A dynamic QSPR framework for understanding drug resistance

Maria G. Khrenova

[1] Lomonosov Moscow State University
[2] Federal Research Centre “Fundamentals of Biotechnology” of the Russian Academy of Sciences

October 19, 2026

Amino acid variations in protein targets and drug resistance

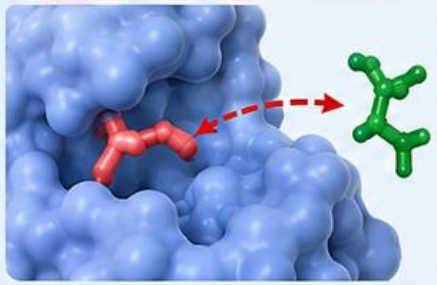
1. Reduced drug binding affinity

Amino acid substitution alters the shape or charge of the drug-binding site, reducing drug binding.

Wild-type (drug sensitive)



Mutant (drug resistant)

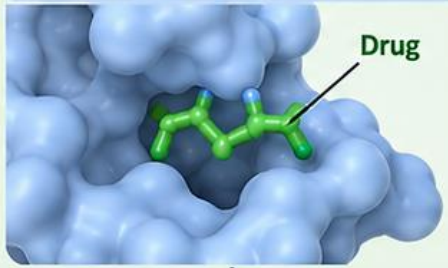


Result:
Weaker drug binding leads to reduced inhibition of the target.

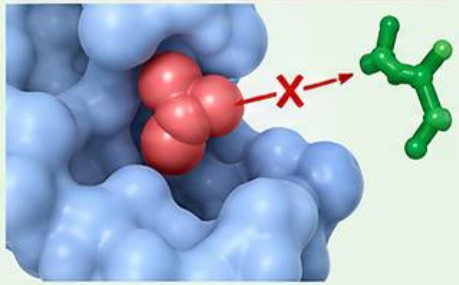
2. Steric hindrance

A bulkier amino acid substitution creates a physical barrier that prevents the drug from binding to its target.

Wild-type (drug sensitive)



Mutant (drug resistant)

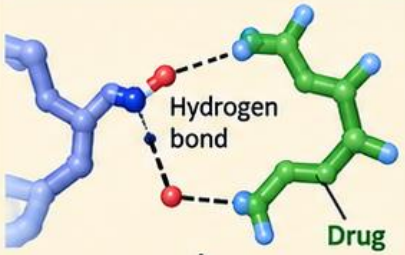


Result:
Drug cannot access the binding site due to a bulkier amino acid.

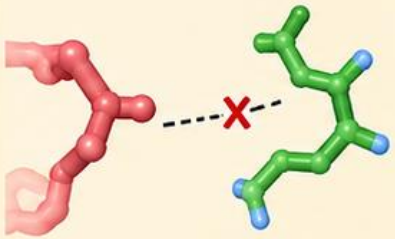
3. Disruption of key interactions

Substitution removes an amino acid that forms critical hydrogen bonds, ionic interactions, or hydrophobic contacts with the drug.

Wild-type (drug sensitive)



Mutant (drug resistant)

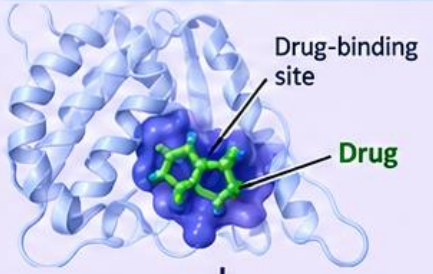


Result:
Loss of key interactions reduces drug binding and effectiveness.

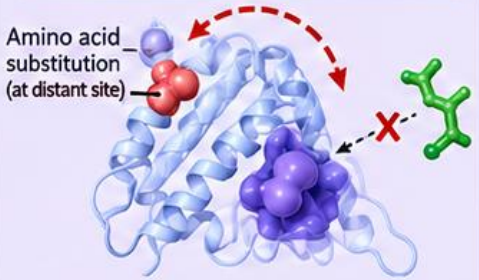
4. Allosteric effects

Amino acid substitution at a distant site induces a conformational change that alters the drug-binding site.

Wild-type (drug sensitive)



Mutant (drug resistant)

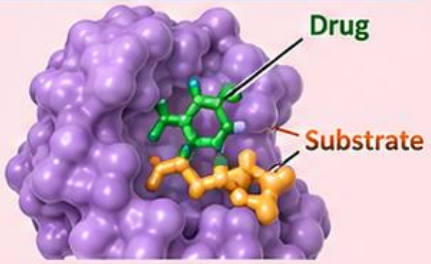


Result:
Conformational change decreases drug binding at the target site.

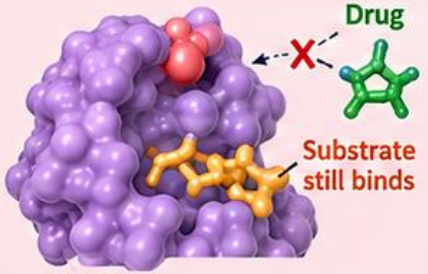
5. Restoration of function in resistance mutations

In some cases, an amino acid substitution reduces drug binding while maintaining or restoring the normal function of the target enzyme or protein.

Wild-type (drug sensitive)



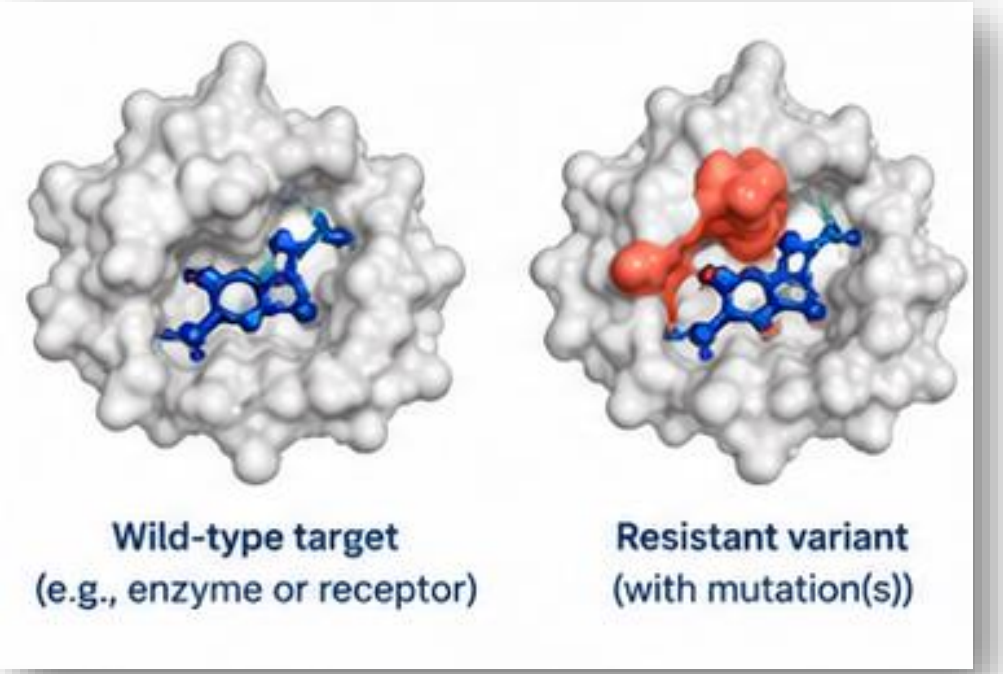
Mutant (drug resistant)

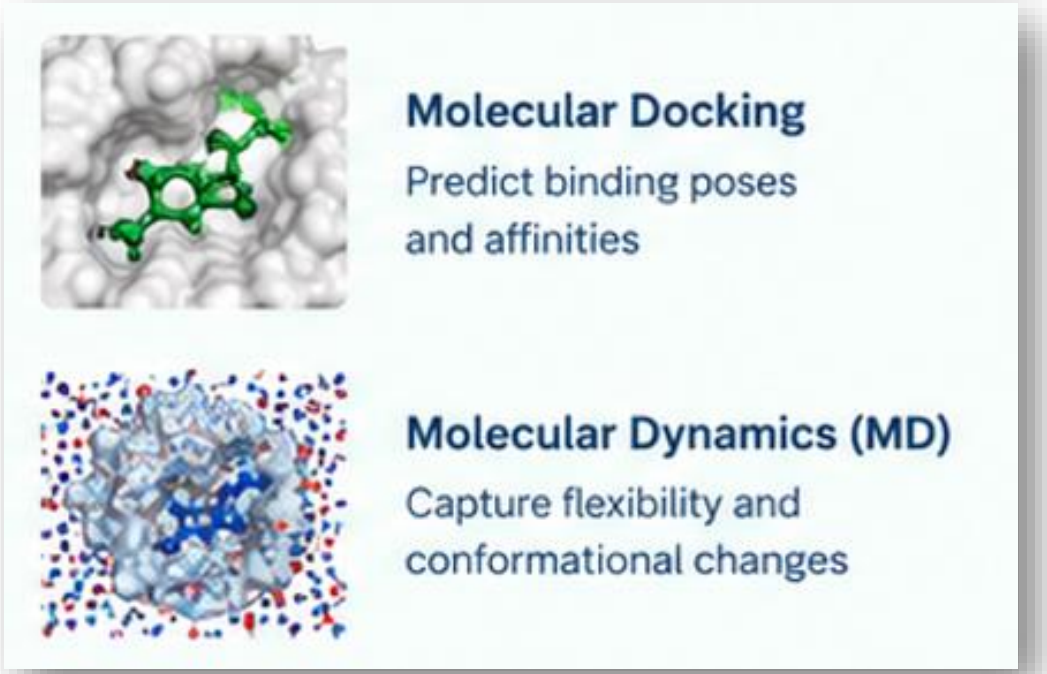


Result:
Drug resistance while preserving essential protein function (e.g., enzymatic activity).

* Created with the ChatGPT

Quantifying drug resistance using molecular modeling tools





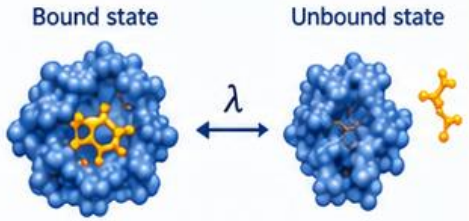
Value of interest is binding energy

Methods in Molecular Dynamics to Calculate Binding Energies

Different approaches • Same goal: estimate the strength of interaction between molecules

1. Thermodynamic Integration (TI)

Compute the free energy difference between bound and unbound states by integrating over a coupling parameter λ.



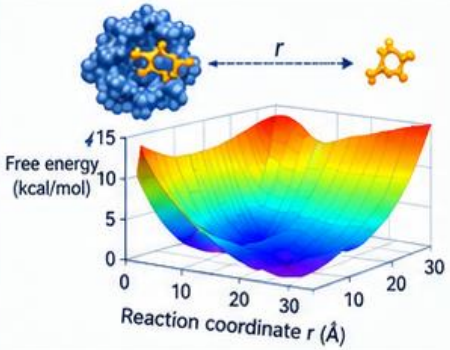
$$\Delta G_{\text{bind}} = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda$$

- Alchemical transformation (gradual decoupling/recoupling)
- Requires multiple λ windows
- Often combined with enhanced sampling (e.g., replica exchange)

Output: ΔG_{bind} (free energy)

2. Potential of Mean Force (PMF)

Calculate the free energy profile as a function of a reaction coordinate (e.g., distance between ligand and protein).



$$\Delta G_{\text{bind}} = -k_B T \ln K_a$$

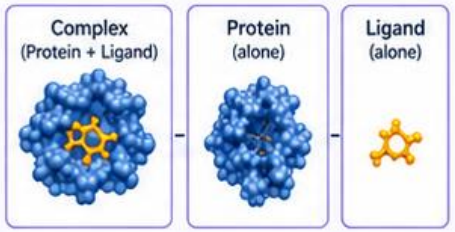
(from PMF minimum)

- Umbrella sampling / metadynamics
- Extract binding free energy from PMF or standard state correction
- Works well for post-binding processes (e.g., unbinding)

Output: ΔG_{bind} (from PMF)

3. MM/PBSA and MM/GBSA

Estimate binding energy from snapshots of MD simulations using molecular mechanics and continuum solvation.



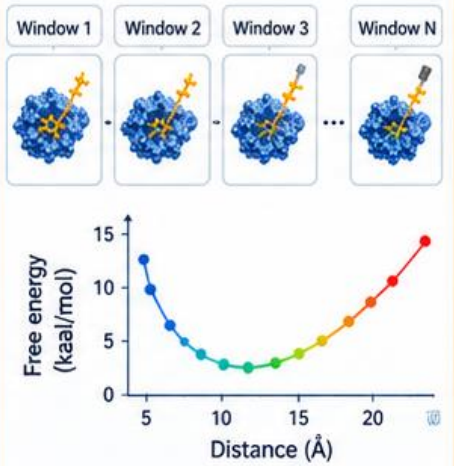
$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{protein}} - G_{\text{ligand}}$$
$$\approx \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\Delta S$$

- Uses MD snapshots (no alchemical steps)
- Solvation via Poisson–Boltzmann (PB) or Generalized Born (GB)
- Entropy often estimated by normal modes or quasi-harmonic analysis

Output: ΔG_{bind} (approximate)

4. Umbrella Sampling

Enhance sampling along a chosen coordinate (e.g., separation distance) to overcome energy barriers and obtain the PMF.

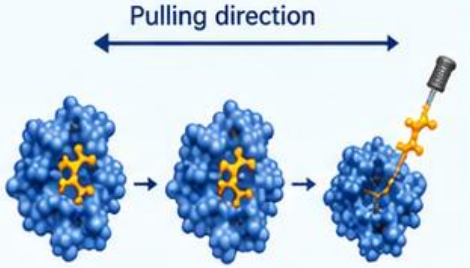


- Apply biasing potentials (umbrellas)
- Reconstruct PMF (e.g., WHAM)
- Widely used for weak and transient interactions

Output: PMF → ΔG_{bind}

5. Steered MD (SMD)

Apply an external force to break the complex and calculate the work required to separate the partners.



$$\Delta G_{\text{bind}} \approx \langle W_{\text{puall}} \rangle - \Delta G_{\text{correction}}$$

- Pull ligand or protein with a spring
- Calculate work from force–distance data
- Fast, but less accurate (requires correction and repeats)

Output: ΔG_{bind} (from work)

Key Takeaways

- ✔ Different methods suit different systems and questions (affinity, kinetics, mechanism).
- ✔ Accuracy depends on force fields, sampling, and analysis.
- ✔ Combining methods (e.g., TI + MM/PBSA) often gives the best results.
- ✔ Molecular dynamics provides a versatile toolkit to calculate binding energies and understand molecular interactions.

* Created with the ChatGPT

Calculated and observed values

Calculations		Experiment
Activation energy	$k \sim A(T) \cdot e^{-\frac{\Delta G(T)^{\neq}}{RT}}$	Rate constant
Gibbs energy of reaction	$K \sim e^{-\frac{\Delta G^{o}(T)}{RT}}$	Equilibrium constant



1 kcal/mol error in energy estimated corresponds to 5-6 fold difference in rate / equilibrium constant

Potential alternatives are QSPR-like relationships

Relations between micro and macro parameters

✓ Regression and classification models

- ❖ QSPR, QSAR – quantitative structure property/activity relationship
- ❖ Predictors are molecular descriptors and physicochemical properties
- ❖ More complex descriptors can be applied: electron density descriptors and dynamic properties.



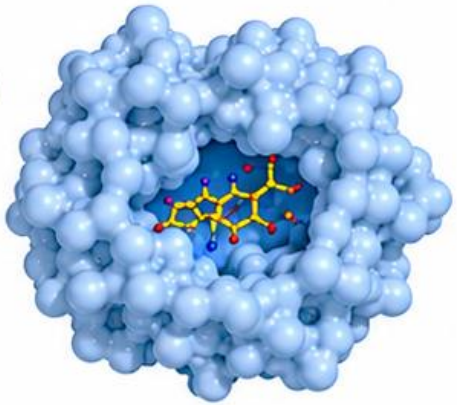
Zefirov N.S.
(1935 – 2017)

Beyond Static Structures:
The Role of Protein–Ligand Dynamics in Molecular Recognition

Same complex. Different motions. Different outcomes.

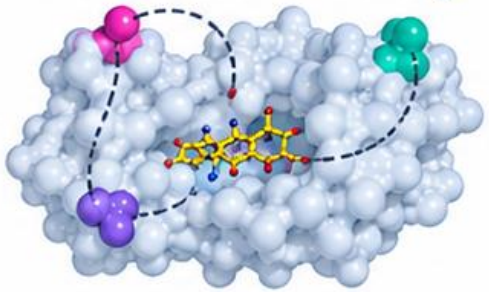
Static structure

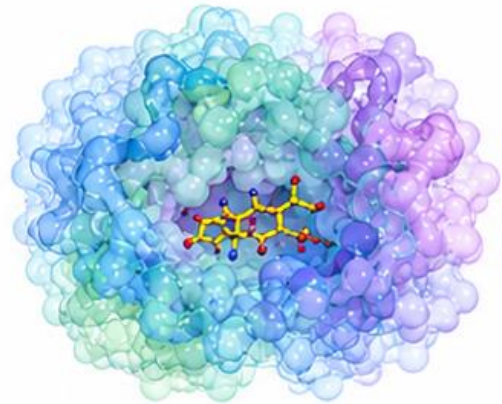
Reveals the binding site and key interactions



Distant amino acid substitutions

Can alter dynamics and binding without direct contact with the ligand

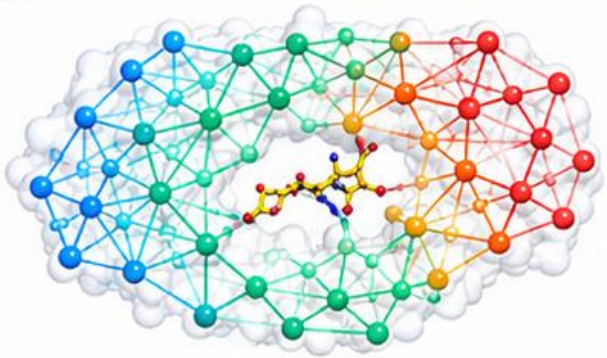




Dynamic view

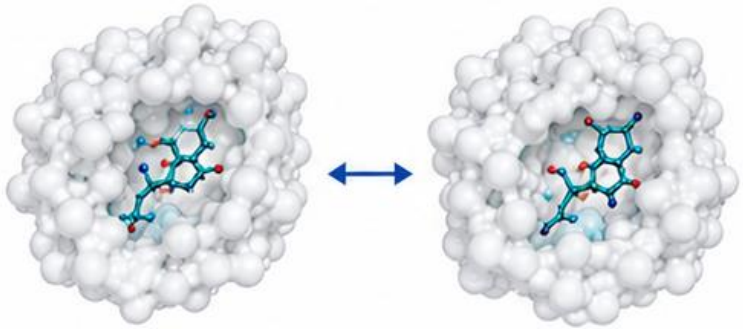
Shows a set of structures from MD simulations, representing different conformations of the protein–ligand complex.

1 Collective motions and efficient binding



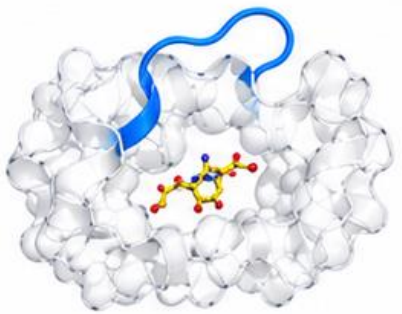
Dynamic network analysis distinguishes collective motions of inhibitor and protein for efficient binding.

2 Covalent inhibitor reactant complexes

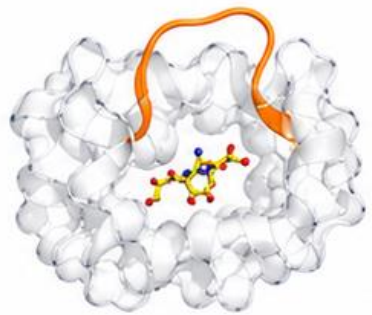


Differences in the dynamics of reactant complexes can influence covalent reaction outcomes.

3 Protein loop dynamics across variants



Variant A

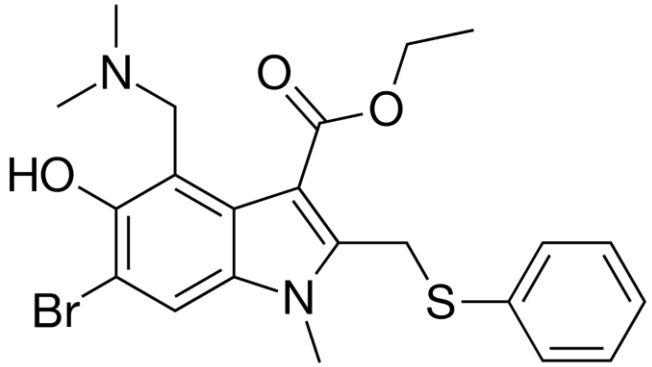


Variant B

Loop flexibility and conformational sampling differ between variants and affect inhibitor binding.

* Created with the ChatGPT

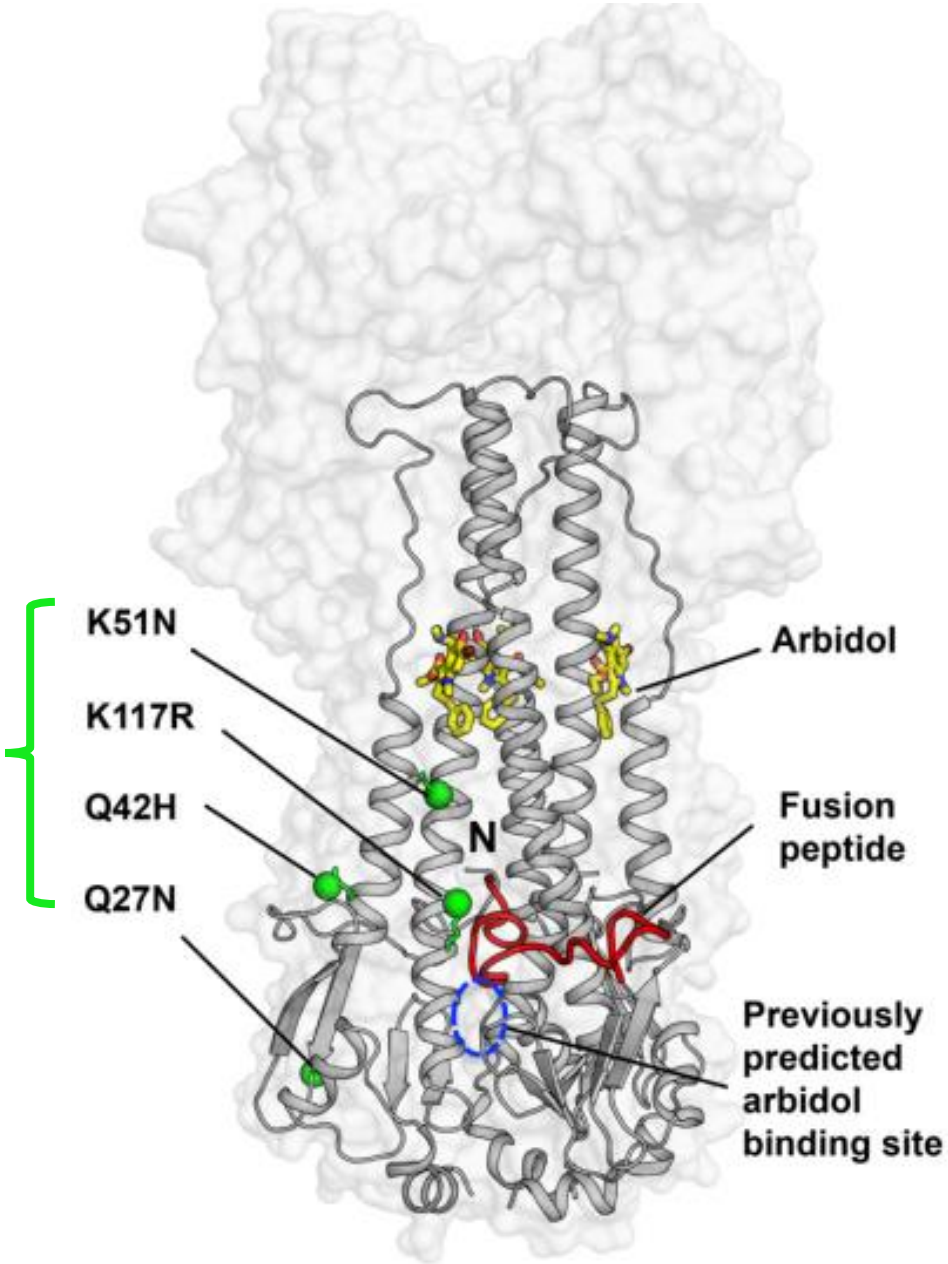
Arbidol (umifenovir) interactions with hemagglutinin



Arbidol is a Russian-made indole-derivative small molecule licensed in Russia and China for the prevention and treatment of influenza and other respiratory viral infections.

Arbidol-resistance mutations K51N, K117R, Q42H, and Q27N in HA2.
Those are far from Arbidol binding site.

Why does the resistance appear?

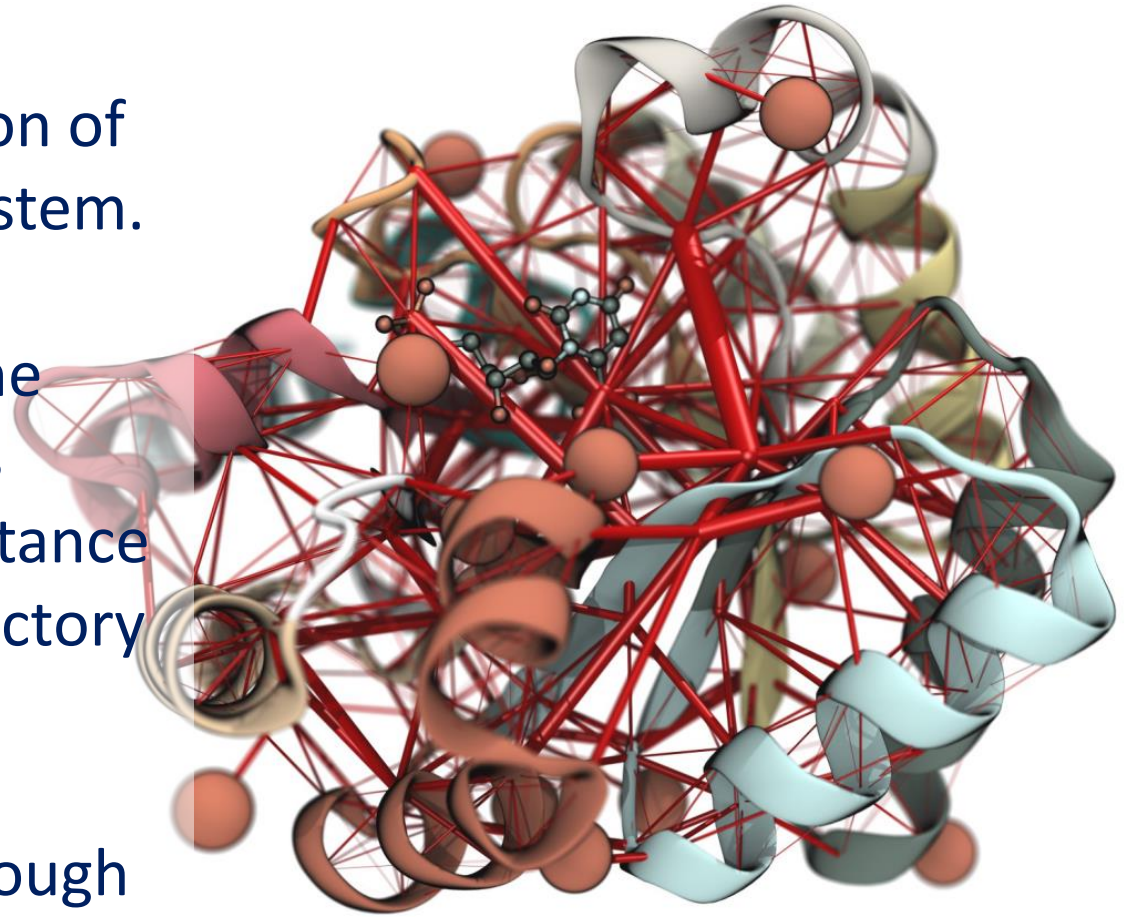


* PNAS (2017) V.114, P. 206-214.

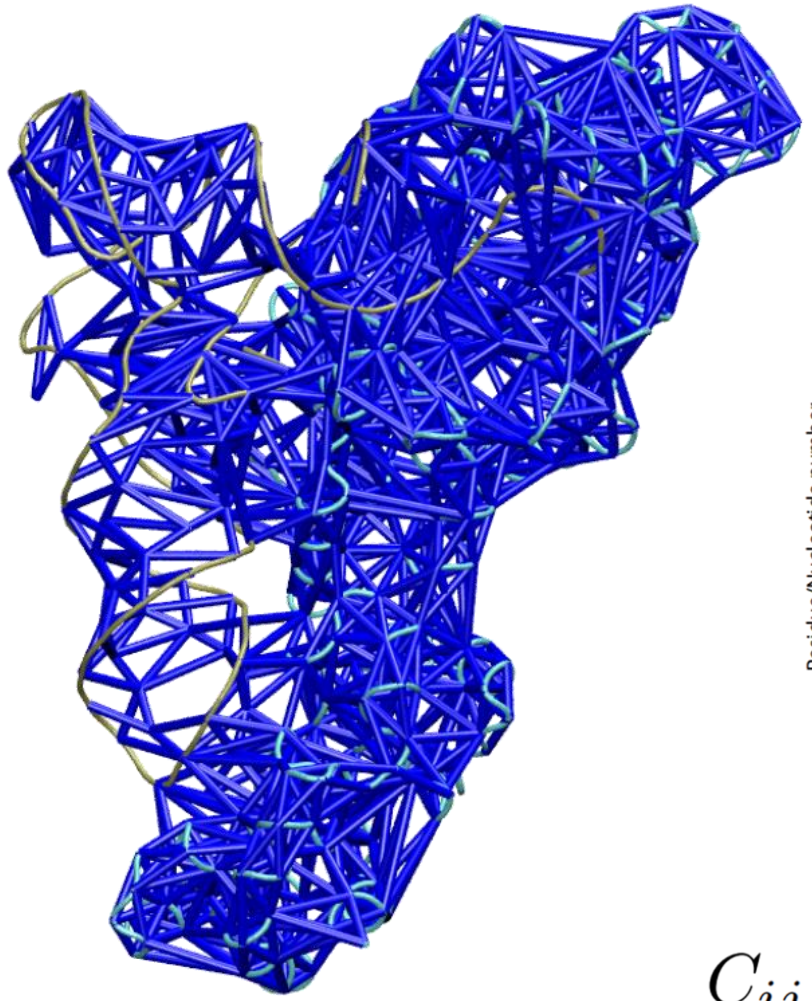
Dynamic network analysis

Dynamic network analysis studies the correlation of movement of representative atoms (or nodes), such as Ca atoms of amino acid residues, and is used to determine the presence and degree of interrelation of movement between different fragments of the system. Only short-range interactions are calculated:

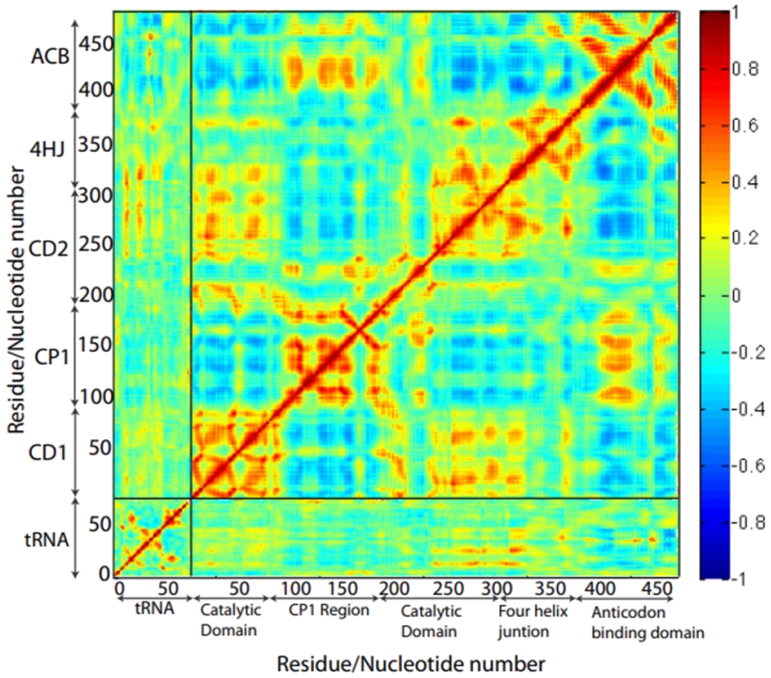
- To determine which nodes are in contact, the shortest distance between any heavy atoms related to two nodes is calculated. If the distance is less than 4.5 Å for more than 75% of trajectory times, such nodes are considered to be interacting.
- Long-range interactions are determined through short-range networks.



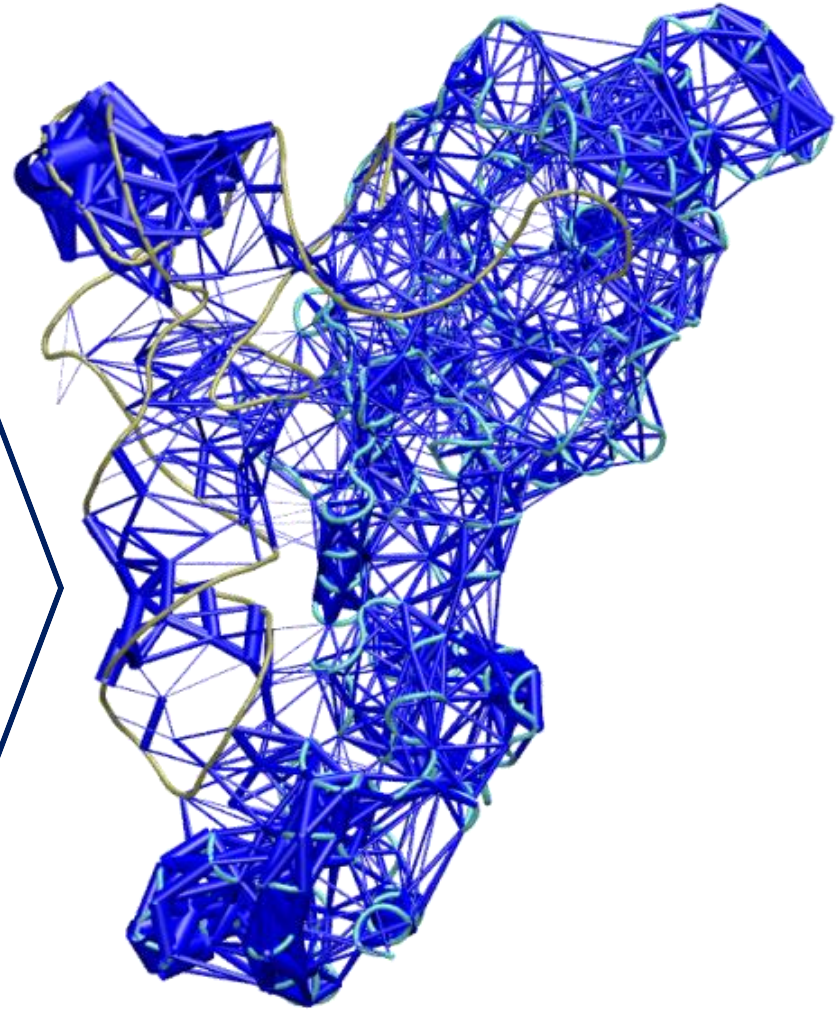
Interacting nodes



The edges between the contacting nodes (without weights)



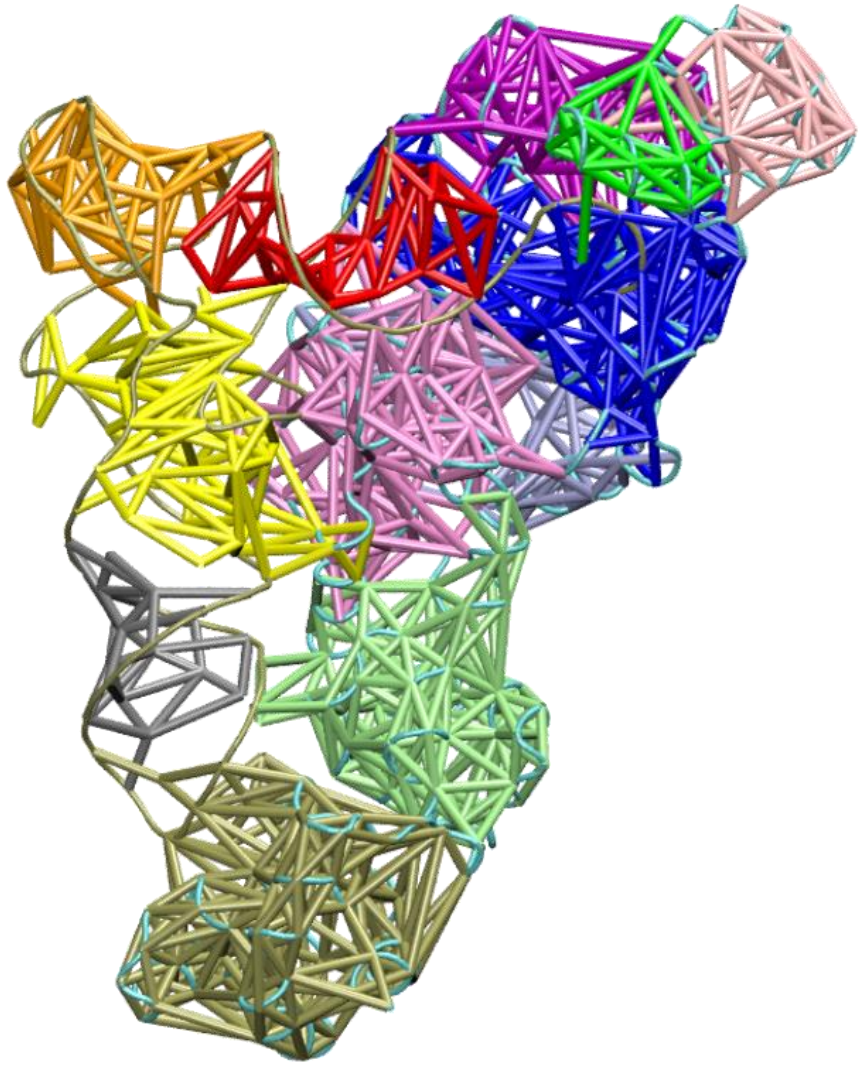
$$C_{ij} = \frac{\langle \Delta\vec{r}_i(t) \cdot \Delta\vec{r}_j(t) \rangle}{(\langle \Delta\vec{r}_i(t)^2 \rangle \langle \Delta\vec{r}_j(t)^2 \rangle)^{1/2}}$$



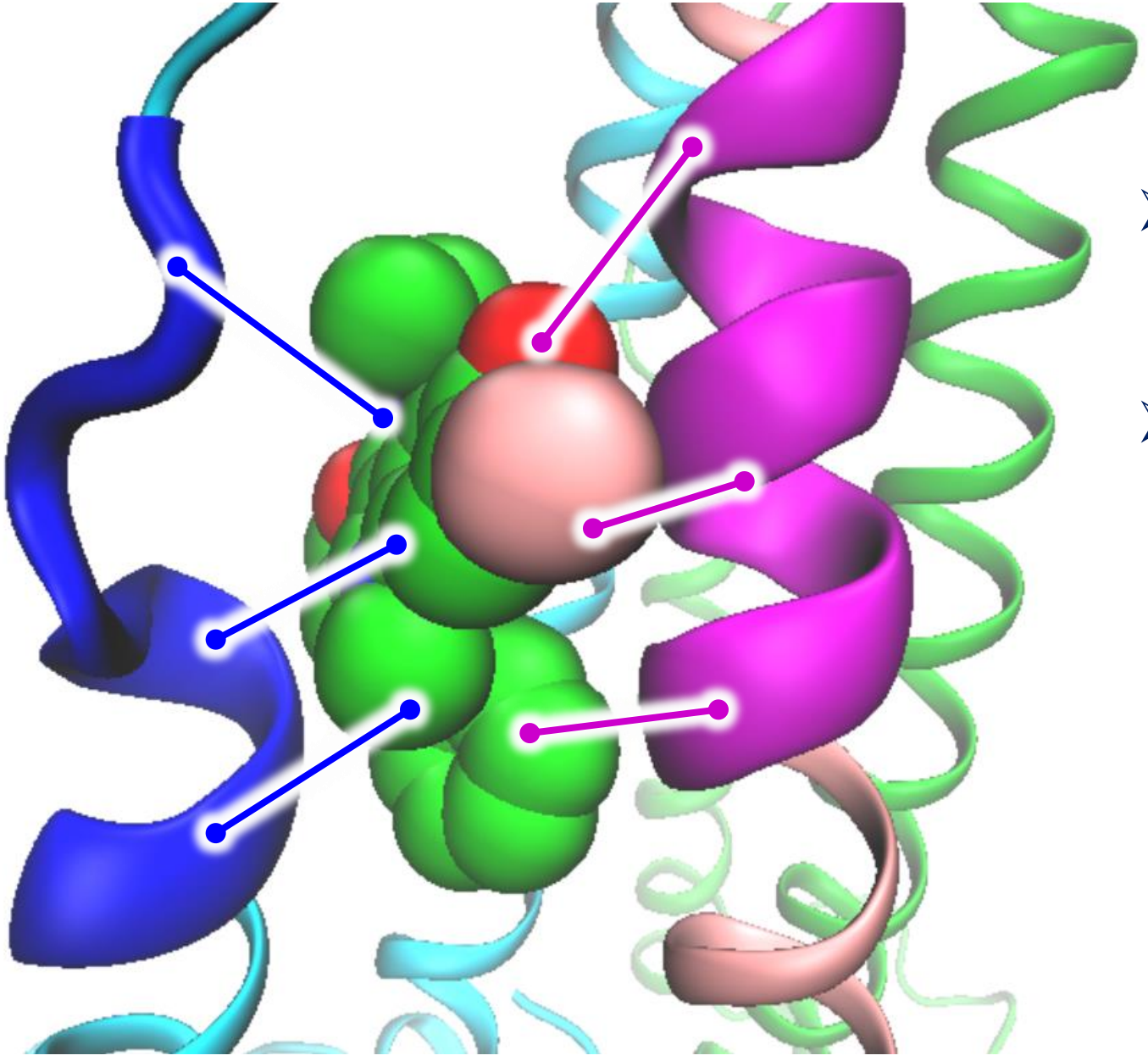
The edges between the contacting with weights

Communities

- Communities are parts of a common network. Nodes within a community are characterized by stronger connections than between communities. Communities are defined according to weights determined through correlation.

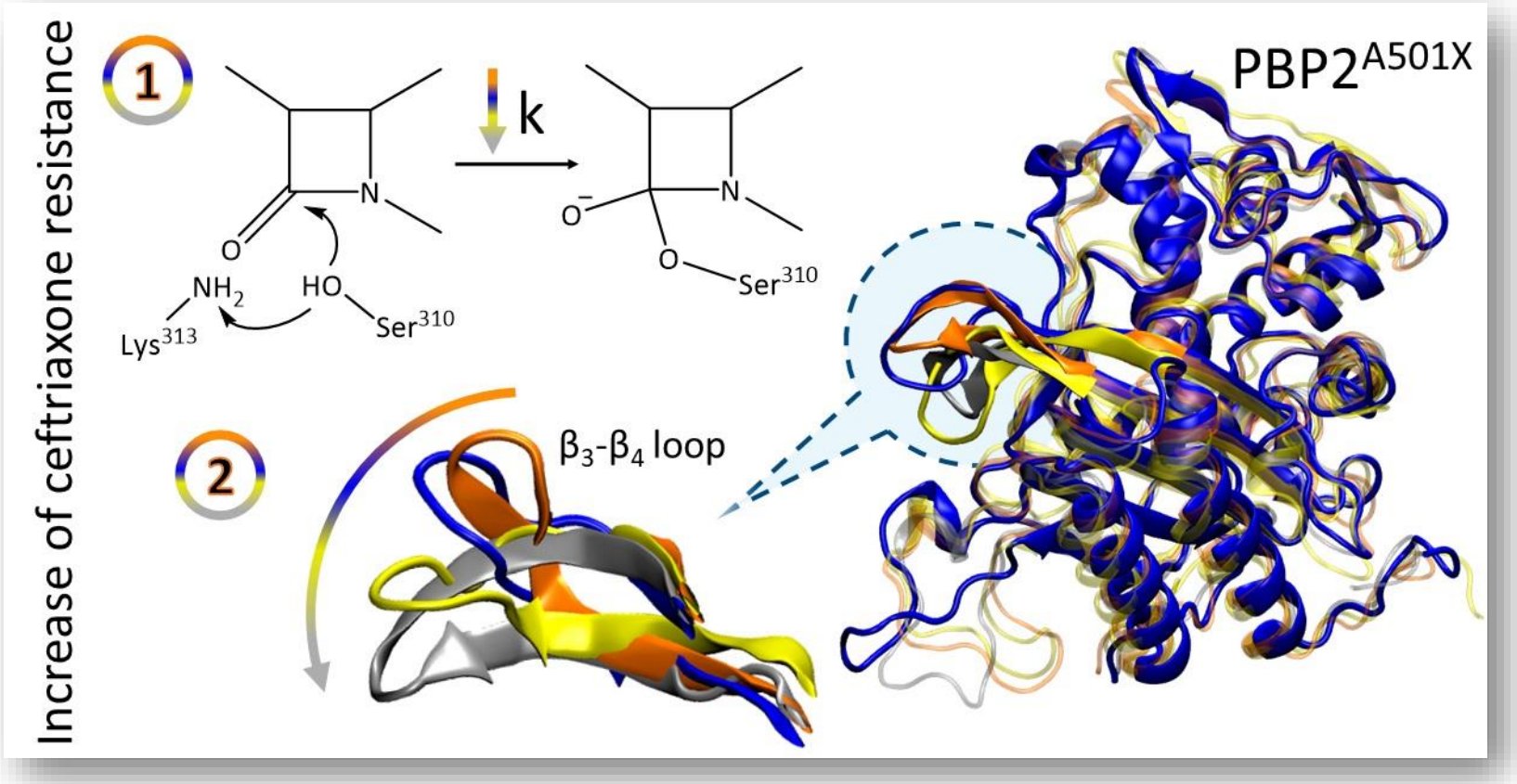


Arbidol and hemagglutinin variants: analysis

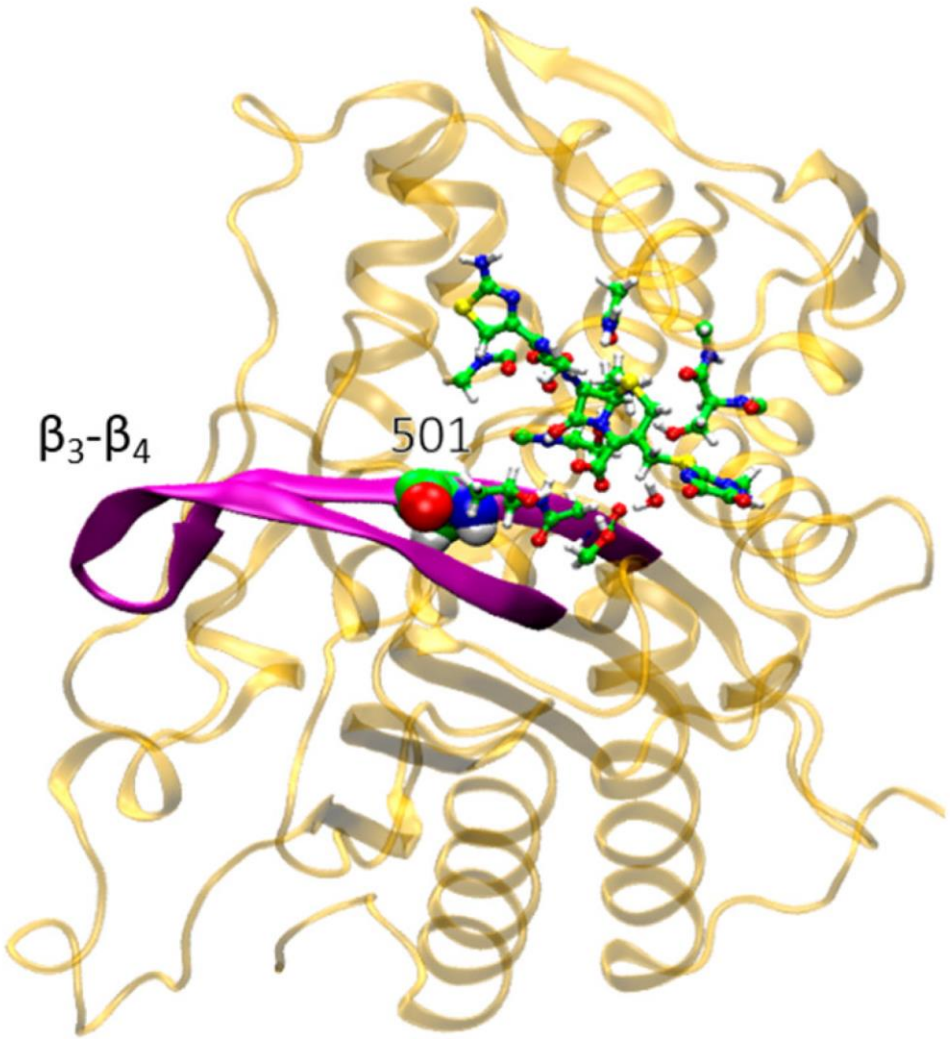


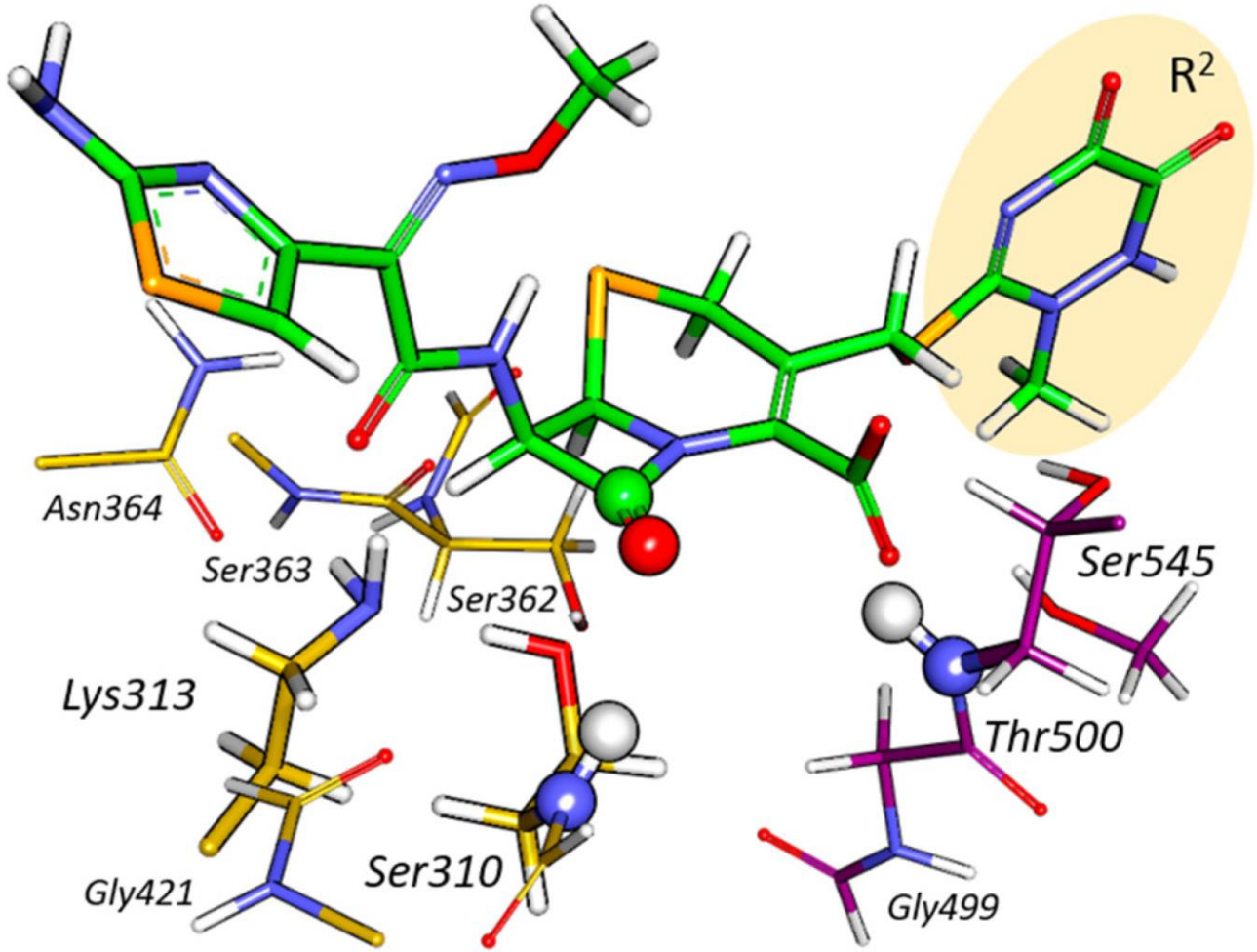
- For WT HA fragments of Arbidol belong to communities with both α-helices being like a “zip”;
- For HA from resistant strains Arbidol belongs to a community with only one α-helix.

Dynamic behavior of penicillin binding protein PBP2 and its A501X variants

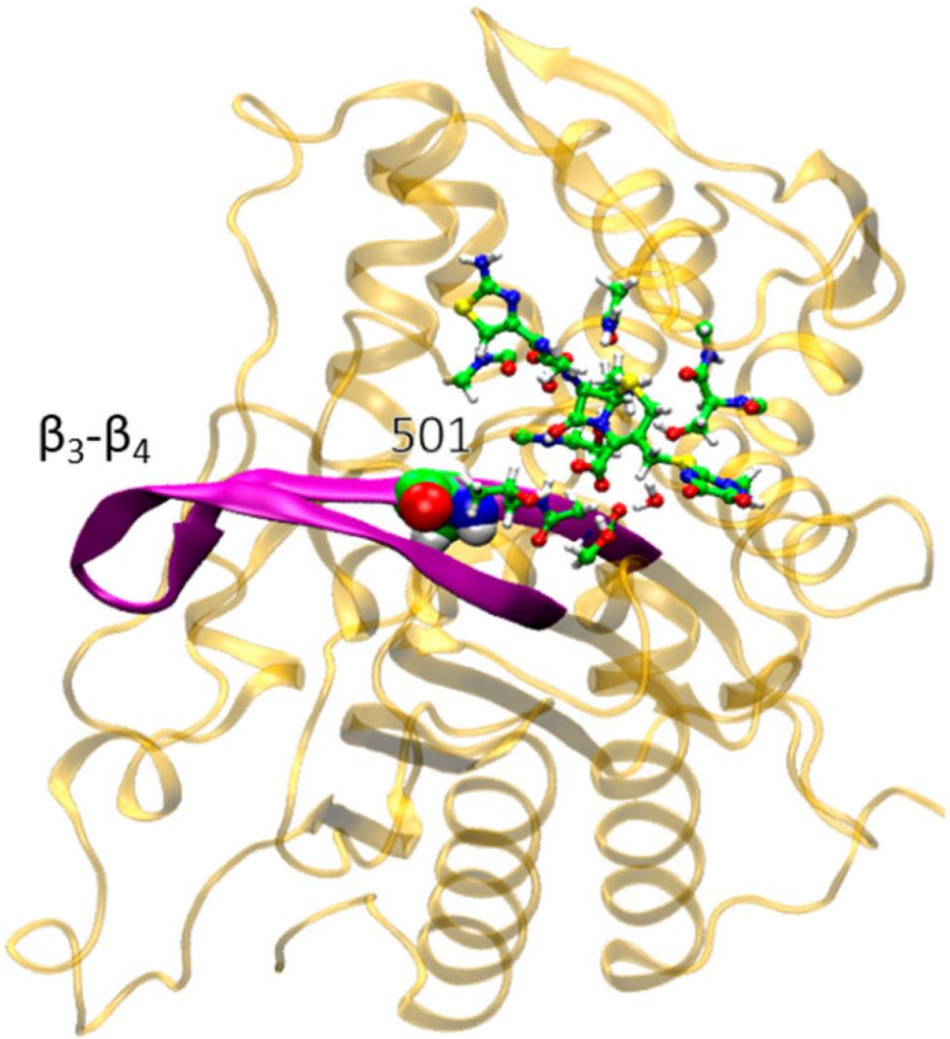


Structure of the PBP2





Penicillin binding protein PBP2



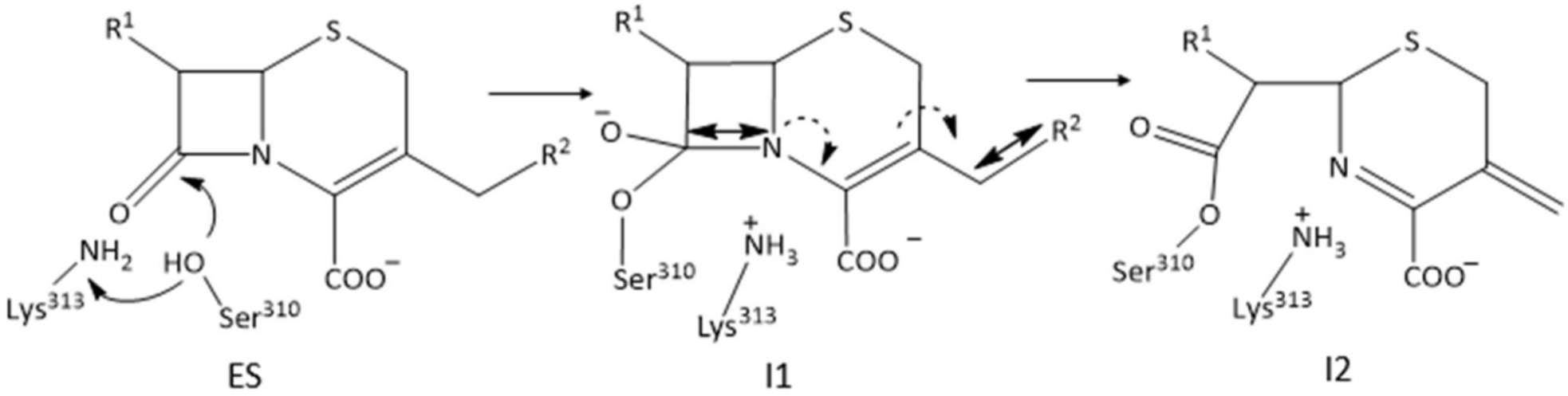
PBP2 with A501X variants demonstrate different catalytic activity *

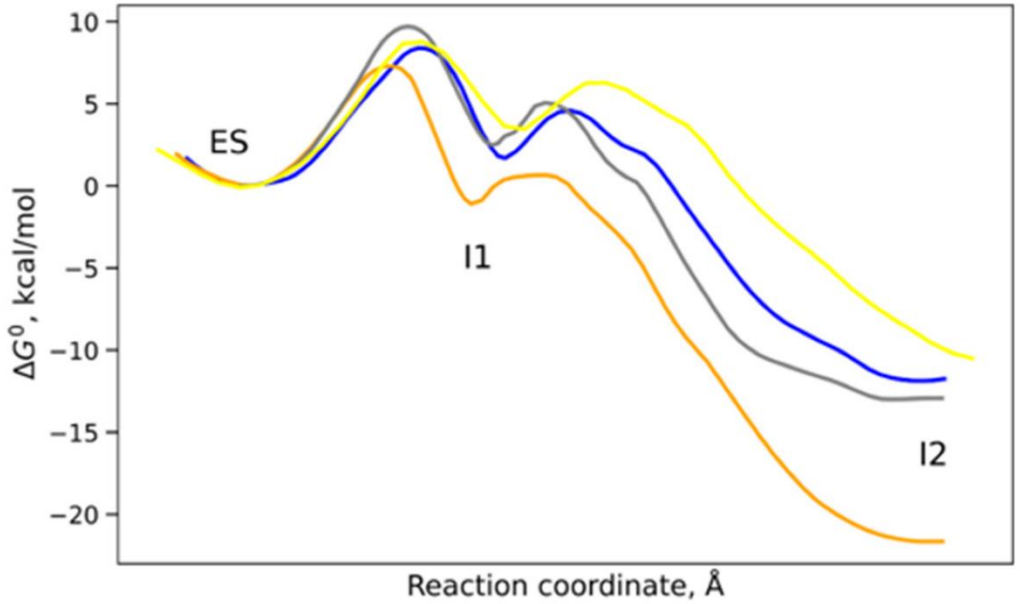
PBP2 variant	Relative k_2/K_s
A501V	1.8
A501	1
A501R	0.5
A501P	0.02

These variations can by due to changes of k_2 and/or K_s values.

* J. Tomberg et al., Biochemistry (2017)

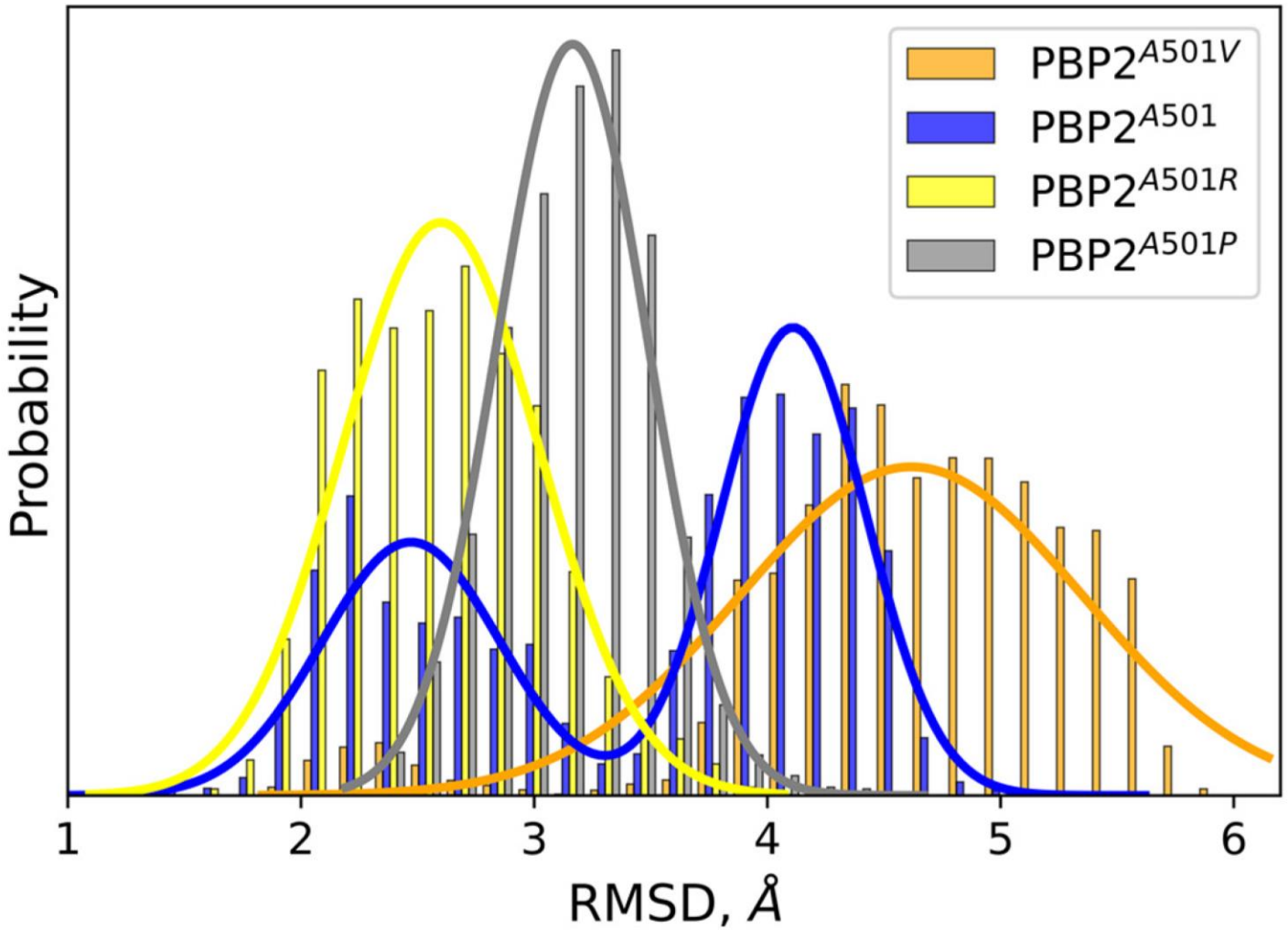
Acylation mechanism in the active site of the PBP2





QM(PBE0-D3/6-31G**)/MM MD

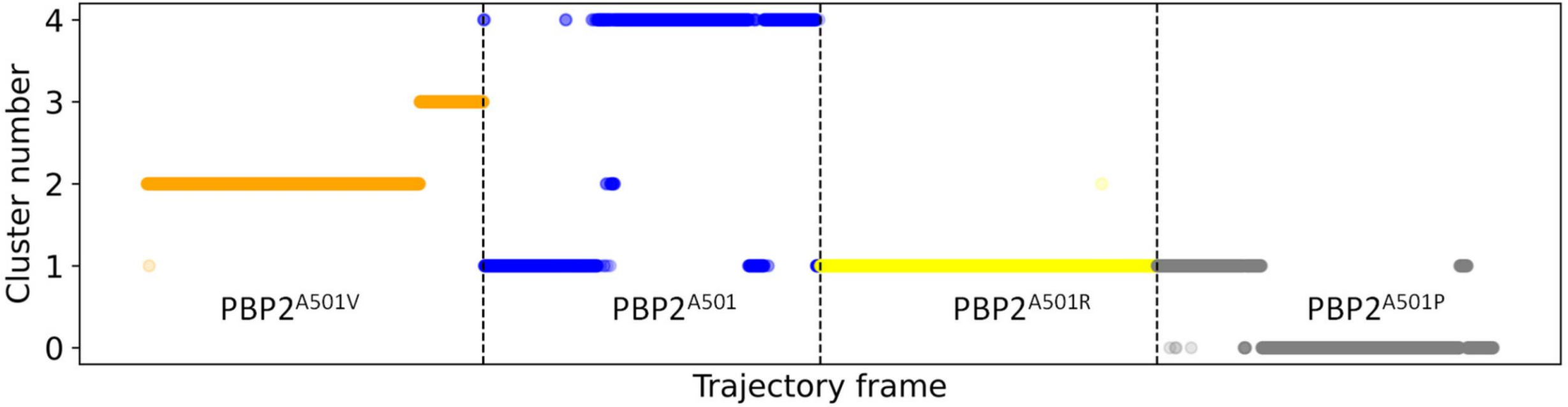
Dynamic behavior of the β_3-β_4 loop

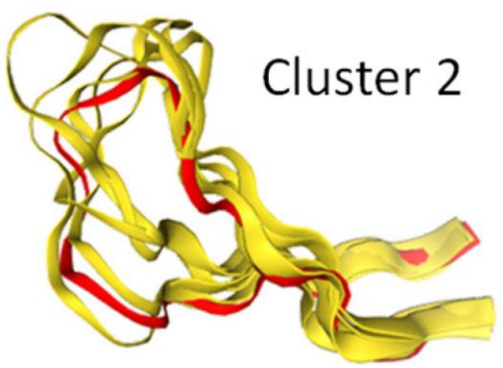


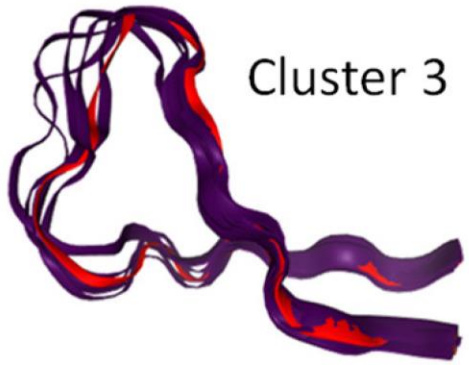
PBP2 variant	RMSD of β_3-β_4 loop mean ± st.dev., Å
$PBP2^{A501V}$	4.61 ± 0.54
$PBP2^{A501}$	4.11 ± 0.10 (w = 0.6) 2.47 ± 0.15 (w = 0.4)
$PBP2^{A501R}$	2.60 ± 0.18
$PBP2^{A501P}$	3.16 ± 0.10

MD: 500ns, CHARMM + TIP3P

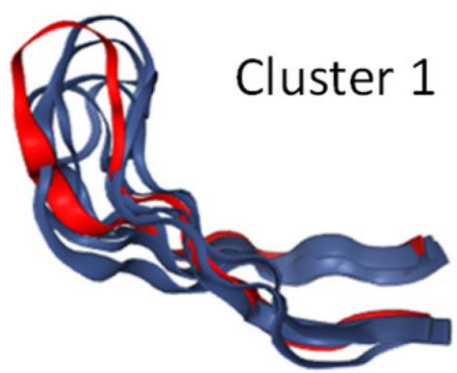
Clustering of β3–β4 loop states

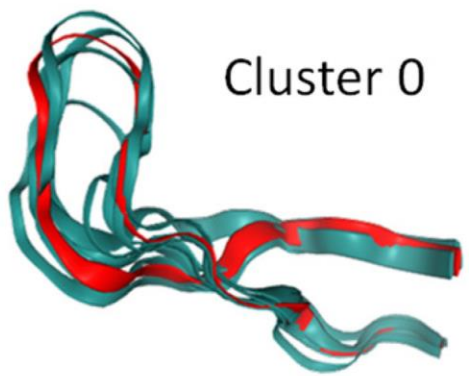






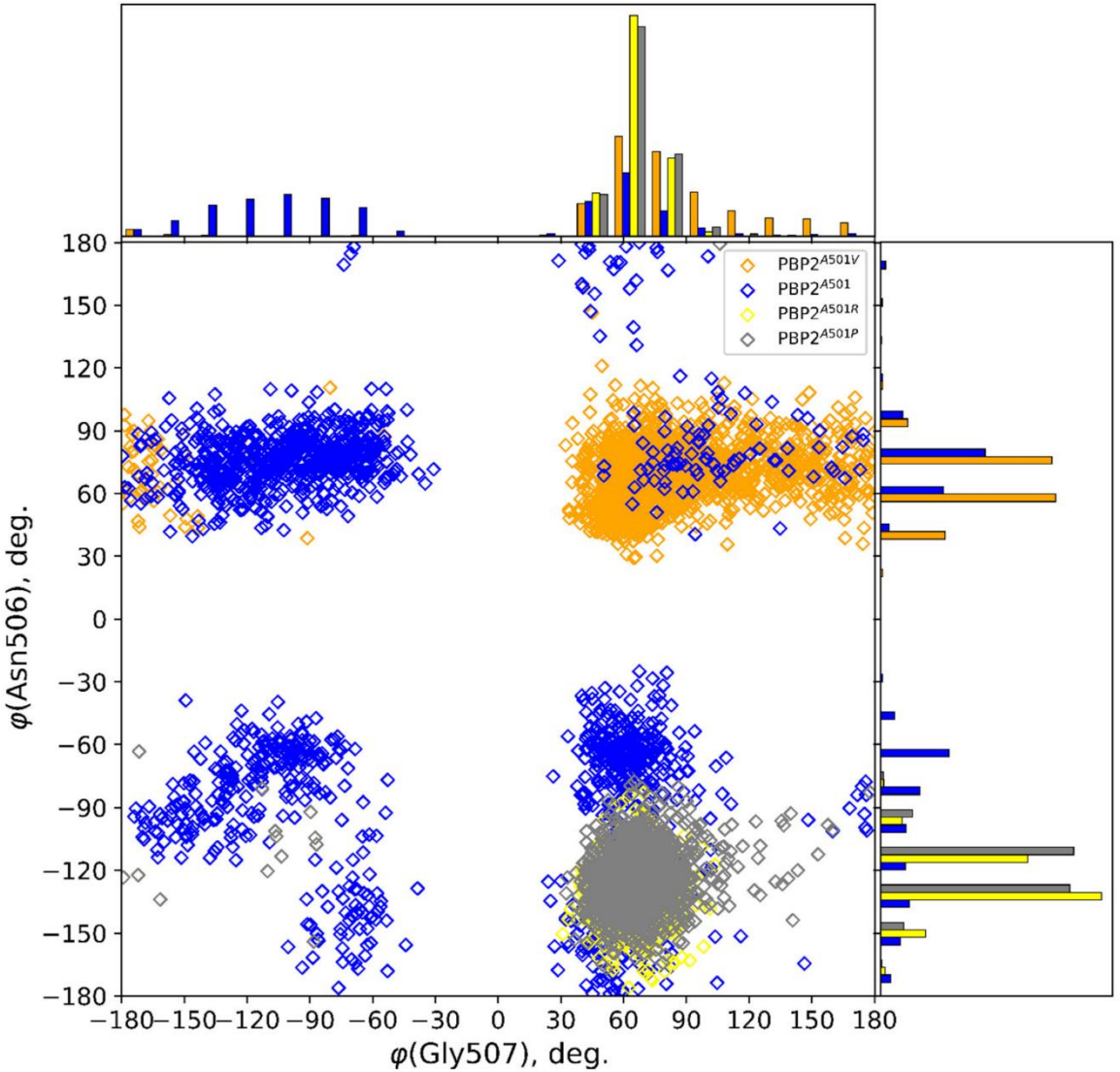


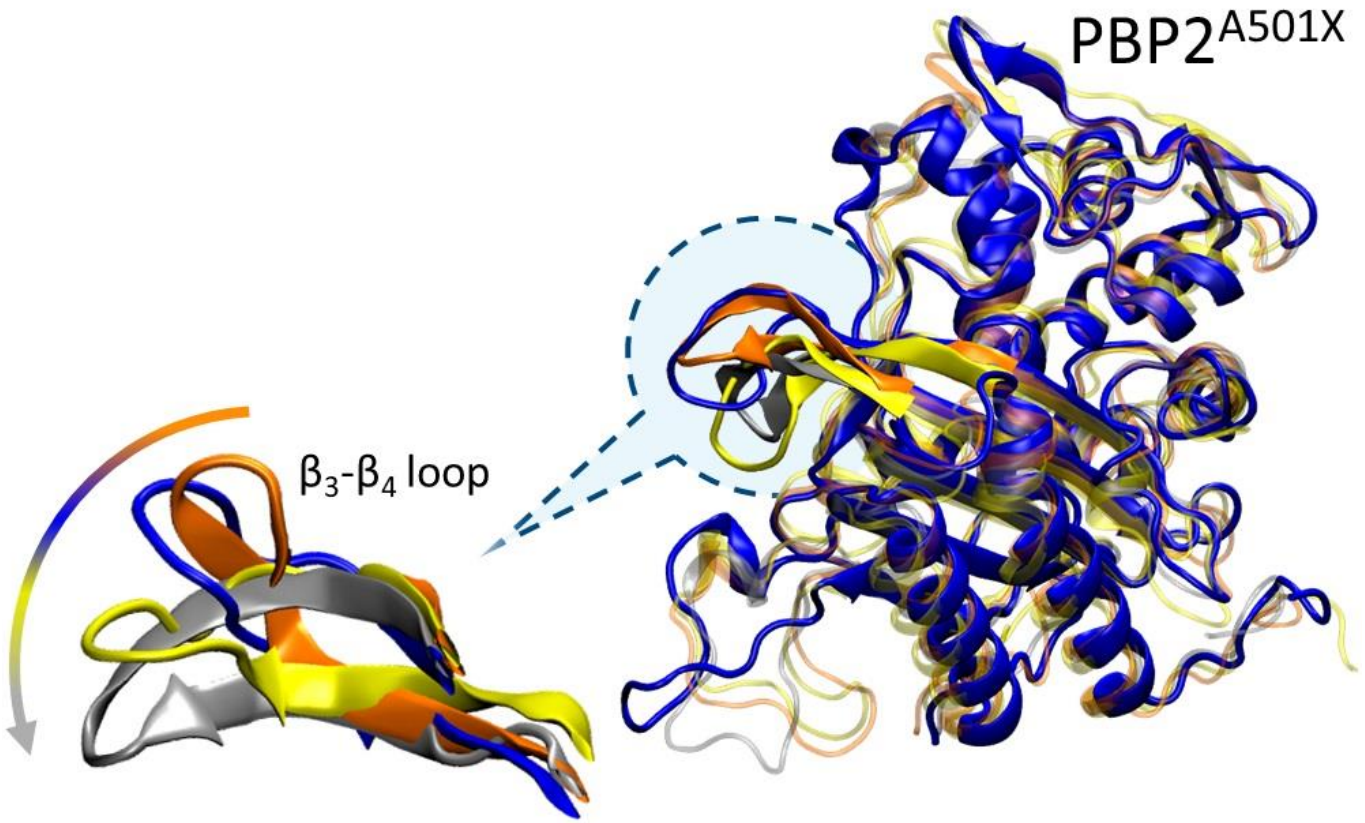




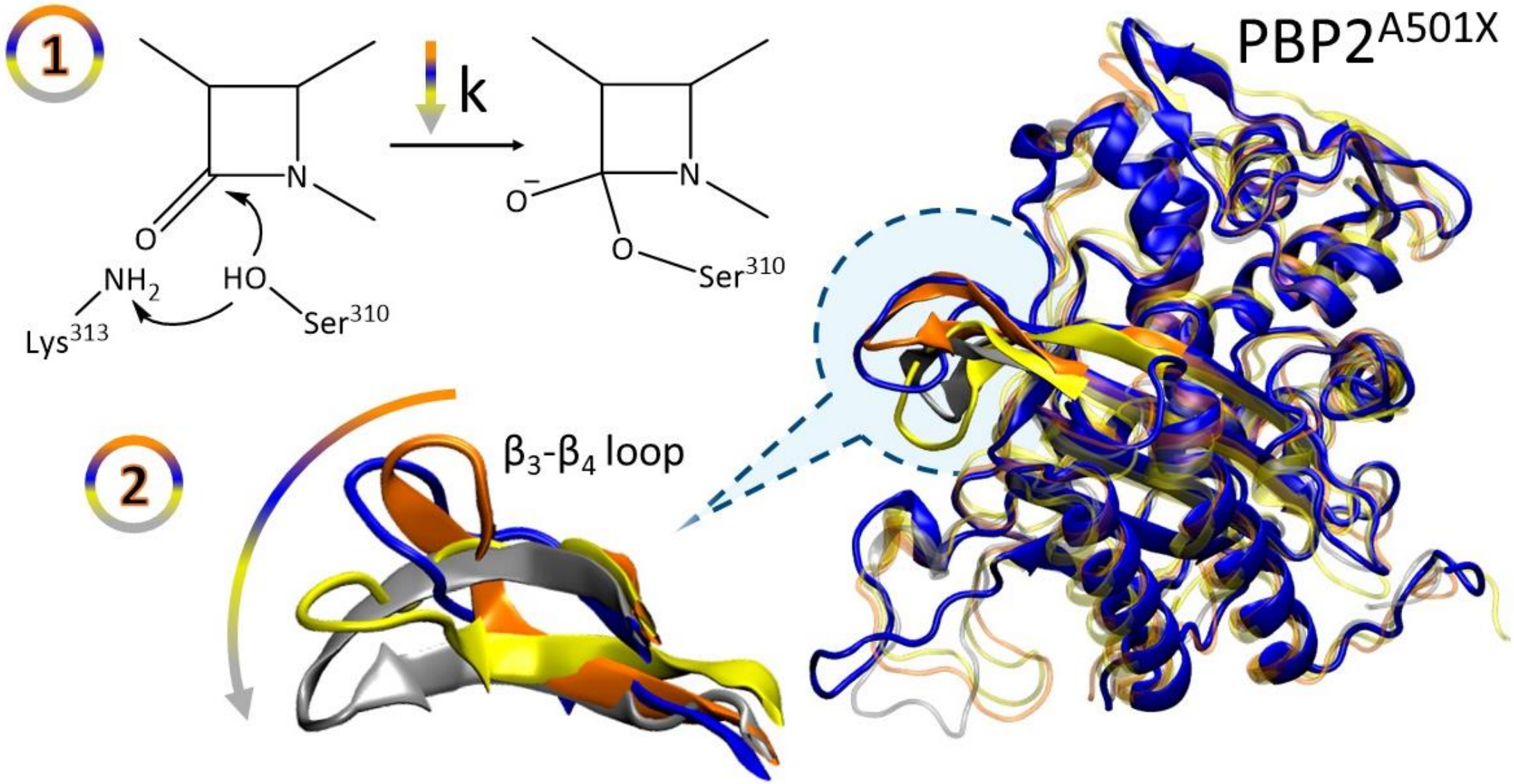
- 500 ns MD trajectories (NAMD software)
- Clustering over backbone dihedrals of the β3–β4 loop, 14 principal components (EnGens service)

Distribution of φ(Gly507) and φ(Asn506) over MD trajectories



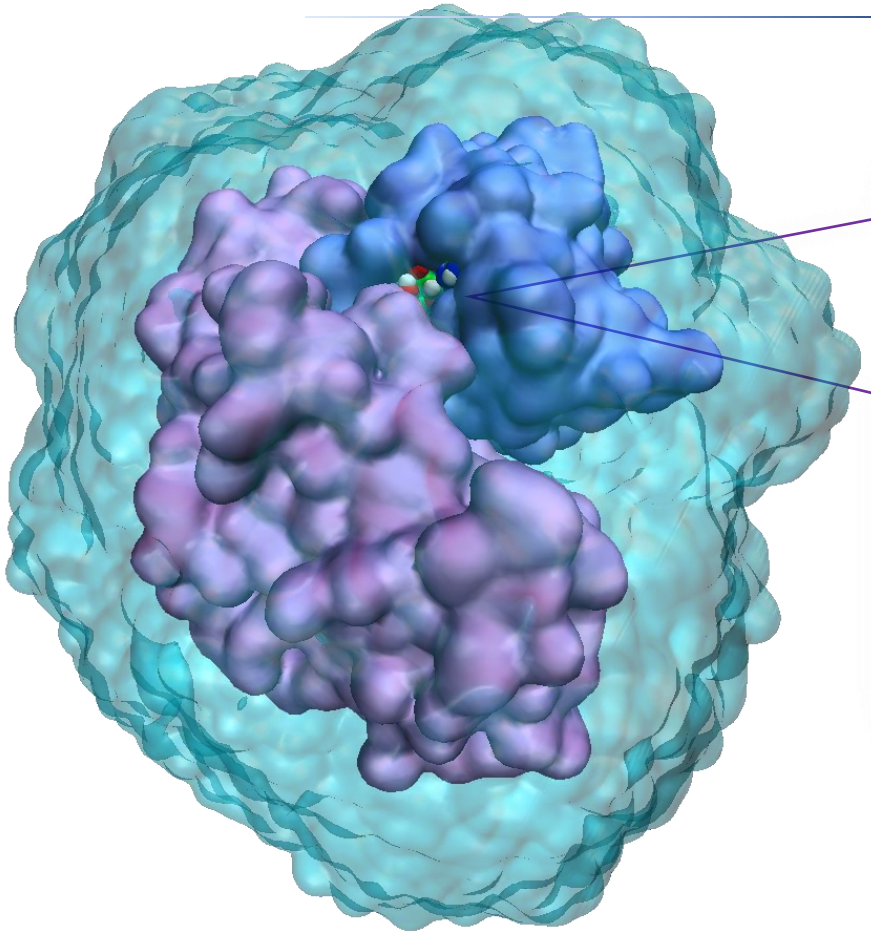


Increase of ceftriaxone resistance is due to both changes of the acylation rate constant and binding constant

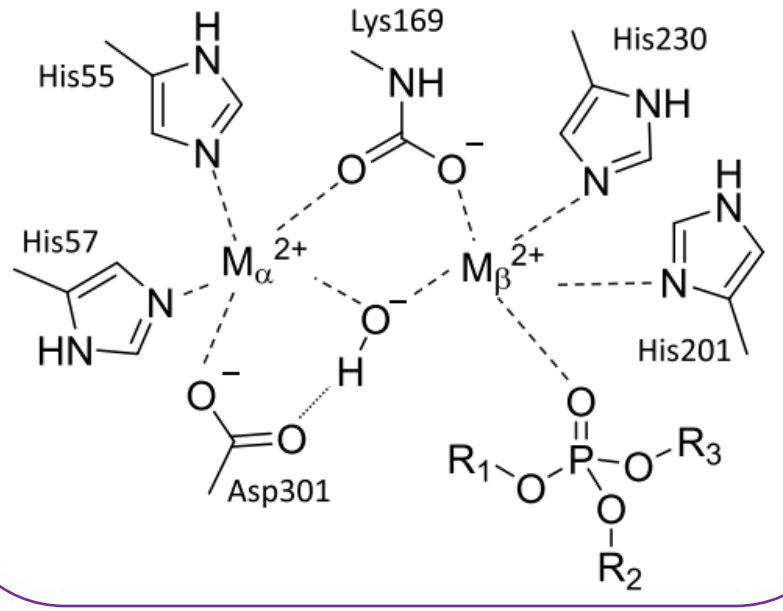


Catalytic efficiency and dynamic behavior of enzyme-substrate complexes: metallo-β-lactamases NDM-1 and L1

Combined quantum mechanics / molecular mechanics method



Active site



<u>System size</u>

QM part:

100 – 200 atoms

MM part:

10^3 – 10^5 atoms



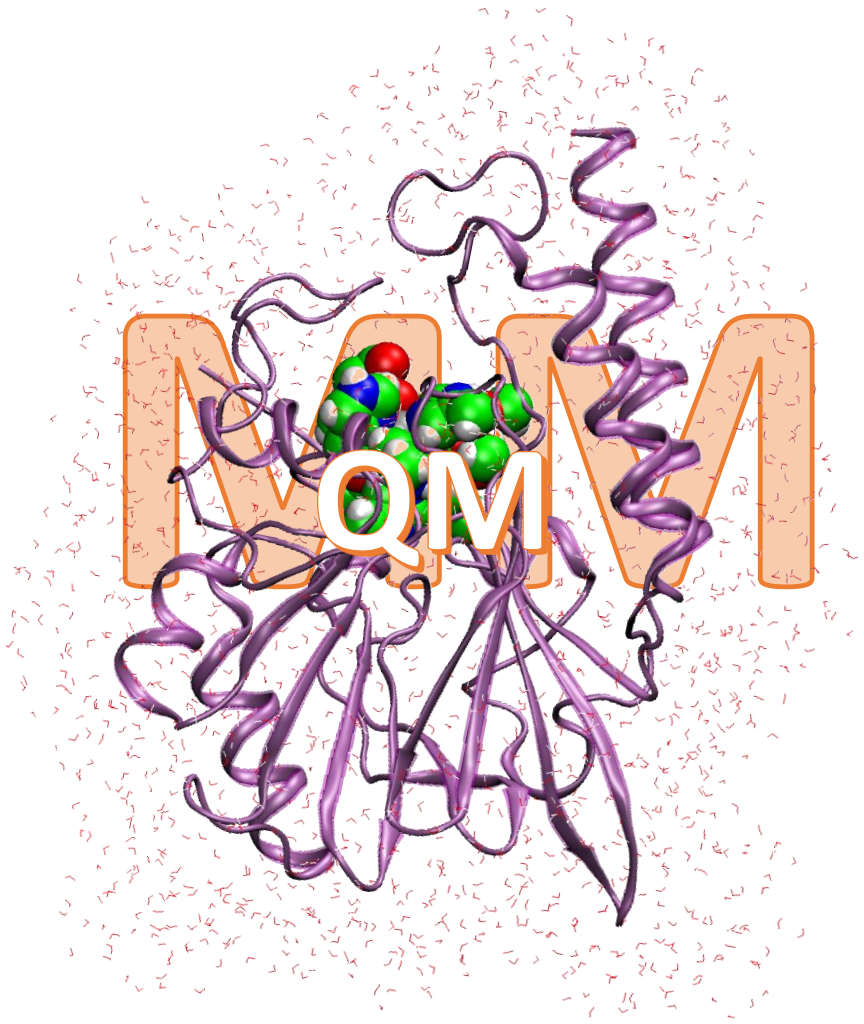
2013

<u>Methods</u>

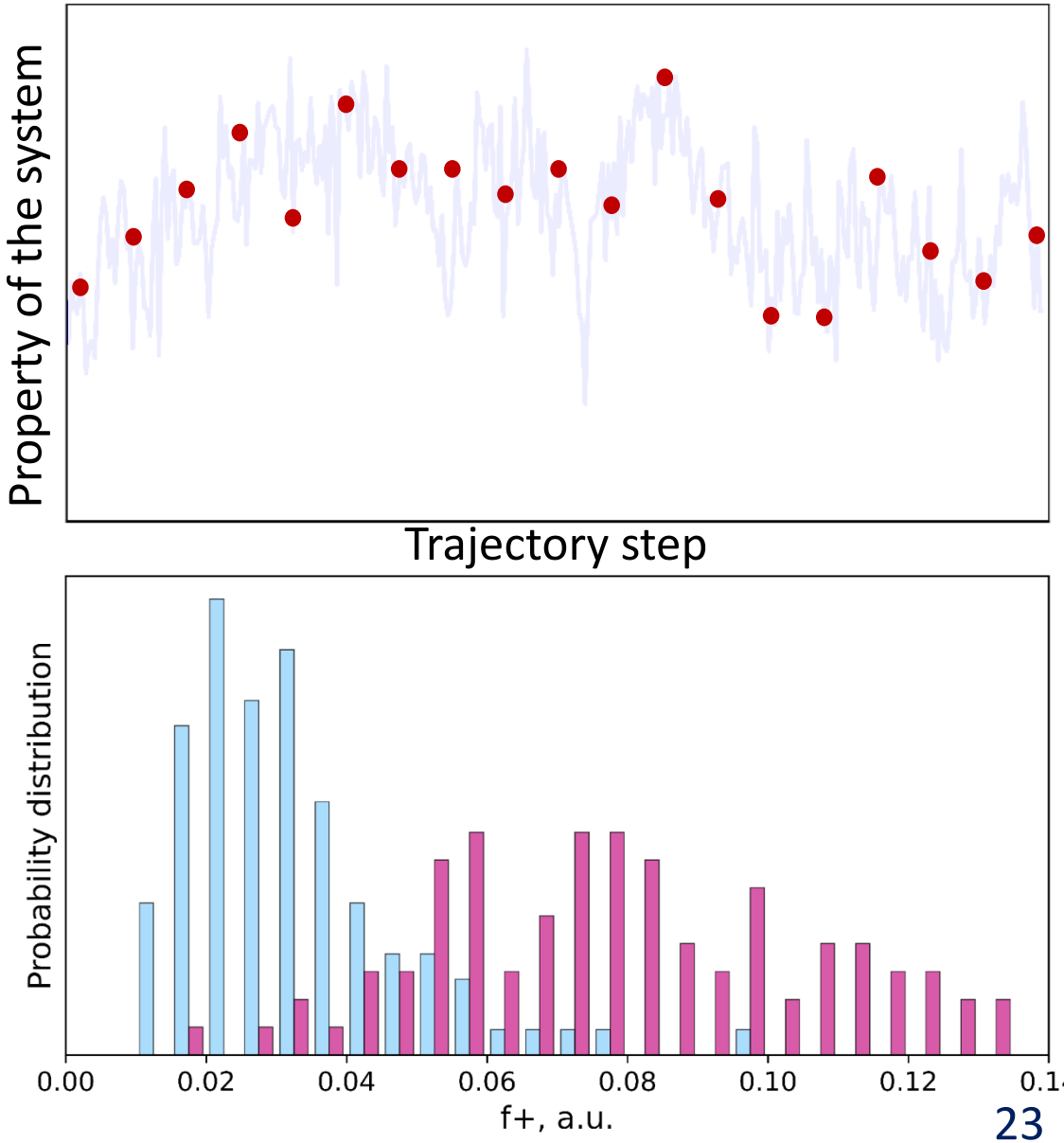
QM subsystem: quantum chemistry methods – DFT, HF…

MM subsystem : classical force fields

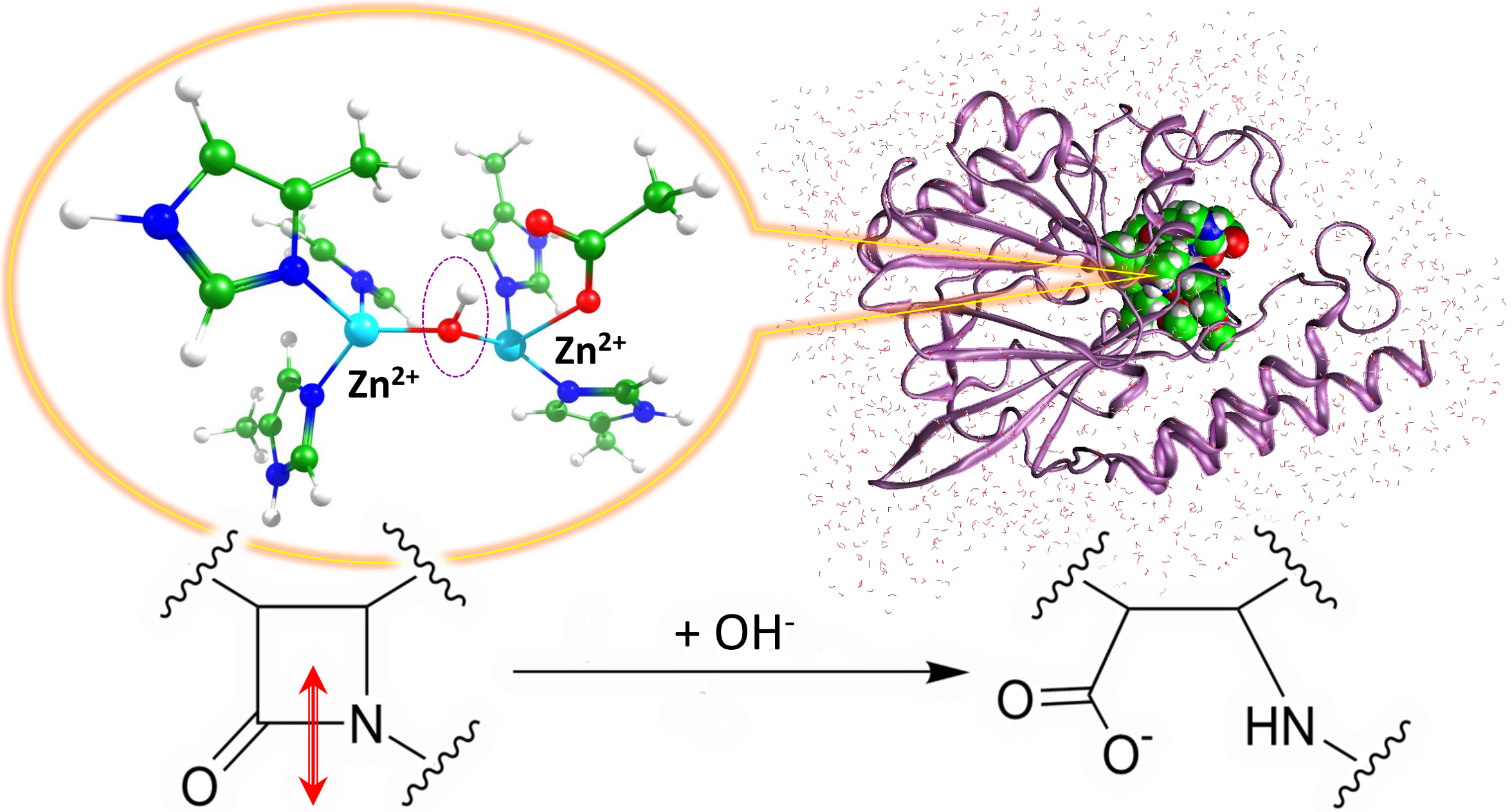
From local minima to trajectories



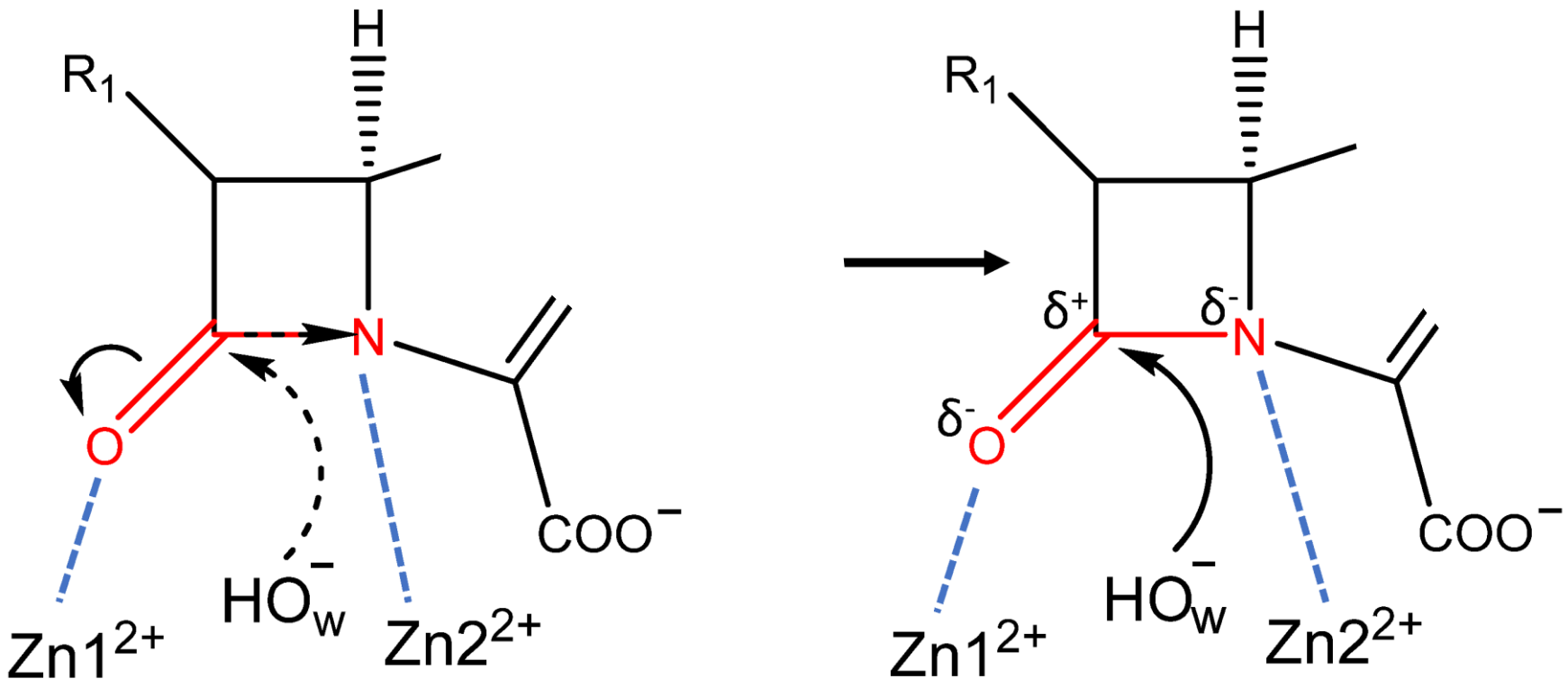
MD



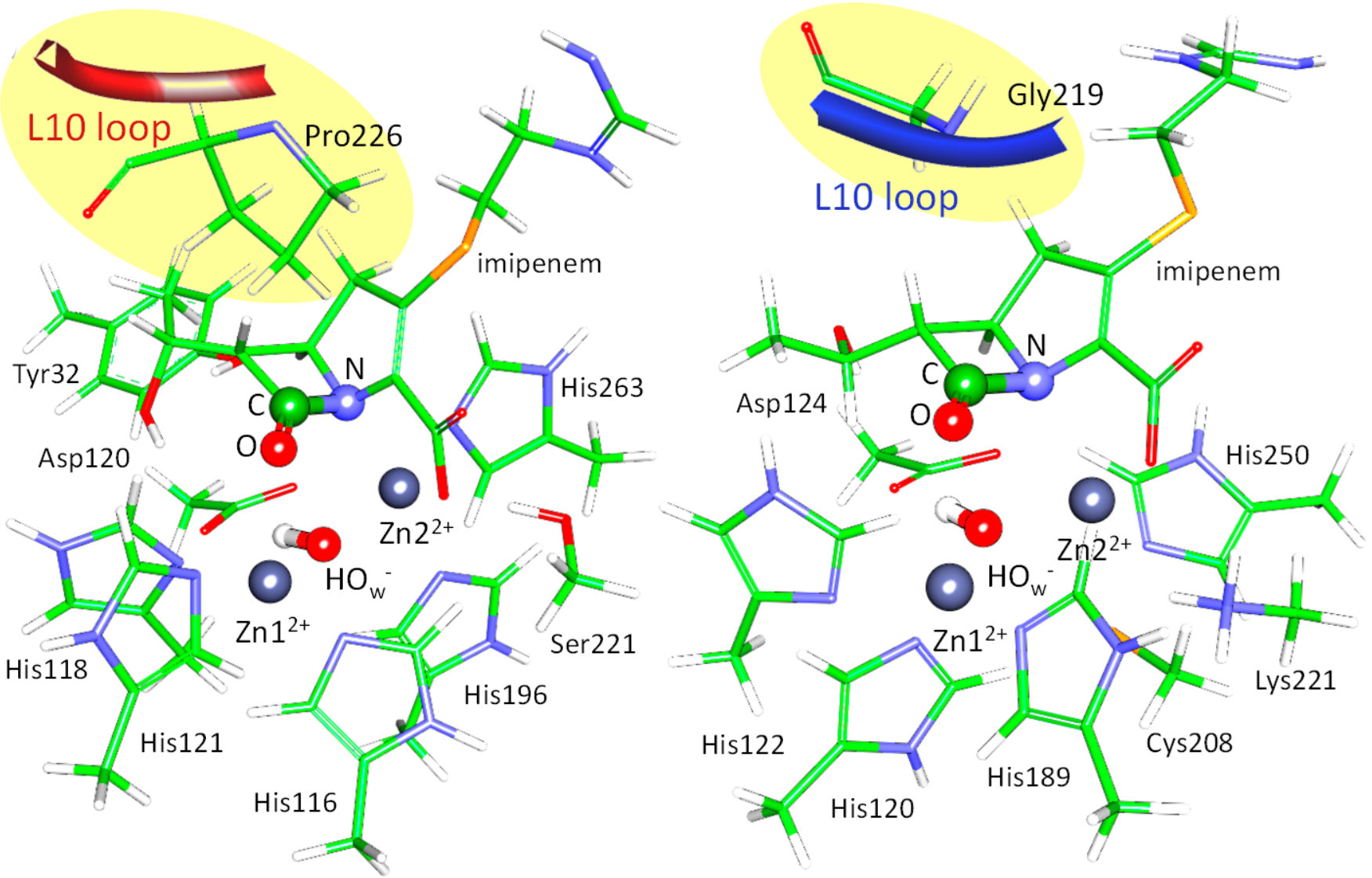
Hydrolysis mechanism in the active site of metallo-β-lactamases



β-lactam activation in the active site of metallo-β-lactamases

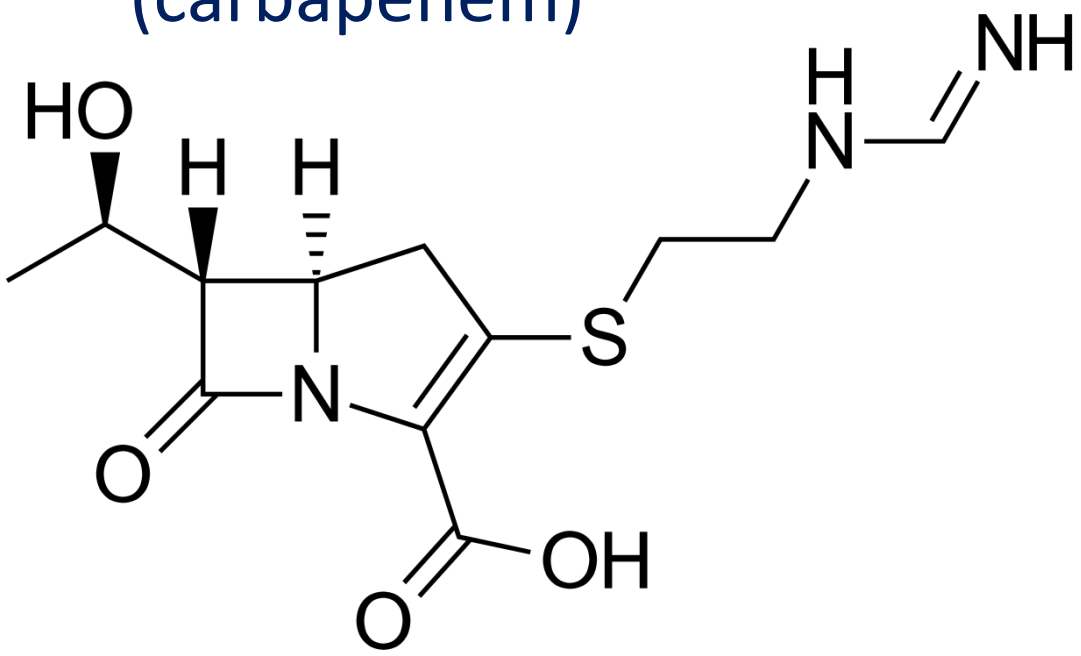


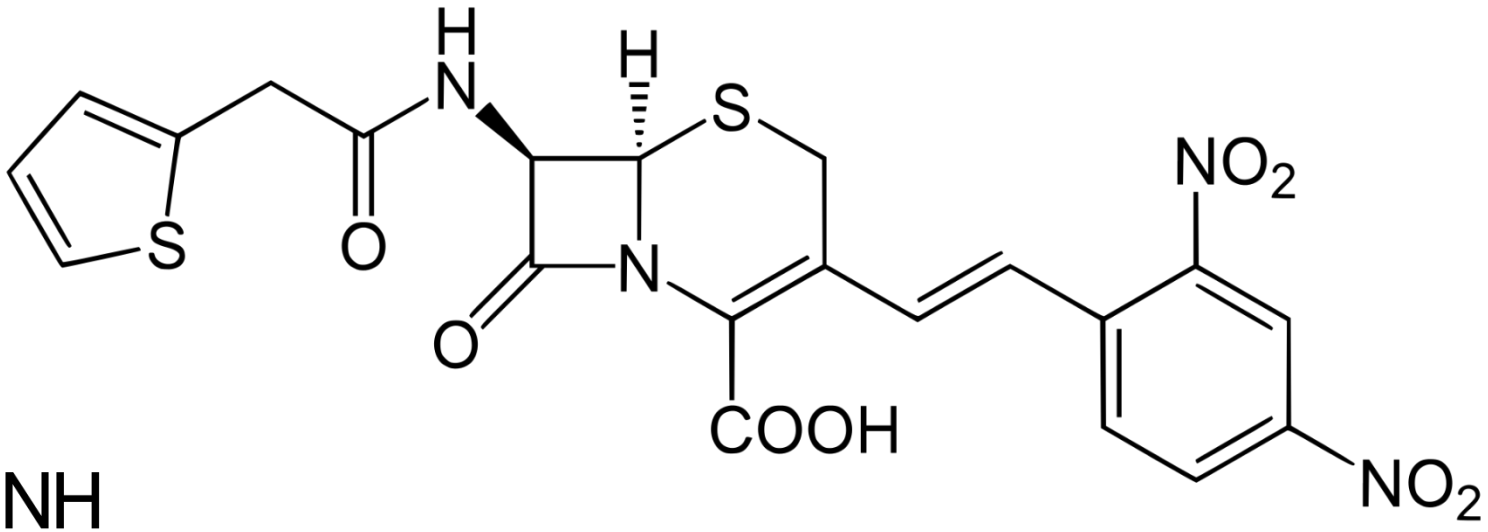
Active sites of metallo-β-lactamases NDM-1 and L1



Substrates from different classes

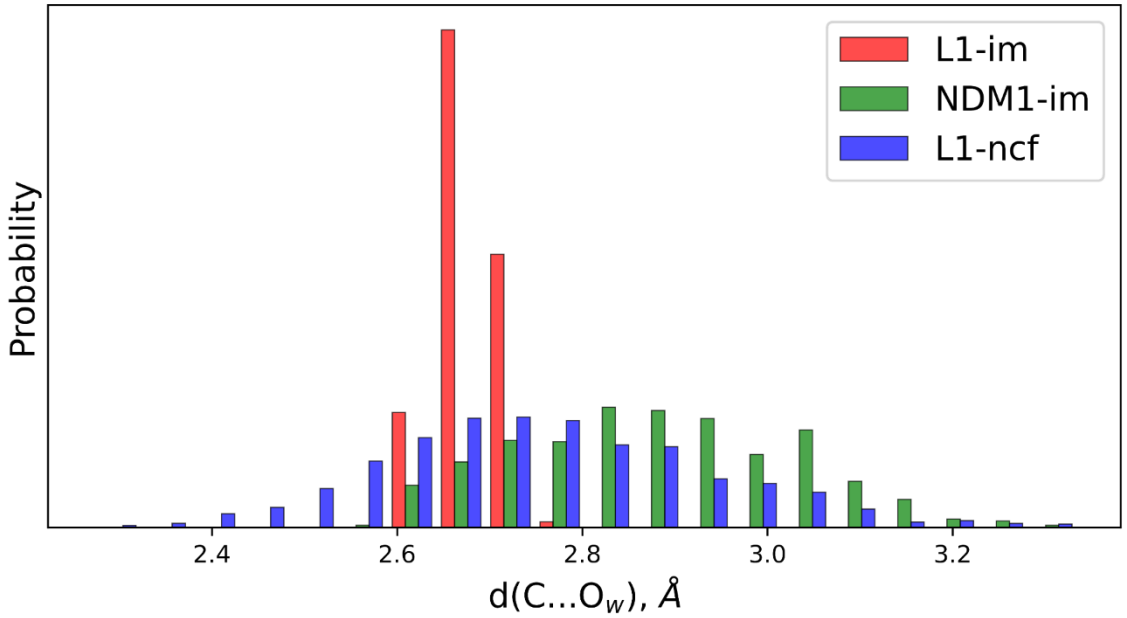
Imipenem
(carbapenem)



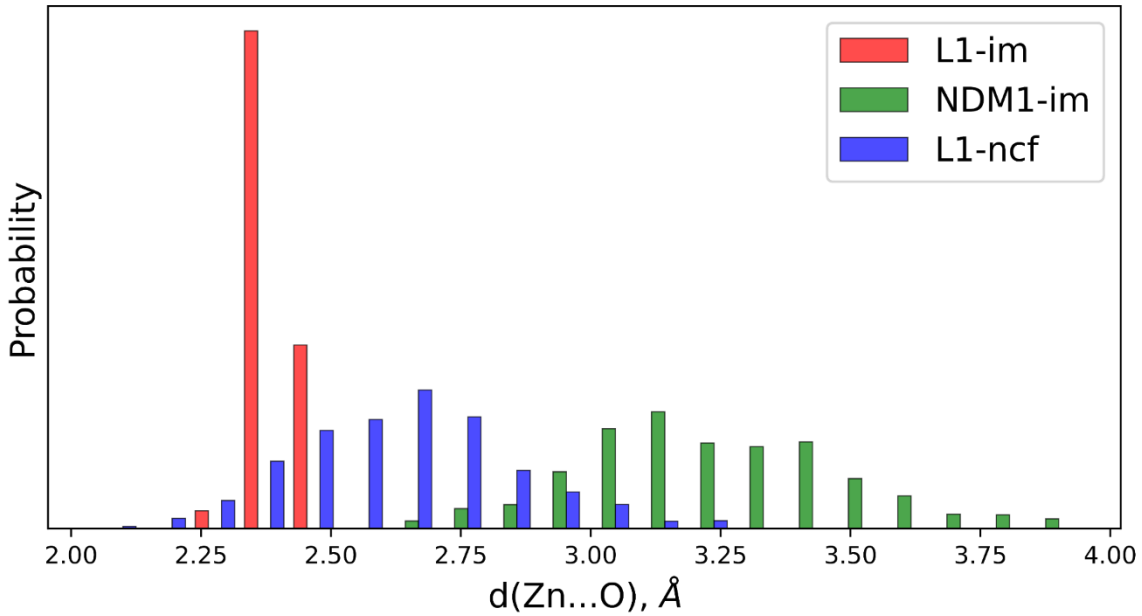


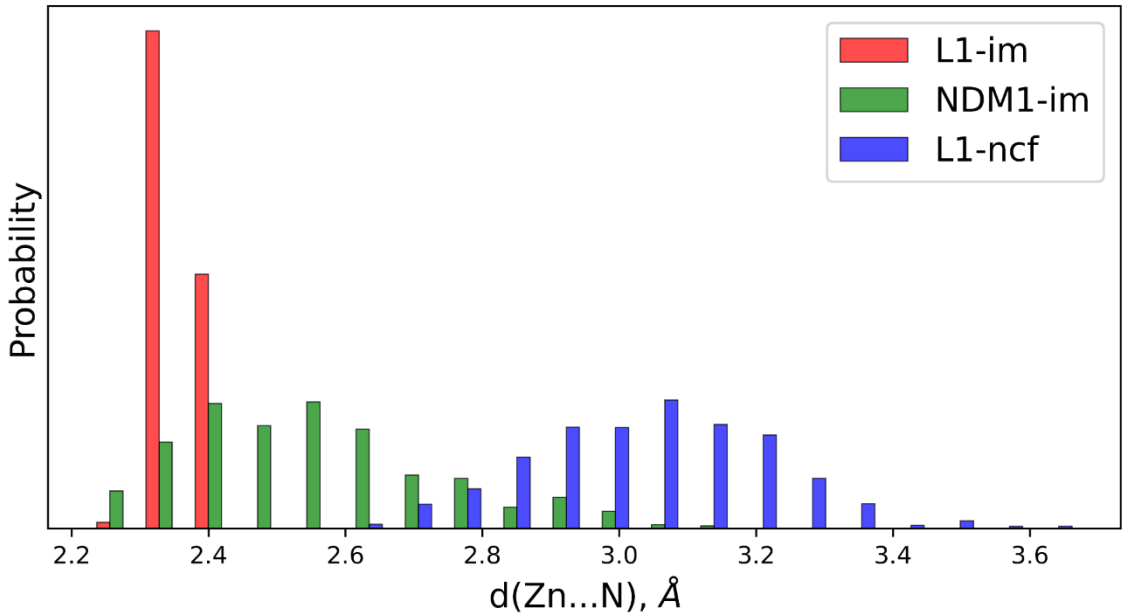
Nitrocefin
(cephalosporin)

ES complex dynamics



	k_{cat}, s^{-1}
L1 – imipenem	384
L1 – nitrocefin	35
NDM-1 - imipenem	64



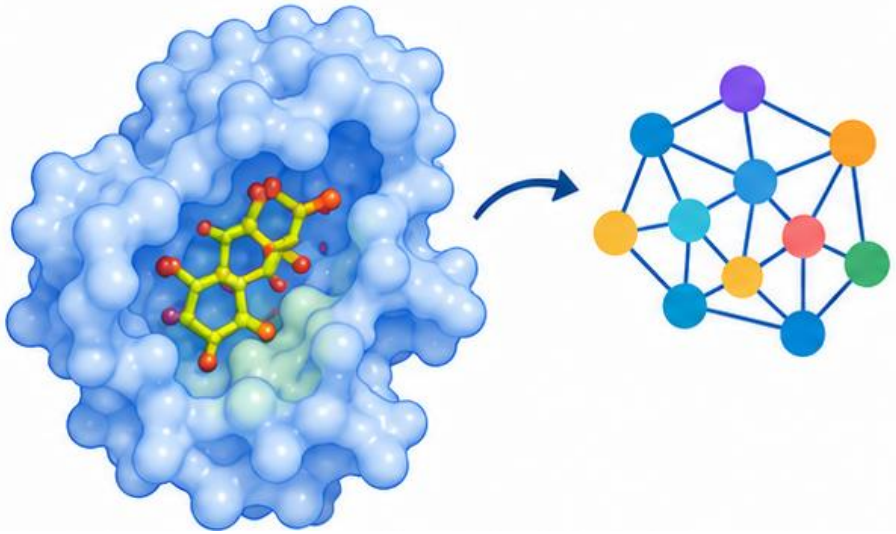


Concluding remarks

QSPR provides a powerful framework
for linking molecular properties and dynamics with macroscopic effects such as drug resistance.

Molecular descriptors and dynamic characteristics
allow the prediction of binding affinity, activity and resistance using regression and classification models.

Network and community analysis
reveals how molecular interactions and dynamic behavior affect function and contribute to resistance mechanisms.



Thank you for your attention!