Übung 6: Revision 1


**Fragenblock 1 (Protein Structure):**

- What order of magnitude is a chemical bond (in Å)?
- On the diagram, mark the two backbone angles which can rotate in a normal protein. You only need do this for one residue.

![Protein Structure Diagram](image)

- Mark the angle which is nearly planar (flat).
- Why can I not have a short α-helix which is only 2 residues long?
- Name a large hydrophobic amino acid, a small amino acid and a polar (but uncharged) amino acid.
- Name the amino acid which often forms covalent bonds from its side-chain.
- If you consider a Ramachandran plot for a protein, there is a region where only one type of amino acid is found, marked on the diagram by the grey oval. Which amino acid is this? Why can it occupy this area?

![Ramachandran Plot](image)

- Why can proline not be part of a perfect α-helix?
- For length N, do all / many / few of the 20^N sequences form specific structures?
- You synthesized a 100 amino acid long protein consisting entirely of aspartate: \((\text{Asp})_{100}\)
  - What are its properties? Is it soluble? Is it acidic? Is it basic?
  - Would it form a compact regular structure?
- If you have a protein of poly-Trp (polytryptophan), would it form a specific structure? How would it behave in solution?
- Why would you want to represent a protein by its surface?
- Why might you draw a protein as a ribbon representation in Chimera?
- What is the biggest chain in the protein data bank?
  - What is the average size of a protein?
Fragenblock 2 (Distance Geometry):

- Draw three atoms with distances between them, which are not possible in 3-dimensional space.
- Aside from experimental distance information, what information does one add to a metric matrix distance geometry calculation, before applying the triangle inequality (bound smoothing).
- Why is the triangle inequality applied twice during a metric matrix distance geometry calculation?
- In the metric matrix distance geometry method, one generates a trial matrix. Imagine you have no experimental errors. All your distance measurements are correct to $10^{-20}$ m. Would you expect the trial matrix to correspond to a single set of 3-dimensional coordinates?
- What is the running time of the bound smoothing step in the metric matrix method? Explain in one sentence.
- You use the metric matrix method to calculate the structure of a protein, but you do not have any experimental data. What would you expect if you generate 20 structures?
- In a distance geometry calculation, I have a set of atoms $i-j-k-l-m-n$. What stops atoms $i$ and $n$ ending on top of each other?
- The following matrix contains upper bounds on distances. Draw a graph that corresponds to this distance matrix.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

Add the missing values to the distance matrix and the corresponding edges to your graph. Is there some value in the distance matrix that can be reduced?

- What is the shortest path between points D and E?
- Name an advantage of the variable target function method, compared to the metric matrix method for distance geometry.
- What is the running time of the variable target function method?
Frageblock 3 (NMR):

- How is uncertainty in protein coordinates from NMR represented?
- Name three elements, with the correct nuclei, which are relevant to biochemistry and NMR.
- When calculating a protein structure based on NMR data, what information does one get from the size of a $J$ (spin-spin) coupling constant? Which atoms are involved?
- Why are only some values of the coupling constant useful?
- Which experimental phenomenon provides most of the structural information for determining a structure by NMR?

Frageblock 4 (Crystallography):

- What is the wavelength of an X-ray (order of magnitude / Potenzordnung) ?
- Explain why one cannot simply apply a Fourier transform to the measured reflections to obtain electron density.
- When one refines a crystallographic structure
  - What are the variables?
  - What is the cost function?
- What is the difference between static and dynamic disorder in protein crystals ?
- In protein crystallography, there is a model for the uncertainty of atomic coordinates. What kind of distribution is it based on ?
- This distribution is based on a model for the energy. What is that model ?
- A crystal structure represents a space and time average. Explain.
- An R-factor in protein crystallography is given by
  $$ R = \frac{\sum_{hkl} \left| F_{obs}^{hkl} \right| - \left| F_{calc}^{hkl} \right|}{\sum_{hkl} \left| F_{obs}^{hkl} \right|} $$

  What is $R$ measuring ? When is the formula used ?
- What is the $R$-factor for a perfect model ?
- What is the difference between $R$ and $R_{\text{free}}$ ?
- To determine phases for a crystallographic data set, one might use molecular replacement. Does it require any modification to a protein sample ? What does one need ?
- What does one need for multiple isomorphous replacement ?
- I have a data set with 1.5 Å resolution, but I claim the uncertainty in coordinates is less than $\frac{1}{2}$ Å. Why might this be reasonable ?