Day 2 – Lecture 2

Allostery and Druggability

Burak Kaynak and Pemra Doruker

Department of Computational and Systems Biology
School of Medicine, University of Pittsburgh, Pittsburgh, PA 15260

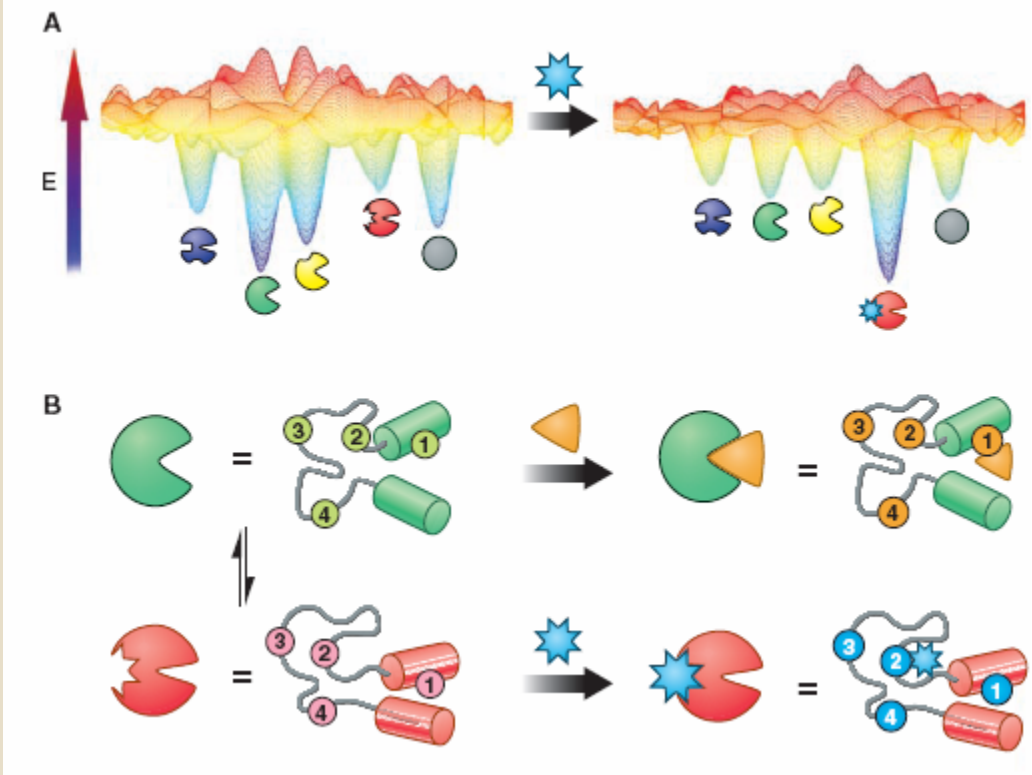
Summary

1. Theory
   a. Gaussian Network Model (GNM)
   b. Anisotropic Network Model (ANM)
   c. Resources/Servers/Databases (ProDy, DynOmics)

2. Bridging Sequence, Structure and Function
   a. Ensemble analysis and functional modes of motion
   b. Combining sequence and structure analyses – signature dynamics
   c. Modeling membrane proteins and lipid environment with ANM

3. Allostery and druggability
   a. Essential site scanning and allosteric pocket prediction
   b. Druggability simulations

Allosteric regulation and drug design



- green: catalytically active state
- others: inactive states
- blue star: allosteric inhibitor
- orange triangle: substrate

- A portion of Free-energy landscape around the native state sampled by the protein
- Protein samples multiple pre-existing conformational states
- Global minimum shifts to favor inactive conformation when an inhibitor binds to an *allosteric site*.
- Current approach in structure-based drug-design

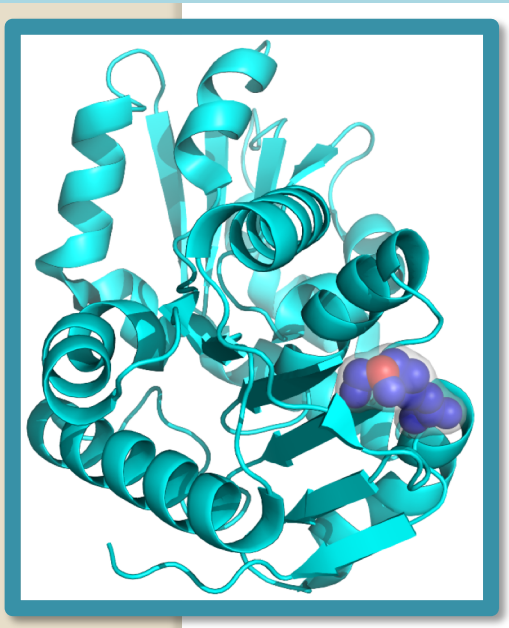
Lee and Craik, Trapping moving targets with small molecules, Science, 2009.

ESSA: Essential site scanning analysis & Allosteric pocket prediction

Kaynak, Bahar, Doruker, (2020) “Essential site scanning analysis: A new approach for detecting sites that modulate the dispersion of protein global motions” Comput. Struct. Biotechnol.

Essential Site Scanning Analysis

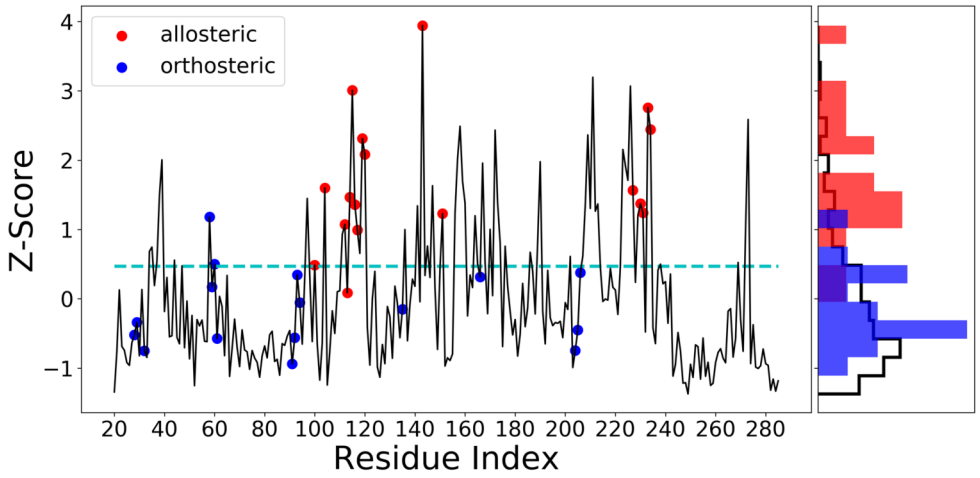
Crowding near each residue



$\langle \Delta\lambda_{1-10}^{(i)} \rangle$

Induced change in global mode dispersion

ESSA profile



C^{α}, other heavy atoms (treated as environment in reduced model technique†)

$$z_i = \frac{\langle \Delta\lambda_{1-10}^{(i)} \rangle - \mu}{\sigma}$$

$\mu : mean$

$\sigma : std$

$$\Delta\lambda_k^{(i)}(\%) = \frac{\left(\lambda_k^{(i)} - \lambda_k\right)}{\lambda_k} \times 100.$$

$k : mode\ index$

$i : residue\ index$

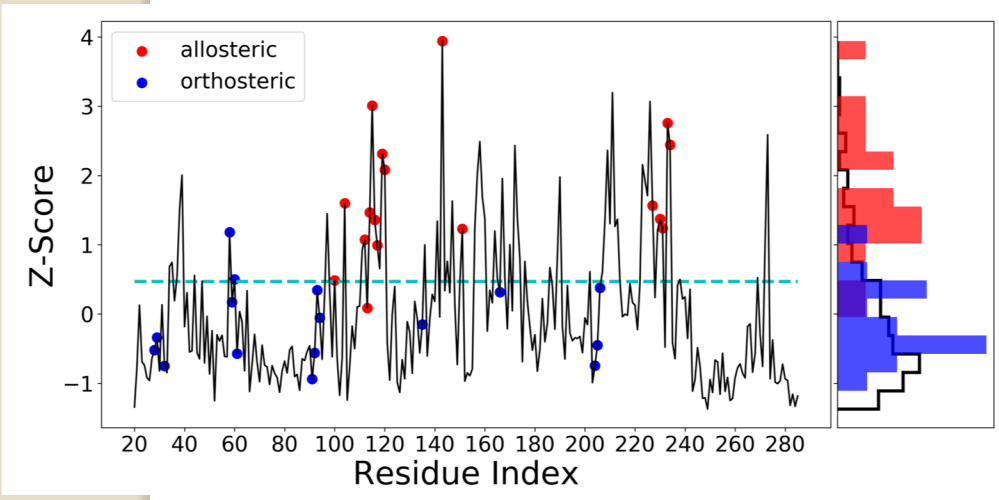
Ligand binding sites of are chosen those whose at least one atom within 4.5 A of a ligand.

† Hinsen et al. *Chem Phys. 261* (1-2) (2000) 25.

Glutamate racemase (PDB id: 2jfn)

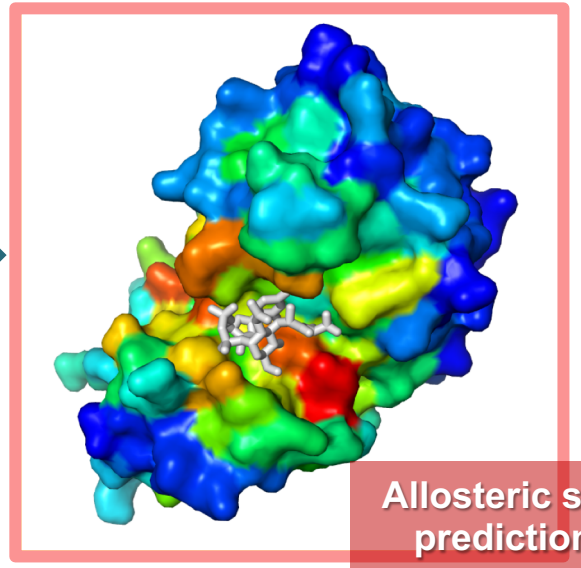
ESSA-based allosteric pocket prediction

ESSA profile



Glutamate racemase (PDB id: 2jfn)

Essential sites



Allosteric site prediction

ESSA ⊕ LHD

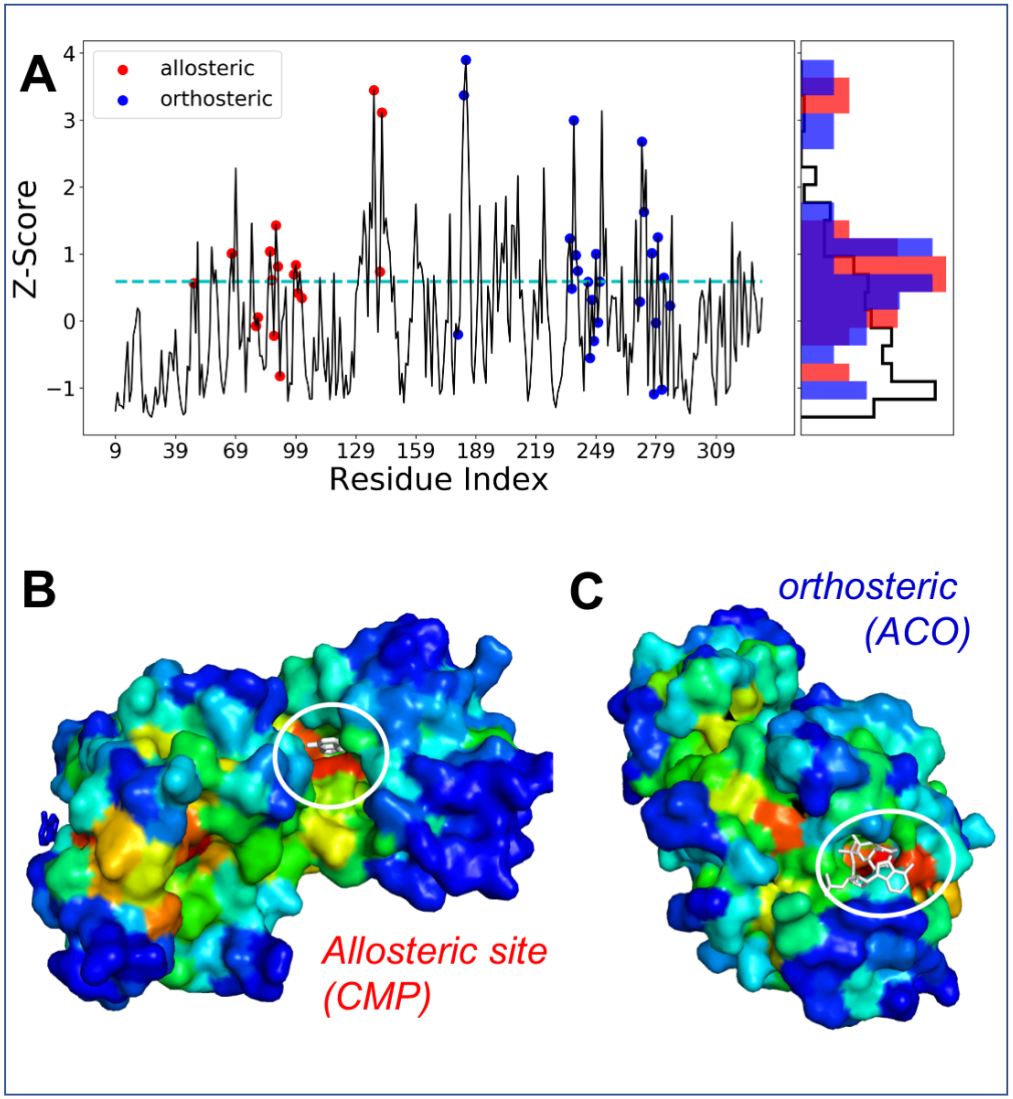
Pocket evaluation by Fpocket‡

Allosteric sites have relatively higher local hydrophobic density (LHD)

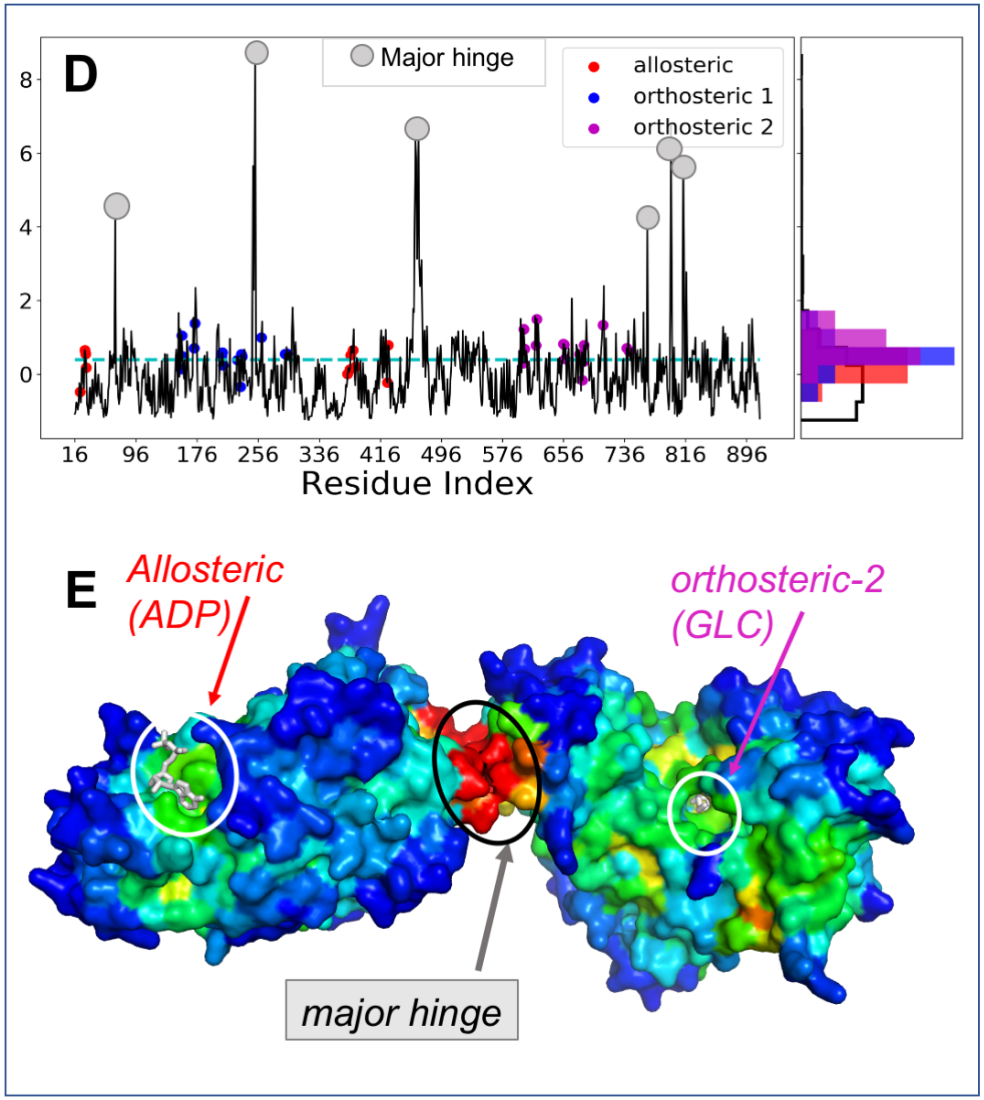
- Song et al. *J Chem Inf Model* 57 (9) (2017) 2358.

‡ Le Guilloux et al. *BMC Bioinformatics* 10 (1) (2009) 168.

Lysine acetyltransferase (4avb)

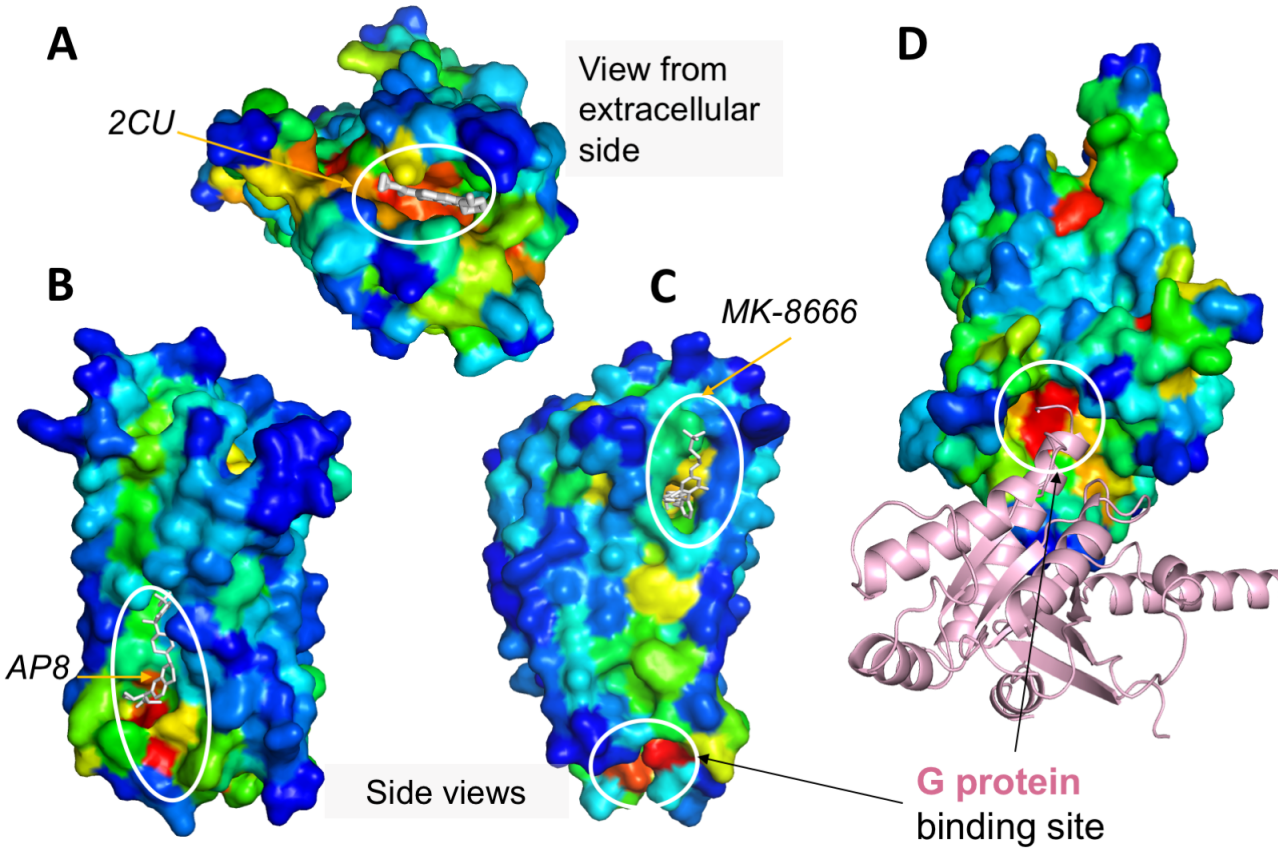


Hexokinase I (1cza)



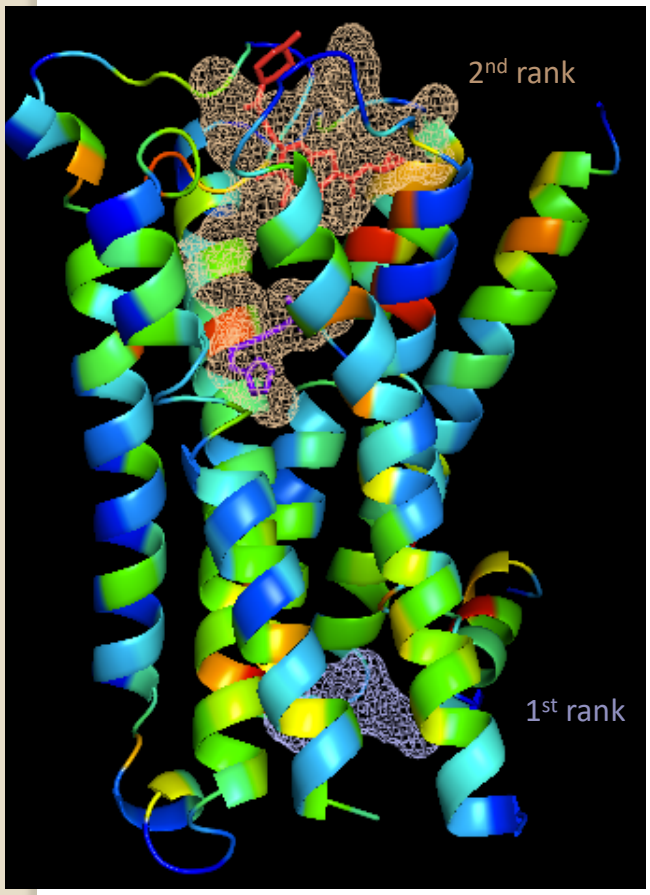
Multiple ligand-binding sides on GPCRs

Muscarinic acetylcholine receptor (4mqt) in A and D



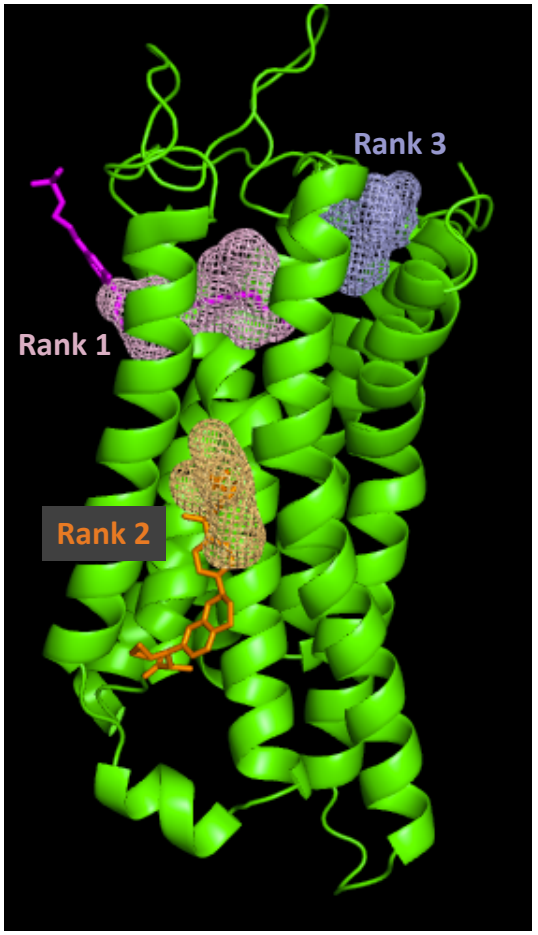
GPR40 (5tzy) in B and C

Muscarinic acetylcholine receptor M2
(Class A, PDB id: 4mqt)



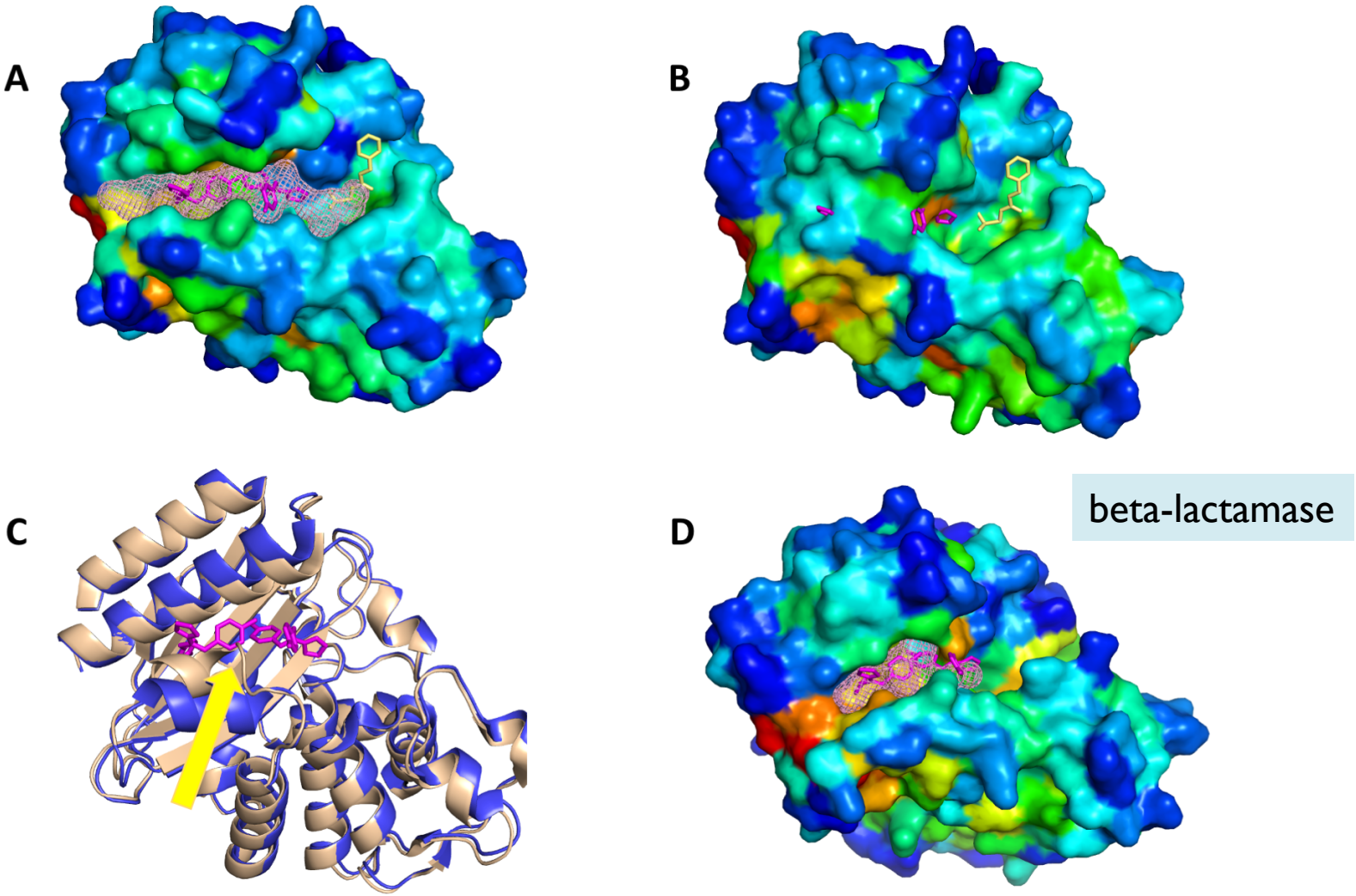
Intrahelical sites
Purple: agonist IXO (ortho site),
Red: PAM 2CU (ligand entry site)

Free fatty acid receptor GPR40
Class A (PDB: 5tzy)



Lipid facing sites
Magenta: agonist MK6
Orange: agoPAM 70S

Cryptic sites by ClustENMD and ESSA



Kaynak, Bahar, Doruker, (2020) “Essential site scanning analysis: A new approach for detecting sites that modulate the dispersion of protein global motions” Comput. Struct. Biotechnol.

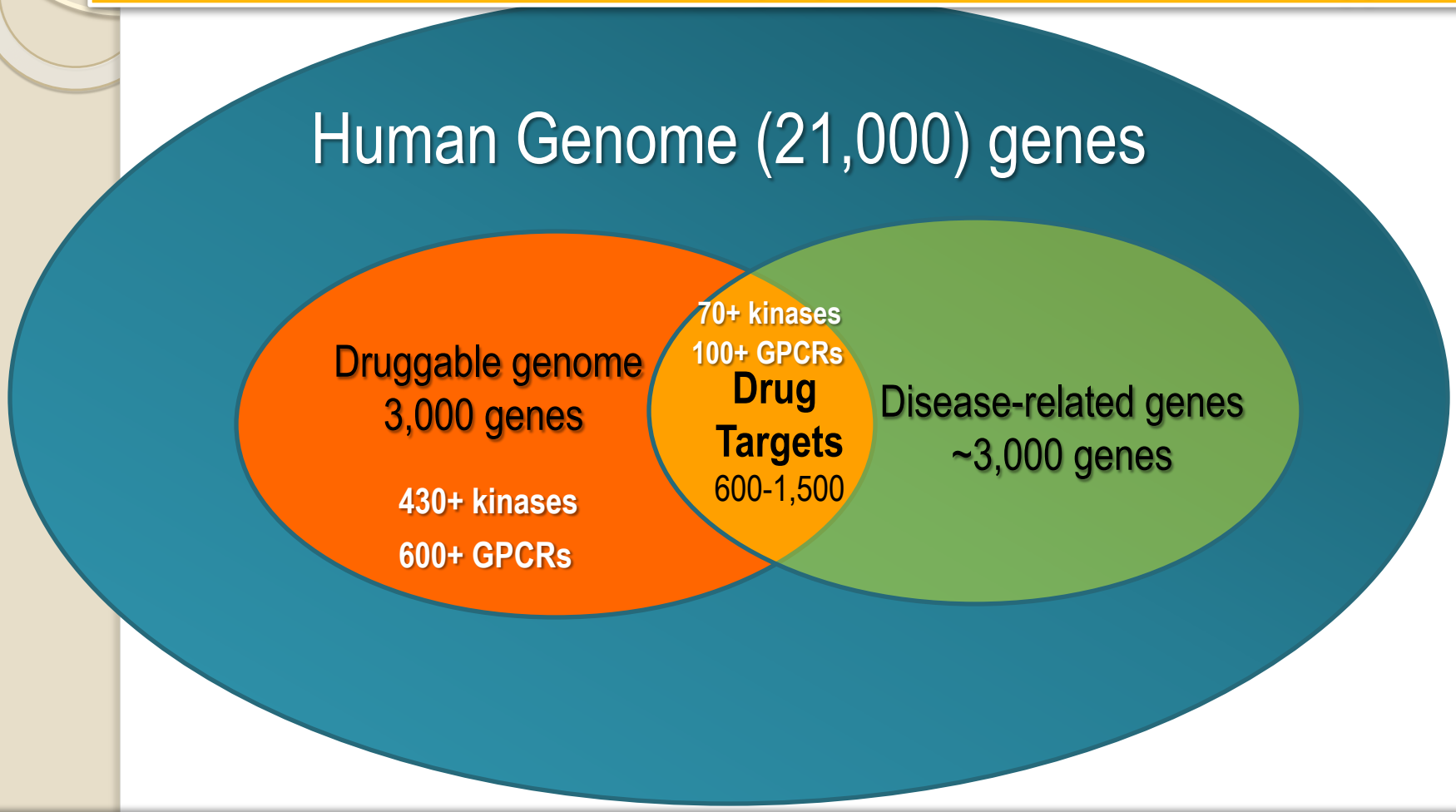
Summary

- Efficient methodology to identify essential sites affecting global dynamics either in response to possible ligand binding or due to participating in global hinge regions.
- Allosteric pocket prediction technique outperforming two widely used ENM-based algorithms, PARS and AllositePro.
- Detection of cryptic sites when combined with ClustENMD
- Identification of multiple ligand-binding sites in homologous structures, e.g. for GPCRs
- Possible application at the proteome level due to efficiency

Druggability

Druggable Genome

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



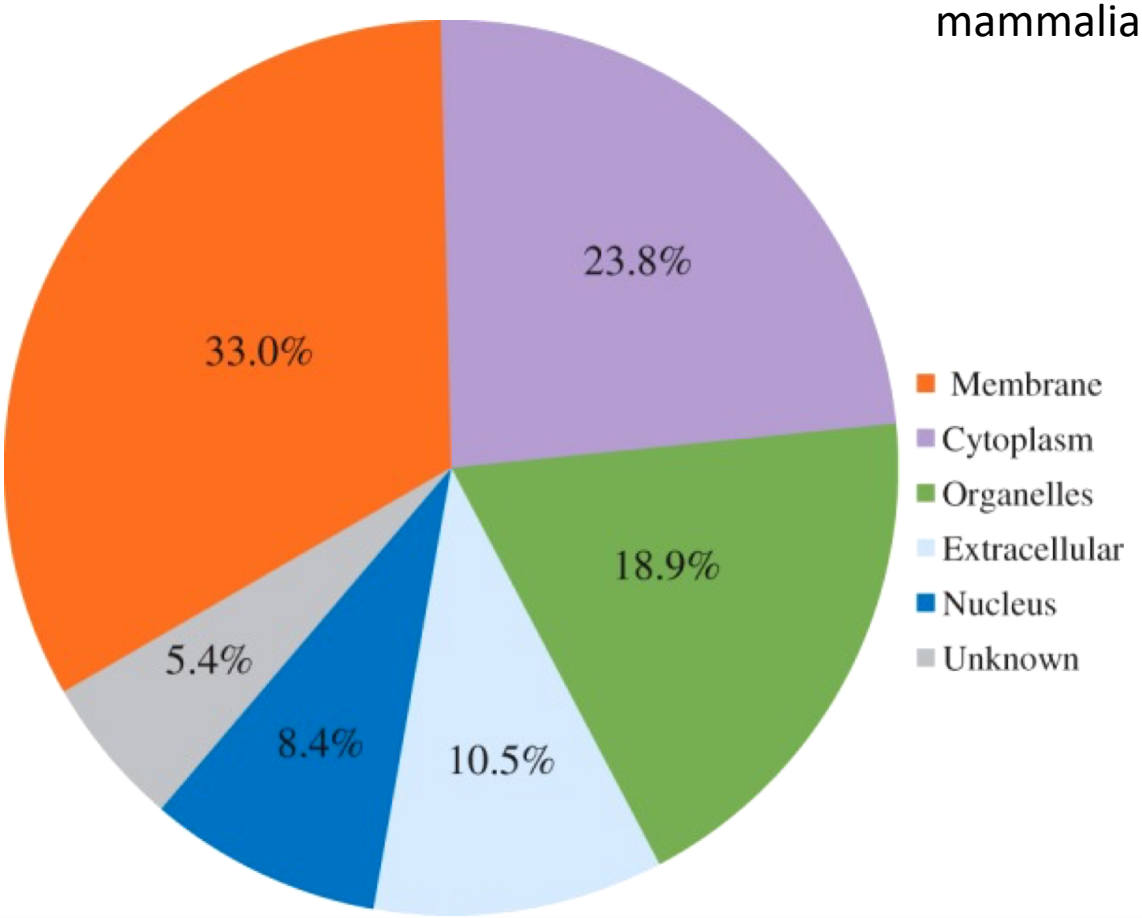
Hopkins and Groom, *Nat Reviews Drug Disc*, 2002

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.

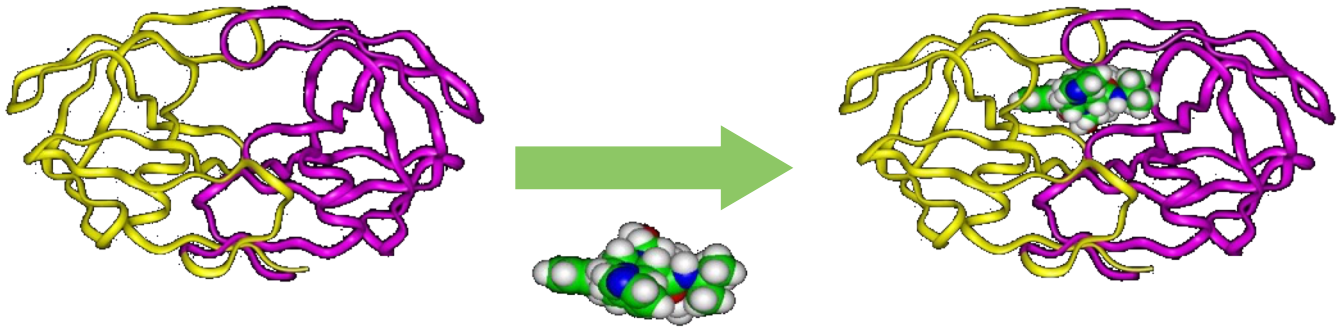


Genome Biol Evol. 2013; 5(7): 1291–1297.

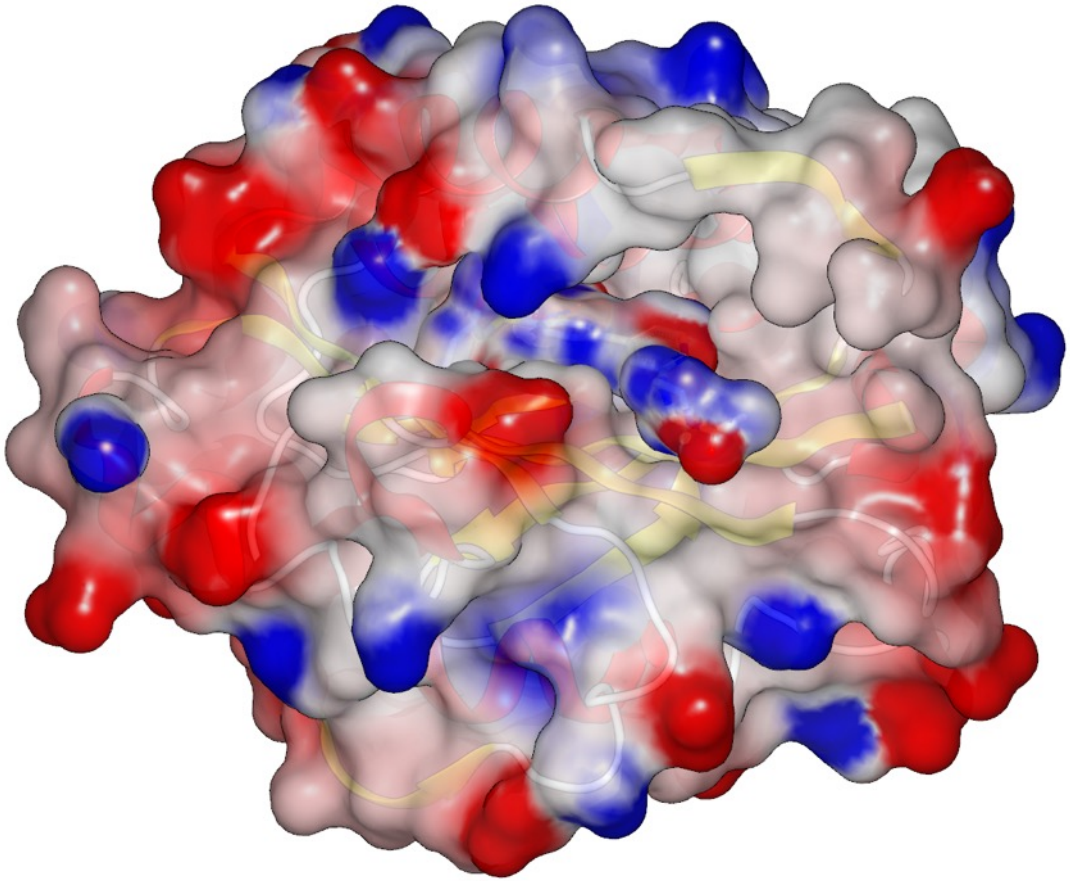
Rational Design of Inhibitors

3D structure of the target is used for

- Molecular insight on function and allostery
- Molecular docking of small molecules/fragments
- Pharmacophore modeling
- Virtual screening

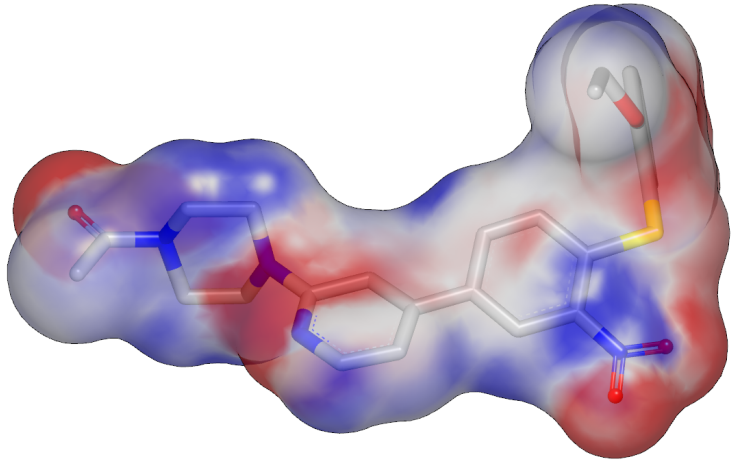


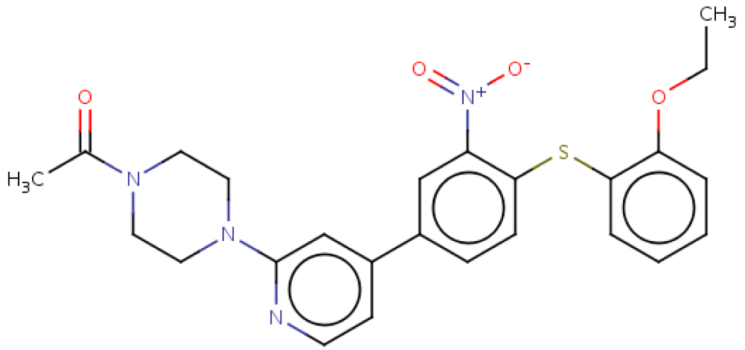
Druggable or not?



Active site druggability:

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

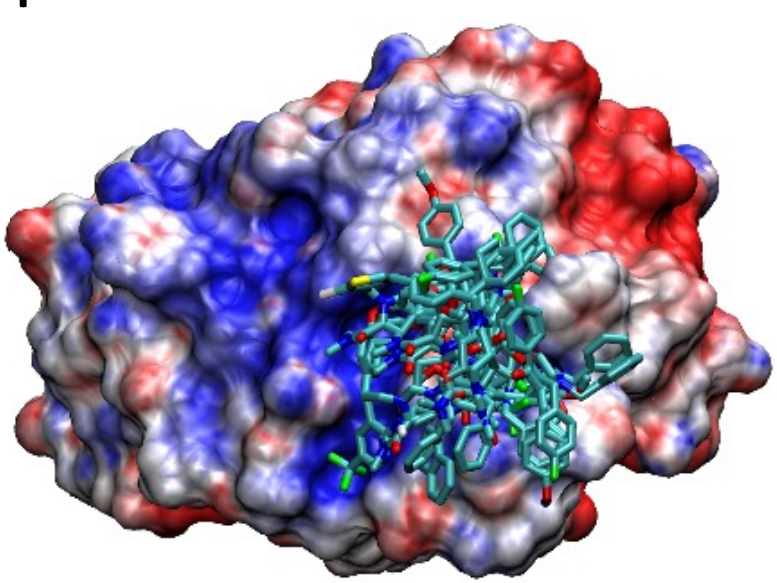




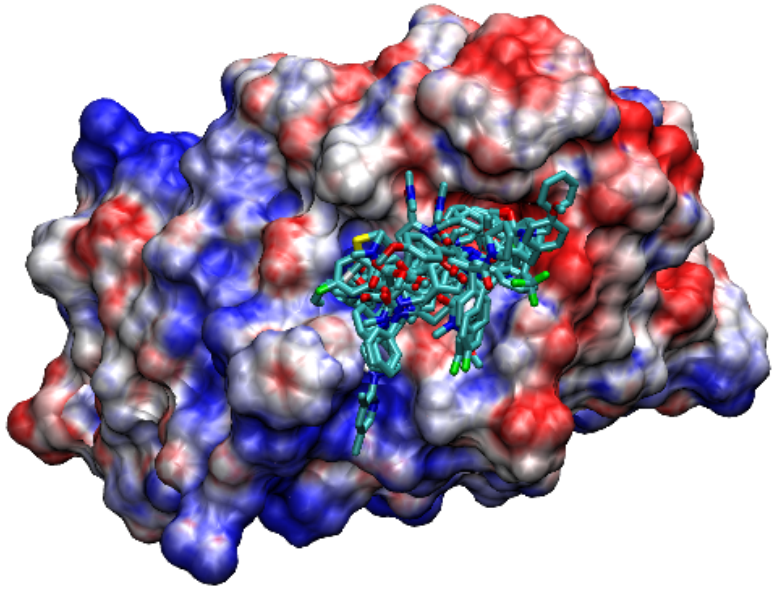
Lfa1 - a leukocyte glycoprotein that promotes intercellular adhesion and binds intercellular adhesion molecule 1

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

Some proteins do not present well-defined pockets



MKP-1



VHR

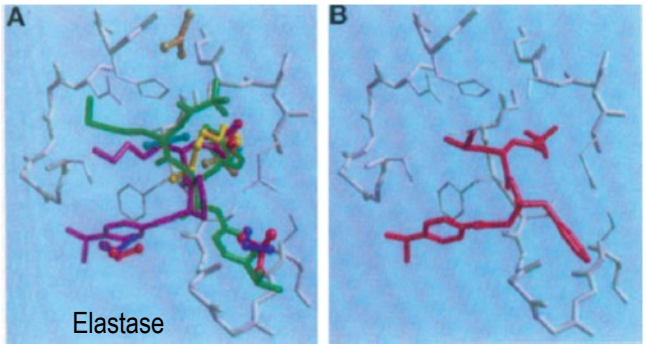
A problem:
Hard to discriminate between different binding compounds/poses for a given target if the surface does not present suitable pockets

"Structurally Unique Inhibitors of Human Mitogen-activated Protein Kinase Phosphatase-1 Identified in a Pyrrole Carboxamide Library."Lazo et al (2007) J Pharmacol Exp Ther.

Druggability from Experiments

- X-ray crystallography

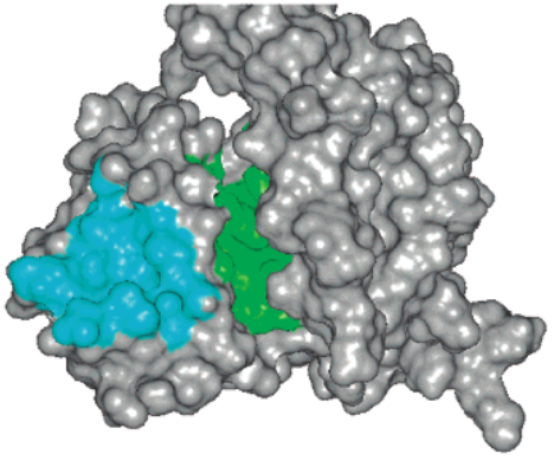
- protein structure is solved in presence of small organic molecules



Mattos and Ridge, *Nat Biotechnology*, 1996

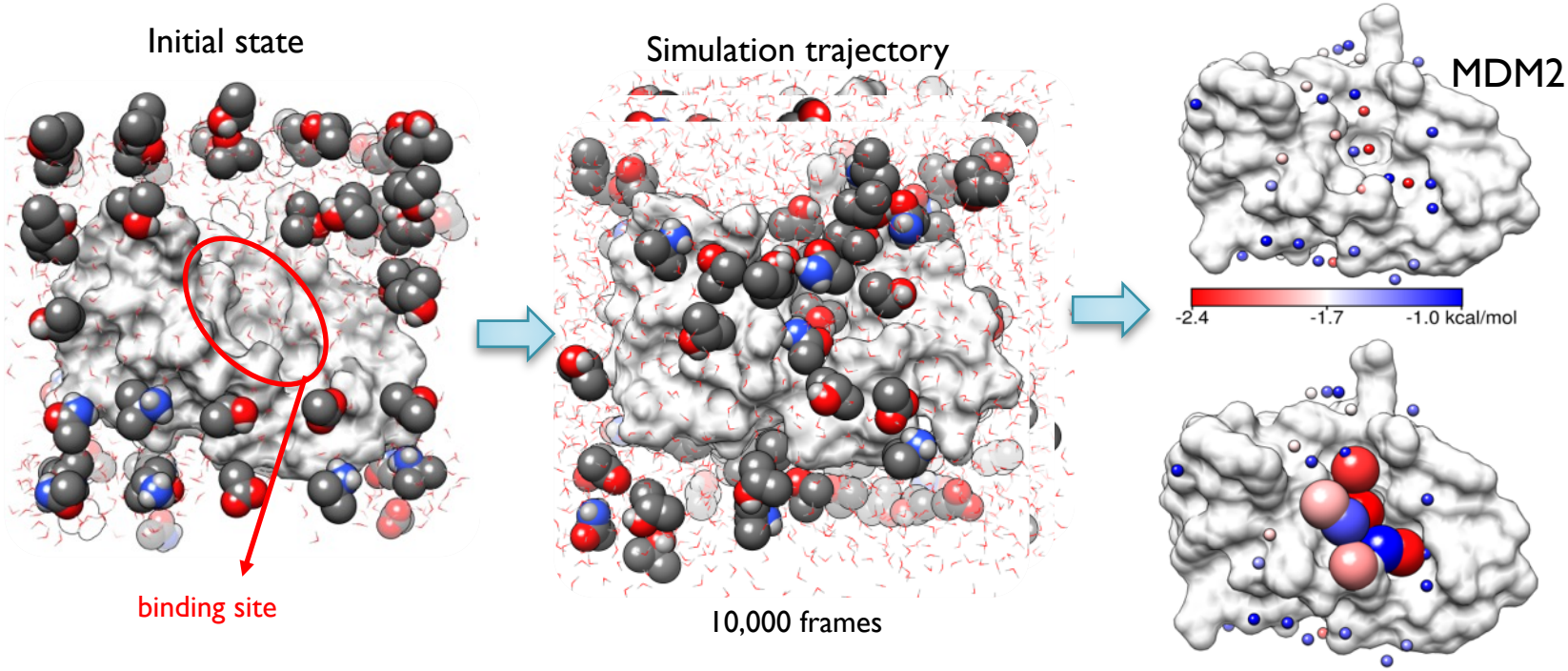
- NMR screening

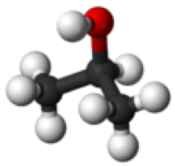
- compounds from a fragment-library are screened as mixtures of 20-30 compounds, druggability is calculated from chemical shift perturbations



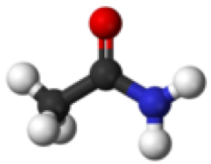
Hajduk et al., *J Med Chem*, 2005

Druggability Simulations

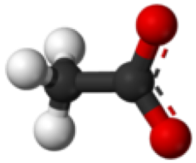




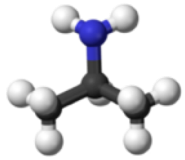
Isopropanol
(observed in 57% of drugs)



Acetamide
(21%)

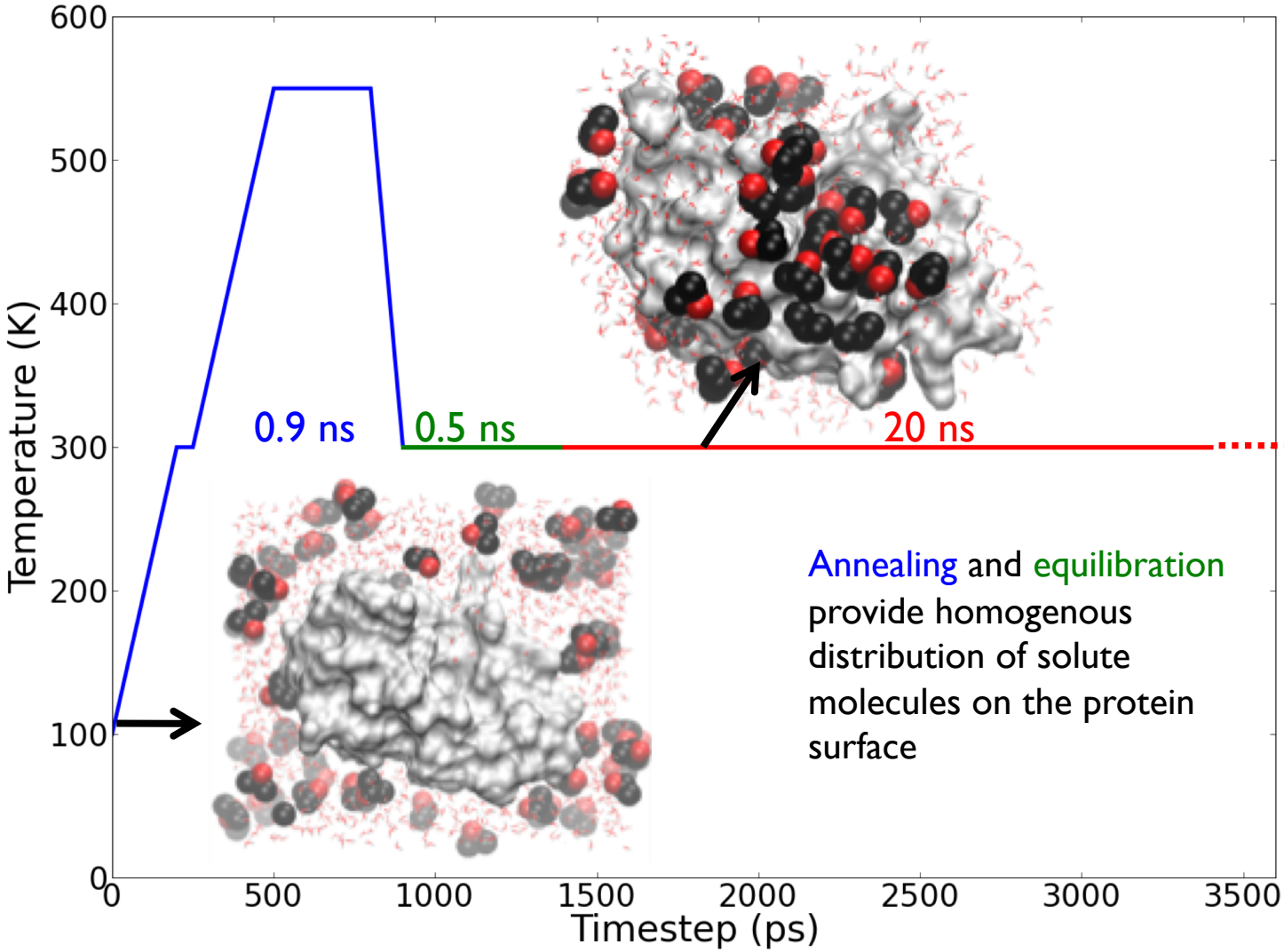


Acetate (-1)
(21%)



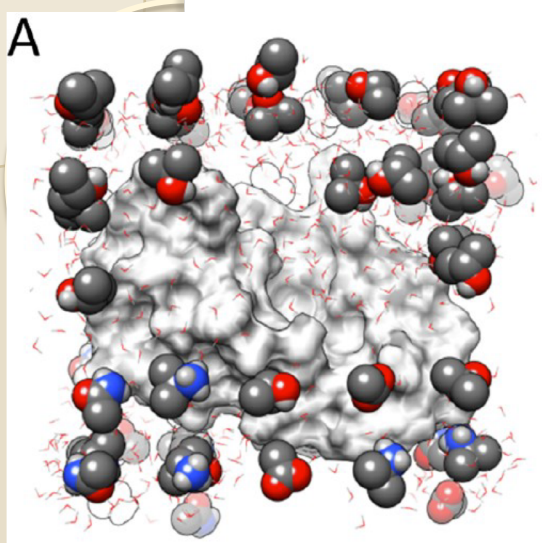
Isopropylamine (+1)
(25%)

Annealing, Equilibration, Simulation

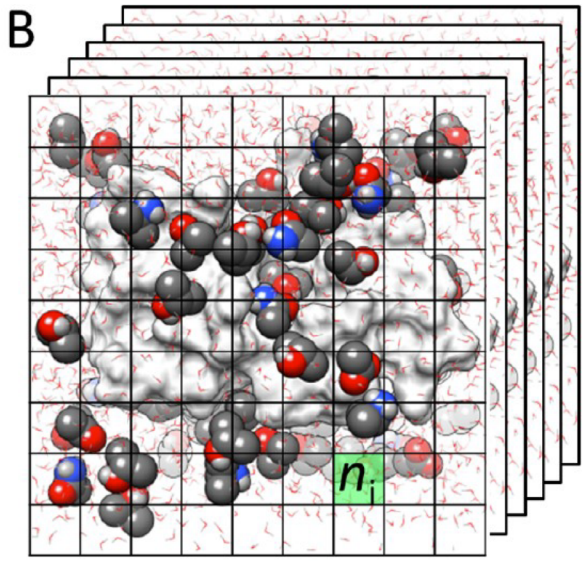


Annealing and equilibration provide homogenous distribution of solute molecules on the protein surface

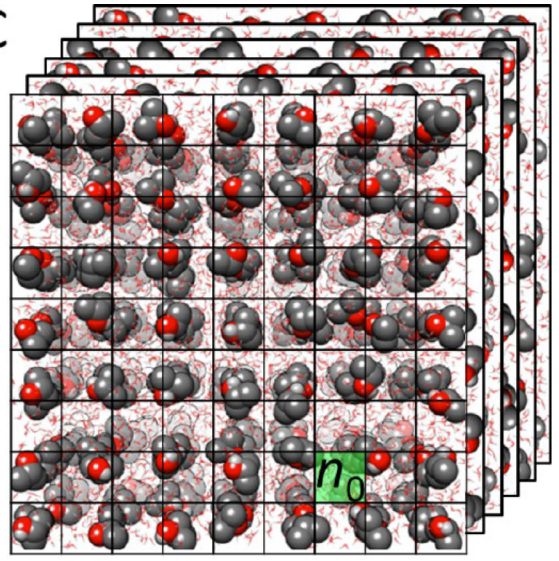
NAMD2 with CHARMM force field was used for simulations.



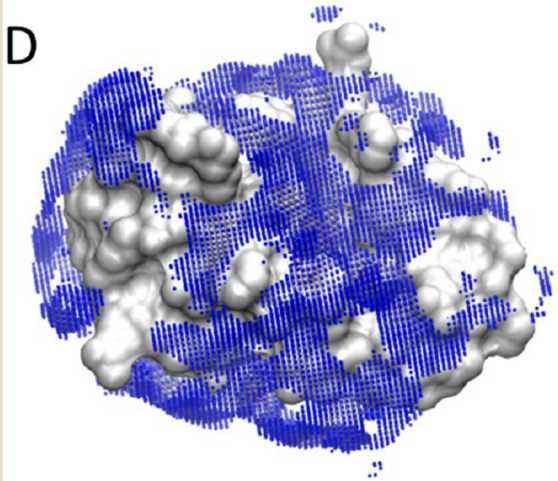
Initial system configuration



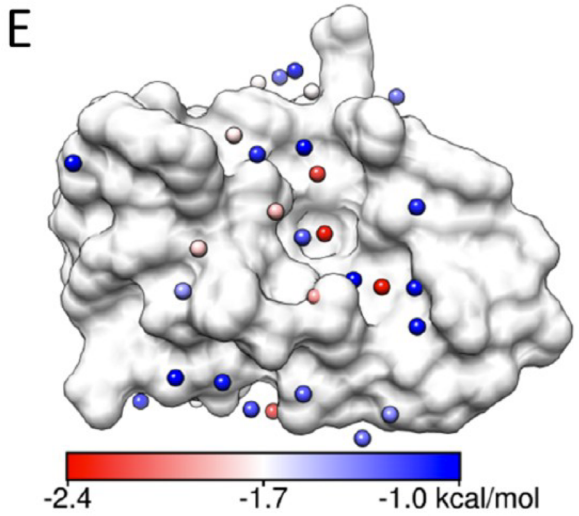
MD simulation and data collection
(8K to 10K frames)



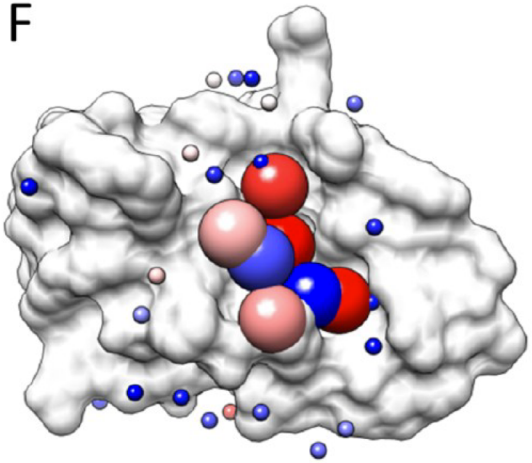
Reference simulation for
calculating enrichment



Probe binding free energy map,
calculated for each voxel using
Eq. 1, n_i (B) and n_0 (C)



Probe interaction spots colored by
binding free energy



A solution obtained from
seven spots (affinity 0.3 nM
and volume 451 Å^3)

$\Delta G_i = -RT \ln(n_i/n_0)$

Hot spots

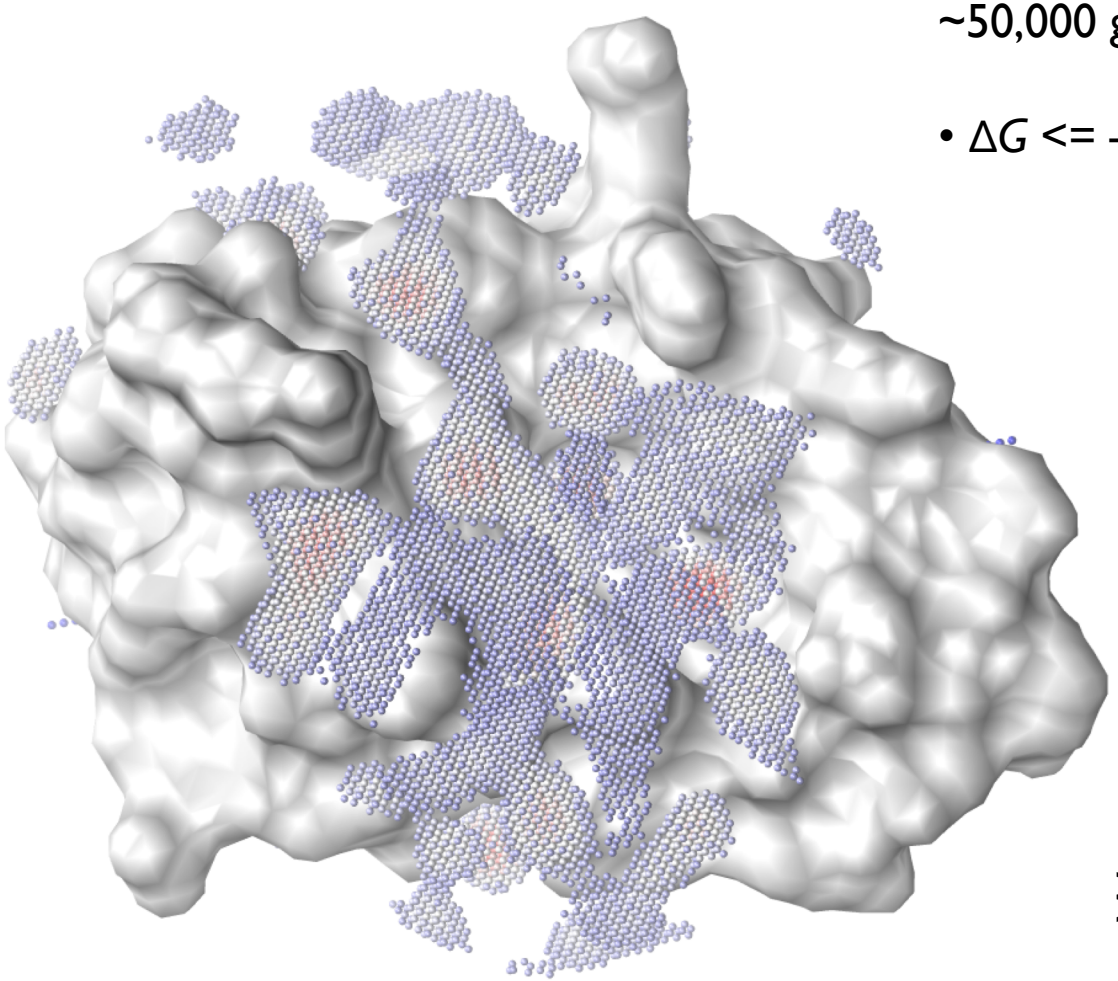
Binding site- Affinity estimation

Isopropanol Binding Spots

ΔG grid is mapped onto the protein structure

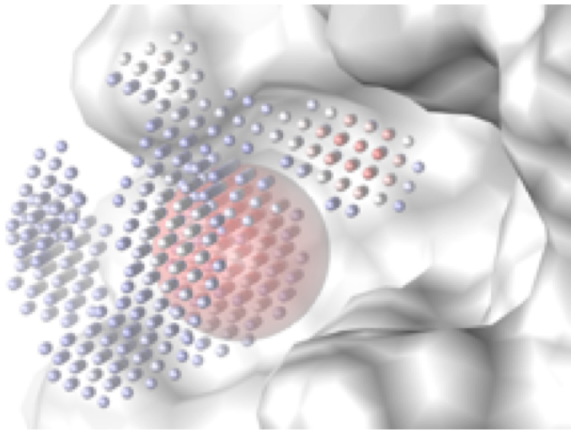
~50,000 grid points with

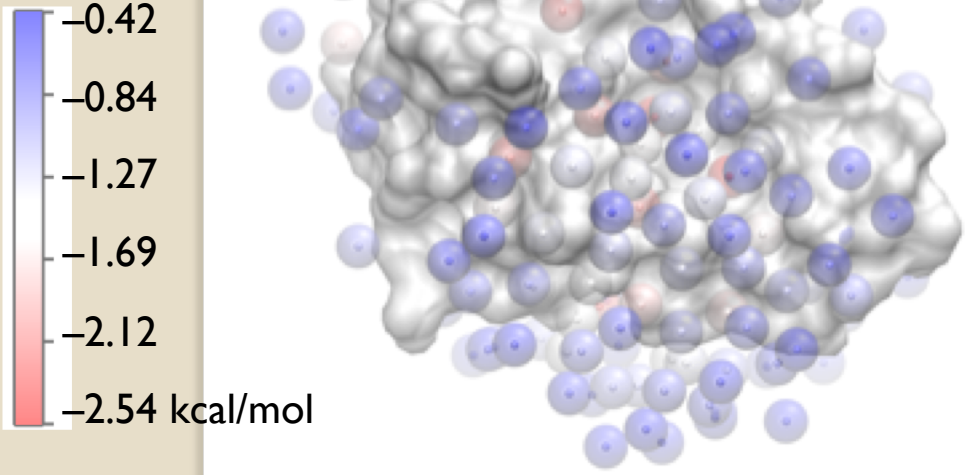
- ΔG <= – 0.416 kcal/mol*
  (2 fold enrichment)

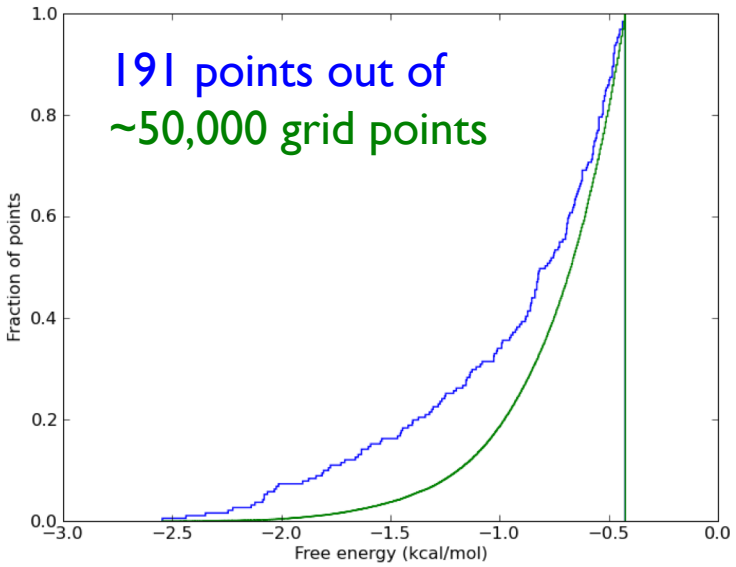


Selecting Isopropanol Binding Spots

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

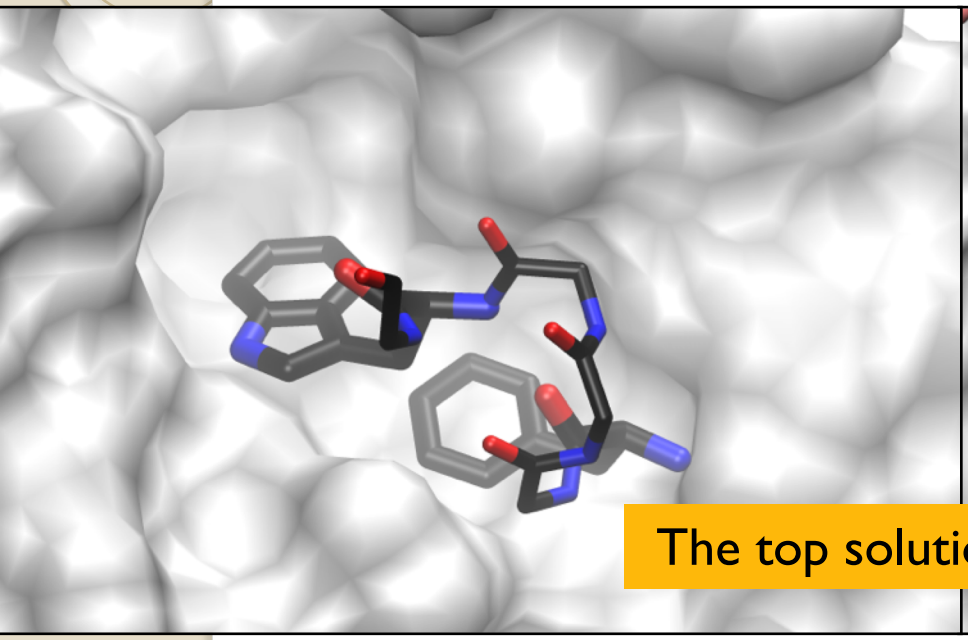




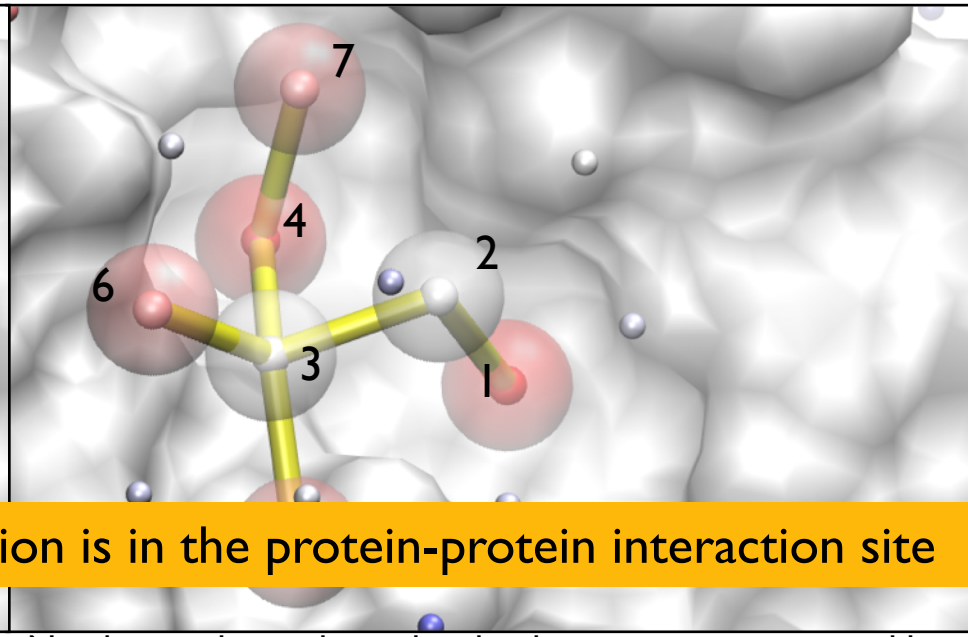


MDM2: p53 binding site

p53 peptide key interactions (X-ray)



Highest affinity solution (7 points)



The top solution is in the protein-protein interaction site

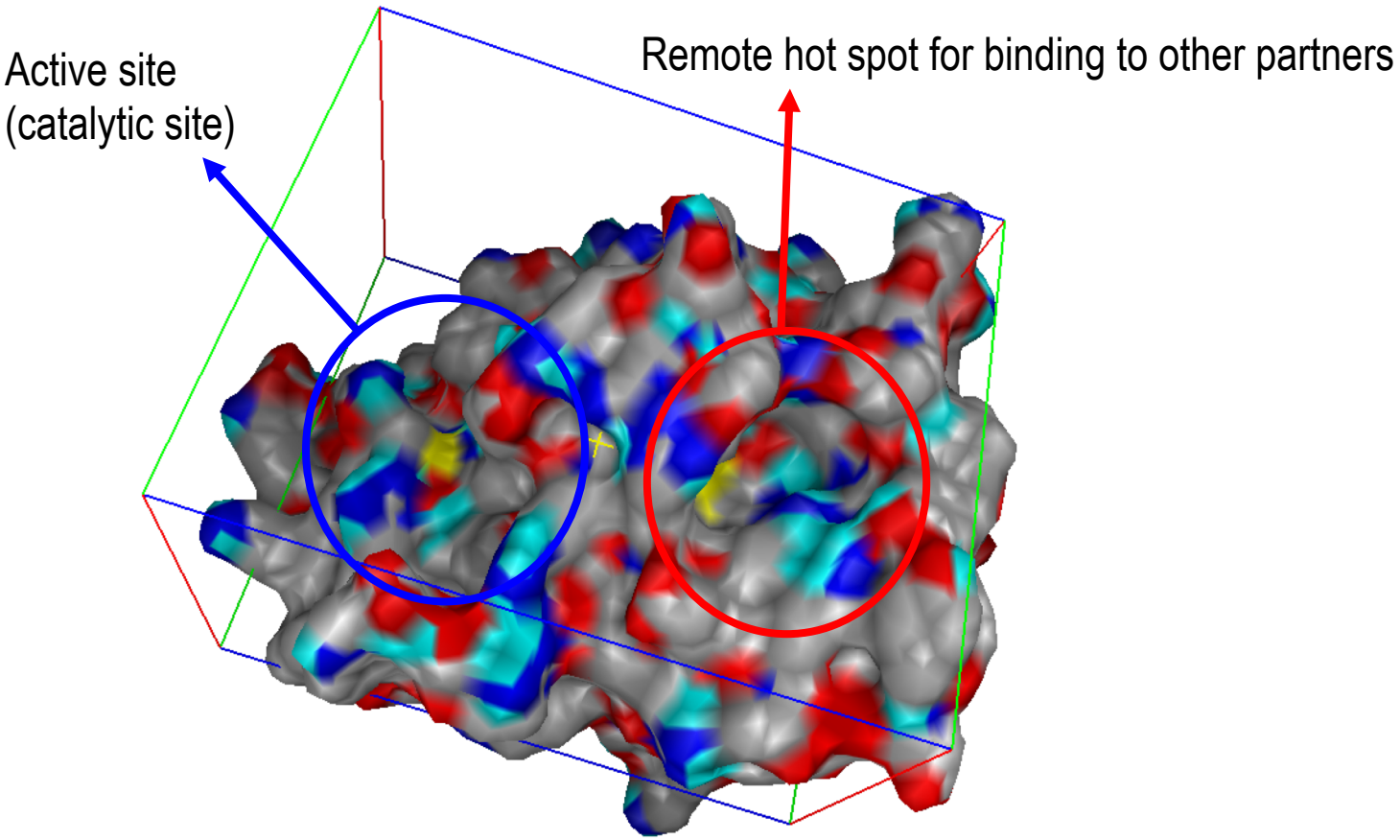
Numbers indicate the order that hot spots were merged by the growing algorithm

Predicted K_d range : **0.3-2.0 nM**

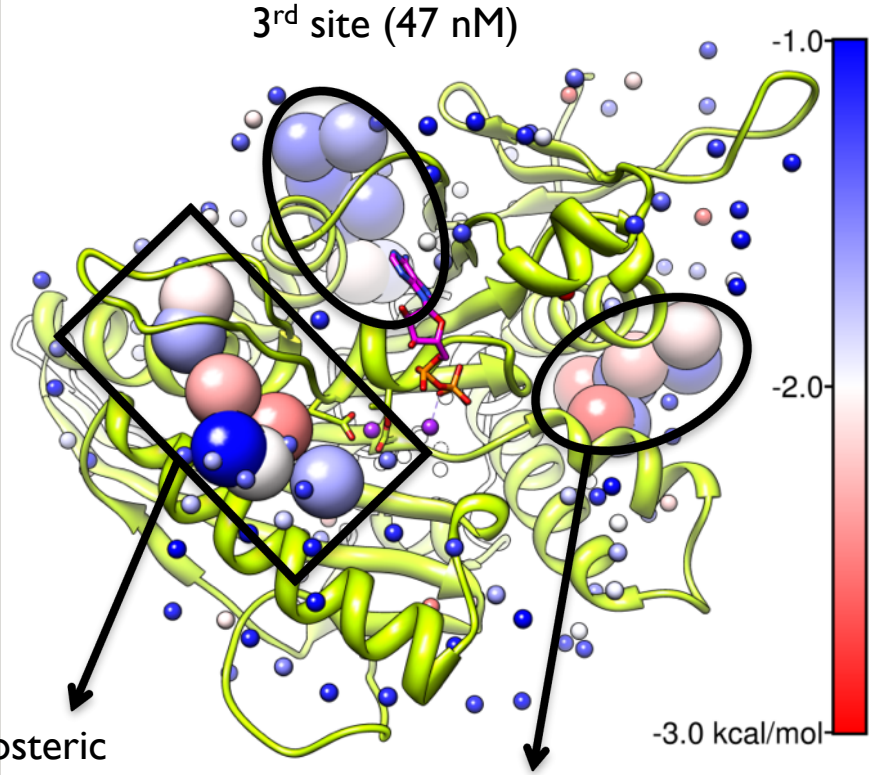
Experimental K_d : **0.6 nM**

(most potent inhibitor by Yu et al, 2009)

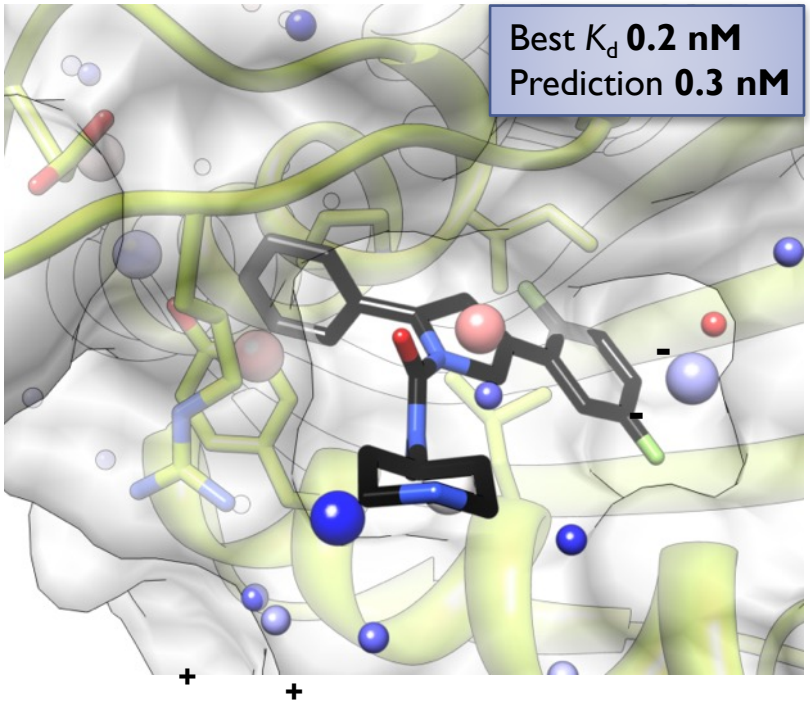
Proteins may have multiple target sites



eg5 Druggable Sites



Allosteric site

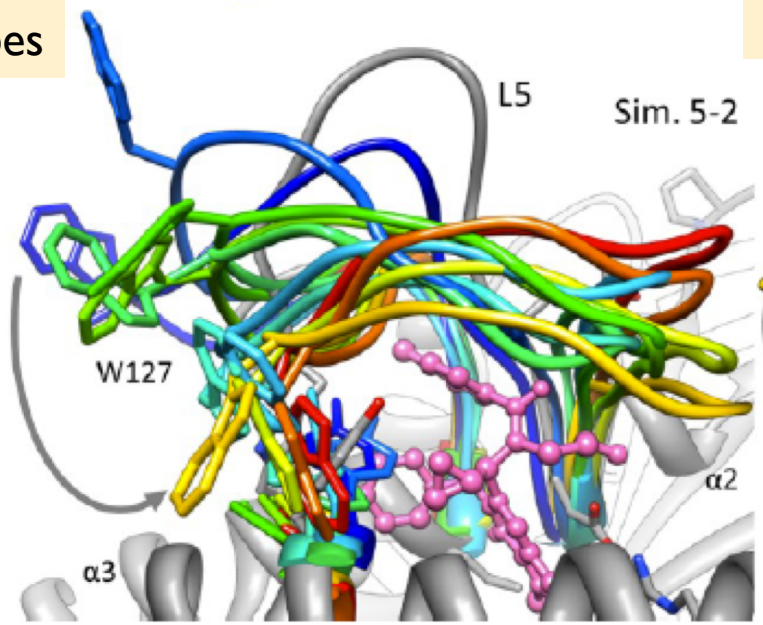


Bioorg. Med. Chem. Lett. **2007,** *17*, 5677-5682; **2006**, 16, 3937–3942

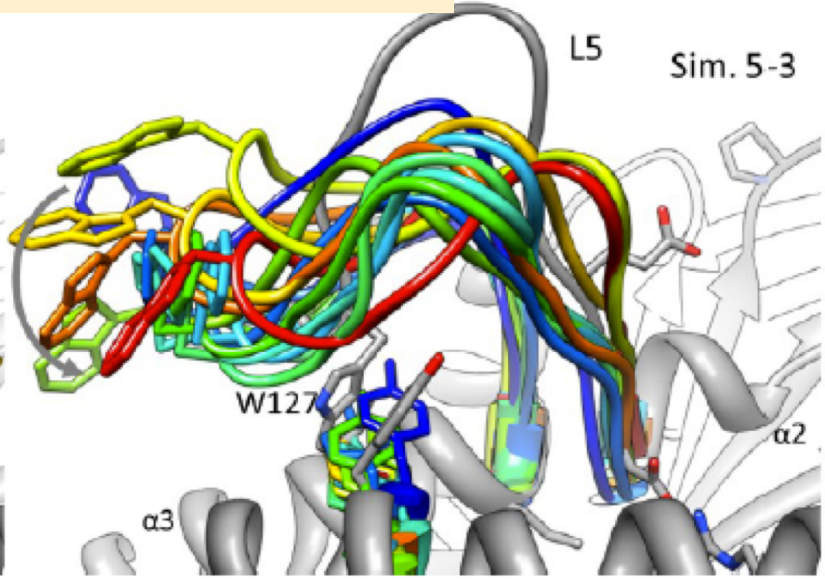
Bakan & Bahar, *J Chem Theory Comput.* 8:2435-2447, 2012

Conformational flexibility-eg5

Probes

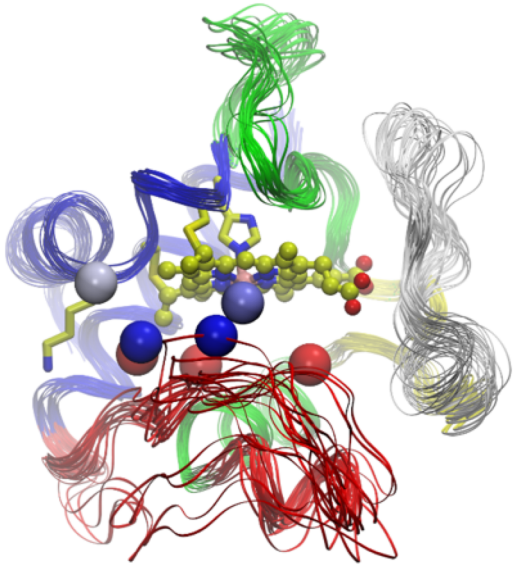


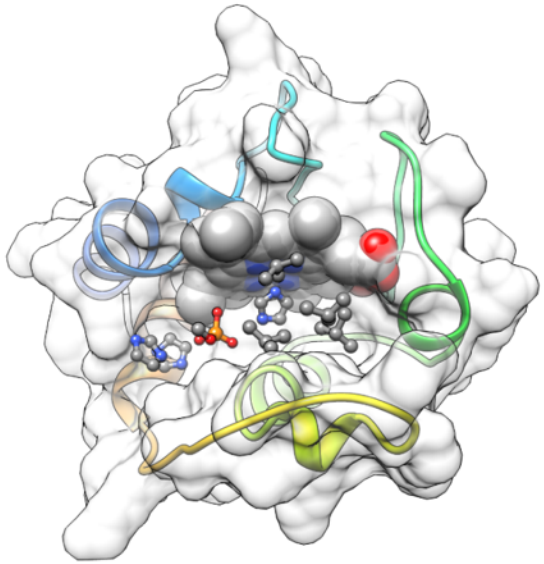
Probe-free simulation



Discovery of inhibitors of cytochrome c peroxidase activity

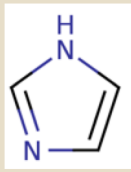
druggability simulations for designing a pharmacophore model



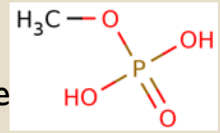


Probes

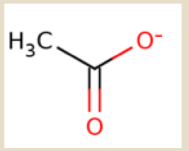
48 imidazole



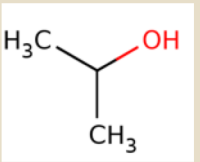
24 methly phosphate



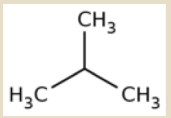
12 acetate



12 isopropanol

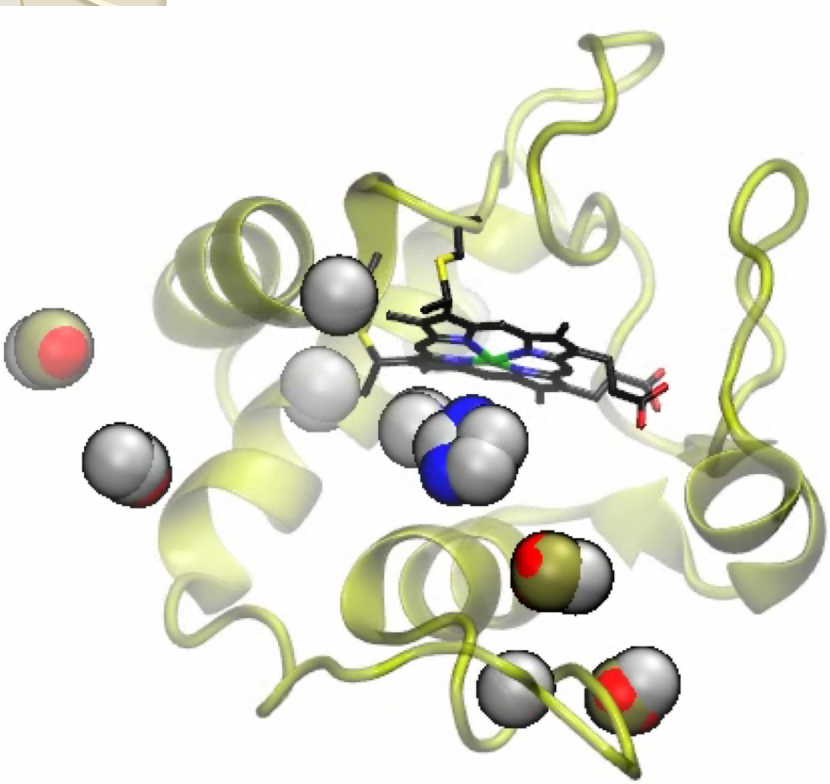


24 isobutane

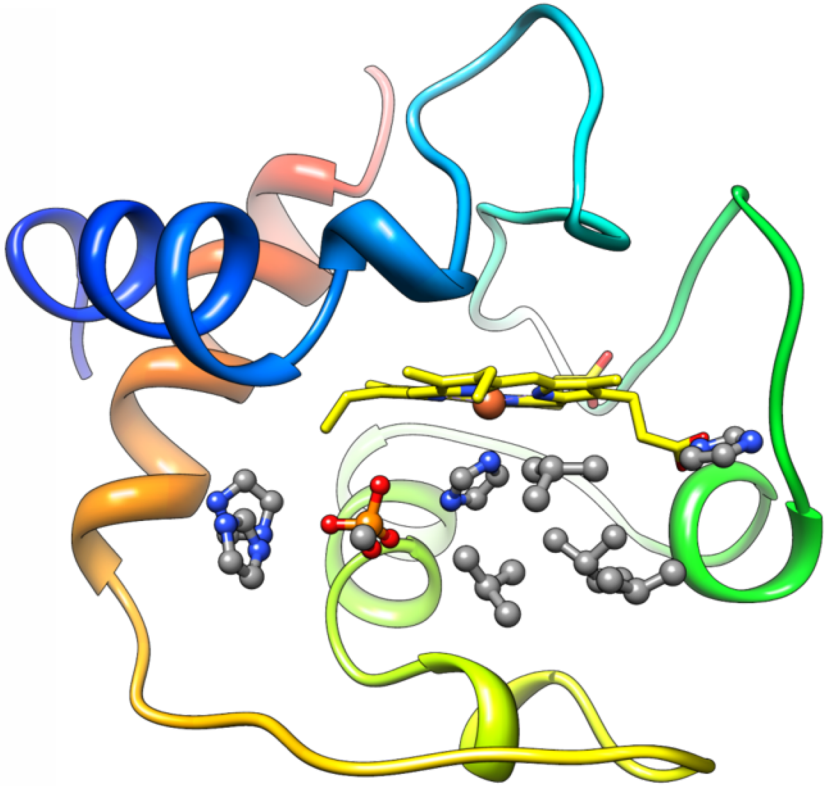


Bakan A, Kapralov AA, Bayir H, Hu F, Kagan VE, Bahar I (2015) Inhibition of Peroxidase Activity of Cytochrome c: De Novo Compound Discovery and Validation *Mol Pharmacol* 88: 421-7

Druggability Simulations



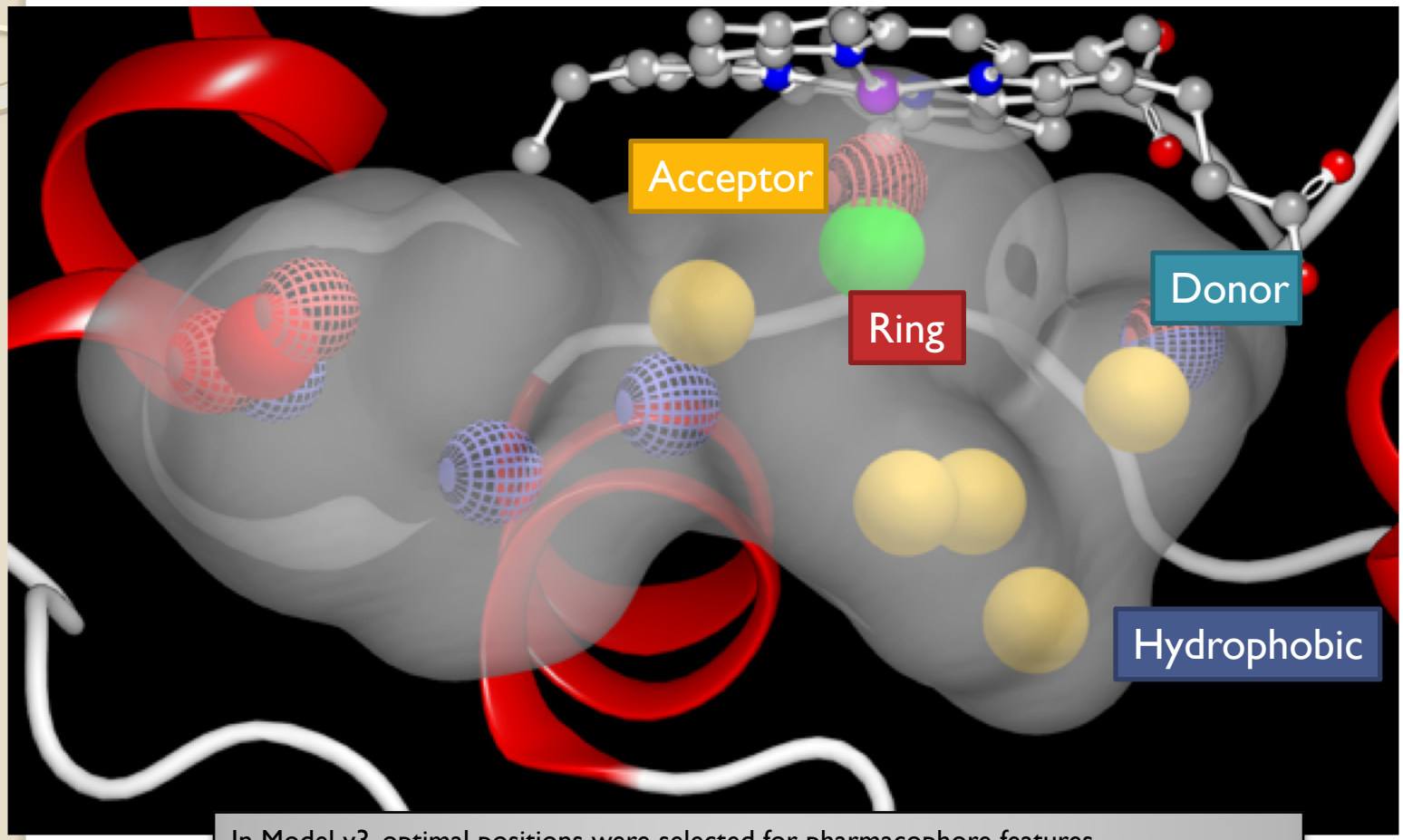
Heme site is the only druggable site with nanomolar achievable affinity



Snapshots from simulations were used to develop a pharmacophore model

Bakan, et al. (2015) *Mol Pharmacol* 88: 421-7

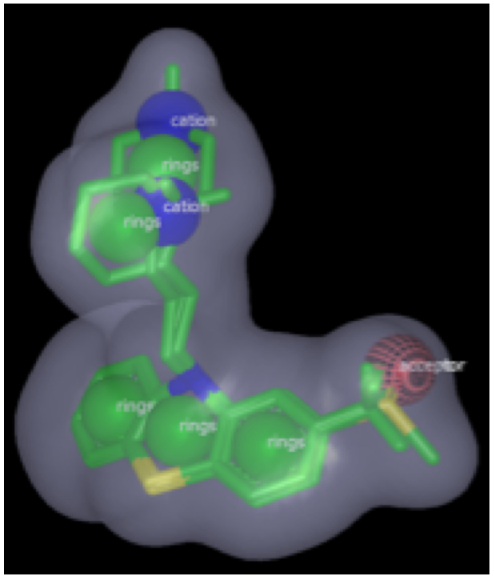
Pharmacophore Model



In Model v3, optimal positions were selected for pharmacophore features

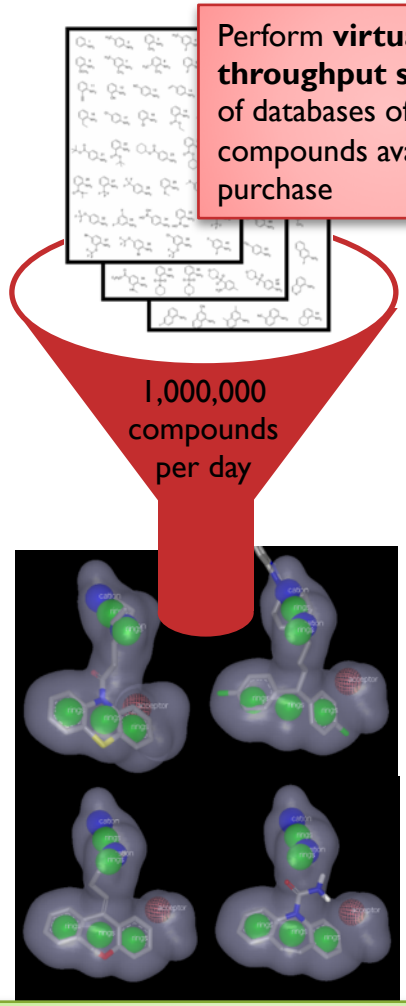
Virtual HTS for hit identification

A **pharmacophore model** describes features common to a set of compounds active against a target protein



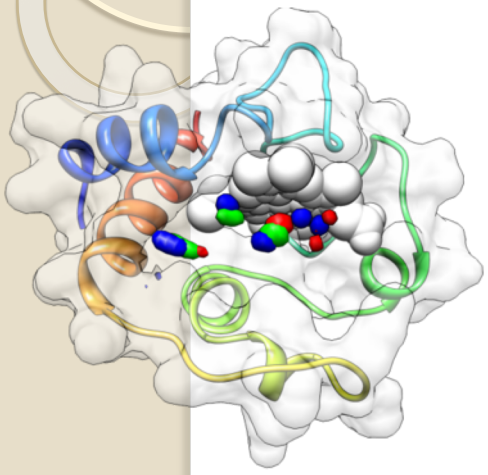
The **virtual HTS pipeline** will allow for identifying *more potent* compounds with similar shapes but diverse chemistry providing us with more choices for chemical synthesis and rational design

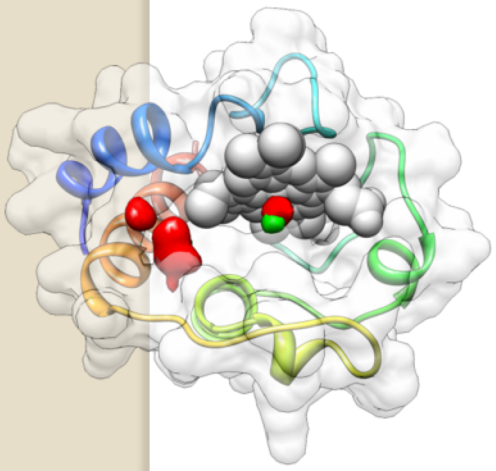
Perform **virtual high throughput screening** of databases of compounds available for purchase



Test virtual hits experimentally and use results to refine the model

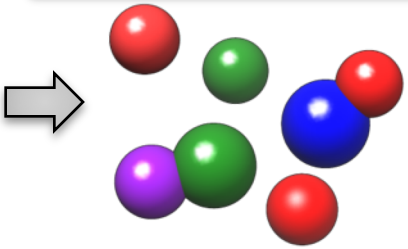
In silico screening



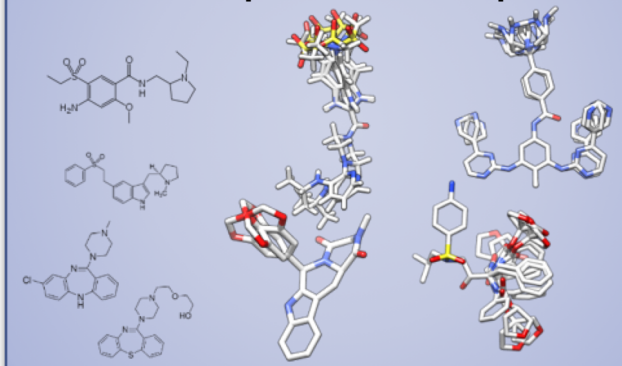


DrugBank ZINC

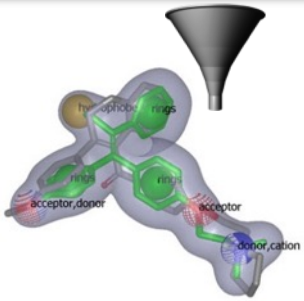
Build a consensus pharmacophore model



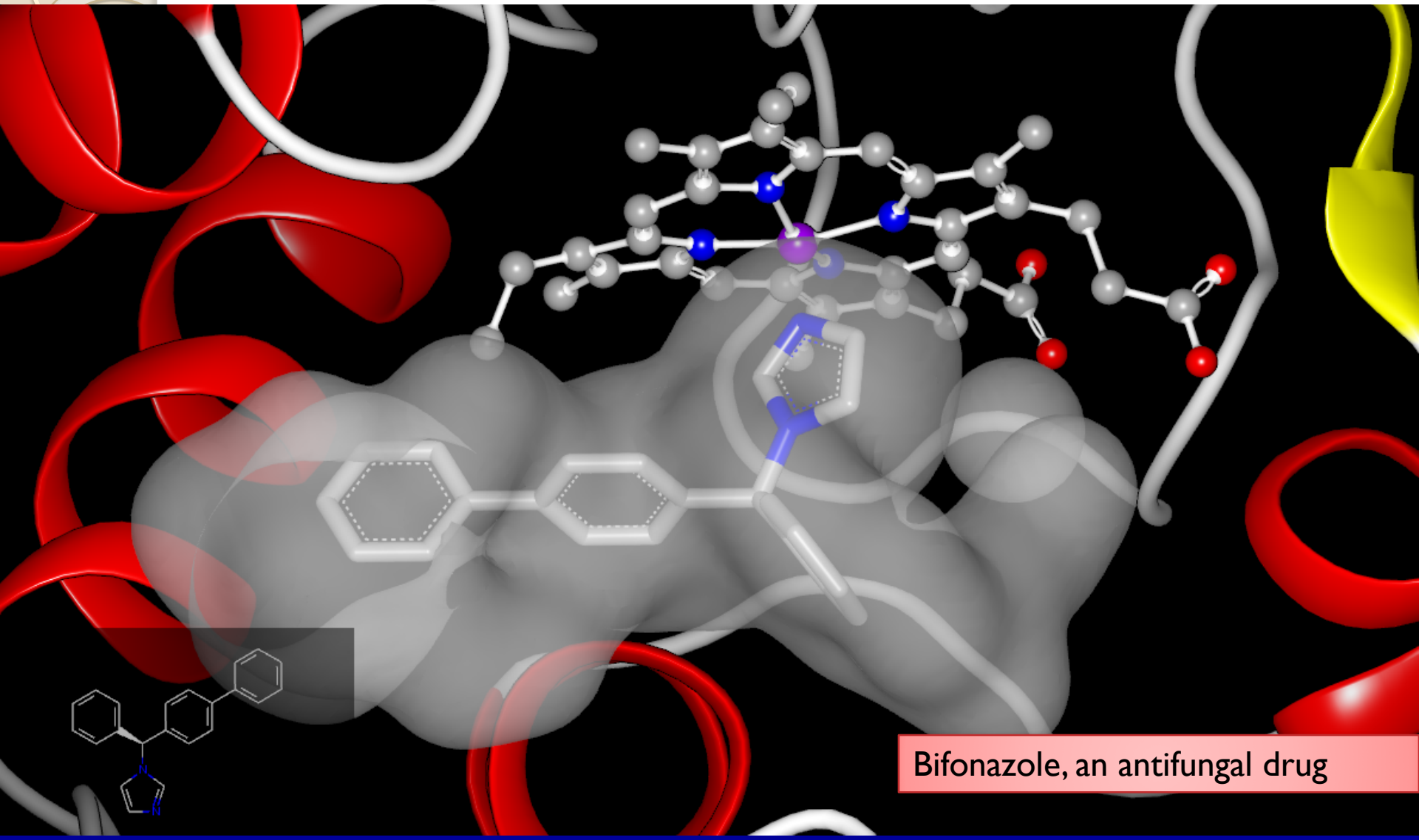
Screen virtual libraries
~6,000 approved or investigational drugs
7.5 million purchasable compounds



Test ~10-20 compounds

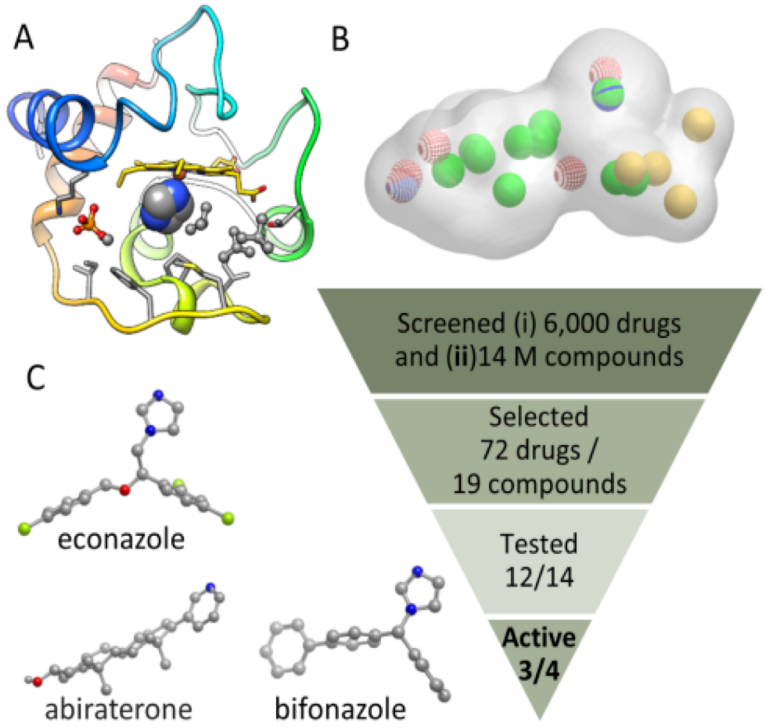


Example *in silico* hit



Bifonazole, an antifungal drug

Discovery of inhibitors of cytochrome c peroxidase activity



Cyt c inhibitor discovery and drug repurposing

A. A snapshot from cyt *c* druggability simulations

B. Pharmacophore model that was built based on tightly bound molecules observed in druggability simulations.

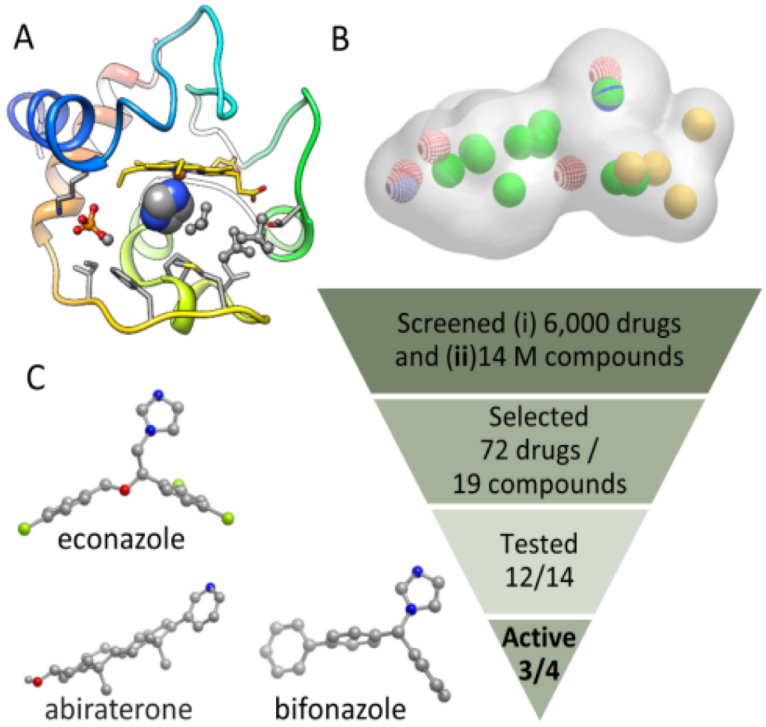
C. This model was used for virtual screening of 6,000 approved and experimental drugs; 72 repurposable drugs were identified, of which 12 have been tested, and 3 turned out to inhibit *cyt* c peroxidase activity, shown in panel **C**.

Additionally, 14 M purchasable drug- and lead-like compounds from the ZINC database were screened, 19 compounds were identified, 14 of which tested, and 4 turned out to be novel inhibitors of cyt *c*.

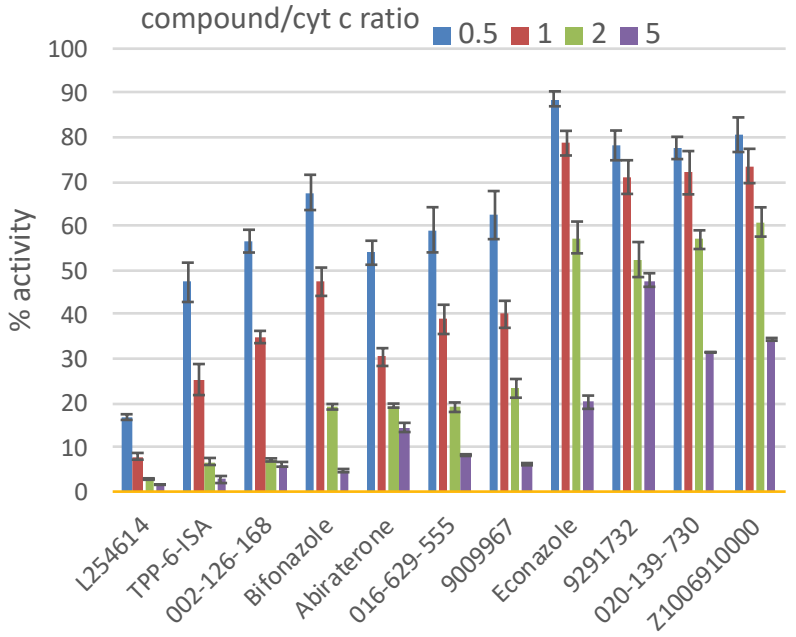
7 novel inhibitors of peroxidase activity of cyt c

Bakan, et al. (2015) *Mol Pharmacol* 88: 421-7

Discovery of inhibitors of cytochrome c peroxidase activity



Exp validation



Peroxidase reaction probed by fluorescence of oxidation product, for cyt c incubated with CL/DOPC liposomes

Overview

Druggability- Purpose

- Assessment of druggable sites
- Identification of allosteric sites that can bind drugs
- Estimation of binding affinity

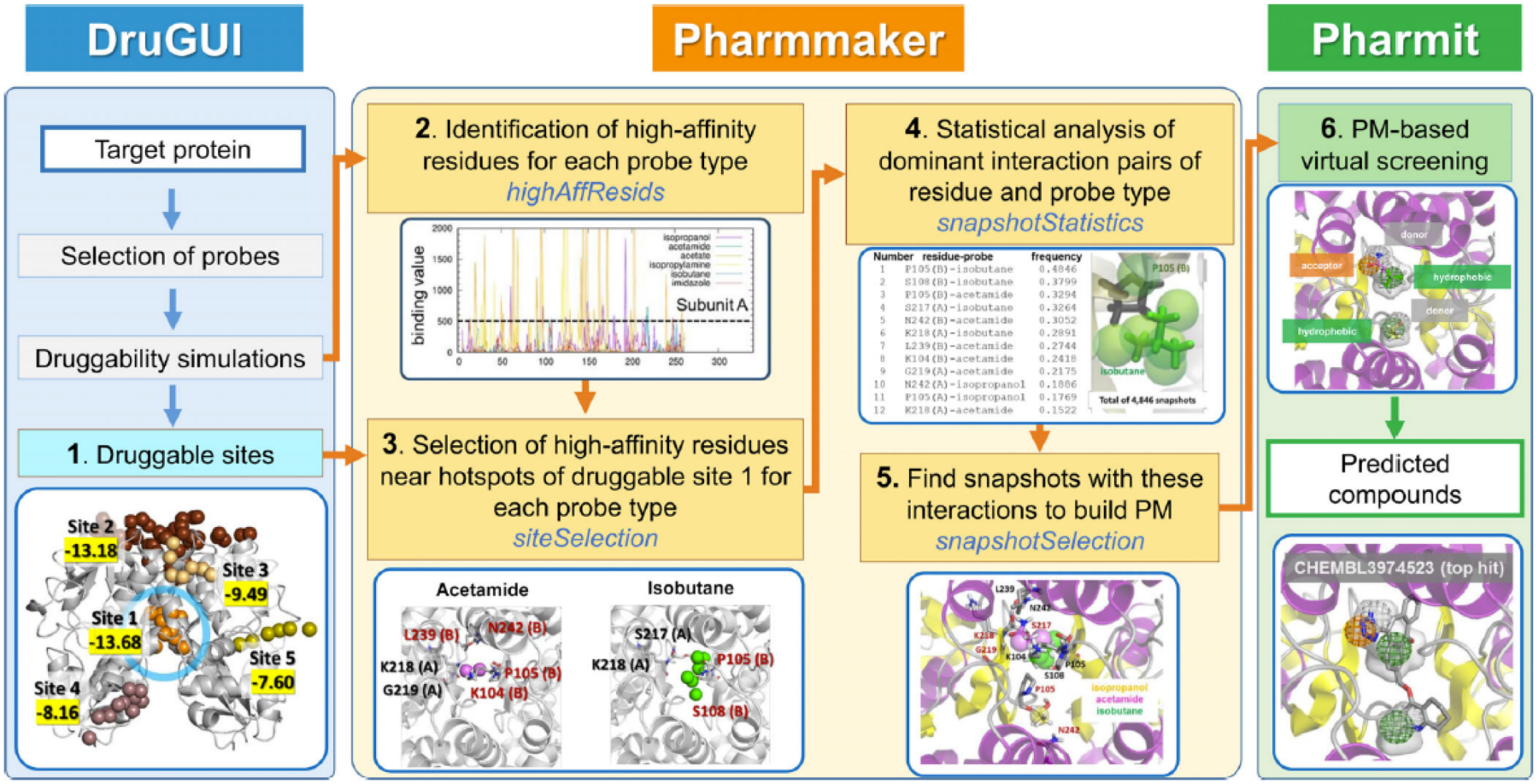
Druggability- Method

- Pre-processing: Define a probe set (customizable)
- Prepare input files for NAMD runs with probe molecules
- Post-processing: Analyze trajectories to make inferences on binding sites and affinities, and to build pharmacophore models

Virtual screening

- Screen pharmacophore model, against libraries of small compounds
- Provide initial hypotheses to be validated by experiments
- Hits confirmed by experiments can be evaluated by atomic simulations (including free energy perturbation methods)

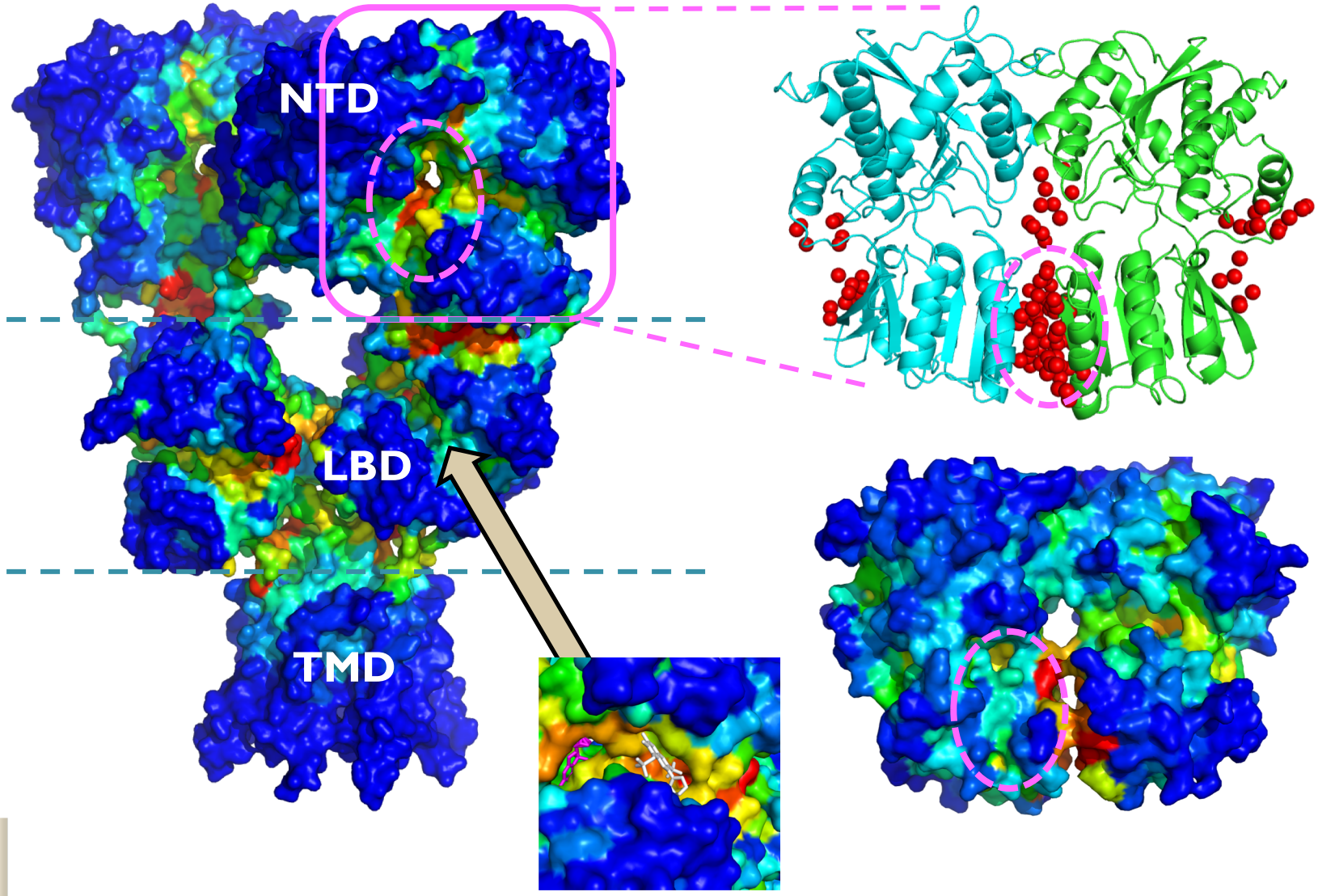
Pharmmaker tool in ProDy



Sunseri and Koes, NAR, 2016

Lee, Krieger, Li and Bahar, *Protein Sci*, 29: 76-86, 2020

Combining ESSA and druggability



Thanks to Organizers and Bahar Lab Members

Support from NIGMS