| | | |
|---|---|---|
| *Piermarco* | *Giustini* | *2056755* |

# Midterm test No. 1

29 / 03 / 2023

Please answer all questions below and submit this document in **PDF format** by **12:30 - 12 April 2022** (two weeks after) to **damiano.piovesan@unipd.it**.

Each student is assigned a different **protein structure (PDB ID)**. The entire exercise is based on the analysis of that structure. Please add your **name**, **surname**, **university ID** and **email** in the **assignment file** here which contains a list of PDB IDs. **Assignment file here**

## Questions

Answer the following questions concisely (**max 500 words in total**).

1. What is the difference between an ionic bond and a covalent bond?

   Ionic bonds occur between metals and nonmetals. They result from the transfer of one or more electrons from the metal atom to the nonmetal atom, creating two ions with opposite charges that attract each other. Covalent bonds, on the other hand, occur between nonmetals. In covalent bonds, two atoms share one or more pairs of electrons, resulting in a stable electron configuration for both atoms.

2. What is the difference between a weak and strong acid?

   Acids are substances that can donate protons (H+) to other molecules. The strength of an acid is determined by its ability to donate H+ ions. A strong acid is one that completely dissociates in water, while a weak acid only partially dissociates. The dissociation of a weak acid is reversible, meaning that some of the H+ ions that were donated to water can be recaptured by the acid molecules. Strong acids have a higher dissociation constant and a lower pKa value than weak acids.

3. What is the isoelectric point of a protein?

   The isoelectric point (pI) of a protein is the pH at which the protein has no net charge. At this pH, the number of positively charged amino acids in the protein is equal to the number of negatively charged amino acids. The pI can be calculated from the amino acid sequence of the protein, and it is an important parameter in protein purification and characterization.

4. What is the “native conformation” of a protein?

   The native conformation of a protein is its biologically active and stable three-dimensional structure. The sequence of amino acids in a protein determines its native conformation, which is essential for its proper function. Proteins can be denatured, or unfolded, by changes in temperature, pH, or other environmental factors, leading to loss of function.

5. What is the difference between the asymmetric unit and the biological entity of a PDB structure?

   The asymmetric unit of a PDB structure is the smallest part of the crystal that can be used to generate the complete structure. The biological entity, on the other hand, refers to the functional

unit of the protein in vivo. The biological entity can include monomers, dimers, or other oligomeric forms, and may differ from the asymmetric unit due to crystal packing or post-translational modifications.

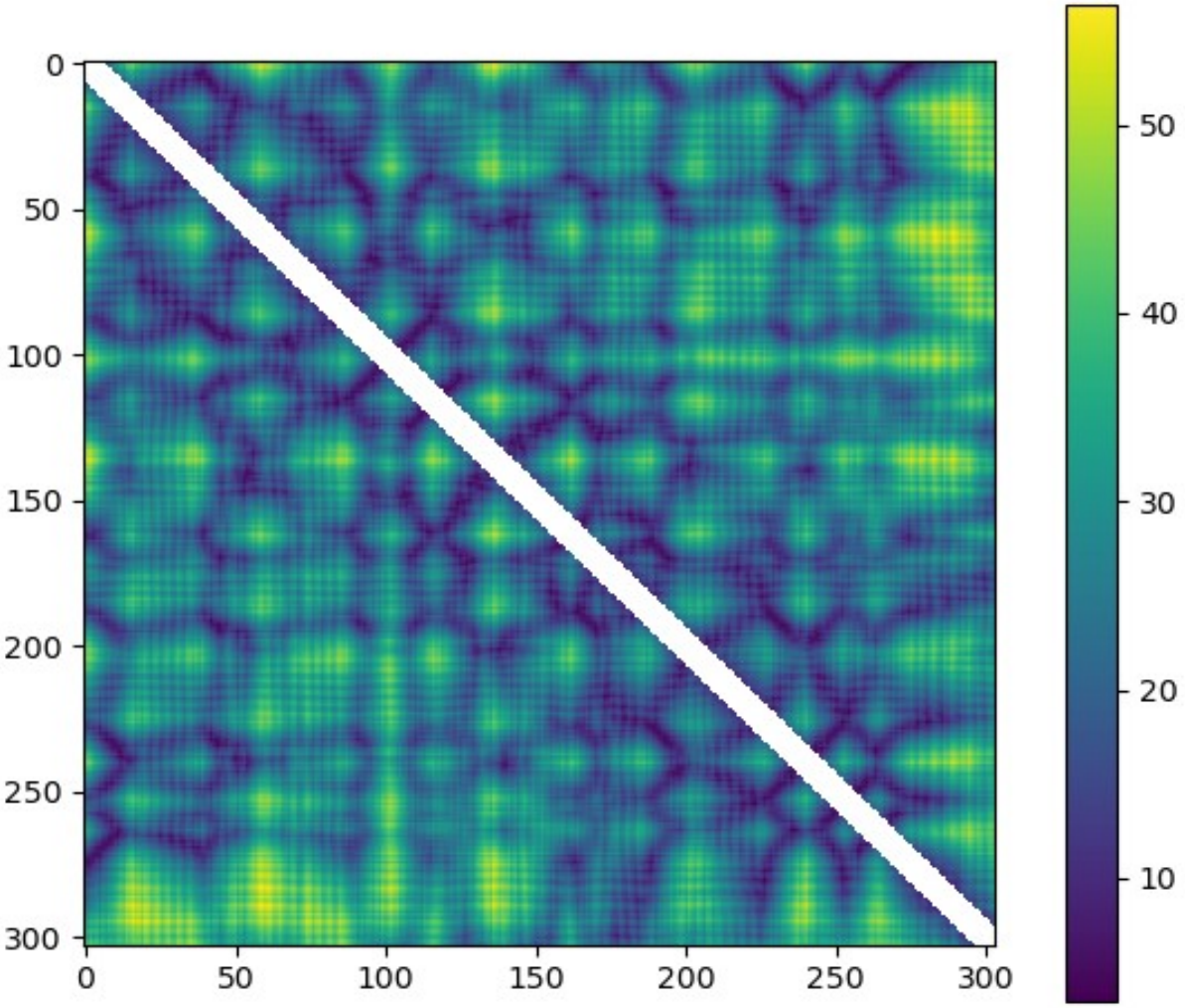
6. What are “missing residues” in a PDB structure?

   Missing residues in a PDB structure refer to amino acids that are not present in the electron density map of the structure. These residues may be missing due to incomplete data or errors in the crystallographic analysis. Missing residues do not necessarily indicate that the protein is incomplete or non-functional, but they may affect its properties or interactions with other molecules.

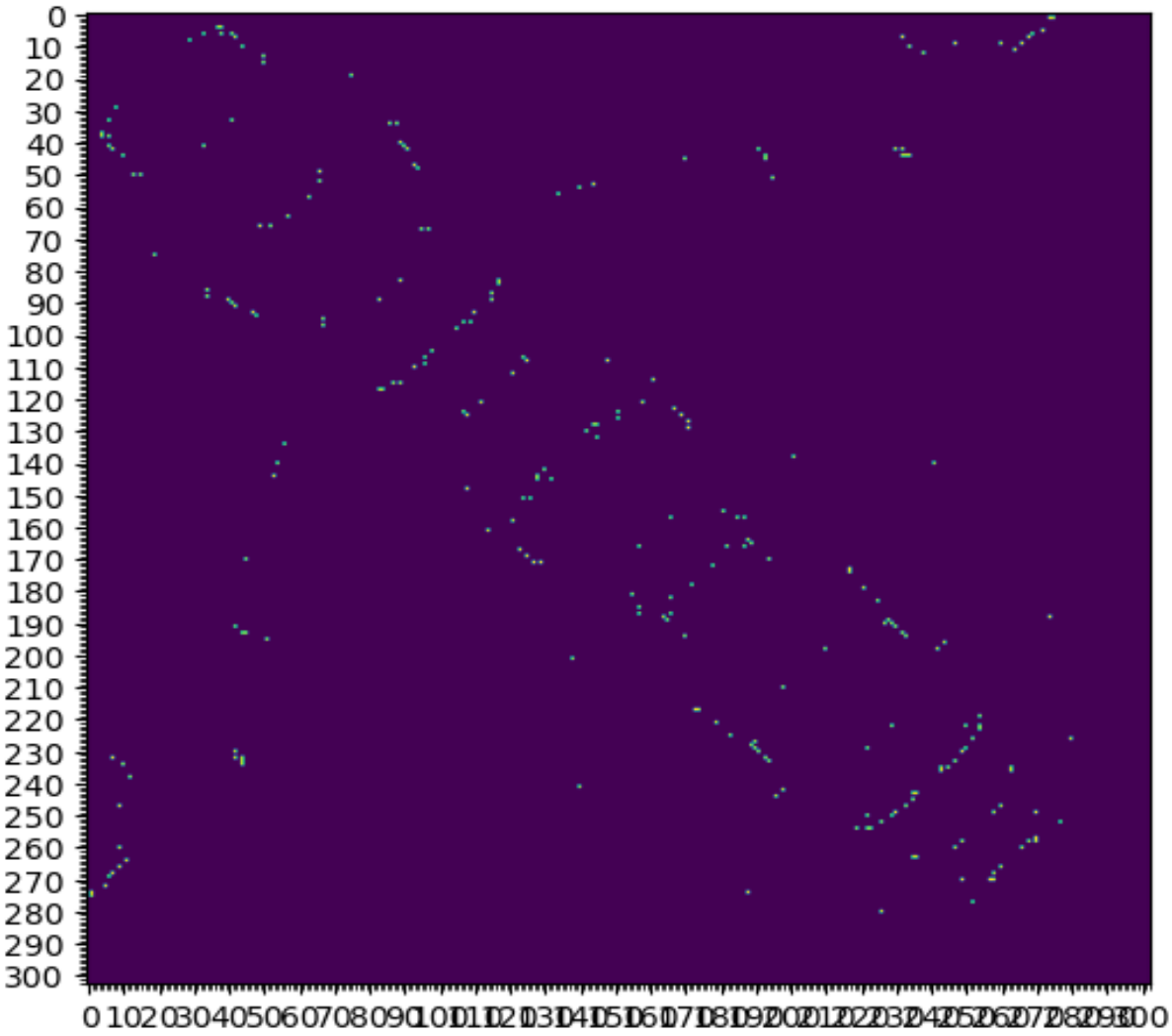
Download the assigned PDB structure and create a new file removing non relevant chains (if it is a complex), solvent molecules and cofactors. For each question **concisely explain all passages (max 5 rows)** necessary to reproduce the results (e.g. parameters, database queries, algorithms, etc.). Optionally, if relevant, you can provide source code (not necessary).
The code is on githhub (https://github.com/Reevoc/Structural-Bioinformatics/tree/plotted)

1. Plot a heatmap representing distances between beta-carbons (CB), consider alpha-carbons when the CB is missing, i.e. for proline. Provide residue indexes along the axes.

   We calculate the heatmap using CB  and if CB is not present we use instead CA also we take in account a separation of at least 6 in absolute value.



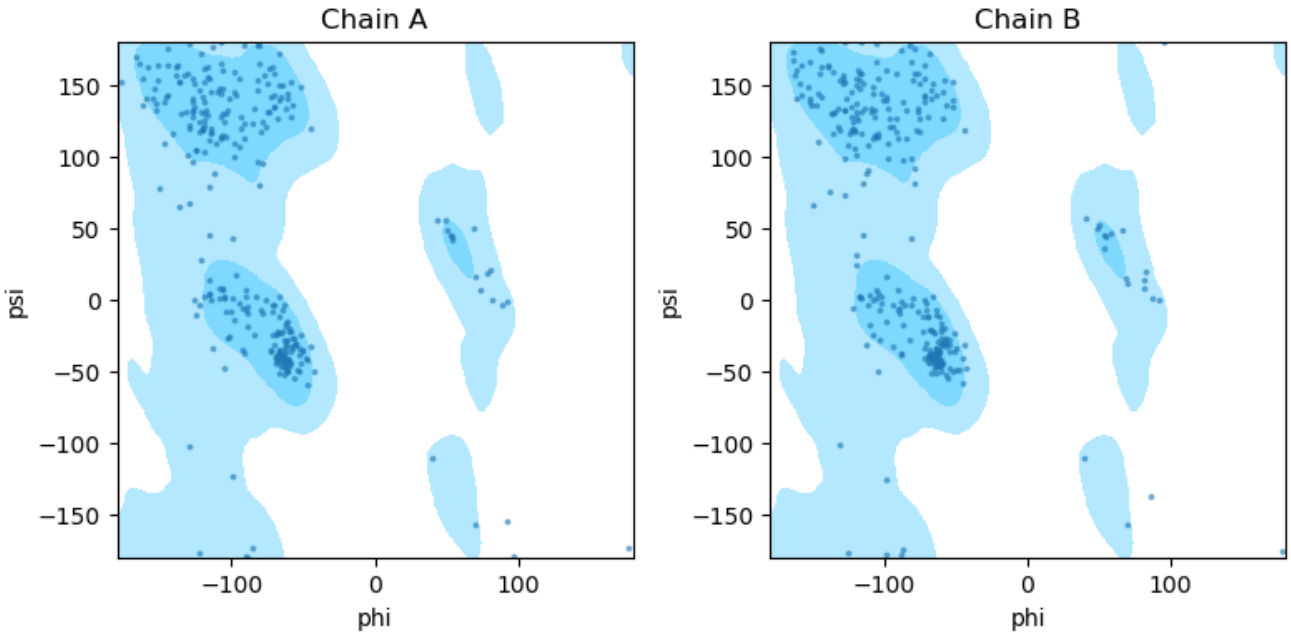
considering a distance threshold of 5 Å.




2. Report the number of residues in contacts for different ranges of sequence separation. Consider the following intervals [0,6], [7,12] and [13,24] and [25,∞].

   In order: [246, 5858, 34062, 42254]

3. Generate the Ramachandran plot of your protein.




4. How many residues are Ramachandran outliers? Consider the Ramachandran regions as shown during the practicals.

   A → 7 outliers

   B → 12 outliers

   Check the region matrix from the ramachandran file provided in stem using the coordinates for the residues (phi, psi); if the value corresponding to those coordinates is 0, it means the residue is outside the permitted region; if it is 1 or 2, it is in the permitted region.