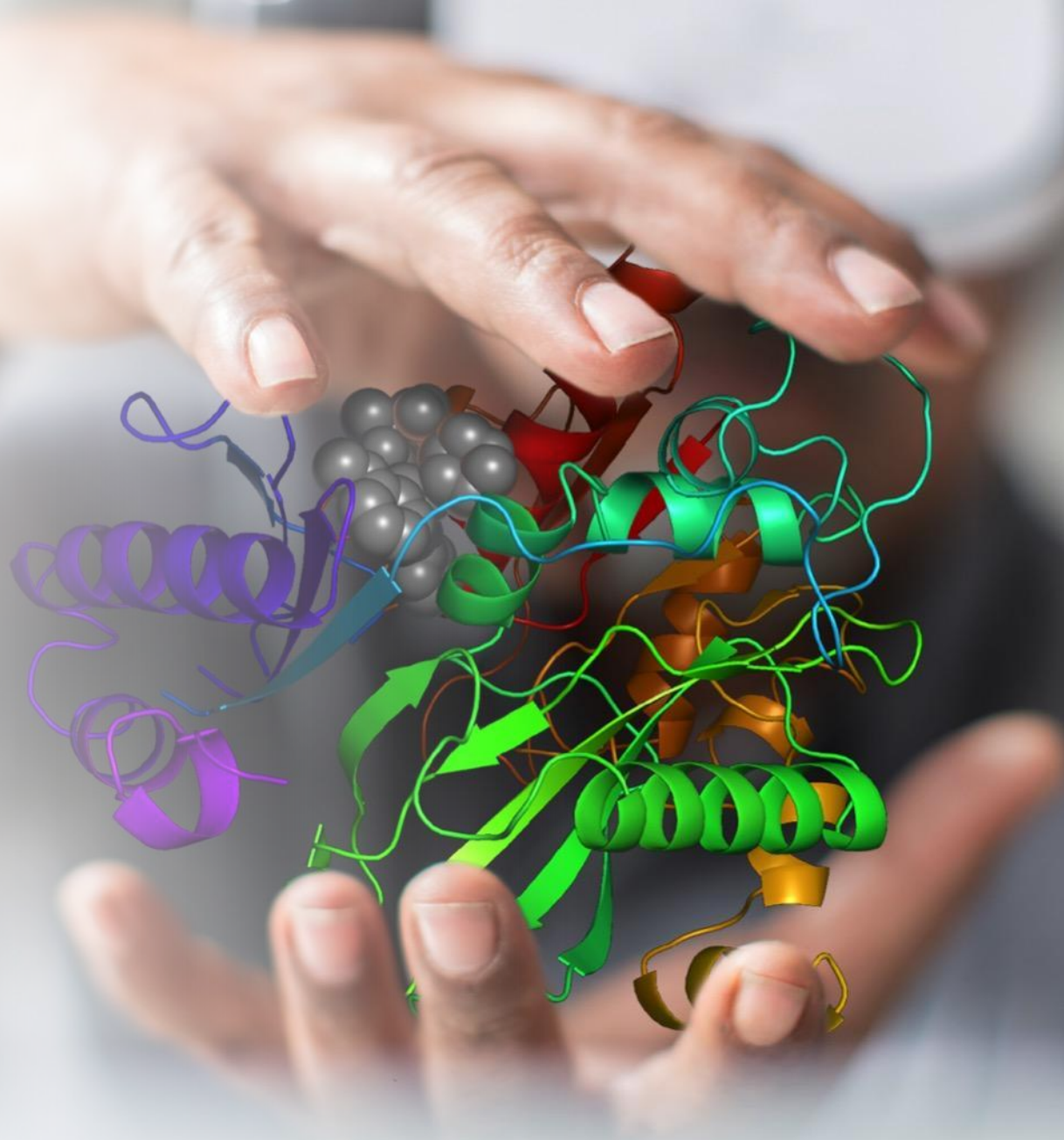
SCHRÖDINGER®

Modeling Protein-Ligand Poses for Hit Identification

Additional Slides

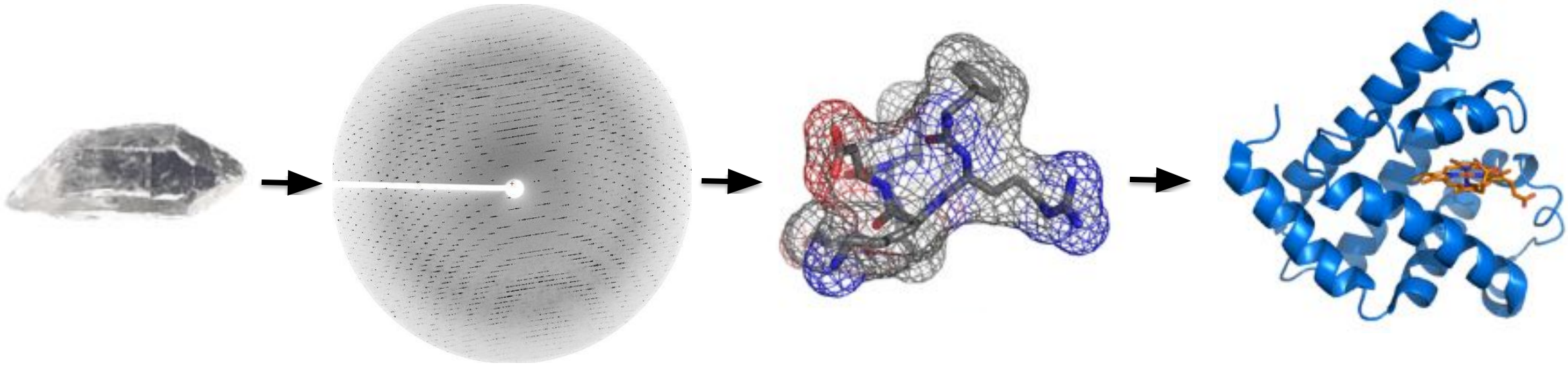


Extra Slides for More Information related to Virtual Screens

SCHRÖDINGER

Transforming drug discovery and materials research

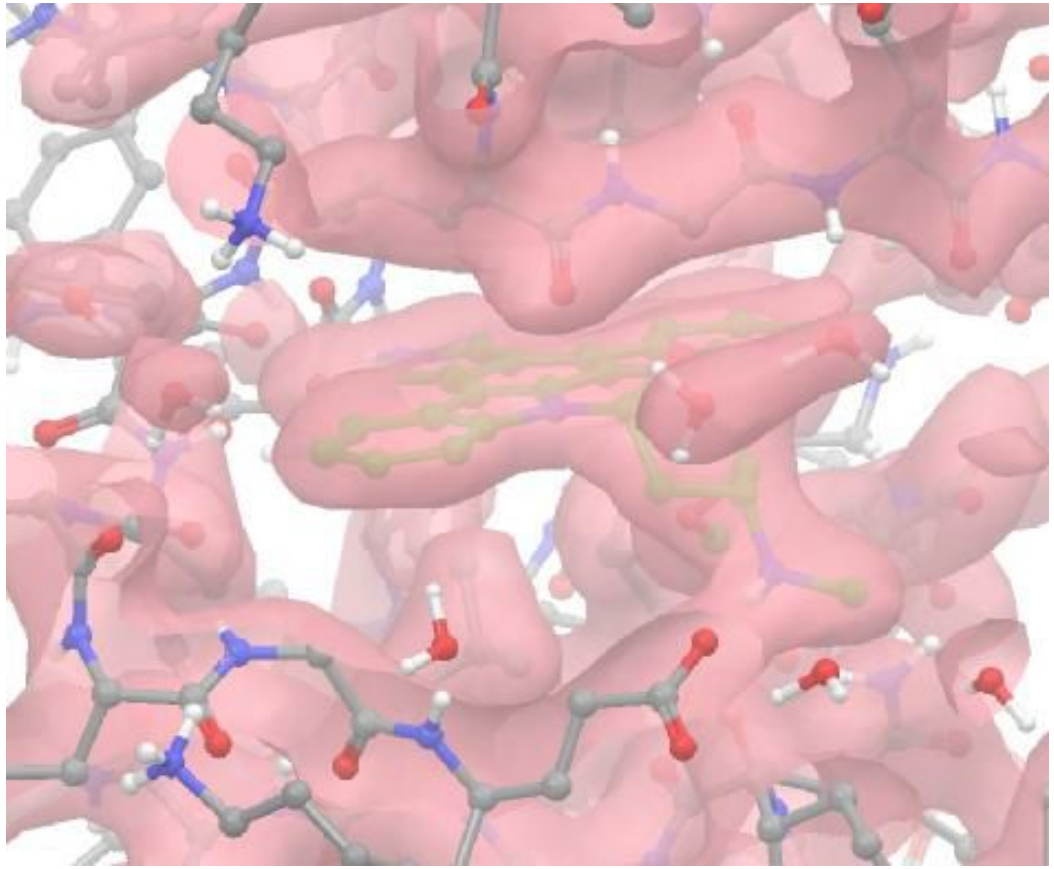
Most SBDD projects utilize crystal structures



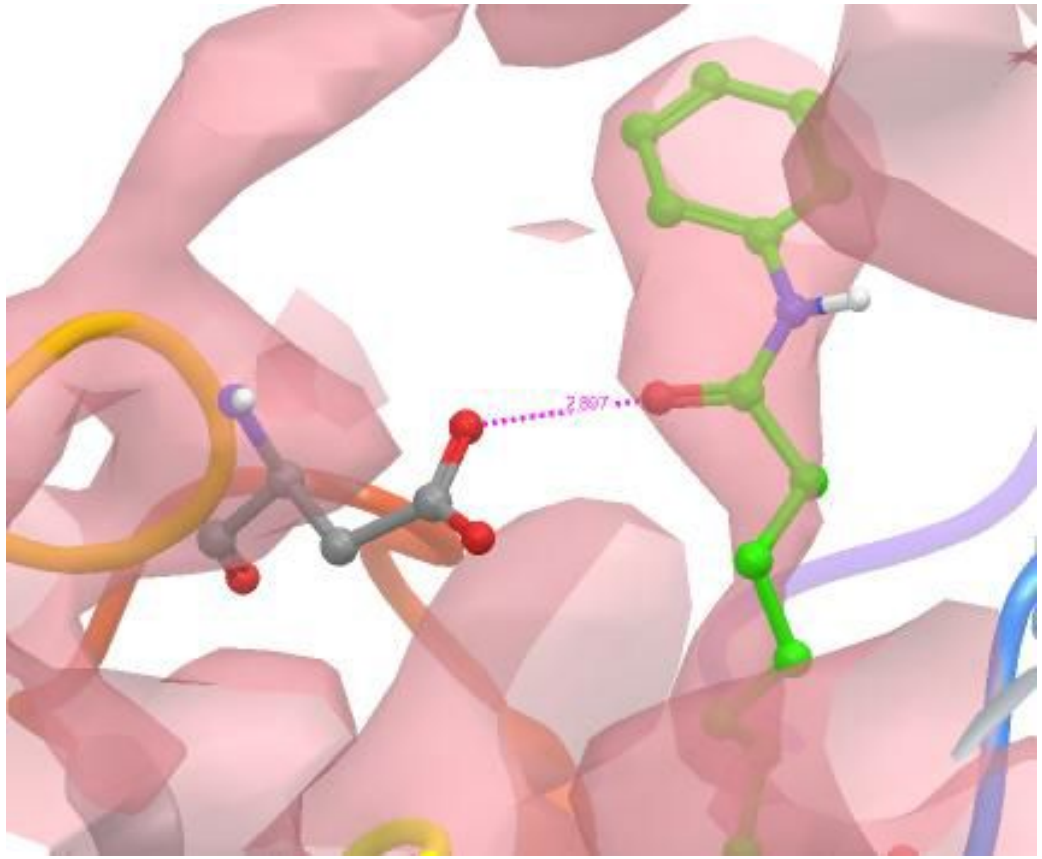
protein crystal → diffraction pattern → electron density map → atomic model

Adapted from: http://www.scistyle.com/

Not all crystal structures are equal



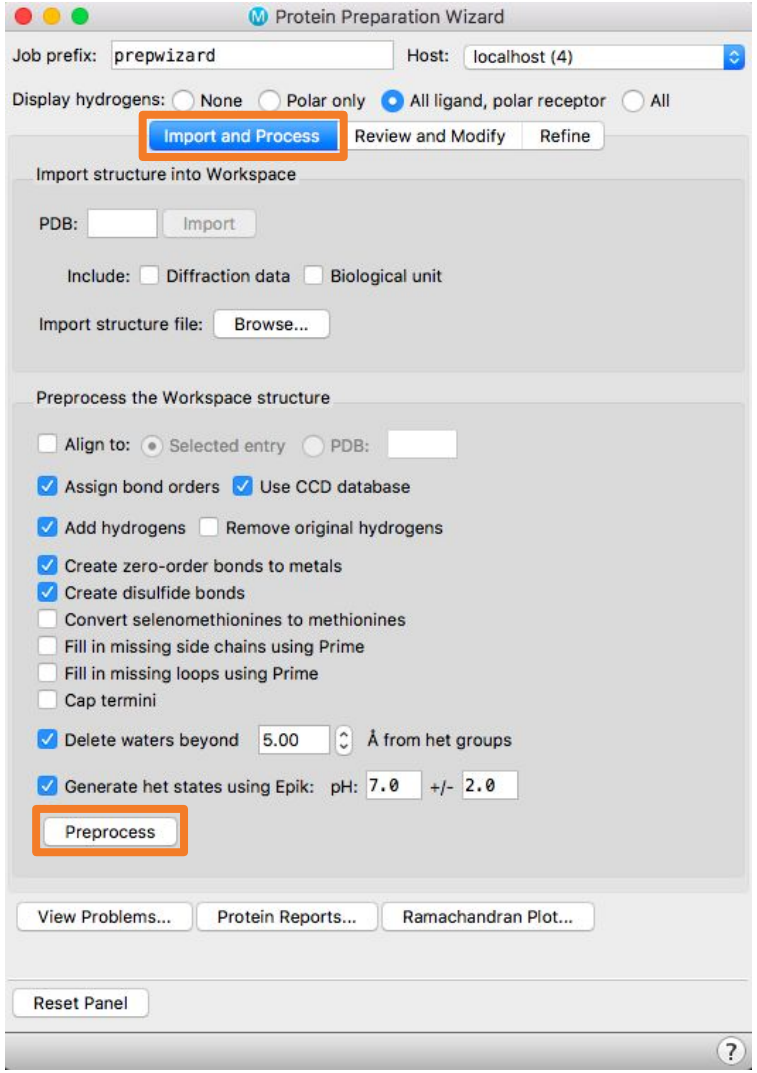
In this case, the ligand density is relatively unambiguous.



In this case the density is missing, which may result in misleading information.

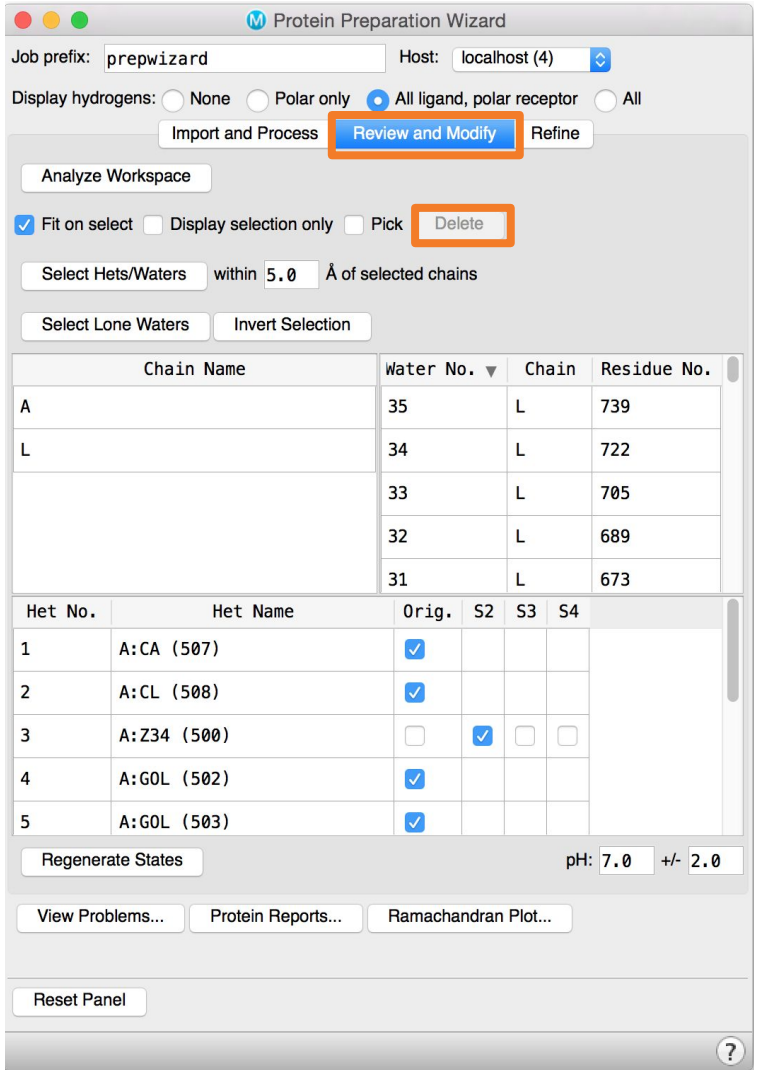
The Protein Preparation Wizard prepares structures for modeling

- **Fix common problems**
  - –Protonation
  - –Missing side chains
  - –Missing loops
- **Remove unwanted molecules**
  - –Counterions, artifacts of crystallography, waters
  - –Biologically relevant?
- **Optimize your structure**
  - –Hydrogen-bond optimization
  - –Restrained minimization



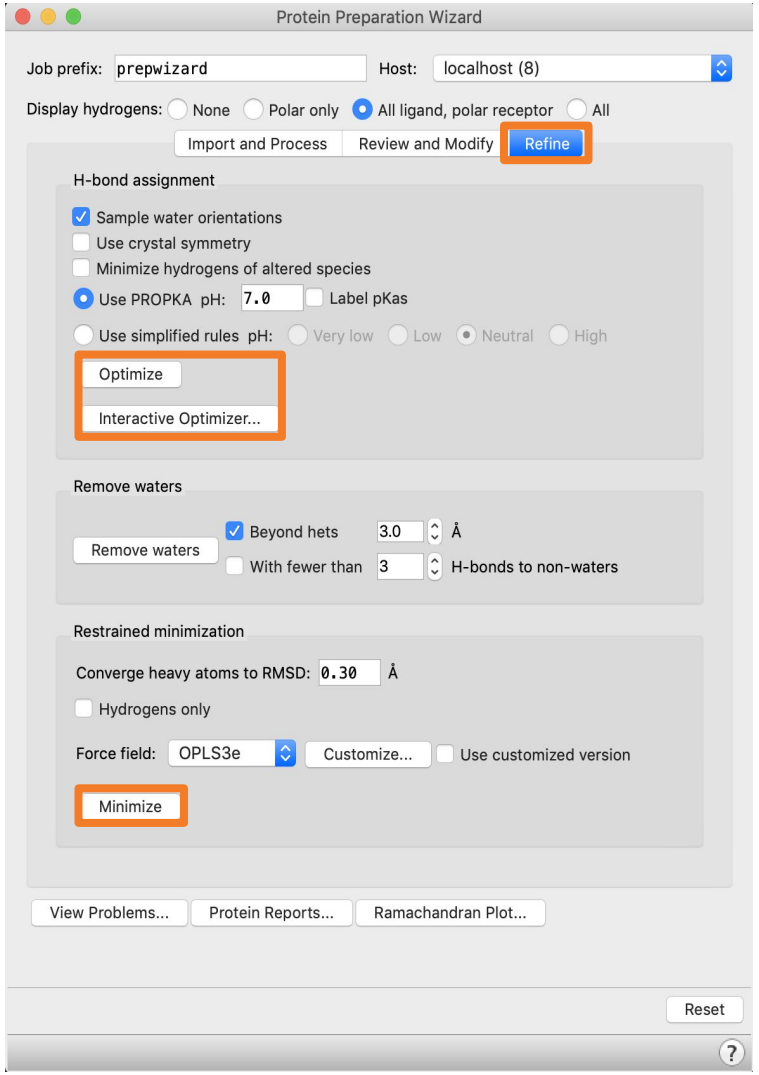
The Protein Preparation Wizard prepares structures for modeling

- **Fix common problems**
  - –Protonation
  - –Missing side chains
  - –Missing loops
- **Remove unwanted molecules**
  - –Counterions, artifacts of crystallography, waters
  - –Biologically relevant?
- **Optimize your structure**
  - –Hydrogen-bond optimization
  - –Restrained minimization



The Protein Preparation Wizard prepares structures for modeling

- **Fix common problems**
  - –Protonation
  - –Missing side chains
  - –Missing loops
- **Remove unwanted molecules**
  - –Counterions, artifacts of crystallography, waters
  - –Biologically relevant?
- **Optimize your structure**
  - –Hydrogen-bond optimization
  - –Restrained minimization



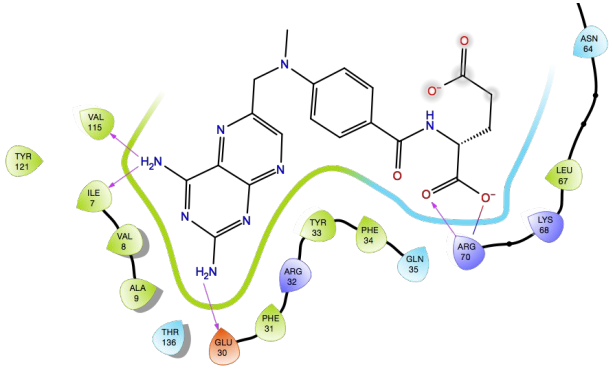
Docking against a well-prepared protein structure is vital

- Structure quality deeply impacts the calculations
  - X-Ray preferred (lower angstrom resolution is typically better, i.e. 2Å is better than 3Å)
  - Homology models can work
  - NMR and Cryo-EM structures are getting better
- Recommendations
  - Use of Protein Preparation Wizard that walks through the preparation process
  - Use of impref = 0.30 to minimize contraction of binding site
  - Retain close crystal waters through complex minimization and then remove them prior to grid generation though **this can be target dependent**

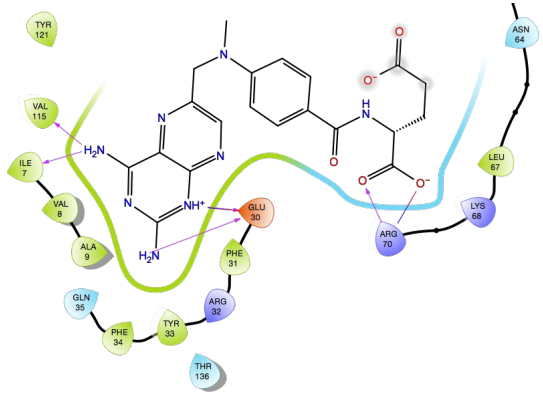
Required inputs for protein-ligand docking: ligands

- Glide will only dock ligand states that are provided
- Recommendations for prepared ligand structures
  - Use LigPrep to generate low energy ionization/tautomeric states for ligands
  - Epik state penalties that estimate free energy required to generate ionization state in water with corrections for interaction with metal sites
  - Typical expansion of compounds by ionization/tautomeric/stereo expansion is 2.5x
  - Increase or decrease pH value and +/- range depending on target physiological location and project goals

State penalty=0.0 kcal/mol



State penalty=1.43 kcal/mol



Methotrexate bound to DHFR (1U72)

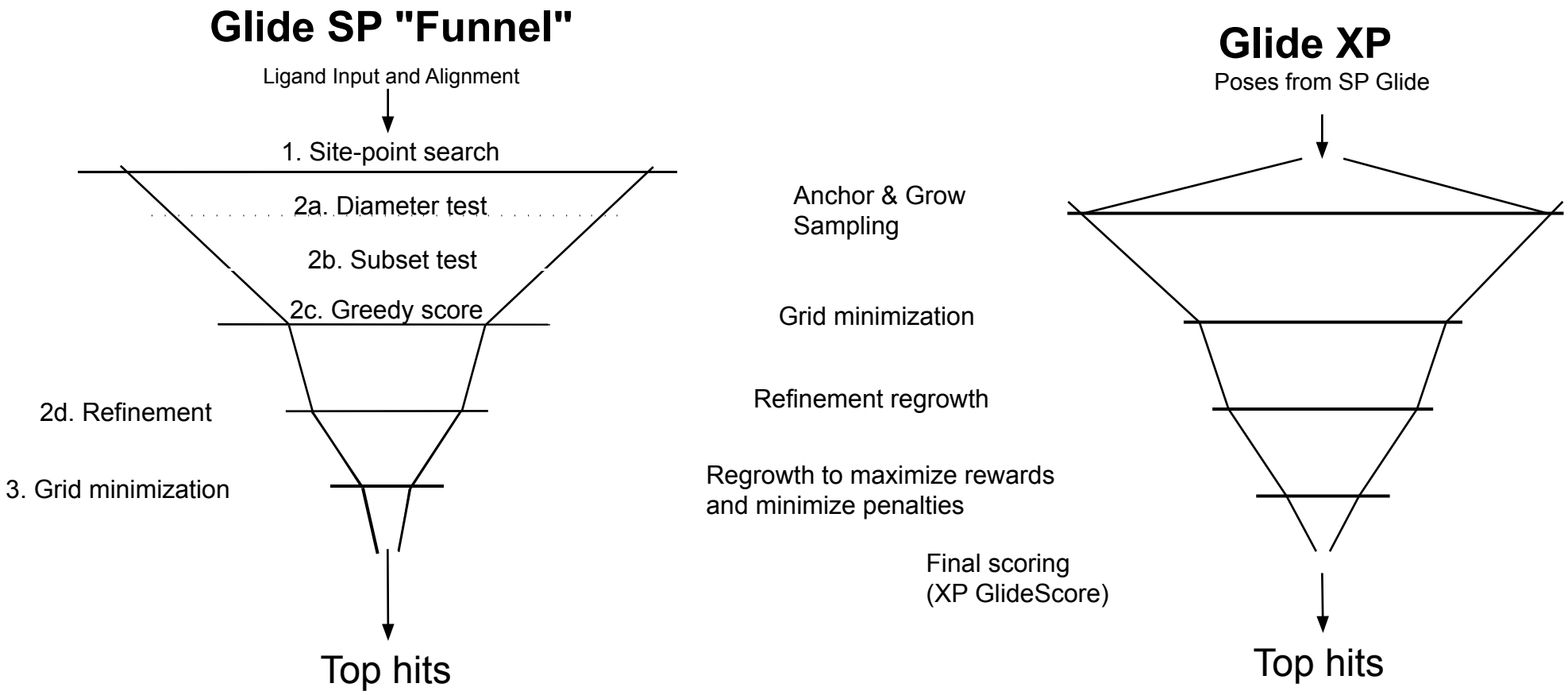
Glide has different scoring functions

Scoring Function	Computing Time	When to Use
SP	5 – 20 sec/molecule	**First pass** virtual screening on large databases/hit generation
XP	3-5 min/molecule	**Refinement** of a smaller dataset for lead optimization

- SP seeks to minimize **false negatives** while XP seeks to minimize **false positives**
- The XP scoring function includes more stringent terms for modeling desolvation, hydrophobic effects, and charged interactions

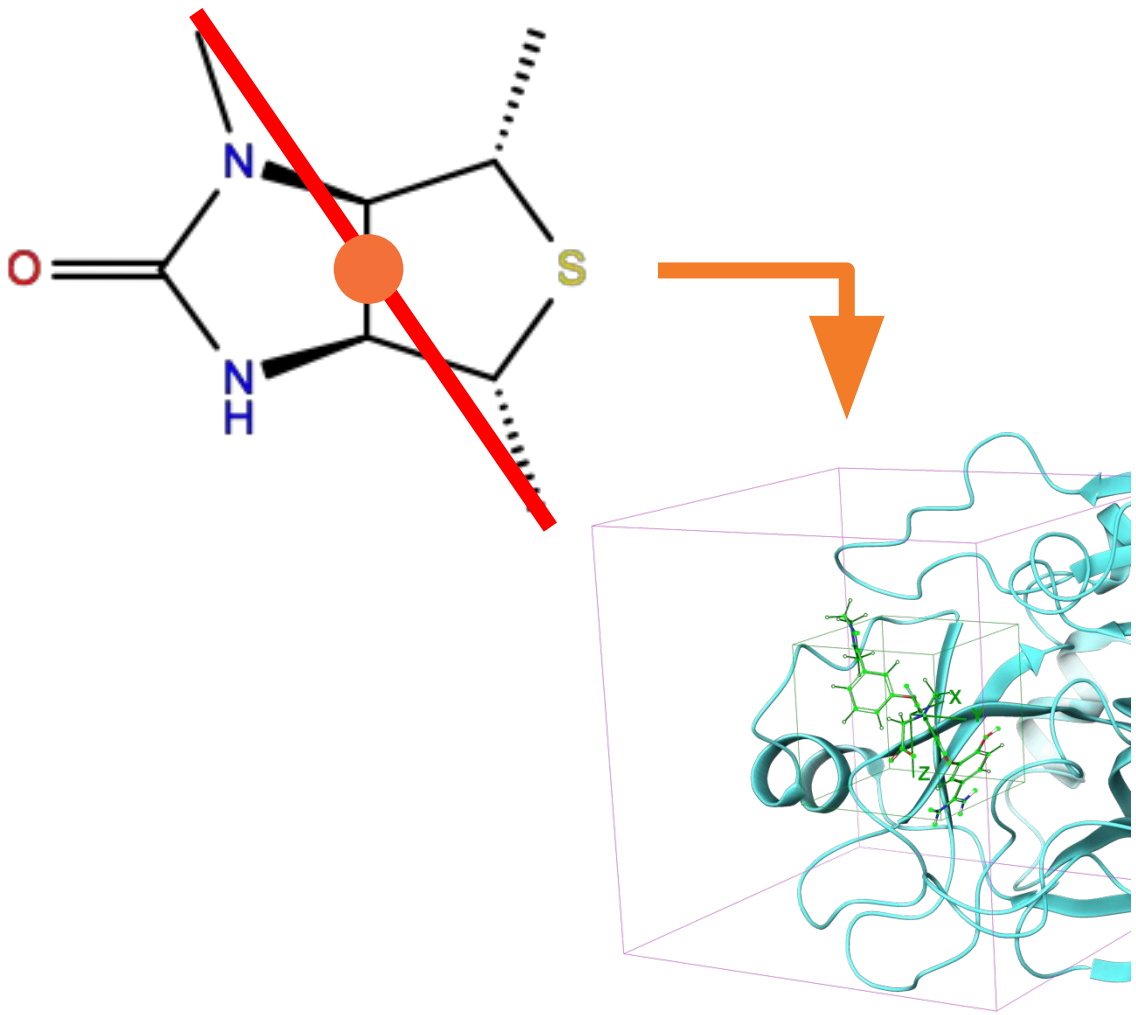
Friesner, R. A. et al. *J. Med. Chem.* **2004**, *47*, 1739-49.

SP versus XP Glide scoring



SitePoint search

- Draw ligand diameter line between two most widely separated atoms in core
- Place midpoint of ligand diameter on all sitepoints (within green box)
- Compare precomputed histograms of distance from sitepoint to receptor with histogram of distance from ligand center to ligand surface



SP GlideScore – An empirical scoring function

$$GlideScore = c * E_{coul} + c * E_{vdW} + E_{lipo} + E_{Hbond} + E_{metal} + E_{rotb} + E_{polar\_phob} + E_{rewards}$$

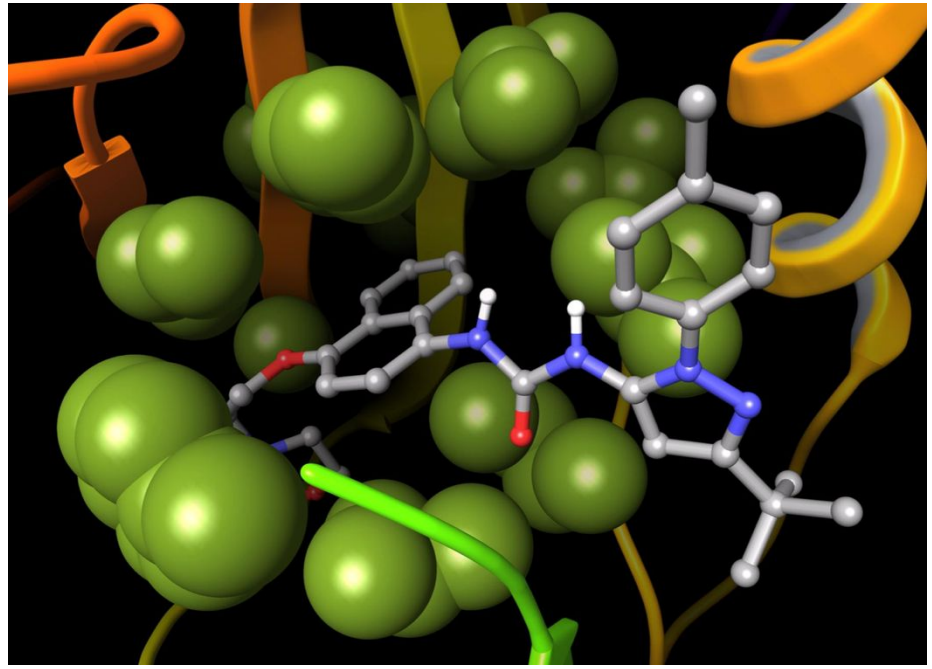
- E_{coul} – gas phase coulomb interaction energy with net ionic contributions reduced by ~50% except for anionic metal-ligand interactions
- E_{vdw} – 6-12 vdW energy with ligand vdW for non-polar interactions reduced by 20%
- E_{lipo} – pairwise hydrophobic interaction term
- E_{hbond} – rewards formation of hydrogen bonds
- E_{metal} – rewards formation of interactions with metal ion
- E_{rotb} – Penalty to avoid correlation of GS with MW
- E_{polar_phob} – Reward non-hbond ligand atom in hydrophobic environment
- $E_{rewards}$ – Primarily XP rewards for phobic enclosure and phobically enclosed H-Bonds also pi-cation rewards

Hydrophobic term

- Generally dominant favorable term in GlideScore

$$E_{phobic\_pair} = \sum_{ij} f\left(r_{ij}\right)$$

- Models effect of displacing receptor water molecules near hydrophobic protein atoms and removing hydrophobic regions of the ligand from water via hydrophobic contacts
- Term includes hydrophobic enclosure recognition



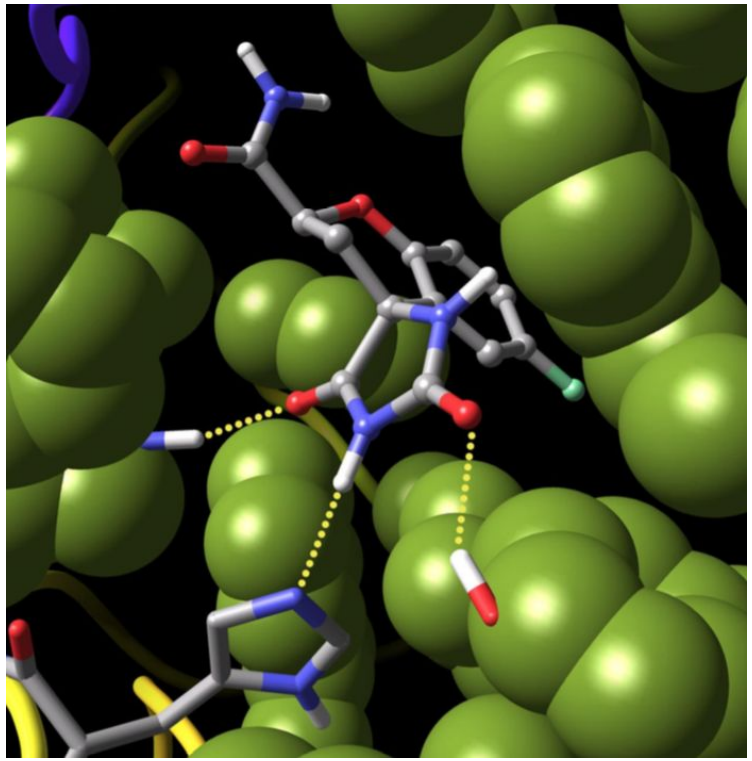
Active for 1kv2 bound to p38. The naphthyl group receives a -4.5 kcal/mol enclosure reward

Hydrogen bond terms always contribute as a reward

- Always favorable

$$E_{HB} = \sum g(\Delta r) h(\Delta \alpha)$$

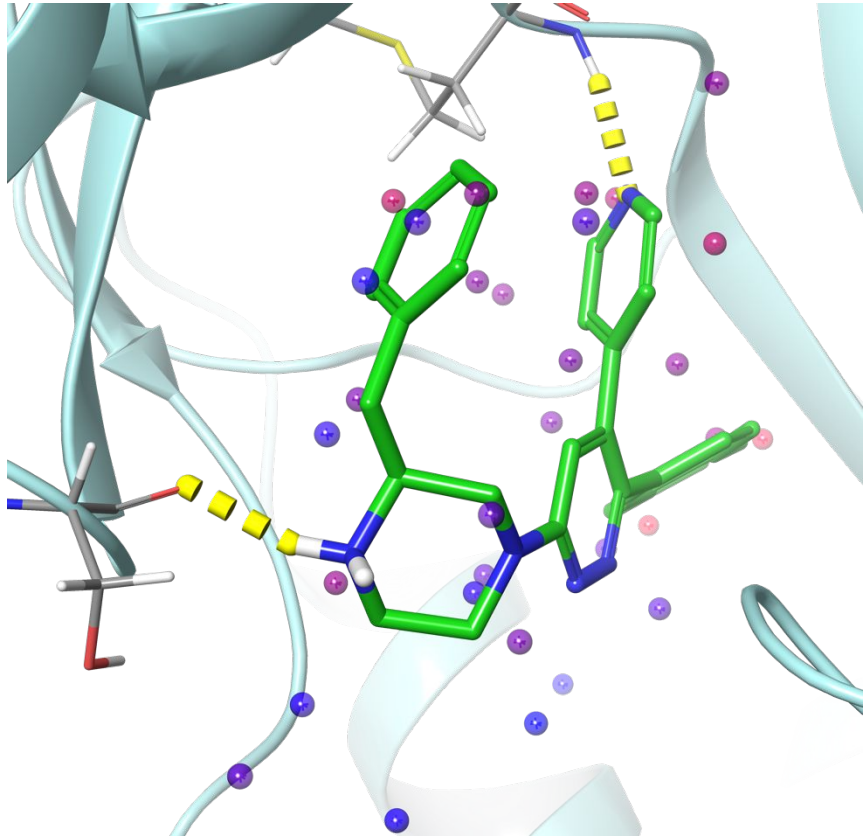
- Differentiates between neutral-neutral, charge-neutral, and charge-charge hydrogen bonds
  - Default maximum reward for formation of hydrogen bonds are -1.0, -0.5, and 0.0 for neutral-neutral, charge-neutral, and charge-charge
- Recognizes formation of single and correlated hydrogen bonds in regions of hydrophobic enclosure



Fidarestat bound to aldose reductase. The triplet of enclosed hydrogen bonds contributes -5.0 kcal/mol

Desolvation Energy Treatment

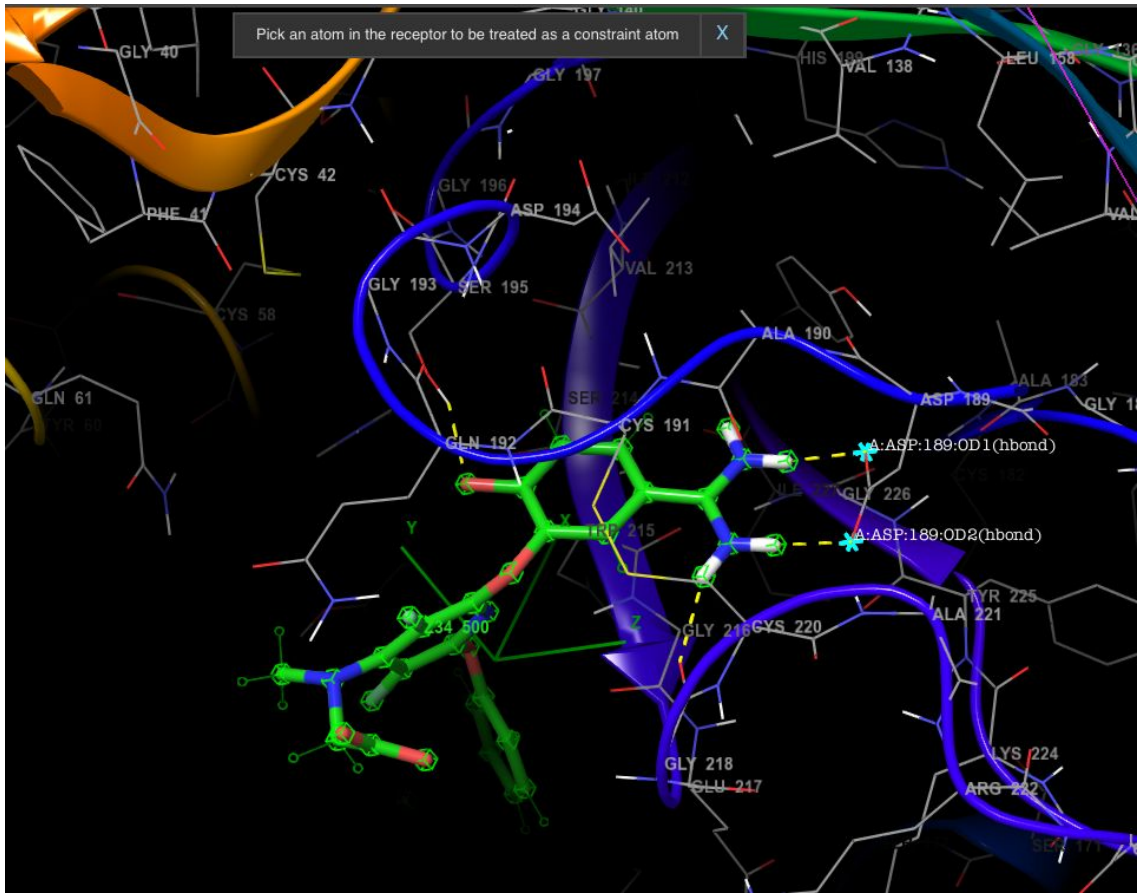
- Glide SP
  - Uses explicit spheres representing waters on grids
  - Measures exposure of various groups to explicit waters
  - No sampling of water molecules
  - Weak penalties
- Glide XP
  - Uses explicit spheres added to voids for docked poses
  - Measures exposure of various groups to explicit waters
  - Accounts for bridging water molecules
  - Stronger penalties
- WScore
  - Drastically improved treatment of desolvation through explicit waters from WaterMap and delocalized desolvation via MM-GBSA



Constraints can add information

- Evaluate evidence that a ligand should bind in a specific area or make a particular contact
  - crystal structure data
  - site-directed mutagenesis data
  - biological data

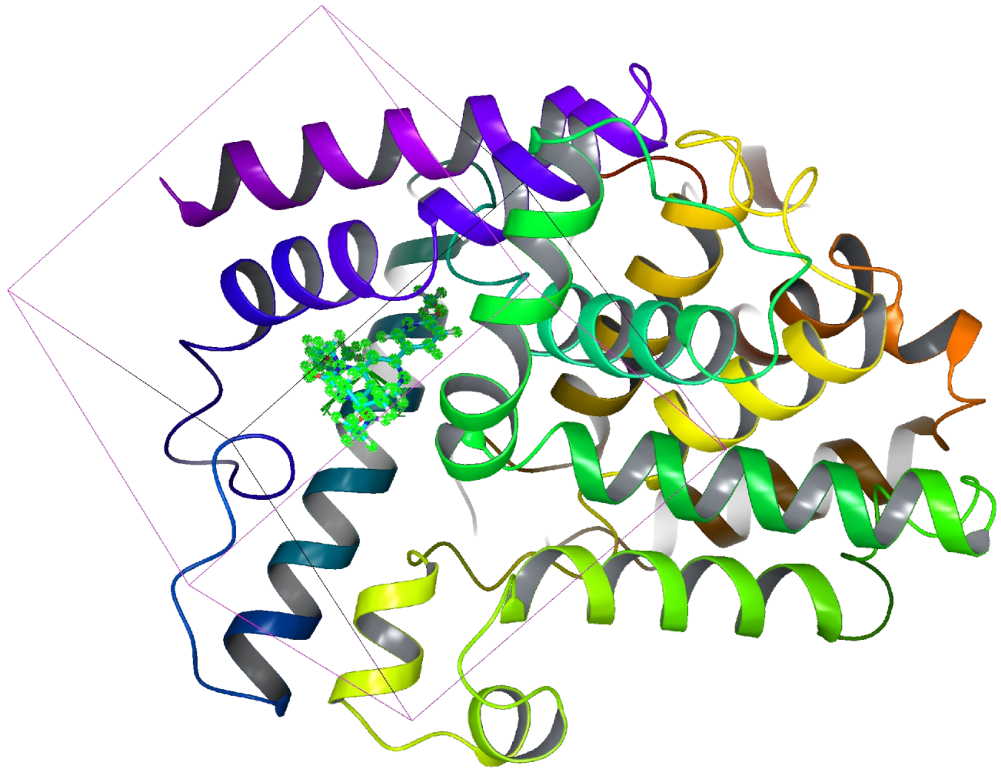
- Types of constraints:
  - Positional and NOE
  - Hydrogen-bonding and metal
  - Metal coordination
  - Core



Checklist and reminders

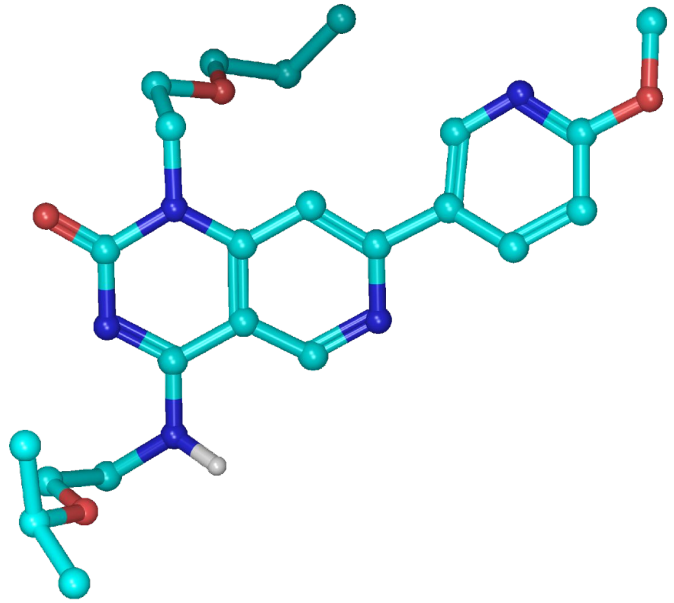
- **Druggability**
  - Alternative binding sites?
- **PDB**
  - Literature checks esp. reference publication (pH, resolution, missing things, chemistry states, motifs and key interactions)
  - Overlay many proteins as possible (water, flexibility, motifs that are consistent across all)
  - Prepare
- **Ligands**
  - Reference data, chirality (active), pka, known actives
- **Docking**
  - With constraints?
  - Dock native ligand check is RMSD high? If so, use constraints
  - Dock known actives/decoys can it distinguish? then screen new ligands for screening

Glide docking is requires two pieces



Receptor grid from a prepared protein

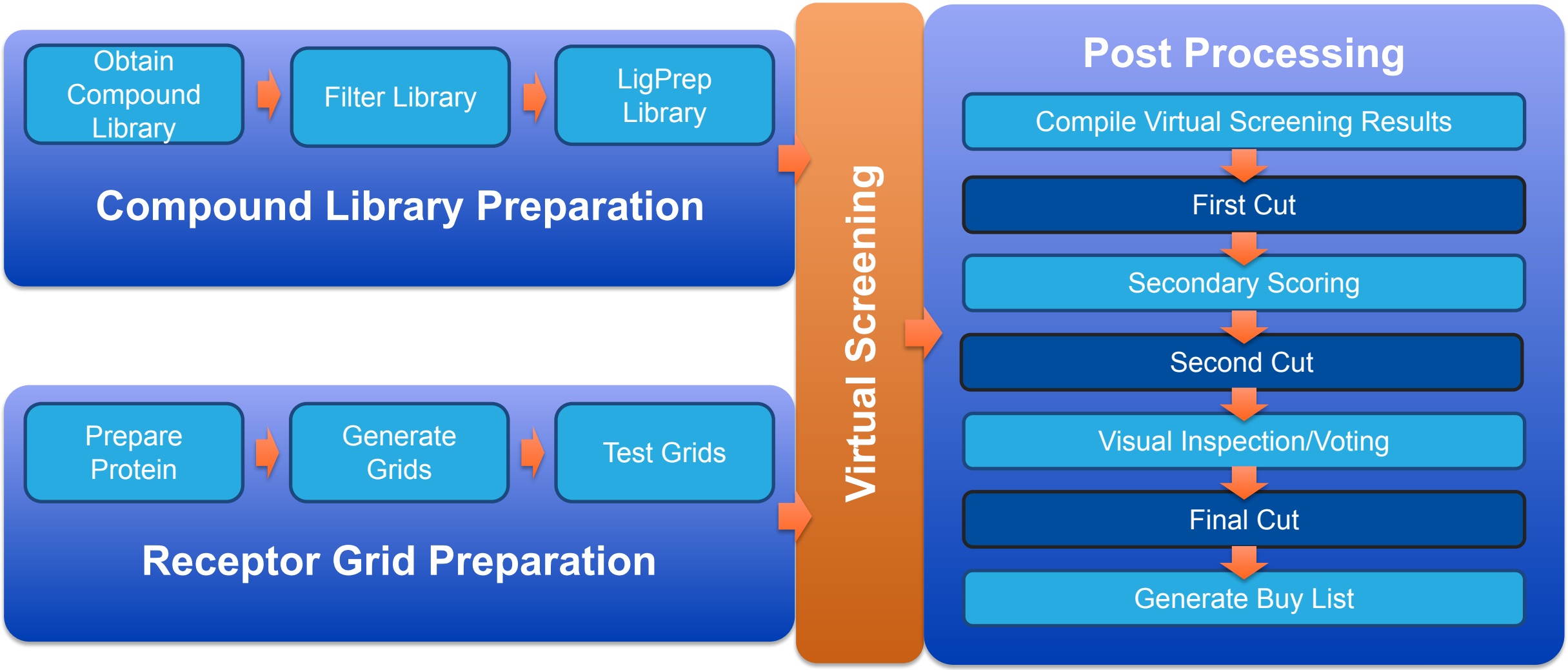
- Protein Preparation Wizard
- Receptor Grid Generation



Prepared ligand(s)

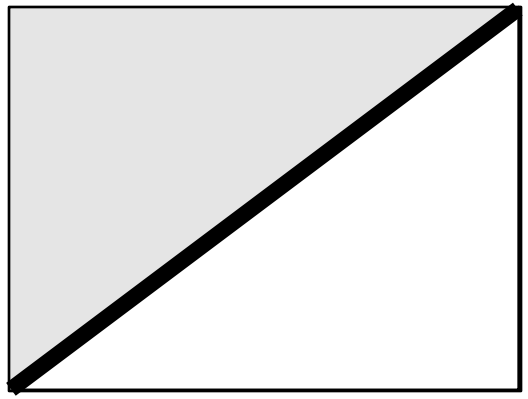
- LigPrep
- Suggested: Filter based on QikProp properties

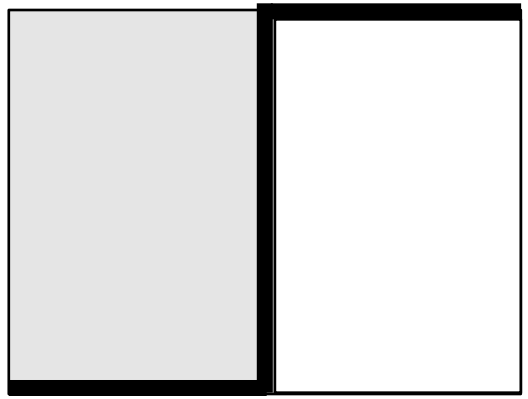
An idealized workflow for a virtual screen using Glide

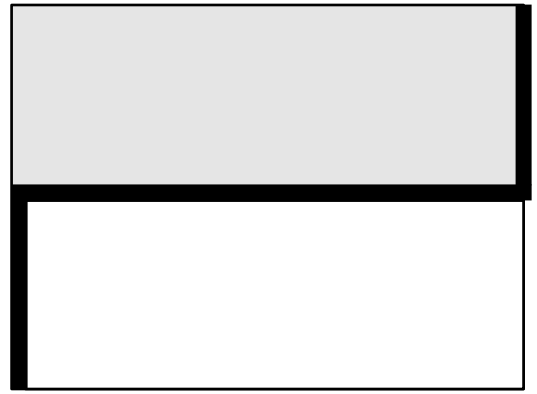


An Aside…Problems with AUC as a metric

- Which of the below is better?

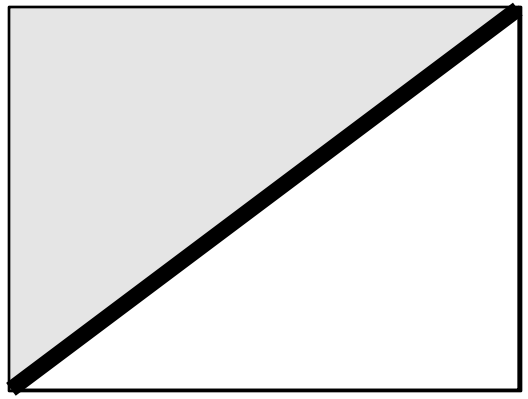




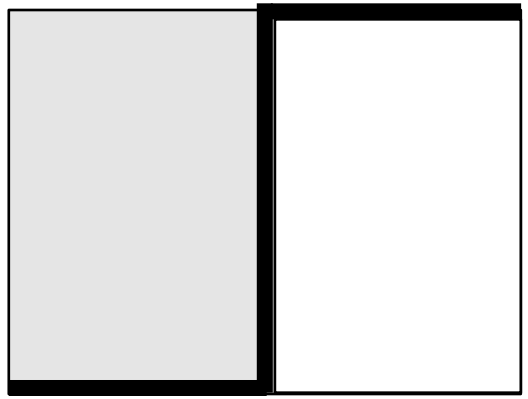


An Aside…Problems with AUC as a metric

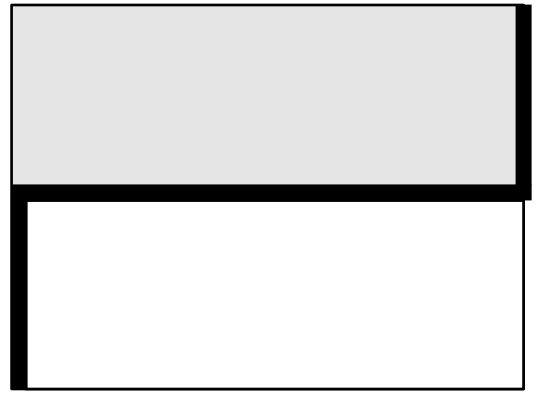
- Which of the below is better?



Random



Worthless



Excellent
(possibly 2 binding modes)