

PRAKTIKUM - FRONTIERS IN APPLIED DRUG DISCOVERY

---

# **PYMOL: ALIGNMENT OF HALOGEN COMPLEX AND PSEUDOATOMS TO A LIGAND**

---

December 10, 2018

Name: Timo Lucas  
Matrikelnummer: 3579800  
Universität Tübingen

## INTRODUCTION

In drug design the formation of halogen bonds between a halogen and a polar atom can contribute to the docking of a ligand to a binding pocket and is therefore actively researched. To investigate the effects of polar atoms in the binding site on the strength of the halogen bond various calculations were made. The result of these calculations were several Pymol sessions containing a halogen complex and a molecule containing a polar atom that it forms a halogen bond with. Negative and positive charges were placed in a grid around the whole complex, measuring the effects of each charge on the strength of the halogen bond. The results of these calculations are presented in pseudo atoms that were colored according to their effect on the binding strength (red: positive effect, blue: negative effect, yellow: neutral). My goal was to write a script that takes this complex and the grid around it and aligns it with a user-defined ligand of a pdb file of choice. After the translation and rotation of the point cloud charged residues within the point cloud are marked for the user. The script is written in python utilizing Pymol.

## THE SCRIPT

The script is called align.py and can be run in Pymol using 'run align.py'. The script will then iterate over all PDB files in the pdbs folder and try to align the predefined halogen complex with the point cloud given as PSE file to the ligand of the PDB file. It finds all halogen bonds present in the PDB file between polar atoms of the ligand's backbone and the halogen complex for each halogen in the ligand. A halogen bond is considered valid if it has a length below 5 angstrom and an angle higher than 140 degrees. Of these halogen bonds the shortest one is chosen for the further calculations. The script continues creating pseudo atoms representing the ligand and its backbone. These pseudo atoms are translated to the origin of the coordinate system, rotating them to exactly fit the halogen complex and backbone of the used PSE file. The translation vector and rotation angles are saved and applied in reverse order on the PSE file as well as the point clouds to align them to the ligand in the binding pocket. After the alignment for each point cloud a pseudo atoms in the center of mass is created using the "center-of-mass.py" script and the distance from each charged atom in the residues around the point clouds are measured. If the distance is below a certain threshold the charged residue is considered to be inside the cloud and is thus reported back. The script also rates the residues due to their influence on the halogen bond displaying if it strengthens, weakens it or has no effect on it. A text file for each session is create in the

session\_info folder containing important information about each run as the used halogen bond, PDB file and the residues in the point cloud. Each session is then also saved in the pses folder.

A helper script was also written that takes a PSE file as input with a charge cloud made up of thousands of individual atom selections. The script then takes these selections and merges them into two new selections, deleting the old ones. This reduces the number of selections to load from several thousands to two, one for the halogen complex and one for the backbone of the ligand. Using the merged PSE file as input for the main script results in a significant speed up.

Possible user modifications include the manual selection of a PDB file at the beginning of the script. It is also possible for the user to use a specific halogen or halogen bond for the alignment. See the comments in the code for specific instructions how to accomplish this. One could also change the distance and angle required for a halogen bond to be considered valid and the distance cutoff of the charged residues that should be considered near each point cloud.

## SCRIPT DETAILS

```
pdb = "5osw"  
hal_bond = 0  
se ="session_chlorine_minus1_merged"
```

At the beginning of the script one can specify a PDB code if one is only interested in this a certain structure. By changing the hal\_bond parameter one can change the halogen bond that is selected for aligning the backbone if there should be more than one. The se parameter is used to specify the session used for computing the alignments.

Important functions:

- rotate\_sel(): Rotates a selection
- zero\_x: Rotates a selection so that the x coordinate of a specified atom becomes 0.
- zero\_y: Rotates a selection so that the y coordinate of a specified atom becomes 0.
- zero\_z: Rotates a selection so that the x coordinate of a specified atom becomes 0.
- find\_halogen\_bonds: Returns a list with polar atoms forming halogen bonds.

- `rate_residues`: Rates a residue according to its effect on the halogen bond
- `align_x_neg`: Aligns the selection to the negative x-axis of the coordinate system. This is used for fast rotation of the pseudoatoms to get the rotation angles.
- `init`: Initialises important settings at the start of each run `align_ligand`: Performs rotation and translation of an object to fit the ligand.

After the function definition the main body of the script starts. There one can find the following two lines:

```
for file in files:
    pdb = file[0:4]
```

Iterate over the files in the `pdb`s folder and choose the respecting file as input. If only one `pdb` file should be examined one has to comment out this line:

```
for file in files:
    #pdb = file[0:4]
```

Then one also needs to break the for loops at the end of the script like this:

```
    if len(aas) > 0:
        cmd.save("pses/"+filename+".pse")
        c_hal+=1
        break
break
```

For the script to function properly it is important that all the selections are actually present and that the atoms are present in the `pdb` file. The cases that there are no charged residues and no halogen bonds are handled with exceptions. If the script should not work for a certain `pdb` file or produce wrong results it might be good to first look at the selections made and if they select what they are supposed to. Two selections in particular caused problems with some `pdb` files, namely: `halligand` (halogen in the ligand) and `c2ligand` (second carbon atom bound to the carbon that is bound to the halogen). For debugging such errors it is useful to comment the code block that cleans up intermediate selections at the end of the script:

```
#clean up selections no longer needed
cmd.hide("lines")
cmd.show("sticks","ligand_circ")
cmd.hide("sticks","circ_complex")
...
```

For changing the information that is saved for each session in the text file one has to adapt the following code block found at the end of the script:

```
with open("session_info/"+filename+".txt","w") as file:
    file.write("File with charge cloud: "+str(se)+ "\n")
    file.write("Used pdb: "+str(pdb)+ "\n")
    file.write("Index of halogen bond: "+str(hal_bond)+ "\n")
    file.write("Index of halogen: "+str(c_hal)+ "\n")
    file.write("Atom ID of halogen: "+str(halligand)+ "\n")
    file.write("List of residues in point cloud: "+str(aas)+ "\n")
```