

Take Home, assigned 10/21/09, due 10/28/09, to be done on your own

Part I. (4 questions, 10 pts. each)

1) The equations that describe the refinement of an X-ray crystal structure or an NMR structure of a protein are shown to the right. In each method of solving protein structures, the refinement incorporates weighting factors (w_{NMR} , w_{Fhkl} and w_{ideal}).

$$\Phi = w_{\text{NMR}} \left(\sum_{\text{violations}} \text{distance restraint} + \sum_{\text{violations}} \text{torsional restraint} \right) + w_{\text{ideal}} E_{\text{total}}$$

$$\Phi = w_{\text{Fhkl}} \sum w'_{\text{hkl}} (|F_o| - |F_c|)_{\text{hkl}}^2 + w_{\text{ideal}} E_{\text{total}}$$

- a. What would be the result to your structural model if the weighting factor- w_{ideal} used in the refinement was **too low**?
- b. During the process of comparative modeling (for ALR, this was done for you by SwissModel), describe how the refinement of the model differs from experimental techniques like NMR and crystallography. Then, also answer a question (similar to part-a): what would be the result to your final comparative model if the weighting factor- w_{ideal} used in the refinement was **too high**?

2) a. What does the “**resolution of a structure**” mean when you are referring to a protein structure solved by X-ray crystallography? Try to be as complete as possible, explaining all the relevant descriptions of resolution. Use what you know about Miller indices, parallel planes, and Bragg’s law to help explain your answer. Also, describe what higher resolution gives you?

b. What does the “**resolution of a structure**” mean when you are referring to a protein structure solved by NMR? Explain differences to crystallography.

3) The expression for the crystallographic R-factor is shown below.

$$\text{R-factor} = \frac{\sum |F_o| - |F_c|}{\sum |F_o|}$$

This equation does not have the phase (α_{hkl}) explicitly shown. However, the R-factor is, without a doubt, the best criteria to evaluate when a crystal structure refinement has converged on the best phases possible. Explain this apparent discrepancy and at the same time demonstrate that you understand how the R-factor is used to guide the completion of a crystal structure of a protein.

4) a. For the crystal structure of IDH (Ceccarelli et al, 2002, J. Biol. Chem. 277, 43454-62), the final converged model had a value for R_{working} of 0.182 and an R_{free} of 0.210. Why is the R_{free} a better indicator of the progress of a structure refinement and the final model? Explain.

b. Explain a way that you could evaluate an NMR protein structure refinement in a way analogous to the use of the crystallographic R_{free} .

Part II. Structural Genomics (4 questions, 15 pts. each)

Suppose you are working for a pharmaceutical company and are given the responsibility of obtaining structural models for 500 specific proteins of company interest. All of these 500 genes have been cloned and are in your lab in a plasmid form. In the questions below, think about coming up with a reasonable model as fast as possible and then following up with better quality models.

- 1) Describe briefly how you would implement homology modeling to obtain a working model for some of the 500 targets. In your description, give a brief description of your selection process, what you would do, the limitations of the method and finally how you could use this approach to give a partial model for some of the more challenging targets.
- 2) For crystallography and NMR you need to have a source of pure and homogeneous protein. Describe how you would setup a high throughput expression and purification of the targets. In your description be sure to answer the following two questions. How would you evaluate the quality of each expressed protein? And how would your approach differ here depending on whether you were going to attempt an NMR or crystal structure solution?
- 3) Describe how you would decide which of the 500 targets you would attempt to solve a crystal structure for. In your description, answer the following questions: What are the experimental limitations of crystallography? What would be your approach to obtain crystals in a high-throughput manner? What would be your approach to obtain phase? How could you use structural information obtained in parts II-1) or II-4) to help with the crystal structure solution?
- 4) Describe how you would decide which of the 500 targets you would attempt to solve an NMR structure for. In your description, answer the following questions: What are the experimental limitations and/or technical challenges of NMR? What types of proteins are particularly good NMR targets due to difficulties from the other two methods? How could you use structural information obtained in parts II-1) or II-3) to help with the NMR structure solution?