Protein Structure Laboratory
Primer on Protein Structure Determination by X-ray Crystallography

Protein crystallography is both an art and a science. Investigators planning their initial venture into this area should be aware of the necessary steps involved in obtaining an X-ray crystal structure. This process can be separated into four categories: Protein Production, Crystallization, Data Collection and Structure Solution. The KU COBRE Protein Structure Laboratory (PSL) is here to assist with any or all of the last three tasks. Consultation and advice are also readily available at no charge.

Contact Information: Dr. Scott Lovell, Director, at 785-864-3772 or swlovell@ku.edu

Protein Production

Although this is not carried out in the PSL, the director has many years of experience in various protein purification methods and works closely with the COBRE Core B (Protein Production Group, http://psf.cobre.ku.edu/ProteinPurif.htm). Clients may also carry out the expression/purification of protein constructs in their own laboratories and deliver samples to the PSL for crystallization experiments. General guidelines for protein samples to be submitted for crystallization experiments are listed below.

- Obtain the protein of interest (at least 1mg), preferably > 95% purity whenever possible

- Concentrate protein for crystallization
  - Typically 8-12 mg/mL is a good starting point
  - Can be carried out in the PSL or client's laboratory

- Provide the following information
  - Physical properties of construct (any information that may be useful for PSL staff)
  - Exact sequence of the expressed construct that will be used for crystallization experiments
  - Biochemical data (ie. exists in a multimeric state, disulfides, etc.)
  - Protein storage buffer

Crystallization

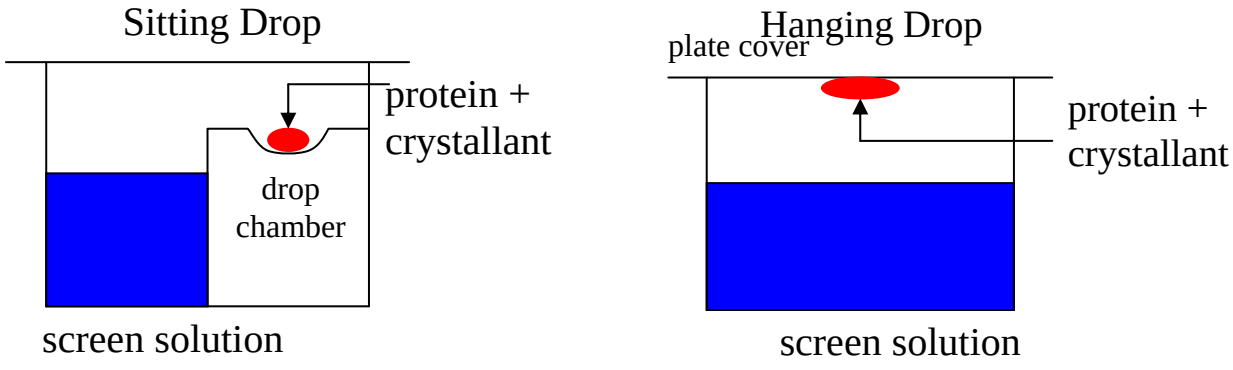
Various methods can be employed for protein crystallization. A protein sample will be initially tested for crystallization using various commercially available sparse-matrix screens or using known conditions if applicable. Even if the crystallization conditions are known for a protein, it is often worthwhile to setup a sparse-matrix screen as new crystallization conditions or crystal forms may be identified. We typically use 96-well sitting drop vapor diffusion plates for screening. These high throughput plates allow one to screen a larger number of conditions quickly using less protein compared to other methods. A 50μL sample of concentrated protein is the minimum amount needed to screen against 96 conditions. Crystallization experiments can be conducted by the PSL staff. However, training will be provided by the PSL Director to the collaborator's personnel if desired.

- Initial screening of a protein for crystallization
  - Sitting or hanging drop vapor diffusion (**Figure 1**)
  - Various crystallization screens
  - Incubate at 20°C and 4°C

- Refinement screening (after initial crystals are obtained)
  - Improve crystals (if necessary) by slight modification of initial conditions
  - Seeding techniques to increase crystal size and/or growth rate
  - Microbatch techniques

- Addition of substrates, ligands, inhibitors etc. to the protein
  - Can often facilitate crystallization of proteins that do not crystallize readily in the apo form

Figure 1. Setup for sitting drop and hanging drop vapor diffusion experiments. Crystallization screen solution is dispensed into the reservoir (blue) of a 96-well crystallization plate. Protein is then added to the drop chamber (sitting drop) or plate cover (hanging drop). An equal volume of the screen solution is then added to the protein and the plate is sealed and incubated at the desired temperature. Crystallization screens are supplied in a 96-well block format and the setup is conducted with multichannel and repeater pipettes. A 96-well screen can be setup in ~5-10 minutes.

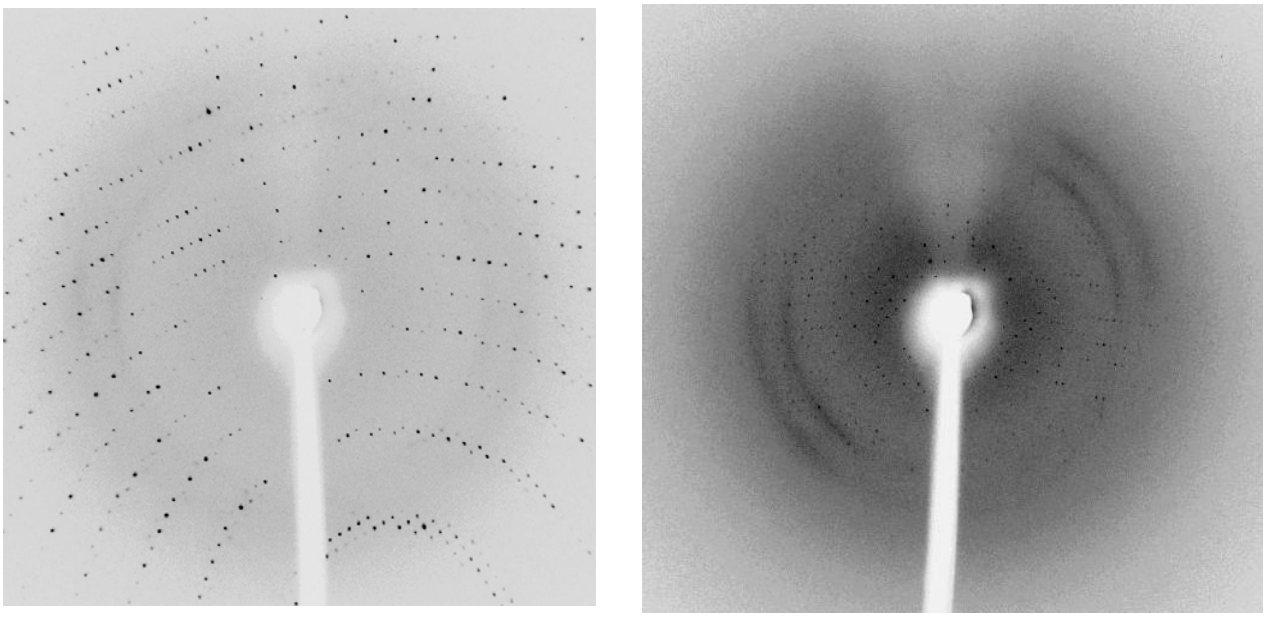


Data Collection

Once crystals have been obtained, they will be tested for diffraction using our in-house X-ray diffractometer (Rigaku RU-H3R generator and R-Axis IV++ image plate detector). Typically, crystals are transferred to a solution containing the components from the original crystallization drop supplemented with a small molecule cryoprotectant (e.g., glycerol, ethylene glycol, PEG 400, etc.). The sample is then mounted on a goniometer in a stream of nitrogen cooled to -180^{o}C. If the cryoprotectant solution is suitable, it will freeze in the stream as a transparent glass. The crystal can then be visualized and centered in the X-ray beam path. One benefit of cryoprotection is that it serves to minimize crystal degradation by the incident X-ray beam. A series of test exposures will be measured to determine the quality of the X-ray diffraction pattern. If the crystal is judged to diffract well, a complete data set will be collected for subsequent structure solution. However, certain cryoprotectants can adversely effect the diffraction of a particular crystal. In such a case, the freezing conditions will need to be optimized. This is carried out by testing various cryoprotectants using different concentrations, transfer methods, equilibration times, etc. Additionally, samples can be mounted directly from their native drop in a capillary and their diffraction examined at room temperature. This allows one to determine if weak diffraction, for example, results from an unsuitable cryoprotectant or a poorly formed crystal for a particular sample. Complete data sets can be collected at room temperature for crystals that do not suffer from extensive radiation damage. Occasionally, samples will also be frozen and stored for data collection at a synchrotron. The PSL Director will consult with clients regarding synchrotron data collection. Data collection training is offered to the client's research personnel upon completion of radiation safety training.

- Identify crystals from screens and test for diffraction in-house
    - Room temperature mount in a capillary
    - Cryoprotection of samples
        - Optimize the freezing conditions for the sample if necessary
- Collect complete data set(s) for the sample
    - A complete data set typically consists of 45 to 360 images as shown in **Figure 2**
    - In-house
    - Synchrotron

Figure 2. Single diffraction images. Examples of strong (left) and weak (right) diffraction of cryoprotected crystals.



Structure Solution

Diffraction intensities measured on multiple images that comprise a complete data set must be integrated and scaled to obtain a unique data set that will be used for structure solution. This step is known as data processing. The processed data is then used for structure solution using various methods. The goal of structure solution is to obtain an electron density map that is used to fit the atomic positions of a structural (phasing) model. For cases where the exact structure or related protein structures are known, solution by molecular replacement can be attempted. With the increasing number of structures deposited to the Protein Databank, the possibility of finding a solution by molecular replacement has increased. In this method, the coordinates of the known structure (search model) are used to estimate the target molecule's position within the crystal lattice. This is carried out by applying a series of rotations and translations such that the diffraction data, which can be calculated from the search model, is in close agreement with the experimental diffraction data (**Figure 3**). For *de novo* structure determination, various methods can be used to solve the structure. Heavy atom derivatives of a protein can be prepared and tested for diffraction in-house. In the case of well diffracting derivative crystals, two heavy atom derivatives and one native data set would be collected and used to solve the structure by multiple isomorphous replacement (MIR). Additionally, heavy atom derivatives, selenomethionine labeled proteins or other derivatives such as ligands that contain a heavy atom can be prepared for multiwavelength anomalous dispersion (MAD) or single wavelength anomalous dispersion (SAD) data collection. For these experiments, the X-ray wavelength is set near the absorption edge of the particular heavy for the data collection. Therefore, it is necessary in most cases to collect diffraction data for MAD and SAD experiments at a synchrotron.

One of the main functions of the PSL Director will be to perform this sophisticated task, which may require considerable time, effort and expertise. Upon request, the Director will teach students who would like to learn crystallographic methods how to process diffraction data and solve crystal structures.

Figure 3. Molecular replacement method. The correct molecular orientation with respect to the crystal lattice is indicated by the target molecule position (X_1). A search model (grey) at position X_2 undergoes a series of rotations ([R]) and finally a translation (T) to position the molecule in the correct orientation.