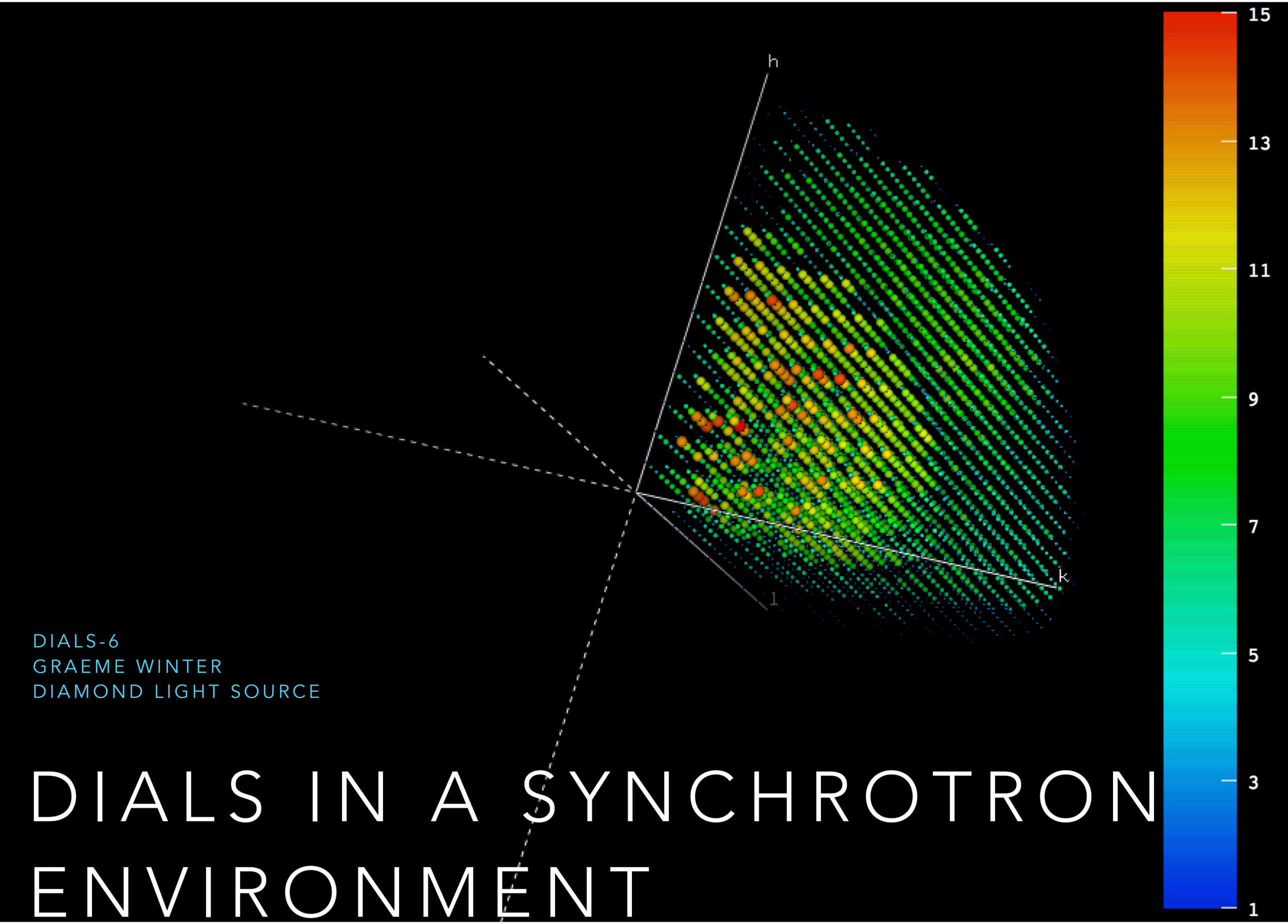


DIALS-6
GRAEME WINTER
DIAMOND LIGHT SOURCE

DIALS IN A SYNCHROTRON ENVIRONMENT

PROJECT SUPPORT

DIALS
Diffraction Integration for Advanced Light Sources

diamond

BERKELEY LAB
Lawrence Berkeley National Laboratory

CCP4

BioStructX

Science & Technology Facilities Council

Research Complex at Harwell

ACKNOWLEDGEMENTS

research papers

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

XDS

Wolfgang Kabsch

Max-Planck-Institut für Medizinische Forschung, Abteilung Biophysik, Jahnstrasse 29, 69120 Heidelberg, Germany

Correspondence e-mail: wolfgang.kabsch@mpimf-heidelberg.mpg.de

Received 19 August 2009
Accepted 9 November 2009

The usage and control of recent modifications of the program package *XDS* for the processing of rotation images are described in the context
include automatic detern
range and recognition a
Moreover, the limitation
number of correction/scal
pixel contents have been
been restructured for pa
and completeness of colle
measurement.

1. Functional specifica

The program package *XL*
developed for the reduct
recorded on a planar det
monochromatic X-rays. I
XDS accepts a sequ
rotation images from a v
and multiwire area dete
metrics and produces a li
of the reflections occurrin
way. The program assum
positive amount of crysta
incident beam and crystal
imposes no limitations
directions of the rotatior
oscillation range covered

research papers

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The finer things in X-ray diffraction data collection

J. W. Pflugrath

Molecular Structure Corporation, 9009 New Trails Drive, The Woodlands, TX 77381, USA

Correspondence e-mail: jwp@msc.com

X-ray diffraction images
sensitive detectors can b
depending on whether the
is greater than or less than
The expectations and cons
and thin images in terms
X-ray background and *I*
software suite for proces
introduced, and results fr
those from another popula

1. Introduction

Two-dimensional position-
for many years in X-ray di
cular, data from crystals of
oligonucleotides and thei
acquired with an area det
obsolete), a multi-wire s
recently commercialized
coupled to a phosphor-coat
detectors, the crystal, cente
oscillated around a single a
~2.0°, while counts from d
for a specified time. At t
detector is read out and th
two-dimensional array with
to a distinct position on

research papers

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The integration of macromolecular diffraction data

Andrew G. W. Leslie

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

Correspondence e-mail: andrew@mrc-lmb.cam.ac.uk

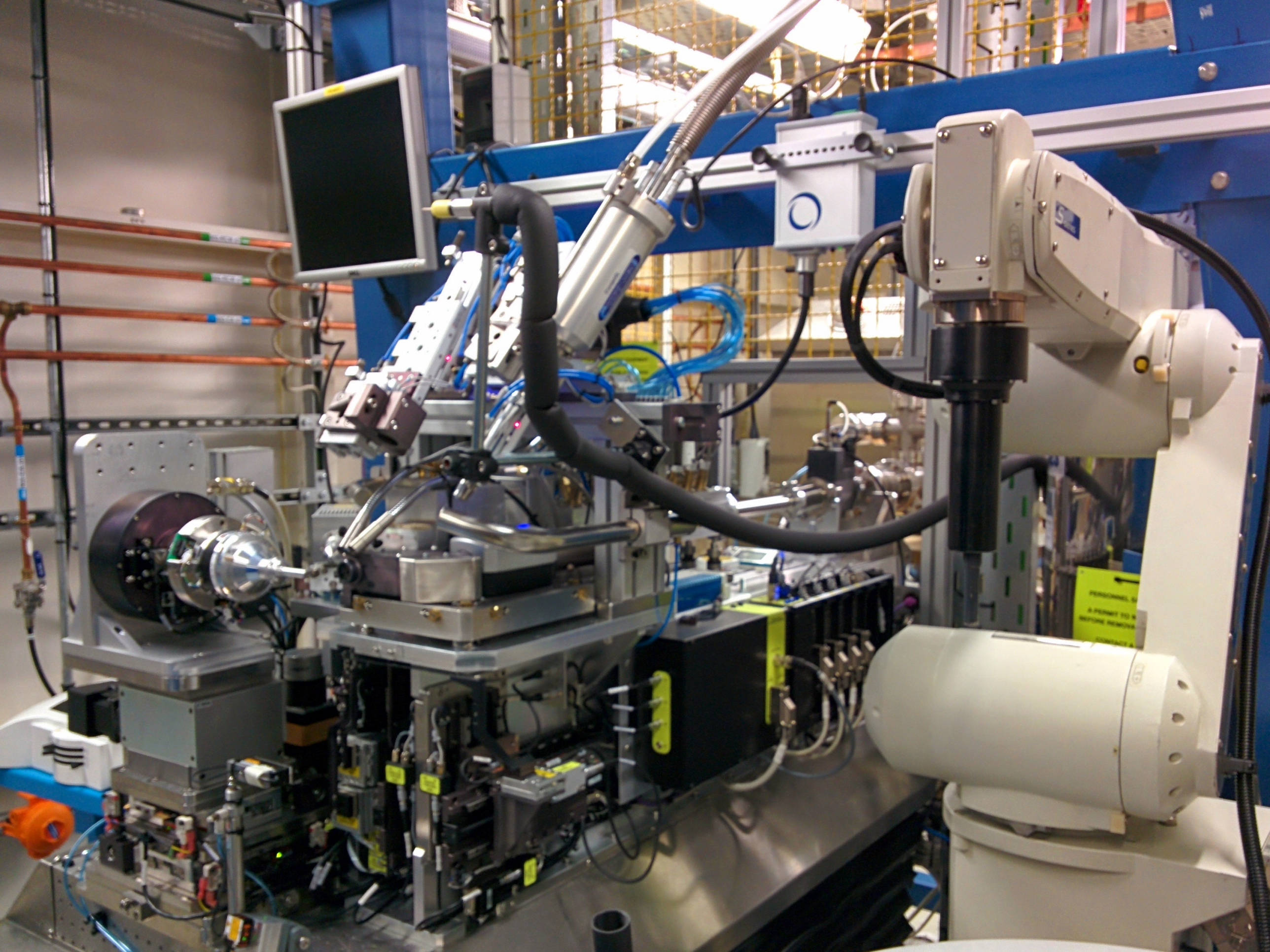
Received 19 May 2005
Accepted 24 November 2005

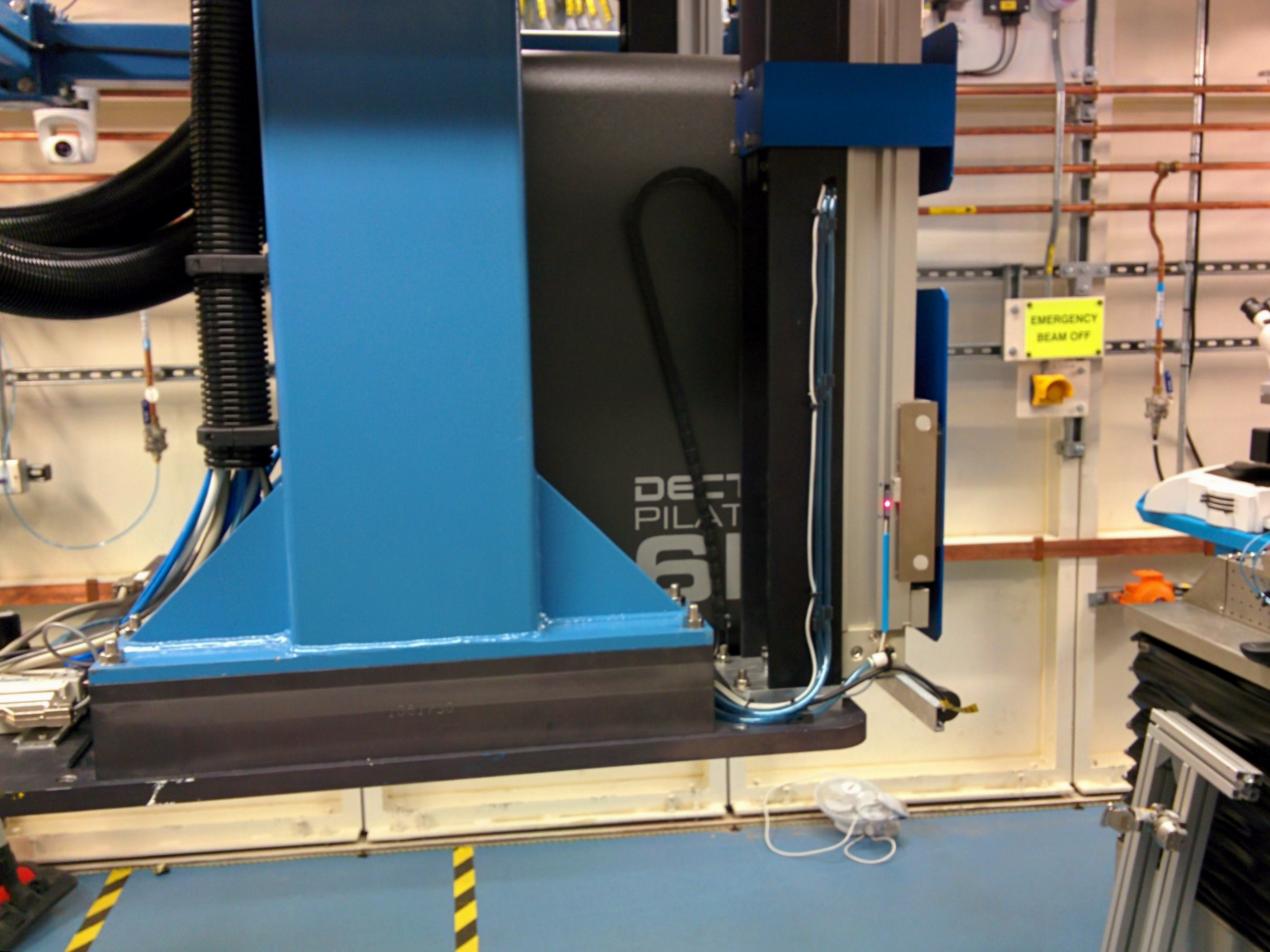
The objective of any modern data-processing program is to produce from a set of diffraction images a set of indices (*hkl*s) with their associated intensities (and estimates of their uncertainties), together with an accurate estimate of the crystal unit-cell parameters. This procedure should not only be reliable, but should involve an absolute minimum of user intervention. The process can be conveniently divided into three stages. The first (autoindexing) determines the unit-cell parameters and the orientation of the crystal. The unit-cell parameters may indicate the likely Laue group of the crystal. The second step is to refine the initial estimate of the unit-cell parameters and also the crystal mosaicity using a procedure known as post-refinement. The third step is to integrate the images, which consists of predicting the positions of the Bragg reflections on each image and obtaining an estimate of the intensity of each reflection and its uncertainty. This is carried out while simultaneously refining various detector and crystal parameters. Basic features of the algorithms employed for each of these three separate steps are described, principally with reference to the program *MOSFLM*.

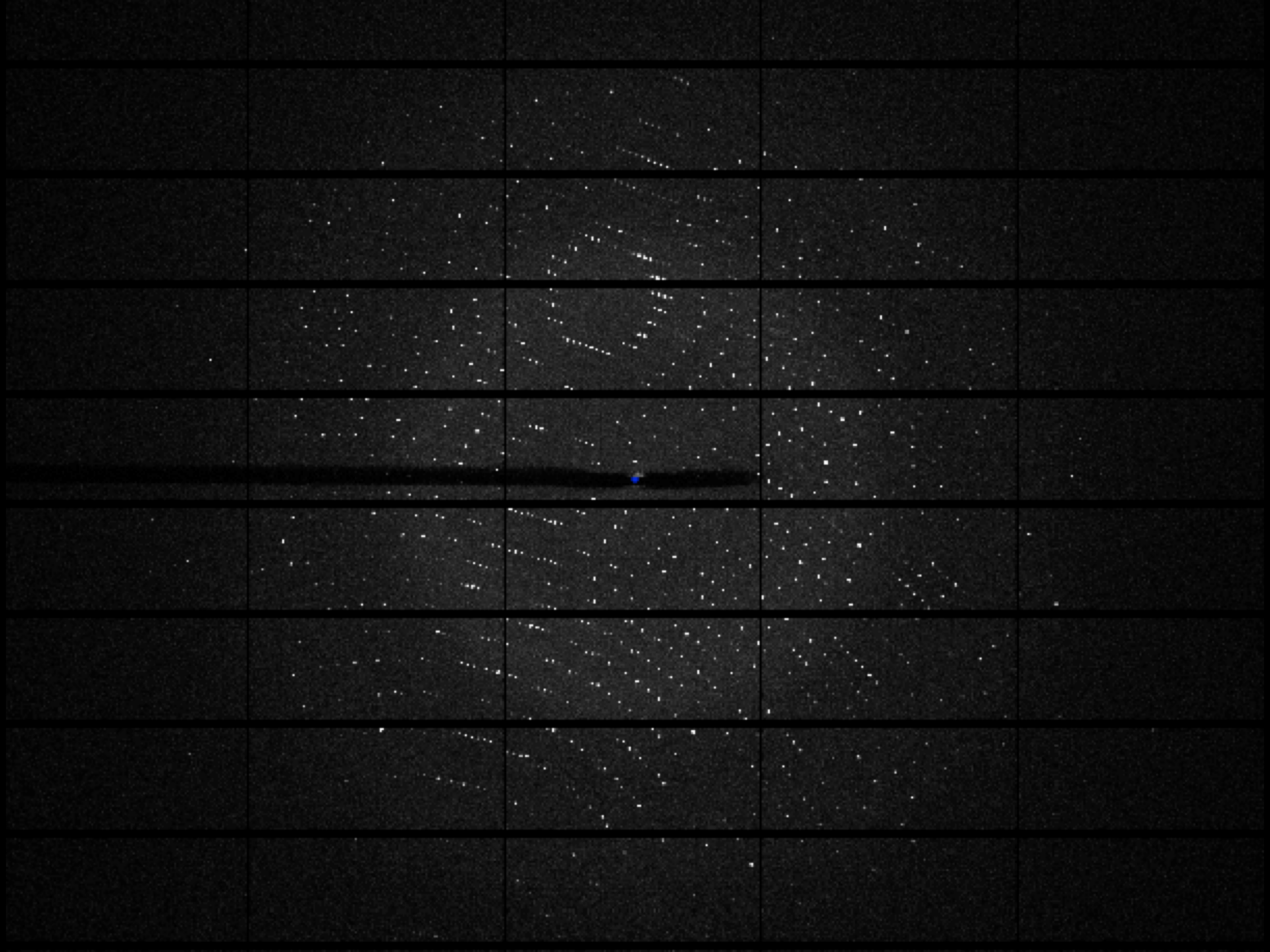
1. Introduction

The collection of macromolecular diffraction data has undergone dramatic advances during the last 15 years with the advent of two-dimensional area detectors such as image plates and CCDs, crystal cryocooling and the availability of intense, monochromatic and highly collimated X-ray beams from

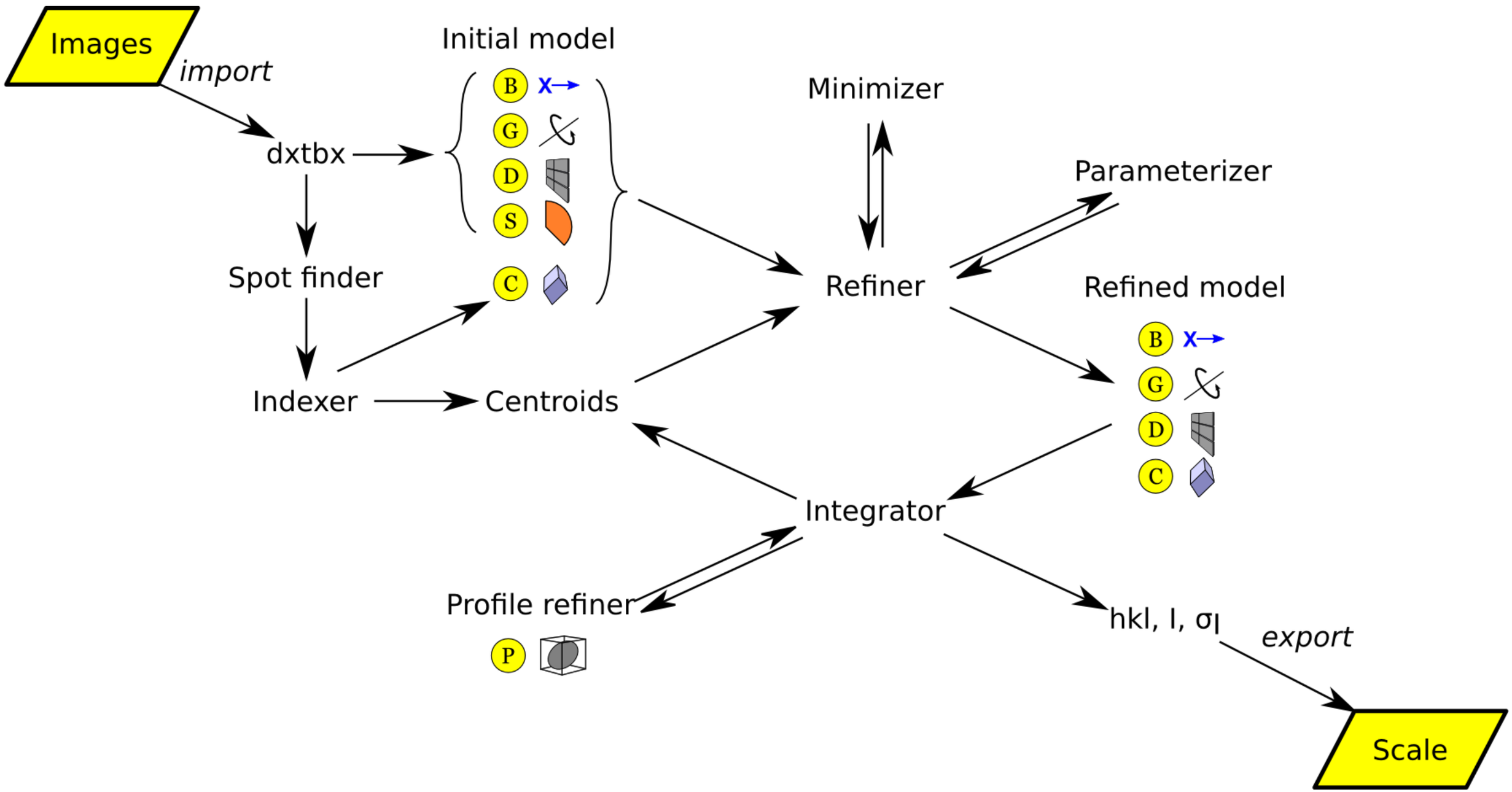
DIALS TOOLBOX




EMERGENCY
BEAM OFF



DIALS


Images
import
dxtbx
Spot finder
Indexer
Centroids
Initial model
B
G
D
S
C
Minimizer
Refiner
Parameterizer
Refined model
B
G
D
C
Integrator
Profile refiner
P
hkl, I, σI
export
Scale

MAIN DIALS PROGRAMS

- dials.import
- dials.find_spots - see Richard's talk
- dials.index - see Richard's talk
- dials.refine_bravais_settings - see Richard's talk
- dials.refine - see David's talk
- dials.integrate - see James's talk
- dials.export_mtz
- (POINTLESS/AIMLESS)

PRINCIPLE USER INTERFACE

example — gw56@cs03r-sc-serv-16:/dls/i04/data/2015/nr12190-10/20150...

```
Finding spots in image 1 to 540...
Extracting strong pixels from images (may take a while)
Extracted strong pixels from images
Merging 8 pixel lists
Merged 8 pixel lists with 922120 pixels
Extracting spots
Extracted 219125 spots
Calculating 219125 spot centroids
Calculated 219125 spot centroids
Calculating 219125 spot intensities
Calculated 219125 spot intensities
Found 1 possible hot spots
Found 1 possible hot pixel(s)
Filtering 219125 spots by number of pixels
Filtered 116321 spots by number of pixels
Filtering 116321 spots by peak-centroid distance
Filtered 116082 spots by peak-centroid distance

---------------------------------------------------------------------------
Saving 116082 reflections to strong.pickle
Saved 116082 reflections to strong.pickle
Time Taken: 21.364578
Graemes-MacBook-Pro-3:example graeme$
```

USEFUL FOR AUTOMATION

- Well defined step by step analysis process
- Cleanly defined command-line interface
- Comprehensive range of capabilities
- Python-compatible data files for results etc.
- Robust, tested

LESS USEFUL FOR AUTOMATION

- Step by step analysis process => lots of work to integrate
- Does not include “intelligent decision making”
- Does not (really) include scaling
- Does not consider data sets with > 1 sweep

BUT WE HAVE XIA2 FOR THIS

INCORPORATING DIALS INTO XIA2

DialsIndexer = dials.import + dials.index + dials.refine_bravais_settings

xia2

DialsRefiner = dials.refine

DialsIntegrater = dials.integrate + dials.export_mtz

Then scale with POINTLESS / AIMLESS

Phase — gw56@cs03r-sc-serv-16:/dls/i04/data/2014/nt5073-4/LIC/2/5 —...

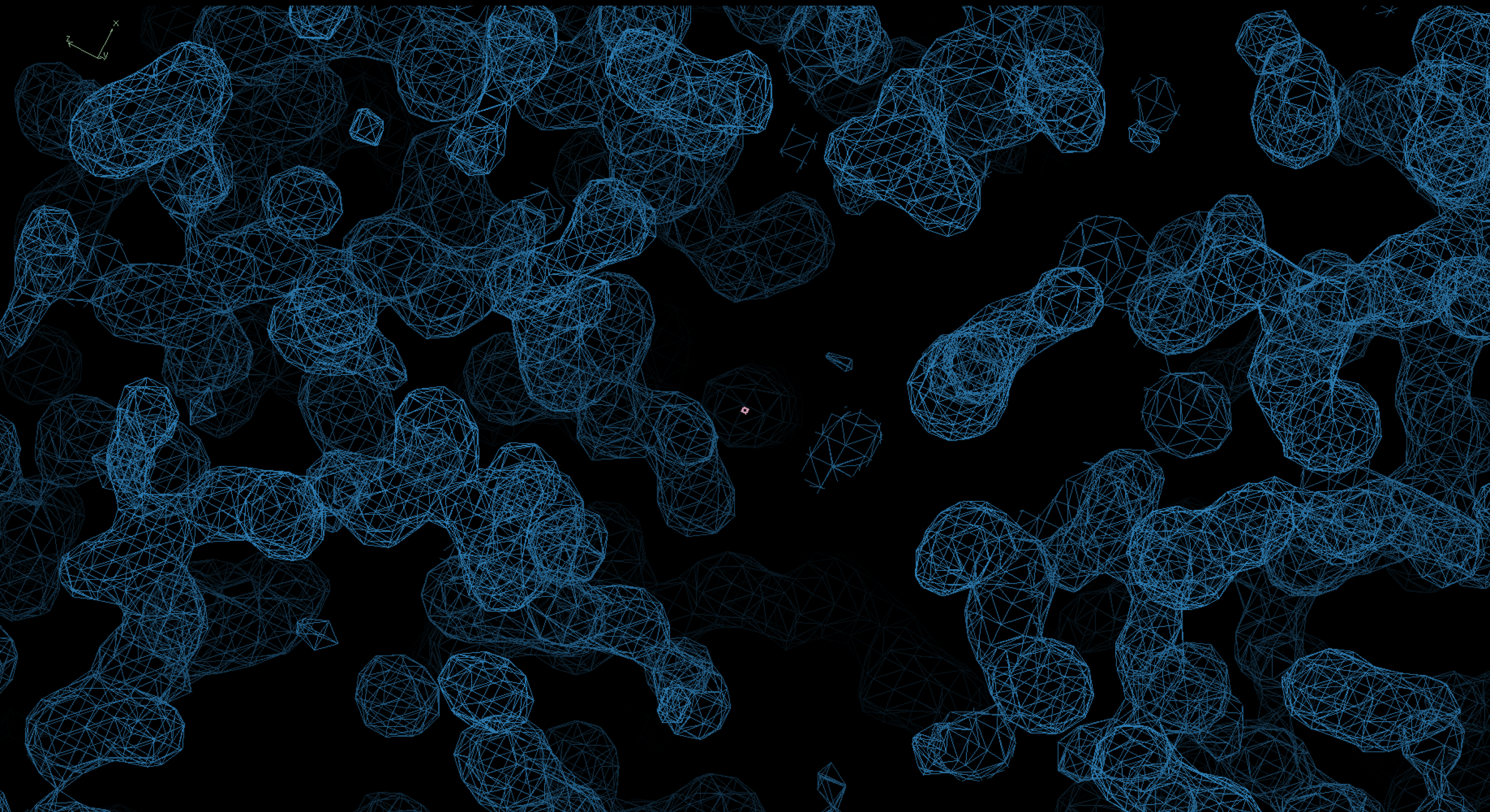
```
Graemes-MacBook-Pro-3:DIALS-6 graeme$ cd LIC25_32/
Graemes-MacBook-Pro-3:LIC25_32 graeme$ xia2 -atom Zn -dials /Volumes/GraemeData/
data/i04-soleil/Lic25_strong/
Environment configuration...
XIA2_ROOT => /Users/graeme/svn/cctbx/modules/xia2
XIA2CORE_ROOT => /Users/graeme/svn/cctbx/modules/xia2/core
Python => /Users/graeme/svn/cctbx/build/../base/Python.framework/Versions/2.7/Re
sources/Python.app/Contents/MacOS/Python
CCTBX => /Users/graeme/svn/cctbx/modules
CCP4 => /Applications/ccp4-6.5
CLIBD => /Applications/ccp4-6.5/lib/data
CCP4_SCR => /var/folders/hs/gvssfcd910s0jczjhjczj6h80000gn/T/tmpreQPsS
Working directory: /Users/graeme/Projects/DIALS-6/LIC25_32
Free space:        218.08 GB
Build: 5267
Contact: xia2.support@gmail.com
XIA2 0.3.8.0
Command line: xia2 -atom Zn -dials /Volumes/GraemeData/data/i04-soleil/Lic25_str
ong/
-------------------- Autoindexing SWEEP1 --------------------
All possible indexing solutions:
tP 121.50 121.50  57.03  90.00  90.00  90.00
oC 171.82 171.84  57.03  90.00  90.00  90.00
oP  57.01 121.46 121.50  90.00  90.00  90.00
```

Phase — gw56@cs03r-sc-serv-16:/dls/i04/data/2014/nt5073-4/LIC/2/5 —...

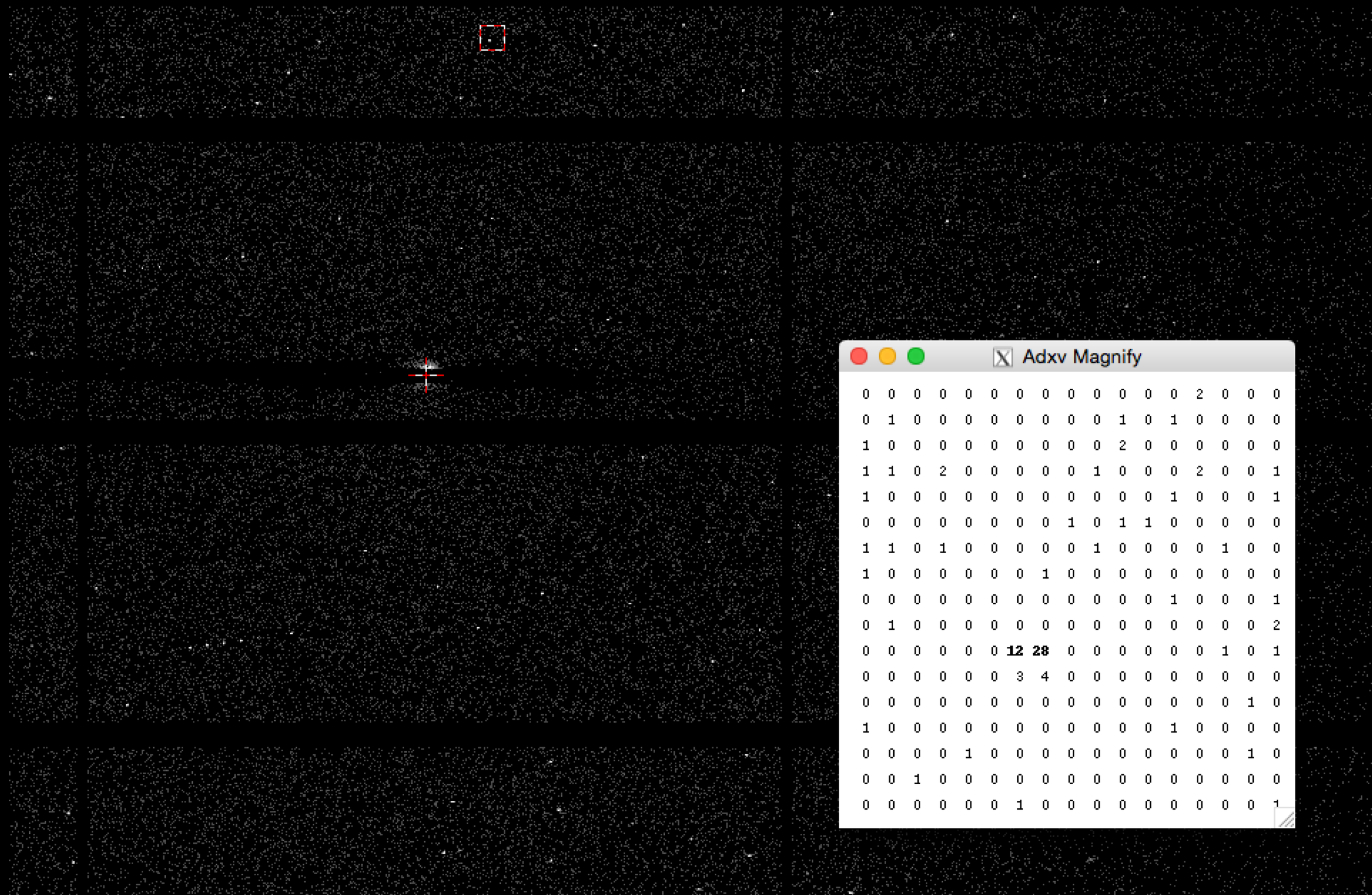
```
For AUTOMATIC/DEFAULT/SAD
High resolution limit
Low resolution limit
Completeness
Multiplicity
I/sigma
Rmerge
Rmeas(I)
Rmeas(I+/-)
Rpim(I)
Rpim(I+/-)
CC half
Wilson B factor
Anomalous completeness
Anomalous multiplicity
Anomalous correlation
Anomalous slope
dF/F
dI/s(dI)
Total observations
Total unique
Assuming spacegroup: P 42 21 2
Unit cell:
121.505 121.505  57.029
```

1.54	6.89	1.54
121.50	121.50	1.58
100.0	99.9	100.0
10.7	9.6	10.4
17.0	39.3	3.4
0.065	0.037	0.609
0.077	0.048	0.683
0.072	0.040	0.675
0.023	0.016	0.210
0.030	0.017	0.289
0.999	0.998	0.893
14.071		
100.0	100.0	100.0
5.5	5.8	5.3
0.498	0.735	0.049
1.251	0.000	0.000
0.081		
1.260		
682919	8089	48088
63570	845	4625

NO INPUT, USEFUL RESULT

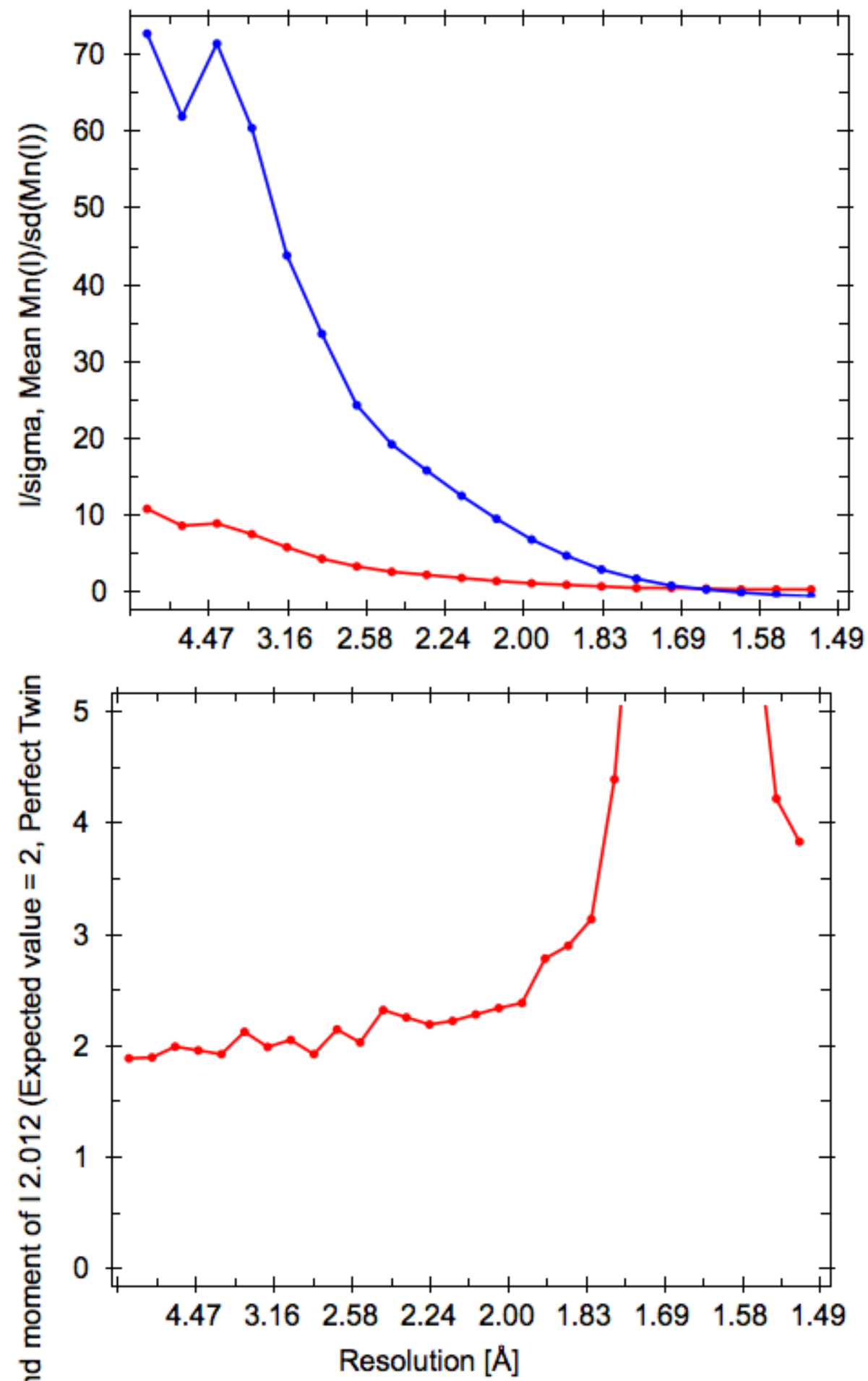


TREATMENT OF WEAK DATA



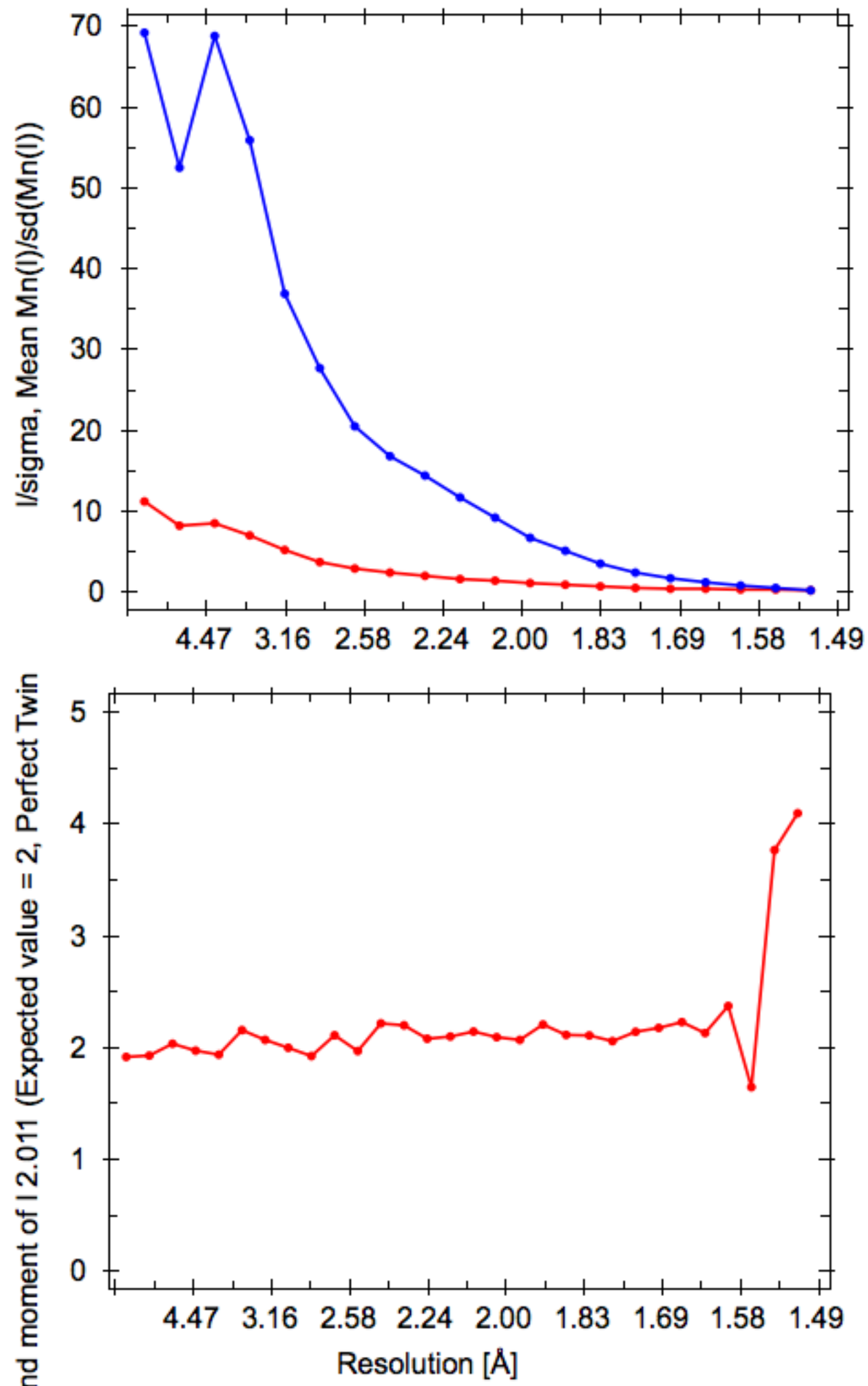
XDS

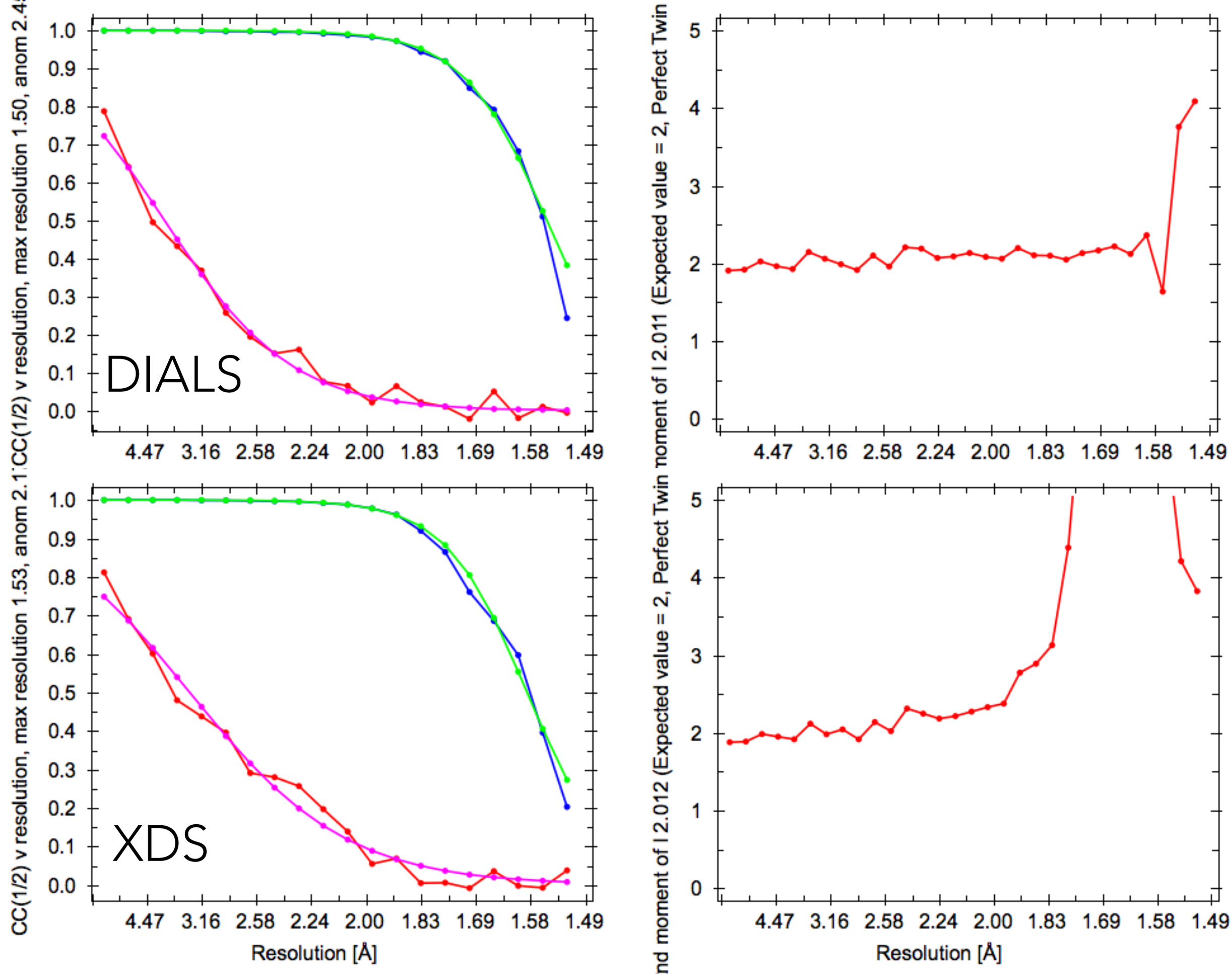
- High multiplicity (60-fold) weak data
- Data much lower intensity than recommended!
- I/sig(I) looks OK
- E^4 looks nasty



DIALS

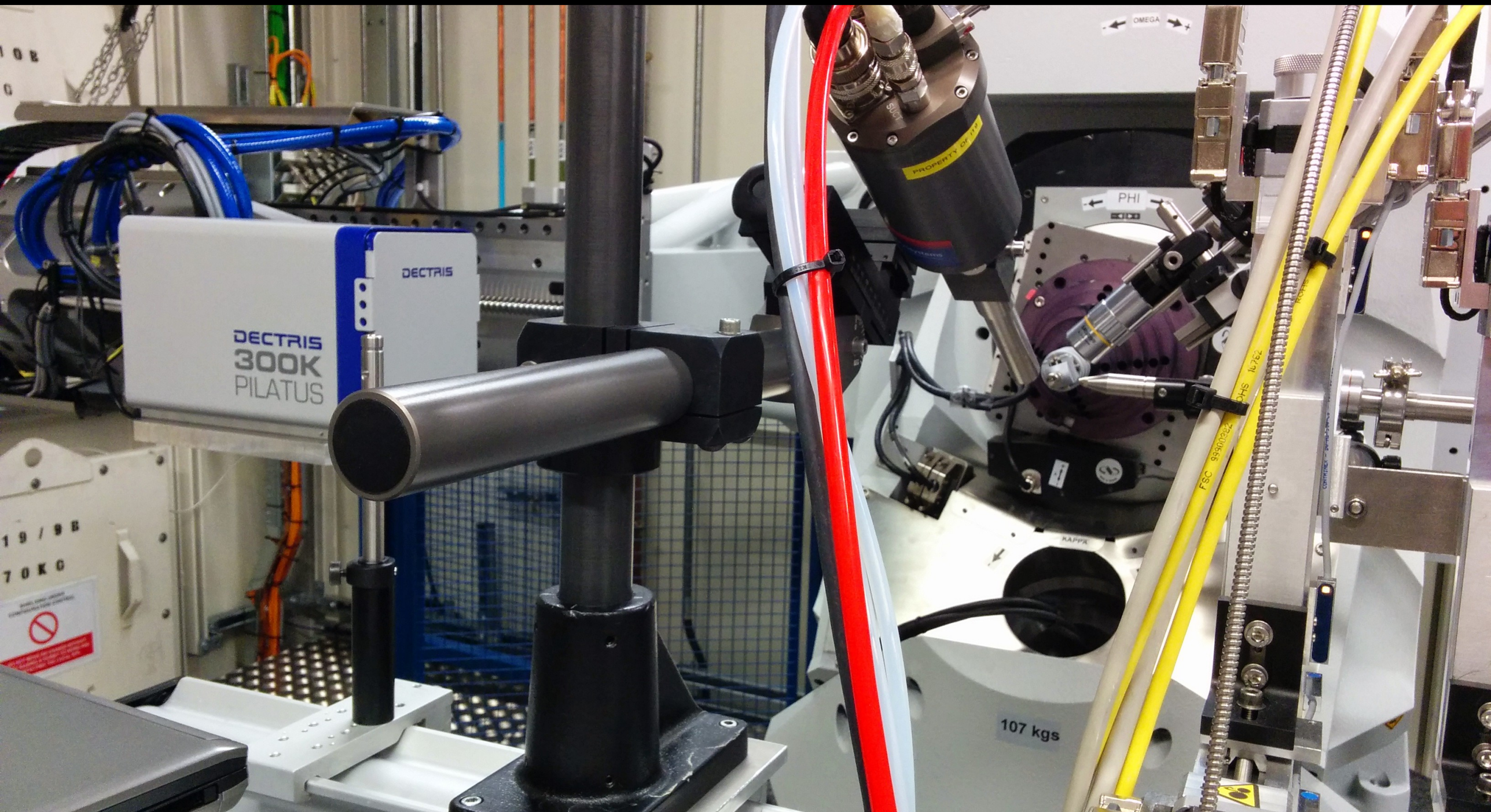
- I/sig(I) etc. look good
- E^4 looks good also to outer shells
- => estimates of $\langle I_{hkl} \rangle$ closer to Wilson distribution
- For high multiplicity data, error estimates critical




CC(1/2) v resolution, max resolution 1.50, anom 2.4
CC(1/2) v resolution, max resolution 1.53, anom 2.1
DIALS
XDS
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
4.47
3.16
2.58
2.24
2.00
1.83
1.69
1.58
1.49
Resolution [Å]
nd moment of I 2.011 (Expected value = 2, Perfect Twin
nd moment of I 2.012 (Expected value = 2, Perfect Twin
5
4
3
2
1
0

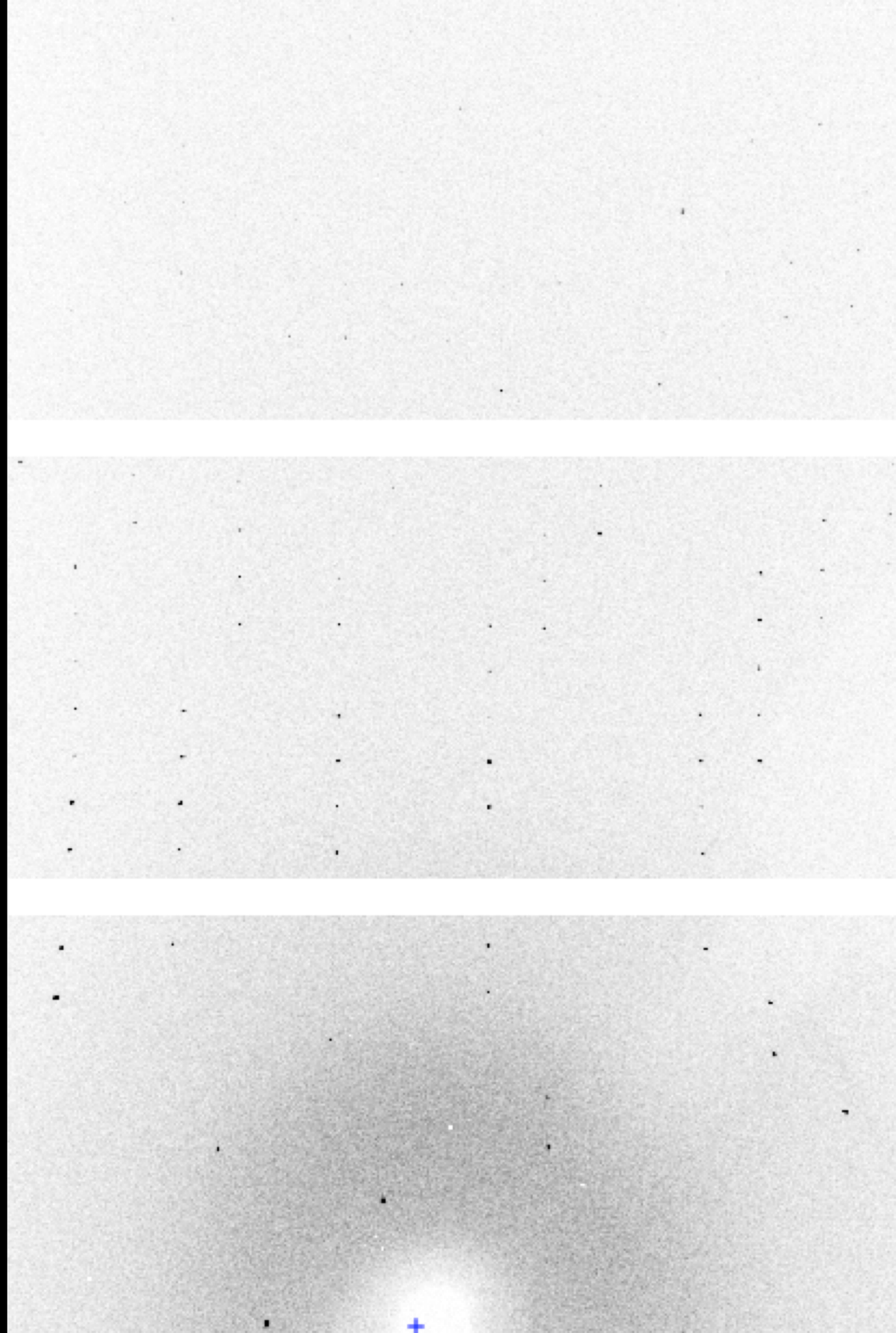
NOT JUST MX...

DIAMOND I19 EH2



CHEMICAL CRYSTALLOGRAPHY

- “Small molecule” crystallography
- Highly absorbing samples
- High energy data collection => DQE important
- More complex experimental geometry



dials-0.01 — gw56@cs03r-sc-serv-16:/dls/i04/data/2014/nt5073-4/LIC/2/5...

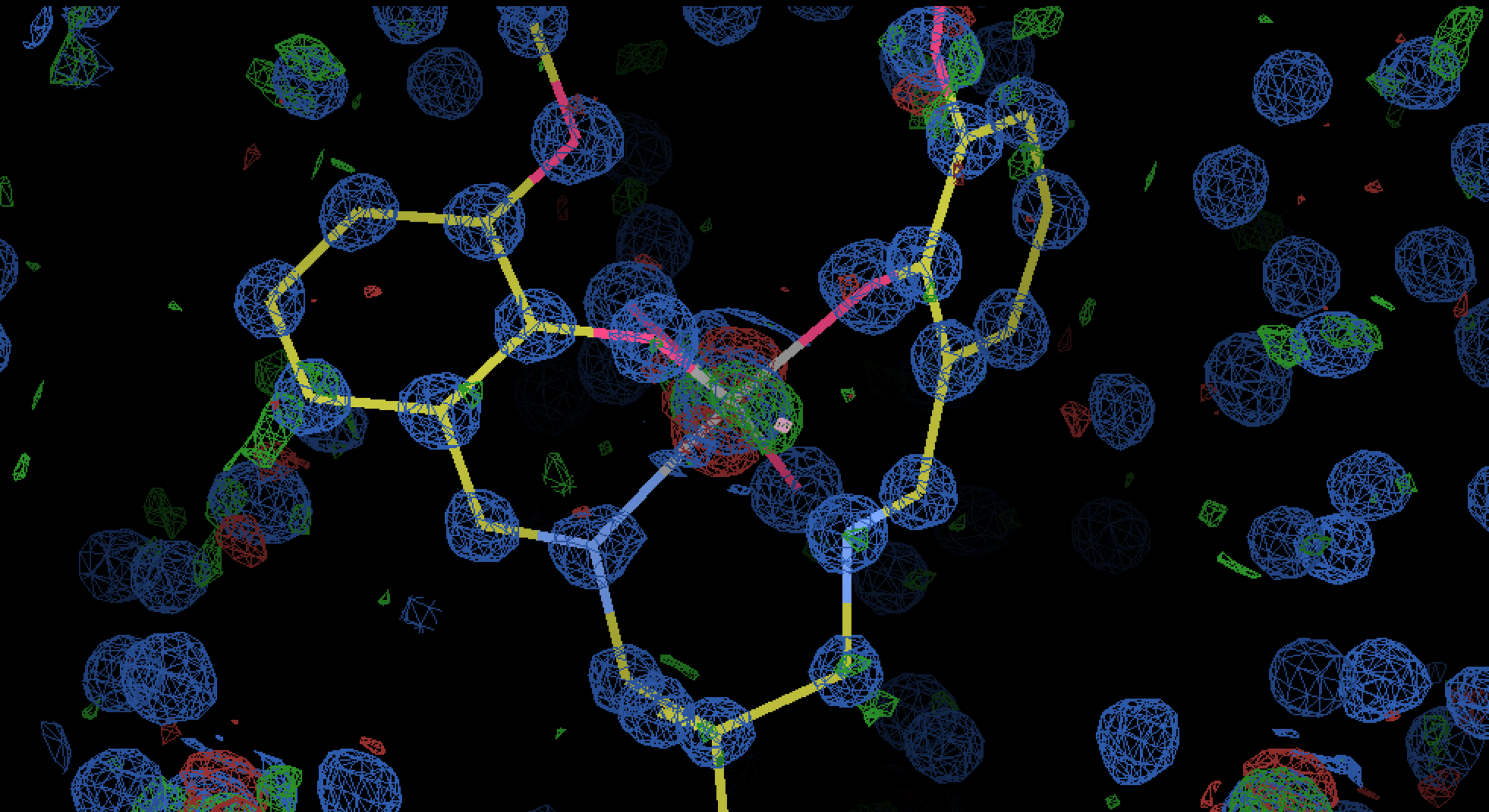
```
XIA2 0.3.8.0
Command line: xia2 reference_geometry=../i19.json trust_beam_centre=true -small_
molecule -dials -atom X surface_tie=0.01 /Users/graeme/data/i19/56_Reza-H2-norma
lhem_100K_3.75mmAl
-------------------- Autoindexing SWEEP1 --------------------
All possible indexing solutions:
oP   9.23  12.60  17.68  90.00  90.00  90.00
mP   9.23  17.68  12.60  90.00  89.98  90.00
aP   9.23  12.60  17.68  90.01  90.01  89.98
Indexing solution:
oP   9.23  12.60  17.68  90.00  90.00  90.00
-------------------- Integrating SWEEP1 --------------------
Processed batches 1 to 325
Standard Deviation in pixel range: 0.182918 0.482826
Integration status per image (60/record):
oooooooooooooooooooooooooooooooooooooooooooooooooooooooooooo
oooooooooooooooooooooooooooooooooooooooooooooooooooooooooooo
oooooooooooooooooooooooooooooooooooooooooooooooooooooooooooo
oooooooooooooooooooooooooooooooooooooooooooooooooooooooooooo
oooooooooooooooooooooooooooooooooooooooooooooooooooooooooooo
oooooooooooooooooooooooooo
"o" => good        "%" => ok        "!" => bad rmsd
"O" => overloaded  "#" => many bad  "." => weak
:
```

dials-0.01 — gw56@cs03r-sc-serv-16:/dls/i04/data/2014/nt5073-4/LIC/2/5...

High resolution limit	0.67	3.00	0.67
Low resolution limit	9.22	9.22	0.69
Completeness	97.8	98.0	76.8
Multiplicity	5.9	8.0	2.2
I/sigma	12.6	34.9	3.6
Rmerge	0.078	0.054	0.228
Rmeas(I)	0.092	0.060	0.310
Rmeas(I+/-)	0.090	0.060	0.302
Rpim(I)	0.034	0.022	0.178
Rpim(I+/-)	0.044	0.025	0.195
CC half	0.995	0.996	0.862
Anomalous completeness	92.9	100.0	39.4
Anomalous multiplicity	3.1	6.3	1.6
Anomalous correlation	0.096	-0.045	0.008
Anomalous slope	1.082	0.000	0.000
dF/F	0.101		
dI/s(dI)	0.962		
Total observations	22964	481	465
Total unique	3924	60	211

```
Assuming spacegroup: P 21 21 21
Unit cell:
9.218  12.598  17.694
90.000  90.000  90.000
:
```

USEFUL RESULTS



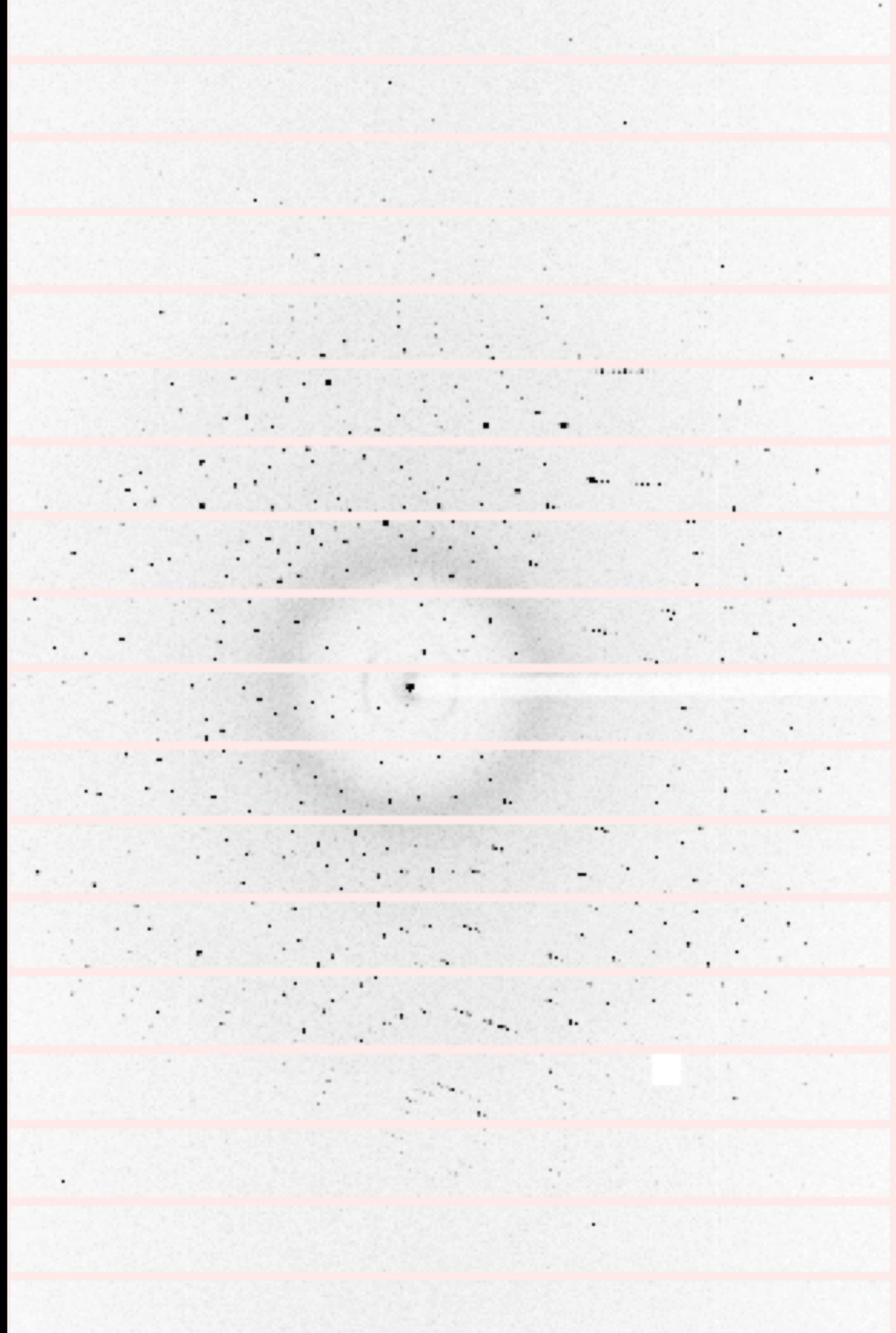
NOT JUST FLAT DETECTORS

NON-PLANAR DETECTORS

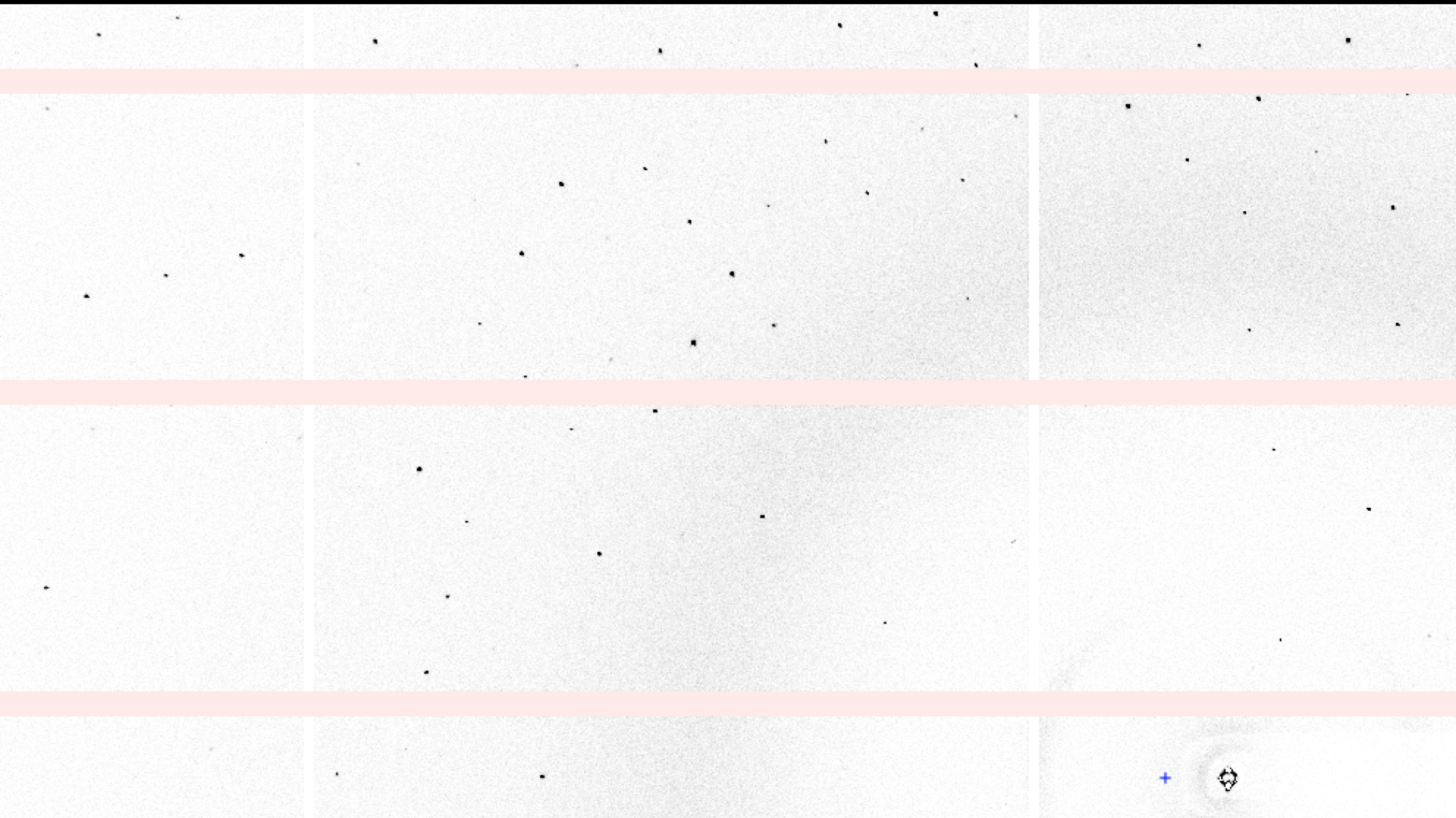
- Long wavelength beamline I23 @ Diamond Light Source
- Custom P12M
- Not yet supported in xia2
- Excellent data seen - experiment in vacuum
- Support simple *via* dxtbx

DATA

- Very low background (in vacuum)
- Very sharp spots despite lack of beam focussing at the time
- Data collection at 9 keV
- Two theta to around 75° hence design of detector
- Currently single axis goniometer



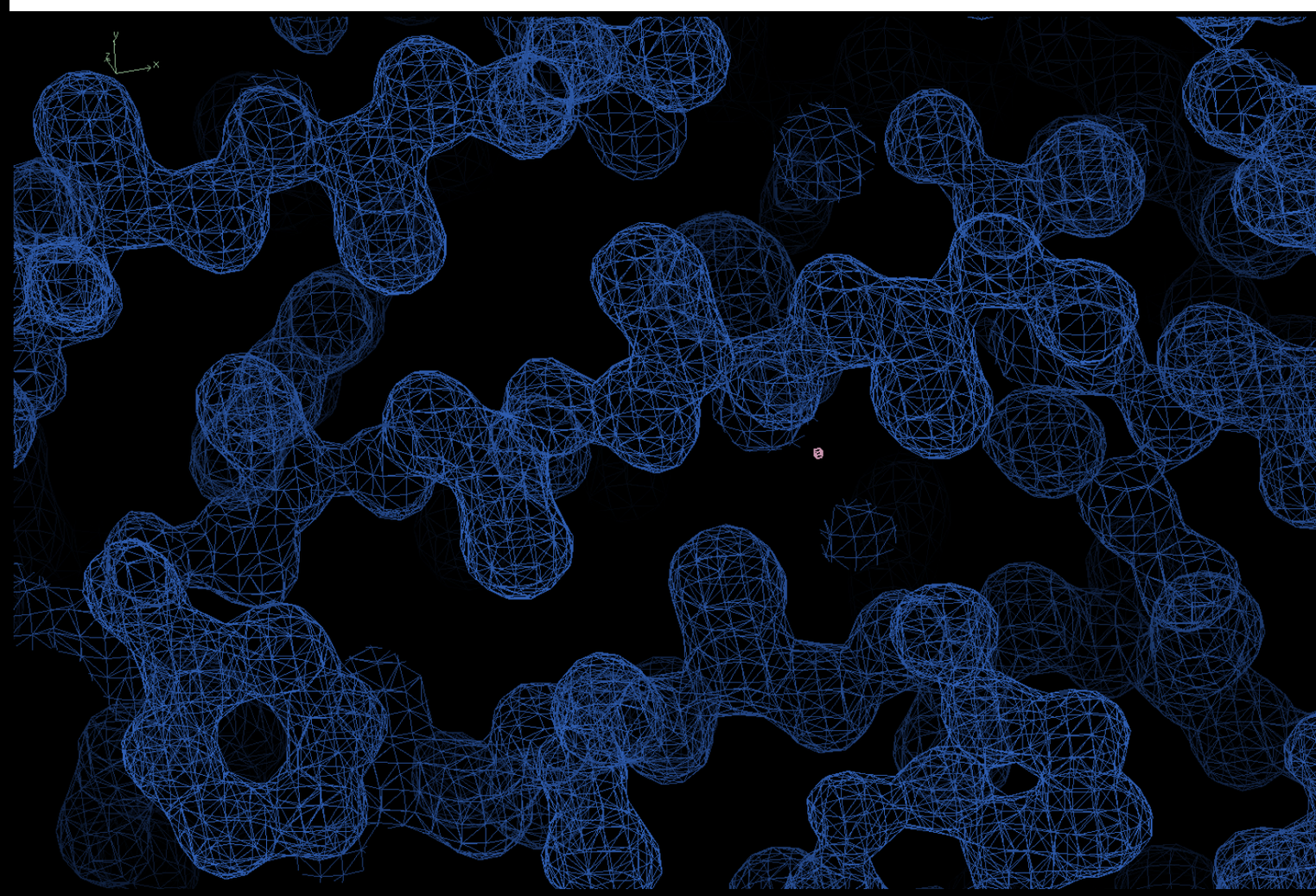
DATA AT LOW RESOLUTION



RESULTS

- Processed using "standard" script
- Scaled with AIMLESS again using standard commands
- Phased with shelxc/d/e - gives excellent maps

	Overall	InnerShell	OuterShell
Low resolution limit	28.93	28.93	1.12
High resolution limit	1.10	6.02	1.10
Rmerge (within I+/I-)	0.047	0.022	0.715
Rmerge (all I+ and I-)	0.050	0.025	0.766
Rmeas (within I+/I-)	0.052	0.024	0.820
Rmeas (all I+ & I-)	0.053	0.027	0.821
Rpim (within I+/I-)	0.022	0.009	0.392
Rpim (all I+ & I-)	0.016	0.008	0.288
Rmerge in top intensity bin	0.023	-	-
Total number of observations	910205	7848	30945
Total number unique	93738	781	4010
Mean((I)/sd(I))	15.3	37.0	2.5
Mn(I) half-set correlation CC(1/2)	1.000	1.000	0.634
Completeness	90.2	99.3	79.0
Multiplicity	9.7	10.0	7.7
Anomalous completeness	88.6	100.0	77.6
Anomalous multiplicity	4.9	6.2	3.9
DelAnom correlation between half-sets	0.109	0.467	0.024
Mid-Slope of Anom Normal Probability	0.917	-	-



NOT JUST DATA PROCESSING

DIALS TOOLBOX

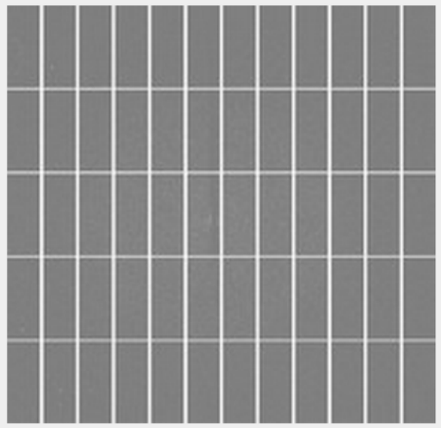
FULL TOOLBOX

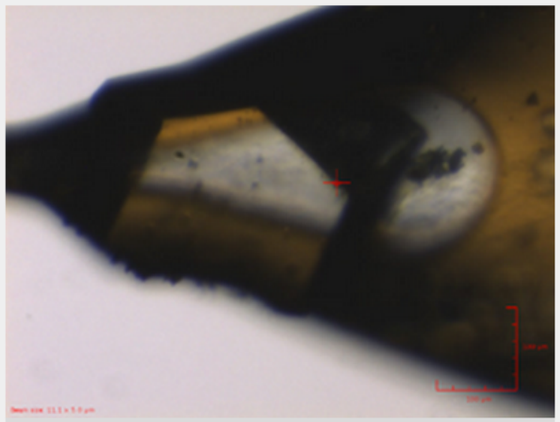
```
dials.analyse_output
dials.background_lookup
dials.check_indexing_symmetry
dials.combine_experiments
dials.compare_mosflm_dials
dials.compare_mosflm_xds
dials.compare_orientation_matrices
dials.compare_reflections
dials.compare_xds_dials
dials.compare_xds_dials2
dials.create_profile_model
dials.detector_max_resolution
dials.discover_better_experimental_model
dials.display_reference_profiles
dials.dqe
dials.estimate_resolution_limit
dials.export
dials.export_mosflm
dials.export_mtz
dials.export_nxmx
dials.export_nxmx_to_mtz
dials.export_spot_xds
dials.export_text
dials.export_xds
dials.extract_shoeboxes
dials.filter_good_intensities
dials.filter_reflections
dials.find_hot_pixels
dials.find_overlaps
dials.find_spots
dials.find_spots_client
dials.find_spots_server
dials.generate_mask
dials.generate_process_test_reflections
dials.generate_test_reflections
dials.image_viewer
dials.import
dials.import_xds
dials.index
dials.integrate
dials.merge_cbf
dials.merge_reflection_lists
dials.plot_reflections
dials.plot_scan_varying_crystal
dials.predict
dials.print_test_reflections
dials.process
dials.python
dials.pythonw
dials.reciprocal_lattice_viewer
dials.refine
dials.refine_bravais_settings
dials.reflection_viewer
dials.reindex
dials.remove_absent_reflections
dials.rs_mapper
dials.show_build_path
dials.show_dist_paths
dials.show_extensions
dials.show_indexed_strong
dials.show_isig_rmsd
dials.show_models
dials.show_profiles
dials.show_spots
dials.simulate
dials.simulate_rs
dials.slice_sweep
dials.sort_reflections
dials.split_experiments
dials.spot_counts_per_image
dials.stereographic_projection
```

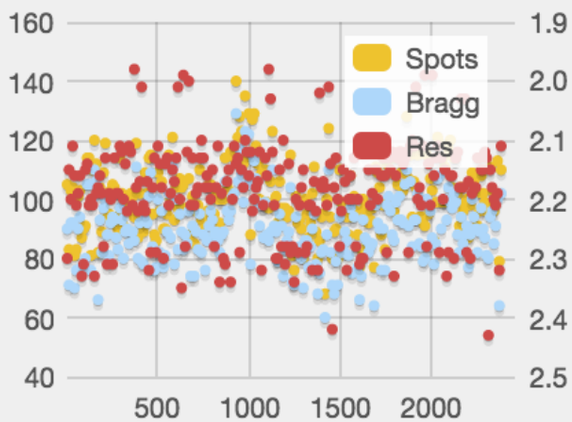
USER FEEDBACK - GRID SCANS

15-05-2015 15:50:35 - GlIso/GlIso_M6S7_1_####.cbf

Sample: Gl_7	Ω Start: 0.0°
Ω Osc: 0.15°	Ω Overlap: 0°
No. Images: 2400	Resolution: 2.27Å
Wavelength: 1.8903Å	Exposure: 0.040s
Transmission: 0.50%	Beamsize: 11x5µm
Type: Data Collection	
Comment: (900,189,-14) EDNAStrategy3: subWedge:1Aperture: Large	







Auto Processing — Fast DP: ✓ Xia2: ✗ ✗

Fast DP | DIALS

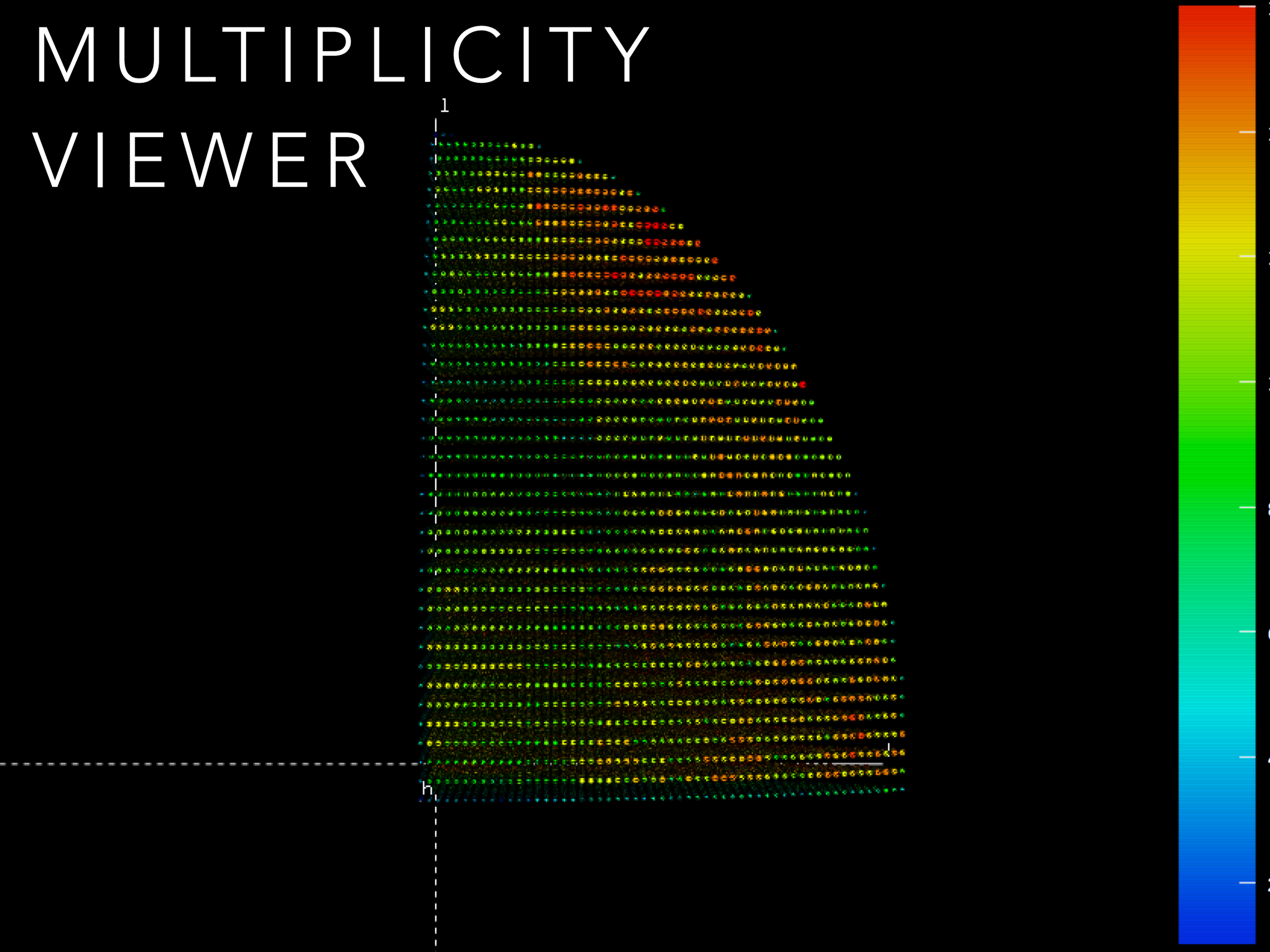
Space Group	A	B	C	α	β	γ
I 2 2 2	93.04	98.22	102.79	90.00	90.00	90.00

MTZ file | Log file | Lookup Cell

Shell	Observations	Unique	Resolution	Rmeas	I/sig(I)	Completeness	Multiplicity	Anom Completeness	Anom Multiplicity
outerShell	947	519	1.90 - 1.95	0.571	2.0	18.5	1.8	4.3	1.5
innerShell	6208	521	8.29 - 71.01	0.082	17.8	99.9	11.9	100.0	7.2
overall	341484	30333	1.90 - 71.01	0.148	11.8	81.4	11.3	78.7	5.8

Downstream Processing — Fast EP: ✓ Dimple: ✓

MULTIPLICITY VIEWER



MISSING REFLECTIONS

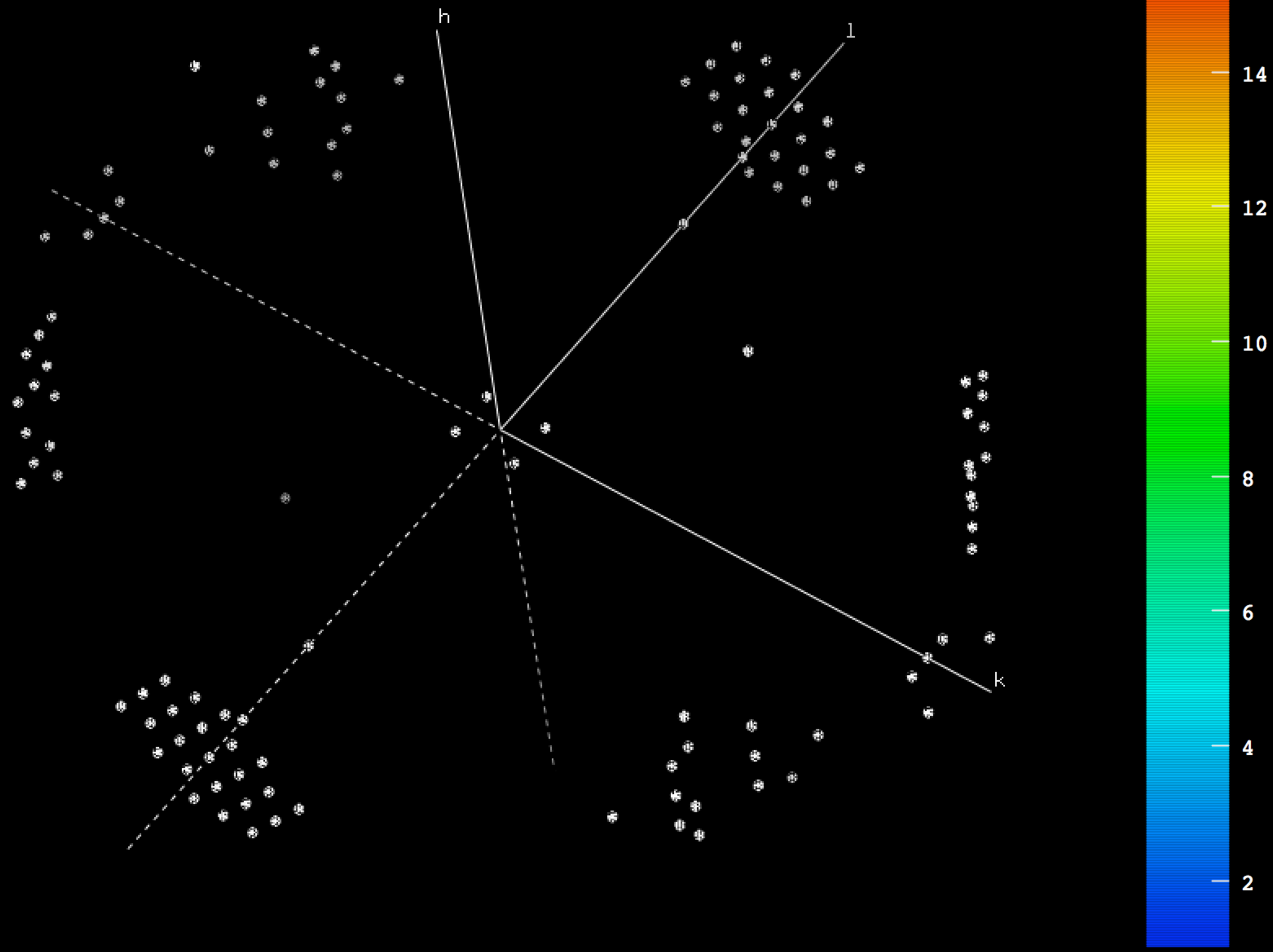


IMAGE VIEWER - RAW

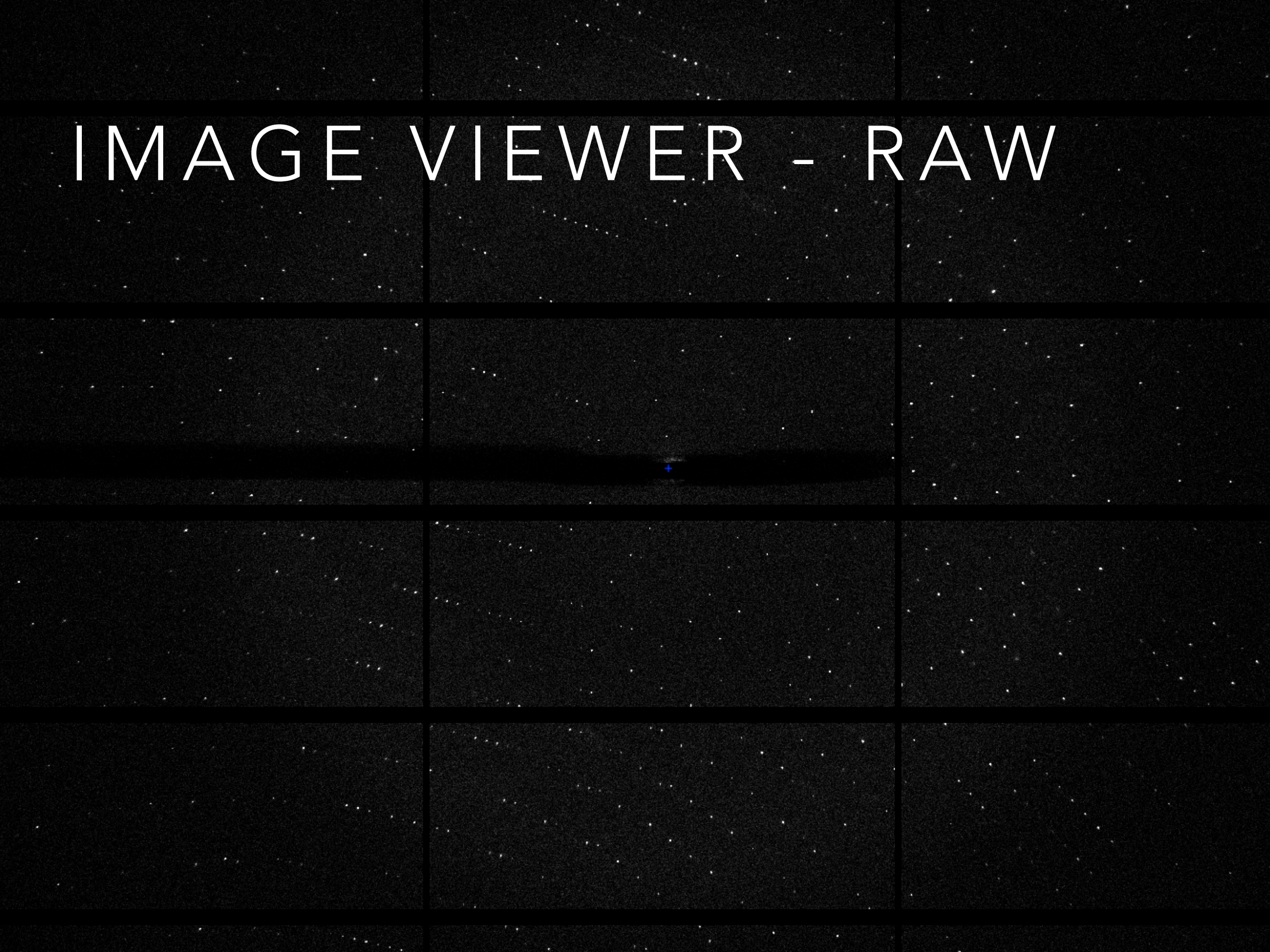
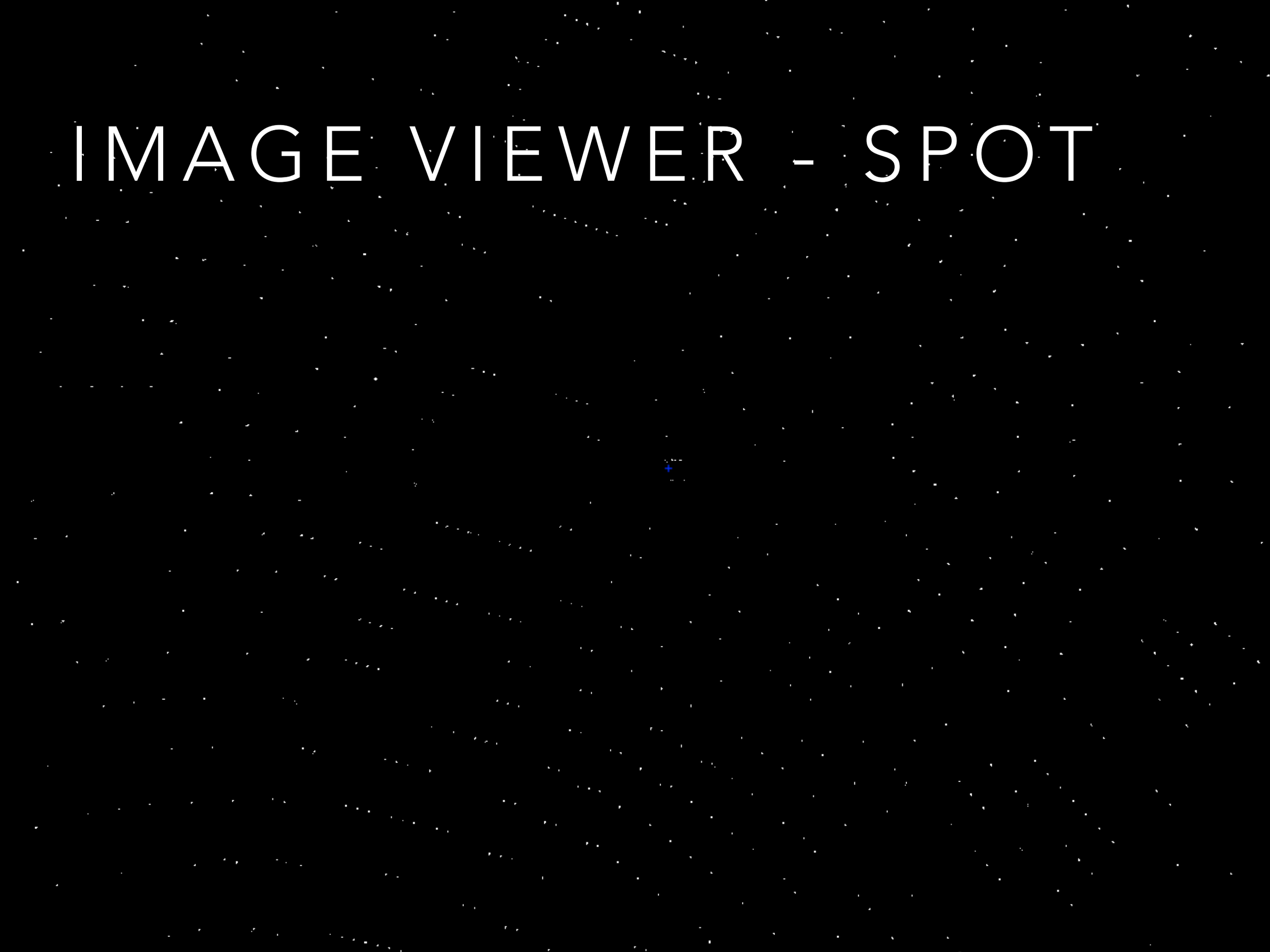


IMAGE VIEWER - SPOT



STATUS

- DIALS now used routinely for automated data processing *via* xia2
- xia2 is the "friendly" DIALS user interface for synchrotron data, and is bundled with DIALS
- DIALS spot finding now back end of per-image analysis used for grid scans & during data collection
- Tools developed using DIALS framework for image inspection, strategy evaluation on I19, multi-crystal analysis *via* BLEND, …

ACKNOWLEDGEMENTS

- Gwyndaf Evans, David Waterman, James Parkhurst, Luis Fuentes-Montero, Richard Gildea, Markus Gerstel, Alun Ashton
- Andrew Leslie, Harry Powell, Phil Evans, Garib Murshudov
- Nick Sauter, Aaron Brewster, Nat Echols, Ian Rees, Nigel Moriarty, Ralf Grosse-Kunstleve, …
- Software users & Diamond scientists

HTTP://DIALS.SF.NET