Modeling ligand binding in protein targets

Scott Lovell, Ph.D.
Modeling ligand binding in protein targets
University of Kansas / SSGCID
July 21, 2026

Introduction: Protein structure and X-ray Crystallography Laboratory (PSXL)

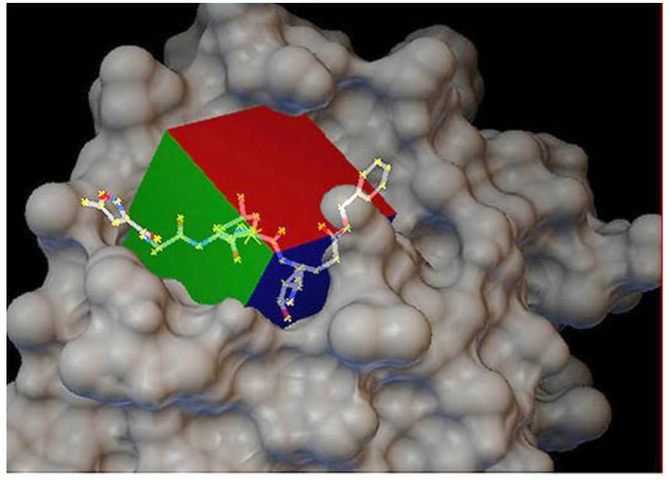
- High throughput structural biology
  - Crystallography and Cryo EM
- Structural Biology Center at KU
  - Protein Production, Computational Chemical Biology, High Throughput Screening, Mass Spectrometry, NMR Infectious Disease Assay Development and Chemical Synthesis

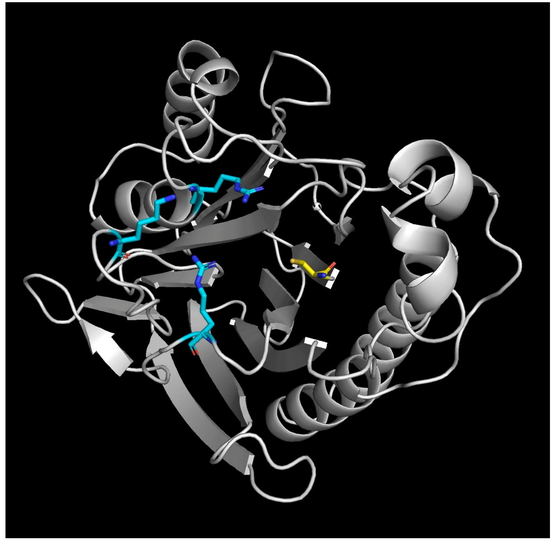
Methods for modelling protein:ligand binding

1. Docking methods: Modelling of a ligand in a protein pocket
2. Virtual screening: Identification of compounds for inhibitor development
3. Molecular dynamics simulations: Sampling conformational states of a time trajectory
4. Case Study: in silico inhibitor development
5. Preparing files for ligand modelling and software applications
6. Ligand modelling examples and analysis in the BioViz lab

General docking approach

- Computationally predict the binding mode(s) of molecules (ligands) to a protein
  - Add hydrogens to protein/ligand
- Setup a grid around the for docking search
- Sample the binding conformations (pose)
- Scoring based on binding free energy

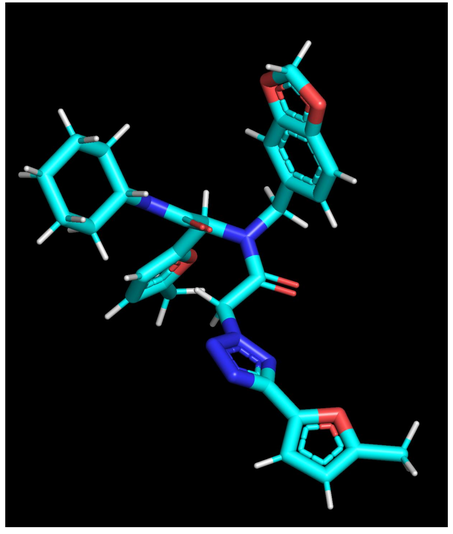




Preparation for ligand modelling

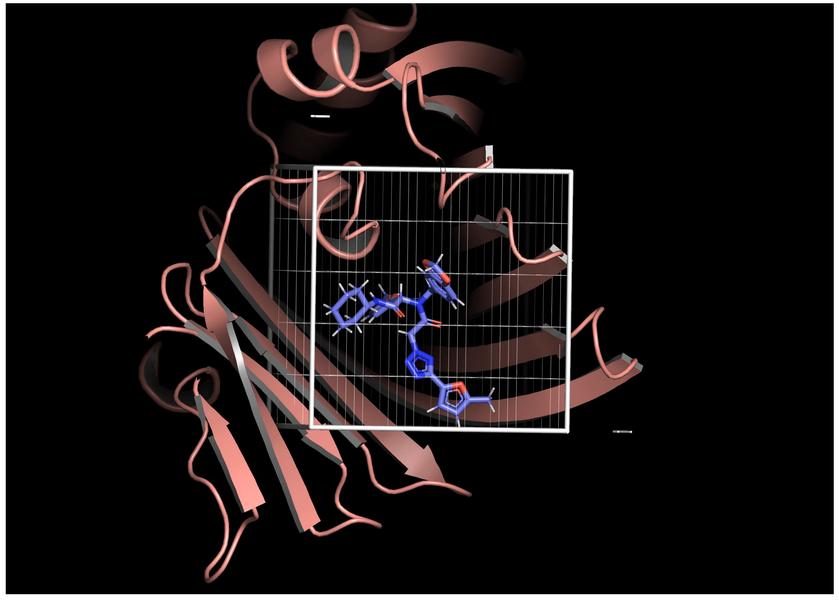
- Large cavity in protein used to test for potential ligand binding
- Protein model loaded into docking application
  - Hydrogen atoms are added in idealize positions
- Ligand molecule is loaded and hydrogen atoms added

Ligand to dock

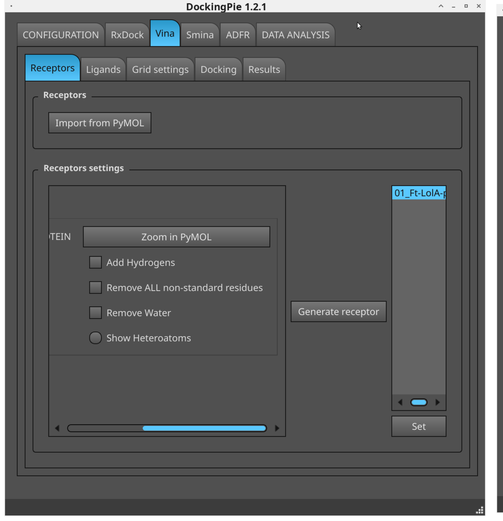


Ligand search parameters

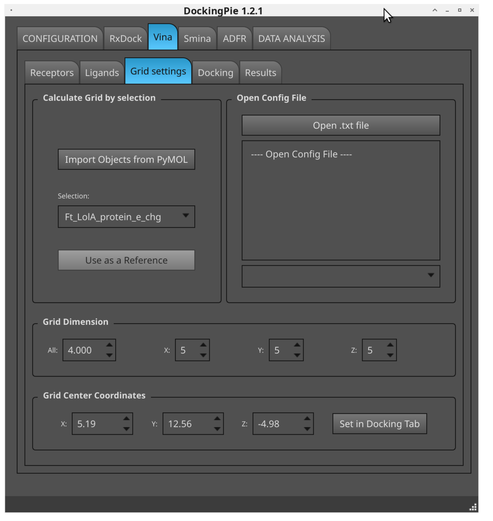
- The search grid is centered near the region of interest (cavity) and set in the x,y,z directions to encompass the desired area



Protein/Ligand Setup

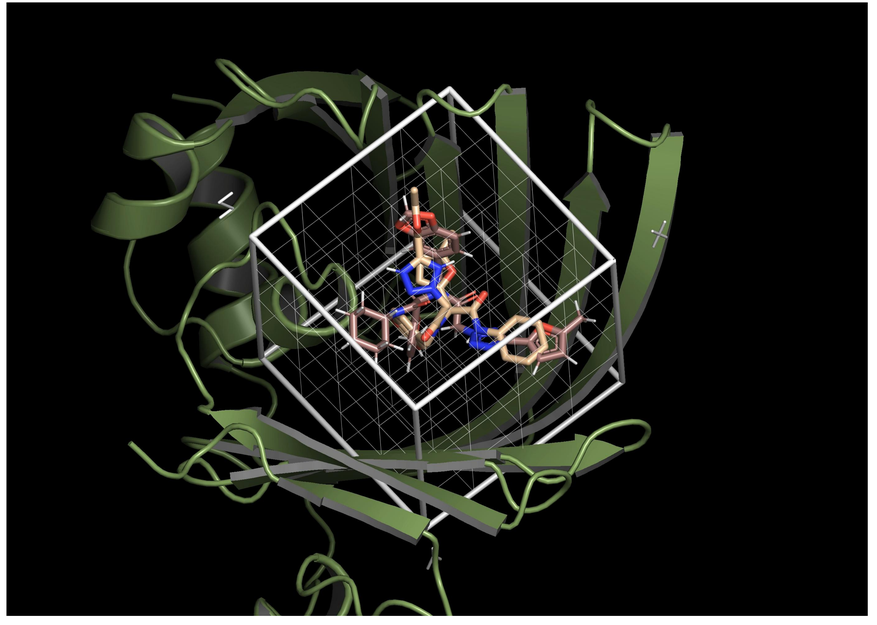


Grid Settings

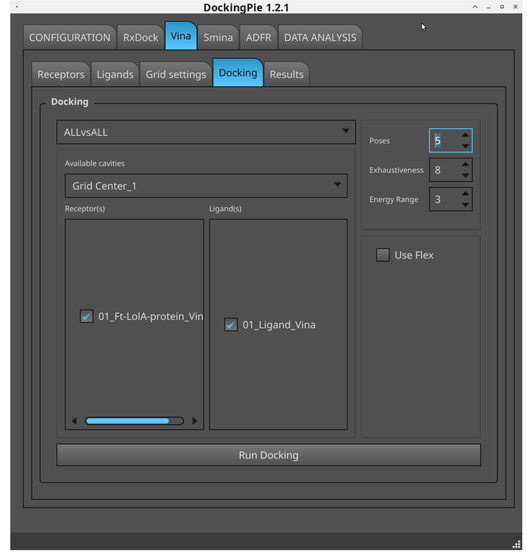


Ligand search parameters and results

- One would typically search for multiple poses within the grid space
- Results are filtered for “top” poses
  - Binding free energy

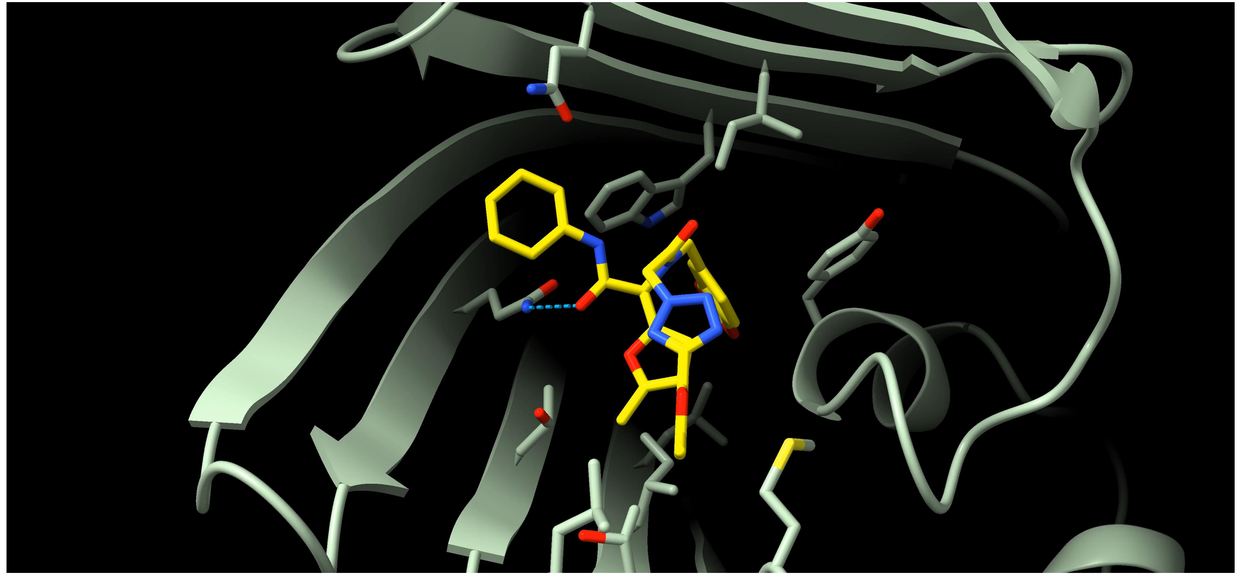


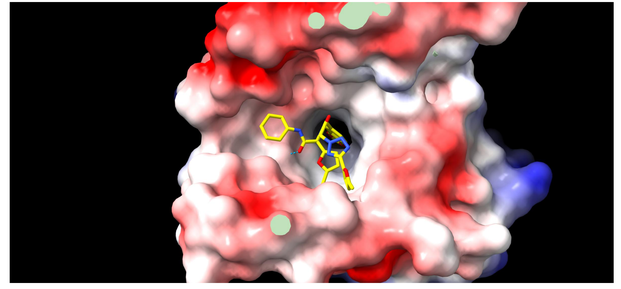
Search Parameters

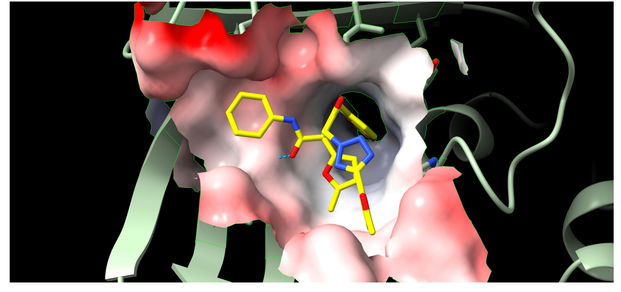


Analysis of the results

- Software: Pymol, CCP4mg, Coot, ChimeraX
- Analyze features such as hydrogen bonds, hydrophobic interactions, electrostatics, pocket environment

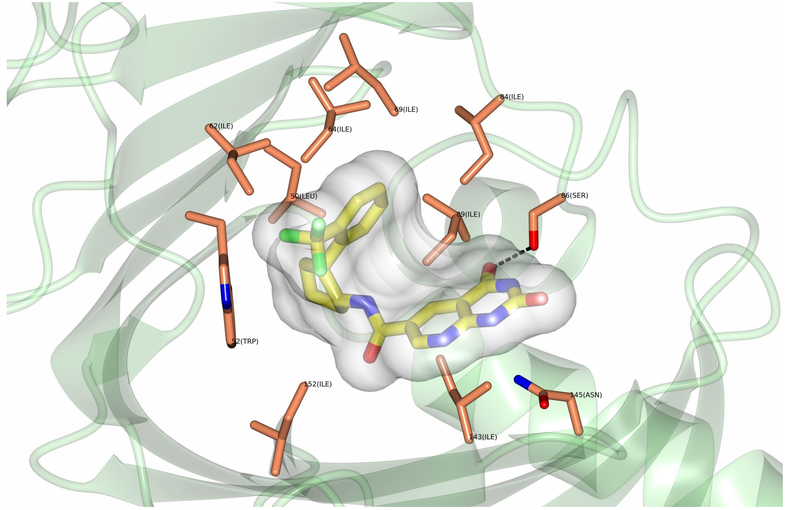


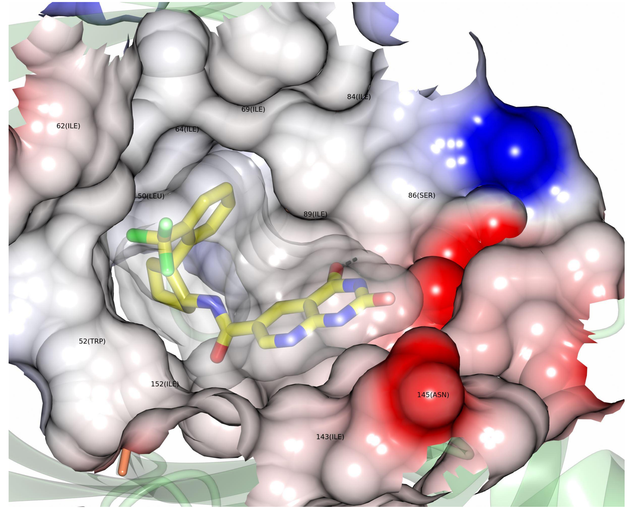




Evaluation of a docked ligand

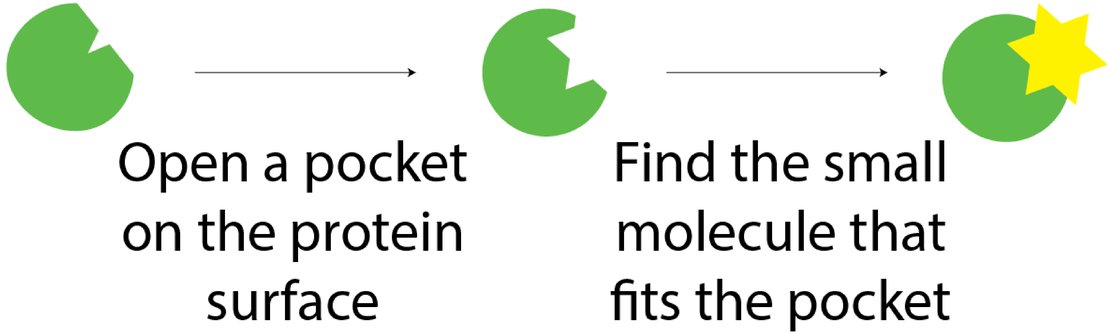
- Hydrogen bond interactions
- Pocket characteristics
  - Hydrophobic and charged regions
- Various software packages for analysis





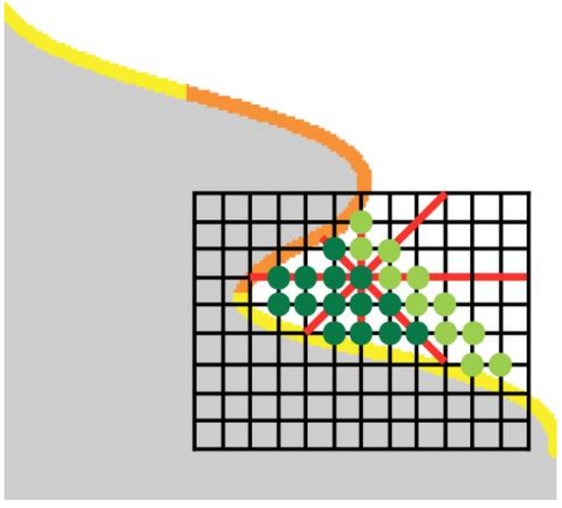
Advanced docking methods

- Docking multiple ligands to a protein target
  - In silico screening protein targets for lead compound identification
- Strategy:
  - Use a structure without a small molecule,
  - predict a conformation with a pocket
  - search for compounds that have shape and chemical complementarity.



Pocket optimization

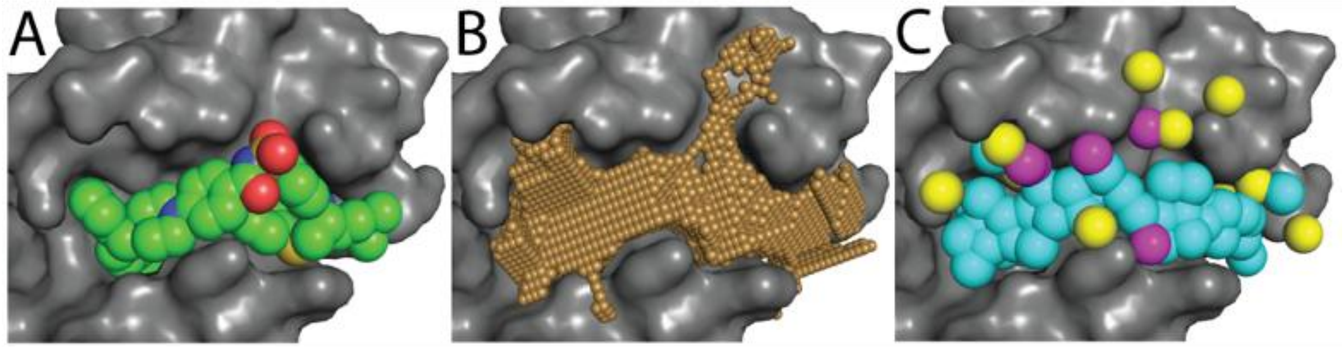
- Apply a biasing potential to drive the ensemble towards pocket containing conformations
  - Biasing potential is a proxy for binding free energy
  - Bias towards low energy conformations
- This can be a structure prediction problem
  - Given a template (unbound) structure one needs to identify conformations that satisfy energetic and pocket constraints



Johnson and Karanicolas, PLoS Comput Biol 2013

Exemplar screening

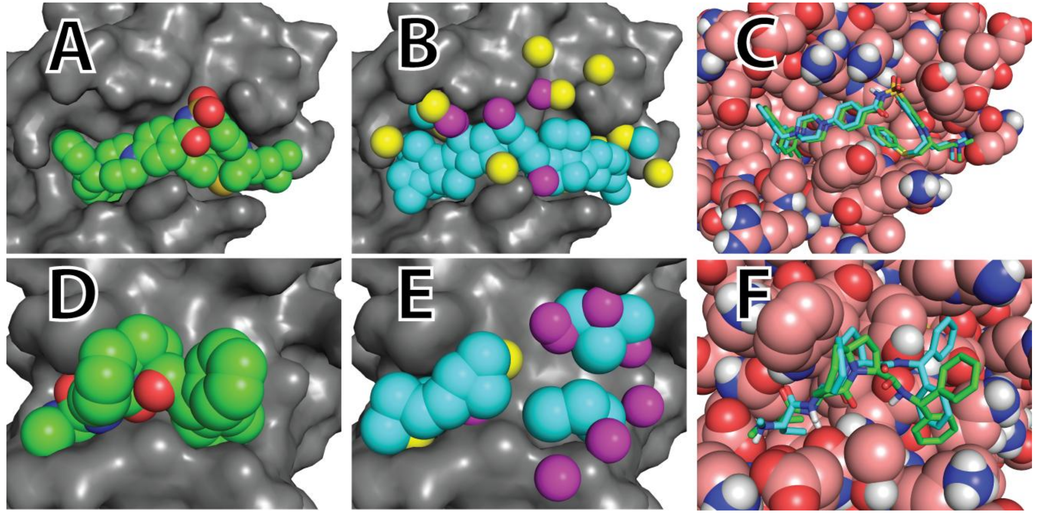
- Identify pockets
- Identify ideal location for donors and acceptors
- Fill pocket with carbons atoms
- Cluster exemplar points



(Johnson and Karanicolas, 2015)

Ligand placement via exemplar

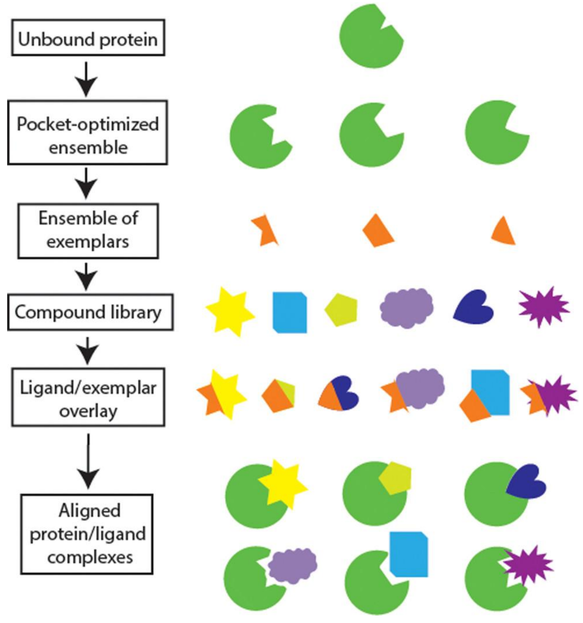
- Rapid Overlay of Chemical Structures (ROCS) developed by OpenEye software
- Fast shape comparison of ligand to exemplar by Gaussian representation of atoms
- Can screen large number of compounds using a computing cluster



Johnson and Karanicolas, JCIM 2016

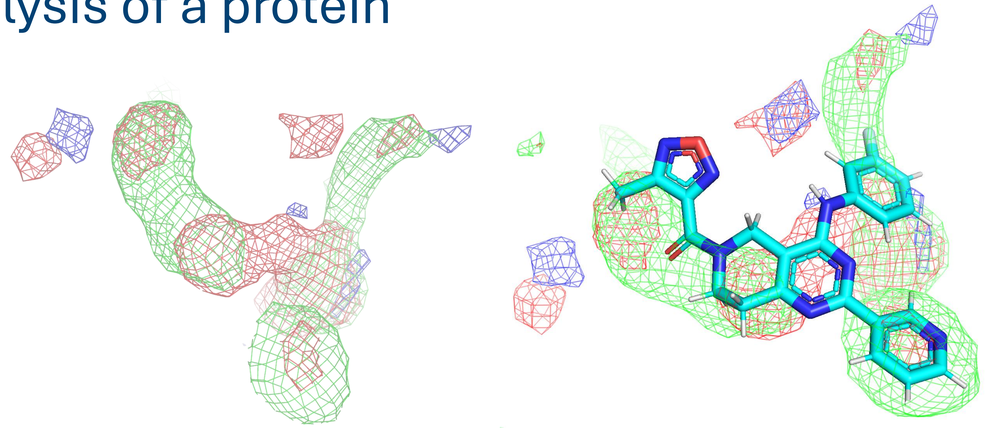
Screening structural ensembles

- Perform pocket optimization
- Assess pocket druggability (select pockets with tighter binding potential)
- ROCS exemplar screen (Rapid Overlay of Chemical Structures)
- VScreenML (active vs decoy classifier)



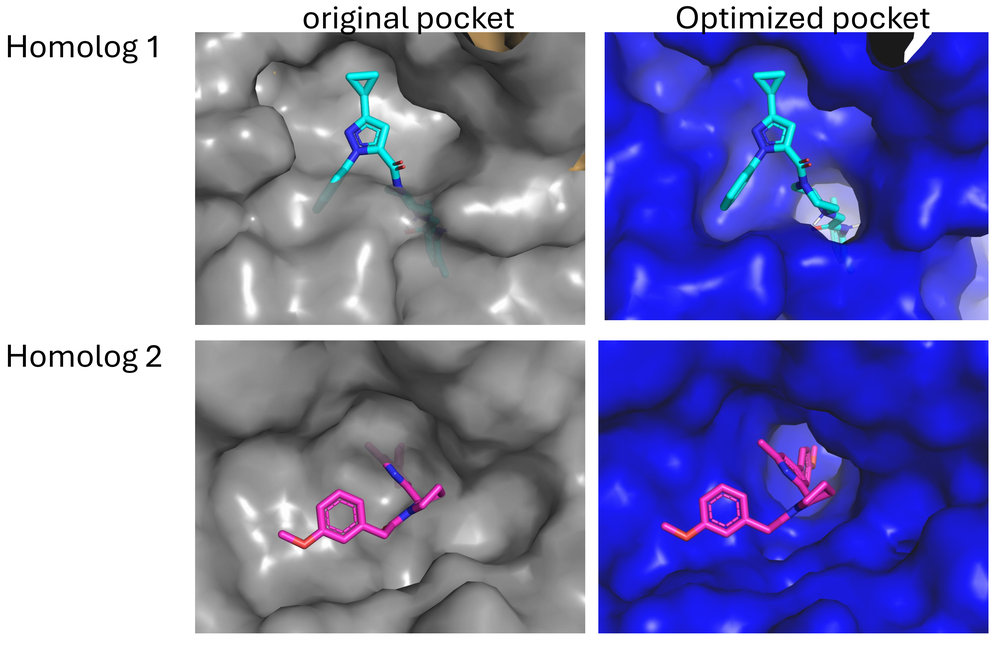
Site Identification by Ligand Competitive Saturation (SILCS) analysis of a protein

- Map of the interior pockets of the protein.
- Zones where the energy of a hydrophobic (green), donor (blue),
- or acceptor (red) are present
- Filter for ligands that fit uniquely to each protein target
- Identify compounds that might uniquely target individual targets (singles)
- Example: pyridine ring and the shape of the ligand similar shape to the pocket



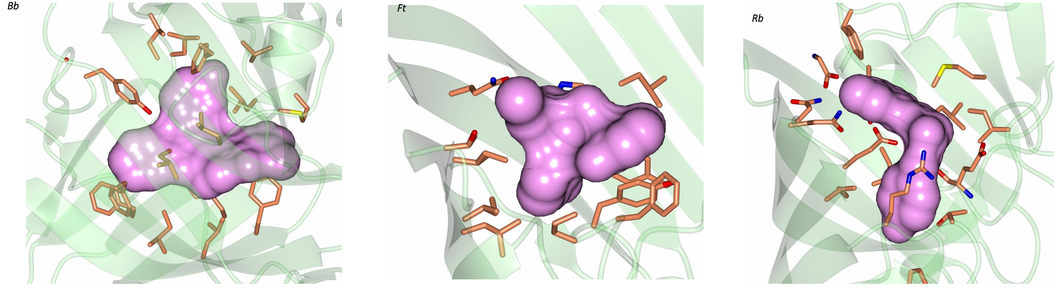
Example: In silico compound library screening

- Multiple protein target homologs
- Screened using the ZINC library
- ROCS and SILCS methods utilized
- Pocket optimization suggests deepening of the pockets during simulations supporting compatibility with binding



Pocket comparison between protein homologs

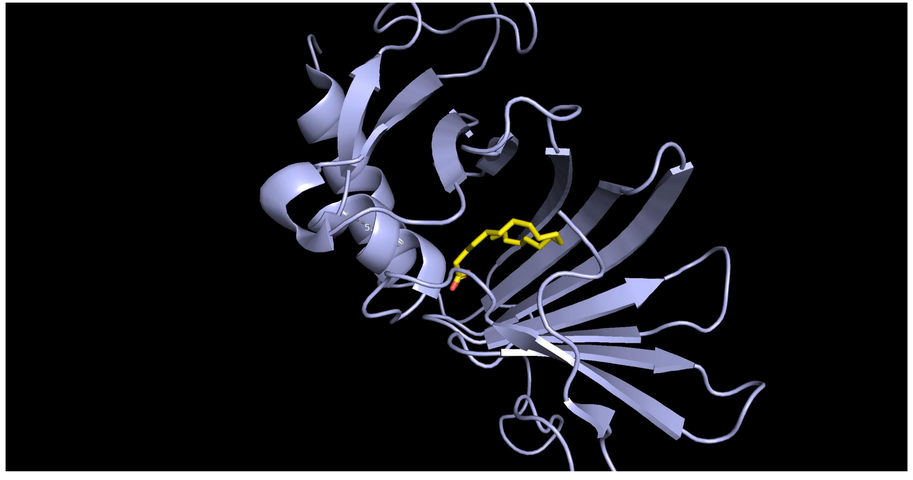
- Map the differences/similarities amongst homologs
- Screen for compounds that might bind to various proteins
- Identify compounds that are selective



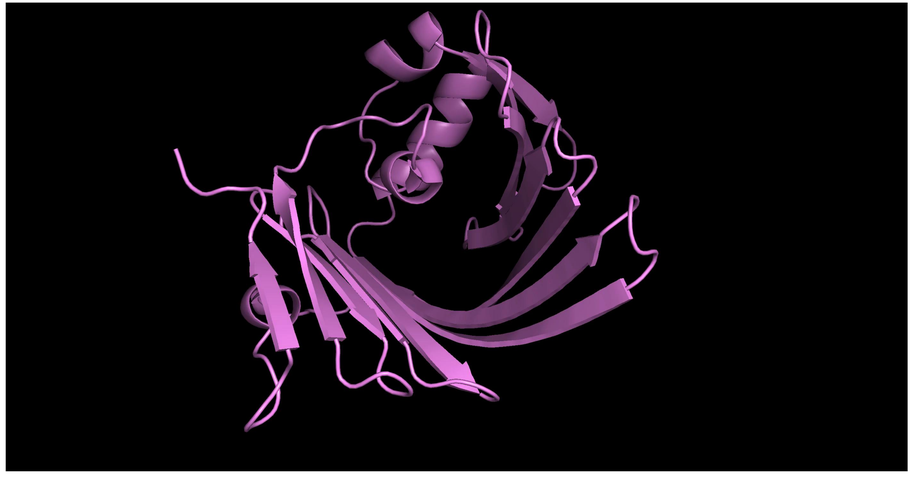
Top hits from in silico screening

- Diverse chemical functionality
- Differences in pocket depth between homologs

Homolog 1

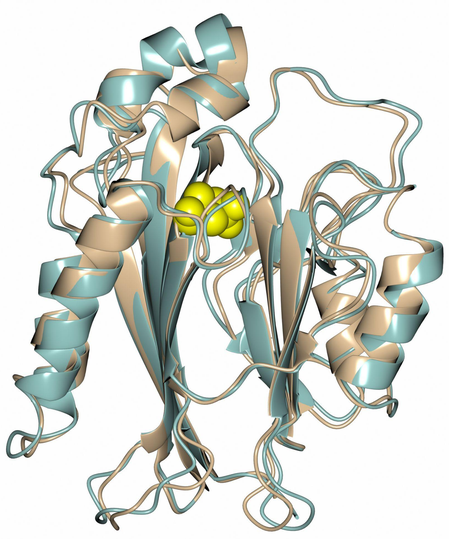


Homolog 2



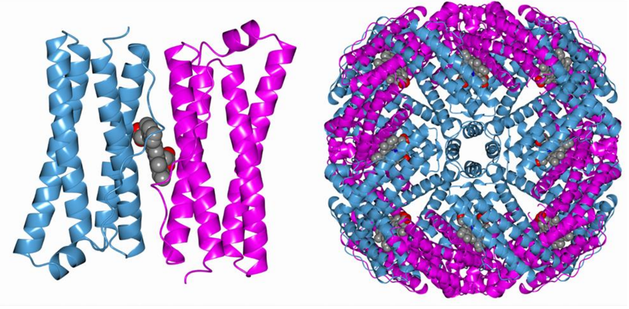
Molecular dynamics (MD) simulations

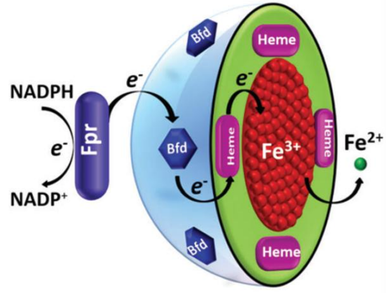
- Modelling protein physical motions over time
  Generates a trajectory of minimized states.
  - Various conformational states sampled
- Can be used to identify dynamic regions near or at a ligand binding pocket

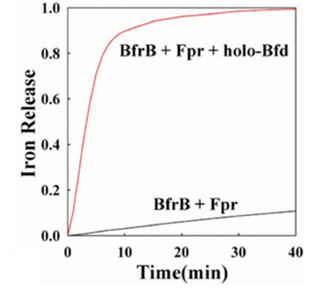


Real world example: Bacterioferritin (BfrB) from *Pseudomonas aeruginosa*

- Iron storage and mobilization
  - Bacterioferritin (Bfr): stores 3,000 Fe ions (Fe^{3+})
  - Fe^{2+} oxidation to Fe^{3+} in the ferroxidase center
  - BfrB subunit (18 kDa): 24 mer assembly
- Bacterioferritin-associated ferredoxin (bfd) binds to Bfr and reduces Fe ions in the core for mobilization



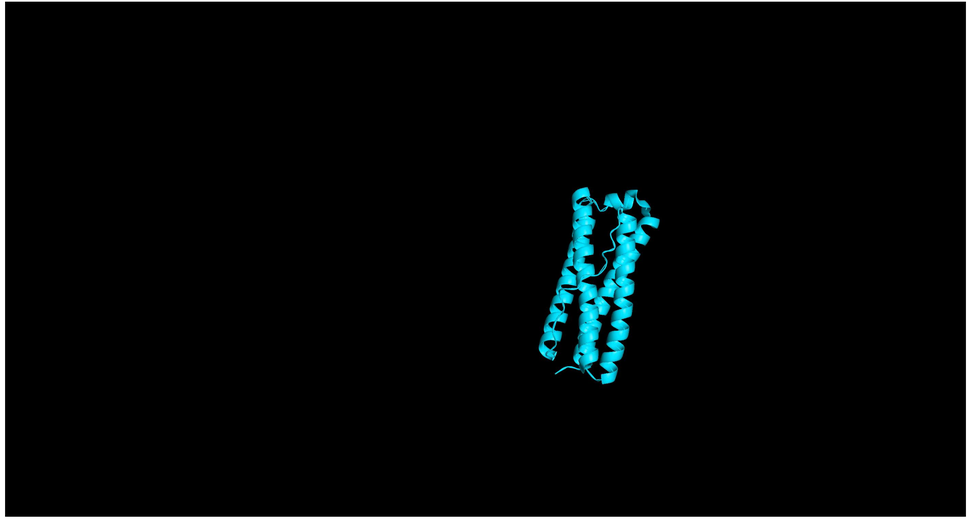




Weeratunga, *etal.*, Biochemistry, 2010.

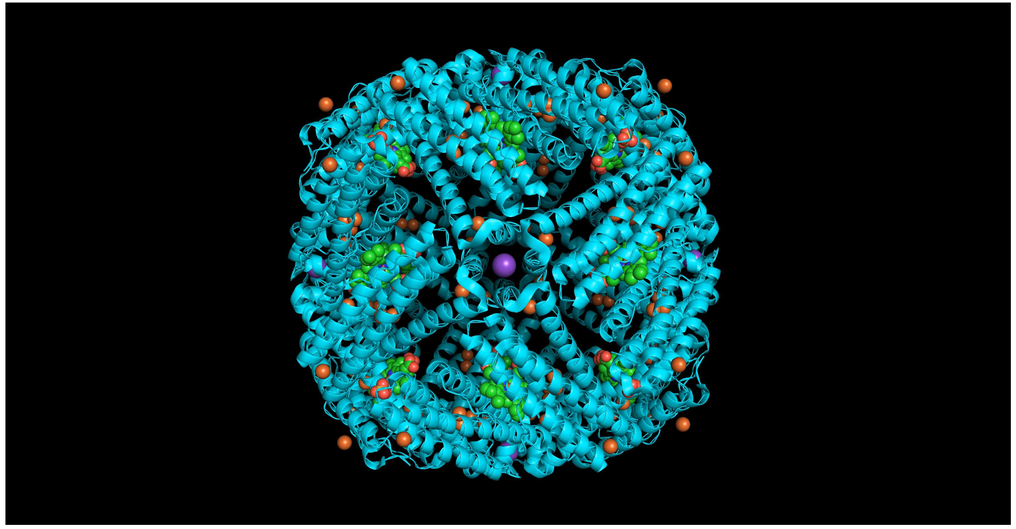
Bfr 24-mer Assembly

- Bfr dimerizes and binds a heme molecule
- The dimers assemble into a 24 meric structure
- The core has a ~90 Å diameter



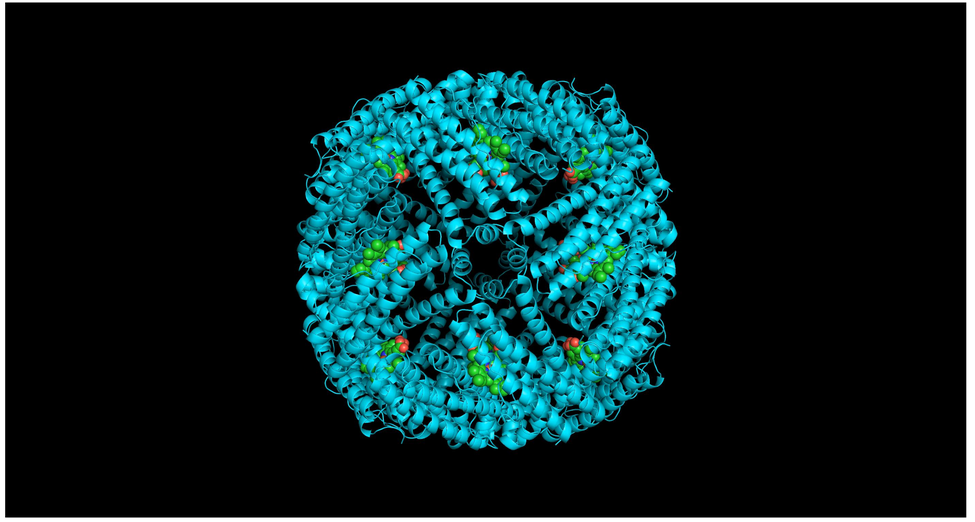
Bfr iron transport to the core

- Fe2+ bind to an external site coordinated by His residues
- The Fe2+ mobilizes to the ferroxidase center of each subunit where oxidation to Fe3+ occurs
- The iron ion then binds to a site at the interior before moving into the core



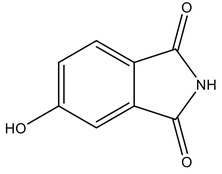
Bfr:Bfd protein complex

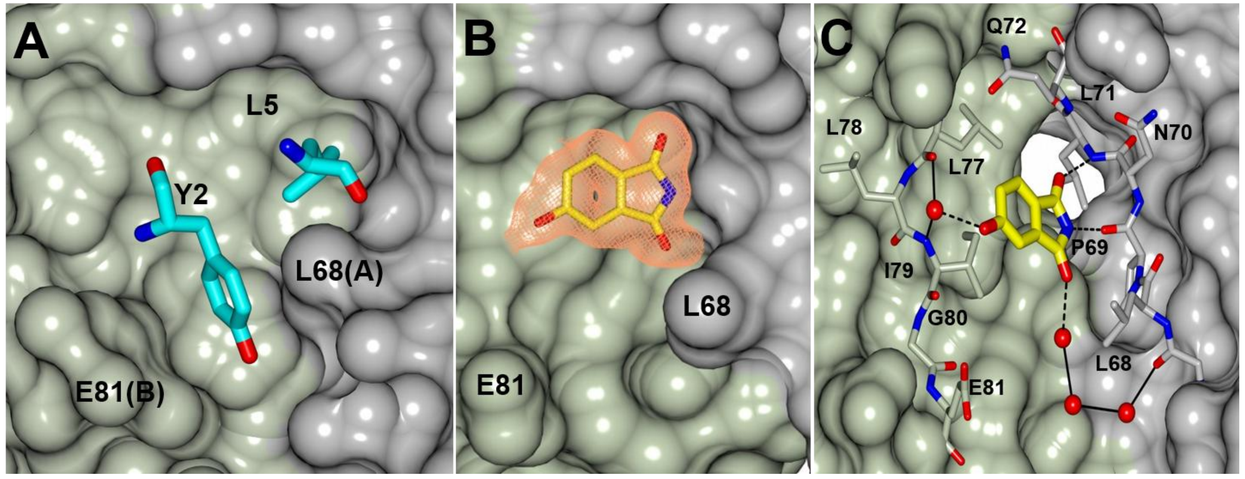
- Bacterioferritin associated ferredoxin (Bfd) binds to a pocket of Bfr
- Reduces Fe^{3+} to Fe^{2+} for mobilization
- Could the pocket be targeted for ligand binding



Computational design of Bfr:Bfd inhibitors

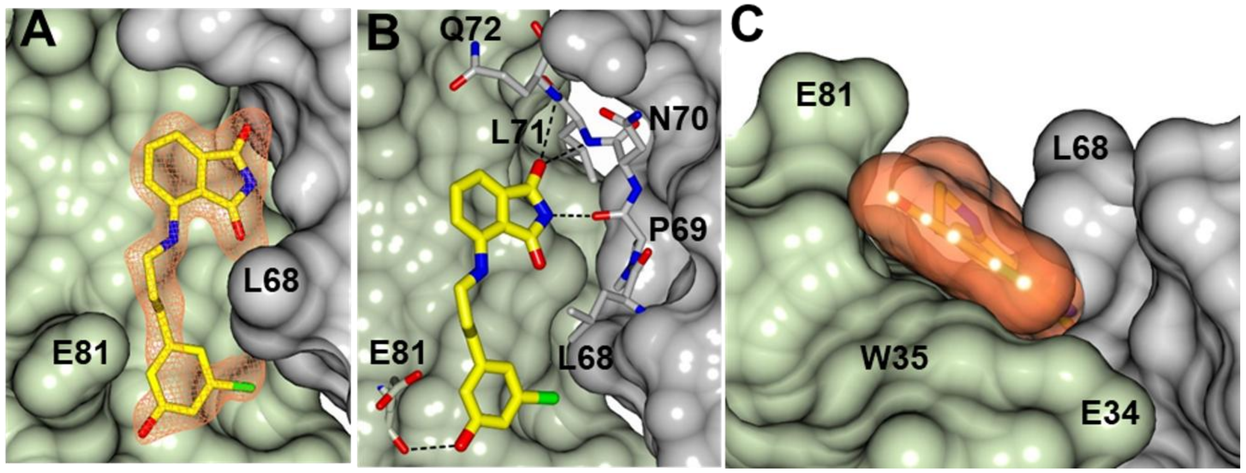
- Y2 of Bfd binds in a pocket formed by L68 and E81 of Bfr subunits in a dimer
- L5 of Bfd binds within a secondary pocket of the Bfr dimers.
- Computational screening (docking) of a virtual library identified 5-hydroxyphthalamide as a top hit





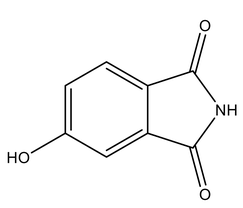
Using a computationally derived hit as the basis for inhibitor development

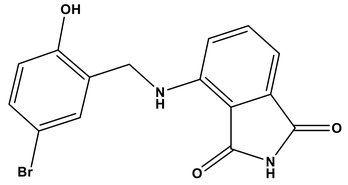
- Adding aryl ring substituents with 1 or 3 carbon linkers improves affinity (6 uM)
- Aryl rings bind in the L68/E81 cleft and mimic Bfd binding

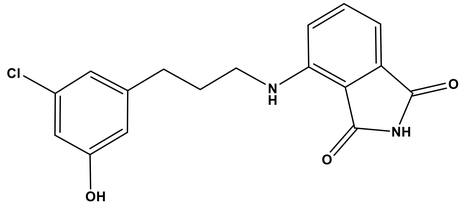


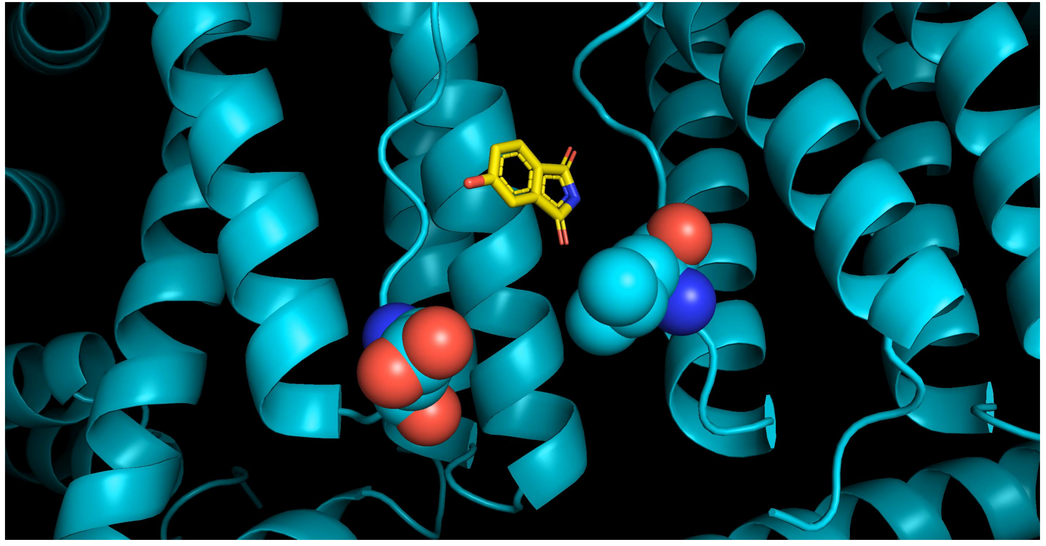
Expanding the computationally derived fragment to potent inhibitors

- The 5-hydroxyphthalamide served as the basis for inhibitor expansion into the aryl ring binding pocket.



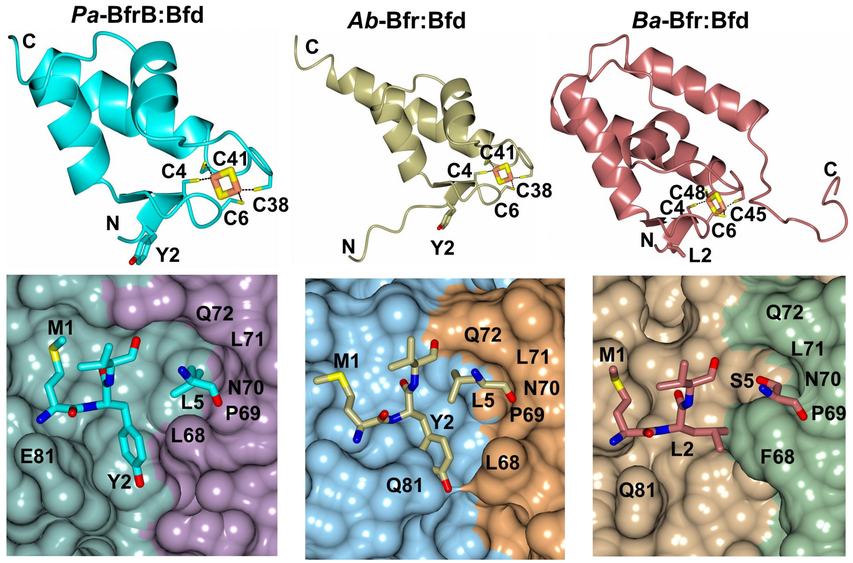






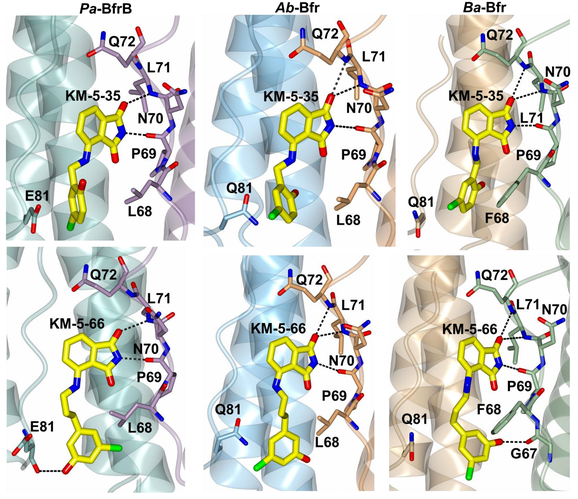
Modelling ligand binding in *Acinetobacter* and *Brucella* Bfr proteins

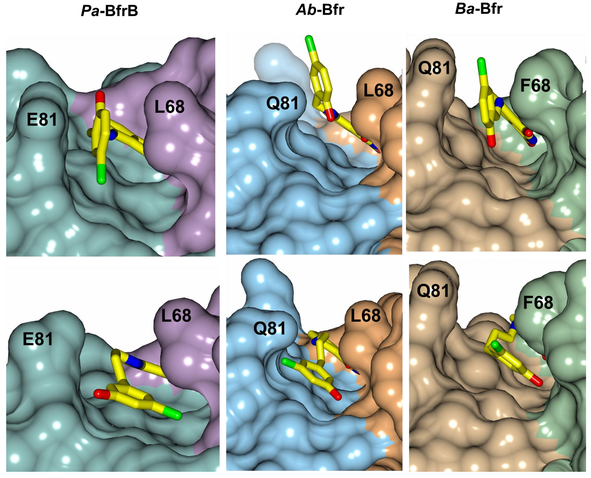
- The Bfd protein structures were modelled using alphafold 3 and fit to the *Acinetobacter* and *Brucella* Bfr crystal structures.
- Similar residues Y2/L5 (*Acinetobacter*) and L2/S5 (*Brucella*) bind to the pockets of each respective Bfr protein.



Docking *P. aeruginoas* inhibitors to *Acinetobacter* and *Brucella* Bfr proteins

- Inhibitors containing a 1 and 3-carbon linker between the phthalimide and aryl ring were docked in the Bfd binding pocket of each Bfr protein
- Similar binding modes are predicted





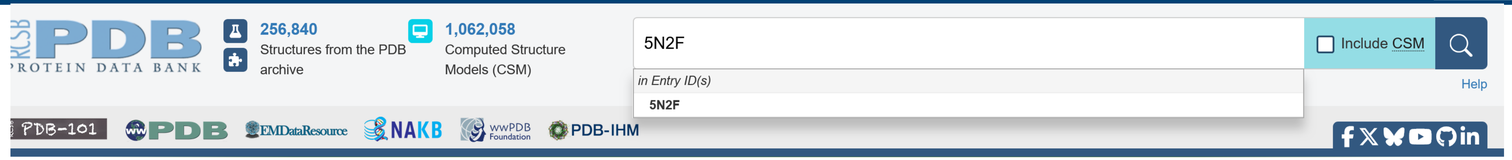
Liu, *etal.*, Proteins, 2026

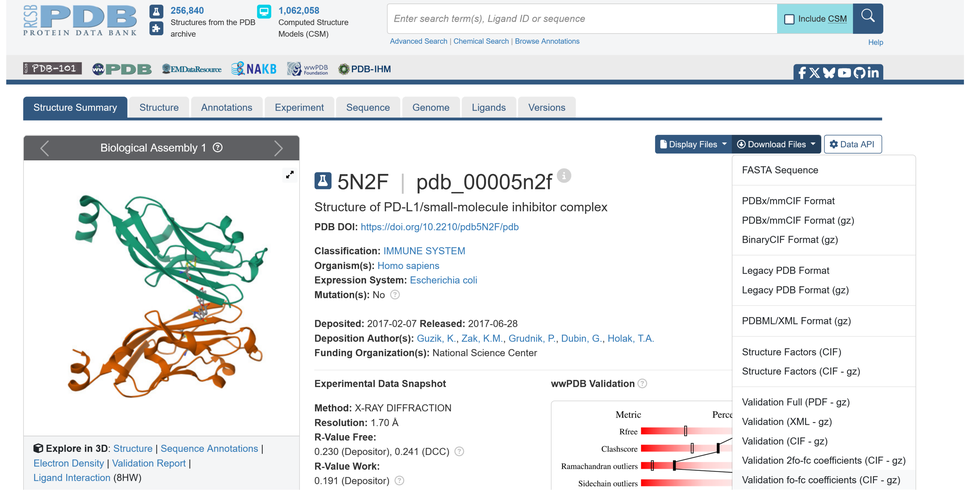
Applications for modelling protein:ligand complexes

- Alphafold 3: Allows modelling of select ligands (eg. NAD, ATP, heme, etc.). DNA, RNA and ions. Small list of ligands to select.
- Autodock vina: Stand alone software for various ligand modelling applications
- Swissdock: Online server that is user friendly. Upload protein and ligand files
- Pymol/docking pie: Separate plugin that provides various docking applications
- DiffDock: AI docking. Some web base interfaces but usually installed locally
- CB-Dock2: Web server based. Automated cavity detection and ligand docking
- Inputs: Need a protein coordinate file and some ligand description file

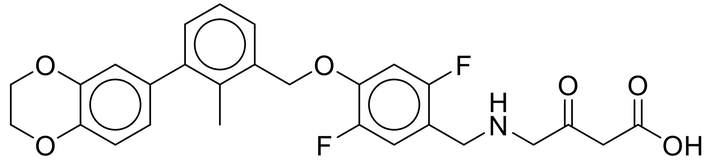
Protein structure file

- PDB file: Downloaded from https://www.rcsb.org/ or https://www.ebi.ac.uk/pdbe/
  - Need the 4-letter code (e.g. 5N2F)
- Alphafold model
- Unpublished crystal structure





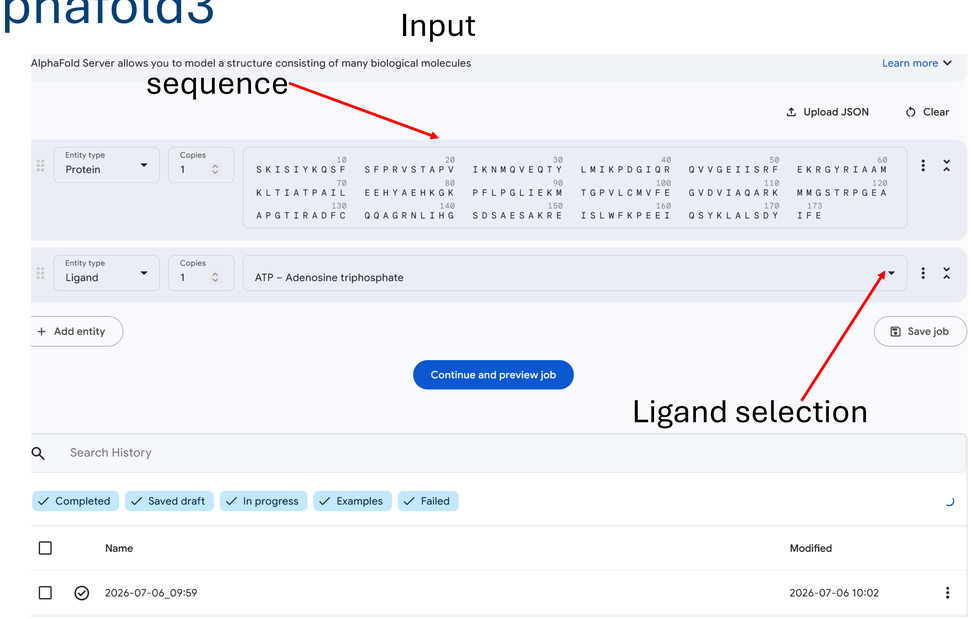
Ligand description file

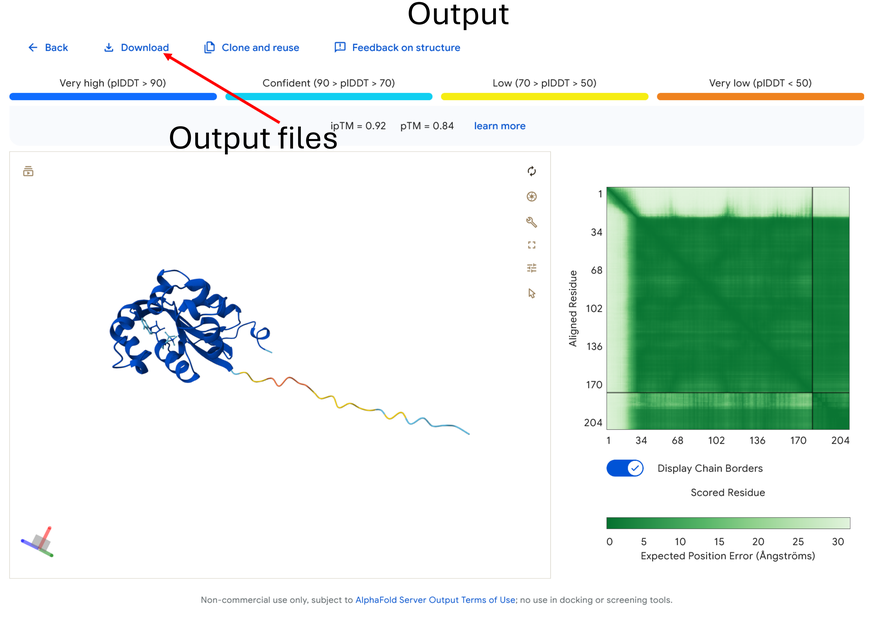


- Structure data file (SDF): text-based coordinates
- Ligand PDB file
- Smiles string: Cc1c(COc2cc(F)c(CNCC(=O)CC(O)=O)cc2F)cccc1c3ccc4OCCOc4c3
- Can be downloaded from the PDB if the ligand is known. (3-letter code: eg. 8HW)
- ChemDraw: Ligands can be constructed and exported in these formats

Structure and ligand with Alphafold3

- The protein sequence is pasted and ligand selected for modelling.
- Output files containing the modeled ligand can be downloaded





Swissdock

- Various options for ligand input
- Protein target added using PDB code or uploading your file
- The search space (grid) is chosen
- Output models can be viewed

SwissDock
SwissDrugDesign

Home About Documentation FAQ Tutorials Command-line Citing Contact **Old version**

Important information: Due to the discontinuation of ChemAxon support for academic websites, SwissDock has been modified to no longer rely on their software.
We sincerely apologize for any inconvenience this transition may cause and thank our users for their understanding.

Docking with Attracting Cavities | **Docking with AutoDock Vina**

Don't know where to start? Try with an example: binding of **SNJ-1715 (CCD ID 0m6) to calpain-1 catalytic subunit (PDB ID 2g8e)**, of **WRR-99 (r99) to cruzipain (1ewl)**, or of **dabrafenib (p06) to B-Raf (5hie)**.

1 - Submit a ligand

Provide a SMILES `N[C@@H](CCSC[C@H]1O[C@H]([C@H](O)[C@@H]1O)n2cnc3c(N)ncnc23)C(O)=O`

. . . or upload a Mol2 file or a PDBQT file [Browse...]

. . . or input, or modify, or check the molecule **using the sketcher**

. . . or use the **advanced search**

✓ Prepare ligand **Reset ligand**

2 - Submit a target

Provide a PDB id *(e.g. 5hie)*

. . . or upload a PDB file or a PDBQT file [Browse...] LemaA.18205.a.B2_AF-Q4QIU2-F1-model_v6.pdb

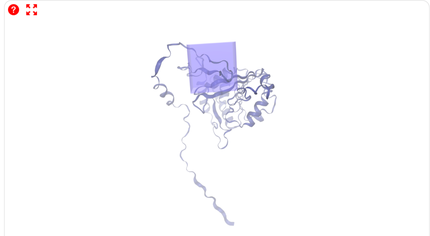
. . . or use the **advanced search**

✓ Prepare target **Reset target**

3 - Define search space

Search box center -8 -13 -4 Å

Search box size 20 20 20 Å



Hydrogen bonds | Ionic interactions | Cation-π interactions | Hydrophobic contacts | π-stacking interactions | Show protein surface | Reset

4 - Select parameters

Sampling exhaustivity 4

✓ Check parameters

This job is estimated to take 0:01:03 (h:mm:ss).

5 - Start docking

Enter an email (optional)

Enter a docking name (optional)

START DOCKING **Reset form**

Query

Ligand	N[C@@H](CCSC[C@H]1O[C@H]([C@H](O)[C@@H]1O)n2cnc3c(N)ncnc23)C(O)=O
Target	LemaA.18205.a.B2_AF-Q4QIU2-F1-model_v6.pdb
Method	AutoDock Vina
Date	July 6, 2026, 5:21 pm UTC

Parameters:

Box center: 11 - 12 - -4 Sampling exhaustivity: 4

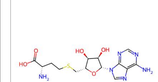
Box size: 22 - 22 - 22

If you publish these results, please, cite the following papers:

Bugnon M, Röhrig UF, Goullieux M, Perez MAS, Daina A, Michielin O, Zoete V. SwissDock 2024: major enhancements for small-molecule docking with Attracting Cavities and AutoDock Vina. *Nucleic Acids Res.* **2024**

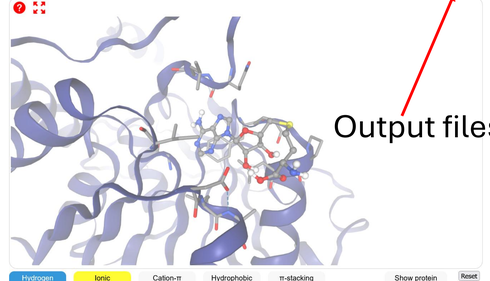
Eberhardt J, Santos-Martins D, Tillack AF, Forli S., AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model.*, **2021**

Ligand



Results

Export your results:



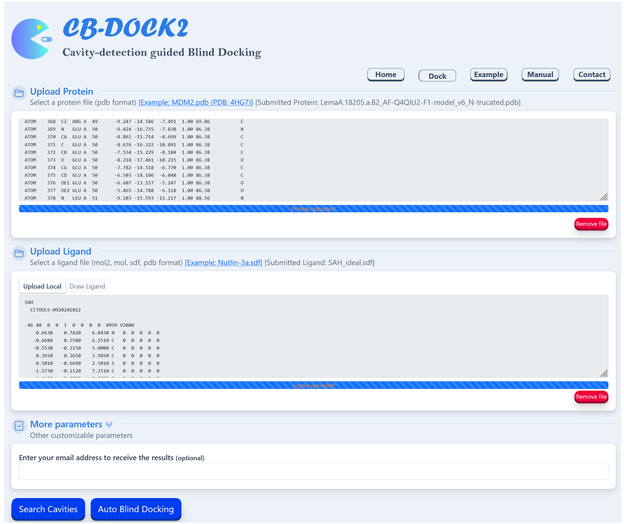
Hydrogen bonds | Ionic interactions | Cation-π interactions | Hydrophobic contacts | π-stacking interactions | Show protein surface | Reset

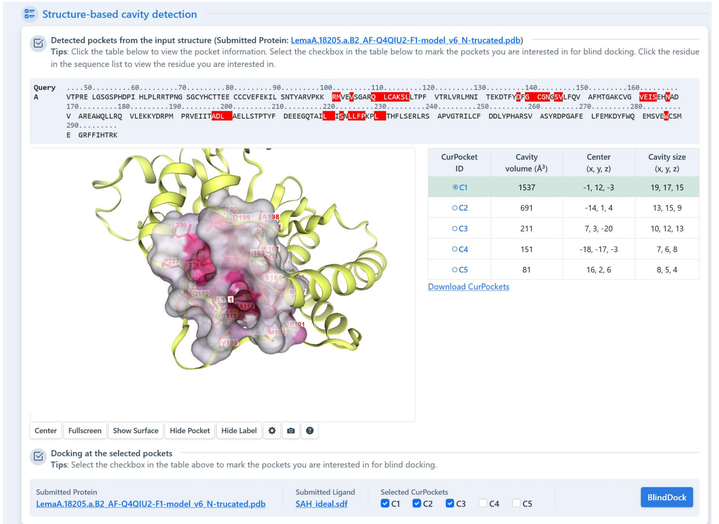
Output files

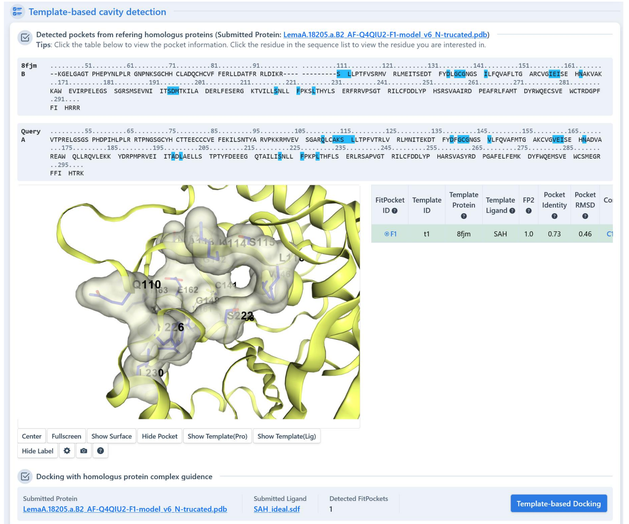
Model	Calculated affinity (kcal/mol)
1	-5.190
2	-5.001
3	-4.873
4	-4.870
5	-4.797

CB-Dock2

- Protein and ligand files are uploaded and prepared automatically
- Cavities are detected
- Docking within the detected cavities

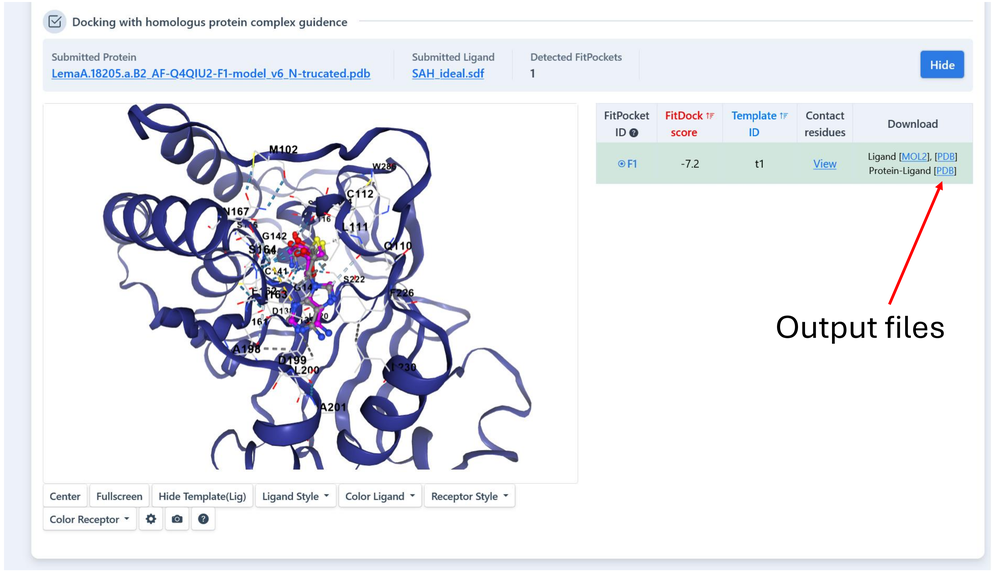






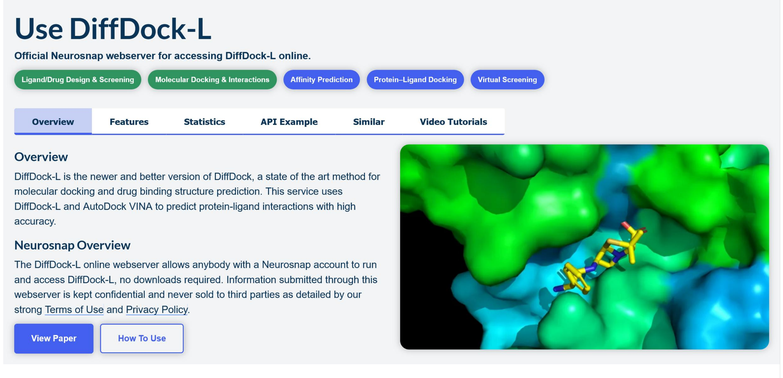
CB-Dock2 Results

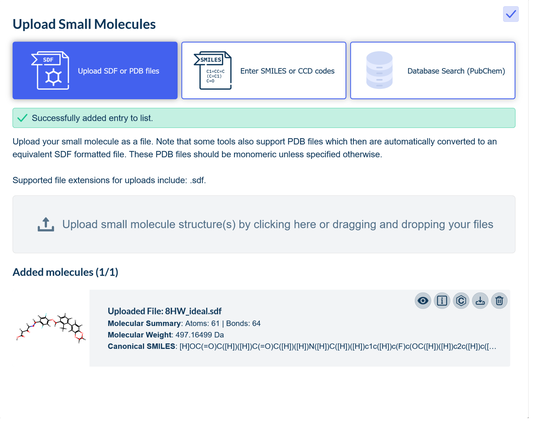
- Interactive analysis of the ligand modelling
- A PDB file containing the protein:ligand can be downloaded



DiffDock-L though Neurosnap website

- Fee for use (credits)
- Various computational applications



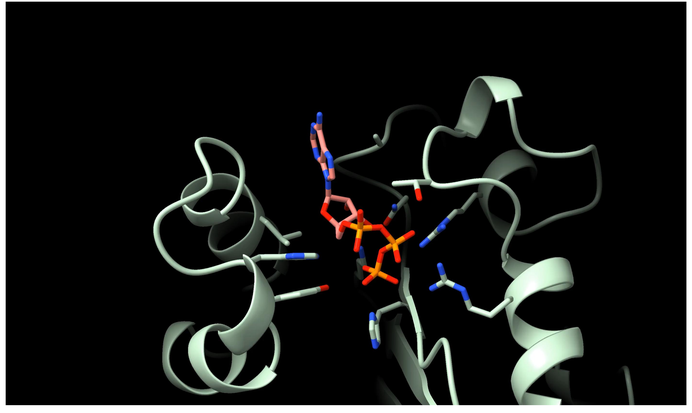


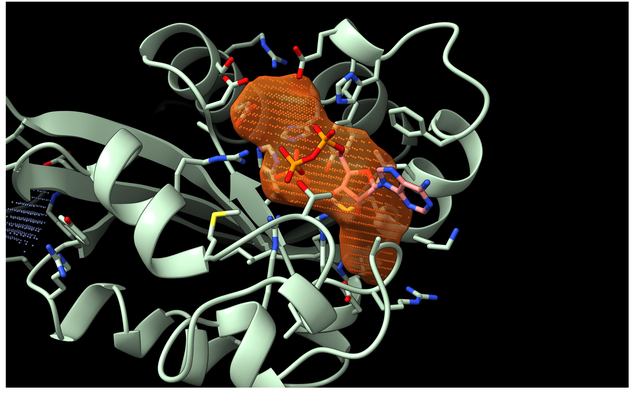
DiffDock-L output

- Results contains 100 ligand models ranked from top to bottom score
- Ligand poses can be viewed though a graphic interface
- Download all or selected models

Protein:ligand complex modelling examples

- Example 1. Nucleoside-diphosphate kinase from *Cryptosporidium parvum*
  - Catalyzes the reversible exchange of phosphates from NTP to NDP (XDP + YTP ↔ XTP + YDP)
  - Uniprot Q5CR64
  - Modelling ATP and GTP bound states using Alphafold 3



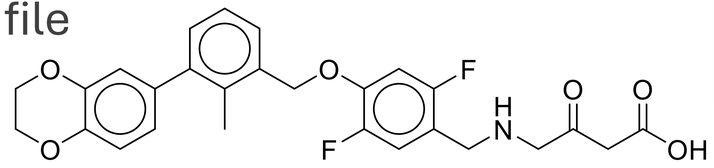


Example 2

- Glyceraldehyde-3-phosphate dehydrogenase from *Neisseria gonorrhoeae*
- Uniprot B4RPP8
- Involved in glycolysis
  - glyceraldehyde-3-phosphate + NAD^+ + Pi $\longleftrightarrow$ 1,3-bisphosphoglycerate + NADH
- Modelling NAD bound form using Alphafold 3

Example 3

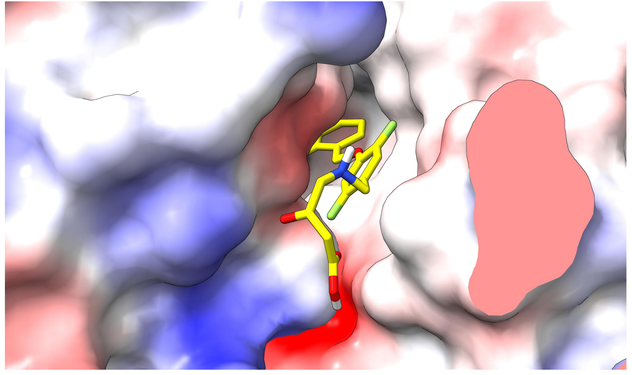
- Programmed cell death-ligand 1 (PD-L1)
- Uniprot Q9NZQ7
- Modelling a novel inhibitor with CB-Dock 2
- Protein model from the PDB input
- Ligand smiles or sdf file



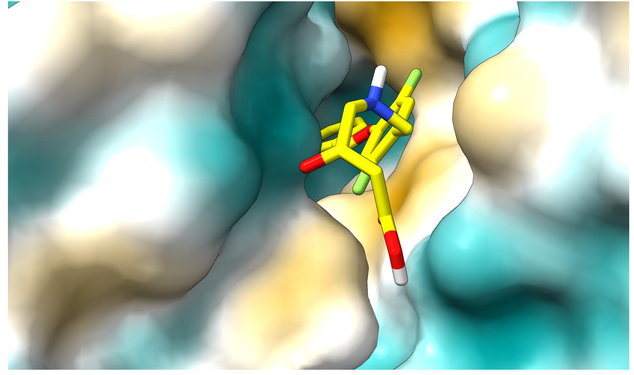
ChimeraX commands for ligand analysis

- ChimeraX commands:
- View ligand (centers on ligand molecule)
- Find cavities (Need PyKVFinder plugin installed)
- hbonds ligand (show hydrogen bond contacts to the ligand)
- coulombic protein (show electrostatic surface of the protein)

electrostatic surface



Lipophilicity surface potential



Questions