

Weak anomalous sample - brief processing notes for iMosflm

Before starting, note the following.

It is assumed that all students will have gone through the online tutorial for iMosflm before even thinking about processing these data! The tutorial is available at

<http://www.mrc-lmb.cam.ac.uk/harry/imosflm/ver105>

(select "documentation" from the left hand menu, and the tutorial specifically for iMosflm, not the "Mosflm User Guide")

The students have been requested to download two SAD datasets collected from a single crystal. Each dataset consists of 360 images, each of 1° oscillation.

The data extend to around 2.5Å, so there should be sufficient to refine the unit cell with post refinement.

The route to process these data is therefore -

```
read images, mask the circular backstop
index & estimate mosaicity
integrate ~5° of data & run Quicksymm
refine cell
integrate
quickscale
```

In more detail

Dataset 1

Read in the 360 images using the normal iMosflm method. Note that the images are of the type "<template>####", which is not automatically recognised by iMosflm's image browser - choose "Files of type: Numbered files (*. [0-9]+)" to see the list.

Go to the indexing pane - unless you have unchecked the options to automatically find spots and automatically index & estimate mosaic spread, the indexing will proceed and you will get a list of "good" solutions; iMosflm will indicate the highest symmetry solution with a low penalty as its "best guess" choice -

Lat.	Pen.	a	b	c	α	β	γ	$\sigma(x,y)$	$\sigma(\phi)$	$\delta(\text{beam})$
aP	0	40.9	53.2	53.3	90.3	90.0	90.1	0.14	0.41	0.50 (0.1)
.										
.										
oP	3	40.8	53.3	53.4	90.0	90.0	90.0	0.20	0.41	0.50 (0.1)
tP	4	53.3	53.3	40.8	90.0	90.0	90.0	0.20	0.41	0.50 (0.1)
mC	151	67.2	67.1	53.3	90.0	90.3	90.0	-	-	-

For really good data, if the $\sigma(x,y)$ for the chosen solution is more than about 0.02mm more than for the triclinic basis solution, the symmetry may be too high. For "okay" data, this criterion can be relaxed, and as long as $\sigma(x,y)$ is no more than about the pixel size bigger than for the basis solution, it is probably okay. In this case I would proceed with the tetragonal solution until evidence indicated otherwise.

Check the predictions on the image display - look at the first image and at least one $\sim 90^\circ$ away - some pathologies are not visible in some crystal orientations.

Mask out the backstop, using the "fit backstop from circle" tool.

Select the "Integration task". All 360 images should be selected automatically; select the first five, either by typing in the selection bar or using the "select images" tool. Integrate the first five images and run "Quicksymm" to check on the symmetry. A strong indication that the symmetry is correct would be if the scores are all close to unity. If the figures are not convincing, integrate a few more images.

Note that there is significant "averaging" in the high resolution profile fitting regions (red spot borders and red boxes). Click and hold the left mouse button down on these boxes to view the unaveraged profiles; there is almost no spot-shaped intensity in most of the boxes. This is a good indication that the resolution limit has been set too high, so consider reducing it to $\sim 2.5\text{\AA}$.

Go to the "Cell refinement" task, and click on "Process". Two segments of data based on the current symmetry and mosaic spread will have been chosen automatically, and when the refinement has converged compare the figures in the bottom panel. The final cell should be close to the initial cell, but, more importantly, the standard deviations should be small.

Select the "Integration task" again. The selection of images chosen before will have been retained, so click on the "select all images" button to include all 360 images in the integration. Click the "Process" button and consider getting a cup of coffee - integration will take a few minutes. The integration should proceed smoothly and there won't be any sharp deviations in any of the plots (except for the RMS residual plots which drop from initial good values to substantially better values, and the $I/\sigma(I)$ plots change between blocks of images).

Select the "Quickscale" task - this launches a Pointless and Scala job - when finished, it will open a web-browser so that you can see the overall statistics. These have been calculated quickly using a small subset of Scala's options, so that you can make sure the images have been processed more-or-less okay.

Dataset 2

Repeat what you have done for dataset 1; the major difference is that the data extend to slightly higher resolution - maybe 0.05\AA better.

Scaling

When processing data with a weak anomalous signal, it is more important to get the scaling right than for most other datasets; the differences that we are looking for between the Bijvoet pairs are often around the same magnitude as the noise level in our experiment, and sometimes smaller. With care in both data collection and processing we can still extract a useful signal.

Also, note that using summation integration intensities seems to give better results than profile-fitted; this is an option in Scala (Mosflm automatically writes both to the MTZ file).

For this, I would run Scala from a command file so that I can easily change parameters. For these data try the following command file, but be prepared to spend time reading and understanding the Scala documentation. I use tcsh or csh, but the modern trend is to use bash or zsh.

```
#!/bin/csh -f
```

```

set run = "ACA10_AF1382_1"

# use Pointless instead of sortmtz

pointless HKLIN ${run}.mtz HKLOUT pointless_${run}.mtz > pointless_${run}.log

scala HKLIN pointless_${run}.mtz \
  HKLOUT scala_${run}_unmerged.mtz \
  SCALES ${run}.scales \
  ROGUES ${run}.rogues \
  ROGUEPLOT ${run}.rogueplot \
  CORREL PLOT ${run}.correlplot \
  ANOMPLOT ${run}.anom \
  NORMPLOT ${run}.norm \
  SCALEPACK ${run}.sca \
  > scala_${run}.log <<EOF
bins 20
run 1 batch 1 to 360
name project AUTOMATIC crystal DEFAULT dataset ${run}
scales rotation spacing 5.000000 secondary 6 bfactor on
cycles 100

# put in values after first run of Scala to see if things improve
# this is right(ish) for the first crystal, wrong values for the second!
#
sdcorrect run 1 full 1.601      7.70   0.0190 part  2.065      6.04   0.0148

# Use "anomalous OFF" for weak signal so we get fewer outliers rejected
#
anomalous off

# write out .sca file for reading into SHELXC - George prefers unmerged data
#
output polish unmerged

# use integrated rather than profile fitted intensities
#
intensities integrated

# may want to use profile fitted weak reflections, summation integration
# strong reflections instead of all integrated or all profile fitted -
# intensities combine

EOF

```

In the above, if you want to merge the two datasets together, change the Pointless run line to -

```
pointless HKLIN ACA10_AF1382_*.mtz HKLOUT pointless_1-2.mtz > pointless_both.log
```

and for the Scala commands, add after the "run 1" line -

```
run 2 batch 1001 to 1360
```

and a second sdcorrect line

```
sdcorrect run 2 full <value> <value> <value> part <value> <value> <value>
```