Bridging Sequence, Structure and Interactions
Environment, Evolution, Druggability,

Ivet Bahar

Department of [1]Computational and Systems Biology
School of Medicine, University of Pittsburgh, Pittsburgh, PA 15260
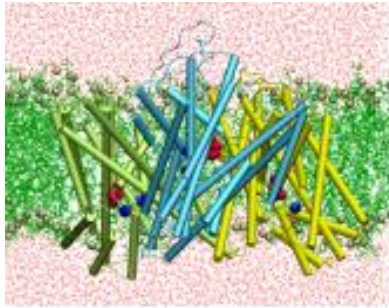
Summary

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 using the ANM
   b. Combining sequence and structure analyses – signature dynamics
   c. Allosteric communication – sensors and effectors

3. **Membrane proteins and druggability**
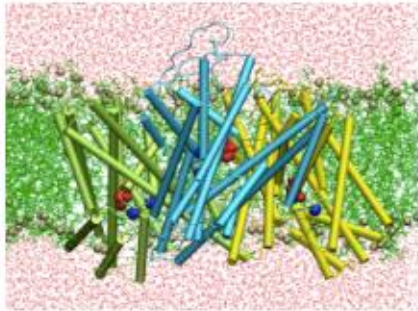   a. Modeling environmental effects using elastic network models
   b. Modeling & simulations of Membrane Proteins with ENMs for lipids
   c. Druggability simulations

ANM for membrane proteins: membrANM

Membrane ANM (membrANM)

Learn how to include the effect of lipid bilayer in ENM study of membrane proteins dynamics.

Go to Tutorial

- Evaluating membrane proteins' dynamics in the presence of lipid bilayer

- Comparing global motions in the presence and absence of membrane

- Understanding mechanisms of protein-membrane remodeling
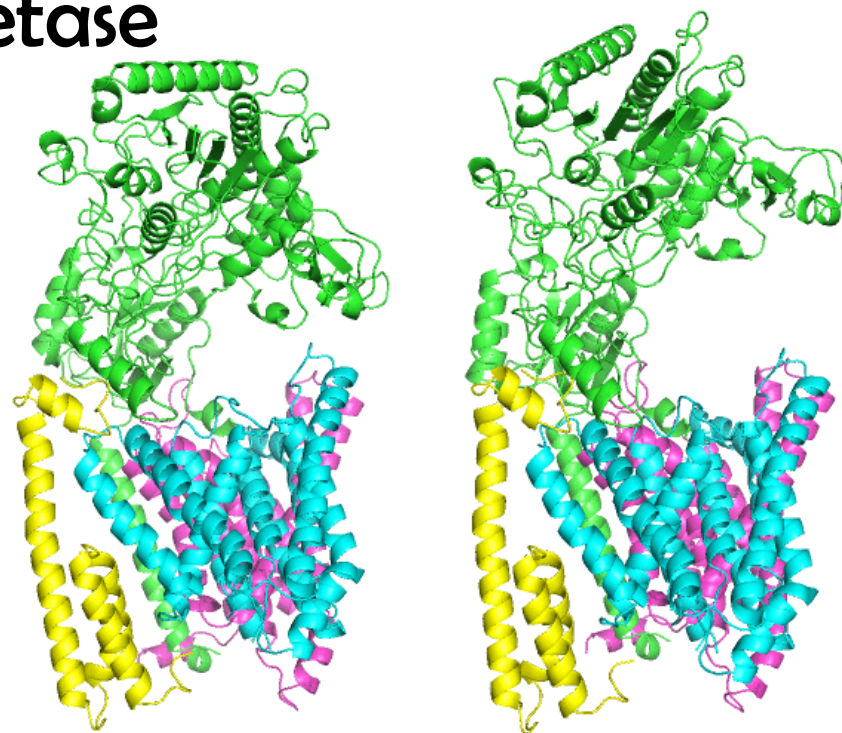
Implemented in ProDy and DynOmics

MembrANM for γ-secretase

The effect of the membrane environment can be incorporated in the ANM analysis by adopting
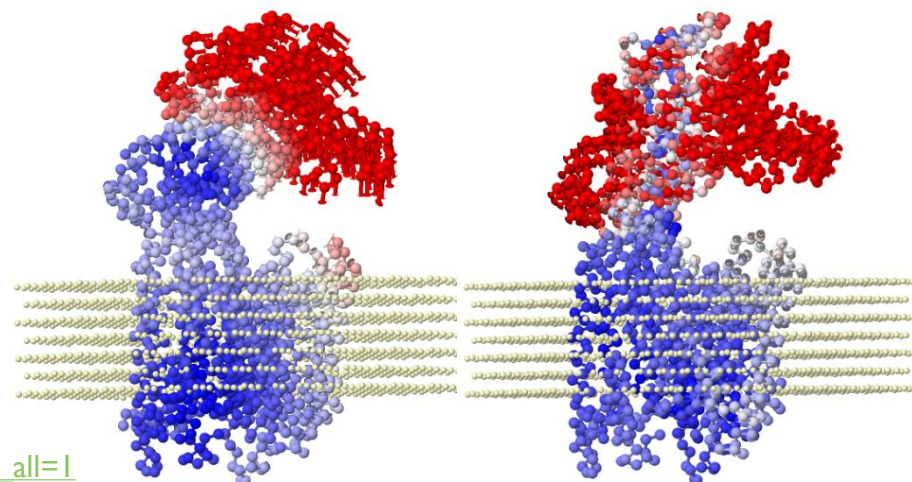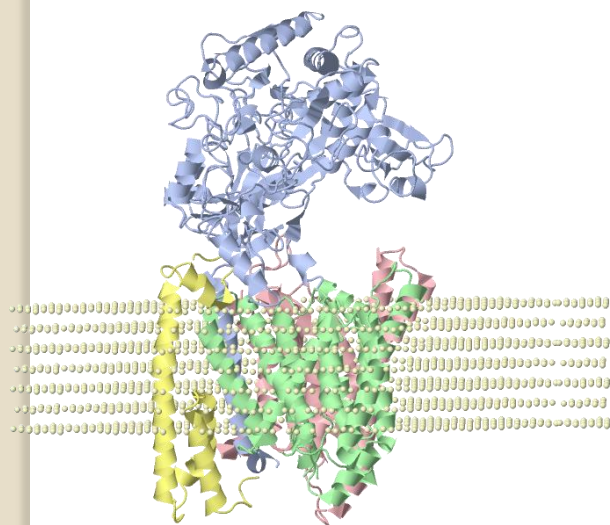- envANM-membrANM or
- substructure-membrANM.

MembrANM results for the γ-secretase (PDB ID: 5FN2) can be viewed on the ENM server:

PyMol movies can be downloaded from DynOmics ENM server by generating "Full Atomic Structures for ANM-Driven Conformers".

Mode 1 - bending   Mode 2 - twisting

http://enm.pitt.edu/oGNM_ANM.php?gnm_id=5FN2&slice_mem=1&sub_all=1

MembrANM for γ-secretase

The effect of the membrane environment can be incorporated in the ANM analysis by adopting
- envANM-membrANM or
- substructure-membrANM.

MembrANM results for the γ-secretase (PDB ID: 5FN2) can be viewed on the ENM server:

PyMol movies can be downloaded from DynOmics ENM server by generating "Full Atomic Structures for ANM-Driven Conformers".
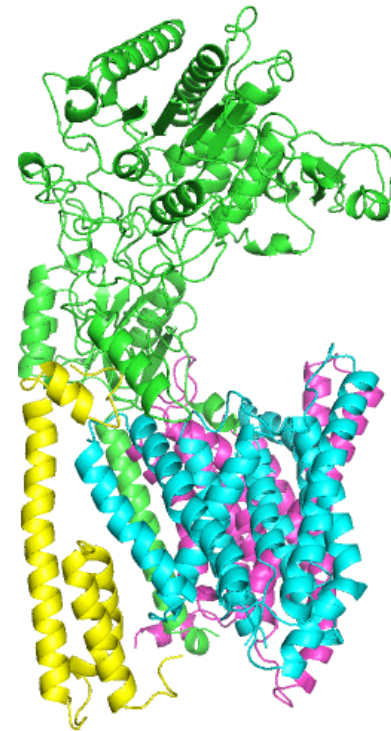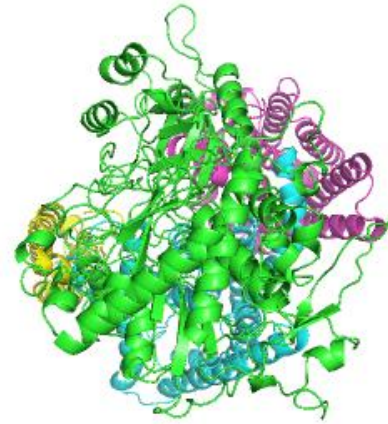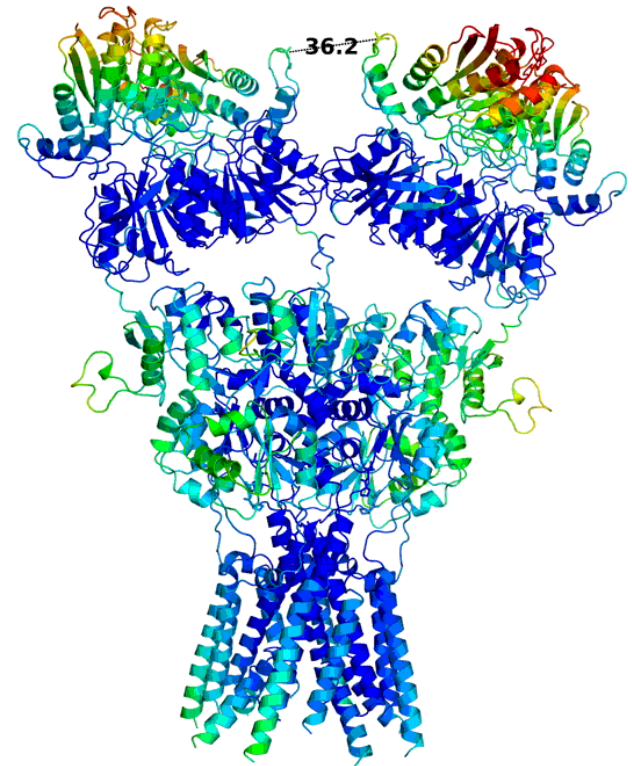

Mode 2 - twisting

Lee et al (2017) Allosteric Modulation of Intact γ-Secretase Structural Dynamics. *Biophys J* 113:2634-2649
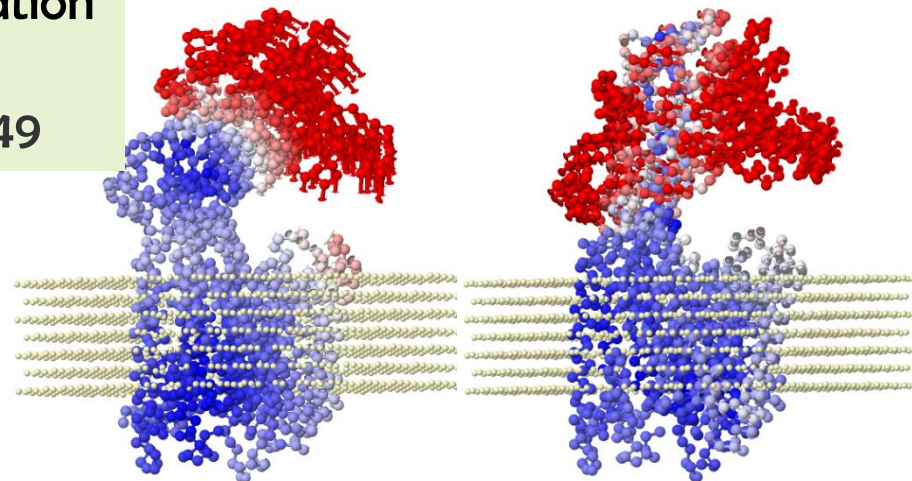
Collective dynamics of AMPA receptors

Experimentally verified by cross-linking experiments. Substitution of cysteines in the presence of an oxidizing agent promotes cross-linking.
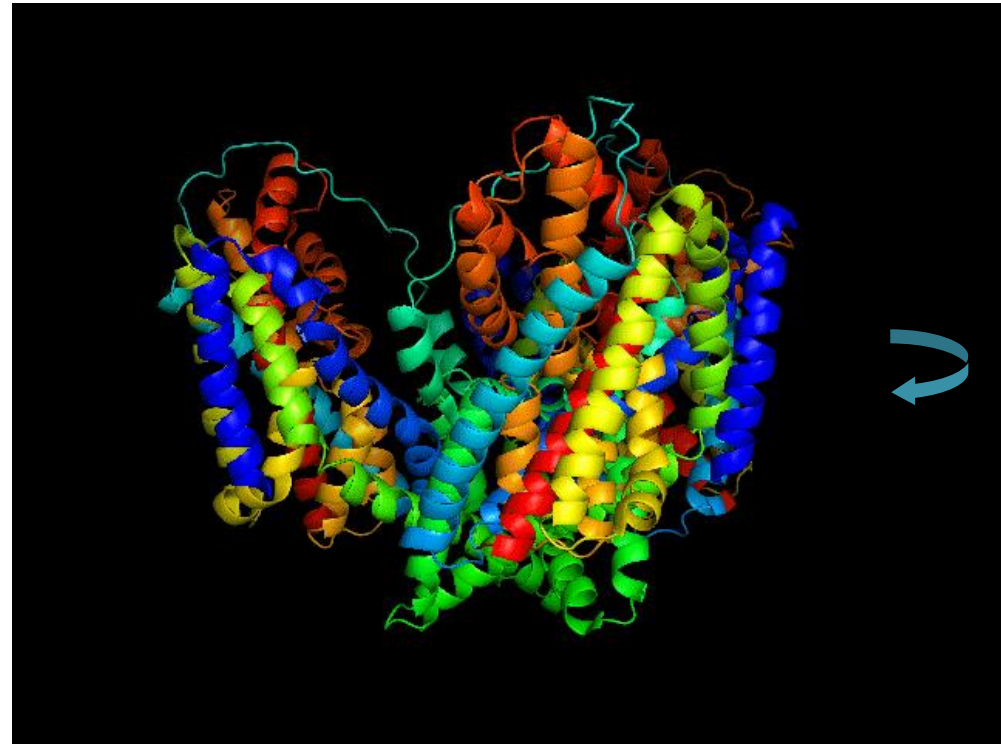

Dutta A, Krieger J, Lee JY, Garcia-Nafria J, Greger IH, Bahar I (2015) Cooperative Dynamics of Intact AMPA and NMDA Glutamate Receptors: Similarities and Subfamily-Specific Differences *Structure* 23: 1692-170

Glutamate transporter - Glt$_{ph}$ structure
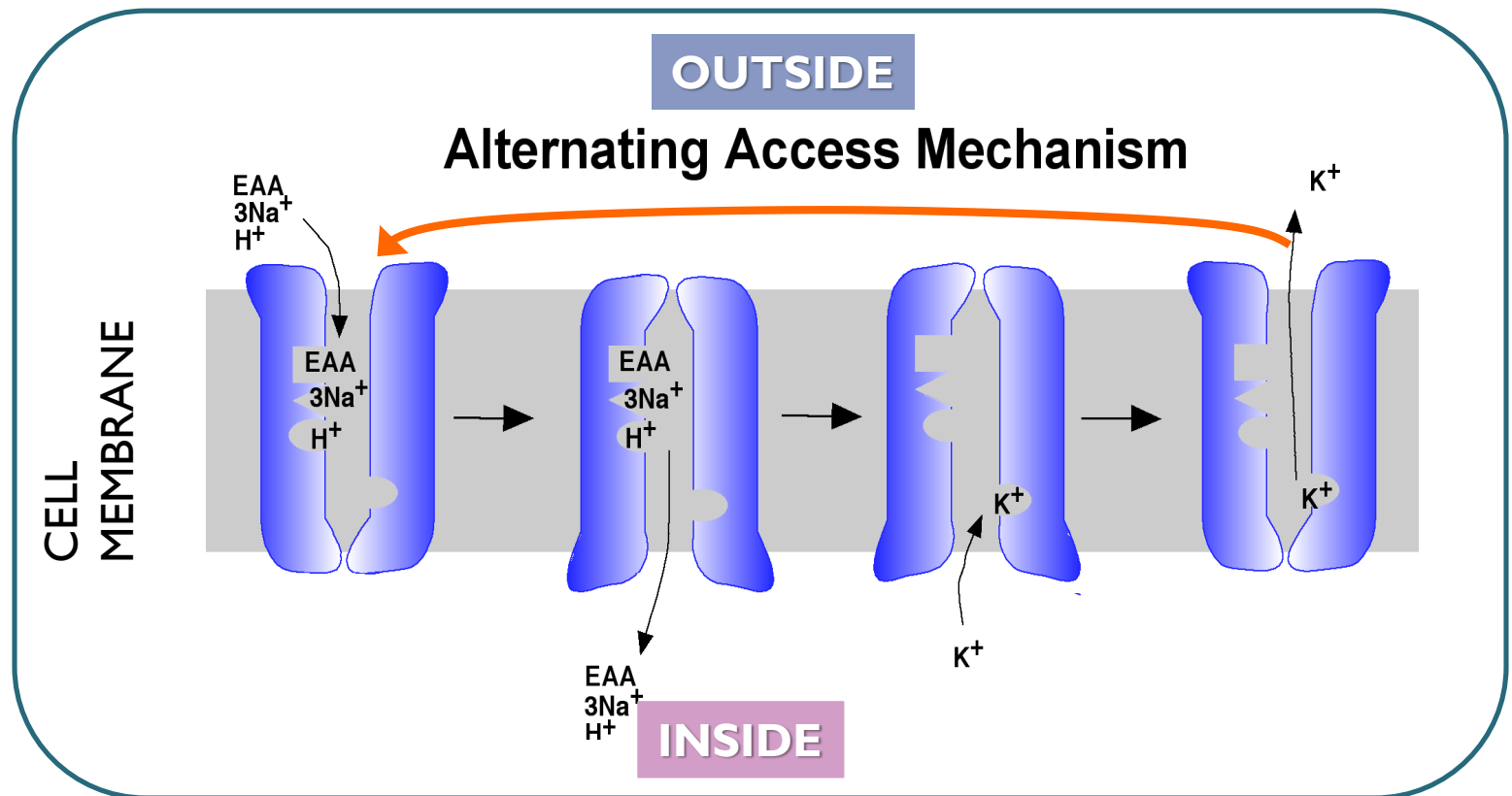

Trimeric aspartate transporter

(also called EAAT – for excitatory amino acid transporter)
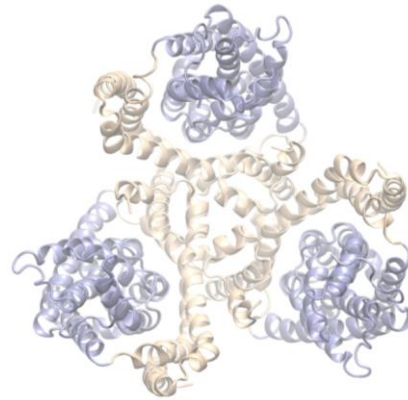
Alternating-access model:

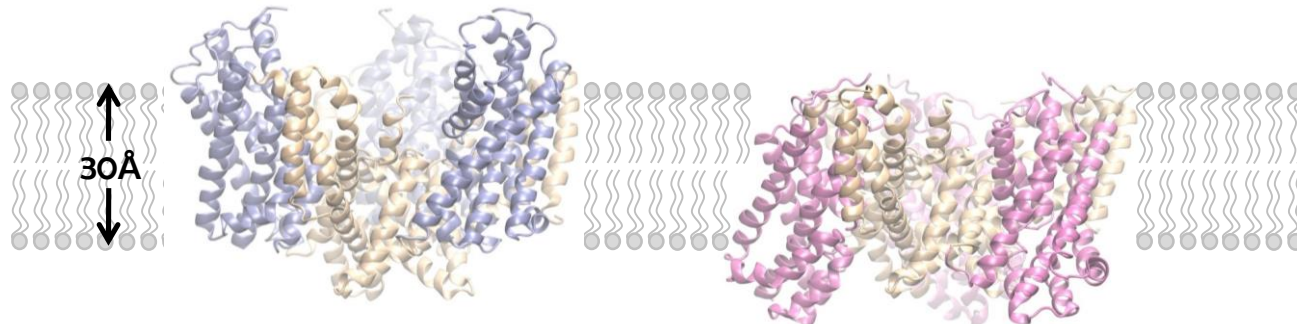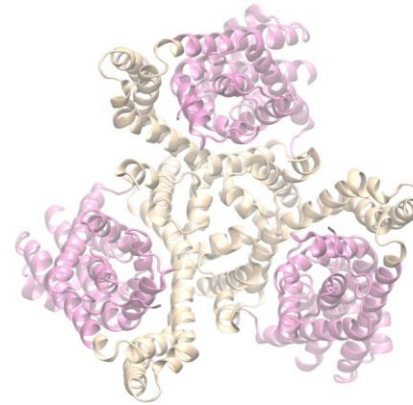Transitions between outward-facing and inward facing conformers

Global transitions

Glutamate transporters – trimeric membrane proteins, each subunit composed of two domains: **scaffolding** + **transport** domains
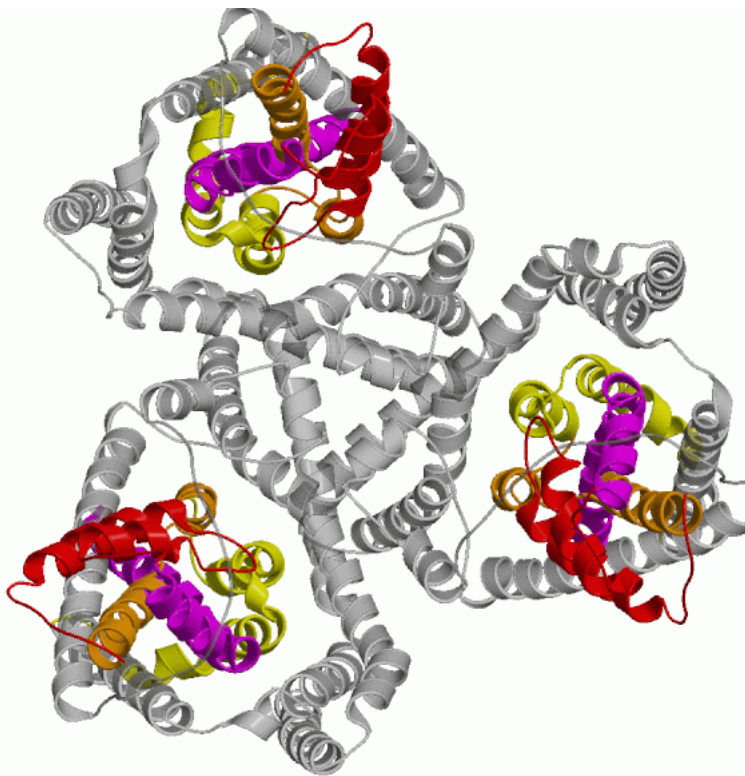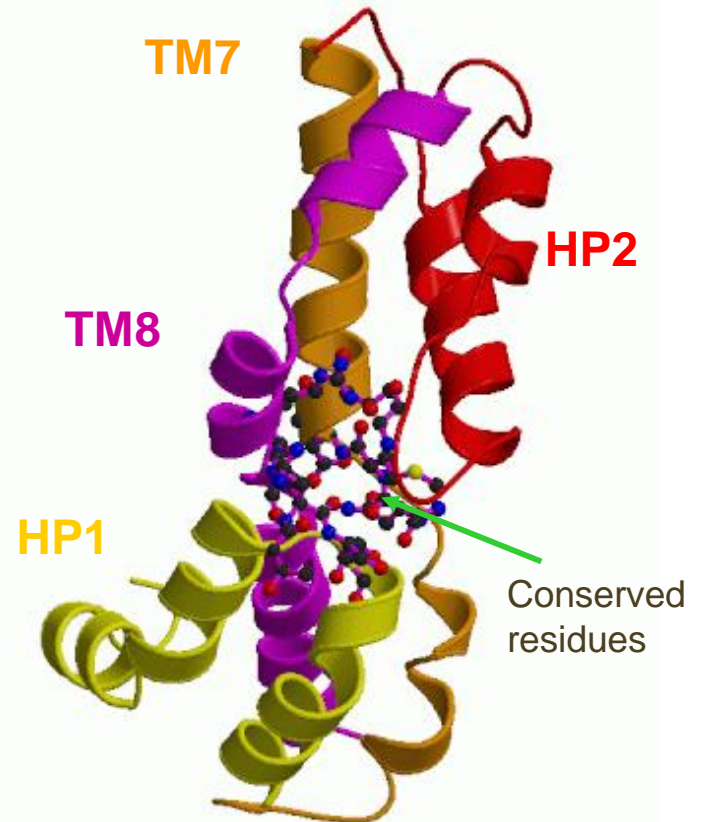
Outward Facing (OF)   Inward Facing (IF)

30Å

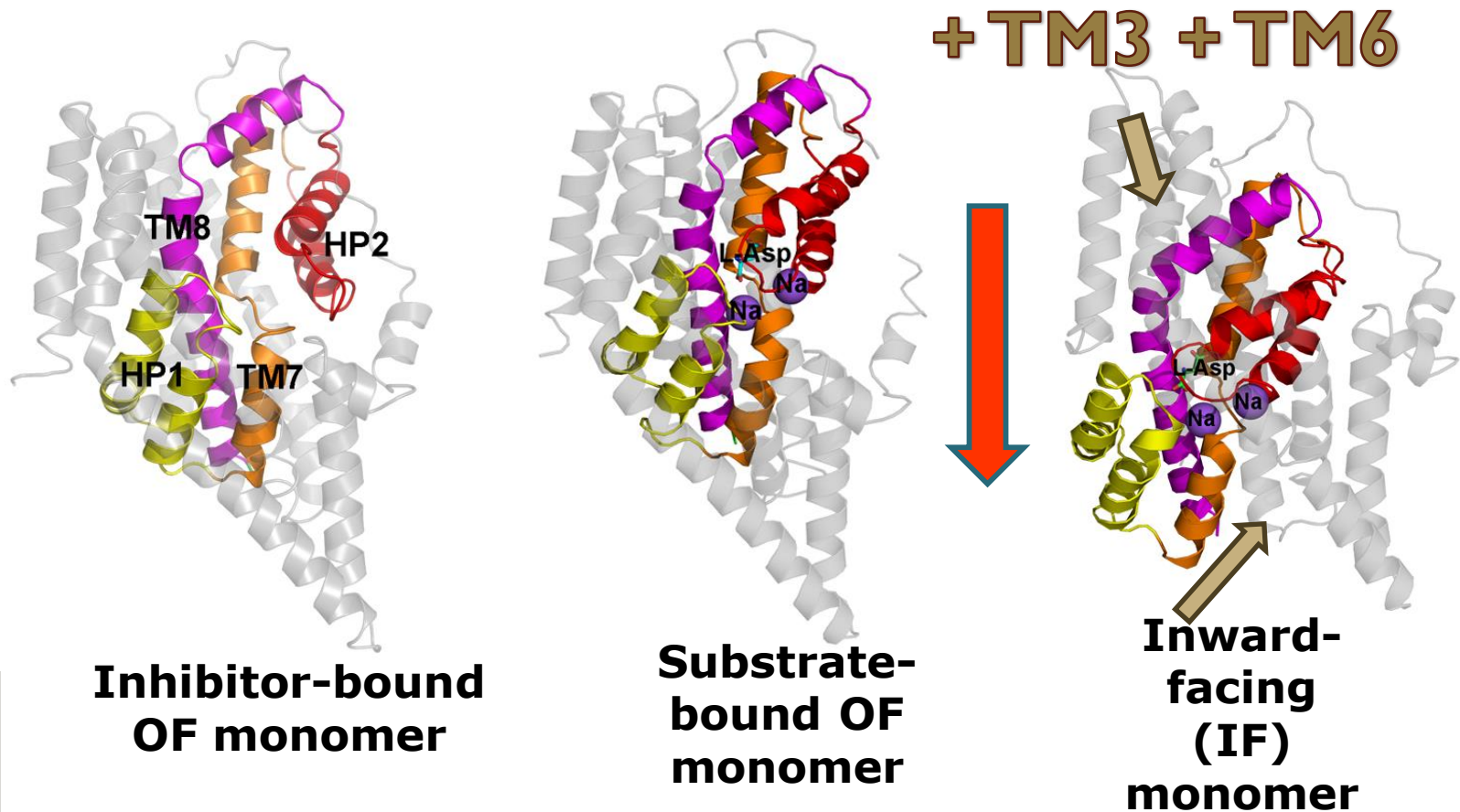# Each monomer is composed of a C-terminal core and a trimerization domain

N-terminal (trimerization) domains are shown in gray; TM1–TM6)

TM7

TM8

HP2

HP1

Conserved residues

C-terminal Core (HP1-TM7-HP2-TM8)

Yernool et al. 2004 *Nature* **431**, 811.

En bloc movement of transport core

+ TM3 + TM6

Inhibitor-bound OF monomer

Substrate-bound OF monomer

Inward-facing (IF) monomer

Translation of transport core by ~ 15Å

Reyes, N., Ginter, C., and **Boudker, O.** (2009) *Nature* **462,** 880-885;
Crisman TJ, Qu S, Kanner BI, **Forrest LR.** (2009) *Proc Natl Acad Sci U S A*. 106:20752-7.

Global transitions

Transport domain undergoes elevator-like motions, while the trimerization scaffolding domains is rigidly affixed to the membrane

Single subunit showing the transport domain moving across the membrane
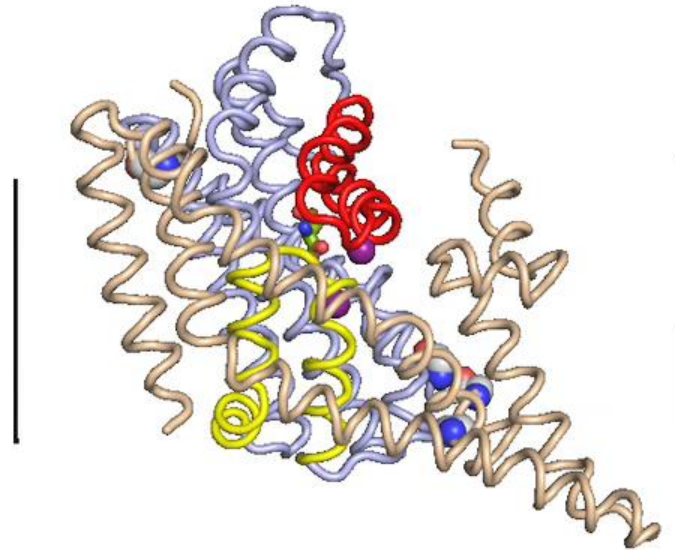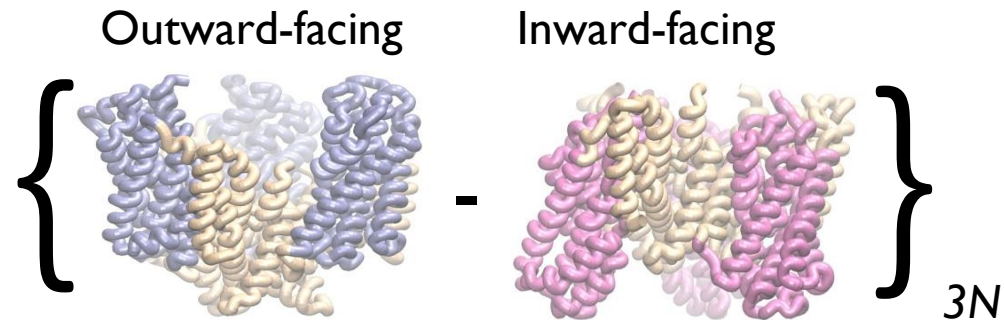
Substates are sampled along soft modes

S1
Outward-facing

Inward-facing

S2
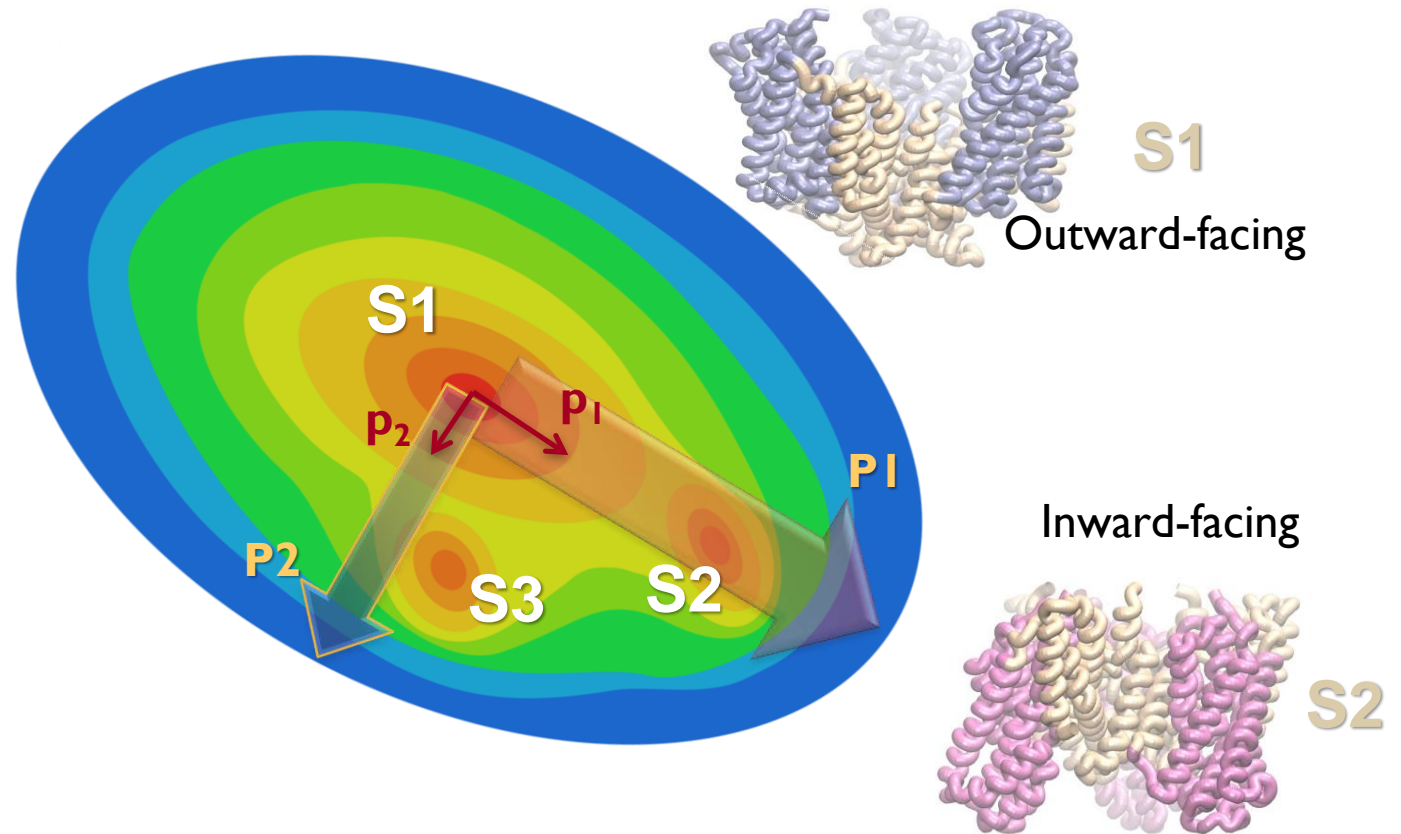
Is this transition along a soft mode?

Correlation between soft modes & observed structural change
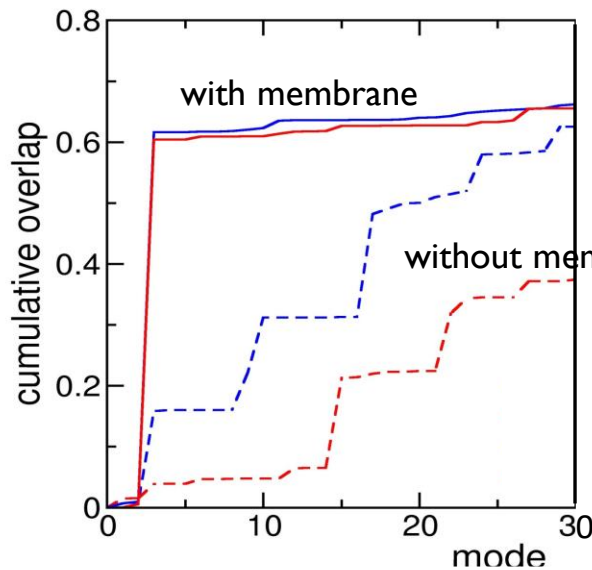
1. Evaluate of the deformation vector $\boldsymbol{d}_{3N}$

Outward-facing   Inward-facing

$$\left\{ \quad - \quad \right\}_{3N}$$

2. Calculate the correlation cosine $\cos(\boldsymbol{v}_i, \boldsymbol{d})$ between each mode vector, $\boldsymbol{v}_i$, and $\boldsymbol{d}$

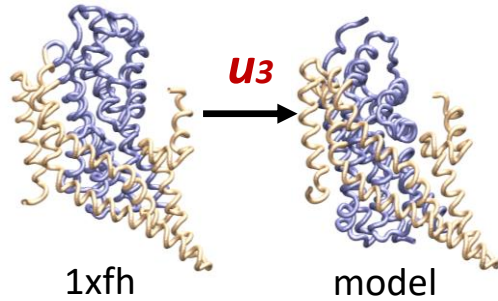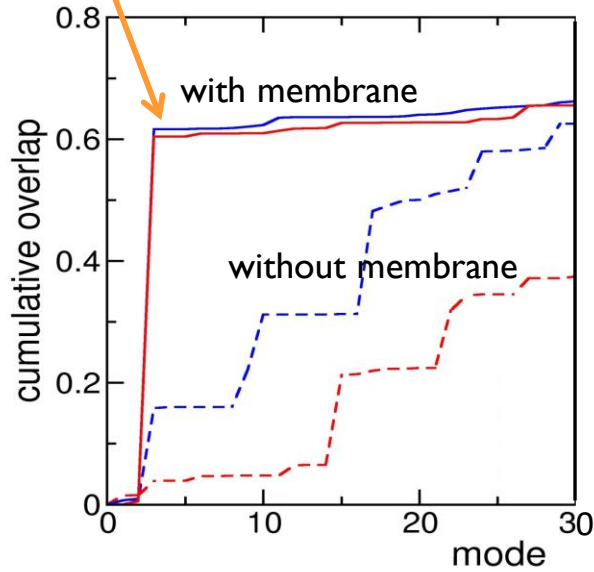3. **Cumulative overlap** $= \ [\sum_i \cos^2(\boldsymbol{v}_i, \boldsymbol{d})]^{\frac{1}{2}}$

Reference:

Lezon TR, Bahar I. (2012) Constraints imposed by the membrane selectively guide the alternating access dynamics of the glutamate transporter Glt$_{Ph}$. *Biophys J.* **102**:1331-40.

Correlation between soft modes & observed structural change

with membrane

without membrane

Dr. Timothy R. Lezon

1. Evaluate of the deformation vector d_{3N}

Outward-facing   Inward-facing

$$\left\{ \text{[Outward-facing]} - \text{[Inward-facing]} \right\}_{3N}$$

2. Calculate the correlation cosine $\cos(v_i, d)$ between each mode vector, v_i, and d

3. **Cumulative overlap** = $\left[\sum_i \cos^2(v_i, d) \right]^{1/2}$

Reference:

Lezon TR, Bahar I. (2012) Constraints imposed by the membrane selectively guide the alternating access dynamics of the glutamate transporter Glt$_{Ph}$. *Biophys J.* **102**:1331-40.

Overlap of > 0.6 achieved with a single mode!

Softest nondegenerate mode out of > 3,000 modes

cumulative overlap

with membrane

without membrane
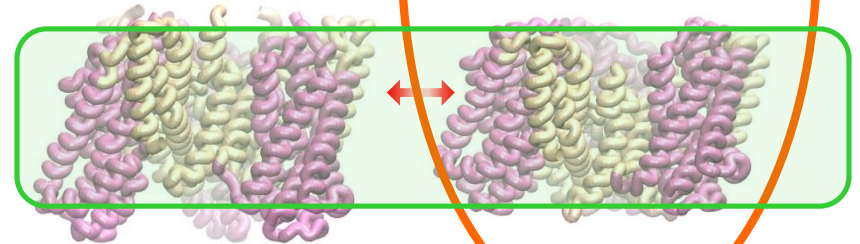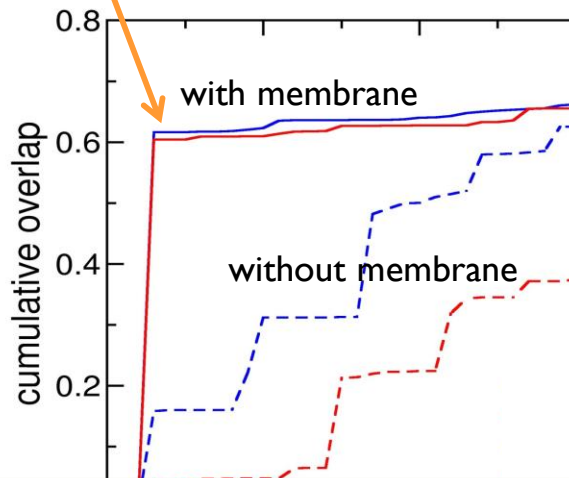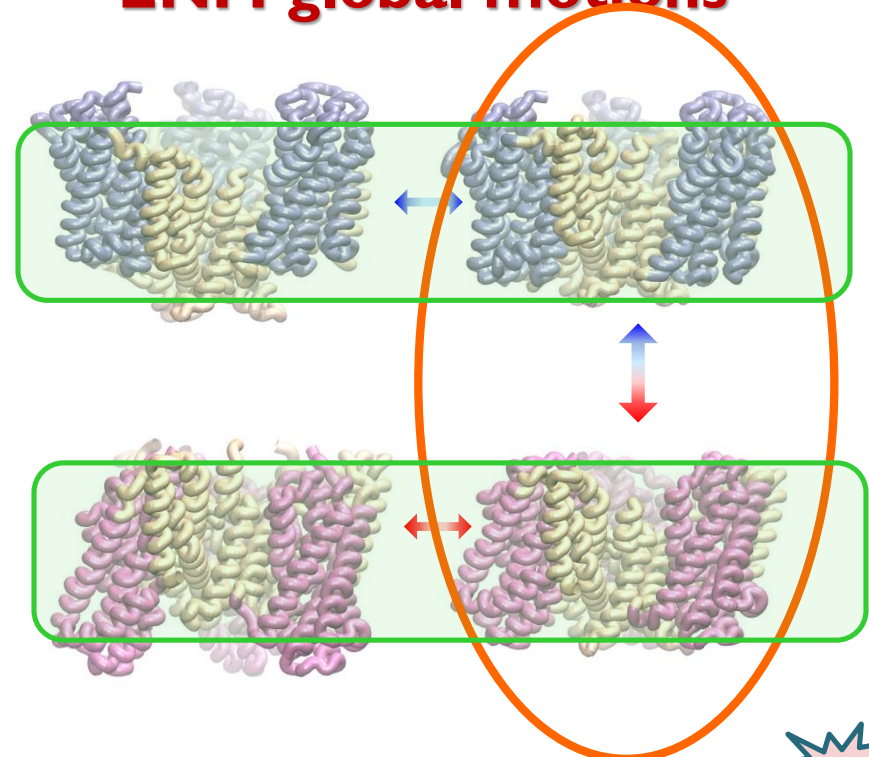
mode

u_3

1xfh → model

ENM global motions

Outward-facing
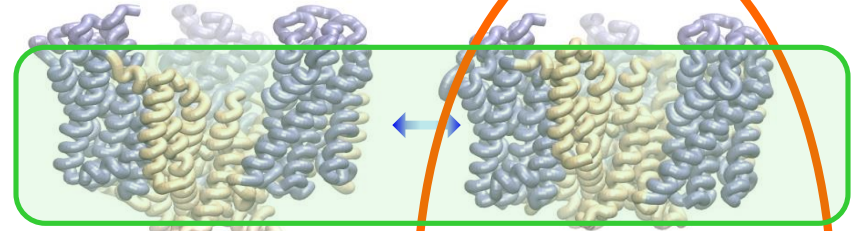
Inward-facing

in the presence of lipid constraints

Reference:

Lezon TR, Bahar I. (2012) Constraints imposed by the membrane selectively guide the alternating access dynamics of the glutamate transporter Glt_Ph. *Biophys J.* **102**:1331-40.

Overlap of > 0.6 achieved with a single mode!

Softest nondegenerate mode out of > 3,000 modes

ENM global motions

If the predicted modes were 'random', each mode would contribute by $1/3N$ to the cumulative overlap, i.e.

$$\cos(v_k \cdot d_{exp}) = (1/3N)^{1/2} = 0.0167$$

in the presence of lipid constraints

Reference:

Lezon TR, Bahar I. (2012) Constraints imposed by the membrane selectively guide the alternating access dynamics of the glutamate transporter Glt_{Ph}. *Biophys J.* **102**:1331-40.

Two approaches for including the lipid bilayer:

Explicit membrane (a network model for the membrane)
Implicit membrane (change in Hessian force constants)
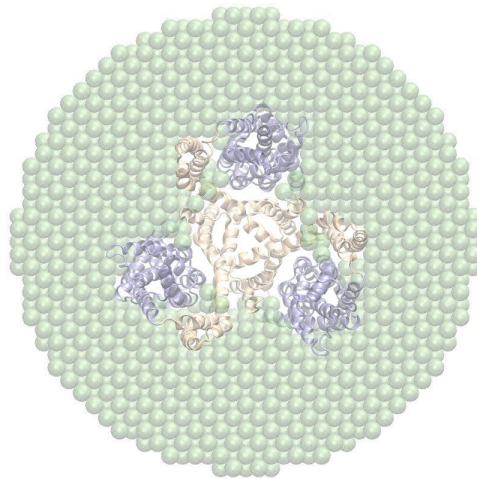
Explicit membrane top view
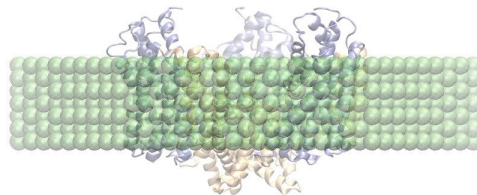
Explicit membrane side view
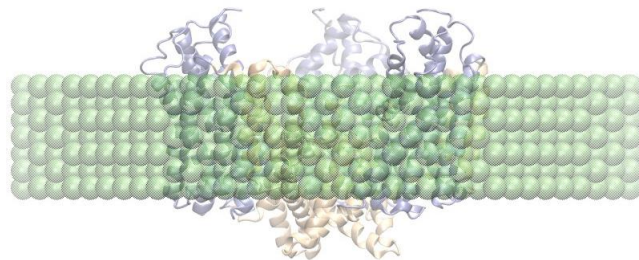
System/environment approximation

"system"

H^{ss}	H^{ES}
H^{SE}	H^{EE}
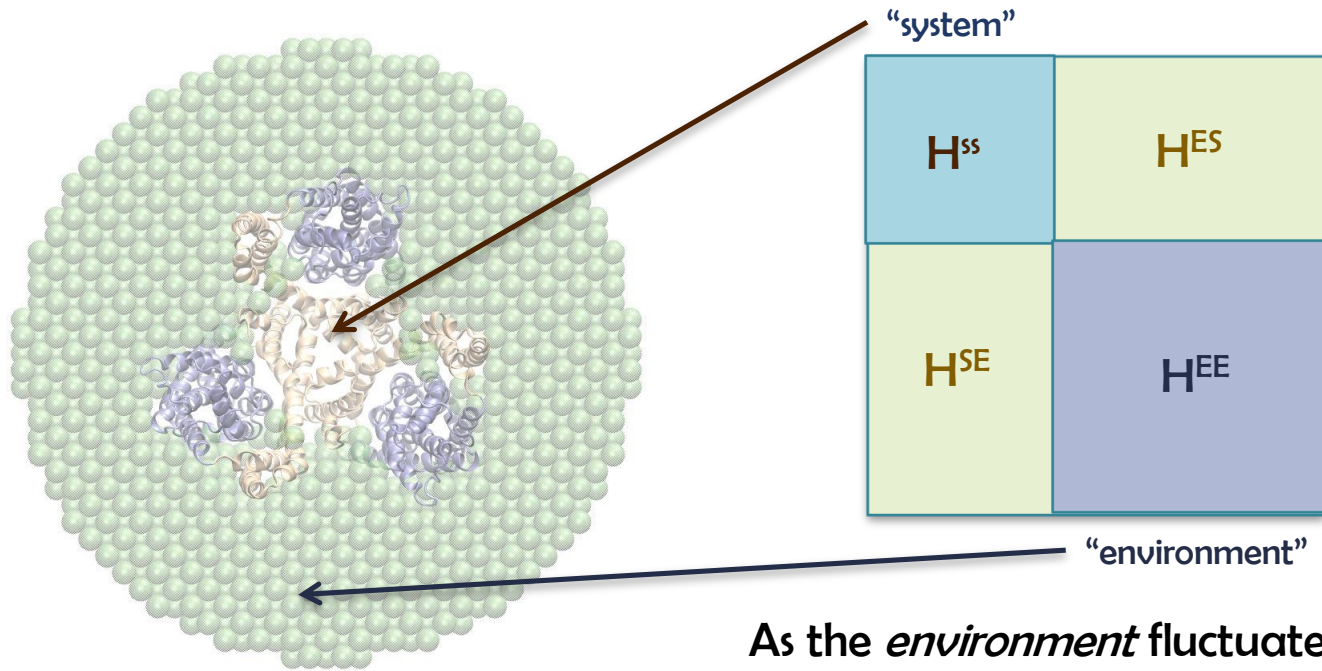
"environment"

As the *environment* fluctuates randomly, the effective motion of the *system* is given by

$$V_{eff}\left(\mathbf{s}\right) = \tfrac{1}{2}\Delta\mathbf{s}^{T}\left(\mathbf{H}^{\mathbf{ss}\,\prime}\right)\Delta\mathbf{s}$$

$$\mathbf{H}^{\mathbf{ss}\,\prime} = \mathbf{H}^{\mathbf{ss}} - \mathbf{H}^{\mathbf{SE}}\left(\mathbf{H}^{\mathbf{EE}}\right)^{-1}\mathbf{H}^{\mathbf{ES}}$$

reduceModel()

Implicit model for membrane effect

$$\mathbf{H_{ij}} = -\frac{\gamma}{\left(R_{ij}^0\right)^2}\begin{bmatrix} \left(x_{ij}^0\right)^2 & x_{ij}^0 y_{ij}^0 & x_{ij}^0 z_{ij}^0 \\ x_{ij}^0 y_{ij}^0 & \left(y_{ij}^0\right)^2 & y_{ij}^0 z_{ij}^0 \\ x_{ij}^0 z_{ij}^0 & y_{ij}^0 z_{ij}^0 & \left(z_{ij}^0\right)^2 \end{bmatrix}$$

Altered radial force constants:

$$\mathbf{H_{ij}} = -\left(R_{ij}^0\right)^{-2}\begin{bmatrix} \left(x_{ij}^0\sqrt{\gamma_x}\right)^2 & x_{ij}^0 y_{ij}^0\sqrt{\gamma_x\gamma_y} & x_{ij}^0 z_{ij}^0\sqrt{\gamma_x\gamma_z} \\ x_{ij}^0 y_{ij}^0\sqrt{\gamma_x\gamma_y} & \left(y_{ij}^0\sqrt{\gamma_y}\right)^2 & y_{ij}^0 z_{ij}^0\sqrt{\gamma_y\gamma_z} \\ x_{ij}^0 z_{ij}^0\sqrt{\gamma_x\gamma_z} & y_{ij}^0 z_{ij}^0\sqrt{\gamma_y\gamma_z} & \left(z_{ij}^0\sqrt{\gamma_z}\right)^2 \end{bmatrix}$$

$$\mathbf{H_{ij}} = -\frac{g}{\left(R_{ij}^0\right)^2}\begin{bmatrix} \left(x_{ij}^0\right)^2 & x_{ij}^0 y_{ij}^0 & cx_{ij}^0 z_{ij}^0 \\ x_{ij}^0 y_{ij}^0 & \left(y_{ij}^0\right)^2 & cy_{ij}^0 z_{ij}^0 \\ cx_{ij}^0 z_{ij}^0 & cy_{ij}^0 z_{ij}^0 & \left(cz_{ij}^0\right)^2 \end{bmatrix}$$
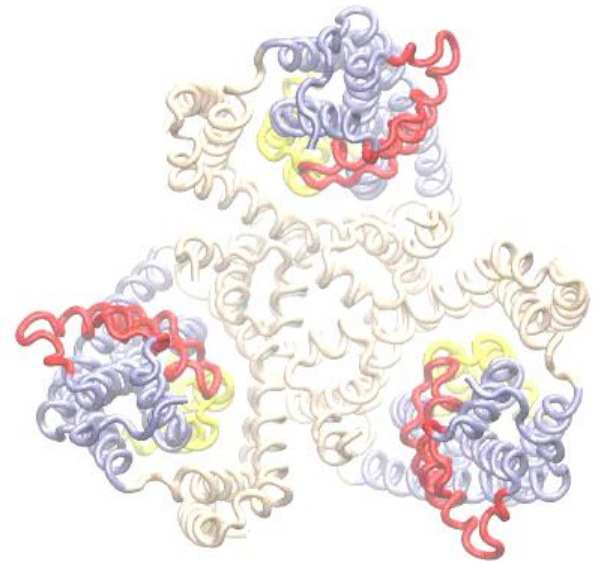
RTB.buildHessian()

Lezon & Bahar. Biophys J 102 (2012).

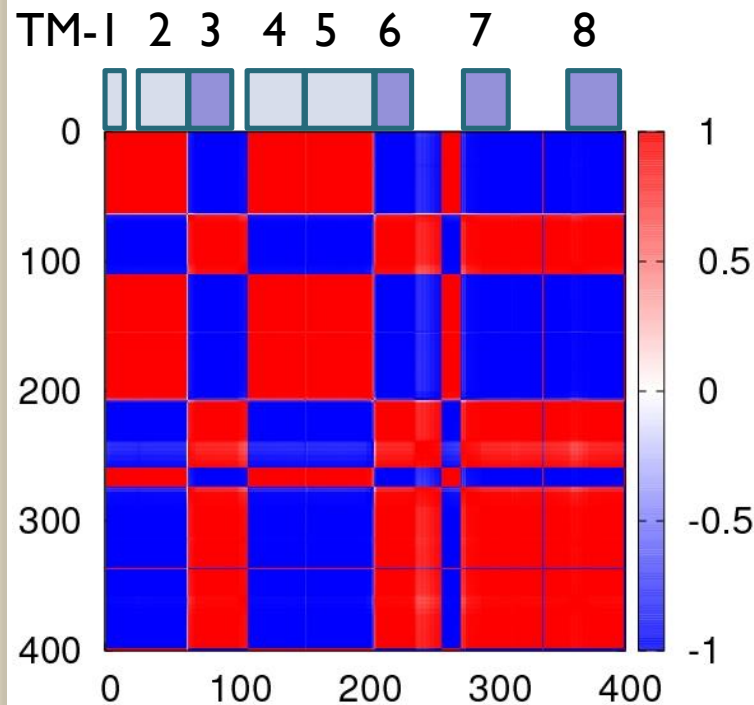# Lipid bilayer favors elevator-like motions

Dr. Timothy R Lezon
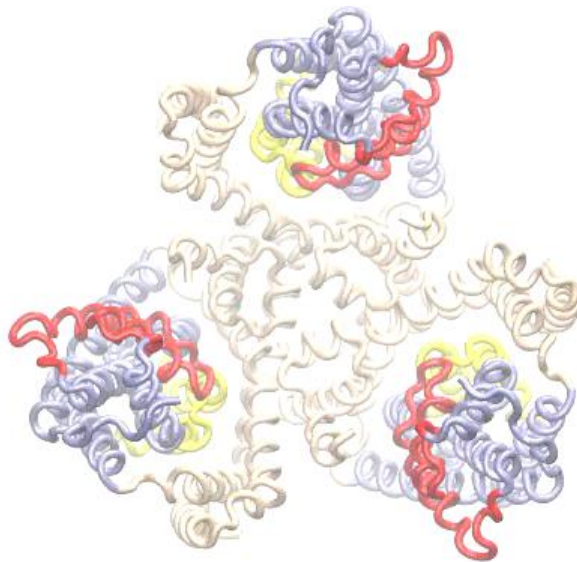
ANM in the absence of membrane
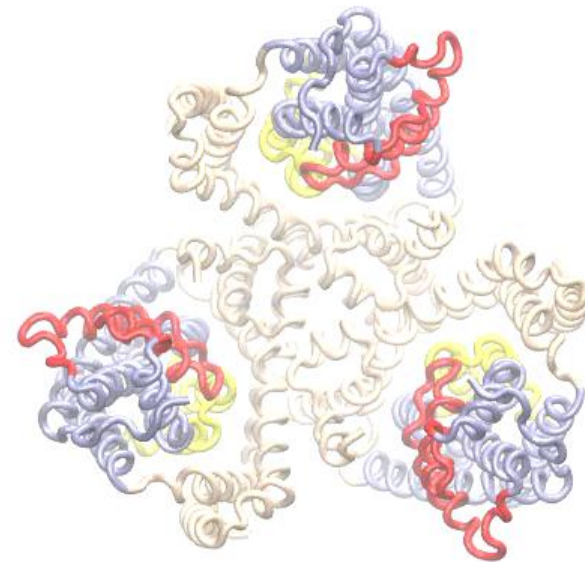
ANM in the presence of membrane

DOMAIN SEPARATION PREDICTED BY GNM

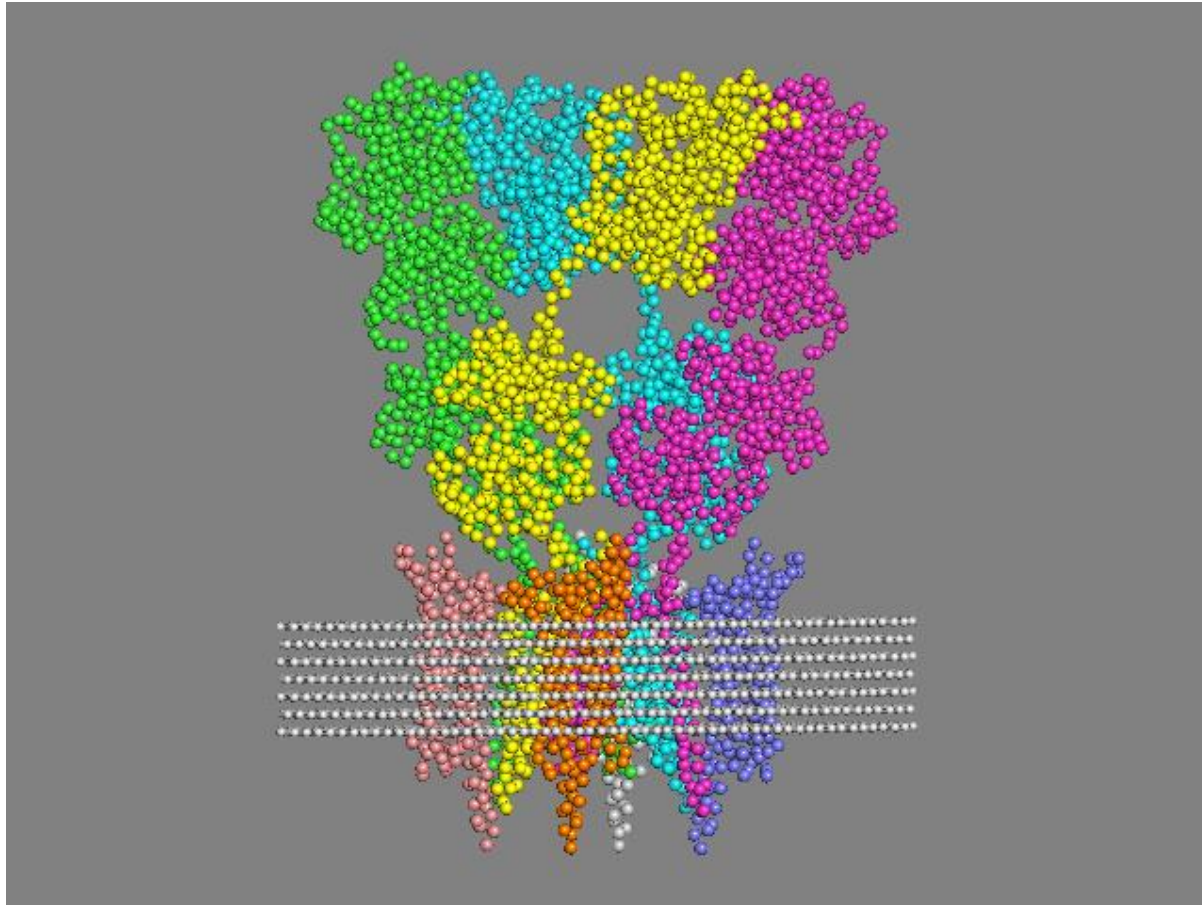**TM3 and TM6 form an integral part of the transport core (TM7, TM8, HP1 and HP2)**
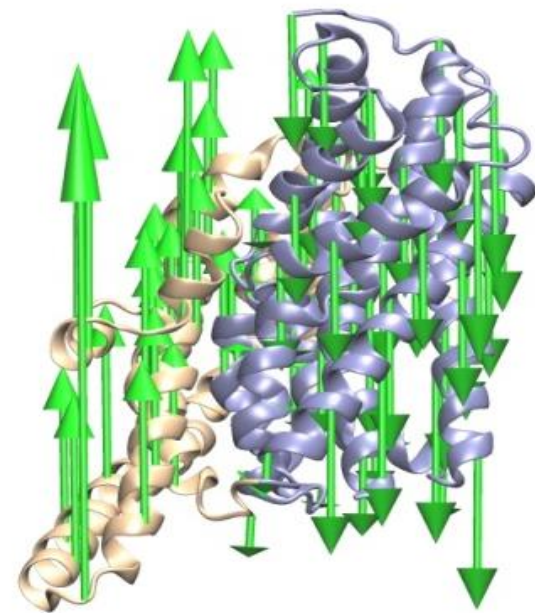
TM1, TM2, TM4, TM5

TM3, TM6, + transport core

Lezon TR, Bahar I. (2012) Constraints imposed by the membrane selectively guide the alternating access dynamics of the glutamate transporter Glt_{Ph}. *Biophys J*. **102**:1331-40.

MembrANM - DynOmics

Coupled motions of AMPAR and lipid bilayer

Motions without membrane

Mode4
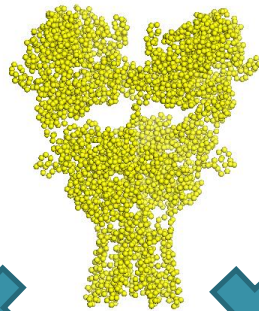
mode5
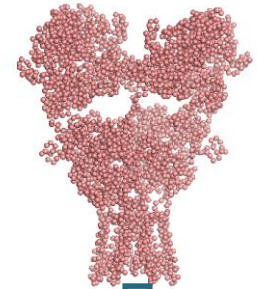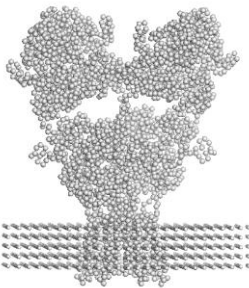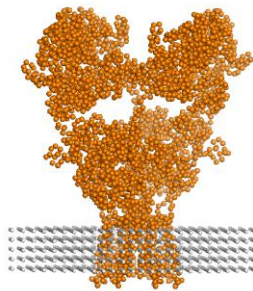
mode6 mode8 New mode mode9 mode10
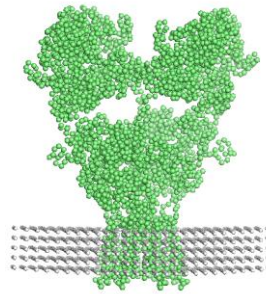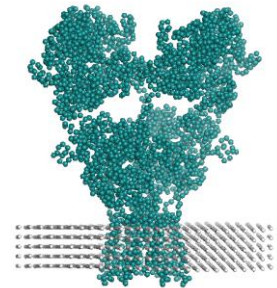
Motions with membrane

Comparison of mode shapes

without membrane (x-axis)
with membrane (ANM-mem) (y-axis)
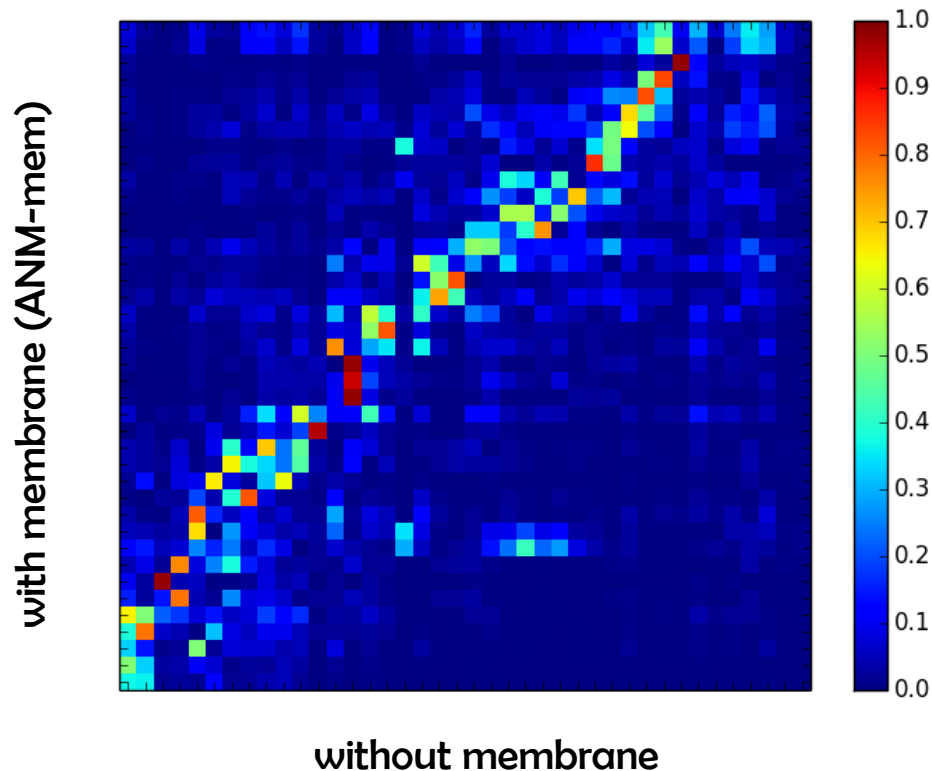
In the presence of membrane:

- Some intrinsic motions are preserved (*red*)
- Others are altered (*yellow*)
- New motions (mainly dominated by membrane fluctuations) emerge (blue rows)
- Diagonal shift due to new modes in the membrane

Druggability

Druggable Genome

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

Human Genome (21,000) genes

Druggable genome
3,000 genes

430+ kinases
600+ GPCRs

70+ kinases
100+ GPCRs
Drug Targets
600-1,500

Disease-related genes
~3,000 genes

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.

- Membrane
- Cytoplasm
- Organelles
- Extracellular
- Nucleus
- Unknown

33.0%
23.8%
18.9%
10.5%
8.4%
5.4%
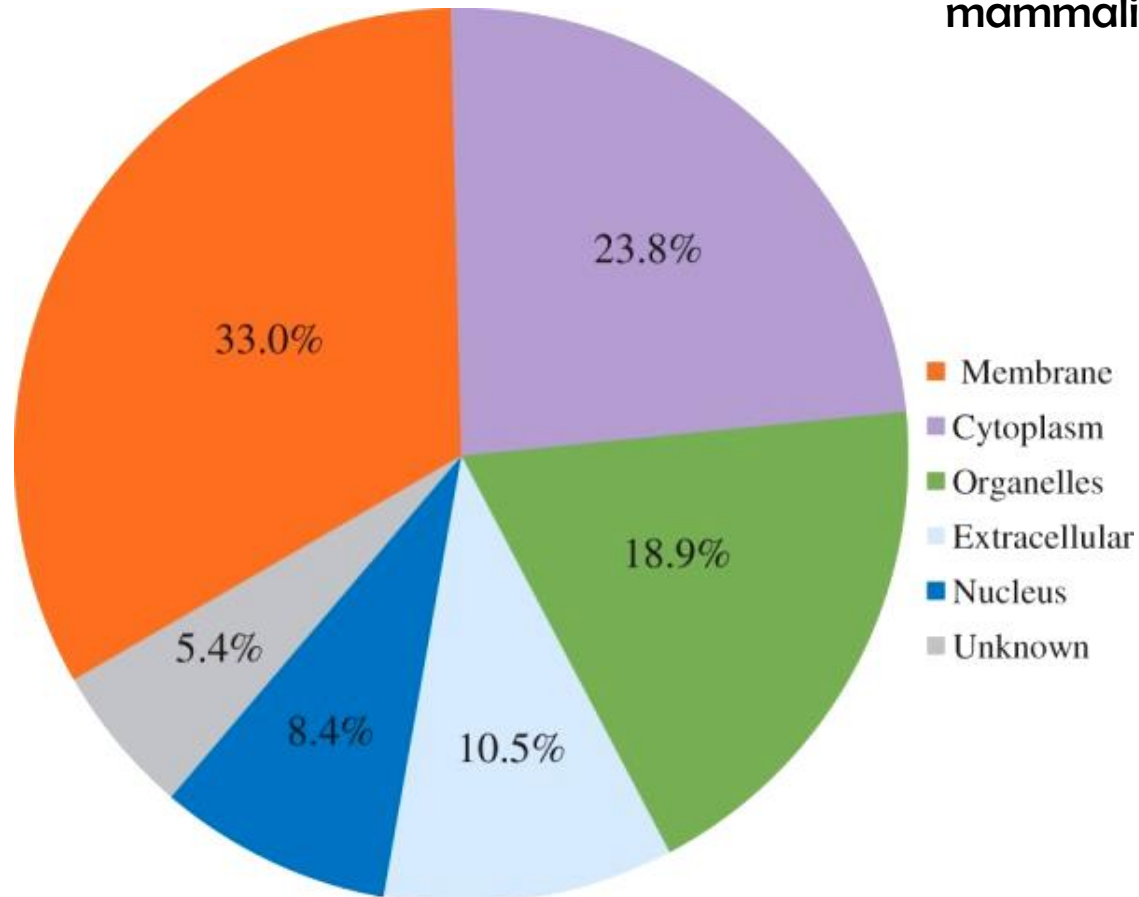
Rational Design of Inhibitors

3D structure of the target is used for

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

Druggable or not?

Active site druggability:
- Best known K_d 18.3 nM
- Simulation   0.03-0.5 nM
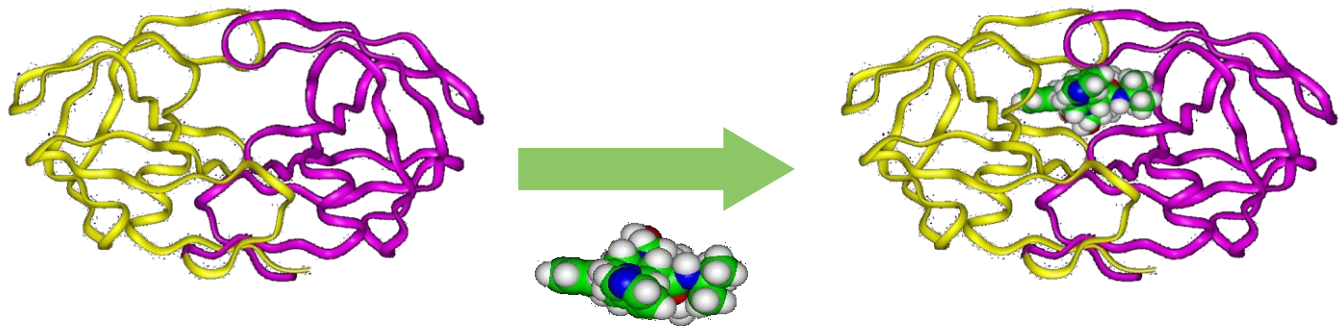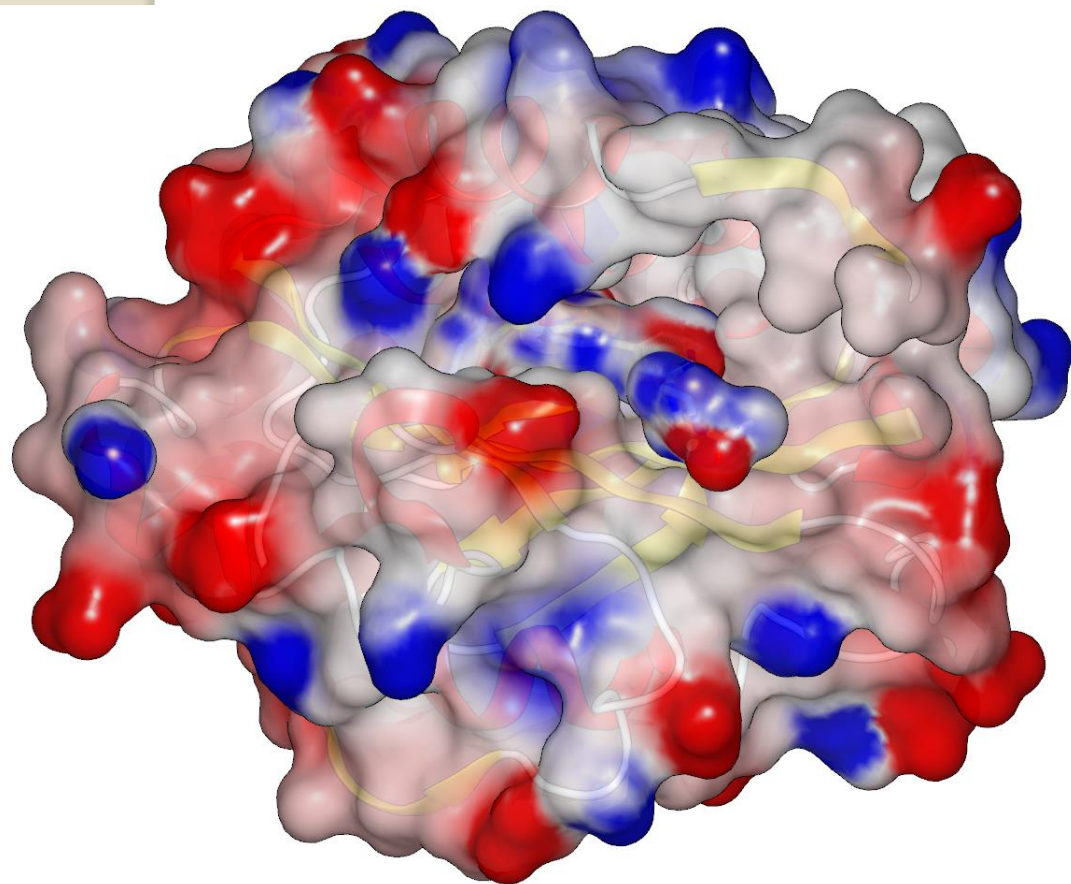
Lfa1 - a leukocyte glycoprotein that promotes intercellular adhesion and binds intercellular adhesion molecule 1
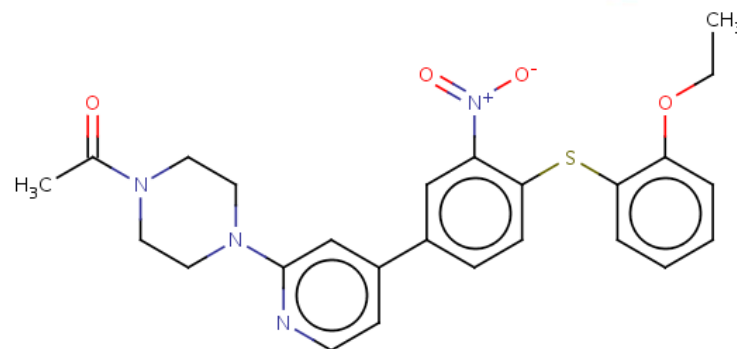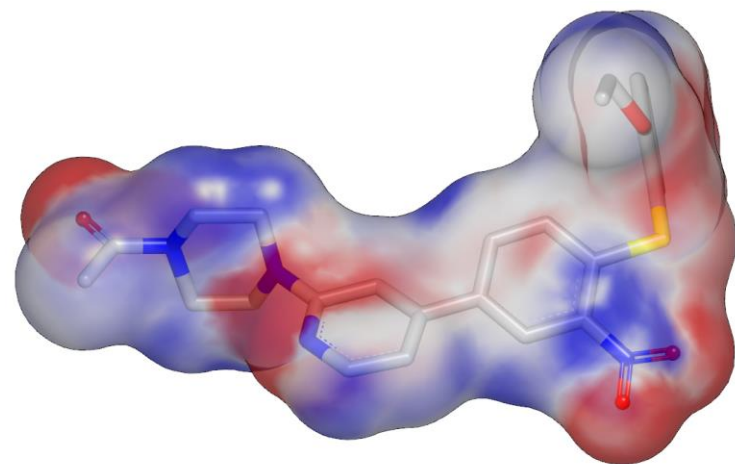
Some proteins do not present well-defined pockets
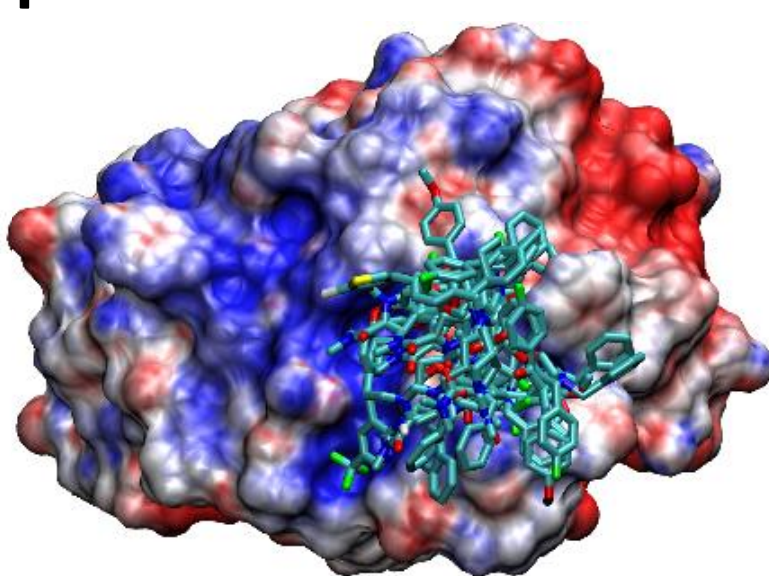
MKP-1
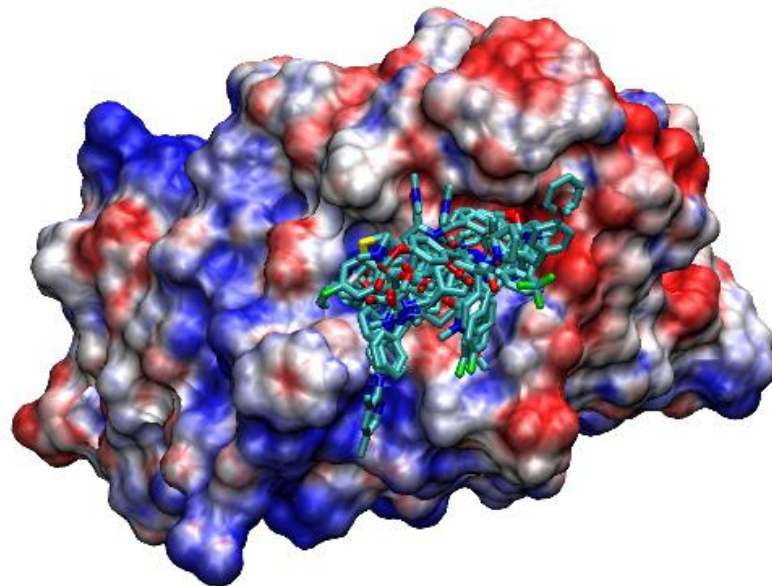
VHR

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
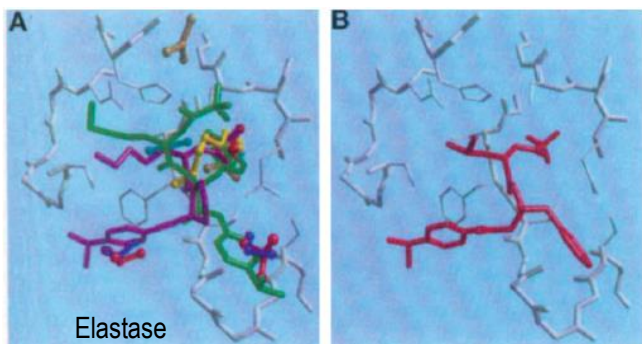
Mattos and Ridge, *Nat Biotechnology*, 1996

NMR screening

- compounds from a fragment-library are screened as mixtures of 20-30 compounds, druggability is calculated from chemical shift perturbations
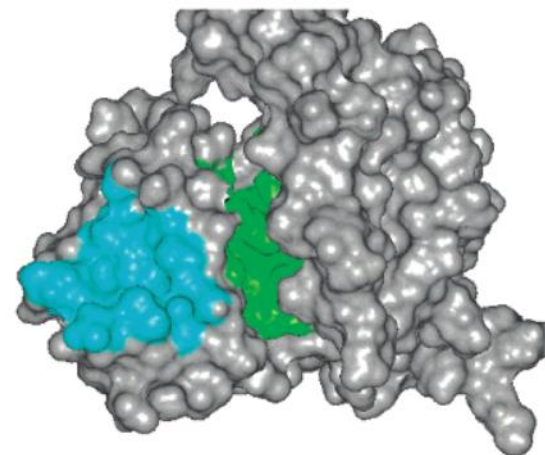
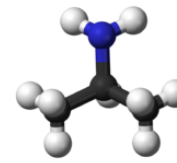Hajduk et al., *J Med Chem*, 2005

Druggability Simulations

Initial state

Simulation trajectory

MDM2

-2.4   -1.7   -1.0 kcal/mol

binding site

10,000 frames
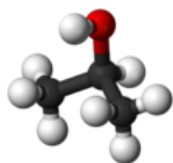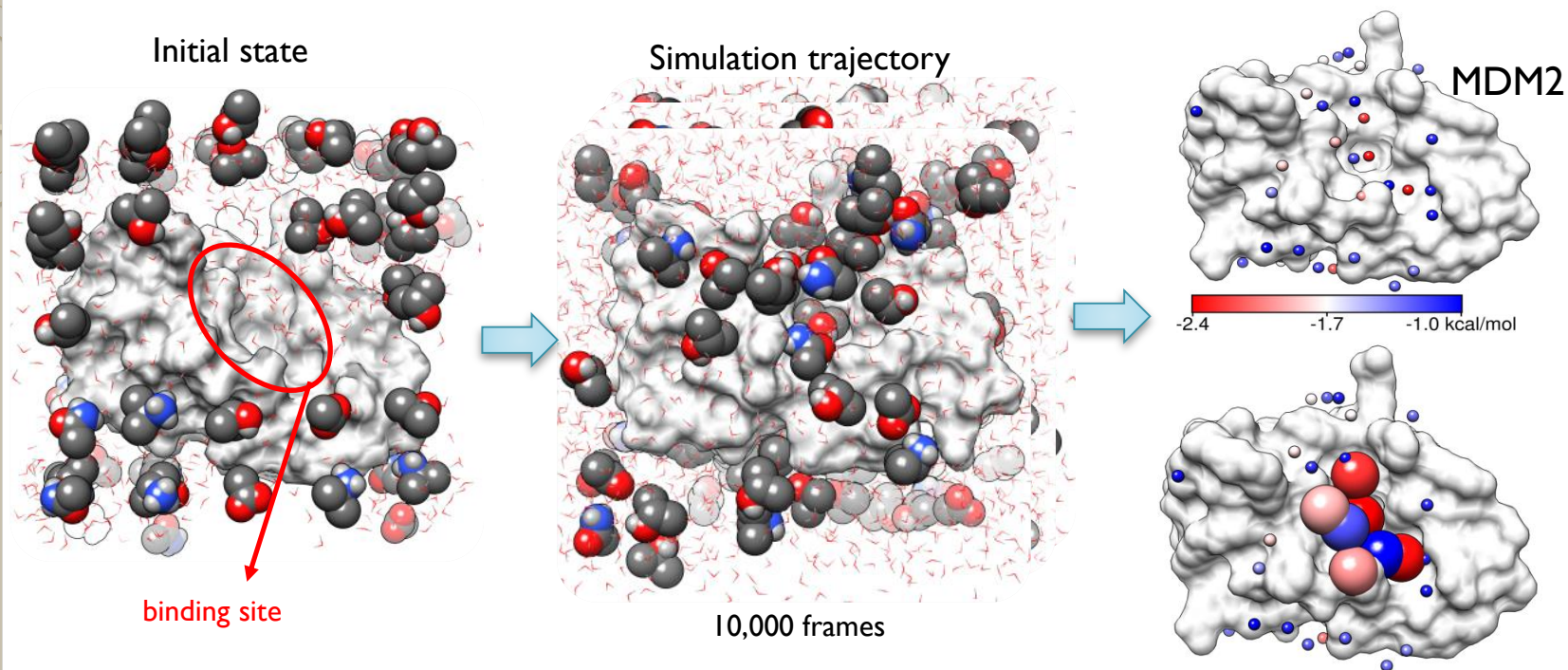
Isopropanol
(observed in 57% of drugs)
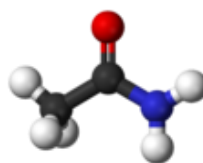
Acetamide
(21%)
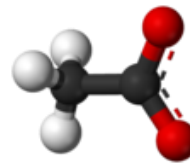
Acetate (-1)
(21%)

Isopropylamine (+1)
(25%)

Methodology Overview

From MD simulations to achievable drug affinities

binding site

10-20 ns

isopropanol
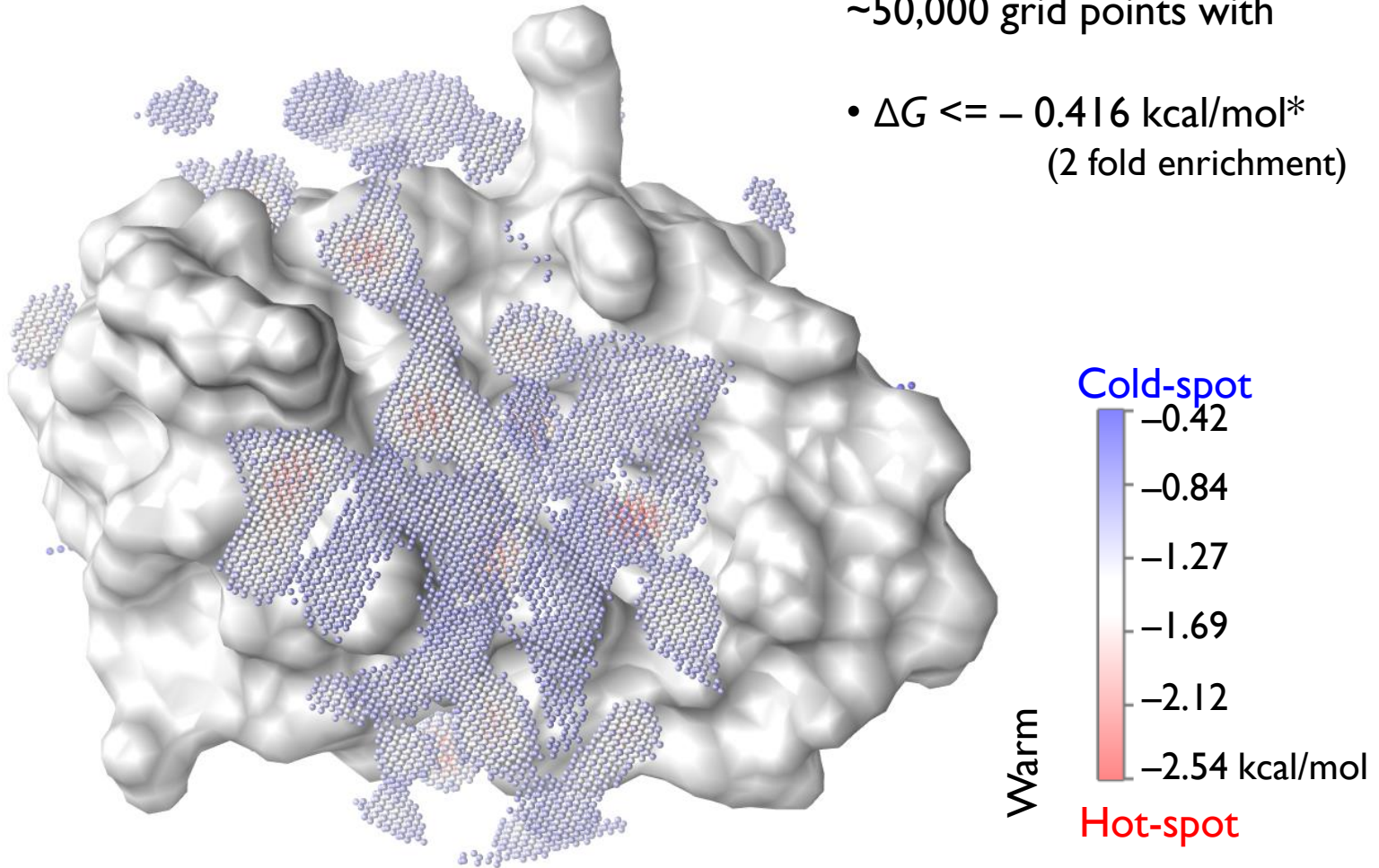
Identify hotspots

Detect binding sites

calculate achievable affinity
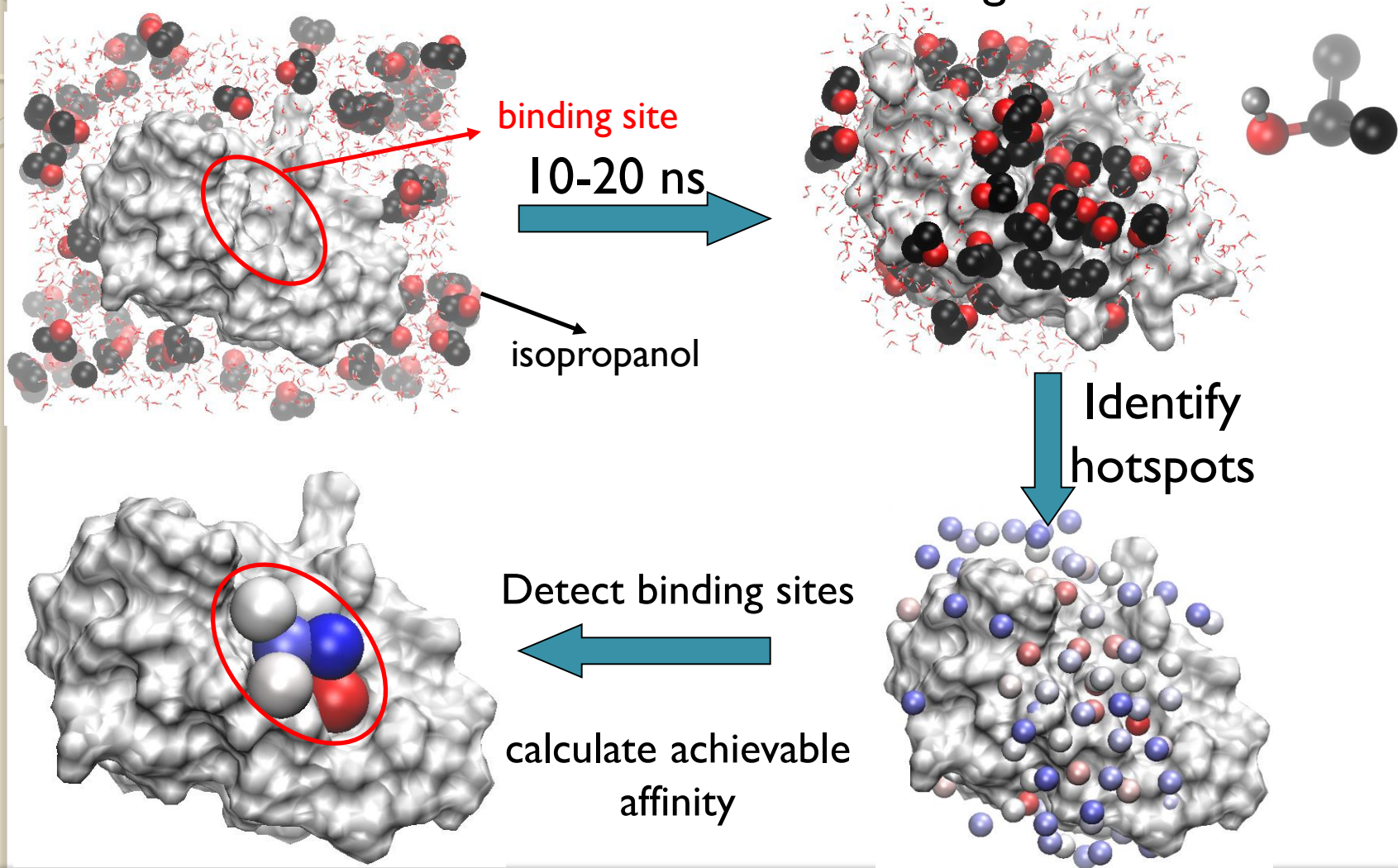
Annealing, Equilibration, Simulation
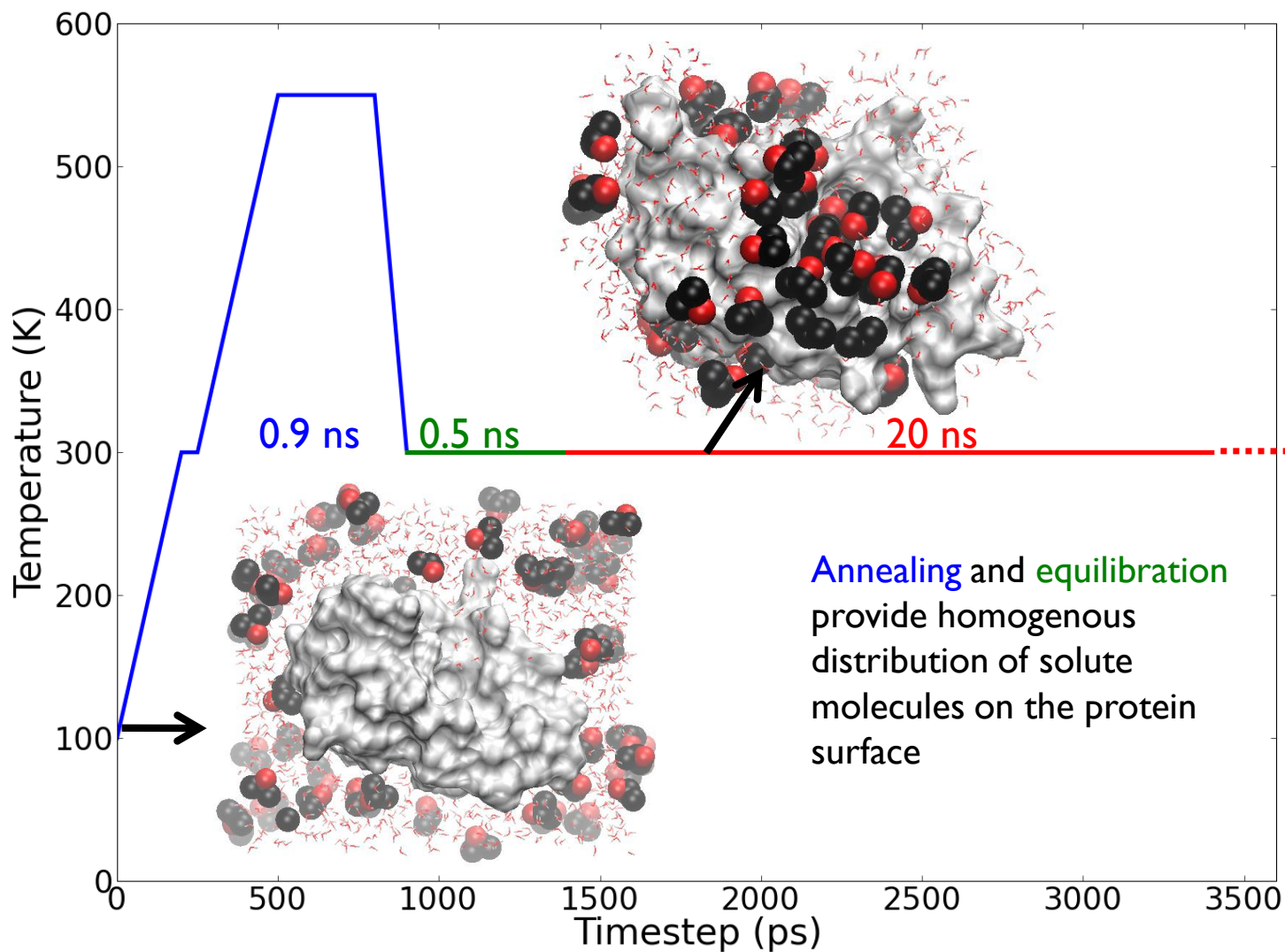
NAMD2 with CHARMM force field was used for simulations.

Isopropanol Binding Spots

ΔG grid is mapped onto the protein structure

~50,000 grid points with

- $\Delta G <= -0.416$ kcal/mol*
  (2 fold enrichment)

Cold-spot
−0.42
−0.84
−1.27
−1.69
−2.12
−2.54 kcal/mol

Warm

Hot-spot

Free Energy of Binding for Isopropanol

Assuming that MD sampling converged to a **Boltzmann ensemble**
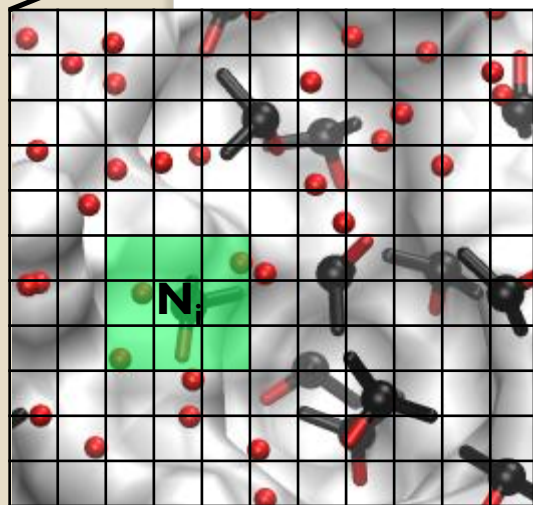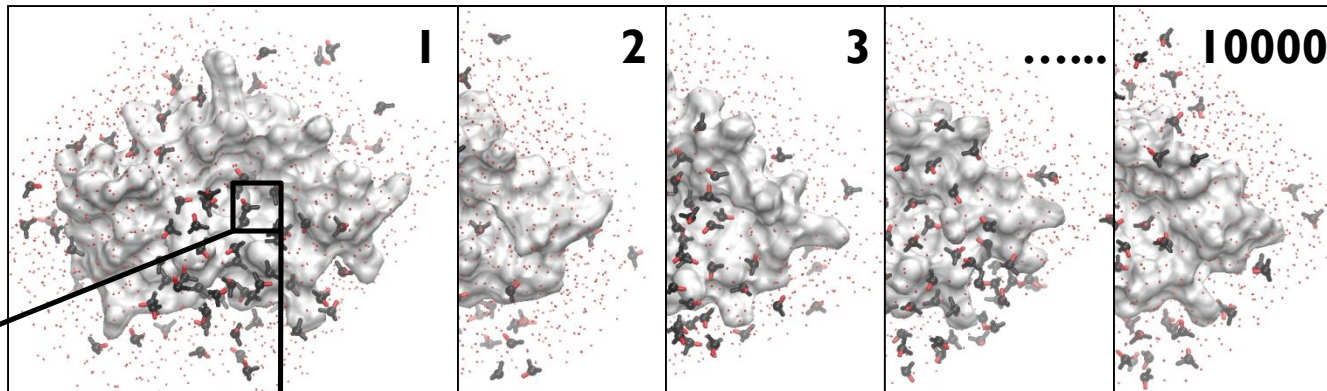
$$\Delta G_i = - RT \ln(N_i/N_0)$$

N_i = <u>observed number of isopropanols</u>
(# of frames) * (# of cubes)

N_0 = <u>total number of isopropanols</u>
total number of frames
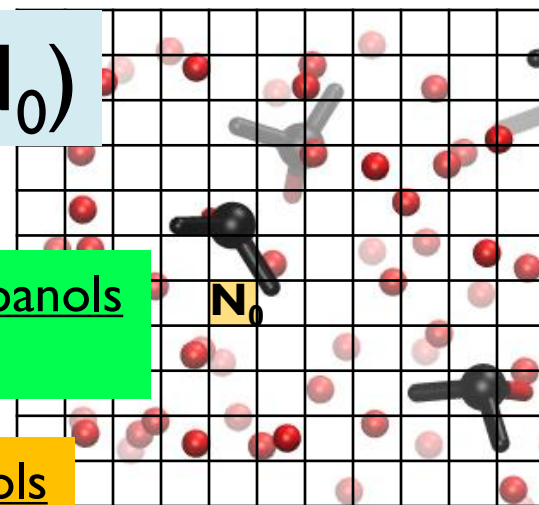
0.5 Å (not to scale)

Selecting Isopropanol Binding Spots

1. Grid element with lowest ∆G 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

191 points out of ~50,000 grid points

−0.42
−0.84
−1.27
−1.69
−2.12
−2.54 kcal/mol
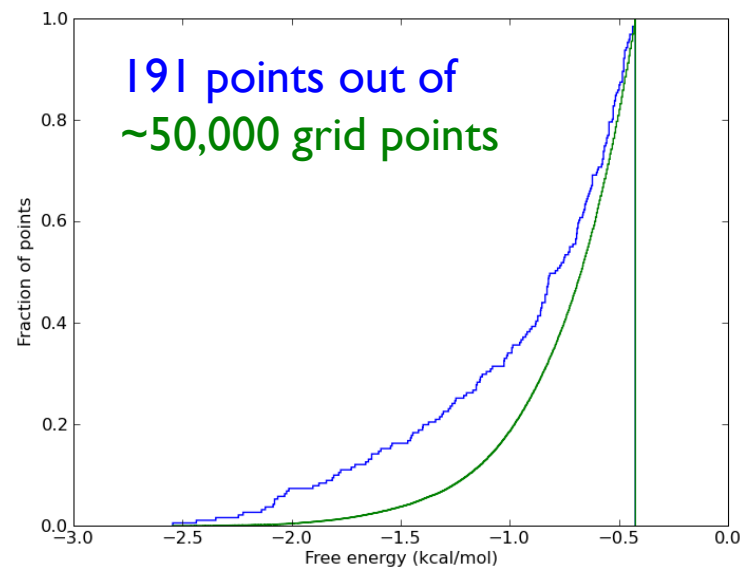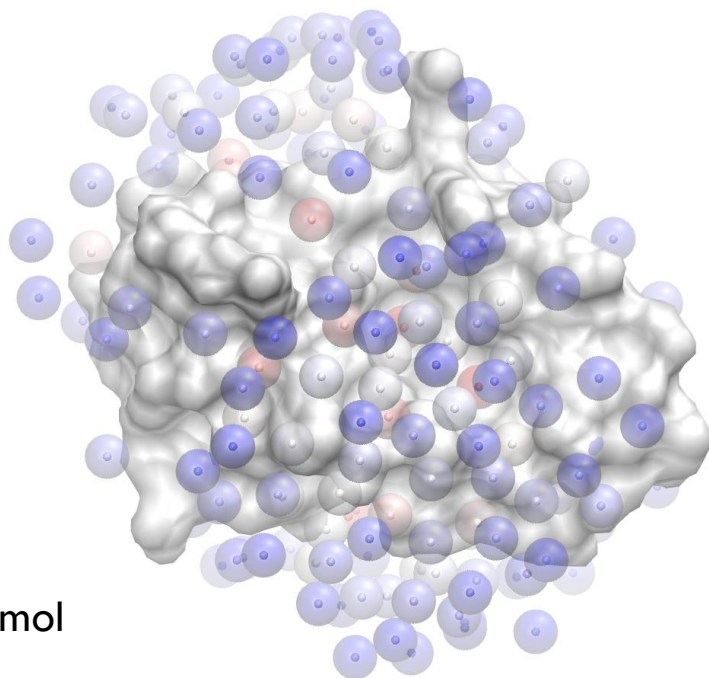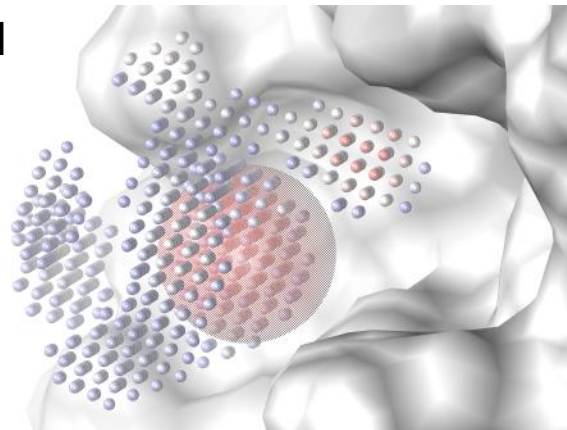
Affinity of a Drug-size Molecule

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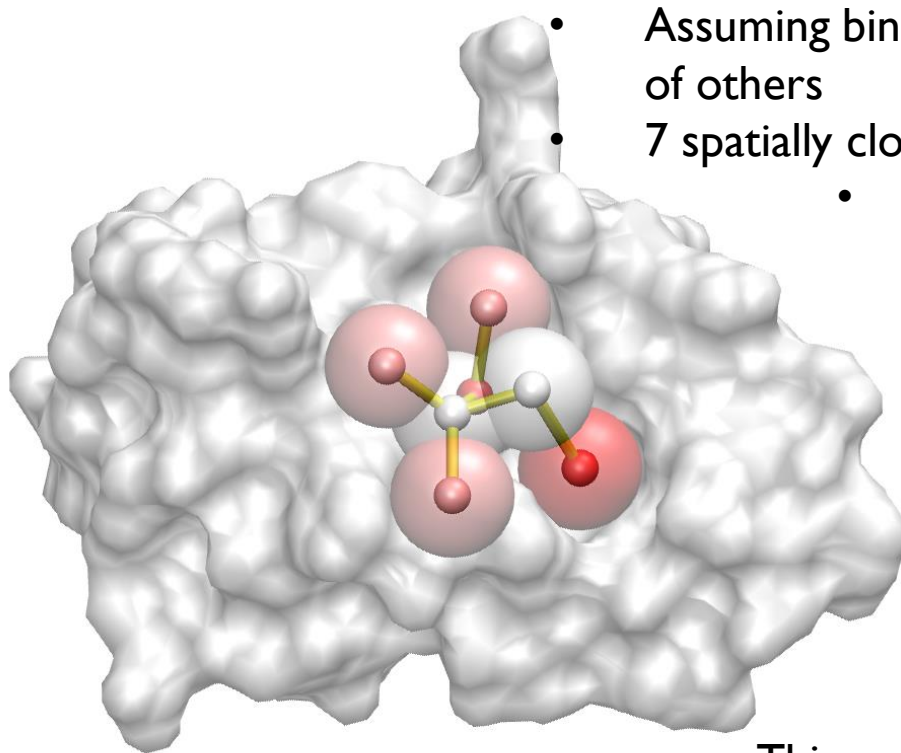
- Assuming binding of an isopropanol is independent of others
- 7 spatially close binding spots are selected
  - The sum of $\Delta G_{binding}$ of individual points is considered as a binding free energy estimate that is achievable by a drug-like molecule

−0.42
−0.84
−1.27
−1.69
−2.12
−2.54 kcal/mol

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)
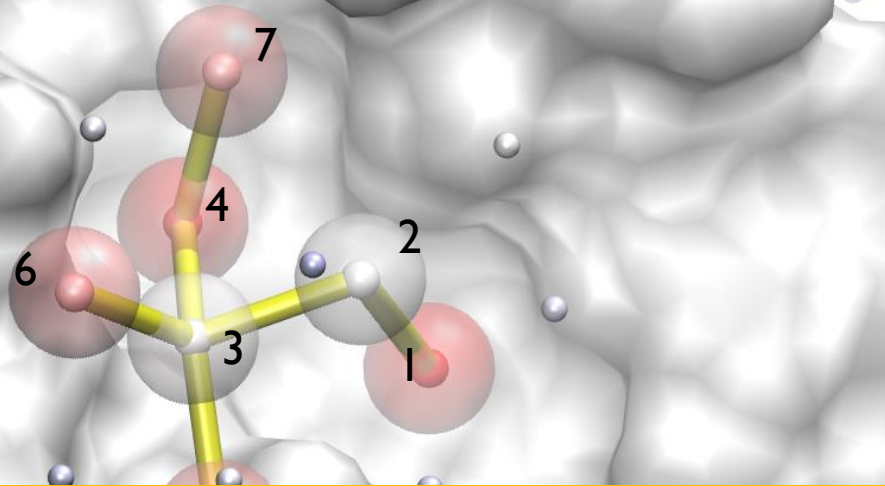
The top solution is in the protein-protein interaction site

Numbers indicate the order that hot spots were merged by the growing algorithm
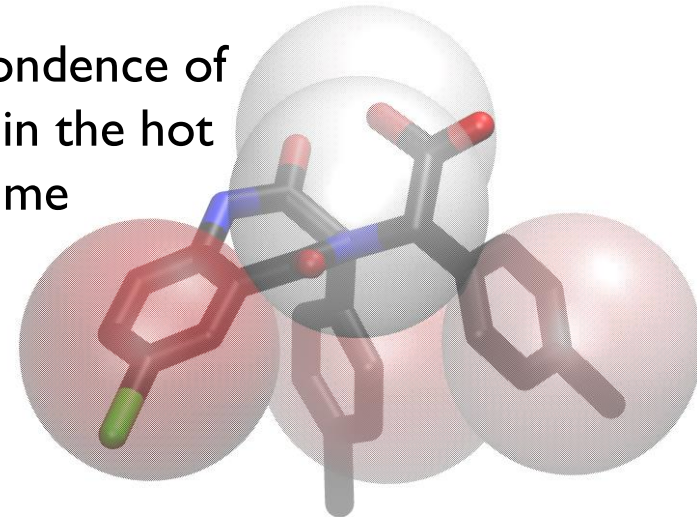
Predicted binding affinity range : **0.05-0.3 nM**
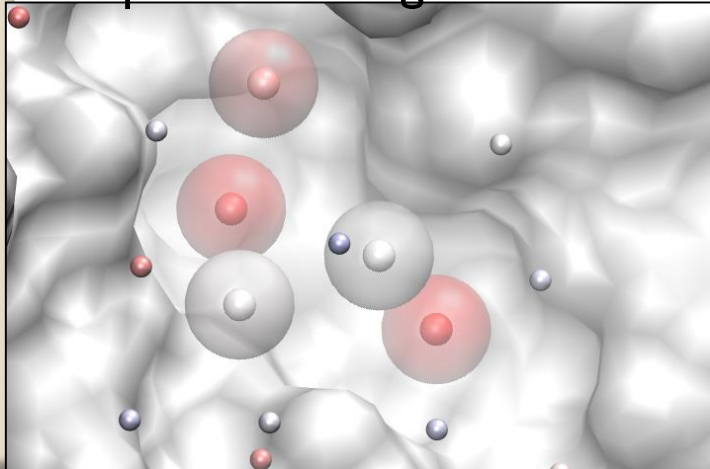Predicted max. affinity by Seco et al. : **0.02 nM**

MDM2: p53 binding site

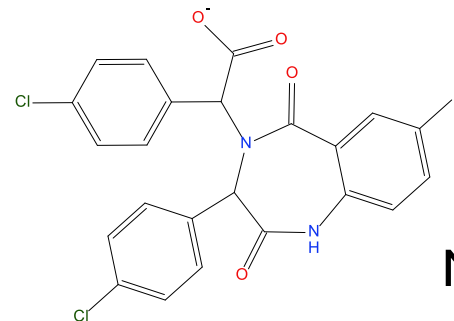An inhibitor that disrupts p53 binding
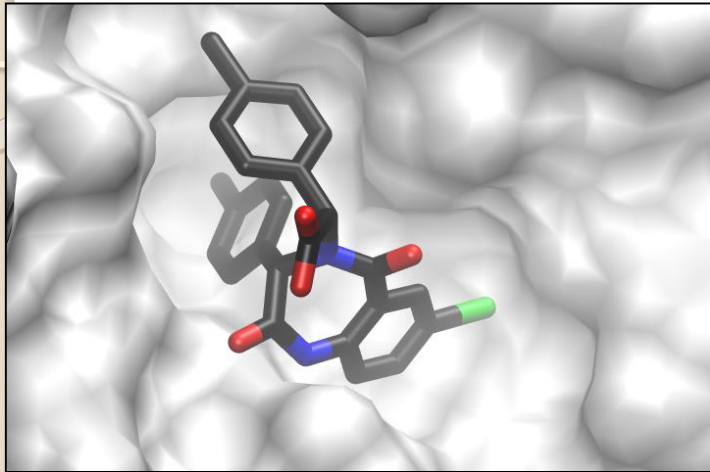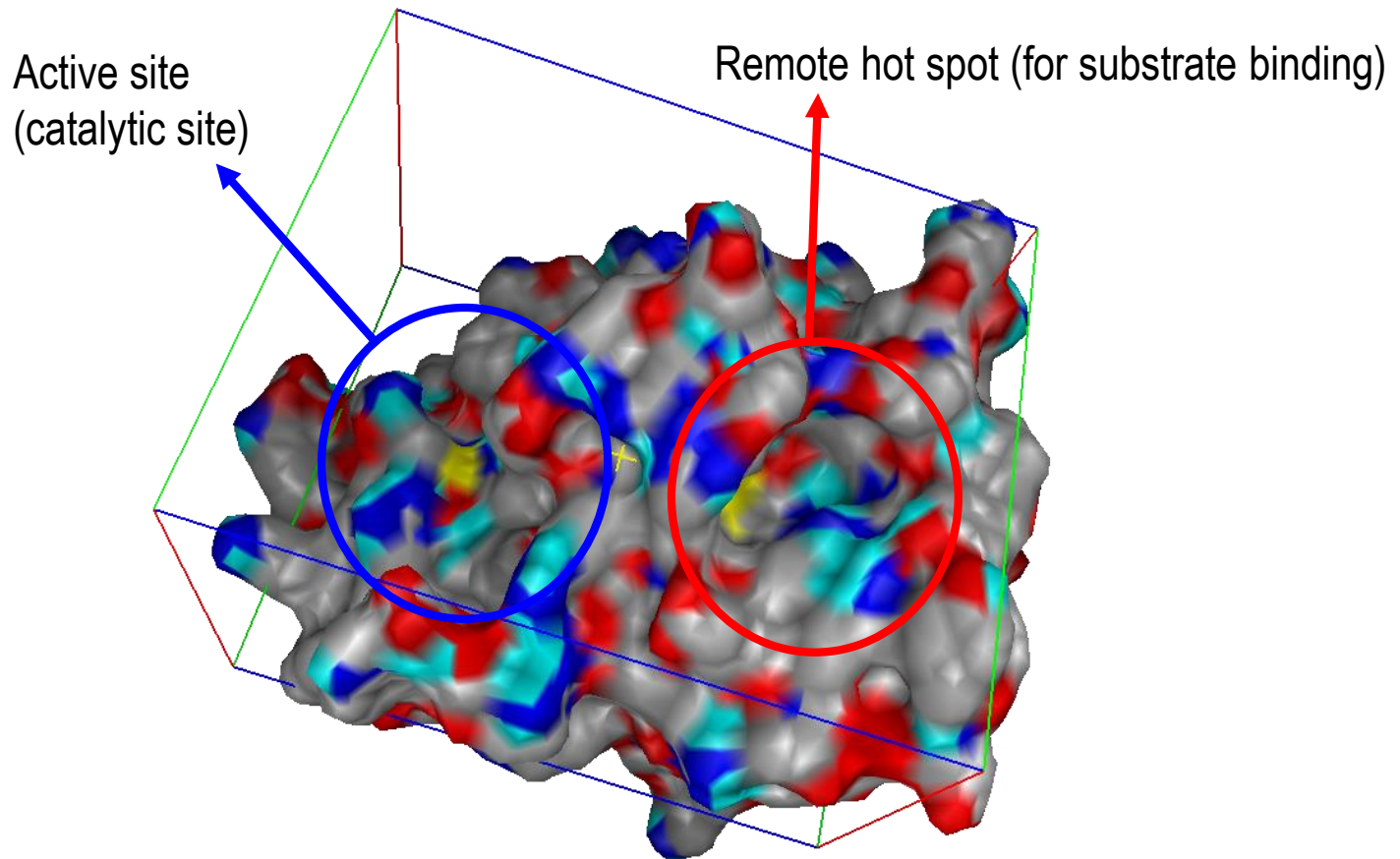

Correspondence of inhibitor in the hot spot volume

Hot spots matching this inhibitor
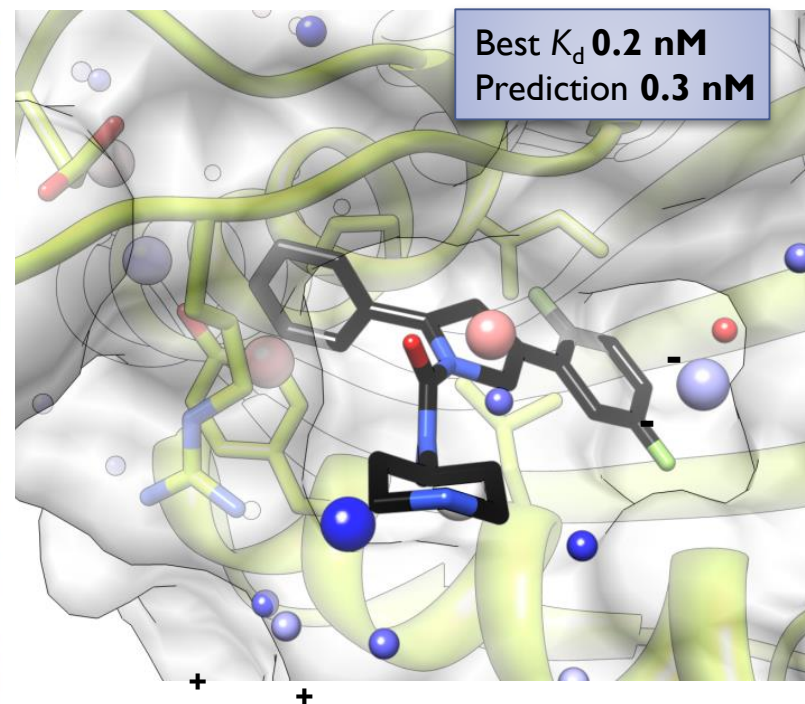

Predicted K_d: **47 nM**
Known K_d : **80 nM**
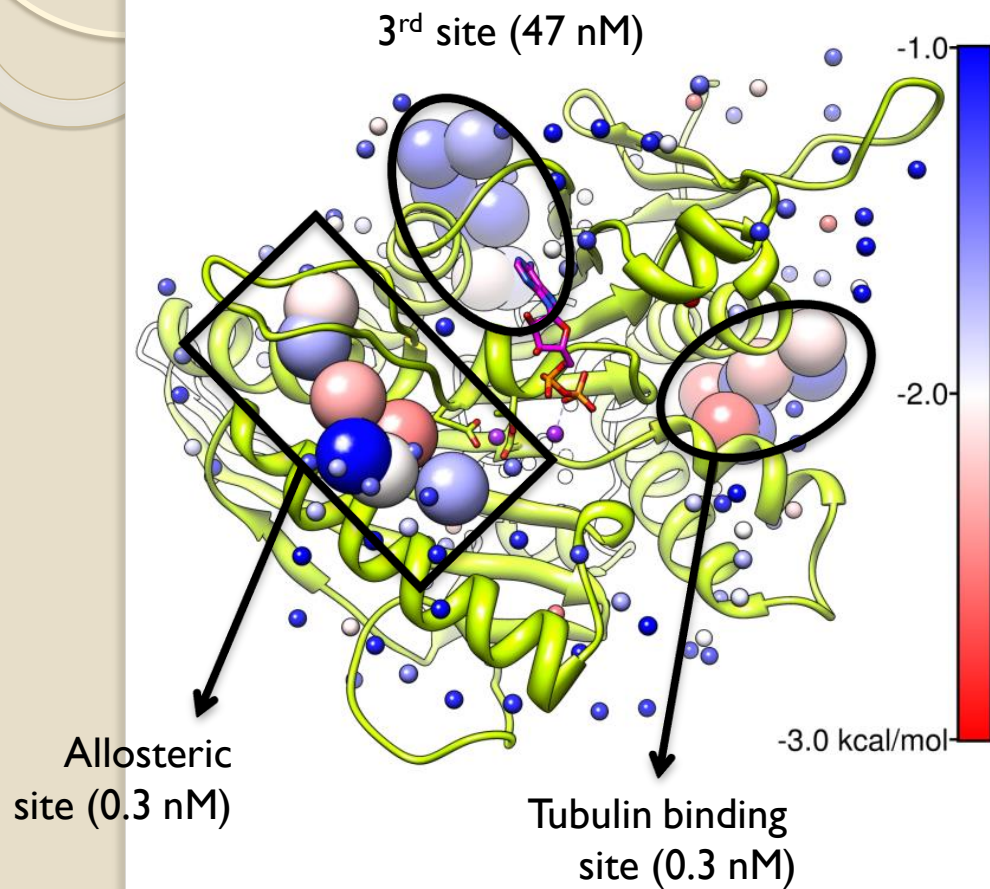

HAC = 32
MW = 580

Proteins may have multiple target sites

Active site
(catalytic site)

Remote hot spot (for substrate binding)

eg5 Druggable Sites

3rd site (47 nM)

-1.0

-2.0

-3.0 kcal/mol

Allosteric site (0.3 nM)

Tubulin binding site (0.3 nM)

Best K_d **0.2 nM**
Prediction **0.3 nM**

−

−

+

+

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

Assessment of druggable allosteric sites

MDM2 p53 site
Best K_d **0.6 nM**
Prediction **0.3 nM**

p53
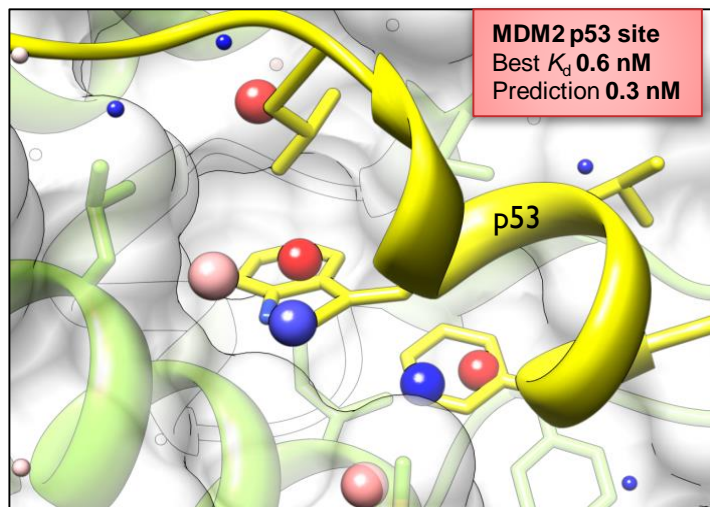
J Med. Chem. **2009,** *52*, 7970-7973

LFA-1 allosteric site
Best IC_{50} **0.35 nM**
Prediction **0.03 nM**

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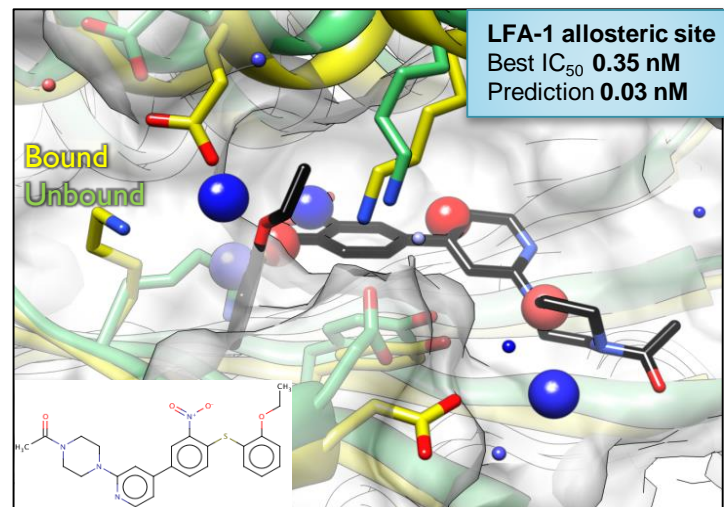
Biochemistry **2004,** *43*, 2394-2404

PTP1B catalytic site
Best IC_{50} **2.2 nM**
Prediction **0.3 nM**
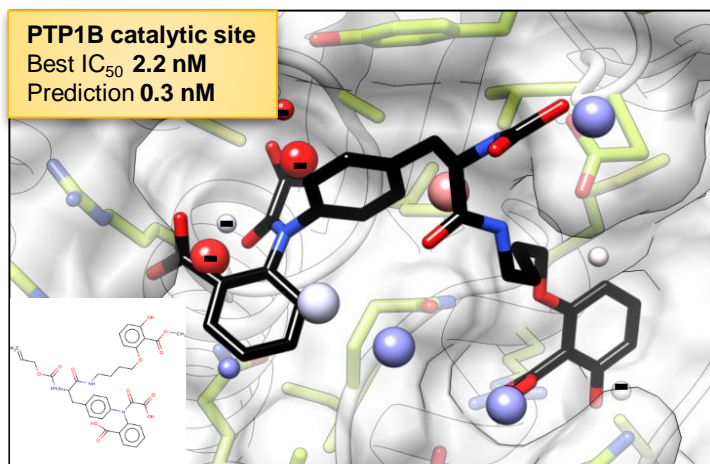
Bioorg. Med. Chem. Lett. **2003,** *13*, 3947-3950

p38 MAPK active site
Best IC_{50} **0.05 nM**
Prediction **0.01 nM**

A

Adenine/ribose
pocket

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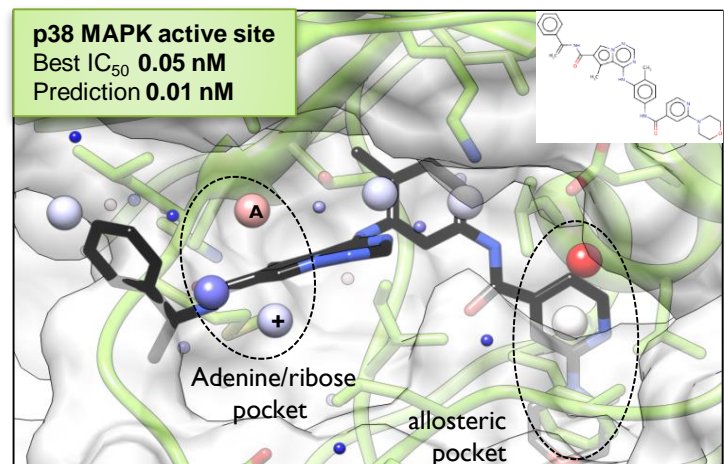
J Med. Chem. **2010,** *53*, 2973-2985

Discovery of inhibitors of cytochrome c peroxidase activity

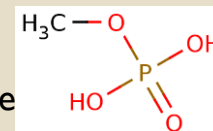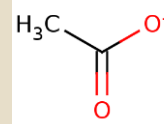druggability simulations for designing a pharmacophore model
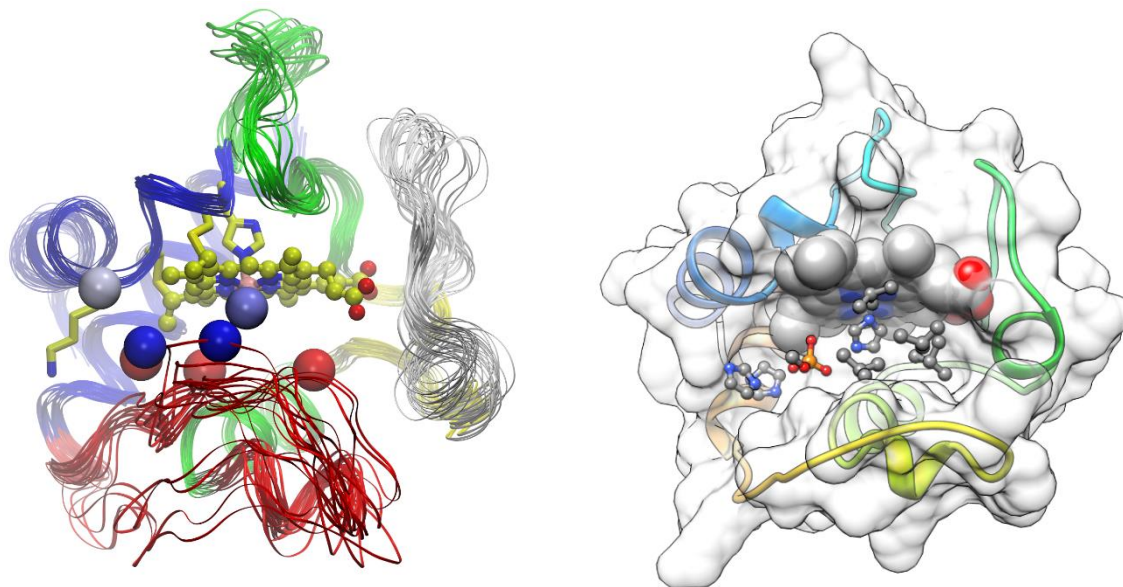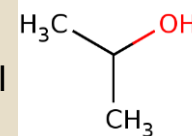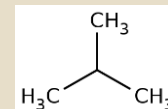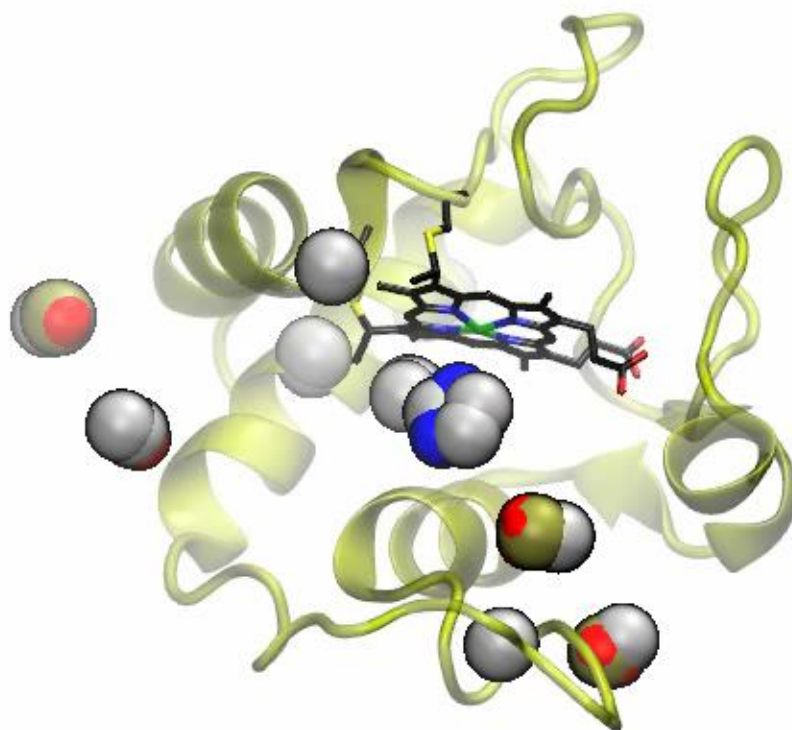
Probes

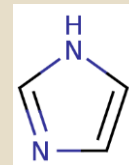48x imidazole
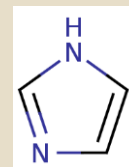
24x methly phosphate

12x acetate

12x isopropanol

24x isobutane

Discovery of inhibitors of cytochrome c peroxidase activity

Probes
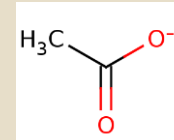
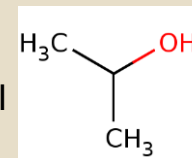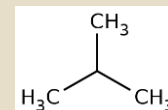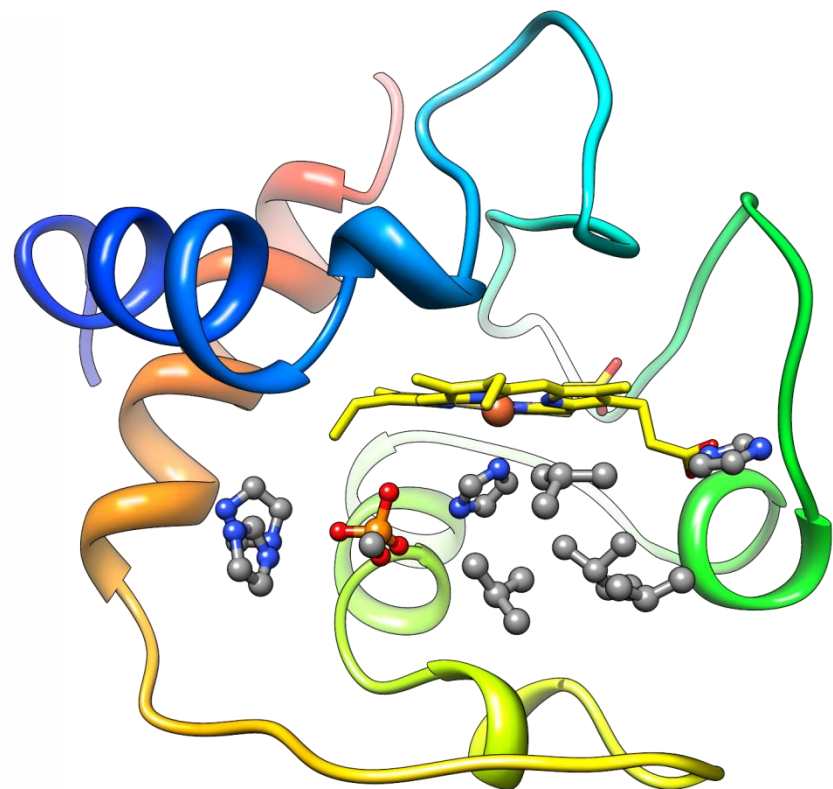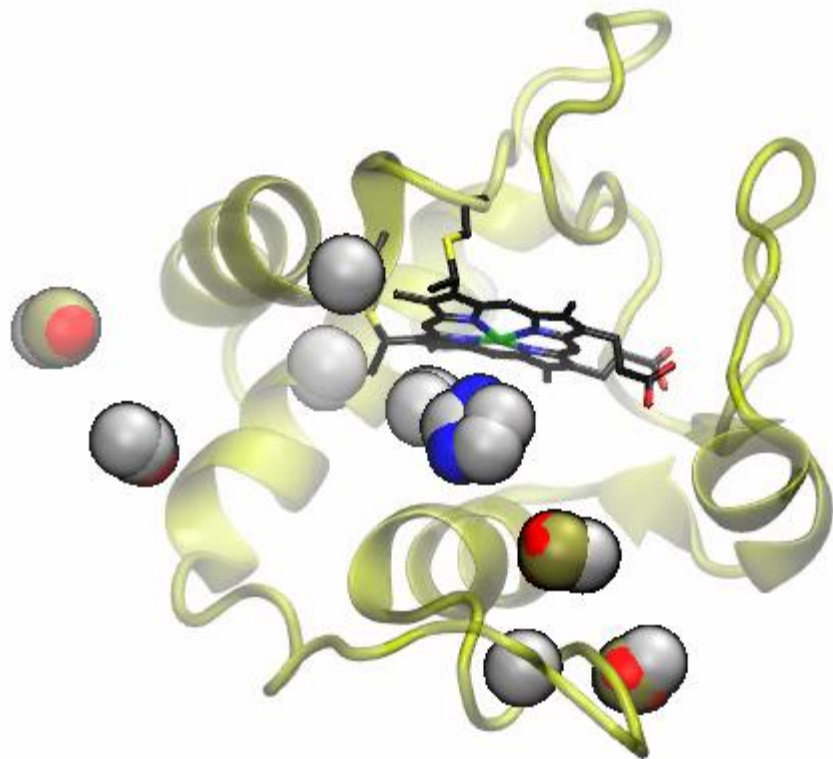48x imidazole

24x methly phosphate

12x acetate

12x isopropanol

24x isobutane

Heme site is the only druggable site with nanomolar achievable affinity
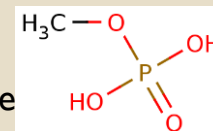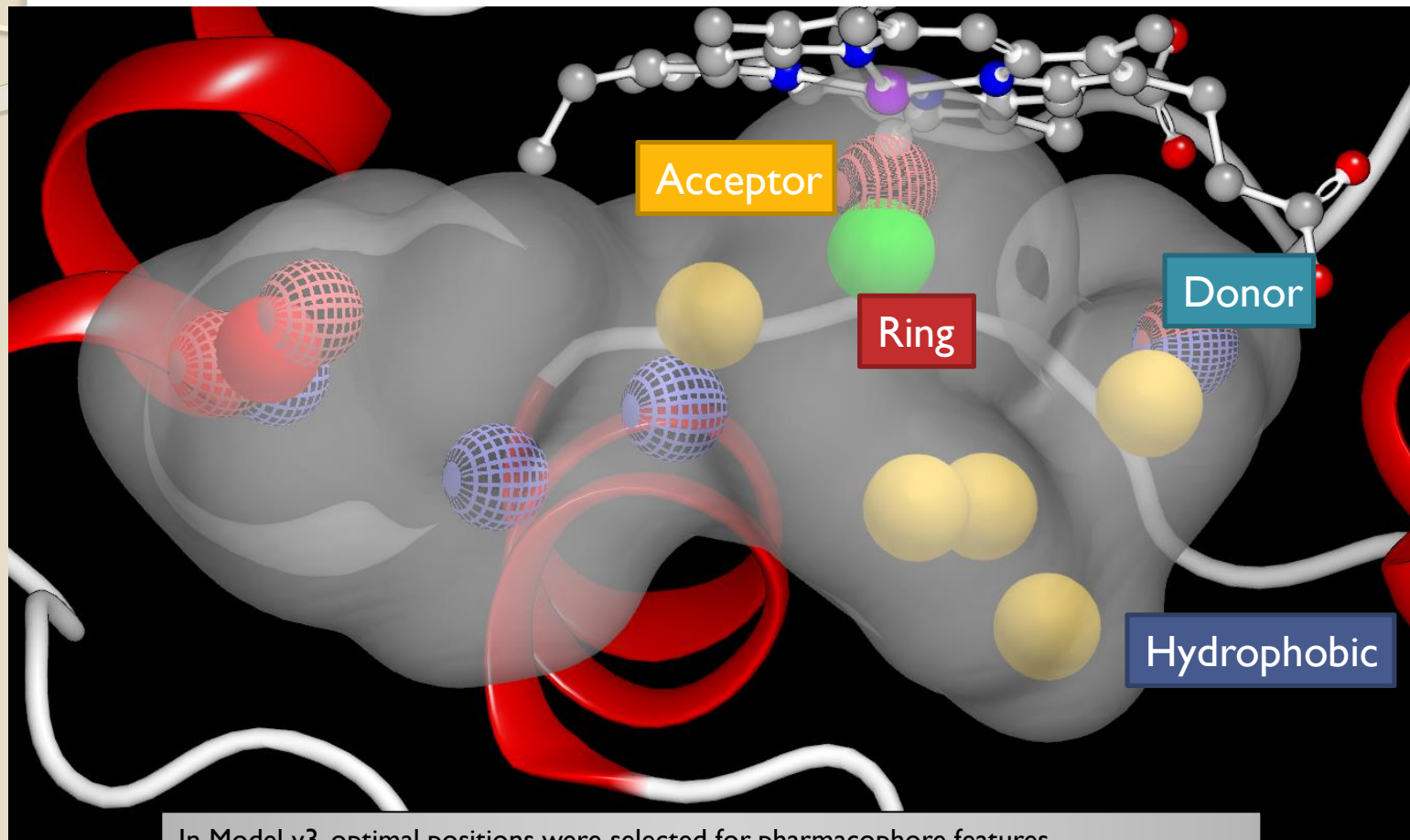
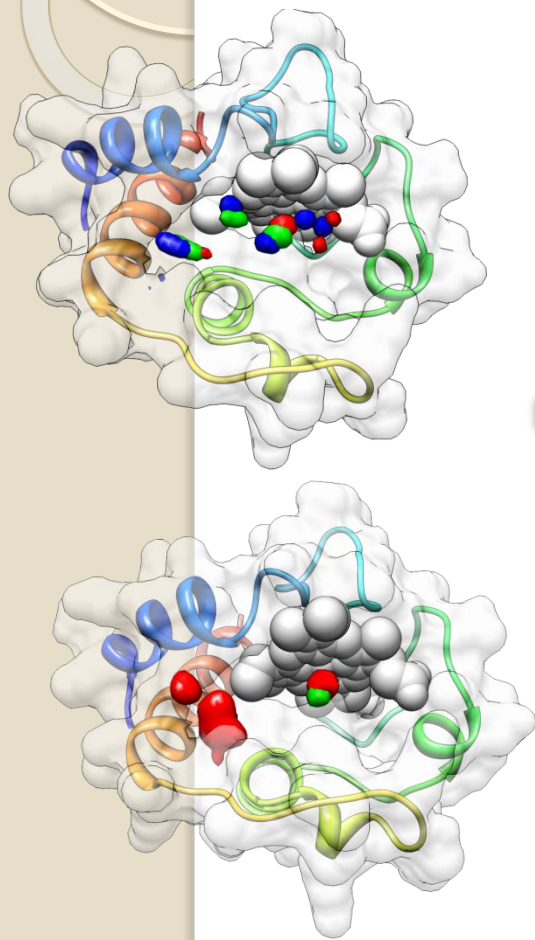# Probe Simulations

Snapshots from simulations were used to develop a pharmacophore model

Pharmacophore Model

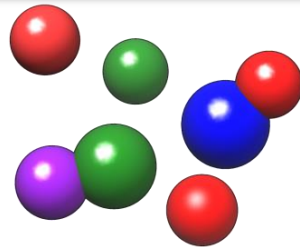In Model v3, optimal positions were selected for pharmacophore features

In silico screening
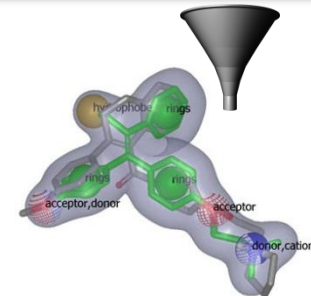

Build a consensus pharmacophore model

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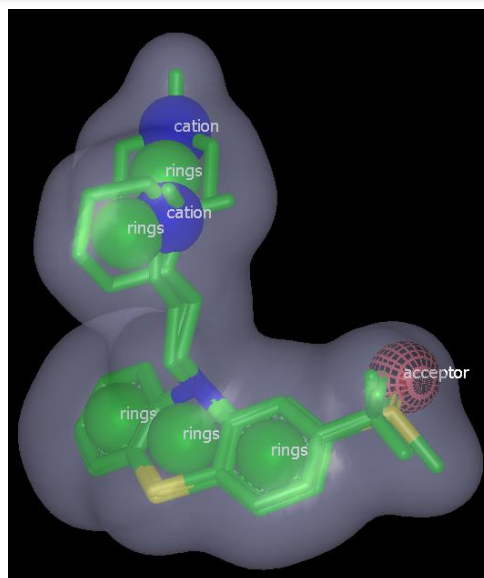~6,000 approved or investigational drugs
7.5 million purchasable compounds

DRUGBANK ZINC

Test ~10-20 compounds

Virtual HTS for hit identification

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

3. Perform **virtual high throughput screening** of databases of compounds available for purchase
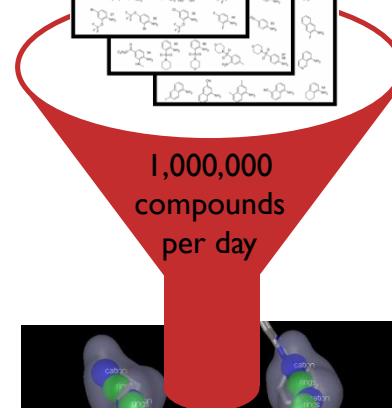
1,000,000 compounds per day

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

4. Test virtual hits experimentally and use results to refine the model

Example *in silico* hit

Bifonazole, an antifungal drug
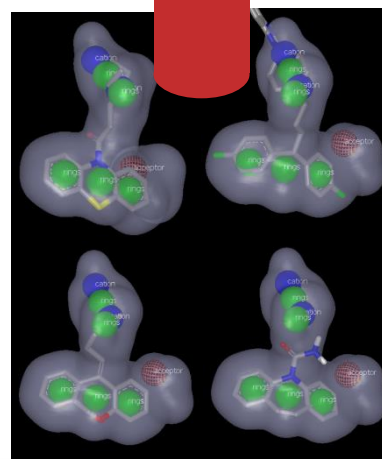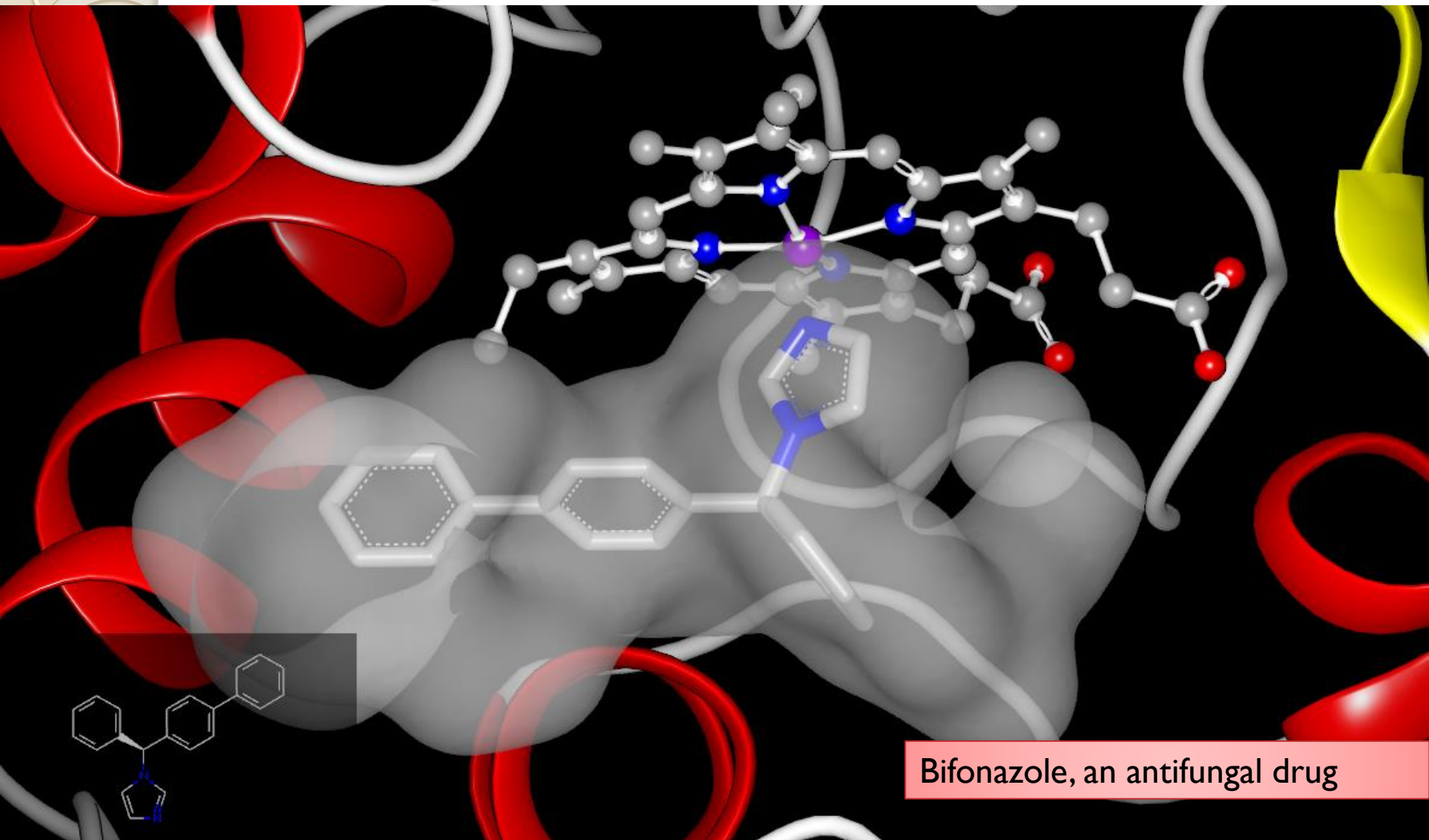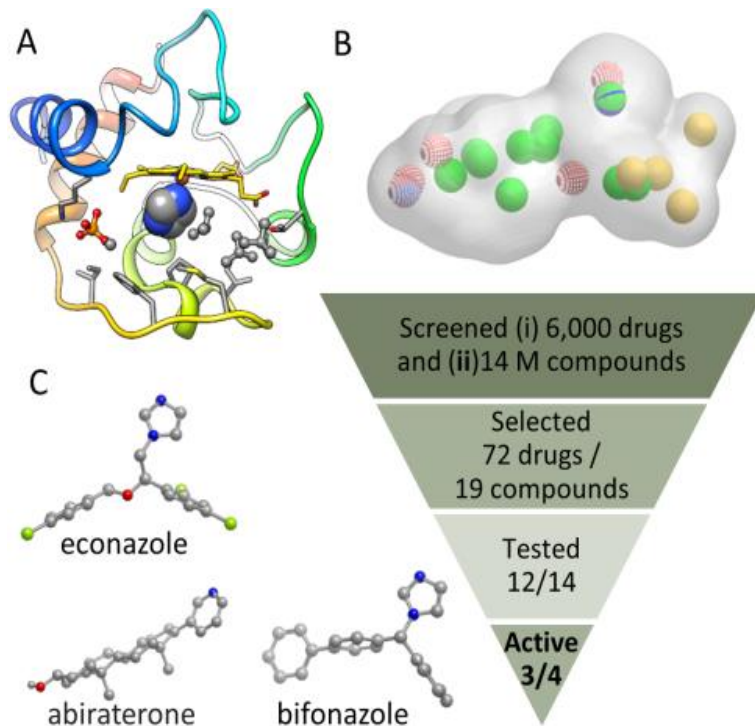
Discovery of inhibitors of cytochrome c peroxidase activity

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

Discovery of inhibitors of cytochrome c peroxidase activity
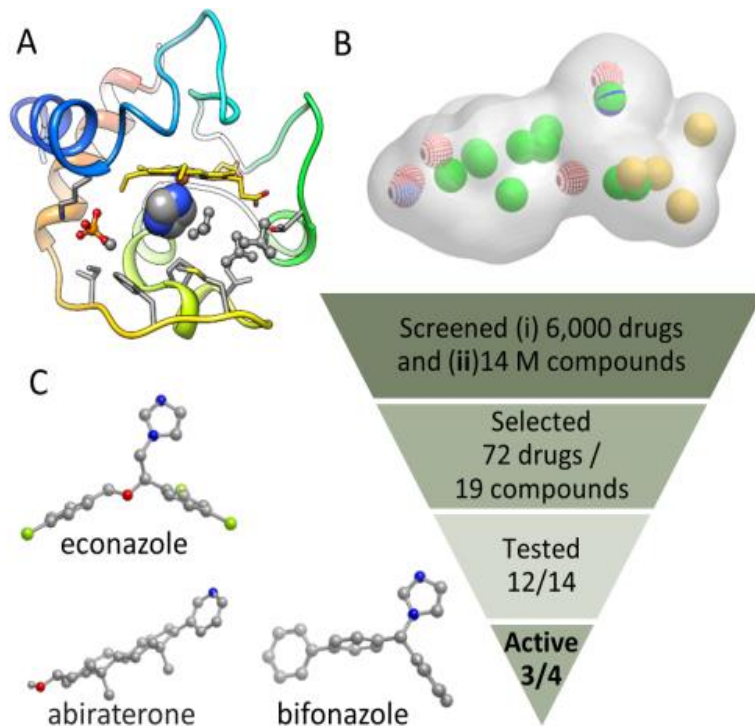
Exp validation
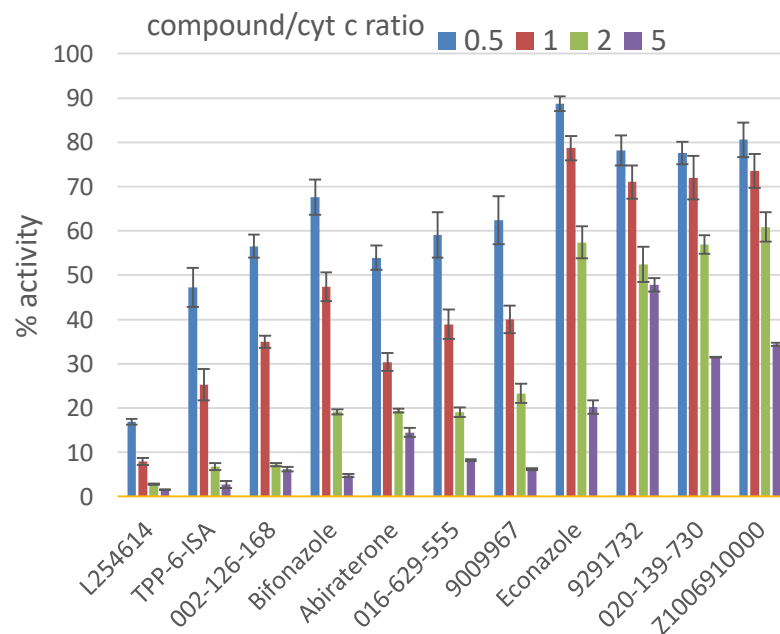
Peroxidase reaction probed by fluorescence of oxidation product, for cyt c incubated with CL/DOPC liposomes

Druggability simulations - summary

Purpose:
For a first assessment of druggable sites
Identification of allosteric sites that can bind drugs
Assessment of achievable binding affinity

How ?
Preprocessing: Using probe set either available, or customizable
Prepare input files for NAMD runs (i.e. ensemble of structures with selected composition of probe molecules)
Post-processing: Analyze trajectories generated by NAMD to make inferences on binding sites and affinities, and pharmacophore models

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