

Project 1. Computer-Aided Drug Design

1. Download `project1.tar.gz` from <https://mulan.swmed.edu/mmms>

2. Select a protein system (Thrombin was used as an example in class)

Protein Class	PDB Code	$-\log k_d$	Resolution
Neuraminidase	2QWG	8.4	1.8
DHFR	1DHF	7.4	2.3
L-arabinose	1ABE	6.52	1.7
Thrombin	1A5G	10.15	2.06
Human oxresin receptor 1	4ZJ8	~10	2.75

1) A clean version of both protein receptor and ligand have been prepared (/home/mmms/project1)

3. Perform 2D-similarity search

- Download a sdf file of the ligand from RCSB Protein Databank (<http://www.rcsb.org/pdb/home/home.do>)
- Perform Similarity searches with OpenBabel. Adjust similarity cutoff (using flag `-at cutoff`) so that 1 to 50 thousands compounds are identified as hits

2) Modify `search.bat` in `Step2_simsearch` and run `search.bat`

3) Dissect `mol2` files by running "`bat`" in the `mol2` subdirectory

4. Prepare necessary files for molecular docking with AutoDock vina

- `Ligand.pdbqt`
- `Receptor.pdbqt`
- `Configure.txt`

MGL maybe used to facilitate the preparation of those files.
Configure.txt should be manually prepared.

5. Docking screening

- Run docking jobs from ajax (129.112.108.164)

Login: `ssh -Y username@129.112.5.8`

`sftp username@129.112.5.8`

- 1) *Run prep.bat in Step3_ligprep to copy files*
- 2) *Go to mol2 subdirectory and run "bat" to prepare ligands*
- 3) *Go to Step4_dock, and ren "prep.bat"*
- 4) *Go to ligdb subdirectory, and run "bat" to prepare subdirectories for multiple-task docking*
- 5) *Go to Step4_dock, run "submit.bat" to submit jobs*

6. Analyze docking results

- Prioritize compounds based on docking scores
- Examine docking poses with PyMOL

7. Write report

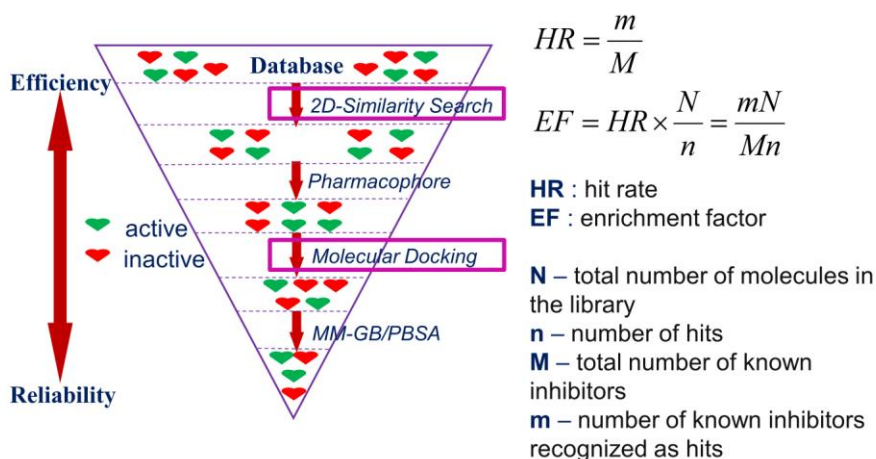
Calculate HR, EF and plot enrichment curves if known inhibitors are included in the screening library. One could search the known inhibitors for specific biological target from ZINC target subset (<http://zinc.docking.org/browse/targets/>).

- Hit rate (HR)
- Enrichment factor (EF)
- Enrichment curve

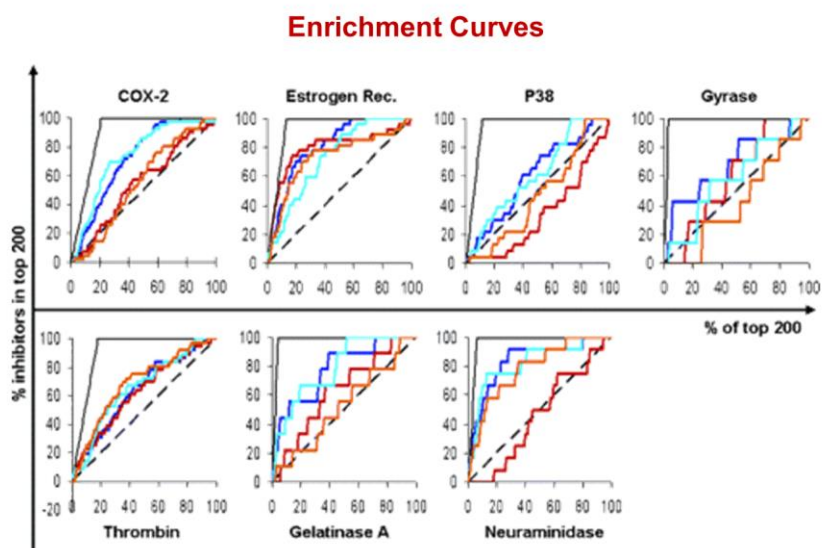
Suggestion: copy the command scripts, input files from the 1A5G subdirectories (/home/mmms/project1/1A5G).

Basic Strategy: two filters (2D-similarity search and modeling docking) are used in hierarchical virtual screenings.

Lead Identification Through Virtual Screening Using A Set of Hierarchical Filters



Sample Enrichment Curves:



random selection (black, dashed)
 ideal performance (black, solid)