Using CCP4 for PX

Martin Noble, Oxford University
and CCP4
Running programs locally or remotely
Overview

Data collection and analysis

Molecular Replacement

Experimental Phasing

Density modification

Model building and refinement

Validation, Analysis, Deposition

Project organisation

Molecular graphics

External programs
Project Organisation

2) Workflows

- Experimental Phasing
  - Data Reduction
  - Experimental Phasing
  - Molecular Replacement
  - Density Improvement
  - Model Building
  - Refinement
  - Structure Analysis
  - Validation & Deposition
  - Map & Mask Utilities
  - Reflection Data Utilities
  - Coordinate Utilities
  - Graphics and Viewing Utilities
Project Organisation

1) Graphical
Project Organisation

3) Data
   - Project
   - Crystal
   - Dataset
Data collection and analysis

MOSFLM → SCALA
Molecular Replacement

[Diagram showing the process of molecular replacement]

Preparing data → Rotation function → Orientation refinement → Translation function → Rigid body refinement

“Known” fraction

[Phaser, BP3]
Amore
Molrep capabilities

++-- Self RF  (FUN=R, without any model)
++-- Standard MR  ++-- Cross RF  (FUN=A or FUN=R )
++-- TF  (FUN=A or FUN=T )
++-- two identical models
++-- Dyad search  ++
++-- (DYAD=D)  !
++-- Multi-copy search  ++-- two different
++-- for MR  models
++-- MOLREP  ++-- Multy-copy for one model
++-- RF and PTF
++-- (DYAD=Y)
++-- Fitting two models  ++
++-- SAPTF, PRF and PTF
++-- (PRF=Y)
++-- RF and PTF
++-- (PRF=N)
++-- Searching in ED map  ++
++-- SAPTF, PRF and PTF
++-- (PRF=Y)
++-- Rotate and position the model (FUN=S)
Molrep

- **scaling by Patterson** origin peak, **soft low resolution cut-off**, anisotropic correction of data
- **standard molecular replacement method** by rotation function (RF), full-symmetry translation function (TF) and packing function (PF)
- allows input of **a priori knowledge** of similarity and completeness of the model (By pseudo B-factor).
- performs rigid body refinement
- can check and manage **pseudo-translation**
- can compute **Self Rotation Function** with PostScript plots
- Spherically averaged phased translation function (**SAPTF**)
- Phased Rotation (**PREF**) and Phased Translation functions (**PTF**)
- (superimposing two models)
- can evaluate R-factor, CC for a proposed rotation and translation
- **multi-copy search**
- can **improve** the model before use, model correction by sequence **alignment**
- can choose from symmetry-related models closest to which was found before
- can use **modified stucture factors** instead of Fobs for RF
- can use **NMR** models

Phaser

- Likelihood targets
  - Multiple search models
  - Test multiple space groups
  - Rapid likelihood target (FRF) prescreen, followed by higher accuracy rescoring
  - Packing function screening of translation function solutions
Complete determination of structures at atomic resolution

Experimental Phasing

Preparing data

Finding sites

Refining sites and phasing

Experimental Phasing

- Merge Datasets (CAD)
- Scale and Analyse Datasets
- Prepare Data for HA Search
- Acorn - ab initio Phasing
- Shelx - Heavy Atom Search
- Rantan - Direct Methods
- Profess - NCS from HA
- Run Mlphare
- Oasis - SAD/SIR phasing
- Generate Patterson Map
- Real Space Patterson Search
Experimental phasing

- Preparing data
- Finding sites
- Refining sites and phasing

<table>
<thead>
<tr>
<th>CCP4</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECALC</td>
<td>XPREP</td>
</tr>
<tr>
<td>REVISE</td>
<td></td>
</tr>
<tr>
<td>RSPS</td>
<td>SHELX</td>
</tr>
<tr>
<td>ACORN</td>
<td>AUTOSHARP</td>
</tr>
<tr>
<td>RANTAN</td>
<td>SOLVE</td>
</tr>
<tr>
<td>MLPHARE</td>
<td>SOLVE</td>
</tr>
<tr>
<td></td>
<td>SHARP</td>
</tr>
<tr>
<td></td>
<td>AUTOSHARP</td>
</tr>
</tbody>
</table>
Using CCP4

Graphical Feedback
Finding Sites (External)

- SHELX
Finding sites

- RSPS
  - Patterson search
    - Harker section searches combined by sum, mean or harmonic mean
    - Cross peaks used to relate multiple sites to the same origin
  - Can exploit NCS
    - Full transformation to generate “pseudo Harker”
    - Can search for pairs of atoms with a known separation

Finding sites

- Rantan
  - Data prepared by (Revise and) ECALC
  - Tangent formula refinement of random phase sets for a large number of triplets

MLPHARE

Phased refinement suitable for SAD, MAD, SIR, SIRAS, MIRAS

- Some maximum likelihood concepts included to minimise phase bias
- Generates coefficients for use in double difference Fouriers for model completion

Z. Otwinowski: Daresbury Study Weekend proceedings, 1991
OASIS

- To break phase ambiguity of SIR (or SAD) experiments by recourse to direct methods and density modification.

Refining sites

2) External interfaces 1

2.1 SOLVE (Terwilliger, solve.lanl.gov)

```plaintext
solve <<EOD
! ccp4 mtz file input
! solve a 2-deriv MIR dataset
logfile mir.logfile
! write out most information to this file.
! summary info will be written to solve.prt
resolution 20 3
! you need resolution, space group and cell
! dimensions read from mtz file.
! get the mtz file information and read it in:
labin FP=FP SIGFP=SIGFP FPH1=FPH1 SIGFPH1=SIGFPH1 DPH1=DPH1 SIGDPH1=SIGDPH1
labin FPH2=FPH2 SIGFPH2=SIGFPH2 DPH2=DPH2 SIGDPH2=SIGDPH2
hklin mir_fbar-cad.mtz
! now we're ready with scaled data
```
Refining sites

- 2) External interfaces
  - 2.1 SHARP
    - (Bricogne, babinet.globalphasing.com)
Density modification

Map → Improved Map

Phases → Improved Phases

Density Improvement
- Cell Content Analysis
- Find NCS from Heavy Atoms
- Run DM
- Run Solomon
- Run DmMulti
DM

- Solvent flattening
- NCS averaging (multi-domain)
- Multicrystal averaging
- Histogram matching
- Skeletonisation
- Sayre phase improvement
- Automatic or Manual mask definition
- Inbuilt operator refinement
- *Multiresolution modification and Perturbation gamma correction

Solomon

- Solvent flipping
- Protein truncation
- Averaging

Model Building

1) Automatic

- ACORN
- ARP/wARP
- FFFear/ FFJoin
Model building: Arp/wArp

☐ 1. Automatic tracing of the density map and model building. (MR-solution refinement and the improvement of MAD and M(S)IR(AS) phases via map interpretation

☐ 2. Free atoms density modification

☐ 3. Building of the solvent structure

Model Building

2) Graphical
   a) Inhouse
      CCP4 Molecular Graphics
   b) Closely linked
      COOT
   c) Well interfaced
      O, mifit
CCP4 Molecular Graphics
Coot wonders

http://www.ysbl.york.ac.uk/~emsley/coot/
Refinement tools
Refinement

- REFMAC strengths
  - Algorithms
  - Monomer libraries (DNA, RNA, Small molecules)
Refinement

- Monomer sketcher