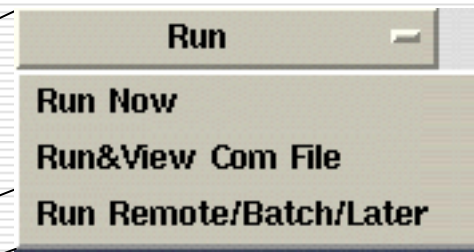
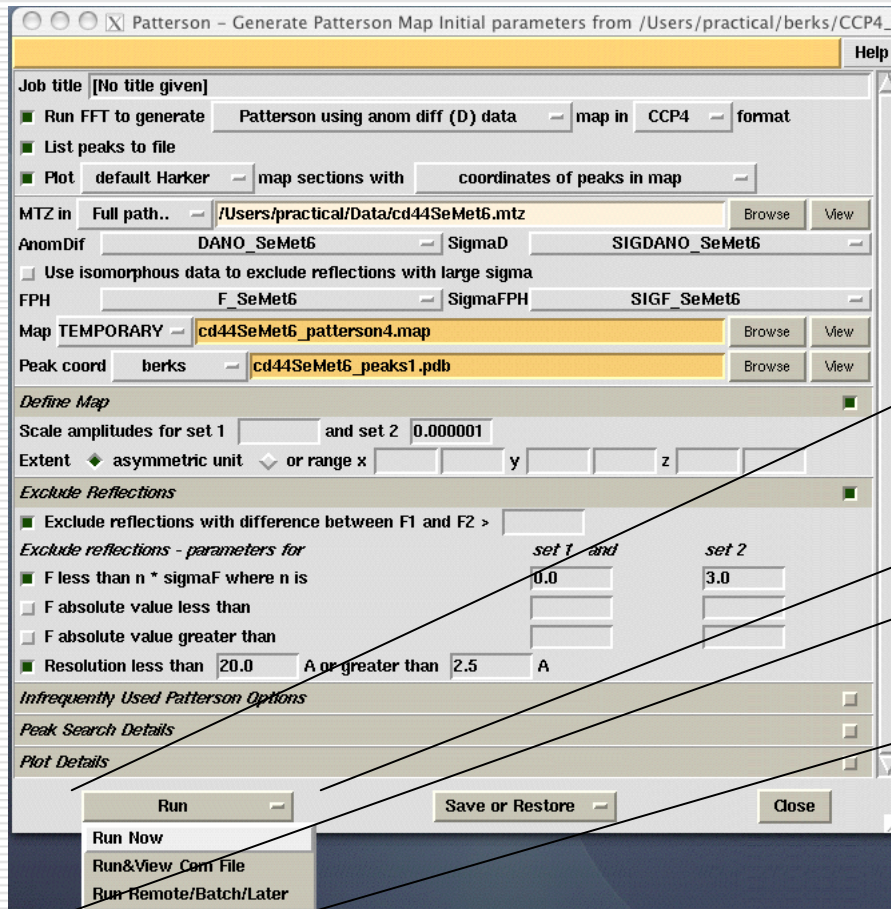


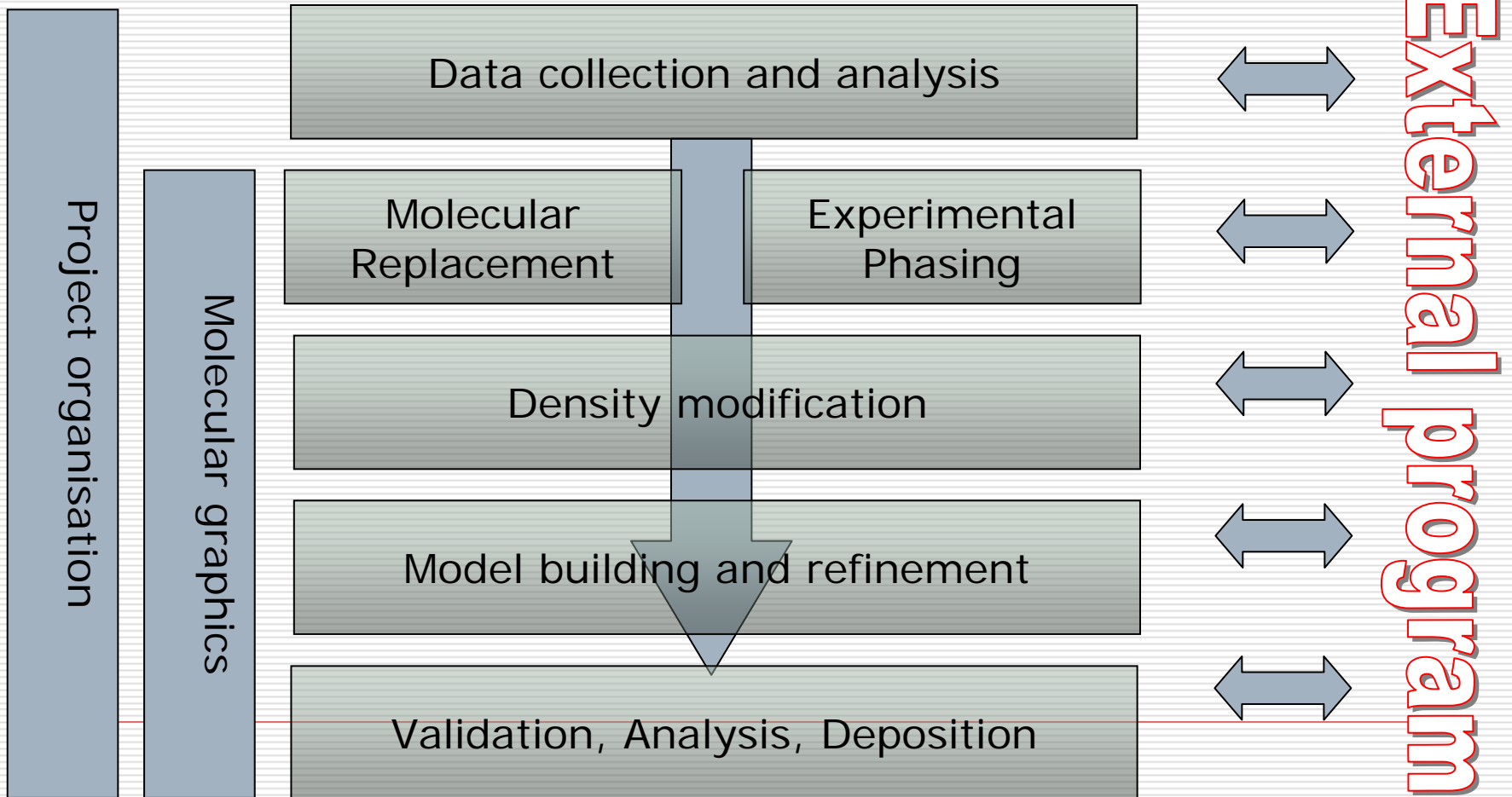
Using CCP4 for PX

Martin Noble, Oxford University
and CCP4

Running programs locally or remotely



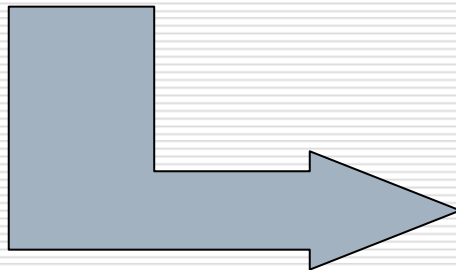
Overview



Project Organisation

□ 2) Workflows

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.



QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Experimental Phasing

Data Reduction

Experimental Phasing

Molecular Replacement

Density Improvement

Model Building

Refinement

Structure Analysis

Validation & Deposition

Map & Mask Utilities

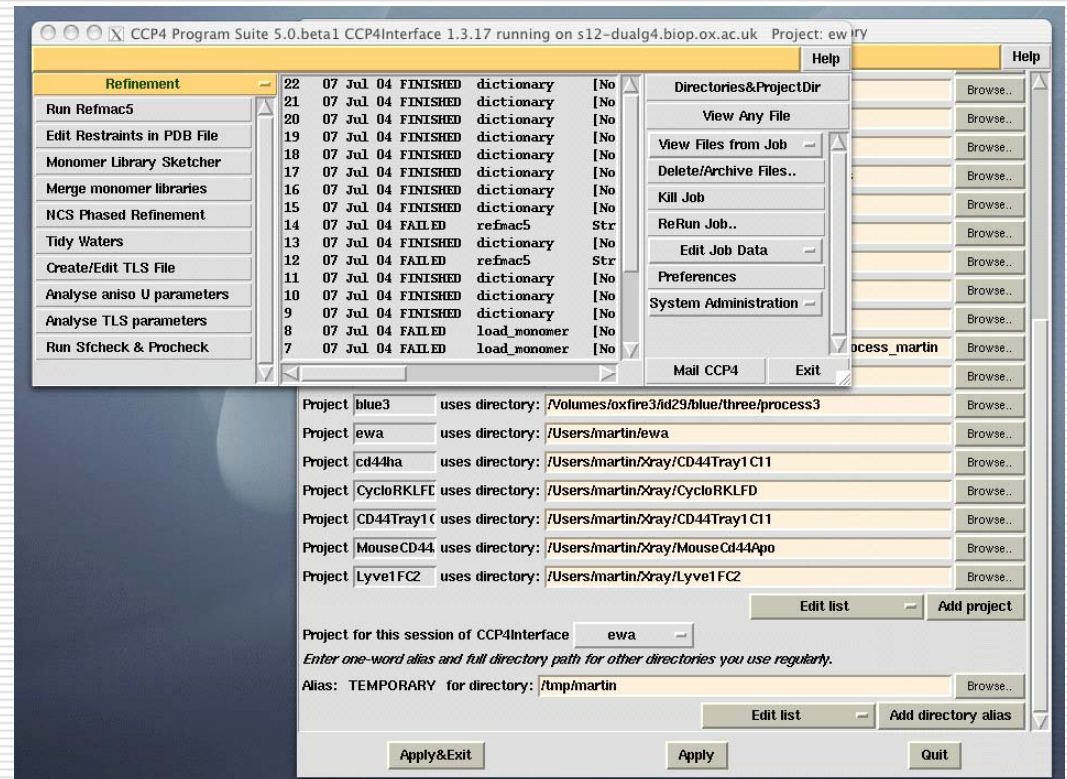
Reflection Data Utilities

Coordinate Utilities

Graphics and Viewing Utilities

Project Organisation

1) Graphical



Project Organisation

- 3) Data
 - Project
 - Crystal
 - Dataset

Scala - Scale Experimental Intensities Initial parameters from /Users/martin/Xray/CD44Tray1C11/C

Job title [No title given]

Customise Scala process (default is to refine & apply scaling)

Separate anomalous pairs for merging statistics

Run Truncate to output Wilson plot and SFs after scaling and output a single MTZ file

Ensure unique data & add FreeR column for 0.05 fraction of data.

Generate Patterson map and do peaksearch to check for pseudo-translations

MTZ in mosflm1.mtz

Override automatic definition of 'runs' to mark discontinuities in data

Exclude data resolution less than 23.421 Angstrom or greater than 1.200 Angstrom

MTZ out CD44Tray1C11.mtz

Convert to SFs & Wilson Plot

Estimated number of residues in the asymmetric unit 151

Use dataset name as identifier to append to column labels

Include the intensities in the output MTZ file(s)

Data Harvesting

Create harvest file in project harvesting directory

Define Output Datasets

The input file contains a single dataset, which will be transferred to the output file.

Crystal CD44Tray1C11_2 belonging to Project MouseCD44

Dataset name Native1

Scaling Protocol

Scale on rotation axis with isotropic Bfactor scaling

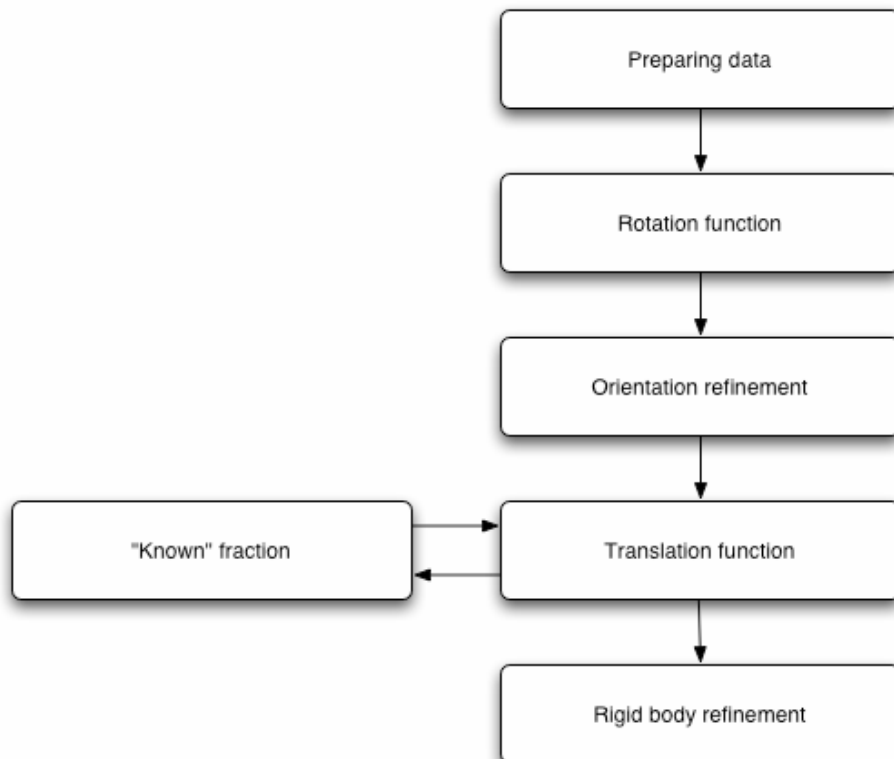
Define scale ranges along rotation axis by rotation interval 5

Independent Bfactors defined by rotation interval 20

Apply tails correction with width 0.01 fraction in peak 0.0 slope 10.0

Observations Used & Handling of Partial

Molecular Replacement

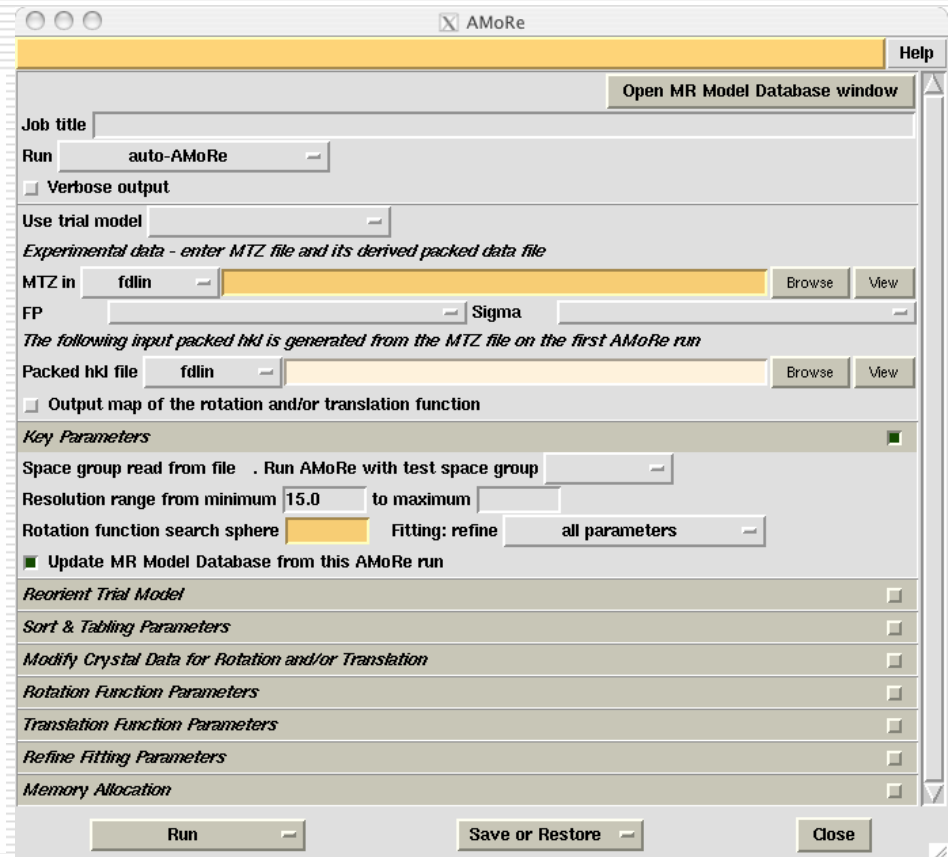
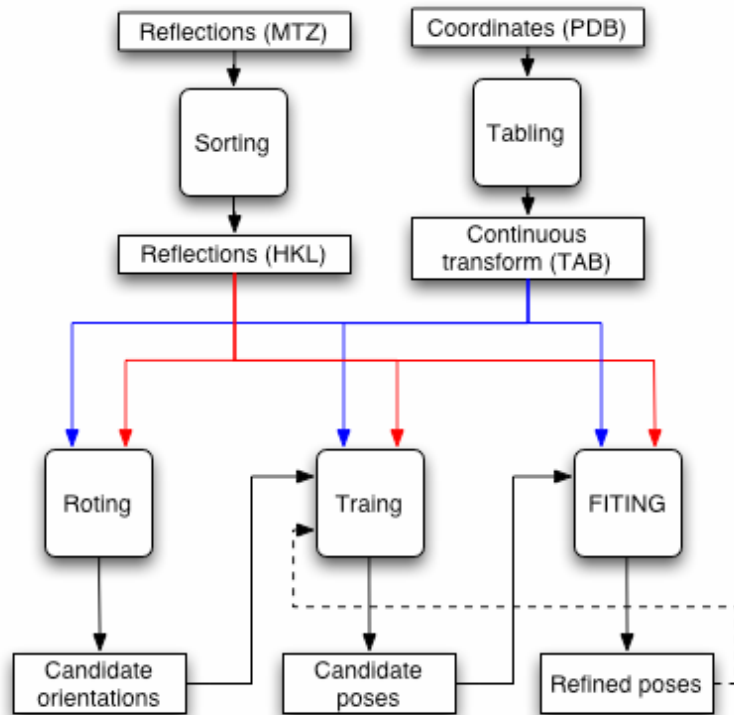


Molecular Replacement

- Cell Content Analysis
- Analyse Data for MR
- Self RF in polars
- Beast - likelihood-based MR
- Molrep - auto MR
- Create Input SFs from Model
- AMoRe Model Database
- AMoRe
- Edit AMoRe Solution File
- Build AMoRe Output Model
- Edit Protein Structure

[Phaser, BP3]

Amore



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([PRF](#)) 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 structure factors](#) instead of Fobs for RF
- ❑ • can use [NMR](#) models

-
- ❑ *J. Appl. Cryst.* (1997). **30**, 1022-1025 **MOLREP: an Automated Program for Molecular Replacement**[A. Vagin](#) and **A. Teplyakov**

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

Phaser - molecular replacement using likelihood-based methods Initial parameters from /User

Job title [No title given]

Mode for molecular replacement auto search (rotate/translate/pack)

Define data

Input data

MTZ in MouseCD44 CD44Tray1C11_p21.mtz Browse View

F F_Native1 SIGF SIGF_Native1

Perform anisotropy correction Resolution range for molecular replacement: 81.923 A to 2.5 A

Space group read from mtz file 'P21'

Run Phaser with alternative space group(s)

Test space group #1 P21

Test space group #2 P2

Edit list Add another space group

Output data

XYZ (PDB) file output on

HKL (MTZ) file output off

Number of top solutions to output as PDB and MTZ files 1 NB: MTZ files not output for all modes

Composition of the asymmetric unit

Component #1 protein Molecular weight 17000 Number of copies in asymmetric unit 1

Edit list Define another protein/NA

Define ensembles (each comprising possible model(s) for a protein/nucleic acid component)

Ensemble # 1

Ensemble id ensemb1 Define ensemble via pdb file(s)

PDB #1 MouseCD44 cd44A.pdb Browse View

Effective RMS error of PDB #1 sequence identity 0.9

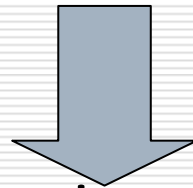
Edit list Add superimposed PDB file to the ensemble

Run Save or Restore Close

Acorn

□ Patterson-based heavy atom substructure determination

□ Molecular replacement of secondary structural elements



Complete determination of structures at atomic resolution

Experimental 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

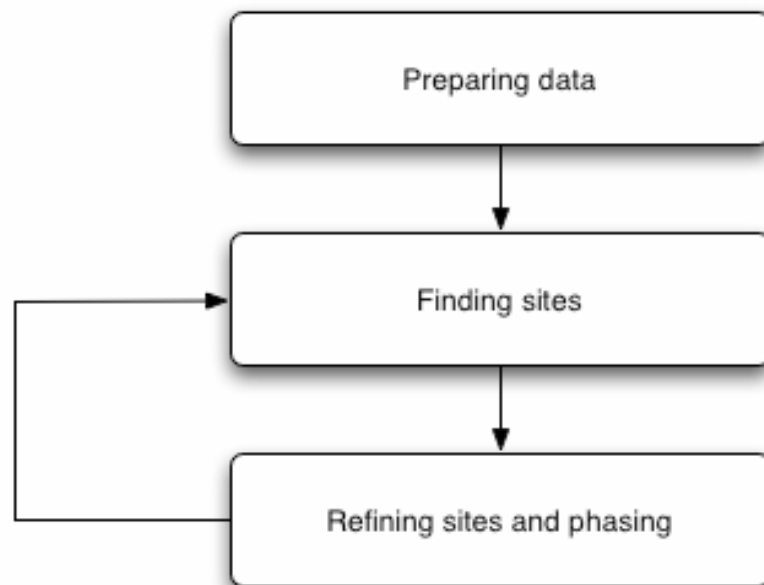
Professs - NCS from HA

Run Mlphare

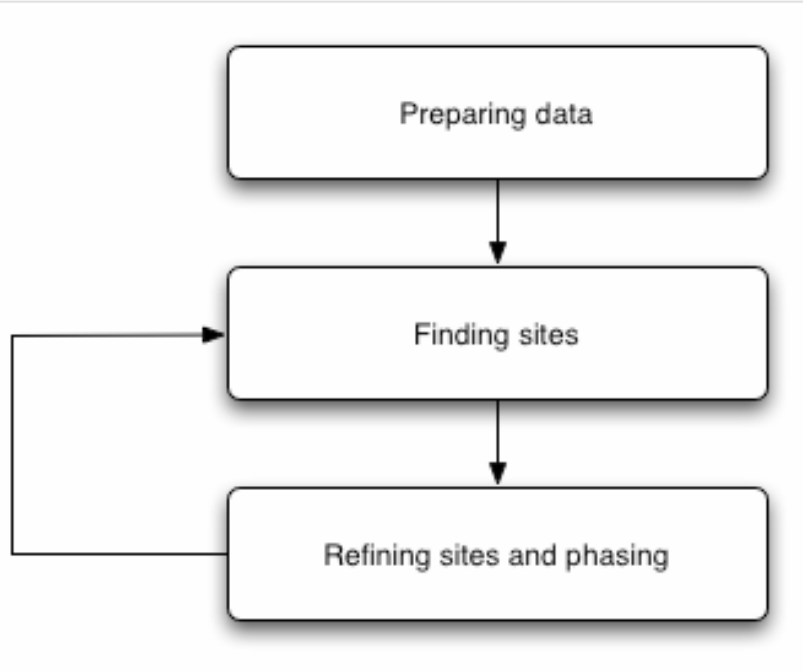
Oasis - SAD/SIR phasing

Generate Patterson Map

Real Space Patterson Search



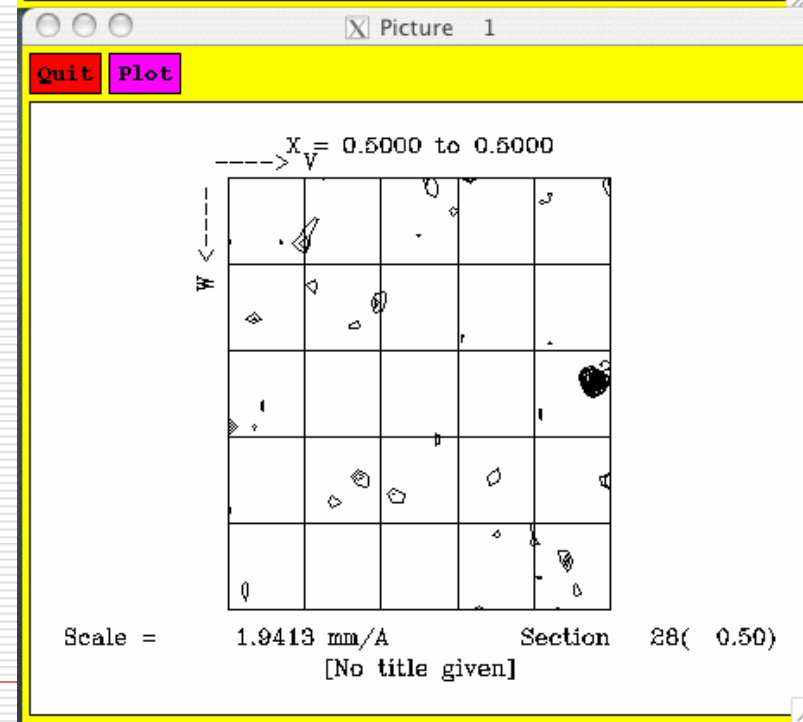
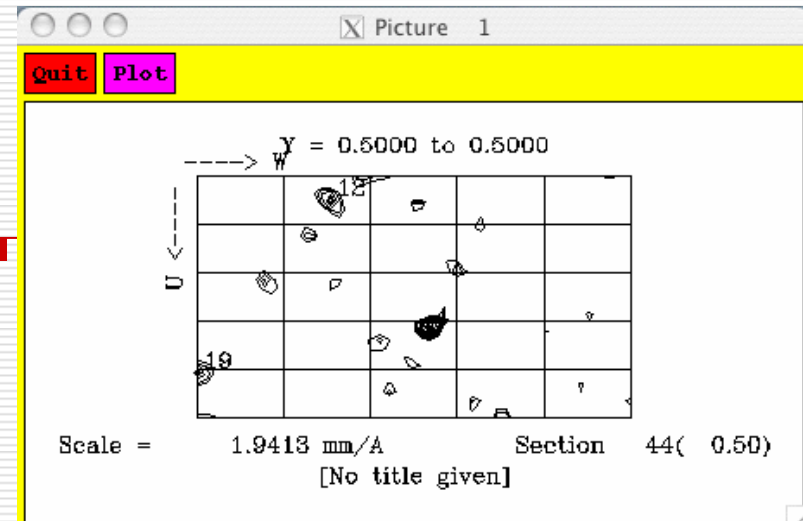
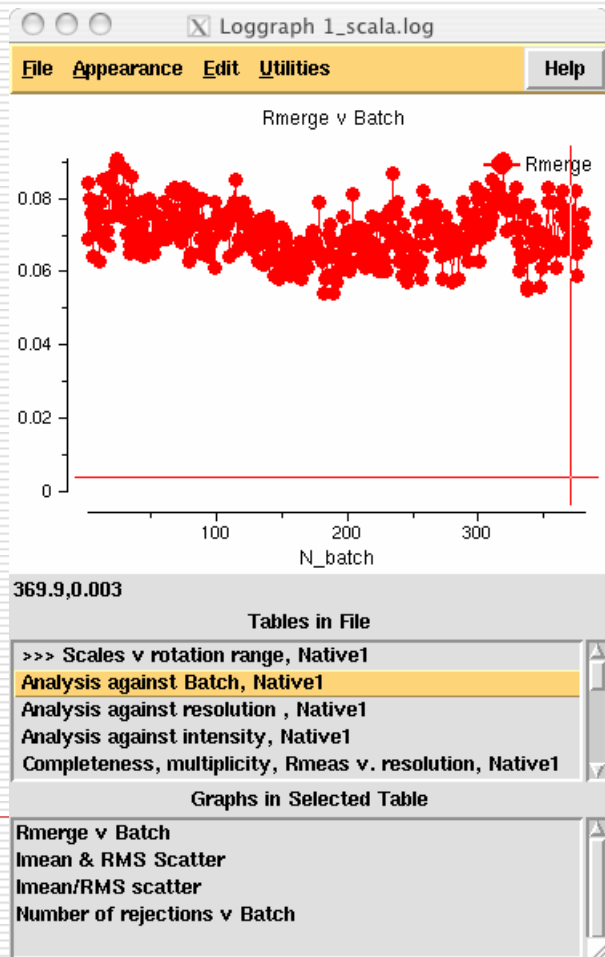
Experimental phasing



CCP4	Others
ECALC REVISE	XPREP
RSPS ACORN RANTAN	SHELX AUTOSHARP SOLVE
MLPHARE	SOLVE SHARP AUTOSHARP

Using CCP4

Graphical Feedback



Finding Sites (External)

□ SHELX

The screenshot shows the 'Shelx - Heavy Atom Search' window. It features a yellow title bar with a 'Help' button. Below the title bar is a note: 'NOTE This task uses the Shelx program which is not distributed with CCP4 - click Help for details'. The interface is organized into several sections:

- Job title:** An empty text input field.
- Try to find heavy atoms by:** A dropdown menu set to 'Patterson search'.
- Input format is:** A dropdown menu set to 'MTZ file', followed by 'using' and another dropdown menu set to 'isomorphous', and finally 'data'.
- MTZ in:** A dropdown menu set to 'berks', followed by a text input field containing 'cd44_3_2_001_scala2.mtz', and 'Browse' and 'View' buttons.
- FP:** A dropdown menu set to 'F_Unspecified', followed by 'Sigma' and a dropdown menu set to 'SIGF_Unspecified'.
- FPH:** A dropdown menu set to 'F_Unspecified', followed by 'SigmaPH' and a dropdown menu set to 'SIGF_Unspecified'.
- Cell Parameters:** A section with a green expand/collapse icon. It includes:
 - Space group:** 'P212121' and **and lattice type:** 'primitive'.
 - Wavelength:** '1.54178'.
 - Set cell a:** '48.9086', **b:** '77.3351', **c:** '87.7317', **alpha:** '90.0000', **beta:** '90.0000', **gamma:** '90.0000'.
 - Cell ESDs a:** '0.001', **b:** '0.001', **c:** '0.001', **alpha:** '0.0', **beta:** '0.0', **gamma:** '0.0'.
 - Search for:** '4' **atoms of:** 'AU'.
- Exclude Reflections:** A section with a grey expand/collapse icon and a checkbox.
- Shelx Patterson Search Parameters:** A section with a green expand/collapse icon. It includes:
 - Try:** An empty text input field followed by 'superposition vectors'.
 - Enter 'known' vectors:** An empty text input field.
 - Edit list:** A dropdown menu.
 - Add Vector:** A button.

At the bottom of the window are three buttons: 'Run', 'Save or Restore', and 'Close'.

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 Fourier for model completion

The screenshot shows the MLPHARE software interface with the following settings:

- Job title: [No title given]
- Use anomalous difference data Add phase info from externally calculated phases
- Use every 10 -th reflection for refinement
- Apply calculated scale to output SFs
- Output Hendrickson-Lattmann coefficients
- Output map(s) in: CCP4 format
- Output: map of FP with final best Phi cross-peaks map(s)
- Generate double difference maps and do peak search for more heavy atoms
- MTZ in: mosflm1_mosflm2_scala1.mtz
- FP: F_SeMet6, SigmaFP: SIGF_SeMet6
- FPH1: F_SeMet6, SigFPH1: SIGF_SeMet6
- DPH1: DANO_SeMet6, SigDPH1: SIGDANO_SeMet6
- Buttons: Edit list, Add Another Derivative
- MTZ out: mosflm1_mosflm2_mlphare1.mtz
- Output label identifier: mlphare1
- Data Harvesting: Create harvest file in project harvesting directory
- Harvest project name: CD44 and dataset name: SeMet6
- Key Parameters: Resolution limit from 20.226 to 2.6
- Angle interval (degrees) for phase probability curve: 12
- Do refinement for 30 cycles
- Describe Derivatives & Refinement:
- Derivative number 1: Phase with&refine derivative
- Use: anomalous data to refine XYZ and alternate Occ & B
- Buttons: Run, Save or Restore, Close

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)

```
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)

autoSHARP Control Panel (user : edl)

User: edl Project ID: MAD

SAD data with 1 wavelength
Entry Point: merged and scaled data

1. General

1.1. Project identifier:

1.2. Title:

1.3.1. Molecular weight: [Da]
- OR -

1.3.2. No. of residues:

1.4.1. No of expected sites: Chemical element:
- OR -

1.4.2. HA sites:

1.5. What to do:

2. Wavelength No. 1

2.1. Dataset identifier: [\(explanation\)](#)

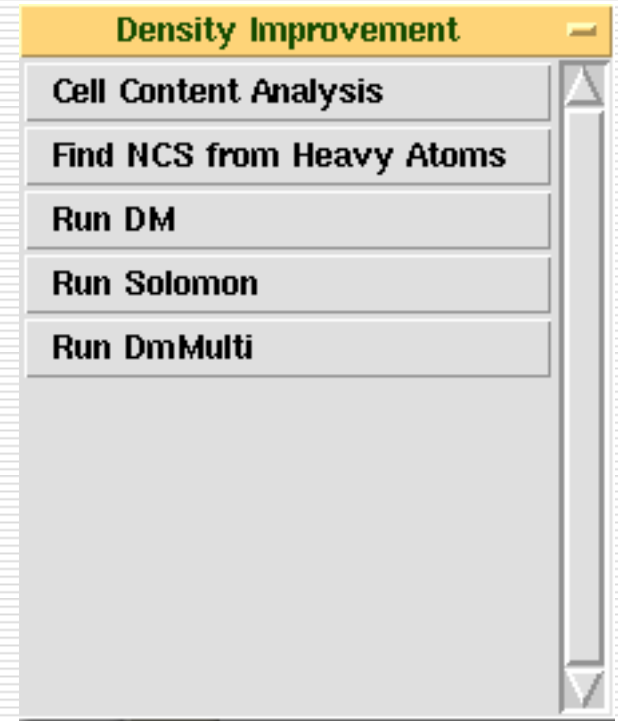
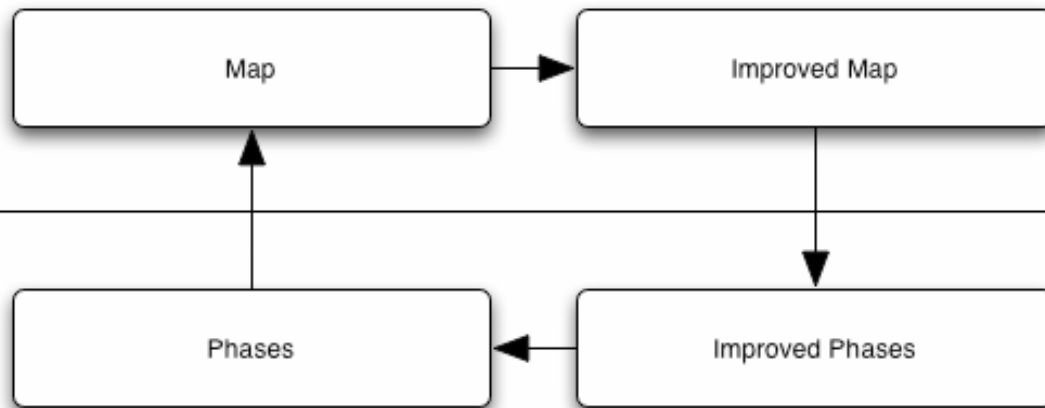
2.2.1. Wavelength: [Å] [\(explanation\)](#)
- OR -

2.2.2. f': f'':

2.3. Datafile: [\(explanation\)](#)

Column labels: FMID: SMID:
DANO: SANO:

Density modification



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

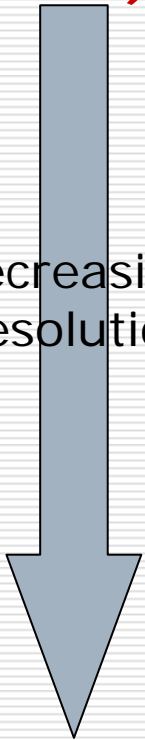
1) Automatic

■ ACORN

■ ARP/wARP

■ FFFear/ FFJoin

Decreasing
Resolution



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

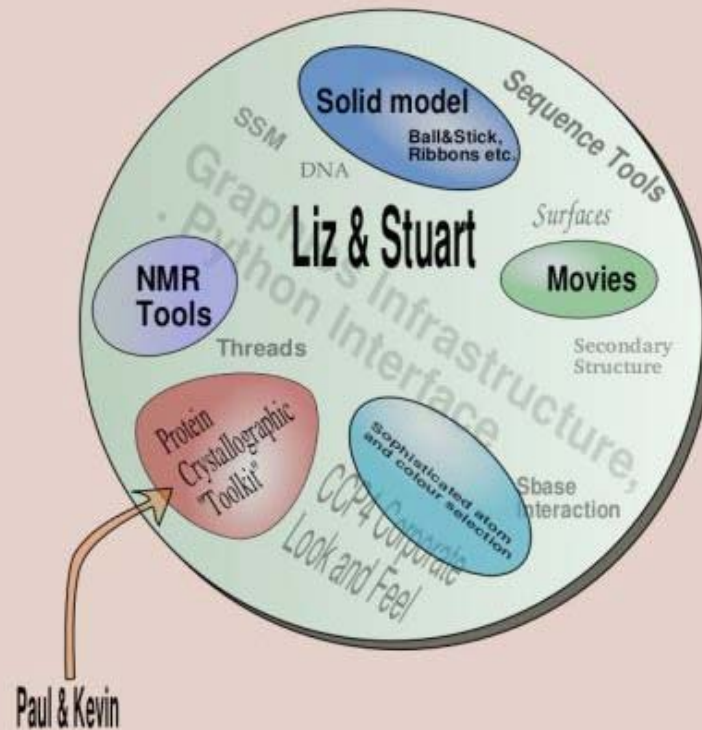
COOT

c) Well interfaced


O, mifit

CCP4 Molecular Graphics

CCP4MG Organisation (at York)



Coot wonders



QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Refinement tools

Refinement

Run Refmac5

Edit Restraints in PDB File

Monomer Library Sketcher

Merge monomer libraries

NCS Phased Refinement

Tidy Waters

Create/Edit TLS File

Analyse aniso U parameters

Analyse TLS parameters

Run Sfcheck & Procheck

Refinement

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

Refinement

□ Monomer sketcher

