# Protein Crystallography

BBMB 334 Lab 4

Whitman College
Program in Biochemistry, Biophysics & Molecular Biology (BBMB)
Biophysics Laboratory
Prof. Douglas Juers

## Part II. X-ray Data Collection

**Purpose:** To collect X-ray diffraction data from your crystals and to understand what the data mean.

**Background:**

Fourier Series

Any periodic function defined on $-L \le x \le L$ can be represented as a sum of sines (or cosines):

$$f(x) = \sum_{n=1}^{\infty} f_n(x) = \tfrac{F_0}{2} + \sum_{n=1}^{\infty} F_n \sin(\tfrac{n\pi x}{L} + \delta_n)$$

where:

$F_n$ is the amplitude
$\delta_\mathrm{n}$ is the phase
$\frac{\mathrm{n}}{2\mathrm{L}}$ is the frequency
$n$ is an integer

With crystallography *f(x)* corresponds to the electron density, or locations of electrons in your material of interest. If you have a crystal of your material, you can use the three dimensional version of the above equation to calculate the electron density in three dimensions, *f(x,y,z):*

$$f(x,y,z) = \sum_{h,k,l}^{\infty} F_{h,k,l} \sin(\tfrac{h\pi x}{L_x} + \tfrac{k\pi y}{L_y} + \tfrac{k\pi z}{L_z} + \delta_{h,k,l})$$

To do the calculation (in 1 dimension), you need to know $L$, and $F_n$ & $\delta_n$, for each term, $n$, in the series. $L$, $F_n$ and $n$ can be determined (relatively) easily from the spots on the diffraction pattern:

> $L$ is related to the separation between the spots.
> $n$ is related to the distance of the spot from the origin
> $F_n$ is related to the intensity of the spot

The phase, $\delta_n$, is more difficult to determine.

Lab Ticket.

1. Draw graphs of the following function for $n=1$ and $n=2$: $f(x) = F_n \sin(\frac{n\pi x}{L} + \delta_n)$. Illustrate how each parameter ($F_n, L$ and $\delta_n$) affects the graph.
2. Draw a graph of $f(x) = \sum_{n=1}^{2} f_n(x) = \frac{F_0}{2} + \sum_{n=1}^{2} F_n \sin(\frac{n\pi x}{L} + \delta_n)$ with the following parameters:

$$F_0 = 10; F_1 = 1; F_2 = 0.5$$
$$L = 2 \text{ Å}$$
$$\delta_1 = \pi; \delta_2 = \frac{\pi}{2}$$

(Please don't type this into a graphing program. Graph each term individually and then think about how they would add together, without using a computer or calculator.)

**Activities**

1. X-ray data collection. See your instructor for details.
   a. We have found that orthorhombic and trigonal trypsin crystals can be effectively cryoprotected by using Paratone N oil (other oils will likely work, as well), and then flash cooled in the cryostream. We try to remove all aqueous solution surround the crystal by dragging the crystal through the oil.
2. Fourier Synthesis exercises – I. Go to: http://phet.colorado.edu/en/simulation/fourier
   a. Click on the "Discrete" Tab and play around with adding sine waves to give a sum.
   b. Switch to the "Wave Game" . Start with level 1 and see how high you can go.

3. Fourier Synthesis exercises II. Download the Excel spreadsheet "Fourier Analysis" from CLEo.
   a. Whale Reconstruction. In one window, open the worksheet "b. Whale Reconstruction". In the other window, open the worksheet "b. Graphs". The worksheet calculates and graphs a Fourier sum of 15 terms (in red on the top graph). The first four rows list $n$, $F_n$, $\delta_n$, and whether that term is included in the sum (1=included, 0=not included). In yellow on the upper graph is the target structure – a whale. Your goal is to adjust $F_n$ and $\delta_n$ to create the whale. The lower graph shows the first few terms in your Fourier series.
      i. Turn on the $n$=0 and $n$=1 terms by setting cells B4 & C4 to 1. Set the amplitude of the $n$=0 term to 0.18. Then make your best approximation of the whale by adjusting $F_1$ and $\delta_1$.
      ii. Turn on the $n$=2 term, and improve your approximation of the whale by adjusting $F_2$ and $\delta_2$.
      iii. Repeat for each term in succession. Print out your best reconstruction of the whale.
   b. Amplitude vs Phase. In the two windows: "b. Amplitude vs Phase" and "b. Graphs". The worksheet plots the whale from part a., a torpedo, and the Fourier sum. Near the bottom of the worksheet are $F_n$ and $\delta_n$ for the whale, the torpedo, and "default" ($F_n = 1$, $\delta_n = 0$ for all $n$).
      i. Copy the $F_n$ and $\delta_n$ rows for the whale to the top $F_n$ and $\delta_n$ rows to create a whale reconstruction.
      ii. Repeat for the torpedo.
      iii. Use the $F_n$ for the whale with the $\delta_n$ for the torpedo and vice versa.
      iv. Use the $F_n$ for the "default" with the $\delta_n$ for the whale, and vice versa.
      v. Repeat iv for the torpedo.
      vi. What can you say about the relative importance of Amplitude vs phase for the Fourier series?
   c. Simulation of Crystallography. In crystallography, a common situation is that you have measured amplitudes, $\{F_{n,obs}\}$, and you have a pretty good idea of what your molecule looks like. So you can calculate what you expect the amplitudes to be, $\{F_{n,calc}\}$ and determine the "R-factor" which is like an error calculation between $F_{n,obs}$ and $F_{n,calc}$:

$$R = \frac{\sum_n \left| F_{n,obs} - F_{c,obs} \right|}{\sum_n F_{n,obs}}$$

Add a calculation of the R-factor for your whale to your spreadsheet, for $F_{obs}$, use the $F_n$ for the whale from part b. For $F_{calc}$, use your estimates from part a.