HKL2MAP 0.4 Tutorial

Fabio Dall’Antonia & Thomas R. Schneider
EMBL Hamburg
July 2015
Steps in Anomalous Phasing

- Data Analysis
- Extraction of Substructure Structure Factors

- Substructure Determination

- Phase calculation
- Phase optimization (resolving ambiguity)
- Electron density map interpretation

- Inspection of Result

SHELXC

SHELXD

SHELXE

Coot
HKL2MAP – version 0.4

- Connects SHELXC/D/E in one Graphical User Interface

- SHELXC
  - XDS-type format added to input data types

- SHELXD
  - Multi-CPU version supported

- SHELXE
  - Heavy-atom sites can be refined (SHELXE –z switch)
  - High-resolution limit for phase extension (FLA)
  - Backbone auto-tracing for iterative phase improvement:

- COOT for visual inspection of map (& poly-ala trace)

- Various bug-fixes / additional graphics
Overall layout of HKL2MAP

The GUI contains 4 expandable/shrinkable panels:
- 3 for the SHELX programs and one for the text output

① - ③ program/module panels

④ log text panel – tabs for each computation step
SHELXC module – input

① always start with a global project name
② select the type of experiment
③ select input files containing the data

④ unit cell parameters and space group are read automatically, but the space group must be confirmed
SHELXC module – runtime

⑤ start the program

⑥ read the program std-output - it is also stored in a log file

⑦ view the graphics
SHELXC module – graphics

The available graphs depend on the type of experiment and on the nature of the data ...

MAD, merged Scalepack input

SAD, unmerged XDS input
SHELXC module – Anomalous signal-to-noise-ratio

\[ \langle d''/\text{sig} \rangle = \left\langle \frac{\Delta F}{\sigma(\Delta F)} \right\rangle \]

- Plot \( \langle \Delta F \rangle / \langle \sigma(\Delta F) \rangle \) as a function of resolution
- It is better to use scaled but unmerged data to do this (in order to not depend on the accuracy of the error estimates from data processing ...)

Plot is for unpublished MAD data (Schneider et al.) on Rabex-5 * Ub complex
Signal-to-noise and multiplicity

\[ \langle d''/\text{sig} \rangle = \left( \frac{\Delta F}{\sigma(\Delta F)} \right) \]

45 / 90 / 135 degrees of data on Zn-free insulin
SHELXD module – input

① Input files (produced by SHELXC) are already set.

② Insert parameters .ins file contents will be adjusted accordingly.
SHELXD module – runtime

③ start the program

④ read the program std-output - it is also stored in a log file

HKL2MAP supports the multi-CPU version of SHELXD

⑤ view the graphics
SHELXD module – graphics

CFOM = CC all + CC weak

Graphs provide quick/clear overview on:
- separation of correct solutions
- success rates by histogram counts
- Ideally, the site occupancy vs. peak number profile matches the expected site number.
- Clear distinction of correct from wrong peaks in terms of an occupancy step is achieved with strong anomalous signal and well-defined derivative sites – that is, (e. g.) with Se-MAD/SAD rather than S-SAD or Halide soaks.
- Occupancy = 0.3 is an empirical cut-off that often proves useful in ambiguous cases (i. e. with flat profiles)
Substructure solution with SHELXD

Iterative dual-space direct methods based on phase refinement in reciprocal space and peak picking in real space are able to locate relatively large numbers of anomalous scatterers efficiently from MAD or SAD data. Truncation of the data at a particular resolution, typically in the range 3.0–3.5 Å, can be critical to success. The efficiency can be improved by roughly an order of magnitude by Patterson-based seeding instead of starting from random phases or sites; Patterson superposition methods also provide useful validation. The program SHELXD implementing this approach is available as part of the SHELX package.
### SHELXE module – input

1. **input files (produced by SHELXC and SHELXD) are already set**

   - Native in: `jia.hkl`
   - Fa in: `jia_fa.hkl`
   - SHELXD out: `jia_fa.res`

2. **select the SHELXE phasing options**
   - (the solvent content is the only essential / non-default parameter)

   - Phase structure based on **refined** sites and modify the electron density for **20** cycles.
   - Use fractional solvent content of **0.65**.
   - native data do not include heavy atoms.
   - Extend diffraction data to [native data]
   - Run **3** cycles of autotracing with [initial search for secondary structure]
   - Interrupt calculations for incorrect enantiomorph after **4** cycles.
   - Invert heavy atom substructure for phasing? try both enantiomorphs

   - Phases ori: `jia_m20_s0.65.phs`
   - Phases inv: `jia_m20_s0.65_i.phs`

---

- **without extra information, one has to start with both …**

- **site refinement can significantly increase the chances of successful phasing**

- **from number of residues in ASU**
The “Contrast vs. Cycle” graph for mere density modification (whether or not auto-tracing was activated) is the first source of information on phasing success. Clear separation of electron density contrast between phase sets starting from the opposite substructure enantiomorphs indicates successful phasing. If this is not achieved within the initial (default: 20) DM cycles, poly-ala auto-tracing is probably required in order to enhance the phasing performance.
Iterative phase improvement by alternating density modification / auto-tracing macro-cycles can help a lot with bad initial phase estimates.

Inspection of the poly-ala model structures by 3D-ribbon graphs is an additional means to identify the correct substructure enantiomorph.

The models are freely rotatable by mouse drag.

CC\_partial > 25% usually indicates successful phasing.
SHELXE module – Coot interface

- Coot can read phs files directly and do the FFT to display maps
- HKL2MAP provides a script (.scm) for quick-loading of PDB model and map and launches Coot on button press
- Thus, the Coot display is just “one mouse-click away”.

jia_m20_s0.65_i.scm