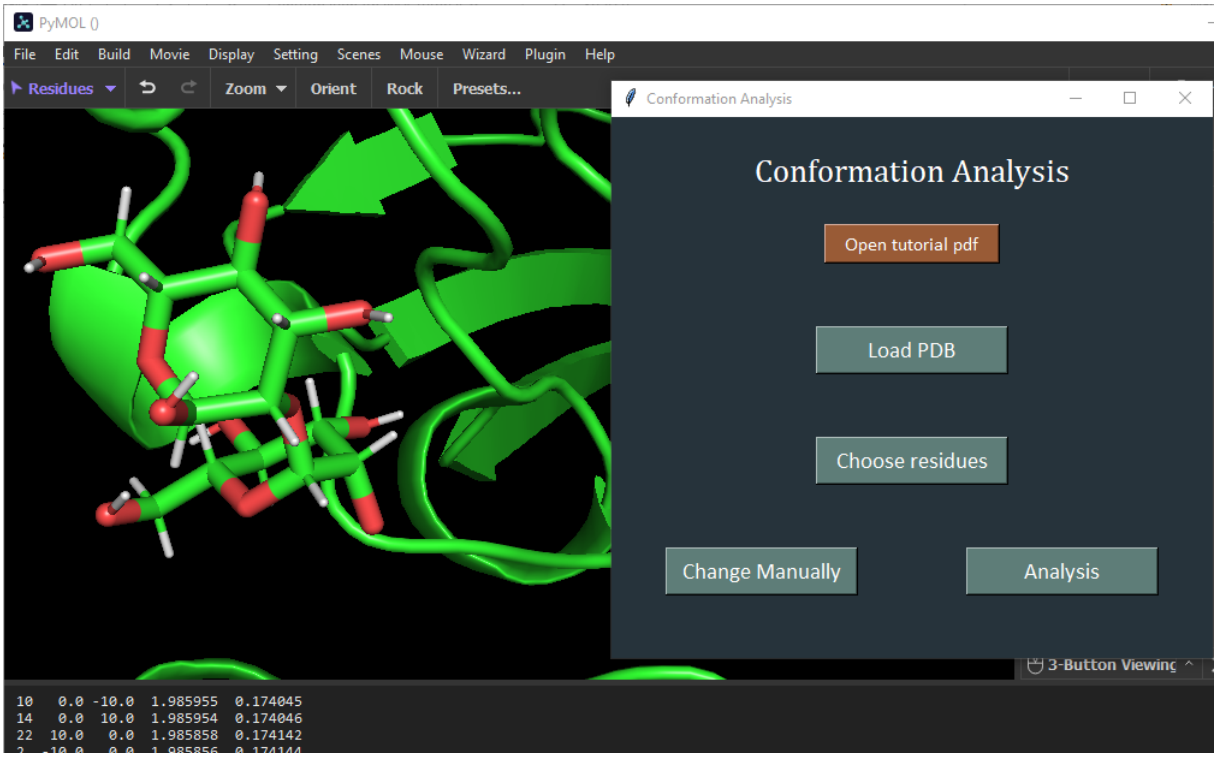
Conformation Analysis

Tutorial PDF



Creator: Matilda Jägare Lindvall

2024

Introduction

Conformation Analysis is a program written during my master's thesis and was initially created to analyse the conformations of substructures of Man-9. Therefore, the program is limited to some extent but I believe it could be used for other purposes as well.

One of the program's limitations is that it only handles single-state structures (the structure is assumed to have one conformation or one major). Hence, there should only be one minimum when looking at the heatmap. Two minima **do not** mean two conformations with a 50/50 distribution!

Moreover, the analysis part of the program is limited to only looking at the dihedral angles φ and ψ. Even if other dihedral angles can be changed, the plot can only be generated for these dihedral angles.

In this tutorial the disaccharide in the PDB file 2RDK[1] will be used as an example.

Have fun!

[1] Fromme R, Katiliene Z, Fromme P, Ghirlanda G. 2008. Conformational gating of dimannose binding to the antiviral protein cyanovirin revealed from the crystal structure at 1.35 Å resolution. Protein Science : A Publication of the Protein Society 17: 939–944.

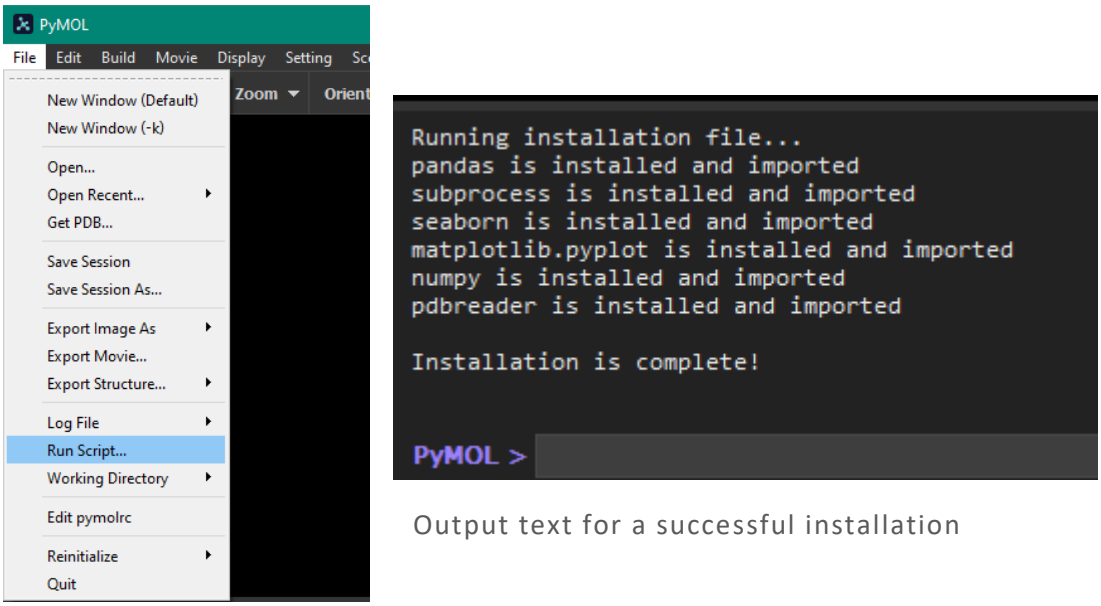
1 Getting Started

Conformation Analysis was created as a script for The PyMOL Molecular Graphics System, Version 3.0.3 Schrödinger, LLC PyMOL. If older versions of PyMOL are used problems can occur when running the program. If that is the case, try a newer version of PyMOL.

Installing packages

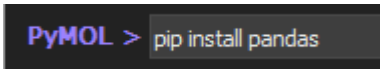
When running the program for the first time a couple of packages need to be installed in PyMOL. This is done by first opening the PyMOL application. Then go to File → Run Script… and choose the script named ‘installation_file’.

If the installation succeeds there will be an output text in PyMOL.



Output text for a successful installation

The required packages are: 'pandas', ‘subprocess’, 'seaborn', 'matplotlib.pyplot', 'numpy', 'pdbreader' and ‘openpyxl’. If an error occurs during installation, the package can be installed manually by writing “pip install *package name*”.

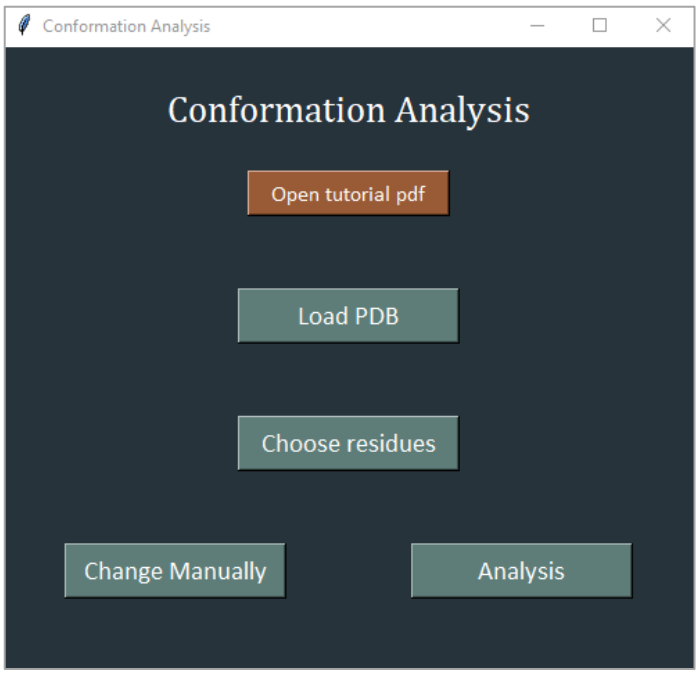


To see which packages that are installed in PyMOL the command “pip list” can be used. This will show the installed packages as a list and the version installed. The packages previously mentioned should be seen in this list after installation.

Run the installation file when using the program for the first time. After that, this step can be skipped. Make sure the installation is complete before continuing!

Starting the program

Start the program by opening the Python file called 'MAIN' in the folder where the scripts are saved. This will open the main window of the program.



Main window in Conformation Analysis

The first button called "Open tutorial pdf" will open this pdf. This tutorial can also be found in the folder with the scripts.

Loading the PDB file

Before the program can be used, a structure has to be loaded into PyMOL. The format has to be PDB (.pdb) and can be loaded into PyMOL by dragging the file and dropping it in the PyMOL window or by going to File → Open… and choosing the file.

Then the **same** file has to be loaded into the program by pressing the button 'Load PDB' and choosing the file. The program will use the PDB file to get the existing atoms from the structure. If different PDB files are loaded into the program and PyMOL the program will not be able to find the correct atoms in the structure!

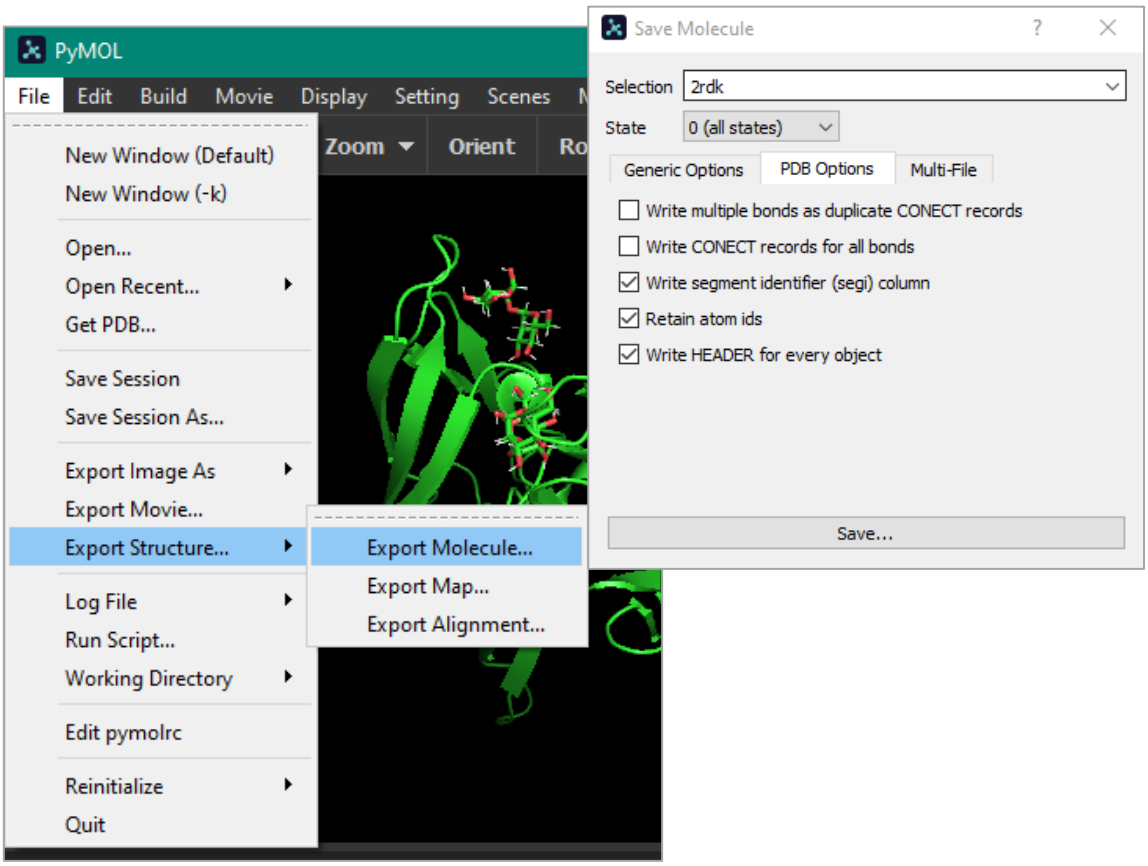
Follow along 1. Loading the PDB

Load the PDB 2RDK into PyMOL and Conformation Analysis.

Adding hydrogens to the structure

Some structures have not added the hydrogens and if the conformation analysis requires hydrogens these can be added with the command **cmd.h_add()** in PyMOL.

The next step is then to save this as a new PDB file. This can be done by going to File → Export Structure → Export molecule…



Choose the structure's name in the pop-up window in the Selection field and *0 (all states)* in the State field. Then uncross the option Original atom order (according to "rank") in the Generic Options tab. Open the other tab PDB Options and cross the boxes as shown in the figure.

Then save the file in PDB (.pdb) format. If hydrogens have been added reload the correct PDB into PyMOL and Conformation Analysis.

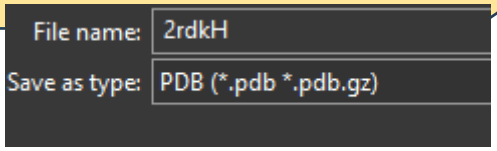
Follow along 2. Adding hydrogens

Add hydrogens to the structure of 2RDK by using the command cmd.h_add().

Note. Here we will only look at the disaccharides so remove the water and the protein.

Save the structure as a new PDB file, in this example, it is saved as 2rdkH.

Reload PyMOL and Conformation Analysis with the new PDB file.



Choose Residues

The next step is to define the residues that will be used.

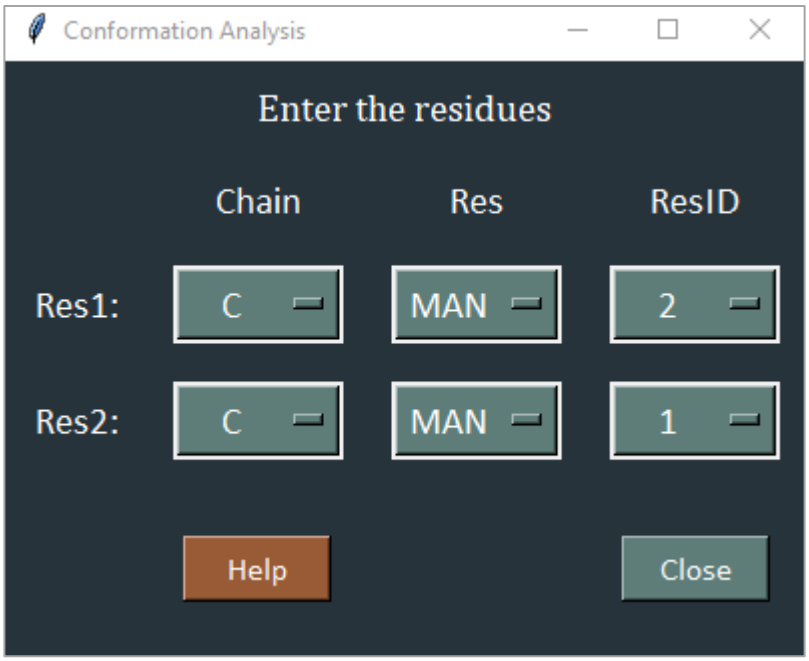
In PyMOL the residue names can be found above the sequence when displayed (Display → Sequence). It will be defined as *chain/residue/residueID* and is needed to select the correct atom. Even if only one of the residues will be used both Res1 and Res2 must be filled in.

Follow along 3. Choosing the residues

In 2rdkH there are two chains to choose between, C or D, corresponding to the two disaccharides. There will only be one residue option, MAN, since the disaccharide consists of two mannoses. With ResID one of the mannoses can be selected either 1 or 2.

Define Res1 as C/MAN/2 and Res2 as D/Man/1.

Note: The reason why we set ResID = 2 for Res1 and vice versa is because of the definition of dihedral angles (see further down).



Choose residue window in Conformation Analysis

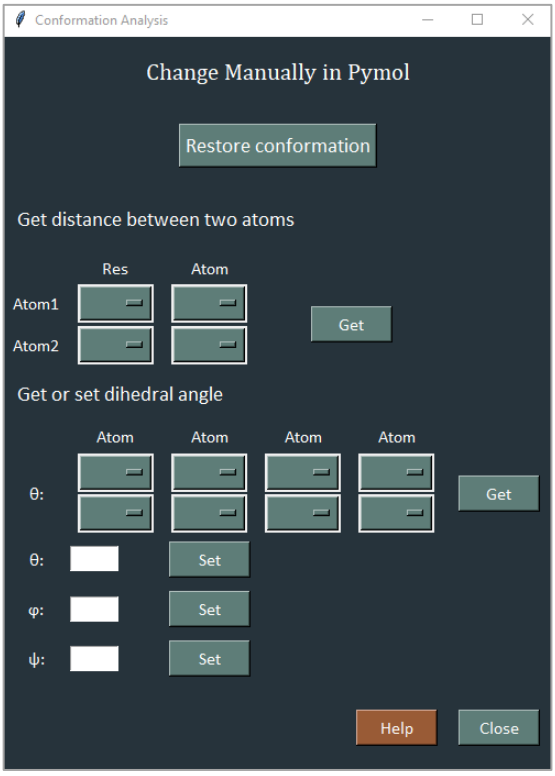
2 Change Manually

The button ‘Change Manually’ will open a new window where you can manually change the structure in PyMOL.

The top button ‘Restore Conformation’ will reload the PDB into PyMOL and reset all changes.

Get the distance between two atoms

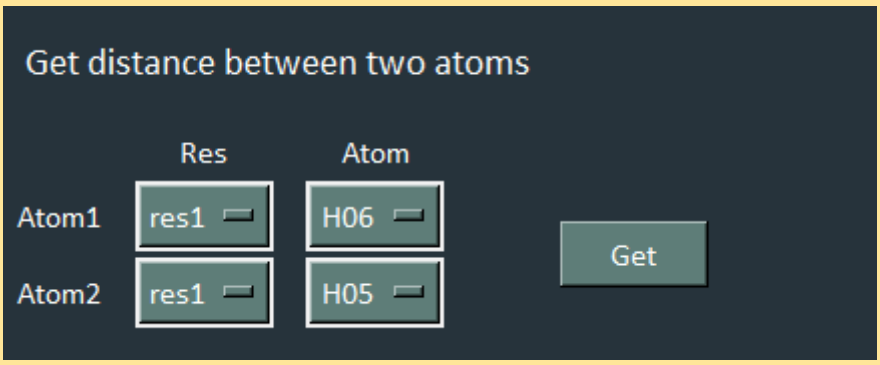
In this section, the distance between two atoms can be measured. This is by defining the two atoms and which residue they belong to.



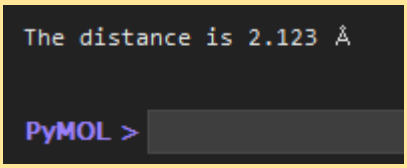
Change manually window in Conformation Analysis

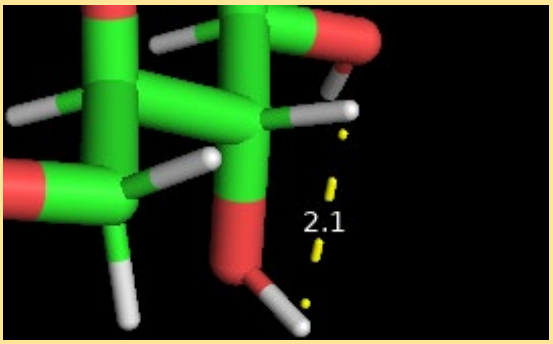
Follow along 4. Get the distance between two atoms

Measure the distance between the hydrogen at O4 and the hydrogen at C4 in Res1. In PyMOL these are labeled: H06 and H05. (Note: The labeling may differ for you!)



The output is seen in PyMOL in two ways, as a text and visually.





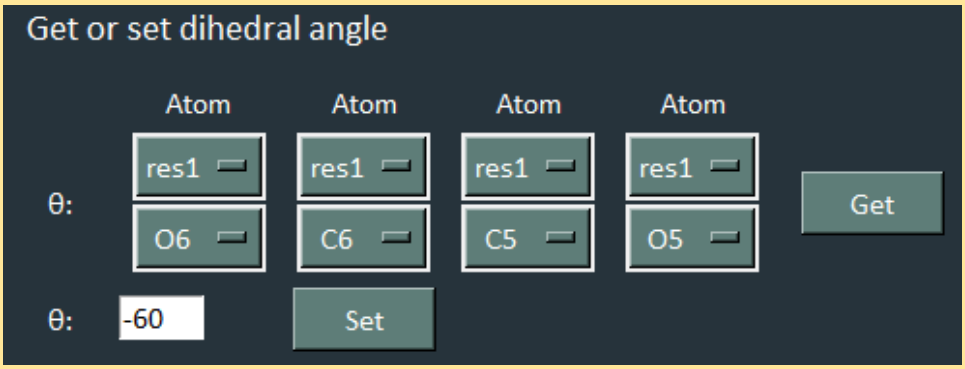
Get or set a dihedral angle

In this section, a dihedral angle can be defined and extracted from PyMOL or changed to a selected value.

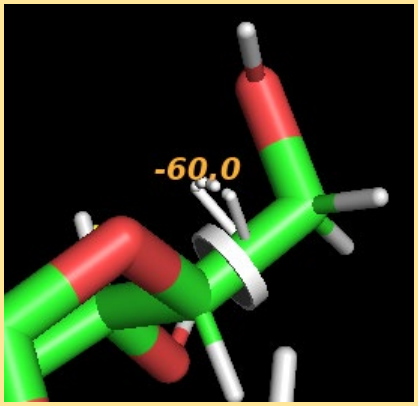
Another option is to set the dihedral angles φ and ψ, but these first have to be defined in the window Analysis.

Follow along 5. Get or set a dihedral angle

Measure the dihedral angle O6-C6-C5-O5 in Res1 and then set it to −60°.



The output will be



3 Analysis

The button 'Analysis' will open a window where a range of conformations can be tested for the dihedral angles φ and ψ.

Define the dihedral angles

First, the torsion angles need to be defined. This is done by choosing the atoms and here it is important to note how you specified the residues previously. Because the definition for these angles is

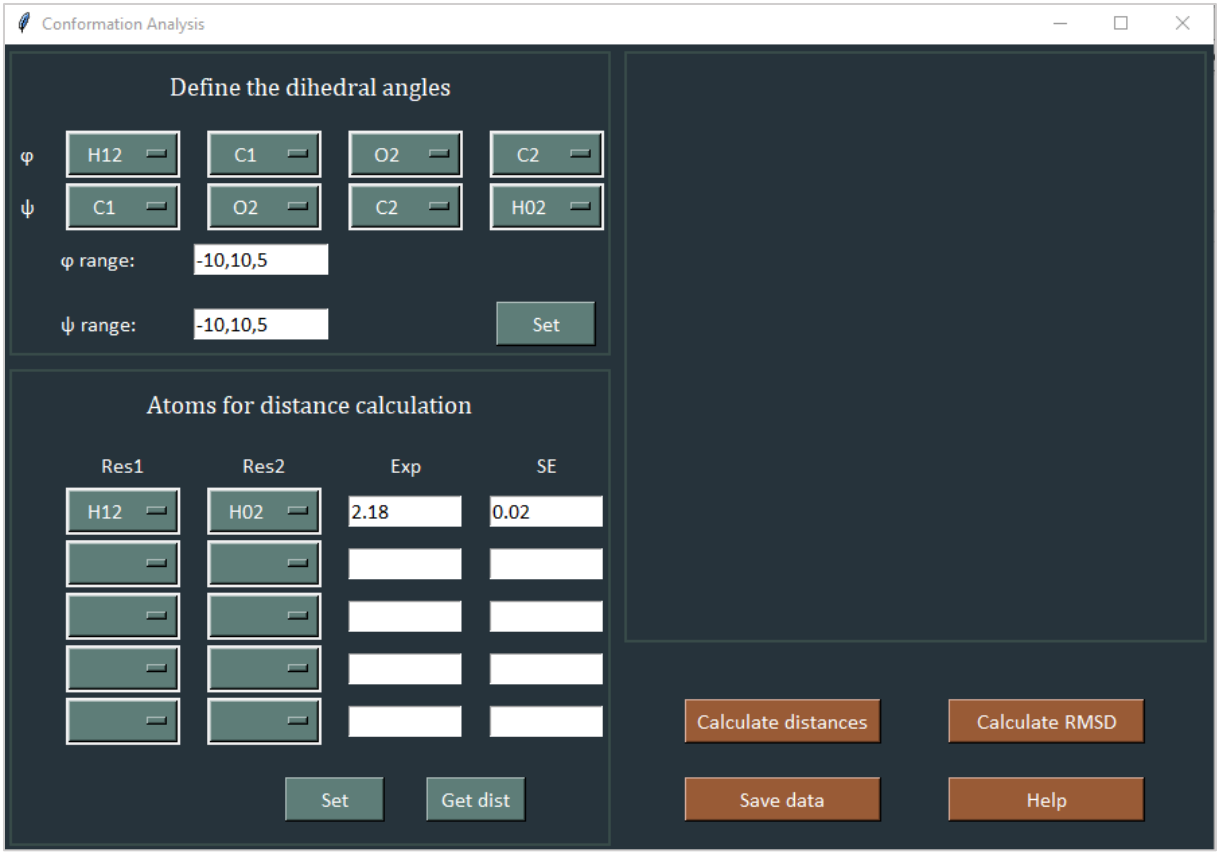
φ = Res1/X-Res1/X-Res2/X-Res2/X

ψ = Res1/X-Res2/X-Res2/X-Res2/X

where X corresponds to a selected atom.

For each dihedral angle, the range has to be selected. This is the interval that the program will use when setting the angles. The input in these fields has to be on the form: **start,stop,step**. Where start is the angle you want to begin with, stop is the angle you want to end with and step is the increment from start to stop.

After every change the 'set' button has to be pressed to apply the changes!



Analysis window in Conformation Analysis

Atoms for distance calculation

In this section, atoms for the distance calculation will be selected. Since this conformation analysis is focused on the glycosidic linkage it is only possible to calculate distances between the two residues. Therefore, the first column corresponds to atoms in Res1, and the second corresponds to atoms in Res2.

The third column is the entry for the experimentally measured distance (e.g. NOE), with the error (SE) in the fourth column. Everything has the unit Å. If there is no error, set SE to 0.

Note: Decimal values have to be written with a dot and not a comma! E.g. 3.15 and not 3,15!

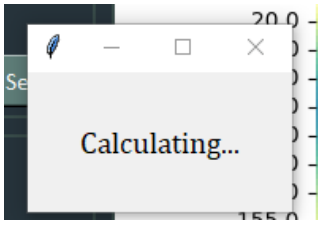
Remember to press ‘set’ to apply the changes.

The button ‘get distance’ can be used to see that the correct atoms were selected in PyMOL.

Calculate distances

The ‘calculate distances’ button will start the calculation of distances when changing the angles. This process may take a while depending on the input range and the speed of your computer. How many runs you do will also affect the speed. Therefore, it can be better to restart the program if a long run will be performed.

A loading screen will appear while the distances are calculated.



Roughly the time will be like this:

Angle range	Number of iterations	Time (s)
(-60,60,5) (-60,60,5)	$25^2 = 625$	5
(-180,180,10) (-180,180,10)	$37^2 = 1369$	20
(-180,180,5) (-180,180,5)	$73^2 = 5329$	120

If the program takes more than 5 minutes, it can be considered to restart it and try a shorter range or longer steps.

Calculate RMSD

When pressing the ‘Calculate RMSD’ button the calculated distances are compared to the experimental data by calculating the root-mean-square

deviation RMSD (root-mean-square deviation). The error (SE) will be used to give the experimental value a margin of error. Therefore, the measured value will be compared to an interval EXP ± SE. In the script, a maximum of 4 Å has been set for the distance. So, if the measured distance in PyMOL exceeds this value the difference will be set to 0.

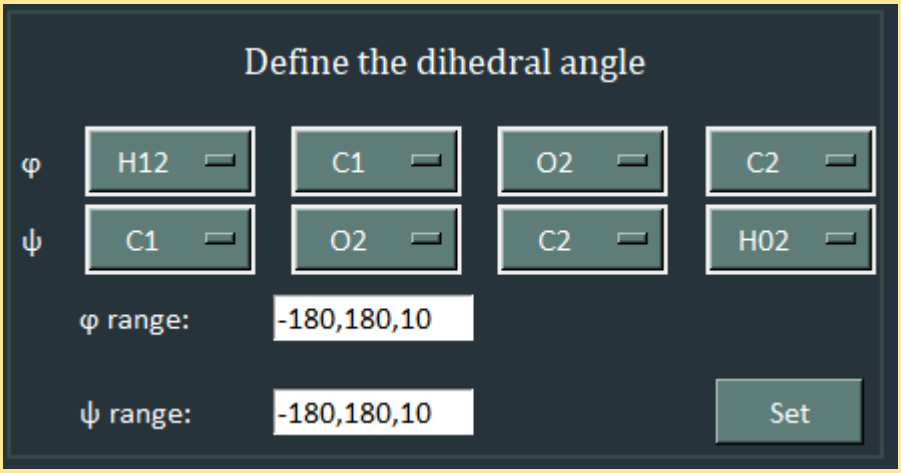
Note: The distance limit 4 Å can be changed in the script 'calcRMSD'

After calculating the RMSD the program will plot the result in a heatmap displayed in the Analysis window. A list sorted after the lowest RMSD will be displayed in PyMOL. The results can be saved by pressing the button 'save data' (saves the data as a table) and pressing the save symbol under the plot.

Follow along 6. Analysis

Perform a conformation analysis and find which value of φ and ψ that give the lowest RMSD.

The dihedral angles are defined as φ= H1-C1-O2-C2 and ψ= C1-O2-C2-H2 with the ranges -180° to 180° and a 5° increase. Insert this in the program and press 'set'.



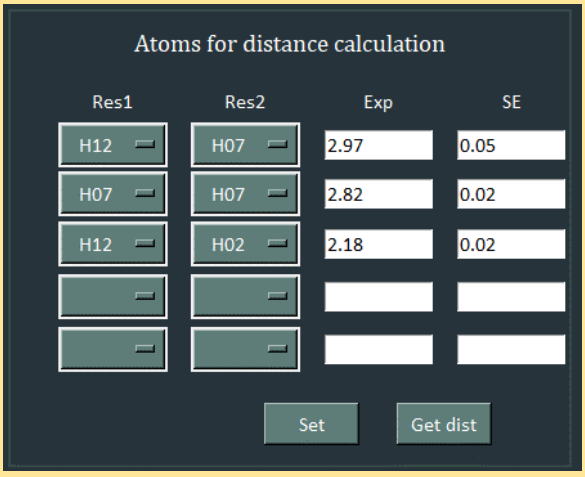
Follow along 6. Analysis cont.

The experimental data is seen in the table below

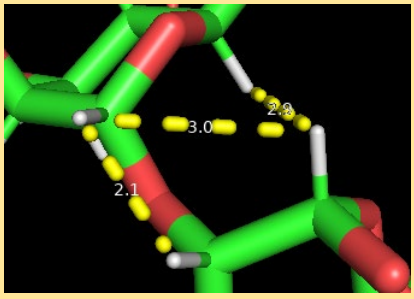
Sugar: Atoms	r	SE_r
H1/H1’	2.97	0.05
H1/H5’	2.82	0.02
H2/H1’	2.18	0.02

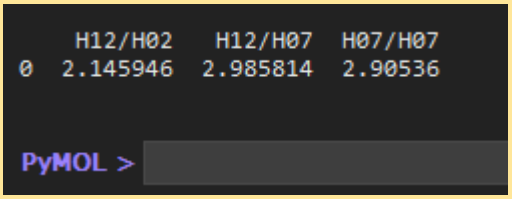
In the program H1’ = H12 (Res1), H1 = H07 (Res2), H2 = H02 (Res2) and H5’ = H07 (Res1).

Insert this in the program and press ‘set’. Then press ‘Get dist’ to visualize the selection in PyMOL.



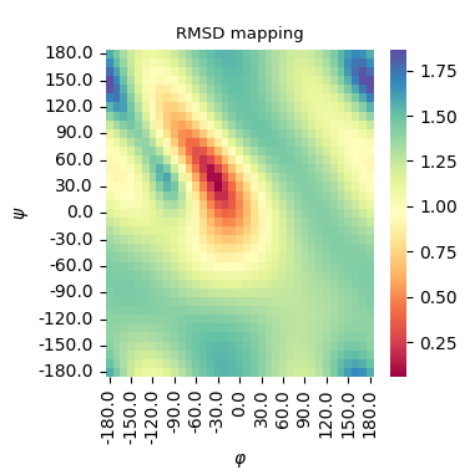
The output from ‘Get dist’





Follow along 6. Analysis cont.

If satisfied with the input press ‘Calculate distances’. When done, press ‘Calculate RMSD’. The output will be a heatmap and a table.




```
Loading...
        phi    psi   H12/H07   H07/H07   H12/H02      RMSD
576   -30.0   30.0  3.109370  2.843638  2.115197  0.057757
577   -30.0   40.0  2.861486  2.997624  2.176440  0.097073
540   -40.0   40.0  3.017629  2.628582  2.223830  0.099920
541   -40.0   50.0  2.771541  2.829270  2.297084  0.102413
612   -20.0   20.0  3.206336  3.091132  2.035155  0.194400
...     ...    ...       ...       ...       ...       ...
32   -180.0  140.0  3.707787  0.288613  4.073692  1.852101
1364  180.0  140.0  3.707815  0.288555  4.073692  1.852131
1328  170.0  150.0  3.874660  0.369786  4.120741  1.855229
1365  180.0  150.0  3.800088  0.306271  4.103322  1.866356
33   -180.0  150.0  3.800061  0.306256  4.103322  1.866359

[1369 rows x 6 columns]

RMSD calculation succeeded!
```

In this example the minimum with the lowest RMSD is at φ =−30° and ψ=30°.

Save the plot and the table.