OpenDiscovery

iscovery.org.uk

.1

areth.price@warwick.ac.uk; a.marsh@warwick.ac.uk

ple 1: TIR1 Ubiquitin Ligase + IAC

ow along with this example by downloading the initial PDB from here. Before starting, watch the video tutorial provided by the authors of AutoDock Vina, here.

g the files

d to split the crystal structure of the complex into separate receptor and ligand files. The first section of a PDB is normally the HEADER, with notes about the structure and how it was acquired. We can easily g
by using `grep` in terminal:

```
TOM\|^TER" 2P1Q.pdb > ubq_lig.pdb
```

symbol is saying "find all the lines which start with ATOM or (\|) TER", and then the filename, then we "pipe" all the output into a new file, `receptor.pdb`.

do the same thing, but with the ligands, which start with HETATM:

```
ETATM" 2P1Q.pdb > IAC.pdb
```

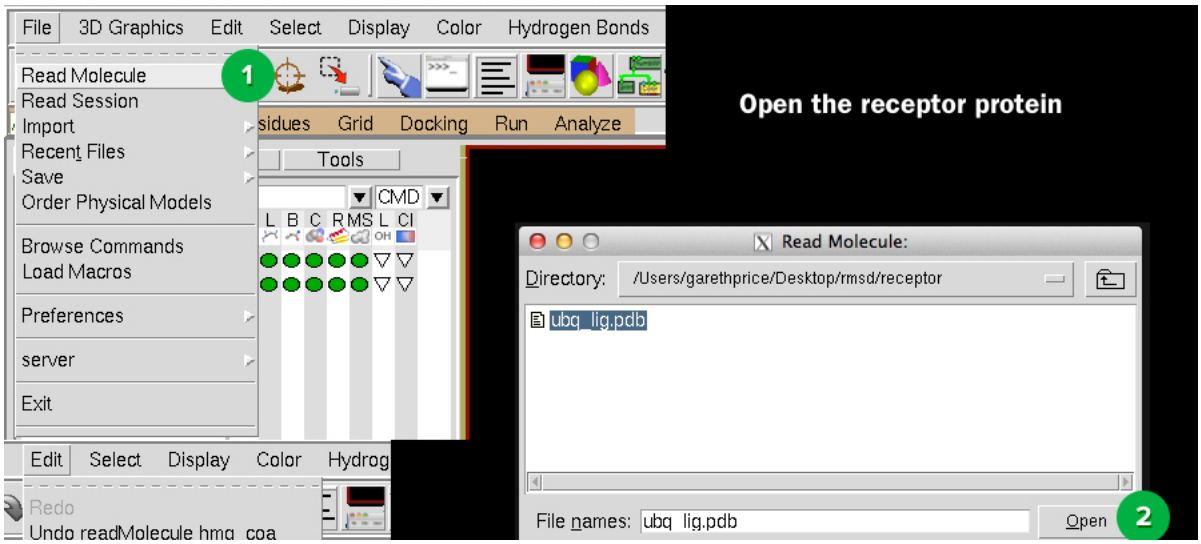
ok int the `IAC.pdb` file, it also includes IHP, an allosteric ligand that binds to another active site (which we are not interested in). To select only IAC, we can do this:

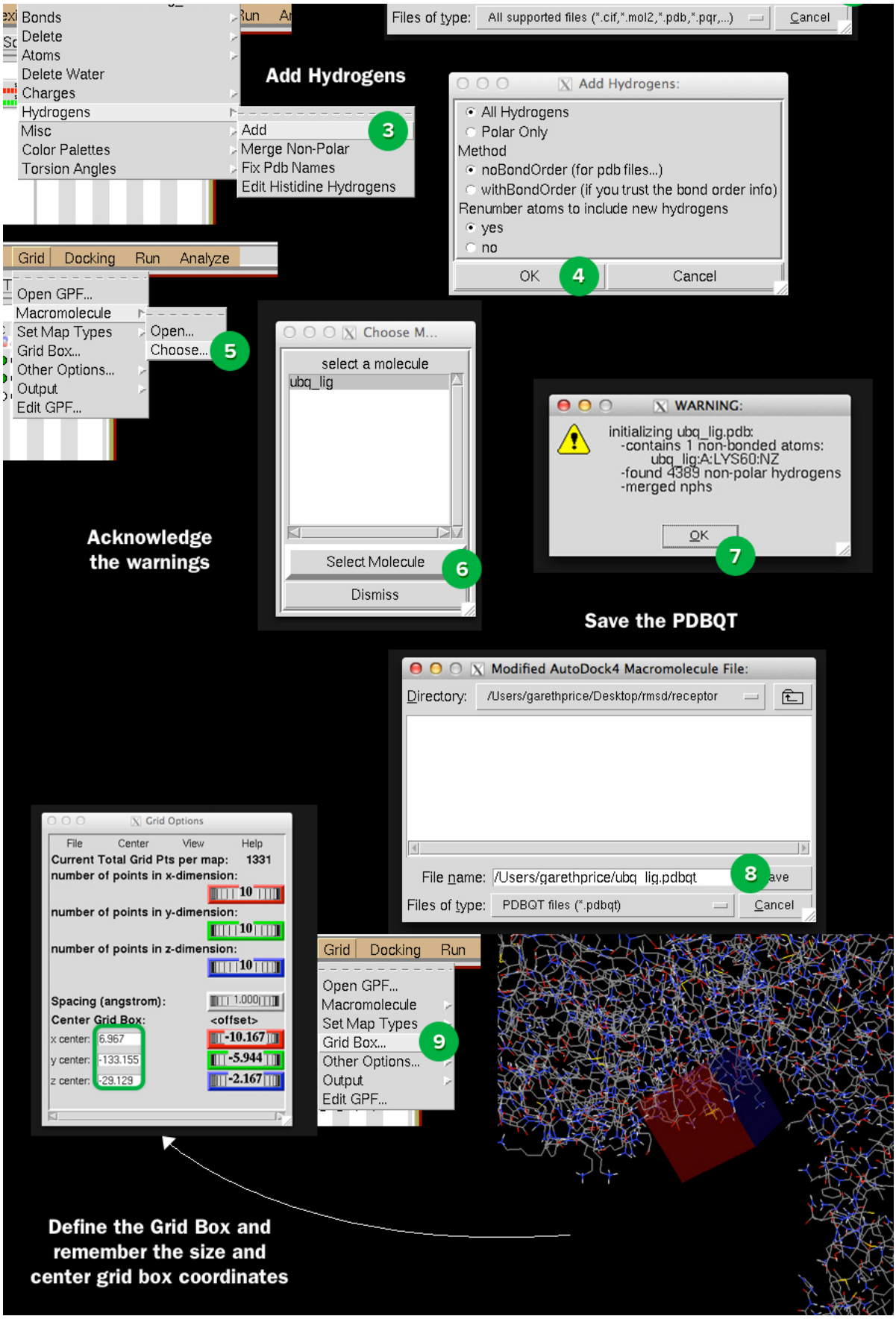
```
ETATM" 2P1Q.pdb | grep "IAC" > IAC.pdb
```

ct all the lines starting with HETATM, then select those which contain IAC, and the make the IAC.pdb file.

`bq_lig.pdb` into a **receptor** folder and `IAC.pdb` into a **pdb** folder. Next, we need to prepare the receptor protein (`ubq_lig.pdb`).

toDockTools, and follow the steps in the image below. Make sure to remember the coordinates and dimensions. Make sure Spacing (Angstrom) is set to 1.000.



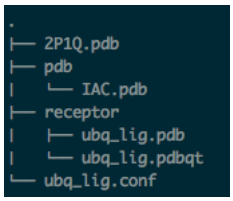

Bonds
Delete
Atoms
Delete Water
Charges
Hydrogens
Misc
Color Palettes
Torsion Angles
Add
Merge Non-Polar
Fix Pdb Names
Edit Histidine Hydrogens
3
Add Hydrogens
Files of type: All supported files (*.cif,*.mol2,*.pdb,*.pqr,...) Cancel
Add Hydrogens:
All Hydrogens
Polar Only
Method
noBondOrder (for pdb files...)
withBondOrder (if you trust the bond order info)
Renumber atoms to include new hydrogens
yes
no
OK
4
Cancel
Grid Docking Run Analyze
Open GPF...
Macromolecule
Set Map Types
Grid Box...
Other Options...
Output
Edit GPF...
Open...
Choose...
5
Choose M...
select a molecule
ubq_lig
Select Molecule
6
Dismiss
WARNING:
initializing ubq_lig.pdb:
-contains 1 non-bonded atoms:
ubq_lig:A:LYS60:NZ
-found 4389 non-polar hydrogens
-merged nphs
OK
7
Acknowledge the warnings
Save the PDBQT
Modified AutoDock4 Macromolecule File:
Directory: /Users/garethprice/Desktop/rmsd/receptor
File name: /Users/garethprice/ubq_lig.pdbqt
8
Files of type: PDBQT files (*.pdbqt)
Cancel
Grid Options
File Center View Help
Current Total Grid Pts per map: 1331
number of points in x-dimension: 10
number of points in y-dimension: 10
number of points in z-dimension: 10
Spacing (angstrom): 1.000
Center Grid Box: <offset>
x center: 6.967 -10.167
y center: -133.155 -5.944
z center: -29.129 -2.167
Grid Docking Run
Open GPF...
Macromolecule
Set Map Types
Grid Box...
9
Other Options...
Output
Edit GPF...
Define the Grid Box and remember the size and center grid box coordinates

file and enter the box coordinates and dimensions such as:

```
= 6.967
= -133.155
= -29.129

10
10
10
```

`ubq_lig.conf` .

have a folder that looks like:



Compare Docked IAC to Crystal Structure

we run ODScreen to dock just the IAC molecule into the active site of the Ubiquitin Ligase protein. Using both the resulting docked *and* crystal ligand-protein complex we can calculate an RMSD (Root Mean S
alue, giving an indication of how well Vina performs.

heck that everything is installed properly. In terminal, `cd` to the Protocol Folder (i.e. the folder with `odscreen.py` etc.) and run `python odcheck.py` . Check that there are no failures, and follow the insta
re are. You can ignore the warnings. Now, we can run the screening protocol. Navigate to the folder where ubq_lig.conf is located (from before). Now we can run `python odscreen.py` . Make sure you use
`odscreen.py` file.

`protocolfolder/odscreen.py -d . -r ubq_lig -i pdb -c ubq_lig.conf`

esult:

```
---------------------------------- #
       OPEN DISCOVERY              #
      Screening Module             #
---------------------------------- #
n:  1.0.1                          #
    www.opendiscovery.org.uk       #
ts: gareth.price@warwick.ac.uk     #
    a.marsh@warwick.ac.uk          #
---------------------------------- #
r: /Users/garethprice/Desktop/rmsd #
    Receptor Name: ubq_lig         #
       Input Type: pdb             #
      Conf: ubq_lig.conf           #
      Exhaustivness: 50            #
---------------------------------- #
Started: Thu, 12 Sep 2013 23:25:55 #
---------------------------------- #


            SMI
     Writing smi/IAC.smi


            MOL
     Writing mol/IAC.mol


            MOL2
    Writing mol2/IAC.mol2


           IMAGES


         MINIMISATION
        Minimising IAC
```

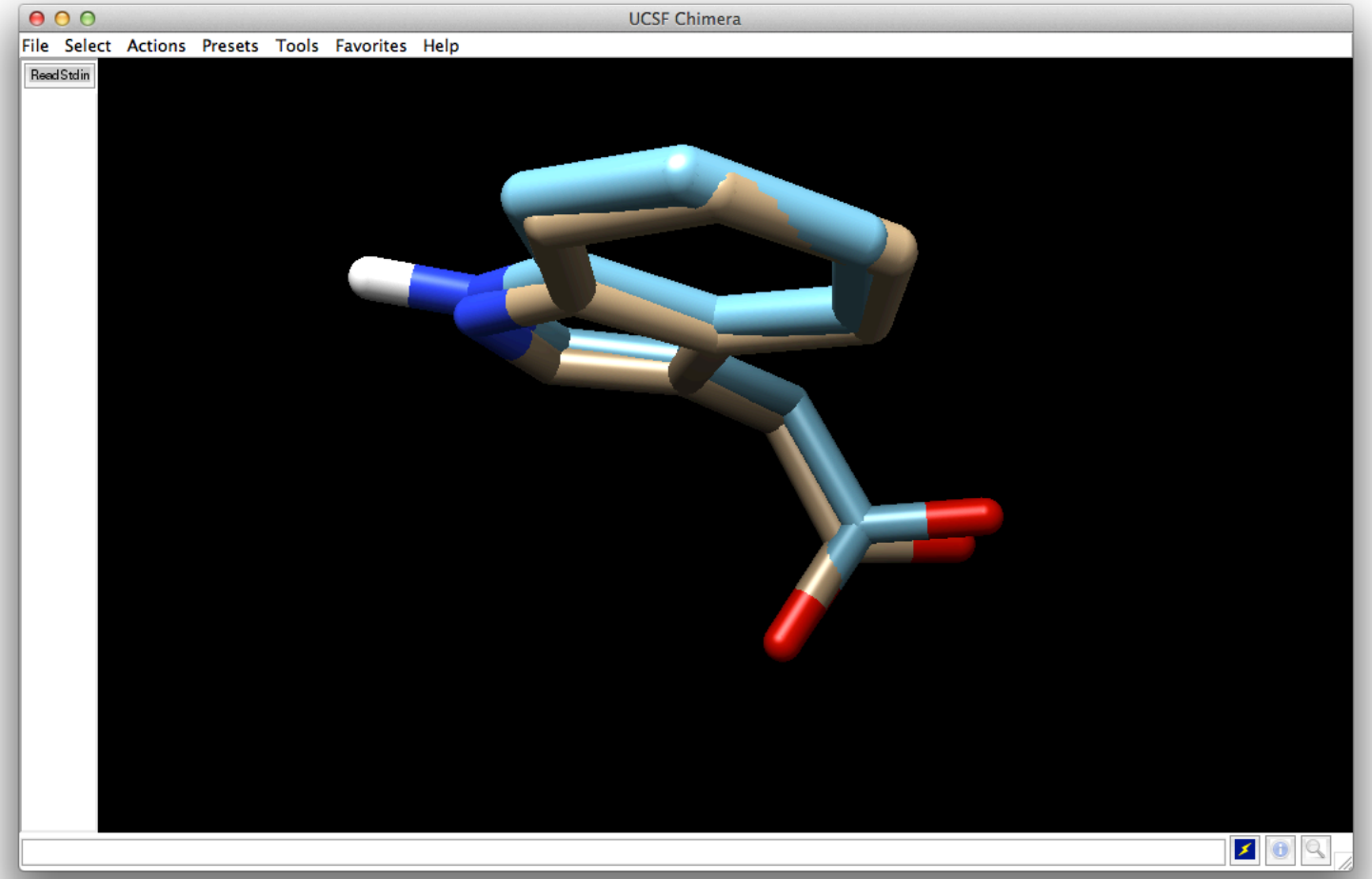
```
       PDBQT PREPARATION
      Writing IAC.pdbqt


          SCREENING
        Processing IAC


          EXTRACTING
   Processing results/IAC/


          PDB -> MOL2
Writing results-mol2/IAC.mol2


          SUMMARISING
       Summarising IAC


       MAKING COMPLEXES
         Writing IAC


----------------------------------- #
           FINSHED                  #
  Time Taken: 54.57 seconds         #
----------------------------------- #
```

g files are:

```
.
├── 2P1Q.pdb
├── combined.mol
├── images
├── mol
│   └── IAC.mol
├── mol2
│   └── IAC.mol2
├── od_log.txt
├── pdb
│   └── IAC.pdb
├── pdb-minimised
│   └── IAC.pdb
├── pdbqt
│   └── IAC.pdbqt
├── receptor
│   ├── ubq_lig.pdb
│   └── ubq_lig.pdbqt
├── results
│   ├── IAC
│   │   ├── IAC_mode_1.pdb
│   │   ├── IAC_mode_10.pdb
│   │   ├── IAC_mode_2.pdb
│   │   ├── IAC_mode_3.pdb
│   │   ├── IAC_mode_4.pdb
│   │   ├── IAC_mode_5.pdb
│   │   ├── IAC_mode_6.pdb
│   │   ├── IAC_mode_7.pdb
│   │   ├── IAC_mode_8.pdb
│   │   ├── IAC_mode_9.pdb
│   │   ├── log.txt
│   │   ├── out.pdbqt
│   │   └── ubq_lig.pdb
│   ├── complexes
│   ├── summary-sorted.csv
│   ├── summary.csv
│   └── summary.txt
├── results-mol2
├── smi
│   └── IAC.smi
└── ubq_lig.conf
```

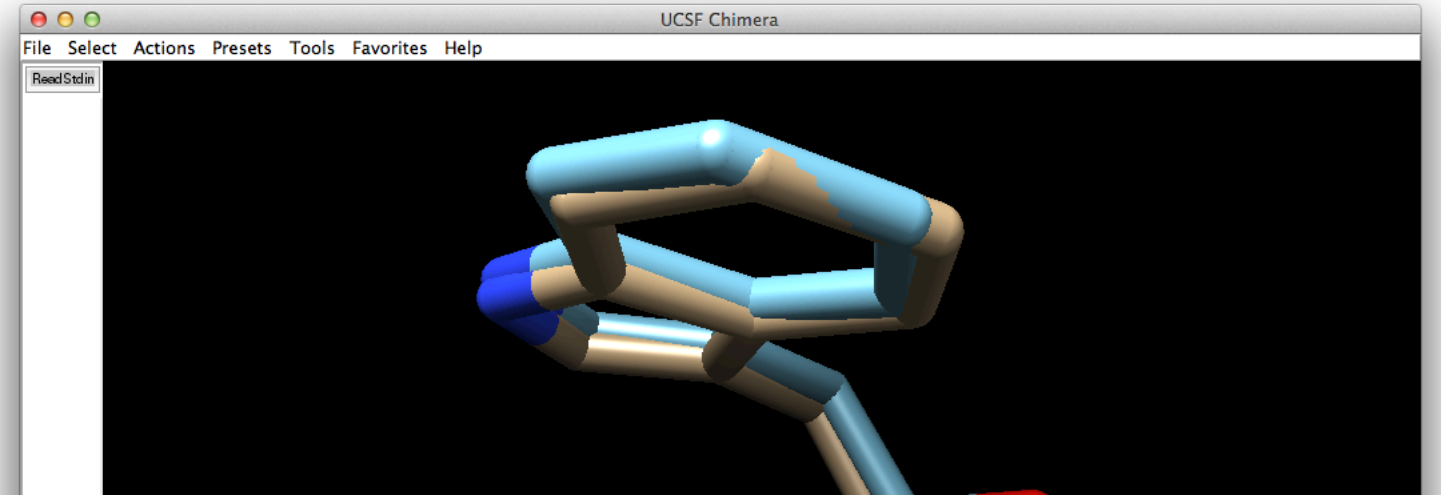
otocol is complete, start Chimera (UCSF), then open both `pdb/IAC.pdb` and `results/IAC/IAC_mode_1.pdb`:

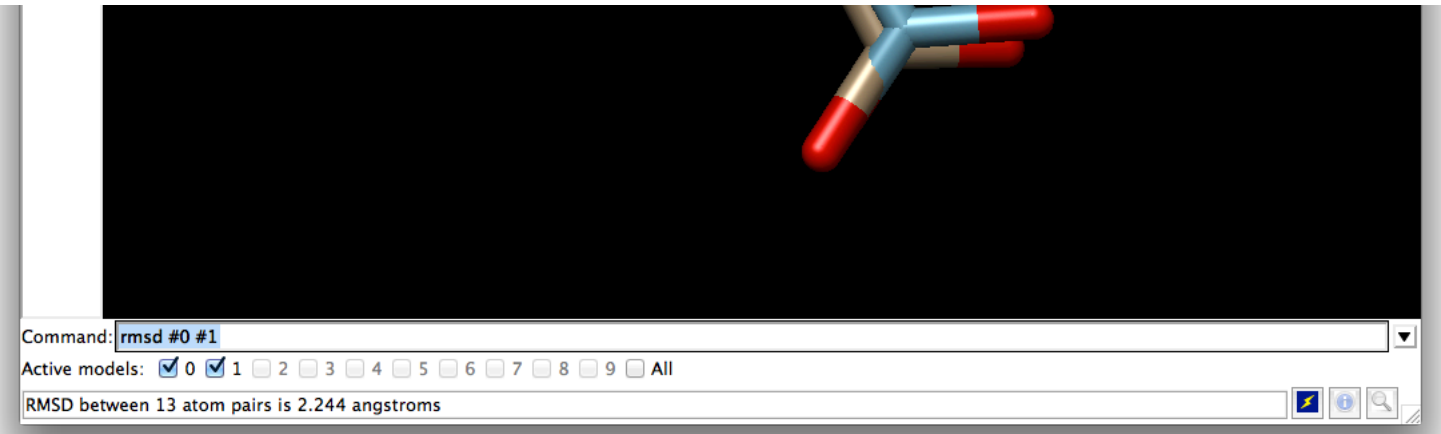
otboot is complete, start Chimera (UCSF), then open both `pub/1AC.pub` and `results/1AC/1AC_mode_1.pub`.



ction tells us that the docking has worked well. We can get a numerical value for RMSD by opening the command line (by Tools->Command Line->Raise). First, select the "Select" toolbar item, then "Chemistr
en "H". Now go "Actions"->"Atoms/Bonds"->"Delete".

and line bar, type `rmsd #0 #1` :

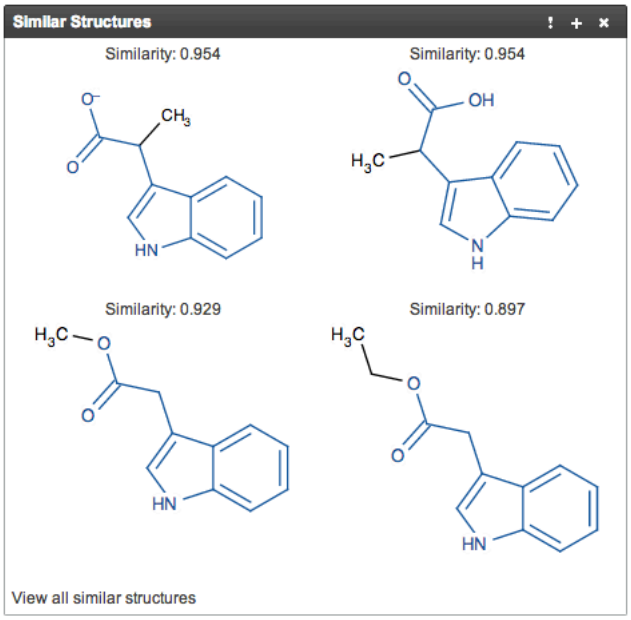




f 2.224 Angstroms is perfectly reasonable.

Generate a library of similar ligands and dock them to the protein

lenty of methods of generating libraries of similar compounds, and you may already have a library of ligands that you are interested in. An easy way is to use Chemicalize. Search for `indolylacetic acid` , ilar structures".



wnload on the top right and choose Smiles as the method of download:

to the folder where your `ubq_lig.conf` file is located. Rename the file to `smiles.smi` , then we can run `odscreen` to automatically go through the whole list and perform the procedure we did in Task d. The file downloaded will contain thousands of chemical ligands– it might be wise to split them into chunks if computer time is limited. Odscreen has a limit of 999 ligands at anyone time, too.

`odscreen` on the smiles.smi file with:

```
protocolrotuer/odscreen -d . -r ubq_rig -l smilestext -c ubq_rig.conf
```

es:

```
---------------------------------- #
       OPEN DISCOVERY              #
      Screening Module             #
---------------------------------- #
n:  1.0.1                          #
    www.opendiscovery.org.uk       #
ts: gareth.price@warwick.ac.uk     #
    a.marsh@warwick.ac.uk          #
---------------------------------- #
r: /Users/garethprice/Desktop/rmsd #
    Receptor Name: ubq_lig         #
    Input Type: smilestext         #
      Conf: ubq_lig.conf           #
      Exhaustivness: 20            #
---------------------------------- #
Started: Thu, 12 Sep 2013 23:46:58 #
---------------------------------- #
cules converted
s output. The first is smi/compound1.smi

    Splitting smiles file


            MOL
  Writing mol/compound1.mol
  Writing mol/compound10.mol
 Writing mol/compound100.mol

. .
```

d, you can perform analysis (visual or otherwise) on the docked complexes. A list (`summary-sorted.csv`) is produced which ranks the ligands based on their calculated free binding energy. This can be op
ormal text editor.