Computational Workshop #2:

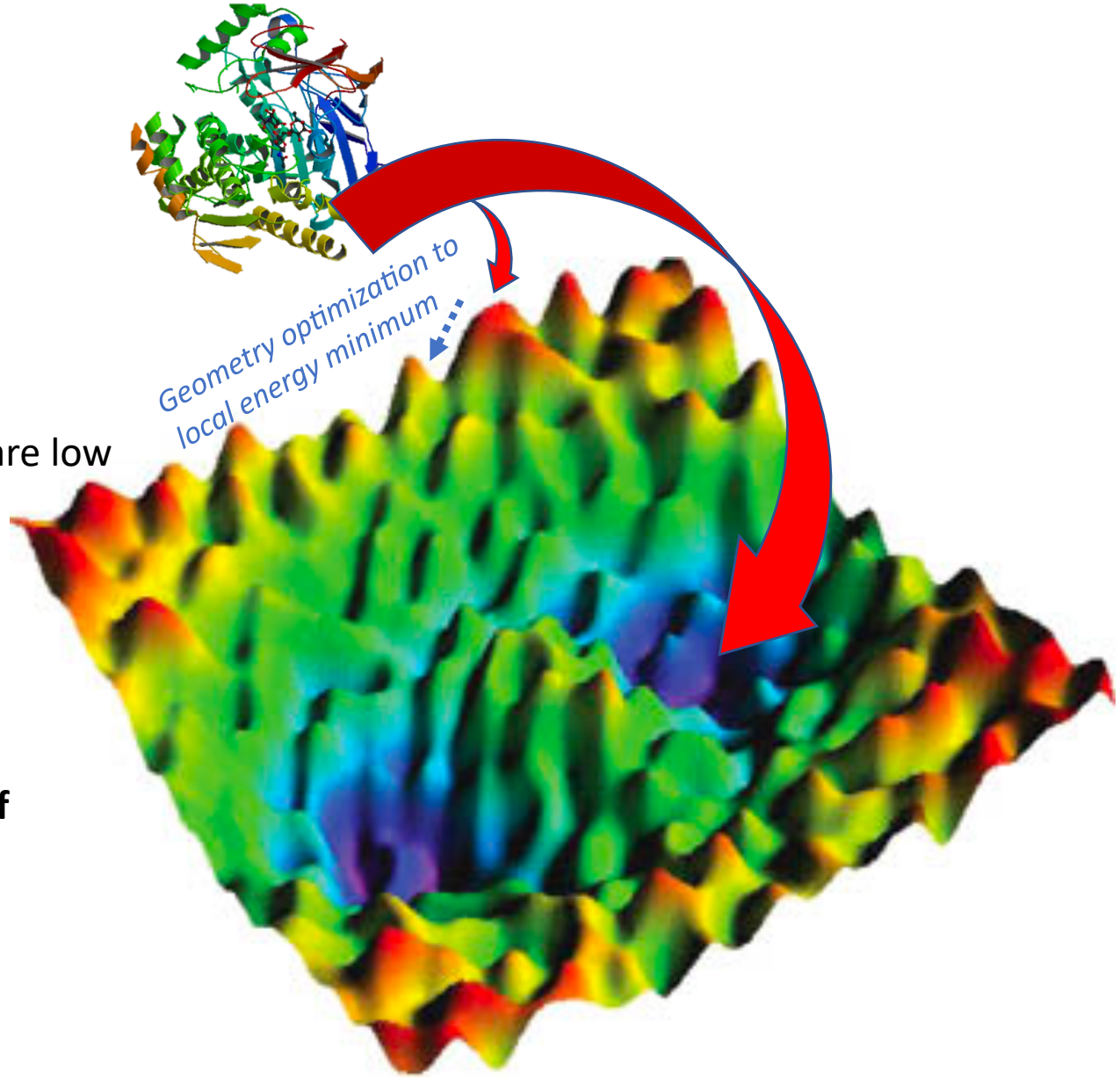
Predicting physicochemical properties, linear response calculations, Monte Carlo simulations and virtual screening

Sampling methods

Biomolecules:

- multidimensional potential energy surface (PES)
- Parts of PES can be relatively flat

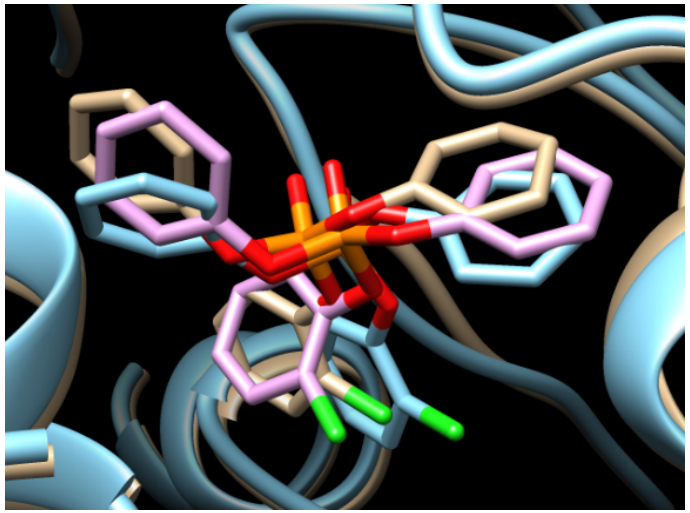
  → Weak noncovalent interactions → barriers are low

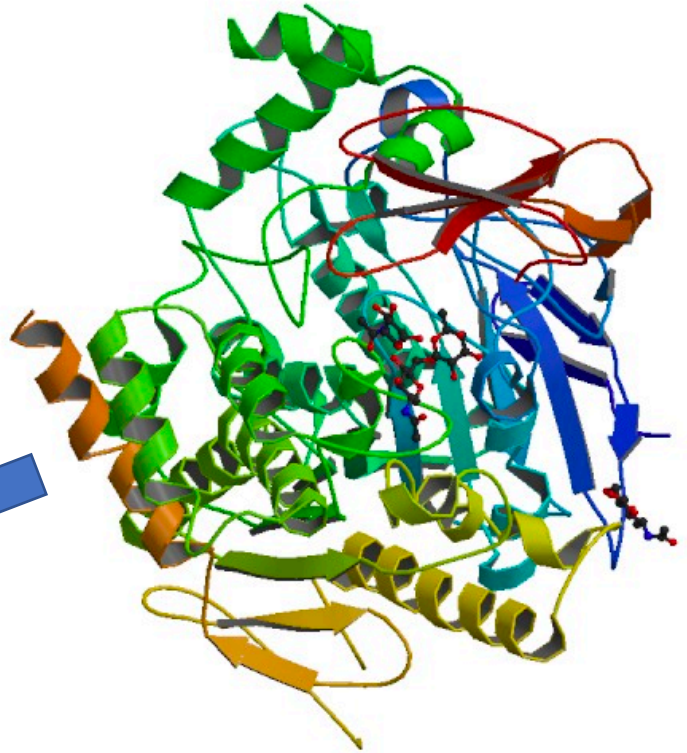


System's properties must be represented by **ensemble of states** and NOT a single state

Sampling methods

Different poses of TPP
bound in acetylcholinesterase





Binding affinities (kcal/mol):
Blue = -5.2 Brown = -8.5 Purple = -7.1

What is the relative population of brown and purple states?

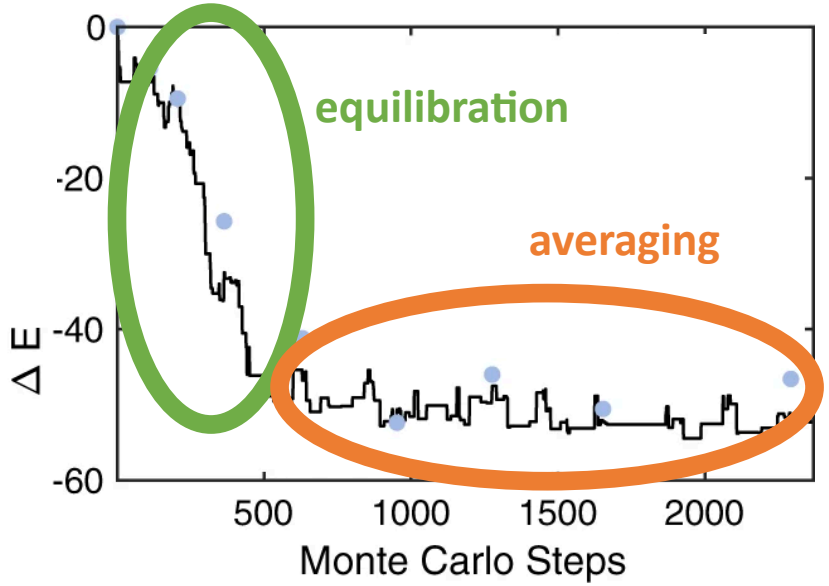
$$n_1/n_0 = e^{-\Delta G/kT}$$

→ If $[AChE]_{tissue}$ ~ 10^{-5} mol/kg then there are ~$1x10^{18}$ molecules in 1kg

→ at 310K, ca. 1/10 will be 'purple' not brown!!!

Monte Carlo simulations

- Generate an ensemble of states by producing random configurations of the system weighted by their Boltzmann probabilities
  - this process would yield an unrealistically large ensemble → **importance sampling (Metropolis method)** produces configurations with a probability proportional to their Boltzmann factors (biases sampling toward low energy structures)



$$\langle A \rangle = \sum_i A_i p_i$$

Linux commands review:

- Linux Cheat Sheets:
  - https://linoxide.com/guide/linux-cheat-sheet.png
  - https://upload.wikimedia.org/wikipedia/commons/7/79/Unix_command_cheatsheet.pdf

File Commands

ls – directory listing
ls -al – formatted listing with hidden files
cd *dir* - change directory to *dir*
cd – change to home
pwd – show current directory
mkdir *dir* – create a directory *dir*
rm *file* – delete *file*
rm -r *dir* – delete directory *dir*
rm -f *file* – force remove *file*
rm -rf *dir* – force remove directory *dir* *
cp *file1 file2* – copy *file1* to *file2*
cp -r *dir1 dir2* – copy *dir1* to *dir2*; create *dir2* if it doesn't exist
mv *file1 file2* – rename or move *file1* to *file2*
if *file2* is an existing directory, moves *file1* into directory *file2*
ln -s *file link* – create symbolic link *link* to *file*
touch *file* – create or update *file*
cat > *file* – places standard input into *file*
more *file* – output the contents of *file*
head *file* – output the first 10 lines of *file*
tail *file* – output the last 10 lines of *file*
tail -f *file* – output the contents of *file* as it grows, starting with the last 10 lines

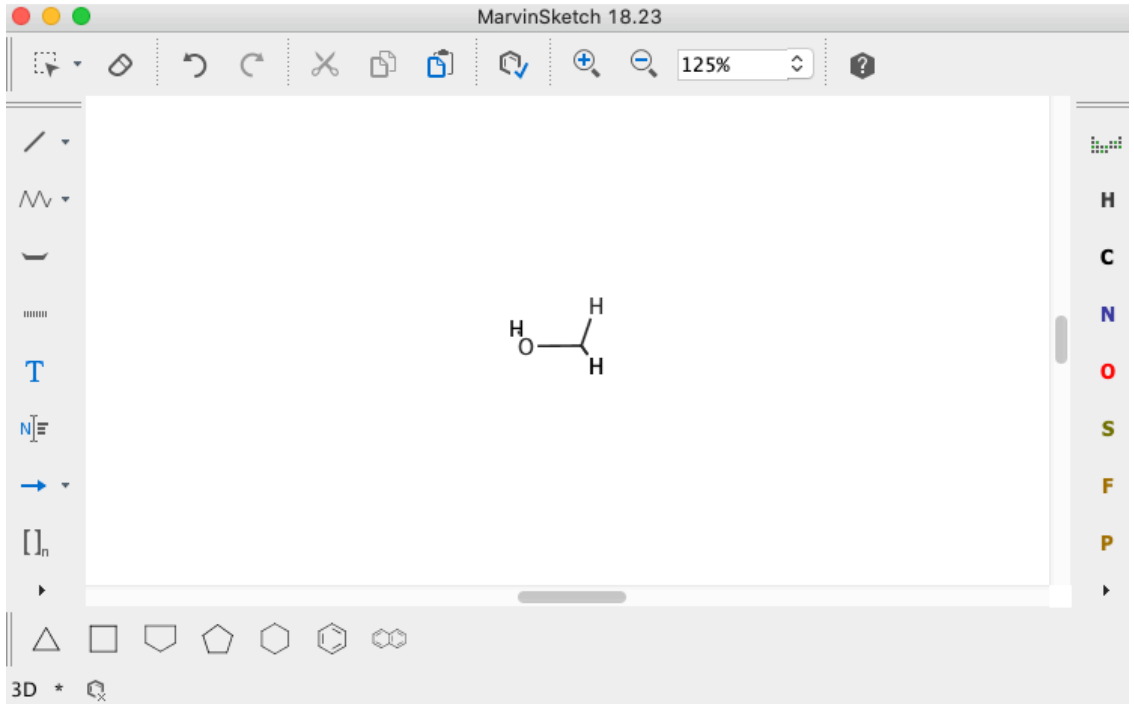
Outline of software for today's webinar

- QikProp – quick property prediction program for drug-like molecules and other commercial chemicals
  - http://gohom.win/ManualHom/Schrodinger/Schrodinger_2012_docs/qikprop/qikprop_user_manual.pdf
- BOSS – Computational Chemistry software
  - Monte Carlo simulations in condensed phases
  - Linear Response calculations
  - Manuals:
    - http://zarbi.chem.yale.edu/doc/boss49.pdf
    - http://zarbi.chem.yale.edu/doc/MCPROman230.pdf
- Virtual screening (Docking+Scoring)
  - Protein-Ligand Docking & Simulations using Vina and Mcule

Starting with the simplest: Qikprop

INPUT prep:

draw any structure in MarvinSketch, hydrogenize and optimize in 3D and save as .pdb file



QikProp

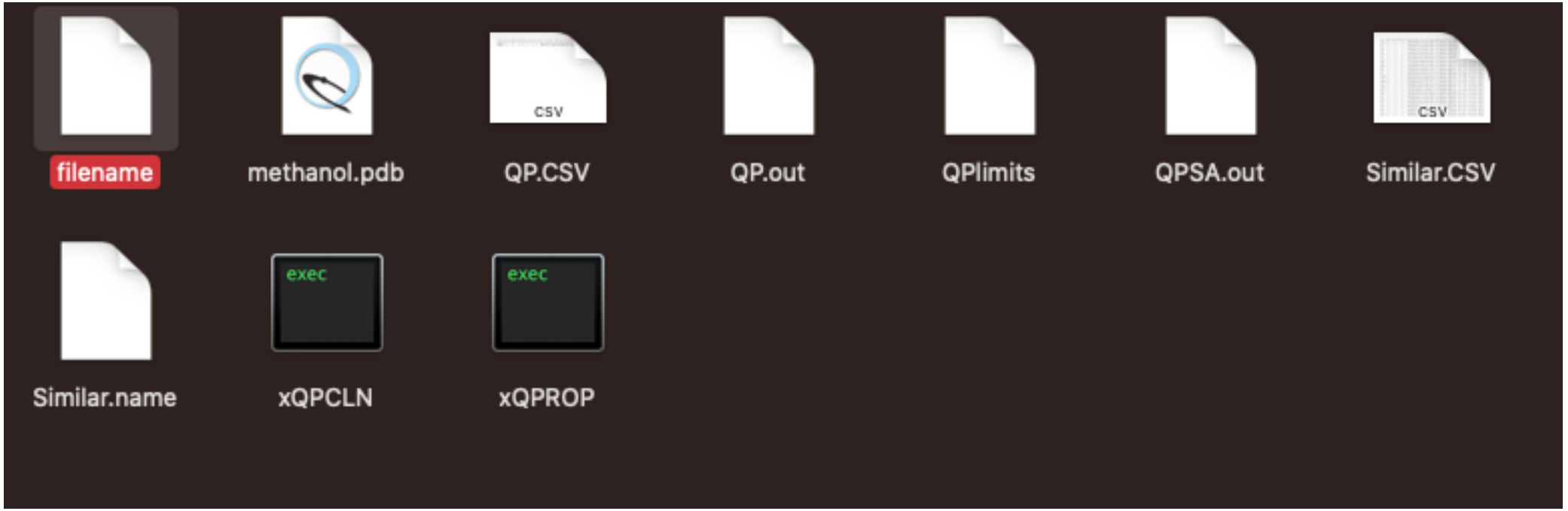
To run a Qikprop calculation execute the following:

./xQPROP filename or **qikprop [options] filename**

Example options are:

`-fast`	Turn on fast mode.
`-nofast`	Turn off fast mode.
`-neut`	Neutralize molecules in Maestro-formatted files before processing, by adding or removing protons (default).
`-noneut`	Do not neutralize molecules in Maestro files before processing.
`-sim`	Generate a list of known drugs most similar to each processed molecule.
`-nosim`	Do not generate of a list of known drugs most similar to each processed molecule (default).
`-simf`	Generate a file of the *n* most similar drug molecules to the last molecule processed. The number of drug molecules is specified by `-nsim`.
`-nsim` *int*	Number of the most similar drug molecules to report. Default: 5.
`-nosa`	Do not write the *jobname*`.qpsa` file.

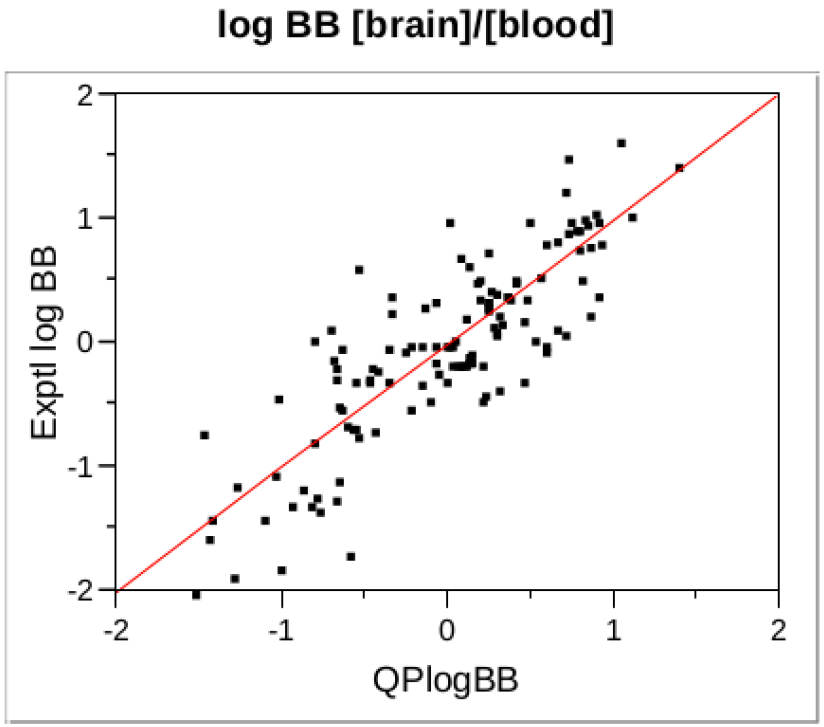
Qikprop Output Files



ANALYZING QikProp output

```
Predictions for Properties:
  QP Polarizability (Angstroms^3)       =    31.919M (  13.0 /  70.0)
  QP log P  for     hexadecane/gas      =    10.213M (   4.0 /  18.0)
  QP log P  for     octanol/gas         =    12.181M (   8.0 /  35.0)
  QP log P  for     water/gas           =     2.935M (   4.0 /  45.0)*
  QP log P  for     octanol/water       =     5.471  (  -2.0 /   6.5)
  QP log P  o/w     by  73% similarity  =     5.471  (  -2.0 /   6.5)
  QP log S  for   aqueous solubility    =    -5.461  (  -6.5 /   0.5)
  QP log S          by  75% similarity  =    -5.461  (  -6.5 /   0.5)
  QP log S - conformation independent   =    -4.028  (  -6.5 /   0.5)
  QP log K hsa Serum Protein Binding    =     0.651  (  -1.5 /   1.5)
  QP log BB for     brain/blood         =    -1.337  (  -3.0 /   1.2)
  No. of Primary Metabolites            =         5  (   1.0 /   8.0)
  Predicted CNS Activity (-- to ++)     =        --
  HERG K+ Channel Blockage: log IC50    =    -3.129  (concern below -5)
  Apparent Caco-2 Permeability (nm/sec) =       236  (<25 poor, >500 great)
  Apparent MDCK   Permeability (nm/sec) =       132  (<25 poor, >500 great)
  QP log Kp for skin permeability       =    -2.098  (Kp in cm/hr)
  Jm, max transdermal transport rate    =     0.008  (micrograms/cm^2-hr)
  Lipinski Rule of 5 Violations         =         1  (maximum is 4)
  % Human Oral Absorption in GI (+-20%) =        88  (<25% is poor)
  Qual. Model for Human Oral Absorption =      HIGH  (>80% is high)

     A * indicates a violation of the 95% range. # stars =  1
    An M indicates MW is outside training range.
```



Linear Fit

Linear Fit

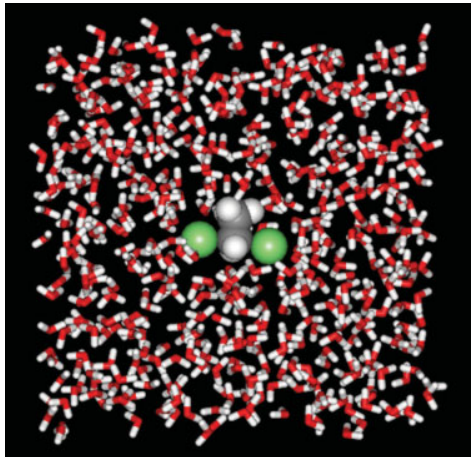
log BB = -0.009813 + 0.9967314 QPlogBB

Summary of Fit

RSquare	0.702763
RSquare Adj	0.700385
Root Mean Square Error	0.415251
Mean of Response	-0.03678
Observations (or Sum Wgts)	127

QP Manual contains linear model stats for all properties predicted

Molecular Modeling of Organic and Biomolecular Systems Using BOSS

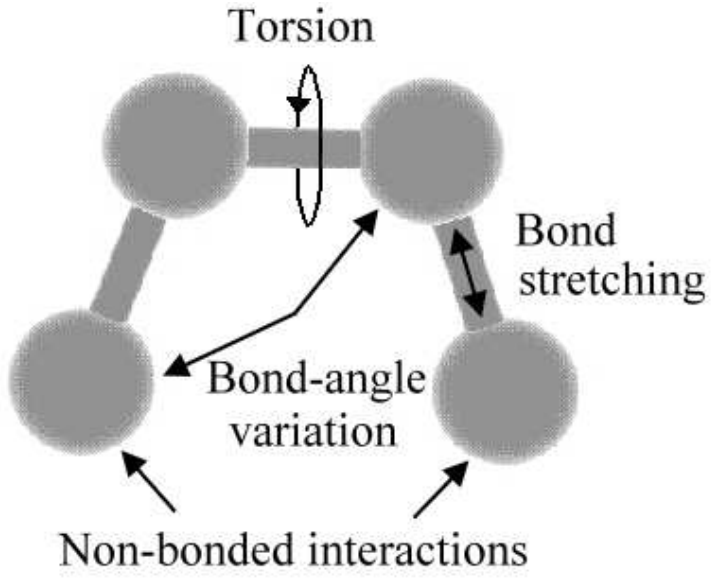


- A program that performs:
  - Monte Carlo (MC) statistical mechanic simulations on chemicals systems
  - Energy optimizations and conformational searching
  - Linear response calculations to predict properties from MC simulations
- QM, MM and QM/MM capability:
  - OPLS-AA and OPLS-UA force fields
  - QM: focus on semiempirical methods: AM1, PM3, PDDG/PM3, PDDG/MNDO
  - Quantum mechanical charges: CM1 and CM3

http://zarbi.chem.yale.edu/doc/boss49.pdf

Force field recap:

- Energies computed as a function of bond distances, angles, torsions and interatomic interactions:



$$E_{\text{bond}} = \sum_i k_{b,i}(r_i - r_{0,i})^2,$$

$$E_{\text{bend}} = \sum_i k_{\vartheta,i}(\vartheta_i - \vartheta_{0,i})^2,$$

$$E_{\text{torsion}} = \sum_i \{V_{1,i}(1 + \cos\varphi_i)/2 + V_{2,i}(1 - \cos 2\varphi_i)/2 + V_{3,i}(1 + \cos 3\varphi_i)/2 + V_{4,i}(1 - \cos 4\varphi_i)/2\},$$

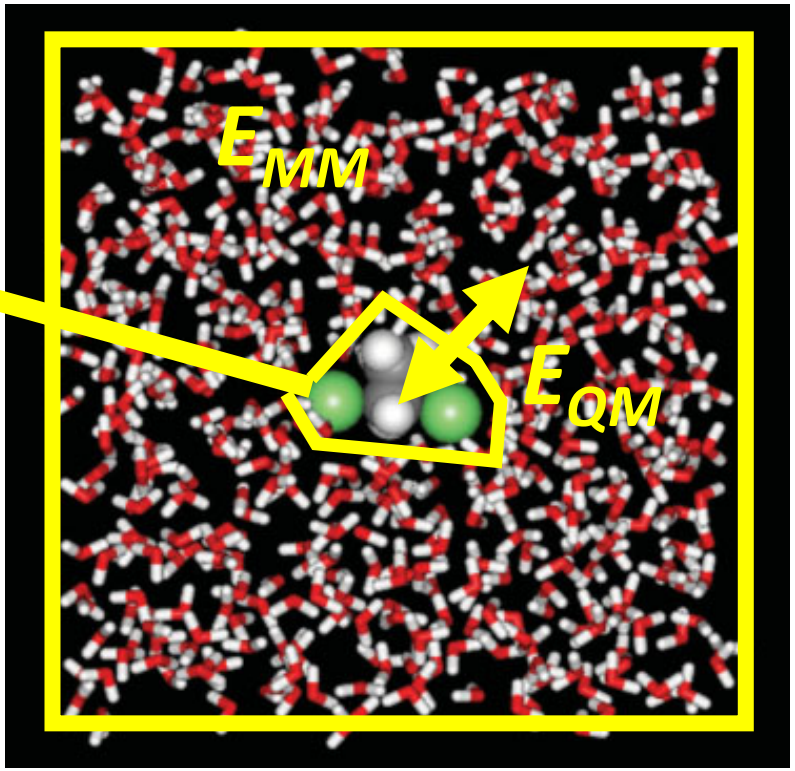
$$E_{nb} = \sum_{i<j} \{q_i q_j e^2/r_{ij} + 4\varepsilon_{ij}[(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^6]\},$$

$$E_{ab} = \sum_i^{\in a} \sum_i^{\in b} \{q_i q_j e^2/r_{ij} + 4\varepsilon_{ij}[(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^6]\},$$

How is system's energy calculated?

$$E_{\text{total}} = E_{\text{QM}} + E_{\text{QM/MM}} + E_{\text{MM}},$$

$$E_{\text{QM/MM}} = \sum_{i}^{\in \text{QM}} \sum_{j}^{\in \text{MM}} \{q_i q_j e^2 / r_{ij} + 4\varepsilon_{ij}[(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^{6}]\}$$



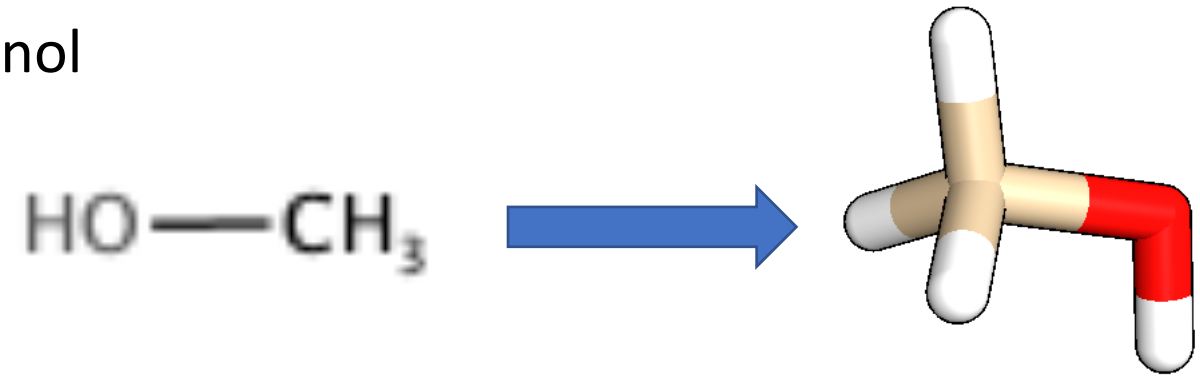
From BOSS output:
total E = ESS + ESX + EXX + EBND + EBC + EANG + EDIH + ENB + ECUT + ESINT

- ESS = the solvent-solvent energy,
- ESX = the solvent-solute energy,
- EXX = the solute-solute (intersolute) energy,
- EBND = the bond stretching energy for the solutes,
- EBC = the energy for the harmonic restraints,
- EANG = the angle bending energy for the solutes,
- EDIH = the torsional energy for the solutes,
- ENB = the >1,3 intramolecular non-bonded energy for the solutes, and
- ECUT = the cutoff correction for the Lennard-Jones interactions neglected beyond the cutoff for non-aqueous solvents,
- ESINT = the intramolecular energy for flexible solvent molecules, = ESBND + ESANG + ESDIH + ESNB

How to run a Monte Carlo simulation in *BOSS*

1. Generate a PDB of the desired solute molecule using MarvinSketch

Methanol



LigParGen

OPLS/CM1A Parameter Generator for Organic Ligands

2. Convert PDB to Z-matrix with ligpargen : http://zarbi.chem.yale.edu/ligpargen/

Downloads

OpenMM

XML | PDB

CHARMM/NAMD

PRM | RTF

GROMACS

GRO | TOP

PQR

PQR

BOSS/MCPRO

ZMAT

CNS/X-PLOR

TOP | PAR

ALL

ZIP

Q

QPARM | QLIB

LAMMPS

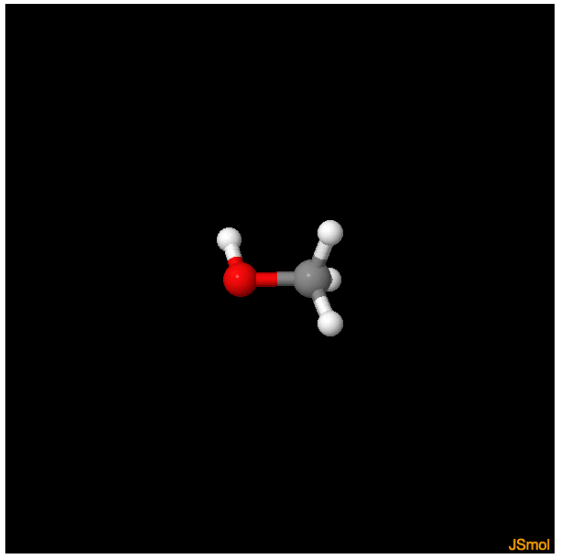
LAMMPS

DESMOND

DESMOND

TINKER

XYZ | KEY

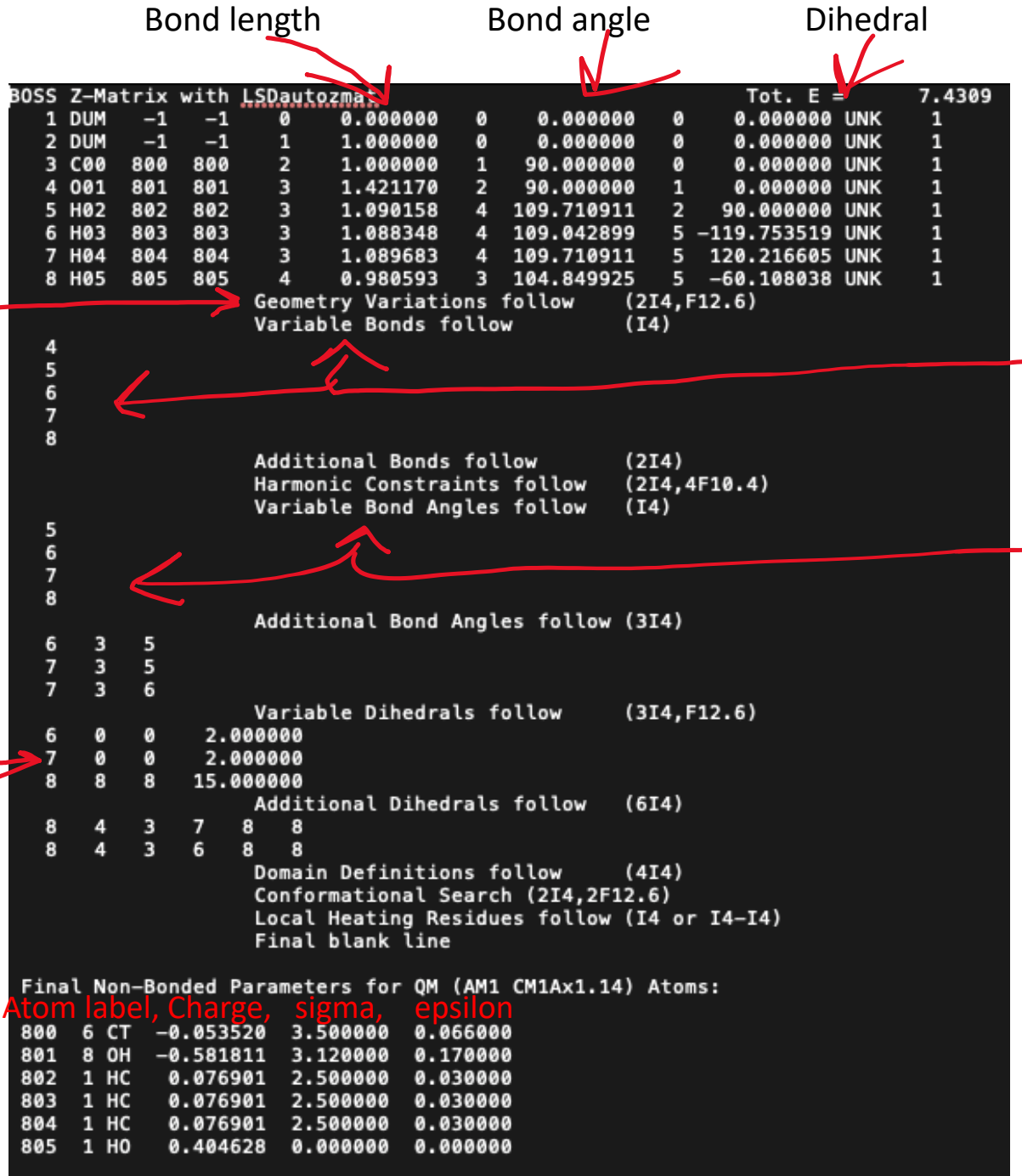



```
BOSS Z-Matrix with LSDautozmat                                  Tot. E =      7.4309
   1 DUM   -1   -1    0    0.000000   0    0.000000   0    0.000000 UNK    1
   2 DUM   -1   -1    1    1.000000   0    0.000000   0    0.000000 UNK    1
   3 C00  800  800    2    1.000000   1   90.000000   0    0.000000 UNK    1
   4 O01  801  801    3    1.421170   2   90.000000   1    0.000000 UNK    1
   5 H02  802  802    3    1.090158   4  109.710911   2   90.000000 UNK    1
   6 H03  803  803    3    1.088348   4  109.042899   5 -119.753519 UNK    1
   7 H04  804  804    3    1.089683   4  109.710911   5  120.216605 UNK    1
   8 H05  805  805    4    0.980593   3  104.849925   5  -60.108038 UNK    1
                  Geometry Variations follow    (2I4,F12.6)
                  Variable Bonds follow         (I4)
   4
   5
   6
   7
   8
                  Additional Bonds follow       (2I4)
                  Harmonic Constraints follow   (2I4,4F10.4)
                  Variable Bond Angles follow   (I4)
   5
   6
   7
   8
                  Additional Bond Angles follow (3I4)
   6   3   5
   7   3   5
   7   3   6
                  Variable Dihedrals follow     (3I4,F12.6)
   6   0   0    2.000000
   7   0   0    2.000000
   8   8   8   15.000000
                  Additional Dihedrals follow   (6I4)
   8   4   3   7   8   8
   8   4   3   6   8   8
                  Domain Definitions follow     (4I4)
                  Conformational Search (2I4,2F12.6)
                  Local Heating Residues follow (I4 or I4-I4)
                  Final blank line

Final Non-Bonded Parameters for QM (AM1 CM1Ax1.14) Atoms:

800  6 CT  -0.053520  3.500000  0.066000
801  8 OH  -0.581811  3.120000  0.170000
802  1 HC   0.076901  2.500000  0.030000
803  1 HC   0.076901  2.500000  0.030000
804  1 HC   0.076901  2.500000  0.030000
805  1 HO   0.404628  0.000000  0.000000
```

Zmat Format



Section to perturb internal degrees of freedom

Bond lengths that will be sampled during the simulation

Bond angles that will be sampled during the simulation

Dihedrals that will be sampled followed by the range of sampling

How to run an MC simulation

3. Edit your parameter (par) file:

Available solvents: Water, Methanol, Acetonictrile, Dimethyl ether, Propane, Chlorofrom, Methylene Chloride, Tetrahydrofuran, Argon, Carbon Tetrachloride, Dimethyl Sulfoxide

```
 parg46: Monte Carlo simulation for a solution
 The principal solvent is
DMSO
 QM Parameters
    CM1   1.20   0   0   0
 The solute format is
ZMAT
```

CM1 = charge model
1.20 = scaling factor (0 for charged systems)
0 0 0 = charge, e.g. for anion this would be -1 -1 -1

Box size (molecule count, NMOL – remove # of solvent molecules that corresponds to # of heavy atoms, i.e. non-Hs, in the solute):

```
 NMOL, IBOX and BOXCUT for principal solvent:
 379 400   0.00
 Second solvent and number of molecules:
NONE     0
```

Temperature (C) and pressure (atm) can be adjusted to mimic experimental conditions:

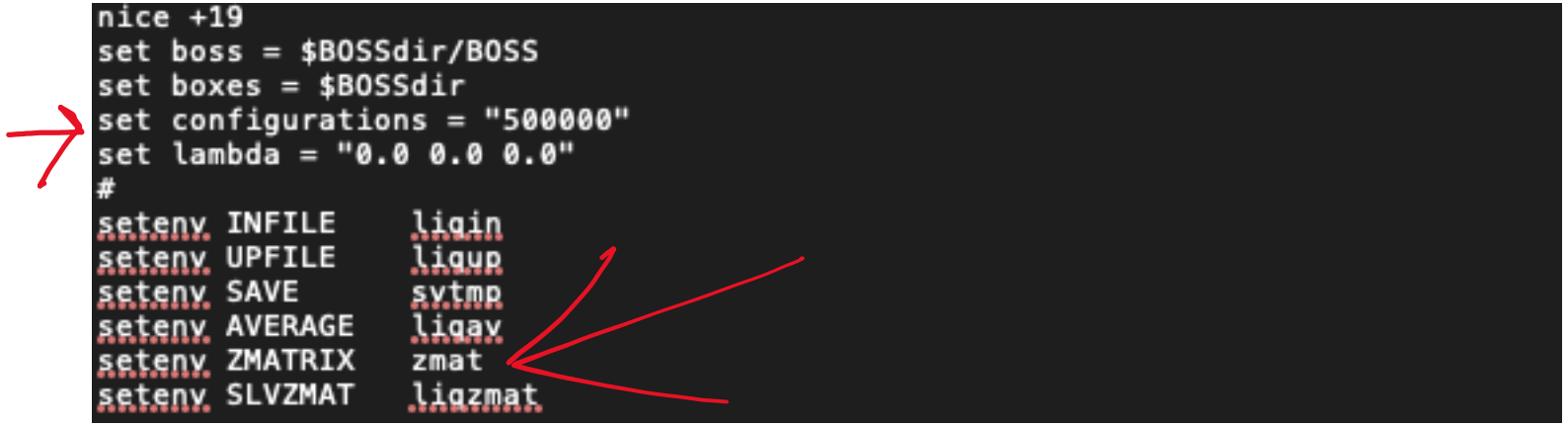
```
 Temperature, pressure, & torsion cutoff:
   25.0000     1.0000     0.0000
```

REFER TO THE MANUAL FOR APPROPRIATE KEYWORDS AND AVAILABLE OPTIONS!

How to run an MC simulation

4. Your command (cmd) file

    Only edit the configurations and ZMATRIX, (SLVZMAT in rare cases for custom solvents)



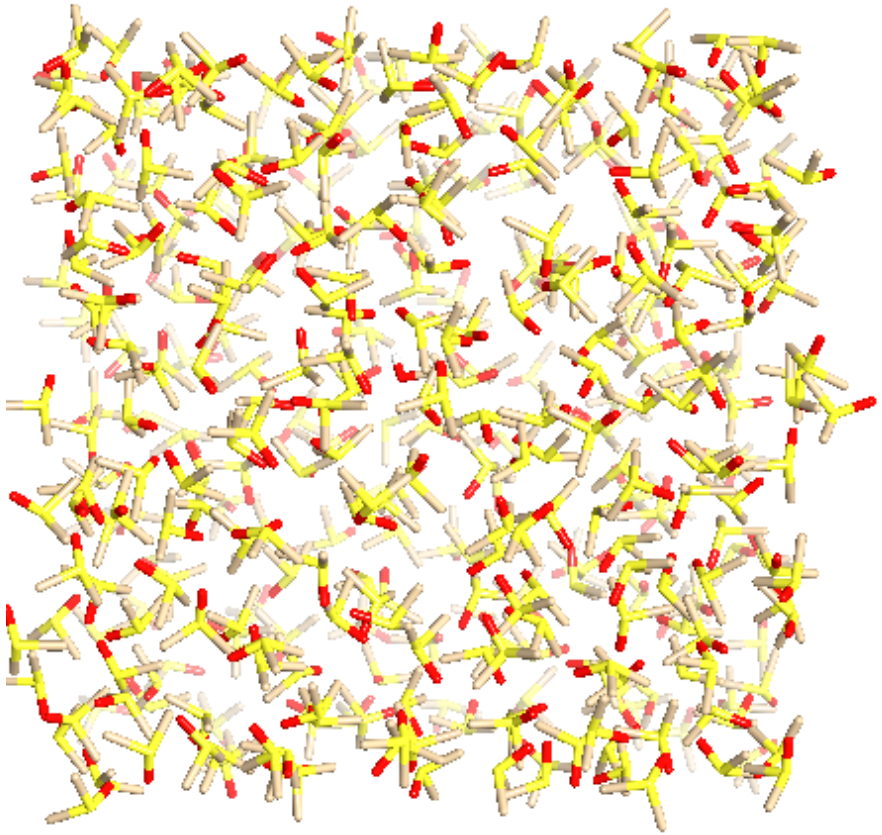
```
nice +19
set boss = $BOSSdir/BOSS
set boxes = $BOSSdir
set configurations = "500000"
set lambda = "0.0 0.0 0.0"
#
setenv INFILE    liqin
setenv UPFILE    liqup
setenv SAVE      svtmp
setenv AVERAGE   liqav
setenv ZMATRIX   zmat
setenv SLVZMAT   liqzmat
```

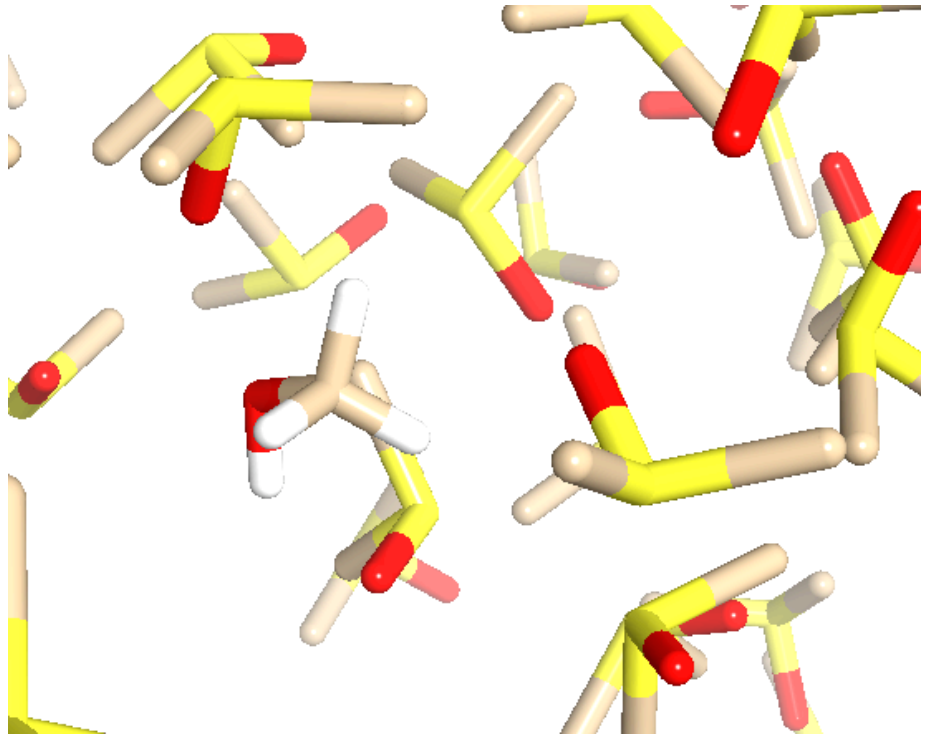
Submit the calculation with: csh liqcmd >& log &

Results

UCSF CHIMERA
an Extensible Molecular Modeling System

plt files can be viewed in CHIMERA (a viewing software from UCSF) https://www.cgl.ucsf.edu/chimera/





Results – ote file = full output

1. Plots showing the history for several variable dihedral angles in the solutes
2. The final total energy and its components along with the parameters for the simulation as above
3. The final coordinates, solvent accessible surface area and volume for the solutes
4. The averages for the thermodynamic properties including the two free energies which are repeated in fuller form below, average numbers of solute-solvent hydrogen bonds, The solute-solvent atom-atom radial distribution functions and their integrals (coordination numbers) that have been requested in the parameter file
5. The solvent-solvent and solute-solvent total energy and energy pair distribution functions
6. The distribution functions for the variable dihedral angles
7. The record of attempted and accepted moves for each solvent and solute molecule
8. The full report on the computed thermodynamic results including the averages for each run, the total averages and the standard deviations calculated from the fluctuations in the averages for each run.

Linear Response Calculations

- Using the same Zmatrix
- Similar to the previous calculations, linear responses can be done in a variety of solvents, but prebuilt linear models are trained on aqueous solutions (TIP4P water model in *BOSS*)
- Uses the same Zmatrix, par, and cmd files – refer to this directory
- To execute just execute

  ./xLMCPHERE ‘filename’

What Linear Response Calculations Do

1. Make a directory for the output files
2. Execute PM3 single-point calculation
3. Optimize geometry
4. Recompute charges with PM3 single-point
5. Run an MC job (what we just discussed)
   - Uses the command file in directory
6. Compute Properties

Linear Response Property Output

./xLMCPHERE ‘filename’

```
                                                    (Range for 95% of Drugs)
Descriptors from MC Simulation in Water:
    Solute-Water  Coulomb Energy          =  -111.050  (-137.5 / -17.3)
    Solute-Water  Lennard-Jones Energy    =   -23.433  ( -36.7 /  -4.2)
    Solute-Solute Coulomb Energy          =     0.000
    Solute-Solute Lennard-Jones Energy    =     0.000
    Solute        Molecular Weight        =   337.350  ( 130.0 / 525.0)
    Solute        Dipole Moment (D)       =     7.306  (   1.4 /  11.4)
    Solute        Total       SASA        =   595.046  ( 300.0 / 865.0)
    Solute        Hydrophobic  SASA       =   366.459  (   0.0 / 475.0)
    Solute        Hydrophilic SASA        =   109.104  (   6.0 / 275.0)
    Solute        Carbon Pi   SASA        =    90.161  (   0.0 / 370.0)
    Solute        Weakly Polar SASA       =    29.321  (   0.0 / 150.0)
    Solute        Molecular Volume (A^3)=  1038.140  ( 500.0 /1600.0)
    Solute-Water  Medium Interactions     =    11.423  (   2.2 /  15.8)
    Solute-Water  Strong Interactions     =     7.946  (   1.2 /  10.9)
    Solute as Donor -    Hydrogen Bonds =     1.004  (   0.0 /   4.2)
    Solute as Acceptor - Hydrogen Bonds =     6.993  (   0.8 /   8.4)
    Solute        No. of Rotatable Bonds=     2.000  (   0.0 /  13.0)
    Solute Globularity   (Sphere = 1)     =     0.833  (  0.75 /  0.95)

Predictions for Properties:
     Polarizability (Angstroms^3)         =    34.969  (  10.0 /  52.0)
     log P for     hexadecane/gas         =     8.666  (   4.0 /  14.4)
     log P for     octanol/gas            =    17.404  (   6.0 /  28.5)
     log P for     water/gas              =    18.147  (   4.0 /  30.0)
     log P for     octanol/water          =     1.767  (  -1.5 /   5.3)
     log S for   aqueous solubility       =    -4.041  (  -6.0 /   0.0)
     log BB for     brain/blood           =    -0.548  (  -3.0 /   1.0)
     log K hsa Serum Protein Binding      =    -0.292  (  -1.5 /   1.2)
     Papp  Caco-2 Permeability (nm/sec)   =   140.557  (< 5 low, >100 high) Boehringe
     Papp  Caco-2 Permeability (nm/sec)   =  1159.179  (<25 low, >500 high) Affymax
     Papp  MDCK   Permeability (nm/sec)   =   881.075  (<25 low, >500 high) Affymax
     log Kp for skin permeability         =    -3.021  (Kp in cm/hr)
     Jm, max transdermal transport rate   =     0.029  (micrograms/cm^2-hr)

     S in micrograms/ml                 =      30.685

     * is a 95% violation
```

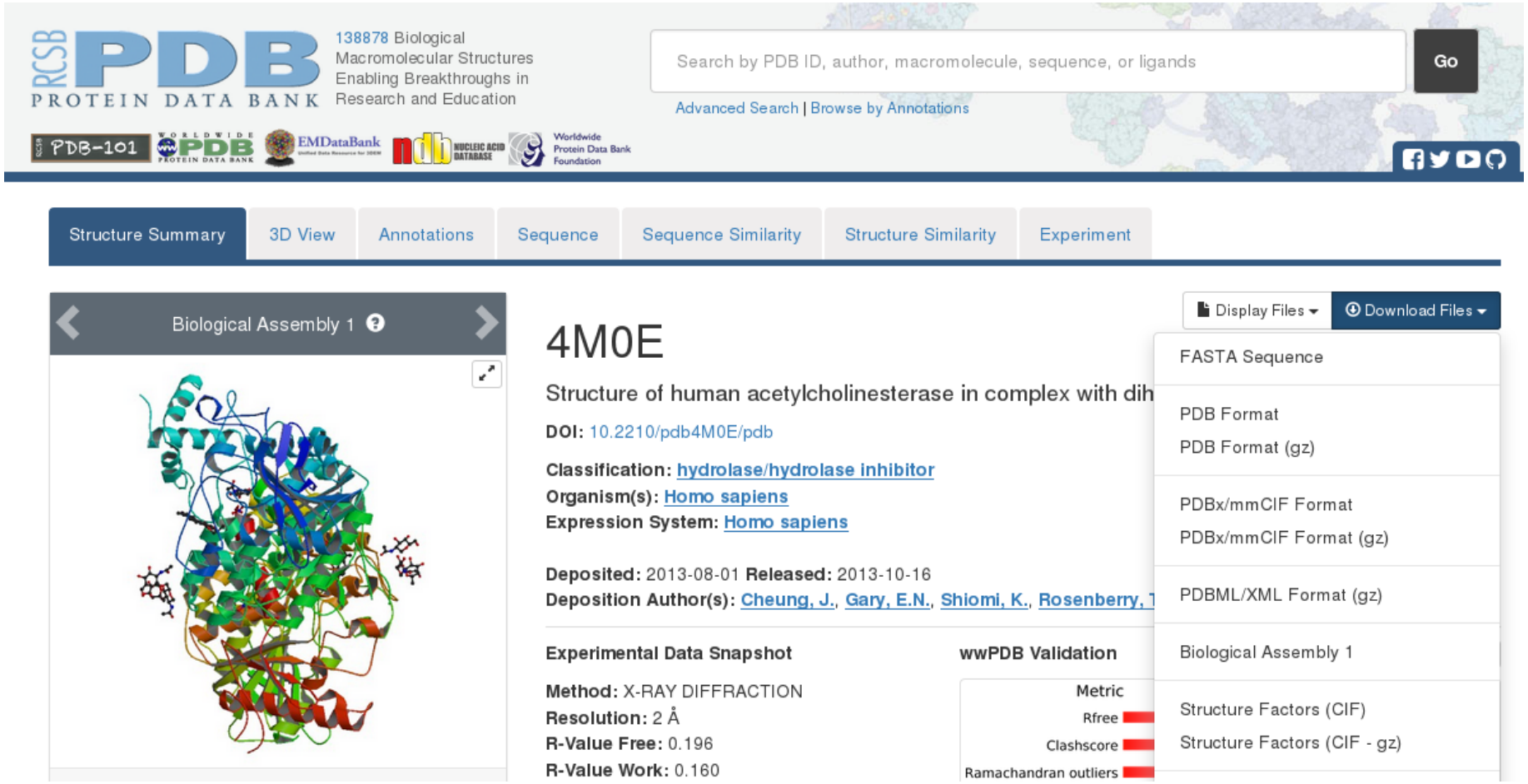
AutoDock

- AutoDock Tools(ADT) is needed for this section it can be downloaded here: http://autodock.scripps.edu/resources/adt

  https://ccsb.scripps.edu/mgltools/
- Works with Mac, Windows and Linux machines
  - Will NOT work with Catilina OS on macs
  - If adt is needed for your work and you only have access to a Mac with Catilina OS you can install a virtual machine on your personal computer: https://www.virtualbox.org/wiki/Downloads to use a different OS compatible with ADT

Molecular Docking with AutoDock Vina

Download 3D structure of Acetylcholinesterase (pdb id: 4m0e) from www.rcsb.org as pdb format (4m0e.pdb)



Prepare your protein for Docking using the Dockprep function of Chimera to add hydrogens, remove solvent ions, excess ligands, cofactors, or subunits, and repair incomplete side chains

Open 4m0e.pdb in chimera:

→Select → Chain → B

#The protein is a dimer you will only need one chain so you are deleting the other

→Select → Residue→all nonstandard

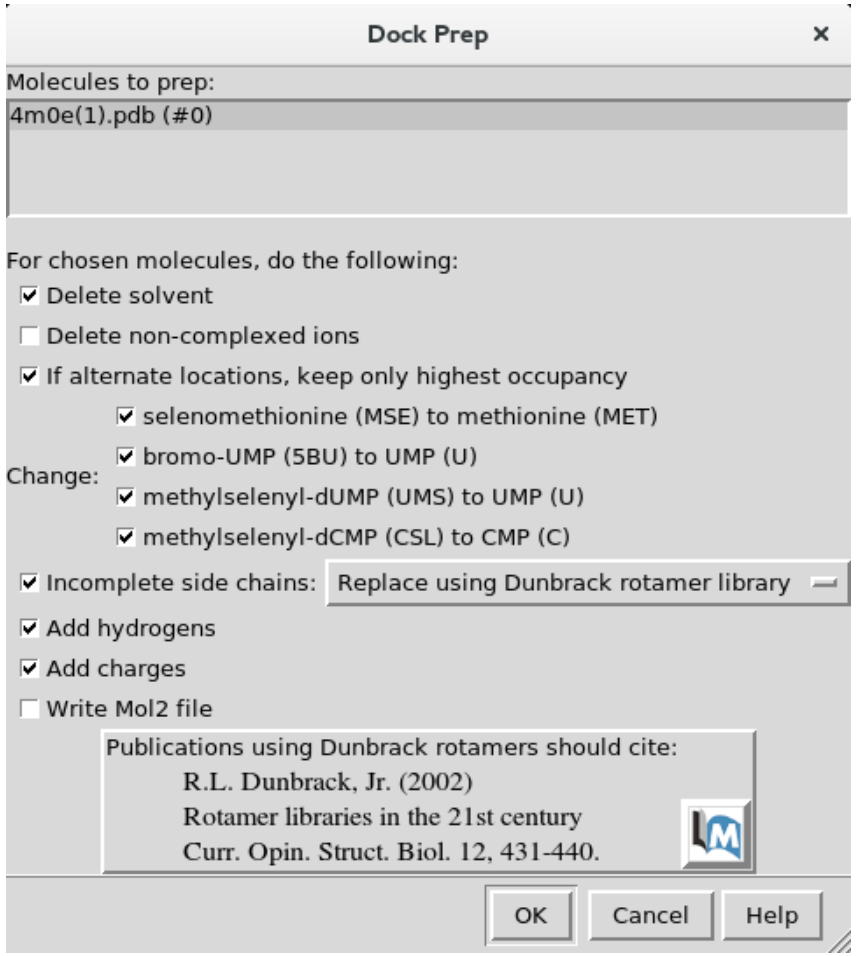
→Actions→ Atoms/Bond→Delete

#This will delete any ligands, ions, etc that are bound to the protein

→Tools→Structure Editing→DockPrep

#This brings up the menu to prepare the structure for docking: Deletes any solvent molecules, adds H's, charge, and fixes incomplete side chains

Uncheck Mol2 file, we will save the structure as a pdb to use in further prep for docking



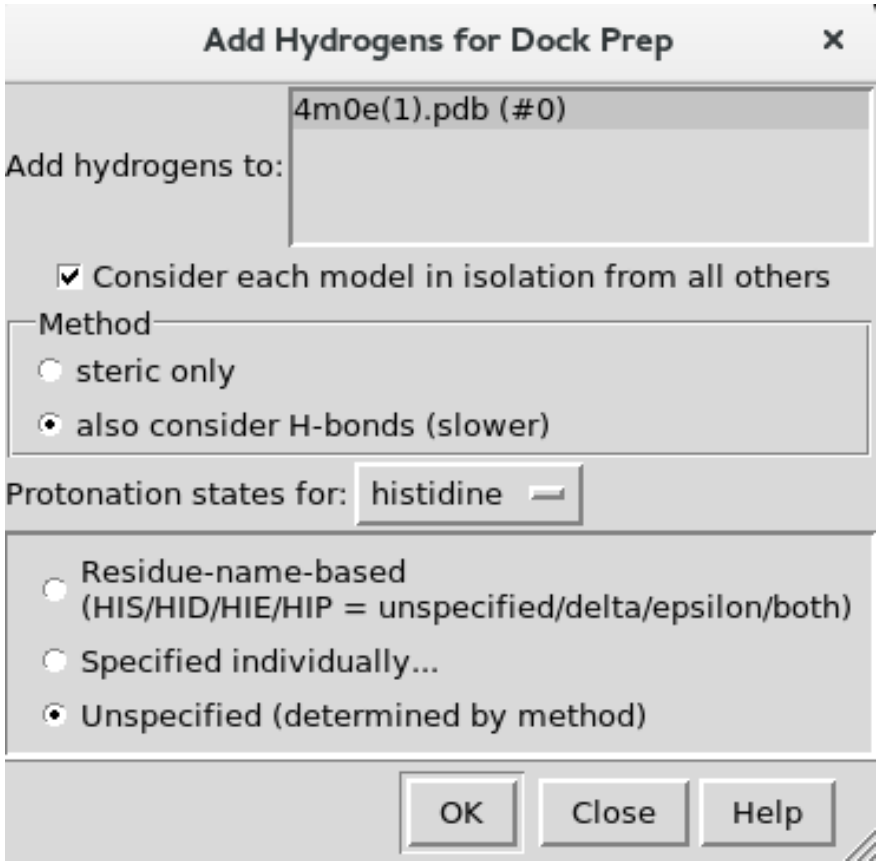
Prepare your protein for Docking using the Dockprep function of Chimera to add hydrogens, remove solvent ions, excess ligands, cofactors, or subunits, and repair incomplete side chains

Change the selection from “Residue-name-based” (default) to “Unspecified (determined by method)”

#Residue-name-based will simply assign a default protonation state based on the name of the residue

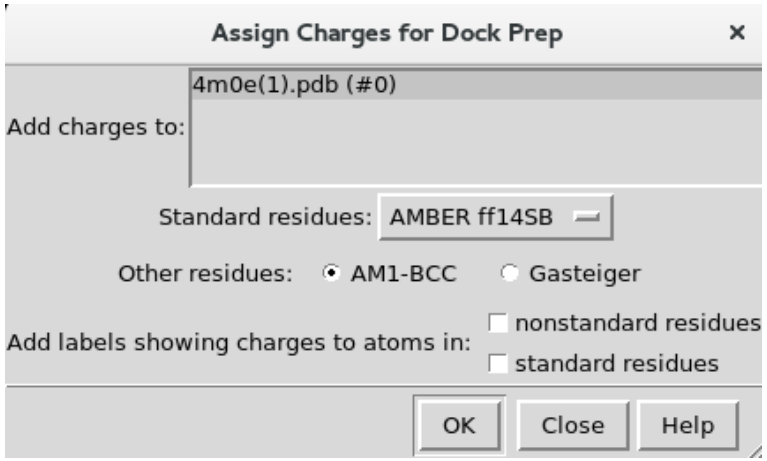
Ex) HIP = doubly protonated histidine

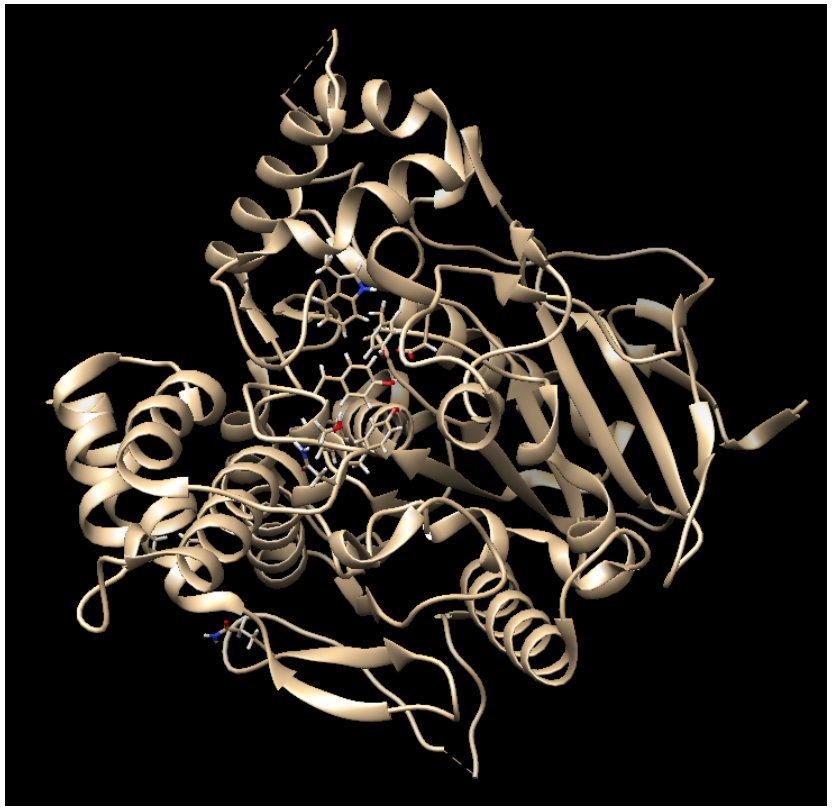
We want to instead calculate the protonation states



Prepare your protein for Docking using the Dockprep function of Chimera to add hydrogens, remove solvent ions, excess ligands, cofactors, or subunits, and repair incomplete side chains

Select the AM1-BCC charges

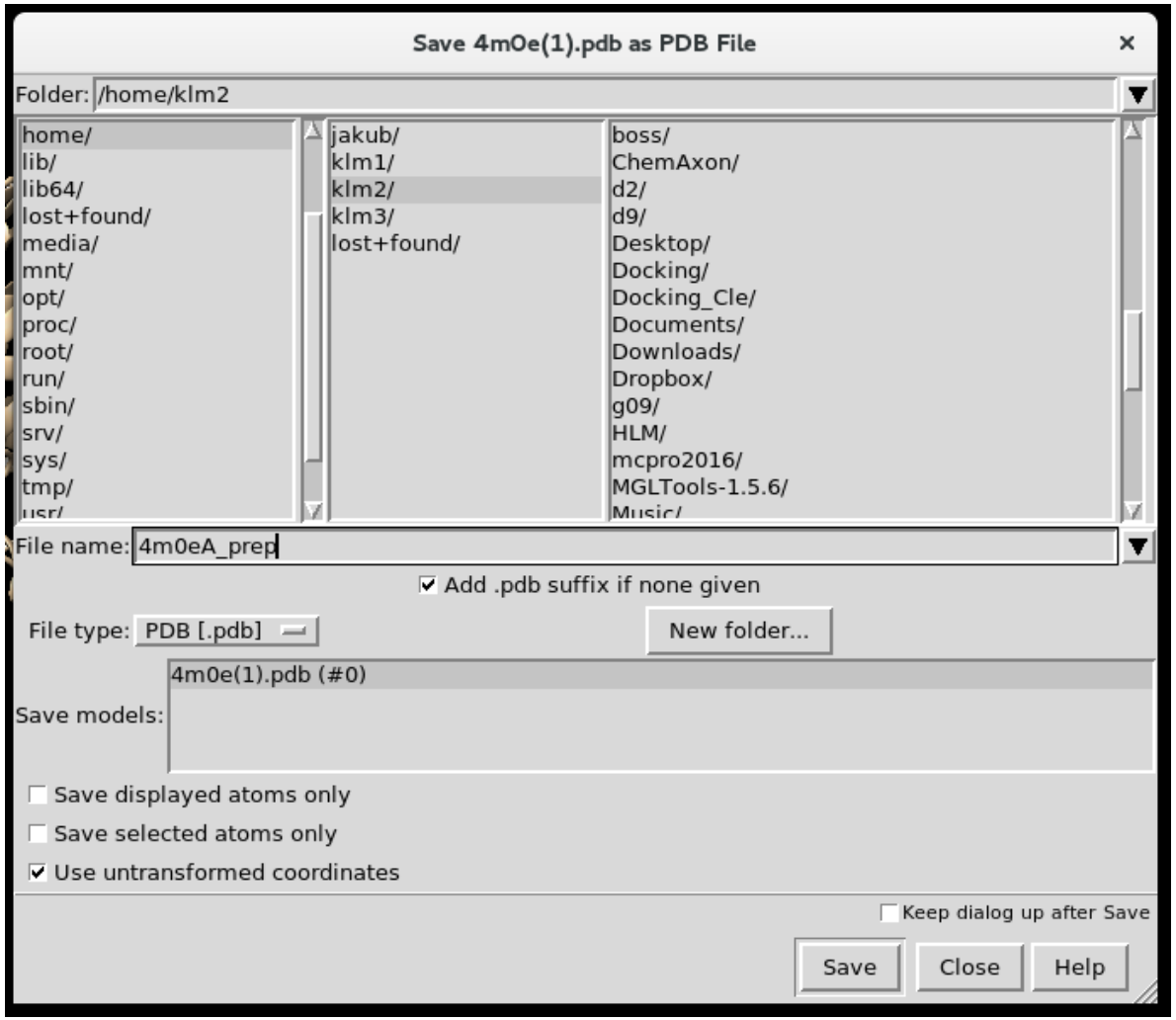




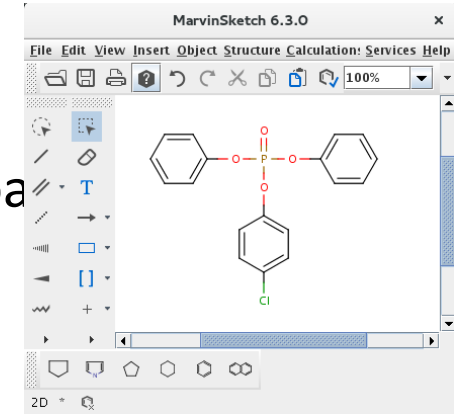
Save your prepped protein as a pdb

File -> Save PDB

Pick a name that indicates the changes/prep you've done

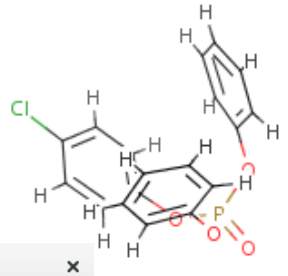


Preparing Ligand with Marvin

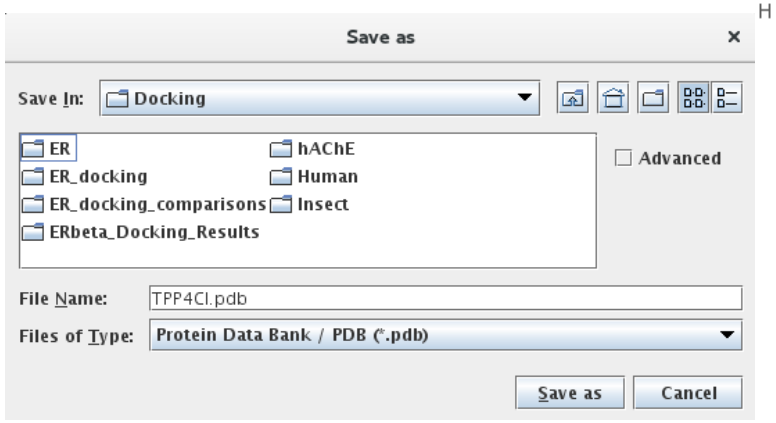


- **Open MarvinSketch and draw your structure:**
  - Triphenyl phosphate with a chlorine substituted in the pa position

- **Use clean in 3D to get a loosely optimized structure:**
  - Structure → Clean 3D → Clean in 3D
  - (it will look crazy because its 3 dimensions shown in 2, don't worry)

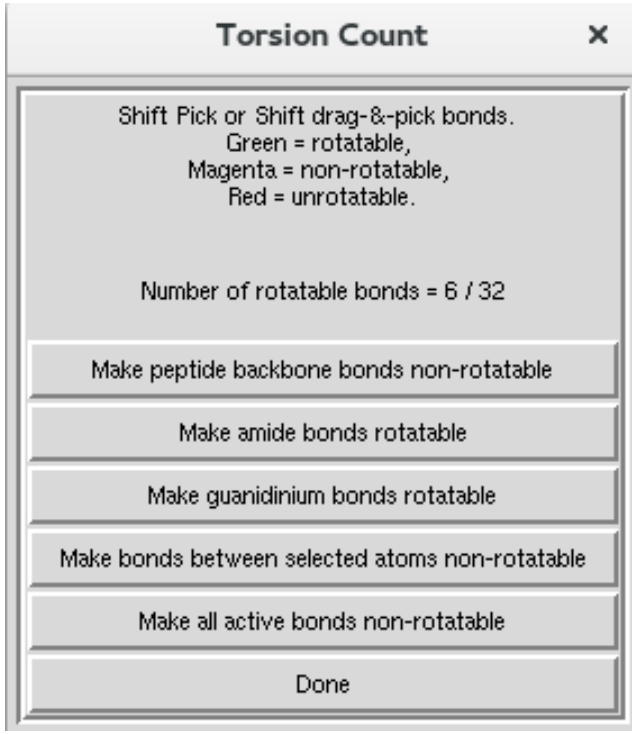


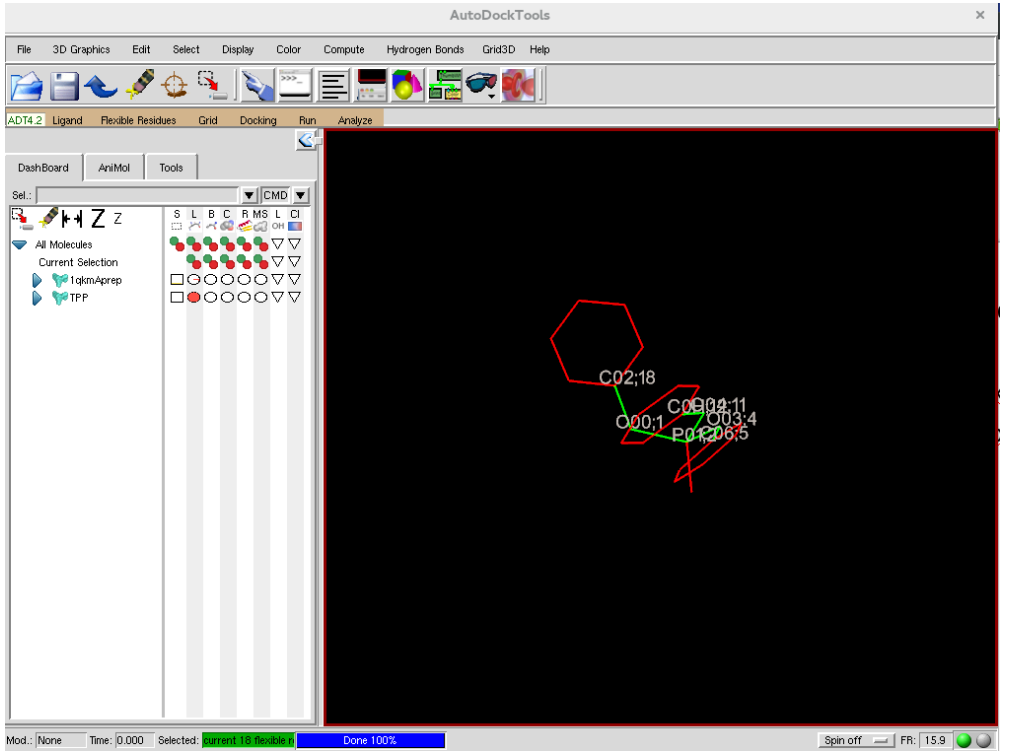
- **Save as a pdb:**
  - File → Save as → “Ligand.pdb”



Use Autodock tools to prepare input file of ligand for docking with Autodock Vina

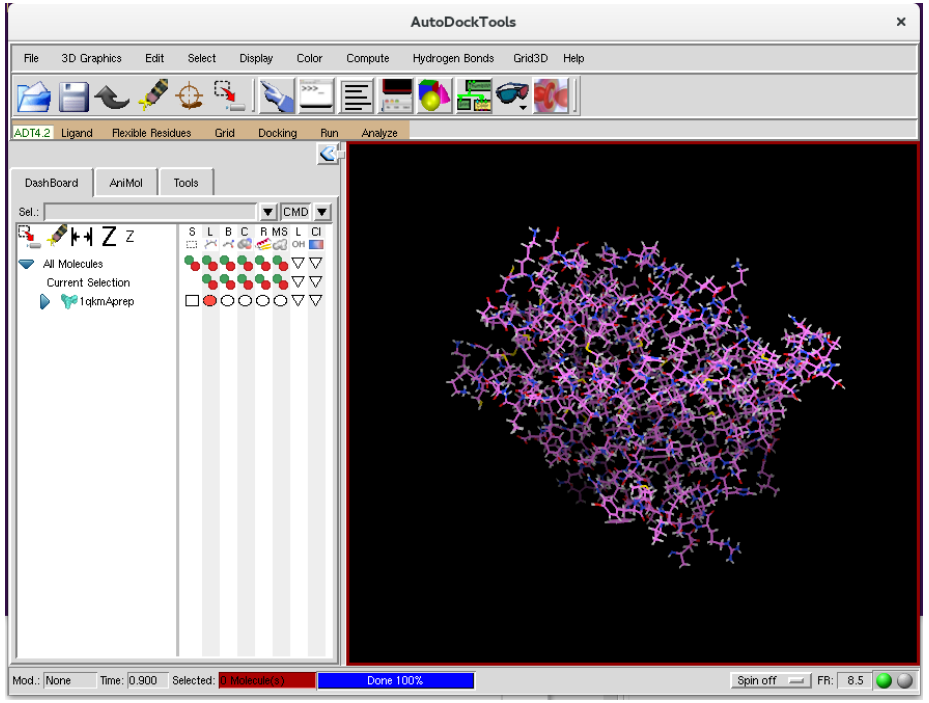
- Open Autodock tools
- Ligand → Input → Open → TPP.pdb
- Ligand → Choose torsions (are they correct?) → Done
- Ligand → Output → TPP.pdb → save TPP.pdbqt
- Close Autodock tools





Use Autodock tools to prepare input file of ligand for docking with Autodock Vina

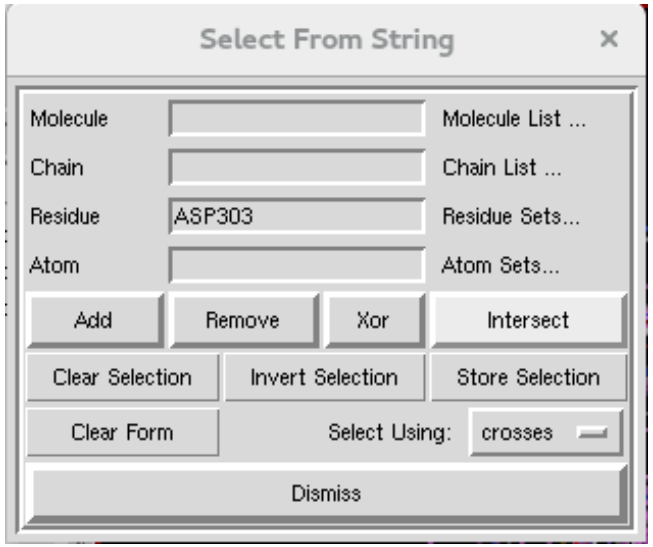
- Open Autodock tools
- Open the pdb of your protein that you prepped in Chimera:
  - File → Read molecule → 4m0eAprep.pdb
  - Edit → Delete water (should already be done)
  - Edit → Hydrogens → Merge non-polar
  - Grid → Macromolecule → choose → 4m0eAprep.pdb
  - (creates 4m0eAprep.pdbqt)

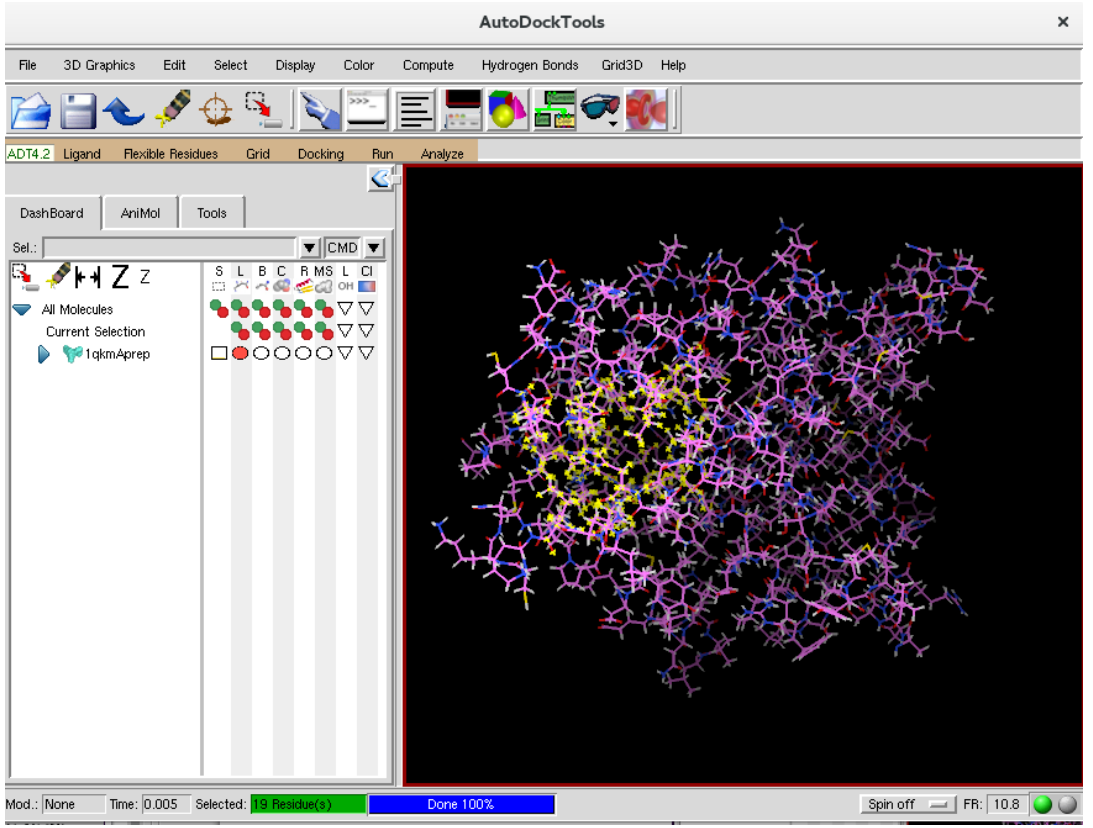


Use Autodock tools to prepare input file of ligand for docking with Autodock Vina

- Select Key Residues that will be allowed to rotate during docking:
  - Select → select from string (for 4m0e specific):

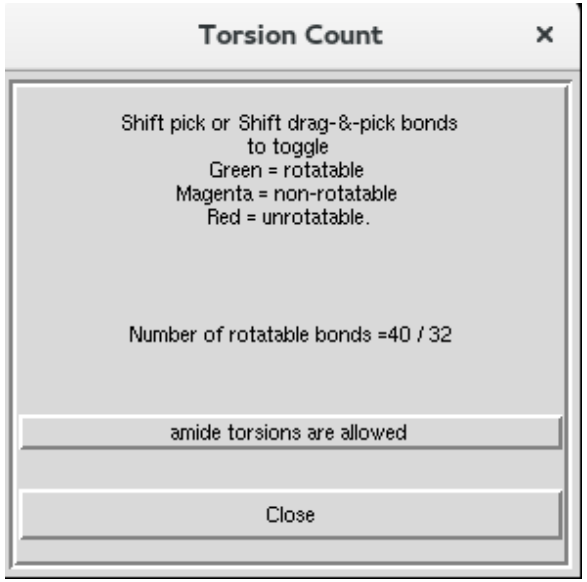
MET85 TRP86 TYR124 TYR133 SER203 GLU202 PHE297 TRP236
PHE295 TYR337 TRP286 HIS447 PHE338 GLU450 TYR449 ILE451

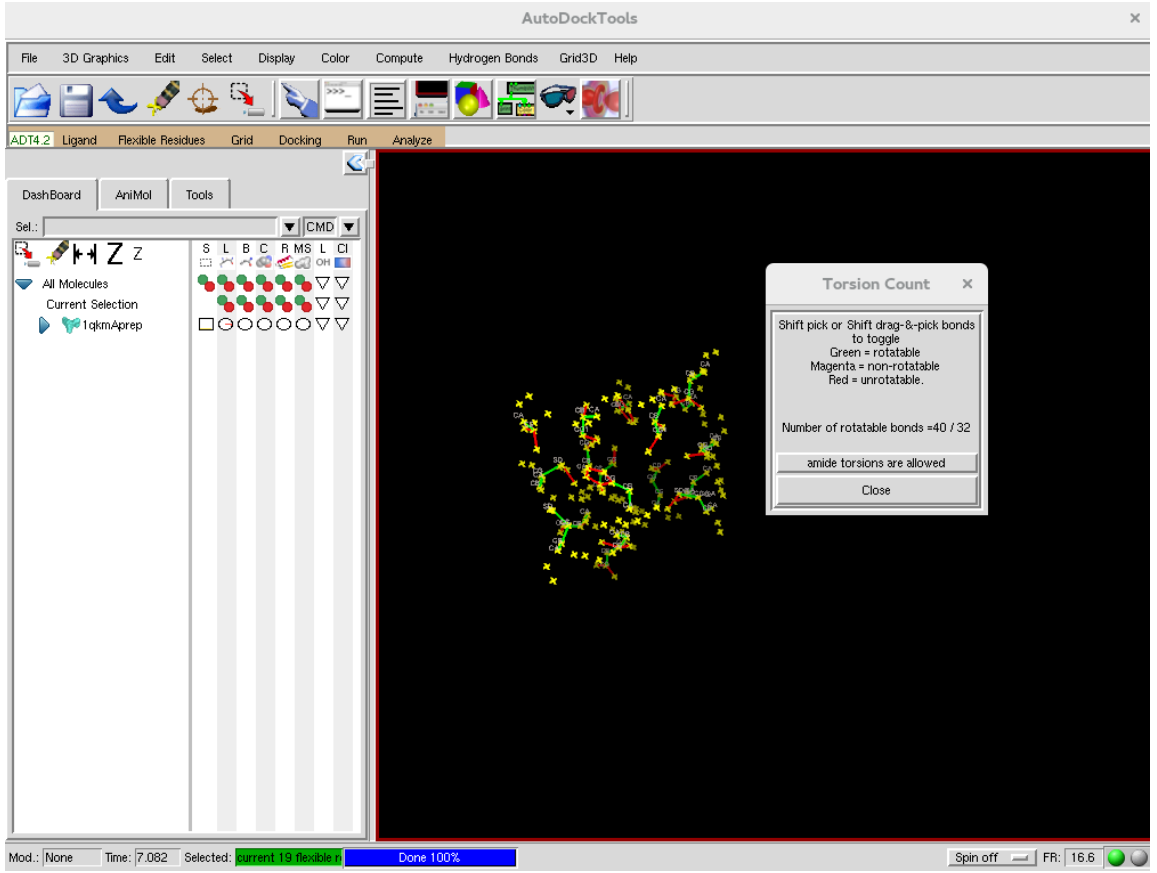




Use Autodock tools to prepare input files from your prepped protein for Autodock Vina

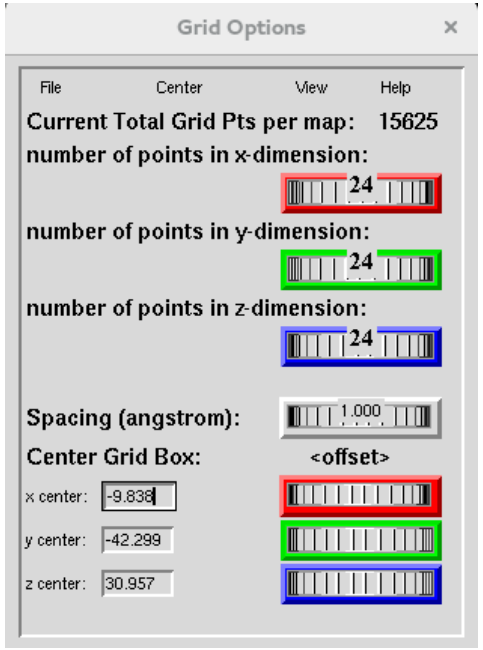
- Flexible residues → Input → Choose macromolecule → 4m0eAprep.pdbqt
- Flexible residues → Choose torsions
- Flexible residues → Output → Save Flexible PDBTQ (4m0eAprep_flex.pdbqt)
- Flexible residues → Output → Save Rigid PDBTQ (4m0eAprep_rigid.pdbqt)

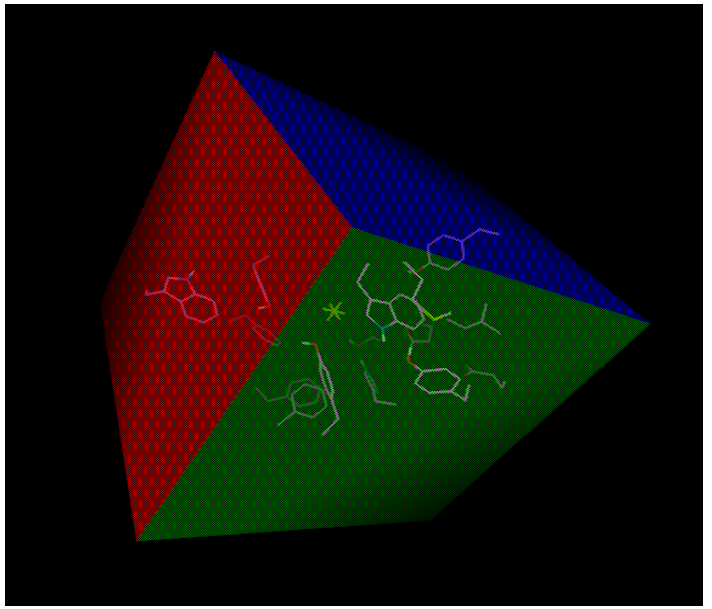




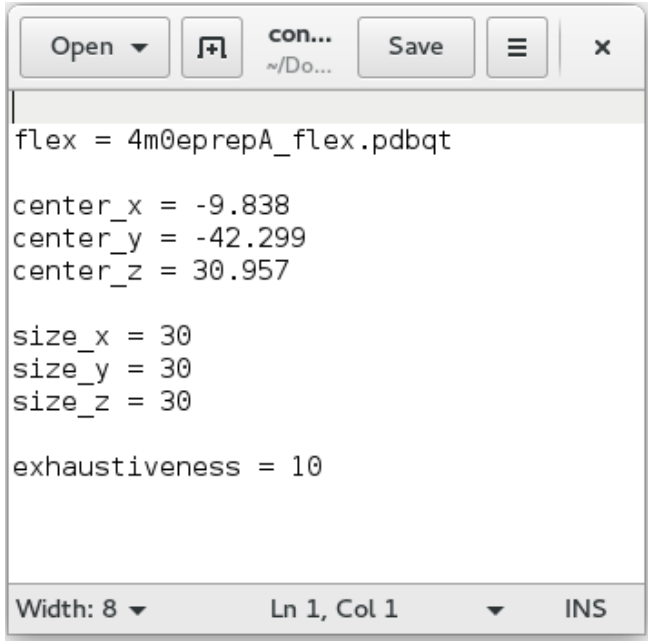
Assigning Dimensions for your Docking Search Space

- **Use the grid box feature visualize what dimensions will encompasses the flexible residues you have selected and be appropriate to search for potential binding poses**
- Grid → grid box
  - Change Spacing to 1.000 (for Å)
  - Adjust coordinates and size so that box encompasses flexible residues (aka binding pocket)
  - Record dimensions and coordinates!!
    - These will define where the docking algorithm should look for potential binding poses
    - You will need to put them into your configuration file
  - Close ADT





Setting up your config file and executing a docking simulation with Vina



config_4m0eA.txt

Create a configuration file in your favorite text editor (as shown on the left)

Assign: -flex file

-grid box coordinates

-grid box size

-exhaustiveness

Run your Docking Simulation in Vina on computing node!!

vina --receptor *4m0eAprep_rigid.pdbqt* --ligand *TPP.pdbqt* --config *config_4m0eA.txt* --log *TPP.log*

Coordinates for each pose and flexible residues will be in TPP_out.pdbqt

Summary tables of the results are found in TPP.log

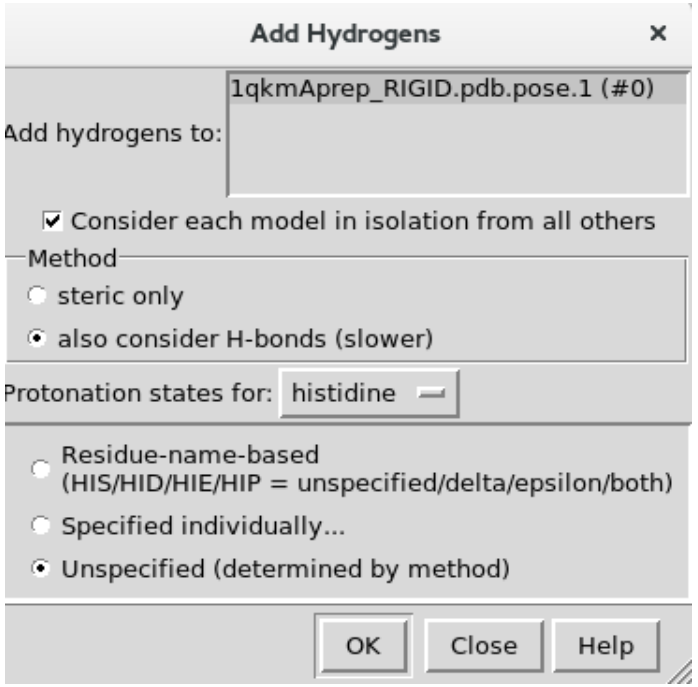
For a summary of all the flags in vina type “vina --help”

Combine the docking poses obtained from TPP4Cl_out.pdbqt with the Rigid receptor (4m0eAprep_rigid.pdbqt) to obtain a structure file for each pose bound to the receptor

- Use the xFLEXRESPREP script to add the coordinates of the ligand and flexible residues to the Rigid pdb for each pose:

./xFLEXRESPREP_v2 TPP_out.pdbqt 4m0eAprep_rigid.pdbqt

Use Chimera to protonate the structure (for docking we merged all non-polar hydrogens)
Open resulting pdb's in Chimera:
-Tools → Structure Editing → AddH
Check unspecified (determined by method)

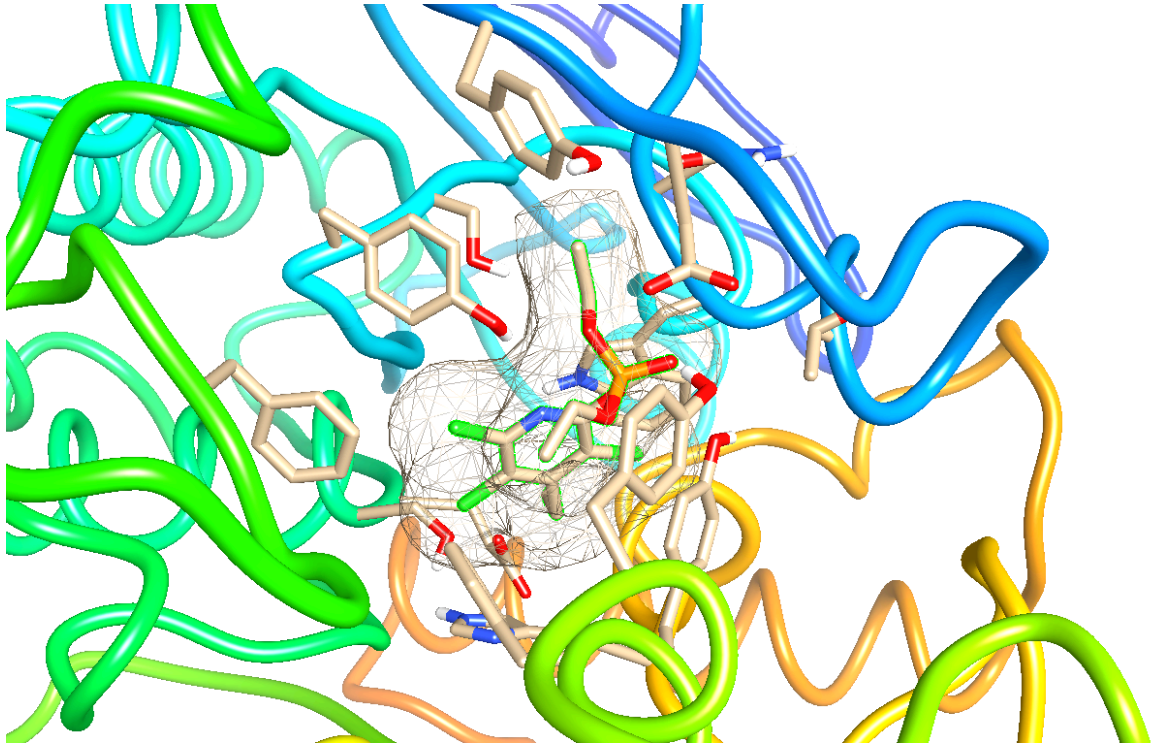


Save structures as .pdb files (you will only need the top pose for this tutorial)

Docking Results

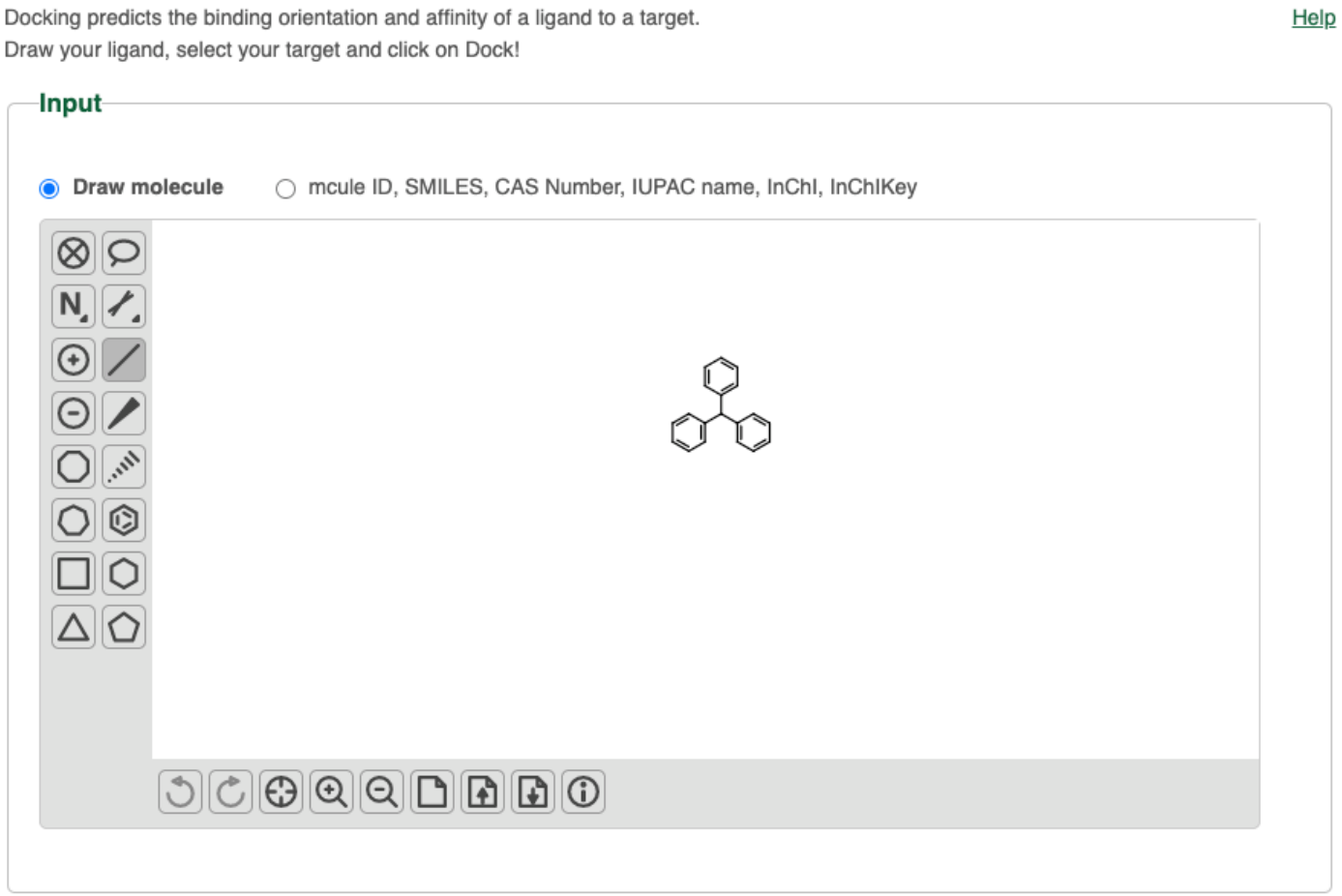
```
Detected 8 CPUs
Reading input ... done.
Setting up the scoring function ... done.
Analyzing the binding site ... done.
Using random seed: -1346177968
Performing search ... done.
Refining results ... done.

mode |   affinity    | dist from best mode
     | (kcal/mol) | rmsd l.b.| rmsd u.b.
-----+------------+----------+----------
   1        -8.1      0.000      0.000
   2        -7.6      1.847      3.401
   3        -7.6      1.133      2.045
   4        -7.6      1.414      2.576
   5        -7.5      1.570      2.927
   6        -7.4      1.272      2.508
   7        -7.3      1.509      2.762
   8        -7.3      1.346      2.383
   9        -7.2      1.390      2.188
Writing output ... done.
```



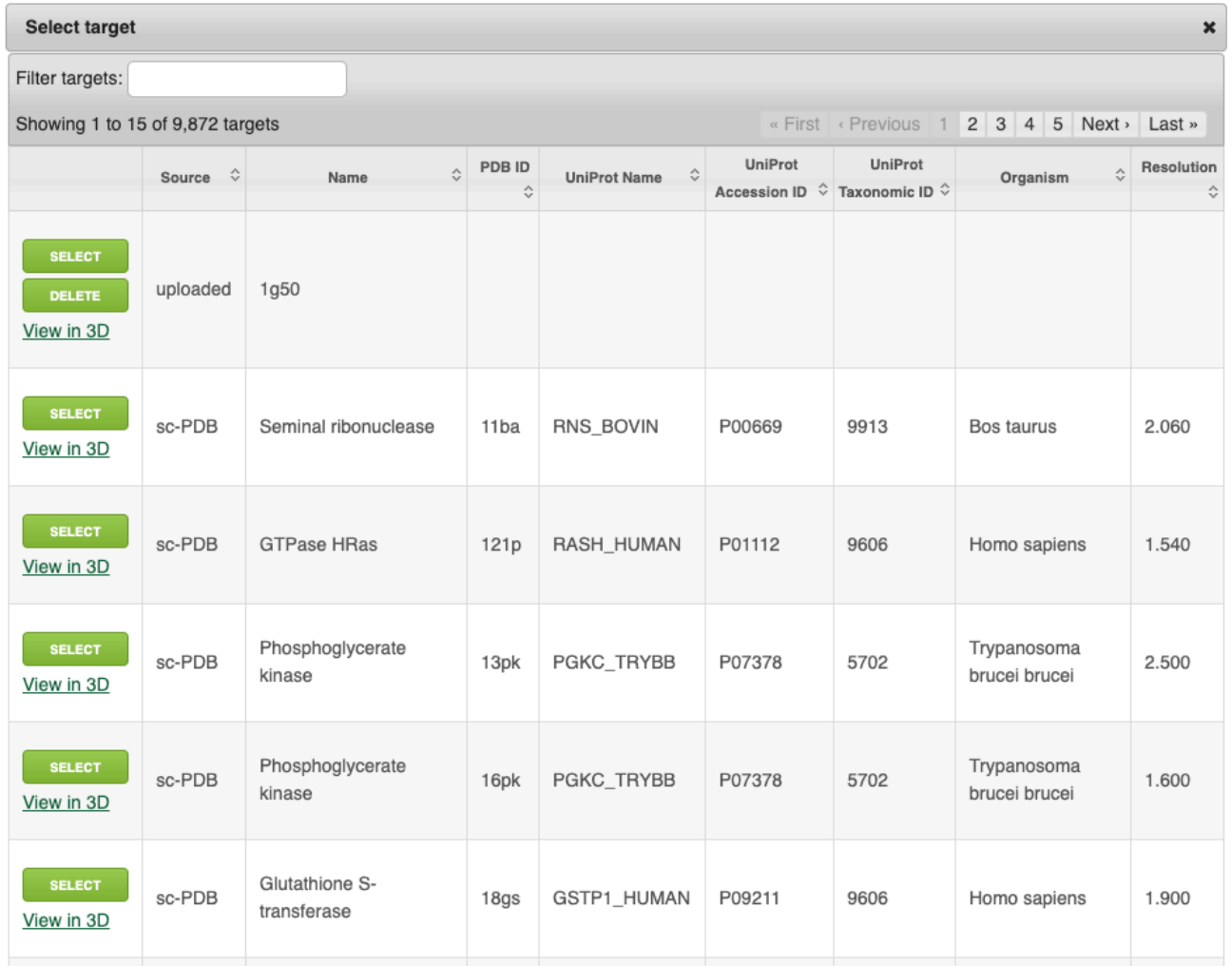
Docking with Mcule
https://mcule.com/apps/1-click-docking/

1. Draw your ligand



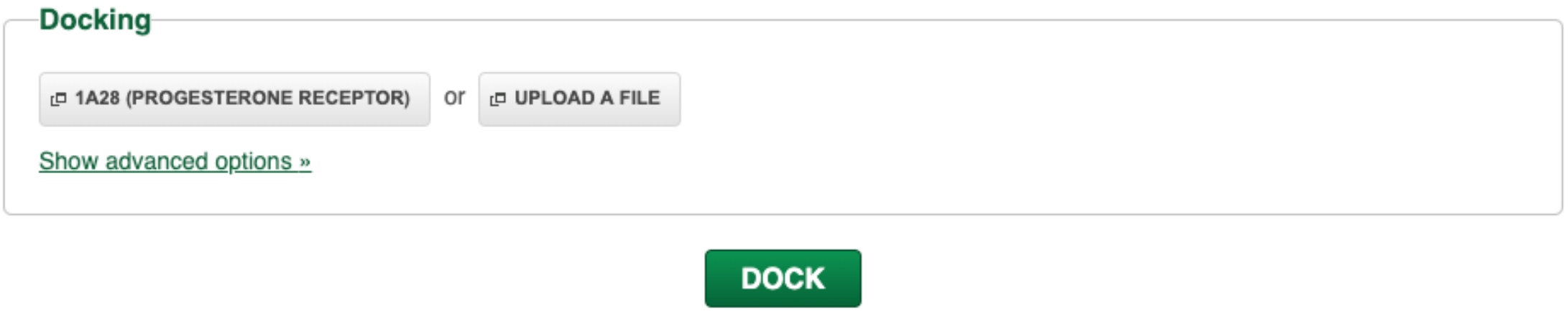
Docking with Mcule

2. Select your target



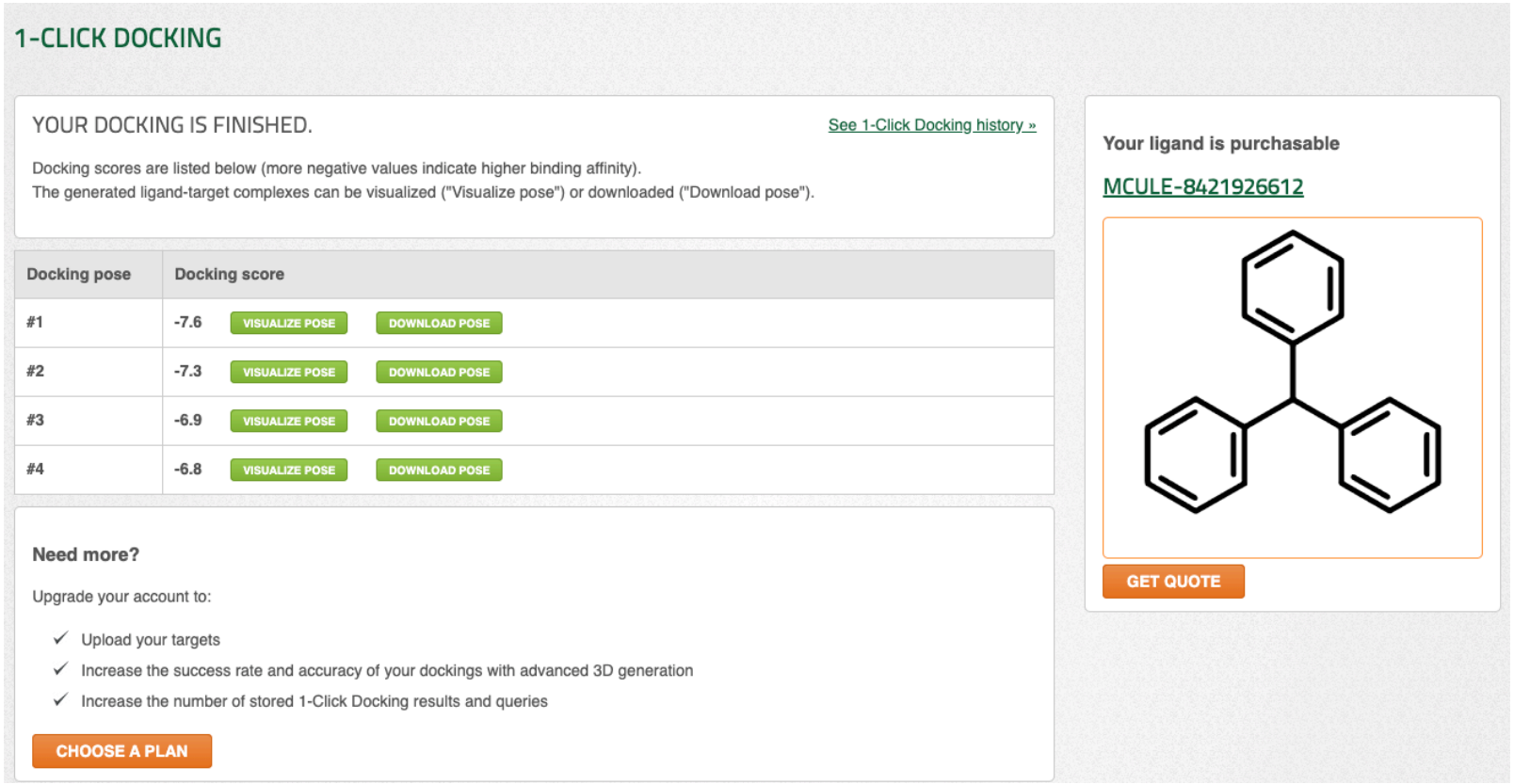
Docking with Mcule

3. Dock your ligand



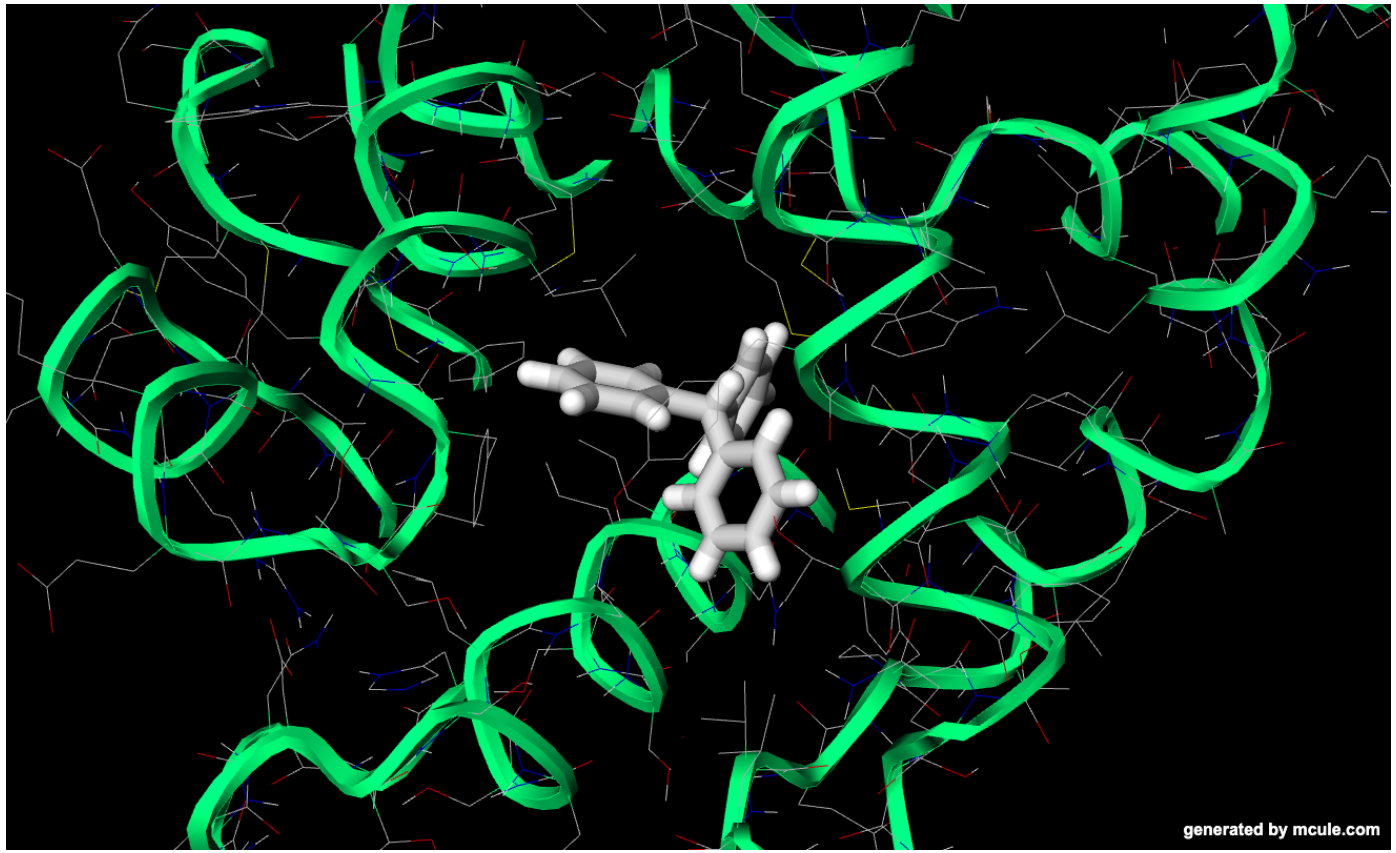
Docking with Mcule

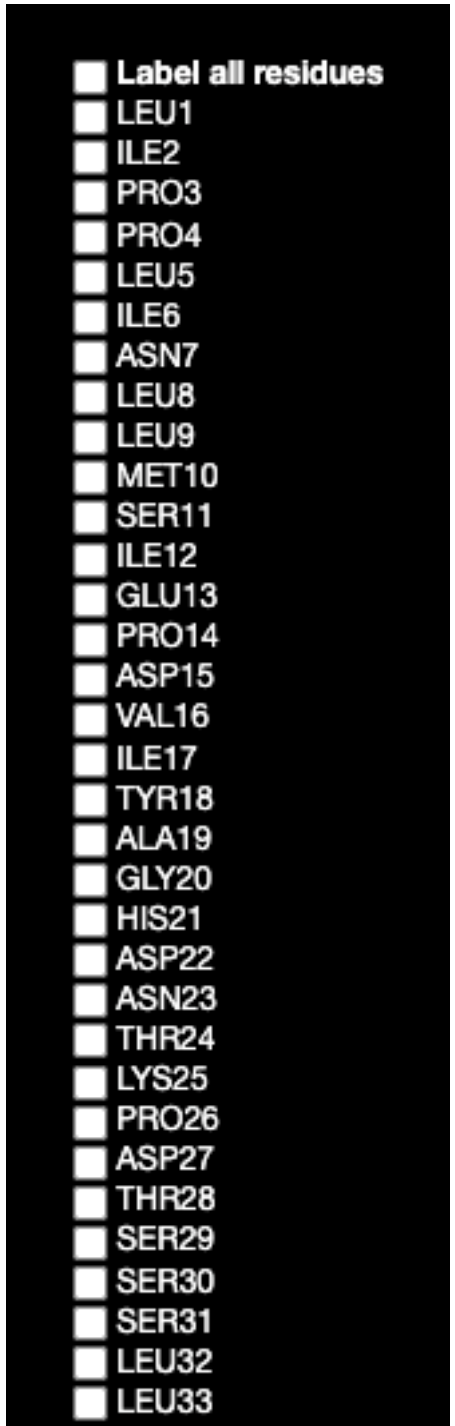
Docking this way takes only a few seconds:



Docking with Mcule

Visualize poses in Mcule or download for later use





Virtual Screening: Vina vs. Mcule vs. other software and approaches

- **Generally, VS is good at finding the right geometries (poses), in agreement with X-ray structures, but not binding energies!**

  (Reasons? poor estimation of entropy and solvent effects – trade offs for speed)

- AutoDock Vina is regarded as more accurate than Mcule however it does take considerably longer to execute and system preparation is more complicated
- Vina gives more poses than Mcule
- For accuracy of virtual screening approaches see:

Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S., A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006,** *49* (20), 5912-31.

Ferreira, L. G.; Dos Santos, R. N.; Oliva, G.; Andricopulo, A. D., Molecular docking and structure-based drug design strategies. *Molecules* **2015,** *20* (7), 13384-421.
https://pubmed.ncbi.nlm.nih.gov/26205061/

Questions?