Hg-SAD example

This example uses data collected in-house on a copper anode to a resolution of 2.1 Å. The protein was soaked for 20 minutes in 1mM $Hg(NO_3)_4$. 80 degrees of data were collected. The space group is cubic, P213. The protein contains 438 amino acids. The structure was actually solved by SIRAS, but this tutorial shows that the Hg derivative alone contains sufficient anomalous phase information to solve the structure.

In this tutorial you will use SHELXC to prepare and analyze the data, SHELXD to find the mercury(s) and SHELXE to carry out density modification to break the phase ambiguity and to give phases and figures of merit to allow a map to be calculated.

P213. a=b=c=125.35 Å. Estimated solvent content 64% with one monomer in the a.u.

Data file: **hg.sca** was created from the SCALA output using mtz2sca.

Statistics from SCALA:

```
=======================================================================
Summary data for Project: H12 Crystal: PSE_Hg2_20mi Dataset: H12

                                          Overall  InnerShell OuterShell

  Low resolution limit                      39.65     39.65      2.21
  High resolution limit                      2.10      6.64      2.10

  Rmerge                                     .042      .025      .246
  Rmeas (within I+/I-)                       .054      .02:      .337
  Rmeas (all I+ & I-)                        .058      .038      .339
  Rpim (within I+/I-)                        .033      .017      .228
  Rpim (all I+ & I-)                         .027      .017      .216
  Fractional partial bias                   -.021     -.019     -.014
  Total number of observations             146358      5976      6849
  Total number unique                       35442      1298      3617
  Mean((I)/sd(I))                            21.6      55.6       2.9
  Completeness                               92.1      98.7      65.5
  Multiplicity                                4.1       4.6       1.9

  Anomalous completeness                     82.4      91.:      32.3
  Anomalous multiplicity                      2.2       2.6       1.2
  DelAnom correlation between half-sets      .226      .533     -.009
  Mid-Slope of Anom Normal Probability      1.121

=======================================================================
```

You could use hkl2map to run SHELXC/D/E and visualise the statistics, or you could create a script using an editor (e.g. call it **pse**, enter the lines below, save it, type **chmod ugo+x pse**, and then enter **./pse** to run it).. The key numbers are NTRY, the number of trials, and FIND, which should be within about 20% of the actual number of anomalous scatterers:

```
shelxc pse  << EOF
SAD hg.sca
CELL 125.35 125.35 125.35 90 90 90
SPAG P213
FIND 1
SFAC Hg
NTRY 100
EOF
```

This will create pse_fa.ins (input for SHELXD), pse_fa.hkl ("heavy atom" structure factors h,k,l,FA, sigFA, α) and pse.hkl (high resolution structure factors, h,k,l,F,sigF)

To run SHELXD, type

shelxd pse_fa

this tries to find the anomalous scatterers. The key numbers are CC All/Weak. These should be greater than 30/15 for a meaningful solution for SAD. This produces pse_fa.res (heavy atom coordinates in shelx format), pse_fa.pdb (heavy atom coordinates in pdb format) and pse_fa.lst (detailed output from SHELXD). Note how the occupancies of the heavy atoms vary.

You can now use SHELXE to use these heavy atom positions to calculate phases, with density modification being used to break the phase ambiguity. You need to try both hands of the heavy atoms. The –h switch indicates that the heavy atoms are in the native data (i.e. pse.hkl), the –b switch creates a file of the refined heavy atom positions, the –s0.64 indicates a 64% solvent content, -m20 indicates 20 cycles of density modification and –i tells shelxe to change the hand of the heavy atoms.

shelxe pse pse_fa –h –b –s0.64 –m20

shelxe pse pse_fa –h –b –s0.64 –m20 -i

the key numbers are the Contrast and the Pseudo-free CC. There should be a clear discrimination between the two runs to indicate the correct hand. The files produced are pse.phs (the protein phases, file containing h,k,l,F,sigF,phi,fom), pse.pha (the anomalous scatter phases), pse.hat (refined heavy atom positions – this can be read into coot), and pse.lst (copy of the output of shelxe). A similar set of files is created for the second run: pse_i.phs, pse_i.pha, pse_i.hat and pse_i.lst.

You can now coot to look at the maps calculated from the phases. Read in the appropriate .hat file first, and then the .phs file.

Is the map interpretable ?

You could try running SHELXE varying the solvent content. Sometimes this helps. You could do more cycles of density modification. You could also experiment with the "free lunch" switch. Even though the data only go to 2.1 Å. Just do this on the correct solution by adding –e1.5 to the SHELXE command line.

You can read the .phs file into CCP4i using “Convert to/modify/extend MTZ” with a user defined format of (3f4.0,f9.2,f8.4,f8.1,f8.4) for H,K,L, FP, FOM, PHI, SIGFP. This can then be used to input to Arp/Warp, DM etc..

Garry Taylor, University of St Andrews (glt2@st-andrews.ac.uk)