

How to go MAD at CHESS

Marian Szebenyi

So, you are a macromolecular crystallographer and have crystals of an interesting protein. You know you can get a set of native diffraction data, but phasing is a problem. You can't do molecular replacement because this is a brand new protein, and you haven't been able to find any usable isomorphous heavy atom derivatives. However, you can make a non-isomorphous derivative, or you know you can get cloned protein incorporating selenomethionine - looks like a case for MAD (multiple wavelength anomalous diffraction).

First question: is MAD phasing feasible for this system? In order for MAD phasing to be successful, the signal from the anomalous scatterers must be sufficiently large relative to the total scattering from anomalous plus non-anomalous scatterers. The anomalous signal from each heavy atom depends on what that atom is, and you must also consider the available x-ray flux at the wavelength of the appropriate absorption edge. At the CHESS F2 station, with the standard Si (111) monochromator crystals, the range of possible energies is 7 - 25 keV (1.8 - 0.5 Å). This allows use of the K edges of elements 26 - 46 (including Fe, Cu, Se, Br, etc.) or the L-III edges of elements 64 and up (including Hg, Pt, U, some lanthanides, etc.). The total signal depends on the number of (well-ordered) anomalous scatterers as well as on their nature. As an example, consider a 100 residue protein containing 2 Se atoms. The total number of electrons is approximately 5200. Of this, the two selenium atoms contribute 68. The single-wavelength anomalous (Bijvoet) signal depends on the imaginary component of the scattering factor, f'' , while the multiple-wavelength (dispersive) signal depends on the difference in scattering factors at the two wavelengths, $f'(1)-f'(2)$. These factors vary with the environment of the Se at-

oms, but typical values are at most 3.7 and 6.6 electrons, respectively [1], assuming suitable selection of wavelengths. The resulting ratio of anomalous signal to total structure factor may then be calculated according to the method of Hendrickson [2], and is at most 4.0% for the Bijvoet and 3.5% for the dispersive case. Table 1 gives the maximum number of residues per heavy atom for a reasonable signal (i.e. 3% Bijvoet differences and 2% dispersive differences) from some representative anomalous scatterers, using f' and f'' values determined in protein crystals. Your crystals may have a larger f'' , and you may be able to get useful information from a smaller signal, if you are very careful in collecting data, but this table is a good place to start in deciding whether a MAD experiment is likely to succeed.

Because the MAD signal is so small, it is especially important to minimize errors in intensity measurements. This means: use your best crystals, freeze them, and be careful in selecting data collection parameters. It is worth spending considerable time optimizing crystallization and freezing conditions (before coming to CHESS, of course). During data collection, select each crystal carefully and optimize exposure time, oscillation range, etc. for it - see article on "Efficient Collection of Oscillation Data" starting on page 48.

Before going to CHESS. Having decided to attempt a MAD experiment at CHESS, follow the usual procedures to request beam time. As with any experiment involving heavy atoms, you must include a "Heavy Atom Compounds Declaration" describing the type and amount of heavy atom compounds that

you will be bringing. The small amount of heavy atoms present in the crystals themselves or in a few milliliters of millimolar soak solutions pose no hazard and no administrative complications beyond supplying the Declaration. If you expect to bring concentrated stock solutions of toxic compounds (not recommended but occasionally necessary), you will also need to complete a "Hazardous Materials Declaration". See the heavy atom guidelines included in the standard CHESS proposal package for more detail.

You will need a sample containing your heavy atom for taking reference energy scans (see below). Check with CHESS first to see if there is an appropriate sample on hand; if not, make one up and bring it along. A simple and satisfactory method is to sandwich a pinch of powdered material (e.g. mercuric acetate for a mercury sample) between two pieces of x-ray transparent sticky tape (Kapton tape is good), and mount the sandwich in a slide mount. The area covered by the powder should be at least a centimeter or so square, so that it is easy to be sure that the x-ray beam is passing through it. If you are bringing a reference sample containing toxic material, such as a mercury compound, be sure to include it on the Heavy Atom Compounds and Hazardous Materials Declarations. If you think you might need to make up another sample while at CHESS and are bringing a bottle of material to do so, include that also. When you are at CHESS, handle any reference samples you bring with care - label them, try not to spill them (know how to clean them up safely if you do spill them), and don't bring any more than necessary in the first place.

Energy scans are essential to select

Table 1. Maximum Amino Acid Residues per Heavy Atom for Good MAD Signal

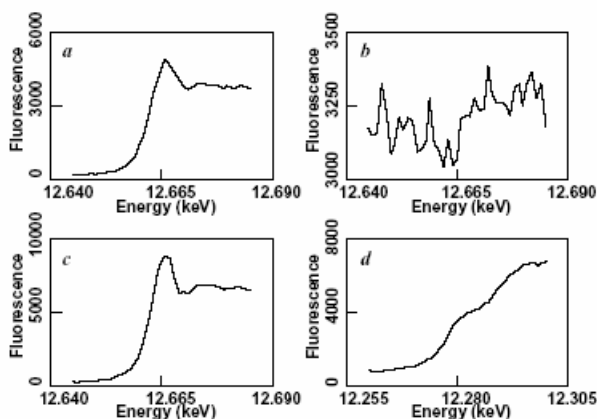
Heavy atom	Absorption edge	Maximum f''	Maximum $\Delta f'$	Max. no. residues per heavy atom
Fe	1.744 Å (K)	5.2	8.0	175
Se	0.980 Å (K)	3.7	6.6	90
Br	0.920 Å (K)	3.4	5.5	75
Hg	1.009 Å (L-III)	10.2	13.3	640
U	0.719 Å (L-III)	11.4	15.4	840

(Table 1) Maximum number of amino acid residues per heavy atom for which the predicted MAD signal will be at least 3% for Bijvoet differences and 2% for dispersive differences, for various anomalous scatterers. The f' and $\Delta f'$ values were taken from References 3 (Fe), 1 (Se), 2 (Br), 4 (Hg), and 5 (U). The maximum number of residues per anomalous scatterer was calculated using a transformation of the equation given by Hendrickson for calculating anomalous signal [2].

(Figure 1) Energy scans taken on the F2 station at CHESS, Jan-Feb 1995. The scan width was 40 eV. The units on the "Fluorescence" axis are arbitrary. (a) Se reference scan. (b) Se crystal scan, crystal too small and/or fluorescence detector badly positioned. (c) good Se crystal scan. (d) Hg reference scan.

the energies (wavelengths) at which to collect data and to check for drifting of the monochromator during the experiment. These are taken by scanning the monochromator through the appropriate energy range and monitoring the x-ray fluorescence from the sample. CHESS staff will set the station monochromator at the absorption edge of the element you are using; scans may then be easily performed in the vicinity of this energy. Reference samples containing high concentrations of commonly used elements are provided (see above for how to make your own if you are using something unusual). The reference sample is mounted in the front of the hutch; the "refscan" command moves it into the beam, scans the energy over a range of about 40 eV (specified using the "madsetup" command), lists and plots the resulting fluorescence spectrum, moves the sample out of the beam, and resets the energy to its initial value. Reference scans are run periodically during the experiment to check, and if necessary recalibrate, the monochromator. In most cases, you will also want to take an energy scan on your crystals, and this is where things can get sticky if you don't plan ahead.

For the crystal scan, you need a large volume of crystal(s). In order to get the best signal-to-noise ratio, the Bicon fluorescence detector should be placed at 90 degrees to both the beam and the spindle axis, i.e. pointing vertically downwards. If you are freezing crystals and prefer to have the cold stream coming in from above, or if there is other equipment in the way, the Bicon may be moved off the vertical, but should stay in the plane defined by the spindle axis and the vertical and should be as close to the vertical as possible. It is not necessary to use frozen crystals for crystal scans, as the environments of the anomalously scattering atoms change very little on freezing. Here is a good place to use those junky crystals that are not suitable for diffraction - pack a bunch of them into a capillary and do a crystal scan with them. Once you have a sample in the beam, the "xtal_scan" command runs the scan and



outputs the results. Panels a - c of figure 1 show scans across the Se edge. In panel a is a reference scan taken on a foil. The scan in panel b was taken using a medium-sized crystal with the detector poorly positioned; the panel c scan used a large crystal sample with the detector correctly placed. Panel d of the figure shows a reference scan for Hg. It is typical of this element that the L-III edge is rather broad, with no "white line". Experimenters have also found empirically that there is very little variation in the edge from one sample to another, so this is a case where you may get away without a crystal scan at all, if you are sure there is mercury in the crystal and are pressed for time.

Energy selection. Select your desired wavelengths from the crystal scan. For a 4-energy experiment, use (1) a point slightly below the edge, (2) the inflection point of the edge, (3) the top of the "white line" peak, and (4) a point about 200 eV above the edge. For a typical 3-energy experiment, omit energy (1). For some elements, such as Hg, it may be better to omit energy (3) instead. In case of severe time limitations, useful data may be obtained using just two energies, e.g. (2) and (3) or (1) and (4). Look at the f' and f'' curves for your element (see Ref. 4 for how to calculate these curves from an x-ray absorption spectrum) to decide what selection of energies will give you the best signal in case you can't collect all the data you would like to. It is generally preferable to have complete data at fewer wavelengths rather than

partial data at more wavelengths.

Planning data collection proper.

Geometry. Once you have selected the wavelengths for data collection, get a good crystal mounted and determine the appropriate oscillation range and exposure time. Typical values are 1 degree and 1 minute, but these depend heavily on your particular crystal. Decide whether you will be using "mirror" or "inverse beam" geometry. In "mirror" geometry, anomalous pairs appear on each image, whereas in the "inverse beam" case anomalous mates appear on pairs of images collected with spindle angles differing by 180 degrees. Mirror geometry requires the crystal to be oriented with a mirror plane perpendicular to the spindle axis. At CHESS, the only means for getting this orientation is the use of the arcs on the goniometer head, so that it may not be possible to get a mirror where you would like it. Installation of a kappa axis goniostat will remove this difficulty in the future. There can also be a problem in scaling images from a precisely oriented crystal together, particularly in the case of a monoclinic crystal of low mosaicity. Nonetheless, mirror geometry is clearly preferable in terms of reducing the number of exposures and insuring good scaling between anomalous pairs, and should generally be used when possible, especially when another data set is available for scaling purposes.

CCD detector. MAD users at CHESS usually use a CCD detector. Considerations in your planning due to the detec-

tor include:

- You will need to take a background image every time you change the exposure time. This doesn't take long, but is important.
- The readout time, including overhead, is about 30 seconds for the 1K x 1K detector and about 15 seconds for the 2k x 2k detector; you must include this time when calculating how many exposures you take per hour.
- You need to allow time periodically to move images off the controlling PC. This cannot be done during data collection.
- The detector area is limited. To get high resolution data, it may be necessary to offset the detector, so that the direct beam position is no longer in the center of the detector. For mirror geometry, a vertical offset, i.e. perpendicular to the spindle, will still allow recording of anomalous pairs on each image; a horizontal offset will eliminate some pairs. In the inverse beam case, corresponding reflections appear in opposite quadrants

CONTINUED FROM PAGE 55

is also necessary to keep an eye on the operation of the spindle motor and the shutter, and of course to halt data collection if the fill ends prematurely. Reference scans should be taken at least once per fill, to check for any monochromator drifting. Keep up with getting data off the PC, corrected, and backed up onto tape. If you have enough people, you can be processing data on the Alpha computers while collection is in progress, just as for monochromatic data.

Summary of what you need to do MAD successfully

- A suitable molecular system with

on the paired images, and any offset will eliminate some of them.

Timing. Data collection proper consists of a series of energy changes, oscillation exposures, and crystal rotations. A sequence for a 3-wavelength experiment using inverse beam geometry, for example, might be:

- go to wavelength 1, take 5 degrees of data, rotate crystal -5 degrees,
 - go to wavelength 2, take 5 degrees of data, rotate crystal -5 degrees,
 - go to wavelength 3, take 5 degrees of data, rotate crystal (180 - 5 = 175) degrees,
 - go to wavelength 1, take 5 degrees of data, rotate crystal -5 degrees,
 - go to wavelength 2, take 5 degrees of data, rotate crystal -5 degrees,
 - go to wavelength 3, take 5 degrees of data, rotate crystal -180 degrees.
- Repeat until data set is complete, crystal dies, or fill ends.

Changing wavelengths takes about 30 seconds. The spindle motor takes about 2 minutes to rotate 180 degrees. Using this

information plus the exposure time and detector readout time, it is straightforward to calculate how much data you can collect in a one hour fill. Figuring about 18 fills per day, will you have time to get your data? If not, consider reducing the number of wavelengths or the resolution, or plan to come back another time.

Monitoring data collection.

You need to monitor the data collection - sorry, it's not safe to go out to dinner and let the computer do the work. It is particularly important to keep track of the oscillation angles and wavelength for each image. The CCD just names its data files in sequence - it has no knowledge of whether the images are good or bad, and it doesn't know what the oscillation range is for an image. It is possible to define macros that will give different names to images taken at different wavelengths, but you still need to verify that image names are correctly matched with wavelength and oscillation parameters. It

CONTINUED ON PAGE 56

enough expected signal for phasing.

- Good freezable crystals, and a large sample for a crystal scan.
- Beam time at CHESS.
- Well-thought out strategy for data collection, with contingency plans.
- Careful experimental technique, with due consideration for safety.
- Good bookkeeping skills.
- A sense of humor to get you through the inevitable equipment failures.

1. Yang W., Hendrickson, W. A., Crouch, R. J., and Satow, Y. "Structure of RNase H Phased at 2 Å Resolution by MAD Analysis of the Selenomethionyl Protein", *Science* **249**, 1368-1405 (1990).

2. Hendrickson, W. A. "Determination of Macromolecular Structures from Anomalous Diffraction of Synchrotron Radiation", *Science* **254**, 51-58 (1990).
3. Hendrickson, W. A., Smith, J. L., Phizackerly, R. P., and Merritt, E. A. "Crystallographic Structure Analysis of Lamprey Hemoglobin from Anomalous Dispersion of Synchrotron Radiation", *Proteins: Structure, Function, Genetics* **4**, 77-98 (1988).
4. Tesmer, J. J. G., Stemmler, T. L., Penner-Hahn, J. E., Davison, V. J., and Smith, J. L. "Preliminary X-ray Analysis of *Escherichia coli* GMP Synthetase: Determination of Anomalous Scattering Factors for a CysteinyI Mercury Derivative", *Proteins: Structure, Function, Genetics* **18**, 394-403 (1994).
5. Glover, I. D., Denny, R. C., and Tame, J. R. H. "Structure Determination of OypA at 2.3 Å Resolution Using Multiple-Wavelength Dispersion Methods", *Acta Crystallographica D - Biological Crystallography* **51**, 39-47 (1995).