Zhoutao Zhang
Noel Kubach

UNIVERSITY OF TUBINGEN
SUMMER 2023

SSBI
ASSIGNMENT 03

Task 1 - Protein Data Bank II

1. What experimental method was used to determine the structure?

   1IGT: X-ray diffraction
   5IRE: Cryogenic electron microscopy

2. Explain why the structure of 5IRE does not contain hydrogen atoms while the structure of 1IGT does.

   The resolution of 5IRE model is 3.8 Å, at this resolution only fold class and some secondary structures are visible, hydrogen atoms are not visible. On the contrary, the resolution of 1IGT is 2.8 Å, which is much better and allows the sidechains and backbone to be placed more correctly. While fitting the peptide to the electron density map, people often do refinement steps with the hydrogen atoms, even though they are not visible in the electron density [5]. This can be done using empirically derived bond lengths to the heavy atoms. In the case of 1IGT the authors probably decided to include the refined hydrogens in the PDB file.

3. The protein of 1IGT reports the so-called R value in the experimental section. Please provide a short description of the R value. How would you conclusively judge the quality of this model?

   The R-value indicates the similarity between experimentally-observed diffraction pattern and the simulated diffraction pattern. The smaller R-Value get, the better experiment fit the simulation. Typically, we get a R-value about 0.20[1].
   The R-Value model has a R-Value of 0.209. The quality of the model is typical.

4. The structure of 1IGT does not only contain protein residues, other molecules can be found in this file as well. Describe, in 2-3 sentences, what these molecules are. What type of interaction do these molecules have with the protein?

   The molecules are ligands. They bind to the protein by ionic bonds, hydrogen bonds and Van der Waals forces. In the case of 1IGT the ligands are olygosaccharides, which can not be described as peptide residues in the PDB file.

Task 2 - Investigating secondary structure elements

The vaules have been calculated using the provided python script. Please refer to the README file for usage documentation.

a) How abundant are sheets, right-handed α-helices, and right-handed 310-helices in these proteins?

Secondary structure	Abundance
Right-handed α-helices	0.620
Right-handed 3_{10}-helices	0.068
Sheet	0.312

b) What are the propensities of the different amino acids to form these structures?

Right-handed α-helices		Sheets	
Amino acid	Propensity	Amino acid	Propensity
ALA	0.105	ALA	0.068
ARG	0.055	ARG	0.048
ASN	0.037	ASN	0.039
ASP	0.049	ASP	0.049
CYS	0.012	CYS	0.022
GLU	0.071	GLU	0.050
GLN	0.038	GLN	0.032
GLY	0.065	GLY	0.071
HIS	0.021	HIS	0.022
ILE	0.060	ILE	0.071
LEU	0.108	LEU	0.093
LYS	0.056	LYS	0.053
MET	0.028	MET	0.023
PHE	0.042	PHE	0.049
PRO	0.032	PRO	0.031
SER	0.058	SER	0.060
THR	0.044	THR	0.068
TRP	0.014	TRP	0.014
TYR	0.034	TYR	0.039
VAL	0.071	VAL	0.100

Helices contains significantly more alanine and slightly more glutamic acid. Sheets contains significantly more valine and slight more threonine. The amount of other amino acids in helices and sheets are similar.

c) For all proteins, compute the distances between backbone N atoms and the O atom of backbone C=O groups four residues earlier

The distances between backbone N atoms and the O atom of backbone C=O groups four residues earlier of a collection of 200 proteins was calculated in this task.
As shown in Figure 1, in all proteins most of these distances distributed between 2 and 12 with two peaks around 3 and 10. In only α-helices, these distances distributed almost only between 2 and 4.

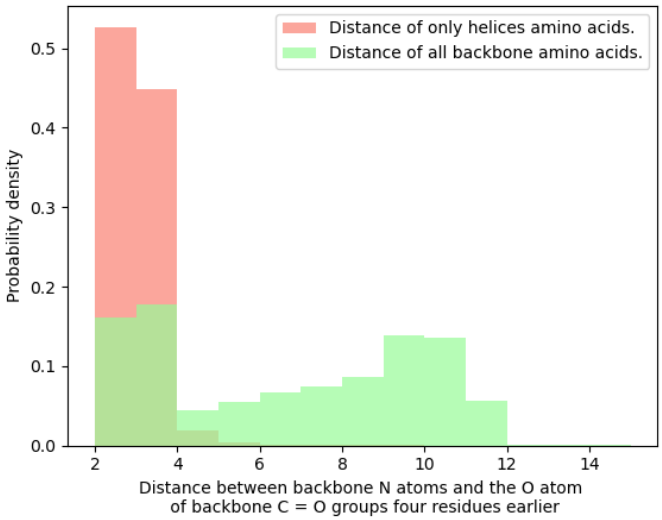
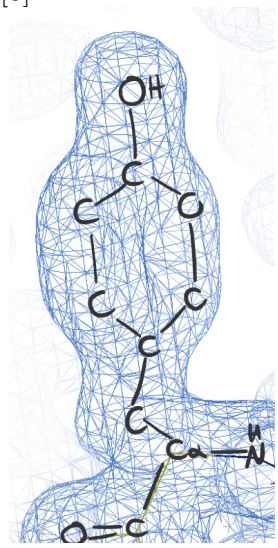


Figure 1: This histogram shows the distribution of distances between backbone N atoms and the O atom of backbone C = O groups four residues earlier. Red bins represent the distribution of distances for only α-Helices. Green bins represent the distribution of distances for all amino acids

Task 3 - Electron Density Map

The given electron density map matches the structure of Tyrosine. One can clearly see the phenol ring, C_α and C_β.
Potential difficulties, by inferring a protein structure from an electron density map are for example wrongly assigned rotamers, defined by the torsion angles Φ and Ψ. and wrongly assigned rotamers. The higher the resolution of the electron density map, the higher is the reliability of the inferred protein structure. Especially unordered and flexible regions in proteins, result in poorly defined regions in the electron density map [8].



Task 4 - Force Fields

The MM3 force field [4] was developed for aliphatic hydrocarbons. While the AMBER force field is described as the sum of the bond stretches, bends, torsion angles, Van der Waals forces and electrostatic interactions, the MM3 force field additionally considers *Stretch-Bend Interactions*, *Torsion-Stretch Interactions* and *Bend-Bend Interactions*.
The AMBER force field is widely used for proteins and different parameter sets have been developed for the use with peptides [2] and proteins [7]. MM3 seems to be suited especially for small organic molecules, since it was developed for this purpose and fitted to the data of these kind of molecules [3][6].

References

[1] Pdb-101: Learn: Guide to understanding pdb data: R-value and r-free. pages https://pdb101.rcsb.org/learn/guide–to–understanding–pdb–data/r–value–and–r–free.

[2] Abil E. Aliev and Denis Courtier-Murias. Experimental verification of force fields for molecular dynamics simulations using gly-pro-gly-gly. *Journal of Physical Chemistry B*, 114:12358–12375, 9 2010.

[3] Norman L. Allinger and Jenn-Huei -H Lii. Benzene, aromatic rings, van der waals molecules, and crystals of aromatic molecules in molecular mechanics (mm3). *Journal of Computational Chemistry*, 8:1146–1153, 1987.

[4] Norman L. Allinger, Young H. Yuh, and Jenn Huei Lii. Molecular mechanics. the mm3 force field for hydrocarbons. 1. *Journal of the American Chemical Society*, 111:8551–8566, 1989.

[5] Richard I. Cooper, Amber L. Thompson, and David J. Watkin. Crystals enhancements: Dealing with hydrogen atoms in refinement. *Journal of Applied Crystallography*, 43:1100–1107, 8 2010.

[6] T L Hill, N L Allinger, M A Miller, F A Van Catledge, J A Hirsch, D H Wertz, O Ermer, and S A Mason. Molecular mechanics. the mm3 force field for hydrocarbons. 3. the van der waals' potentials and crystal data for aliphatic and aromatic hydrocarbons. *ACS Publications*, 111:53, 1989.

[7] Viktor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65:712–725, 11 2006.

[8] Audrey L Lamb, T Joseph Kappock, and Nicholas R Silvaggi. You are lost without a map: Navigating the sea of protein structures. *Biochimica et biophysica acta*, 1854:258–68, 4 2015.