

BMB/Bi/Ch 173 – Winter 2018

Homework Set 6.1 – Assigned 2-13-18, due 2-20-18 by 10:30 a.m

TA – Taylor Stevens tstevens@caltech.edu

Office hours – Friday Feb 16 3:00pm - 5:00pm (in SFL 220) and Monday Feb 19 4:00pm - 5:30pm (in SFL 229), or by appointment

1) From diffraction to electron density. (27 points)

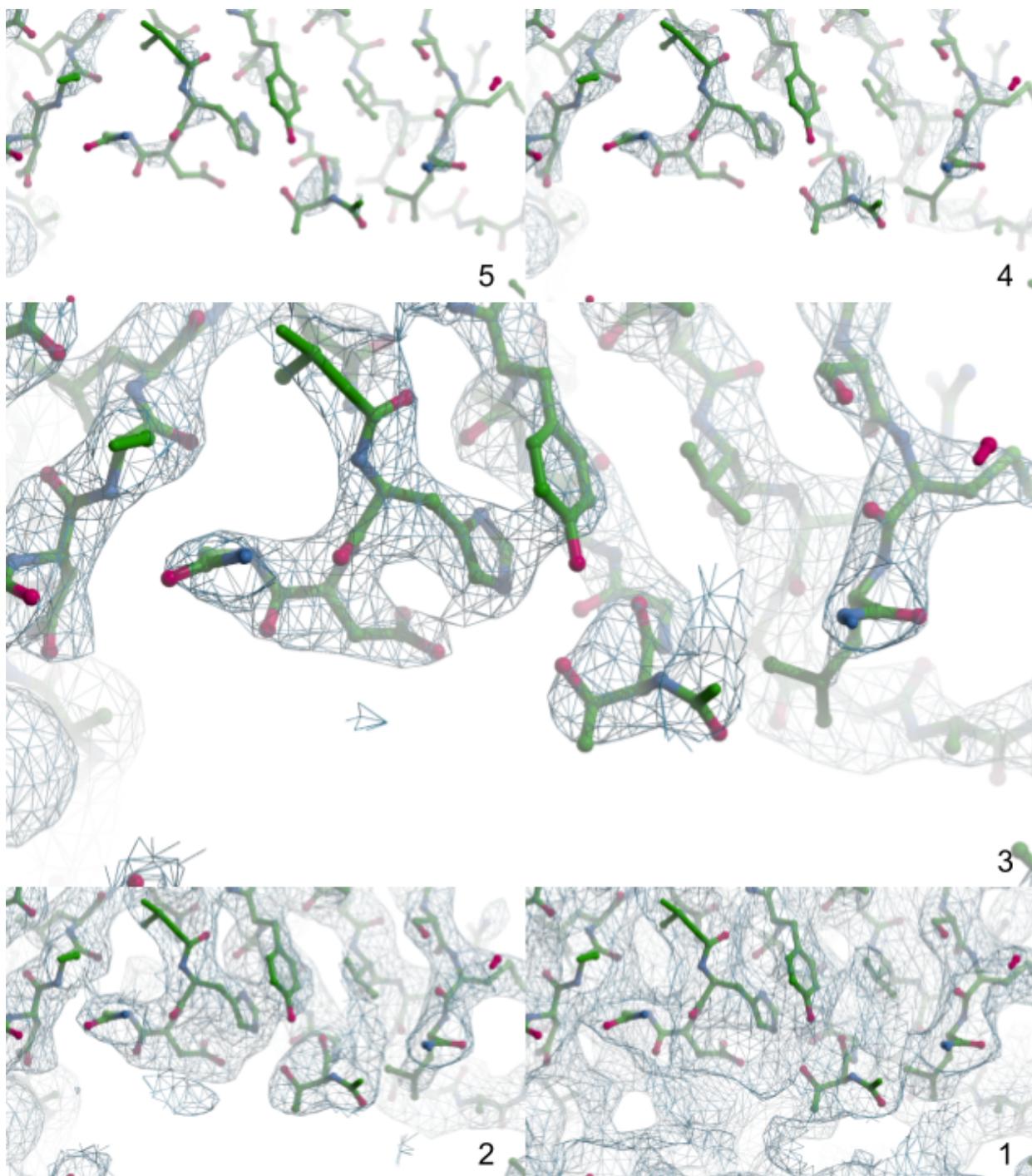
a) A diffraction pattern gives us two important types of information needed to compute the electron density map of the crystal unit cell. Describe both types of information. What information is needed that is missing from the diffraction pattern? (6 points)

- From the spacing of reflections in a diffraction pattern we can derive the dimensions of the crystal unit cell.
- From the intensity of each diffraction spot we can derive the amplitude of each related fourier component.
- We cannot compute the phase of each fourier component from a diffraction pattern.

b) What is the quantity we measure from a single diffraction spot? Describe how it is measured. How does this quantity relate to $\mathbf{F}(\mathbf{hkl})$ (the structure factor)? (9 points)

- The quantity we measure from a single diffraction spot is the intensity of the fourier component ($I_{\mathbf{hkl}}$).
- The intensity is measured by integrating over the entire diffraction spot in each frame where it occurs.
- The square root of the intensity is equal to the magnitude of the structure factor vector.

c) Electron density maps are typically displayed with a surface or mesh representation that forms a shell around the structure. A cutoff value must be specified to generate this representation. To illustrate the effect of the cutoff value, figures of a structure and electron density map with decreasing cutoffs are shown below: (5 is the highest cutoff value, and 1 is the lowest)



What is the benefit of displaying electron density in this way? What would make a particular cutoff value appropriate for displaying electron density? Explain why it's important to avoid using an inappropriate cutoff value. (12 points)

- Displaying electron density in this way outlines regions of high electron density, which tells us where atoms are likely to be located in our map.

- An appropriate cutoff value will outline as many atoms in the structure as possible without outlining 'noisy' regions which contain no atoms in reality
- If the cutoff value is too high you will miss details while trying to build your structure into the density. If the cutoff value is too low you may be led to build structural features into the map that don't exist in reality.

2) Heavy metals: it's just a phase. (73 points)

a) MAD and MIR are two different experimental phasing methods, and the M stands for multi- in each case. What are there multiples of in MAD and in MIR? (4 points)

- In MAD the data is collected at multiple different wavelengths
- In MIR the data is collected from multiple different crystals with different derivatives

b) List two advantages of MAD phasing methods over MIR phasing methods. List one advantage of MIR phasing methods over MAD phasing methods: (6 points)

MAD:

- MAD only requires a single crystal while MIR requires multiple derivative soaked crystals
- MAD with selenomethionine can validate the solution to a low resolution structure by telling you the specific location of the selenomethionine residues within your map
- MAD generally produces more accurate phases than MIR

MIR:

- MIR can be done at a home source while MAD requires access to a synchrotron

c) Caveats: (12 points)

i) What is the major caveat of isomorphous replacement related phasing methods?

Structures must be isomorphous, which means that the incorporation of heavy metals into the structure must not perturb the structure, which is not always the case.

ii) What is one potential caveat and one limitation of molecular replacement?

One potential caveat is introduction of bias from the search model. One limitation is that you need a search model.

iii) It is common to use protein that has incorporated multiple selenomethionines in place of methionine residues for phasing via anomalous diffraction. What feature must a protein have in order be labelled with selenomethionine? If a protein does not naturally contain this feature, we must engineer it into the protein somehow. Explain how this could be a problem.

The protein must contain methionine residues (not counting in disordered loops). If a wild type protein does not have methionine residues we must mutate some other residues to methionine and this may disrupt the proteins stability.

d) To demonstrate the process of determining phases via MIR we will go over a simplified example in 2D. Let's assume we have already used Patterson methods to find the location of three mercury atoms in a unit cell (in fractional coordinates*): (0,0), (0.3,0.7), (0.4, 0.1). Calculate the amplitude and phase (in radians) of $F_{\text{Hg}}(5\ 6)$, the structure factor of the mercury atoms alone. For this reflection, use a value of 70 for the atomic scattering magnitude of mercury (f_j). You can use the equation below to calculate scattering from j number of atoms:

$$F(hk) = \sum_j f_j e^{2\pi i(hx_j + ky_j)}$$

where f_j is the atomic scattering magnitude (70), h and k are the Miller indices (5,6) and x and y are the positions of each atom. To express $F_{\text{Hg}}(5\ 6)$ as a single complex number, use Euler's formula:

$$e^{ix} = \cos(x) + i\sin(x)$$

To solve for the amplitude and phase, plot $F_{\text{Hg}}(5\ 6)$ on the complex plane. The phase of $F(hk)$ is the angle between the x-axis and the vector, and the amplitude of $F(hk)$ is the magnitude of the vector. (24 points)

*Fractional coordinates describe the positioning of atoms in terms of the unit cell dimensions, e.g. an atom that is directly in the center of the unit cell will have the fractional coordinates (0.5, 0.5, 0.5)

$$F(5\ 6) = 70 \cos[2\pi(5 \times 0 + 6 \times 0)] + 70 \cos[2\pi(5 \times 0.3 + 6 \times 0.7)] + 70 \cos[2\pi(5 \times 0.4 + 6 \times 0.1)] \\ + 70i \sin[2\pi(5 \times 0 + 6 \times 0)] + 70i \sin[2\pi(5 \times 0.3 + 6 \times 0.7)] + 70i \sin[2\pi(5 \times 0.4 + 6 \times 0.1)]$$

$$F(5\ 6) = -8.26 - 107.7i \text{ (12 points)}$$

$$|F(5\ 6)|^2 = (-8.26)^2 + (107.7)^2$$

$$|F(5\ 6)| = 108 \text{ (6 points)}$$

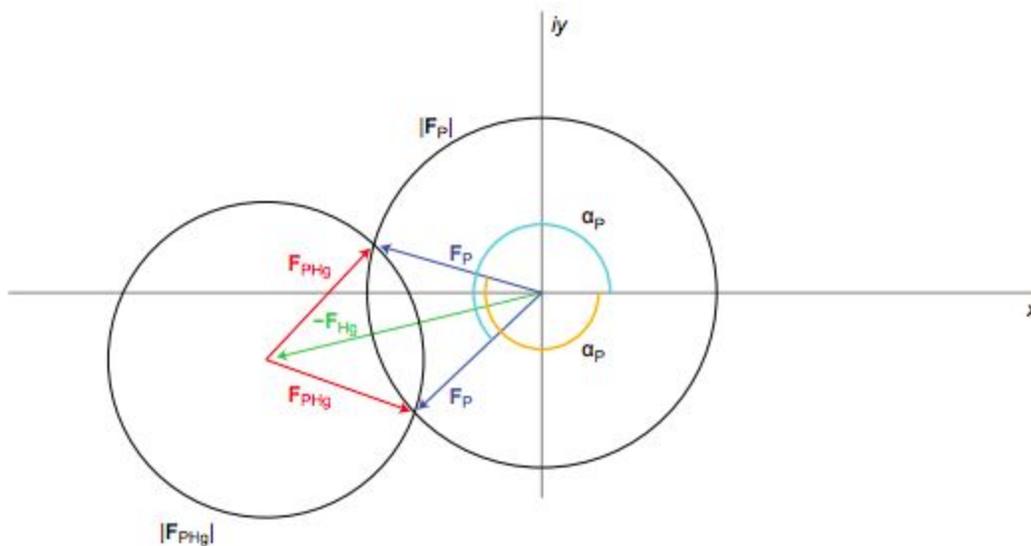
$$\alpha(5\ 6) = \tan^{-1}(107.7/8.26) + \pi = 4.64 \text{ radians (6 points)}$$

e) The amplitude of the native structure factor at $(h,k) = (5,6)$ ($|F_{\rho}(5,6)|$) is 66.57, and the amplitude of the mercury derivative structure factor ($|F_{\rho\text{Hg}}(5,6)|$) is 60. Draw a Harker construct to demonstrate the two possible phase angles for the $F_{\rho}(5,6)$. Calculate the two possible angles using the following equation, where α is the phase angle in radians:

$$\alpha = \alpha_{Hg} \pm \cos^{-1} \left[\frac{(|F_{PHg}|^2 - |F_P|^2 - |F_{Hg}|^2)}{(2|F_P||F_{Hg}|)} \right]$$

Your Harker construct should illustrate F_P , the structure factor of the native protein, F_{Hg} , the structure factor of the mercury atoms, and F_{PHg} , the structure factor of the mercury derivative protein. (24 points)

To draw the Harker construct we first draw a circle with a radius of $|F_P|$ around the complex plane. This circle represents possible points where the vector of F_P will end depending on its phase. Next we draw the vector of $-F_{Hg}$. Starting from the end of this vector, we draw another circle with a radius of $|F_{PHg}|$. Since we are starting this off at $-F_{Hg}$ and $F_{PHg} - F_{Hg} = F_P$, this circle represents possible points where the vector of F_P will end depending on its phase. Since the two circles we have drawn should only overlap at two points, this allows us to narrow down the phase of F_P to two possible values. (12 points)



$$\alpha_P = 4.64 \pm \cos^{-1} \left[\frac{(60^2 - 66.57^2 - 108^2)}{(2 \times 66.57 \times 108)} \right]$$

$$\alpha_P = 4.64 \pm 2.62 = 7.26 \text{ or } 2.02 \text{ (12 points)}$$

f) What additional experiment can you do to determine which of these two phases is correct? (3 points)

A second heavy atom derivative can be created. This will create a third circle on the Harker construct, and the correct phase will be represented by the intersection of three circles. The second heavy atom derivative must use an element other than mercury.

g) (5 points extra credit) Explain how the above process can be applied to the whole diffraction dataset to generate an initial electron density map.

For each diffraction spot measured you must perform the above process with two derivatives to get the phase. Then you combine all the phases and amplitudes you have measured with fourier synthesis to generate an electron density map.