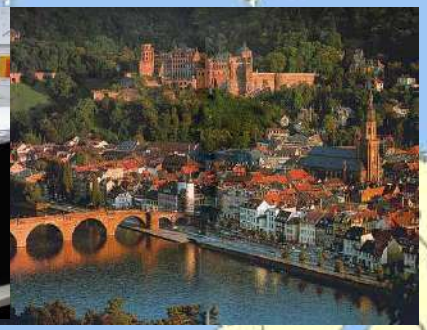
# Εισαγωγή στις ΜΔ Προσομοιώσεις και το Σχεδιασμό Φαρμάκων μέσω υπολογιστή

**Ζωή Κούρνια**

*Ίδρυμα Ιατροβιολογικών Ερευνών, Ακαδημία Αθηνών*

**4 Μαΐου 2020**

*zcournia@bioacademy.gr* *twitter: @zoecournia*



**2001-2006**
**Doctoral Student in Chemistry**
**Heidelberg University, Germany**

**1996-2001**
**Chemistry, University of Athens**



Ph.D in Chemistry (2006)
U Heidelberg

2006-2009 PostDoc
Chemistry Dept,
Yale University
New Haven, USA





BRFAA
2009-2015
Investigator D'

2015-2019
Investigator C'

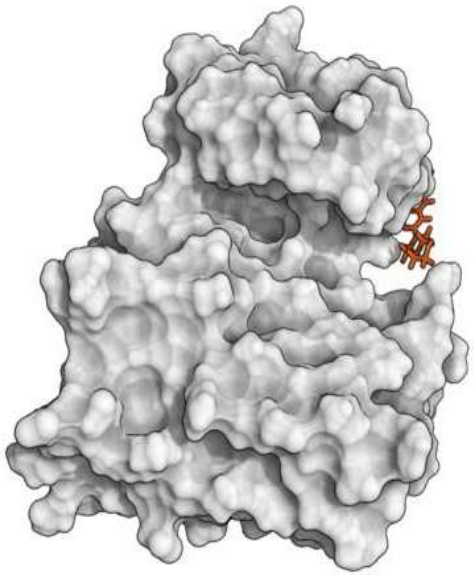
2019-date
Investigator B'

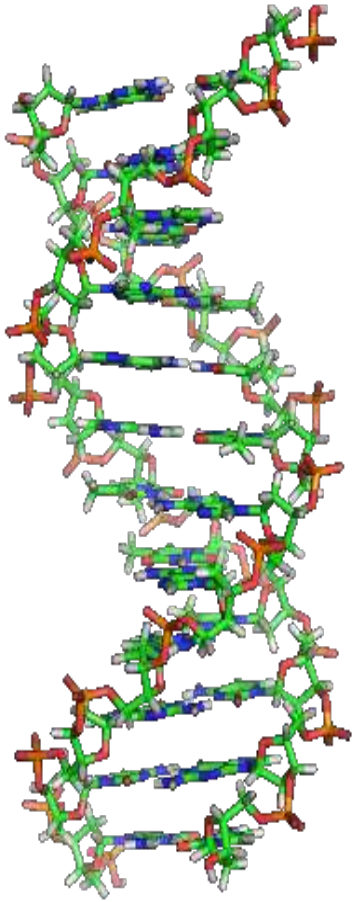
Instructor, MSc Program
Data Science &
Information Technologies

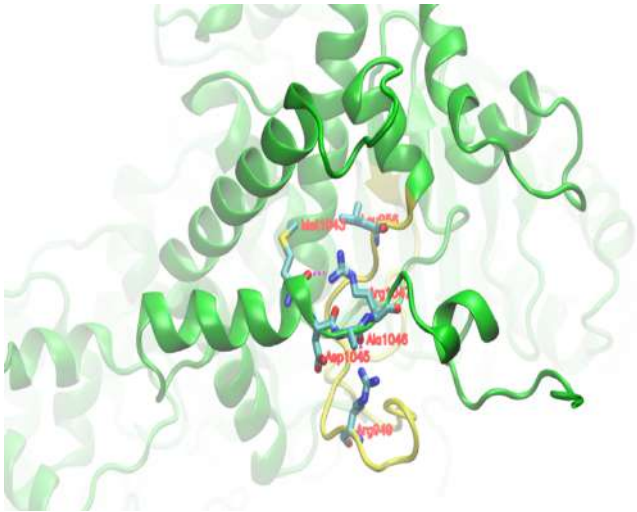
# Computational Drug Design

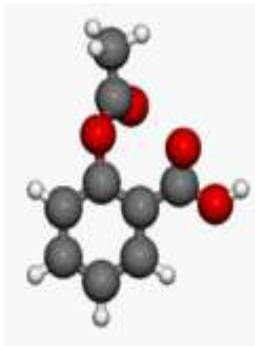
# Cournia Lab

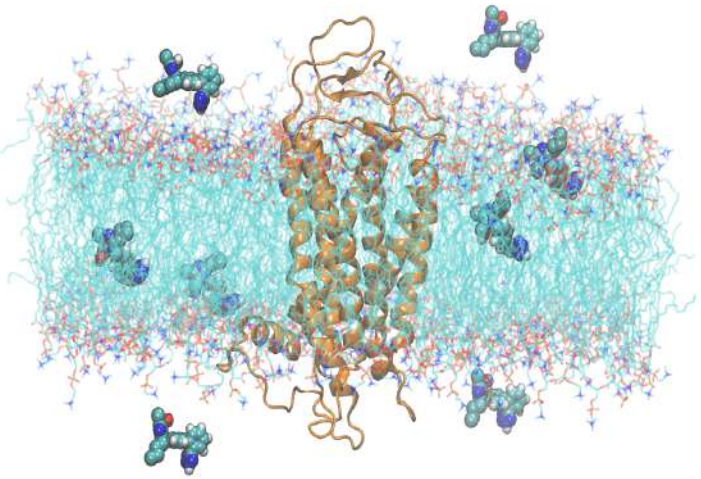
IIBEAA
ΙΔΡΥΜΑ ΙΑΤΡΟΒΙΟΛΟΓΙΚΩΝ ΕΡΕΥΝΩΝ
ΑΚΑΔΗΜΙΑΣ ΑΘΗΝΩΝ



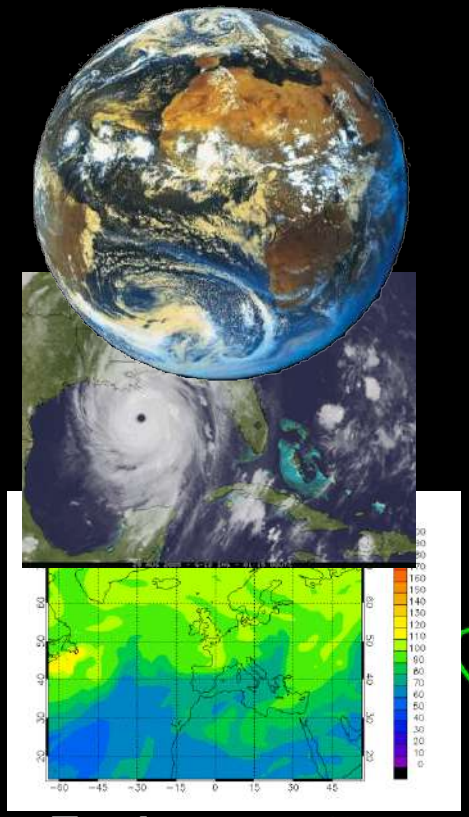




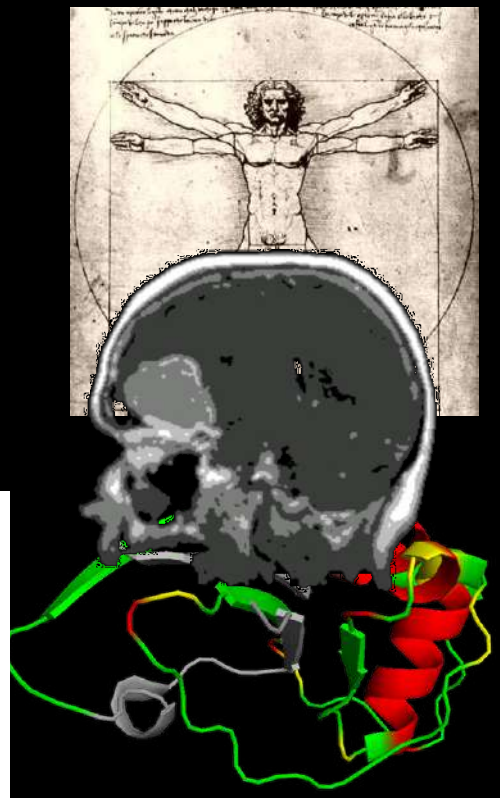




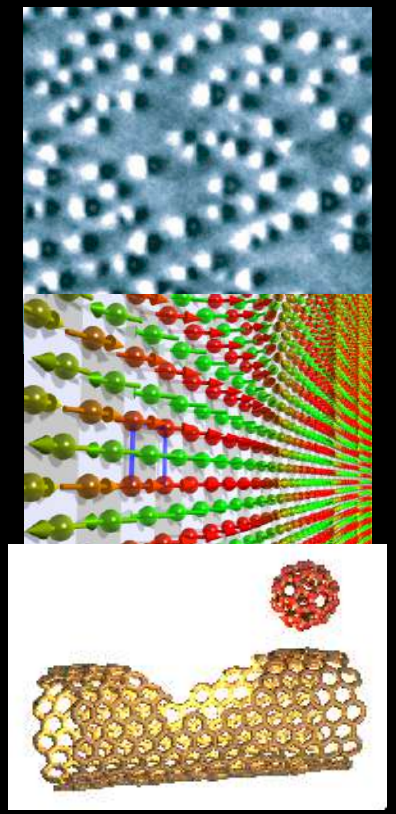
# Supercomputing Drives Science through Simulation



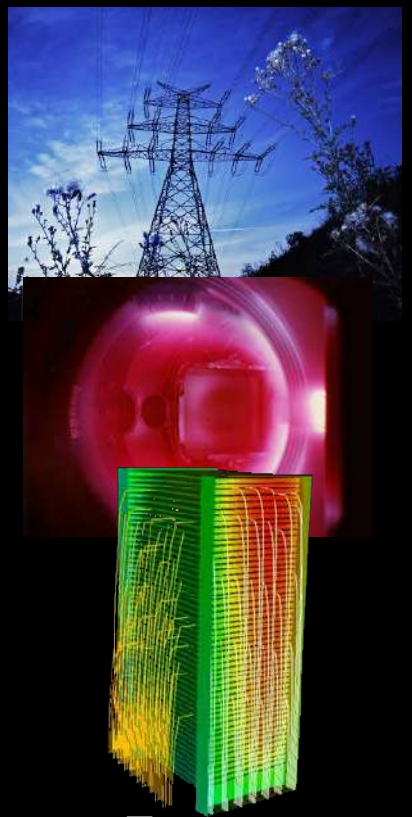
**Environment**
Weather/ Climatology
Pollution / Ozone Hole



**Finding Cures**
Medicine
Biology



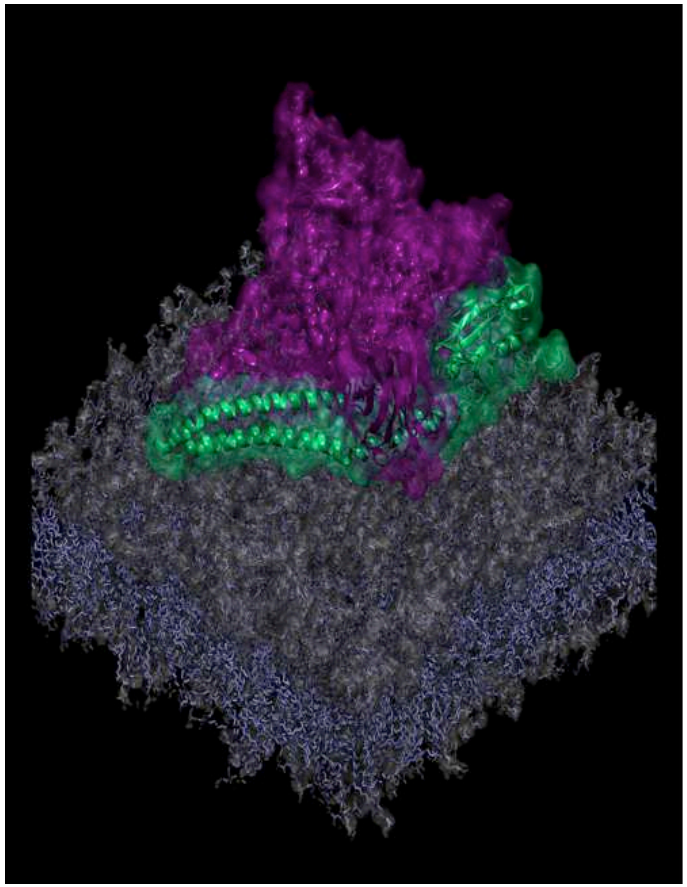
**Materials/ Inf. Tech**
Spintronics
Nano-science



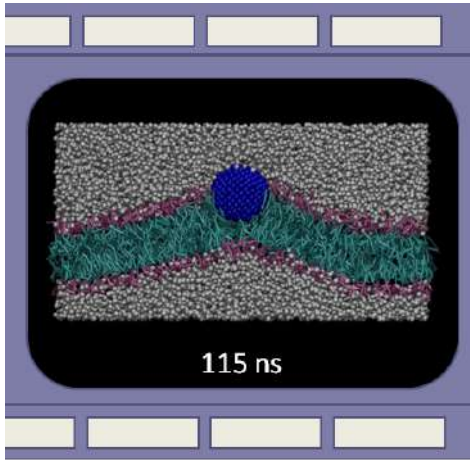
**Energy**
Plasma Physics
Fuel Cells

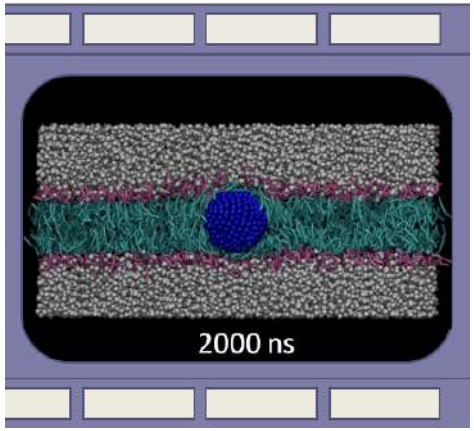
# Computer-aided drug and drug delivery design

## Protein biophysics

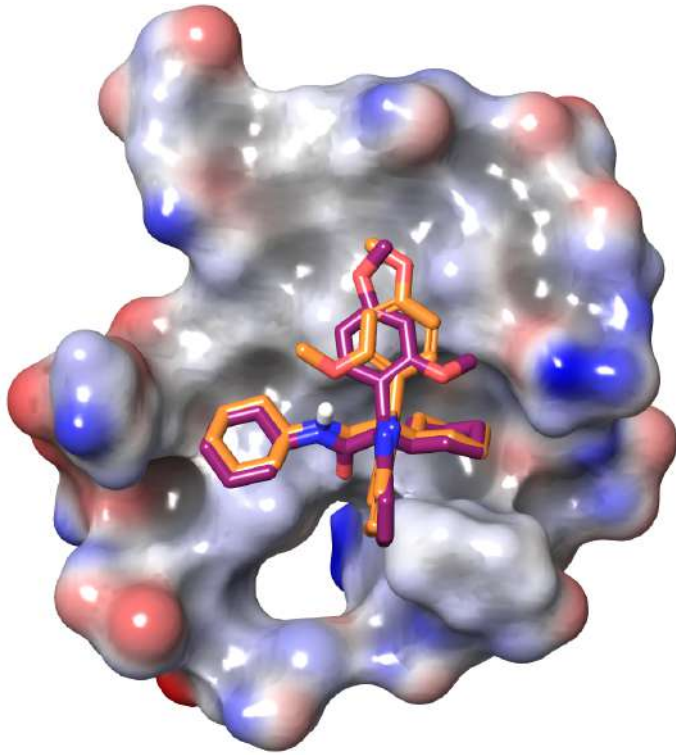


## Drug delivery systems



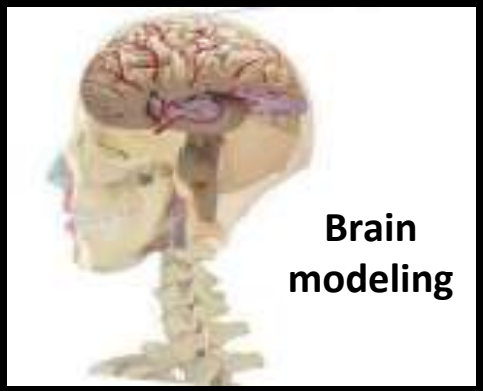



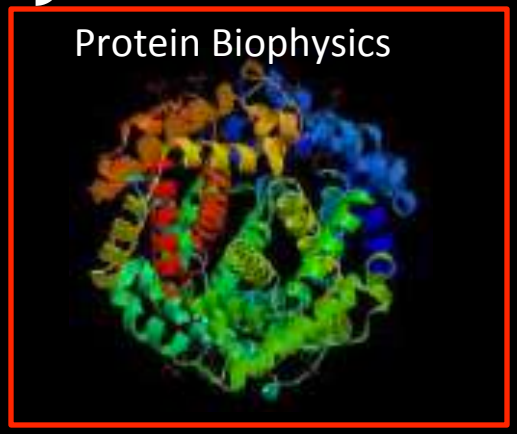

## Computer-aided drug design

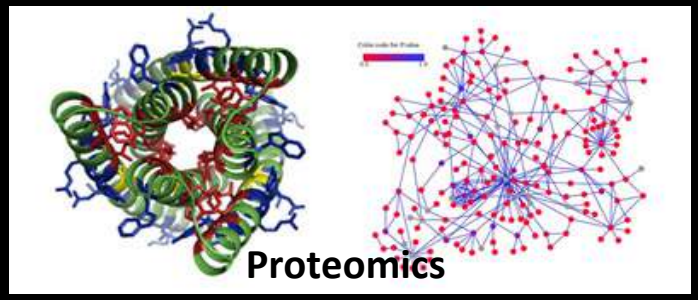



# HPC Systems

# Key areas of biomedical research where HPC is key

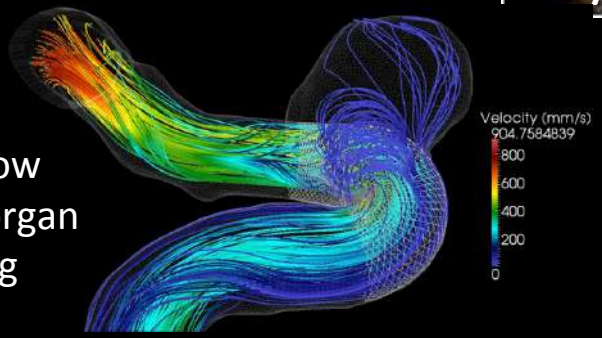



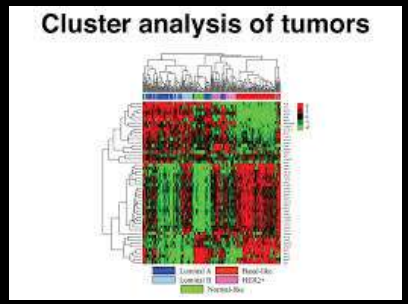



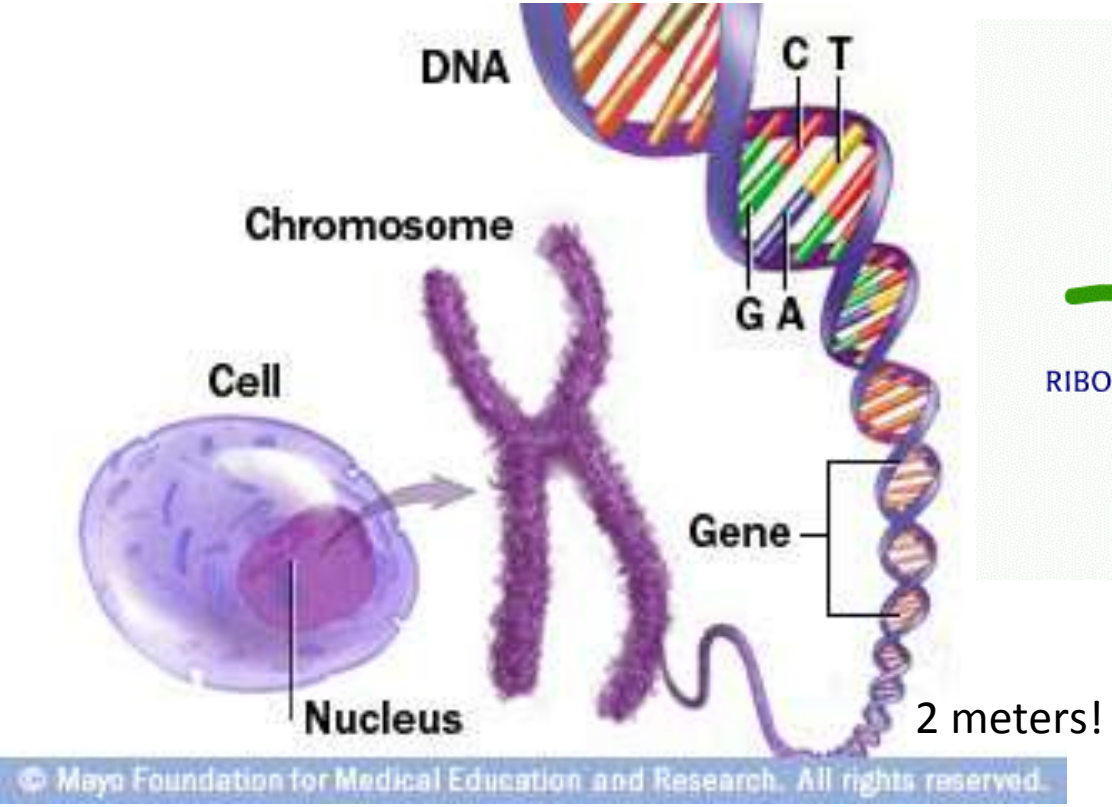


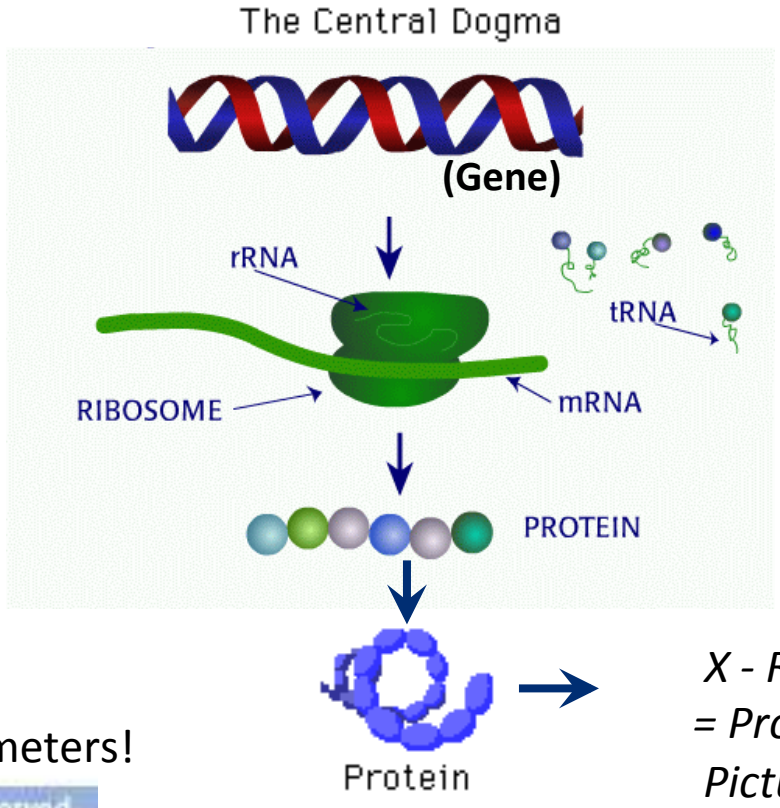

Blood flow
Whole organ
modeling



# From DNA, to genes and proteins



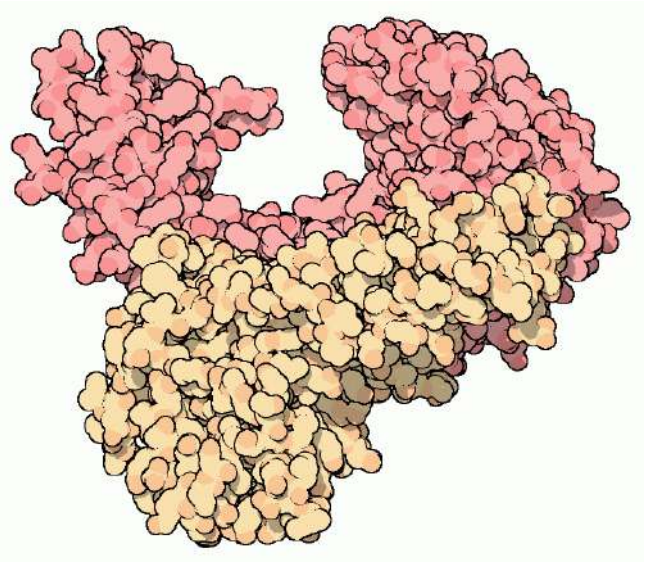

*X - Rays = Protein Pictures*

**20.000 genes in the nuclei of our cells**
**→ PROTEINS**

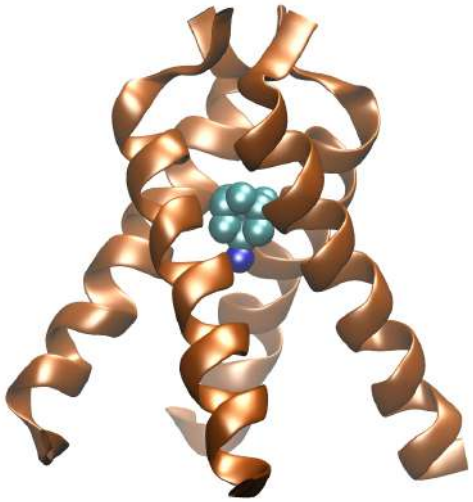
- **Proteins are the expression of genes in functional molecules**
- **Proteins perform essential functions in the cell**

# Some proteins need to be stopped!

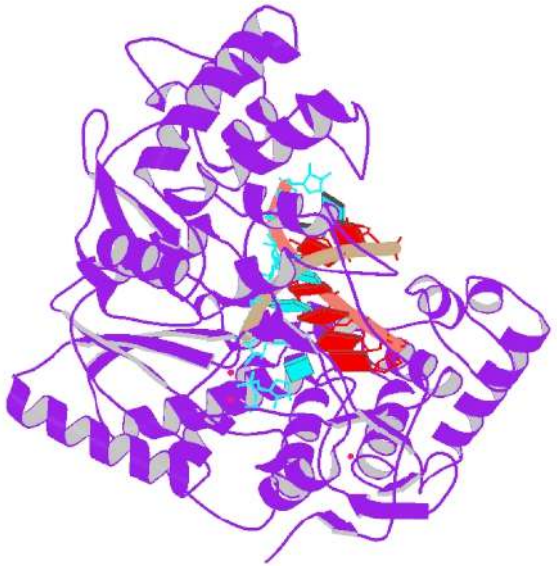
**HIV-RT**



**M2TM**
**(Influenza virus)**



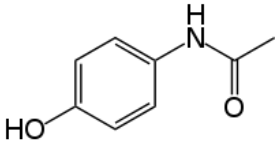
**NS5B**
**(Hepatitis C)**

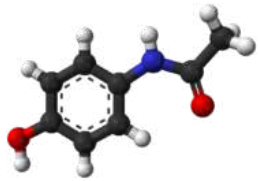


# Drugs block or activate diseased proteins

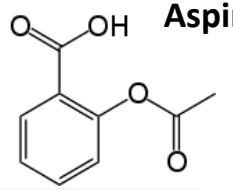
**Normally they are small organic molecules**
- **Therapy**
- **Relief**
- **Prevention**
- **Quality of life improvement**
- **Life expectancy prolongation**

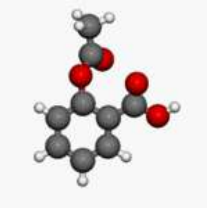
**Paracetamol (Depon)**

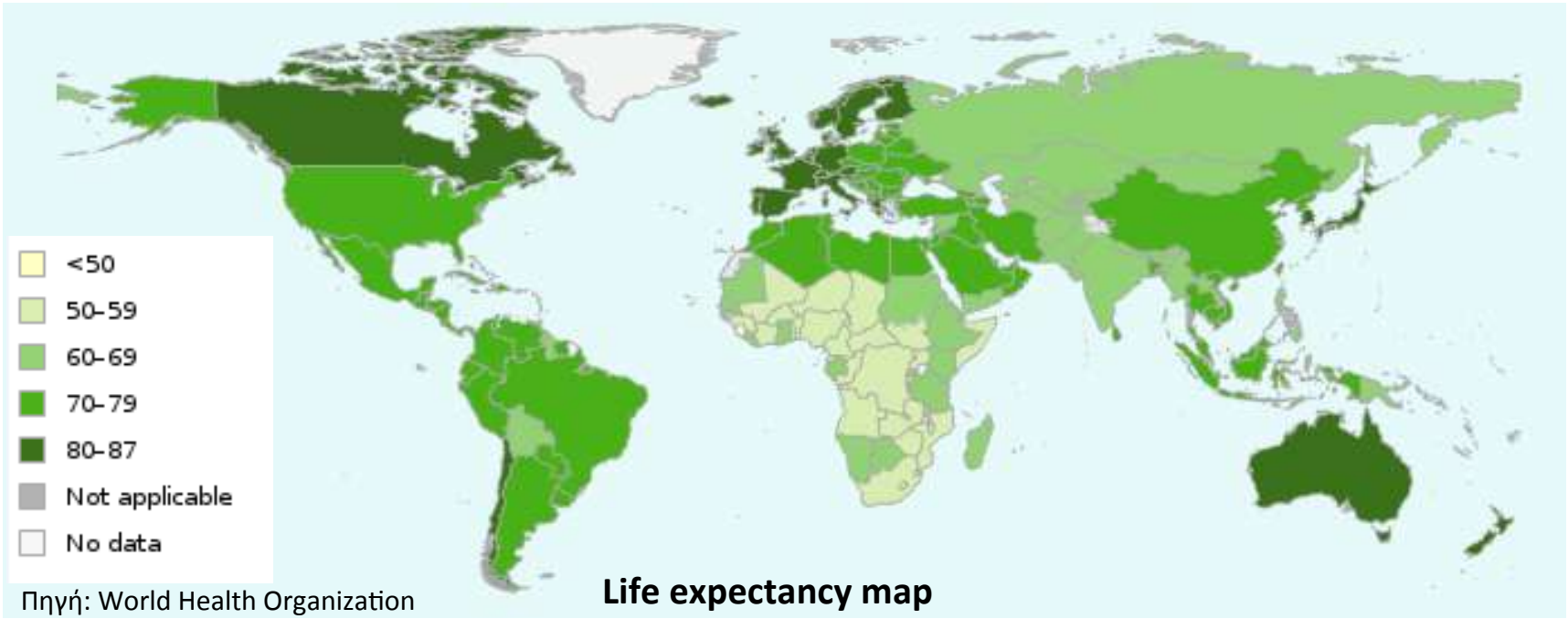


**Aspirin**



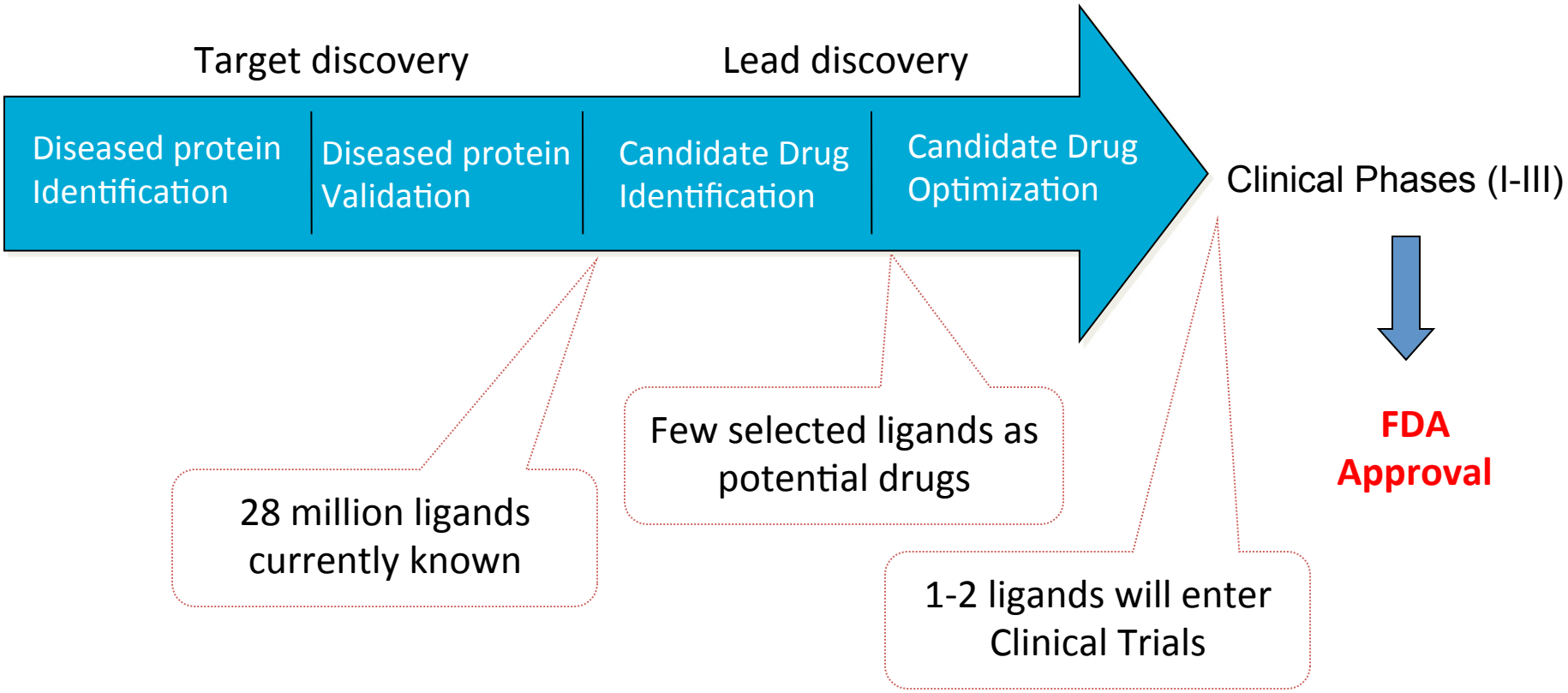






Πηγή: World Health Organization

**Life expectancy map**

# Phases of Pharmaceutical Development

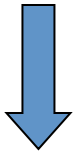



Duration: 12 – 15 years, Cost: ~ 1 billion US $

# Traditional Drug Discovery

- Random screening of hundreds of thousands of molecules with High Throughput Screening (HTS) for combating the pathogen
- Random discoveries (i.e. penicillin, viagra)
- Trying out existing drugs and modifications

- Estimated number of small molecules that can act as drugs **$10^{66}$**
- Estimated number of atoms in the world **$10^{50}$**

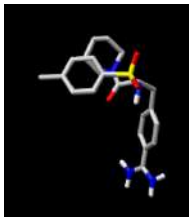


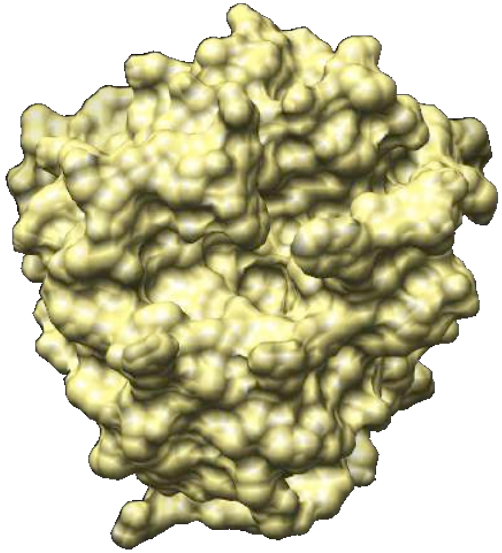
**Structure-based approaches + Targeted Therapy**

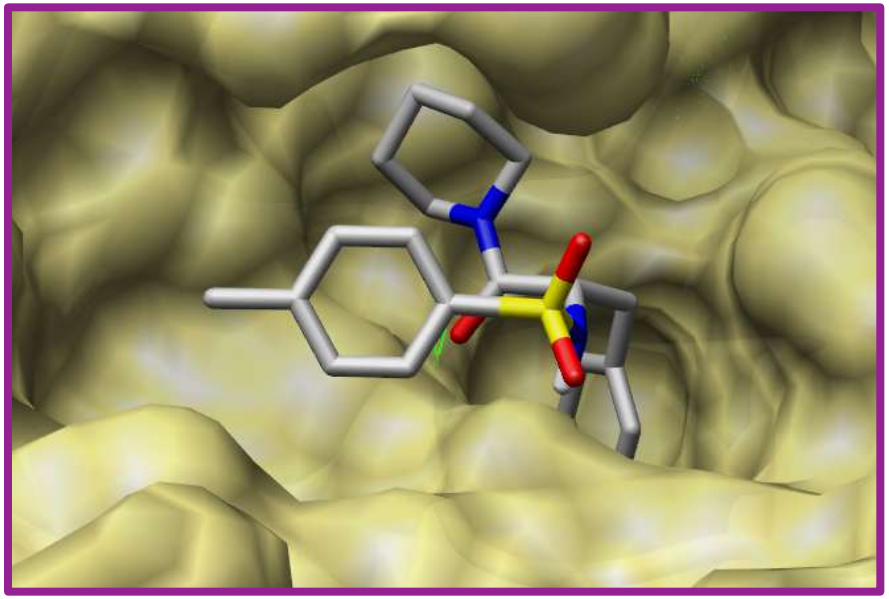
# *Rational* Drug Discovery

- Identify important genes for a diseases
- Targeting/inactivating genes (proteins) of the pathogen with small molecules = drugs

**TARGETED THERAPY!**

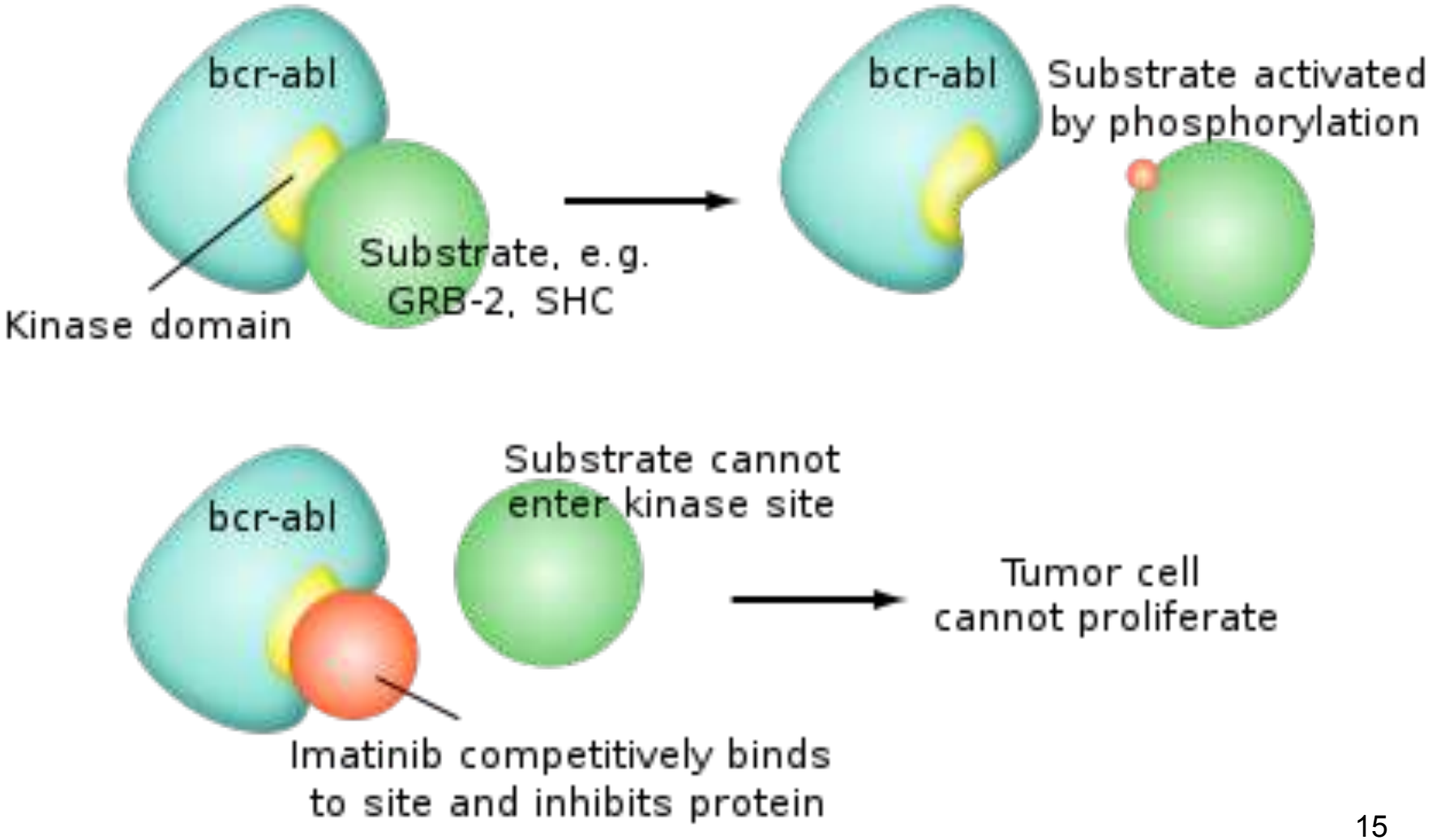



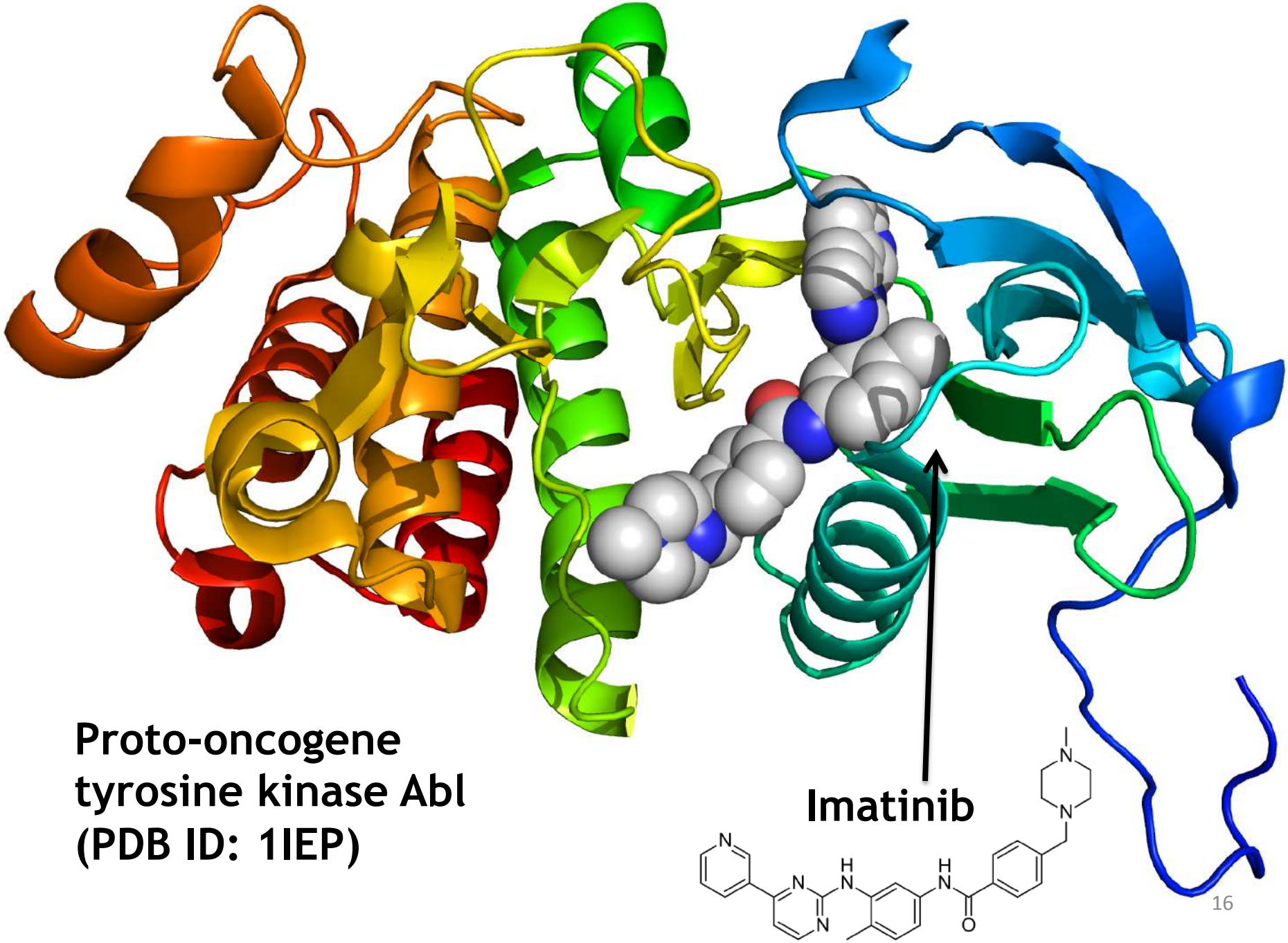




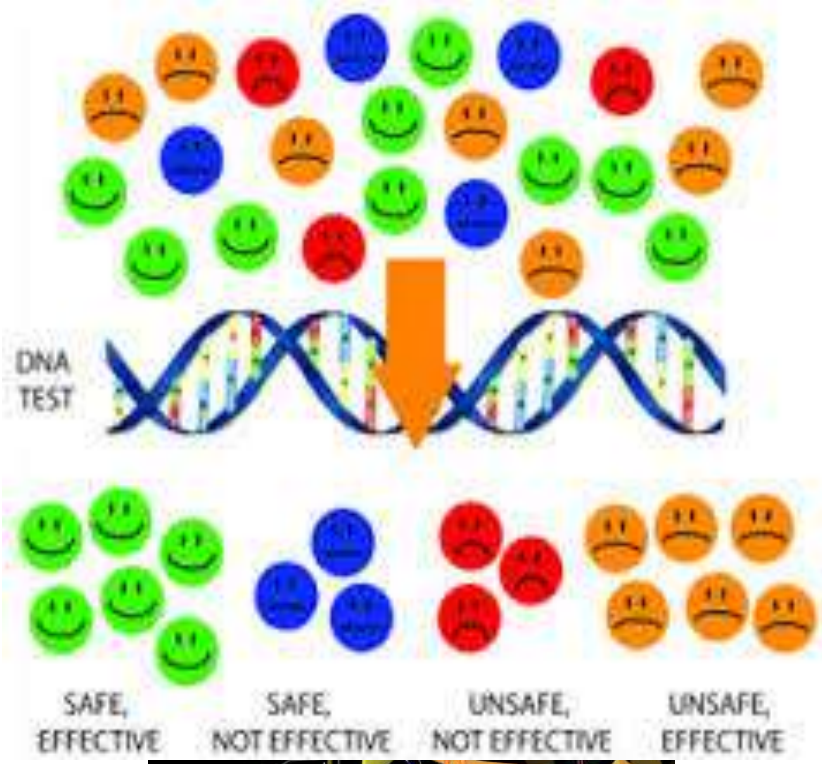
Curr Opin Drug Discov Devel. 2002 May; 5(3): 355–360

# Constitutively active chimeric oncogene Bcr-Abl Tyrosine Kinase in CML




Proto-oncogene
tyrosine kinase Abl
(PDB ID: 1IEP)
Imatinib

# What is personalized therapy?




**Lung Cancer**

↓

genotyping

↓

4% of patients with
non-small cell lung carcinoma
Rearrangement in **ALK protein**

↓

carcinogenesis

↓

Drug design for this specific
subset of patients

↓

**Crizotinib for ALK+**
**lung cancer patients**

# Drug Design: Outline of the process

1. Protein Structure
2. Computer Simulations
3. Finding binding sites where drug binds
4. Design of chemical compound suitable to bind on the specific protein (interactions)
5. Choose compounds / Organic Synthesis

DESIGN

6. Assaying compounds in vitro (without cells)
7. Cell-based in vitro assays
8. Calculate efficacy of molecule – candidate drug

IN VITRO

9. Pharmacokinetics/ Pharmacodynamics in healthy animals
10. Check toxicology in healthy animals
11. Efficacy studies in mouse xenografts
12. Efficacy studies in animal models of the disease

IN VIVO

# Pre-clinical and clinical stages

## Pre-clinical studies

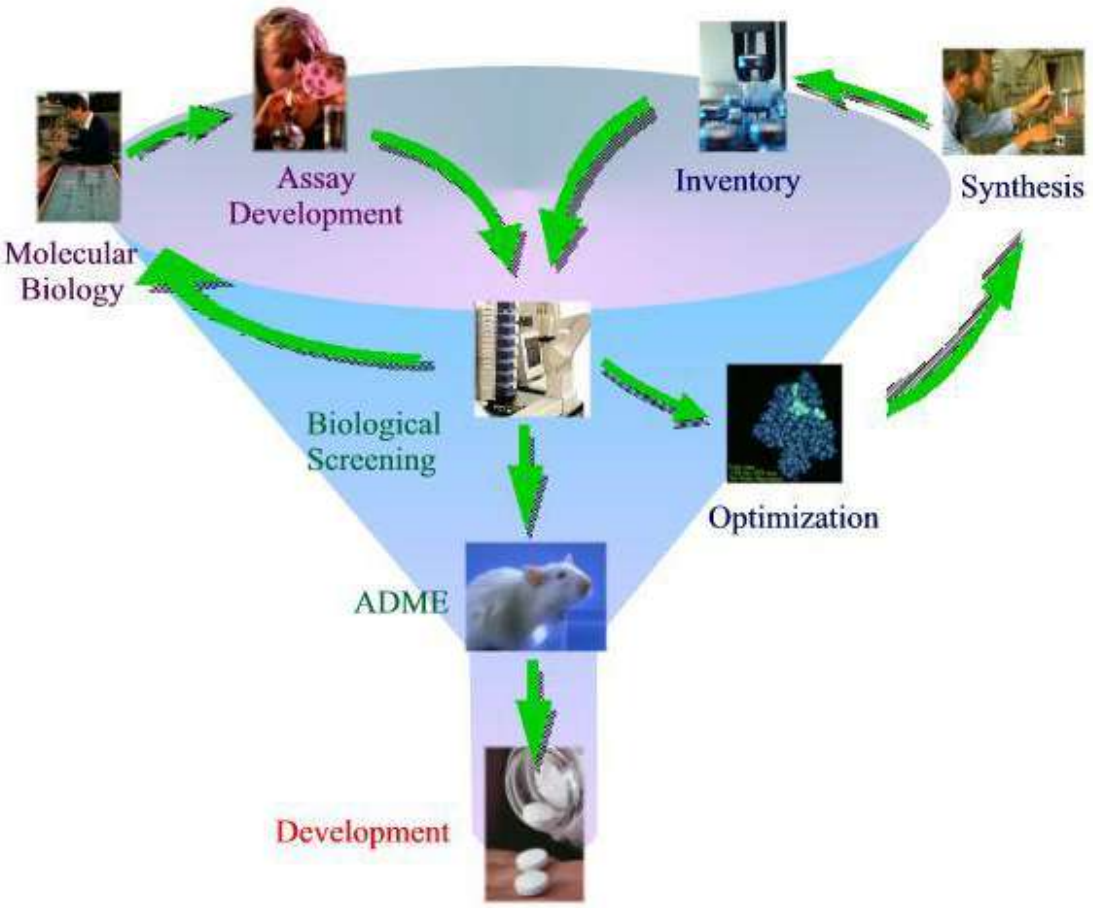



Image source: Akos

## Clinical Trials

**Phase I**

- 5-40 healthy volunteers, months
- Dosing and safety studies
- ~70% Success Rate

**Phase II**

- 100 - 300 patients, 2 years
- Efficacy studies
- ~30% Success Rate

**Phase III**

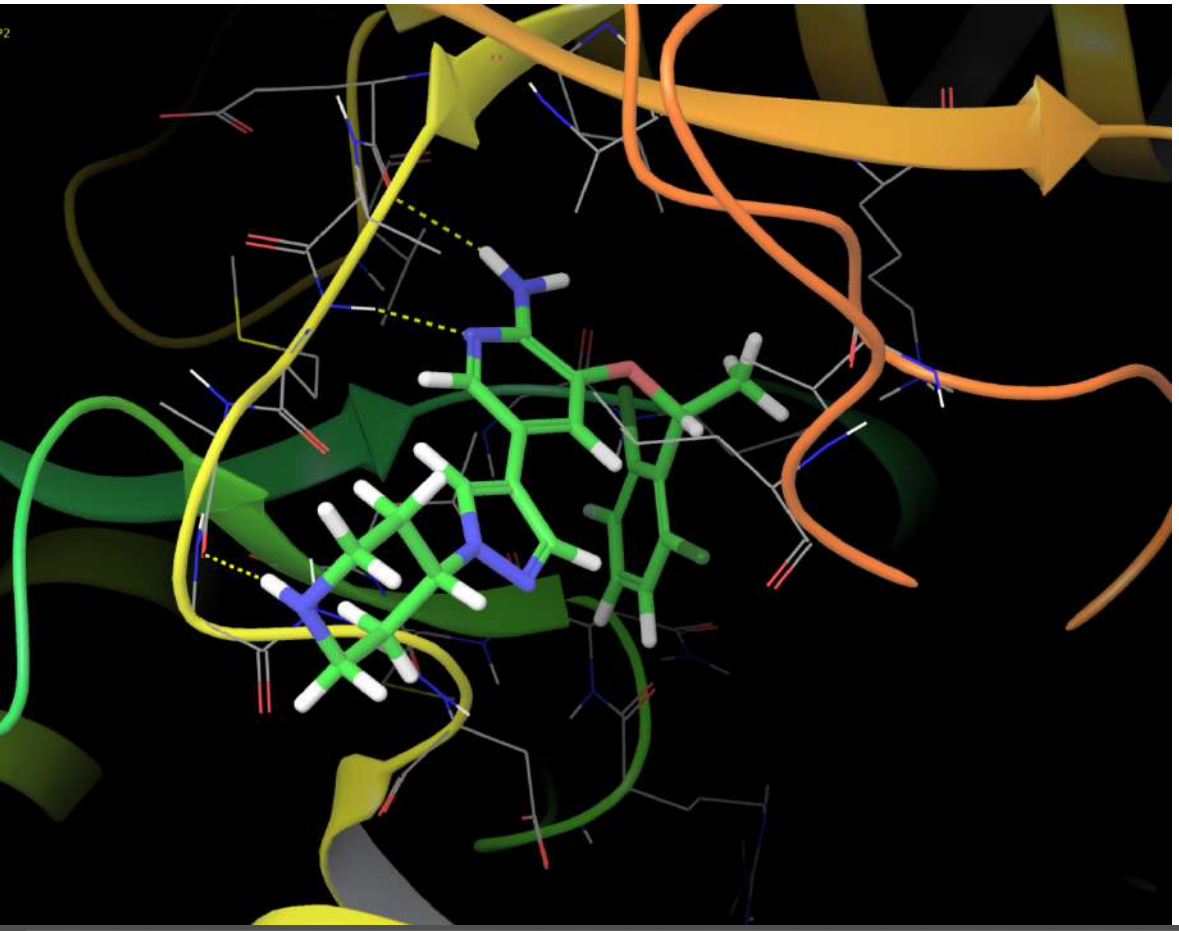
- 1000 - 3000 patients, 1-4 years
- Large-scale Efficacy, Dosing, and Safety Studies
- ~25% Success Rate

**Phase IV**

- Marketing
- Long term drug effects

# Drugs bind on protein pockets through intermolecular interactions

Structure of the anaplastic lymphoma kinase (ALK)
Complexed with the drug crizotinib – (PDB ID: 2XP2)



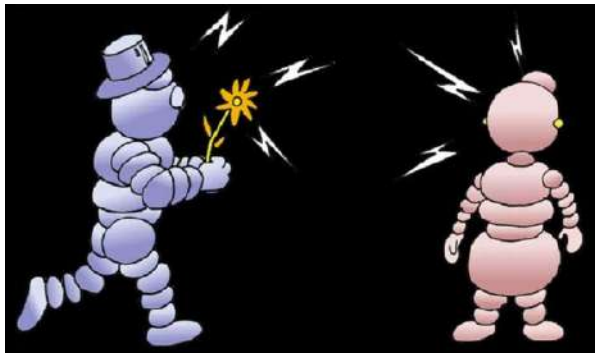
**Protein-Ligand interactions:**

**Intermolecular Interactions (Enthalpy)**

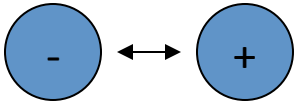
***Hydrogen Bonds***
***Electrostatic Interactions***
***van der Waals Forces***
***π – π Interactions***

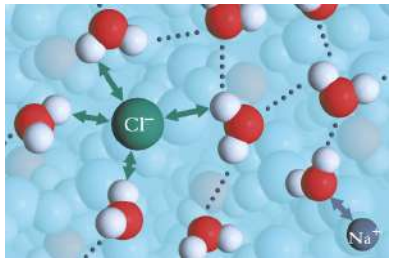
**Entropy**

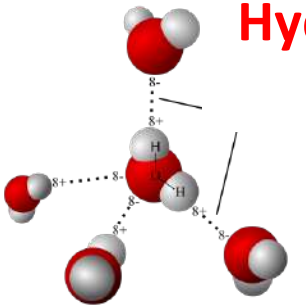
# Intermolecular Interactions



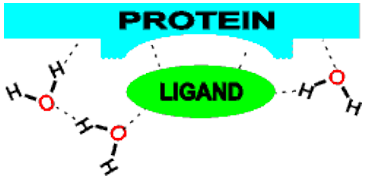
## Electrostatic Interactions

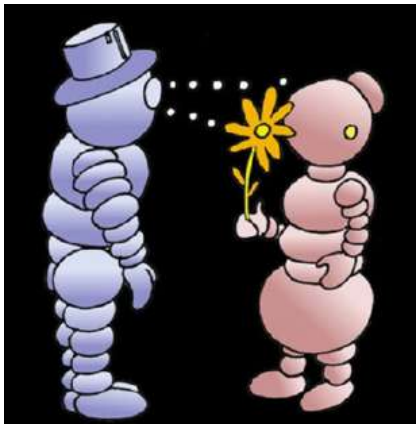


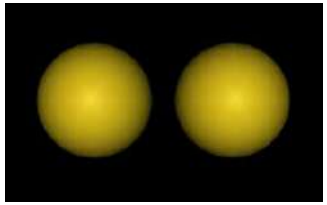



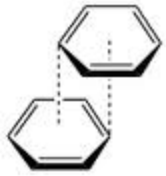
## Hydrogen Bonds

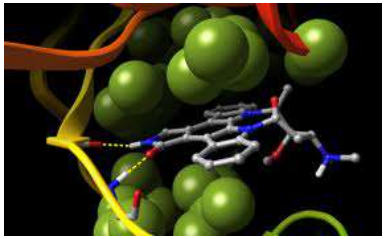


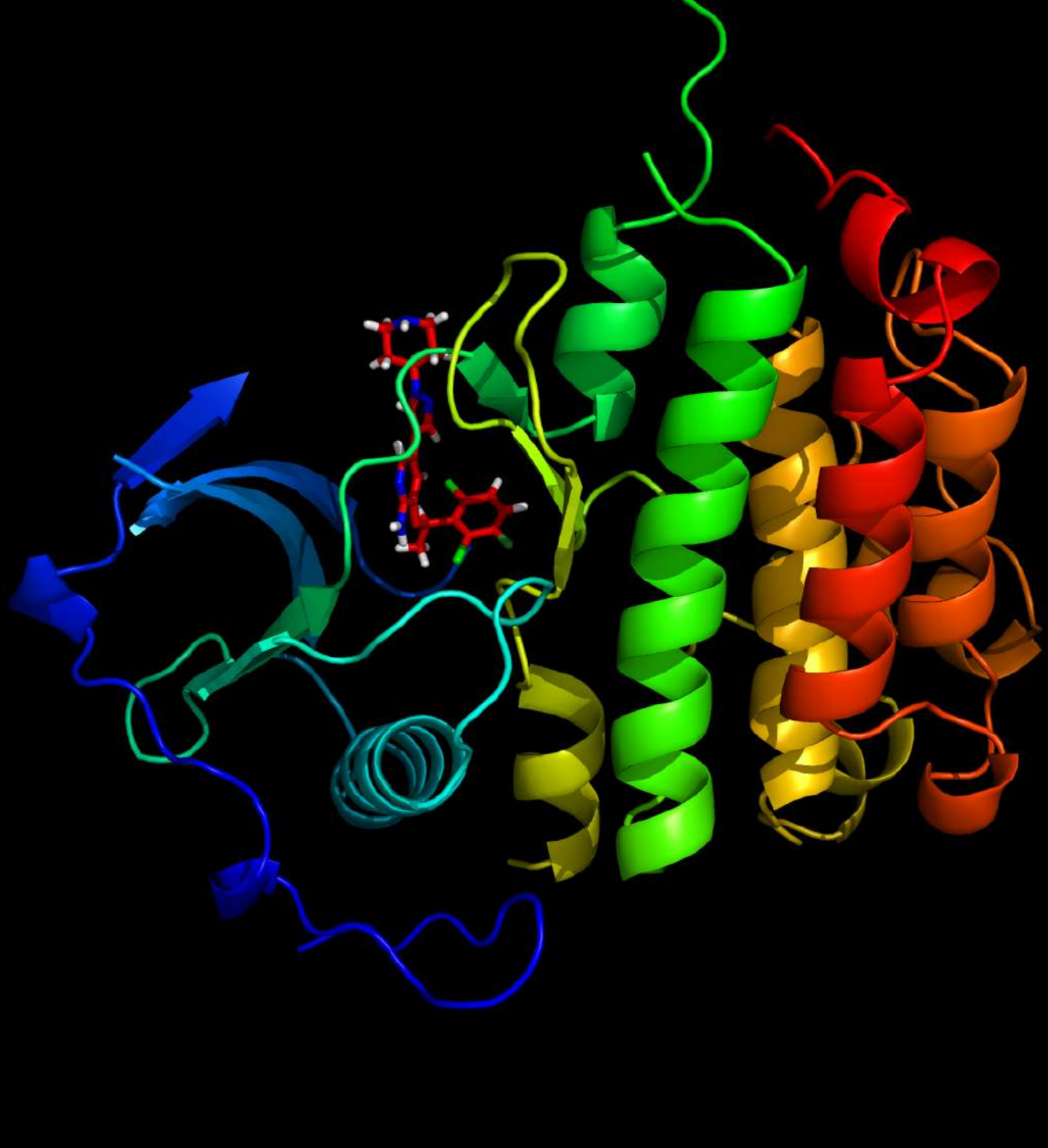
## van der Waals Forces
## π – π, cation - π interactions



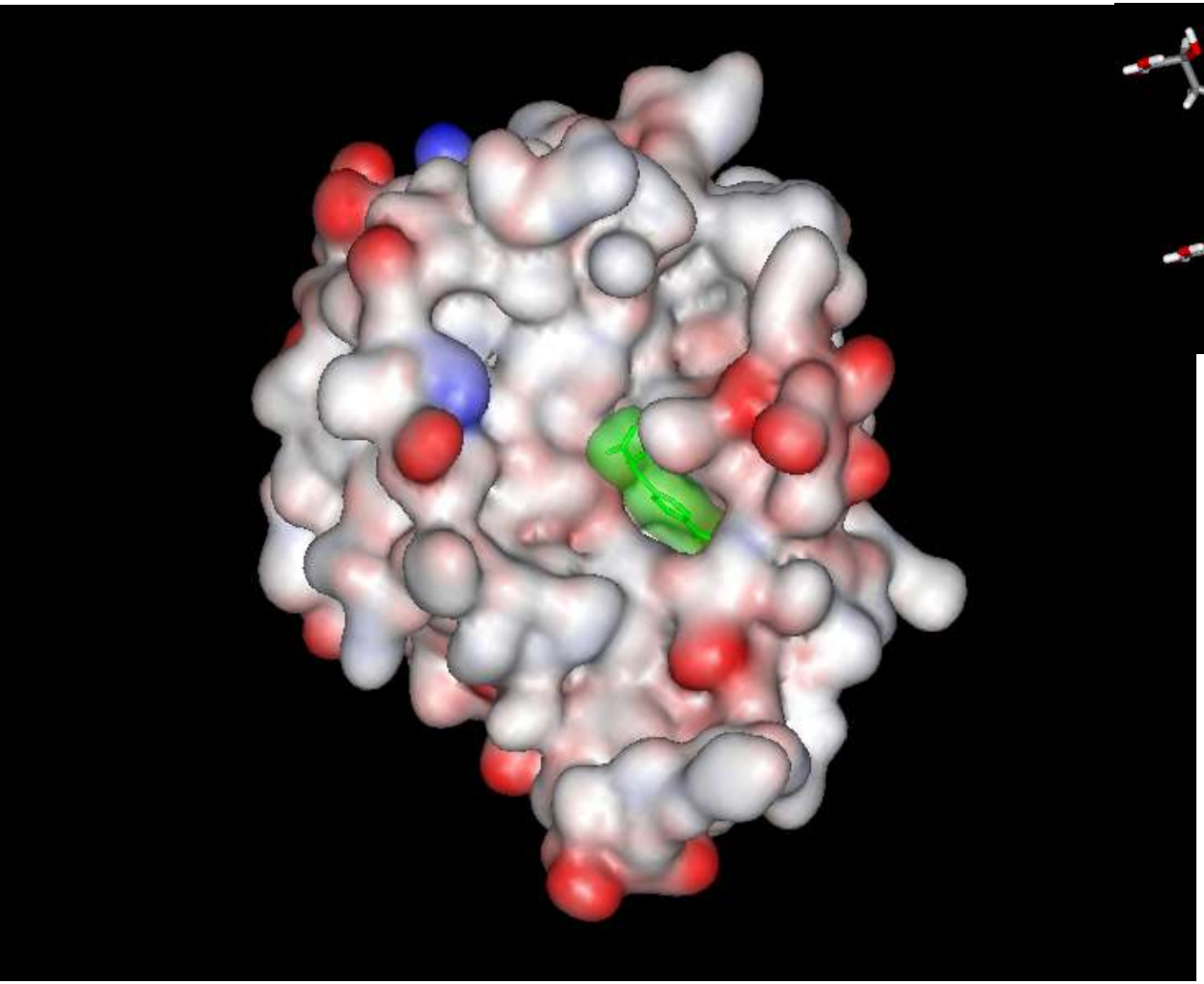


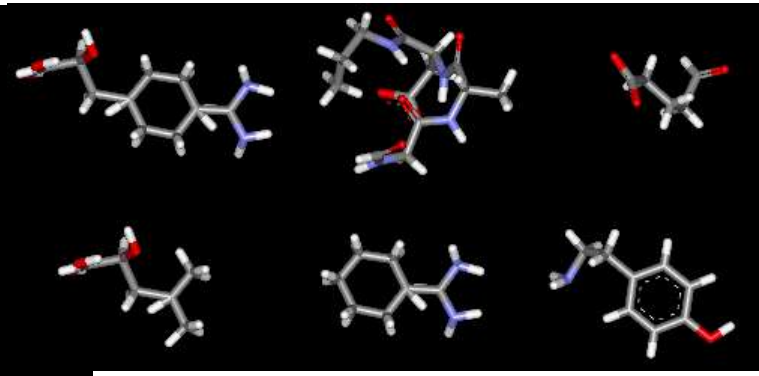


https://www.youtube.com/watch?v=u49k72rUdyc



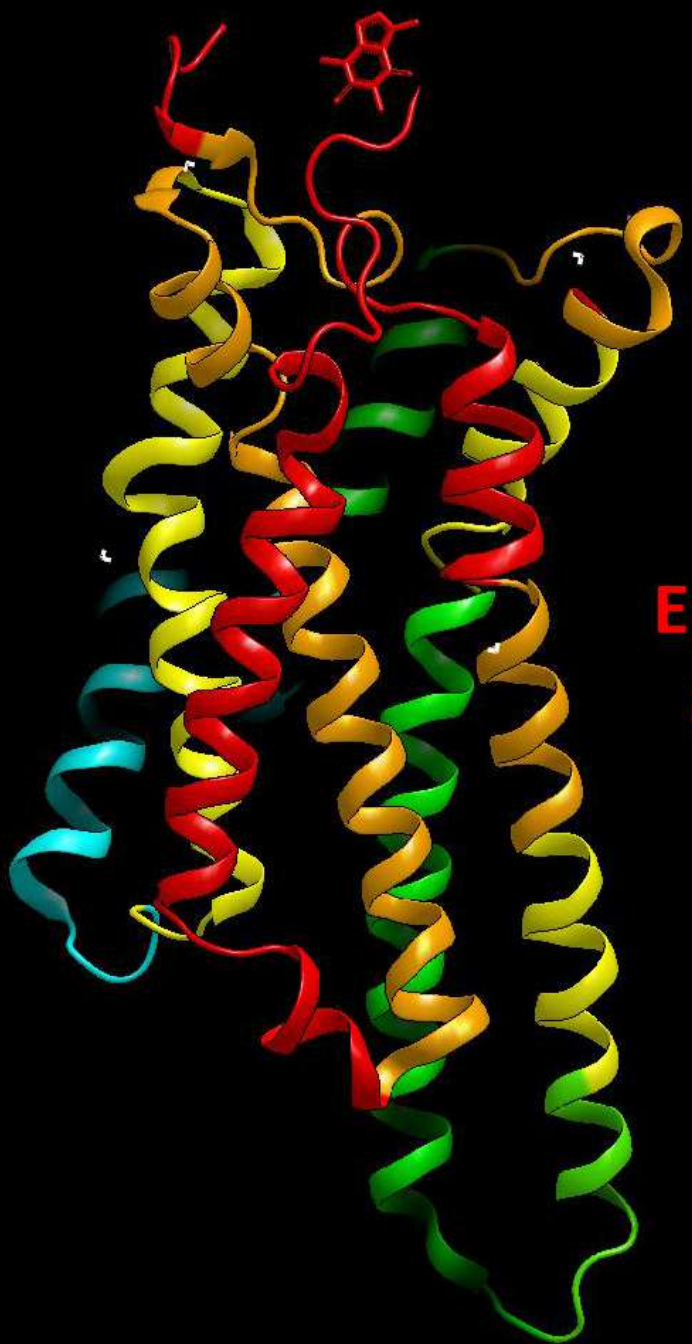
# Predicting the protein-ligand complex



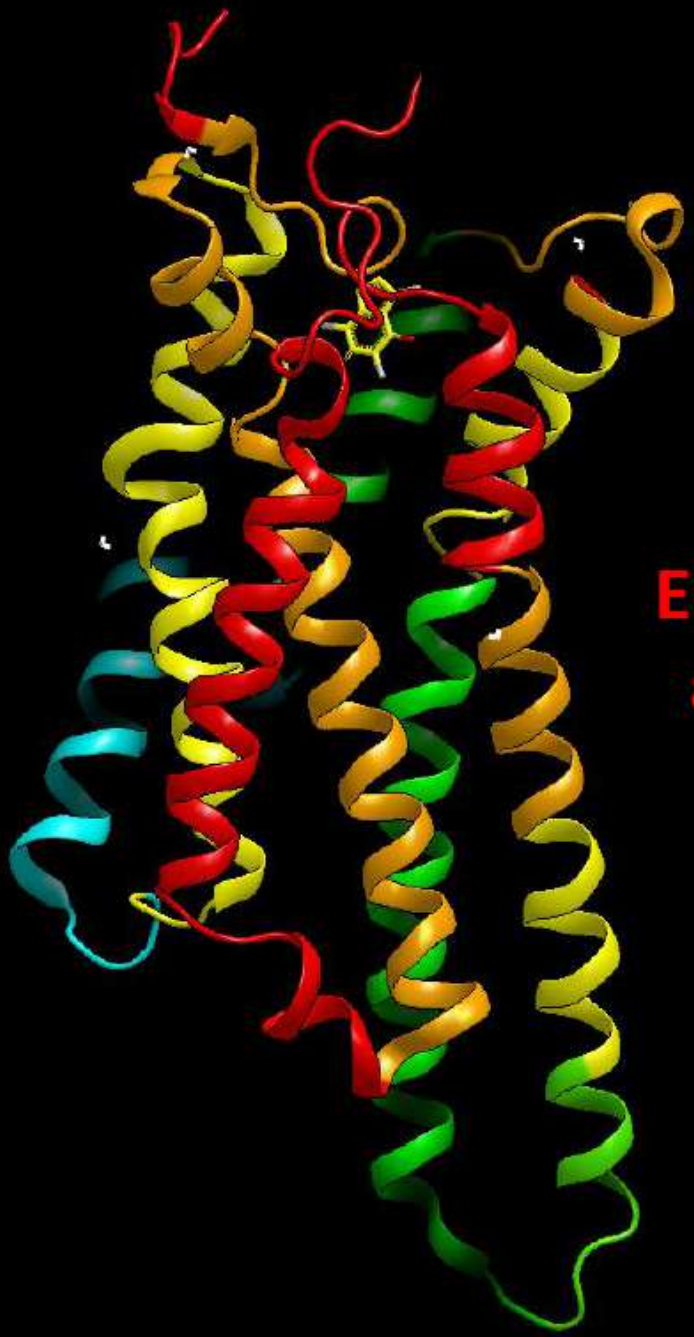


*Docking / Virtual Screening*

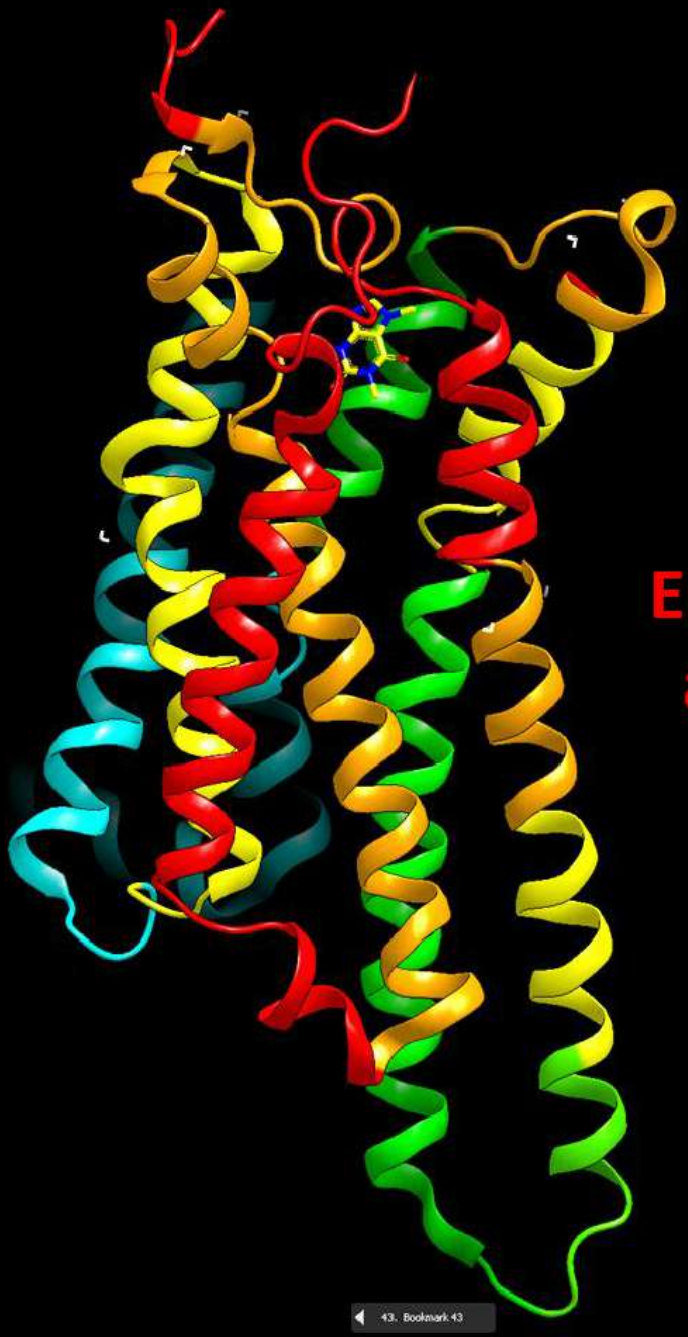
https://www.youtube.com/watch?v=u49k72rUdyc



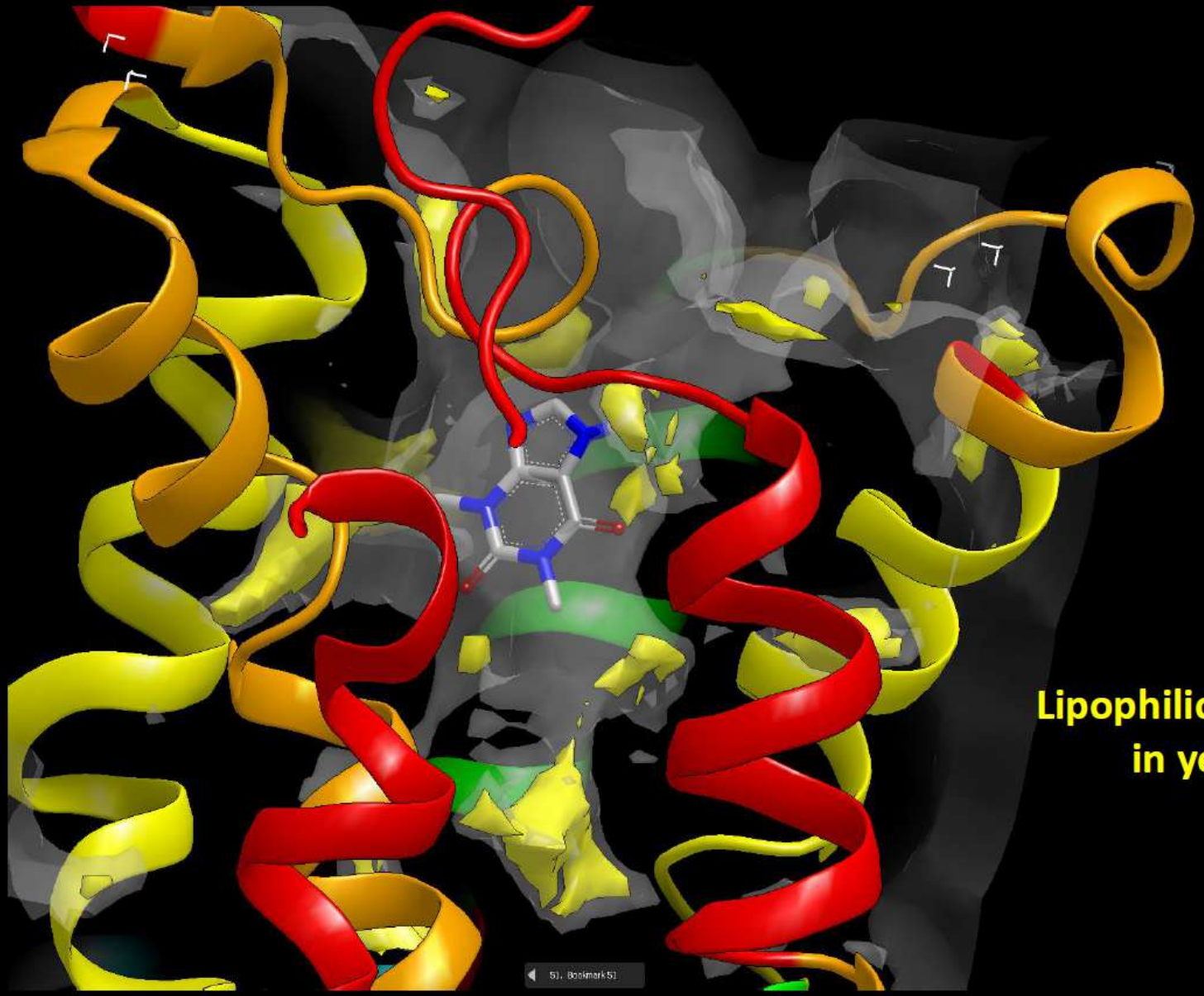
Entry of caffeine into the adenosine $A_{2A}$ receptor



Entry of caffeine into the adenosine $A_{2A}$ receptor

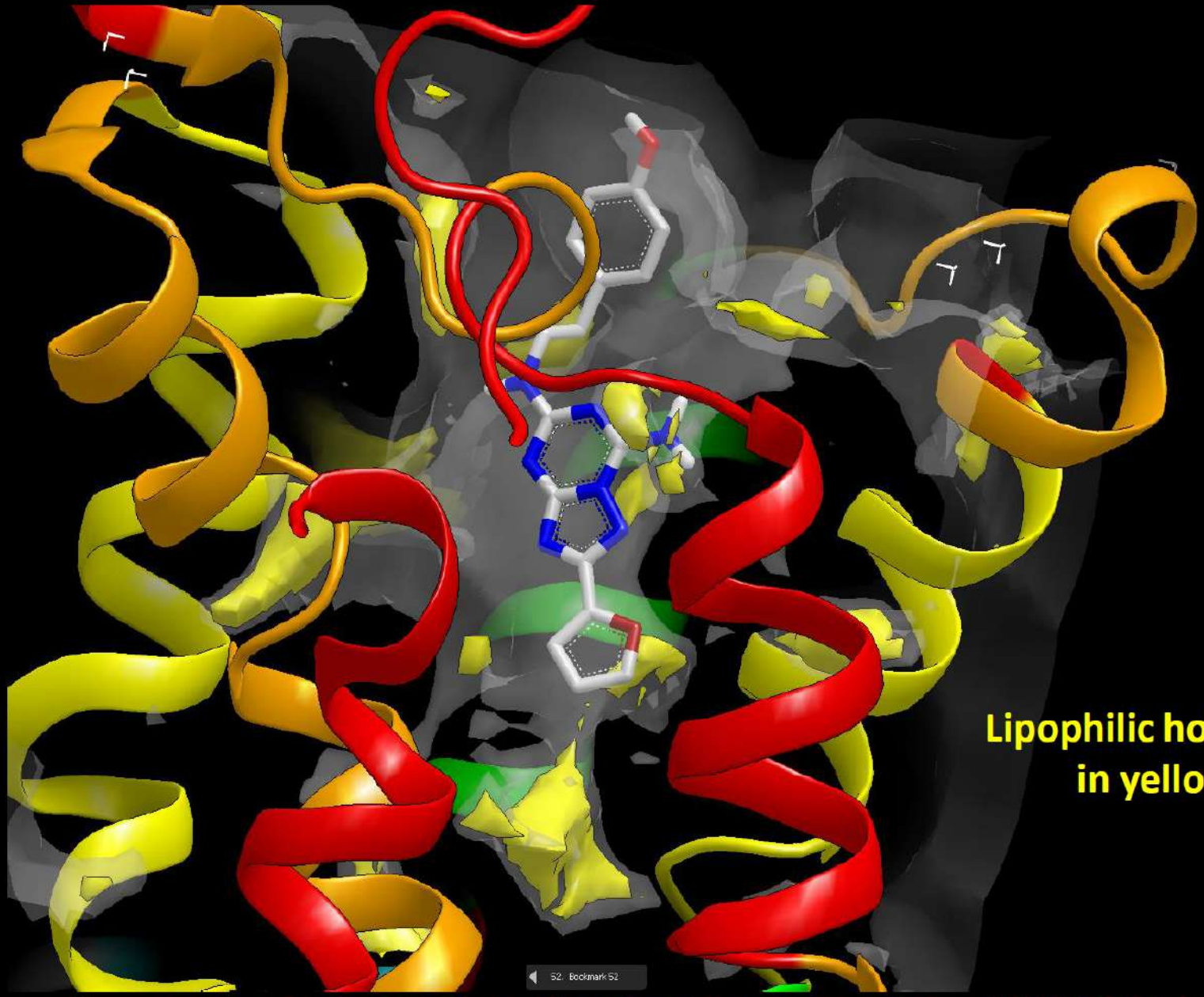



Entry of caffeine into the adenosine $A_{2A}$ receptor



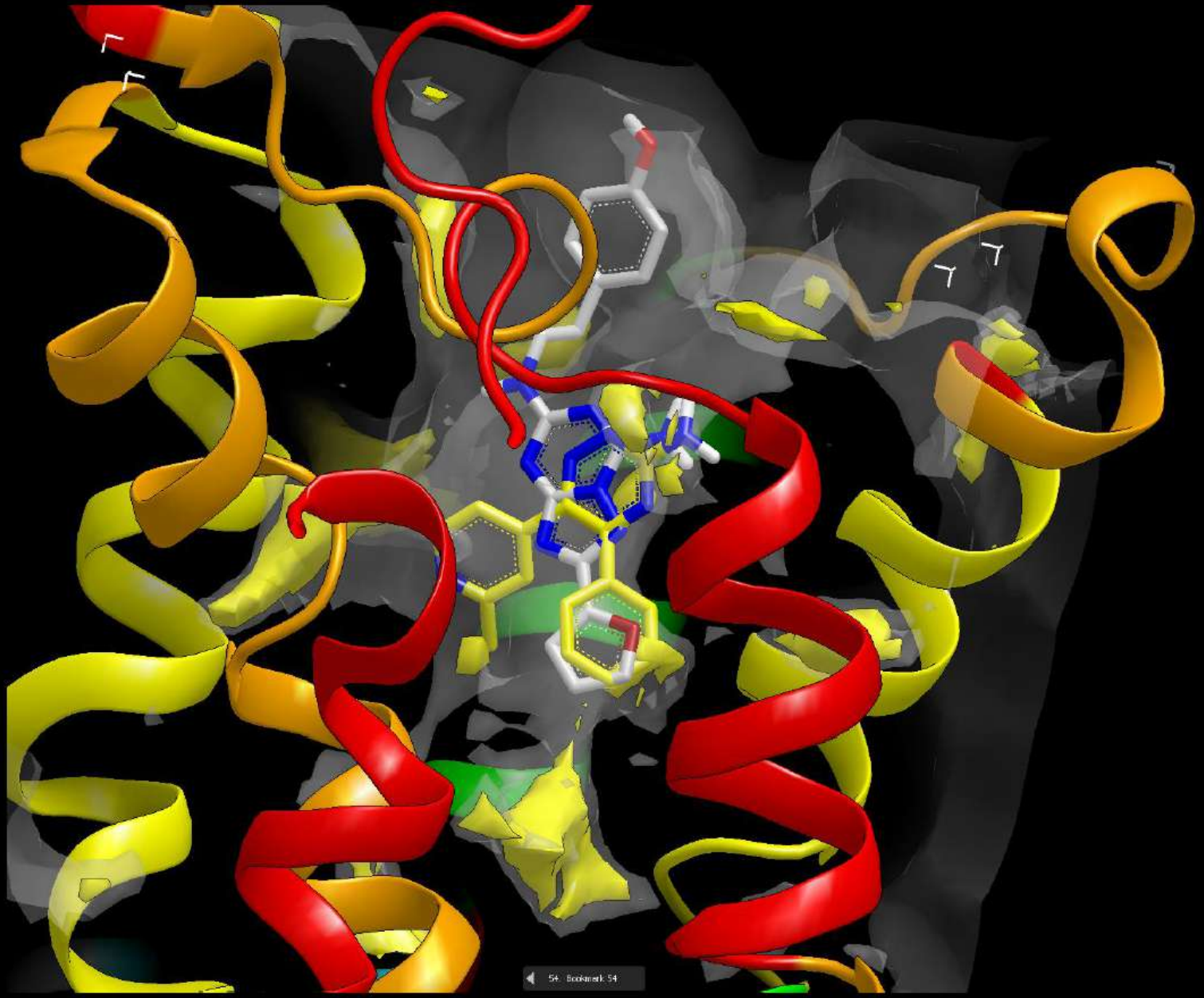

Lipophilic hotspots
in yellow

# The caffeine binding pocket in the $A_{2A}$ receptor
# A Neutral antagonist



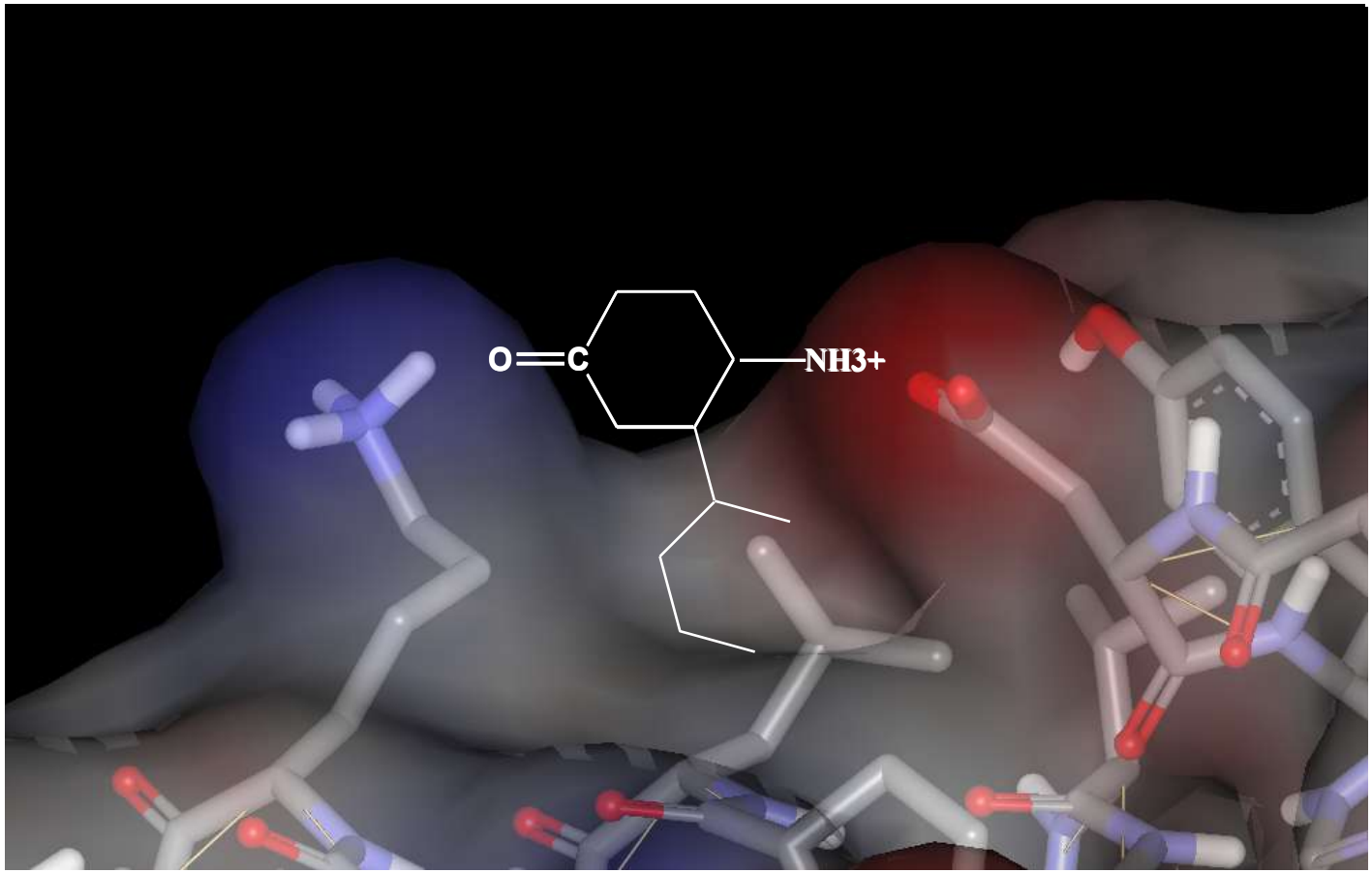

# $A_{2A}$ receptor bound to the inverse agonist - ZM241385

© 2014 Heptares Therapeutics



Overlay of ligands bound to the $A_{2A}$ receptor

# De novo computer-aided drug design



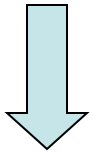

- Design of inhibitors from scratch based on 3D structure of protein
- With a ligand-growing program analogs are built inside protein binding site

# Scoring functions for molecular docking

**Scoring Functions → approximate the free energy of binding of a small molecule binding to a target**

- The algorithm performs a conformational search in the binding pocket of the target
- Some of the ligand conformations are rejected because of high-energy clashes with the protein

- The remainder conformations have to be assessed or ranked
- Different Ligands have to be ranked relative to each other



***Scoring Functions***
(approximate free energy of binding)

$$\Delta G_{bind} = \Delta G_{solv.} + \Delta G_{conf.} + \Delta G_{int.} + \Delta G_{rot.} + \Delta G_{t/r} + \Delta G_{vib.}$$

# Scoring functions for molecular docking

***SP***

$$\Delta G_{bind} = \Delta G_0 + \Delta G_{hbond} \Sigma f(\Delta R, \Delta\alpha) + \Delta G_{ionic} \Sigma f(\Delta R, \Delta\alpha) + \Delta G_{lipo} A_{lipo} + \Delta G_{rot} N_{rot}$$

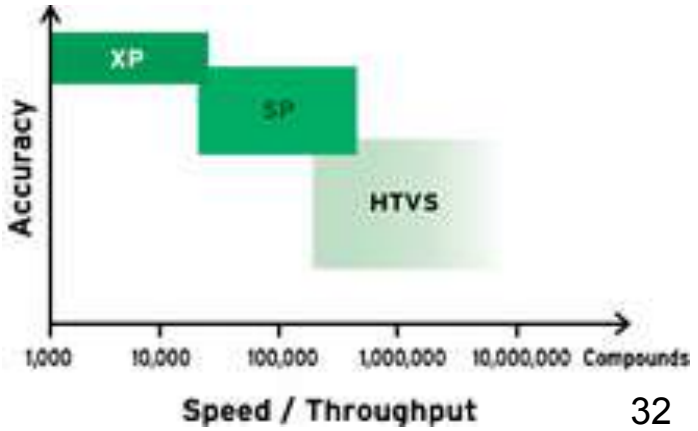
***XP***

$$\mathbf{XP\ GScore} = \boldsymbol{E}_{\mathbf{coul}} + \boldsymbol{E}_{\mathbf{vdW}} + \boldsymbol{E}_{\mathbf{bind}} + \boldsymbol{E}_{\mathbf{penalty}}$$

$$\boldsymbol{E}_{\mathbf{bind}} = E_{\text{phobic\_pair}} + E_{\text{hyd\_enclosure}} + E_{\text{hb\_nn\_motif}} + E_{\text{hb\_cc\_motif}} + E_{\text{hb\_pair}} + E_{\text{PI}}$$

$$\boldsymbol{E}_{\mathbf{penalty}} = E_{\text{desolv}} + E_{\text{ligand strain}}$$

• Parameters in scoring functions are being estimated based on training sets




Friesner, R.A.,., *et al* (2004) J Med Chem, 47, pp. 1739-1749
Friesner, R.A., *et al*., (2006) J Med Chem, 49, pp. 6177-6196

# Binding site Prediction

Crystal structure

↓

Molecular Dynamics

↓

Cluster conformations from trajectory (*RMSD*)

↓

Binding site prediction **(Q-Site Finder, PCA)**

# Virtual Screening

Zinc Database | Maybridge Hitfinder

↓

**Ligand conformations**

↓

*Glide Docking & Scoring*

**2,1 M cmpds** **SP** **24,000 cmpds**

↓

**40,000** **XP** **10,000**

↓

**1,000 (Top *Glide* ***Hits based on G-score)**

↓

***postprocessing***

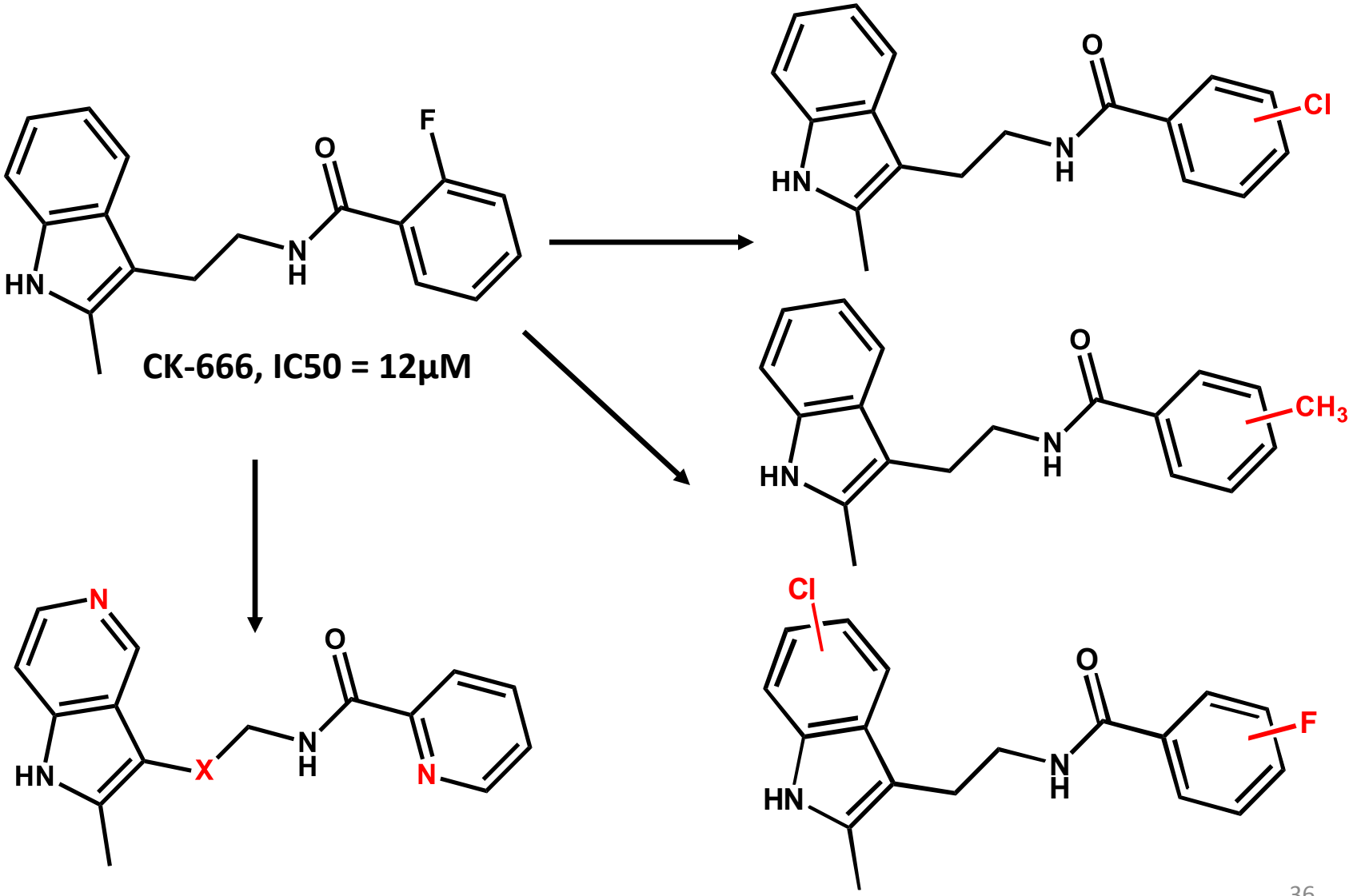
**30 compounds purchased and assayed in vitro**

# How are compounds selected for assaying?

- Estimation of binding affinity through docking score
- Conformation of ligand inside the protein (binding pose):
  - -Look out for bad van der Waals contacts
  - -Cis-trans amides
  - -E-Z esters
- Identification of unwanted or toxic moieties on a compound
- Identification of metabolic liabilities (benzylic hydrogens, p-position on benzene, sites of glucuronidation, etc…)
- Calculation of physicochemical properties (lipophilicity, cell permeability, solubility, etc). Choose molecules with drug-like Pchem profile.
- Clustering
- Chemical intuition

# Early pre-clinical phases where computation is key

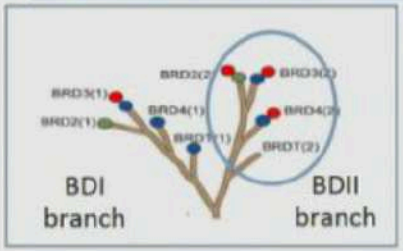
# Compound optimization for potency, selectivity, metabolism & physicochemical properties

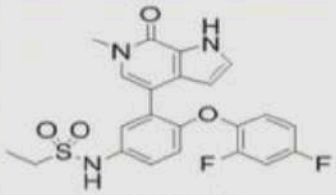


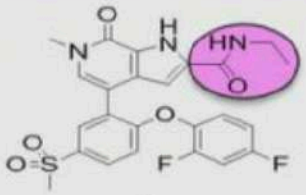
# Discovery of ABBV-744: A first-in-class highly BDII-selective BET bromodomain inhibitor

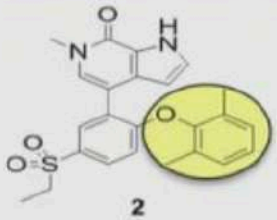
## Hypothesis: BDII-selective BET Inhibitors May Exhibit a Wider TI

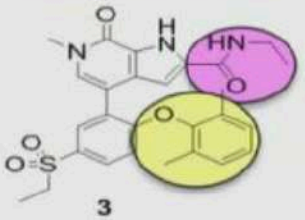
## Discovery of the BDII-selective Tool Compound 3



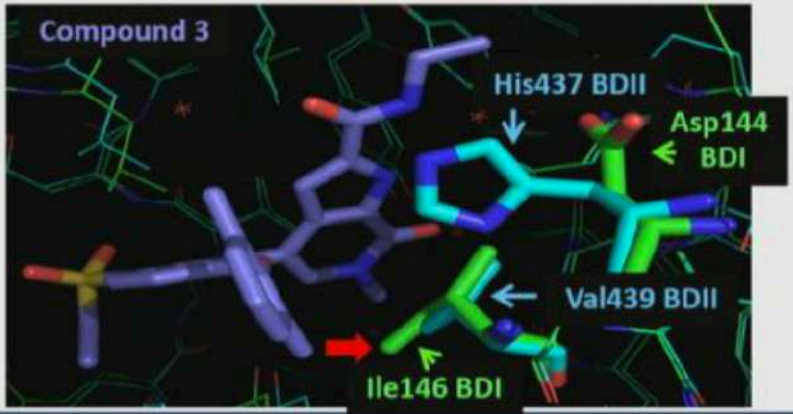


ABBV-075
mivebresib


1


2


3

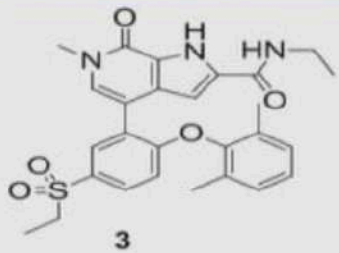
| | ABBV-075 | 1 | 2 | 3 |
|---|---|---|---|---|
| BDII selectivity | 2x | 17x | 7x | 110x |
| BRD4 BDI (nM) | 2.4 | 20 | 46 | 120 |
| BRD4 BDII (nM) | 1.2 | 1.2 | 6.9 | 1.1 |

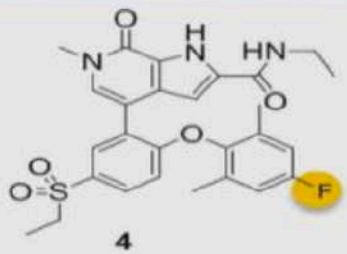


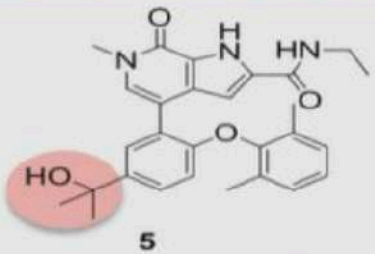

- Ethyl amide buries His437 (BDII) but not Asp144 (BDI)
- 2,6-Disubstituted phenyl clashes with Ile146 (BDI) (red arrow) but not with Val439 (BDII)
- Combination of ethyl amide and 2,6-dimethylphenyl enhances BDII-selectivity

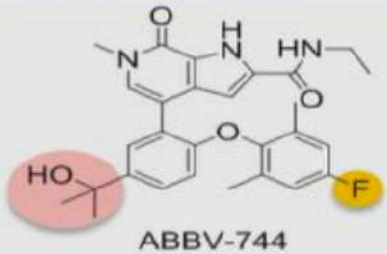
# Discovery of ABBV-744: A first-in-class highly BDII-selective BET bromodomain inhibitor

## Elaboration of the Series to Discover the Clinical Asset ABBV-744

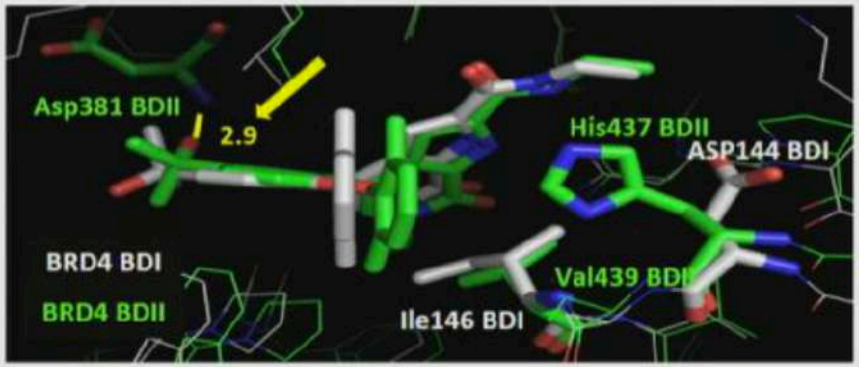


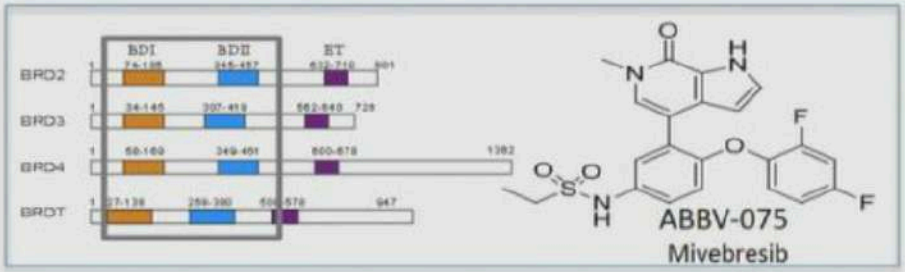

| | 3 | 4 | 5 | ABBV-744 |
|---|---|---|---|---|
| BRD4 BDII selectivity | 110x | 140x | 400x | 325x |
| Microsomal $Cl_{int,u}$ (L/hr/kg) (Mouse/Rat/Human) | 42/64/12 | 55/28/10 | 170/65/37 | 135/30/31 |



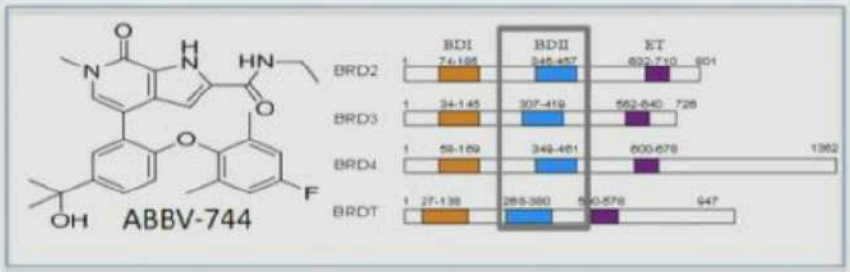

- 4-F on aryl ether provides metabolic stability (rat)
- Tertiary alcohol provided better selectivity and physical properties than sulfone
- Tertiary alcohol accepts H-bond to NH of Asp381 in BDII; no interaction in BDI

# Discovery of ABBV-744: A first-in-class highly BDII-selective BET bromodomain inhibitor

## Affinity and Selectivity of AbbVie BET Bromodomain Inhibitors




Pan-BET inhibitor: equal affinity for all 8 bromodomains




~300-fold selective for BDII vs. BDI bromodomains

| | Biochemical | | | | | | | | | NanoBRET cellular | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | BRD2 $K_i$ (nM) | | Select. | BRD3 $K_i$ (nM) | | Select. | BRD4 $K_i$ (nM) | | Select. | BRD4 $EC_{50}$ (nM) | | Select. |
| | BDI | BDII | (fold) | BDI | BDII | (fold) | BDI | BDII | (fold) | BDI | BDII | (fold) |
| ABBV-075 | 13 | 3.7 | 3 | 6.3 | 1.8 | 4 | 2.8 | 1.3 | 2 | 34 | 13 | 3 |
| ABBV-744 | 1160 | 4.6 | 250 | 3140 | 4.9 | 640 | 520 | 1.6 | 325 | 21,000 | 28 | 750 |

# Computing is transforming biomedical research




Exascale

LCF
Mira 786k cores
Titan 280k cores
+ GPUs

2015
Enveloped virus
200 mil+ atoms
1-100 μs

Ranger
60k CPUs

2007
ribosome
2 mil atoms
100s ns

Compute Power

LeMieux
3k CPUs

HP 735
12 CPUs

SGI Origin
128 CPUs

2002
ATPase
500k atoms
10s ns

1993
protein
10k atoms
100s ps

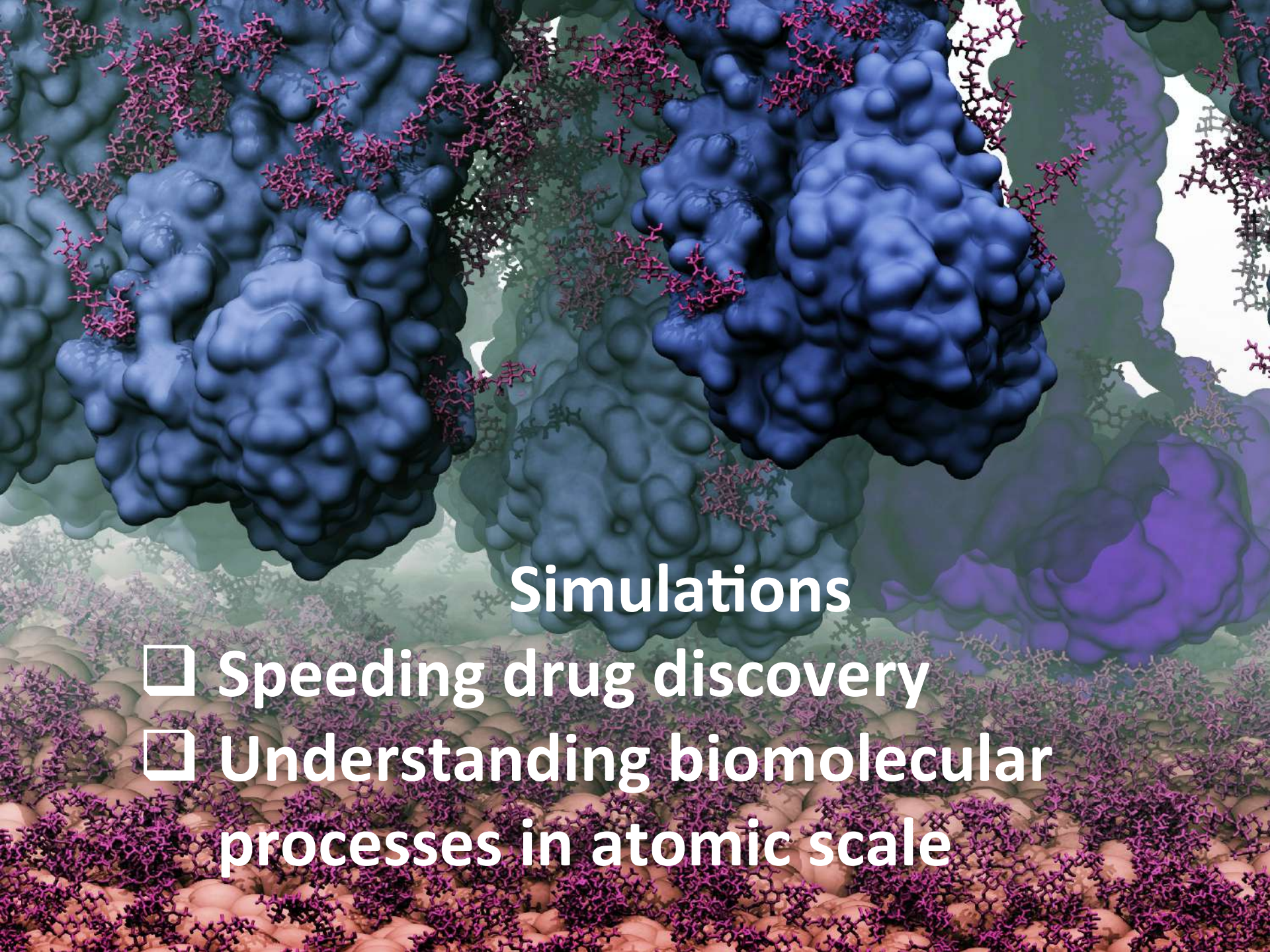
1997
ion channel
100k atoms
1 ns

time

**Performance (in FLOPS):**

| | |
|---|---|
| Megaflop | $10^6$ |
| Gigaflop | $10^9$ |
| Teraflop | $10^{12}$ |
| Petaflop | $10^{15}$ |

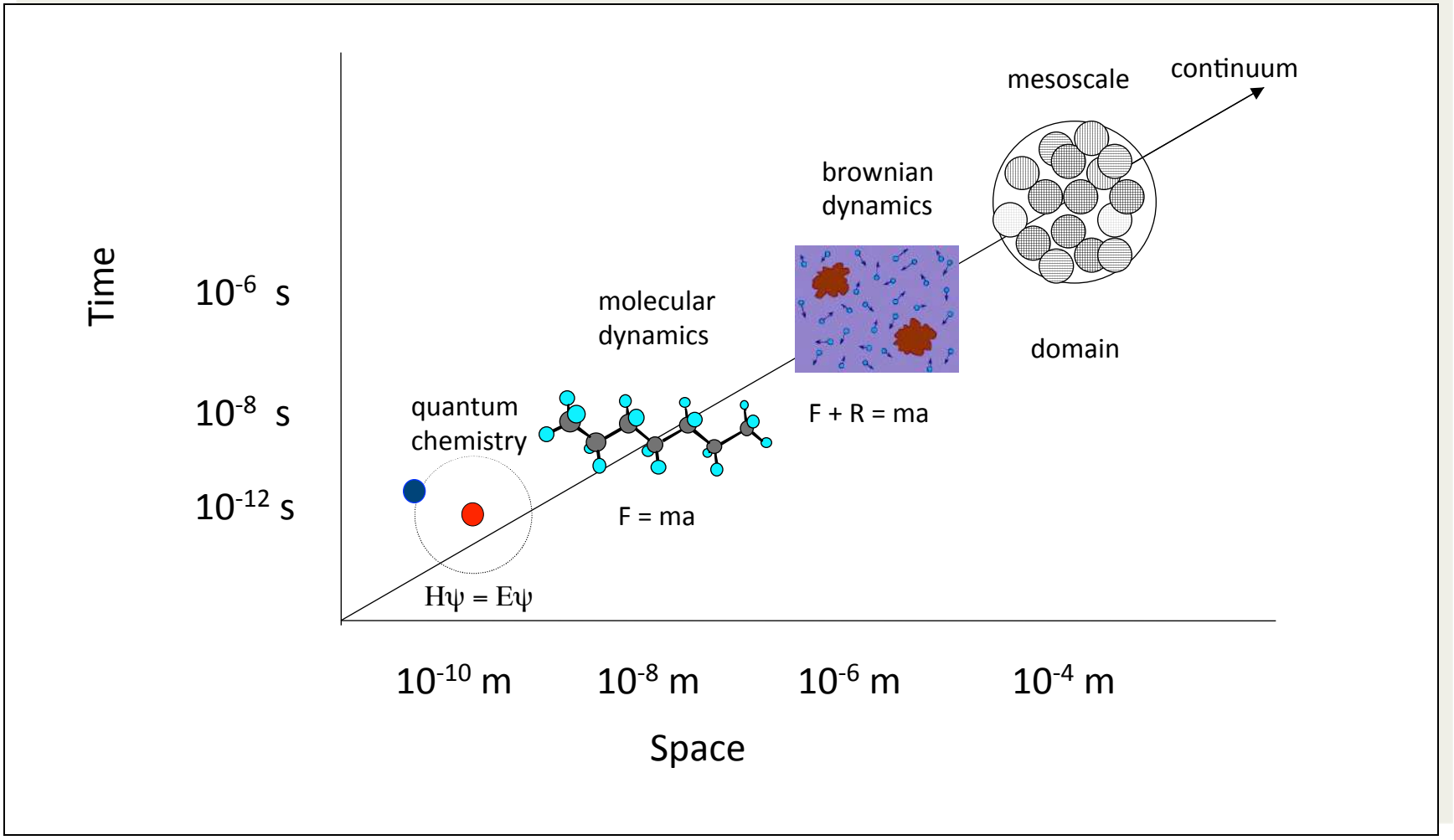
Source: adapted from Prof. Rommie Amaro



## Simulations

- Speeding drug discovery
- Understanding biomolecular processes in atomic scale

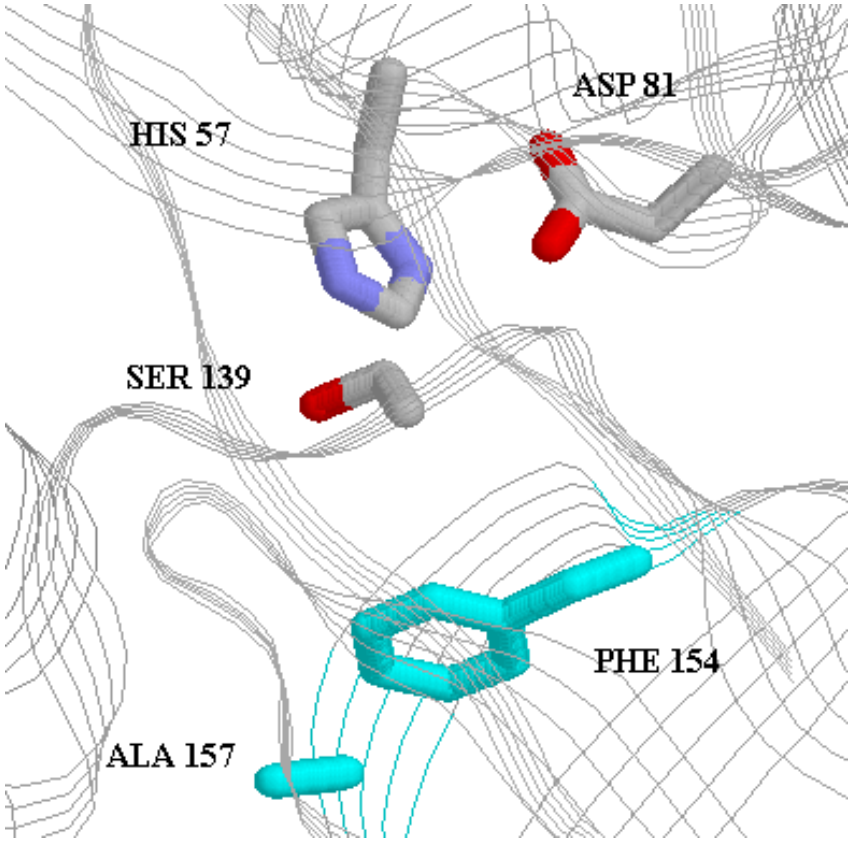
# Molecular Simulations across scales



# Molecular Modeling

*dynamics*

Structure --------> function

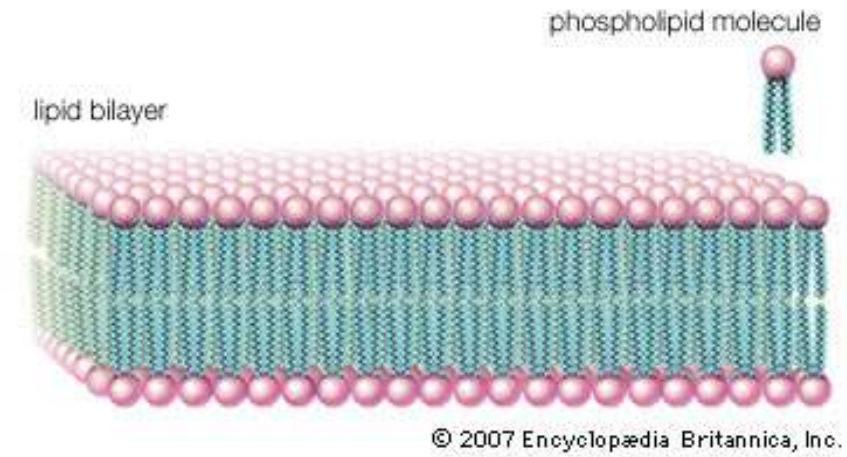



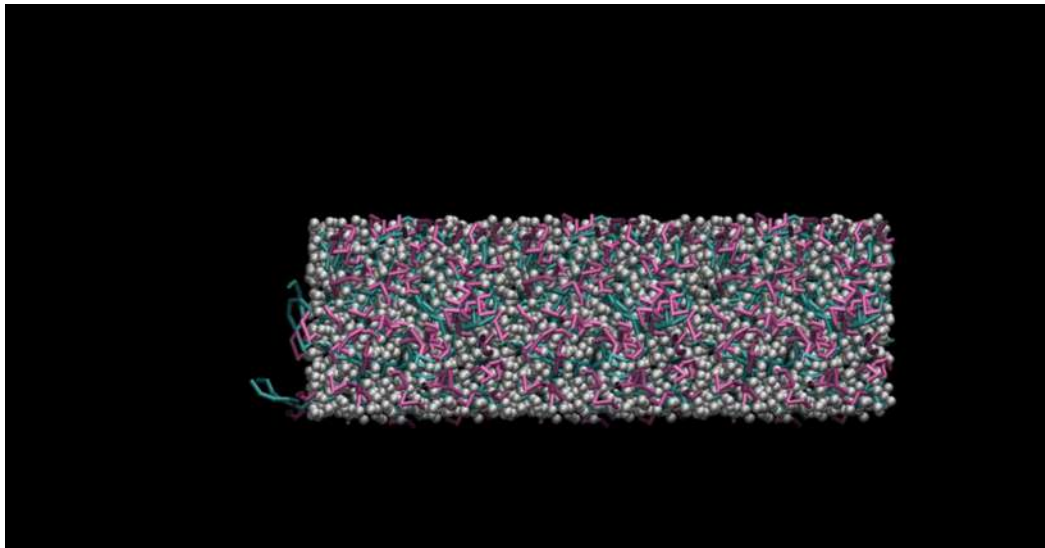

## Molecular Dynamics

- molecular/atomic level picture of structure and dynamics
- property prediction
- ion transport
- solvent effects
- protein stability / conform. changes, …

# Molecular Dynamics

- A computational method which describes equilibrium and dynamics properties of a biological system
- Generates configurations of the system by integration of Newton's law of motion –calculate the time dependence of the molecular system
- Generates information at the microscopic level –atomic positions and velocities and connects to macroscopic properties through Statistical Mechanics
- Connects structure and function by providing additional information to experimental techniques through the system dynamics

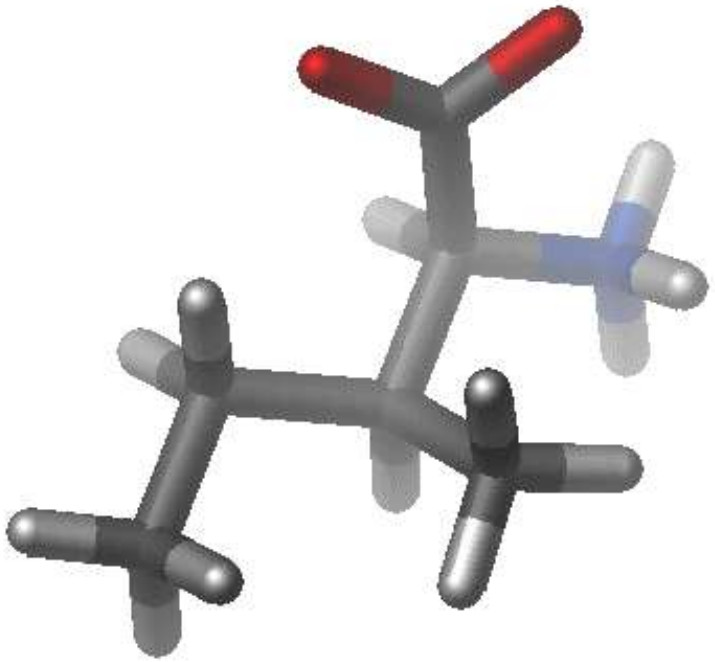


# Statistical Mechanics

- In Molecular Dynamics simulations we explore the **macroscopic** properties of a system through **microscopic** simulations
- The connection between microscopic simulations and macroscopic properties is made via **statistical mechanics**, which studies a macroscopic system from a molecular point of view
- The distribution of the system within the ensemble follows the **Boltzmann** distribution
- **Ensemble:** collection of all possible systems which have different microscopic states but identical macroscopic or thermodynamic state

# Biomolecular Simulations



# The Nobel Prize in Chemistry 2013

© Nobel Media AB

**Martin Karplus**

Photo: Kellana via Wikimedia Commons

**Michael Levitt**

Photo: Wikimedia Commons

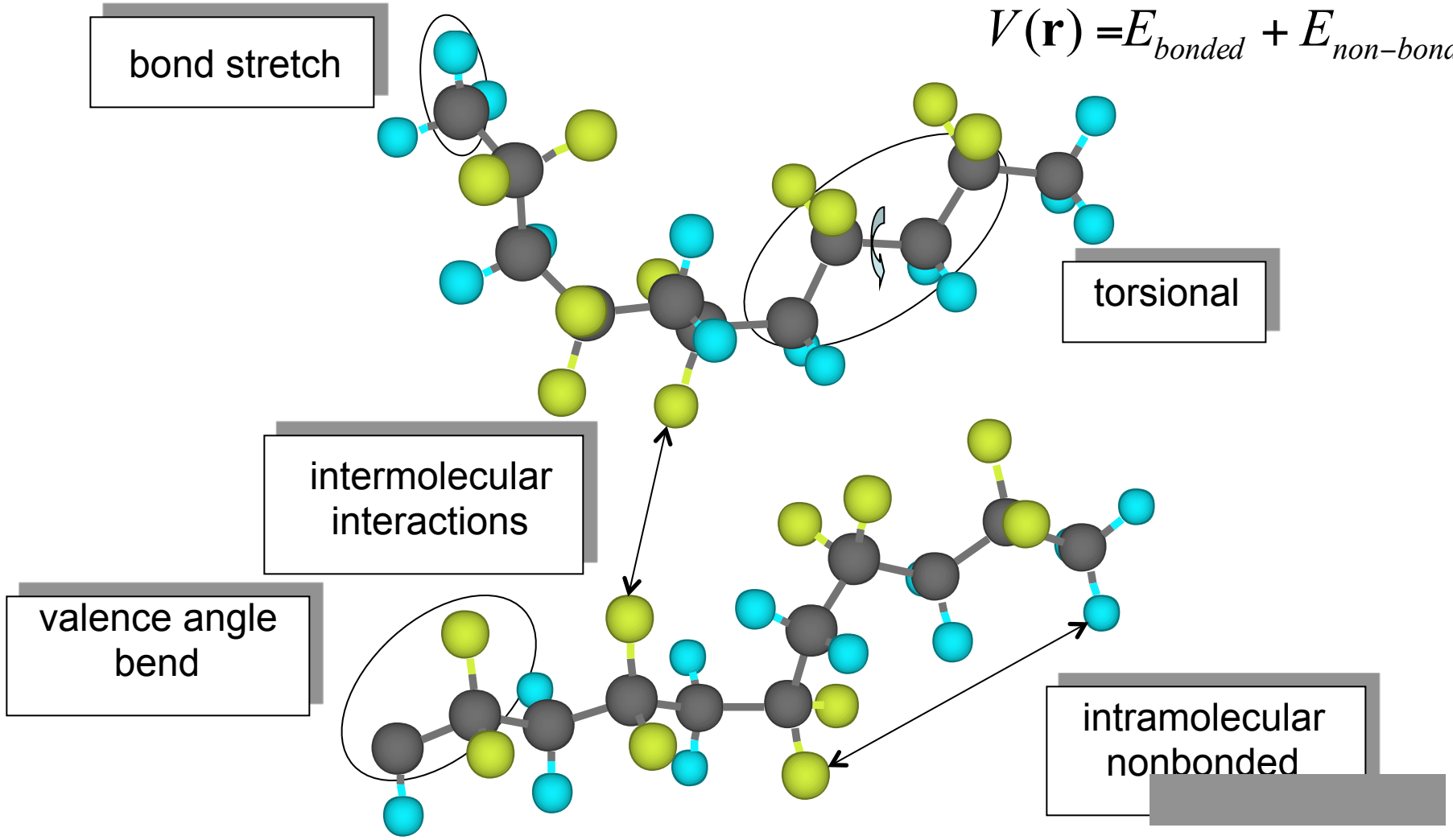
**Arieh Warshel**

**The Nobel Prize in Chemistry 2013 was awarded jointly to Martin Karplus, Michael Levitt and Arieh Warshel *"for the development of multiscale models for complex chemical systems"*.**

# The Potential Energy Function (Force Field)

The energy of the system is represented by the Hamiltonian: $H = K + V = \frac{1}{2} m\mathbf{v}^2 + V(\mathbf{r})$

$$V(\mathbf{r}) = E_{bonded} + E_{non-bonded}$$

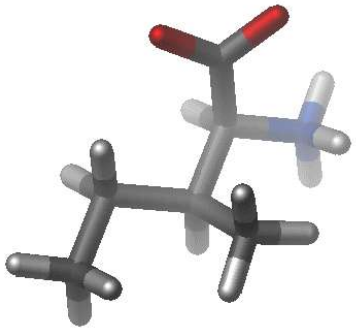


# Modeling the Potential E: Bond stretch potential

- Molecules undergo vibrational motion, which is modeled as a harmonic potential according to HOOKE's law

$$F = -kx = -\nabla V(x)$$

$$V(x) = E_{bond-stretch} = \sum_{1,2\,pairs} k_b (b - b_0)^2$$



- $K_b$ represents the force constant and $b_0$ represents the equilibrium value around which the bond oscillates

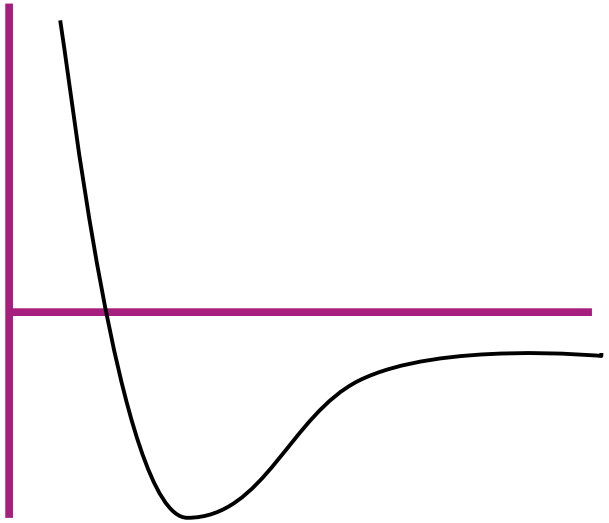
- This harmonic potential is valid only for deviations of 0.1 Å or less

# Harmonic vs Morse potential

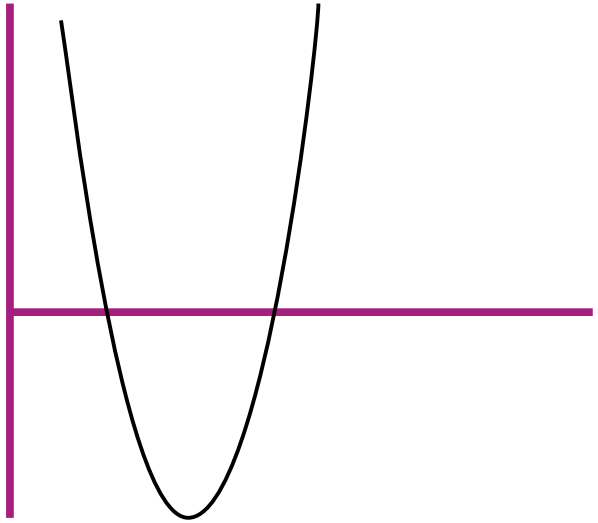
- The Morse term is more accurate, however it is generally not used in MD simulations since it requires 3 parameters to be specified for each bond

$$v(l) = D_e \left\{ 1 - \exp\left[ -a(l - l_0) \right] \right\}^2$$

- The Morse potential would allow a bond to stretch to an unrealistic length and break



Morse potential for a C-H bond



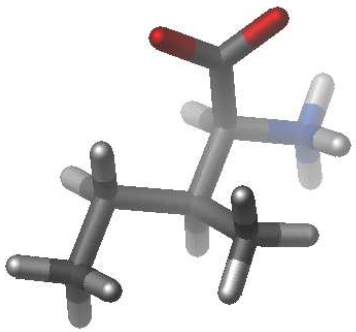
Harmonic potential for a C-H bond

# Bond angle potentials

- Describe the deviation from an ideal bond angle geometry

$$E_{bond-bend} = \sum_{angles} K_\theta (\theta - \theta_0)^2$$

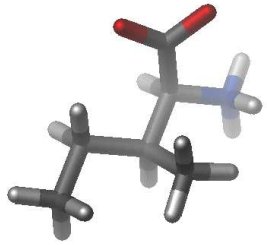
- $K_\theta$ represents the angle bending constant, $\theta_0$ represents the deviation from the ideal bond angle



# Torsion angle potentials

- This terms models the steric barrier between atoms separated by 3 covalent bonds

$$E_{rotate-along-bond} = \sum_{1,4\ pairs} K_{\phi}\left(1-\cos\left(n\phi\right)\right)$$



- The motion associated is rotation, described by a dihedral angle around the middle bond

- The potential is assumed to be periodic and expressed as a cosine function

- $K_{\varphi}$ represents rotation constant, n represent the periodicity of the rotational barrier and $\varphi$ the dihedral angle

# Electrostatic interactions: The Coulomb potential

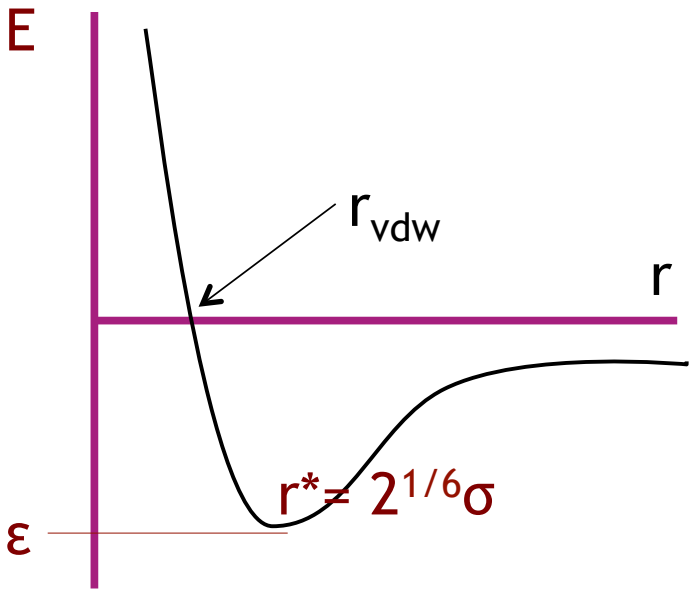
- Electrostatic interaction decays slowly with distance, considered long range interactions. Can be modeled by Coulomb's law.

$$E_{electrostatic} = \sum_{i,j} \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}}$$

- $r_{ij}$ represents the distance between two atoms having charges $q_i$ and $q_j$

- $\varepsilon_0$ represents the vacuum permittivity, a number relating the ability of a material to carry current

# The van der Waals potential: Lennard-Jones

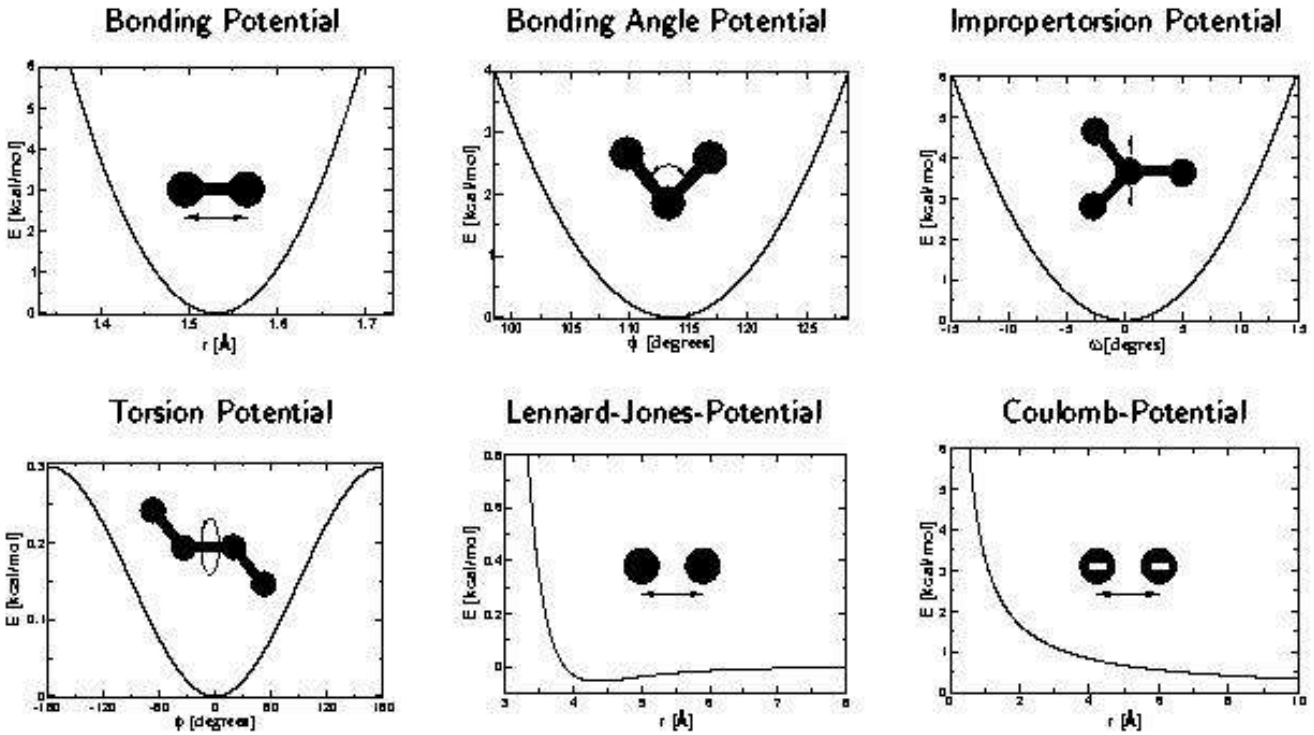



VdW energy best described by a Lennard-Jones potential

$$E_{vdw} = \sum_{i,j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

- Expresses the interaction energy between two atoms
- Contains an attractive part and a repulsive part
- Attractive forces due to London forces (dipole –dipole interaction)
- Repulsive part due to Pauli-exclusion principle and inter-nuclear repulsion
- ε is the depth of the potential well, σ is the finite distance at which the inter-particle potential is zero

# The Potential Energy Function (Force Field)




$$E = \frac{1}{2} m\mathbf{v}^2 + V(\mathbf{r}) \qquad \mathbf{F}_i = -\nabla V(\mathbf{r}) \qquad V(\mathbf{r}) = E_{bonded} + E_{non-bonded}$$

$$E_{bonded} = \sum_{bonds} k_b (b - b_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} k_\phi (1 + \cos[n\phi - \delta]) + \sum_{impropers} k_\omega (\omega - \omega_0)^2$$

$$E_{non-bonded} = \sum_{i,j} 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \sum_{i,j} \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}}$$

# MD Simulations study structure + dynamics

Is there a fast and efficient way to study the structure and dynamics of biomolecules in atomic-level detail?

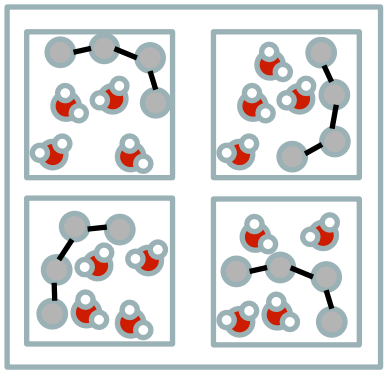
⇩

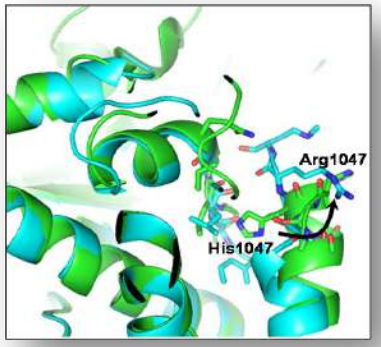
## Molecular Dynamics simulations

**Step 1.** Model the potential energy and use coordinates from experimental structures and assign initial velocities (Etotal = Epotential + Ekinetic)

**Step 2.** Integrate Newton's second law and get the new velocities (**v**) of the system and the new coordinates (**r**) of the atoms

**Step 3.** Macroscopic properties can be expressed through **v** and **r** *via statistical mechanics*

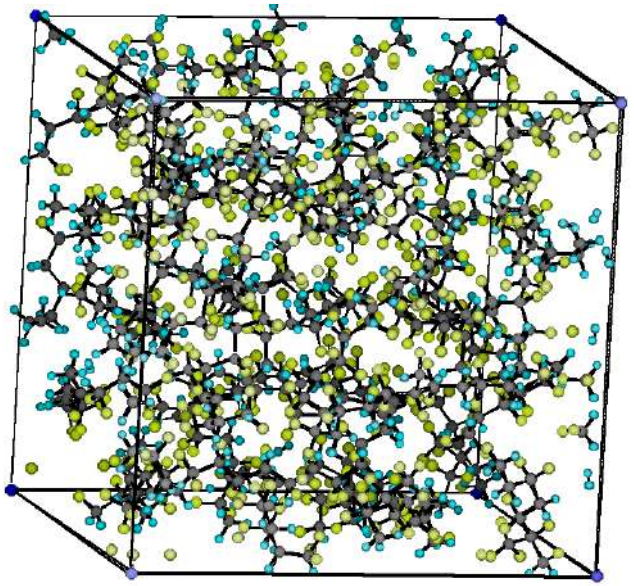





System of interest

# Μοριακή Δυναμική Προσομοίωση

**Δυνατότητες**



- Περιγραφή συστήματος σε ατομικό επίπεδο
- Συσχέτιση δομής και λειτουργίας συστήματος
- Υπολογισμός δυναμικής πρωτεϊνης-φαρμάκου
- Επίδραση του διαλύτη, υπολογισμός διάχυσης κ.ά.

# Παραδοχές Μοριακής Δυναμικής

- Προσέγγιση Born-Oppenheimer
- Βαρείς πυρήνες → Μοντελοποιούνται σαν σημειακές μάζες και η κίνησή τους περιγράφεται κλασικά
- Δημιουργία & σπάσιμο δεσμών δεν μπορούν να μοντελοποιηθούν
- Εργοδική υπόθεση → Σύνδεση προσομοίωσης με εργαστήριο
- Τα άτομα αλληλεπιδρούν με κλασικά δυναμικά για τα οποία χρησιμοποιούμε εμπειρικές παραμέτρους
- Περιοδικές οριακές συνθήκες

# MD Formalism

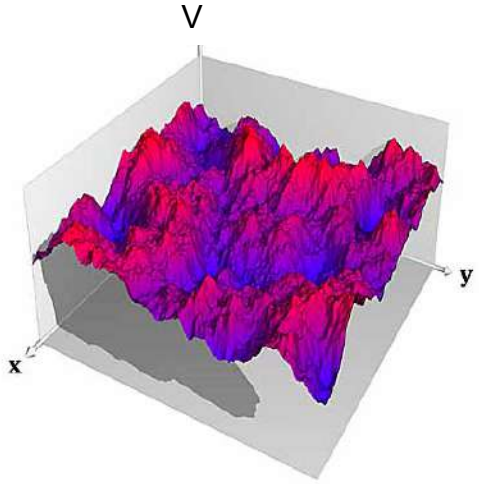
- Initial coordinates are taken from experimental structures and velocities from a distribution, e.g. Maxwell-Boltzmann

- Newton's equation of motion

$$\mathbf{F}_i = m_i \mathbf{a}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2}$$

- The force can be written as the gradient of the potential energy

$$\mathbf{F}_i = -\nabla_i V(\mathbf{r})$$




- Combine the two equations to get

$$\frac{dV(\mathbf{r})}{d\mathbf{r}} = -m_i \frac{d^2 \mathbf{r}_i}{dt^2}$$

- A trajectory is obtained by solving this differential equation

# How to integrate Newton's equation of motion?

- The potential energy is a function of the atomic positions of all the atoms in the system.
- Due to this complexity there is no analytical solution
- Use algorithms to obtain the positions, velocities, accelerations at a later time $t + \delta t$ to a sufficient degree of accuracy
- $\delta t$ is limited by the fastest vibration of the system, ie. the C-H bond ($\delta t$ = 1 fs = $10^{-15}$ s)
- An estimate of the positions, velocities, etc may be obtained with **Taylor's expansion**

$$\underset{\text{new position}}{\mathbf{r}(t+\delta t)} = \underset{\text{old position}}{\mathbf{r}(t)} + \underset{\text{old velocity}}{\delta t \mathbf{v}(t)} + \frac{1}{2}\underset{\text{acceleration}}{\delta t^2 \mathbf{a}(t)} + \ldots$$

$$\underset{\text{new velocity}}{\mathbf{v}(t+\delta t)} = \underset{\text{old velocity}}{\mathbf{v}(t)} + \underset{\text{acceleration}}{\delta t \mathbf{a}(t)} + \ldots$$

# Examples of numerical algorithms: Verlet

- Common use is the **VERLET** algorithm.
- For a differential equation of second order of the type $\frac{d^2\mathbf{r}(t)}{dt^2} = V(\mathbf{r}(t))$

with initial conditions $\mathbf{r}(t_0) = \mathbf{r}_0$ and $\frac{d\mathbf{r}(t_0)}{dt} = \mathbf{v}_0$ , an approximate numerical solution $\mathbf{r}_n \approx \mathbf{r}(t_n)$ at the times $t_n = t_0 + n\delta t$ may be obtained by the method:

- set $\mathbf{r}_1 = \mathbf{r}_0 + \mathbf{v}_0\delta t + ½\, V(\mathbf{r}_0)\delta t^2$
- for $n$ = 1,2 iterate:

$$\mathbf{r}_{n+1} = 2\mathbf{r}_n - \mathbf{r}_{n-1} + \mathbf{v}(r_n)\delta t^2$$

- In MD, each position is determined from the current position and position at time $t - \delta t$

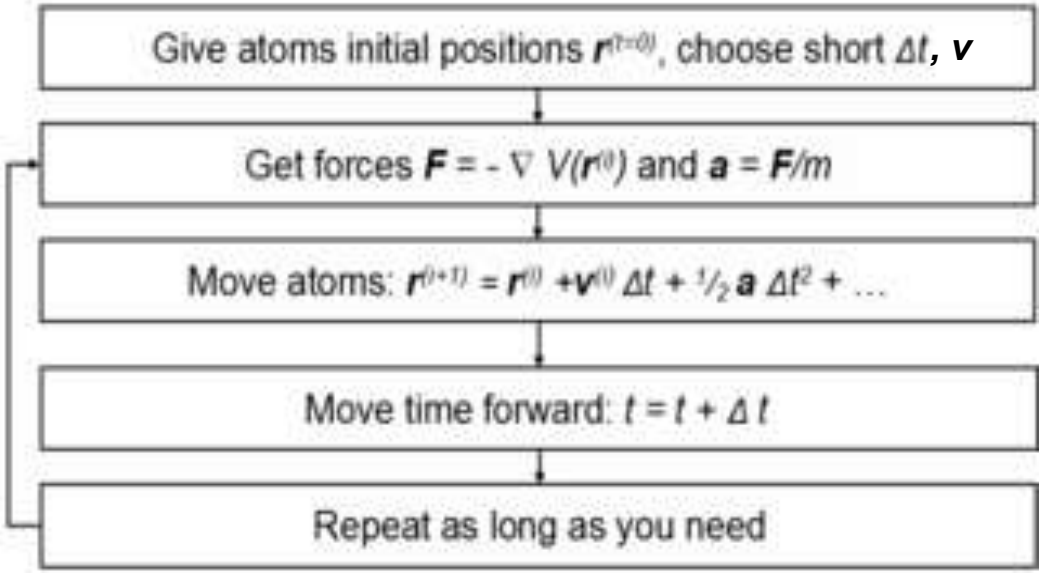
$$\mathbf{r}(t+\delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\delta t) + \mathbf{a}(t)\delta t^2 + \ldots$$

- Vecocities calculated from

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\delta t) - \mathbf{r}(t-\delta t)}{2\delta t}$$

# Molecular Dynamics Simulations

- Integration broken down to many small stages: **δt**
- The total force on each particle in the configuration at a time t is the vector sum of its interactions with other particles.
- From the force determine the acceleration of the particles and combine it with positions and velocities at time t to calculate at time t +δt
- The force is constant during the time step



# Στατιστική Μηχανική

- Σε μία Μοριακή Δυναμική Προσομοίωση διερευνάται η σχέση μεταξύ μικροσκοπικών και μακροσκοπικών ιδιοτήτων

- Η σχέση γίνεται μέσω της **στατιστικής μηχανικής**, η οποία μελετά τα συστήματα σε μοριακό επίπεδο

- Η κατανομή του συστήματος στο στατιστικό σύνολο ακολουθεί την κατανομή **Boltzmann**

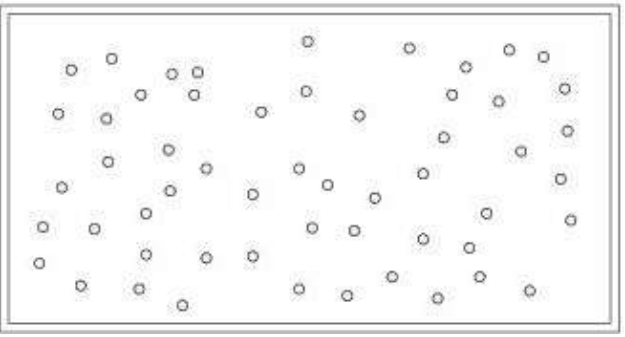
- **Θεμελιώδης έννοια - στατιστικό σύνολο:** το σύνολο όλων των πιθανών συστημάτων που έχουν διαφορετικές μικροσκοπικές καταστάσεις αλλά ίδια μακροσκοπική ή θερμοδυναμική κατάσταση

# Στατιστική Μηχανική & Χώρος Φάσεων

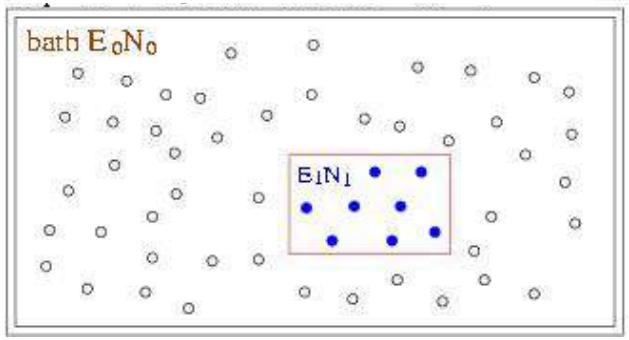
Ένα στατιστικό σύνολο είναι το σύνολο των μικροσκοπικών καταστάσεων για δεδομένη μακροσκοπική κατάσταση

Χρησιμοποιείται για να υπολογιστούν οι ιδιότητες του θερμοδυναμικού συστήματος από τους νόμους της Κλασικής ή της Κβαντικής Μηχανικής

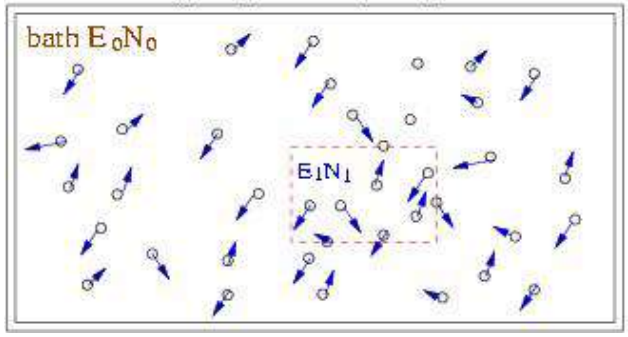
Μικροκανονικό, NVE

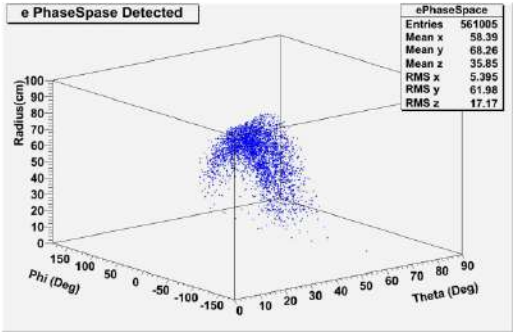


Κανονικό, NVT




Μεγαλοκανονικό, μVT




Είναι ένα σύνολο από αντιπροσωπευτικά σημεία στο χώρο των φάσεων διάστασης 6N

Όπου οι μικροκαταστάσεις κινούνται δημιουργώντας μία ***δυναμική τροχιά*** καθώς οι θέσεις και οι ορμές των ατόμων που εξελίσσονται στο χρόνο

# Η συνάρτηση καταμερισμού σε θερμική ισορροπία

Νόμος του Boltzmann $\frac{n_i}{n_j} = e^{-(\varepsilon_i - \varepsilon_j)/kT}$ $n_i$, $n_j$ πληθυσμοί ενεργειακών καταστάσεων

Για το χαμηλότερο ενεργειακό επίπεδο: $n_i = n_0 e^{-\beta\varepsilon_i}, \beta = 1/kT \quad n_0 = \frac{N}{\sum e^{-\beta\varepsilon_i}}$

q → Μέτρο για το πλήθος των ενεργειακών σταθμών που είναι διαθέσιμες σε συνθήκες θερμικής ισορροπίας

$$q = \sum e^{-\beta\varepsilon_i} \qquad n_i = \frac{N e^{-\beta\varepsilon_i}}{\sum e^{-\beta\varepsilon_i}}$$

**Συνάρτηση καταμερισμού ανά σωματίδιο**

Τα παραπάνω ισχύουν όταν οι ενεργειακές καταστάσεις είναι διακριτές (Κβαντική Στατιστική Μηχανική). Στην Κλασική Στατιστική Μηχανική η θέση και η ορμή μεταβάλλονται με συνεχή τρόπο οπότε οι μικροκαταστάσεις δεν μπορούν να μετρηθούν

$$\mathrm{Q} = \mathrm{C} \iint \mathrm{d}\mathbf{r}^{\mathrm{N}}\, \mathrm{d}\mathbf{p}^{\mathrm{N}} \exp\left(-\frac{\mathrm{H}(\mathbf{r}^{\mathrm{N}}, \mathbf{p}^{\mathrm{N}})}{\mathrm{kT}}\right)$$

# Θερμοδυναμικές συναρτήσεις

Μικροκανονική Συνάρτηση Καταμερισμού, N,V,E σταθερά

$$Q(N,V,E) = \frac{1}{h^{3N}N!}\sum(N,V,E) = \frac{1}{h^{3N}N!}\int d^{3N}\mathbf{q}\,d^{3N}\mathbf{p}$$

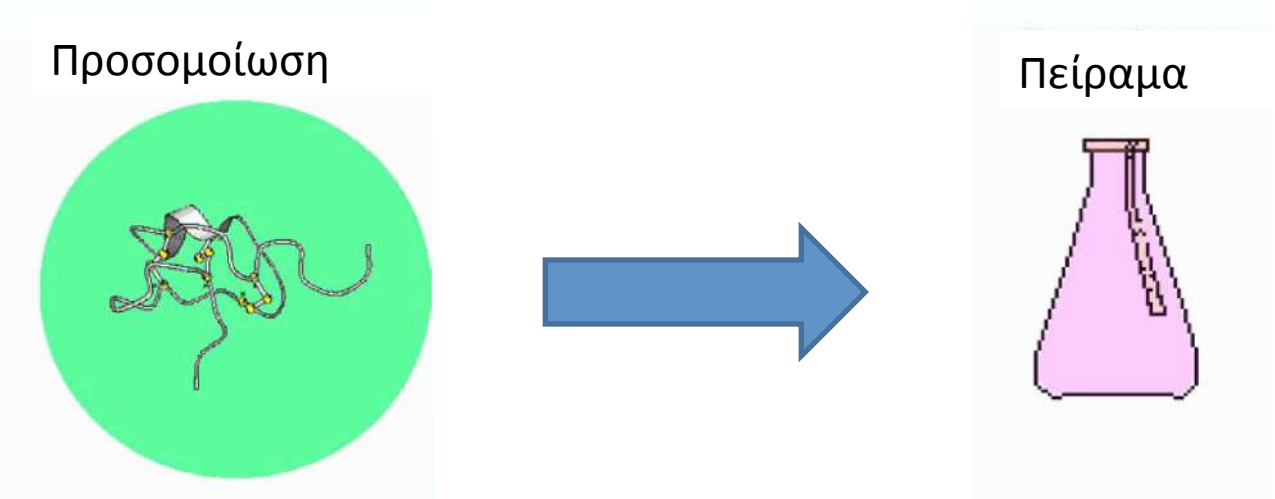
$$E = \frac{N}{q}\sum_i \varepsilon^i e^{-\beta\varepsilon^i} \qquad U - U_0 = -\left(\frac{\partial \ln Q}{\partial \beta}\right)_V \qquad S = \frac{U - U_0}{T} + k\ln Q$$

Σύνδεση με μακροσκοπική θερμοδυναμική: $S(N,V,E) = k_B \ln Q(N,V,E)$

Από την παραπάνω σχέση εξάγονται οι θερμοδυναμικές ιδιότητες του συστήματος

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_{N,V} \Rightarrow \frac{1}{T} = k_B\left(\frac{\partial \ln Q}{\partial E}\right)_{N,V}$$

# Στατιστική Μηχανική




Μικροσκοπική περιγραφή

Μακροσκοπική περιγραφή

***Κβαντική Μηχανική***:
Ιδιοτιμές Ε και ιδιοσυναρτήσεις
$\Psi(r_1,r_2,\ldots,r_N)$ από εξίσωση Schrodinger

***Μοριακή Μηχανική***:
Κινητική και Δυναμική ενέργεια Ε(**r,v**)

***Θερμοδυναμική***:
Σχέσεις του συστήματος σε θερμοδυναμική ισορροπία ή εκτός ισορροπίας

Χρήση στατιστικής Μηχανικής για να περιγράψουμε τις θερμοδυναμικές ιδιότητες

# Μοριακή Δυναμική

**Περιγραφή της κίνησης:**

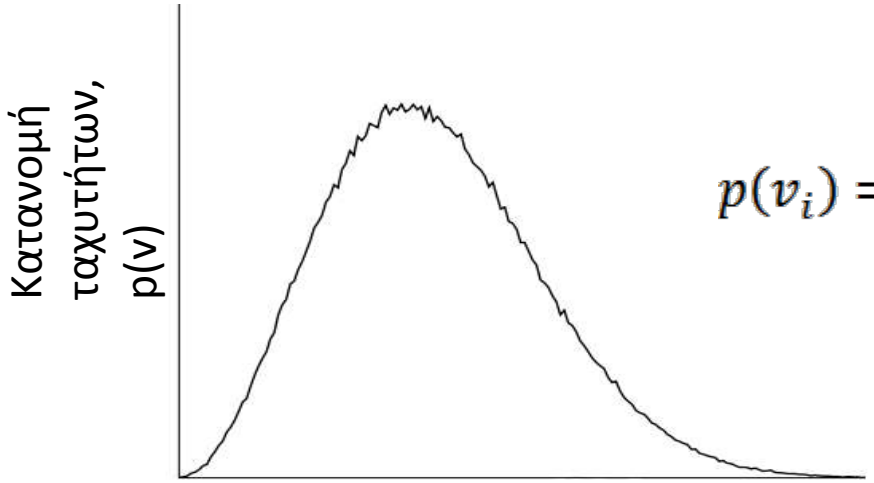
- Η Κλασική Μηχανική περιγράφει την κίνηση των ατόμων → 2ος νόμος του Νεύτωνα:

$$\boldsymbol{F_i} = m_i \boldsymbol{a_i}$$

- Ολική ενέργεια του συστήματος - Χαμιλτωνειανή

$$H = T(|\boldsymbol{v}|) + V(|\boldsymbol{r}|) = \frac{1}{2} m|\boldsymbol{v}|^2 + V(|\boldsymbol{r}|)$$

- Αρχικές ταχύτητες γνωστές από κατανομή Maxwell




$$p(v_i) = \sqrt{\frac{m_i}{2\pi k_B T}} e^{\left(-\frac{m_i v_i^2}{2k_B T}\right)}$$

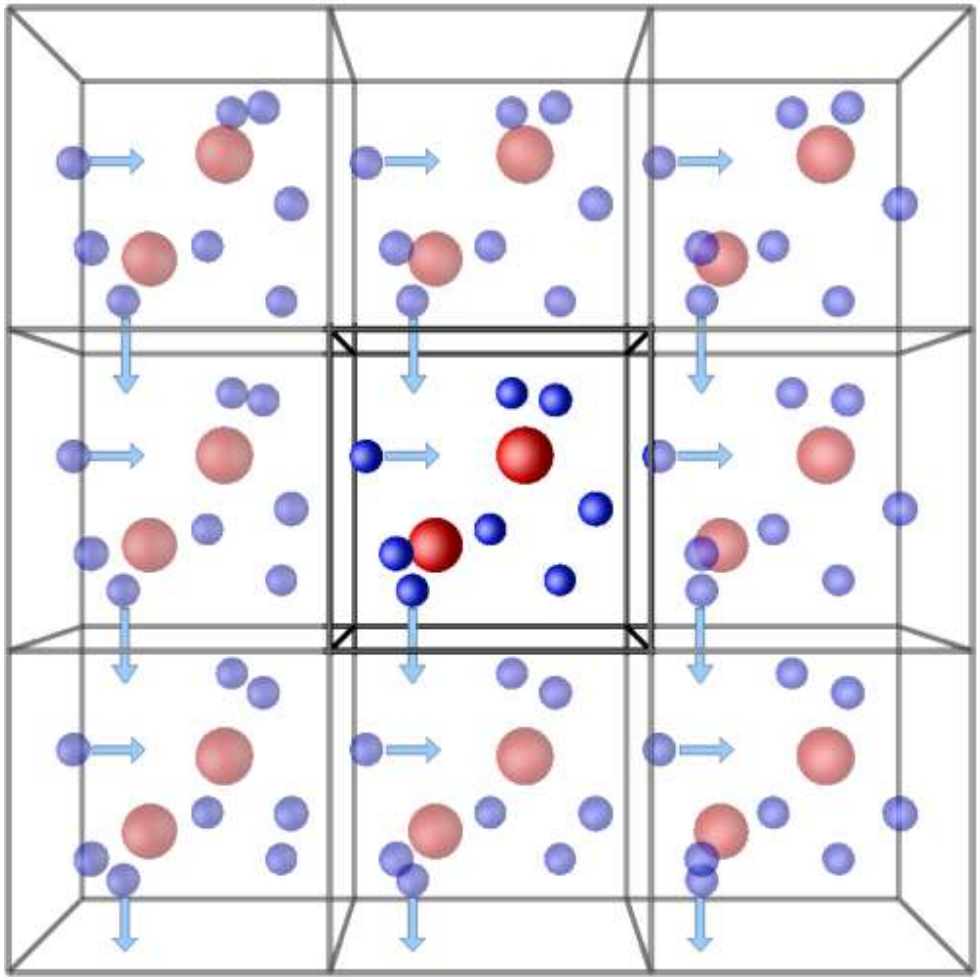
# Περιοδικές οριακές συνθήκες

Κουτί προσομοίωσης → Προσομοιώνεται ένα τμήμα του πραγματικού συστήματος → Απλούστερος υπολογισμός

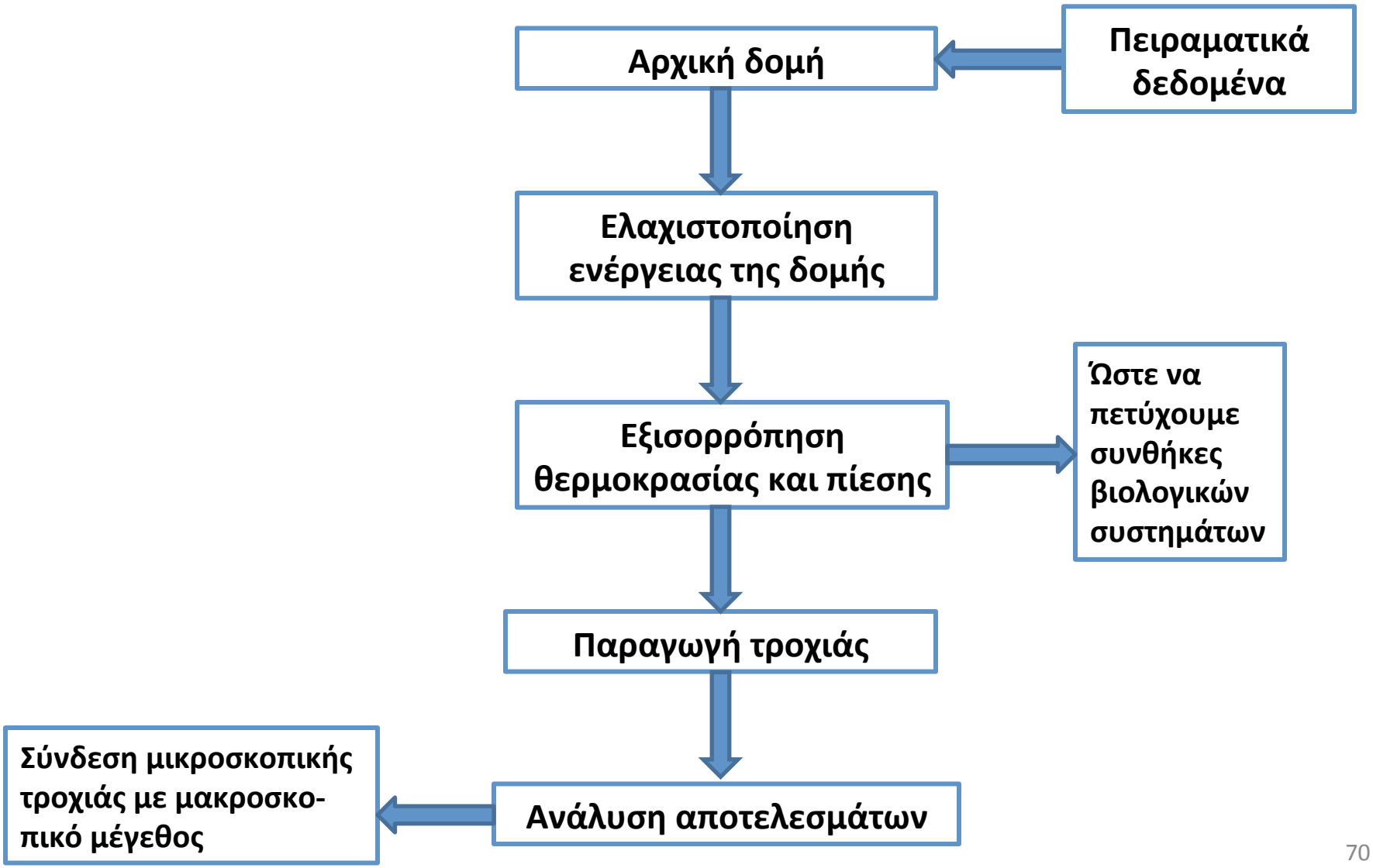
❑ Όταν άτομο φτάσει στο όριο → Επανεμφανίζεται από την αντίθετη πλευρά

❑ Αναπαράσταση της συνεχούς συμπεριφοράς του υγρού – διαλύτη

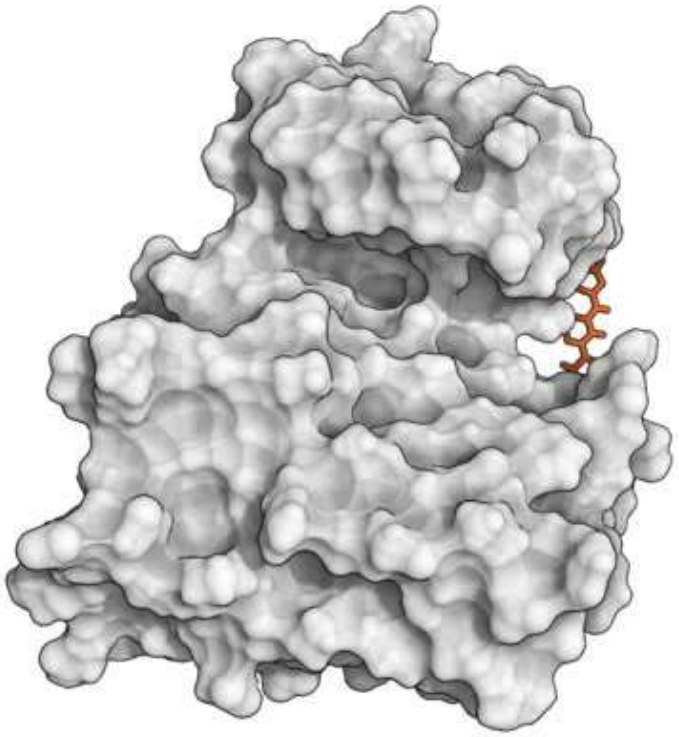
❑ Αποφυγή επιφανειακών φαινομένων



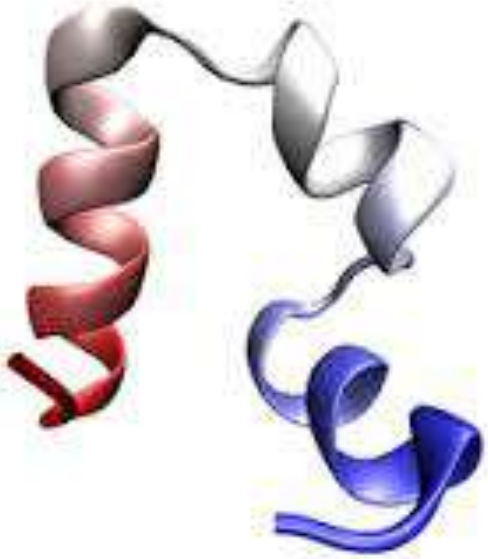
# Πώς λειτουργεί ένα πρόγραμμα Μοριακής Δυναμικής;



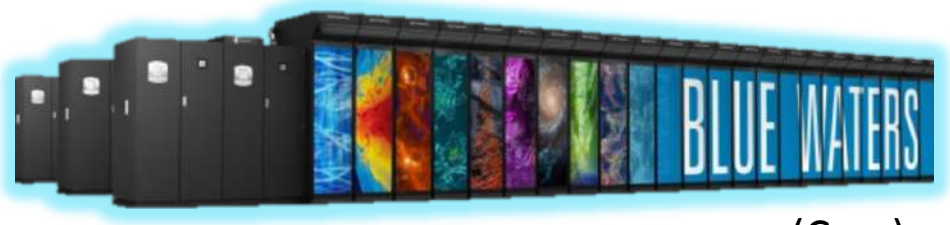
# Examples of MD simulations of proteins



Shan et al (2011)
**Cancer drug dasatinib binding on Src kinase**



Schulten et al (2012)
**Folding of the Villin Headpiece protein**


(Cray)

Academy of Athens

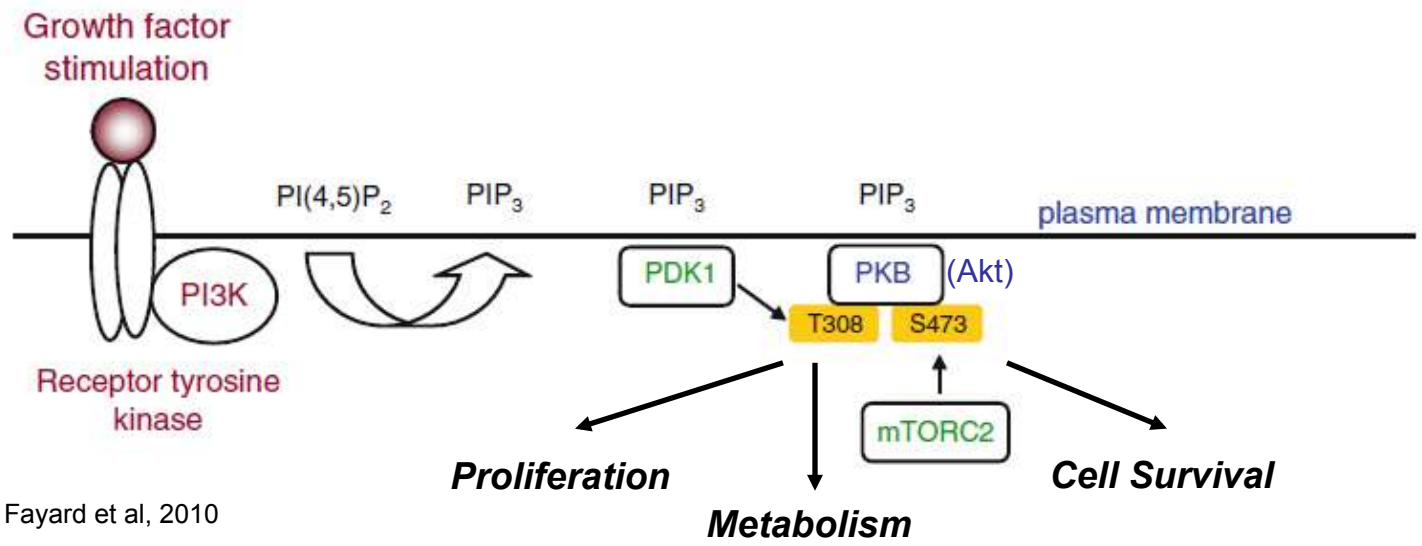
Biomedical Research Foundation

# Computer-aided Drug Design: *Targeting the mutant PI3Kα*

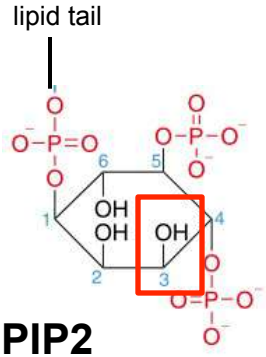
Zoe Cournia
zcournia@bioacademy.gr

**12 December 2014**

# PI3Kα is a lipid kinase that promotes cell survival




Fayard et al, 2010



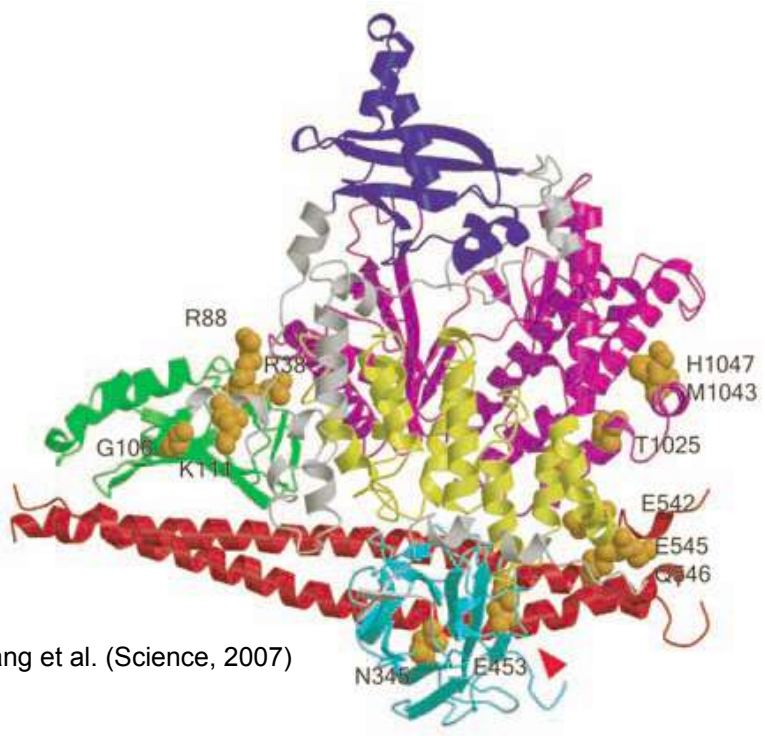

- Active PI3Ka phosphorylates PIP2 to PIP3 at the plasma membrane.

- PIP3 recruits Akt close to PDK1.

- Co-localization of these proteins leads to phosphorylation of residues, which in turn leads to proliferation, growth, survival.

# PI3Kα: most commonly mutated kinase in cancer

- PI3Ka is a membrane-associated lipid kinase
- Involved in cell growth, proliferation, differentiation
- Most commonly mutated kinase in the human genome ⇒ cancer



Huang et al. (Science, 2007)

80% of all mutations:

Glu545Lys

His1047Arg

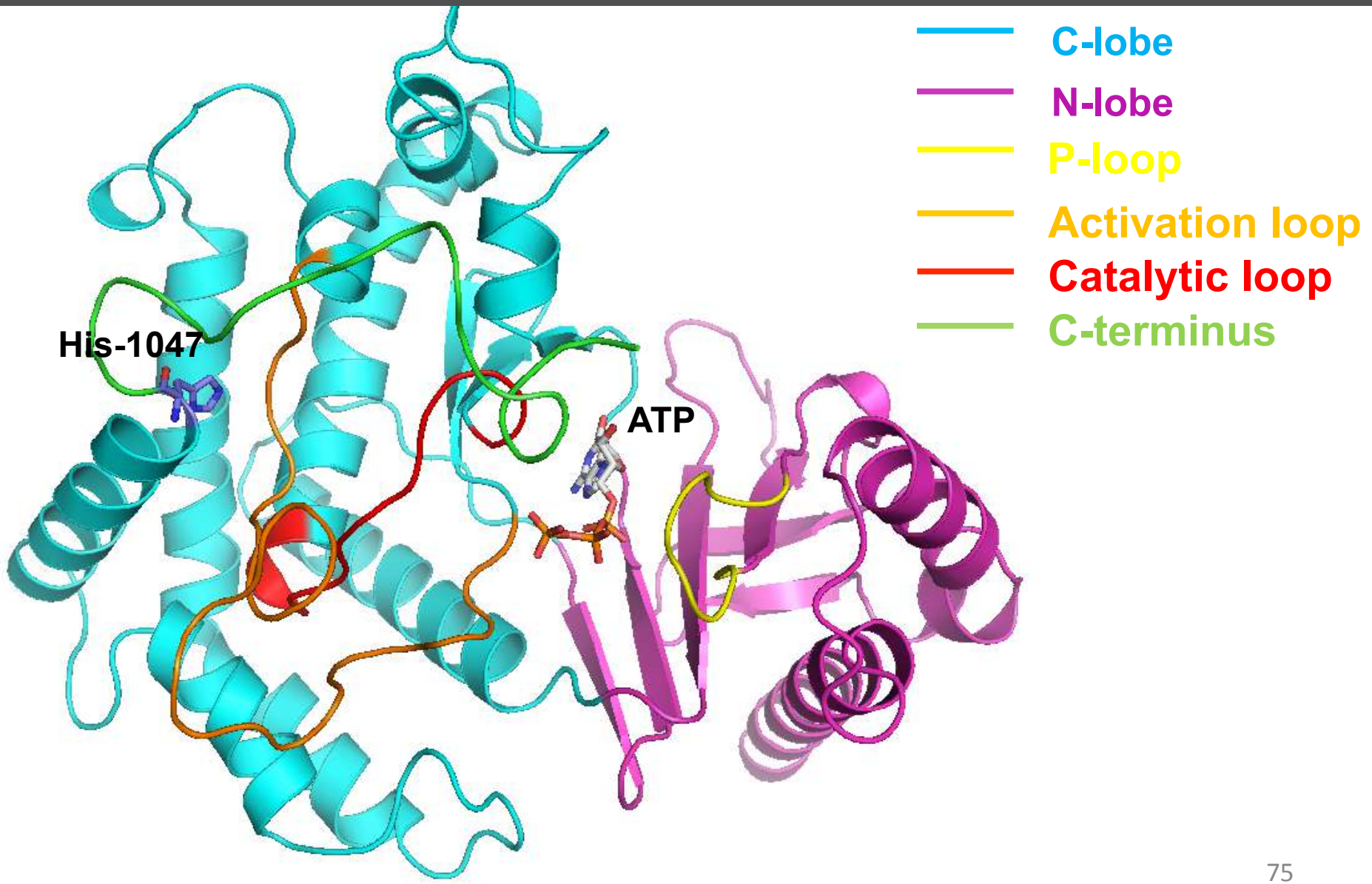
30% of breast cancer patients

**Mechanism of overactivation?**

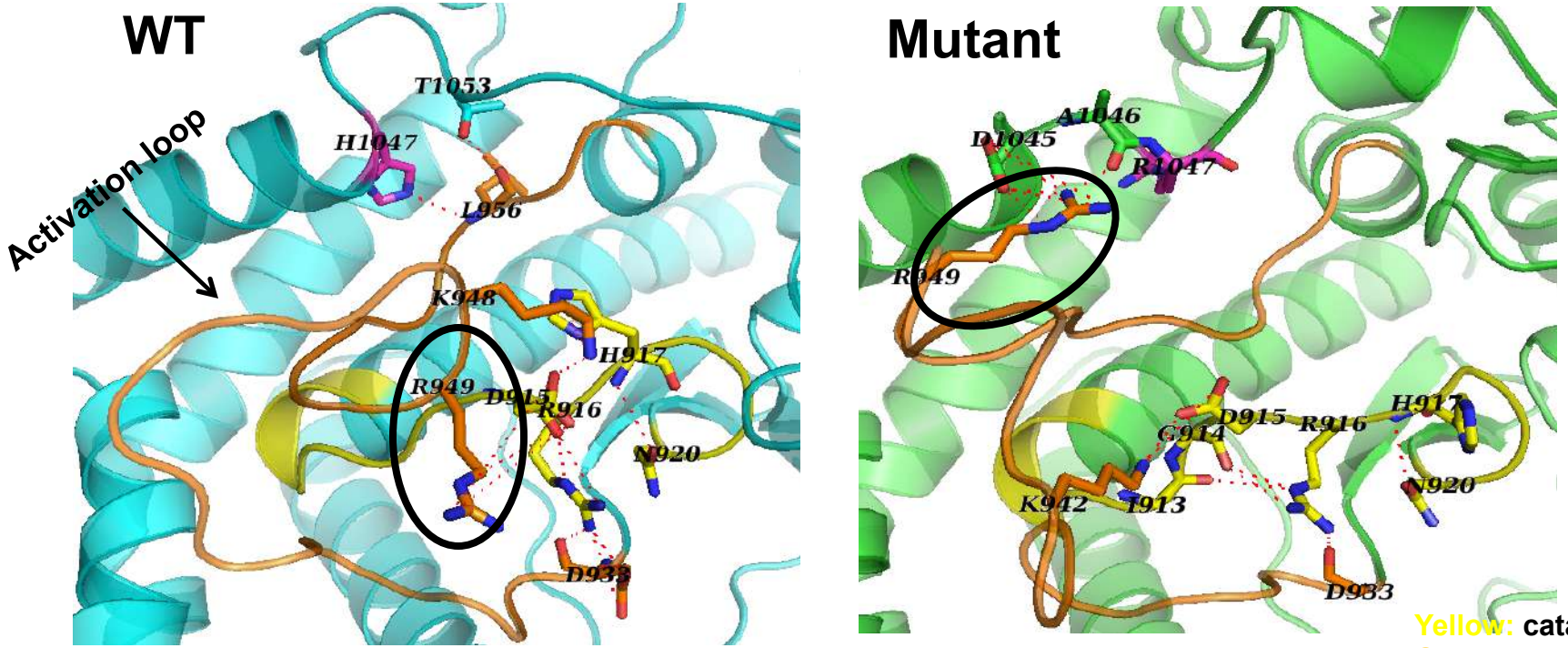
**Mutant and isoform specific therapies?**

**MD Simulations**
**Virtual screening**
**Property prediction**
***In vitro* & *In vivo* assays**
**Lead Optimization**

# Kinase Domain Organization



# Hydrogen Bond Analysis

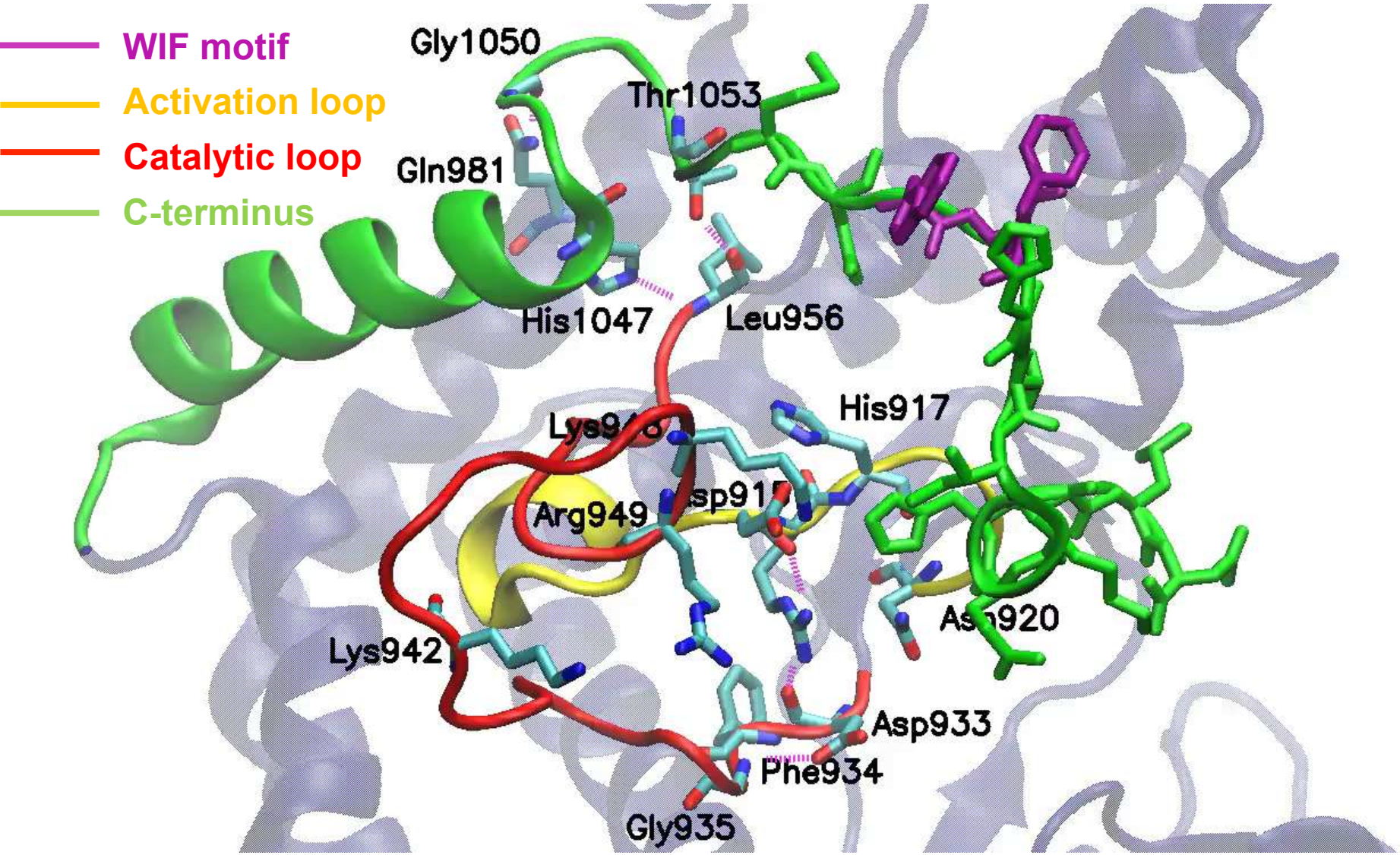



Yellow: catalytic loop
Orange: activation loop
Magenta: residue 1047

- The Hbond between activation loop Leu956 and His1047 breaks
- The α-helix of H1047 partially unfolds in the presence of 1047R
- Displacement of Arg949 creates a different Hbond network in the mutant, which changes the activation and catalytic loop positions

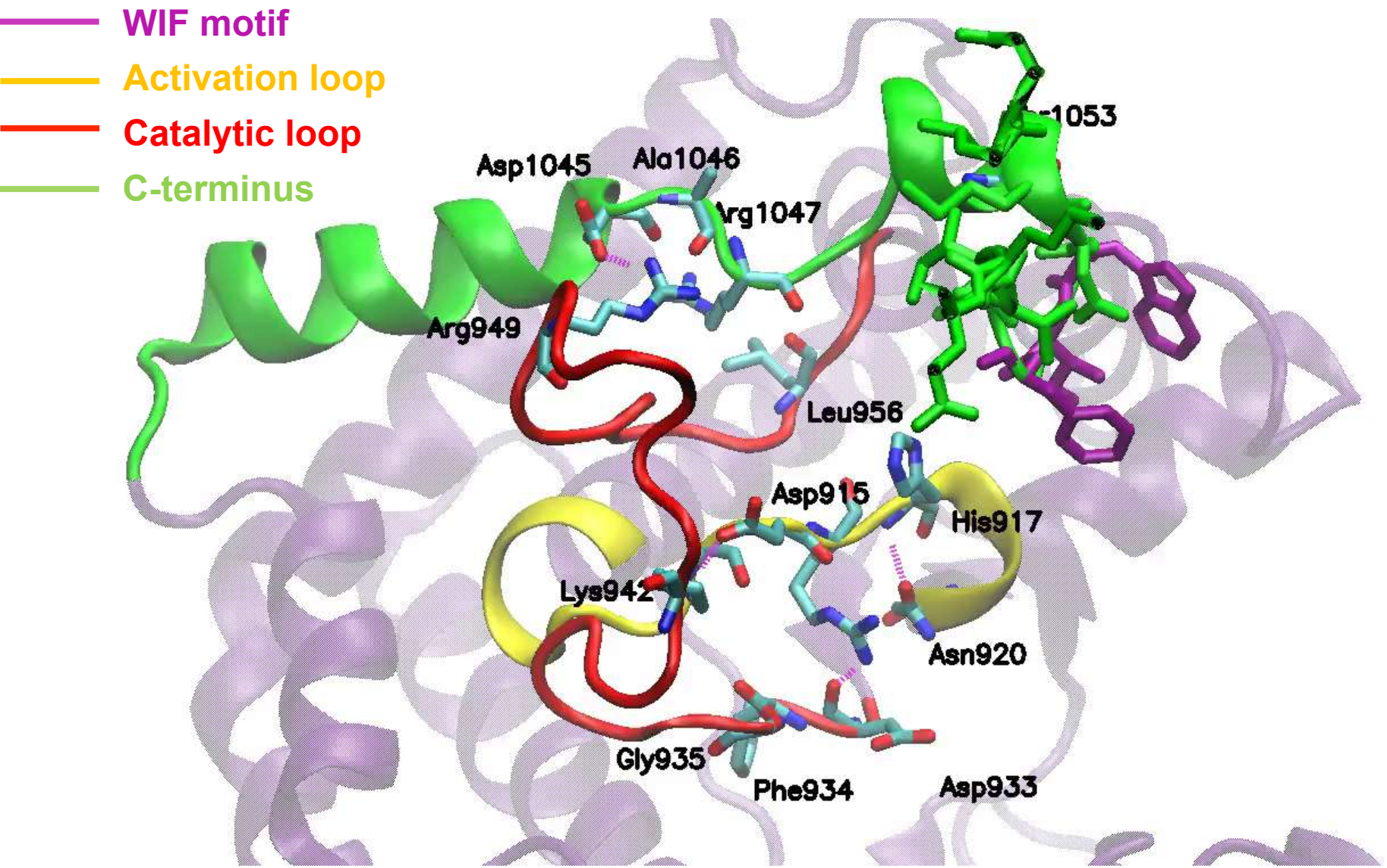
**H917, RESPONSIBLE FOR ATP HYDROLYSIS, IS ORIENTED TOWARD THE CATALYTIC POCKET IN THE MUTANT AND AWAY FROM THE POCKET IN THE WT**

# WT H-bond network



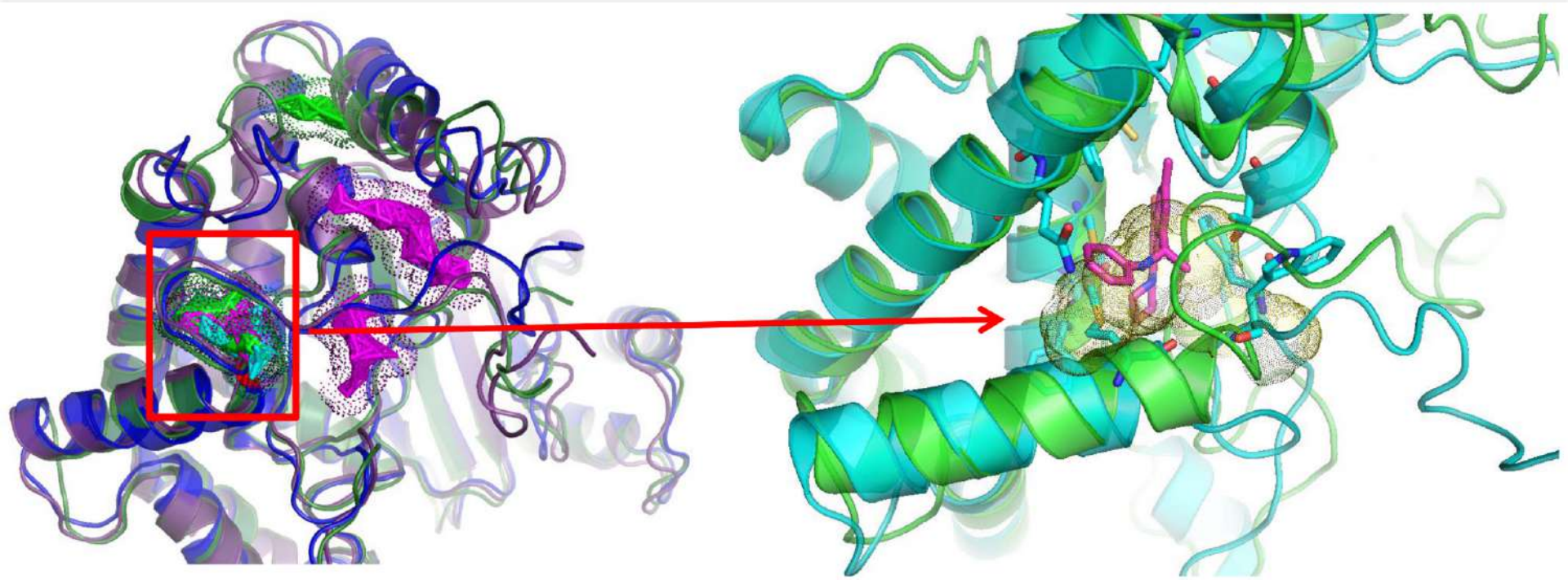

**His-917 points away from the active site, while the C-terminus prevents the catalytic loop from reaching the ATP-binding site.**

# Mutant H-bond Network is altered




**His-917 points towards the active site, while the C-terminus does not interfere with the access of the catalytic loop to the ATP-binding site.**

# Binding site identification on PI3Kα conformers



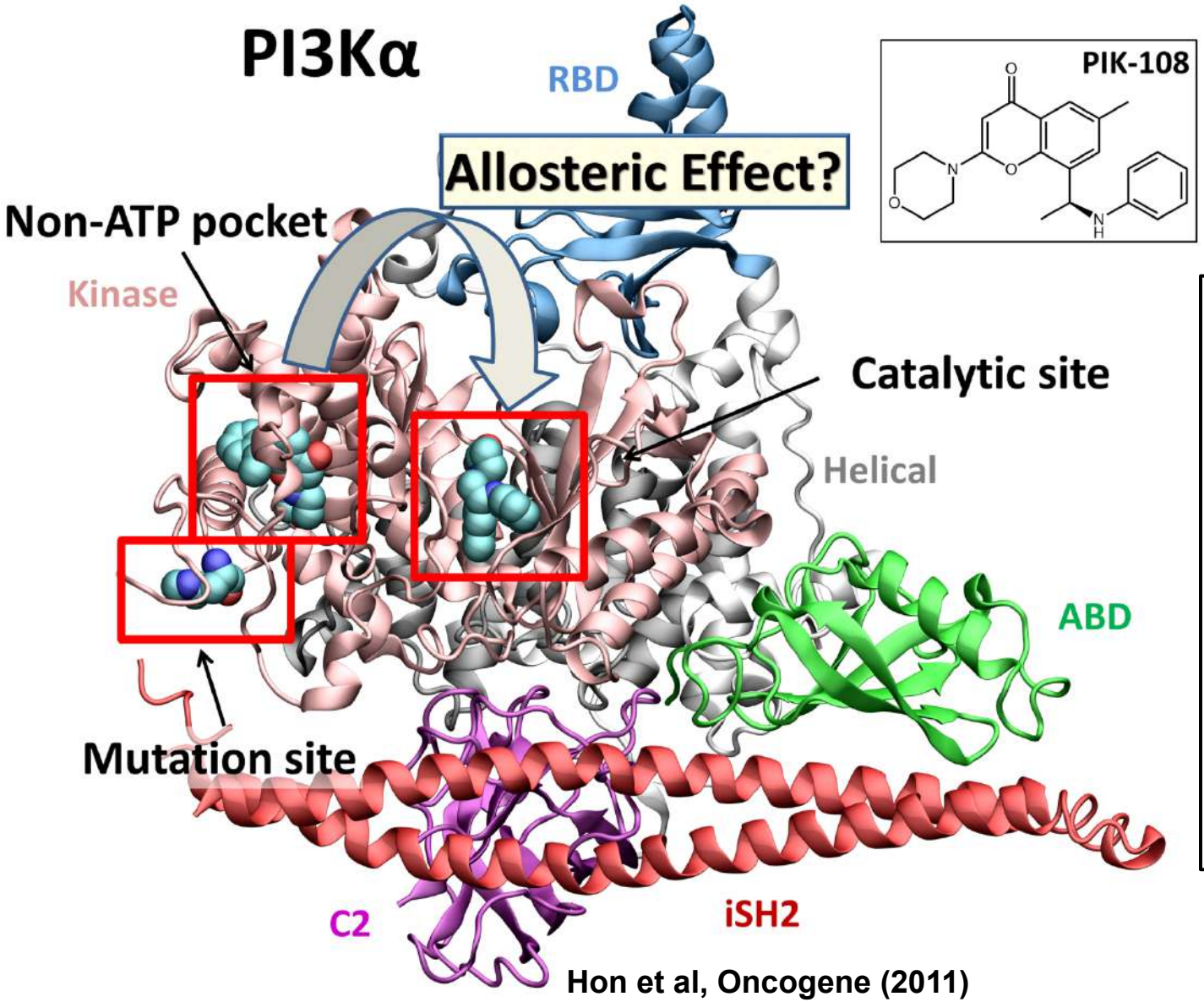
**Binding site prediction on PI3Kα representative structures**

**Blue: WT Crystal Structure by Hon et al (2011)**

**Green: Cluster conformation from MD**
**Dots: Predicted binding site**

**Does this binding site also exist in the mutant form and can it be exploited for selective drug design?**

# Non-ATP PI3Kα binding pocket discovered




Hon et al, Oncogene (2011)

- Active site and non-ATP pocket occupied by **PIK-108**
- MD simulations of WT, H1047R apo and holo forms (100ns production run)
- Is the non-ATP pocket allosteric?

# Allosteric or Noncompetitive Inhibition

**The inhibitor binds itself to a site other than the active site (allosterism), thereby changing the conformation of the active site. The substrate still binds but there is no catalysis.**

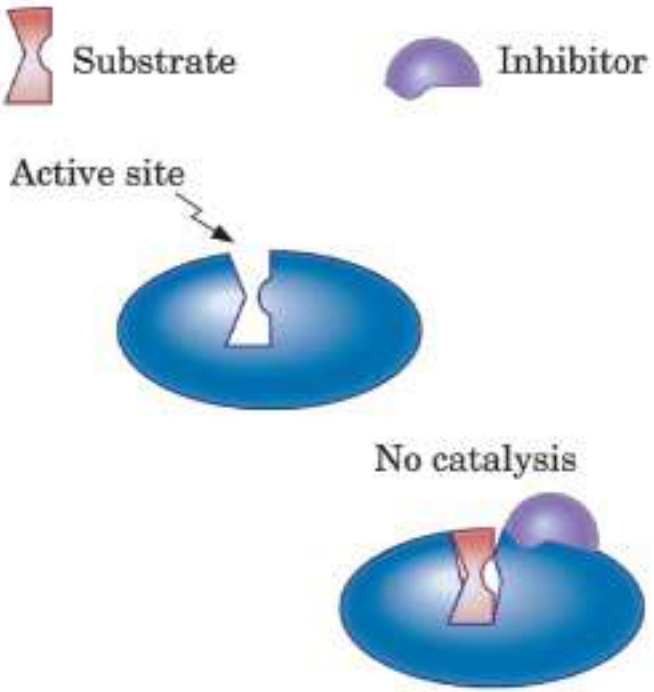
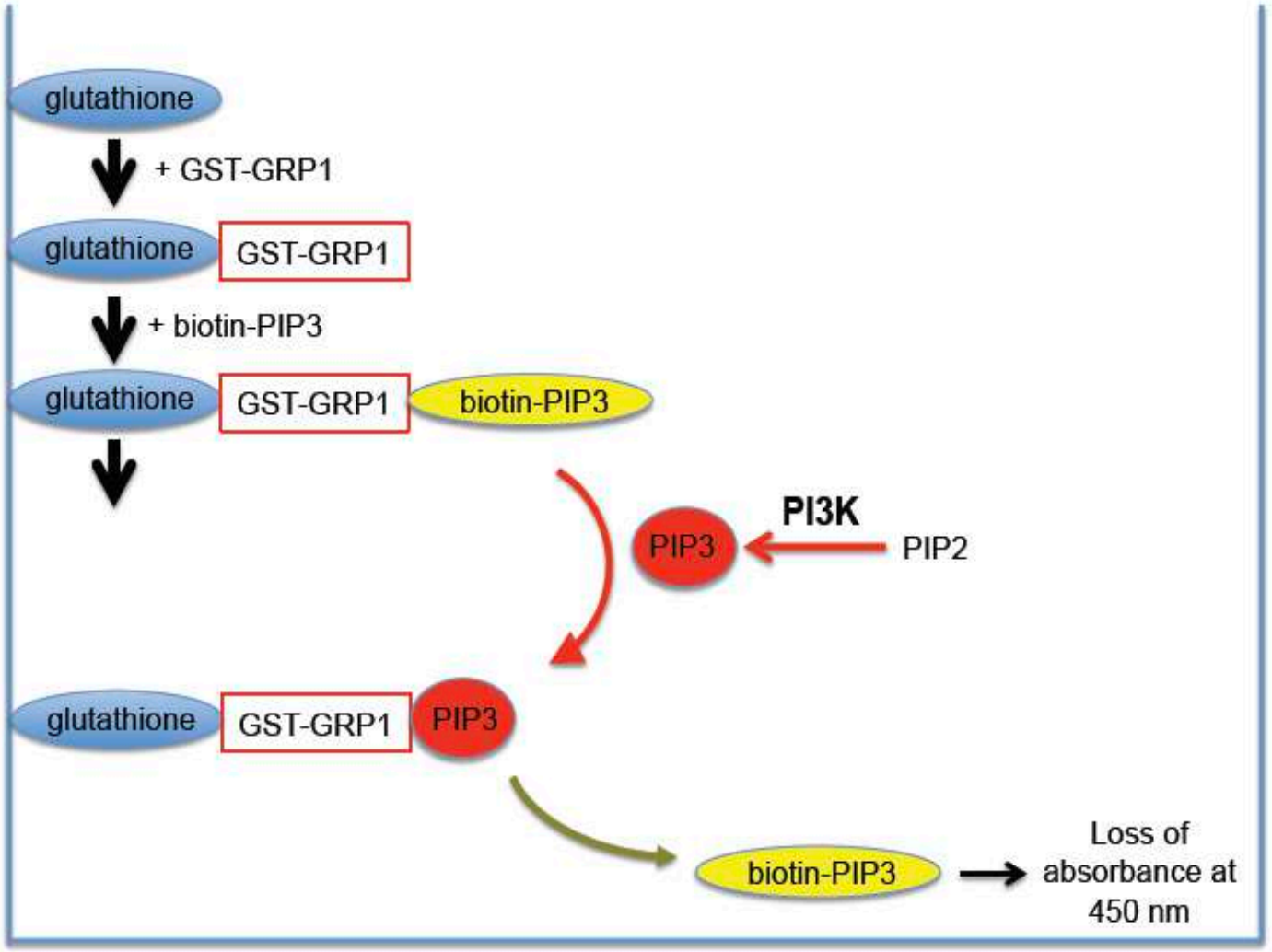



Image credit: K. Dunlap

# In vitro cell-free assay with cancer liposomes




Christoforidis lab, University of Ioannina, in vitro assays
Couladouros lab, University of Athens, synthesis of PIK-108

# Binding site Prediction

Crystal structure

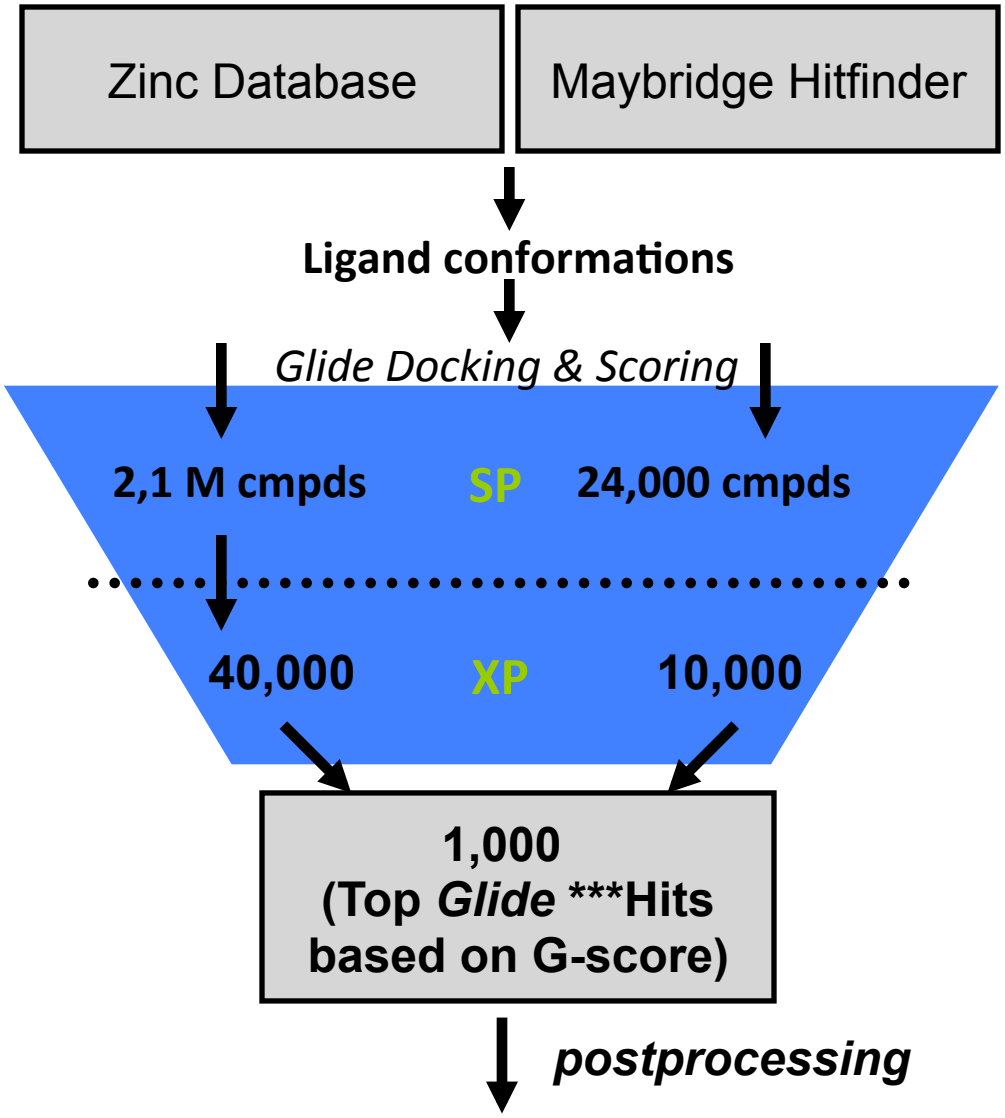
Molecular Dynamics

Cluster conformations from trajectory (*RMSD*)

Binding site prediction and *confirmation for allosteric pockets* **(Q-Site Finder, PCA)**

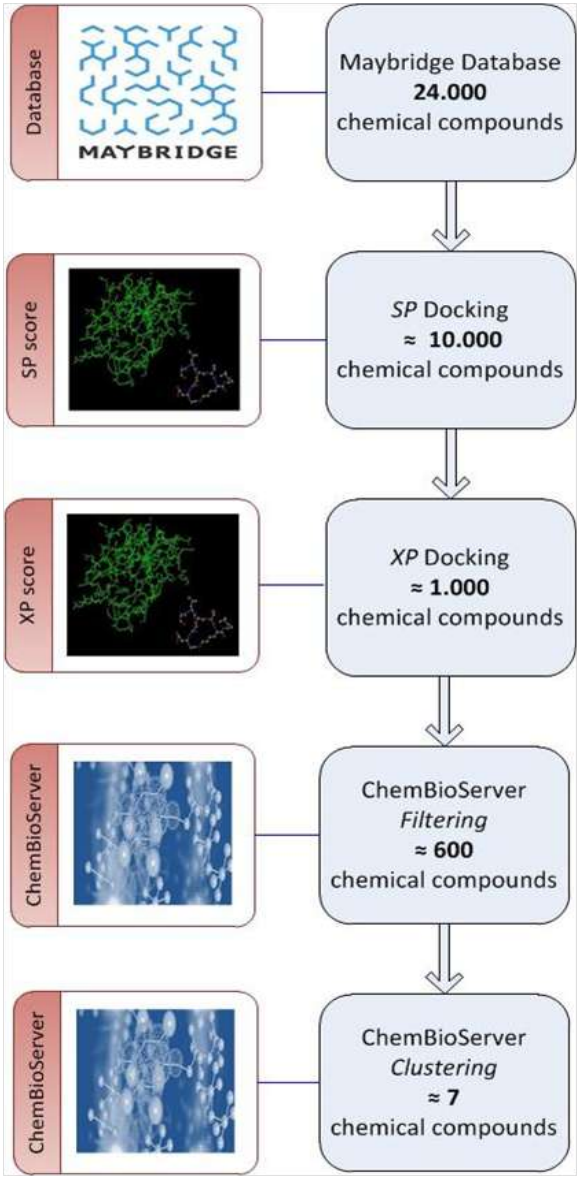
Lionta et al, *Curr Top Med Chem* (2014)

# Virtual Screening




**30 compounds purchased and assayed in vitro**

# How are compounds selected for assaying?



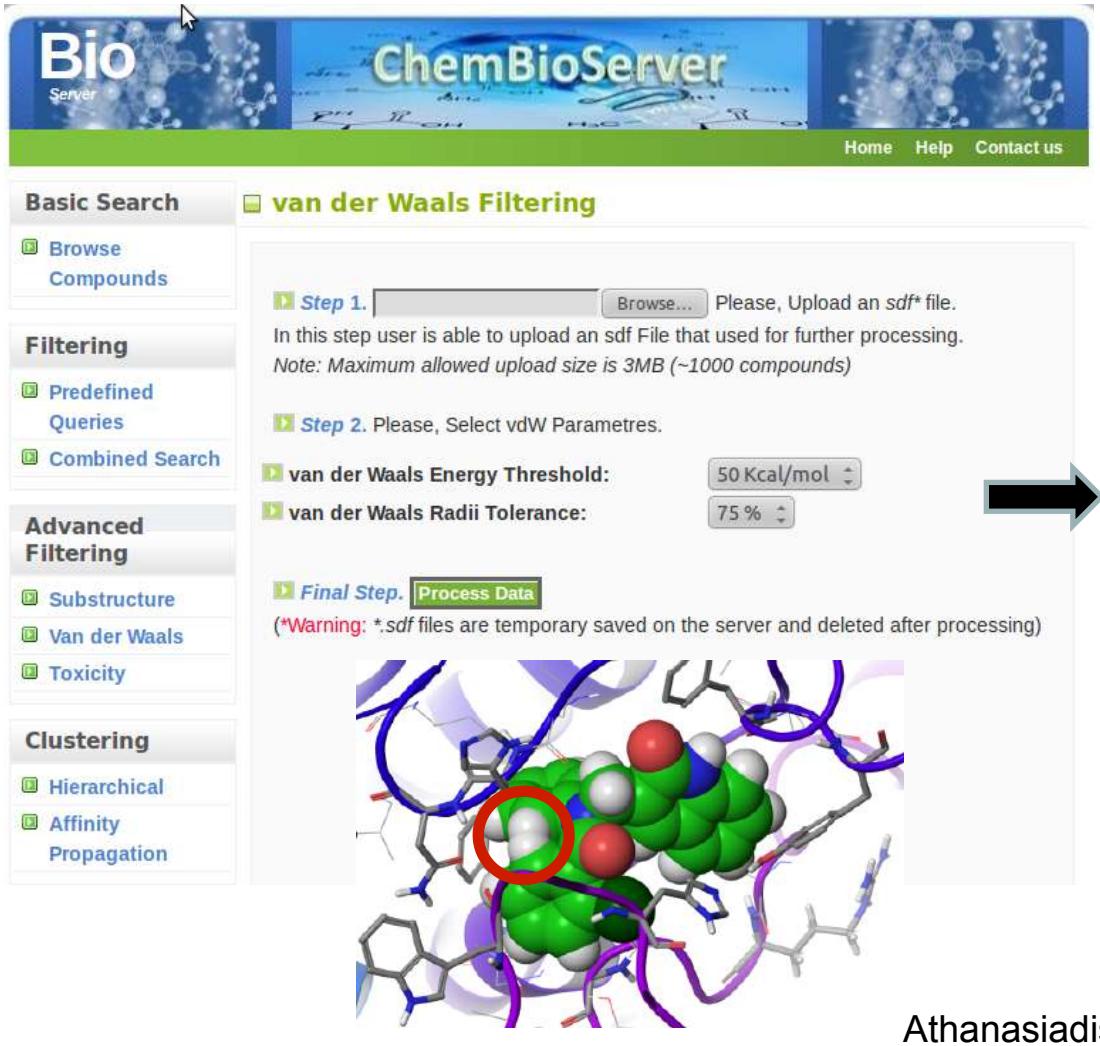

- **Library docking using Glide SP, XP**
- **1000 Top-scored XP compounds**
- **Postprocessing with ChemBioServer**
- **Calculate ADME/tox properties**
- **Check for bad vdW contacts**
- **Hierarchical Clustering**
- **Affinity Propagation (exemplars)**
- **Visualization: check for compound conformations**

**http://chembioserver.vi-seem.eu**

Athanasiadis, Cournia, Spyrou, *Bioinformatics* (2012)

# Pre/Postprocessing with ChemBioServer

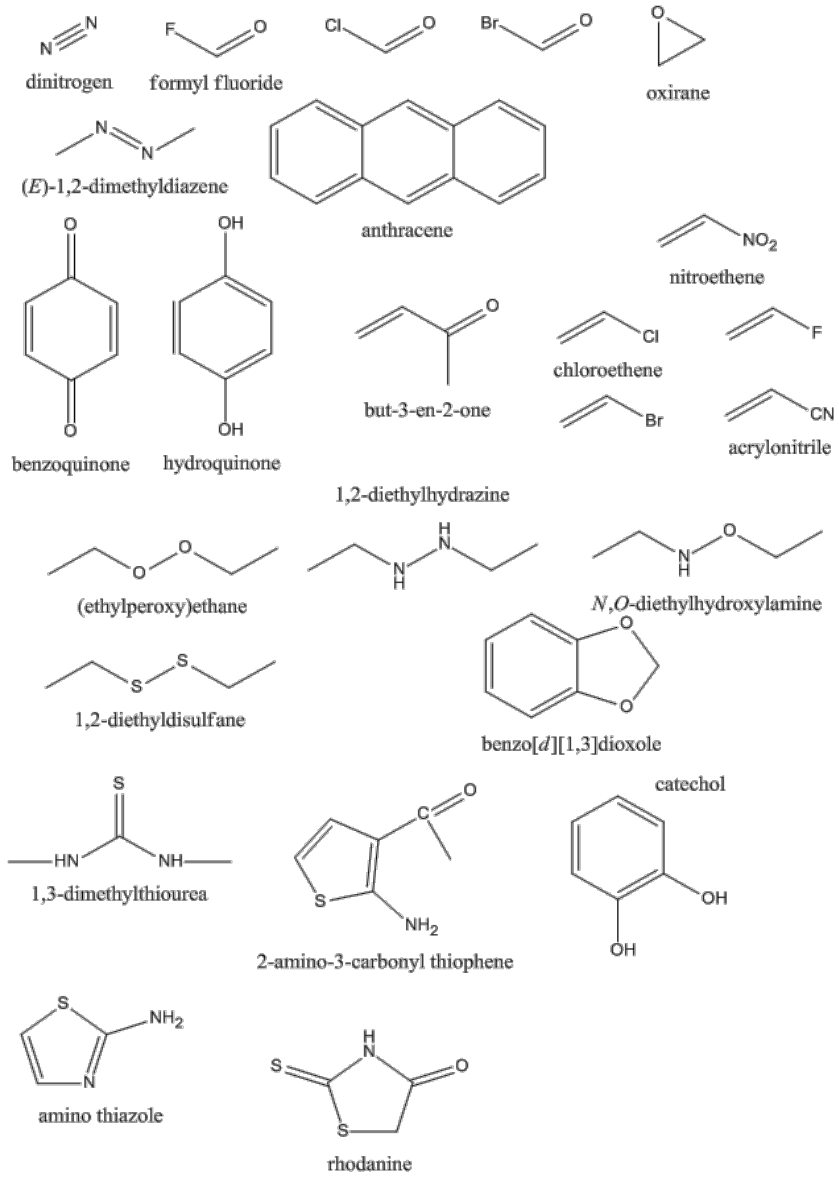
ChemBioServer post-processes virtual screening results



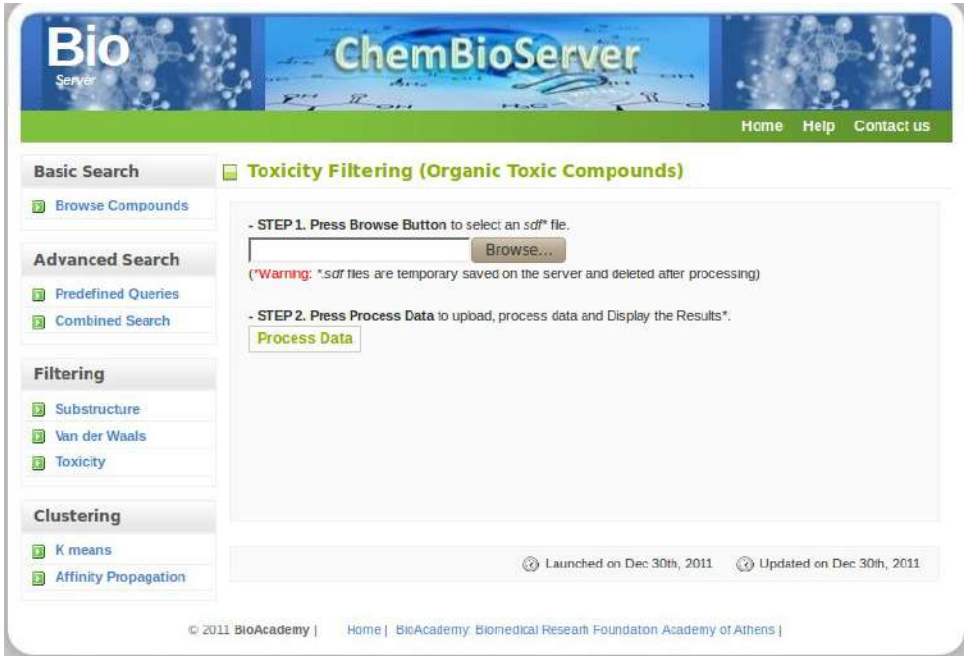

| Compound ID | VDW Energy Test | VDW Distance Test |
|---|---|---|
| Compound: 1 AW 00785 | - PASS AW 00785 - Browse List For Details... | - FAIL AW 00785 - Browse List For Details... |
| Compound: 2 AW 00788 | - PASS AW 00788 - Browse List For Details... | - FAIL AW 00788 - Browse List For Details... |
| Compound: 3 AW 00785 | - PASS AW 00785 - Browse List For Details... | - FAIL AW 00785 - Browse List For Details... |
| Compound: 4 AW 00939 | - PASS AW 00939 - Browse List For Details... | - FAIL AW 00939 - Browse List For Details... |
| Compound: 5 AW 00694 | - PASS AW 00694 - Browse List For Details... | - FAIL AW 00694 - Browse List For Details... |
| Compound: 6 CD 10205 | - PASS CD 10205 - Browse List For Details... | - PASS CD 10205 - Browse List For Details... |
| Compound: 7 GK 02096 | - PASS GK 02096 - Browse List For Details... | - FAIL GK 02096 - Browse List For Details... |
| Compound: 8 HTS 01561 | - PASS HTS 01561 - Browse List For Details... | - FAIL HTS 01561 - Browse List For Details... |
| Compound: 9 MWP 00404 | - PASS MWP 00404 - Browse List For Details... | - FAIL MWP 00404 - Browse List For Details... |
| Compound: 10 NRB 02577 | - PASS NRB 02577 - Browse List For Details... | - FAIL NRB 02577 - Browse List For Details... |

Athanasiadis, Cournia, Spyrou, Bioinformatics (2012)

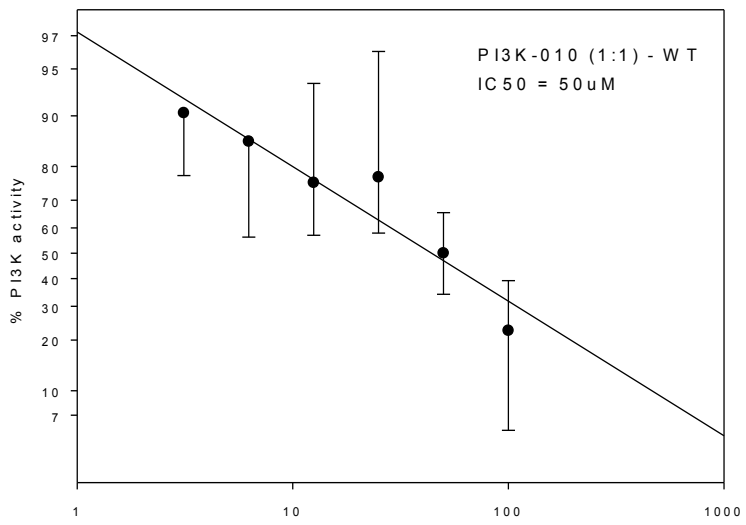
# Pre/Postprocessing with ChemBioServer

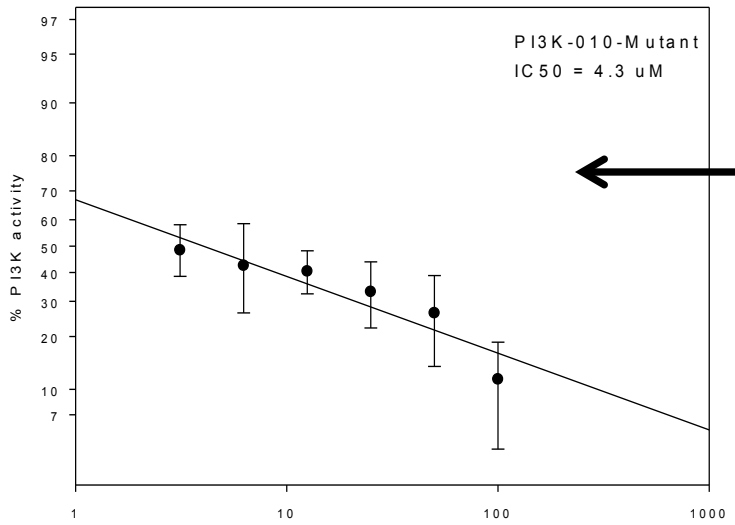



**http://chembioserver.vi-seem.eu**

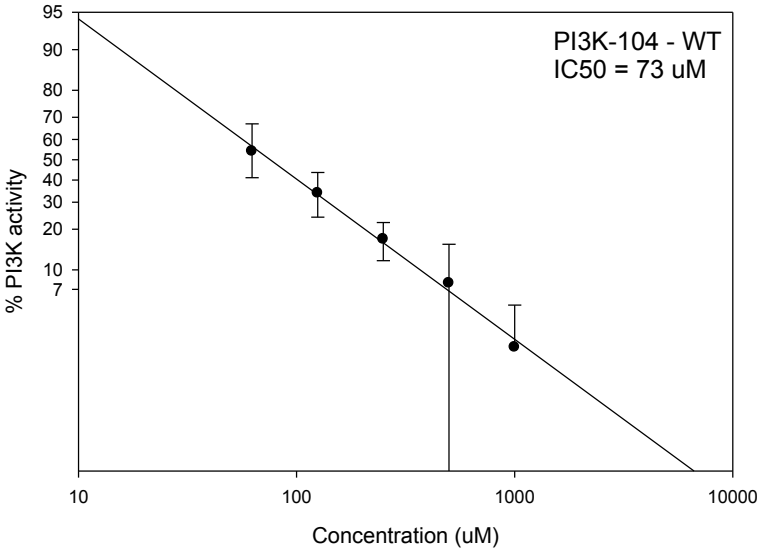


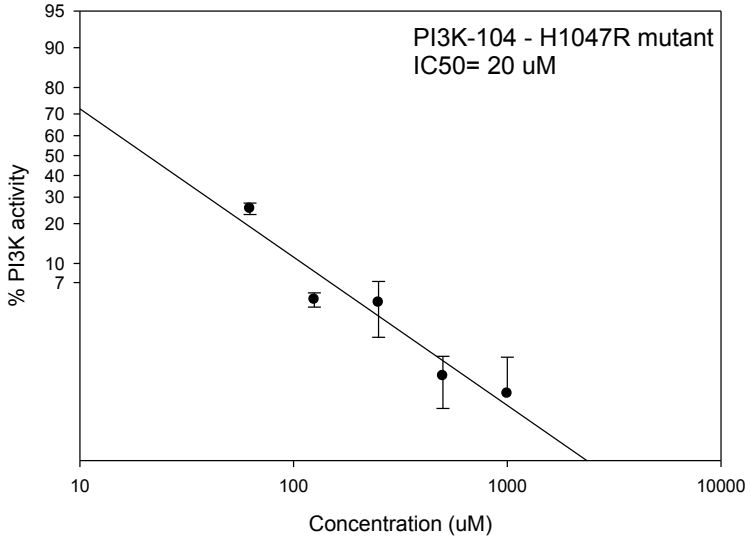
# In vitro cell-free assay with cancer liposomes



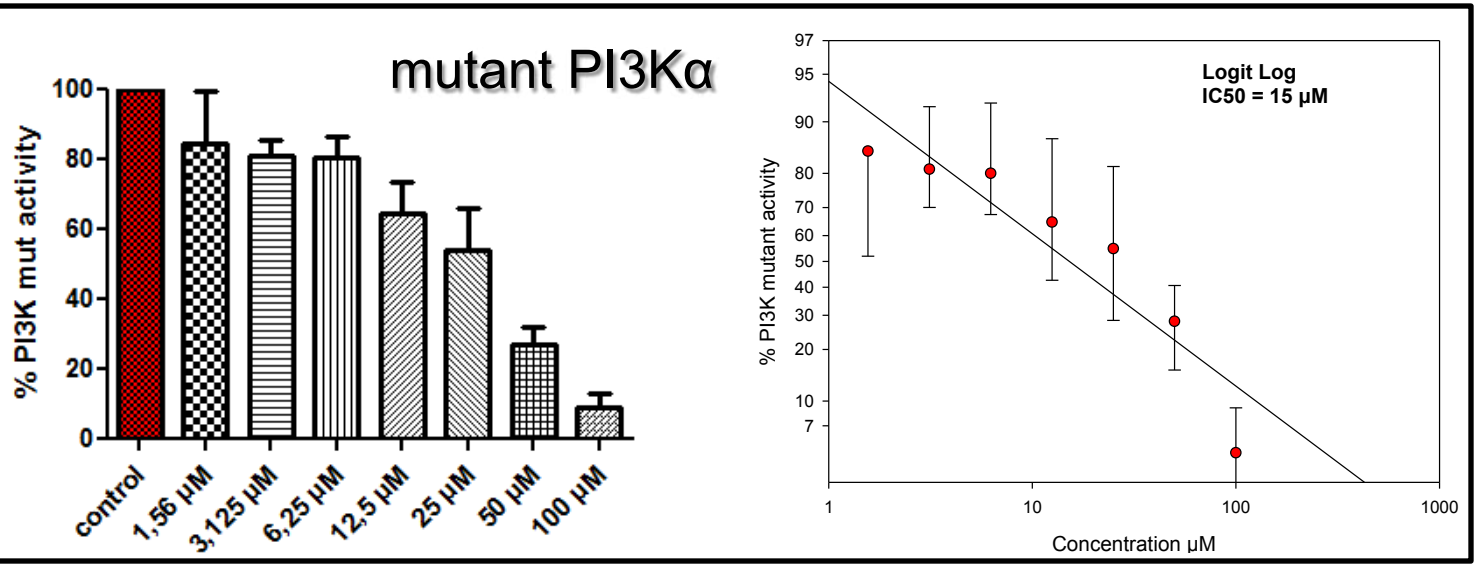

**11-fold selectivity of the mutant vs the WT**



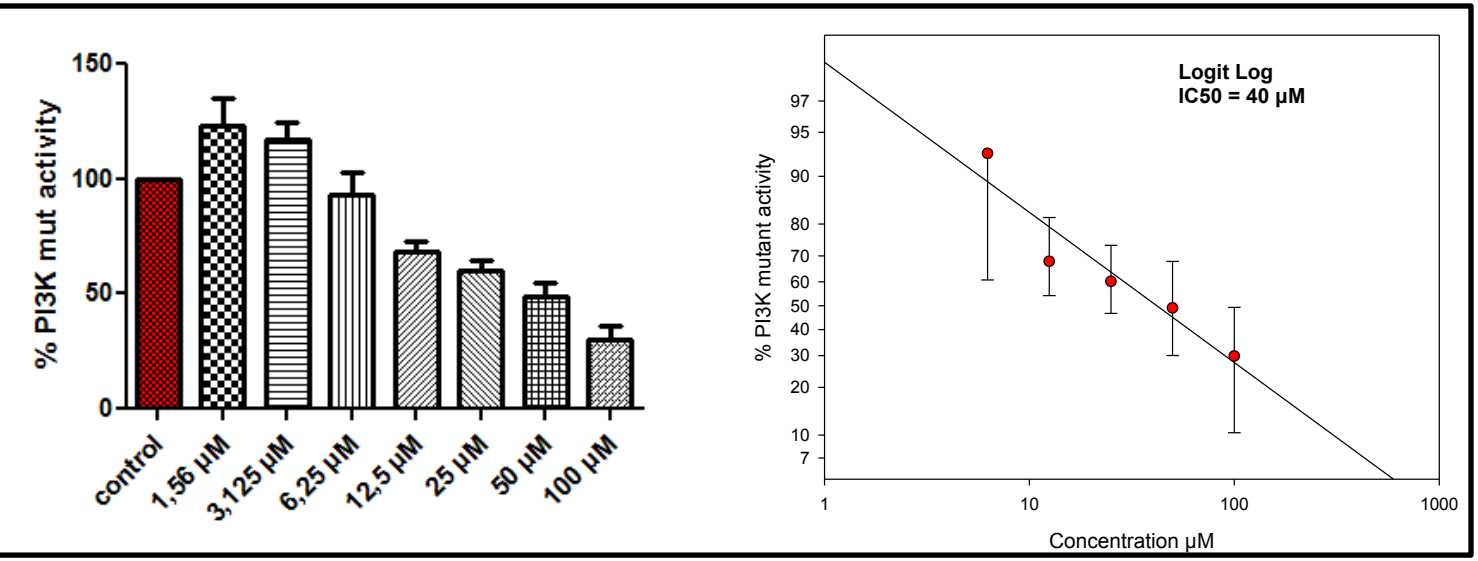

IC50 = the concentration of the compound required to inhibit the protein by 50%

# Is PI3K-010 an allosteric (non-competitive) inhibitor?




**Low ATP (100µM):**
**IC50 = 15 µM**

**High ATP (2mM):**
**IC50 = 40 µM**

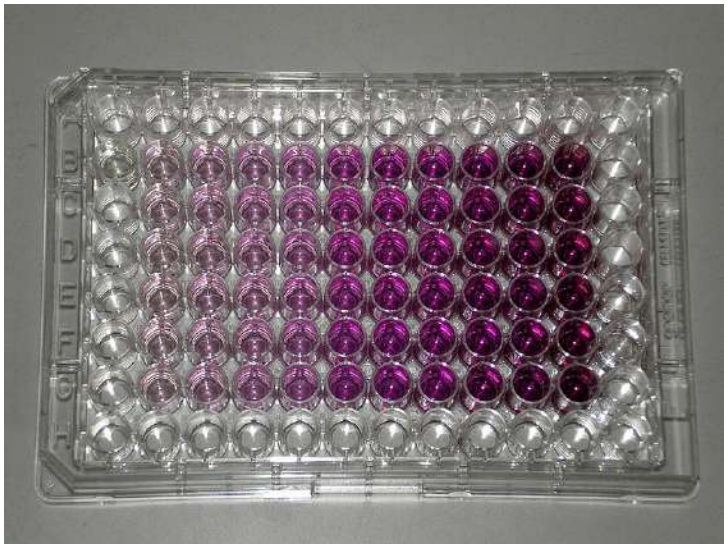



**PI3K-010 IC50 is not influenced by ATP concentration**

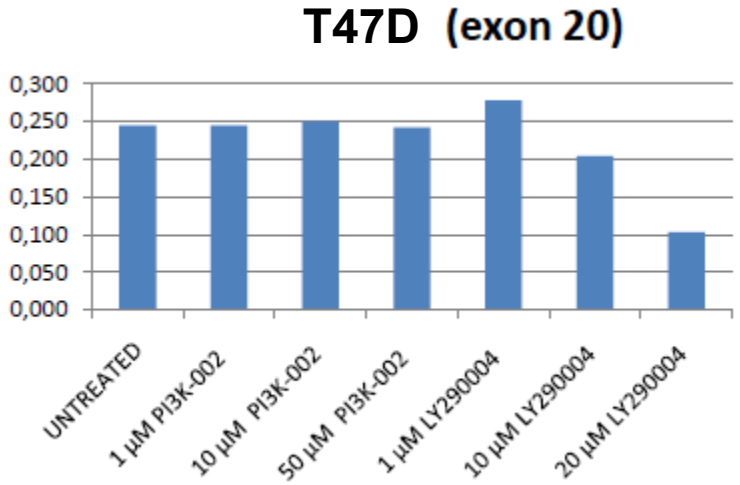
**Could be considered allosteric**

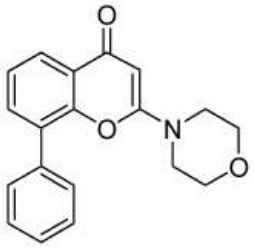
**2 experiments low ATP, 4 experiments high ATP**

# Cell-based MTT assay



- Cell viability assay
- MTT (a yellow tetrazole) is reduced to purple formazan in living cells
- Initially 96-well plates, now 48-well plates/ seeding with 10000 cells
- Four cell lines were used: Three mutant and one control WT

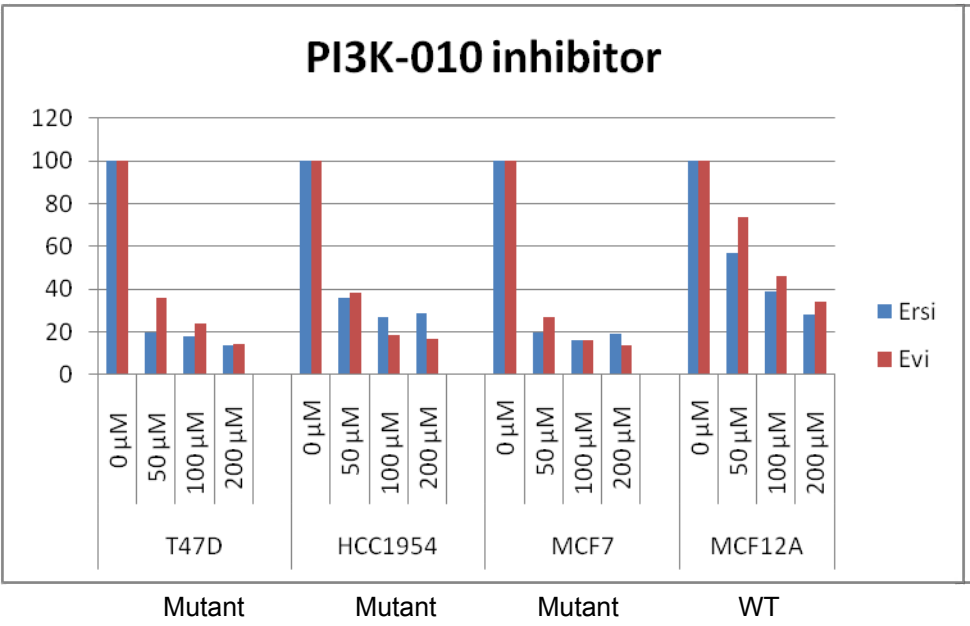


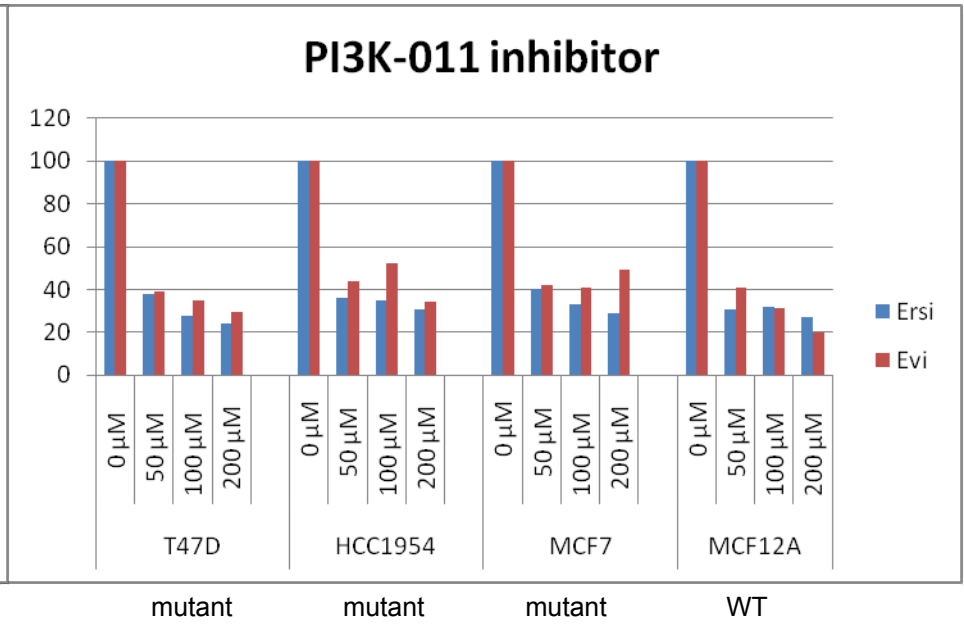




LY294002 (

- **LY290004 is a known PI3Kα inhibitor (control experiment)**
- **Compounds PI3K-001 – 011 were assayed**

# MTT assay on mutant and WT PI3Kα



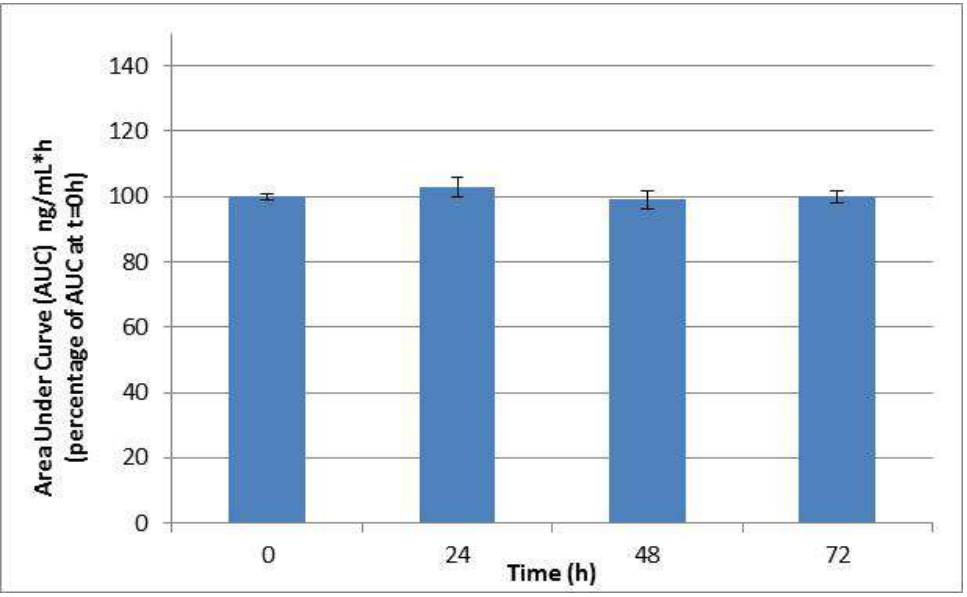




- 7-fold Mutant-specific inhibition in cells bearing the H1047R mutation
- IC50 WT = 7uM
- IC50 H1047R = 1uM

(Cournia and Efstratiadis labs, BRFAA)

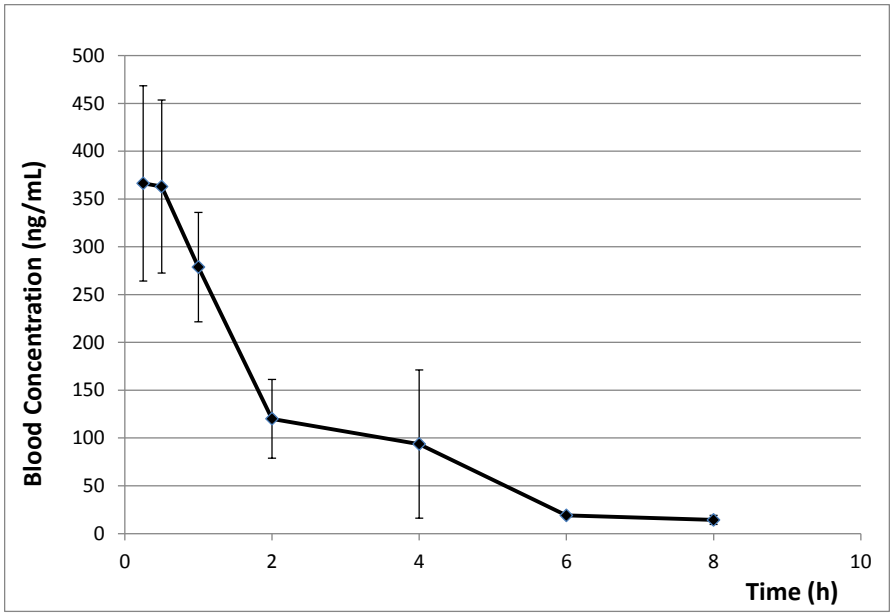
# Pharmacokinetic experiments on PI3K-010

*Stability of compound PI3K010 in cell conditioned- medium*




(Tamvakopoulos lab, BRFAA)

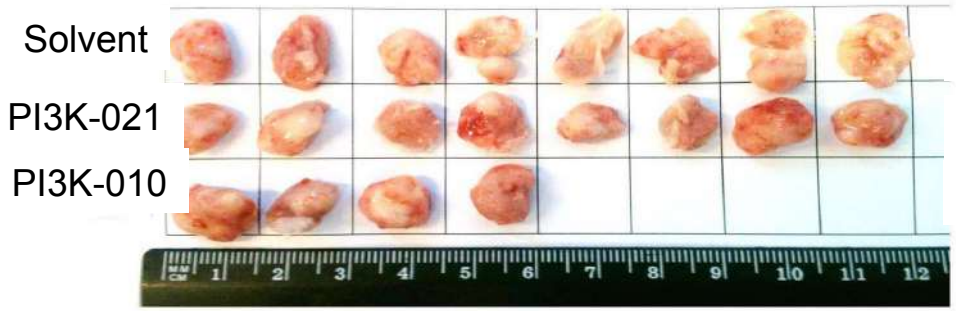
*Mean blood concentrations of PI3K010 in corn oil following oral dosing in mice (10 mg/Kg).*



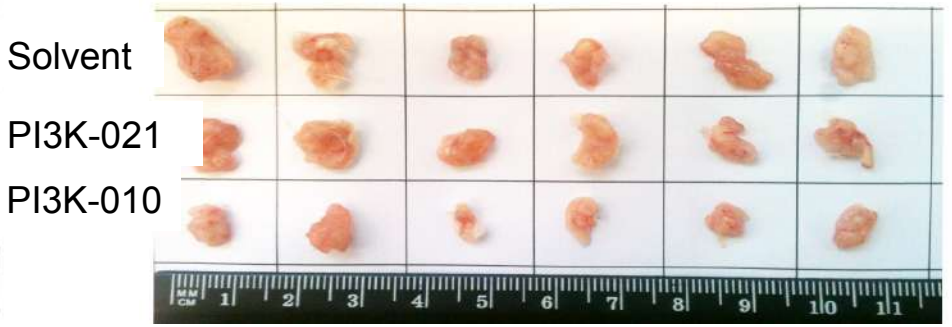

Cmax of 396 ng/mL (~ 1 μM)
4 h post-dose - average concentrations of 100 ng/mL (~ 0.3 μM).

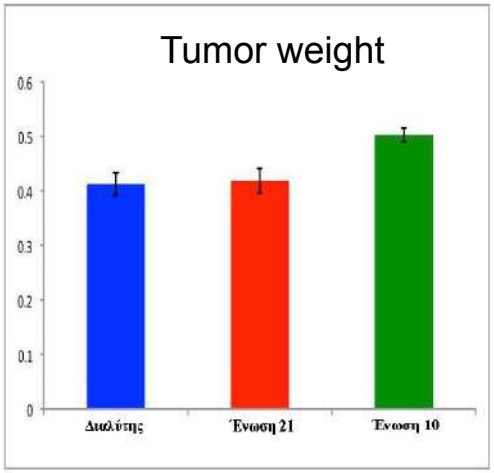
# Preclinical study of PI3K-010 (xenografts)

**MDA-231-MB (PI3Kα WT)**

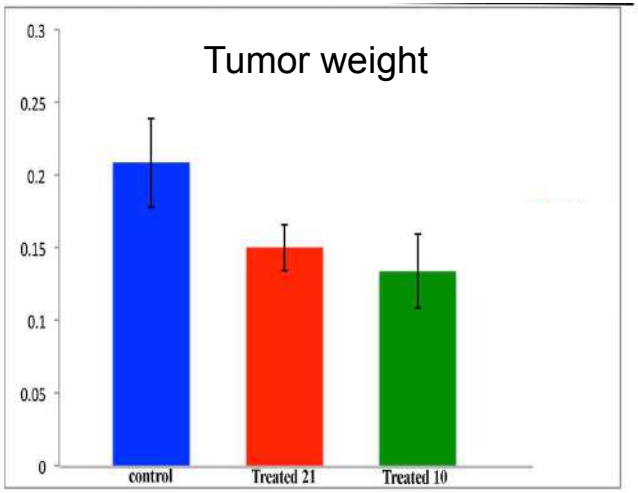



**HCC1954 (H1047R PI3Kα mutant)**



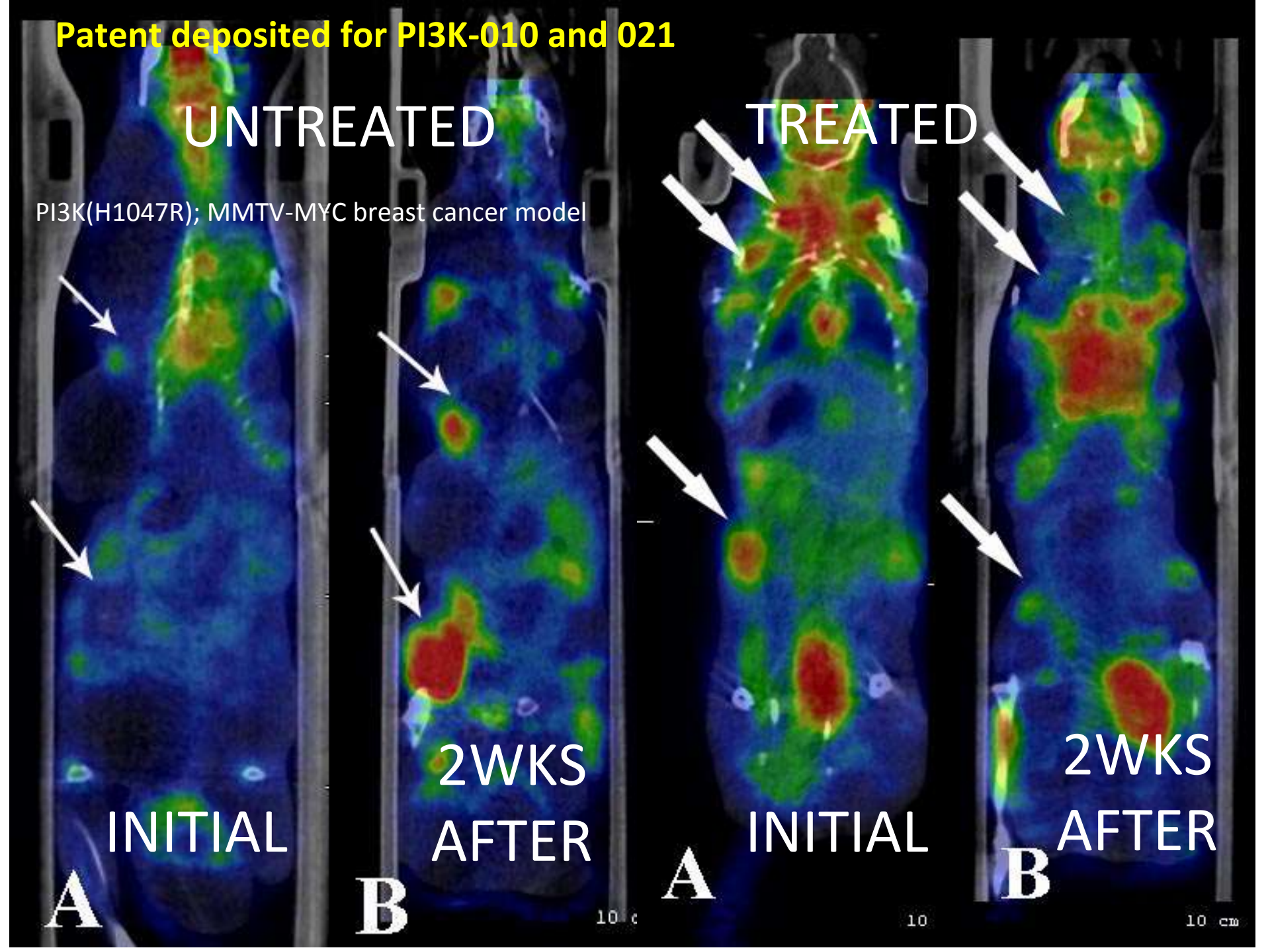




Solvent
PI3K-021
PI3K-010




(D. Stellas, Klinakis & Efstratiadis labs)

***PI3K010 in corn oil following oral dosing in mice (100 mg/Kg).***


Patent deposited for PI3K-010 and 021
UNTREATED
TREATED
PI3K(H1047R); MMTV-MYC breast cancer model
INITIAL
2WKS AFTER
INITIAL
2WKS AFTER
A
B
A
B
10
10
10 cm

# Lead optimization of PI3K-010

**Synthesis of analogs**

**Compound PI3K-021**

**In vitro cell-free assay**

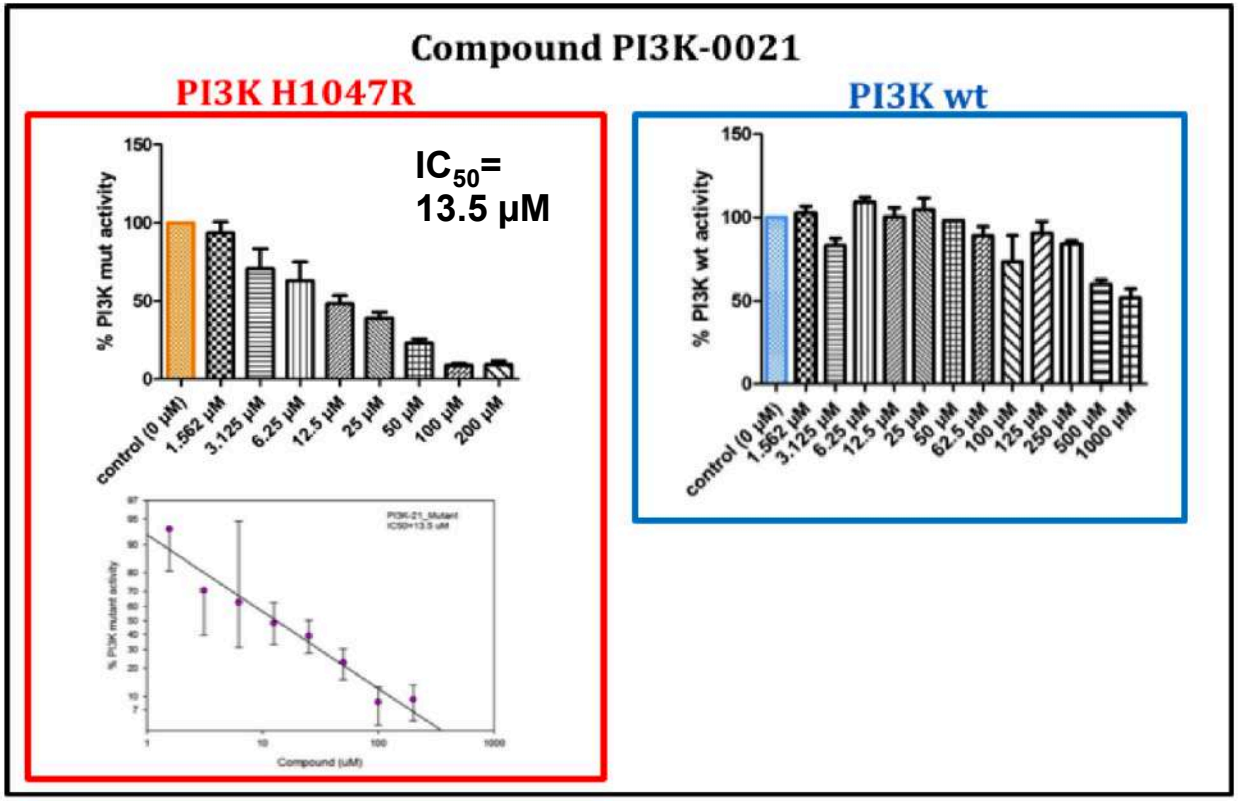
**$IC_{50}$ WT: > 1000 μM**

**$IC_{50}$ Mutant: 13.5 μM**

**Selectivity > 100 fold**

Solubility issues with PI3K-021

Optimization of pchem properties




Sent to S. Gabelli Johns Hopkins U
for X-ray last week

# Διαθέσιμες Πτυχιακές

- Βελτιστοποίηση δραστικότητας υποψηφίων φαρμάκων
- Μοριακές Δυναμικές Προσομοιώσεις για αντι-καρκινικούς στόχους με στόχο τη μελέτη της δομής και δυναμικής μεταλλάξεων
- Σχεδιασμός αναστολέων για αντι-καρκινικούς στόχους
- Εφαρμογές τεχνητής νοημοσύνης στο σχεδιασμό φαρμάκων

# Project Team

**BRFAA**
**Cournia lab (MD, drug design, cells)**
**Dr. Evi Gkeka**
**Dr. Hari Leontiadou**
**Thomas Evangelidis**

**Efstratiadis & Klinakis labs (cells+mice)**
**Dr. Ersi Tsellou**
**Dr. Dimitris Stellas**

**NCSR Demokritos**
**Couladouros lab**
**Anna Kapela**
**Maria Ouzouni**

**University of Thrace**
**Agianian lab**
**Dr. Maria Pavlaki**

**University of Ioannina**
**Christoforidis lab (cell-free assays)**
**Alexandra Papafotika**
**Dr. Vasiliki Lazani**

American Association for Cancer Research