DIALS: software for ED data processing

David Waterman
CCP4 / SCD / UKRI-STFC

NCCAT MicroED Symposium
November 2025

CCP4

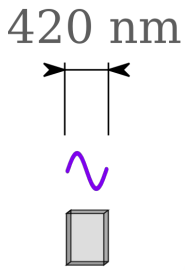
DIALS
Diffraction Integration for Advanced Light Sources

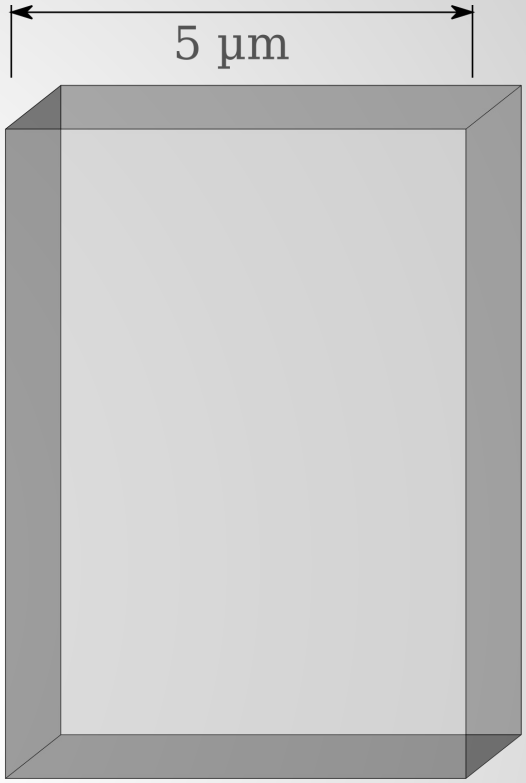
Introduction

MicroED is a method for crystallographic rotation data collection from *vanishingly small* crystals in an electron microscope

Software for MicroED data processing can be divided into specialised packages, and those borrowed from X-ray crystallography

Specialised	Borrowed/adapted
ADT3D	MOSFLM
RED	DIALS
PETS2	XDS
	CrysAlis^{Pro}





CCP4

DIALS
Diffraction Integration for Advanced Light Sources

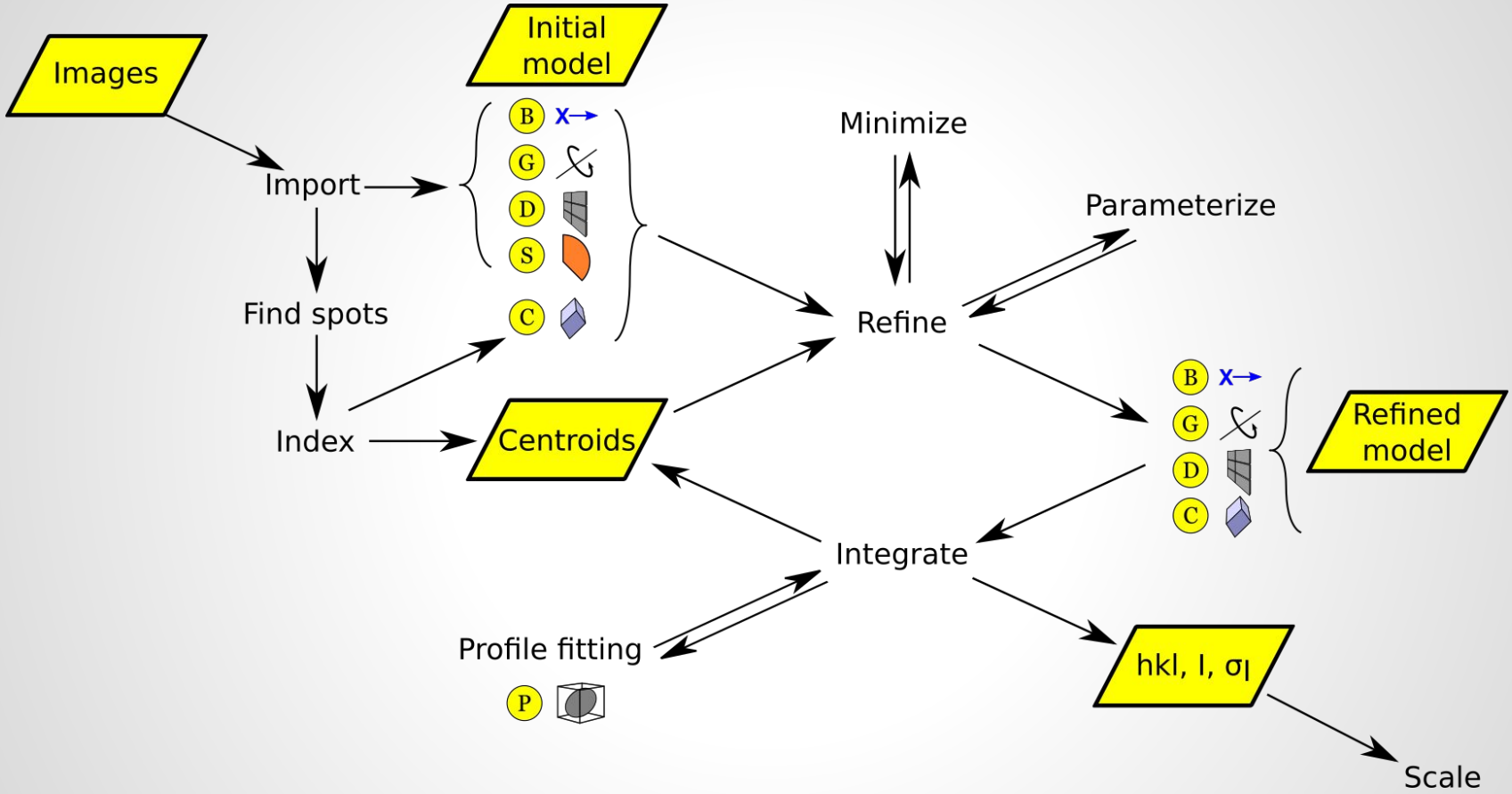
What is DIALS?

Diffraction **I**ntegration for **A**dvanced **L**ight **S**ources

- A decade and a half of development
- International open-source software collaboration
- General framework and toolkit
- Single crystal rotation and stills, MX and CX
- X-rays, electrons, neutrons

https://dials.github.io/

DIALS data processing flowchart



CCP4

DIALS
Diffraction Integration for Advanced Light Sources

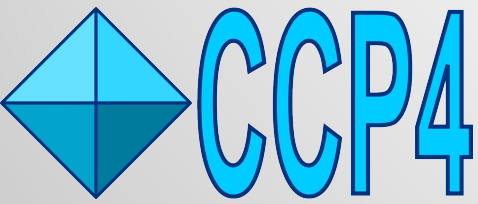
DIALS CLI

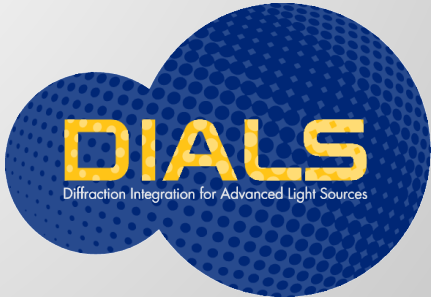
Main programs:

dials.import
dials.find_spots
dials.find_rotation_axis
dials.index
dials.refine_bravais_settings
dials.reindex
dials.refine
dials.integrate
dials.symmetry
dials.two_theta_refine
dials.scale
dials.export

```
$ dials.find_spots imported.expt

...

Histogram of per-image spot count for imageset 0:
646 spots found on 72 images (max 25 / bin)
                                    *
    *                               *
    **                        *     **
    **                    * * *     **
    ***        ***  *  ***** *      ****  *
    ***        ***  ** ***** * *  * ****  *  *      *
   ****     ***************** * ** ******* *****   * **   **
** ****    ****************** ***************** * ** ****
********   ************************************************
************************************************************
1                           image                          72

------------------------------------------------------------
Saved 646 reflections to strong.refl
```

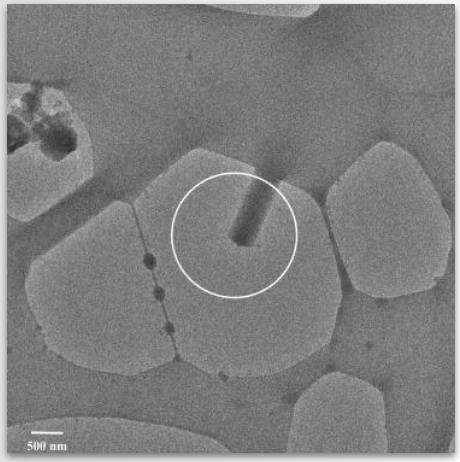
About 90 **dials.*** commands in version 3.25.0





DIALS for MicroED

Since 2018 various features for MicroED have been added to DIALS



research papers

STRUCTURAL BIOLOGY

Electron diffraction data processing with *DIALS*

Max T. B. Clabbers,[a] Tim Gruene,[b] James M. Parkhurst,[c] Jan Pieter Abrahams[a,b] and David G. Waterman[d]*

Received: 26 March 2018
Accepted: 21 May 2018

Edited by R. J. Read, University of Cambridge, England

Keywords: electron microscopy; electron crystallography; protein nanocrystals; diffraction geometry; *DIALS*.

Supporting information: the article has supporting information at journals.iucr.org/d

Electron diffraction is a relatively novel alternative to X-ray crystallography for the structure determination of macromolecules from three-dimensional nanometre-sized crystals. The continuous-rotation method of data collection has been adapted for the electron microscope. However, there are important differences in geometry that must be considered for successful data integration. The wavelength of electrons in a TEM is typically around 40 times shorter than that of X-rays, implying a nearly flat Ewald sphere, and consequently low diffraction angles and a high effective sample-to-detector distance. Nevertheless, the *DIALS* software package can, with specific adaptations, successfully process continuous-rotation electron diffraction data. Pathologies encountered specifically in electron diffraction make data integration more challenging. Errors can arise from instrumentation, such as beam drift or distorted diffraction patterns from lens imperfections. The diffraction geometry brings additional challenges such as strong correlation between lattice parameters and detector distance. These issues are compounded if calibration is incomplete, leading to uncertainty in experimental geometry, such as the effective detector distance and the rotation rate or direction. Dynamic scattering, absorption, radiation damage and incomplete wedges of data are additional factors that complicate data processing. Here, recent features of *DIALS* as adapted to electron diffraction processing are shown, including diagnostics for problematic diffraction geometry refinement, refinement of a smoothly varying beam model and corrections for distorted diffraction images. These novel features, combined with the existing tools in *DIALS*, make data integration and refinement feasible for electron crystallography, even in difficult cases.

1. Introduction

Electron diffraction (ED) allows the structural analysis of nanometre-sized samples of crystalline material. Since the minimal radiation dose is proportional to the sample volume, electron diffraction of organic and macromolecular compounds was long limited to two-dimensional samples (Unwin & Henderson, 1975; Henneller, 2017). In contrast to X-ray crystallography, the three domains, inorganic, organic and macromolecular electron crystallography, developed rather independently of each other (Vainshtein, 1964; Dorset, 1995; Kolb *et al.*, 2007; Glaeser *et al.*, 2007; Zou *et al.*, 2011). Physical and instrumental limitations, such as miniature sample size or dynamic scattering effects and lens distortions, affect data precision. However, several studies show that the model accuracy compares with that of X-ray structures (Weirich *et al.*, 1996; Magnuski & Kolb, 2014; Dorset, 1995; Palatinus *et al.*, 2017). Only about one and a half decades ago, electron diffraction of three-dimensional crystals was pioneered with automated diffraction tomography (ADT) and was further refined with rotation electron diffraction (RED).

- Distortion correction map
- Beam drift modelling
- Refinement diagnostics
- Image format readers
- Spotfinding for Ceta-D etc.
- Rotation axis determination
- Serial ED indexing
- Beam centre determination
- …

General strengths of DIALS include:

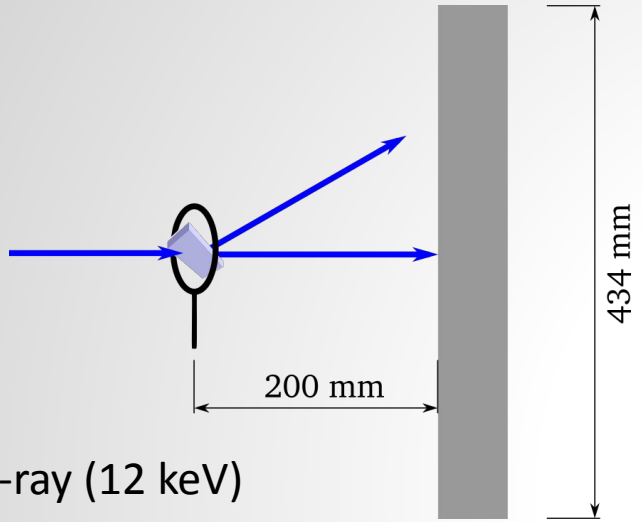
• Open-source extensible framework • Scripting • Multi-crystal analysis

• Familiarity for Structural Biologists • Support from DLS/CCP4/LBNL

CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Diffraction geometry: from MX to ED



X-ray (12 keV)

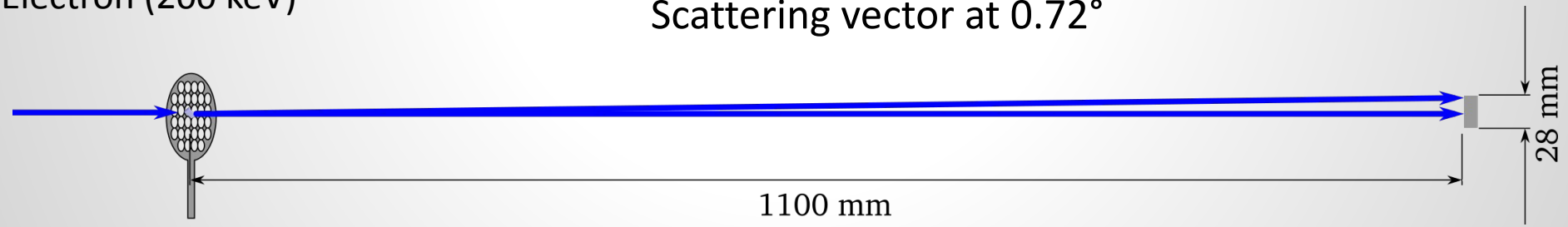
Typical MX geometry, Pilatus 6M detector

Scattering vector at 29.9°

Electron (200 keV)

Real ED example, Timepix quad

Scattering vector at 0.72°



CCP4

DIALS
Diffraction Integration for Advanced Light Sources

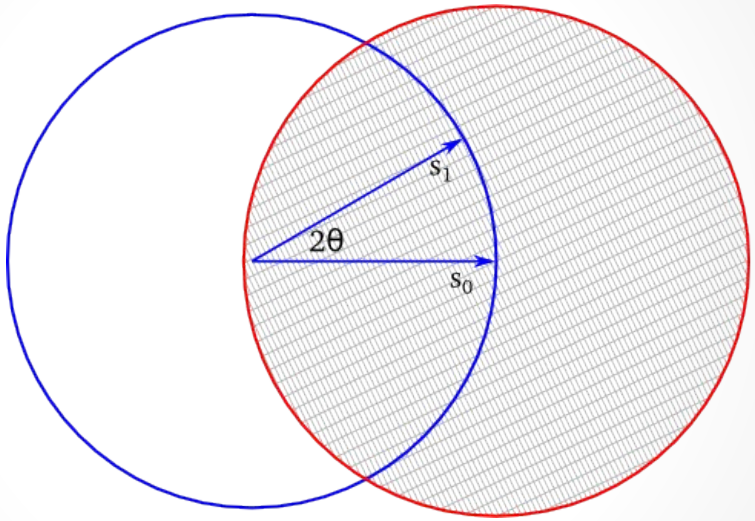
Ewald construction

X-ray diffraction

Photon energy 12 keV (λ = 1.03 Å)

Scattering vector at 2 Å

2θ = 29.9°



CCP4

DIALS
Diffraction Integration for Advanced Light Sources

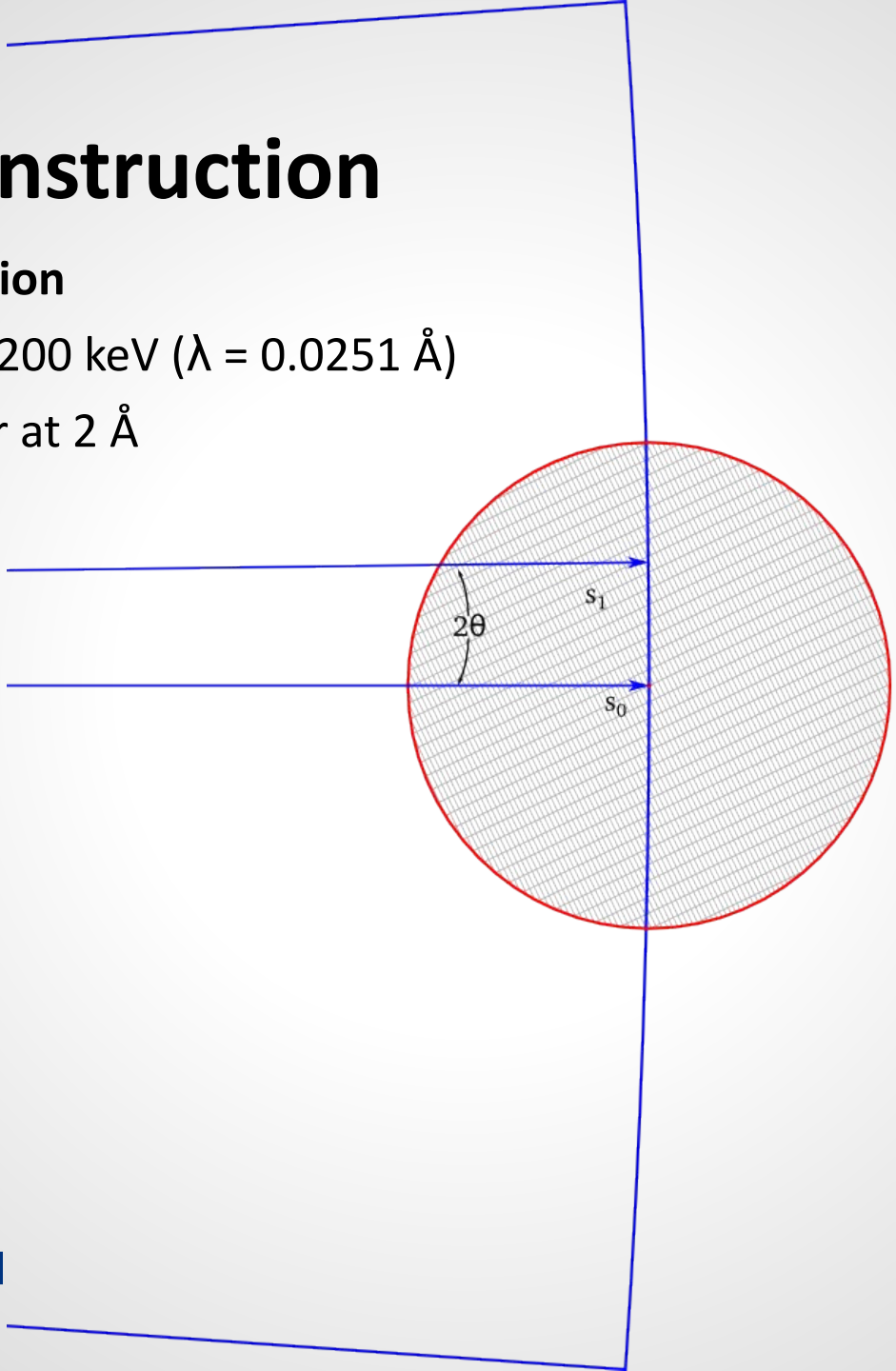
Ewald construction

Electron diffraction

Electron energy 200 keV (λ = 0.0251 Å)

Scattering vector at 2 Å

2θ = 0.72°



CCP4

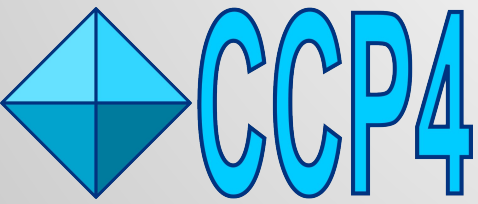
DIALS
Diffraction Integration for Advanced Light Sources

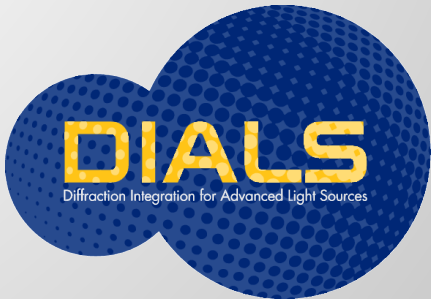
Challenges

Almost flat Ewald sphere, high detector distance and low diffraction angle

- Joint refinement of detector and unit cell may not be possible
- Refined detector and unit cell parameters may be poor
- Indexing from a single image is challenging
- It may even be difficult to determine the direction of rotation
- The beam centre may be unknown and can drift
- Lens distortions may introduce systematic error in observed positions

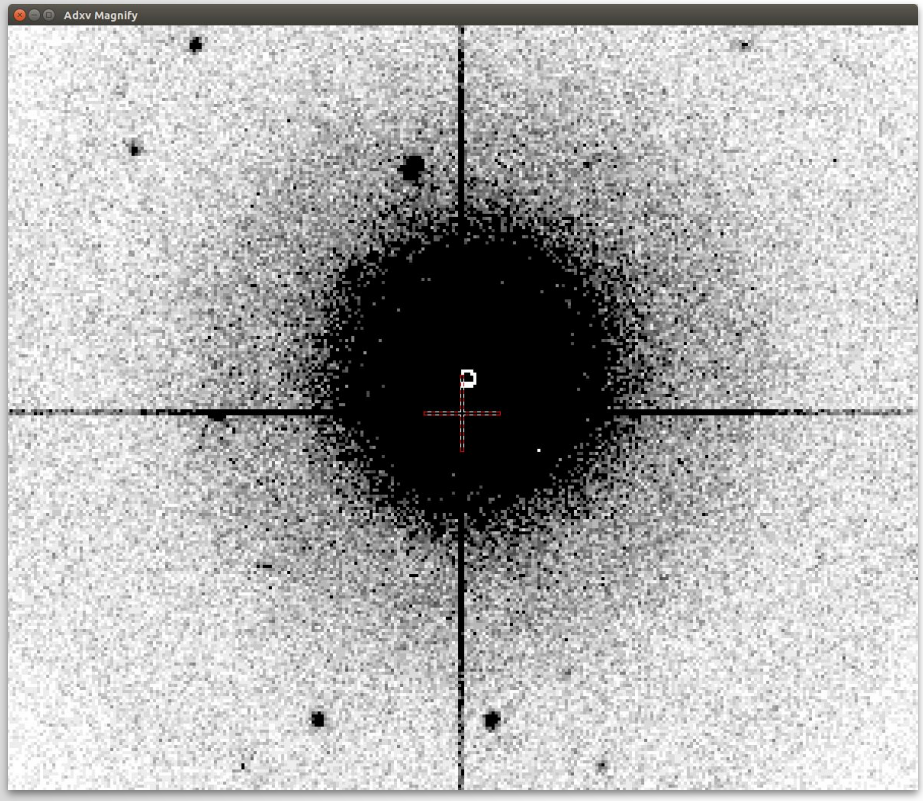
The biggest headache remains handling image formats!

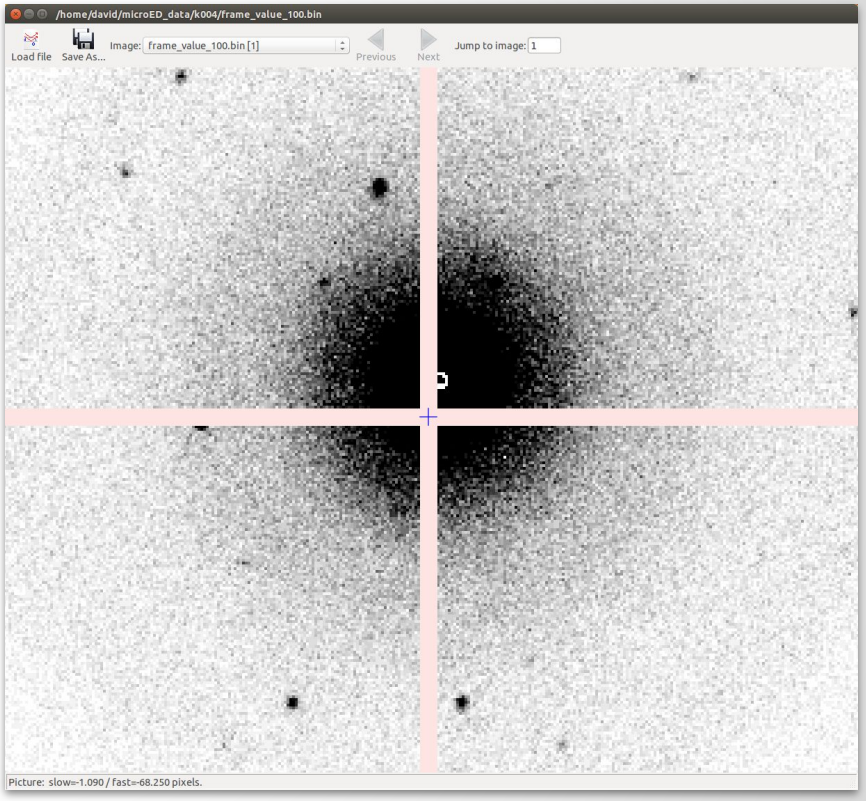




Importing data

Plug in a new dxtbx `Format` class to recognise new image file formats





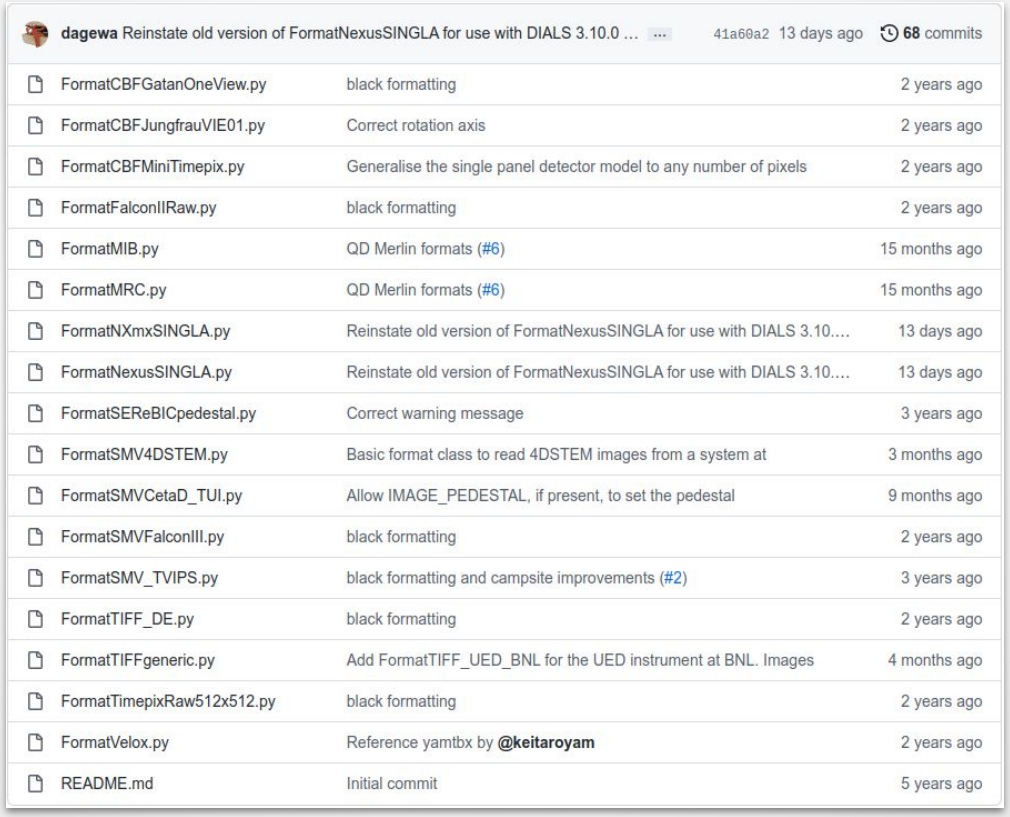
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Image format plugins

```
dxtbx.install_format --user [/path/to/format/class.py] [URL]
```

https://github.com/dials/dxtbx_ED_formats



dagewa Reinstate old version of FormatNexusSINGLA for use with DIALS 3.10.0 …	41a60a2	13 days ago	68 commits
FormatCBFGatanOneView.py	black formatting		2 years ago
FormatCBFJungfrauVIE01.py	Correct rotation axis		2 years ago
FormatCBFMiniTimepix.py	Generalise the single panel detector model to any number of pixels		2 years ago
FormatFalconIIRaw.py	black formatting		2 years ago
FormatMIB.py	QD Merlin formats (#6)		15 months ago
FormatMRC.py	QD Merlin formats (#6)		15 months ago
FormatNXmxSINGLA.py	Reinstate old version of FormatNexusSINGLA for use with DIALS 3.10....		13 days ago
FormatNexusSINGLA.py	Reinstate old version of FormatNexusSINGLA for use with DIALS 3.10....		13 days ago
FormatSEReBICpedestal.py	Correct warning message		3 years ago
FormatSMV4DSTEM.py	Basic format class to read 4DSTEM images from a system at		3 months ago
FormatSMVCetaD_TUI.py	Allow IMAGE_PEDESTAL, if present, to set the pedestal		9 months ago
FormatSMVFalconIII.py	black formatting		2 years ago
FormatSMV_TVIPS.py	black formatting and campsite improvements (#2)		3 years ago
FormatTIFF_DE.py	black formatting		2 years ago
FormatTIFFgeneric.py	Add FormatTIFF_UED_BNL for the UED instrument at BNL. Images		4 months ago
FormatTimepixRaw512x512.py	black formatting		2 years ago
FormatVelox.py	Reference yamtbx by **@keitaroyam**		2 years ago
README.md	Initial commit		5 years ago

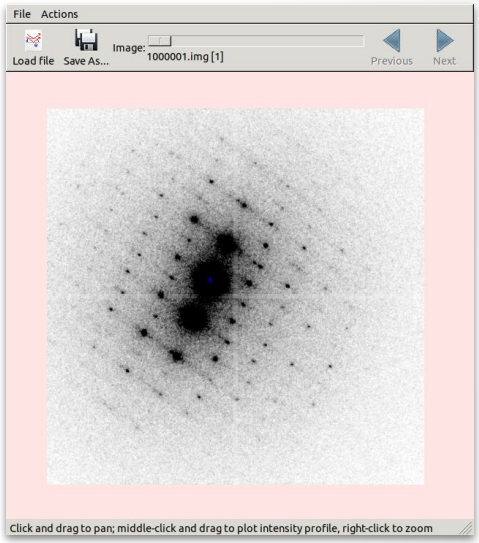
CCP4

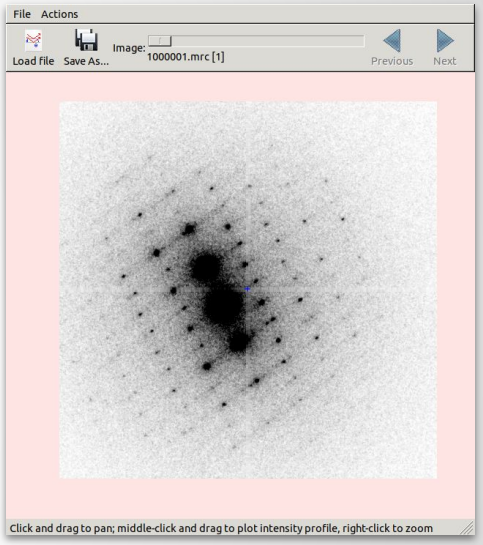
DIALS
Diffraction Integration for Advanced Light Sources

Data format standardisation

The format zoo

CAU-36 example





CCP4

DIALS
Diffraction Integration for Advanced Light Sources

A solution to the format woes?

- EM formats often lack important diffraction metadata
- This is a roadblock for autoprocessing
- Perhaps we can follow the path taken in the X-ray (MX) world:

  HDF5 ← NeXus ← NXmx

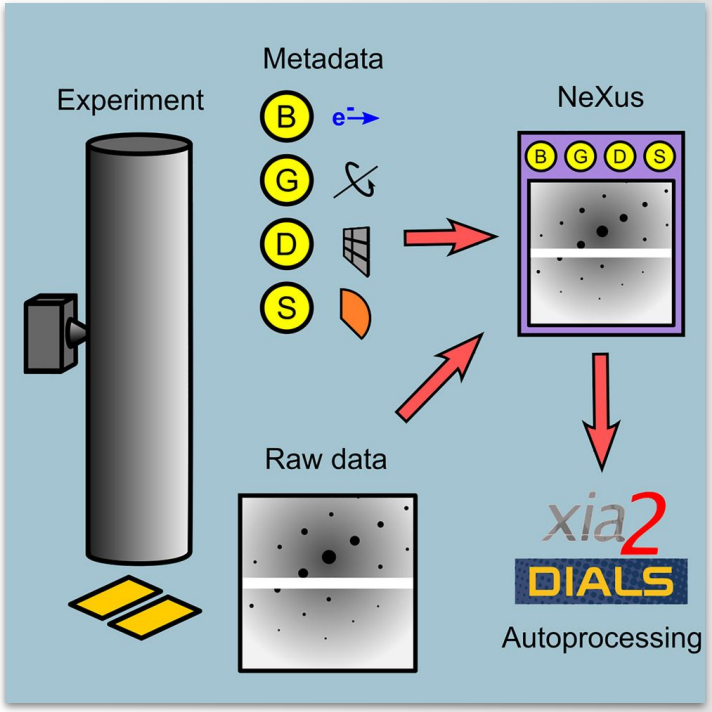
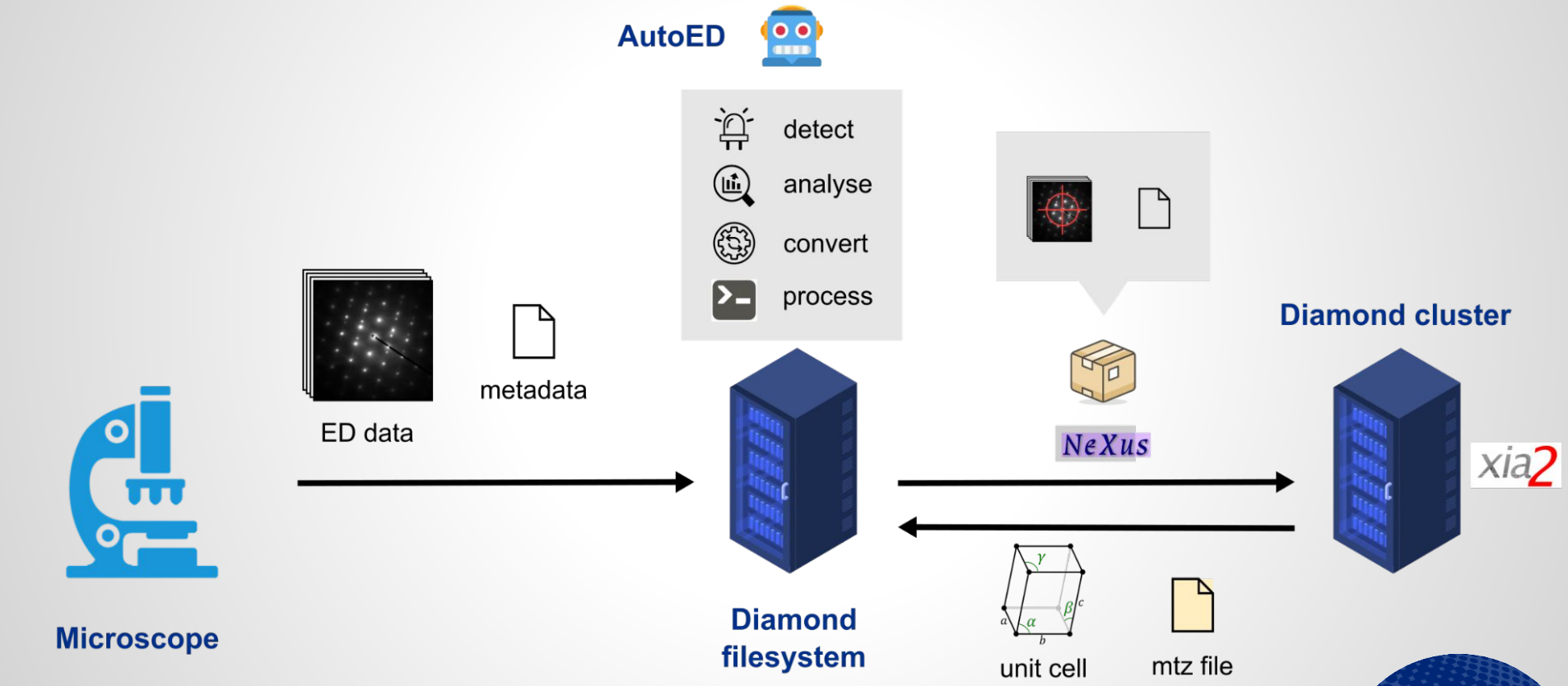


Figure from Waterman et al. (2023). A standard data format for 3DED/MicroED. Structure. 10.1016/j.str.2023.07.004.

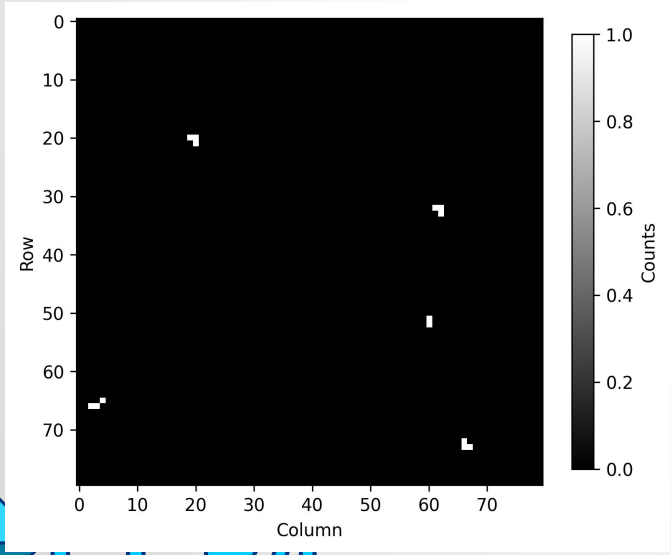
Automatic data processing at eBIC

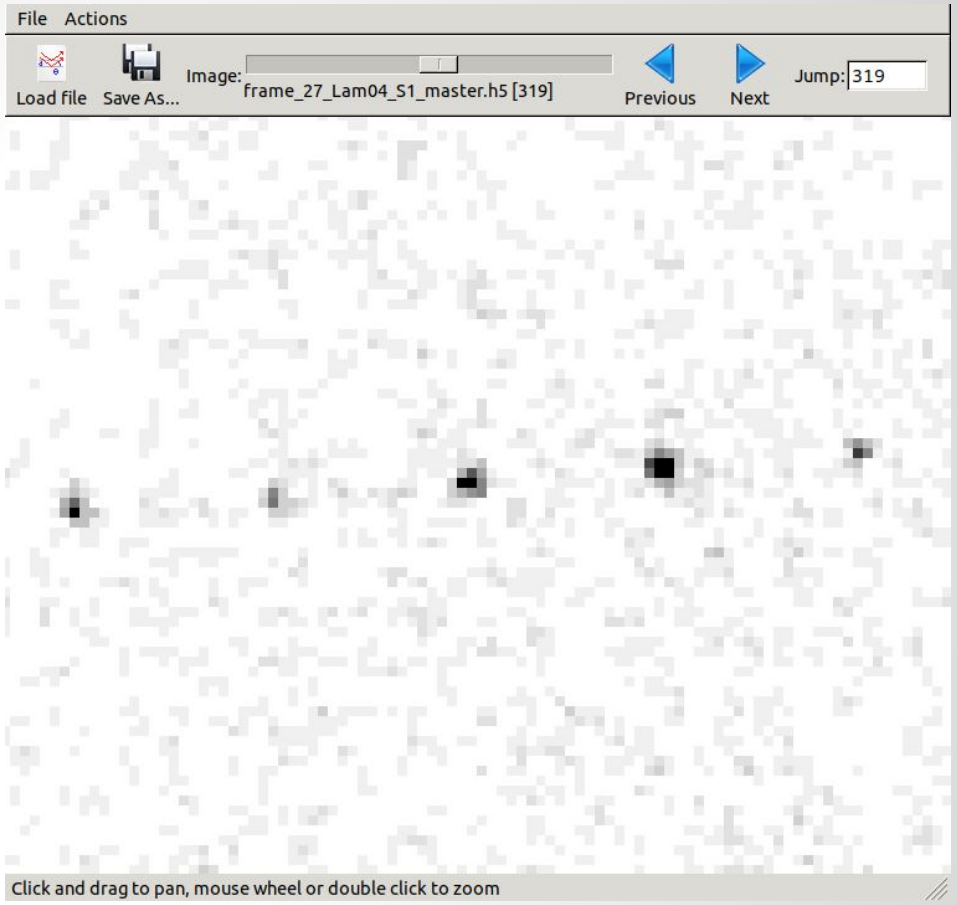


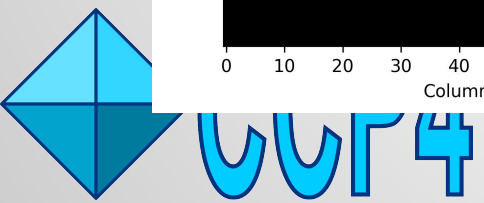
Spot-finding

- The default method for spot-finding is optimised for X-ray photon counting detectors (HPADs)
- This works well for HPADs used for electron diffraction too

…though they are *not* counters!



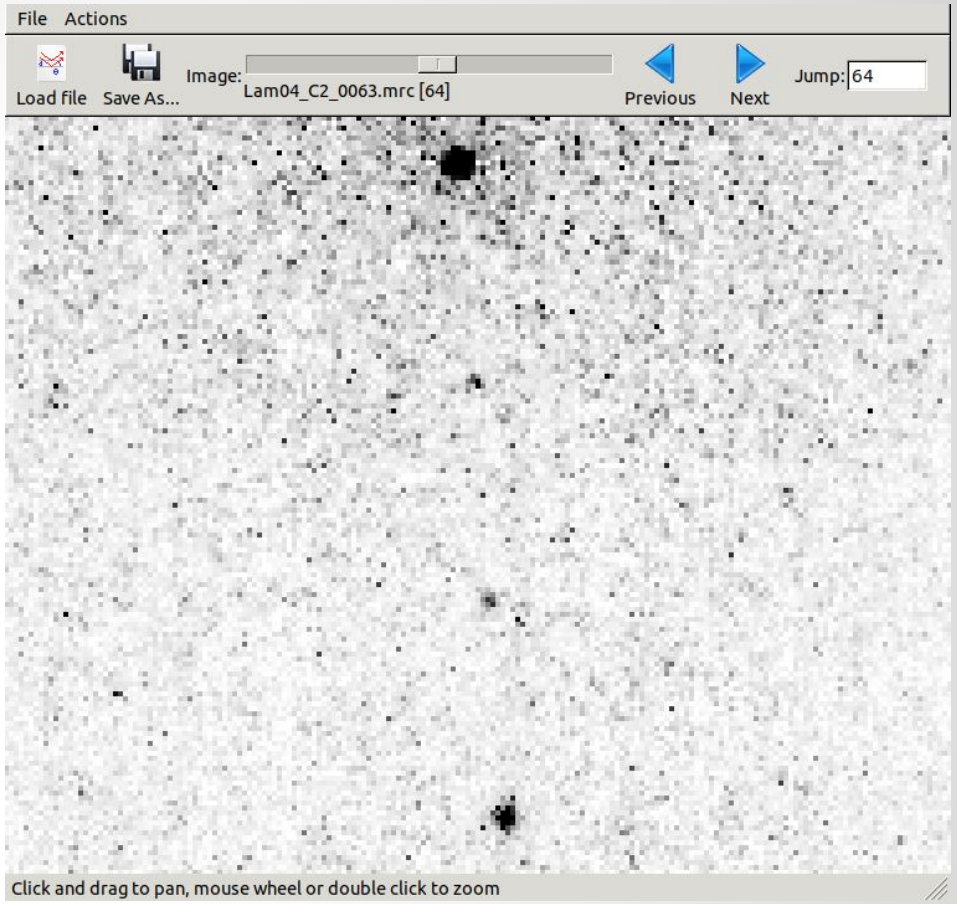




DIALS
Diffraction Integration for Advanced Light Sources

Spot-finding

- It does not work so well for integrating detectors with non-Poisson response and imperfect calibration

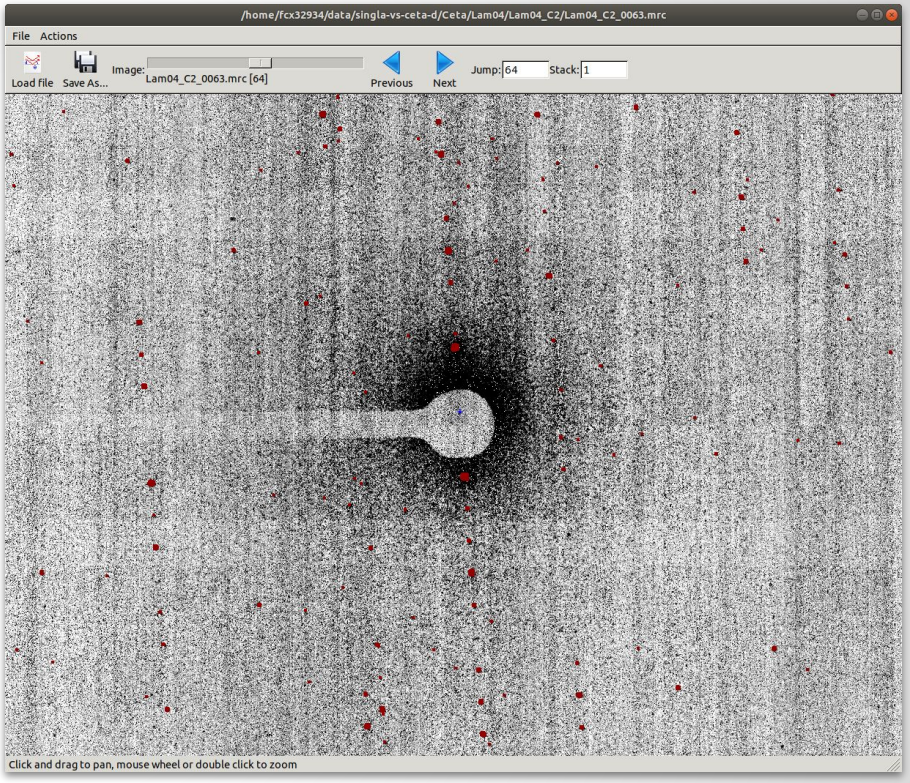


CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Spot-finding

- Zooming out, a lot of noise peaks are visible
- This makes indexing very difficult
- The `radial_profile` algorithm is designed to cope better in such cases
- It is based on the statistics of the observed pixel values, rather than a specific model for the detector response
- Explore settings interactively with the `dials.image_viewer`



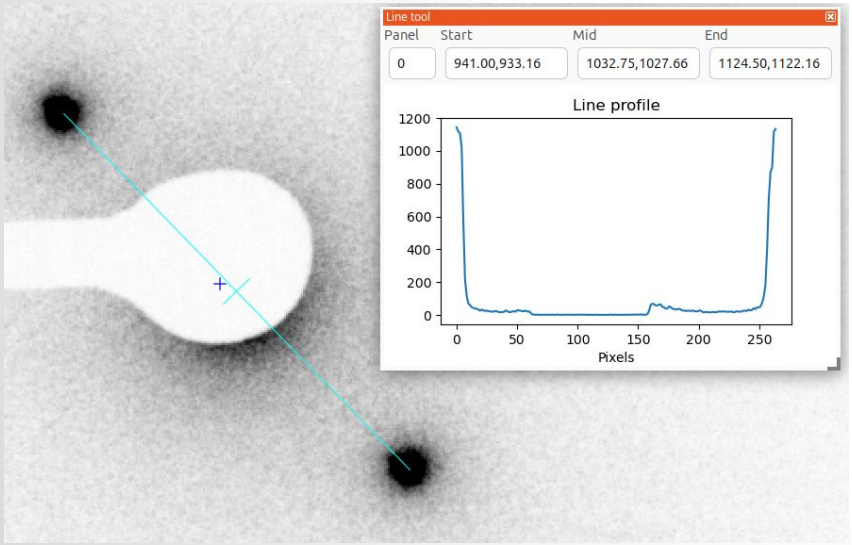
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

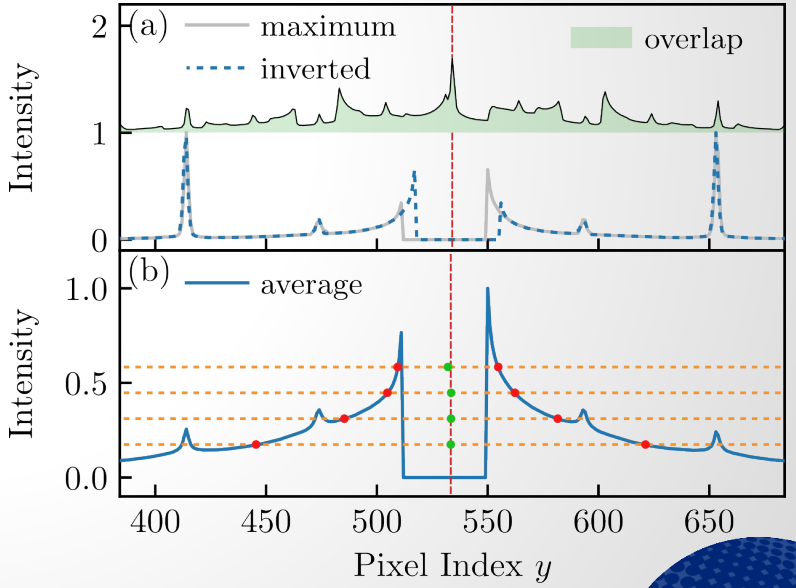
Beam centre determination

The problem is historically common in MX, though much less so at modern beamlines. However, with 3D ED the issue has become prevalent again!

Line tool in the image viewer: Interactively determine the bisection of a line between two points



Inversion and midpoint algorithms: designed for use with 3D ED data with either high background or visible Friedel pairs

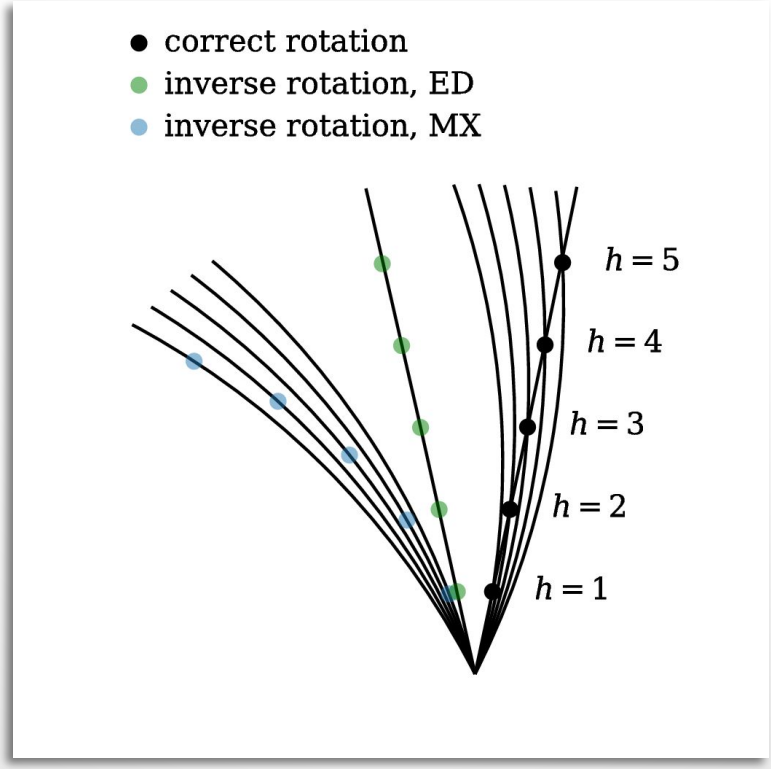


CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Rotation axis determination

- Not so straightforward as with an X-ray beamline!
- Rotation axis orientation varies with lens settings in the microscope
- Can't even easily distinguish between correct axis and its inversion

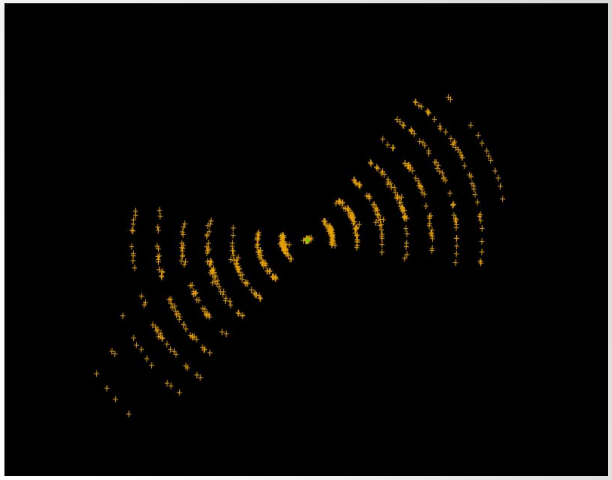


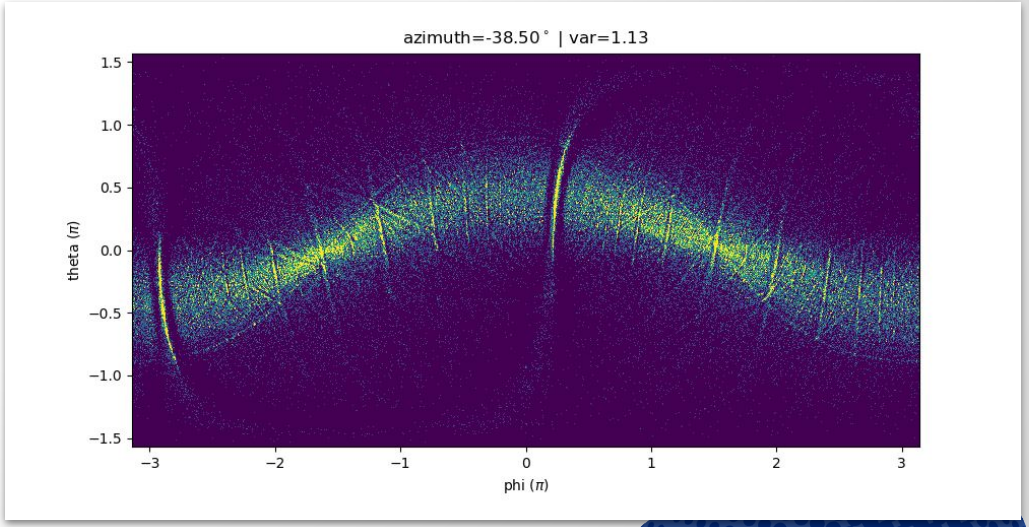
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Rotation axis determination

- An incorrect rotation axis makes it impossible to find the reciprocal lattice
- `dials.find_rotation_axis` performs a coarse and then fine search to locate the rotation axis with sub degree precision
- A cylindrical projection of difference vectors is used to assess each step



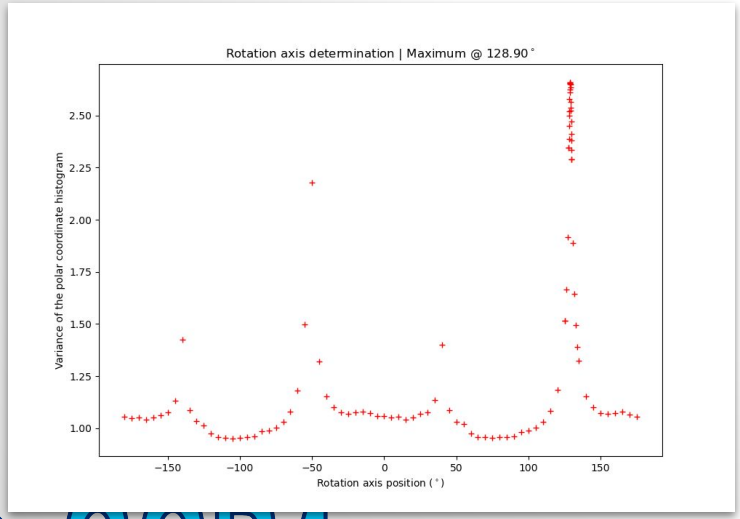


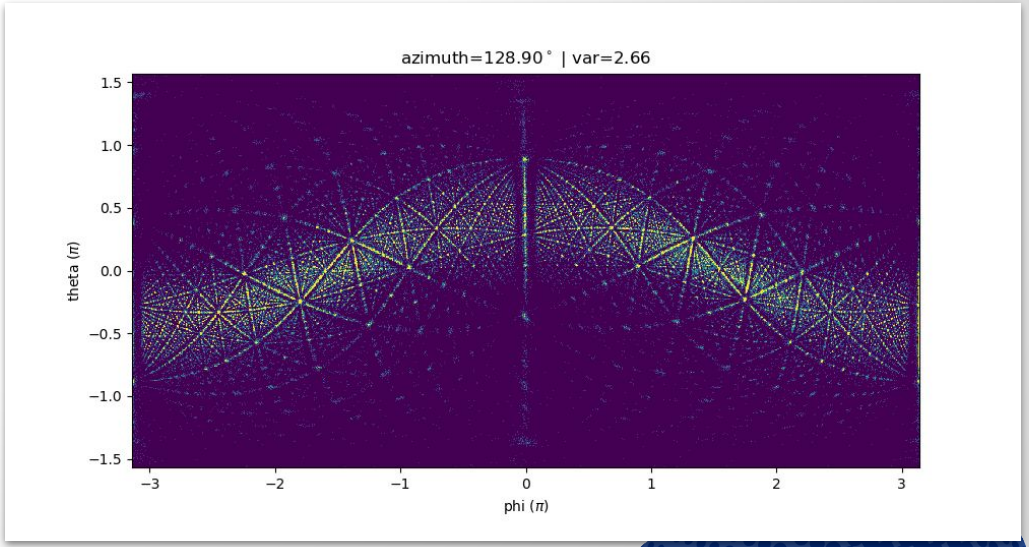
Algorithm described in Kolb *et al.* MRS Symp. Proc. Vol. 1184 (2009)

DIALS
Diffraction Integration for Advanced Light Sources

Rotation axis determination

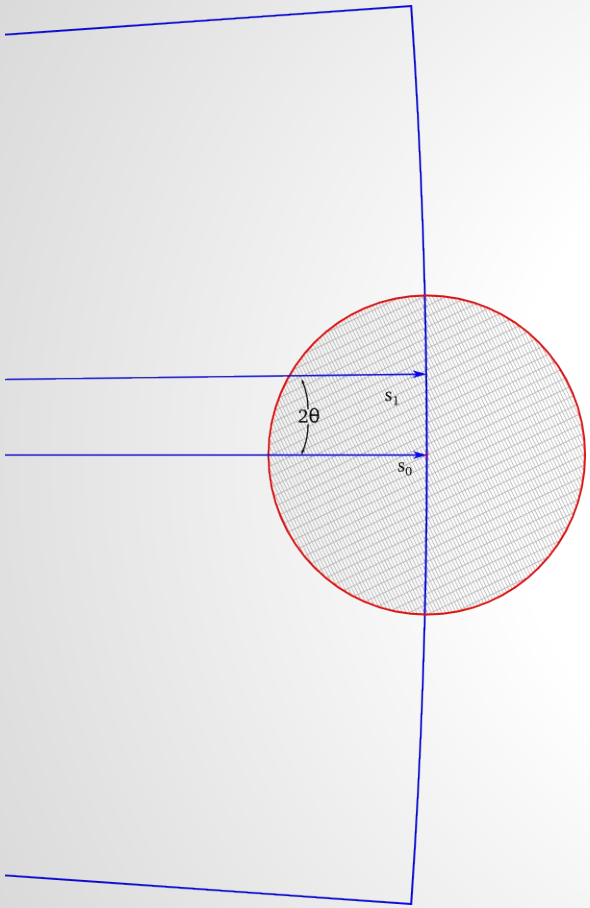
- An incorrect rotation axis makes it impossible to find the reciprocal lattice
- `dials.find_rotation_axis` performs a coarse and then fine search to locate the rotation axis with sub degree precision
- A cylindrical projection of difference vectors is used to assess each step





Algorithm described in Kolb *et al.* MRS Symp. Proc. Vol. 1184 (2009)

Refinement of detector distance



- In MicroED the Ewald sphere is flat
- It is usually impossible to refine the effective detector distance and the unit cell simultaneously
- Typically, fix detector distance throughout processing
- Rely on good calibration, though hard to get <1% error
- Leads to error being expressed in the unit cell parameters

CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Restrained refinement

Flat Ewald sphere → extremely high correlation between cell and detector distance

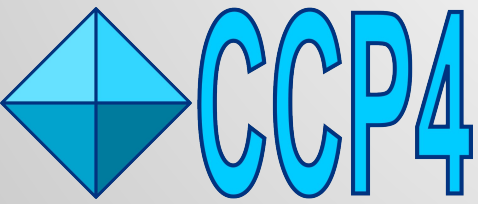
Fix detector distance to properly calibrated value...

...or if the cell is known, provide a restraint to the target unit cell

```
refinement.parameterisation.crystal.unit_cell
{
  restraints
  {
    tie_to_target
    {
      values=99.06,31.01,54.30,90,107.7,90
      sigmas=0.01,0.01,0.01,0.01,0.01,0.01
    }
  }
}
```

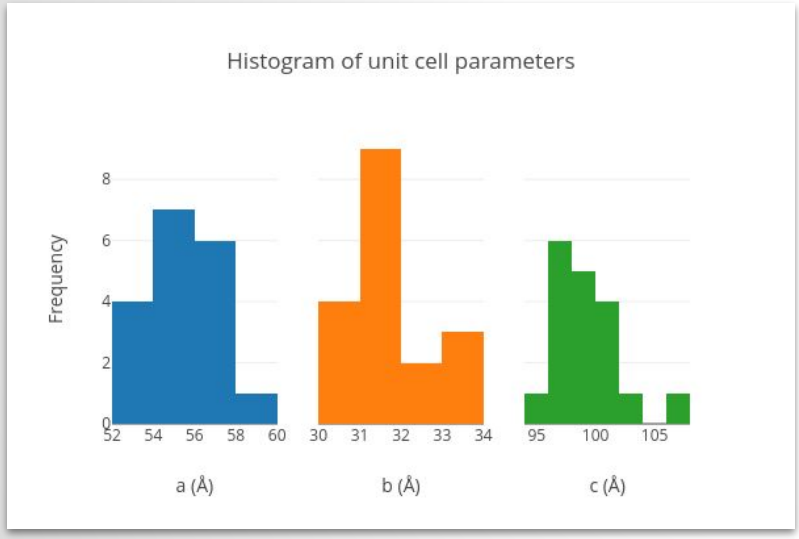
$$R_a = \frac{(a - a_t)^2}{\sigma_a^2}$$

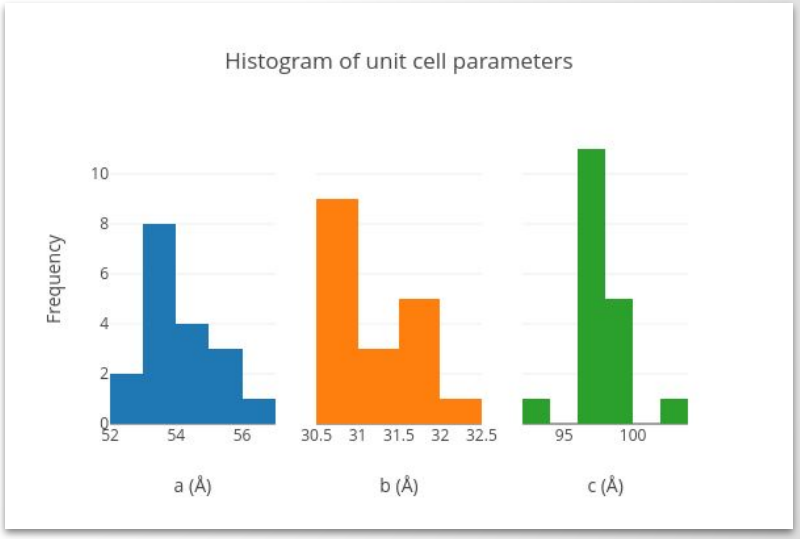
https://dials.github.io/documentation/tutorials/3DED/MyD88.html



Effect on unit cell clustering

- Restraints tighten up the distribution of unit cells
- This affects the results of unit cell clustering in programs such as BLEND, or dials.cosym
- Is it a better description of the population of cells?



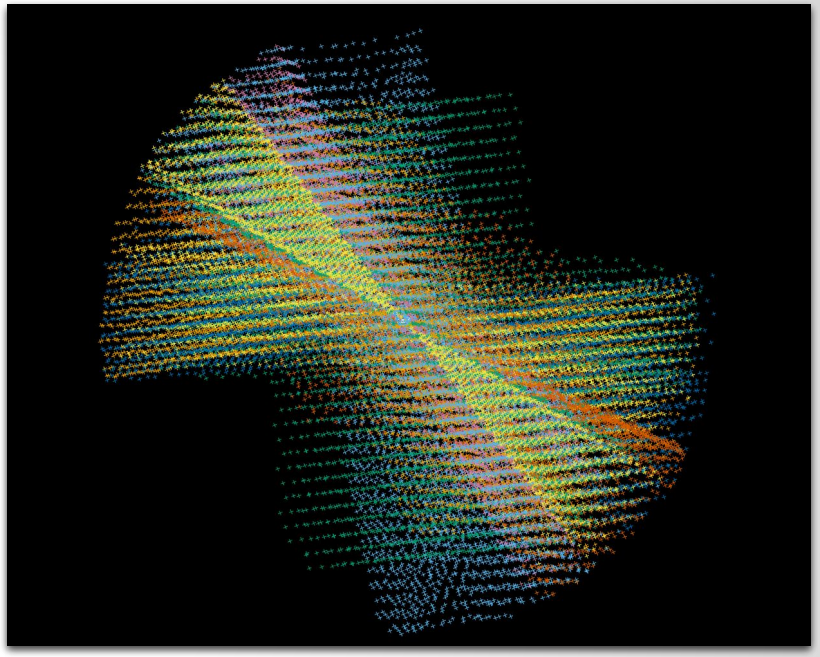


CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Multi-crystal methods

- DIALS can handle multiple crystals simultaneously
- Multiple lattice indexing is often useful for nanocrystal data
- Joint refinement improves unit cell estimation
- Joint scaling helps improve data completeness and averages out systematic errors
- The reciprocal lattice viewer can give a rough idea of completeness prior to scaling



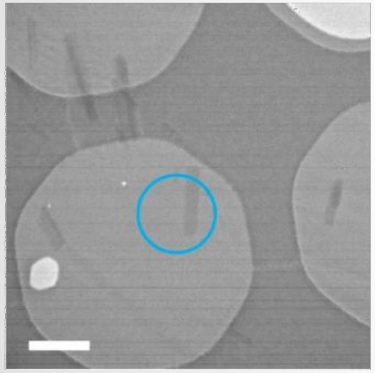
☑ Show in crystal frame

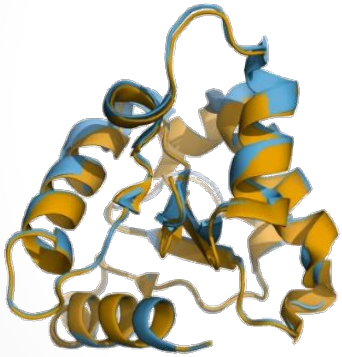
CCP4

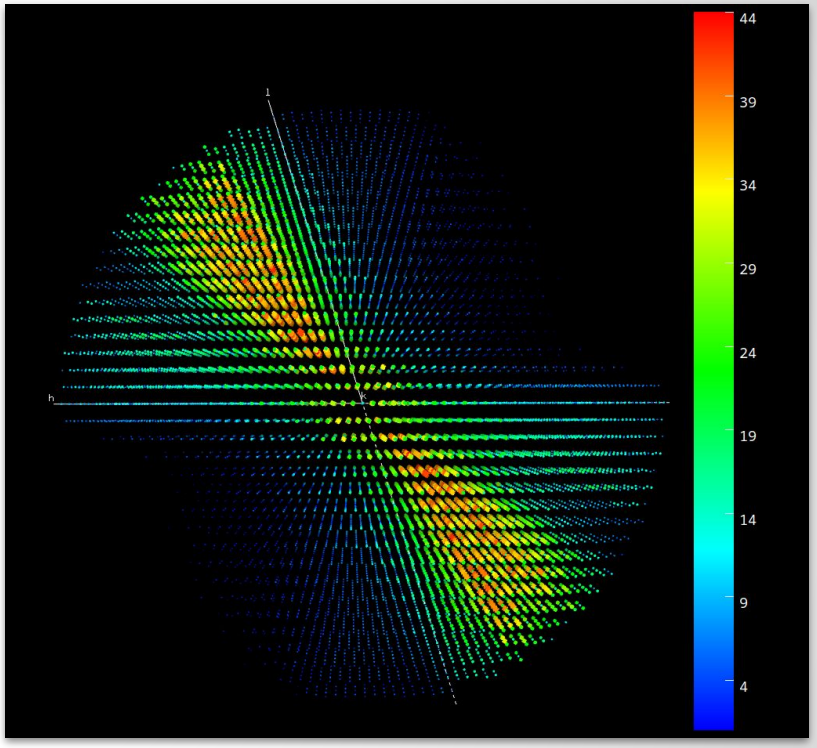
DIALS
Diffraction Integration for Advanced Light Sources

Multi-crystal methods

- After joint scaling, multiplicity can be assessed with the `cctbx.multiplicity_viewer`







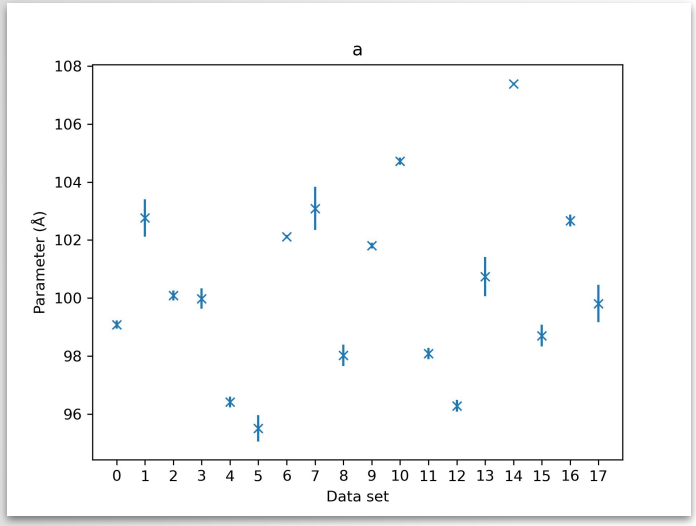
MyD88^{TIR} figures adapted from Clabbers et al. (2021). Nat Commun 12, 2578. 10.1038/s41467-021-22590-6.

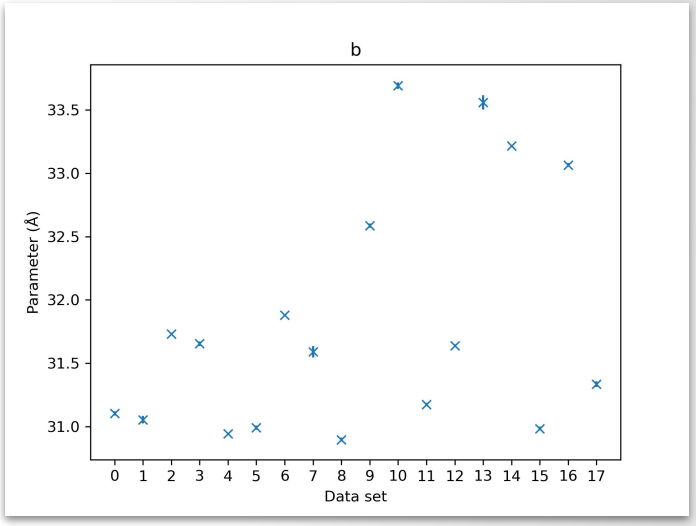
CCP4

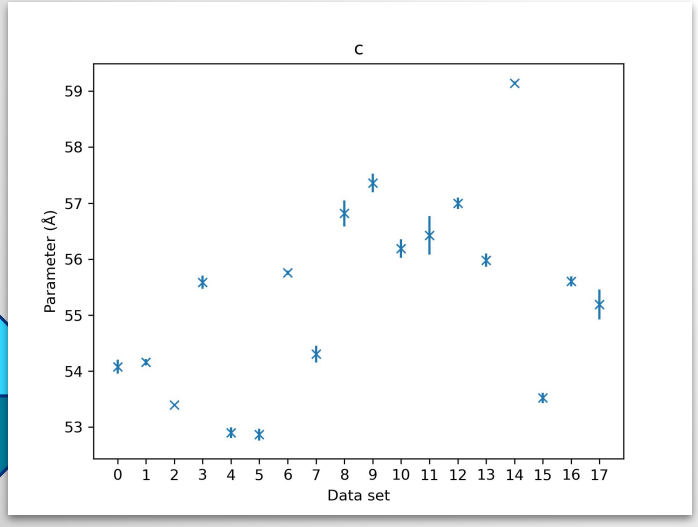
DIALS
Diffraction Integration for Advanced Light Sources

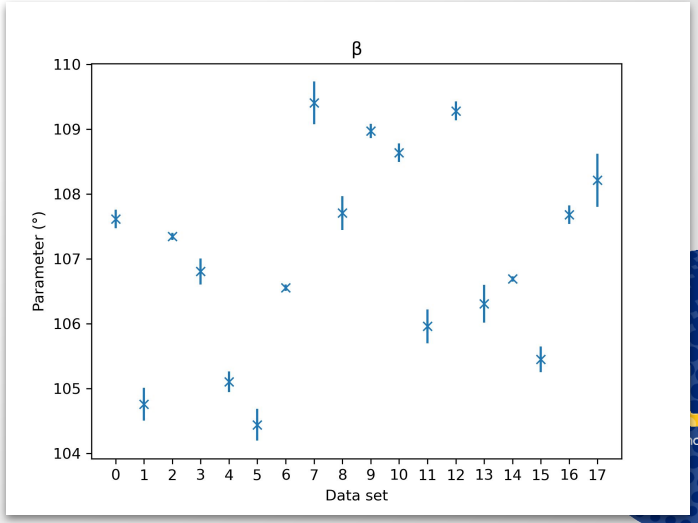
Multi-crystal unit cell refinement

Independent unit cell variation can be large



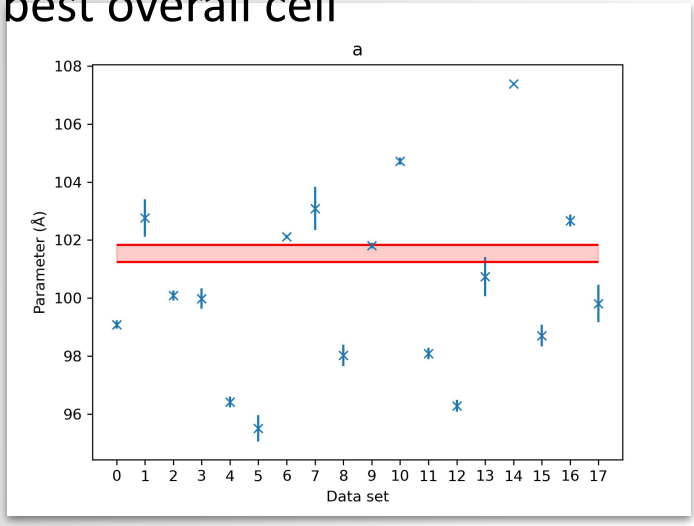


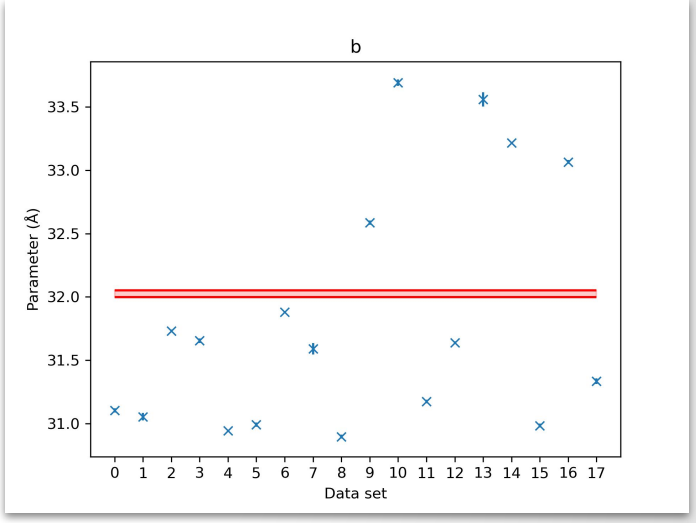


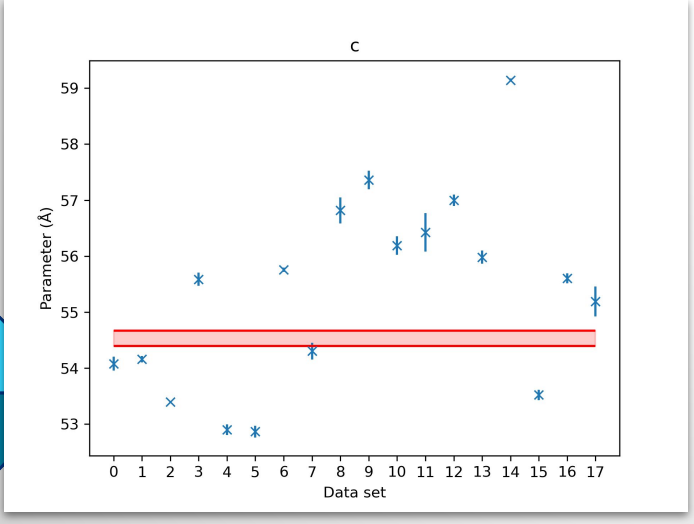


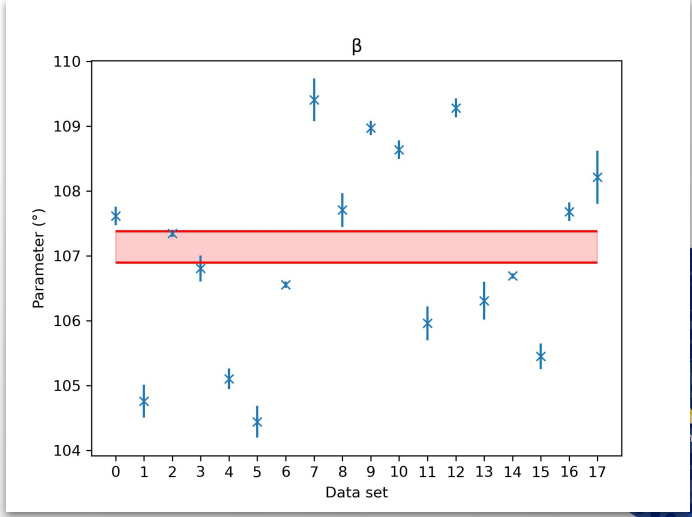
Multi-crystal unit cell refinement

Joint refinement against 2θ target (`dials.two_theta_refine`) provides a best overall cell



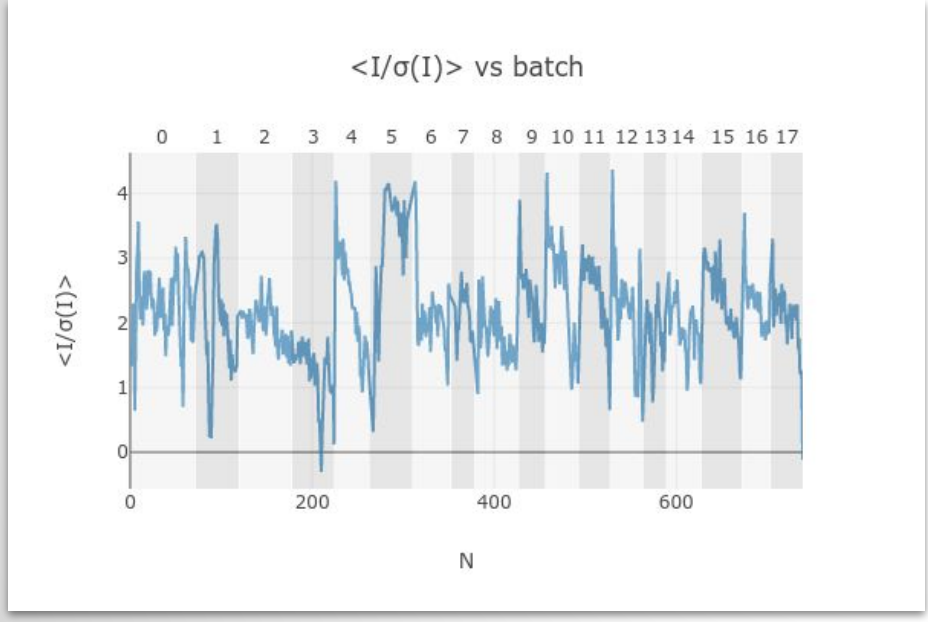


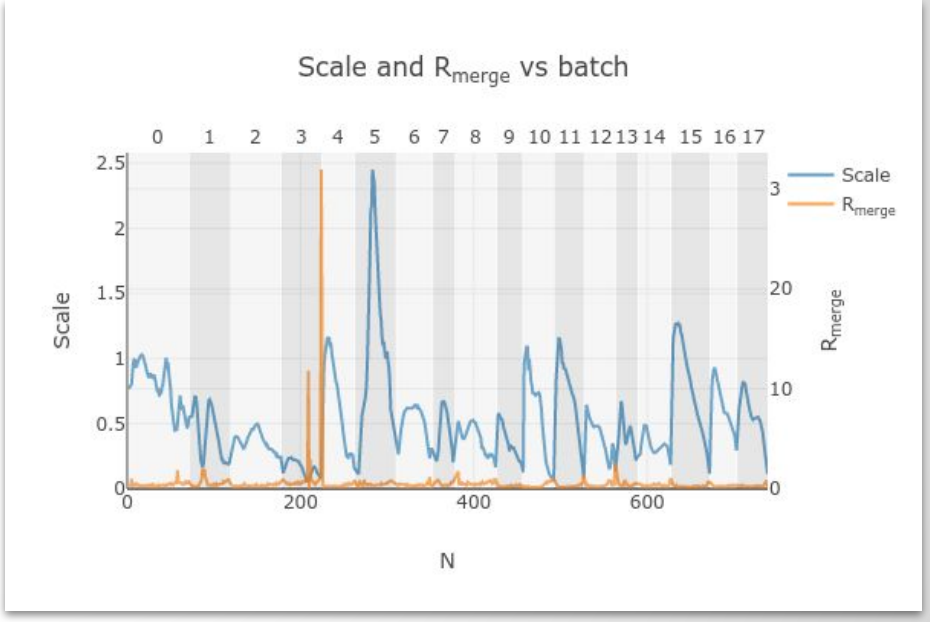




Multi-crystal scaling with dials.scale

- Rapid re-scaling when adding or subtracting individual datasets
- Automated adjustment of parameters, reflection selection
- Successive cycles of scaling and filtering using $CC_{1/2}$



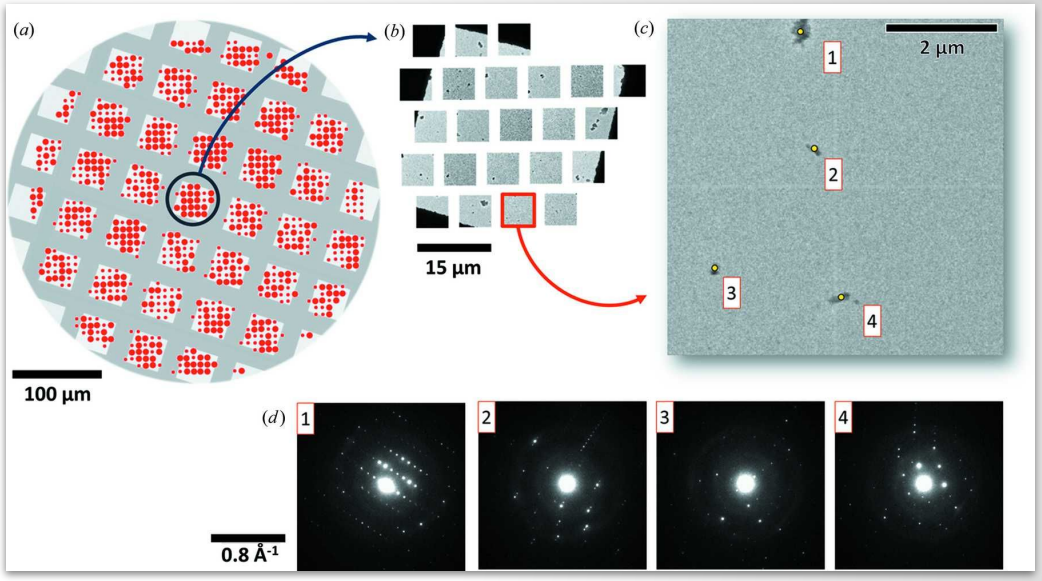


CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Serial ED

- Beam-sensitive samples
- More materially (and shot) efficient than XFEL
- Higher resolution than conventional 3DED
- Time-resolved studies via UED
- Various modalities: snapshot, discrete tilts, narrow rotation wedges



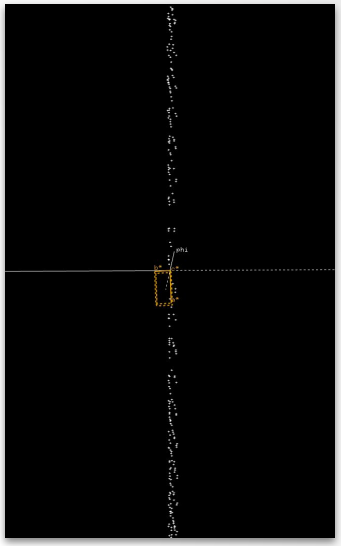
Smeets *et al.* (2018) J. Appl. Cryst. **51**

CCP4

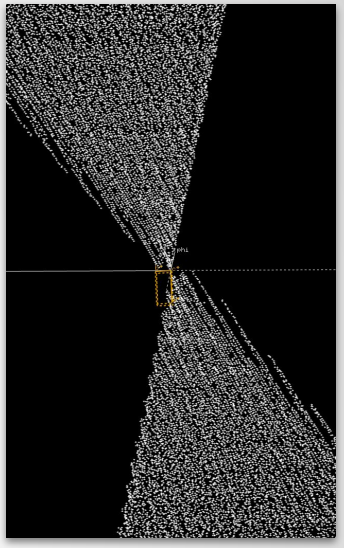
DIALS
Diffraction Integration for Advanced Light Sources

Serial ED indexing

- Each pattern has only 2D information
- Determination of 3 basis vectors from a single pattern is not generally possible
- Using *prior knowledge of the unit cell* we can fit the best orientation of that cell to the pattern, where we have two observed directions and one inferred
- 3 methods in DIALS can be used for this:
  - `real_space_grid_search`
  - `low_res_spot_match`
  - `pink_indexer`



single image



56° wedge

CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Example: triclinic lysozyme, Glacios + SINGLA

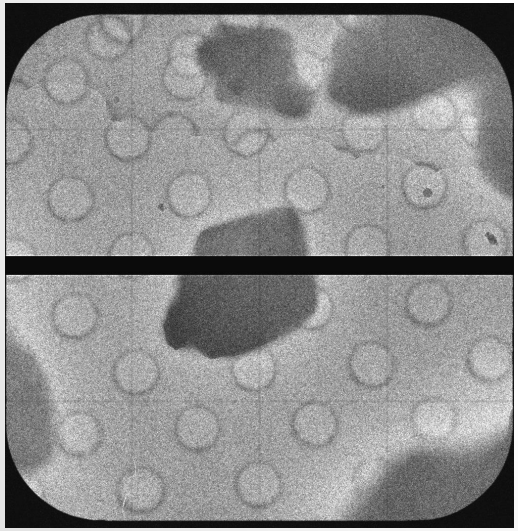
a = 25.20, b = 30.90, c = 33.27

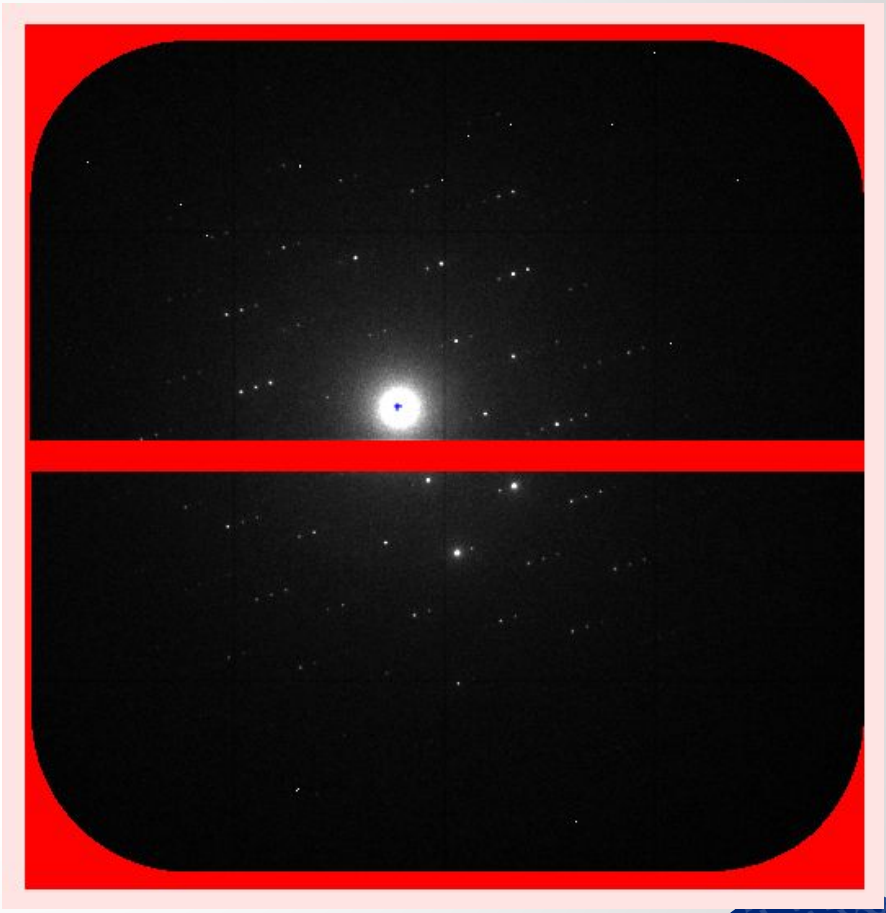
α = 88.06, β = 71.04, γ = 69.69

$P1$

200 keV (0.02508 Å)

Unit cell restraint required

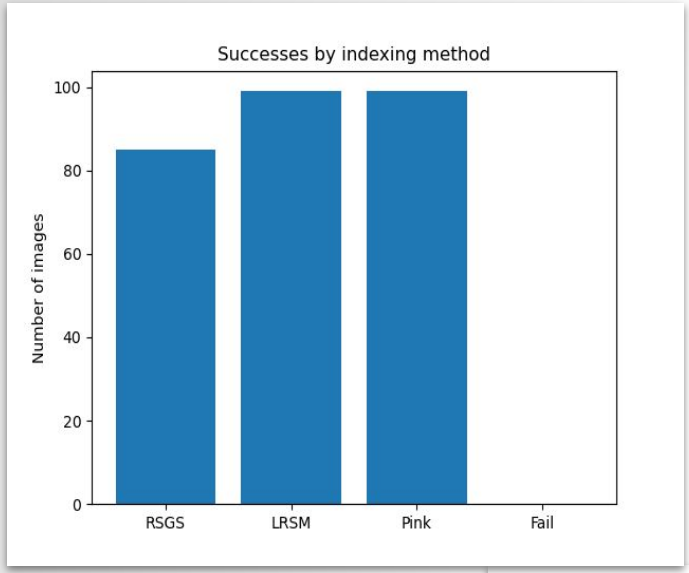


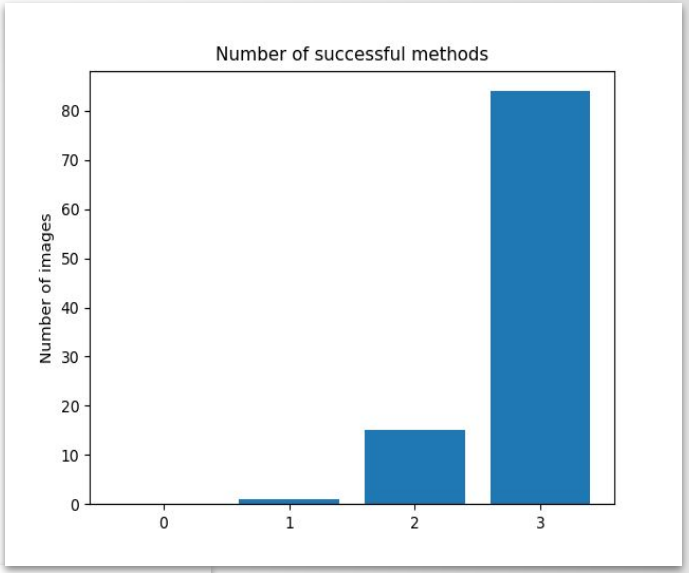


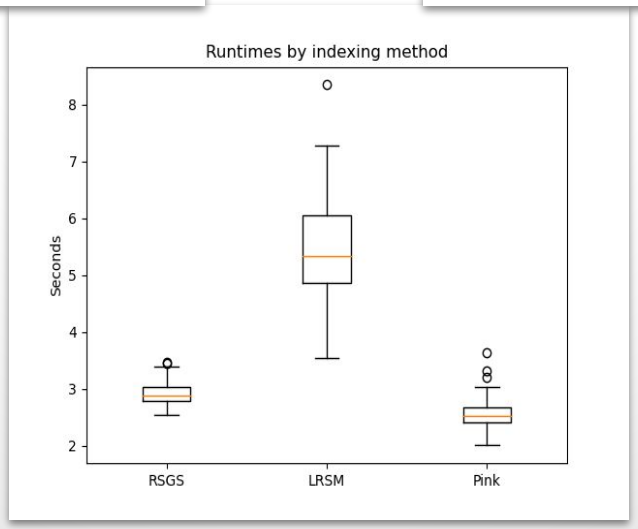
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Example: triclinic lysozyme on Glacios







CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Acknowledgements

All past and present DIALS developers for the framework

Data and discussions with many users and developers of MicroED / 3DED worldwide. In particular, for this talk I thank:

Tim Gruene and Zou lab for Timepix zeolite data

Max Clabbers, Hongyi Xu and Xiaodong Zou for the MyD88^{TIR} tutorial data

David Owen at DLS/eBIC for SINGLA and Ceta-D lamella data

Lei Wang and David Owen for the triclinic lysozyme data

Robert Bücker provided inspiration for the `radial_profile` spot-finder

Stef Smeets, Lukáš Palatinus and Ute Kolb for rotation axis search algorithm

Marko Petrovic wrote AutoED and new beam centre determination methods

Funding: