Lezione 8

Andrea Ilari

Structure-based drug design

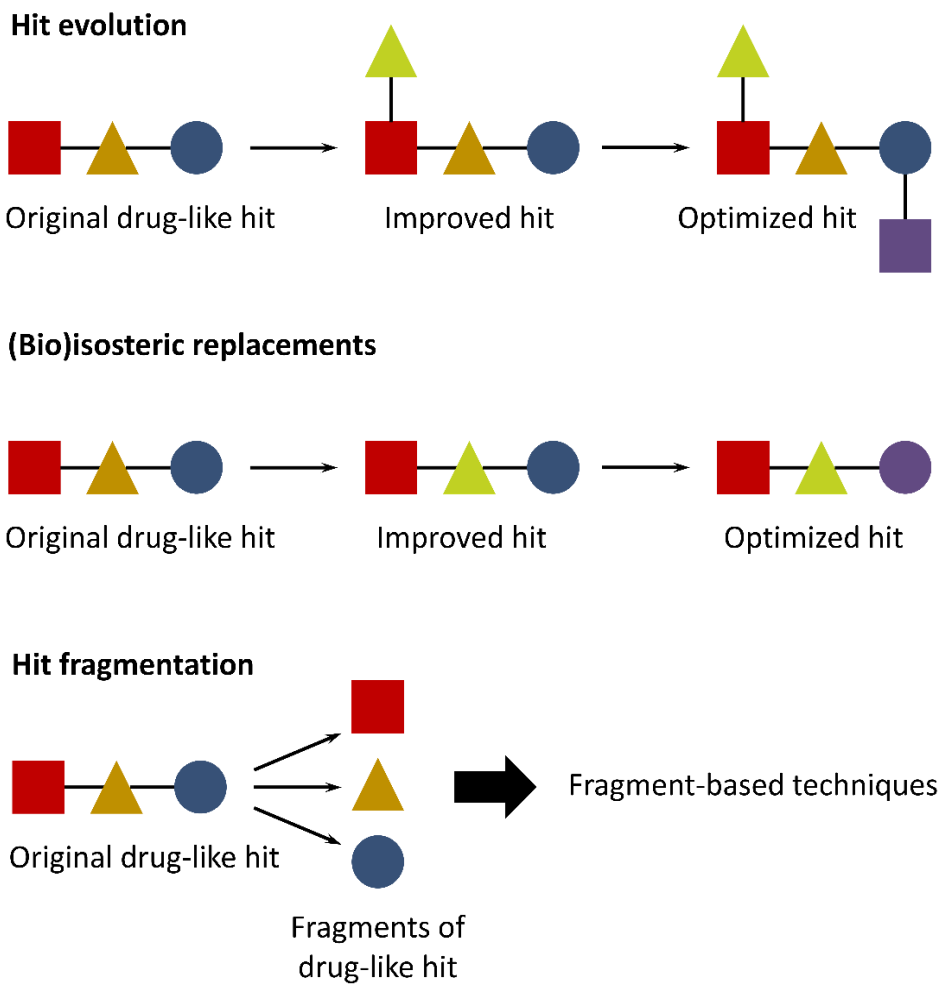
Structure-based drug Design

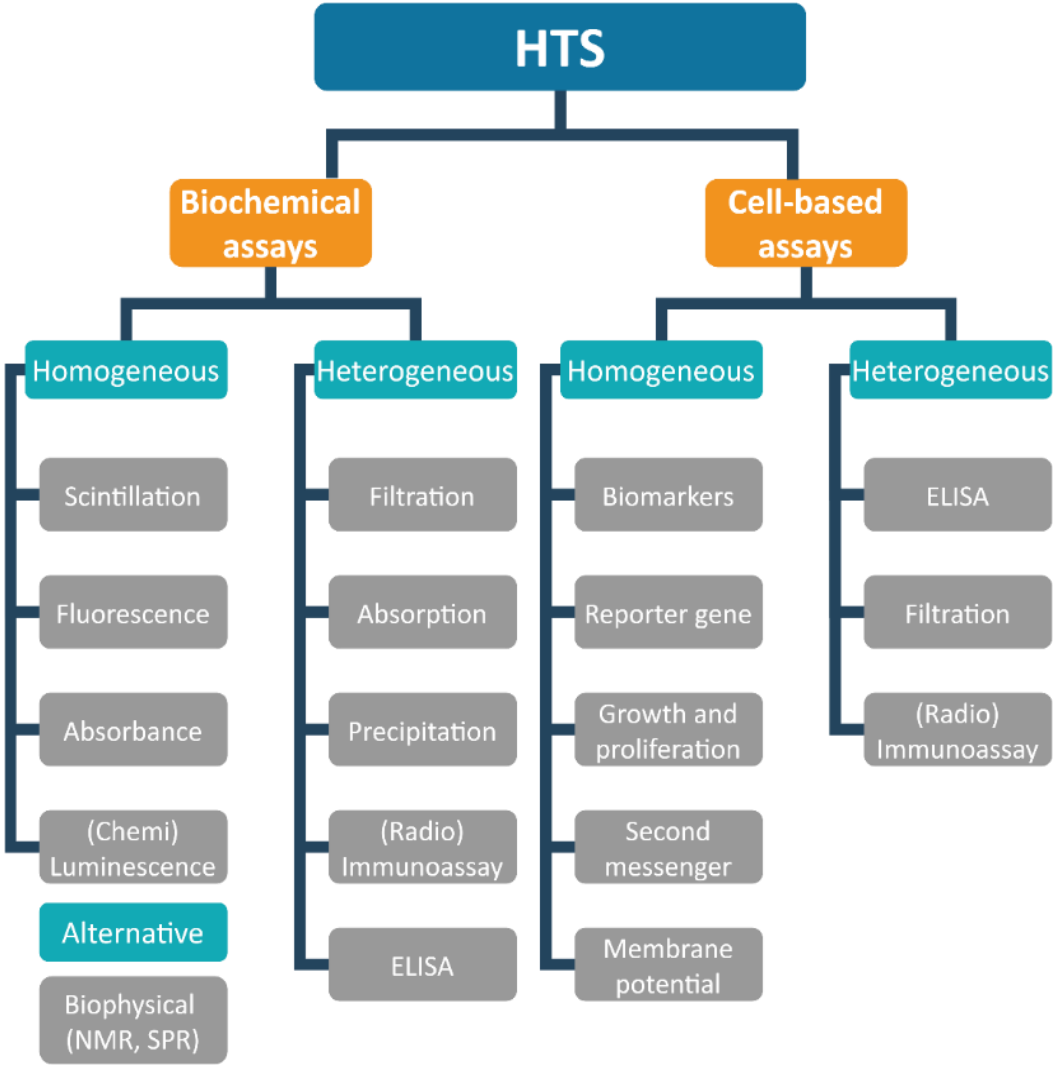
1. HighThroughput screening and lead optimization

2. Virtual screening- ligand binding- lead oprimization

3. Fragment-based drug sesign

Possibili strategie per la sintesi di lead compounds

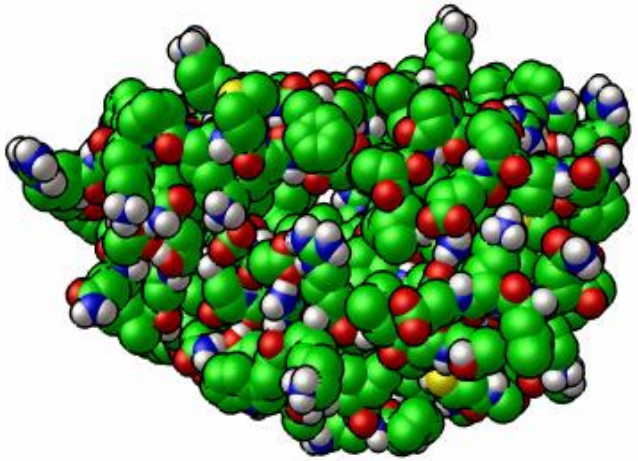



HTS
Biochemical assays
Cell-based assays
Homogeneous
Heterogeneous
Homogeneous
Heterogeneous
Scintillation
Fluorescence
Absorbance
(Chemi) Luminescence
Alternative
Biophysical (NMR, SPR)
Filtration
Absorption
Precipitation
(Radio) Immunoassay
ELISA
Biomarkers
Reporter gene
Growth and proliferation
Second messenger
Membrane potential
ELISA
Filtration
(Radio) Immunoassay

The Question

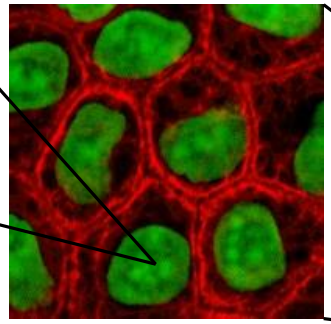
Identify a molecule able to bind to a target and provide a biological function

K_i (score)



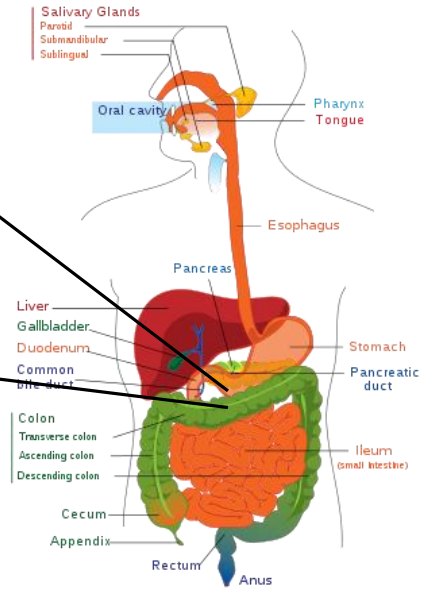
receptor

Unusual elements (Pt, Ru, U...)
Reactive chemical groups
Over/Under-functionalization
Partition coefficient (logP)



cell

Absorption
Distribution + **Tox**
Metabolism
Excretion



organism

What are we looking for

Hit low/medium target affinity

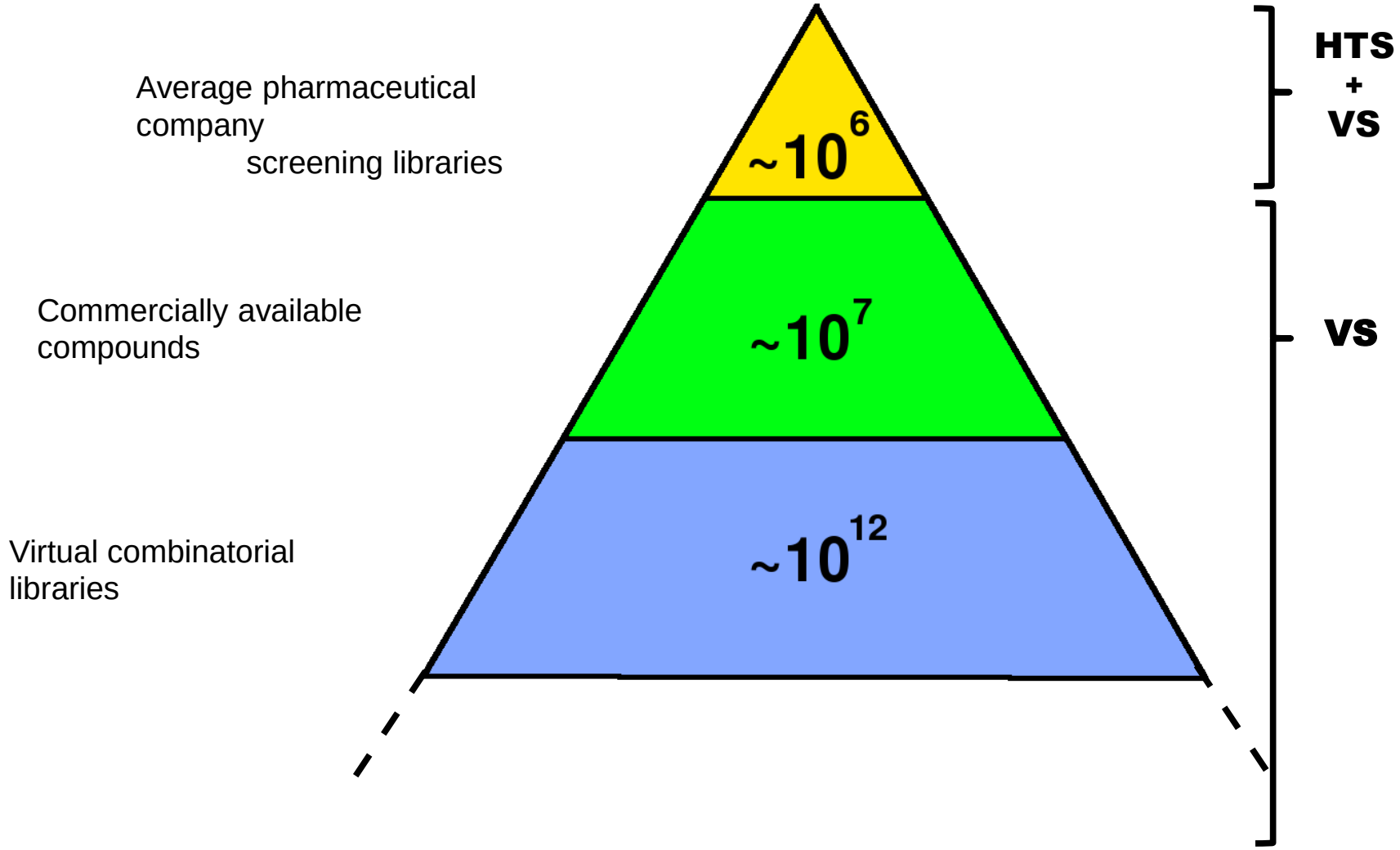
Lead sub-optimal target binding affinity

To be chosen for further development a **lead** compound should have the following properties:

- relatively simple chemical features (suitable for combinatorial/med-chem optimization, no/few chiral centers)
- well-established SAR (structure-activity-relation) series (similar compounds/chemical groups should present similar activity)
- good ADME properties
- novel structure (or favorable patent situation)

The Chemical Space

Where to look for



The Chemical Space

Filtering rules…

- Use filtering rules whenever possible
- Fragment molecules vs drug-like molecules (better coverage of chemical space)
- High-throughput screening efforts using Lipinski-filtered libraries led to few μM hits
- LR5 requires sampling a huge chemical space to be effective!
- Leads are not drugs
- *"Rules are for the obedience of the fools and the guidance of wise men"* (Harry Day, 1898-1977)

(Le regole sono per l'obbedienza degli stolti e la guida dei saggi)

Drugs 'Rule of Five' (Lipinski rule)

Hydrogen bond donors <= 5
Hydrogen bond acceptors <= 10
Molecular weight <= 500 Da.
$LogP_{w/o}$ < ~5

Fragments 'Rule of Three' (Congreve's rule)

Molecular weight <= 300 Da.
HB donor/Acceptors <= 3
C*log*P <= 3
HBD and HBA <= 3

log (P) = il logaritmo del coefficiente di ripartizione ottanolo/acqua, il quale esprime una misura della liposolubilità della molecola

Lipinski et al. **2001**, *Adv. Drug Deliv. Rev.,* 46
Congreve et al. **2003**, *Drug Discov. Today,* 8

The Chemical Space

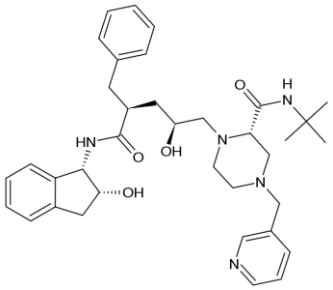
…and exceptions

The **nature and location of the target** must be considered for properties profile:

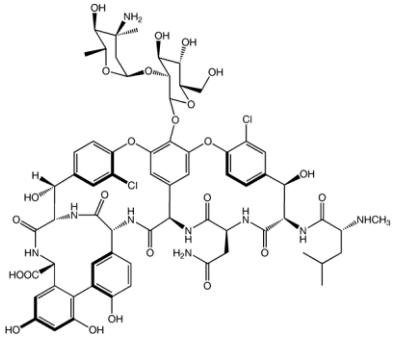
CNS molecule
(lipofilic blood-brain-barrier)

gastro-intestinal antibiotic
(highly soluble)

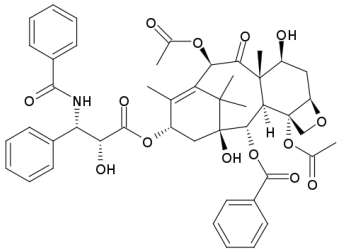
Natural compounds, pre-drugs, reactive molecules, anti-cancer agents…



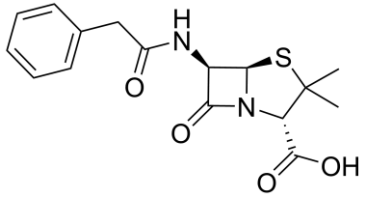
Indinavir



Vancomycin



Paclitaxel



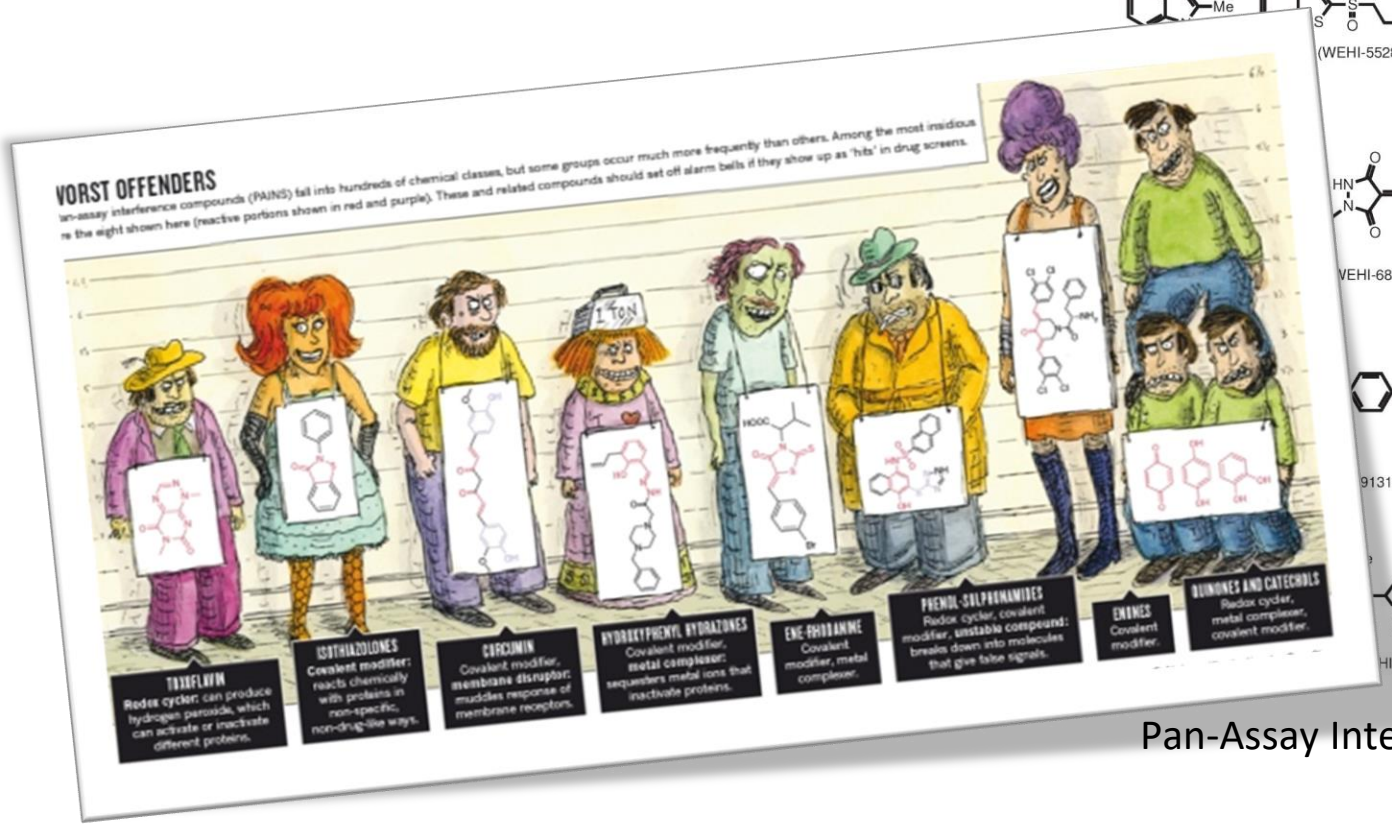
Penicillin

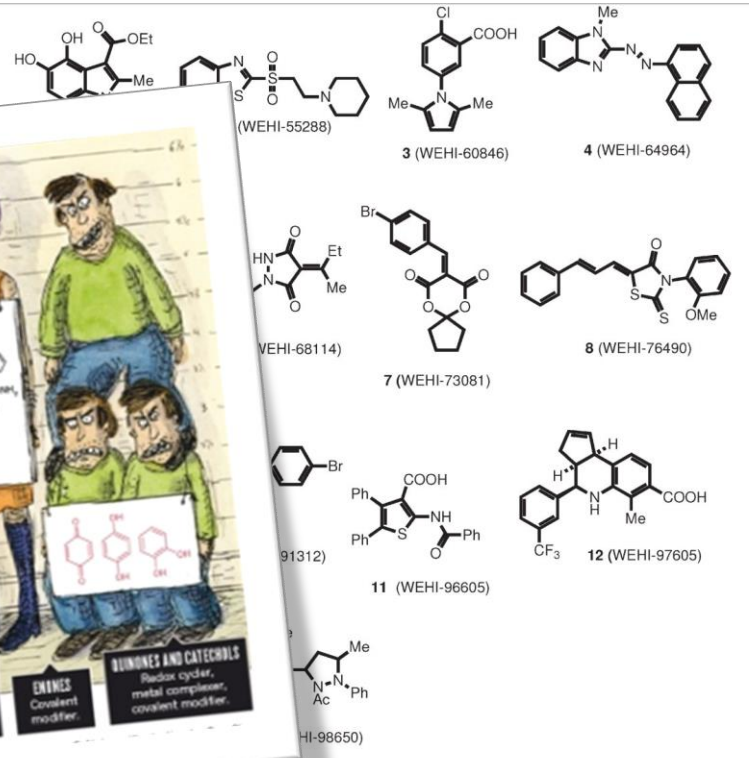
Scoring and assay issues

False negative & false positive

Pan-Assay Interference Compounds: sono composti che danno risultati positivi negli HTS ma che interagiscono con molti target diversi
E quindi non sono buoni lead compounds

FALSE POSITIVE





Pan-Assay Interference Compounds (PAINS)

Baell, J. B. et al. **2010**, J.Med.Chem, 53

Osservabile Fisico

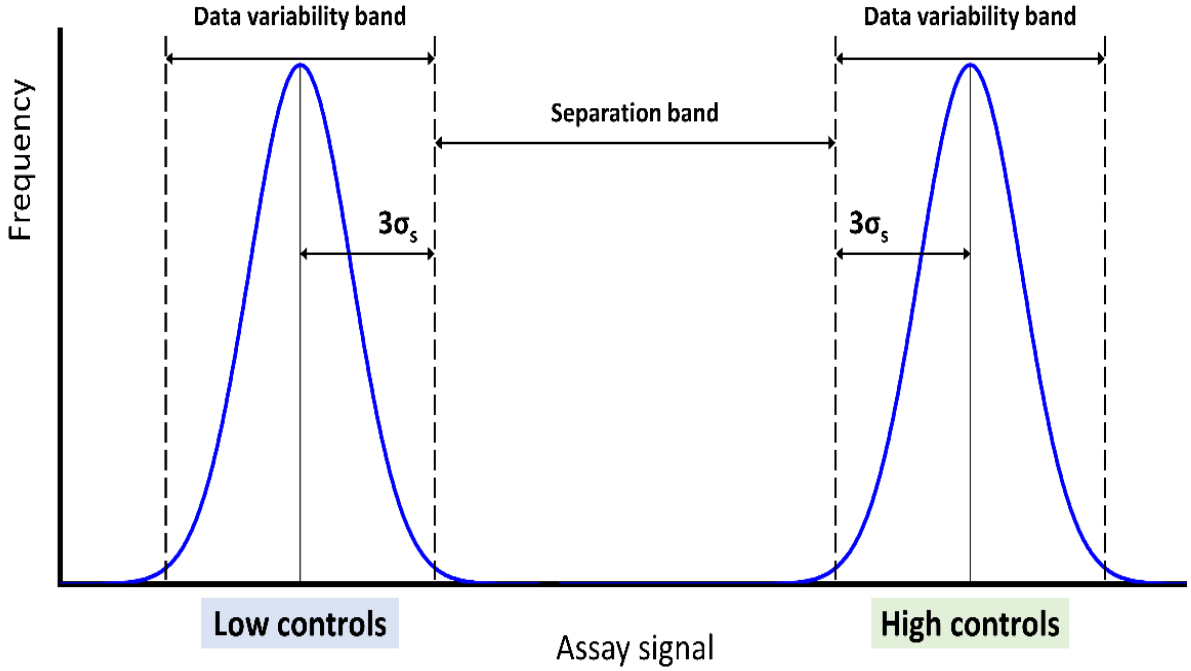
- Cromoforo- fluorescenza e assorbimento
- Indice di Rifrazione – Surface Plasmon Resonance
- Chemical shift - NMR

Cosa si Misura

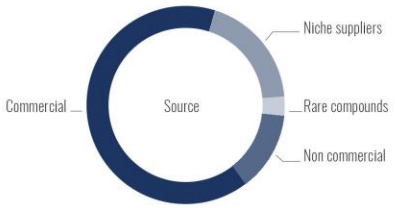
- IC_{50}= concentrazione di un inibitore enzimatico (farmaco, tossina o veleno, ecc.) necessaria per inibire il 50% del bersaglio in esame (diminuire la pendenza dI/dt del 50%).
- Ki = costante di inibizione

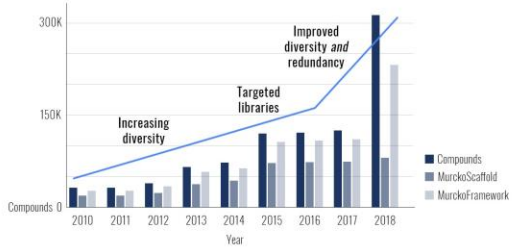
Ottimizzazione dell'HTS

$$Z' = 1 - \frac{\left(3(SD\ of\ High\ Control) + 3(SD\ of\ Low\ Control)\right)}{|Mean\ of\ High\ Control - Mean\ of\ Low\ Control|}$$



Compounds collections





> 320,000 compounds available for HTS

> 98% free of PAINS / undesirable compounds

Commercial and exclusive compounds

Curated collection

High diversity

HTS

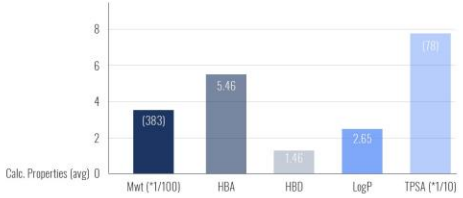
Balanced redundancy

Focused subsets

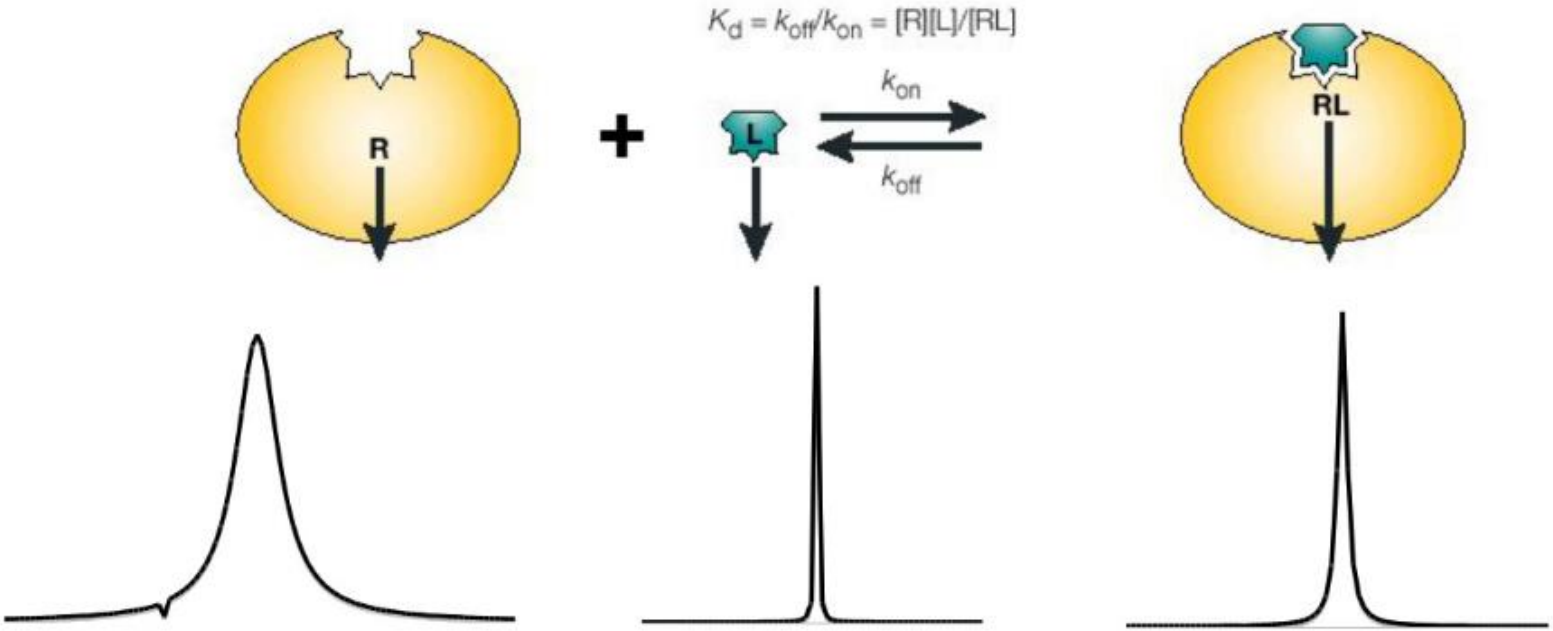
Attractive properties

- PTM and kinase subsets
- Safe-In-Man subset
- RNA-binding subset
- Cancer cpd subset
- Natural product subset
- FBDD collection

Fully automated, acoustic compound transfer



Basics of Ligand-Observed NMR Binding

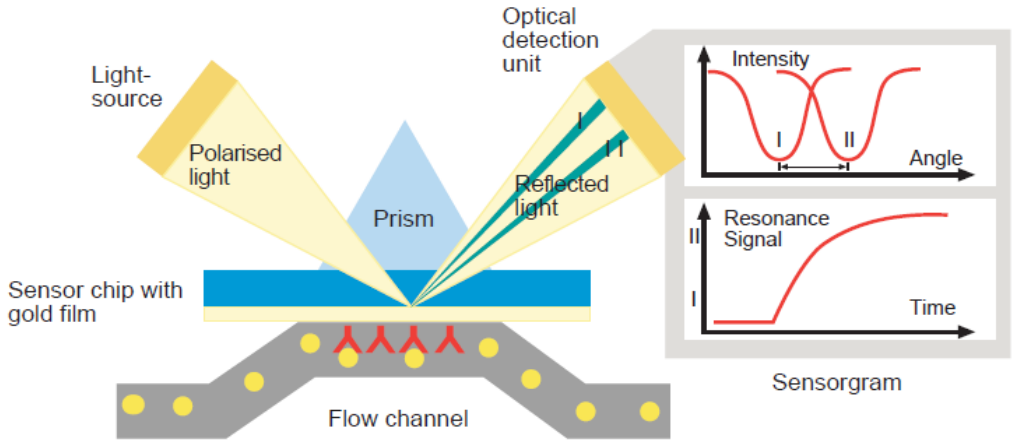


Slow Tumbling
Many Interactions
Slow Diffusion
Broad Peaks

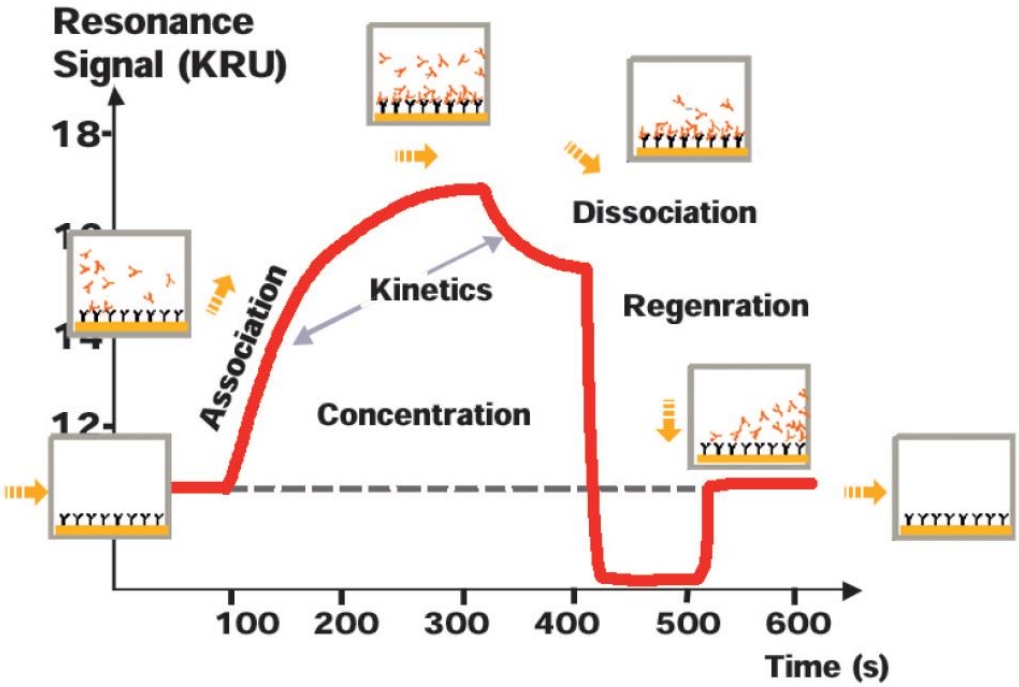
Fast Tumbling
Few Interactions
Fast Diffusion
Sharp Peaks

Surface Plasmon Resonance

L'SPR misura I cambiamenti nella massa dello strato vicino alla superficie del sensor chip attraverso la misura dell'indice di rifrazione.



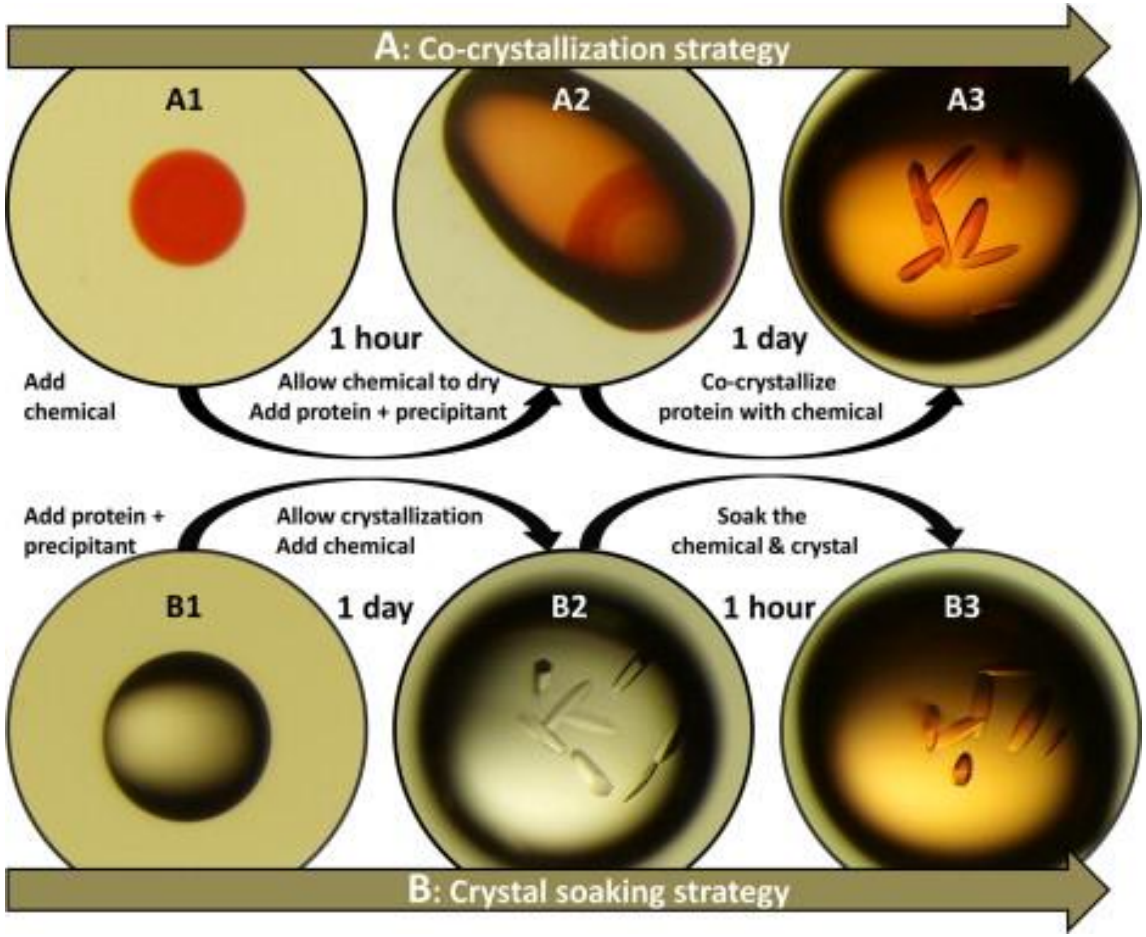
Quando la molecola si lega ad un taget immobilizzato la massa cresce e con essa l'indice di rifrazione. Quando viceversa la massa decresce when they dissociate the mass falls. This simple principle forms the basis of the sensorgram – a continuous, real-time monitoring of the association and dissociation of the interacting molecules.
The sensorgram provides quantitative information in real-time on specificity of binding, active concentration of molecule in a sample, kinetics and affinity.



Lead optimization attraverso la cristallografia

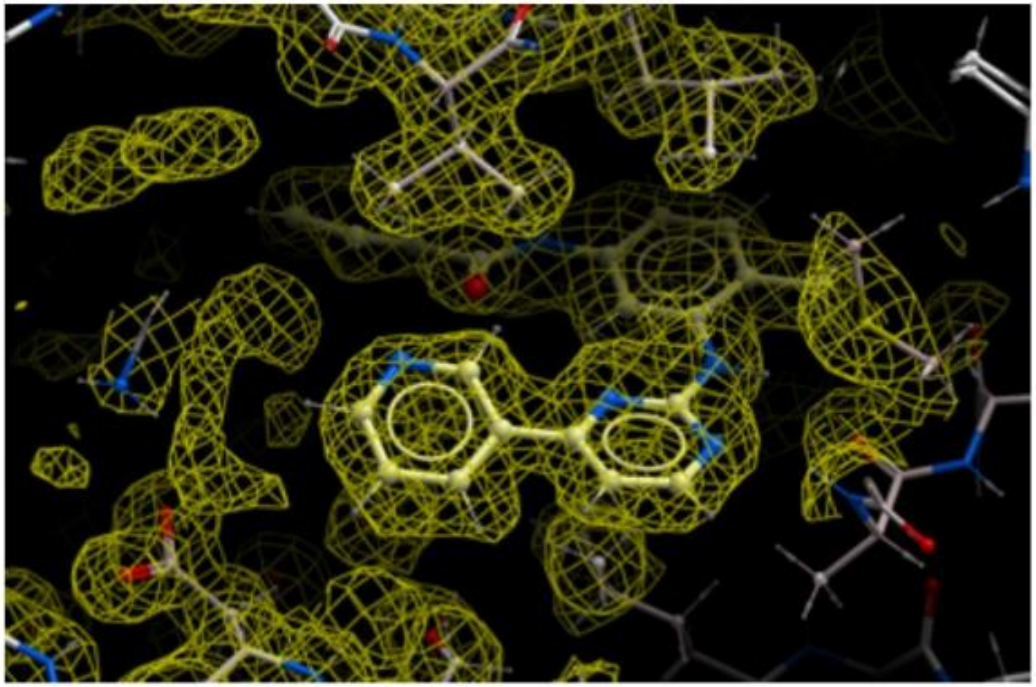
Metodi per ottenere i cristalli per risolvere la struttura del complesso dell'Hit trovato Attraverso l'HTS.

1. Soaking
2. Cocritallizzazione

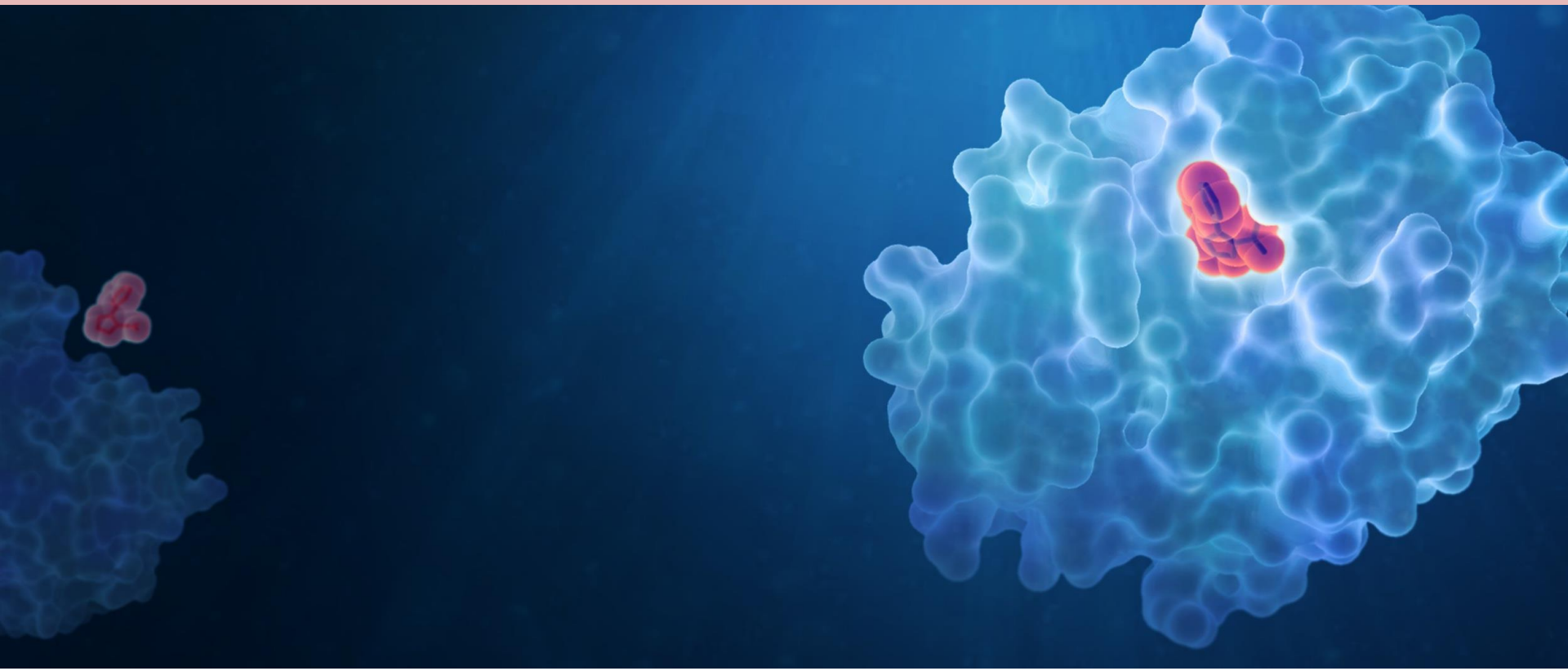


X-Ray Crystallography

- Three-dimensional structural information is essential
  - Enables optimization of compounds
  - NMR can give binding site, but crystallography positions ligands
- Typically preference for a 'soakable' system
  - Co-crystallography does work



Docking & Virtual Screening



Introduction
to Virtual Screenings

Stefano Forli, PhD
Assistant Professor of ISCB
Molecular Graphics Laboratory
Dept. Integrated Structural and
Computational Biology
The Scripps Research Institute

Molecular Graphics Lab

AutoDock

The Scripps Research Institute®

Screening & Virtual Screenings

virtual
High-Throughput Screening

Cheap (saves both money and time)

Can be easily automated

Dramatic reduction of compounds to test and false negative

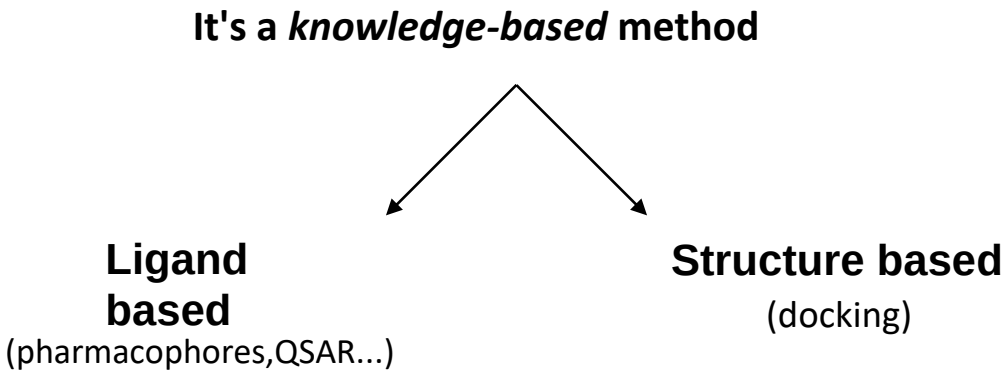
Played crucial role in development of 20 drugs (more than HTS!)

Irwin J., et al. **2016**, *J.Med.Chem, 59*

Virtual Screening

Definition

"Search for compounds with a defined biological activity using a computational model"

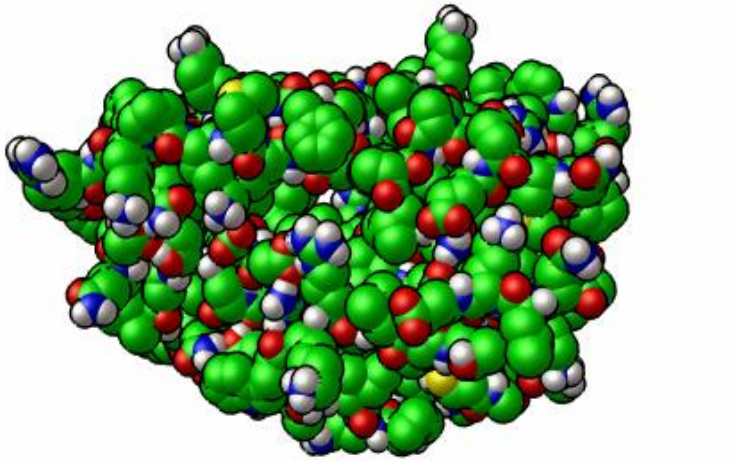


Horvat, D., **1997**, *J.Med.Chem,* 40

The Question

Identify a molecule able to bind to a target and provide a biological function

K_i (score)

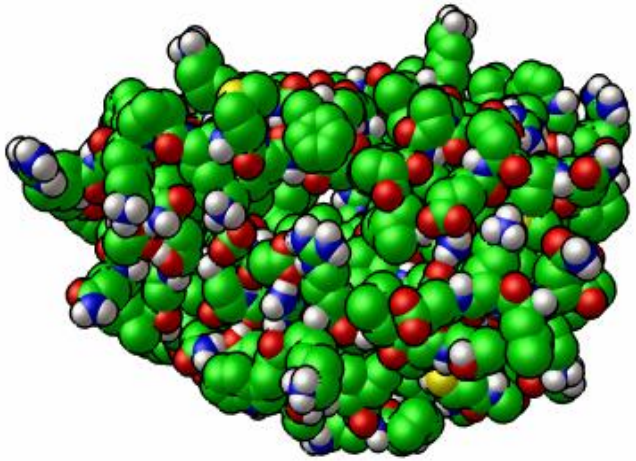


receptor

The Question

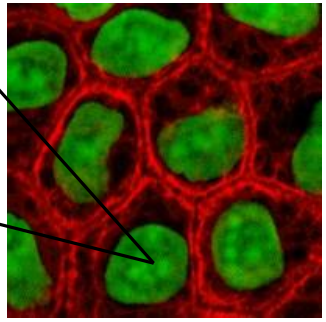
Identify a molecule able to bind to a target and provide a biological function

K_i (score)



receptor

Unusual elements (Pt, Ru, U...)
Reactive chemical groups
Over/Under-functionalization
Partition coefficient (logP)



cell

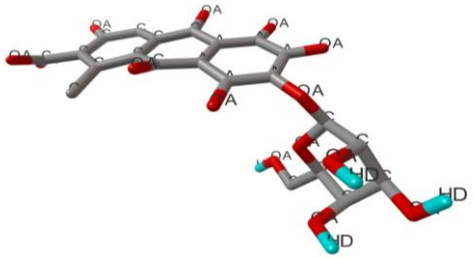
Ligand libraries

Preparation

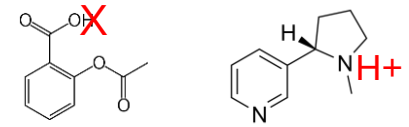
Prepare your molecules…

- Generate accurate geometries (bond lengths and angles!)
- Generate tautomer and protomer models
- Partial charges (->depending on scoring function!)

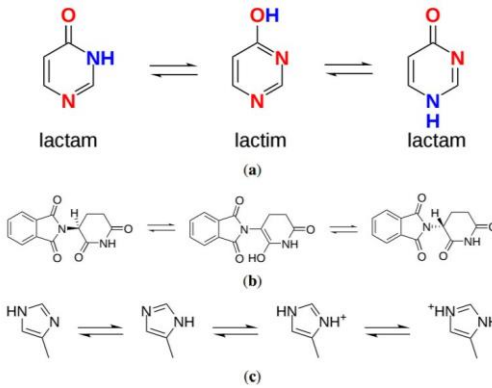
ACCURATE GEOMETRY



TAUTOMERS AND PROTOMERS



Aspirin **Nicotine**



Ligand libraries

Preparation

Prepare your molecules…

Manually

Avogadro [OPEN SOURCE]
http://avogadro.openmolecules.net/wiki/Main_Page

Automatically

Corina [COMMERCIAL, DEMO AVAILABLE]
http://www.molecular-networks.com/online_demos/corina_demo.html

ProDRG2 server [FREE BUT LIMITED]
http://davapc1.bioch.dundee.ac.uk/prodrg/index.html

ChemAxon Marvin [FREE FOR ACADEMICS]
http://www.chemaxon.com/product/marvin_land.html

OpenBabel [OPEN SOURCE]
http://openbabel.org/wiki/Main_Page

Ligand libraries

Database & Repositories

PubChem

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

Search by structures, names, synonyms or keywords
Biological data available for each compound
2D/3D structures

Z I N C

http://zinc.docking.org/

A free database of commercially available compounds for virtual screening

109 commercial compound suppliers+FDA approved drugs,metabolites...

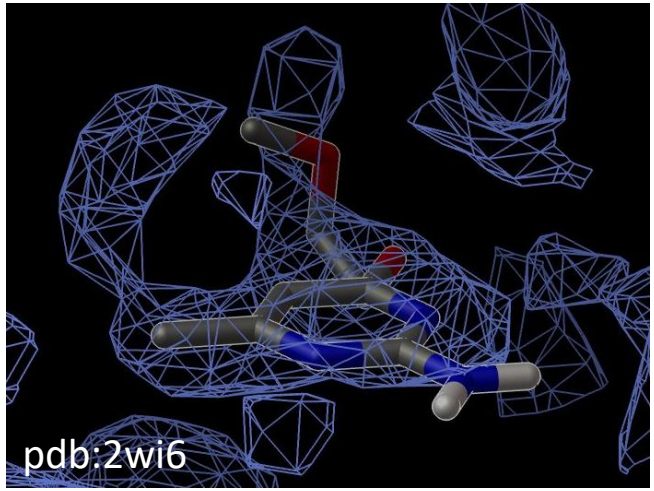
$30x10^6$ compounds (non-unique)
- different tautomers/protonation states
- reliable 3D structures
- pre-filtered for not allowed chemical groups and too complex structures

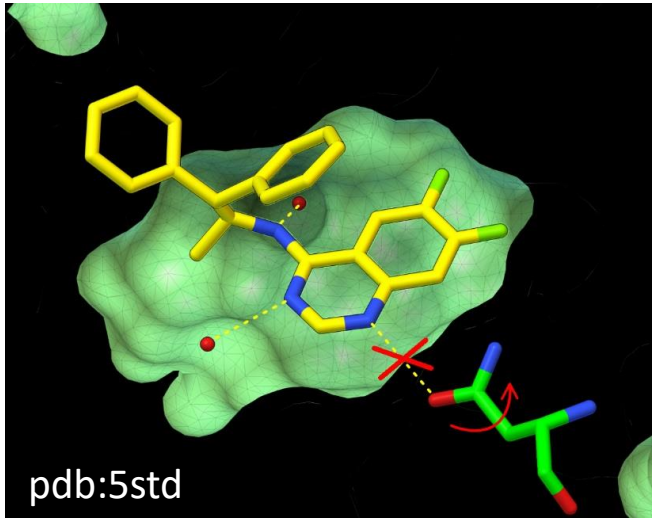
Irwin, J. et al. **2005**, *J.Chem.Inf.Model.* 45

Receptor

"Never trust a structure"

- **Inspect R-factor and resolution**
  - prefer lower values
- **Inspect B-factor, occupancy and alternate locations**
  - hints on dynamic behavior of side chains
- **Inspect density maps whenever you can**
  - undefined/non-interpreted densities
  - look for errors
- **Fix missing atom/residue/loops**
  - …if necessary
- **Inspect and correct special features**
  - non-natural AA
  - cofactors
  - salts and other additives
  - waters…?
- **Consider multiple states**
  - *apo*/*holo*
  - active/inactive
  - …





Virtual Screening

Summary

Prepare both target and ligand libraries with care

- Filter unusual elements (or AutoDock atom types)
- Reliable 3D geometries
- Protonation states and tautomers

Reduce the space of your search

- diversity sets
- generic filtering (reactive, insoluble, etc..)
- target specific filtering (lipophilic binding site vs. hydrophobic binding site)

Use all available information to select results

- mutagenesis, SAR...

Try to sample different conformations states of the protein

- reduce false negative
- ensemble dockings (e.g. relaxed complex scheme)

Use reference compounds whenever available

- Useful for comparing results with ligands with known activity

Be ready to accept the inevitable

- Not all targets are druggable

Scoring and assay issues

False negative & false positive

FALSE NEGATIVE

“You are not pregnant!”

FALSE POSITIVE

“You are pregnant!”

Ellis, P. “The essential guide to Effect Sizes: Statistical Power, Meta-Analysis and the Interpretation of Research Results”, **2010**, *Cambridge University Press*

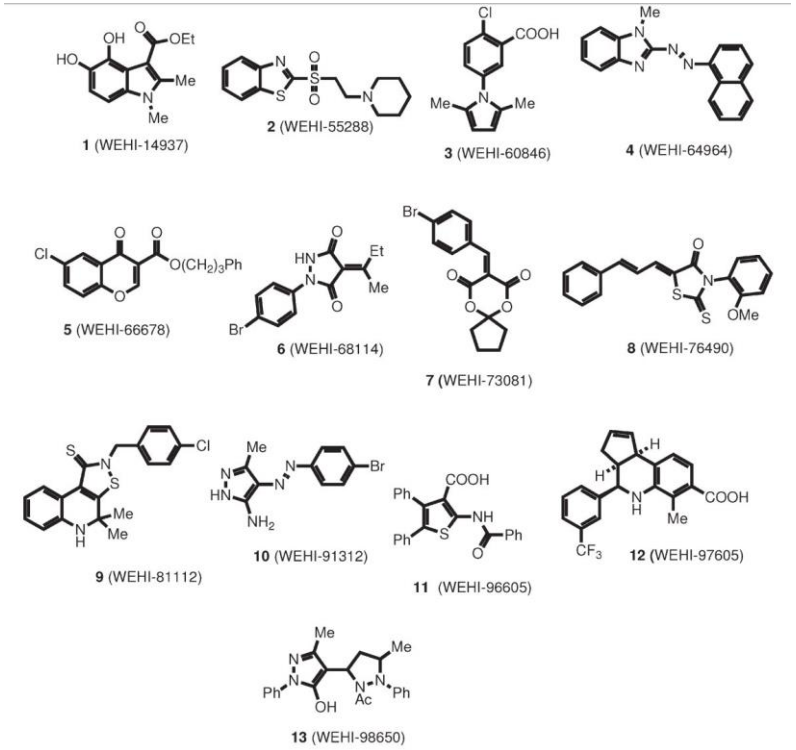
Scoring and assay issues

False negative & false positive

FALSE NEGATIVE

- Scoring function/search limitations
- Waters, protonation, …
- *Apo* vs. *holo*
- Target flexibility
- Crystallography artifacts

FALSE POSITIVE



Pan-Assay Interference Compounds (PAINS)

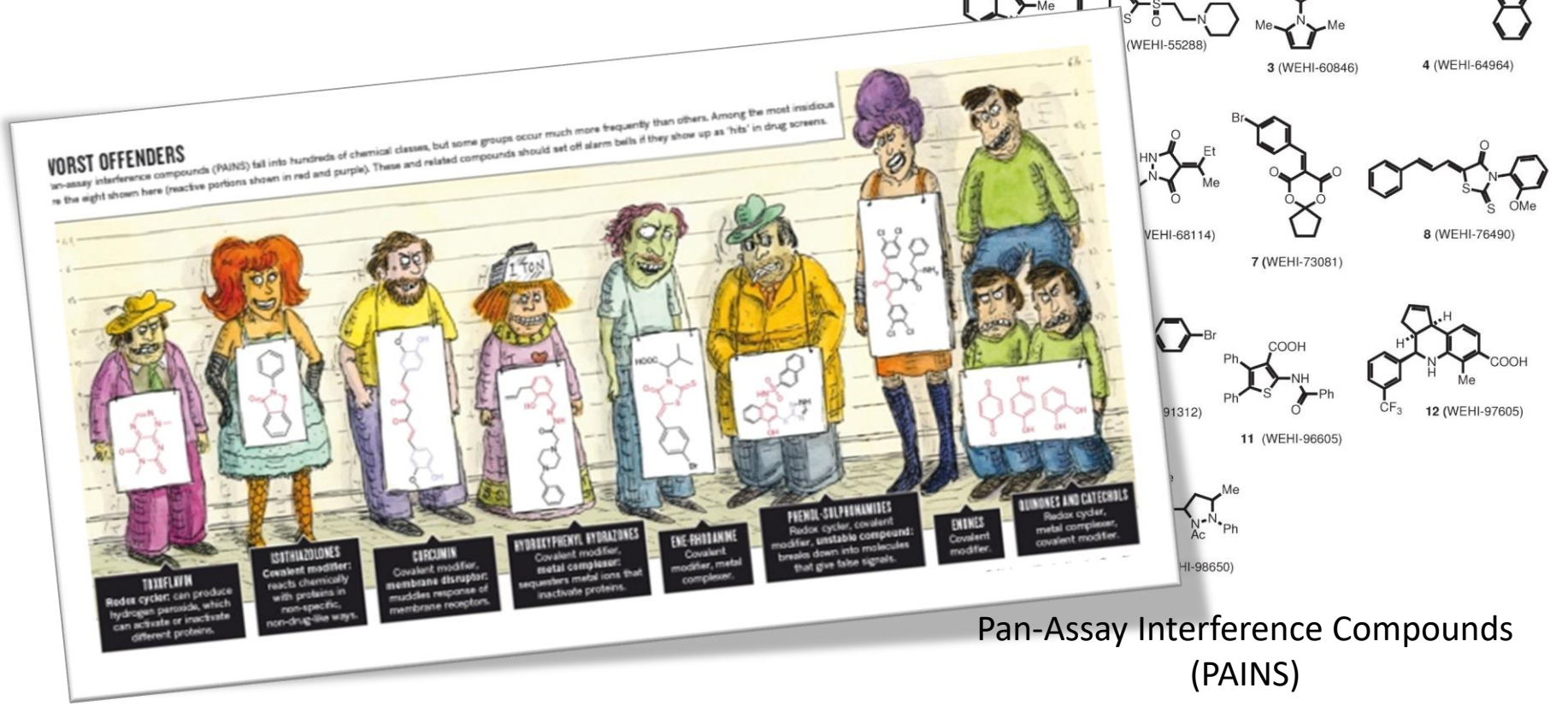
Baell, J. B. et al. **2010**, J.Med.Chem, 53

Scoring and assay issues

False negative & false positive

Pan-Assay Interference Compounds: sono composti che danno risultati positive negli HTS ma che interagiscono con molti target diversi
E quindi non sono buoni lead compounds

FALSE POSITIVE



Pan-Assay Interference Compounds (PAINS)

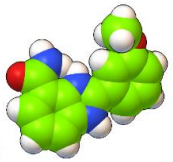
Baell, J. B. et al. **2010**, J.Med.Chem, 53

Virtual Screening

Recipe

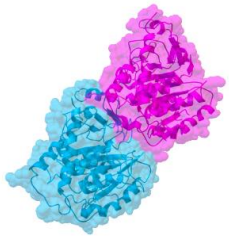
1. Ligand

3D geometry
protomers/tautomers
charges
physico-chemical properties



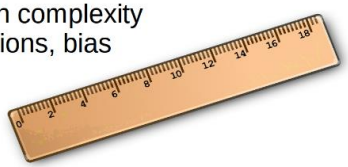
2. Receptor

Structure quality
Structure clean-up
Ions, co-factors, waters
conformation
binding site selection



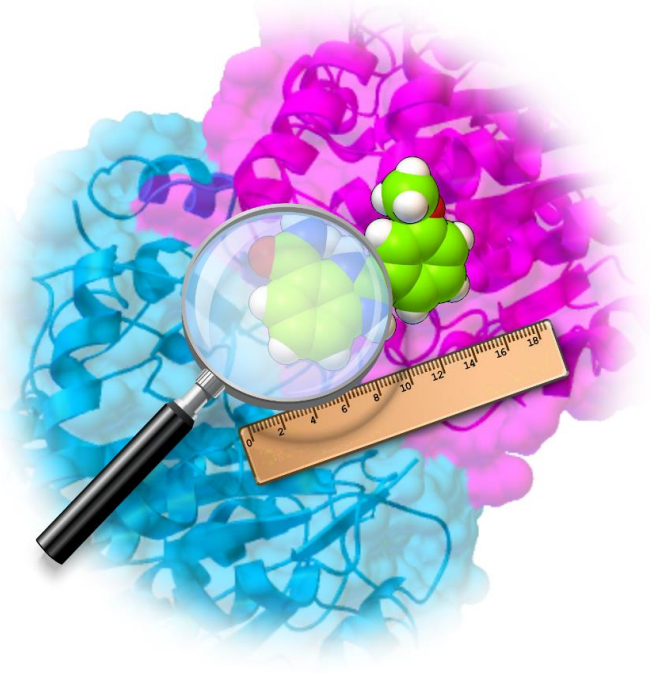
3. Scoring function & search method

search complexity
limitations, bias



4. Analysis

hit selection and validation



IBM World Community Grid

distributed grid project for scientific research

runs on volunteers computers & cell phones

free computer power

AutoDock4 and AutoDock Vina supported

FightAIDS@Home project generated *20 billion* ligand-target dockings

fightAIDS@home powered by AUTODOCK

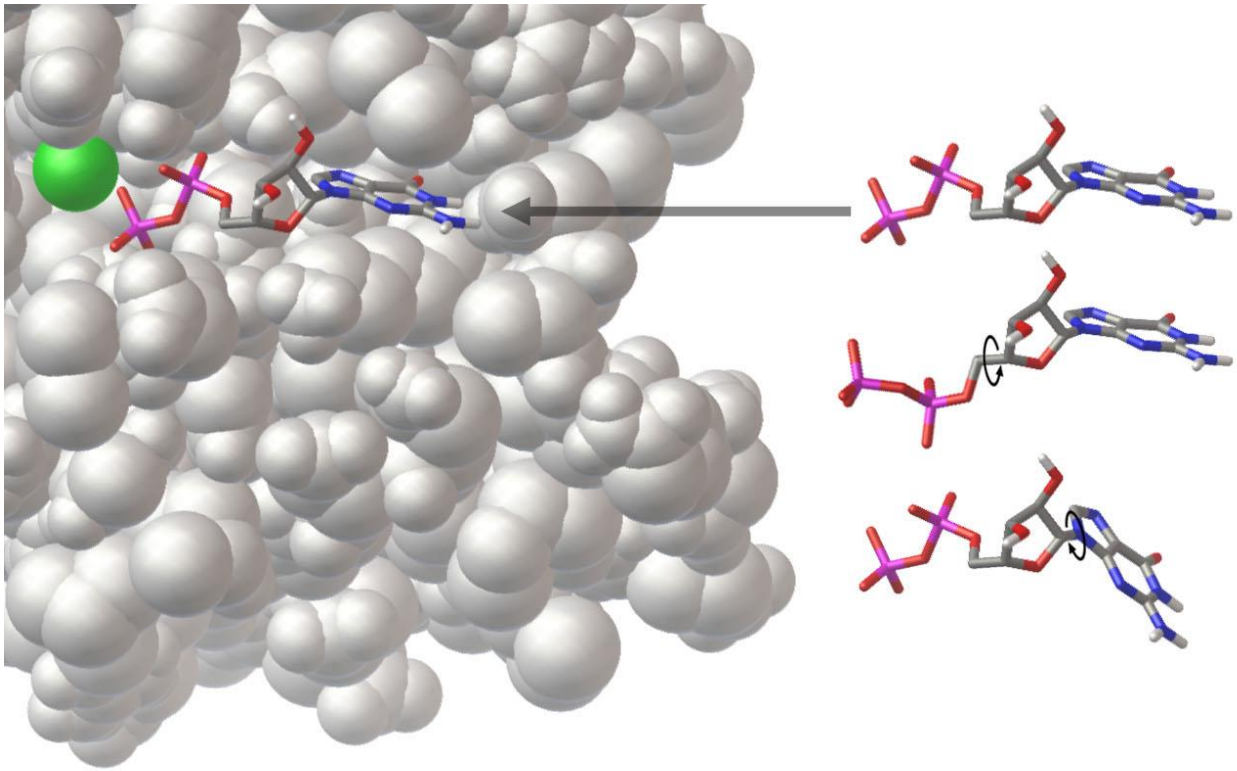
The Olson Laboratory

The Scripps Research Institute

Using AutoDock 4 and AutoDock Vina with AutoDockTools

Ruth Huey
Stefano Forli
Michel Sanner
David S. Goodsell

What is Docking?



Given the 3D structures of two molecules, determine the best binding modes.

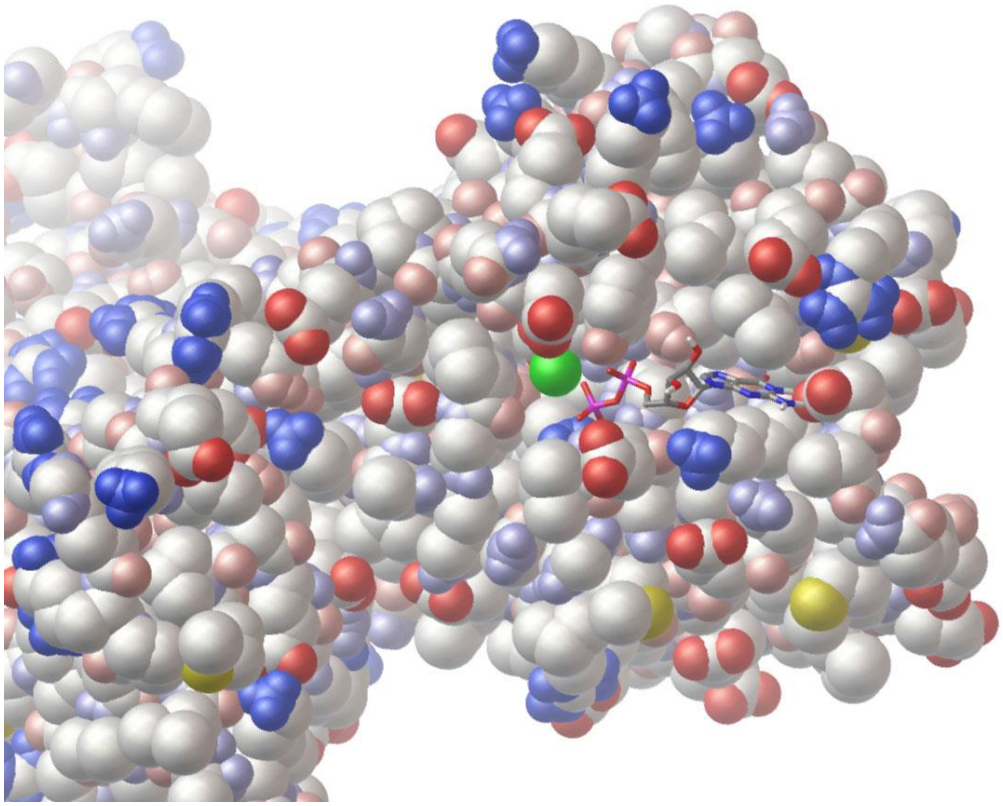
Key aspects of docking…

- Scoring Functions
  - *Predicting the energy of a particular pose*
  - *Often a trade-off between speed and accuracy*
- Search Methods
  - *Finding an optimal pose*
  - *Which search method should I use?*
- Dimensionality
  - *Can we trust the answer?*

AutoDock History

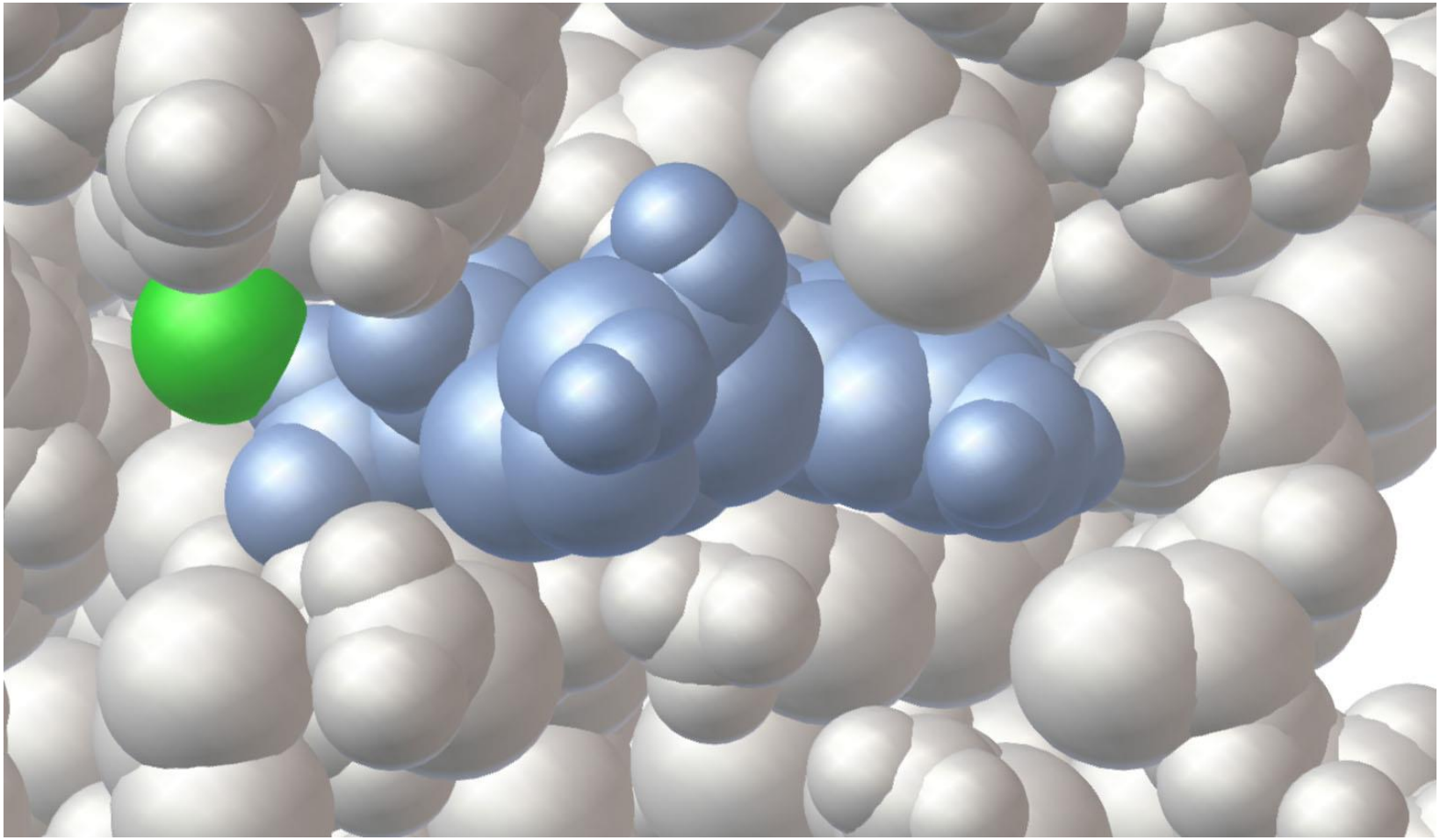
- 1990 - AutoDock 1
  - *First docking method with flexible ligands*
- 1998 - AutoDock 3
  - *Free energy force field and advanced search methods*
  - *AutoDockTools Graphical User Interface*
- 2009 - AutoDock 4
  - *Current version of AutoDock*
  - *Many parameters available to user*
- 2009 - AutoDock Vina
  - *Rewritten by Oleg Trott, new approach to scoring and search*
  - *One step solution to docking*
- 2011 - Raccoon
  - *GUI for virtual screening and analysis*

Scoring Functions



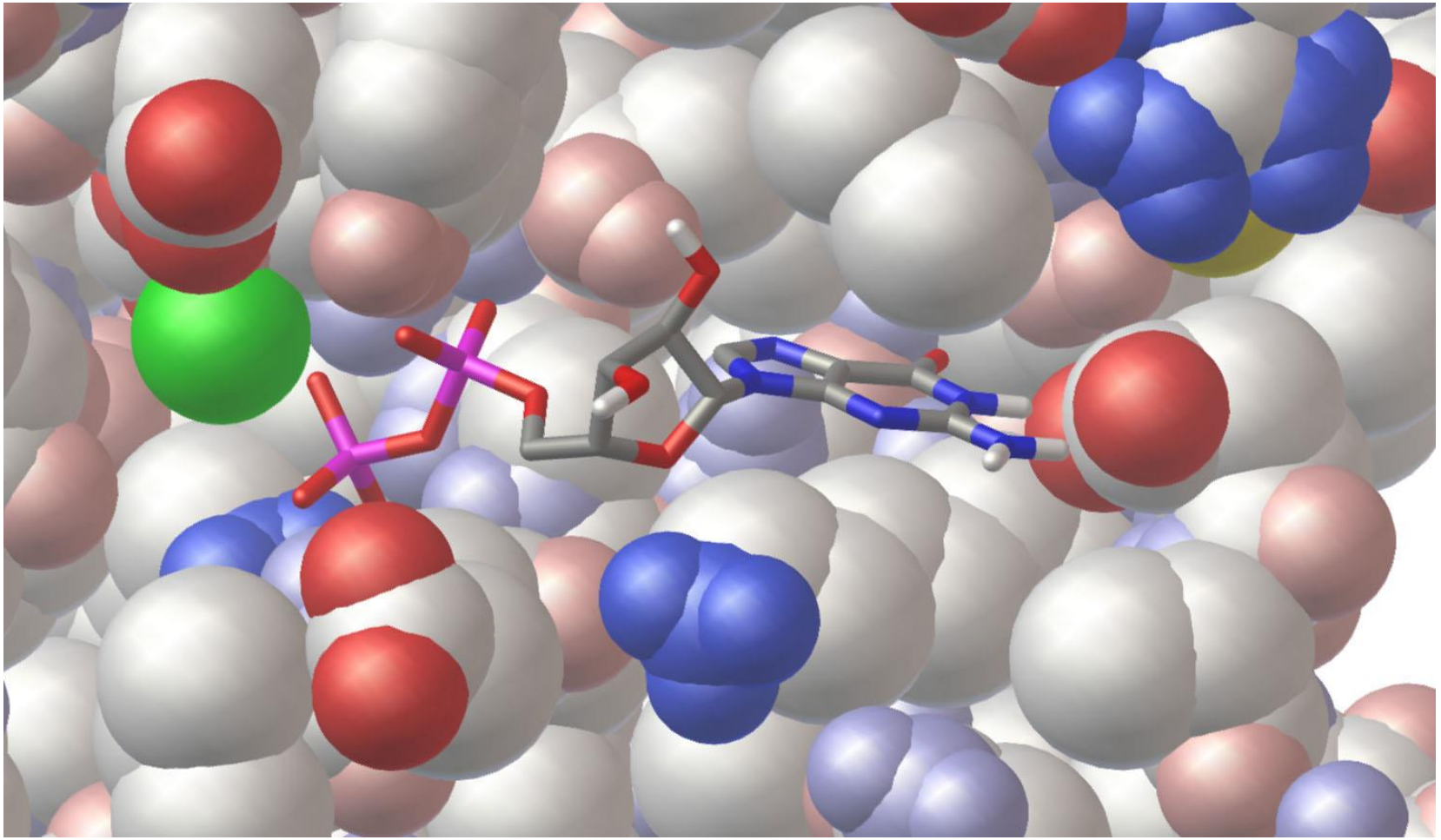
$$\Delta G_{binding} = \Delta G_{vdW} + \Delta G_{elec} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta G_{tors}$$

Dispersion/Repulsion



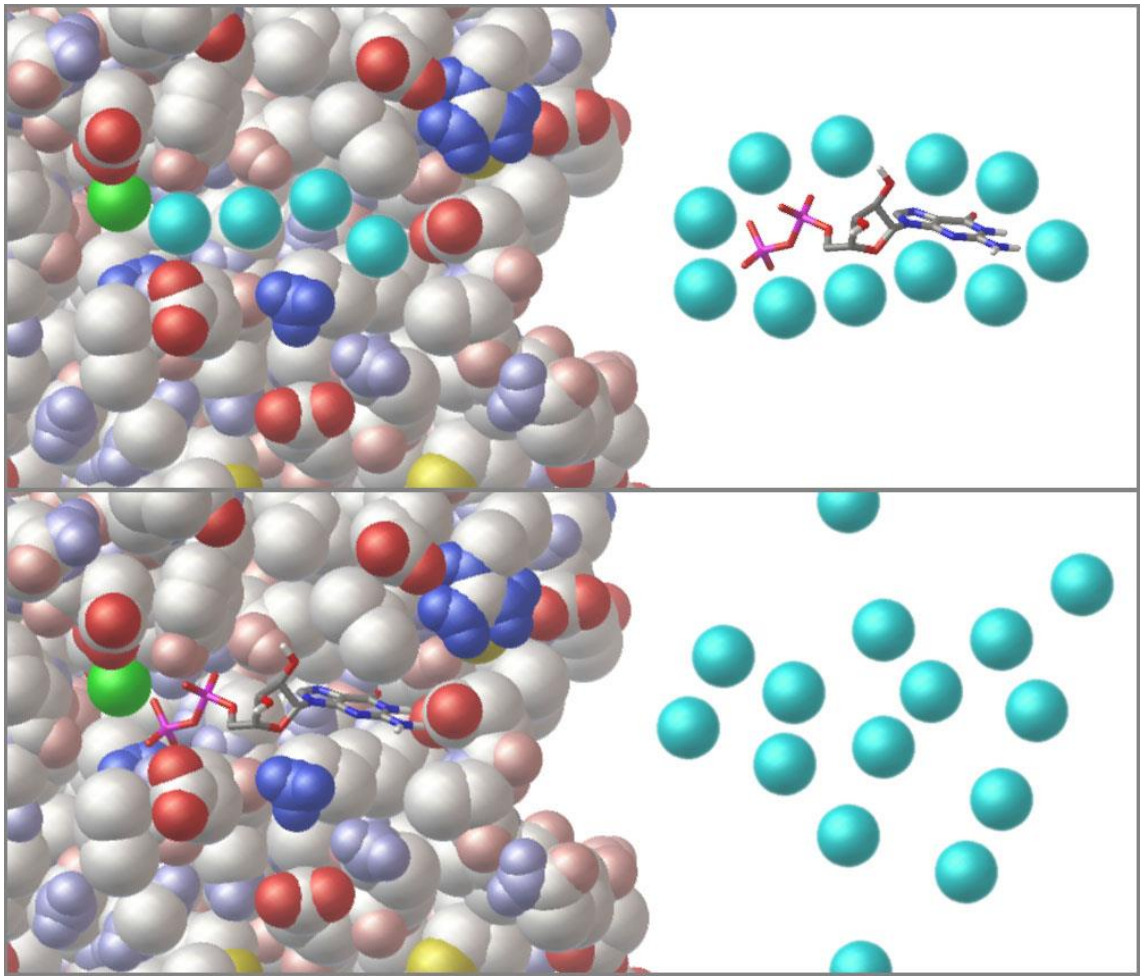
$$\Delta G_{binding} = \Delta G_{vdW} + \Delta G_{elec} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta G_{tors}$$

Electrostatics and Hydrogen Bonds



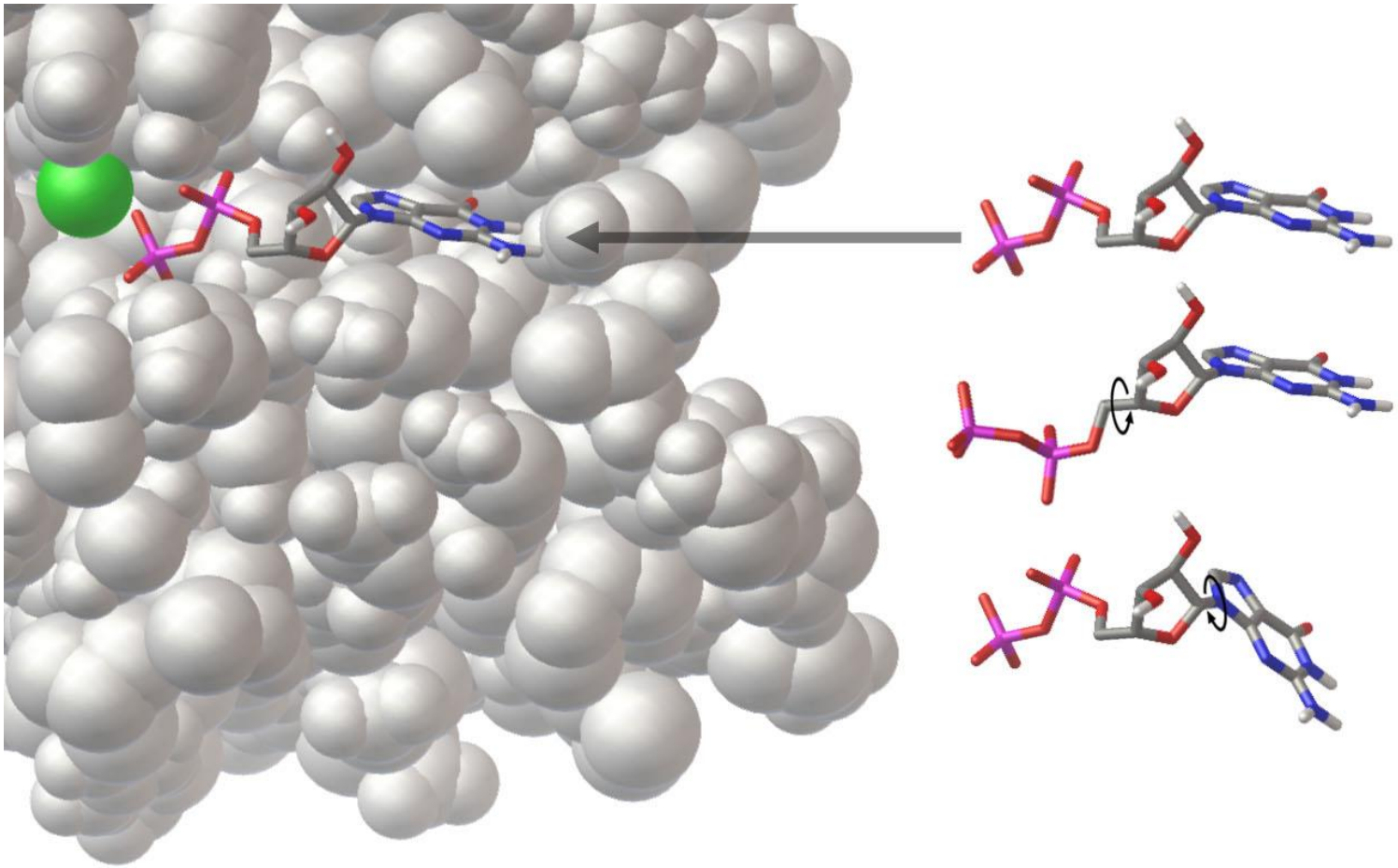
$$\Delta G_{binding} = \Delta G_{vdW} + \Delta G_{elec} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta G_{tors}$$

Desolvation



$$\Delta G_{binding} = \Delta G_{vdW} + \Delta G_{elec} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta G_{tors}$$

Torsional Entropy



$$\Delta G_{binding} = \Delta G_{vdW} + \Delta G_{elec} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta G_{tors}$$

AutoDock Empirical Free Energy Force Field

$$W_{vdw}\sum_{i,j}\left(\frac{A_{ij}}{r_{ij}^{12}}-\frac{B_{ij}}{r_{ij}^{6}}\right)+$$

$$W_{hbond}\sum_{i,j}E(t)\left(\frac{C_{ij}}{r_{ij}^{12}}-\frac{D_{ij}}{r_{ij}^{10}}\right)+$$

$$W_{elec}\sum_{i,j}\frac{q_i q_j}{\varepsilon(r_{ij})r_{ij}}+$$

$$W_{sol}\sum_{i,j}\left(S_i V_j + S_j V_i\right)e^{(-r_{ij}^2/2\sigma^2)}+$$

$$W_{tor}N_{tor}$$

- Physics-based approach from molecular mechanics
- Calibrated with 188 complexes from LPDB, K_i*'s* from PDB-Bind
- Standard error = 2.52 kcal/mol

AutoDock Vina Scoring Function

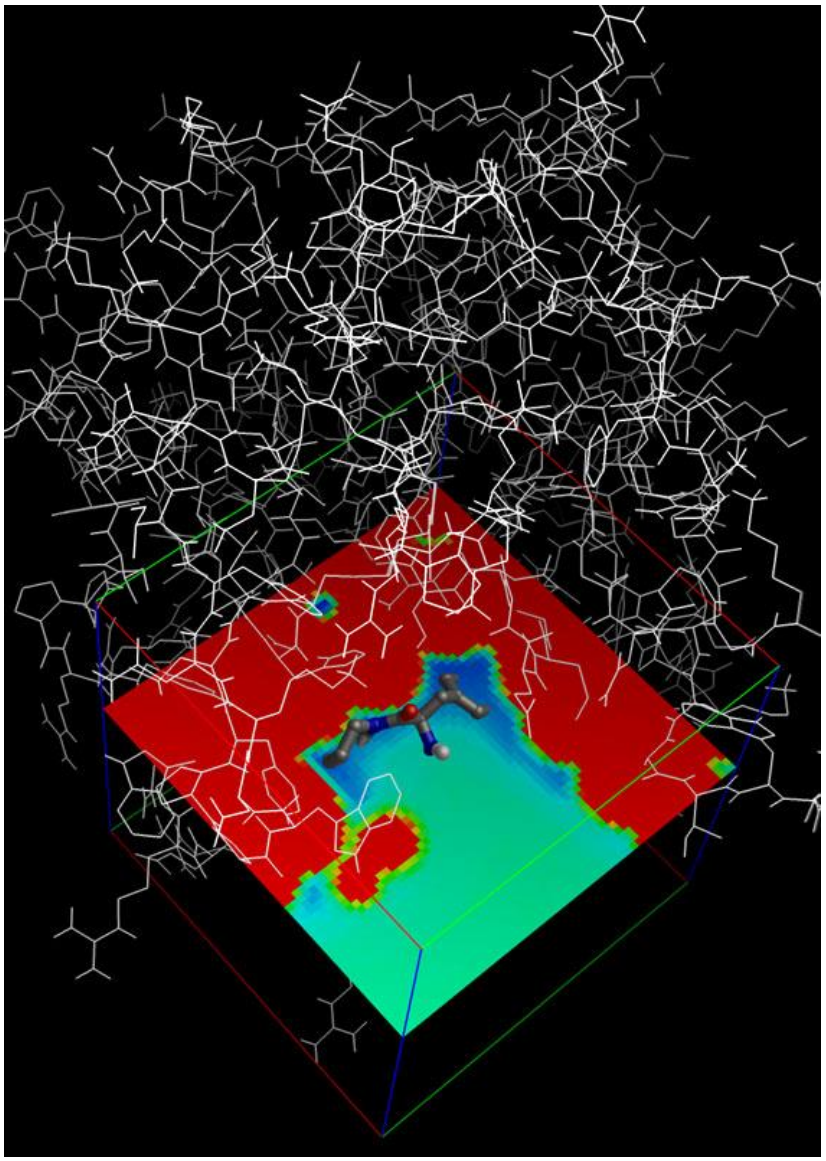
Combination of knowledge-based and empirical approach

$$\Delta G_{binding} = \Delta G_{gauss} + \Delta G_{repulsion} + \Delta G_{hbond} + \Delta G_{hydrophobic} + \Delta G_{tors}$$

- ΔG_{gauss}
  Attractive term for dispersion, two gaussian functions
- $\Delta G_{repulsion}$
  Square of the distance if closer than a threshold value
- ΔG_{hbond}
  Ramp function - also used for interactions with metal ions
- $\Delta G_{hydrophobic}$
  Ramp function
- ΔG_{tors}
  Proportional to the number of rotatable bonds

- Calibrated with 1,300 complexes from PDB-Bind
- Standard error = 2.85 kcal/mol

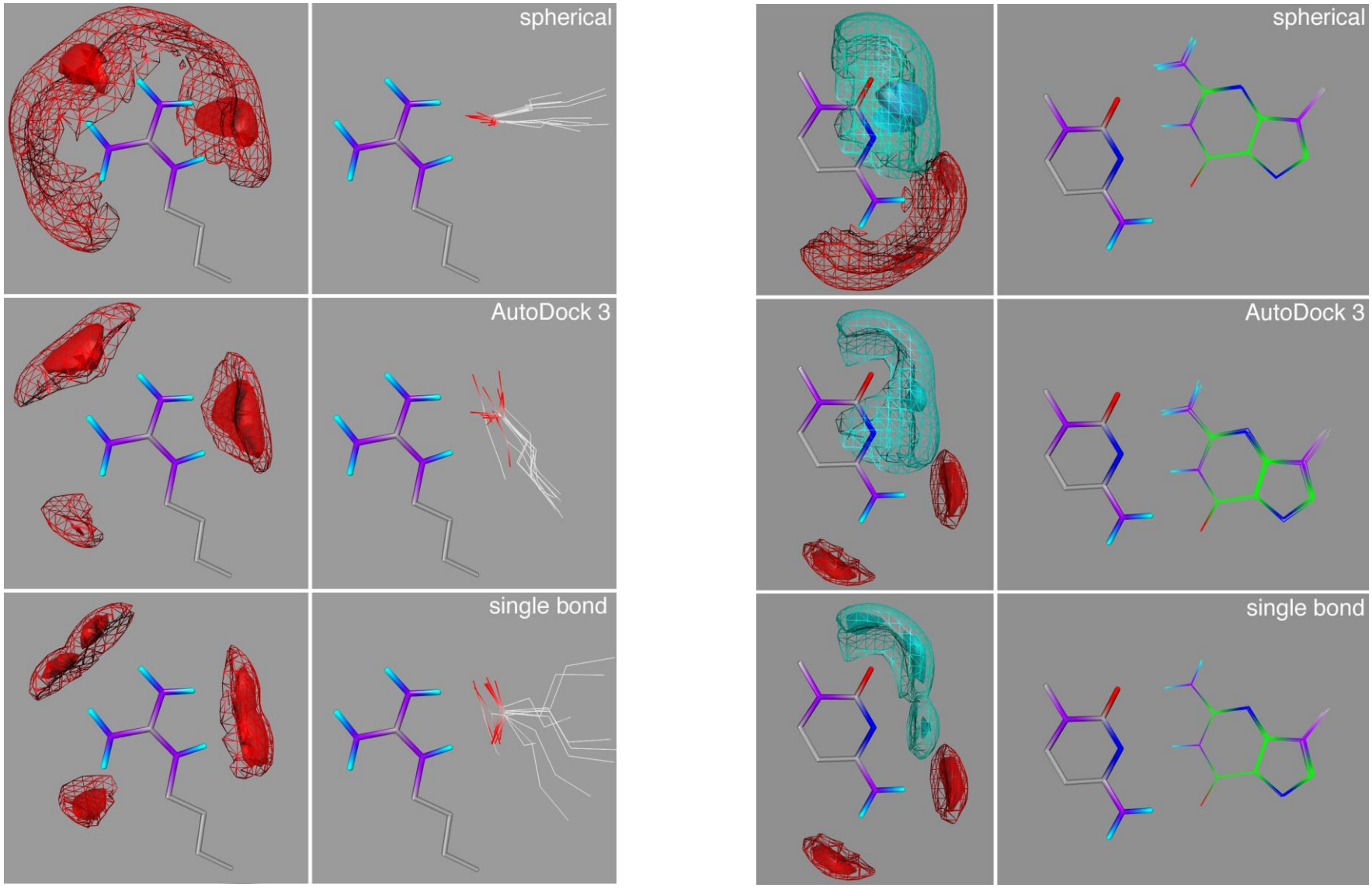
http://vina.scripps.edu



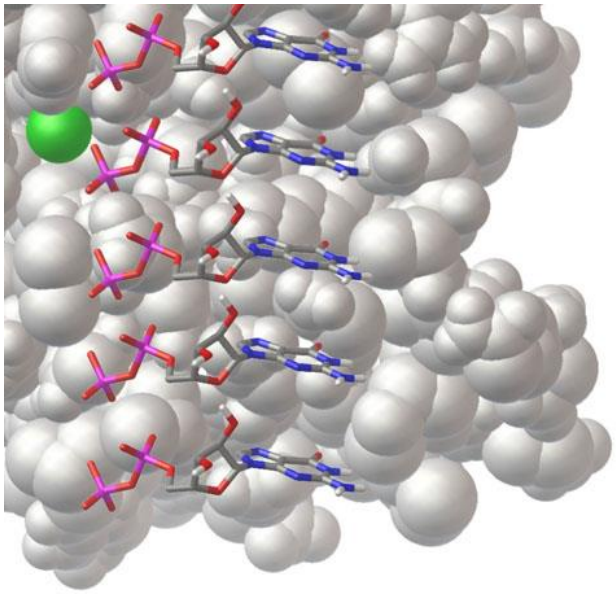
Grid Maps

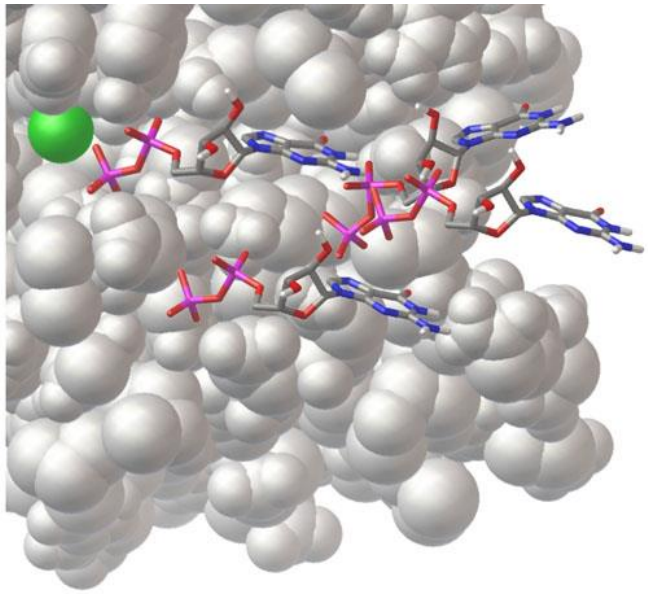
- Precompute interactions for each type of atom
- 100X faster than pairwise methods
- Drawbacks: receptor is conformationally rigid, limits the search space

H-bond Directionality



Two Kinds of Search





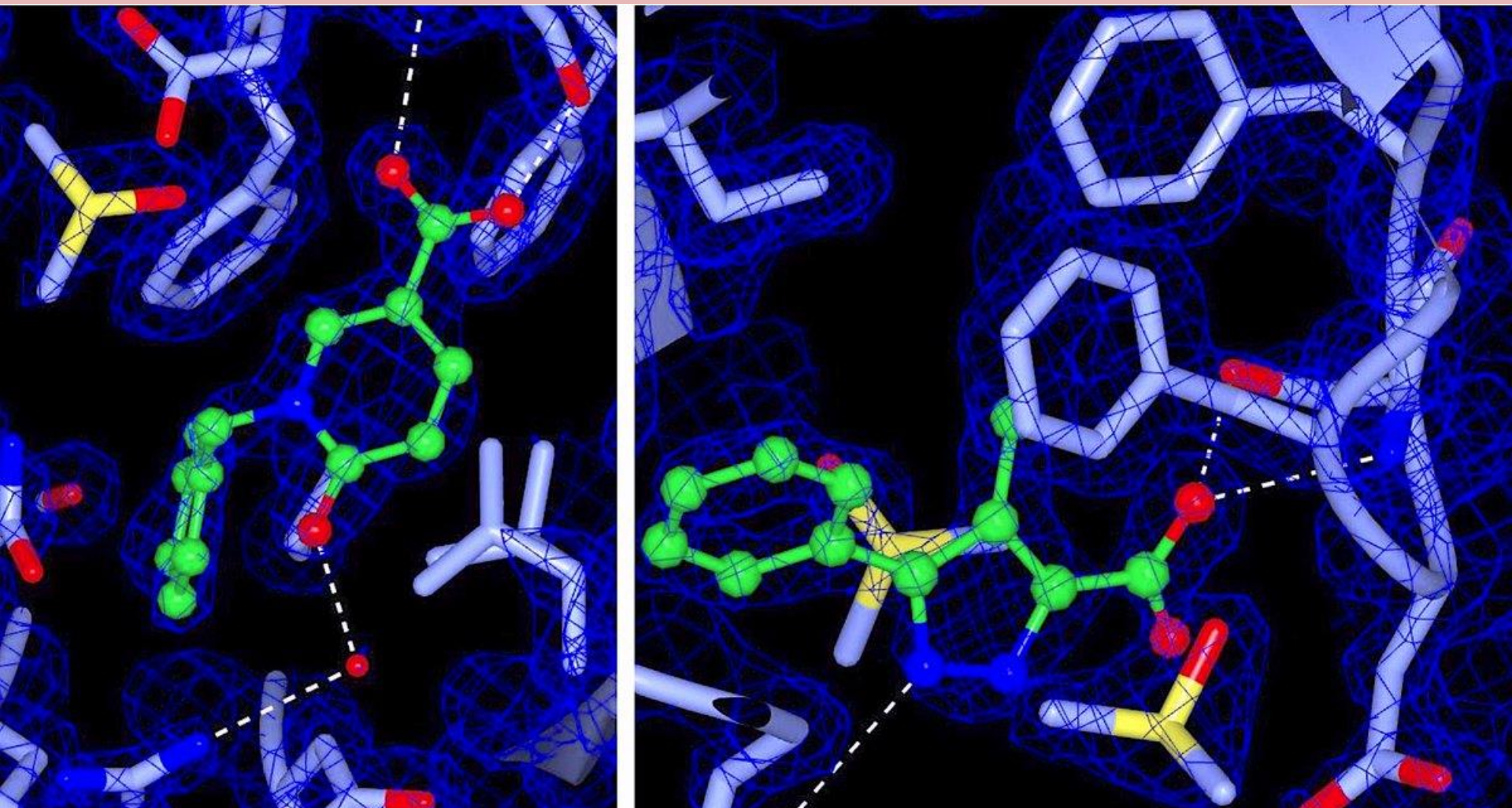
Systematic

- Exhaustive, deterministic
- Outcome is dependent on granularity of sampling
- Feasible only for low-dimensional problems

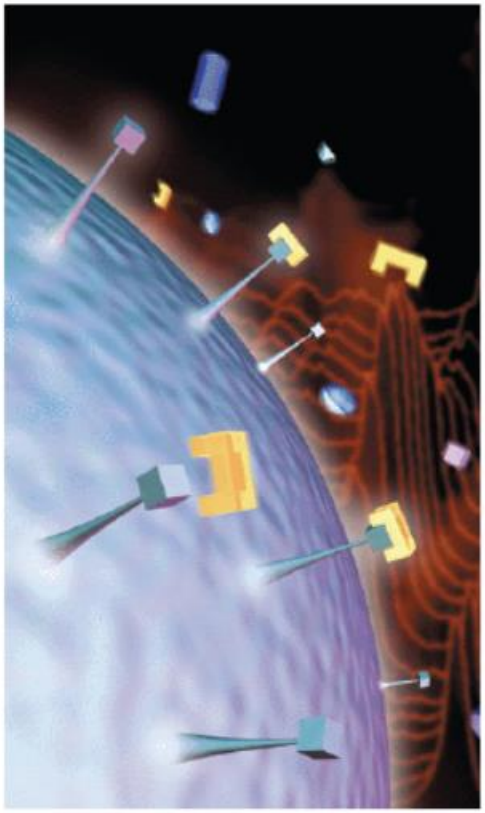
Stochastic

- Random, outcome varies
- Must repeat the search or perform more steps to improve chances of success
- Feasible for larger problems

Fragment Based Drug Discovery (FBDD)

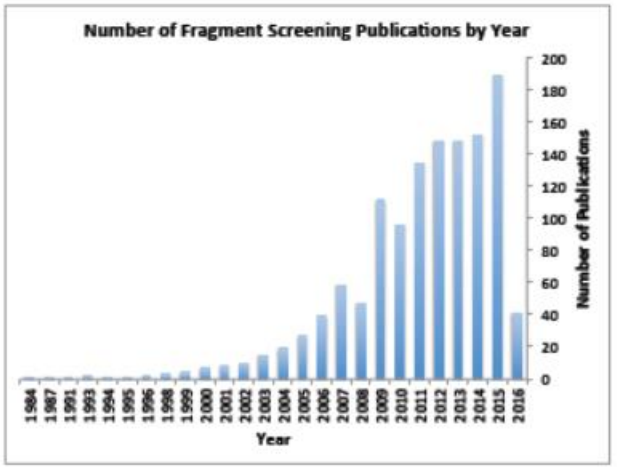


BROAD INSTITUTE



Fragment-Based Drug Discovery: A Primer

March 2016

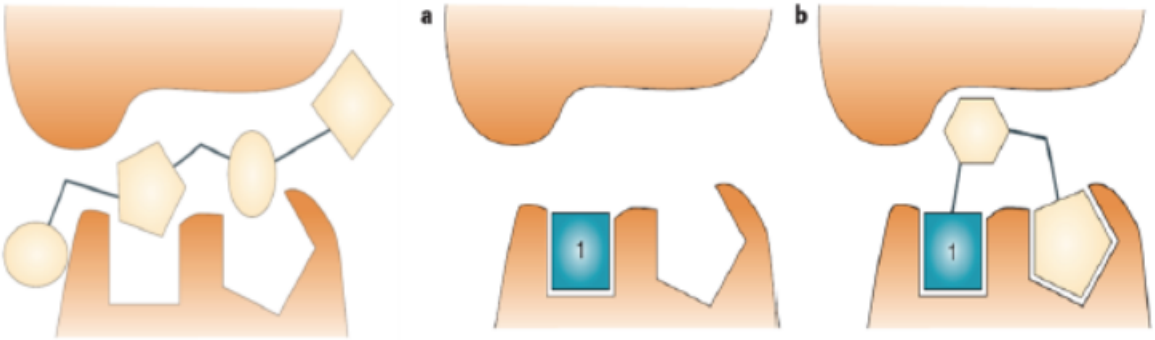


https://www.broadinstitute.org/videos/broade-fragment-based-drug-discovery-101

Youtube: https://youtu.be/UiSqOCgznew

Why Fragment Screening?

1. Fragments allow diverse sampling of 'Chemical Space'
   - Helps get out of 'local minima' – New binding modes and binding sites
   - Better Ligand Efficiency
   - Needle in a smaller haystack
2. HTS tends to lead to worse drug-like properties
   - Flat, aromatic, hydrophobic: Poor solubility, more promiscuity, higher toxicity
   - High Concentration Screening (HCS) assays can lead to artifacts
   - Opinion: Canonical fragment libraries still often too hydrophobic, monolithic

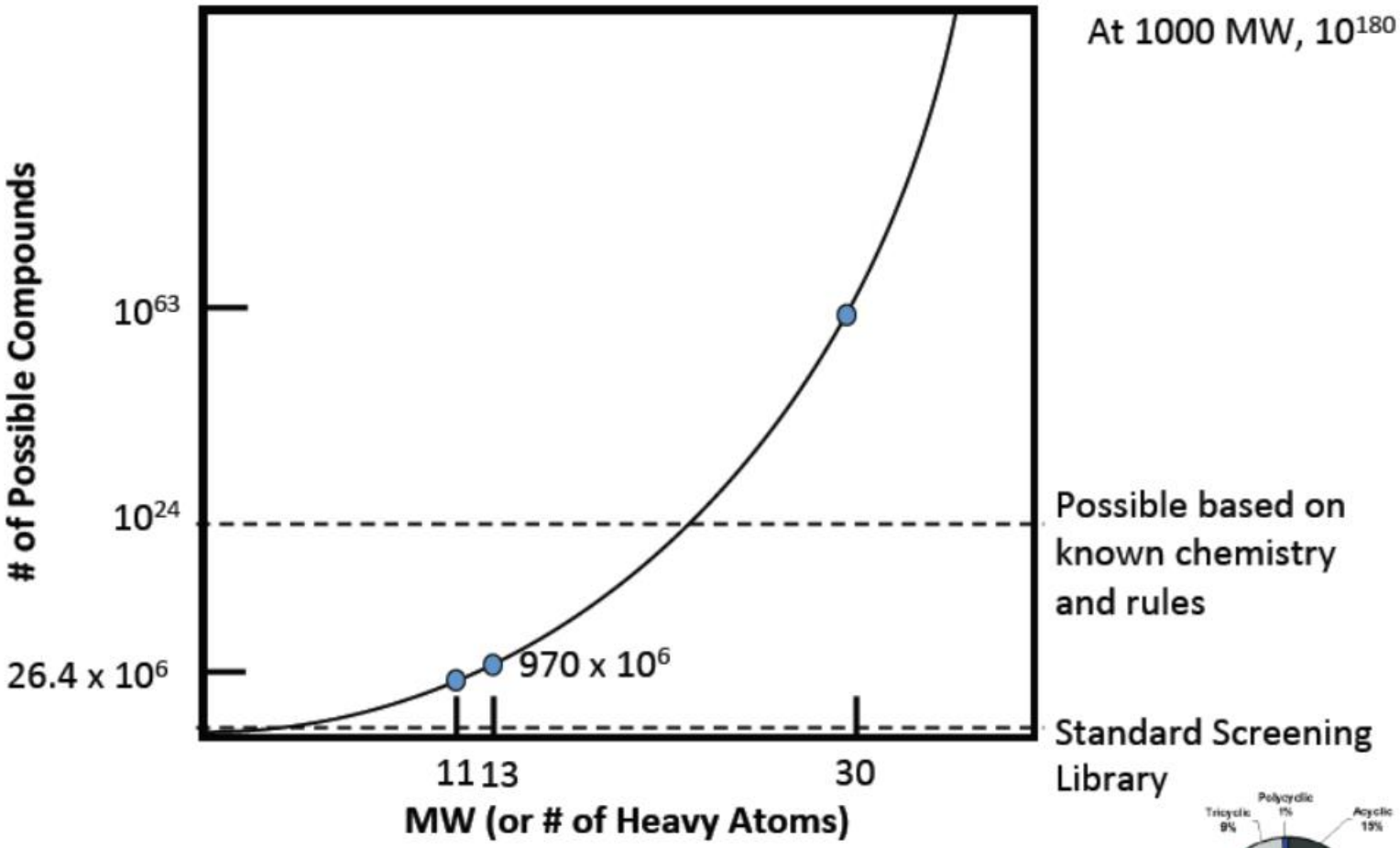


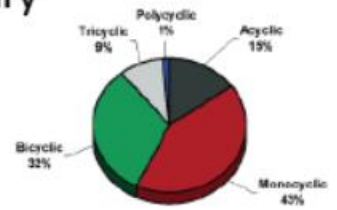
Classical HTS Hit Structure Guided Evolution

Weaker binders
Additivity
Sensitive assay
- SPR
- NMR
- Others

Rees et al. *Nat. Rev. Drug. Disc.* **2004** 3: 660-672.

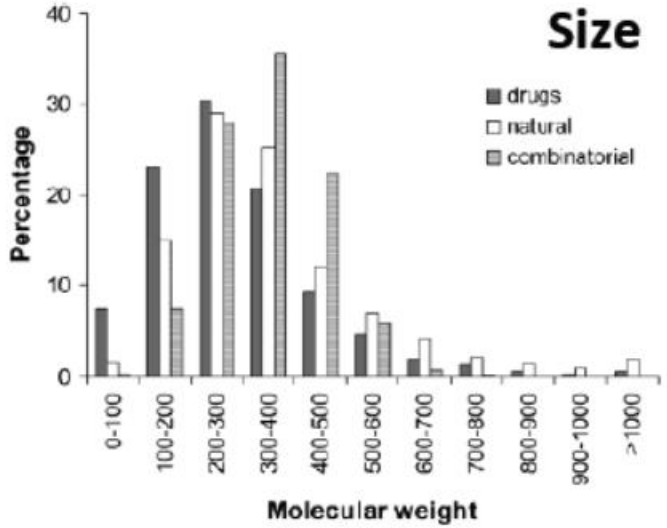
Chemical Space: Possible Compounds

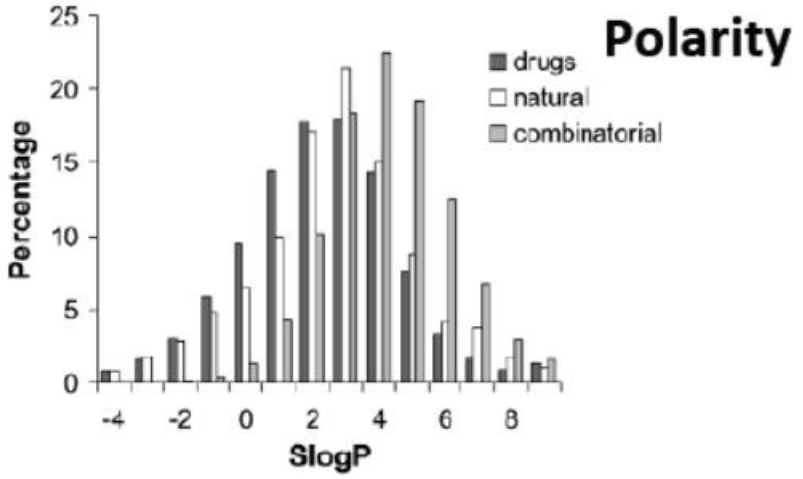




Fink and Reymond, *J. Chem. Inf. Model.* **2007** 47: 342-353.
Bohacek, et. al., Med. Res. Rev. **1996** 16: 3-50.

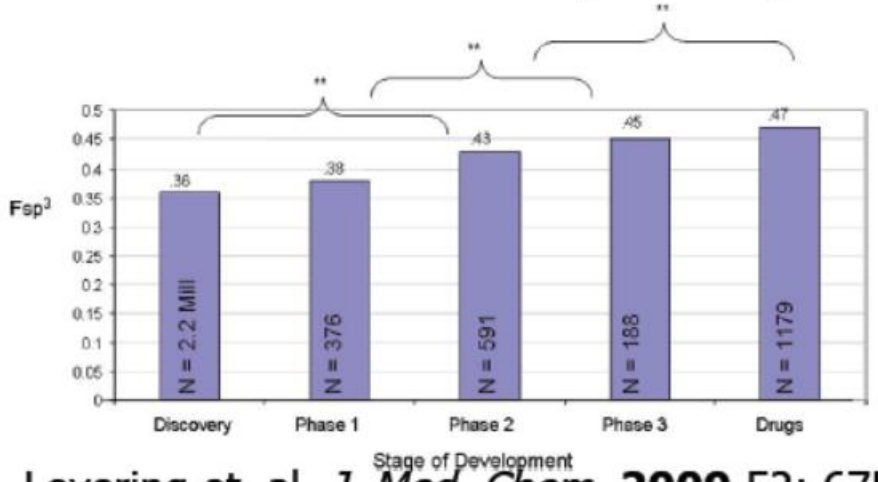
Druglike Properties are Also Important

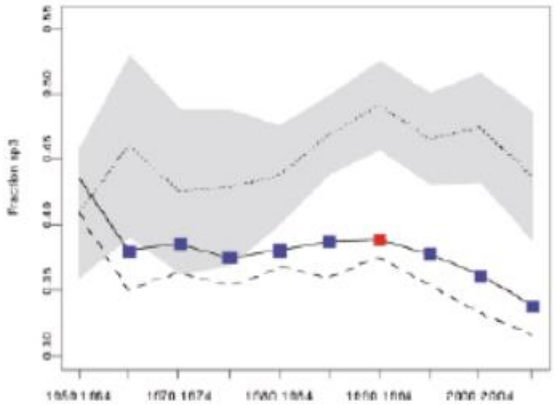




Feher and Schmidt. *J. Chem., Inf. Comp. Sci.* **2003** 43: 218-227.

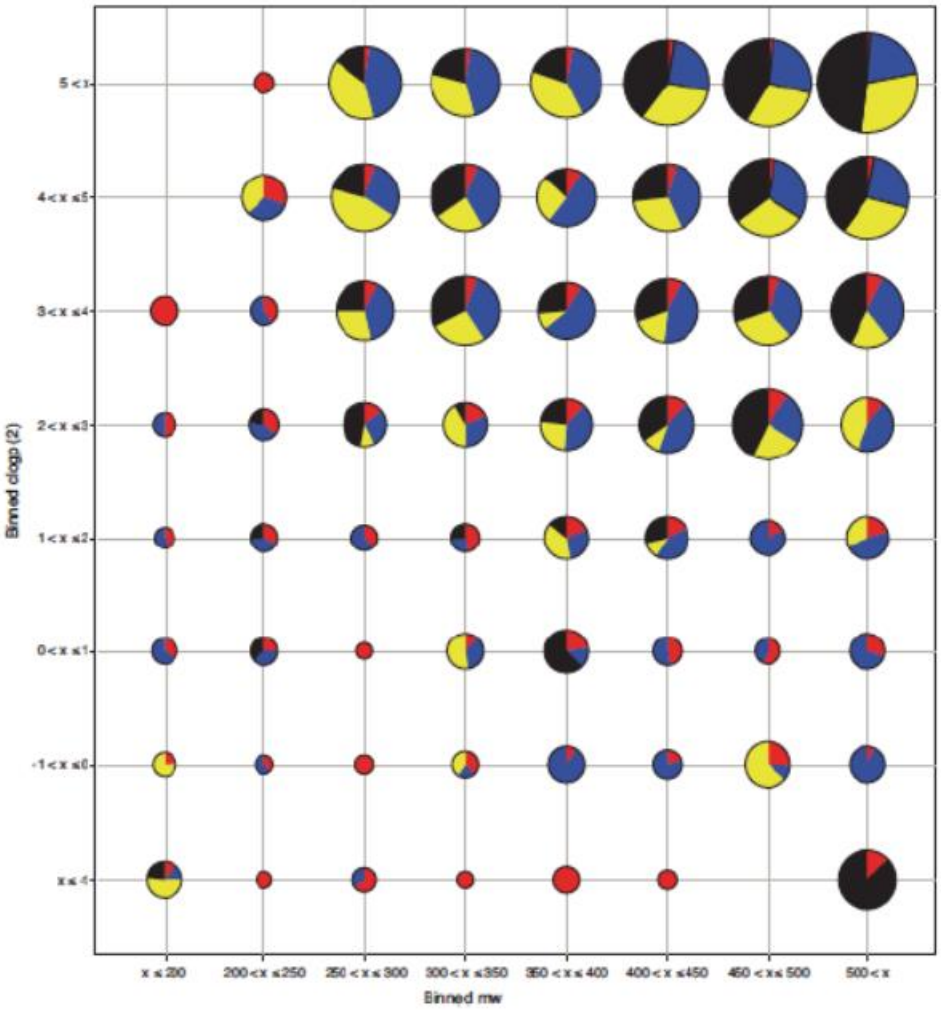
Three Dimensionality





Lovering et. al. *J. Med. Chem.* **2009** 52: 6752-6756.

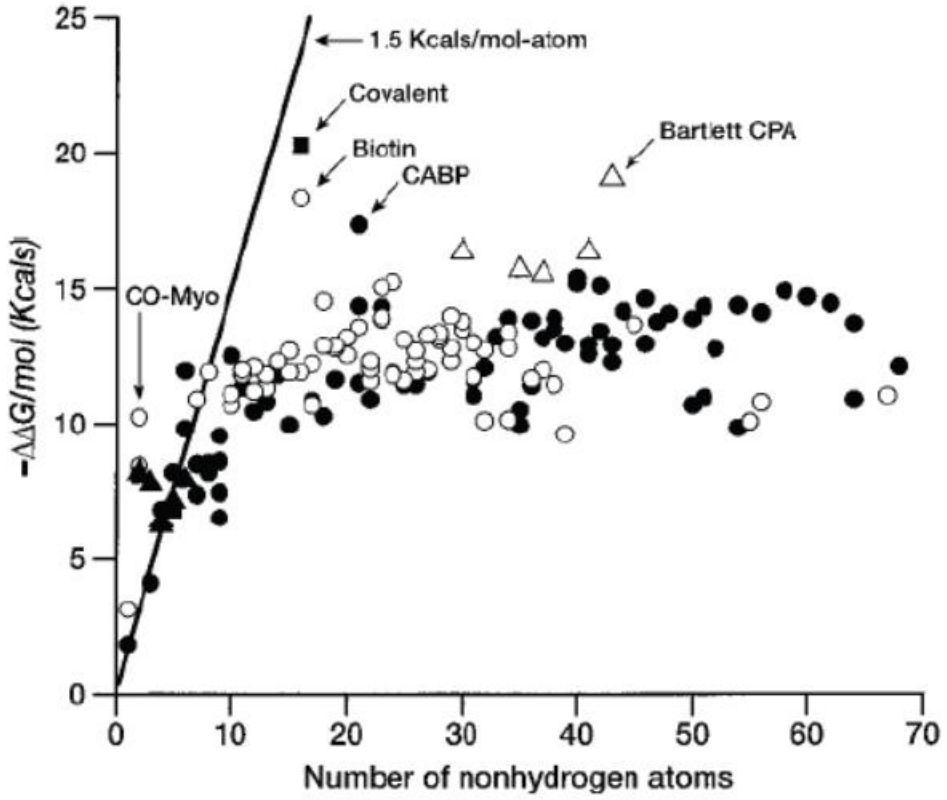
Physical Properties and Hit Frequency



Less polar compounds hit more, and are more likely to be promiscuous

Leach and Hann. *Curr. Opin. Chem. Biol.* **2011** 15: 489-496.

Maximal Affinity of Ligands



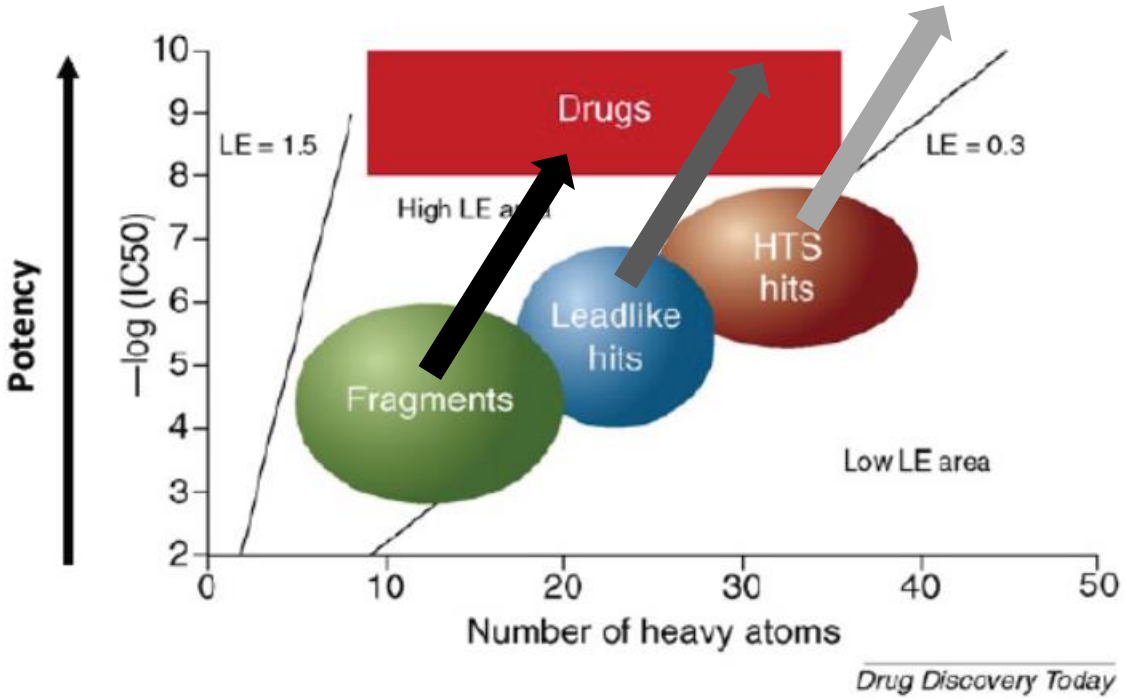
Kuntz, et. al. *Proc. Natl. Acad. Sci. USA* **1999** 96: 9997-10002.

Ligand Efficiency is Important

Ligand Efficiency: The amount of binding energy gained from each heavy atom

For every 64 mass units, there is typically a ten-fold improvement in potency

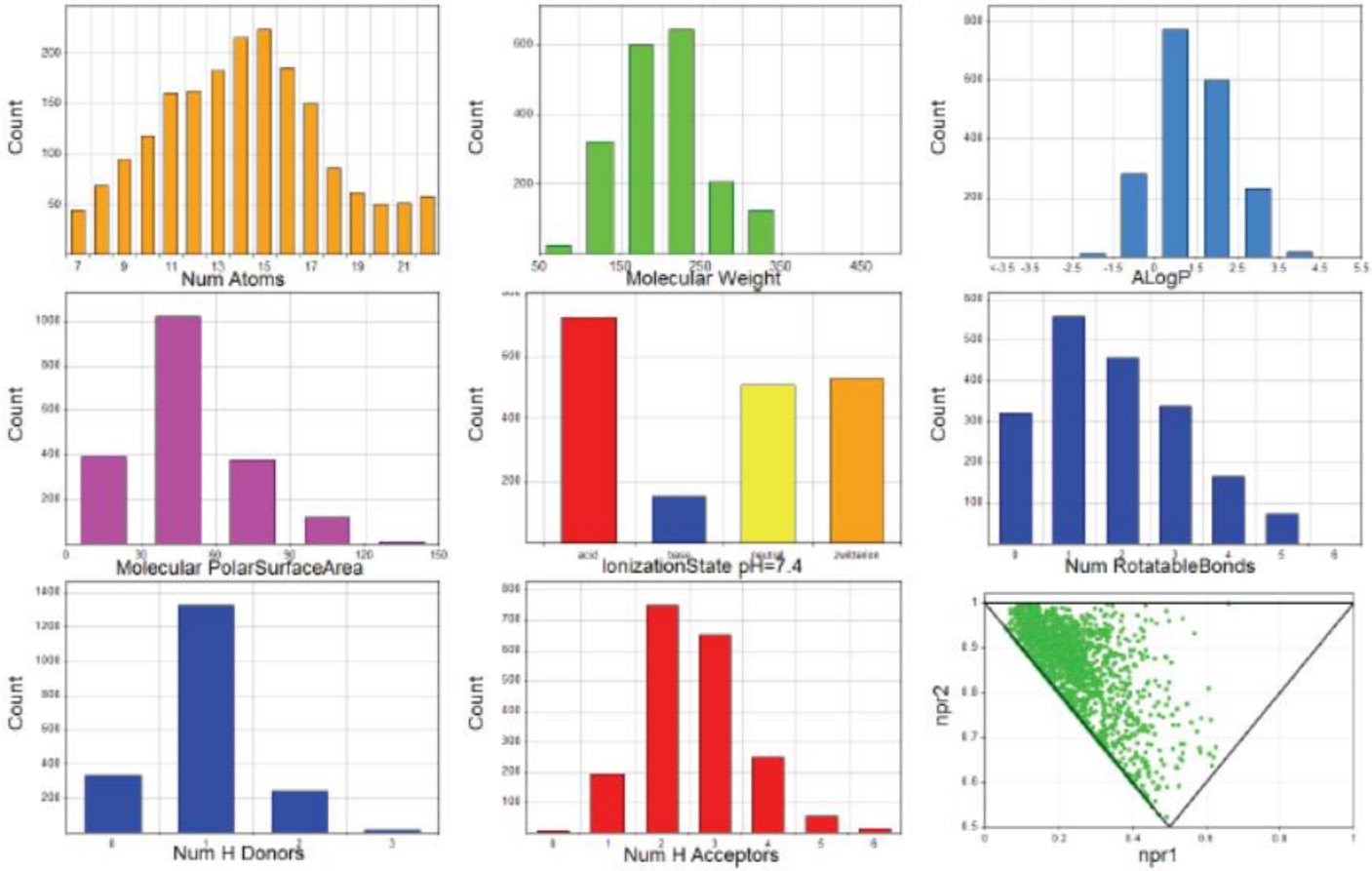
$$LE = \frac{-\Delta G}{HAC} = \frac{-RT \ln IC_{50}}{HAC}$$



Typical trajectory from hit to lead to candidate

3rd Dimension: Molecules also typically get less polar

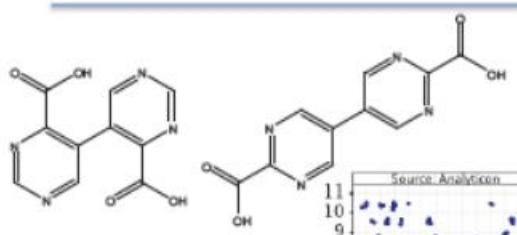
Physical Properties: Diverse Library

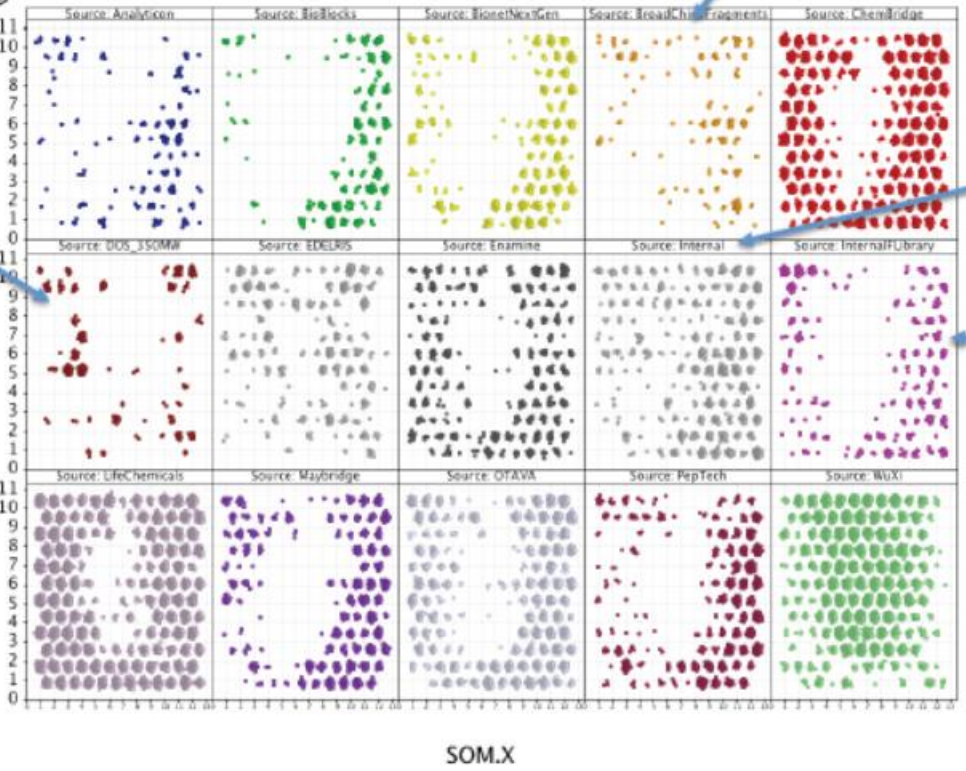


Many people use 'The Rule of Three'

Thanks: Patrick McCarren

Self-Organizing Maps and Chemical Space





Chiral Skeletons

Diverse Vendor Library

DOS Fragments

^{19}F Library

Use chemical fingerprints to cluster compounds and assess diversity

Overall coverage is good, suggests some areas to enrich

Thanks: Patrick McCarren

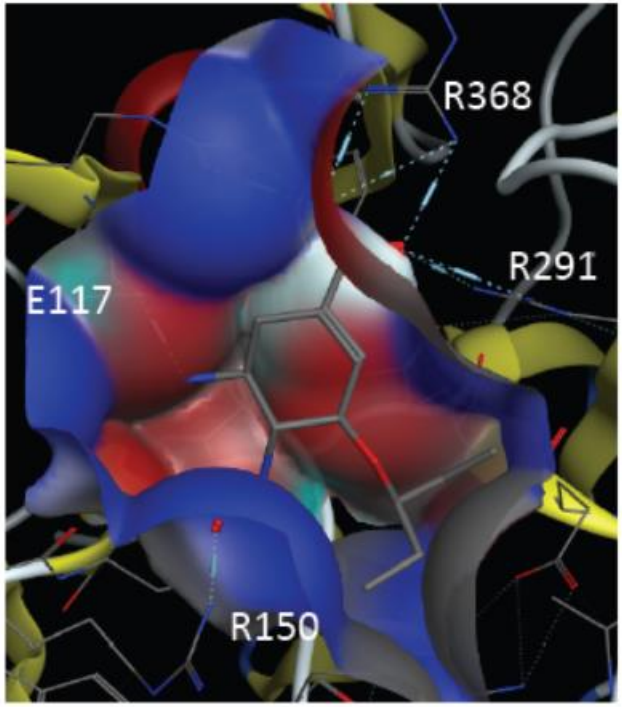
Commercial Fragment Libraries

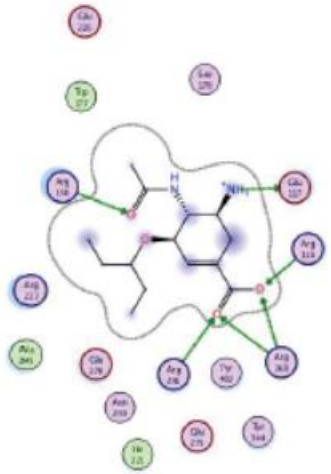
Vendor	# Frags	Comment	Vendor	# Frags	Comment
Analyticon	50 scaff.	FRGx (NP)	Key Organics	26,000	Sub: 1,166
Asinex	22,524	RO3	Life Chemicals	31,000	Sub: 3,000
Biofocus	1,500	Sol., 3D	Maybridge	2,500	Sub: 1,000
ChemBridge	6,500	RO3, Sol	Otava	12,486	RO3
ChemDiv	14,294	RO3	Prestwick	910	RO3
DomainEx	1,000	Diverse, Sol	Selcia	~2,000	CE Comp.
Edelris	>500	3D	TimTec	3,200	RO3
Emerald	1,500	Frag of Life	Vitas-M	18,932	RO3
Enamine	26,543	RO3	WuXi		RO3
Infarmatik	1,700	3D	Zenobia	1,000	RO3

http://www.cambridgemedchemconsulting.com/resources/hit_identification/fragment_collections.html

Why Structure-Based Design?

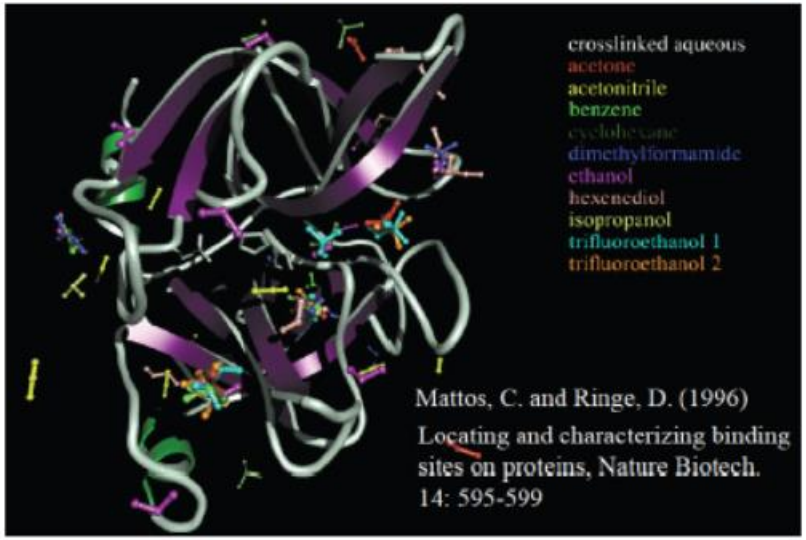
- Weaker compounds are more sensitive to subtle changes
- Useful to guide design: potency and selectivity
- Practical: Accelerates timeline of optimization
- Examples: Tamiflu, Tipranavir, Gleevec



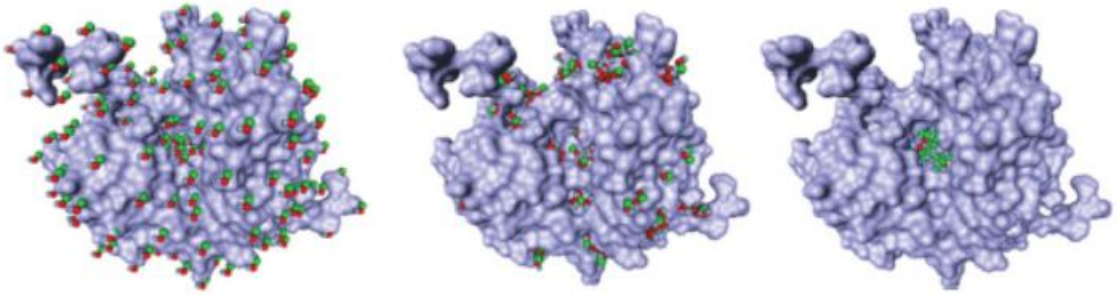


Druggability Assessment

- Experimental
  - NMR (Solvent-induced shifts)
  - X-Ray (MSCS)
  - Fragment Screening hit rates



- Computational
  - Assess surface concavity purely geometrically
  - Hotspot analysis (FTMap) – http://ftmap.bu.edu



Brenke, et. al. *Bioinformatics.* **2009** 25: 621-627.

Biophysical Methods

- Fragment-sized molecules are generally weaker
  - μM to mM
  - Biophysical methods capable of robustly detecting and quantifying
- Many interesting targets have no assay
  - May have no known function or assay
  - Protein-Protein Interactions
  - Scaffolding Proteins
  - Intrinsically Disordered Proteins
- High ligand concentrations cause artifacts
  - Fluorescence, redox cycling, aggregation (compound or protein), …

Fragment-Based Discovery Workflow