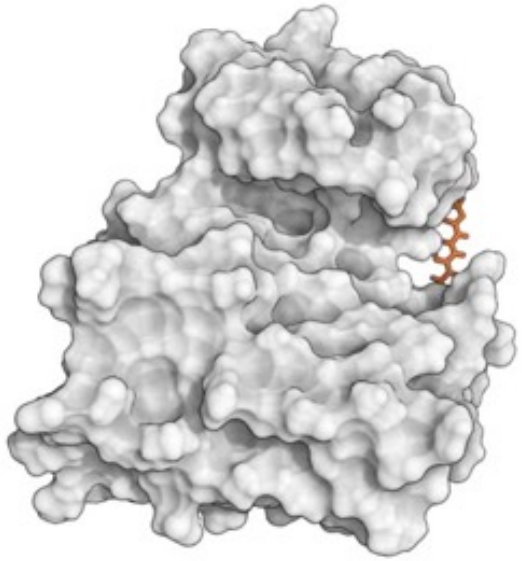
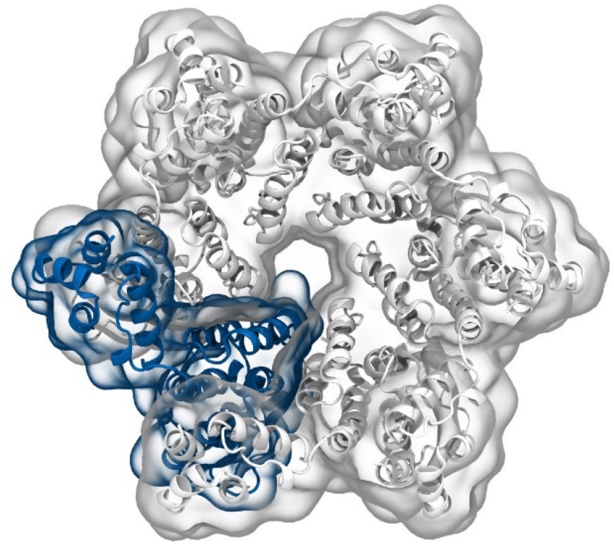
Simulation of Biomolecules



Docking

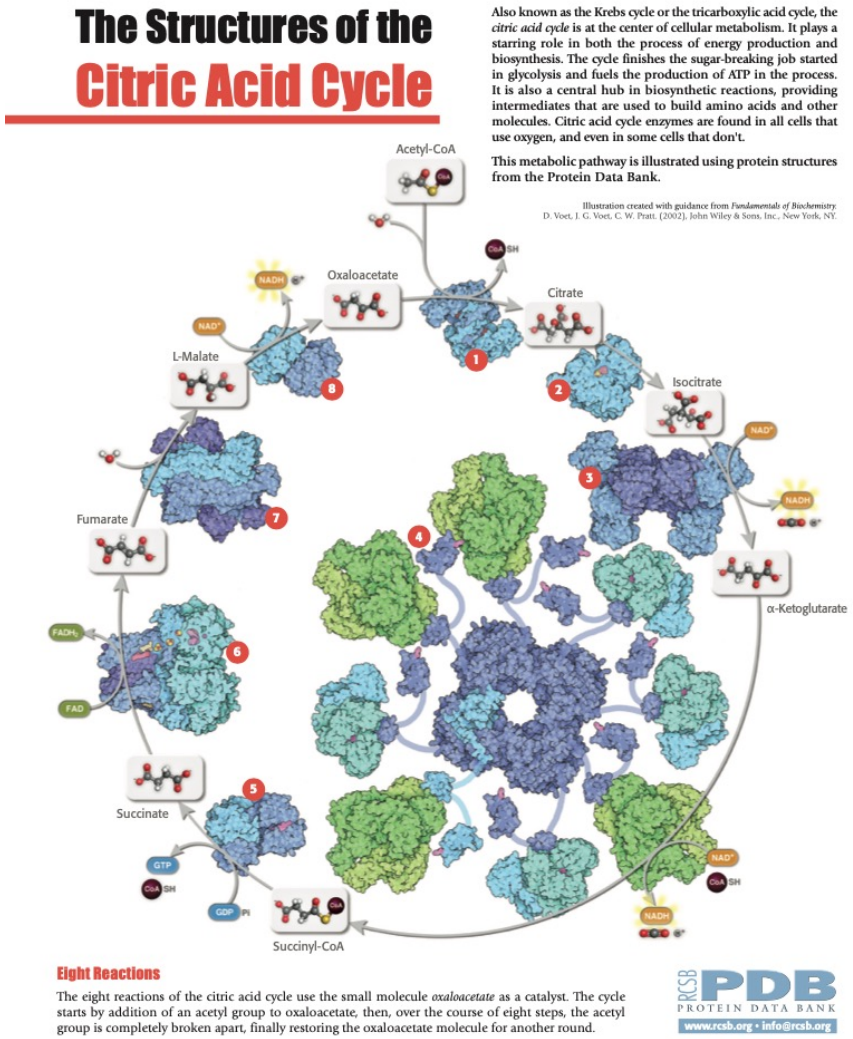
2024 CCP5 Summer School



Dr Matteo Degiacomi
Durham University
matteo.t.degiacomi@durham.ac.uk

Dr Antonia Mey
University of Edinburgh
antonia.mey@ed.ac.uk

Life is built on protein and small molecule interactions



Small molecules are:

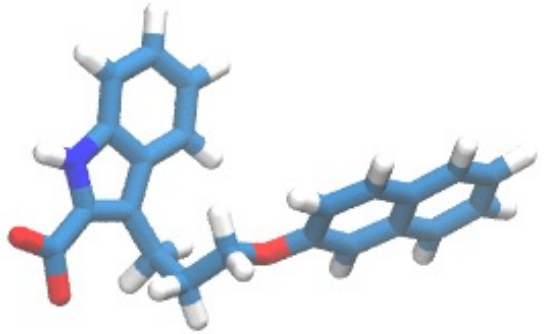
- substrates of enzymes
- Inhibitors or activators
- Co-factors

And play an important role in life.
Accurate interaction prediction is essential.

What is docking?

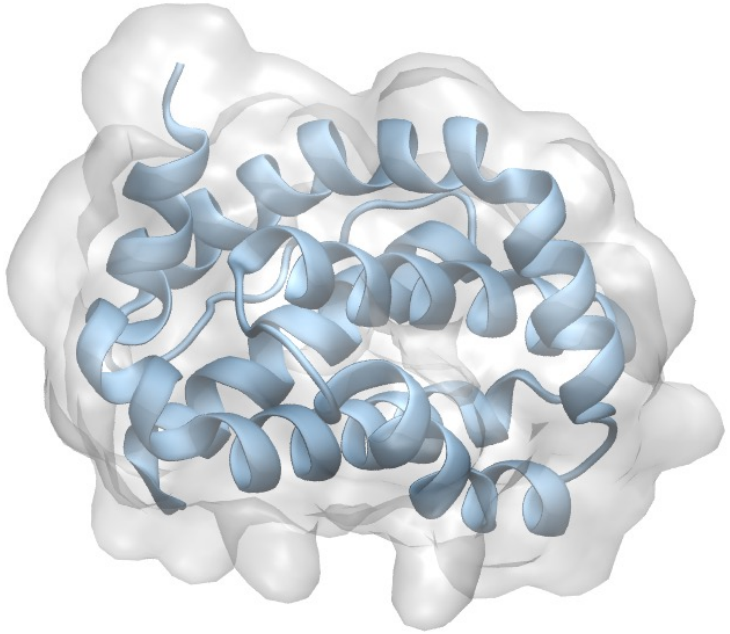
The process of predicting a stable 3D geometry of an interacting pair of molecules – **a binding mode/pose.**

Ligand:



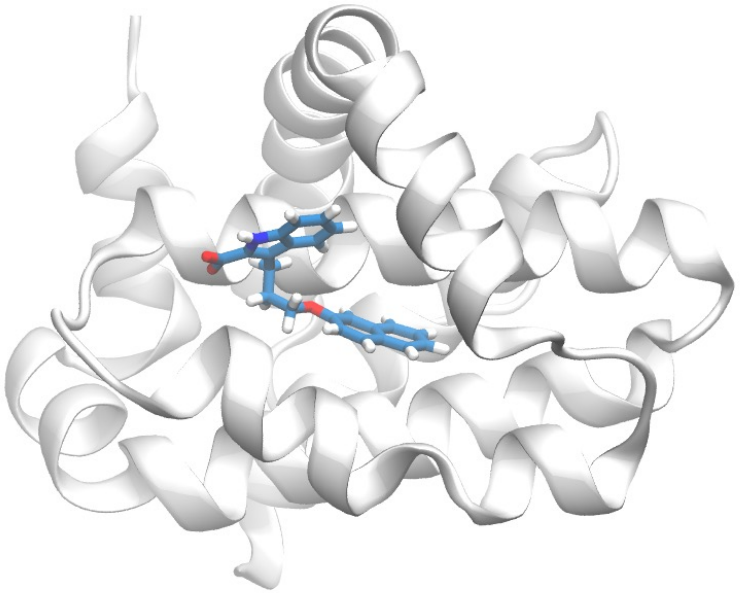
+

Receptor:



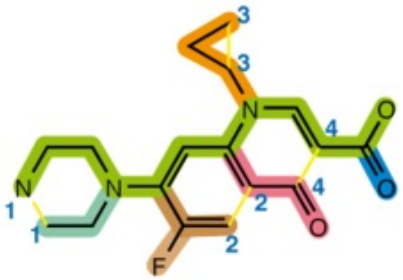
=

Binding pose:



Typical workflow

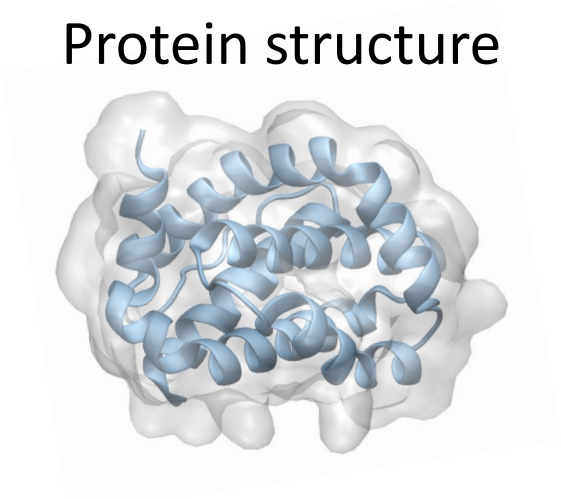
1-D or 2-D ligand structure

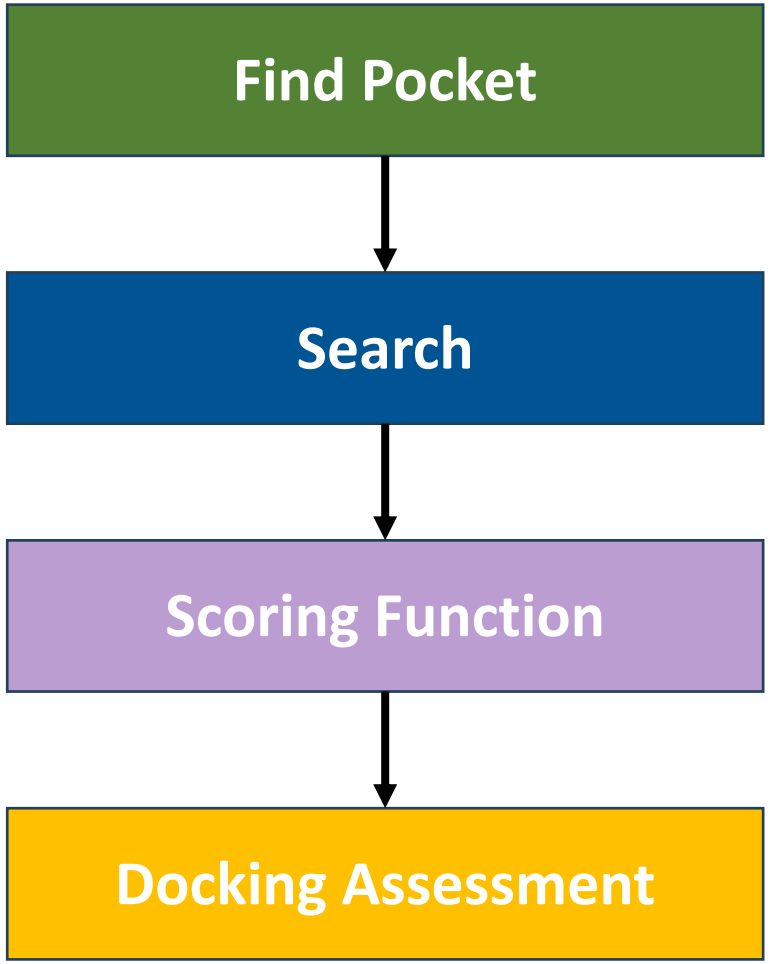


N1CCN(CC1)C(C(F)=C2)=CC(=C2C4=O)N(C3CC3)C=C4C(=O)O

+

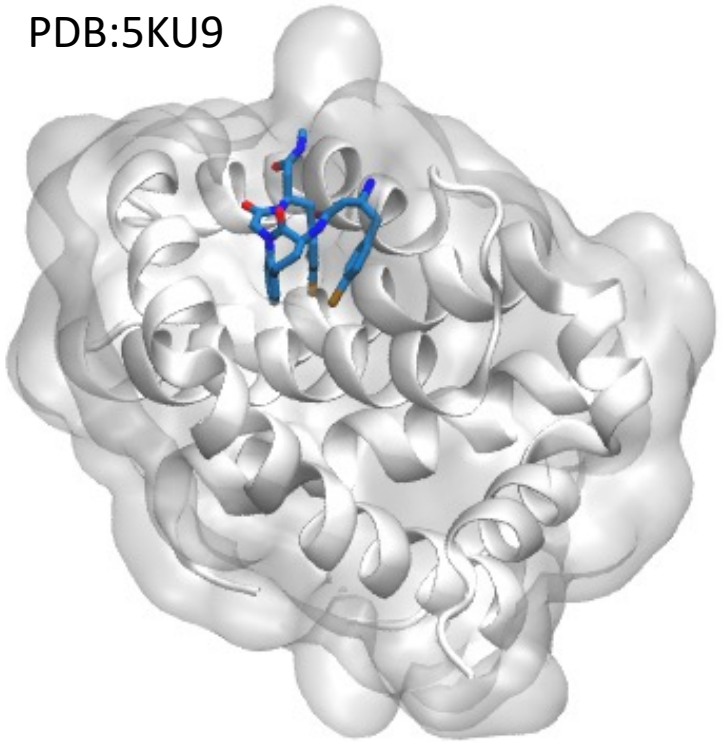
Protein structure



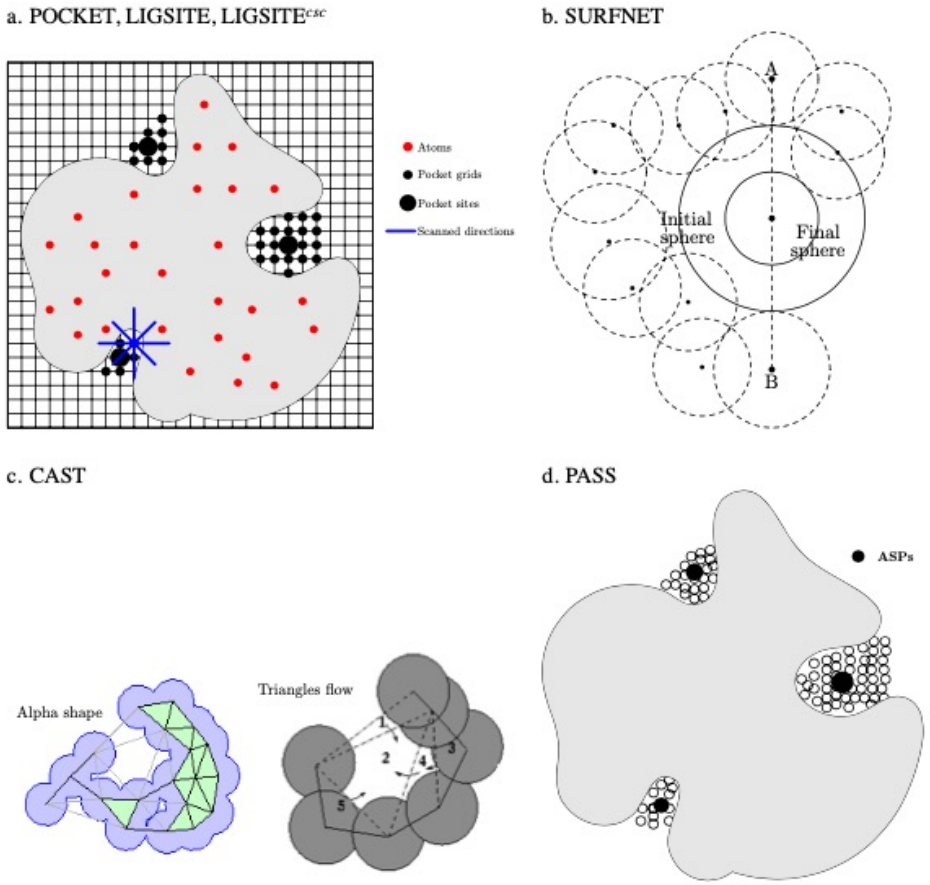


Finding a pocket

PDB:5KU9

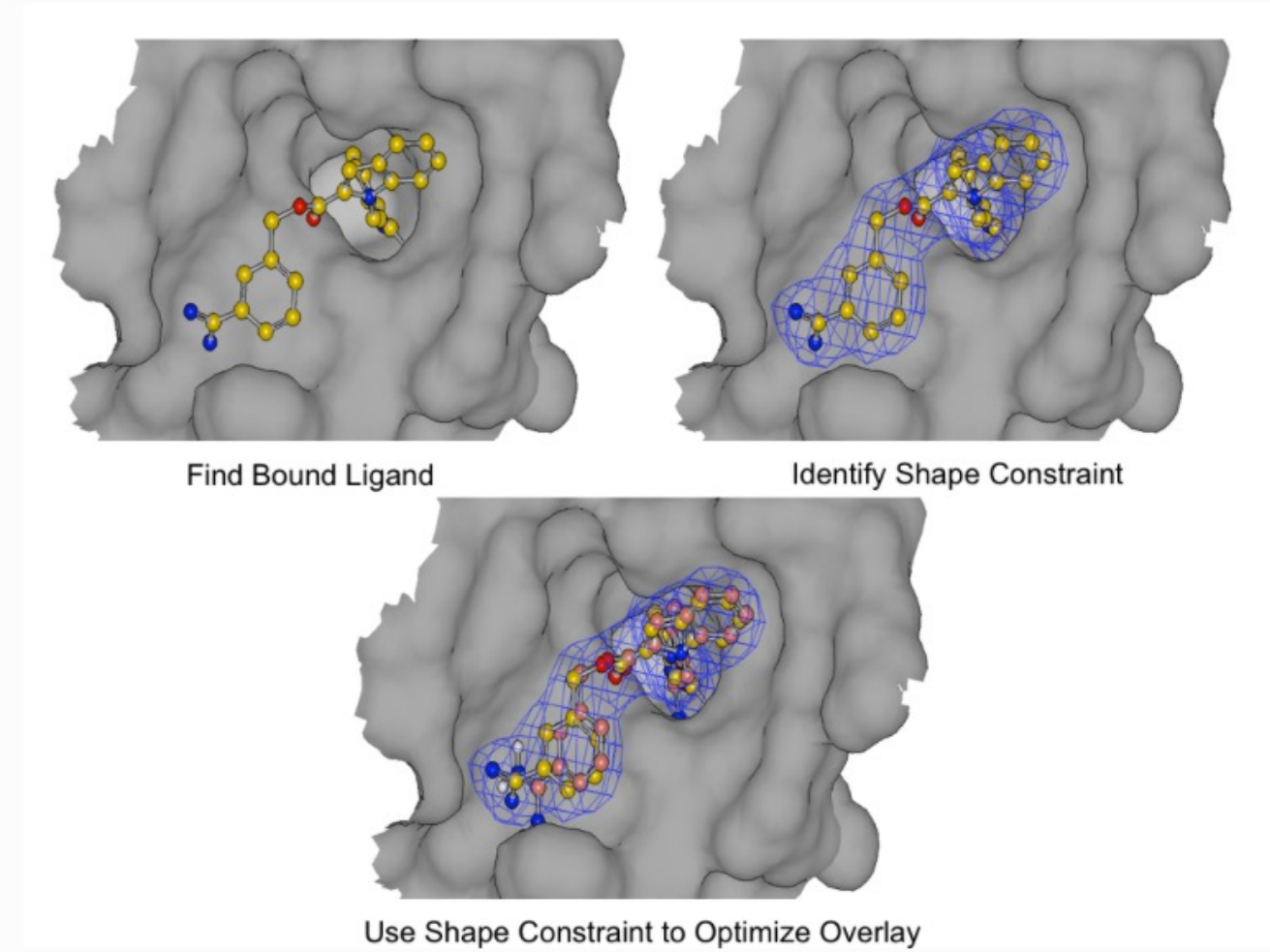


Using a reference atomic structure with an existing molecule bound



Using a pocket finding algorithm

Shape based methods



With an existing ligand it is possible to match the shape and optimize the overlay

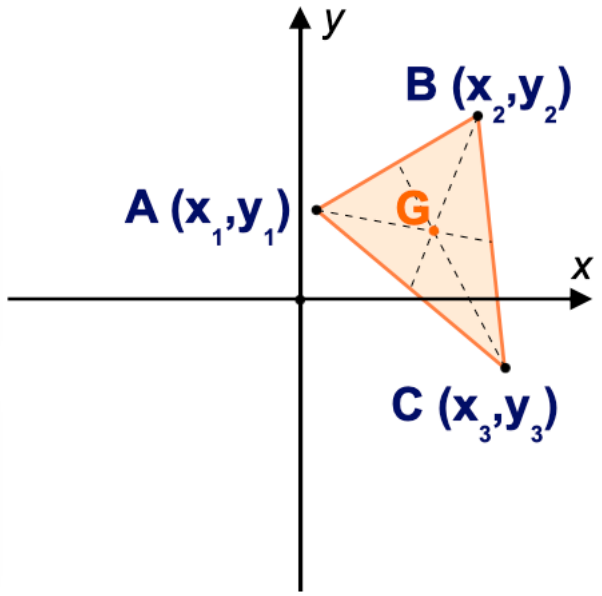
Fast and robust

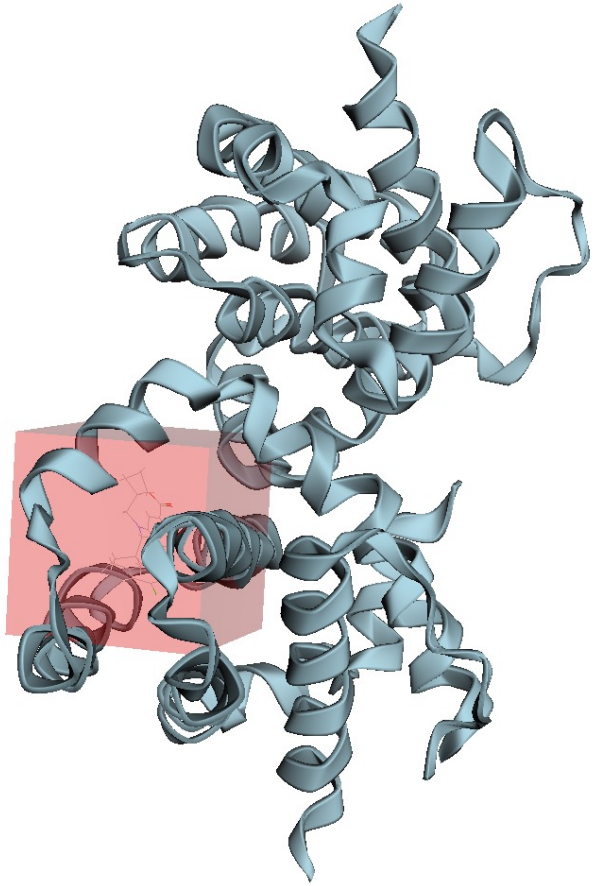
Ligand changes are not taken into account

https://docs.eyesopen.com/applications/oedocking/theory/posit_theory.html#posit-heat-map

Finding the docking grid area

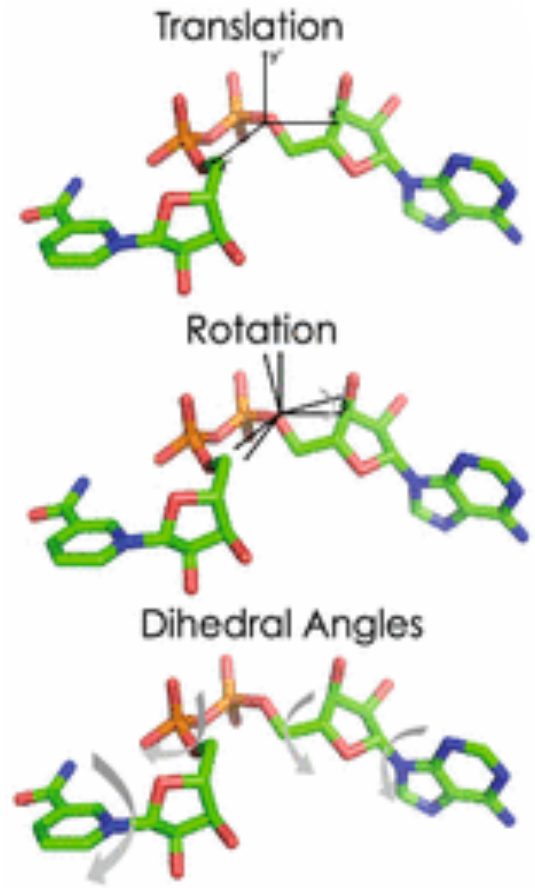
Often you have a ligand template or binding site residue to help with designing the docking grid



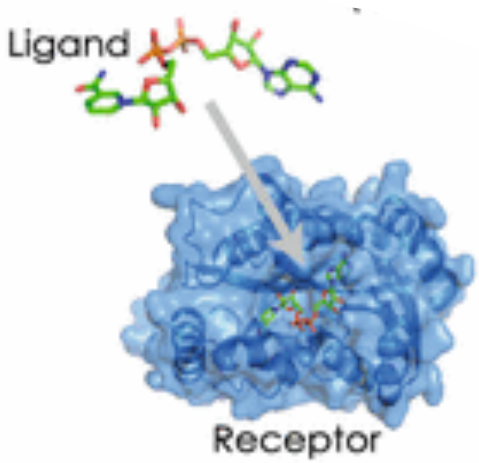


Use centre of geometry of a molecule from X-ray

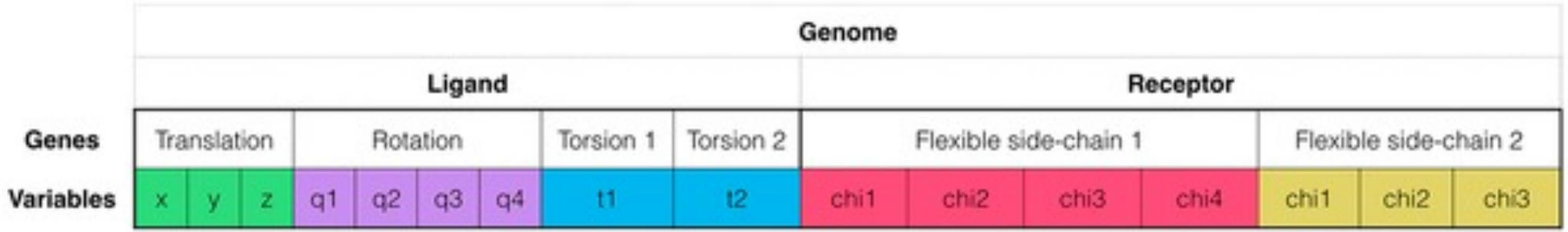
Genetic algorithm for ligand conformers

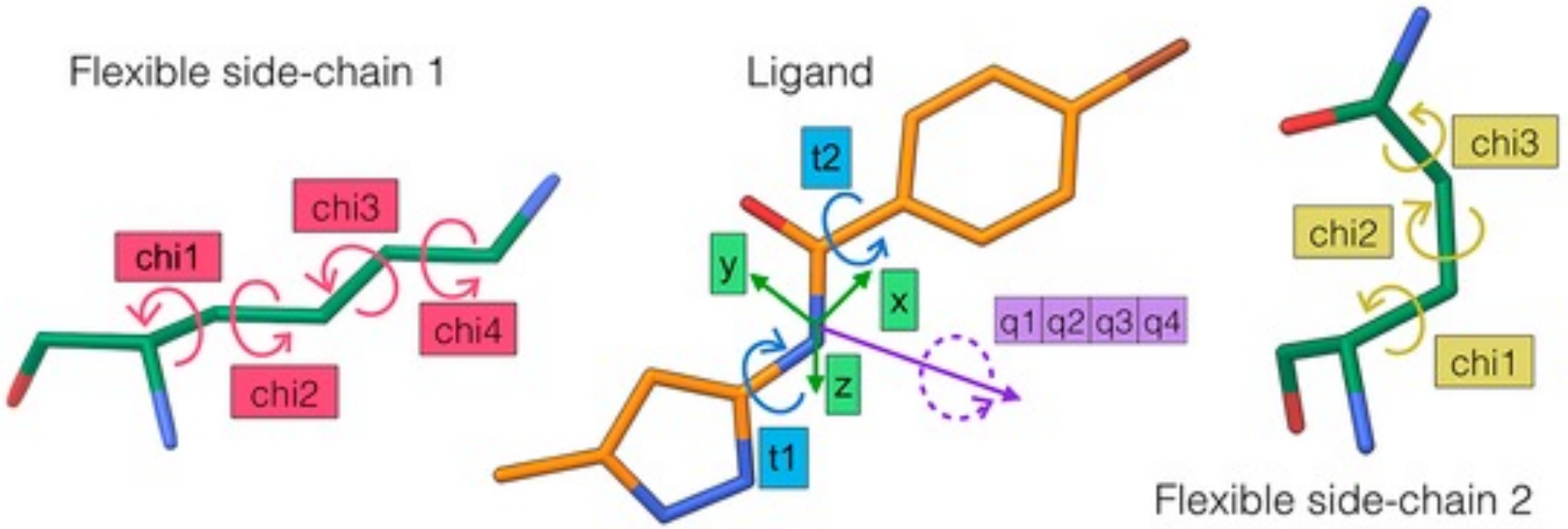


Testing different arrangements of the ligand in the rigid binding site of the protein

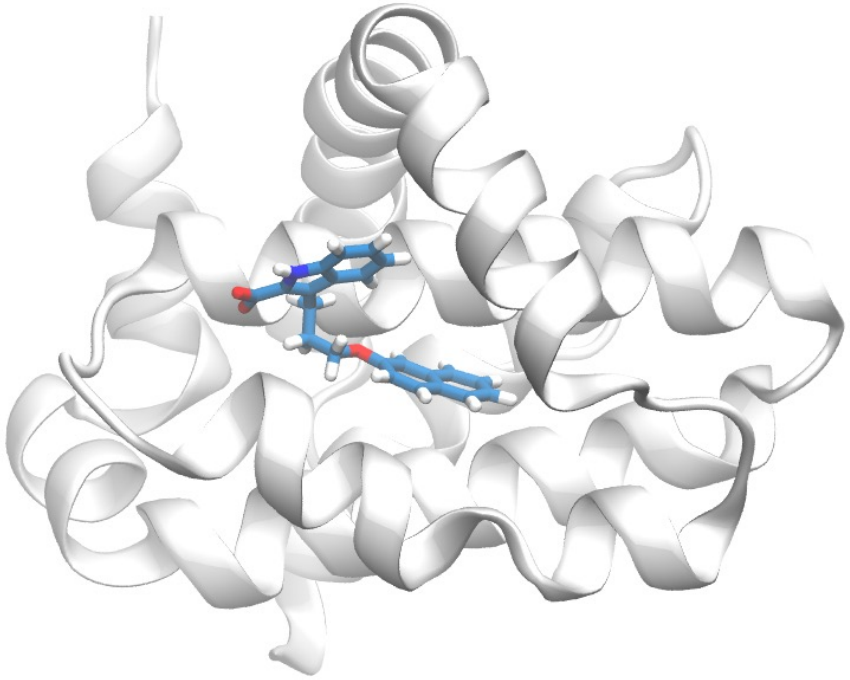


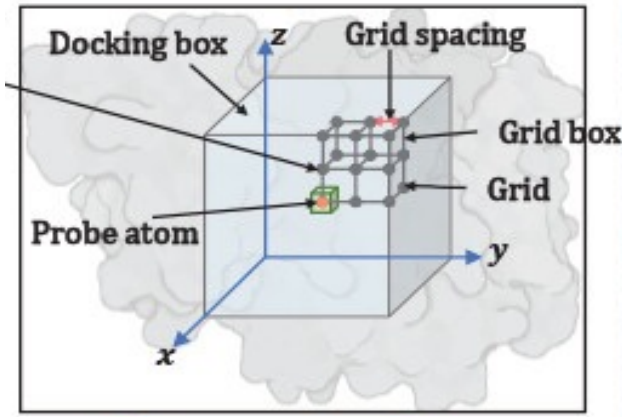
Allowing protein and ligand flexibility is often better



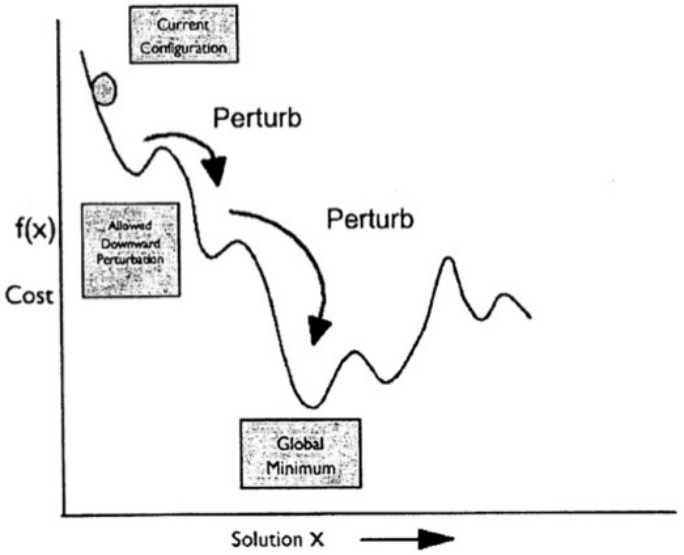


Flexibility increases compute time





Simulated annealing

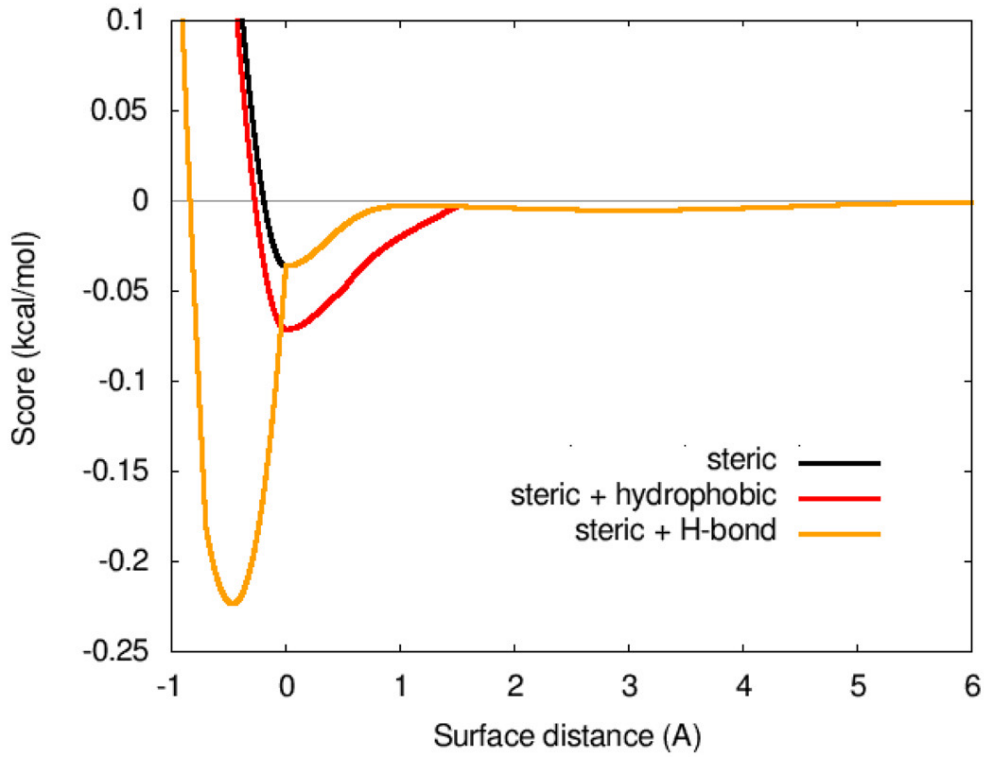


N=T360/i
N: number of conformations
T: number of rotable bonds
I: incremental degrees

→

Typical drug molecule
10 rotable bonds
30º increments (discrete)
10^{12} plausible conformations!

Scoring functions



Scoring functions can be used beyond shape optimization to optimize ligand and protein interactions

$$\Delta G = \left(V_{bonded}^{L-L} - V_{unbonded}^{L-L}\right) + \left(V_{bonded}^{R-R} - V_{unbonded}^{R-R}\right) + \left(V_{bonded}^{R-L} - V_{unbonded}^{R-L} + \Delta G_{conf}\right)$$

$$V = W_{vdw} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + W_{hbond} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) + W_{elec} \sum_{i,j} \frac{q_i qj}{\epsilon(r_{ij}) r_{ij}} + W_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{\frac{-r_{ij}^2}{2\sigma^2}}$$

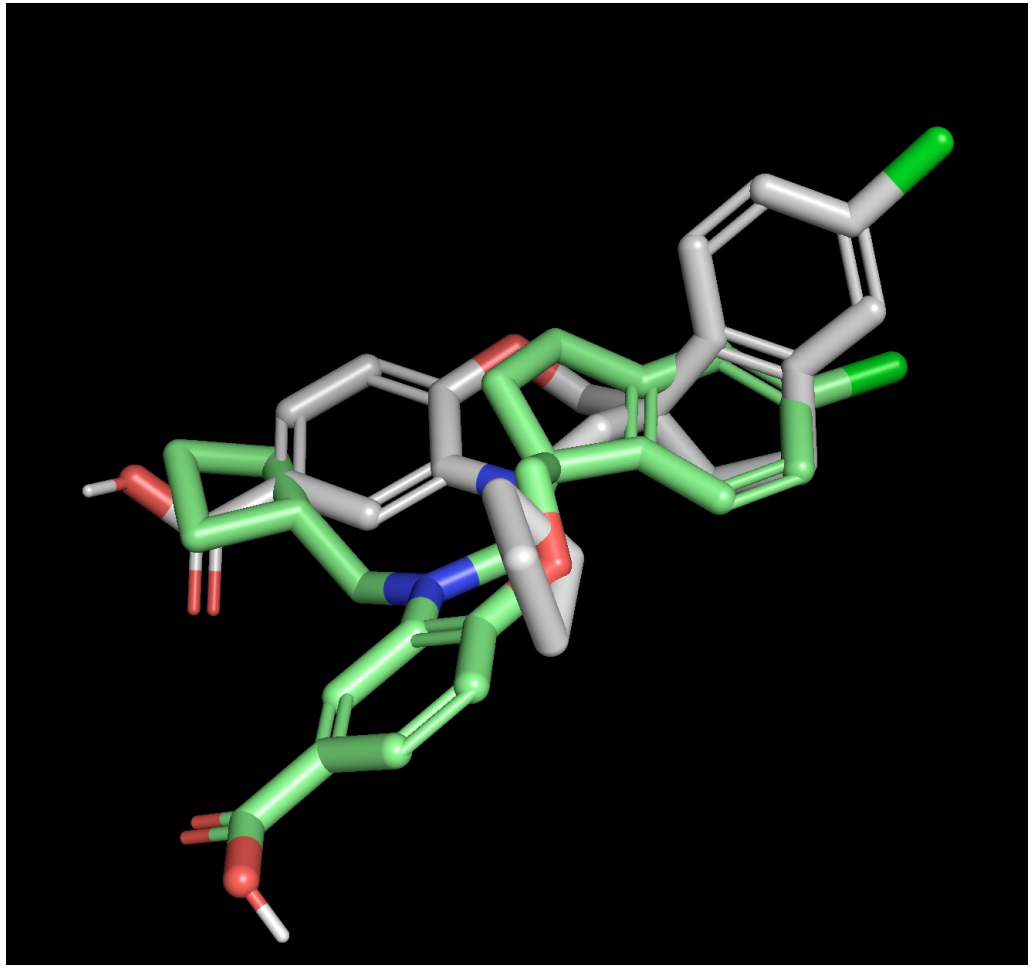
$$\Delta G_{conf} = W_{conf} N_{tors}$$

J Comput Chem. 2010 Jan 30; 31(2): 455–461.

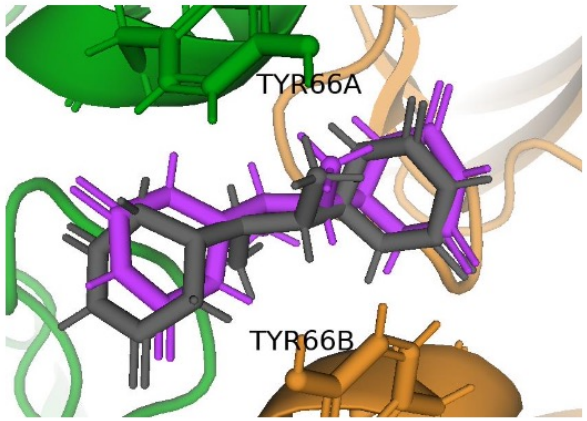
Soft Comput **23**, 4155–4176 (2019).

Typical docking output generates multiple poses

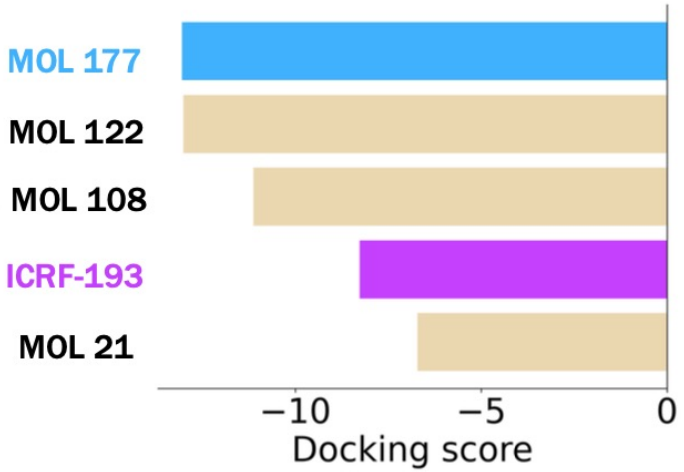
mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-8.36	0	0
2	-8.08	2.899	6.789
3	-7.985	3.643	7.852
4	-7.914	3.415	5.21
5	-7.765	2.167	2.826

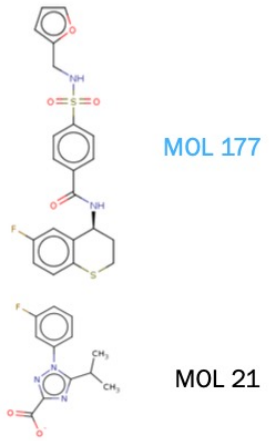


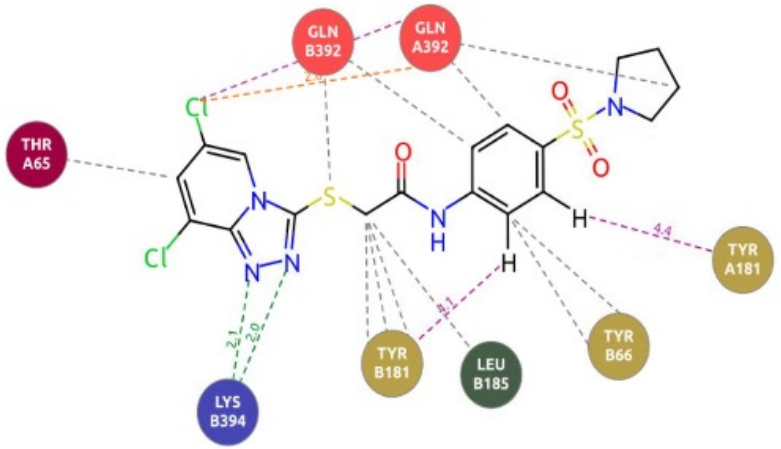
Evaluating Docked structures



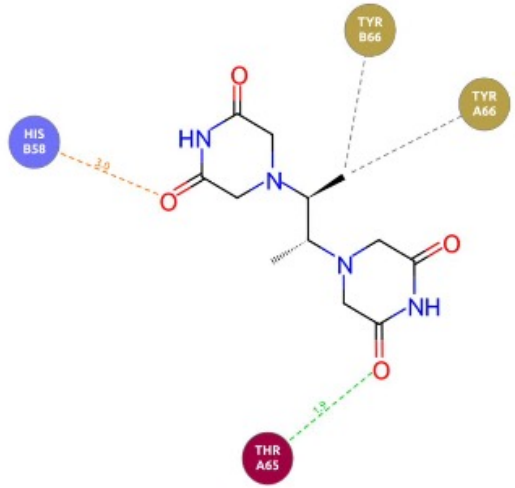
Crystal structure ICRF-193





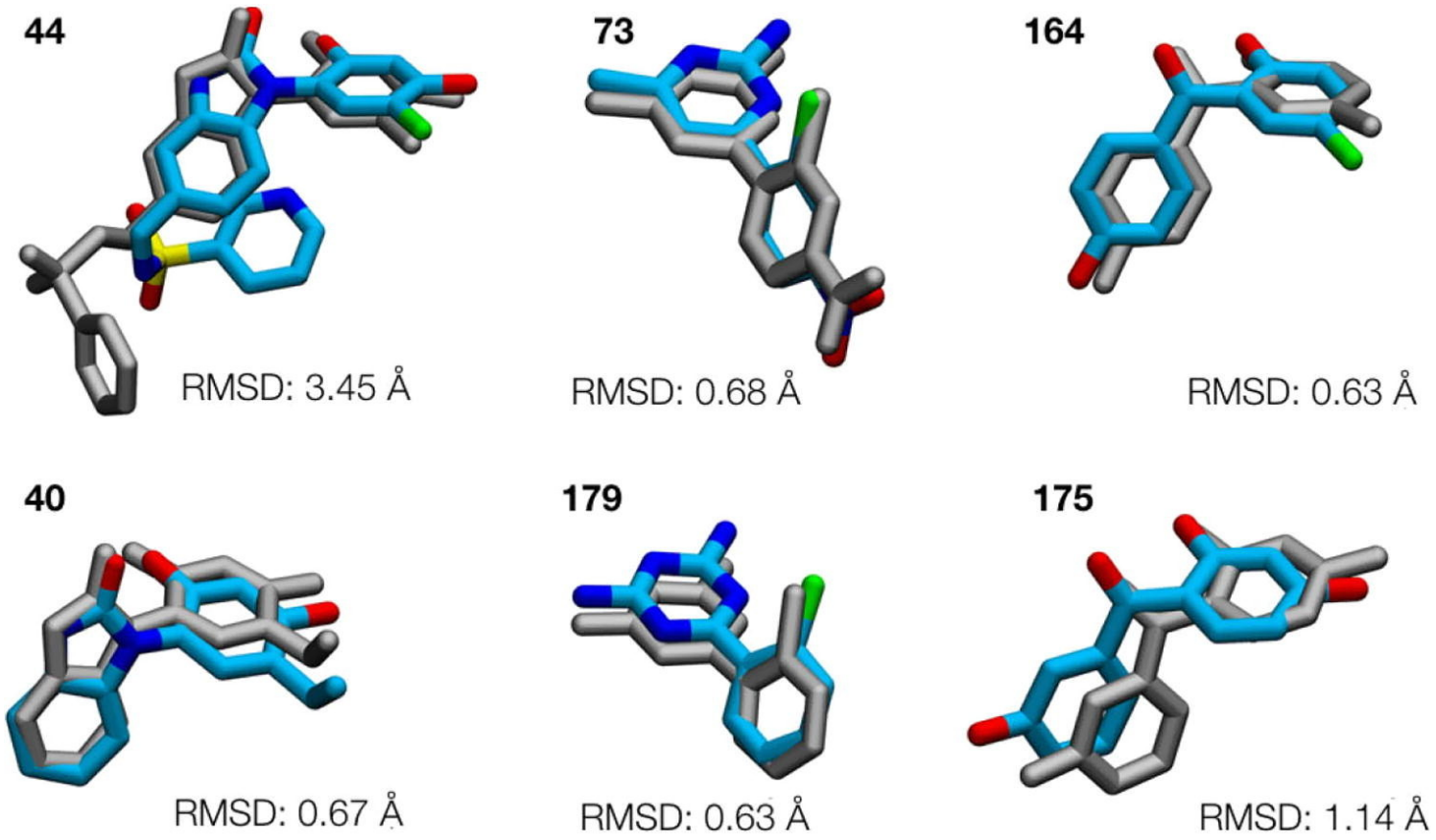


Contact map MOL-177



Contact map ICRF-193

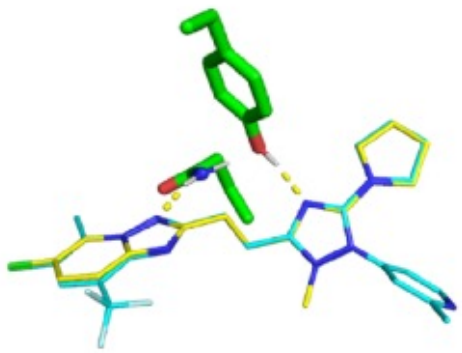
Evaluating the binding mode/pose



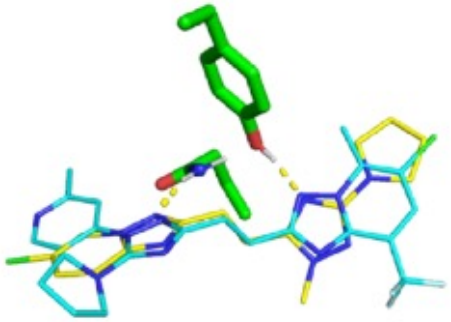
Bioorganic & Medicinal Chemistry
Volume 24, Issue 20, 15 October 2016, Pages 4890-4899

Template docking and cross docking improves docking

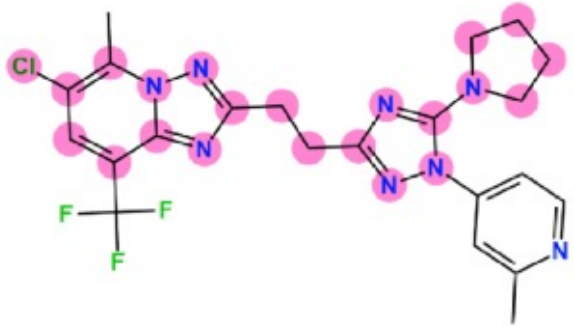
(a)



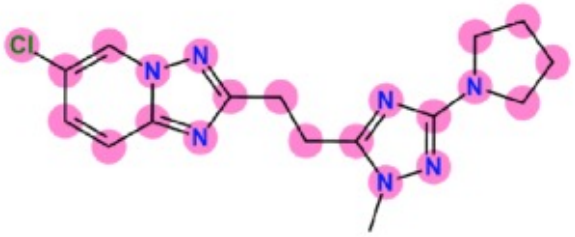
(b)



(c) Ligand compound 40:



(d) Template compound PDB code 5sej:

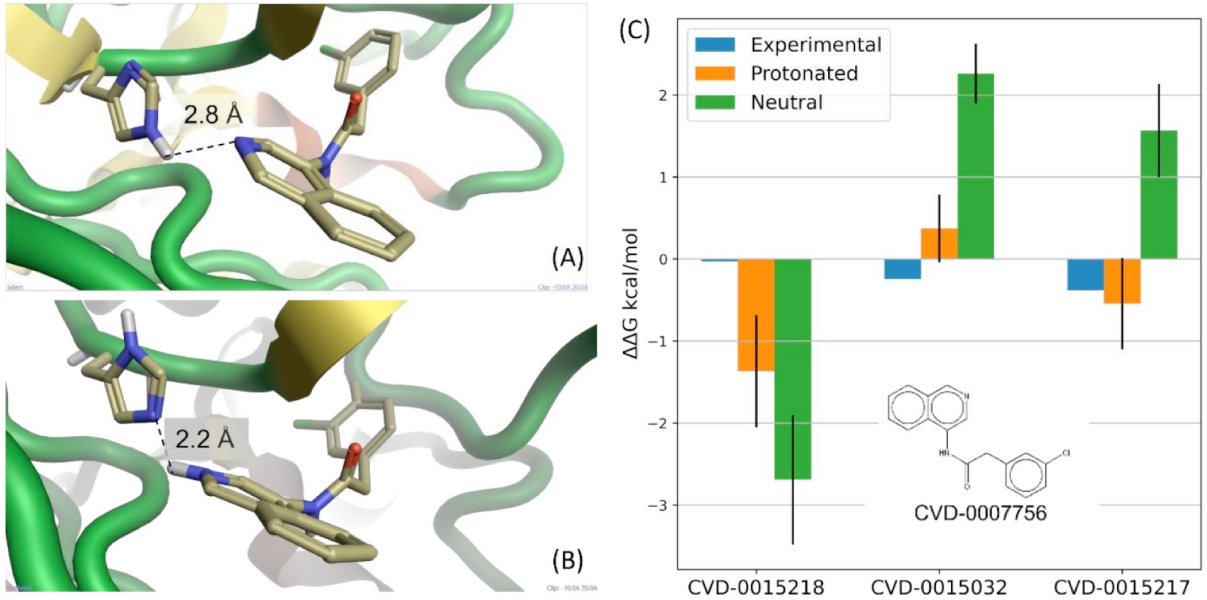


- Dock the same ligand into multiple protein structures (X-ray, MD)
- Generate multiple ligand conformers and dock into multiple structures
- Use existing ligand data as a template or guide e.g. through Maximum common substructure

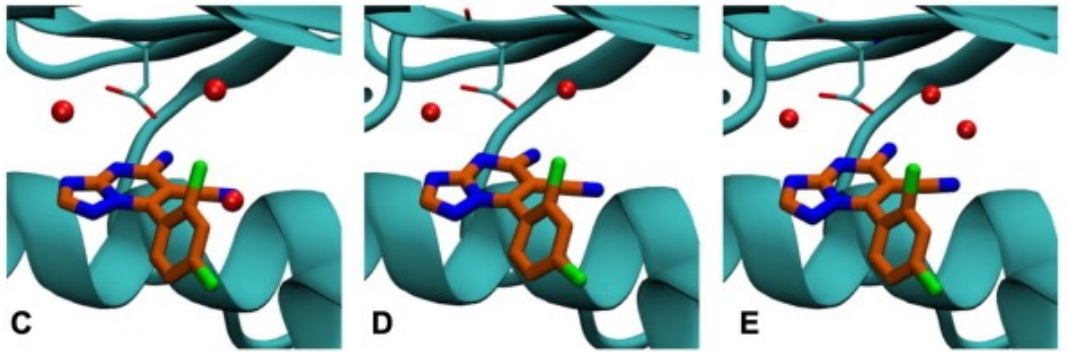
J Comput Aided Mol Des **36**, 753–765 (2022)

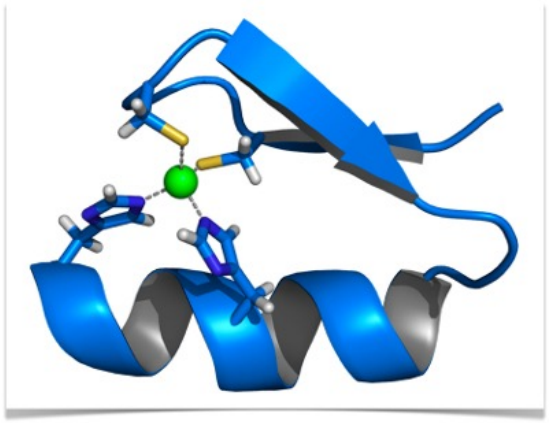
Things to worry about

pKa of ligands and binding site protonation need to be considered



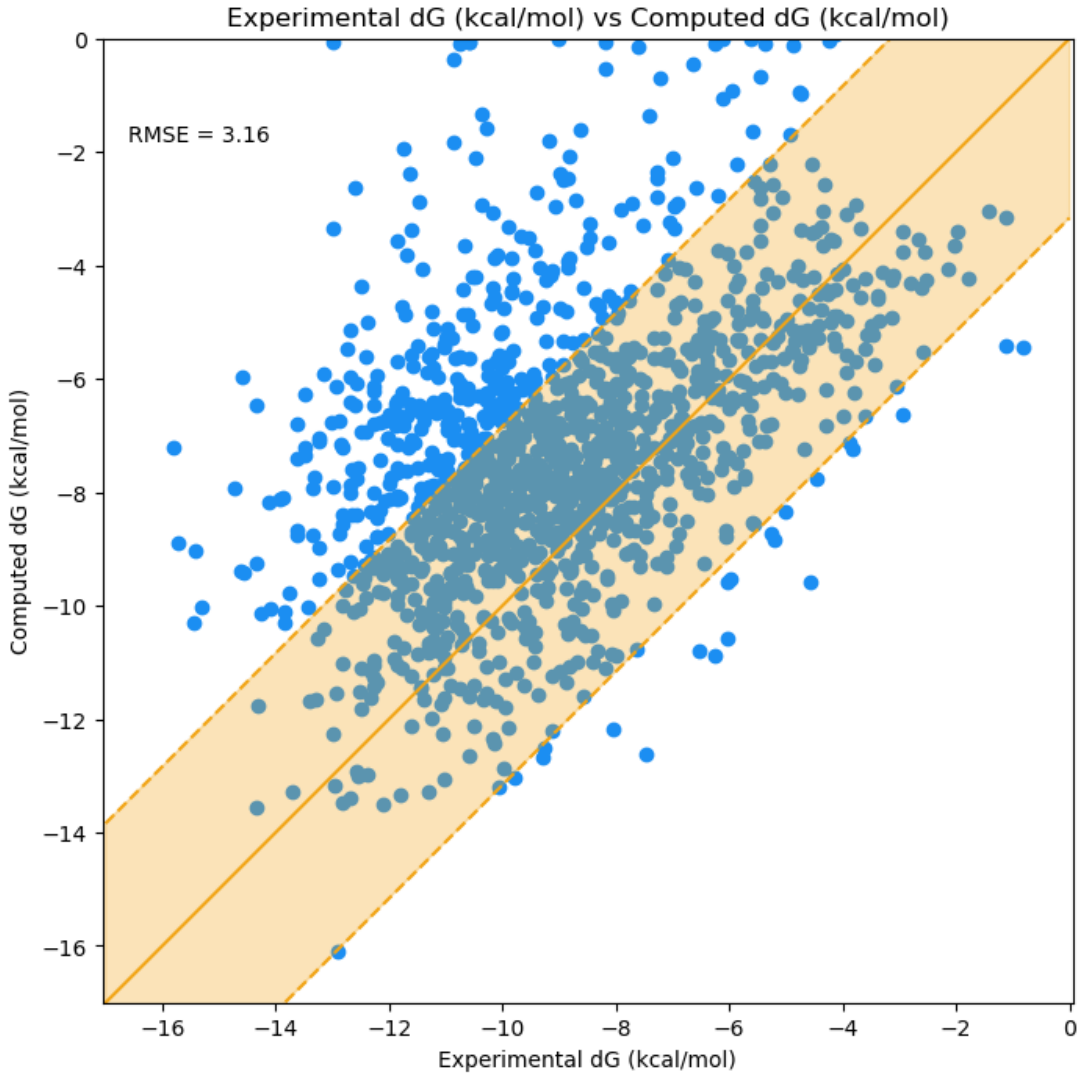
Structural waters are important



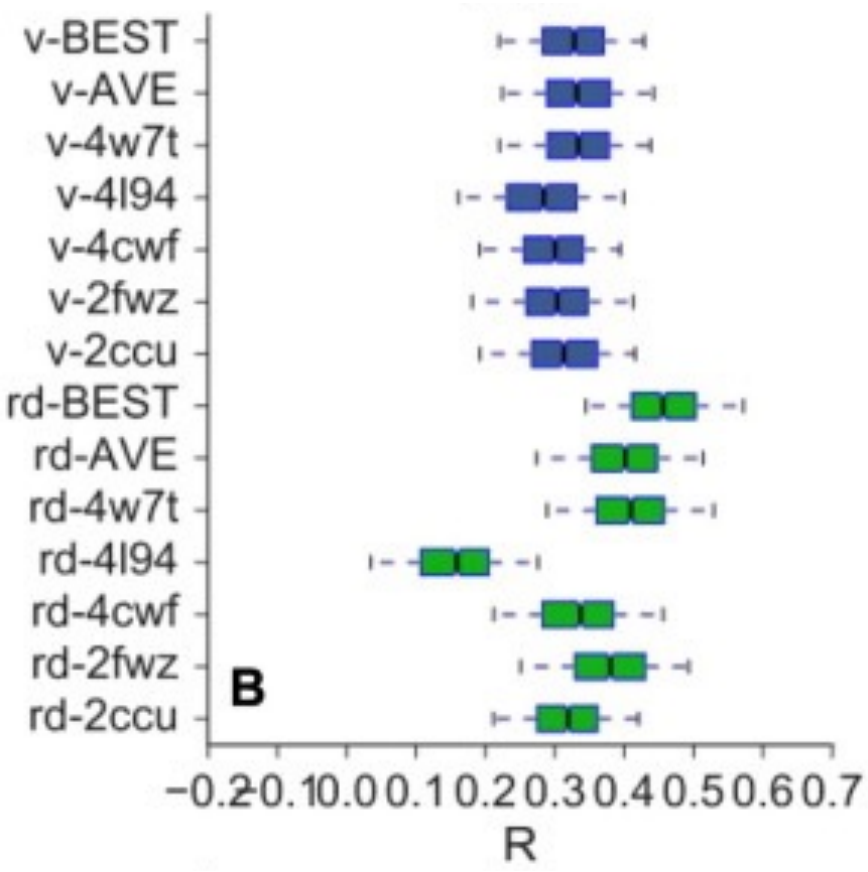


Co-factors such as ions and other molecules are imporant

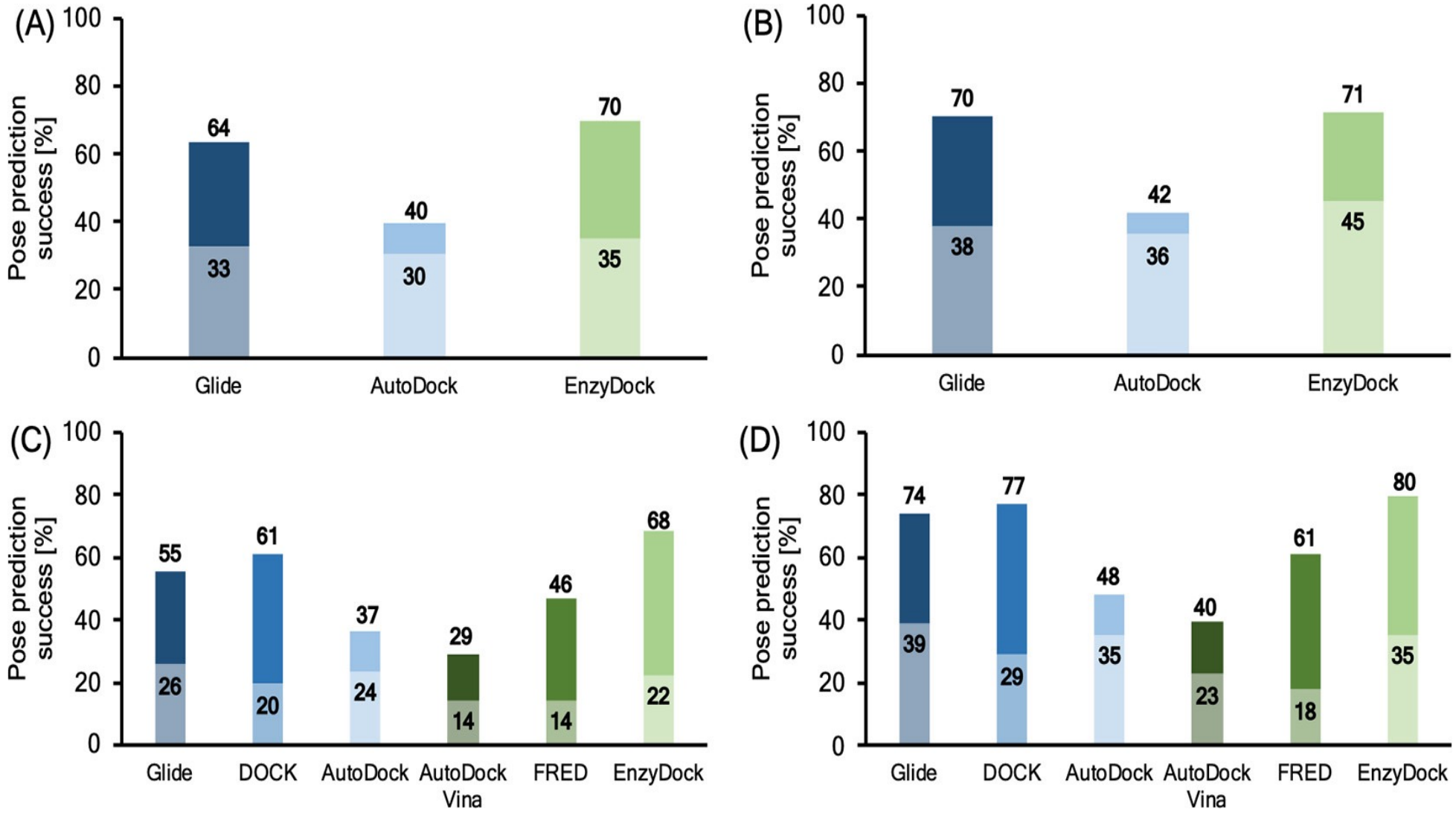
Comparing against experimental ΔG



Bulk assessment of correlation coefficient

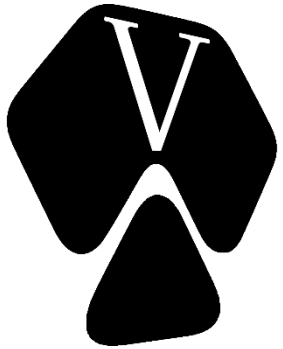


Recent docking benchmark



J. Chem. Inf. Model. 2021, 61, 6, 2957–2966

What tools exist for molecular docking?



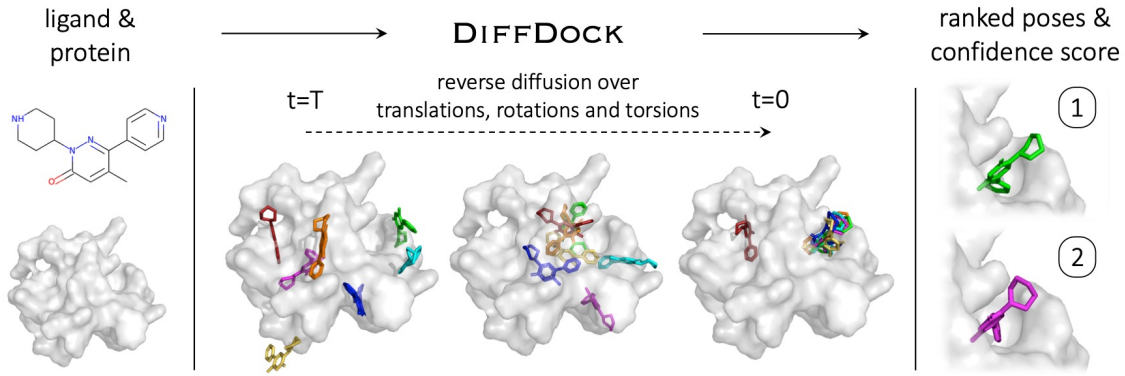
Autodock Vina

Glide

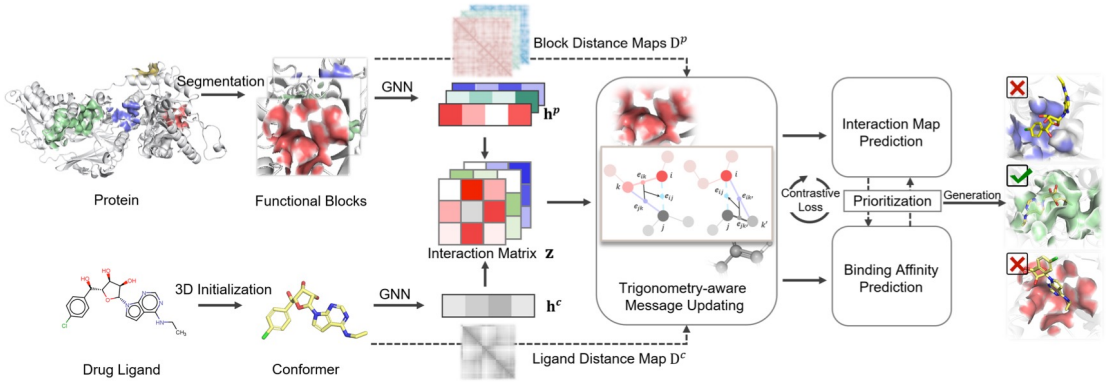
And others….

ML-based docking

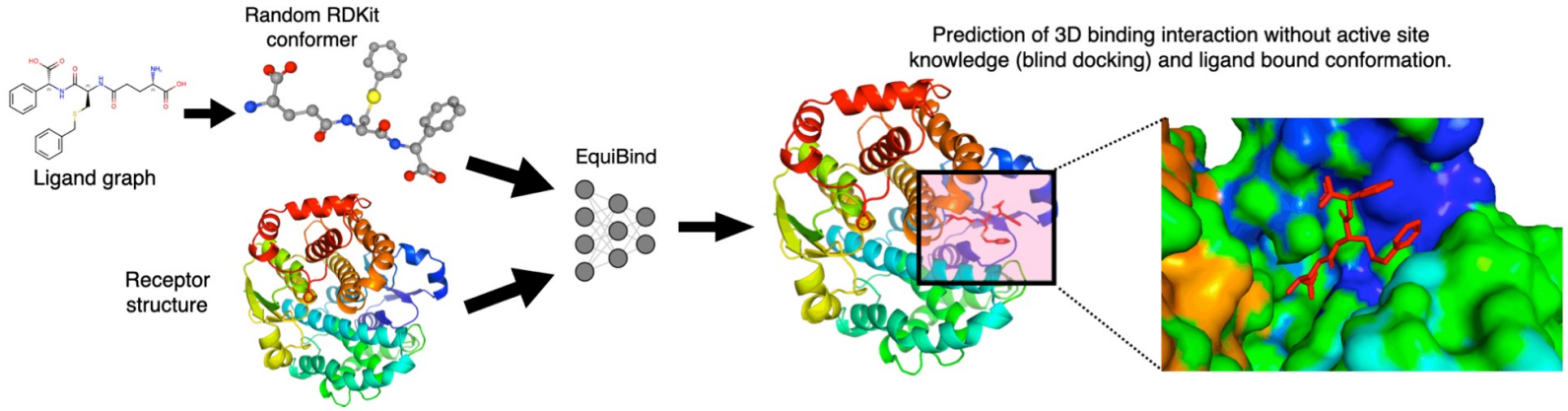
Diffdock



TankBind



Equibind



Gnina

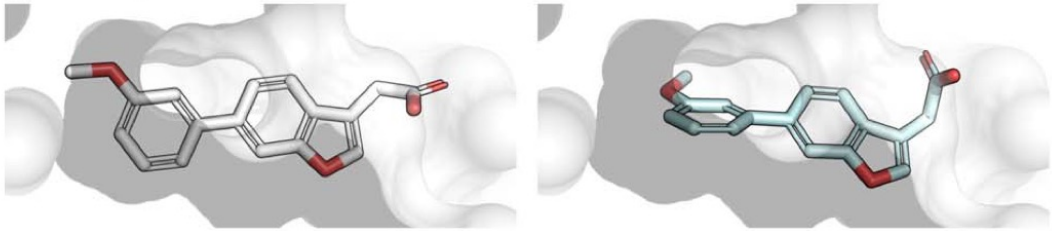
Deepdock

....

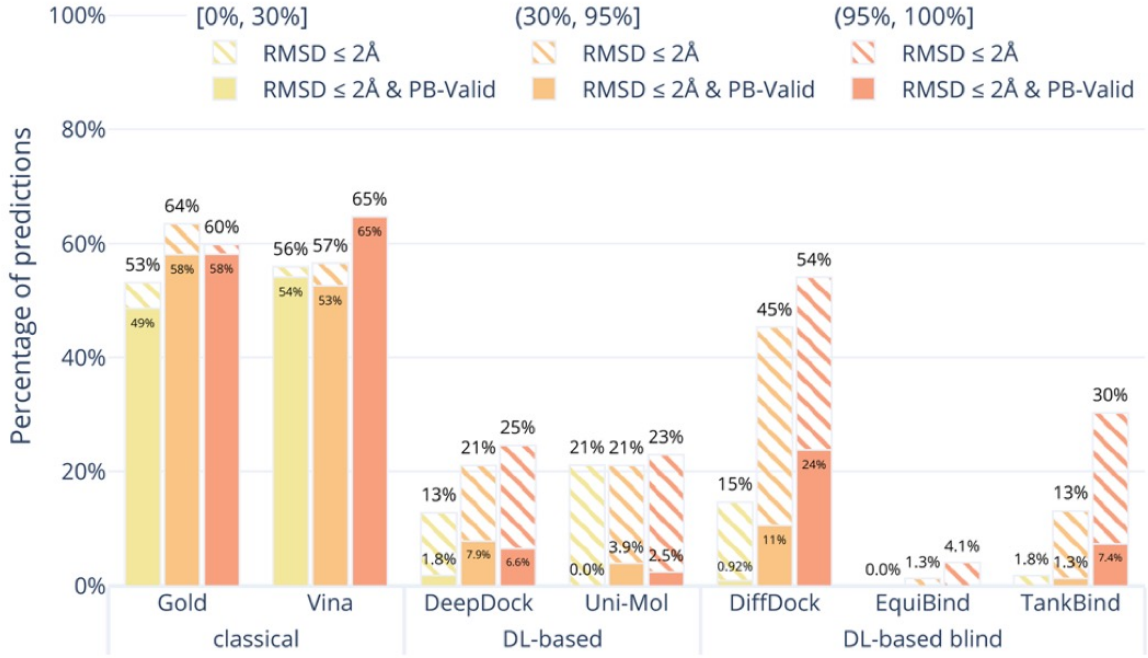
Evaluating the binding mode/pose of ML tools

(f) Double bond not flat. TankBind prediction for ligand DBQ of protein-ligand complex 1U4D. RMSD 1.7 Å.

(g) Energy ratio too high. AutoDock Vina prediction for ligand IFM of protein-ligand complex 7LOU. RMSD 1.9 Å.



(h) Clash with protein. DiffDock prediction for ligand XQ1 of protein-ligand complex 7L7C. RMSD 1.6 Å.



PoseBuster: ***Chem. Sci.***, 2024,**15**, 3130-3139


Break Time!