Prof. Dr. Oliver Kohlbacher
Prof. Dr. Manfred Claassen
Anke King, Rosanna Krebs,
Dr. Kyowon Jeong, Hadeer Elhabashy, Dr. Samuel Wein
Faculty of Science
Department of Computer Science and
Institute for Bioinformatics and Medical Informatics (IBMI)

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

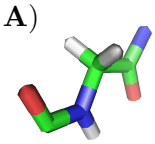
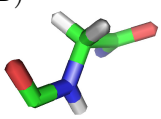
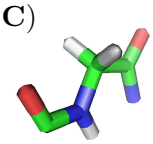
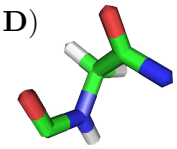
2nd Assignment - Structure and Systems Bioinformatics

Hand in: 2023-05-11 10:00 CEST (source code and pdf in a single archive, uploaded via ILIAS)

Task 1 Torsion Angles (26 P)

The backbone of a polypeptide consists of peptide bonds between amino acids, bonds between the C_α atoms and the amide nitrogen, and bonds between the C_α atoms and the carboxyl groups. The latter two are freely rotatable. The corresponding torsion angles are named Φ and Ψ, respectively.

a) The image below shows four glycine conformations with different Φ/Ψ angles. Associate the given Φ/Ψ angles with the displayed glycine conformations. (**6** P)

A)  **B**)  **C**)  **D**) 

1) 180°/0°
2) 180°/60°
3) 180°/180°
4) 60°/60°

b) Amino acids differ in their flexibility due to the rotatable bonds of their side chains. The torsion angle of the side chains & χ_x, x=1,2,3,4 & 5, can be defined as (atom1-atom2-atom3-atom4) where, the bond between (atom2-atom3) represents the axis of angle and the zero value of torsion angle occurs when atom4 is in cis configuration to atom1. Similarly, define the torsion angles, if exist, for the following amino acids (ALA, PRO, TYR, LYS, and ARG). Please refer to chemical structures and atoms nomenclature found in the lecture slides. (**10** P)

Amino acid	χ_1	χ_2	χ_3	χ_4	χ_5
MET	N-CA-CB-CG	CA-CB-CG-SD	CB-CG-SD-CE	-	-
ALA					
PRO					
TYR					
LYS					
ARG					

c) Similar to the Ramachandran plot for the angles Φ/Ψ of the protein backbone, torsion angles of the side chains also have preferred values. What distinctive conformations does the torsion angle χ_1 of the bond C_α-C_β in glutamate take, from most to least preferred? **Please cite your sources.** (**10** P)

Task 2 Protein Data Bank I (**20 P**)

The Protein Data Bank (PDB) is a resource/database, which contains experimentally validated (NMR, Cryo-EM, X-ray crystallography) structures of proteins. The entries have a unique identifier and are stored in the PDB format, which can be used among others for visualization. (https://www.rcsb.org/) Look up the proteins with the PDB identifier 1IGT and 5IRE. Download and open the PDB files in an appropriate structure viewer of your choice (e.g. PyMOL, BallView, or Chimera(https://www.cgl.ucsf.edu/chimera)). **Please provide short answers of around 1-2 sentences most each.**

1. How many amino acids does each primary structure contain? (**2** P)
2. What is the prevailing secondary structure of each protein? (**2** P)
3. How many chains do the proteins consist of? (**2** P)
4. What is the role of the quaternary structure for the 5IRE protein entry? (**2** P)
5. For both proteins, find three hydrogen bonds and three disulfide bonds. State the residue number of involved residues and give an estimate of the bond lengths. (**6** P)
6. How many disulfide bonds do you find in each protein? Which residues form these bonds? (**4** P)
7. What is the relationship of the disulfide bond type to structure stability? (**2** P)

Task 3 Ramachandran Maps (**40 P**)

Write a Python program that computes the Ramachandran maps from a given set of proteins. Compute the Ramachandran maps of three proteins of your choice (e.g., the two from Task 2 and one more). Please consider the following:

- The program input should be a list of PDB files that you want to calculate the Ramachandran maps on. Please use a command line parser for providing input and output files (like in the Nussinov assignment). Upon running for provided input PDBs, it has to output a pdf file containing the Ramachandran maps.

  The program has to have the following command syntax:

  ```
  <last_name>_ramachandran.py -i <PATH_TO_PDB_FILE_1>
  <PATH_TO_PDB_FILE_2> ... <PATH_TO_PDB_FILE_n>
  -o <PATH_FOR_PDF_OUTPUT>
  ```

- You can use Python libraries to parse the PDBs and get access to the structure elements (e.g., residues, residue atoms, atom coordinates).
- **However, you have to compute the required torsion angles yourself.** You are not allowed to use external libraries/methods for this step.
- For the actual plots, you are free to use any Python plotting library of your choice. We highly recommend Matplotlib, please refer to the documentation of the scatter plot here: https://matplotlib.org/api/_as_gen/matplotlib.pyplot.scatter.html.
- In your hand-in, provide the code of your program, the README.txt, the chosen protein structures as .pdb-files and the resulting PDF containing the Ramachandran plots.

Additionally, shortly describe what you observe in the Ramachandran plots and whether the observations make sense to you in regard to the protein structures.

Task 4 Unusual Structure Motifs (**14 P**)

Polyglycine can form right-handed as well as left-handed α-helices. **For the following questions, please provide short answers of around 1-2 sentences each, and provide reasoning in your answers.**

- What properties make the left-handed conformation feasible, compared to other amino acids? (**4** P)
- What would you expect to happen if a glycine residue in a long polyglycine helix was exchanged to a different proteinogenic amino acid? (**6** P)
- Would D-alanine prefer left- or right-handed helices? (**4** P)

Questions can be directed to ssbi-ss23@informatik.uni-tuebingen.de or the ILIAS course forum. We highly encourage you to use ILIAS for communication.