# Seminar 2 – Basic operations with receptors, redocking, docking

Author(s): Jan Zitko ([jan.zitko@faf.cuni.cz](mailto:jan.zitko@faf.cuni.cz)) and Martin Juhás ([juhasm@faf.cuni.cz](mailto:juhasm@faf.cuni.cz))

## Required software

* Python3: <https://www.python.org/downloads/>
* AutoDock Vina: <https://github.com/ccsb-scripps/AutoDock-Vina/releases>
* Chimera v1.16 (already installed in Seminar 1): <https://www.cgl.ucsf.edu/chimera/download.html>

## THEORY AND INTRODUCTION

### How to obtain receptors

All publicly available experimental protein structures are contained within wwPDB (<https://www.wwpdb.org/>) database (approx. 198K experimental structures) which can be searched through, e.g.:

* RCSB PDB: <https://www.rcsb.org/>
* PDBe: <https://www.ebi.ac.uk/pdbe/>
* PDBj: <https://pdbj.org/>

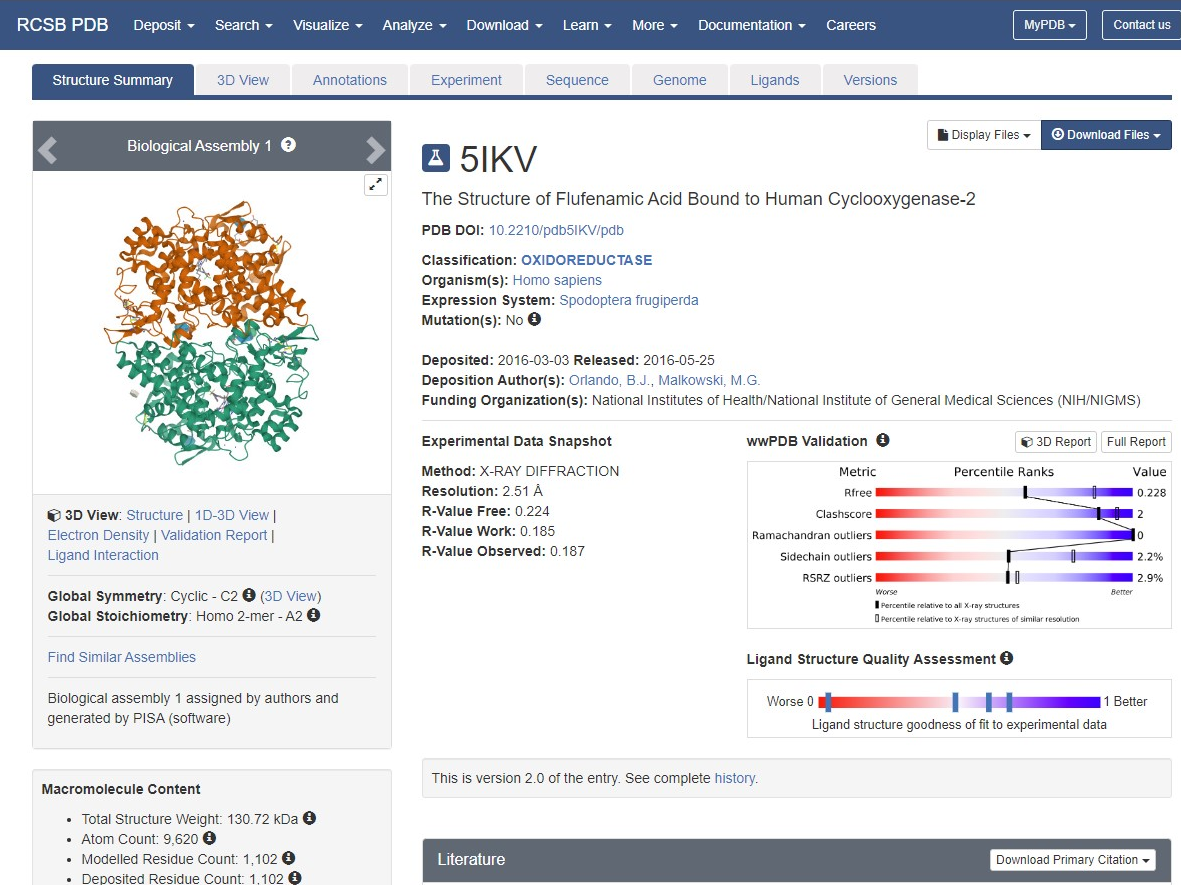
### Methods of structure determination

(details, e.g.: [PDB-101: Methods for Determining Structure (rcsb.org)](https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/methods-for-determining-structure))

* X-ray crystallography – the most common method, constitutes >85% of PDB structures. A high-quality crystal of the protein is needed.
* Cryo-electron microscopy (Cryo-EM) – in contrast to protein crystallography, it visualizes a single molecule of the protein
* NMR spectroscopy – determines the structure of proteins in solution
* Homology modelling – computational prediction of the 3D structure

### Assessing the quality of the PDB structures for docking

(details, e.g.: [Assessing the Quality of 3D Structures (rcsb.org)](https://www.rcsb.org/docs/general-help/assessing-the-quality-of-3d-structures))



Most important quality determinants (with recommended values):

* **Resolution** – defines overall structure quality, indicates how well two adjacent atoms can be distinguished
  + resolution < 2.5Å
* **R-value** – defines the agreement between the experimental diffraction data and simulated experimental data calculated based on the 3D model of the biomolecule
  + R-factor (R-value Work) < 0.2
  + R-free (R-value Free) < 0.22
* **Sidechain** and **Ramachandran outliers** – distortion of bond lengths, angles or torsion angles
  + The lower the better (<1%)

It is always important to visually check the agreement of the PDB structure and the electron density map (EDM) **in and near the active site**.

## PRACTICAL TRAINING

### Task 1 - Processing the PDB structure for docking (work in Chimera)

1. Obtain PDB structure of human COX-2, (**PDB ID: 5IKV**):

* Download directly from RCSB PDB ([RCSB PDB: Homepage](https://www.rcsb.org/))
* Use Chimera: **File-Fetch by ID...**: 5ikv

1. Remove solvents, ions, metals, or other crystallization additives:
   * NOTE: Solvent (water) if present in PDB is often removed before docking. However, specific molecules important for the ligand-protein interactions (i.e. crystallographic waters) should be retained (no such water is present in COX)
   * Using Tools-General Controls-Model Panel-Select Chain(s)… select chains B, C, D (by ID) (Apply) and delete them (Actions-Atoms-Delete).

NOTE: If you do not see **Select Chain(s)…** in the right panel, click **all** to show all available options

* + Delete all non-standard residues except FLF (flufenamic acid) and COH (the co-factor, [COX mechanism](https://genomebiology.biomedcentral.com/articles/10.1186/gb-2004-5-9-241/figures/3)). You can proceed e.g. using **Select-Residues**...

1. Correct any potential errors in the structure (missing hydrogens, missing sidechains etc.) using **DockPrep wizard** (**Structure Editing** submenu), which will guide you through all the necessary steps. Skip charges calculation (AutoDock Vina does not use calculated charges).
   * NOTE: In some PDB structures, flexible portions of the protein can be missing, this might interfere with DockPrep and must be repaired beforehand. Chimera offers an interface to Modeller (<https://salilab.org/modeller/>, free academic license) for homology modelling. For more details, check the following video: [plato.cgl.ucsf.edu/chimera/videodoc/Modeller/](http://plato.cgl.ucsf.edu/chimera/videodoc/Modeller/)
   * Save the resulting file as: 5ikv\_DockPrep.pdb
2. With complex prepared, save ligand (residue FLF) and receptor (protein + heme-like co-factor COH) separately as two PDB files:

* select required atoms and in **File-Save PDB…** tick **Save selected atoms only**
* Produce 2 files: 5ikv\_DockPrep\_lig.pdb and 5ikv\_DockPrep\_rcp.pdb

NOTE: As with most other chemical file formats, PDB is human readable and can thus be edited also manually in the text editor to contain only some atoms

1. Close Chimera session and open the previously saved files. This is required to obtain two separate entries for the ligand and receptor for redocking later.

### Task 2 - Redocking

Redocking of the original (co-crystallized) ligand is an important step in docking experiments. It checks the suitability of the used docking algorithm – if the docking can reproduce the co-crystallized pose, it is suitable (for compounds of similar structural type).

1. Open AutoDock Vina toolbar (**Tools-Surface/Binding Analysis-AutoDock Vina**).
2. Define the output filename (e.g. redock). The output of the docking will be saved as \*.pdbqt. PDBQT format is the input and output format specific to AutoDock Vina. It is an extension of a common PDB format.
3. Define the receptor and ligand by choosing the proper entries in the **AutoDock Vina** panel.
4. Define the **Receptor search volume options** as follows – center defines the coordinates of the centre of the box, size is the dimension of the box in Å. The box should cover the whole (supposed) binding site of the protein.

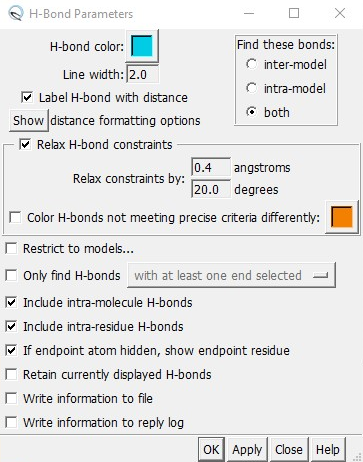
Center: 165.4 185.5 190.5

Size: 27 19 22

1. Leave Receptor options and Ligand options to defaults. In Advanced options change Number of binding modes to 5 (for each ligand, the five best poses by score will be retained and written)
2. Choose vina binary (locate the vina\*.exe file on your system) and click **Apply** (docking will start in the background)
3. Upon successful completion, a **ViewDock** window appears. By clicking on the entries, observe the calculated poses. Notice that docked poses lack non-polar hydrogens due to Vina united scoring function (more info see FAQ section [Autodock Vina 1.2.0 documentation](https://autodock-vina.readthedocs.io/en/latest/faq.html#frequently-asked-questions)).

(if the results are not opened automatically, results can be accessed by opening **Tools-Surface/Binding Analysis-ViewDock** window and choosing **redock.pdbqt** file)

1. Calculate HBonds using **HBonds-Add count to Entire Receptor** menu, use the following settings for **H-Bond Parameters**:



1. Fill the following table with the two best poses obtained:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Pose number | Score | RMSD l.b. | No. of HBonds between ligand and receptor | RMSD to original ligand\* |
|  |  |  |  |  |
|  |  |  |  |  |

\* (optional, see below for instructions)

ADVANCED USAGE TIPS: Agreement of the poses with the co-crystallized ligand upon redocking is judged by the RMSD (root-mean-square deviation). This can be calculated in Chimera followingly:

* To calculate RMSD e.g. between the first docked pose and the original co-crystalized ligand, run this command in the Chimera command line, changing **ID\_pose** and **ID\_ligand** with the correct ID found in the **Model Panel** window:

rmsd #ID\_pose@C=,N=,O=,Cl=,F=,Br=,I= #ID\_ligand@C=,N=,O=,Cl=,F=,Br=,I=

(*e.g.* rmsd #1.1@C=,N=,O=,Cl=,F=,Br=,I= #2@C=,N=,O=,Cl=,F=,Br=,I=)

* Try calculating the RMSD of the best obtained pose and the co-crystallized ligand.

1. Visualize the overlay of the best-docked pose and the original ligand and export it as an image/figure. (if desired, the colors of the molecules can be changed in the **Model Panel**)

### Task 3 – Docking of diclofenac

Similarly to Task 2, dock diclofenac (prepared in the previous seminar) to the receptor. To do so, you have to have the following entries loaded into the Chimera: The prepared receptor (can be reloaded from your file 5ikv\_DockPrep\_rcp.pdb if needed, and the prepared ligand (can be opened from Seminar 1 – mol2 file containing the **3D** (!) structure of diclofenac). Continue with the **AutoDock Vina** panel.

1. Save the docking result as diclofenac\_dock.pdbqt.
2. Report your docking results in the following table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Pose number | Score | RMSD l.b. | RMSD u.b. | No. of HBonds between ligand and receptor |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| 3 |  |  |  |  |
| 4 |  |  |  |  |
| 5 |  |  |  |  |

1. Save the image of the best docked pose (Pose 1) for diclofenac.

## TAKE-HOME-MESSAGE

Docking in AutoDock Vina includes:

* Obtaining 3D coordinates of the receptor (e.g. from RCSB PDB database)
* Correcting errors in receptor structure, adding missing residues or side chains
* Obtaining 3D coordinates of the ligand
* Converting both ligand and receptor to PDBQT format (can be done in batch using specialized tools like ADFR software suite: https://ccsb.scripps.edu/adfr/downloads/)
* Defining binding site search volume
* (Re)Docking
* Analysing the results visually and by score

## SUBMISSION CHECKLIST

As a result of this practical training, you are supposed to submit:

* This document with the filled (two) tables
* Structure files (2): 5ikv\_DockPrep\_lig.pdb; 5ikv\_DockPrep\_rcp.pdb
* Results of the docking (2): redock.pdbqt; diclofenac\_dock.pdbqt
* Two images: 1. Image showing the overlay of the best redocked pose with the original ligand. 2. Image showing the best docked pose of diclofenac.