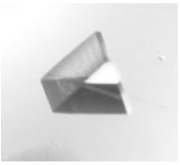
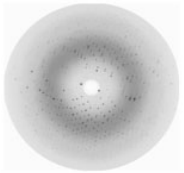
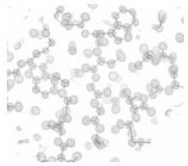
Data Processing

HZB Helmholtz Zentrum Berlin



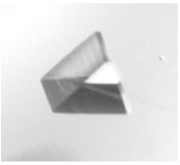



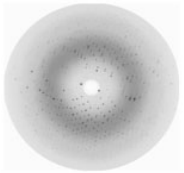
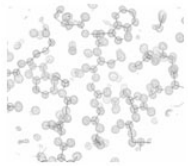
XDSAPP and future developments

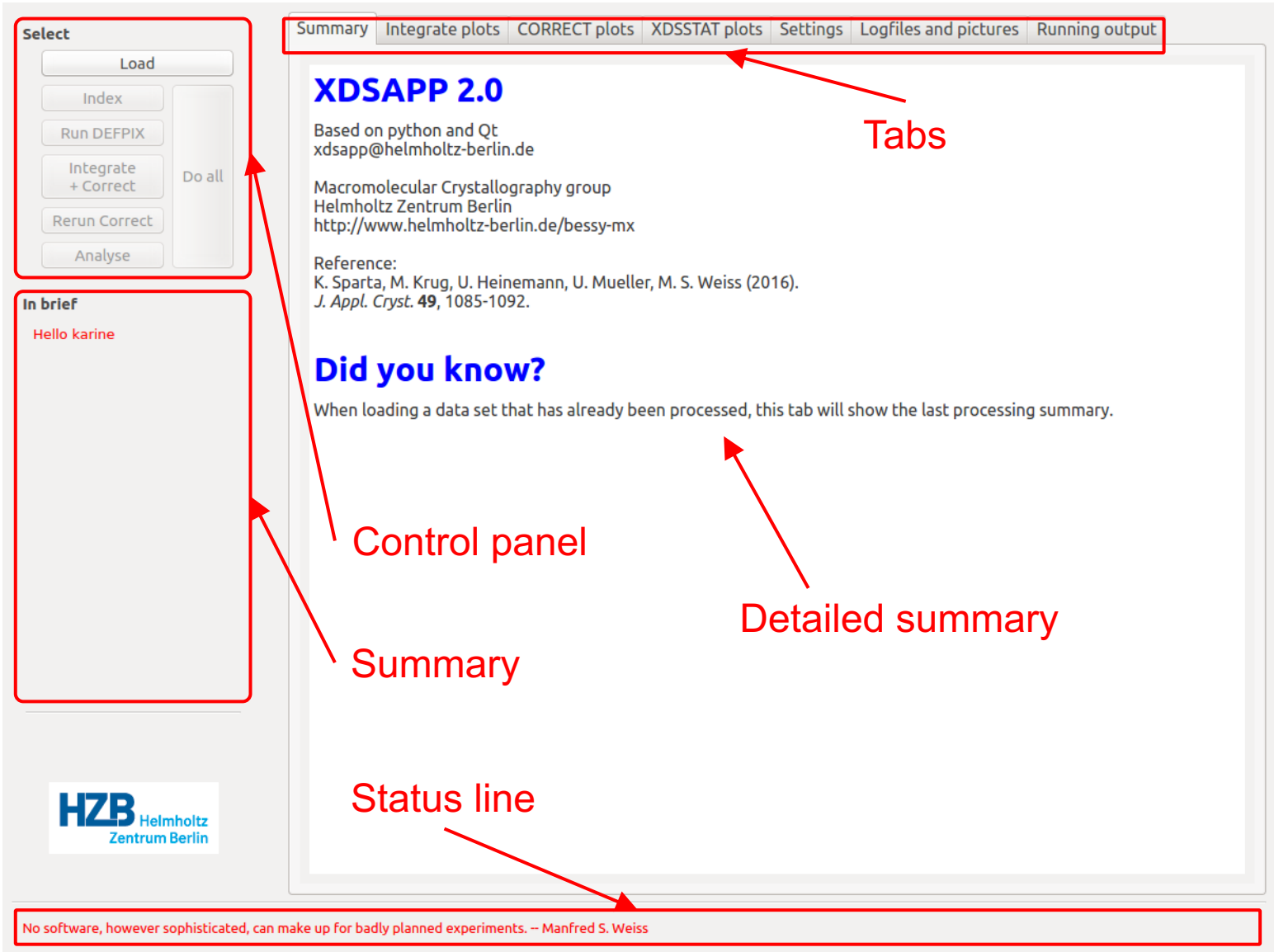
Karine Röwer

Helmholtz-Zentrum Berlin für Materialien und Energie
Macromolecular Crystallography (NP-GMX)
Albert-Einstein-Str. 15
D-12489 Berlin, Germany
karine.sparta@helmholtz-berlin.de

XDSAPP layout

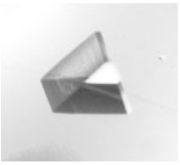


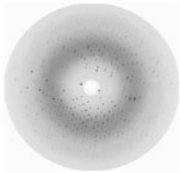
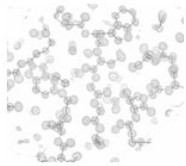





K. Sparta *et al.* (2016). *J. Appl. Cryst.* **49**, 1085-1092.

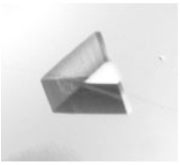
XDSAPP features

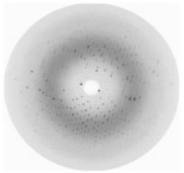
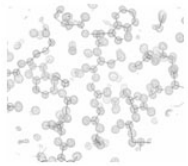


- Processing of diffraction data using XDS, CCP4 and Phenix
- Live graphical representation of data statistics
- Automated decision-making
  - Space group determination
  - Detection of anomalous signal
  - Resolution cutoff
- Conversion of intensities to various formats: mtz, hkl, cns
- Live mode for processing during data collection

XDSAPP workflow



Index

- XYCORR
  INIT
  COLSPOT
- IDXREF
- DEFPIX

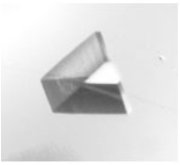
Integrate + CORRECT

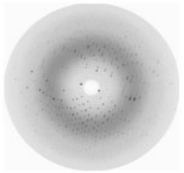
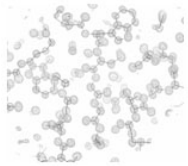
- INTEGRATE in P1
  CORRECT
- Determination of
  - maximum resolution
  - anomalous signal
  - point group and Bravais lattice
- Smart reintegration cycles
- Check aliens, ice rings
- Space group determination
  - pointless
  - CORRECT

Analyze

- XDSSTAT
- Data conversion
  - MTZ
  - Shelx
  - CNS format
- Twinning analysis
  - SFCHECK
  - phenix.xtriage
- Enantiomorphic space groups
- Quality analysis

Indexing



Select

Load
Index
Run DEFPIX
Integrate + Correct
Rerun Correct
Analyse
Do all

In brief

Prol_wt_Mg_LP_X9_3m15s_

Images:
1 to 1600

Wavelength:
0.918409 Å

Spots indexed:
5065 of 7410 (68.4%)

Finished

Summary | Integrate plots | CORRECT plots | XDSSTAT plots | Settings | Logfiles and pictures | Running output

Indexing results for data set Prol_wt_Mg_LP_X9_3m15s_1

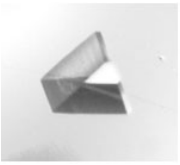
No. of subtrees:	15
No. of strong spots in subtree 1:	2982
No. of strong spots in subtree 2:	2
Quality value:	1.1
Delta:	0.1
DH, DK, DL:	0.01, 0.01, 0.05
Selected origin	0 0 0
Computed (given) beam position on detector:	1240.4, 1250.2 (1239.0, 1248.0)
Std. of spot position (refined):	0.30 (0.35) px
Std. of spindle position (refined):	0.11 (0.14) deg.
Detector distance given (refined):	283.565 (282.39)
Spots indexed:	5065 out of 7410 (68.4 %)

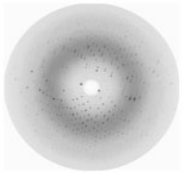
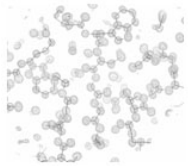
Possible lattices are:

aP	0.0	74.4	74.4	216.9	90.0	90.0	88.2
aP	0.2	74.4	74.4	216.9	90.0	90.0	91.8
mP	1.3	74.4	216.9	74.4	90.0	91.8	90.0
mC	1.3	106.8	103.5	216.9	90.0	90.0	90.0
oC	2.0	103.5	106.8	216.9	90.0	90.0	90.0
mC	2.0	103.5	106.8	216.9	90.0	90.0	90.0
mP	16.0	74.4	74.4	216.9	90.0	90.0	91.8
mP	16.7	74.4	74.4	216.9	90.0	90.0	91.8
oP	16.9	74.4	74.4	216.9	90.0	90.0	91.8
tP	17.6	74.4	74.4	216.9	90.0	90.0	91.8

Idle since 15:51:44.

Reintegration cycles



Select

Load

Index

Run DEFPIX

Integrate + Correct

Rerun Correct

Analyse

Do all

In brief

Prol_wt_Mg_LP_X9_3m15s_

Space group:
C222

Resolution:
1.15Å

Rmeas total:
9.5 %

Mosaicity:
0.082°

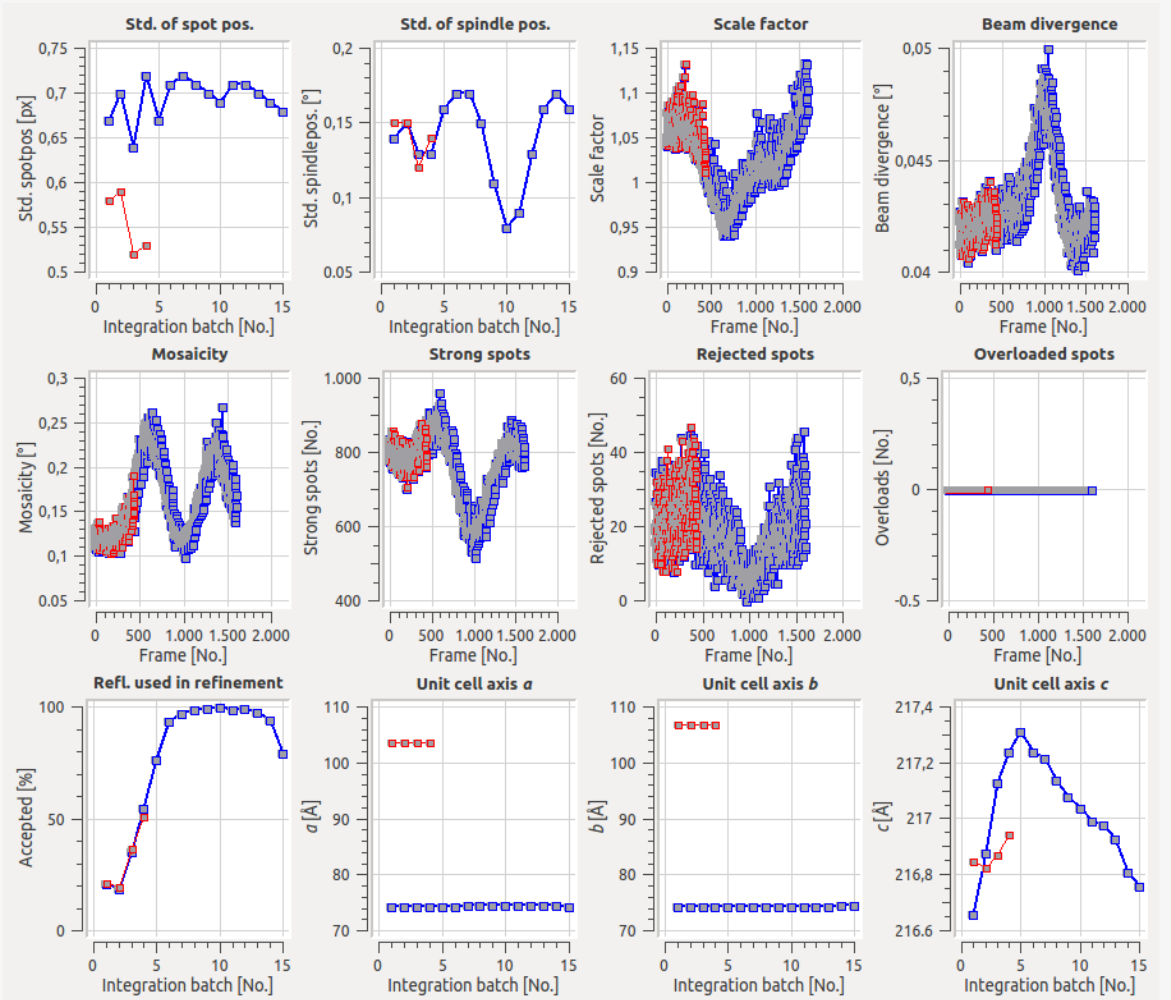
Completeness:
78.5 %

I/sigI:
7.87

Isa:
55.57

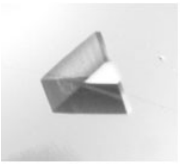
Reprocessing

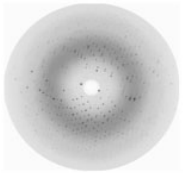
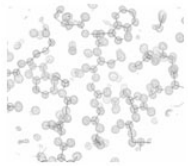
Summary | Integrate plots | CORRECT plots | XDSSTAT plots | Settings | Logfiles and pictures | Running output



Processing of images 429 - 534, data set Prol_wt_Mg_LP_X9_3m15s_1, reintegration step 1 of max. 3

Data quality indicators (CORRECT)



Select

Load
Index
Run DEFPIX
Integrate + Correct
Rerun Correct
Analyse
Do all

In brief

Prol_wt_Mg_LP_X9_3m15s_

Space group:
C222

Resolution:
1.15Å

Rmeas total:
9.5 %

Mosaicity:
0.082°

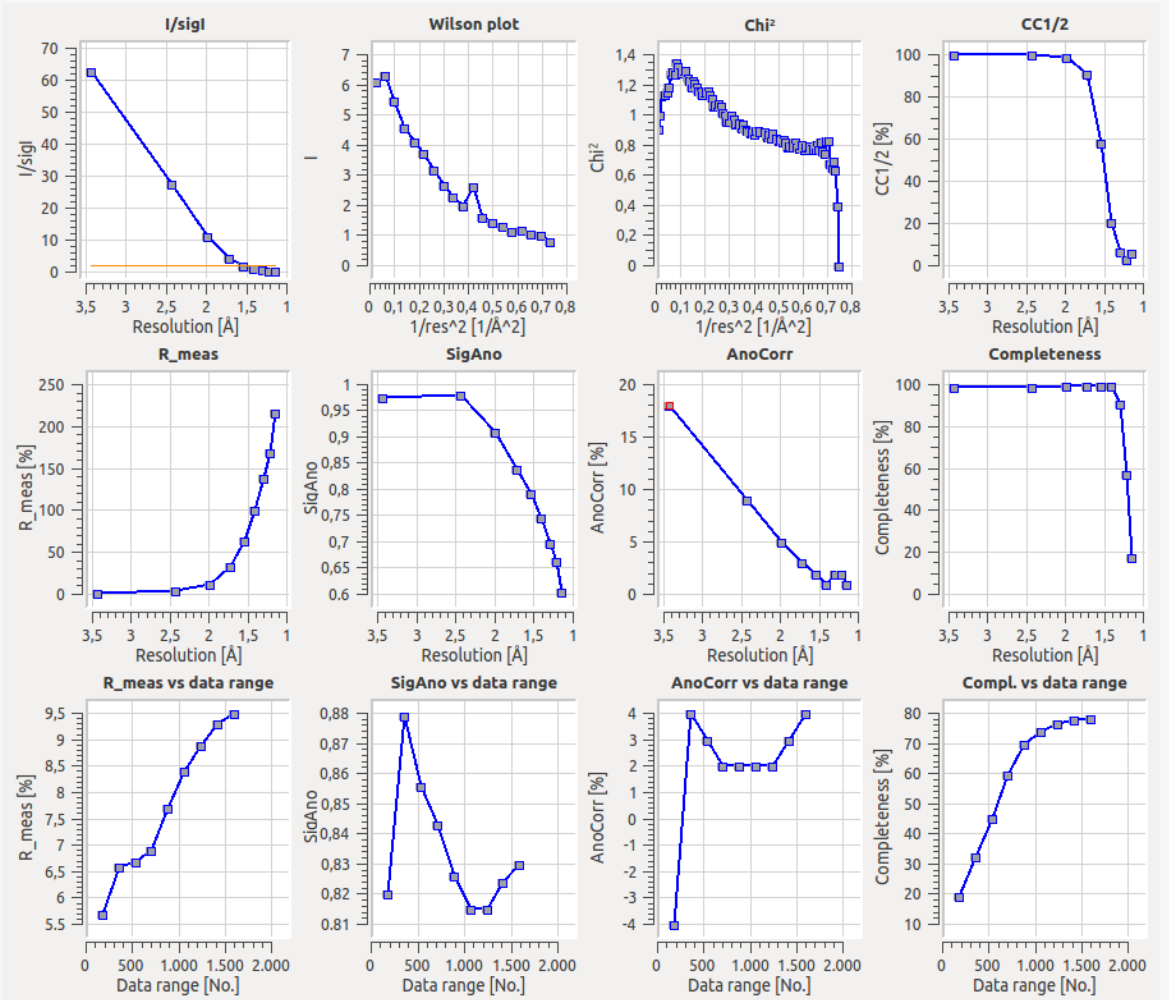
Completeness:
78.5 %

I/sigI:
7.87

Isa:
55.57

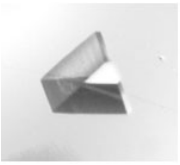
Reprocessing

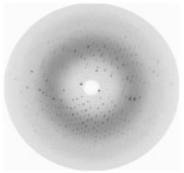
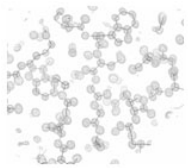
Summary | Integrate plots | CORRECT plots | XDSSTAT plots | Settings | Logfiles and pictures | Running output



CORRECT Data set Prol_wt_Mg_LP_X9_3m15s_1

Analyse

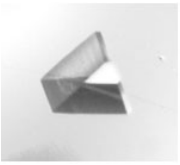


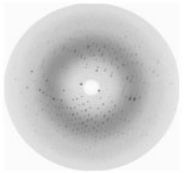
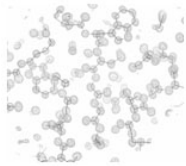



To be performed after the last CORRECT run

- XDSSTAT
- XDSCONV
- SFCHECK
- Quality analysis
  - Indexing quality
  - Twinning (POINTLESS, PHENIX.XTRIAGE)
  - Pseudo translation
  - Ice rings
- Things you may want to check

Analyse



Select

Load
Index
Run DEFPIX
Integrate + Correct
Rerun Correct
Analyse
Do all

In brief

Prol_wt_Mg_LP_X9_3m15s_

Space group:
C222(1)

Resolution:
1.46Å

Rmeas total:
7.0 %

Mosaicity:
0.065°

Completeness:
99.8 %

I/sigI:
17.29

Isa:
50.99

Finished

Summary | Integrate plots | CORRECT plots | XDSSTAT plots | Settings | Logfiles and pictures | Running output

Processing summary for data set Prol_wt_Mg_LP_X9_3m15s_1

Results in: /home/karine/Prolidase/xdsapp/Prol_wt_Mg_LP_X9_3m15s_1

Things you may want to check:
Twinning

Table 1 like summary

Space group	C222(1) (20)
Unit cell parameters [Å]	103.61 106.84 217.00 90.0 90.0 90.0
Resolution limit [Å]	47.93-1.46 (1.55-1.46)
No. of reflections	1227068
No. of uniques	205595
Multiplicity	5.97
I/sigI	17.29 (2.01)
R_meas [%]	7.0 (92.9)
Completeness [%]	99.8 (99.3)
B(Wilson) [Å²]	24.3
Mosaicity [deg]	0.065
CC(1/2)	100.0 (52.8)
ISa	50.99

Result files generated for CCP4, CNS, SHELX

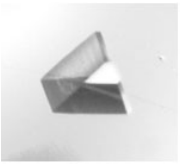
Prol_wt_Mg_LP_X9_3m15s_1_I.mtz
Prol_wt_Mg_LP_X9_3m15s_1_F.mtz
Prol_wt_Mg_LP_X9_3m15s_1_F_plus_F_minus.mtz
Prol_wt_Mg_LP_X9_3m15s_1.hkl
Prol_wt_Mg_LP_X9_3m15s_1.cv

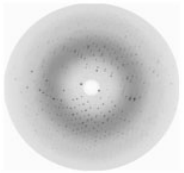
Intensity checks

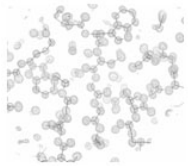
SFCHECK	
Perfect twinning test	2.0715
Pseudo translation	not detected

Idle since 16:34:33.

Analyse







Select

Load
Index
Run DEFPIX
Integrate + Correct
Rerun Correct
Analyse
Do all

In brief

Prol_wt_Mg_LP_X9_3m15s_

Space group:
C222(1)

Resolution:
1.46Å

Rmeas total:
7.0 %

Mosaicity:
0.065°

Completeness:
99.8 %

I/sigI:
17.29

Isa:
50.99

Finished

Summary | Integrate plots | CORRECT plots | XDSSTAT plots | Settings | Logfiles and pictures | Running output

Intensity checks

SFCHECK	
Perfect twinning test	2.0715
Pseudo translation	not detected
Phenix.xtriage	
$\langle I^2 \rangle / \langle I \rangle$	2.033 (untwinned 2.000, perfect twin 1.500)
$\langle F \rangle^2 / \langle F^2 \rangle$	0.790 (untwinned 0.785, perfect twin 0.885)
$\langle \lvert E^2-1 \rvert \rangle$	0.740 (untwinned 0.736, perfect twin 0.541)
Pointless	
|L|-test	0.442 (untwinned 0.500, perfect twin 0.375)
L^2-test	0.271 (untwinned 0.333, perfect twin 0.200)

Space group determination by Pointless

Selected space group	C 2 2 21 (20)
Total probability	0.961
Systematic absences probability	0.992
Next probable space group	none (-)
Total probability	-
Systematic absences probability	-

Indexing summary

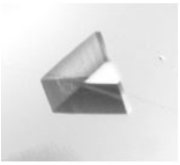
No. of subtrees	15
No of strong spots in root subtree	2982
No. of strong spots in subtree 2	2
Quality value	1.1
Delta	0.1
DH, DK, DL	0.01, 0.01, 0.05
Selected origin	0 0 0
Computed (given) beam position	1240.4, 1250.2 (1239.0, 1248.0)
Std. of spot position (refined)	0.30 (0.35)
Std. of spindle position (refined)	0.11 (0.14)
Detector distance (refined)	283.565 (282.39)
No. of spots indexed (total)	5065 (7410) (=68.4 %)

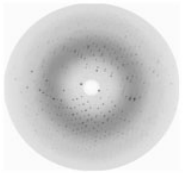
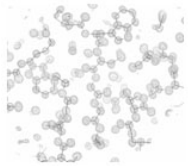
Experimental settings

Idle since 16:34:33.

More statistics: XDSSTAT

HZB Helmholtz Zentrum Berlin



Select

Load
Index
Run DEFPIX
Integrate + Correct
Rerun Correct
Analyse
Do all

In brief

Prol_wt_Mg_LP_X9_3m15s_

Space group:
C222(1)

Resolution:
1.46Å

Rmeas total:
7.0 %

Mosaicity:
0.065°

Completeness:
99.8 %

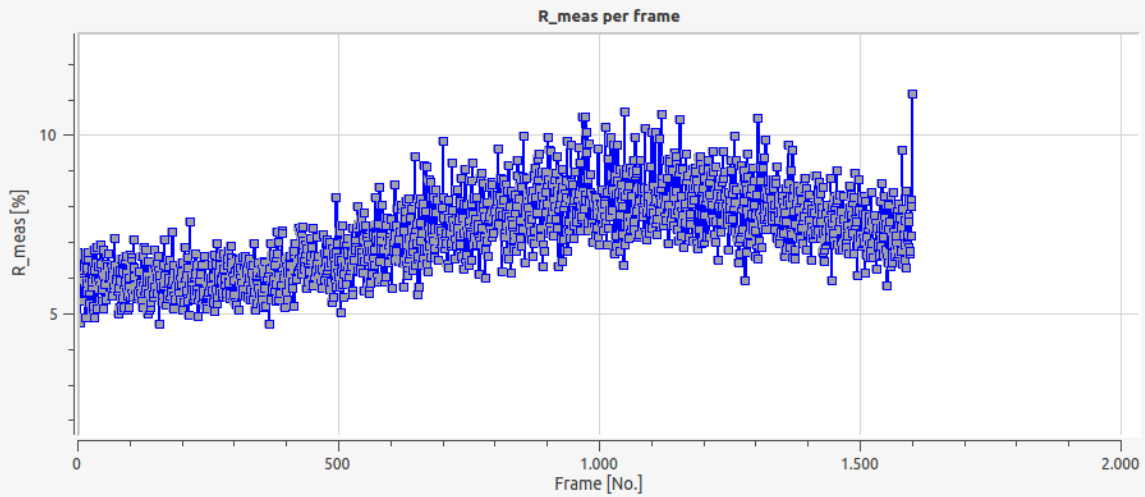
I/sigI:
17.29

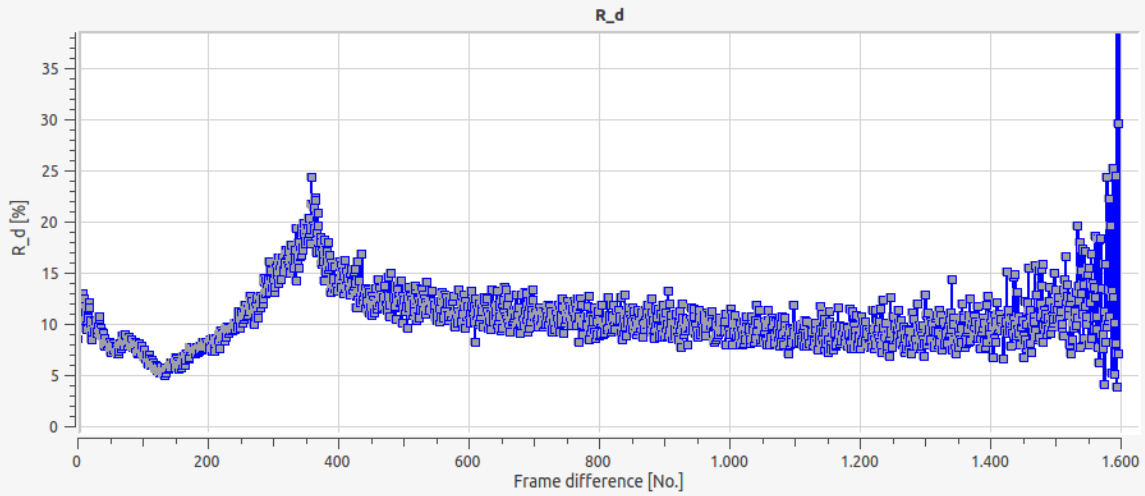
Isa:
50.99

Analysing intensities

HZB Helmholtz Zentrum Berlin

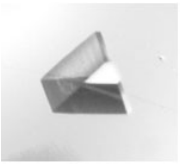
Summary | Integrate plots | CORRECT plots | XDSSTAT plots | Settings | Logfiles and pictures | Running output

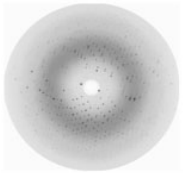
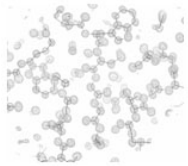




PHENIX.XTRIAGE Data set Prol_wt_Mg_LP_X9_3m15s_1

XDSAPP in command line mode



Command line version of XDSAPP at the beamline, with full automatic server and database connection. Possibility to process all data sets at once.

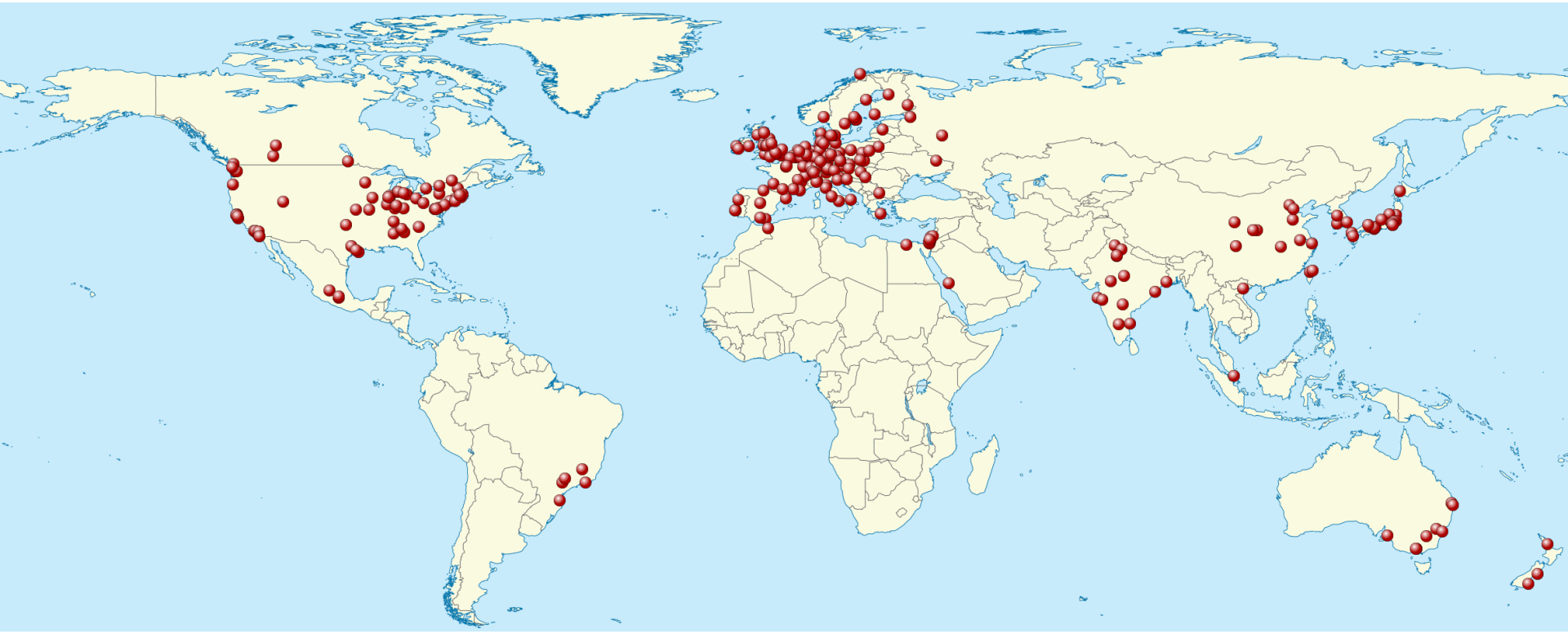
Particularly useful for fragment screening projects.

```
Terminal
File  Edit  View  Search  Terminal  Help
[sparta@hkl4] 109 (karine/pil_thau): xdsapp --cmd --all
/soft/pxsoft/run/xdsit/xdsappbeta/xdsit.py --all
[sparta@hkl4] 110 (karine/pil_thau):
xds assumed to be in PATH
xdsstat found in PATH
ccp4 programs assumed to be in PATH
phenix assumed to be in PATH
16 cpus used
3 jobs
xdsapp is looking for datasets
init
work_dir is: /inhouse/psfrdat/karine/pil_thau/xdsit/t000_02/thau000_02_1
images: 1 - 1100
work_dir is: /inhouse/psfrdat/karine/pil_thau/xdsit/t000_02/merged10/thau000_02_m10
images: 1 - 110
work_dir is: /inhouse/psfrdat/karine/pil_thau/xdsit/t000_05/thau000_05_1
images: 1 - 1100
487  495     0 2528
['', ' UNTRUSTED_RECTANGLE= ', '487 495 0 2528', ' ', ' ']
981  989     0 2528
['', ' UNTRUSTED_RECTANGLE= ', '981 989 0 2528', ' ', ' ']
1475 1483     0 2528
['', ' UNTRUSTED_RECTANGLE= ', '1475 1483 0 2528', ' ', ' ']
1969 1977     0 2528
['', ' UNTRUSTED_RECTANGLE= ', '1969 1977 0 2528', ' ', ' ']
0 2464   195  213
['', ' UNTRUSTED_RECTANGLE= ', '0 2464 195 213', ' ', ' ']
0 2464   407  425
['', ' UNTRUSTED_RECTANGLE= ', '0 2464 407 425', ' ', ' ']
0 2464   619  637
['', ' UNTRUSTED_RECTANGLE= ', '0 2464 619 637', ' ', ' ']
```

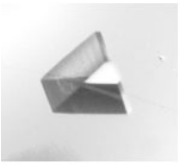
XDSAPP users community

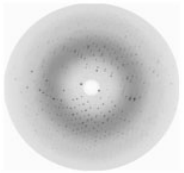
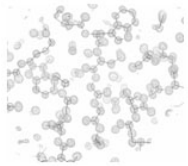
HZB Helmholtz Zentrum Berlin

821 downloaders, distributed over 488 institutes and 43 countries.



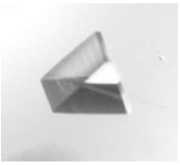
XDSAPP next features

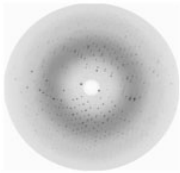
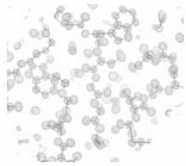


- Improved indexing procedure
- Resolution cutoff based on significant CC1/2
- Handling of Eiger data sets

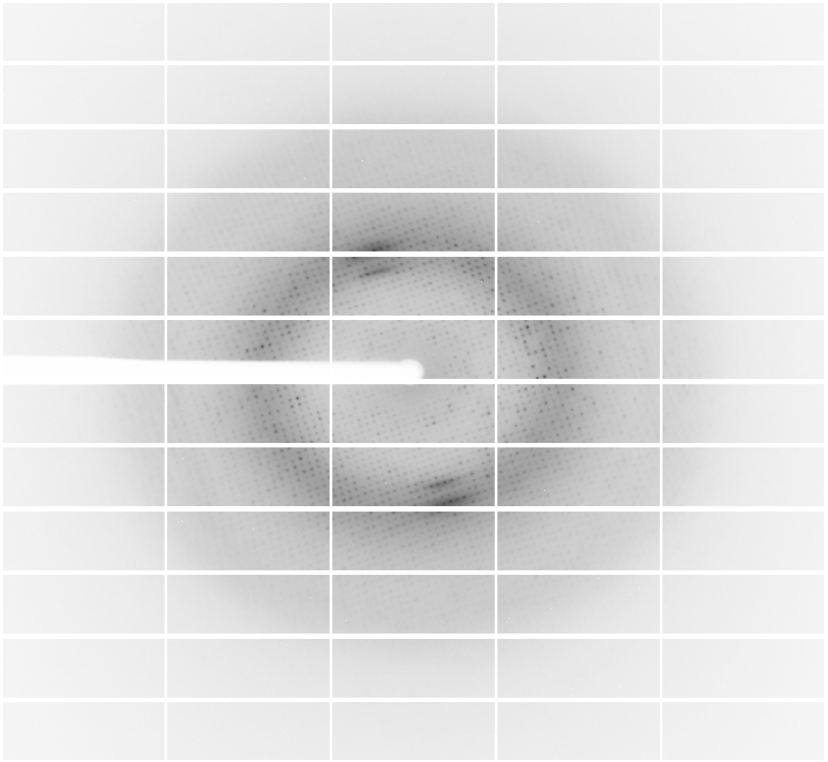
Improved performance



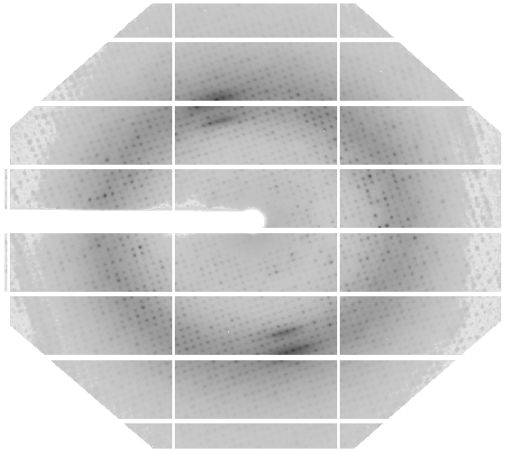



Regions of interest for COLSPOT and IDXREF

Limit resolution in BKGINIT.cbf

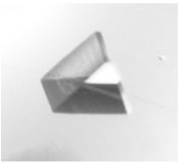


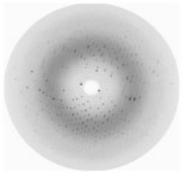
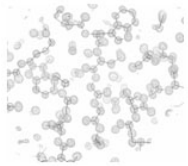
No resolution limit



~ 2.5Å maximum resolution

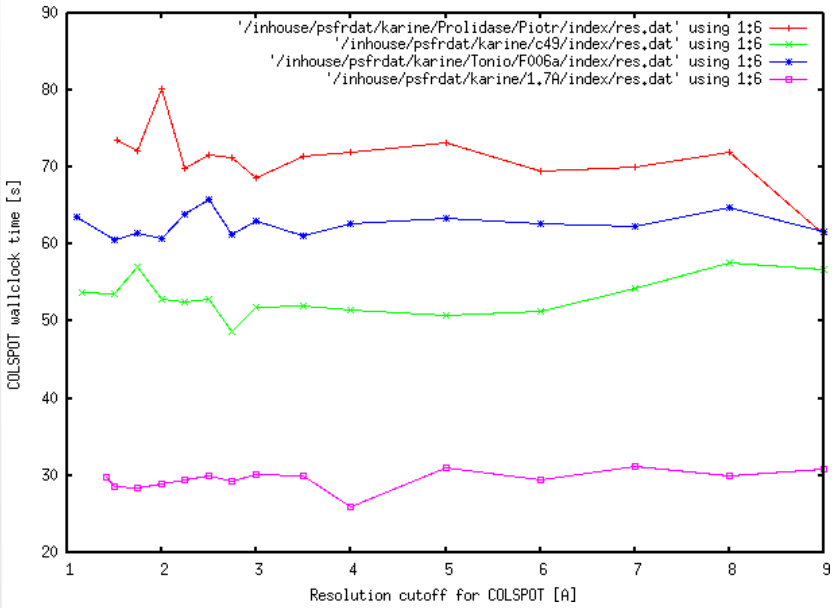
Improved performance



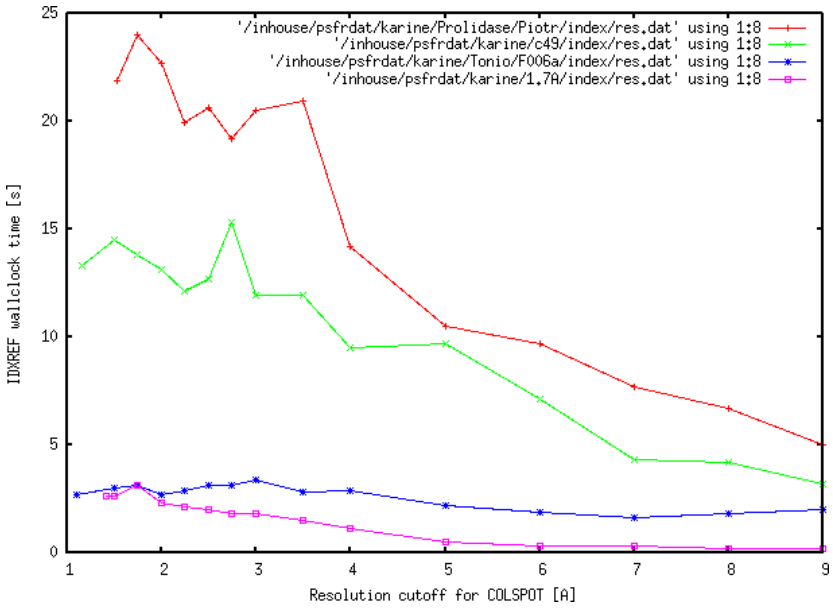



Regions of interest for COLSPOT and IDXREF

Time performance



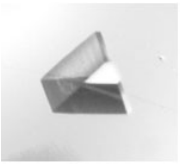
COLSPOT

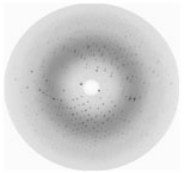
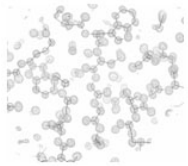


IDXREF

Time performance only improved for IDXREF: collect all spots

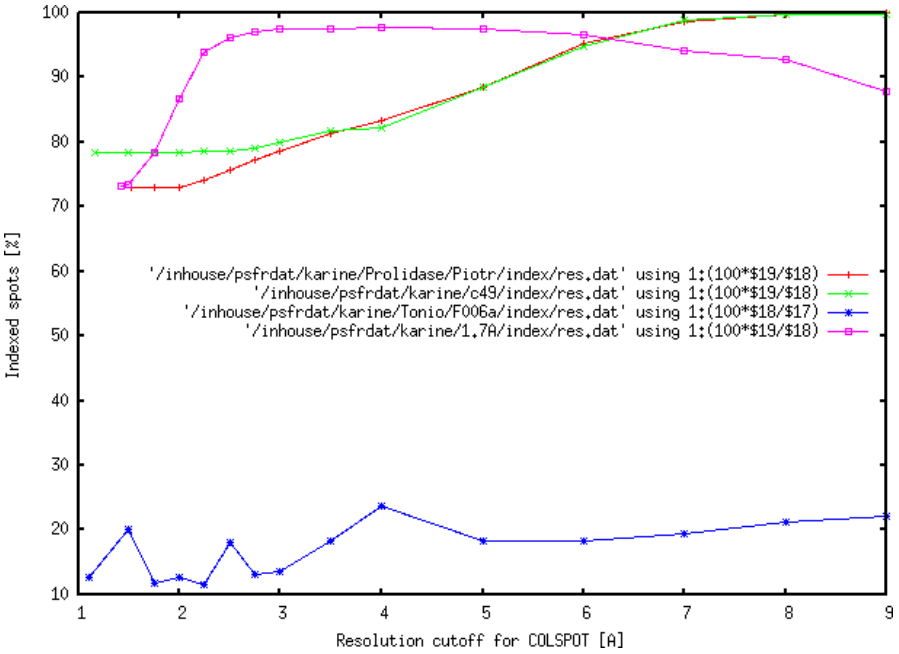
Improved performance



Regions of interest for IDXREF

Better indexing through resolution-based selection of spots for IDXREF



Indexed spots

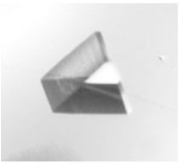
Useful for pathological crystals

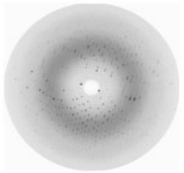
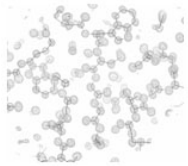
- Mosaic spread
- Split crystals

Easier indexing at low resolution

Successful indexing crucial for further processing steps

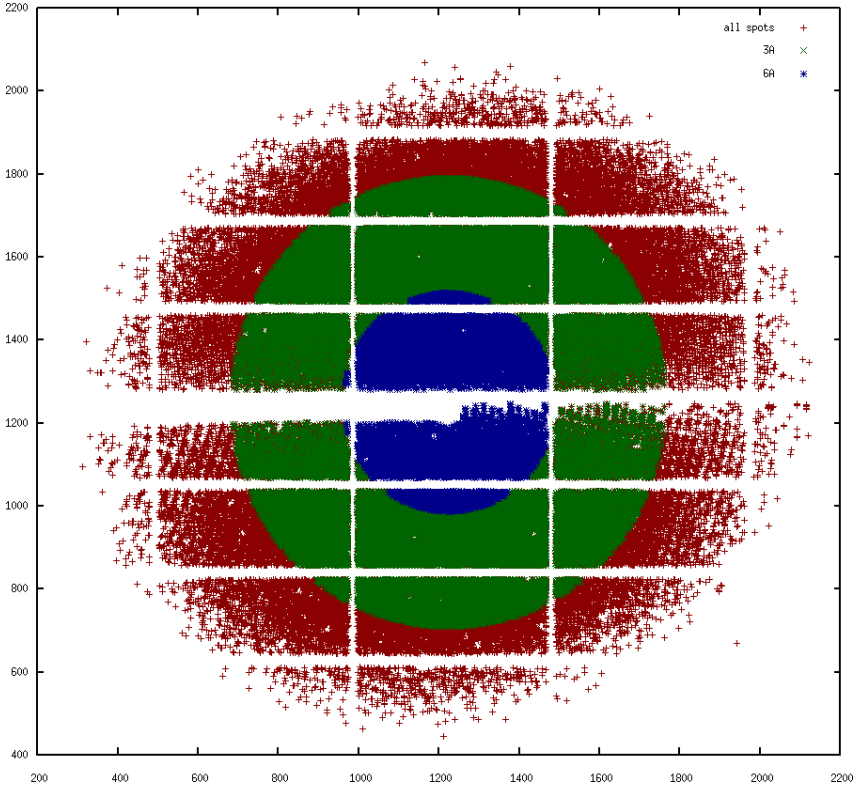
Improved performance



Regions of interest for COLSPOT and IDXREF

Better indexing



Spots used for IDXREF

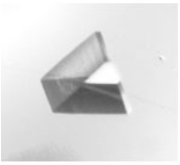
Useful for pathological crystals

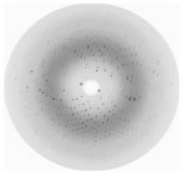
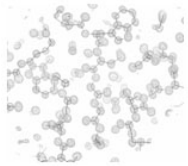
- Mosaic spread
- Split crystals

Easier indexing at low resolution

Successful indexing crucial for further processing steps

Eiger data sets



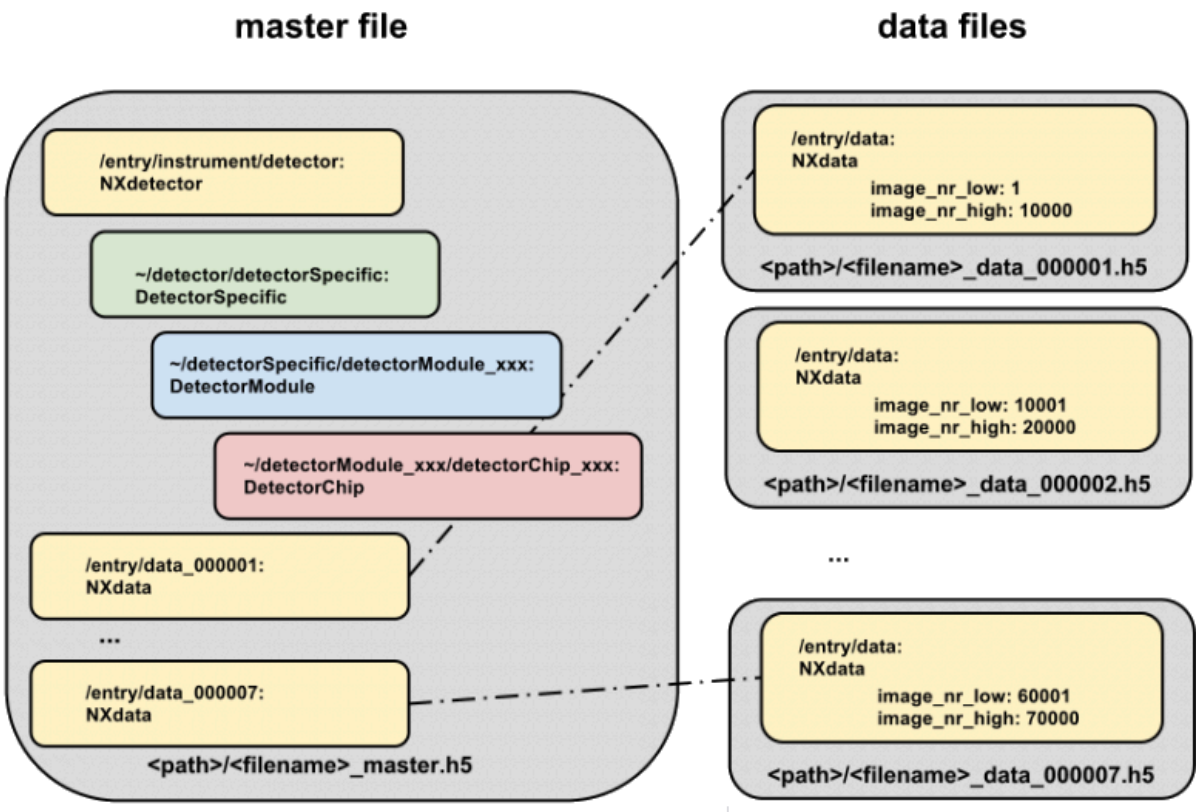
Even faster data collection

Finer φ-slicing

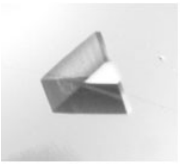
180 000 frames, 360GB / minute

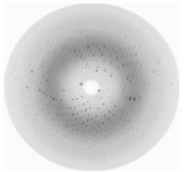
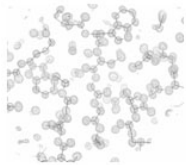
No native reading of .h5 files in XDS

Conversion to .cbf format



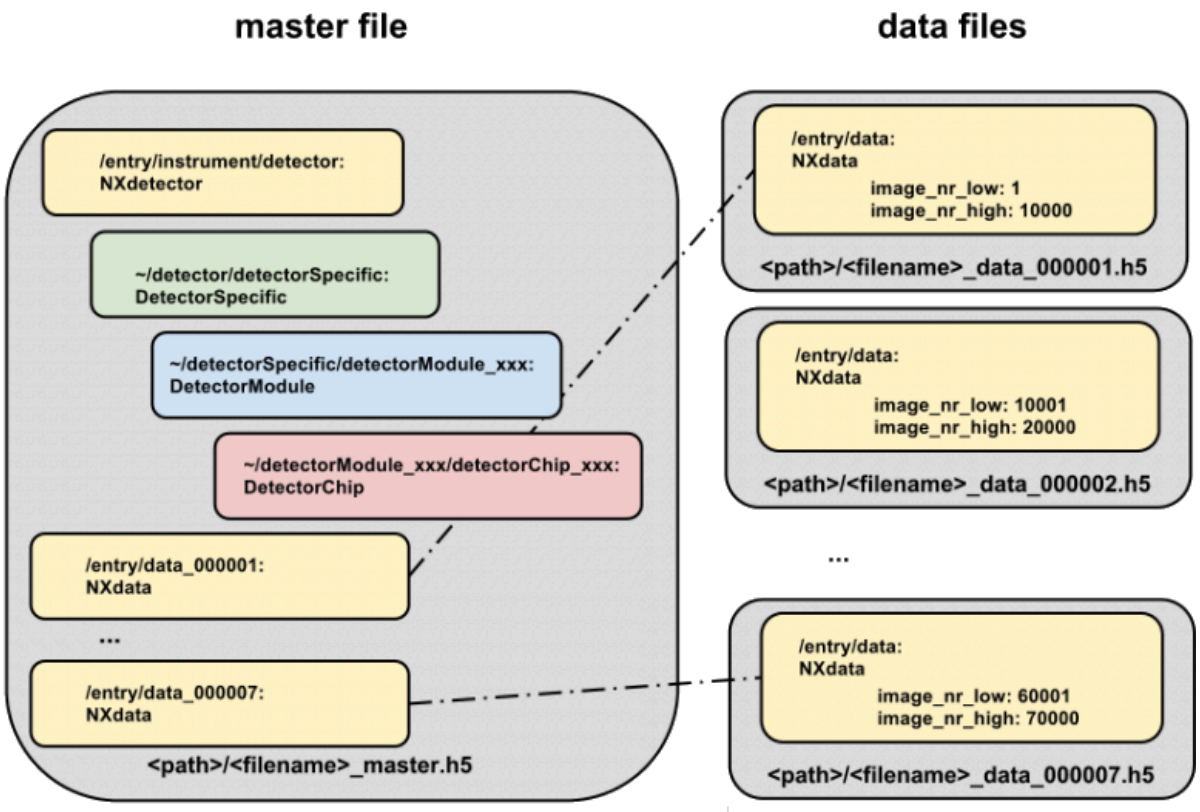
Eiger data sets



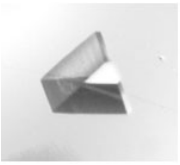



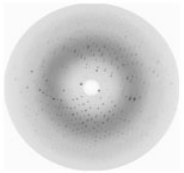
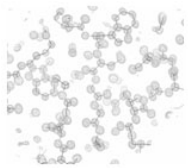
No native reading of .h5 files in XDS

Conversion to .cbf format



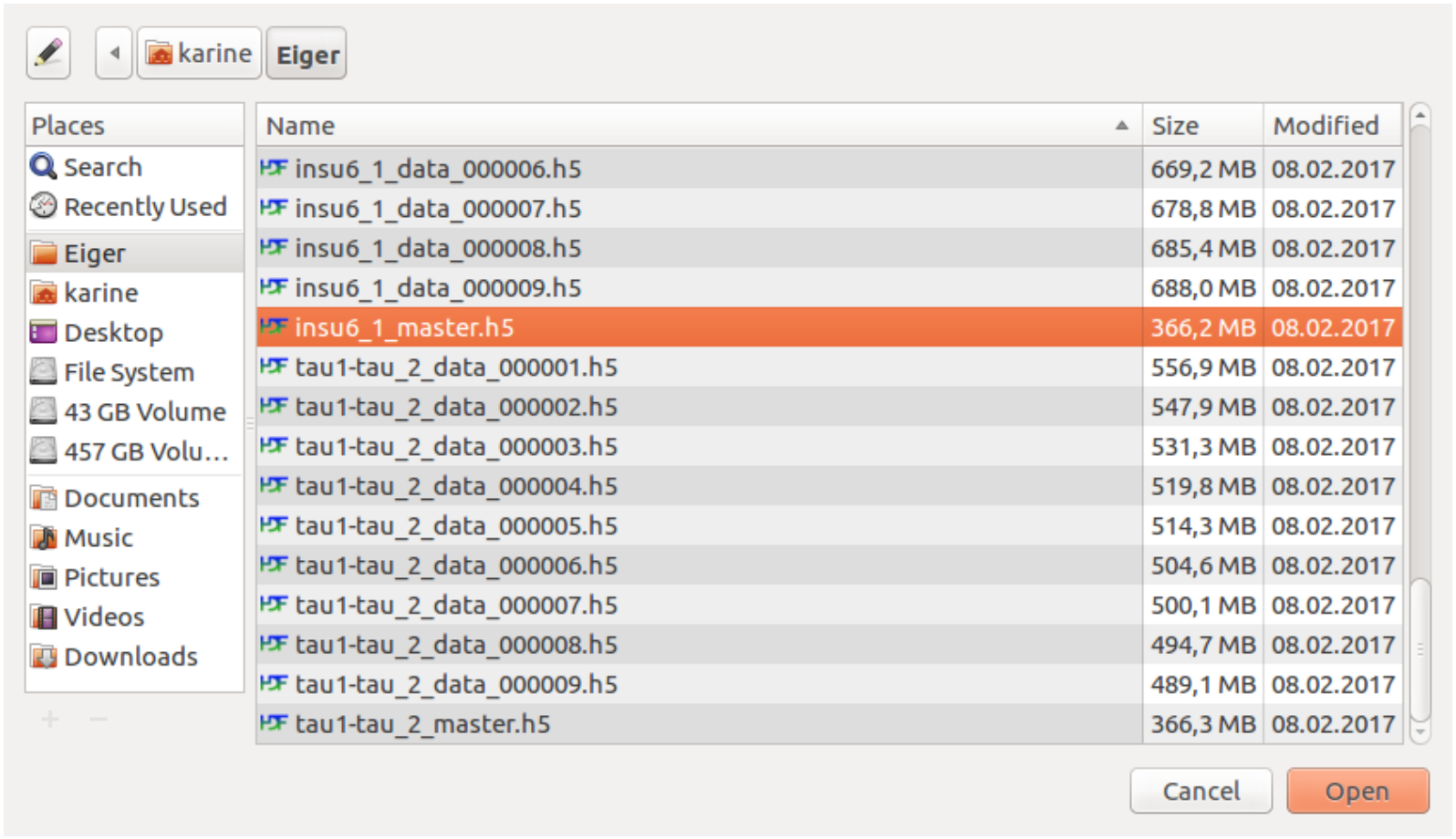
Eiger data sets



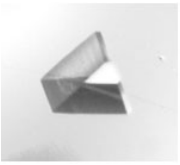



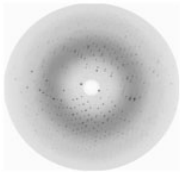
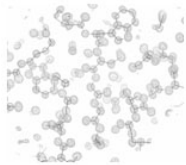
Solution compatible on all operating systems: eiger2cbf

Select either _data_ or _master file



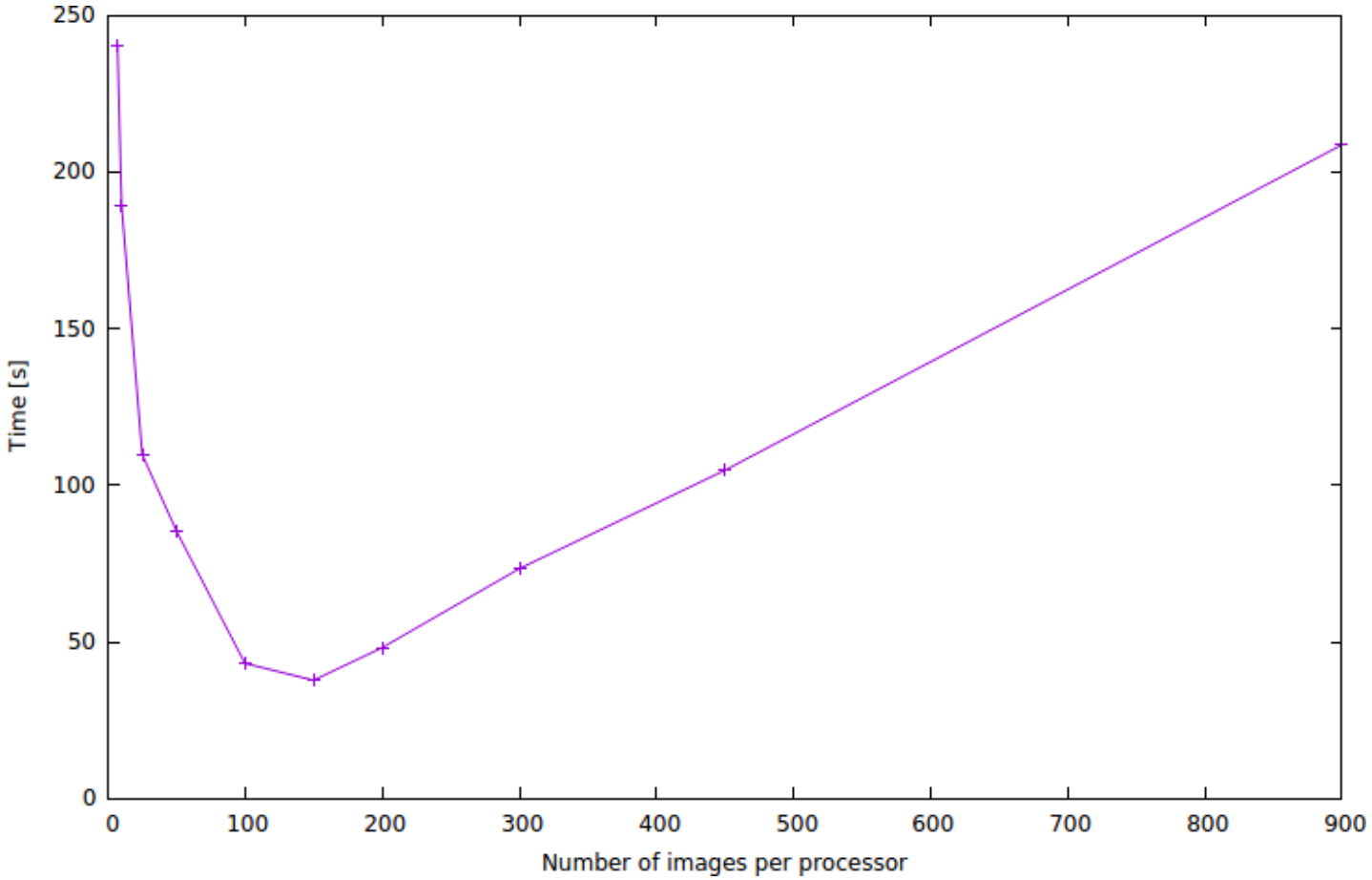
Eiger data sets



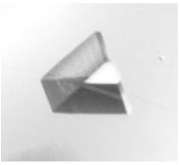



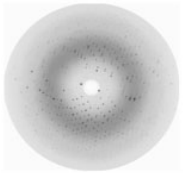
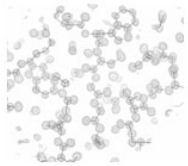
Solution compatible on all operating systems: eiger2cbf

Speed-up: 1 CPU per _data_ file



Eiger data sets



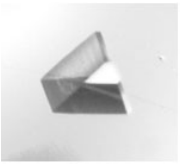



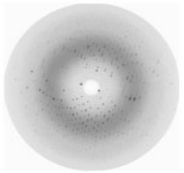
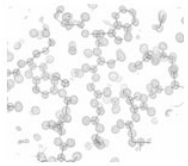
Possibility to load already converted .cbf files from eiger2cbf, thanks to modified generate_XDS.INP script by Olof Svensson (ESRF)

Tasks left

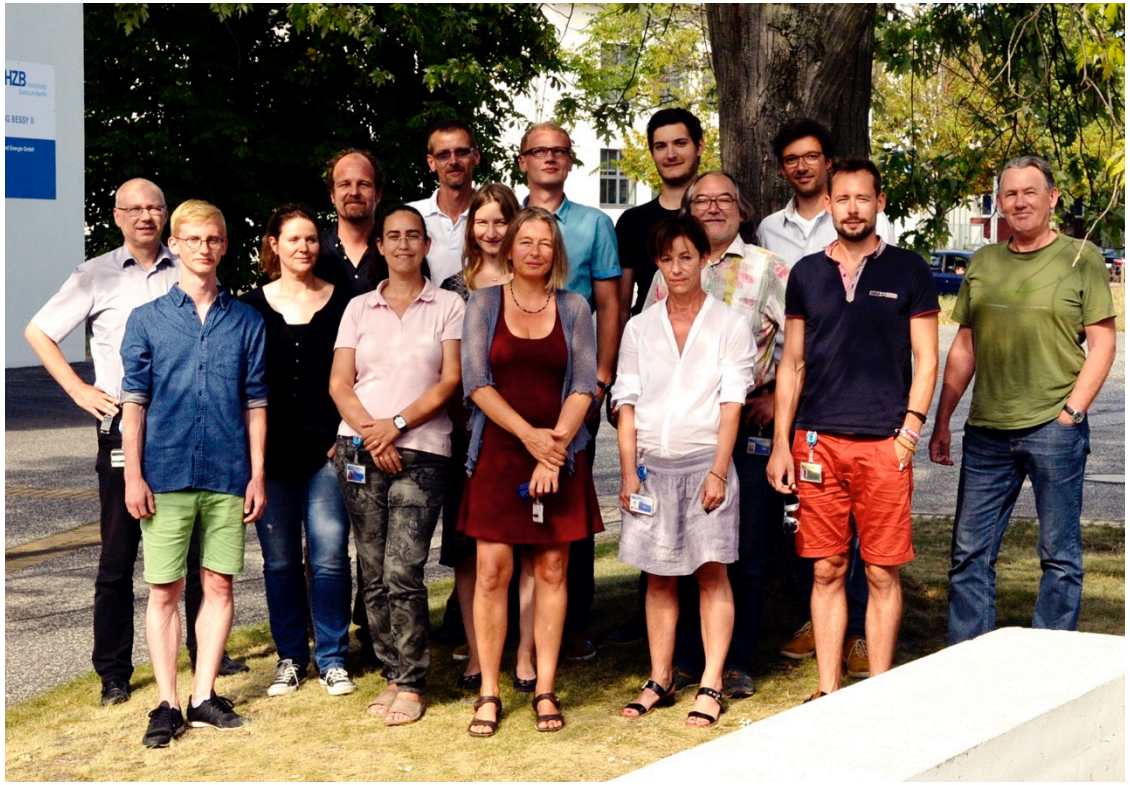
- Get rid of explicit data conversion in XDSAPP by using .h5 files directly in XDS, Linux only (H5ToXds)
- Delete converted .cbf files after processing
- Make the command line version of XDSAPP recognize Eiger data sets

Acknowledgment



Thank you for your attention



- The HZB MX-group
- Micheal Krug
- Beamline users
- Kay Diederichs Universität Konstanz
- Gerhard Klebe‘s group, Universität Marburg
- Röntgen-Ångstrom-Cluster
- Bernhard Lohkamp, Karolinska Institutet