b-it Bonn-Aachen International Center for Information Technology

Molecular modeling and drug design

Docking – Cathepsin K

Danqi Wang

05 Aug 2021

Outline

- Information about Cathepsin K
- Structure selection, analysis and preparation
- Self-docking
- Benchmarking study
- Cross-Docking
- Conclusion and discussion

Cathepsin K(CatK)

- **CatK:** One of cysteine proteases

  Highly expressed in osteoclasts

  Unique and potent collagenase activity[1].
- **Function**

- In Bone: osteoclast-mediated bone resorption without perturbing the bone formation [2].
- Beyond Bone: Central Nervous System (CNS)[3], Cardiovascular System[4], Respiratory System[5]

Cathepsin K(CatK)

- **Druggability**
  - **Disease related protein**
    - Osteoporosis: Excess osteoclast activation and bone resorption
    - Inhibitor: Anti-osteoporosis drug [6]
  - **Active site**
    - V-shaped cleft located at the top of the molecule and contains the catalytic triad Cys^{25}, His^{159},Asn^{175}.

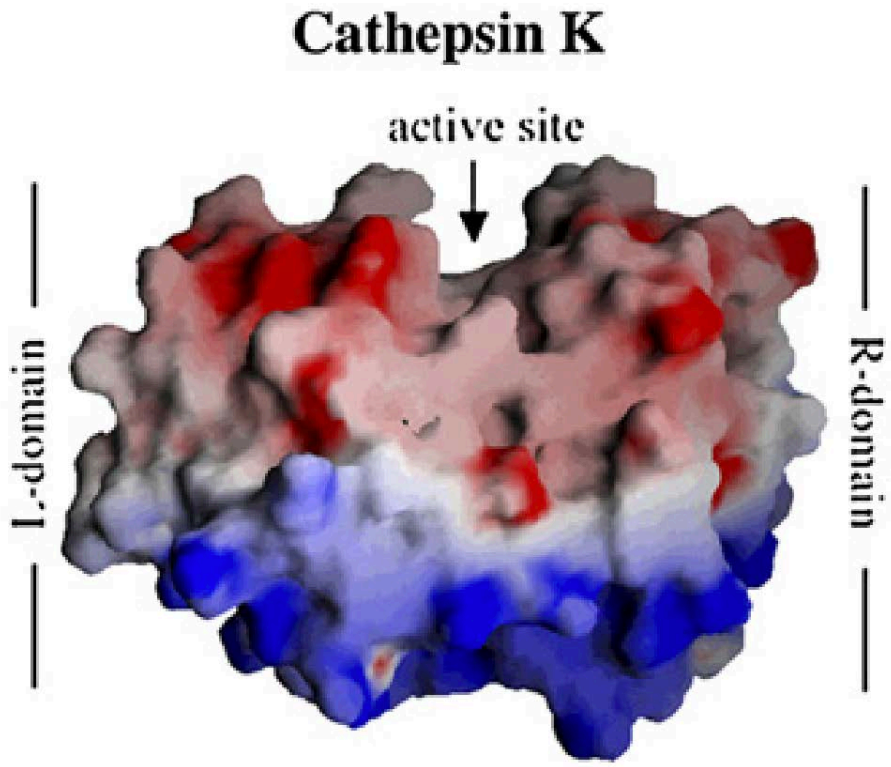


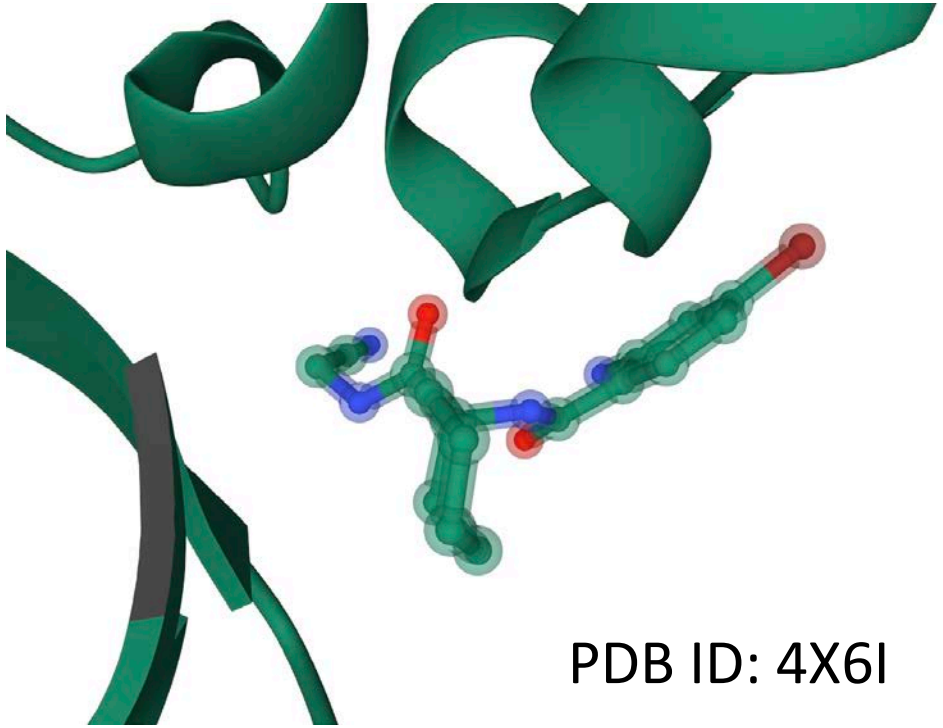
Fig 1. Surface potential of Cathepsin K [7].
Negative potential are colored red and positive potential is blue.

Inhibitors of CatK

- **No FDA approved drug till now[1]**
  - Previously used drugs in clinal trial:
    - Odanacatib (ODN) developed by Merck & Co[8].
    - Balicatib(AAE581) developed by Novartis [9].

    } Terminate in clinical trial because of adverse effect
  - Undesirable CatK inhibition in non-bone sites[2]
- **Academic study**
  - Over one thousand candidate compounds target at cathepsin K (ChEMBL)

1. Structure selection



PDB ID: 4X6I

Fig 2. 3D presentation of protein-ligand complex

- Released time: 2015-09-30
- Method: X-ray diffraction
- High resolution: 1. 87Å
- **Small molecule: 3Y1 (**C_{16} H_{19} Br N_4 O_2)
- Ligand is accommodated in pocket
- Ligand is not covalently bound to residues

ID	Binding Affinity (Sequence Identity %)
3Y1	Ki: 10 nM Binding MOAD

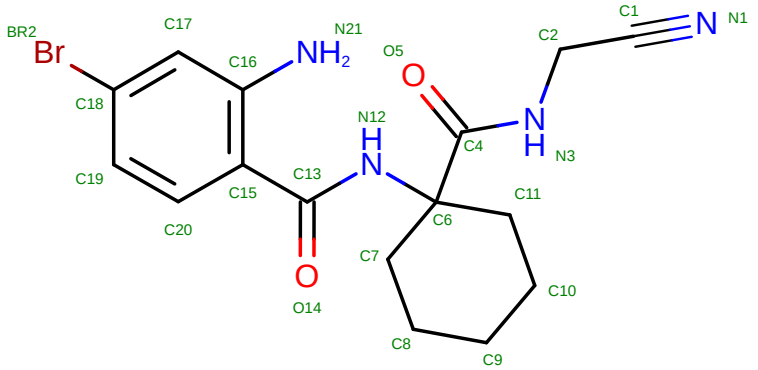


Fig 3. 2D Diagram of compound 3Y1

2. Structure analysis and preparation

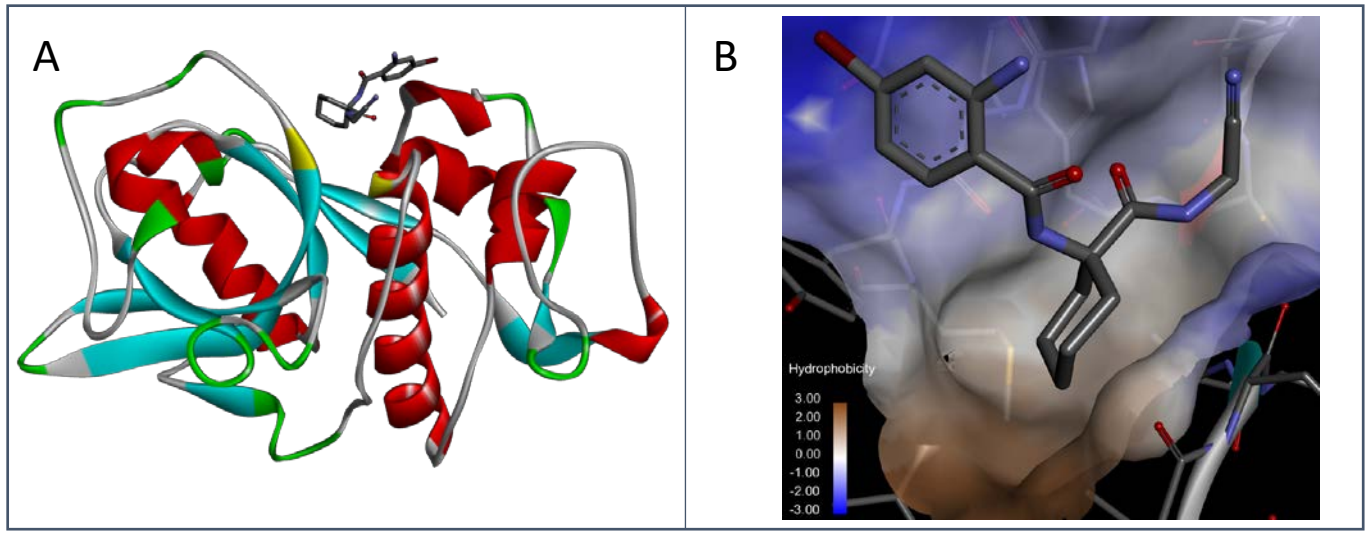


Fig 4. Ligand locates in active site (by Discovery Studio Visualizer)

- Separate receptor and ligand (ChimeraX)

  Delete water molecule, sulfate ion

  Export .*pdb* files for ligand and receptor

3. Self-docking

3.1 **Ligand preparation**

- Place hydrogens
- Default setting: amide bonds not to be rotatable;
  merges both non-polar hydrogens and lone pairs

3.2 **Receptor preparation**

- Place hydrogens
- Default setting: addition of gasteiger charges; remove lone pairs

3. Self-docking

3.3 **Docking**

- **Calculate the center of ligand**
  - Fix the coordinates of binding center
- **Call *vina.exe* in command line based on *configuration file***
  - Conformation search and scoring
  - *Given parameter: prepared receptor, ligand; center of ligand, searching area, number of modes, exhuastiveness, energy range*
- **Call vina_split.exe to generate each docking mode**

3. Self-docking

3.4 **Docking result** — Score vs. RMSD

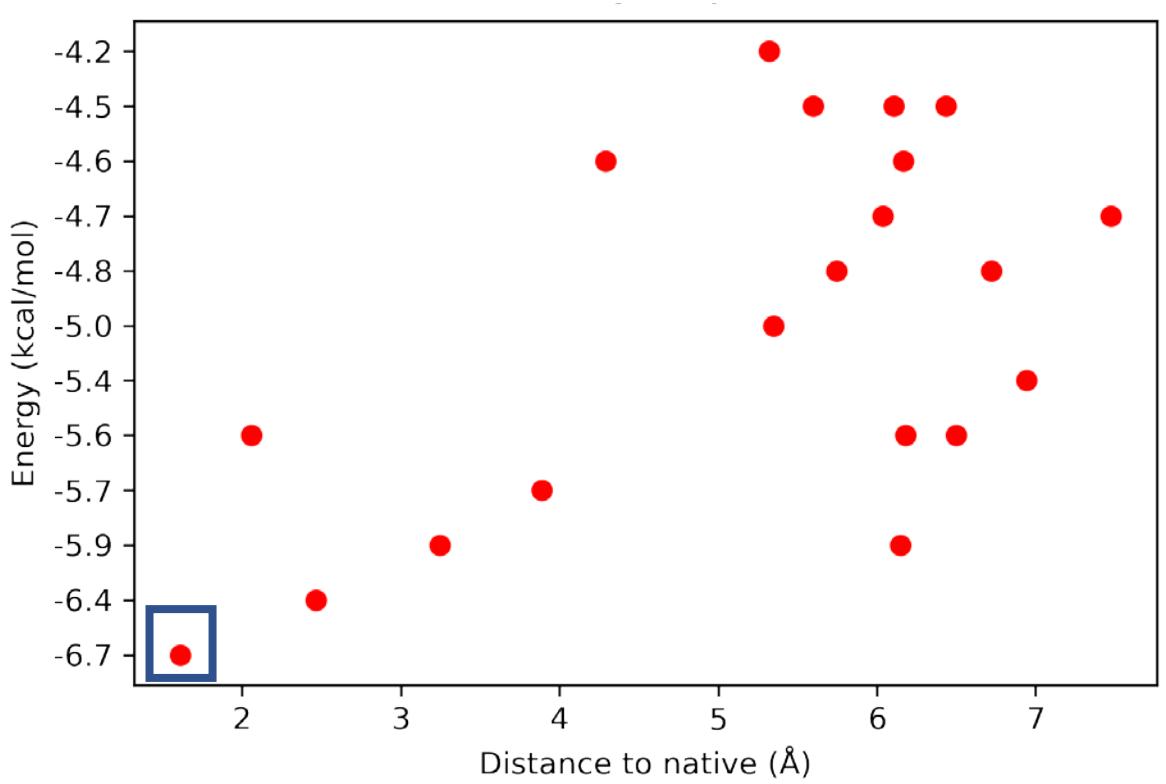
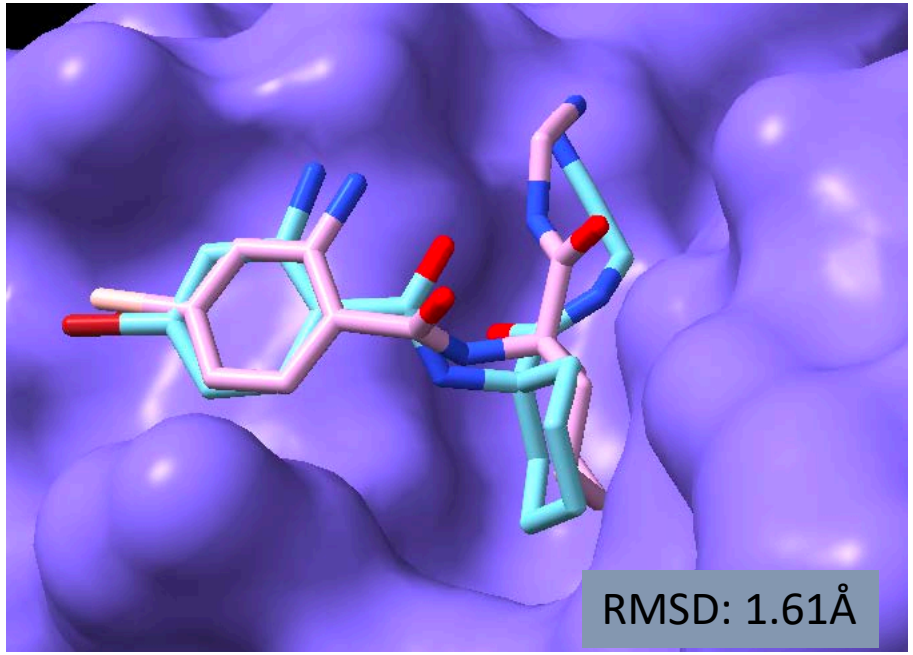


Fig 5. RMSD plot against the corresponding energy

- Best mode:
  RMSD 1.61Å
  Energy -6.7kcal/mol

3. Self-docking

3.5 **Visualization**



Pink: pose with highest score

vs.

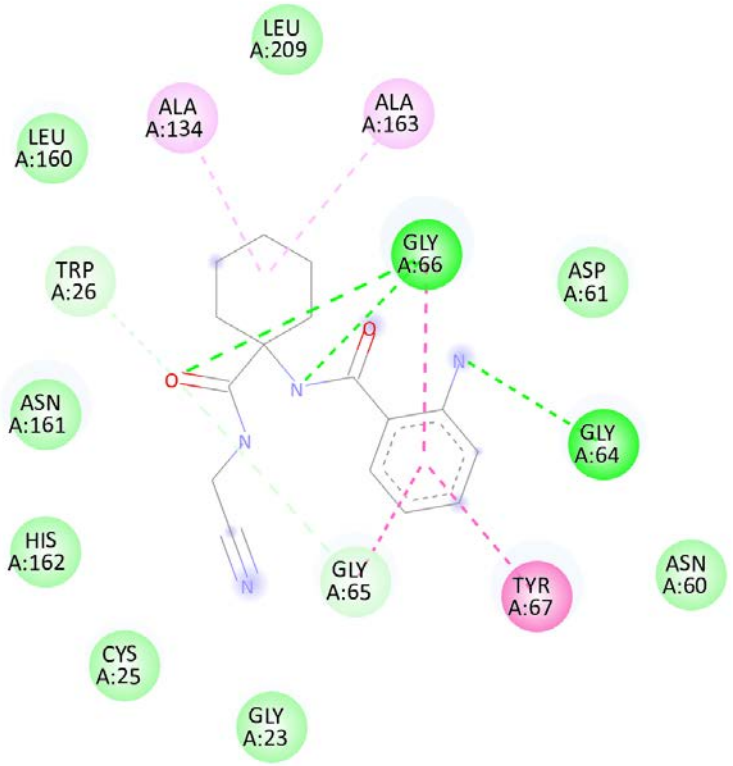
Blue: pose in crystal structure

Fig 6. Native conformer (blue) and re-docking conformer (pink) against 4K6I

3. Self-docking

3.6 **Receptor – ligand interactions**

A. Native pose



B. High score docking pose

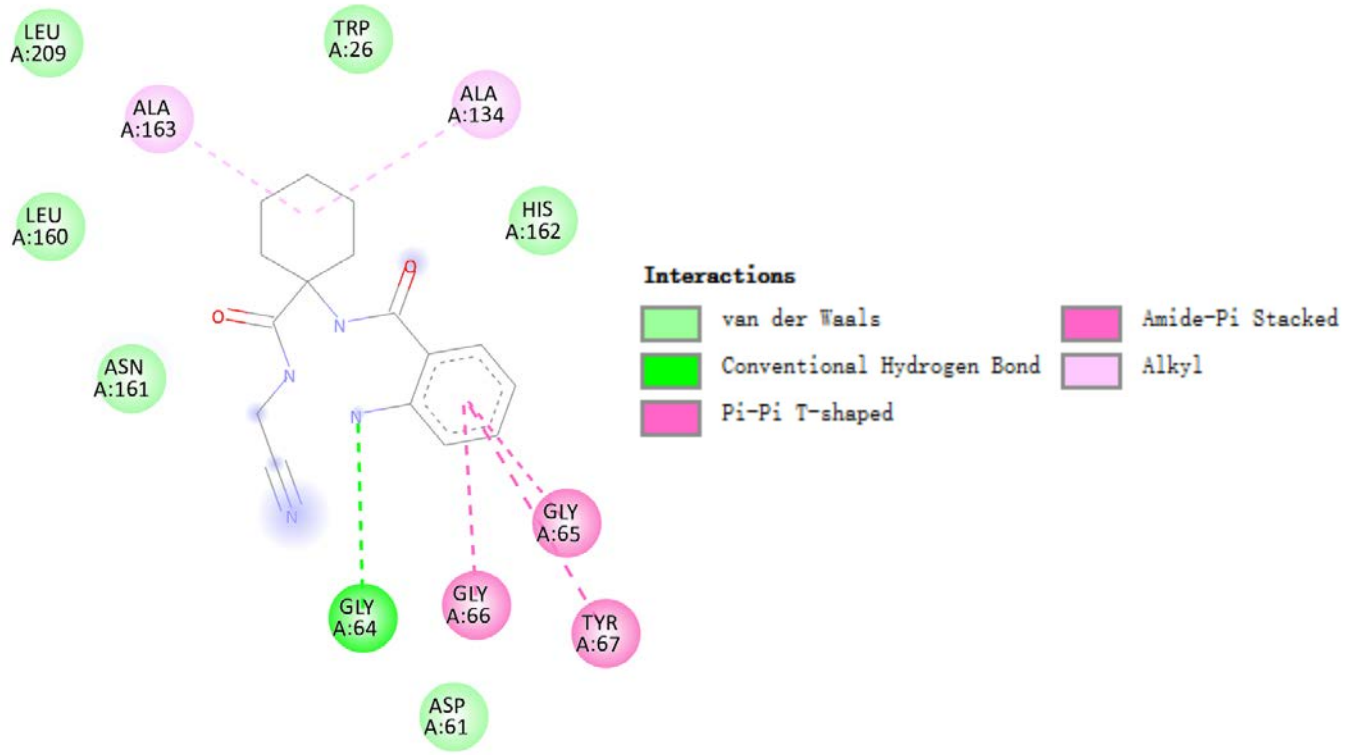


Fig 7. 2D Diagram of protein- ligand interactions

4. Benchmarking study

- 4.1 **Select 10 active compounds in ChEMBL**

Active compounds										
ID ChEMBL	481611	371064	191881	381715	378188	370139	361507	204079	30356	1087285
IC_{50}	0.2 nM	5.0 nM	0.018 µg/ml	48.0 nM	0.84 µg/ml	36.0 nM	7.9 nM	400 nM	38 nM	3.02 nM
Max Phase	3	2	-	-	-	-	-	-	-	-

4. Benchmarking study

- 4.2 **Select 10 inactive compounds in ChEMBL**

Inactive compounds										
ID CHEMBL	52333	3142836	3733999	1672424	1236936	274120	191588	165345	230479	3354496
pIC_{50}	-	-	-	-	< 4	< 4	< 4	< 4	< 4	< 4
Target at CatK	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Benchmarking study

- 4.3 Generate 3D conformation *(.pdb* file) of each compounds
- 4.4 Repeat docking
- 4.5 Select the pose with highest score of each ligand

4. Benchmarking study

- 4.6 Comparison between active and inactive compounds

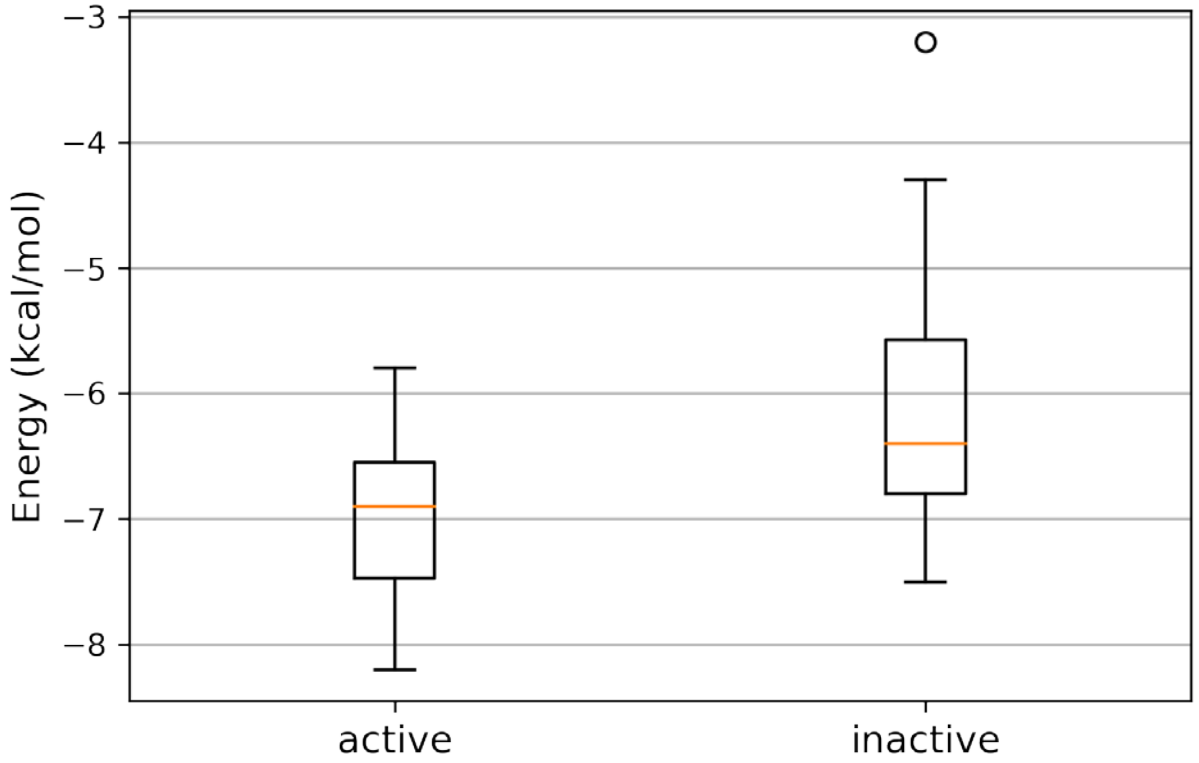


Fig 8. The comparison of scores of active and inactive compounds with best pose

4. Benchmarking study

- 4.7 Receiver operating characteristic (ROC) curve

# Matches	FPR	TPR
1	1.0	1.0
2	0.9	1.0
3	0.8	1.0
4	0.8	1.0
5	0.7	1.0
6	0.5	0.8
7	0.3	0.7
8	0.2	0.4
9	0.0	0.1
10	0.0	0.0

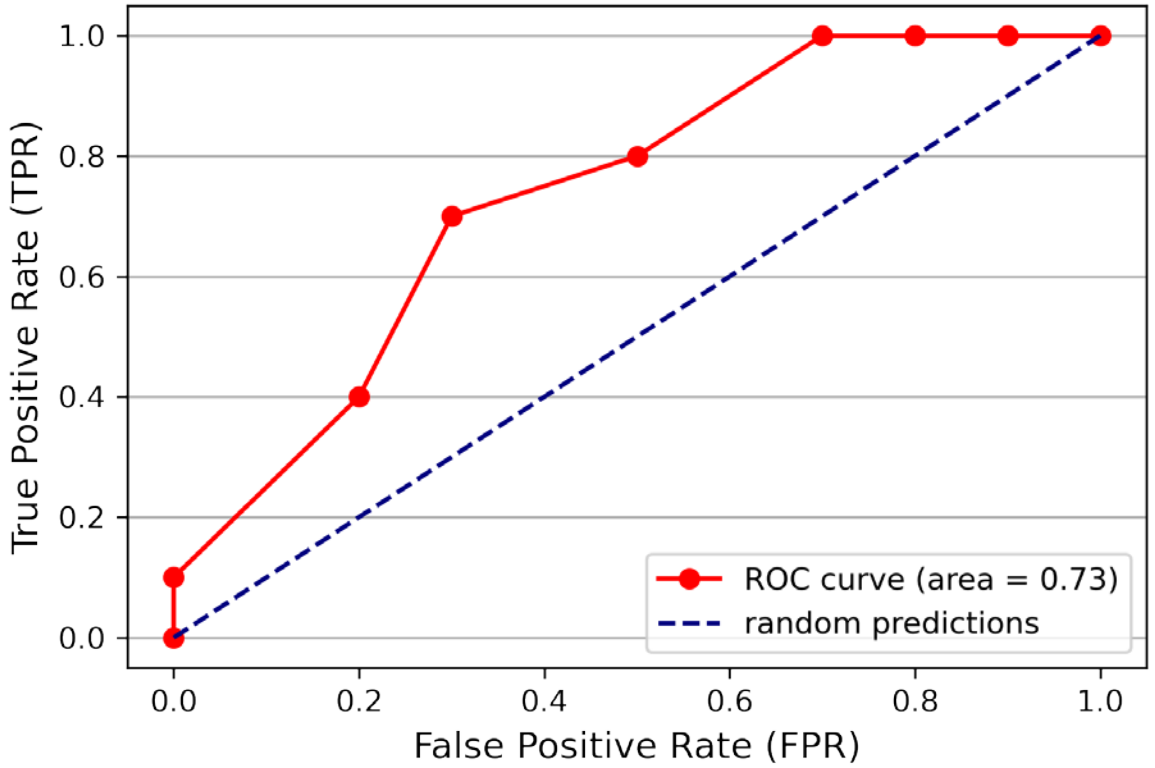


Fig 9. ROC curve of benchmarking study

5. Cross-Docking

- **Why to perform cross-docking?**
  - Receptor conformation changes by ligand
  - Estimation for performance on native-ligand
- **Procedure:**
  - Select new model for cross-docking:

	Self -docking	New structure used for cross docking
Target	4X6I (1.87Å)	3KX1 (1.51Å)
Ligand	3Y1	KX1

  - Superpose protein structure (RMSD: 0.324), get separate *.pdb* files
  - Repeat docking procedure

5. Cross-Docking

- 5.1 Score vs. RMSD plot

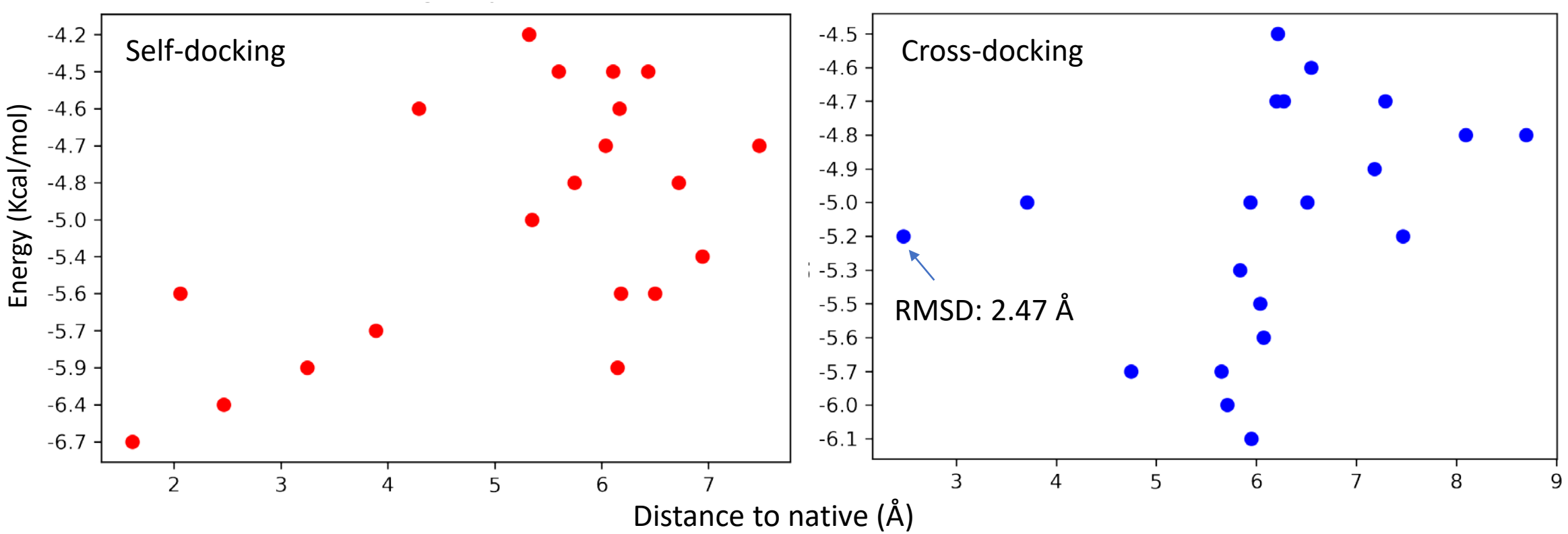


Fig 10. Scatter plot of RMSD against energy

5. Cross-Docking

- 5.2 **Visualization**

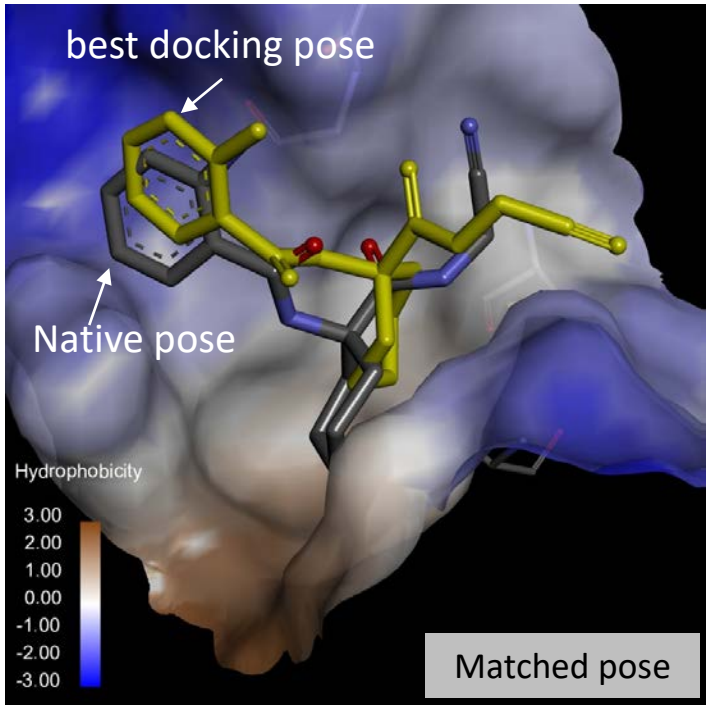
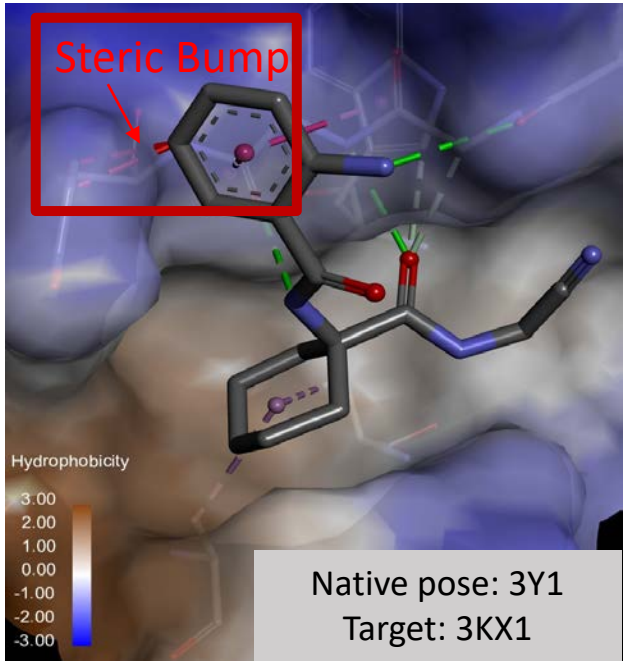


Fig 11. Native pose and best docking pose in the active site of new target 3KX1

5. Cross-Docking

- 5.3 Unfavorable bump in the native pose of ligand to new target



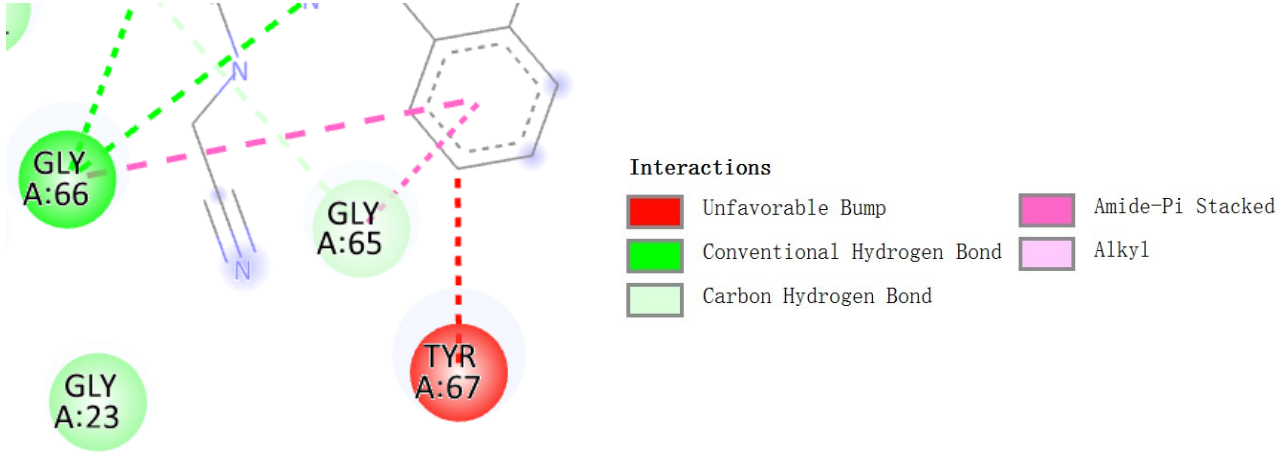


Fig 12. protein-ligand interaction of native pose (3Y1) to new target 3KX1

6. Conclusion and discussion

- 1. **Self docking: reliable posing mode**
  - Pose with high score: lowest RMSD
  - Docking is effective to find best binding pose conformation and orientation.
- 2. **Benchmarking study**
  - Force-field scoring function cannot discriminate active and inactive poses.
  - AUROC score: 0.73
    - Increase the magnitude of test size
    - Optimize parameter of algorithm (i.e. energy range, scoring function)
- 3. **Cross-docking**
  - Conformational change results to the change in thermodynamics, and will effect the docking result.

Reference

- [1] Rongchen Dai *et al*. Cathepsin K: The action in and beyond bone. Front. Cell Dev. Biol. 2020 ;89:127-134.
- [2] Fabien Lecaile *et al.* Biochemical properties and regulation of cathepsin K activity. Biochimie. 2008; 90:208-226.
- [3] Bernstein HG *et al.* Cathepsin K and schizophrenia. 2007 ;61(4):252-3.
- [4] Lange AW *et al.* NFATc1 expression in the developing heart valves is responsive to the RANKL pathway and is required for endocardial expression of cathepsin K. *Dev Biol.* 2006; 292(2):407-17.
- [5] Bühling F. *et al.*Lysosomal cysteine proteases in the lung: role in protein processing and immunoregulation. *Eur Respir J.* 2004; 23(4):620-8.
- [6] Matthew T Drake et al. Cathepsin K Inhibitors for Osteoporosis: Biology, Potential Clinical Utility, and Lessons Learned. Endocrine Rreviews, 2017; 38(4):325-350.
- [7] M.E. McGrath, J.L. Klaus, M.G. Barnes, D. Bromme, Crystal structure of human cathepsin K complexed with a potent inhibitor, Nat. Struct. Biol. 1997;4:105-109.
- [8] Tanaka M *et al.* Antiresorptive effect of a cathepsin K inhibitor ONO-5334 and its relationship to BMD increase in a phase II trial for postmenopausal osteoporosis.*BMC Musculoskelet Disord.* 2017; 18(1):267.
- [9] Peroni A *et al*. Drug-induced morphea: report of a case induced by balicatib and review of the literature. *J Am Acad Dermatol.* 2008; 59(1):125-9.

Thanks for your attention!