Andrew Torda
Björn Hansen

Zentrum für Bioinformatik
Übung zur Vorlesung
Grundlagen der Strukturanalyse
Wintersemester 2016/2017

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

30. Januar 2017

Übung 9: Revision 2

Dies ist die zweite von zwei Übungen, welche Ihnen die Prüfungsvorbereitung erleichtern soll. Auf den folgenden Seiten finden Sie typische Fragen, wie sie in einer Klausur gestellt werden könnten. Dies ist aber kein Fragenkatalog, sondern nur eine kleine Sammlung möglicher Prüfungsfragen. Im Gegensatz zu dieser Übung werden die Prüfungsfragen der Klausur auf Deutsch gestellt.

Bitte beantworten Sie mit Hilfe Ihrer Übungs- und Vorlesungsunterlagen alle gestellten Fragen und bringen Sie Ihre Antworten für die Besprechung in der Übung am **30.01.2017** mit. Sie haben während dieser Übung außerdem die Möglichkeit, über die Fragen dieses Übungszettels hinaus alles zur Sprache zu bringen, was Ihnen in Übung und Vorlesung bisher unklar geblieben ist. Setzen Sie sich deshalb schon vor dieser Übung mit dem GST-Lehrstoff auseinander und bereiten Sie entsprechende Fragen vor. Bei sehr speziellen Fragen Ihrerseits empfiehlt es sich, diese bereits vor der Übung an ***hansen@zbh.uni-hamburg.de*** zu schicken.

Fragenblock 5 (Structure Analysis and Comparison):

- Why is it fundamentally difficult to superimpose two protein structures if they are not the same size?
- I have two models of one protein, but they are rather different.
  Describe an algorithm with pseudo-code to find the more similar regions of the structures.
- I have two proteins and an effective algorithm to find the common region between two protein structures. When I run the program I find the following alignment:

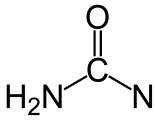
	residues			
protein 1	1-10	11-60		61-90
protein 2		1-50	51-70	71-100

  So, for example, residues 11-60 in protein 1 are aligned to 1 to 50 in protein 2.
  Sketch a diagram of domain structure that would give this alignment.
- You would like to align protein structures of different sizes and you would like to turn the problem into a classic dynamic programming formulation. Describe one method for this.
- Similarity of protein structures is often measured using the root mean square difference of coordinates. Draw an example to show why this may not be a good measure.
- Describe a measure of protein similarity which is quantitative (in Å), but is not the root mean square difference (*rmsd*) of Cartesian coordinates.
  Why may it be better than *rmsd* of Cartesian coordinates.
- Over the course of evolution, which changes faster – protein sequence or structure? Give a reason why this may be the case.
- Some protein structure classifications impose a hierarchy on proteins.
  Why may this be a reasonable thing to do?
- Give an argument why a hierarchical classification may not be appropriate for many proteins.

Fragenblock 6 (Modelling):

- You have built an initial structural model for a sequence. You have a very simple model for the energy of the system. Describe a method to find a reasonable arrangement of side-chains.
- You want to use distance geometry to generate possible conformations of a loop in a protein. You have endpoints for the loops. Describe how you would cast this into a problem suitable for the metric matrix method.

Fragenblock 7 (Protein Stability):

- As a chemist, you would expect *ΔG* to refer to some reaction.
  What is the relevant reaction in protein folding?
- In the reaction above, I imagine there is something called the "unfolded state".
  Why is this a simplification?
- Forget entropy. What would be the balance of energies which make a protein stable?
- I can measure the stability of a protein. I change the pH of the system and the protein becomes more stable.
  Give one example of contributions (e.g. chemical interaction) to the *ΔG* which could explain this.
- I have a small molecule which causes a protein to unfold. According to all evidence, the small molecule does not interact with the native protein. How could the small molecule be causing a change in stability?
- If I say *ΔG* is for the reaction A→ B is 10 kJ mol^{-1}, do I have more A or B at equilibrium? Explain your answer!
- Why might  destabilise a protein.
- I should be able to calculate ΔG from $\Delta G = RT \ln \frac{[folded]}{[unfolded]}$.
  Why is this hard for a solution of native protein?

Fragenblock 8 (Protein Motions):

- Given the motion coordinate of a particle in a harmonic oscillator is $x(t) = A\cos(\omega t + \delta)$ and given that kinetic energy is ½ mv^2, write an expression for the kinetic energy of a harmonic oscillator. Is the energy constant with time? If not, is energy still conserved?
- I consider the motions within a protein, treating them as harmonic oscillators. I claim that most particles in a protein have similar kinetic energy. Consider the expression for kinetic energy. The relationship of kinetic energy, frequency and amplitude is given by

  $$E_{kin} = \frac{1}{2}mv^2 = \frac{1}{2}mA^2\omega^2 \sin^2(\omega t + \delta)$$

  Are the larger amplitude motions associated with the low or high frequencies? Explain.
- In a harmonic oscillator, the force depends on the coordinates x as in

  $$m\frac{d^2x}{dt^2} = -kx\ .$$

  Show that $x(t) = A\cos(\omega t + \delta)$ is a valid solution.
- If I have a two-state system, what does the frequency of the motions mean?
- Why does the frequency of motions increase with increasing temperature in a two-state model?
- Why does the frequency of the motions not increase in a harmonic oscillator model?
- A crystallographer does not usually speak about harmonic oscillators. They normally use a wave equation, $y(x) = F\cos\left(\frac{2\pi}{\lambda}x + \alpha\right)$. How does this correspond to the harmonic oscillator equation $x(t) = A\cos(\omega t + \delta)$? What are the meanings of α and λ?