BC530 2017 X-ray Crystallography Unit Homework

Question 1: Crystallization

You are trying to crystallize a protein that comes out of solution at low salt + high PEG concentrations. [PEG = polyethyleneglycol, a common crystallizing agent]. In initial trials you are able to produce only tiny clusters of needles. These are too small to use for data collection, so you decide to crush them and use them as seeds for crystal growth in new trials ("microseeding").

a) Sketch a possible phase diagram for the solubility of this protein, with PEG concentration along the X axis and protein concentration on the Y axis.

b) Draw a trajectory on your phase diagram that shows how you hope large crystals will grow starting from the fragments used for micro-seeding.

Question 2: Resolution and Bragg's Law

Diffraction patterns

Unlike laboratory X-ray sources, synchrotron radiation source can usually be tuned to select whatever X-ray wavelength is needed. You freeze and mount a protein crystal in the beam of a synchrotron radiation source. Then you take two diffraction pictures without moving the crystal. The first picture uses X-rays with 1.5Å wavelength; the second picture uses X-rays with wavelength 1.0Å.

a) Describe (or sketch) the expected difference in the two diffraction images.

b) You collect two data sets, one using $\lambda = 1.5\mathring{A}$ and one using $\lambda = 1.0\mathring{A}$, and refine the structural model separately for each data set. Do you expect the resolution of the two refined models to be the same or different?

c) You mount another crystal from the same crystallization conditions, but this time before freezing it you soak it for several hours in a solution containing a known inhibitor of this protein. Assuming that the inhibitor does bind to the protein, and the crystal survives, in what way to you expect diffraction images from this second crystal to differ from those taken from the first, unsoaked crystal? Why?

Question 3: Evaluating structures in the PDB

Every Tuesday the PDB (Protein Data Bank) highlights new structures released that week. Last week's collection included 4 structures of the same protein bound to various inhibitors [*Discovery of N-substituted 7-azaindoles as PIM1 kinase inhibitors*, Barberis *et al.* (2017) Bioorg. Med. Chem. Lett. 27: 4730-4734]. Note that I have not read this paper and don't expect you to either, although you are certainly welcome to do so. The PDB codes are

5TEL: https://www.rcsb.org/pdb/explore/materialsAndMethods.do?structureId=5TEL

5TEX: https://www.rcsb.org/pdb/explore/materialsAndMethods.do?structureId=5TEX

5TOE: https://www.rcsb.org/pdb/explore/materialsAndMethods.do?structureId=5TOE

5TUR: https://www.rcsb.org/pdb/explore/materialsAndMethods.do?structureId=5TUR

a) Based on the summary information on the PDB site, would you expect all 4 structures to be of equal quality and thus equally informative about inhibitor binding modes? Do the various quality scores in the validation reports bear this out?

b) Can you point to any particular weak points or red flags in the validation reports?

Question 4: Another PDB structure

Here's another one released last week by the PDB:

5CVZ: https://www.rcsb.org/pdb/explore/materialsAndMethods.do?structureId=5CVZ

a) Would you judge the overall quality good or bad? Why?

b) Can you point to anything unusual in the validation scores (not necessarily good or bad, just atypical)? Can you explain why it might be that way?

c) This structure has a huge unit cell (226Å x 226Å x 226Å) but there is only one chain of 141 residues in the structural model. Why? Does this influence your judgment of model quality?