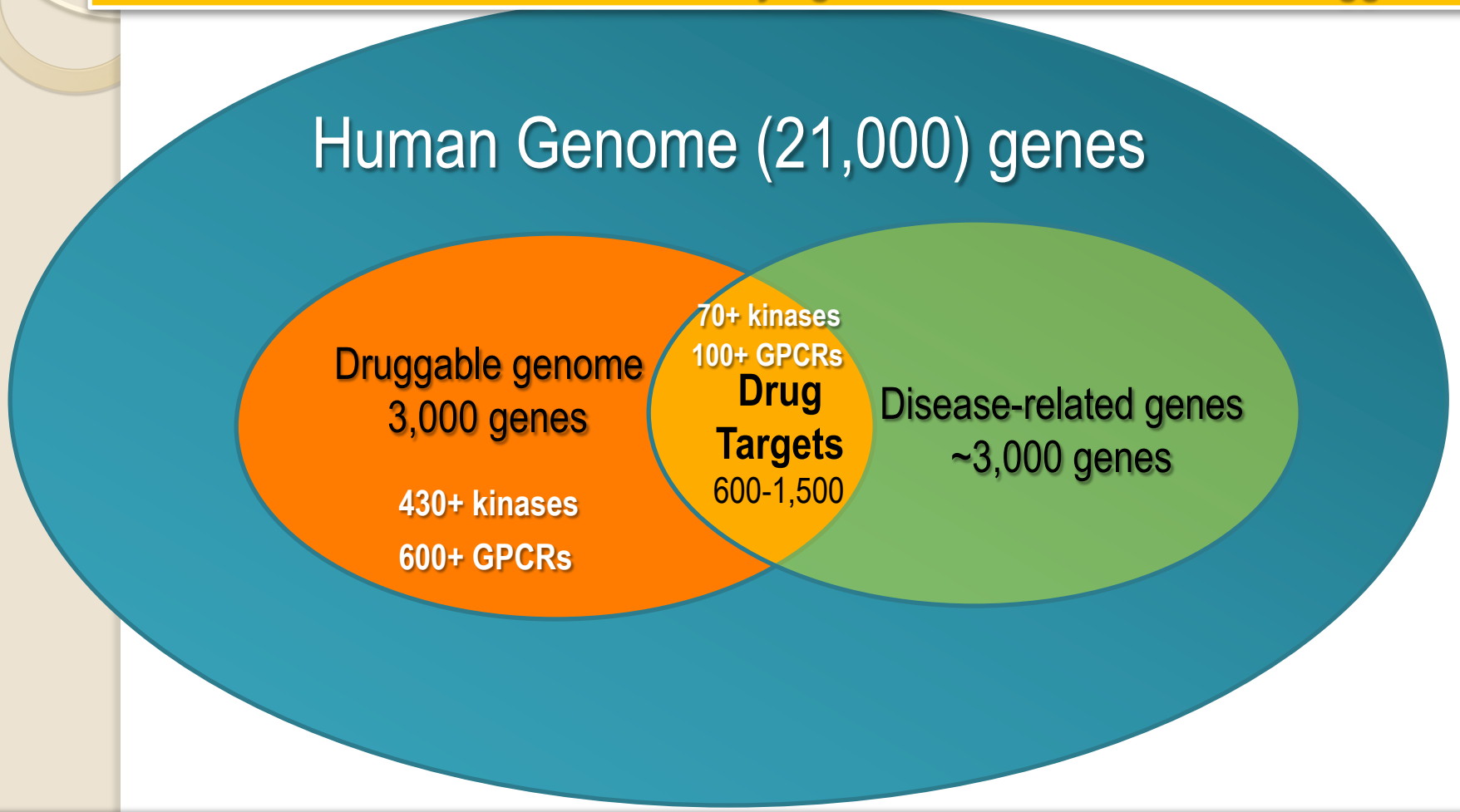
SUMMARY

1. Structure-encoded **flexibility** of drug targets and significance in drug discovery and design
2. **Druggability** assessment: a first step before selecting a target
3. Modularity and **promiscuity** of proteins and **quantitative systems pharmacology** methods

Druggable Genome

A small subset of are 'disease-modifying' – and not all of them are druggable



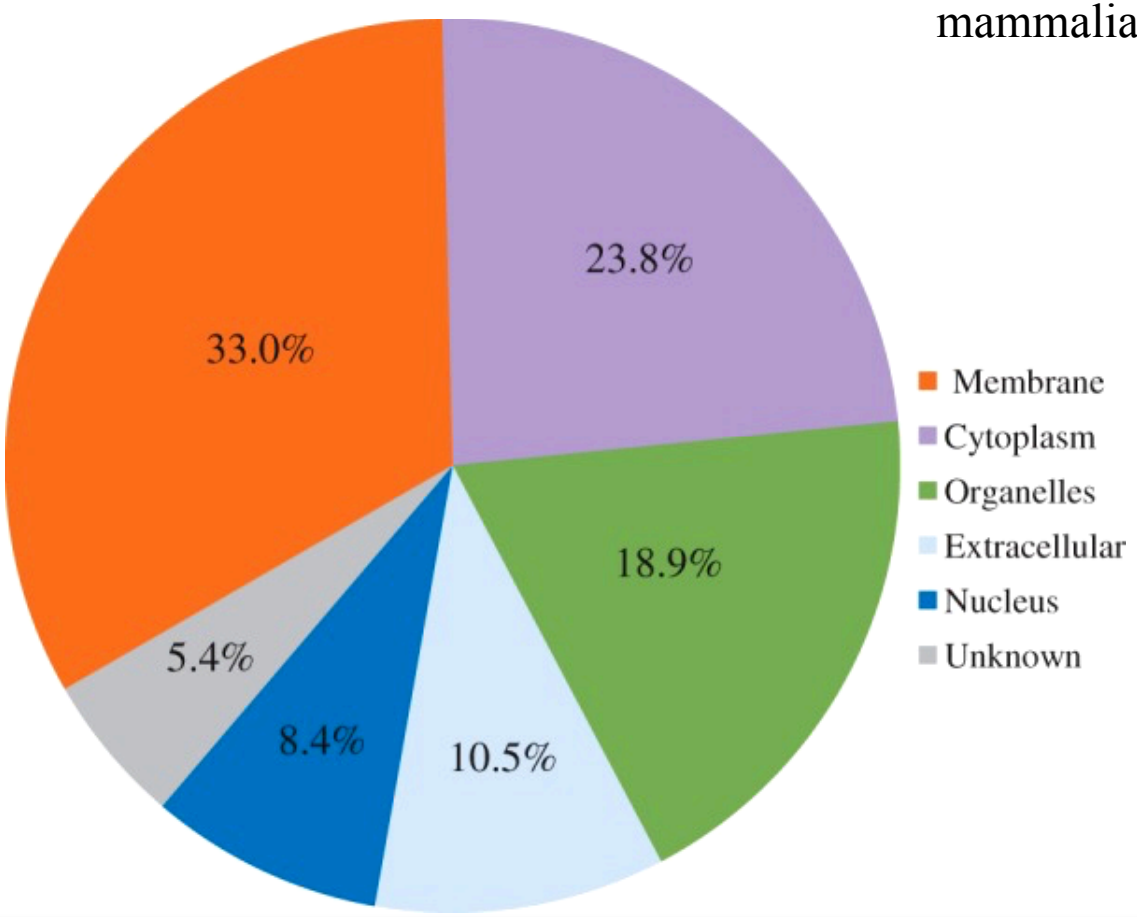
Hopkins and Groom, *Nat Reviews Drug Disc*, 2002

A few numbers…

- Only 2% of human proteins interact with currently approved drugs.
- 10-15% of human proteins are disease-modifying
- 10-15% are druggable
- 5% are both disease-modifying and druggable

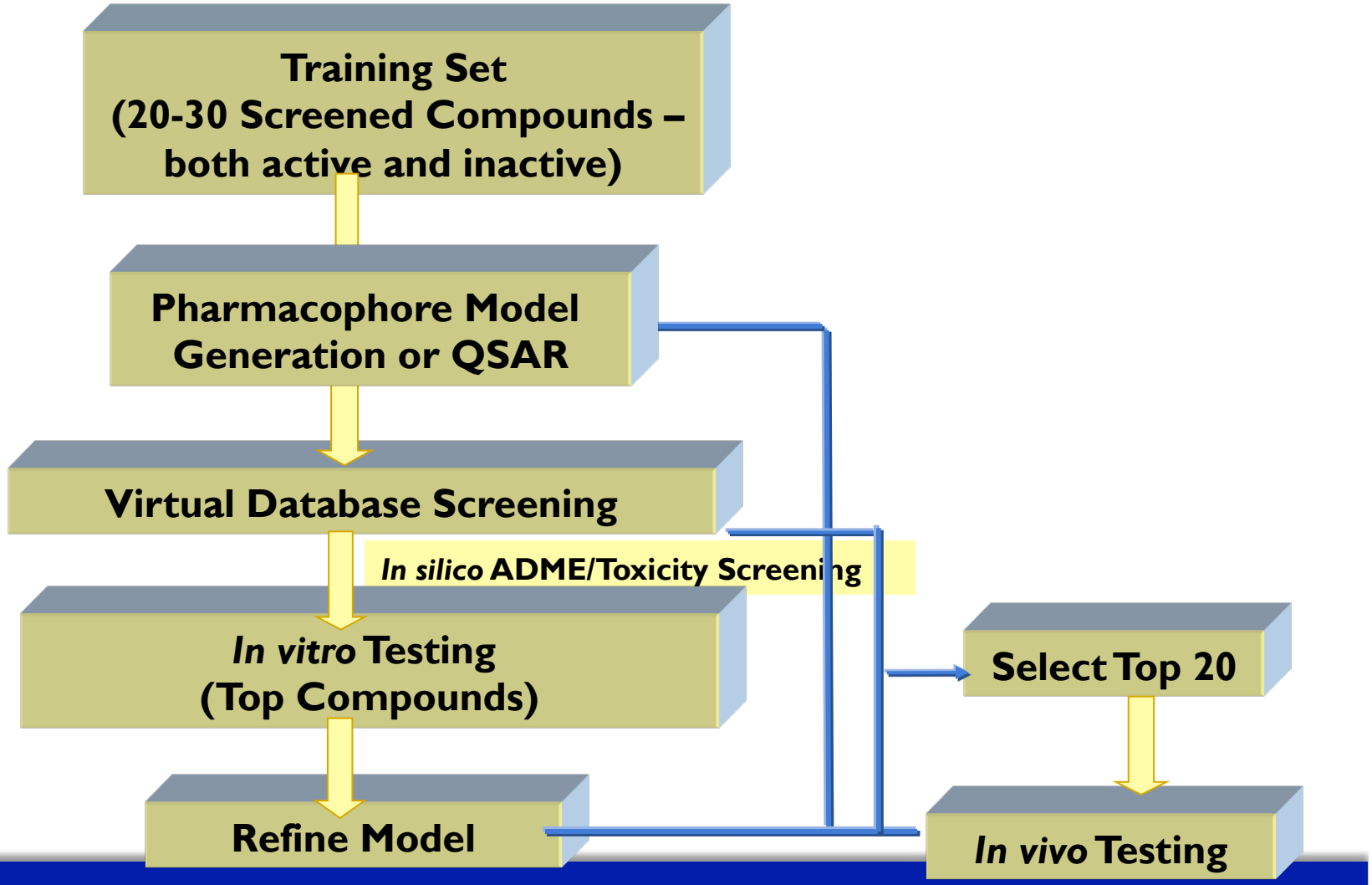
Subcellular distribution of 1,362 druggable targets

among four mammalian species.

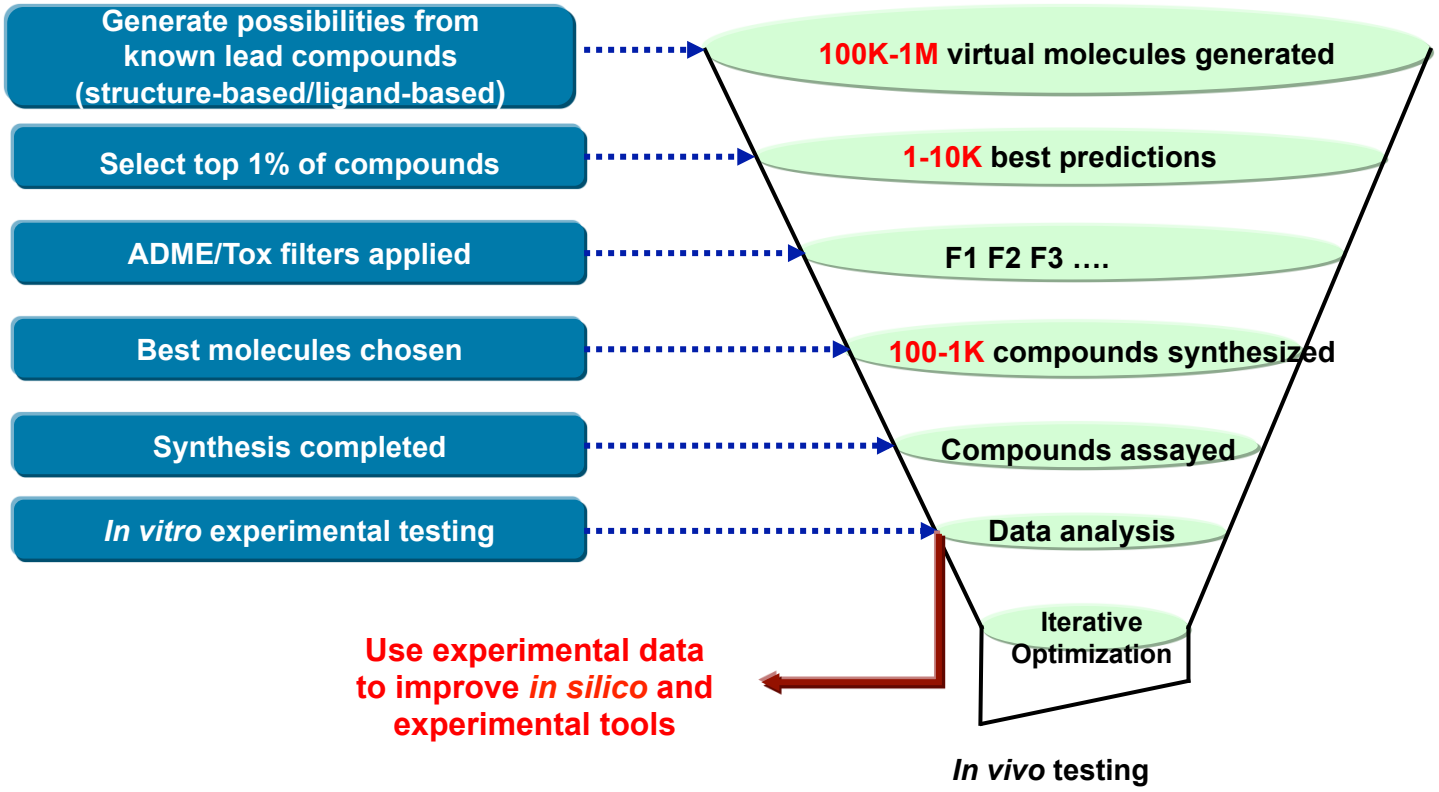


Genome Biol Evol. 2013; 5(7): 1291–1297.

Ligand-Based Strategy



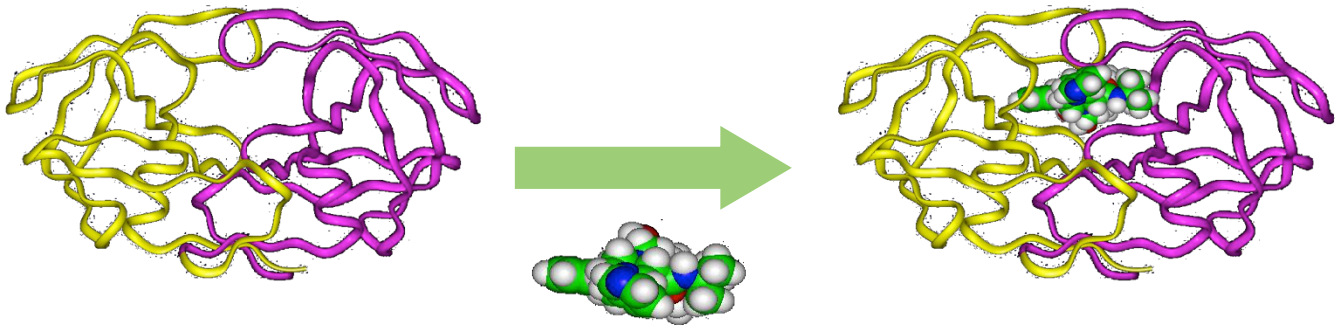
Screening Cascade



Rational Design of Inhibitors

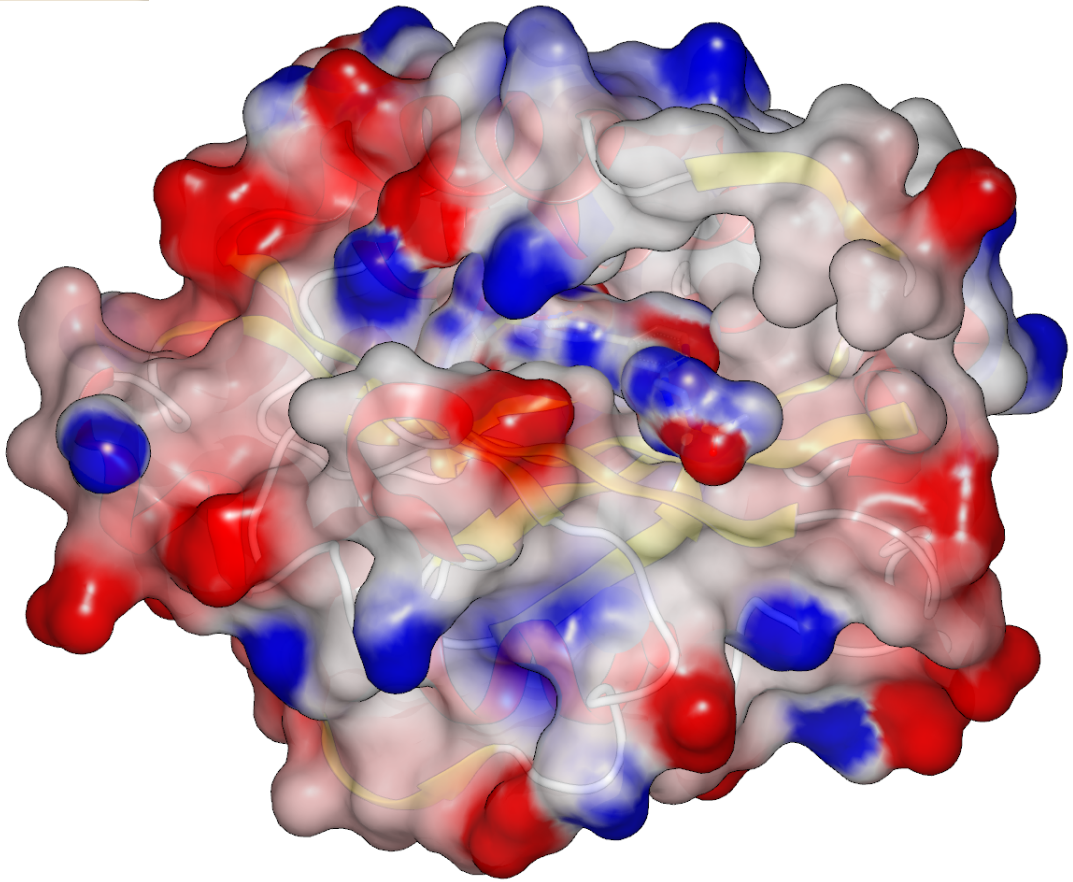
3D structure of the target is used for

- Visual inspection/molecular graphics
- Docking (of both small molecules or fragments thereof)
- De novo methods
- Receptor properly mapping + database searching



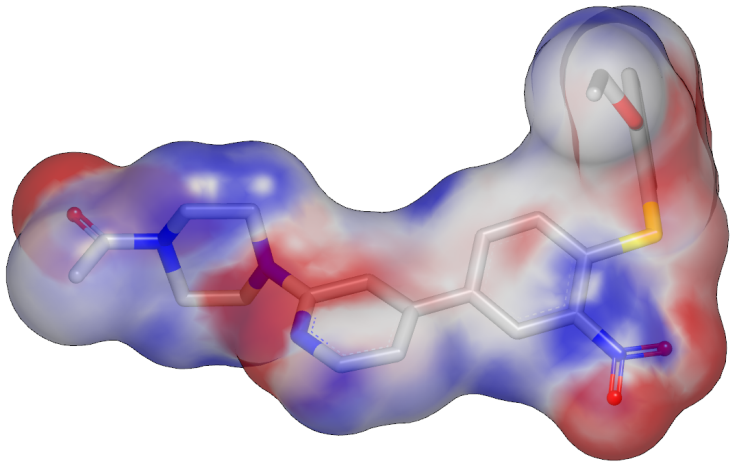
2

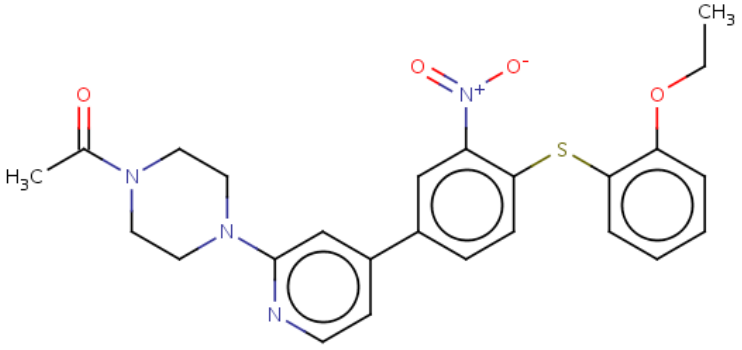
Druggable or not?



Active site druggability:

- Best known K_d 18.3 nM
- Simulation 0.03-0.5 nM

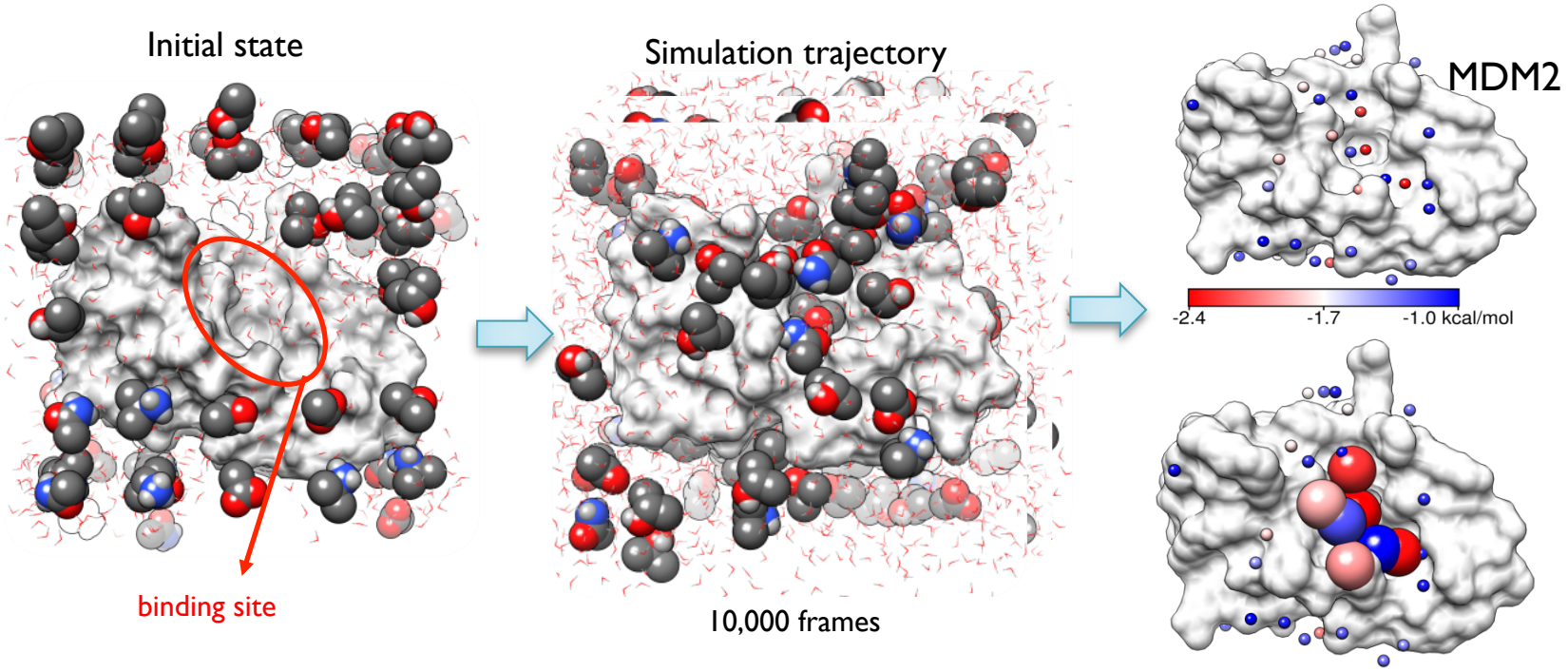


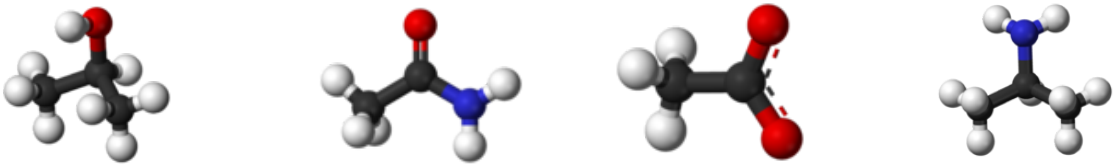


Lfa1 - a leukocyte glycoprotein that promotes intercellular adhesion and binds intercellular adhesion molecule 1

2

Druggability Simulations





Isopropanol
(observed in 57% of drugs)

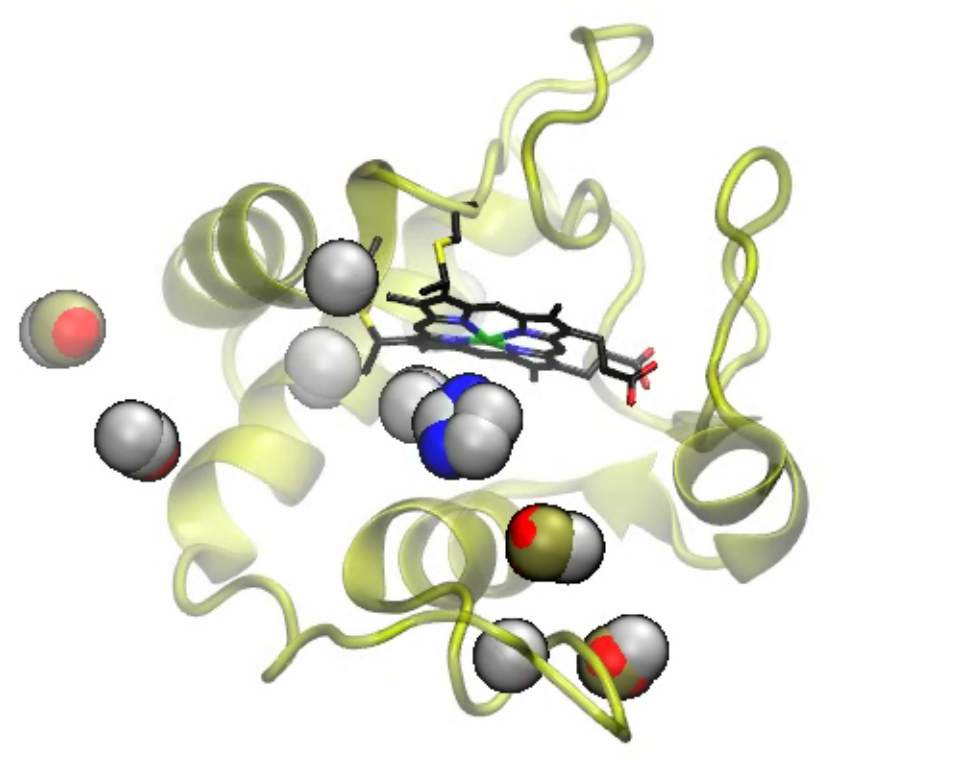
Acetamide
(21%)

Acetate (-1)
(21%)

Isopropylamine (+1)
(25%)

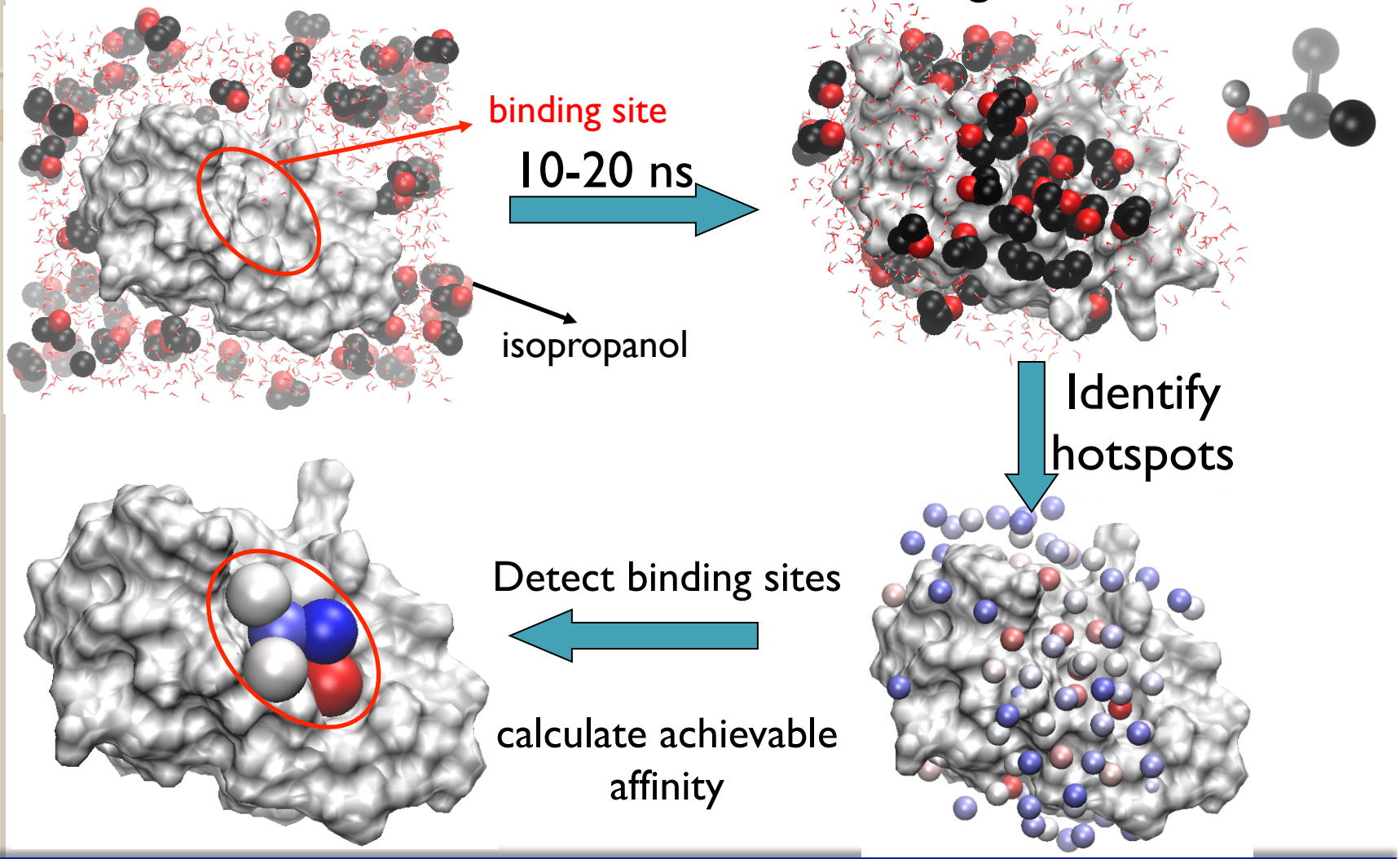
Bakan & Bahar, *J Chem Theory Comput*. 8:2435-2447, 2012

Cytochrome c druggability

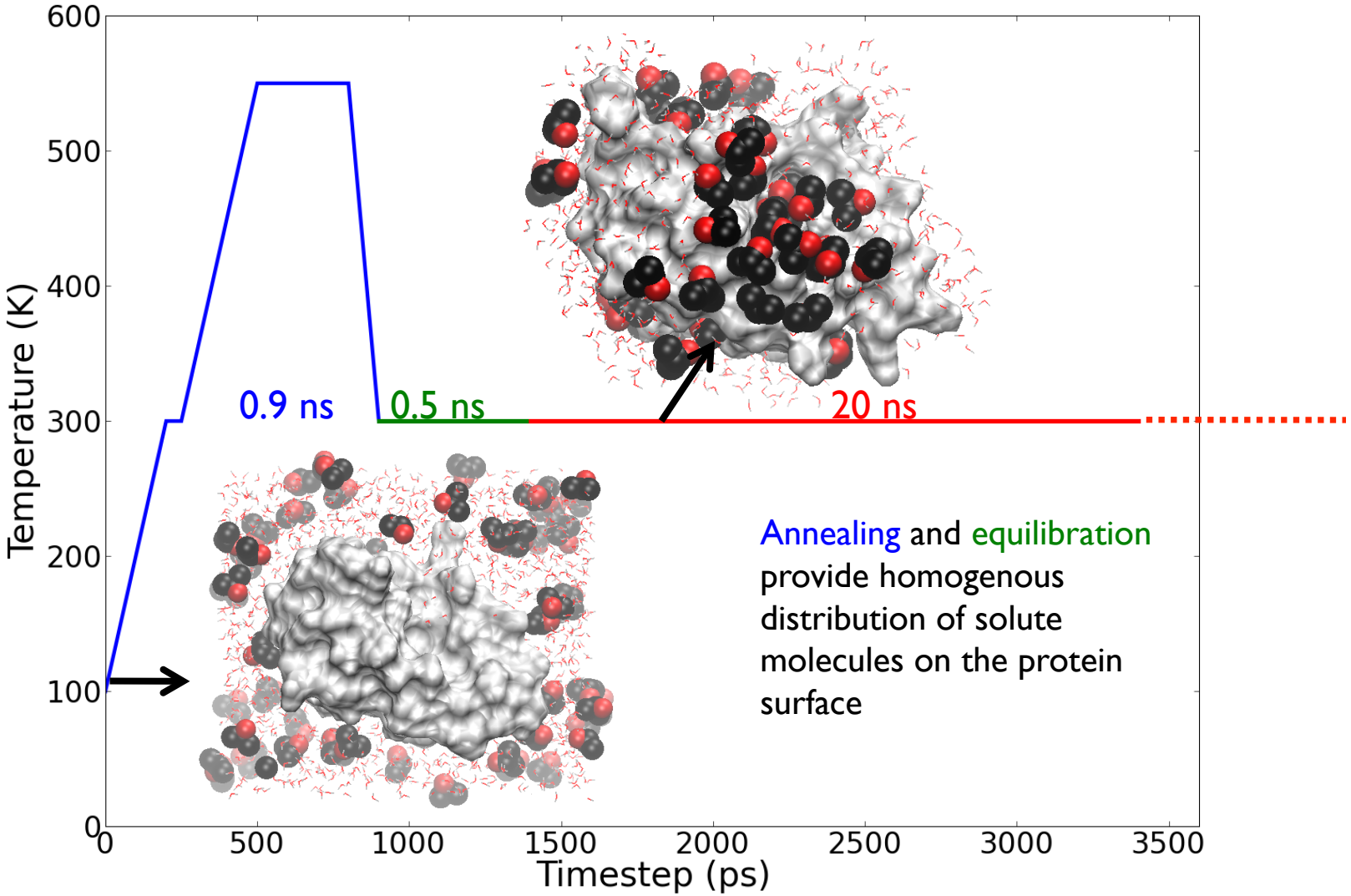


Methodology Overview

From MD simulations to achievable drug affinities



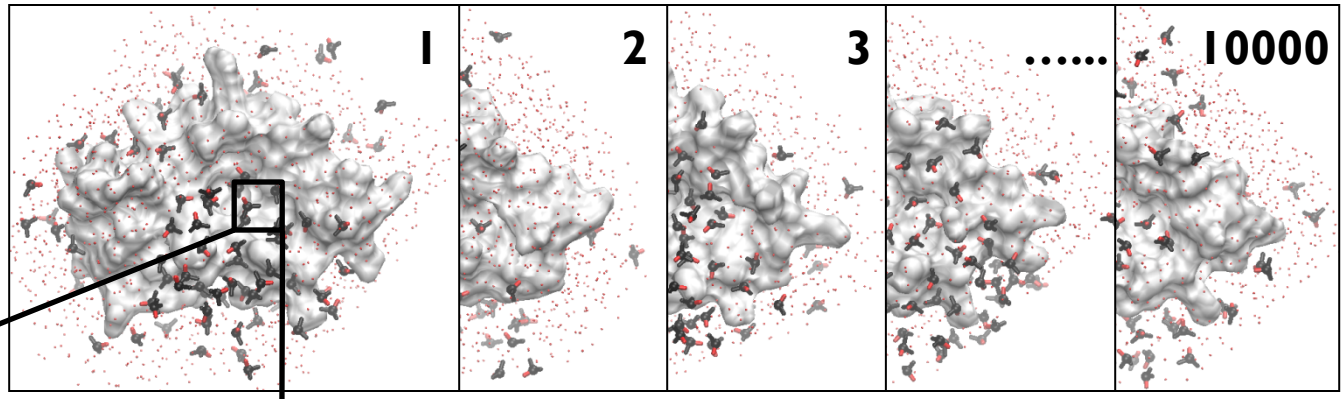
Annealing, Equilibration, Simulation

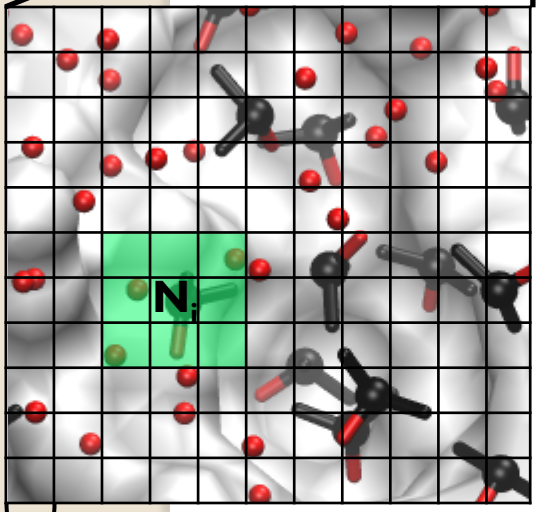


NAMD2 with CHARMM force field was used for simulations.

Free Energy of Binding for Isopropanol

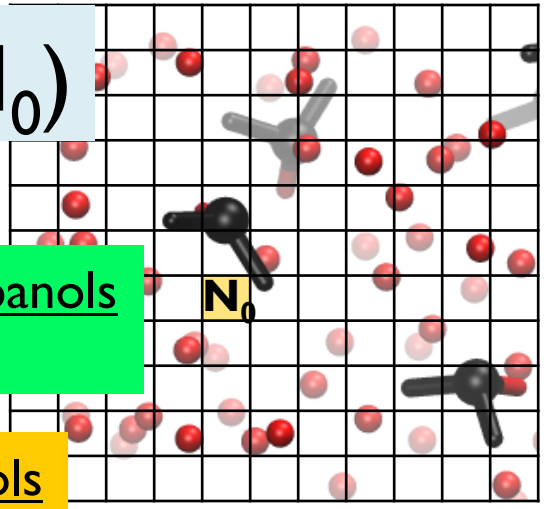
Assuming that MD sampling converged to a **Boltzmann ensemble**





0.5 Å (not to scale)

$$\Delta G_i = -RT \ln(N_i/N_0)$$



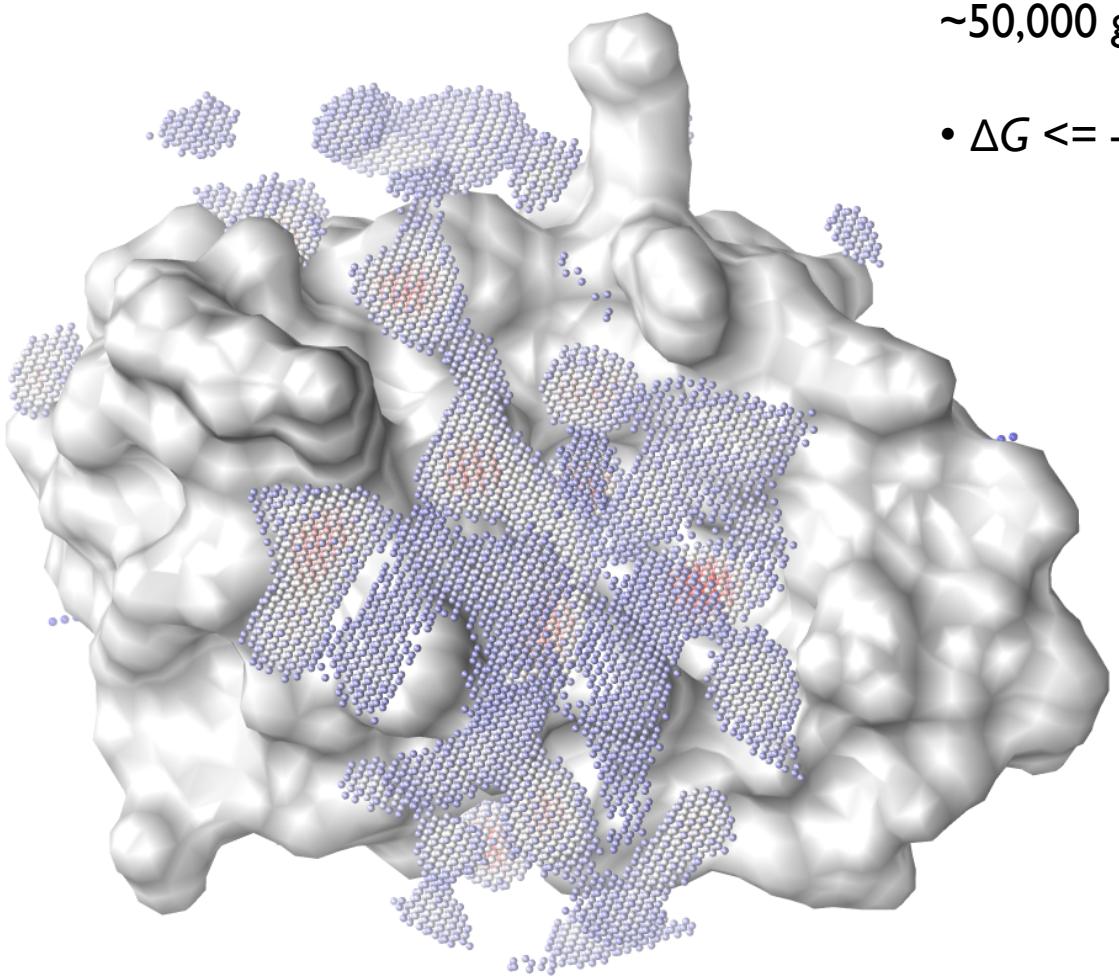
N_i = observed number of isopropanols
(# of frames) * (# of cubes)

N_0 = total number of isopropanols
total number of frames

N_i corresponds to the central highlighted grid element;
number of cubes is introduced if multiple cubes are occupied by a single isopropanol

Isopropanol Binding Spots

ΔG grid is mapped onto the protein structure



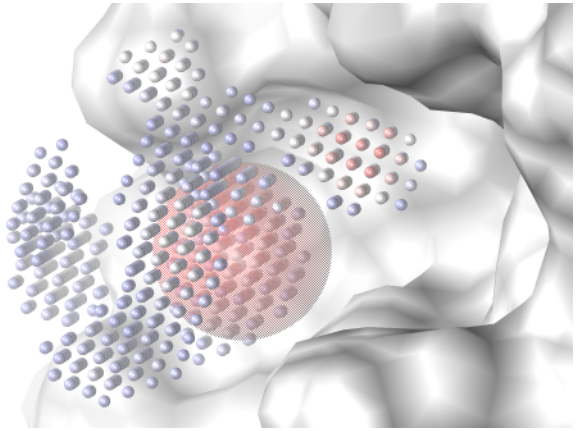
~50,000 grid points with

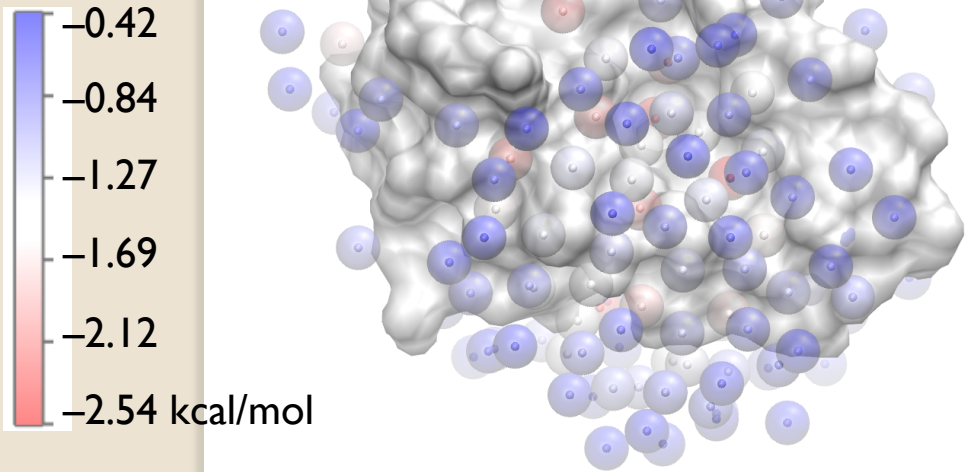
- ΔG <= – 0.416 kcal/mol*
  (2 fold enrichment)

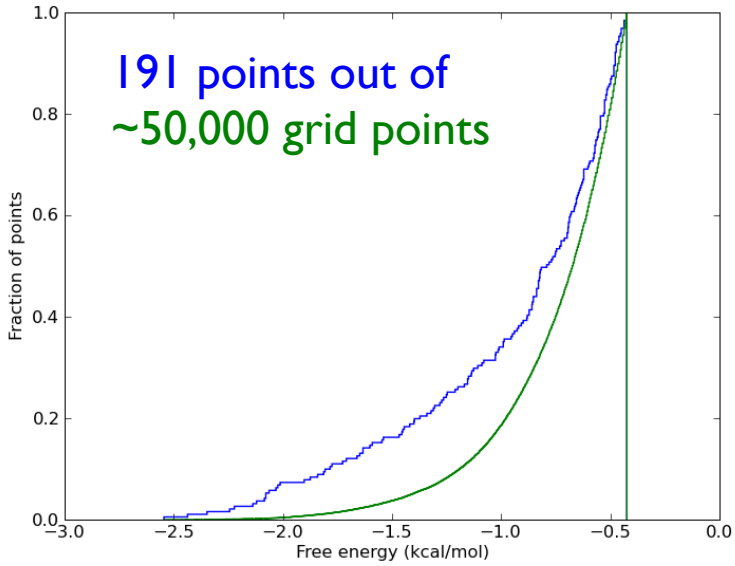
Cold-spot
–0.42
–0.84
–1.27
–1.69
–2.12
–2.54 kcal/mol
Warm
Hot-spot

Selecting Isopropanol Binding Spots

1. Grid element with lowest ΔG value is selected
2. Other elements within **4 Å** are removed (elements inside the red sphere –>)
3. 1 and 2 are repeated until no more points are left to remove

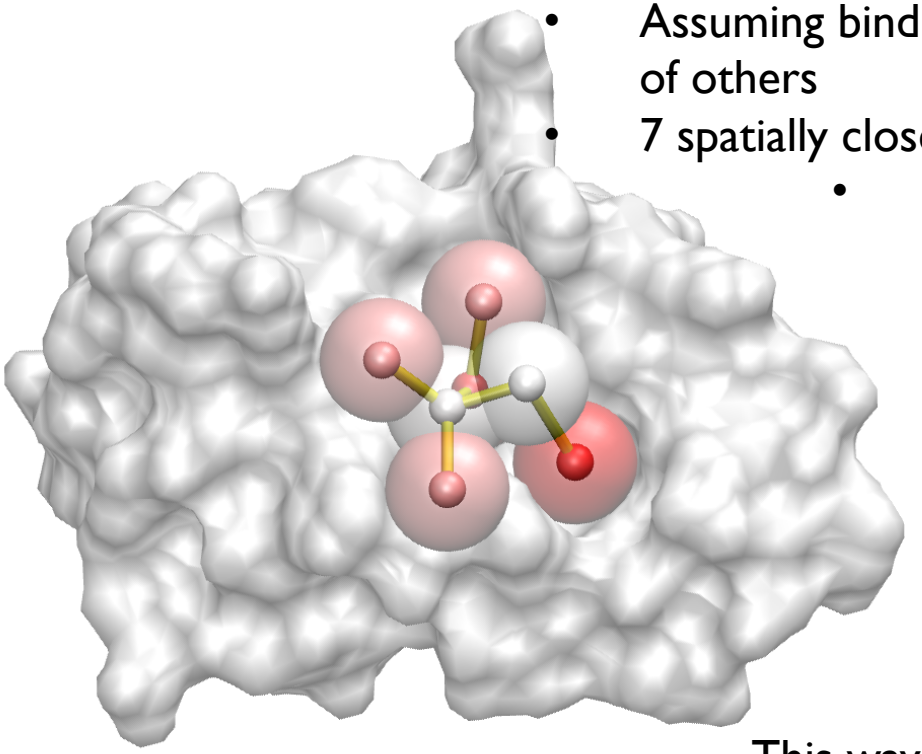






Affinity of a Drug-size Molecule

A heuristic approach for calculating achievable free energy of binding



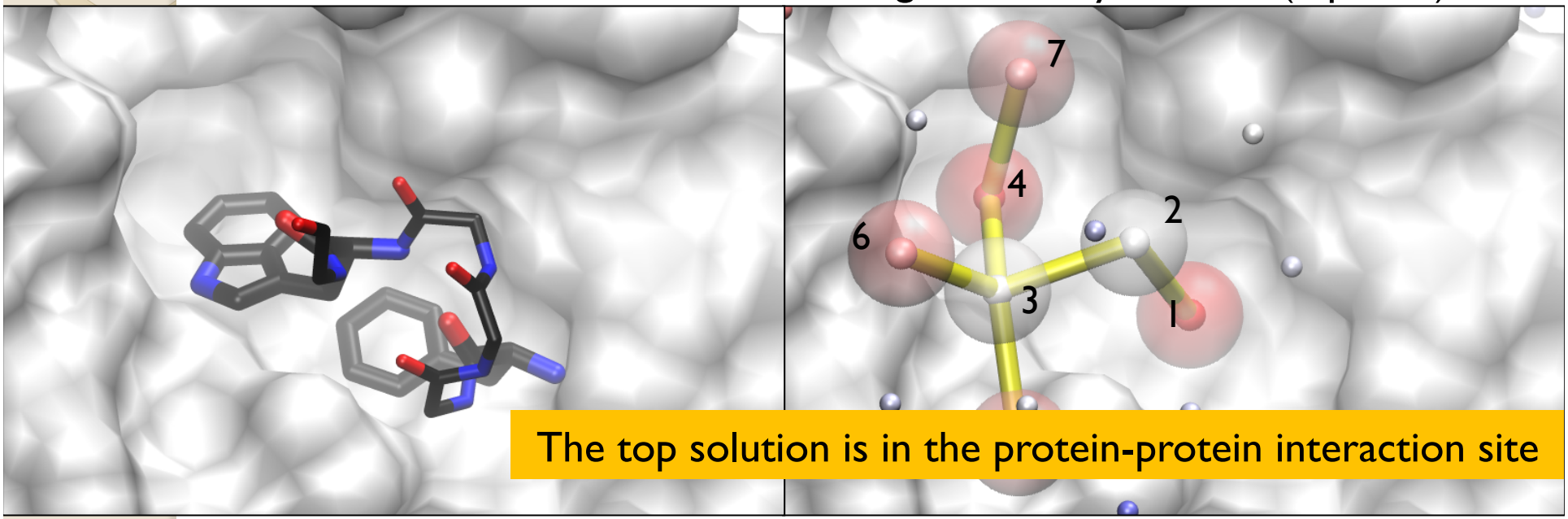
- Assuming binding of an isopropanol is independent of others
- 7 spatially close binding spots are selected
  - The sum of $\Delta G_{binding}$ of individual points is considered as a binding free energy estimate that is achievable by a drug-like molecule

This way, the highest affinity we can observe is 5 fM (10^{-15}).

MDM2: p53 binding site

p53 peptide key interactions (X-ray)

Highest affinity solution (7 points)



The top solution is in the protein-protein interaction site

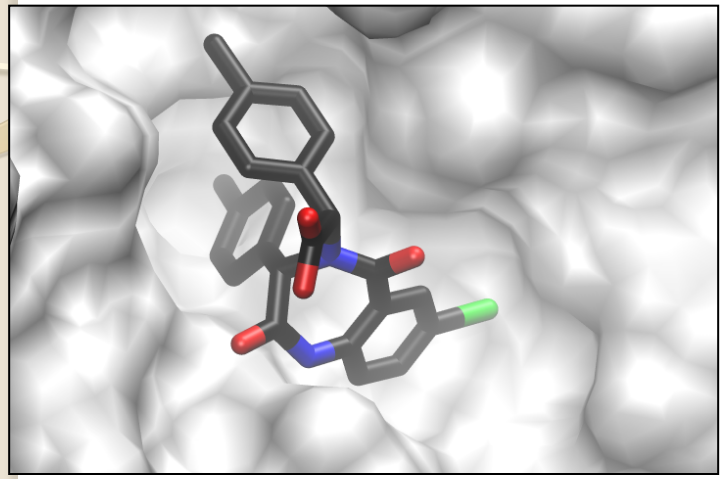
Numbers indicate the order that hot spots were merged by the growing algorithm

Predicted binding affinity range : **0.05-0.3 nM**

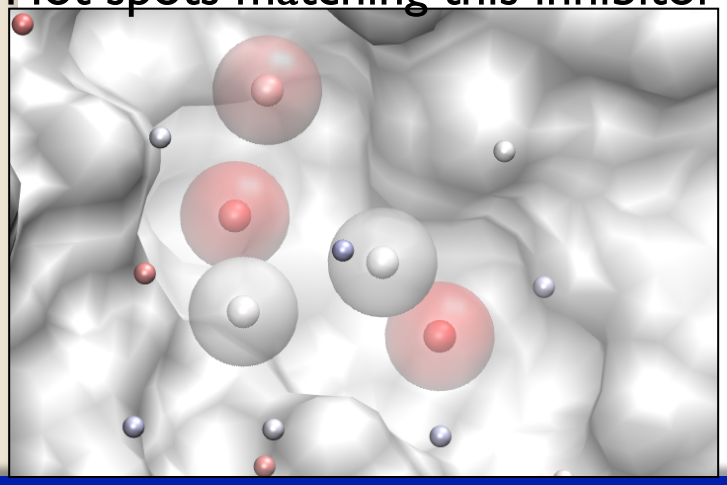
Predicted max. affinity by Seco et al. : **0.02 nM**

MDM2: p53 binding site

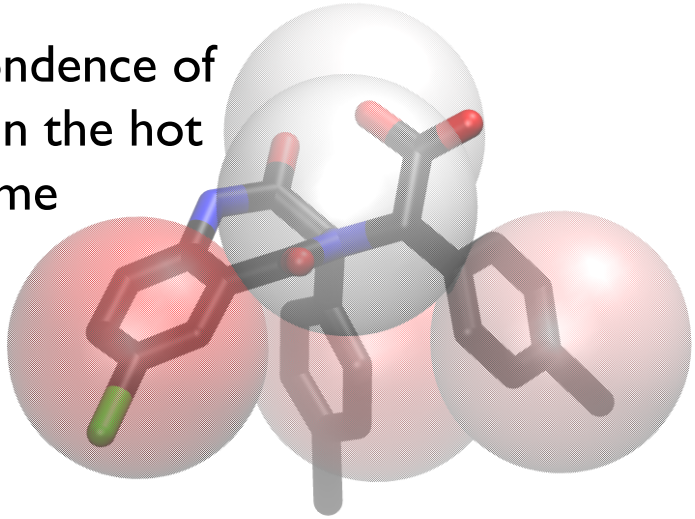
An inhibitor that disrupts p53 binding



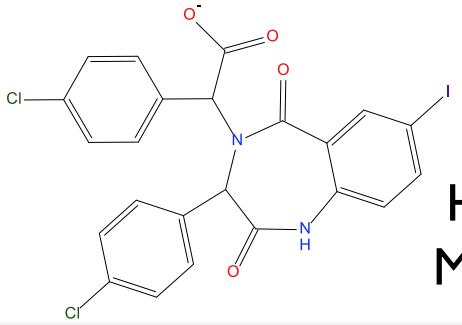
Hot spots matching this inhibitor



Correspondence of inhibitor in the hot spot volume

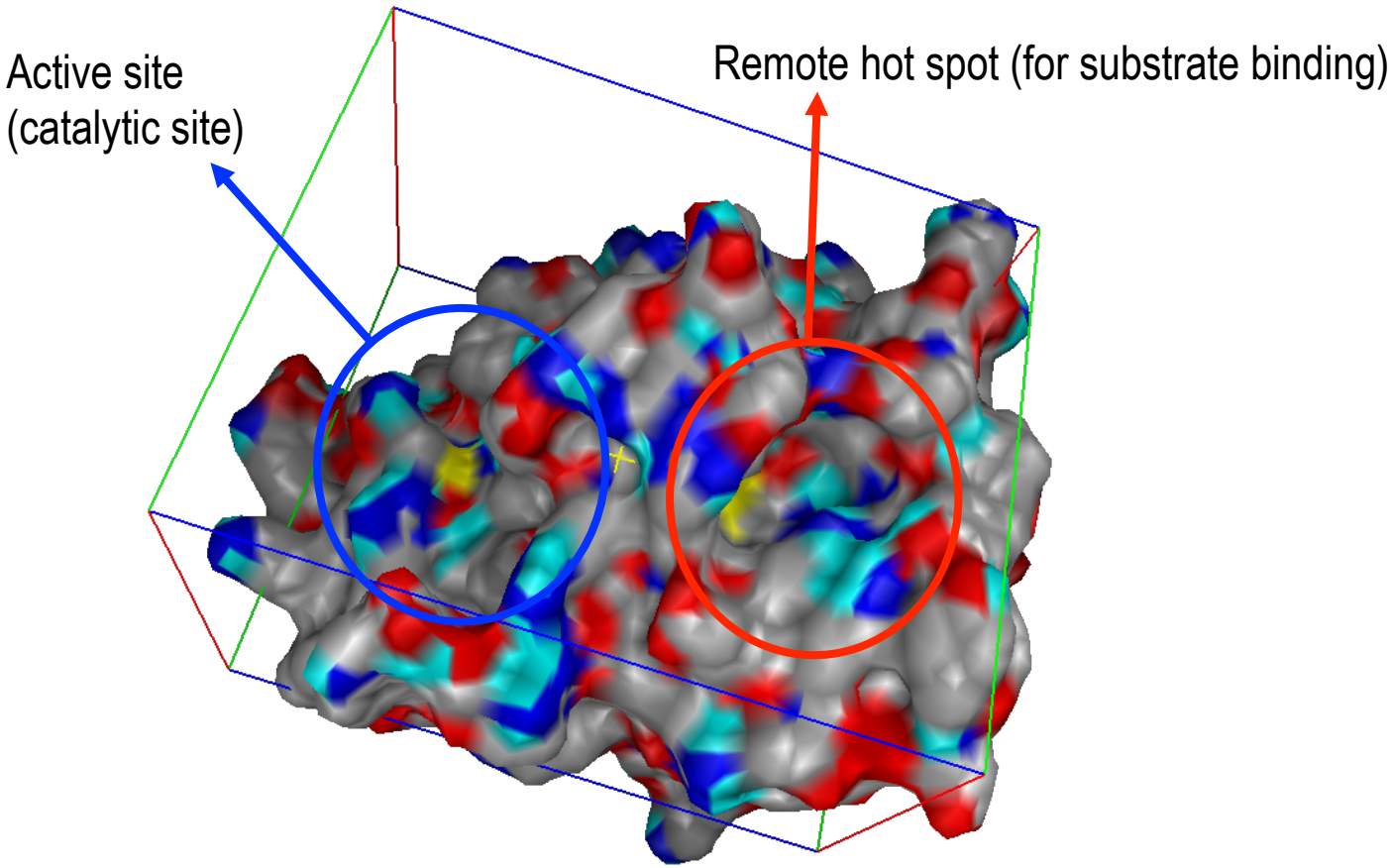


Predicted K_d: **47 nM**
Known K_d : **80 nM**

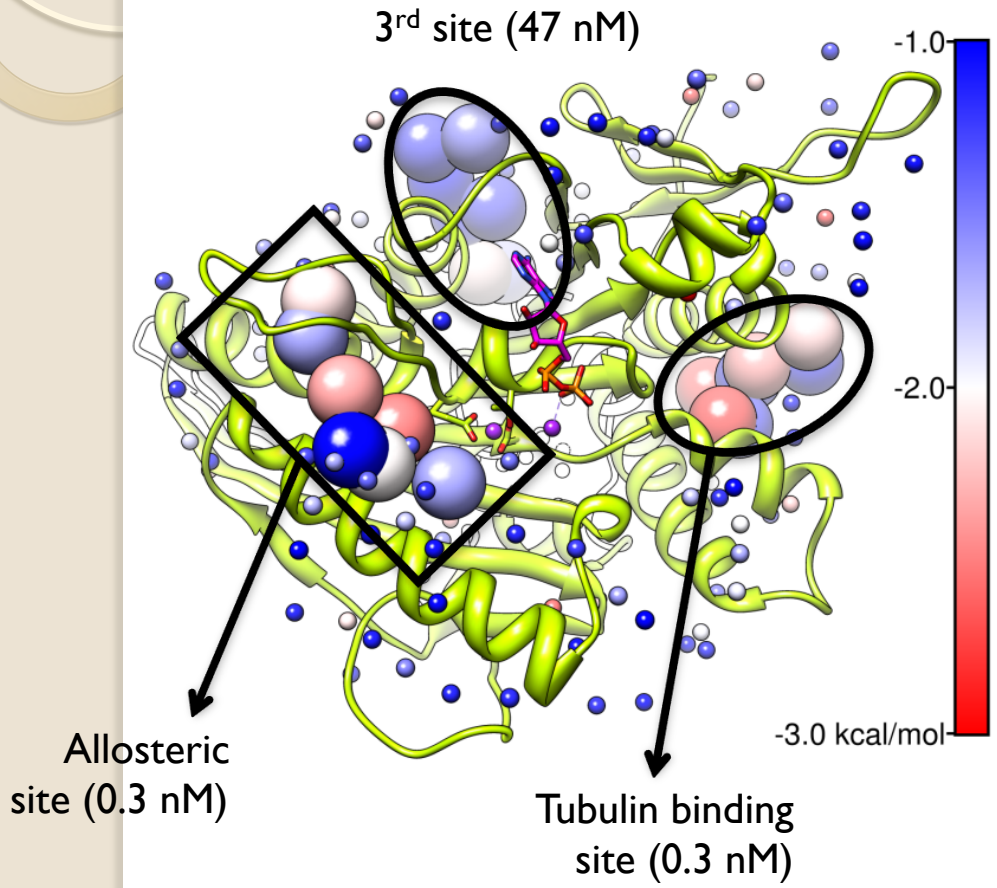


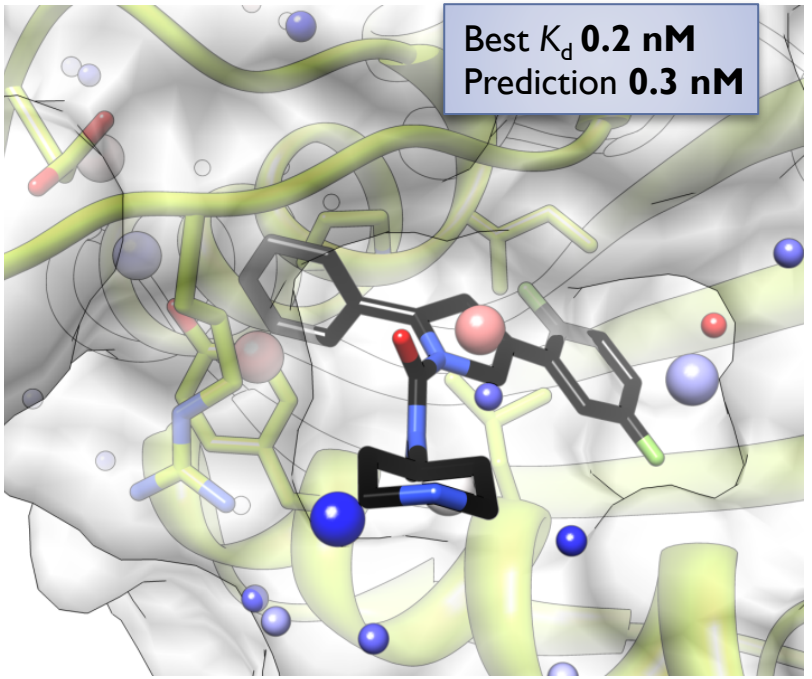
HAC = 32
MW = 580

Proteins may have multiple target sites



eg5 Druggable Sites



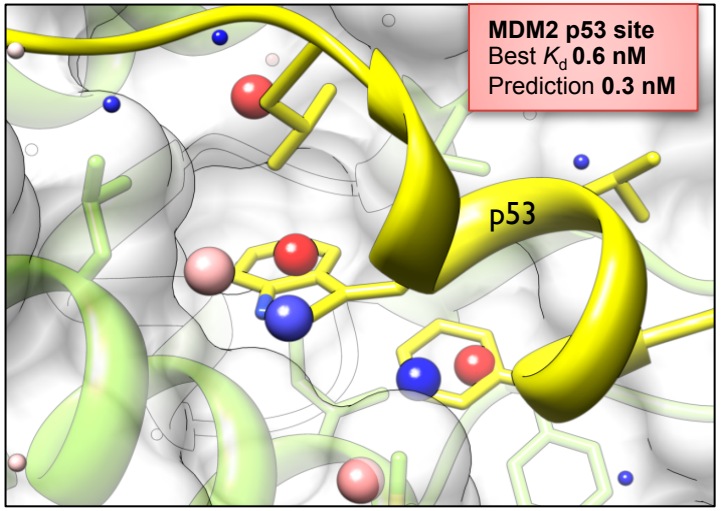


Bioorg. Med. Chem. Lett. **2007,** *17*, 5677-5682

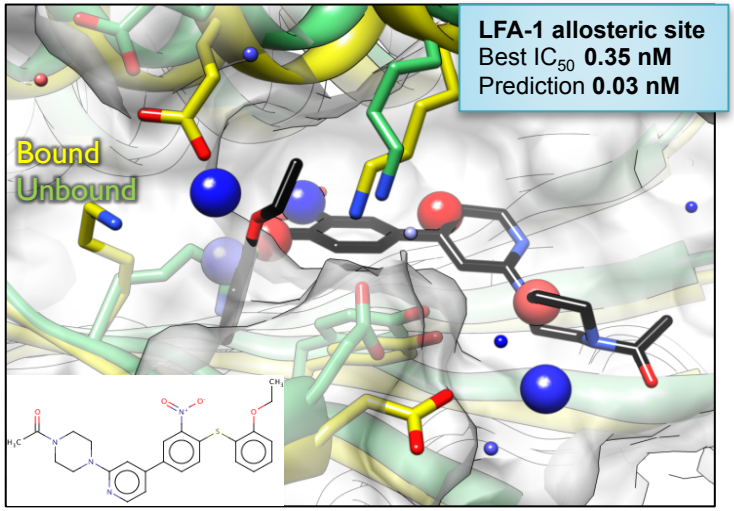
Bakan & Bahar, *J Chem Theory Comput.* 8:2435-2447, 2012

2

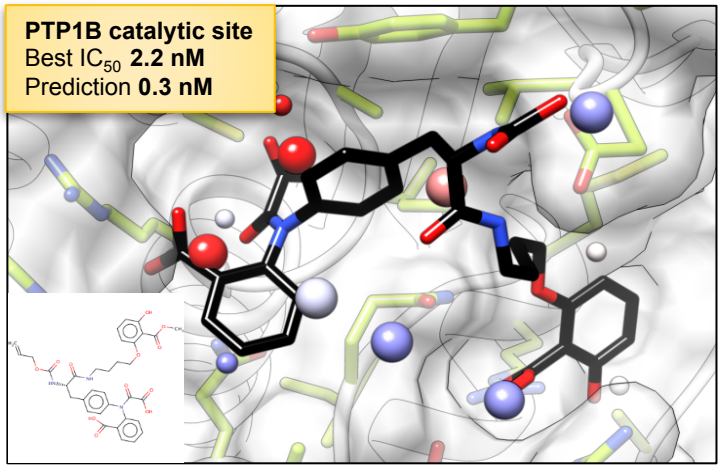
Assessment of druggable allosteric sites



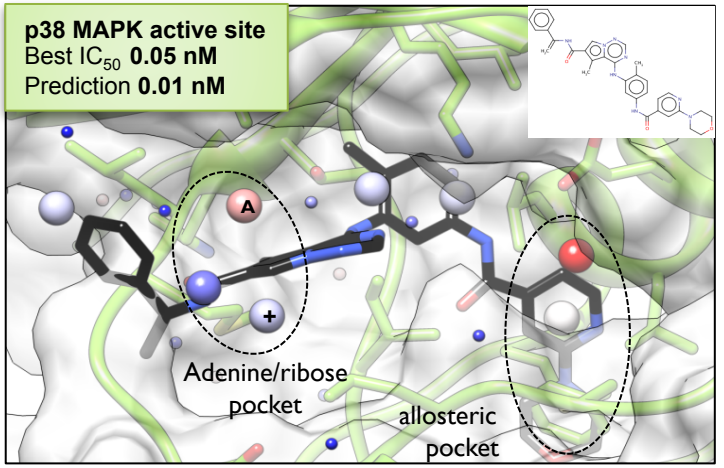
J Med. Chem. **2009,** *52*, 7970-7973



Biochemistry **2004,** *43*, 2394-2404

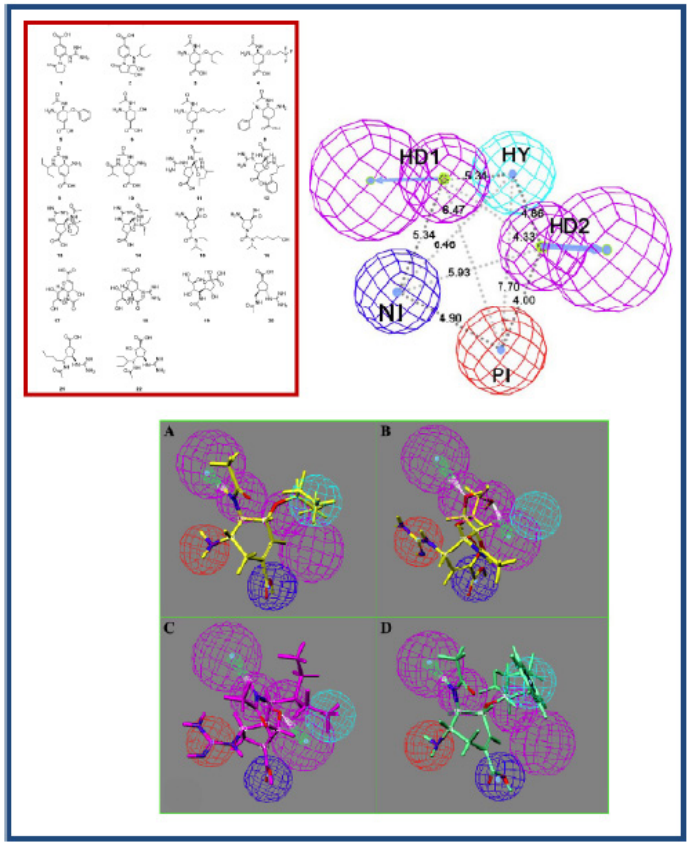


Bioorg. Med. Chem. Lett. **2003,** *13*, 3947-3950



J Med. Chem. **2010,** *53*, 2973-2985

Probe distributions are used for building pharmacophore models



Zhang et al (2006) Bioorganic & Medicinal Chemistry Letters 16, 3009

Pharmacophore Model: PM

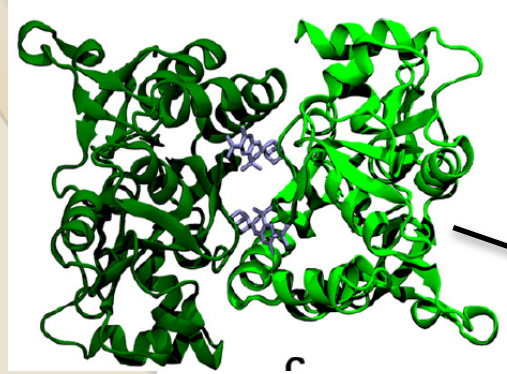
Starting point: a series of hits

Method: clustering, identifying commonalities, assigning weights

Used for screening
- approved drugs
-libraries of small compounds

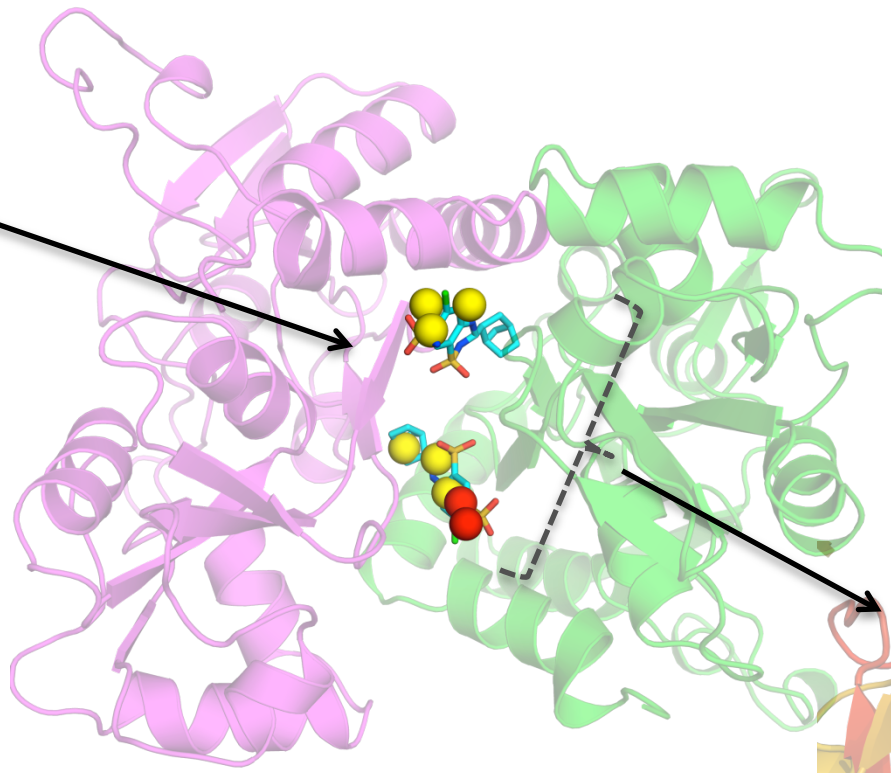
To identify repurposable or new drugs

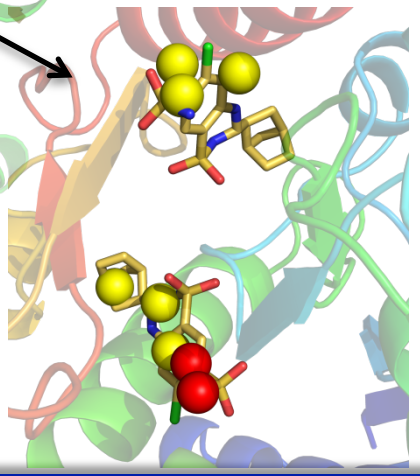
2 Probes capture allosteric modulator site of AMPAR LBD Dimer



Experimental Results

Pohlsgarrd et al (2011).
Neuropharmacology. 60,135-150.

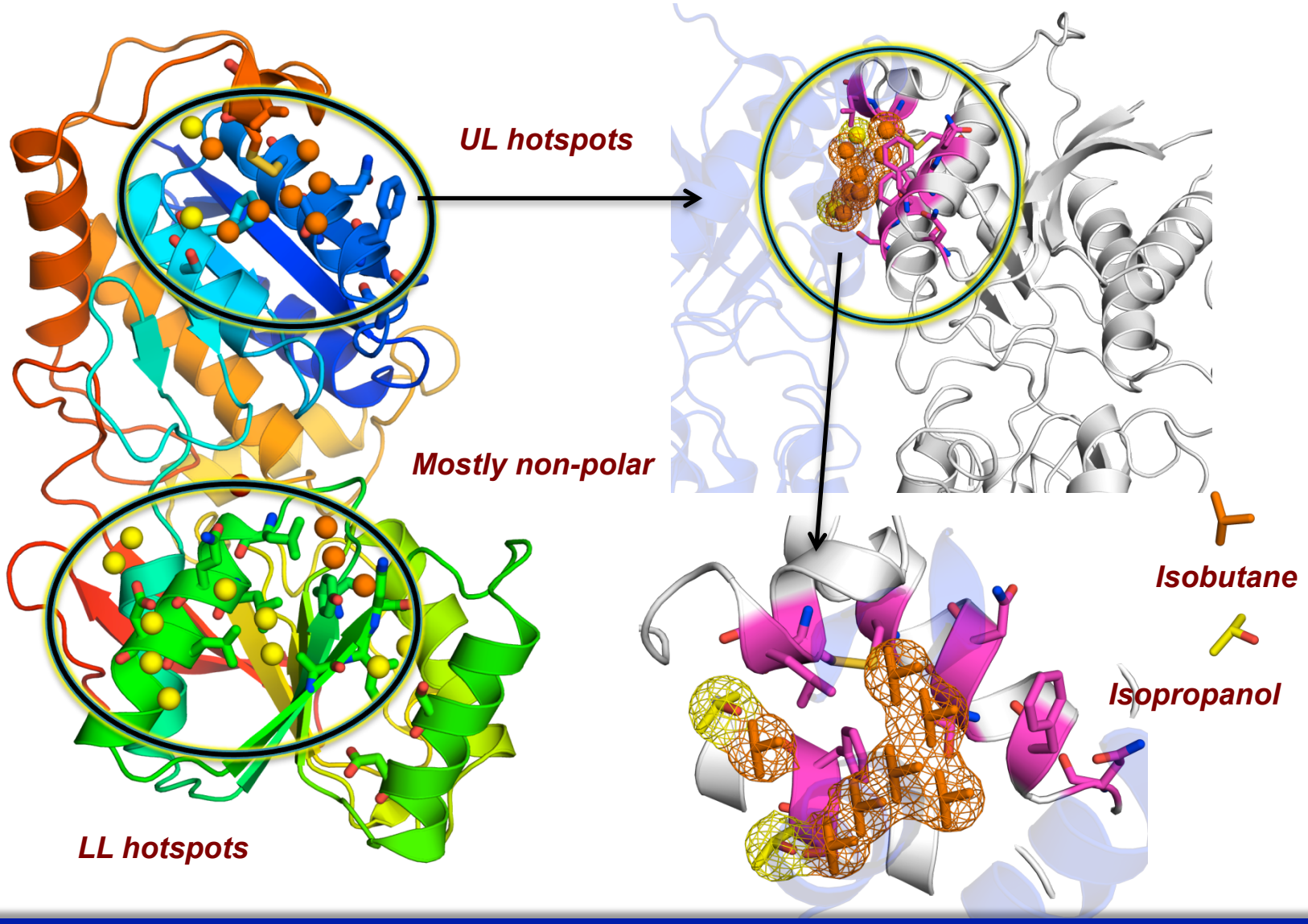




Computational modeling detects experimentally observed binding site

Dutta, Greger, Bahar, manuscript in preparation.

2 Interfacial regions captured in AMPAR NTD



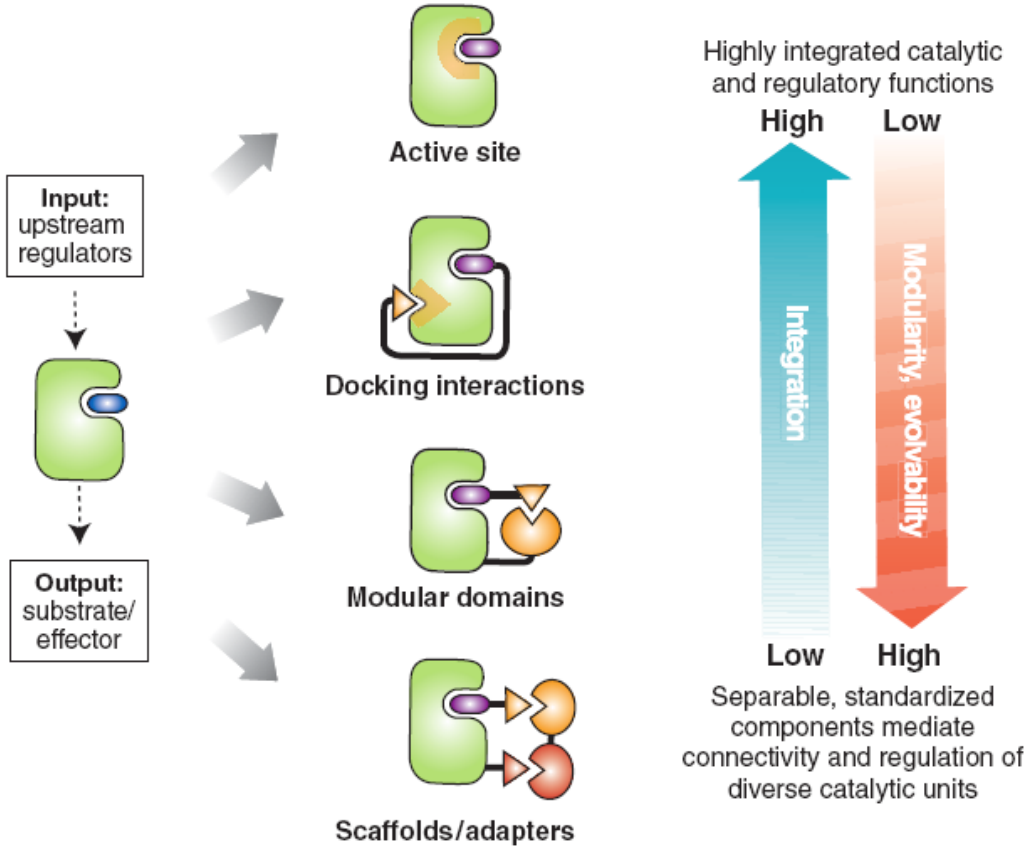
Dutta, Greger, Bahar, manuscript in preparation.

SUMMARY

1. Structure-encoded **flexibility** of drug targets and significance in drug discovery and design
2. **Druggability** assessment: a first step before selecting a target
3. Modularity and **promiscuity** of proteins and **quantitative systems pharmacology** methods

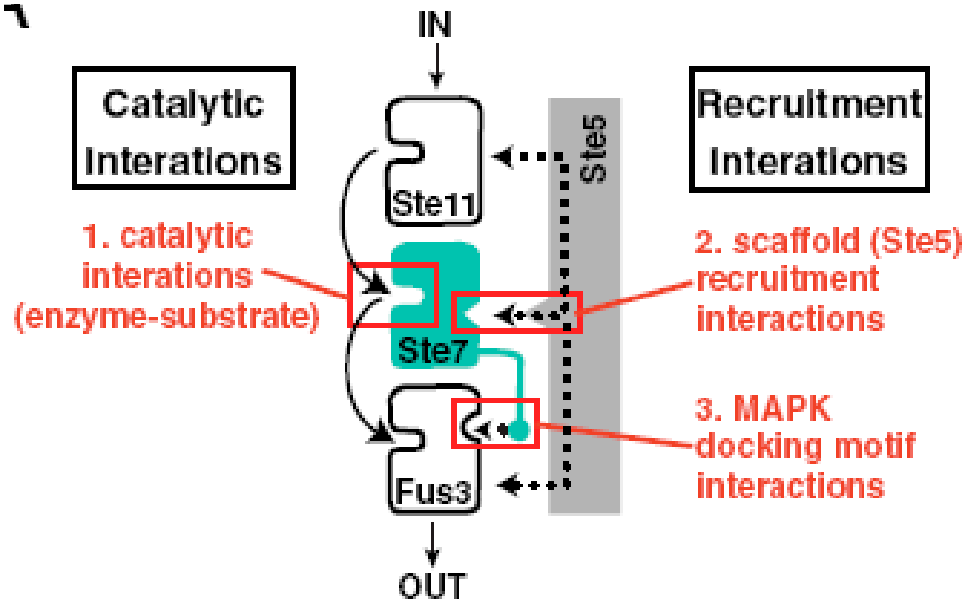
3

Diversity & complexity of phenotypes arise from **combinations of proteins & modular domains**



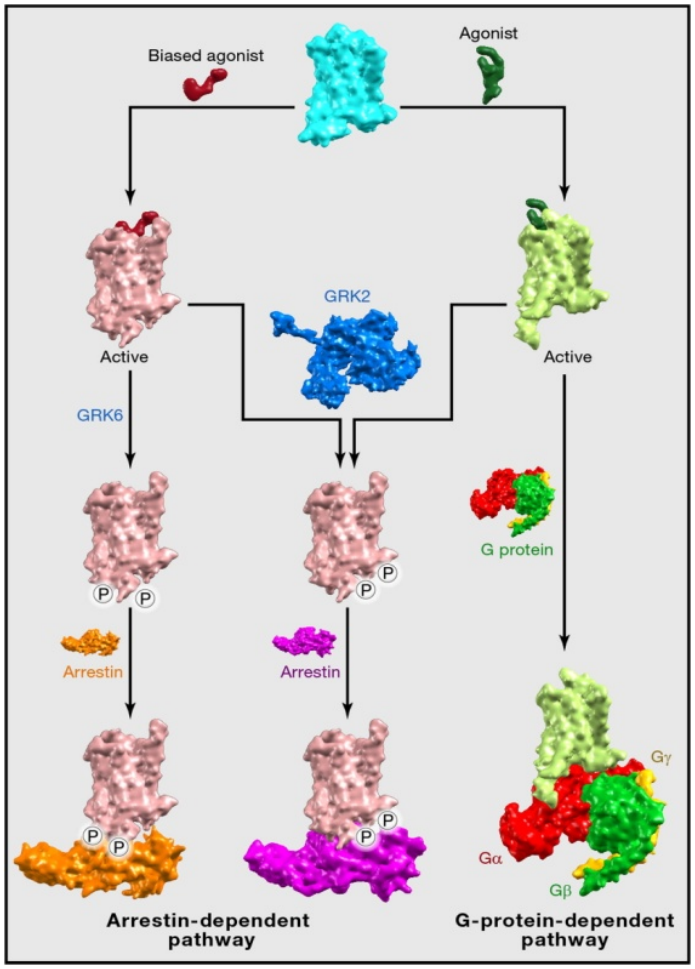
Bhattacharyya et al. *Annual Rev Biochem* 2006

Significance of targeting a specific site, not only a target protein



3

Allostery Can Diversify Cellular Signaling Pathways through a Single Receptor



GPCRs use **conformational selection** to shape signaling.

Two (different) conformations of GPCR bind two (different) agonists, which branch into two pathways

Nussinov & Tsai (2013) "Allostery in Disease and Drug Discovery" *Cell* **153**, 293-305.

Protein Promiscuity

Many proteins are involved in multiple pathways.

Depending on the targeted **surface** region, or on the accessible **structural change/dynamics**

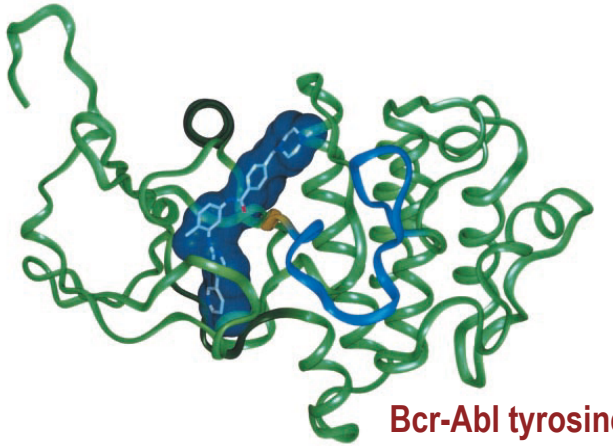
the interactions with different (or multiple) upstream or downstream partners/substrates may be affected,

which in turn would impact different (or multiple) pathways, and may result in various phenotypes

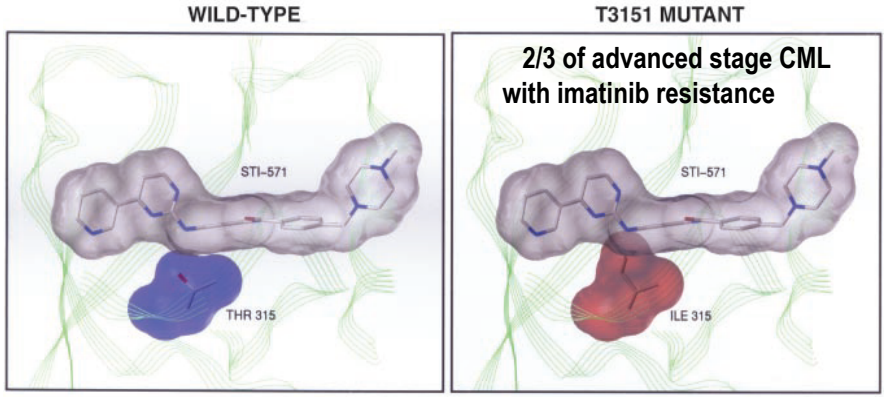
Assessment of druggable allosteric sites

Imatinib (Gleevec)

Imatinib was developed for chronic myelogenous leukemia (CML), but was also used for gastrointestinal stromal tumors (GISTs) and some other diseases.



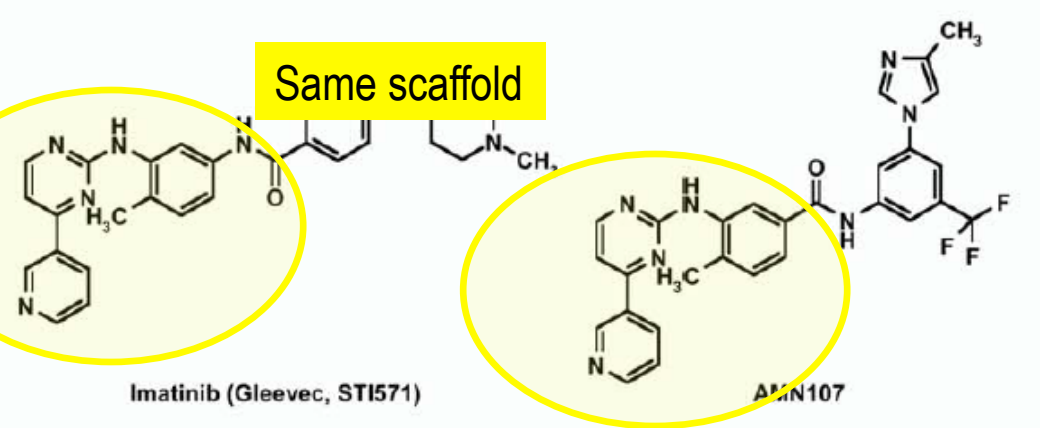
Bcr-Abl tyrosine kinase



IC_{50} ~200 nM

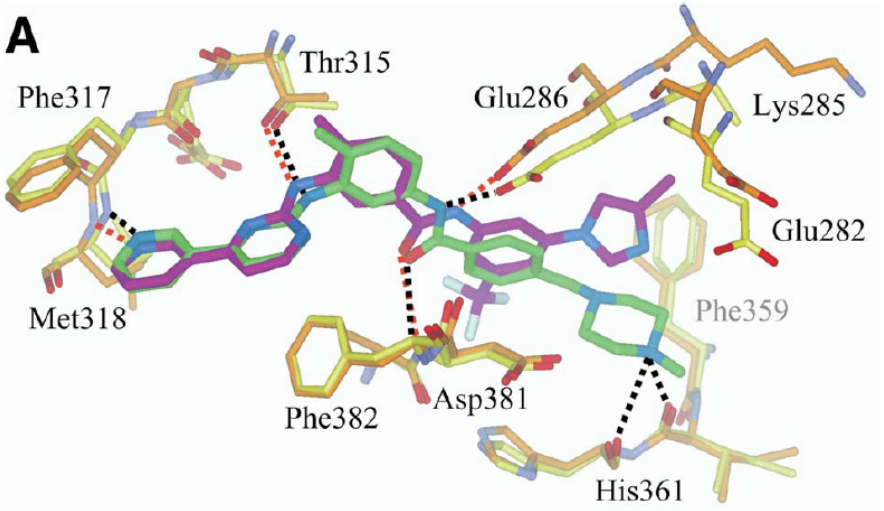
IC_{50} > 10,000 nM

Imatinib vs Nilotinib



Novartis, approved in 2001

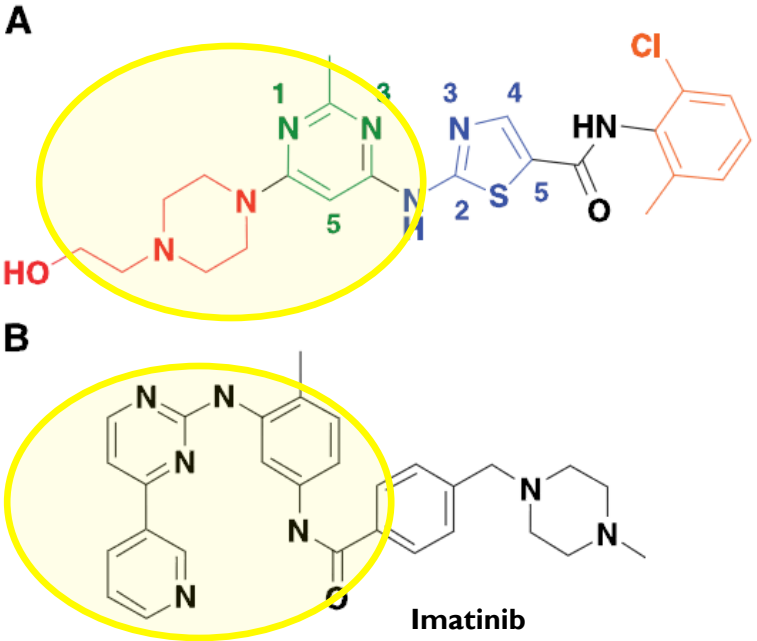
Novartis, approved in 2003

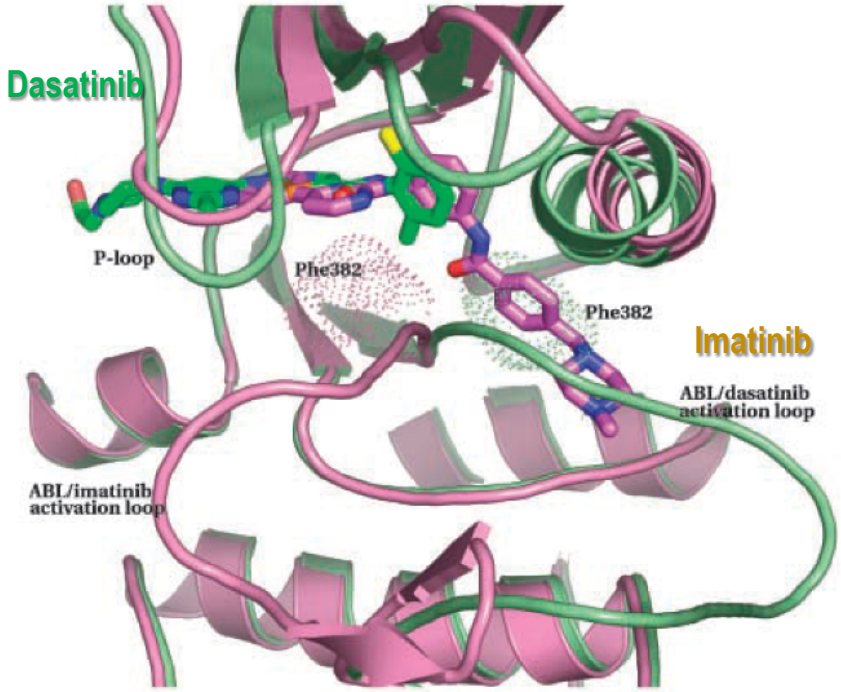


Cancer Cell. 2005, 7:129-41.

Dasatinib addresses imatinib resistance mutations, but fails with mutant T315I

Dasatinib
Bristol Myers Squibb, approved in 2011





Scaffold hopping via pharmacophore modeling

Cancer Res. 2006 , 66: 5790-7.

5

Vol 463|28 January 2010|doi:10.1038/nature08675 nature

ARTICLES

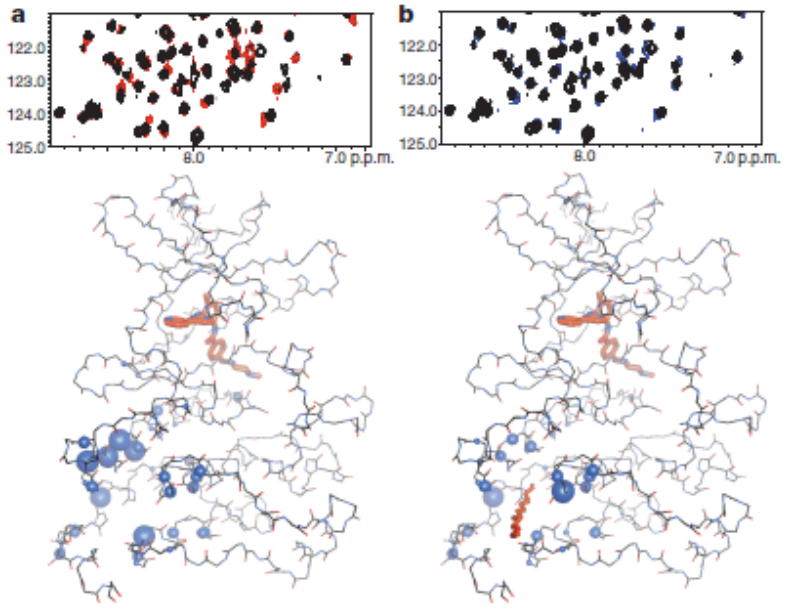
Targeting Bcr–Abl by combining allosteric with ATP-binding-site inhibitors

Jianming Zhang[1]*, Francisco J. Adrián[2]*, Wolfgang Jahnke[3], Sandra W. Cowan-Jacob[3], Allen G. Li[2], Roxana E. Iacob[4], Taebo Sim[1,5], John Powers[6], Christine Dierks[2], Fangxian Sun[2], Gui-Rong Guo[2], Qiang Ding[2], Barun Okram[7], Yongmun Choi[1], Amy Wojciechowski[1], Xianming Deng[1], Guoxun Liu[2], Gabriele Fendrich[3], André Strauss[3], Navratna Vajpai[8], Stephan Grzesiek[8], Tove Tuntland[2], Yi Liu[2], Badry Bursulaya[2], Mohammad Azam[6], Paul W. Manley[3], John R. Engen[4], George Q. Daley[6], Markus Warmuth[9] & Nathanael S. Gray[1]

GNF-2 binds to the myristate-binding site of Abl, leads to changes in the structural dynamics of the protein, and thus inhibits allosteric interactions!

5

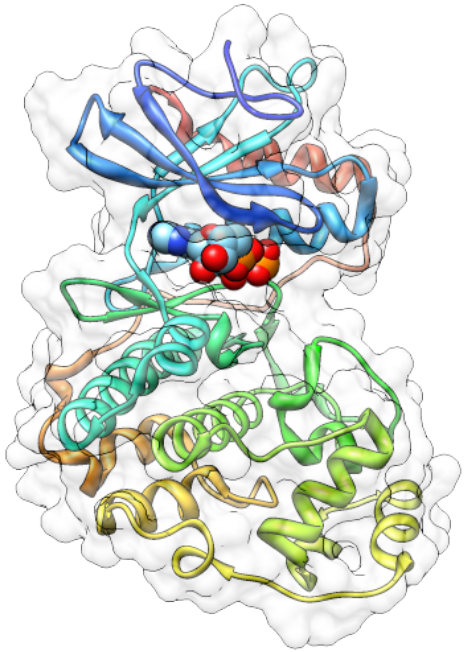
Polypharmacological strategy: Inhibition of allosteric interaction site in addition to catalytic site



Evidence for GNF-2 binding to the myristate pocket of Abl. HSQC spectrum of Abl/imatinib with (red) and without (black) GNF-2 (top) shows chemical shift changes induced by ligand binding. Mapping of chemical shift changes to structure (PDB 1OPK8) identifies the myristate pocket as the GNF-2 binding site. **b**, Same as **a** except myristic acid used instead of GNF-2.

Simultaneously targeting of

- the ATP binding site (by Gleevec)
- the myristate pocket (by GNF-2)

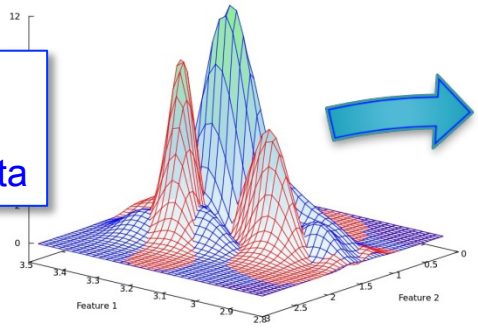


Khateb et al. BMC Cancer. 2012 Overcoming Bcr-Abl T315I mutation by combination of GNF-2 and ATP competitors in an Abl-independent mechanism.

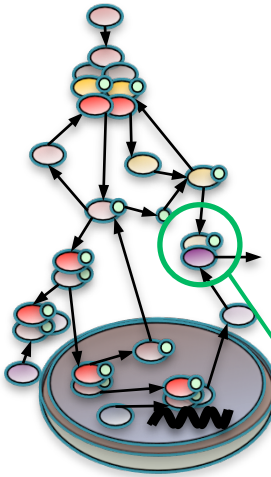
Quantitative Systems Pharmacology:

Integrating Quantitative Models with Experimental Data for Drug Discovery

Signals extracted from the data

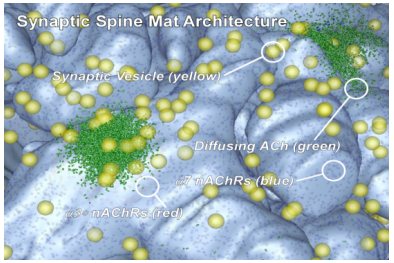


Information

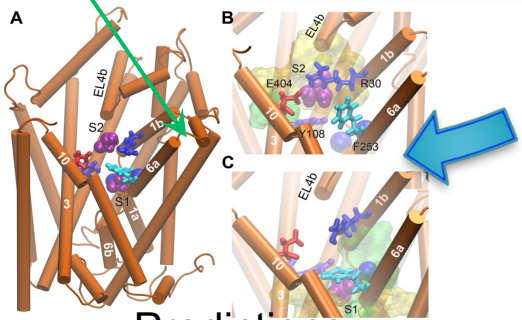


Refined cellular pathways /processes

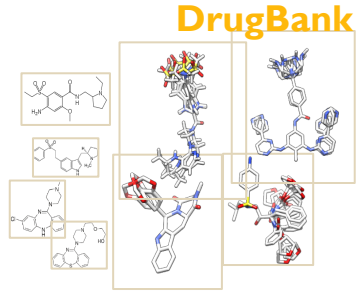
Models

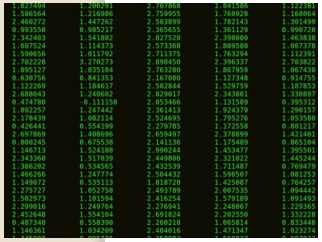


Predicted Drug binding to targets



Predictions

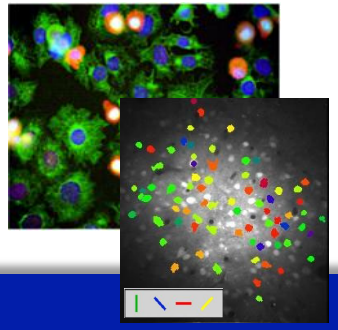




Data

High-Content Scre
Targeted Libraries

Measurements of biological functions



3D or 4D images

Experimental testing

Clinical Trials

Nobeli et al., Nature Biotechnology (2009)

A Role for Computational Biology

Box 2 A role for computational biology?

Predicting protein promiscuity is a problem of daunting complexity for bioinformaticians. Indeed, earlier work has shown that bioinformatics methods need improving to reliably uncover promiscuous reactions[117], and our own *in silico* work in protein function prediction[118,119] has curbed our optimism. Even so, we do not doubt that certain areas of bioinformatics research will be important for progress in this field. Some of these avenues have already been pursued by computational biologists, but referencing individual studies is outside the scope of this review.

Data analysis. There is a great deal of data on protein promiscuity to be found in function-related databases. Collating information on promiscuous proteins would be a necessary first step, and existing enzyme, pathway or ontology databases can provide a lot of information on proteins with multiple EC numbers, reactions or substrates, or function categories for genes. Clues as to moonlighting might also be found from expression data. Unexpected expression patterns that do not correlate with our knowledge of existing networks are often indicative of a moonlighting function for a protein. Such data would be complemented by data on alternative splicing of a single gene, which will give hints to any additional roles in the cell.

Sequence-based methods. Large protein families with many relatives may indicate a trend toward promiscuity. Is there a correlation between number of orthologs and number of paralogs and how could it be explained?

Structure-based methods. Analyzing binding site characteristics could reveal those that make proteins more amenable to promiscuity.

Docking profiles. Probing the binding site with panels of selected ligands or other proteins can assess how restrictive the site is toward different types of molecules.

Flexibility. *In silico* studies of the flexibility of proteins can reveal how this may contribute to recognizing multiple partners.

Redundancy in pathways. The evidence of redundancy in metabolic and regulatory networks should be examined carefully, as it may also contain evidence for protein functional promiscuity.

Calculation of promiscuity indices. This could be based on *in silico* or experimental data and could help rank proteins and their partners according to their interaction promiscuity.

Mapping of small-molecule space to protein space. This would reveal any preferences of protein families for sets of chemical groups and possibly allow the engineering of mutatants capable of binding small molecules from neighboring parts of the chemical space.

These are only some possible directions that could be explored to improve our chances of successfully exploiting promiscuity. Experimental verification of any rules learned and predictions made will be indispensable.