Outline

- Objectives of practical activities
  - Introductory exercises
  - Project
- Understanding PDB data
- Understanding Bio.PDB basics

Objective

- Solve a Biophysics question using basic scripting
  - Examples.
    - Evaluate influence of amino acids residues in a protein-protein complex interaction energy
    - Perform a in-silico Ala scanning experiment.
    - Evaluate structure quality before simulation

- Objectives Session I.
  - Become familiar with structure data (PDB) and Bio.PDB
  - Evaluate programming skills

Obtaining data

- PDB is main source of data for structural Bioinformatics and BioPhysics

- Selecting 3D structures
  - Contents of the structure
  - Quality

- Most used formats are
  - PDB: Traditional and widely used, but near to be obsolete
  - mmCIF: Official PDB format, complete but very complex, less popular

PDB format

- Structure levels
  - Model > Chain > Residue > Atom
- Identifiers
  - Fixed for amino acids and nucleic acids
  - Less Standard for Hydrogen atoms!
  - Insertion codes, alternative locations, ...
- Residue Numbering
  - Reason of most important headaches in structural bioinformatics !!!
- Coordinates
  - Arbitrary axis
- Occupancy, B-Factor
  - Less used but relevant in some cases
  - Columns often re-used for other data (charges, atom-types, vdw radii, etc.)

Questions 1:

- Identify PDB structures for analyzing the effects of lysine acetylation on binding of p53 to DNA

- Identify PDB structures to analyze conformation changes in adenylate kinase

- Identify PDB structures to analyze binding of SARS-Cov-2 Spike and human angiotensine-converting enzyme.

Questions 2:

- Find structures for thymidine kinase from human herpes viruses
- Which are the main diferences in composition?
- Identify ligands
- Locate missing residues, missing side chains
- Locate atoms with alternative locations (Visualize them in pymol)
- Identify regions for large fluctuations (B-factor)