700.
New approaches to molecular cancer therapeutics.
Collins I, Workman P.

NOVARTIS

Criteria for selecting chemical probes



Nat Chem Biol. 2013 Apr;9(4):195-9
Target validation using chemical probes.
Bunnage ME, Chekler EL, Jones LH.

Genuine on target activity necessary

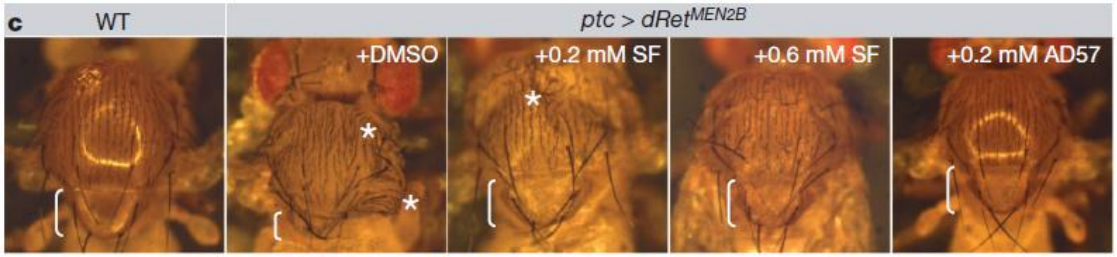
Drugs are NOT necessarily chemical probes

NOVARTIS

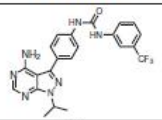
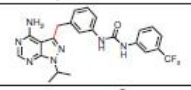
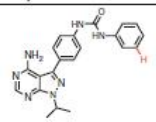
Explore polypharmacologies

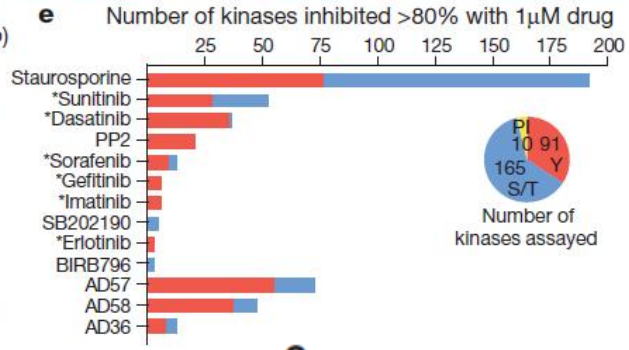
Chemical genetic discovery of targets and anti-targets for cancer polypharmacology

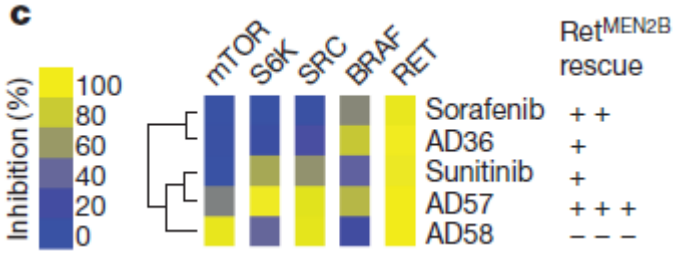
Arvin C. Dar[1]*, Tirtha K. Das[2]*, Kevan M. Shokat[1] & Ross L. Cagan[2]



d

Name	Structure	Ret^{MEN2B} rescue	IC_{50} Ret (nM; *in vitro*)
AD57		Strong (+++)	2 +/– 0.5
AD36		Weak (+)	7 +/– 3
AD58		Negative (– – –)	49 +/– 10





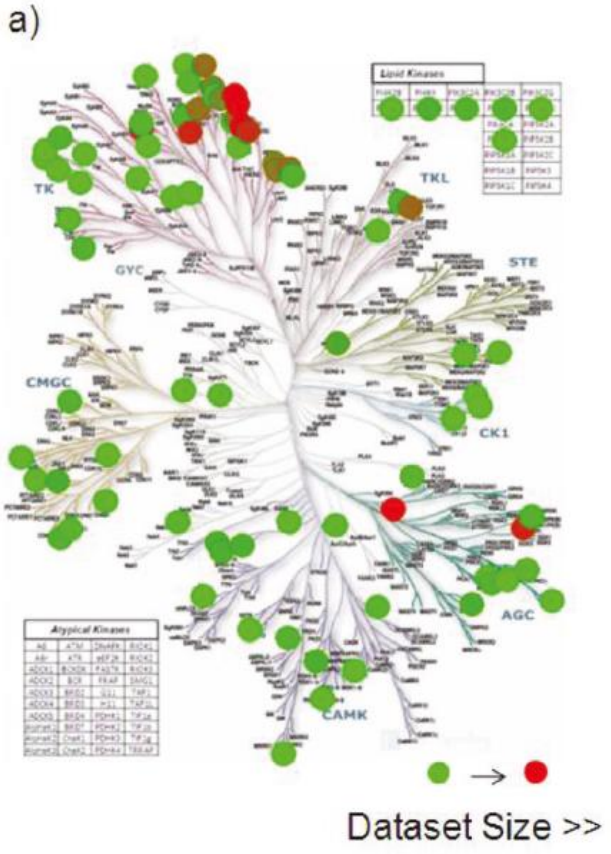
Requires many compounds with similar polypharmacologies

NOVARTIS

Finetuning the polypharmacology

Profile QSAR

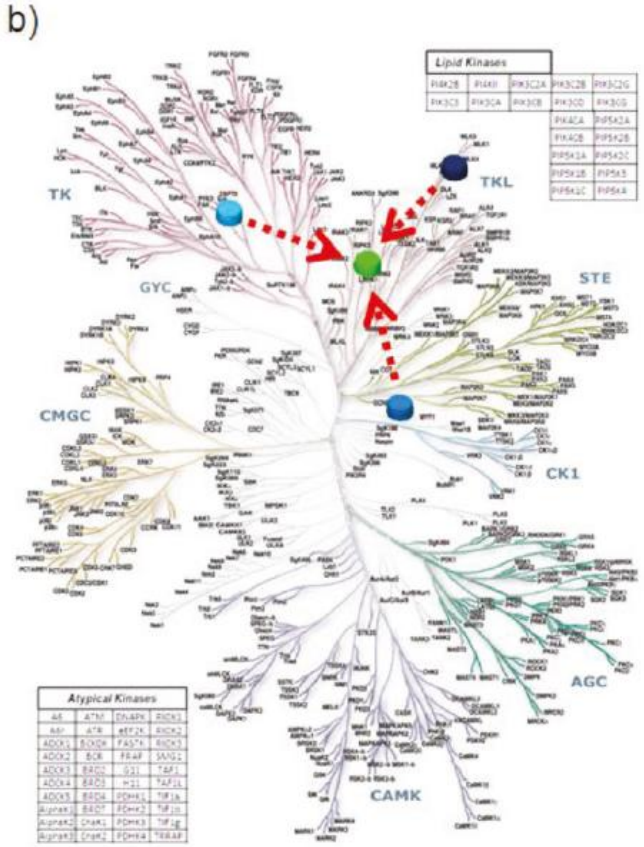
Models trained on known SAR data



Eric Martin, Prasenjit Mukherjee, David Sullivan, and Johanna J. Chem. Inf. Model. 2011, 51, 1942

Kinase Kernel Models

Interpolation between known kinases



Eric Martin and Prasenjit Mukherjee J. Chem. Inf. Model. 2012, 52, 156

NOVARTIS

Overview

- Focused screening
- Where to focus
- Focus on biology
- Conclusions

Conclusions

- **Only select compounds that would be followed up, if they were a hit**

- Focused set for phenotypic screens can be used for MOA identification / exploration

Backups

Assay compatible chemical space

Chemical and Biological Properties of Frequent Screening Hits.

Che J, King FJ, Zhou B, Zhou Y., J Chem Inf Model. 2012, 52, 913.

Table 1. Key MOAs of HTS Frequent Hits

MeSH	description	$\log_{10} P$	hit rate (%)
82004639	emetics	−14.0	83
82014344	trypanocidal agents	−11.2	31
82007476	ionophores	−15.0	28
82000903	antibiotics, antineoplastic	−27.2	27
82011500	protein synthesis inhibitors	−34.5	20
82000563	amebicides	−10.0	19
82002400	cathartics	−9.8	18
82002316	cardiotonic agents	−24.1	16
82003049	coccidiostats	−7.7	16
82000969	antinematodal agents	−9.4	15
82019384	nucleic acid synthesis inhibitors	−11.0	14
82000964	antimetabolites, antineoplastic	−12.7	14
82047428	protein kinase inhibitors	−5.7	14
82005456	fluorescent dyes	−9.0	13

Table 2. Key GO Groups and Pathways for Cellular Frequent Hits

description	$\log_{10} P$
cell cycle G1-S growth factor regulation	−44.7
signal transduction TGF-beta, GDF, and Activin signaling	−42.8
inflammation MIF signaling	−41.5
development VEGF signaling and activation	−39.1
CYPs-exo	−38.0
GO:0004674: protein serine/threonine kinase activity	−36.7
development EGFR signaling pathway	−35.0
development VEGF signaling via VEGFR2-generic cascades	−32.7
development hemopoiesis, erythropoietin pathway	−29.8
cell cycle G1-S interleukin regulation	−29.2
development A2A receptor signaling	−27.8
signal transduction Erk interactions: Inhibition of Erk	−26.7
CYPs C/EBP-mediated regulation	−26.2
immune response gastrin in inflammatory response	−26.2
signal transduction ERBB-family signaling	−26.1
inflammation IL-10 anti-inflammatory response	−25.7
development regulation of angiogenesis	−25.2

Many examples might still be interesting in some scenarios

NOVARTIS

NOVARTIS

NOVARTIS