Data reduction and preliminary analysis

Ivan Rajkovic

BioSAXS workshop, March 28, 2016

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

BL4-2 beam layout

Selecting the right distance

Make sure you put the right info on you proposal

Communicate the Rg and other info to the beamline scientists BEFORE the experiment begins

For 11 keV

Distance [m]	Q-range [Å^{-1}]	Max Rg [Å]
3.5	0.0035 - 0.25	~130
2.5	0.005 - 0.35	~95
1.7	0.008 - 0.5	~65
1.1	0.012 - 0.85	~42
0.7	0.018 - 1.2	~26

$q_{min} * Rg = 0.5$

$q * Rg < 1.3$ (Guinier approx.)

Changing the distance

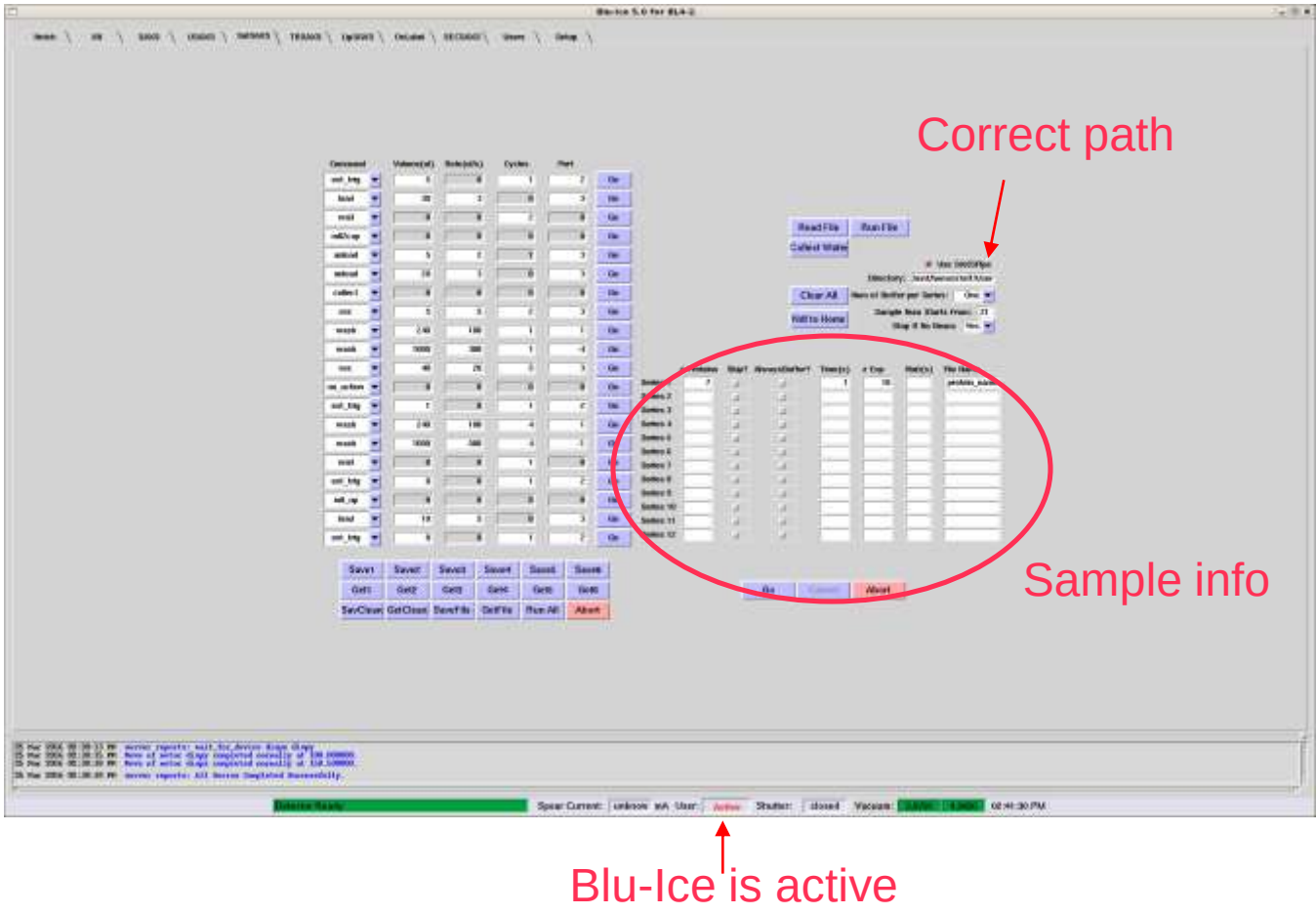
Not like a crystallography beamline!

Steps include:

- Vent vacuums
- Move sample stage
- Insert new pipes
- Pump down vacuums
- Change wavelength
- Align beam to center of instrument
- Align slits - s0, s1 and guard slit
- Set-up autosampler again
- Calibrate detector
- Take test data
- 1-2 hours

Controlling the beam line – Blu-Ice

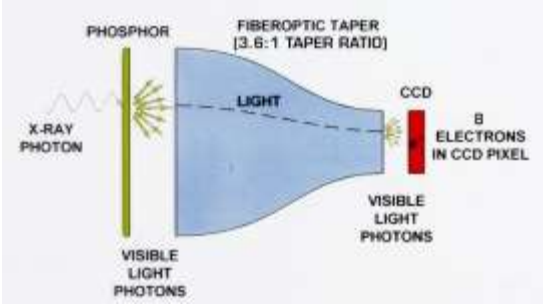
- On your first visit your be allocated linux, windows and Blu-Ice accounts
- Use a right click to open a terminal window on the main Red Hat computer
- Type ‘go’ on the command line to run Blu-Ice
- Select the SolSaxs tab



Collecting data - CCD Detectors

X-Ray CCD detectors work by:

a) A phosphor screen (to convert X-rays to light)
b) A fiber-optic taper (to demagnify the light image down to the size of the CCD chip)
c) A CCD chip to detect the light image as an electric charge on a pixel

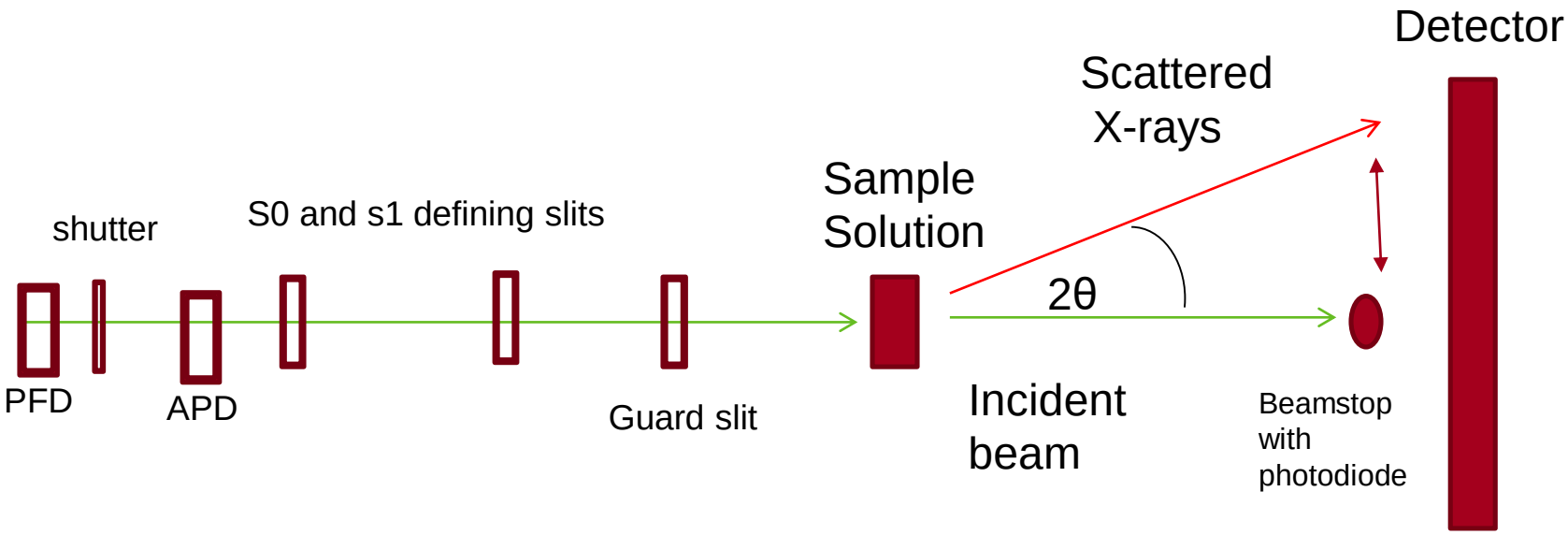


Depending on the 'binning' the count from a number of pixels will be averaged to produce the final reading

Low binning gives higher resolution, but lower accuracy

We generally use the highest possible binning – 8x8 on our Rayonix MX225-He

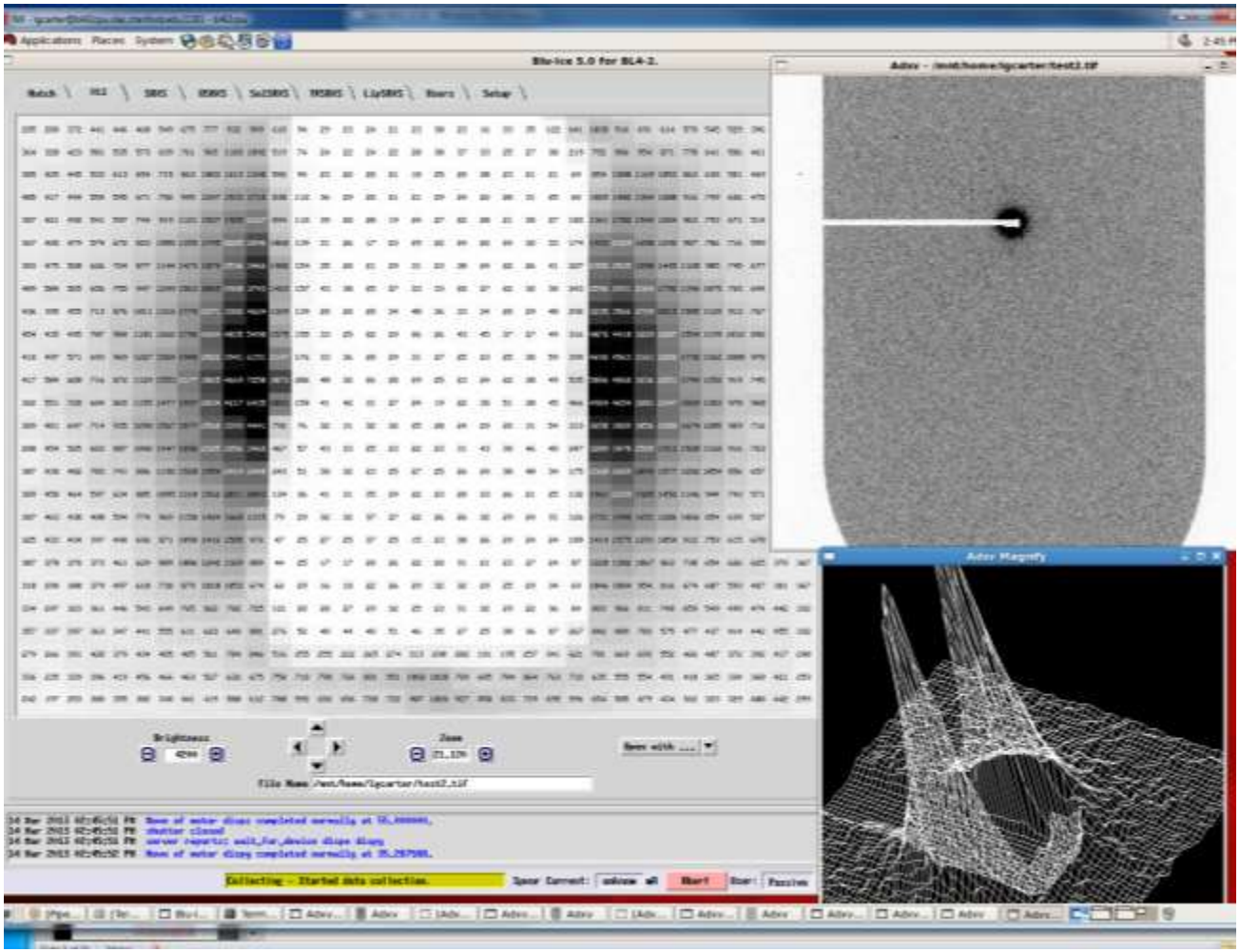
Beam layout



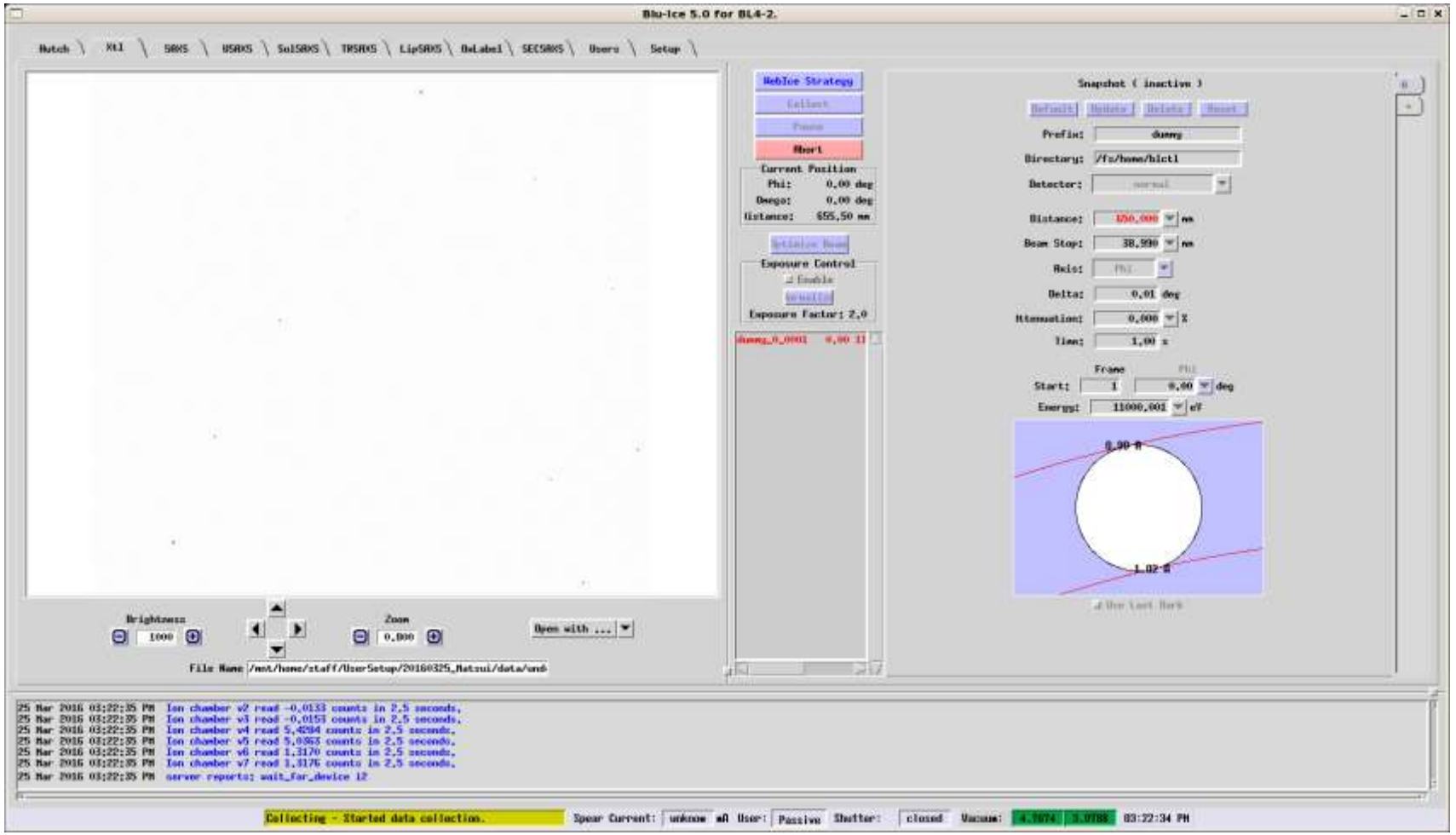
The intensity at q is measured, where q is the angular dependence defined as:

$$q = 4\pi \sin\theta / \lambda$$

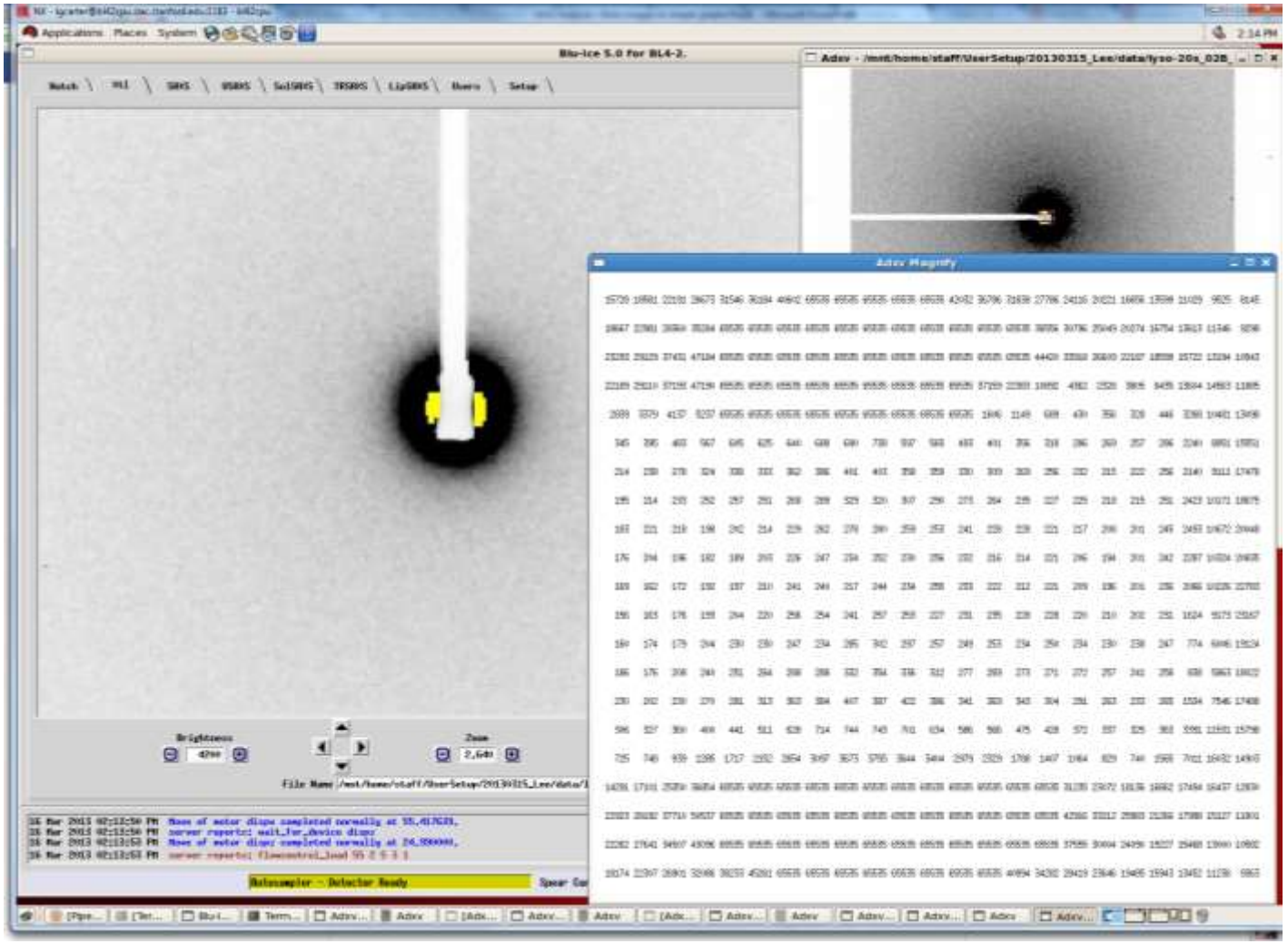
Typical images – nothing in the beam



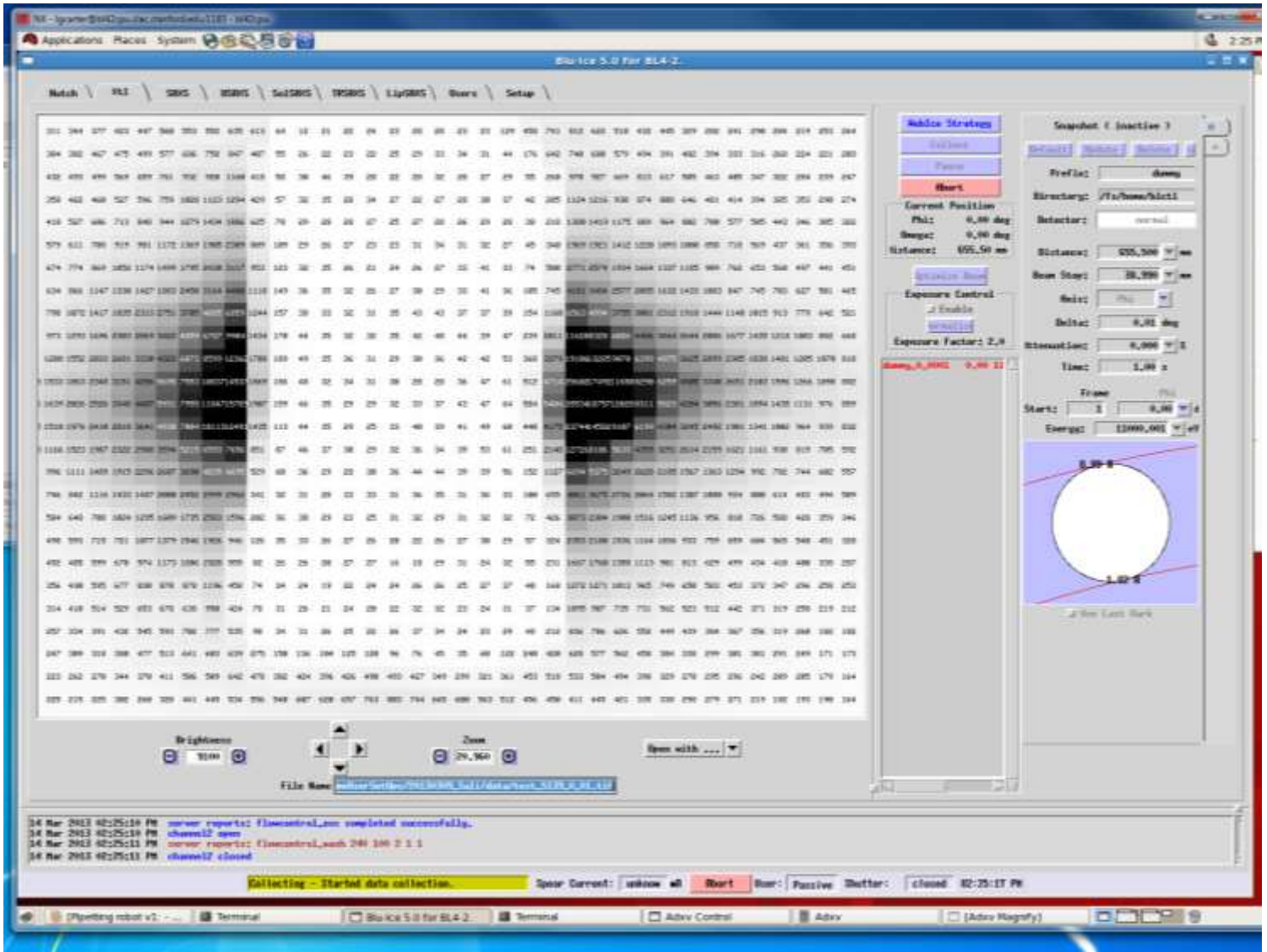
No beam


Blu-Ice 5.0 for BL4-2.
Hutch
XtI
SAXS
USAXS
SolSAXS
TRSAXS
LipSAXS
BeLabel
SECSAXS
Users
Setup
WebIce Strategy
Collect
Pause
Abort
Current Position
Phi: 0.00 deg
Omega: 0.00 deg
Distance: 655.50 mm
Exposure Control
Enable
Exposure Factor: 2.0
Snapshot (inactive)
Prefix: dummy
Directory: /fs/home/blctl
Detector:
Distance: 150.000 mm
Beam Stop: 38.990 mm
Axis: Phi
Delta: 0.01 deg
Attenuation: 0.000 %
Time: 1.00 s
Frame
Phi
Start: 1 0.00 deg
Energy: 11000.001 eV
0.90 A
1.02 A
Brightness
1000
Zoom
0.000
Open with ...
File Name: /mnt/home/staff/UserSetup/20160325_Matsui/data/und
25 Mar 2016 03:22:35 PM Ion chamber w2 read -0.0133 counts in 2.5 seconds.
25 Mar 2016 03:22:35 PM Ion chamber w3 read -0.0153 counts in 2.5 seconds.
25 Mar 2016 03:22:35 PM Ion chamber w4 read 5.4284 counts in 2.5 seconds.
25 Mar 2016 03:22:35 PM Ion chamber w5 read 5.0353 counts in 2.5 seconds.
25 Mar 2016 03:22:35 PM Ion chamber w6 read 1.3170 counts in 2.5 seconds.
25 Mar 2016 03:22:35 PM Ion chamber w7 read 1.3176 counts in 2.5 seconds.
25 Mar 2016 03:22:35 PM server reports: wait_for_device 12
Collecting - Started data collection.
Spear Current: unknow mA User: Passive Shutter: closed Vacuum: 4.7074 3.0788 03:22:34 PM

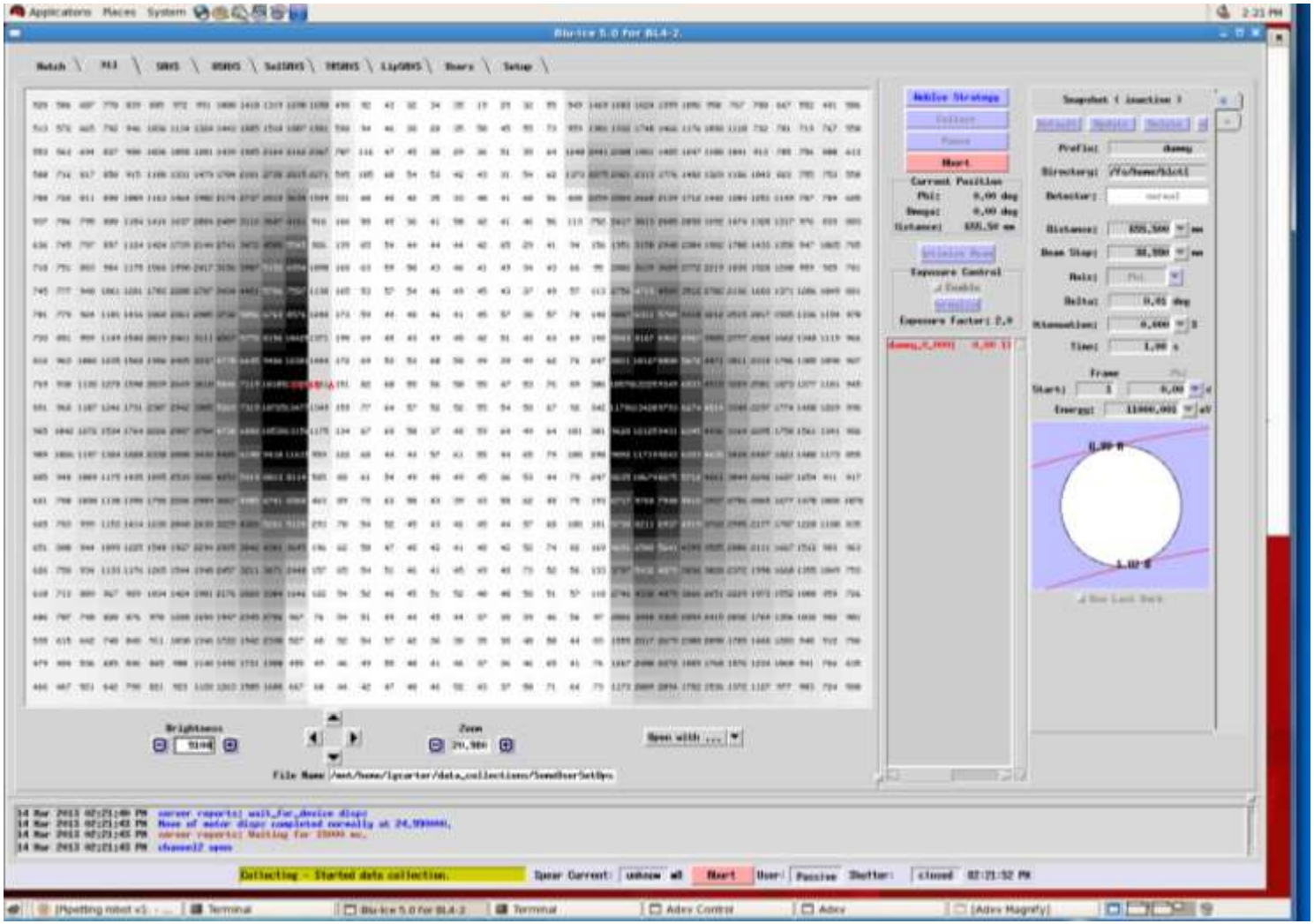
Overloaded pixels



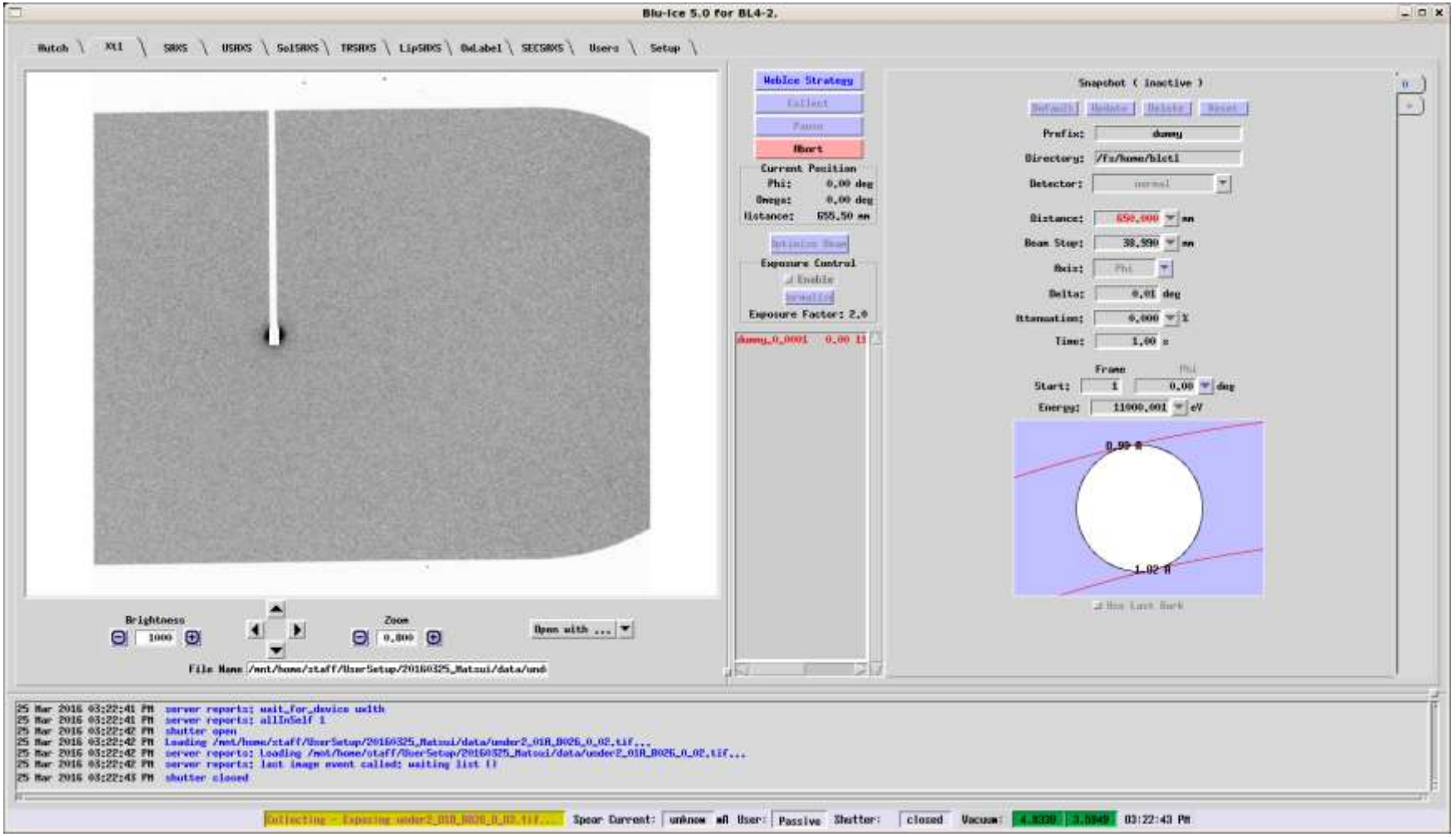
Misaligned beam



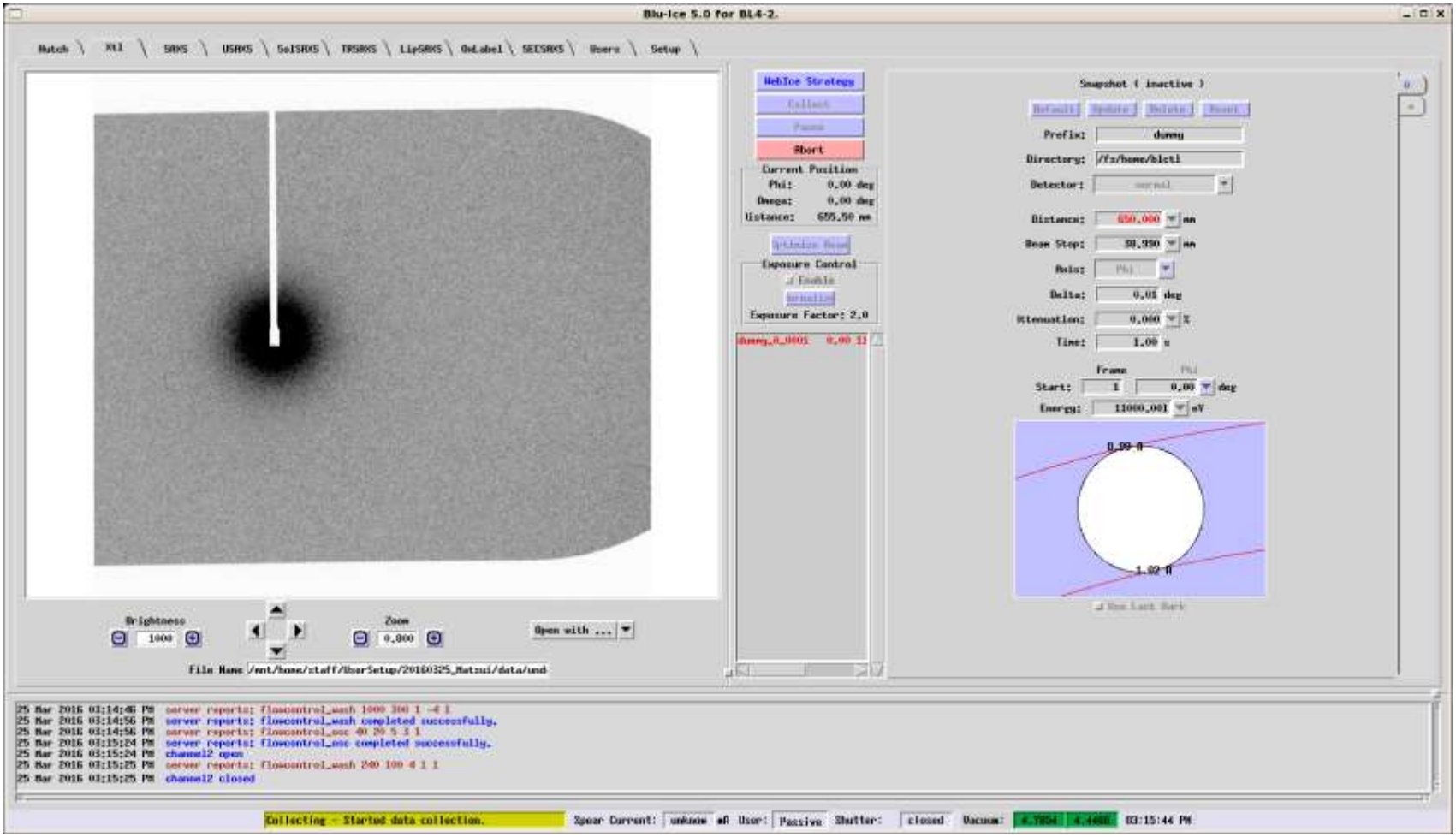
High background – dirty capillary or windows



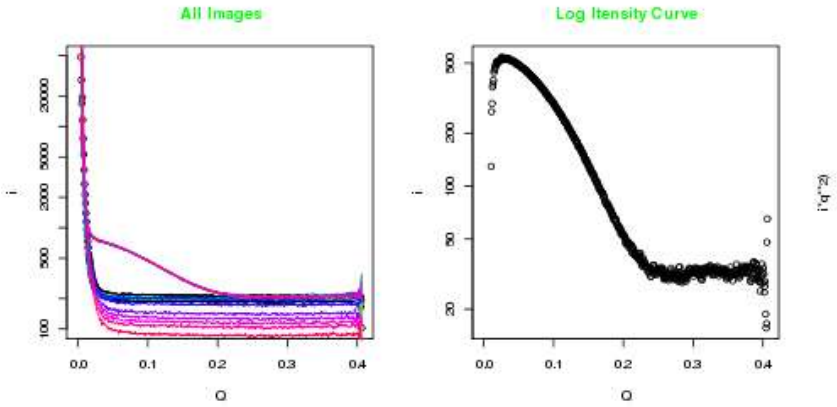
Typical buffer image

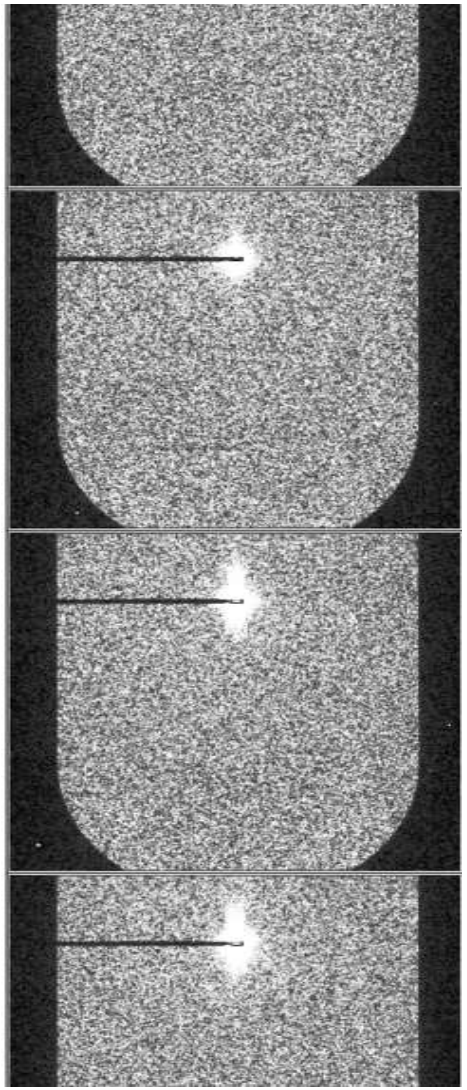


Typical sample image

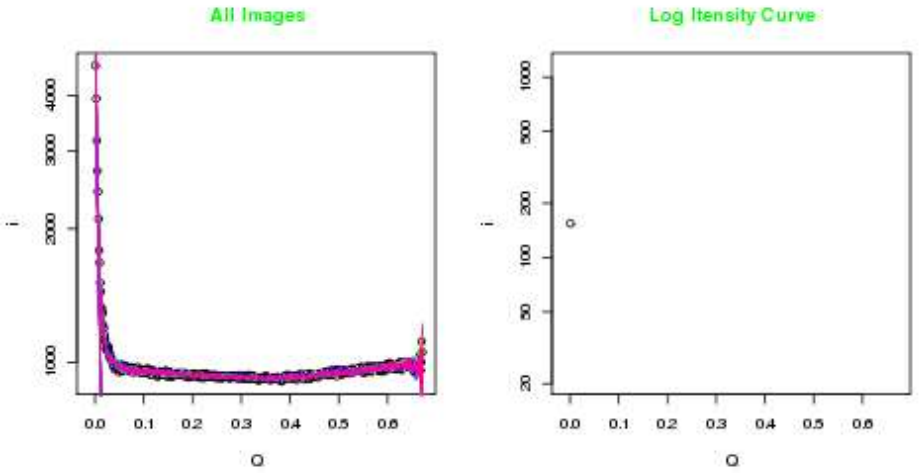


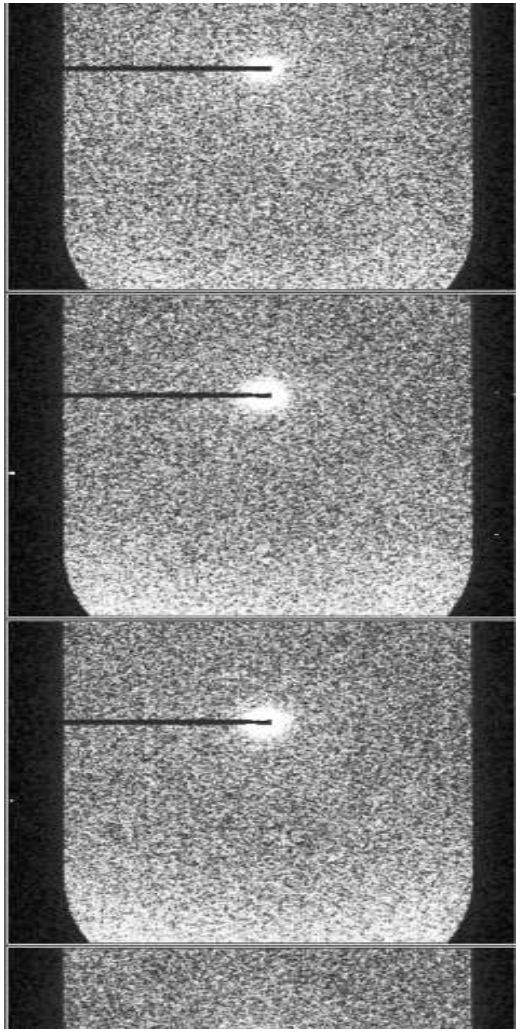
Air bubbles!



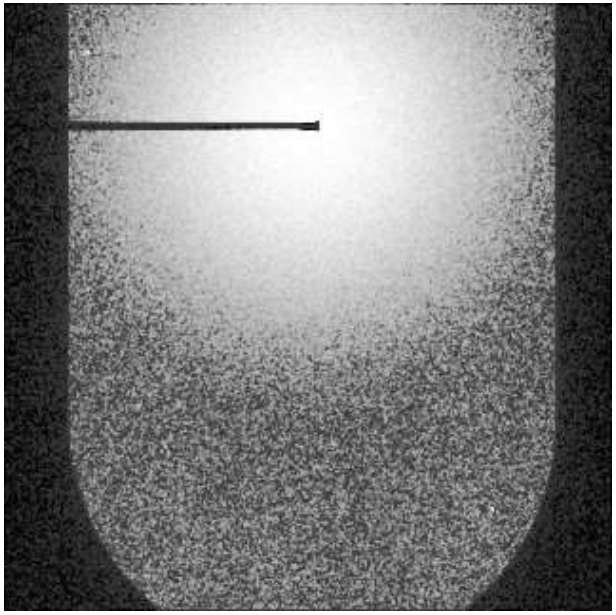


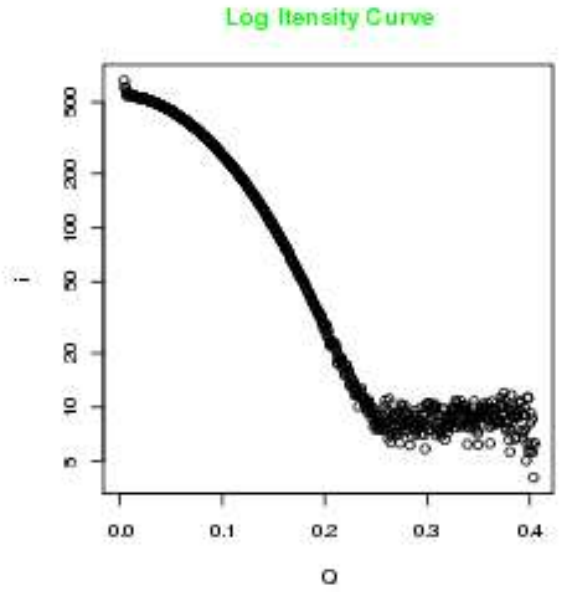
No SAMPLE!


All Images
Log Itensity Curve
i
Q
0.0
0.1
0.2
0.3
0.4
0.5
0.6
1000
2000
3000
4000
20
50
100
200
500
1000



The SAXS profile: 2D image -> 1D graph





What we need to calculate the profile?

1. The masking out of the beamstop

2. The beam center

3. A mapping of the pixels to a known q value

4. A value to scale each image by

Silver Behanate

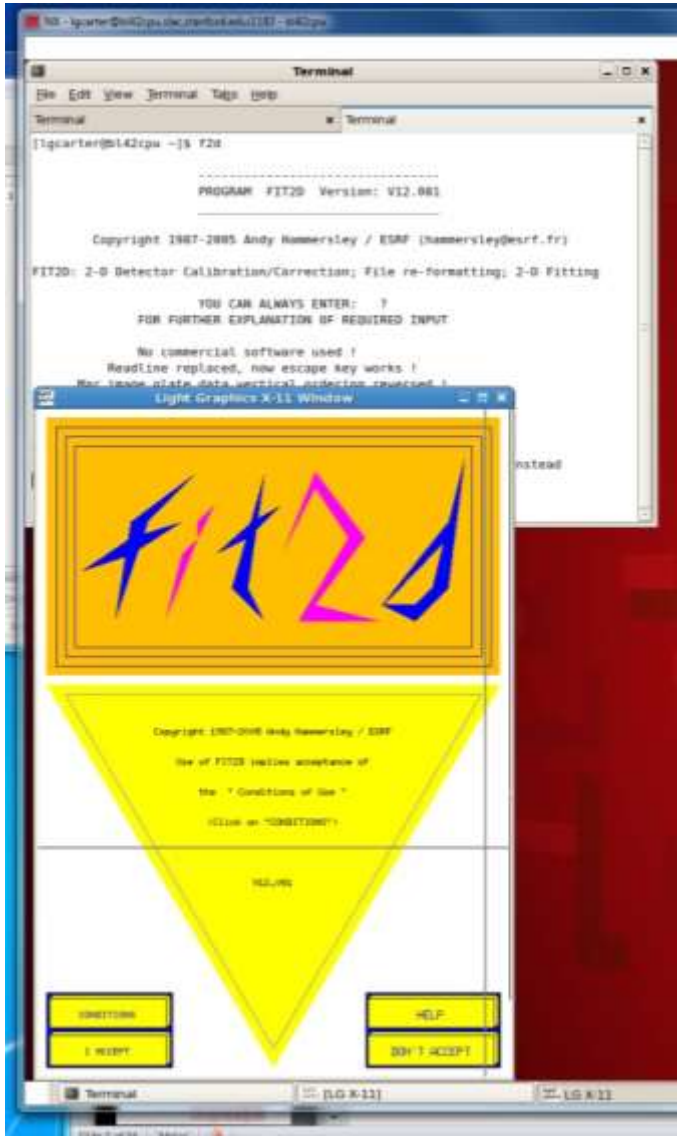
- Silver behenate is a silver salt of the long-chain fatty acid behenic acid
- Powder diffraction from Silver behenate can be used to find the beam center and map pixels to q
- powder diffraction rings of Silver behenate occur at known intervals:

Reciprocal space

	$2\pi/$ Å		
58.380 Å	→	0.107	Å^{-1}
29.190 Å		0.215	
19.40 Å		0.328	

$AgC_{22}H_{43}O_2$

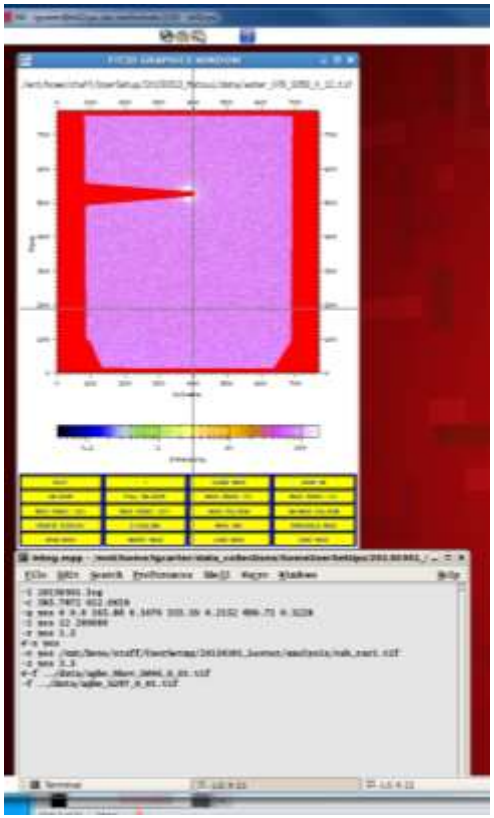
Using FIT2d to examine images



Free software provided by the ESRF

Can be run on bl42cpu by typing f2d

Creating a mask using fit2d



Preferably done on a water image or similar with lower contrast

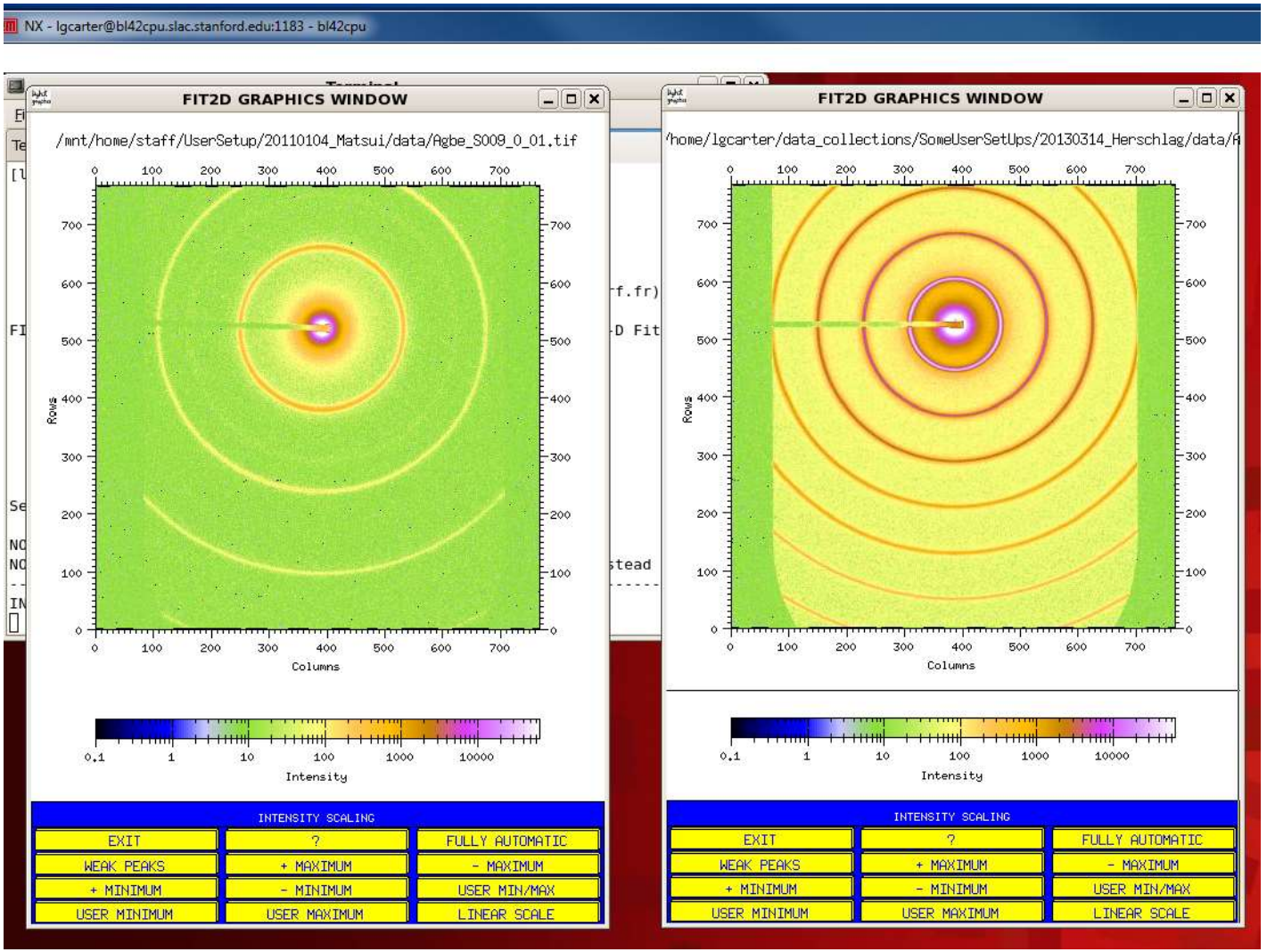
Using the Mask tool in the powder diffraction menu the sides of the detector and beamstop are masked out

The mask is then saved in tif format

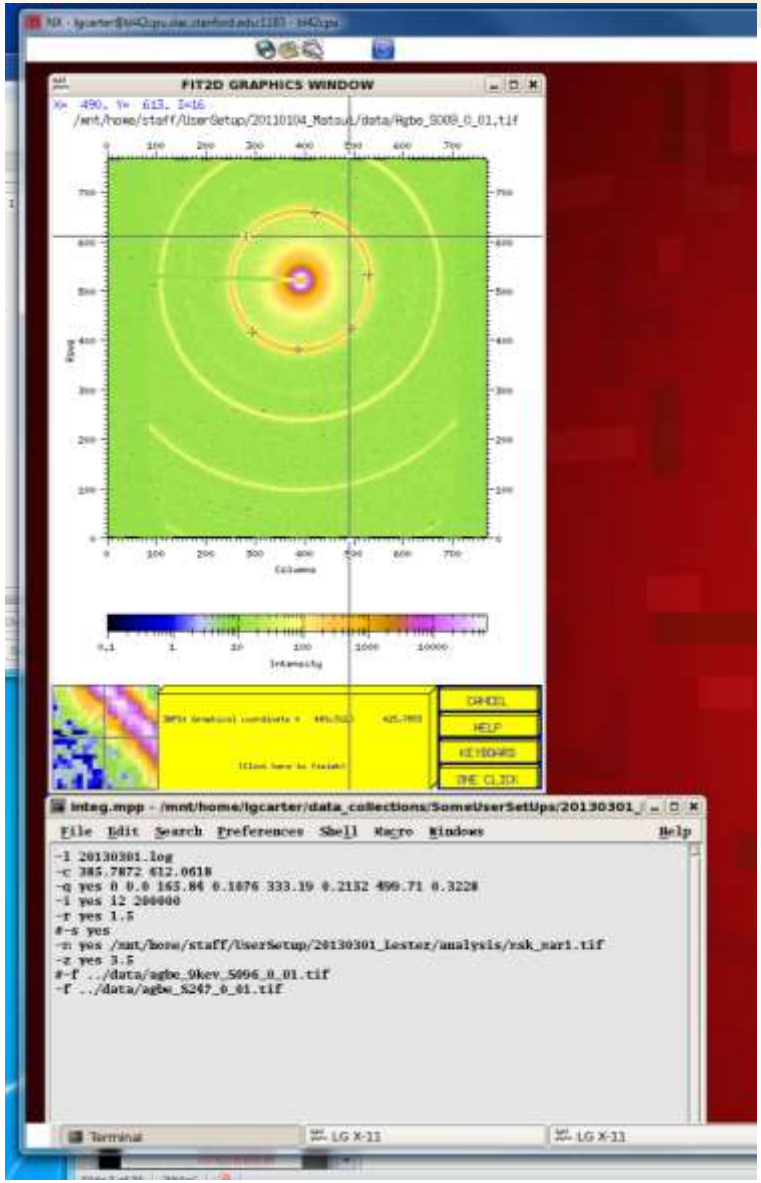
Silver behenate scattering

2.5 m

1.1 m

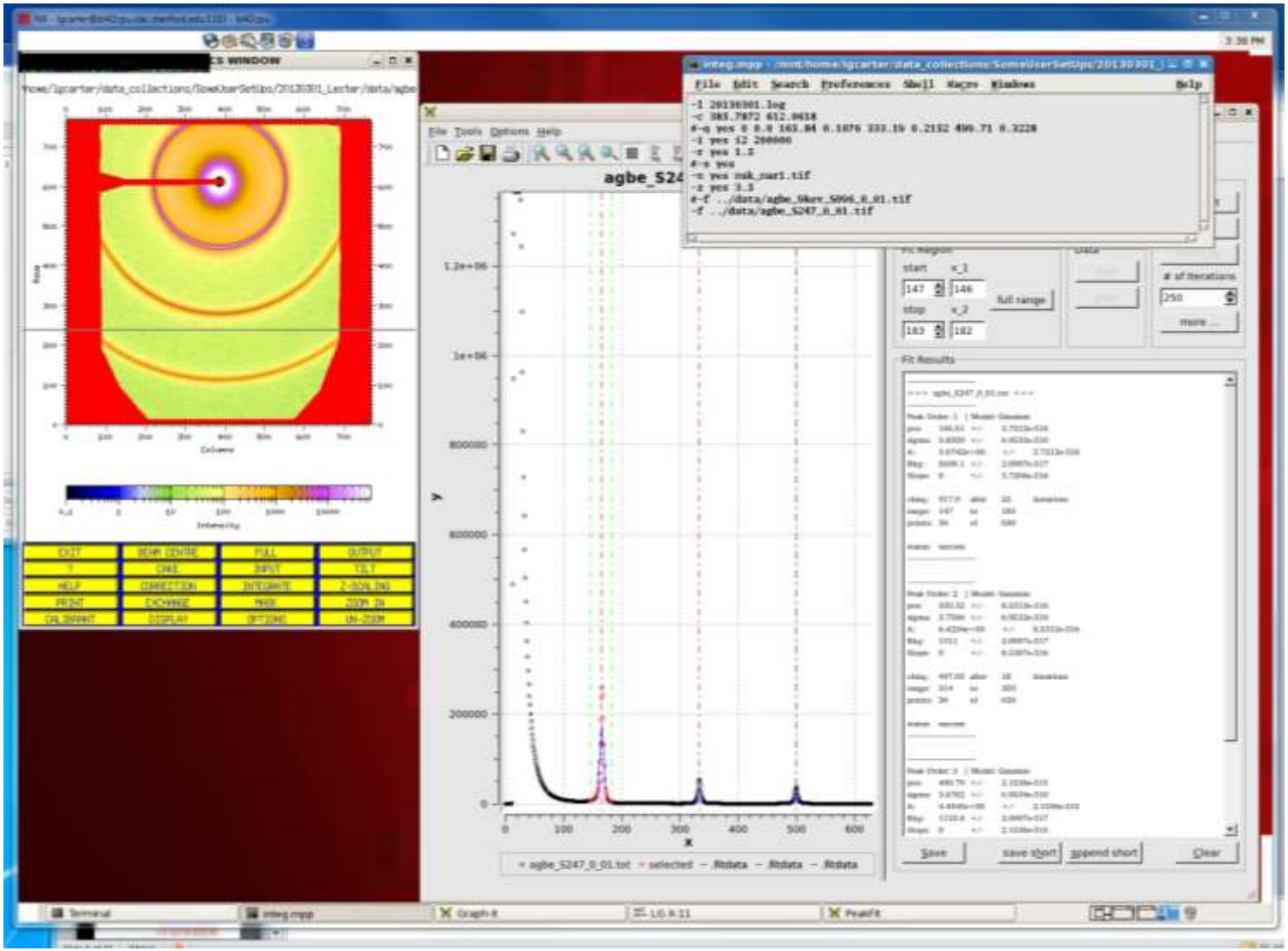


Finding the beam center using AgBE and fit2d



- Once an image is loaded in it can be Z-scaled to make it clearer
- Under the powder diffraction menu is ‘beam center’ option
- This scan be used to select several points around the first ring, which can then be used to find the image center

Calibrating the q-space



Sastool

Sastool is our in house tool for convert images to intensity profiles

The is also a Windows GUI version available

Both can be downloaded from:

- http://ssrl.slac.stanford.edu/~saxs/analysis/sastool.htm

Sastool

- reads the tiff raw image file
- obtains and subtracts offsets, normalizes for beam intensity
- converts 2-D grid data to 1-D distance from beam center
- converts distance from pixels to q-space
- calculates statistics for each image frame (average and standard deviation) (this generates a .dat file for each frame)
- compares statistics and use them to include or reject frames in the sum of the whole series (this generates a .tot file for each series)
- and subtracts the buffer intensity from the sample intensity (generating the final .sub file)

Main Sastool parameters

SLAC

```
-l sample.log              # The log file name.
-c 343.4691 250.5907 # Beam center coordinates in pixels.
-q yes 0 0.0 102 0.1076 205 0.2152 308 0.3228          #  Conversion from pixel to q
-i yes i1 1000             # Data normalization factor
-z yes 3.5                 # Remove zingers
-r yes 2.0                 # Rejection of individual frames from the summation of the whole series
-s yes                     # Subtract buffer intensity from individual sample frame.
-m yes ./mask.tif          # Use mask file for data integration
-f ../sample03_0_01.tif ../buffer03_0_01.tif       #  Files for SasTool to process.
```

Typical Sastool run

Usually you will have a terminal (right click desktop) open and cd-ed into the analysis directory

You can run gedit in the background to edit integ.mpp:

- gedit integ.mpp &

You will need to type in the buffer and sample pairs at the end of the file

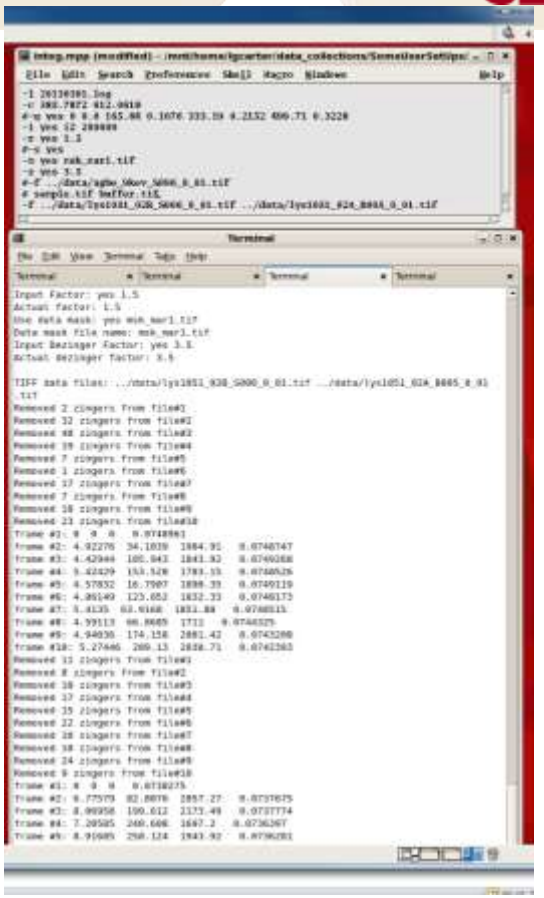
- -f sample.tif buffer.tif

These can be found by searching data directory:

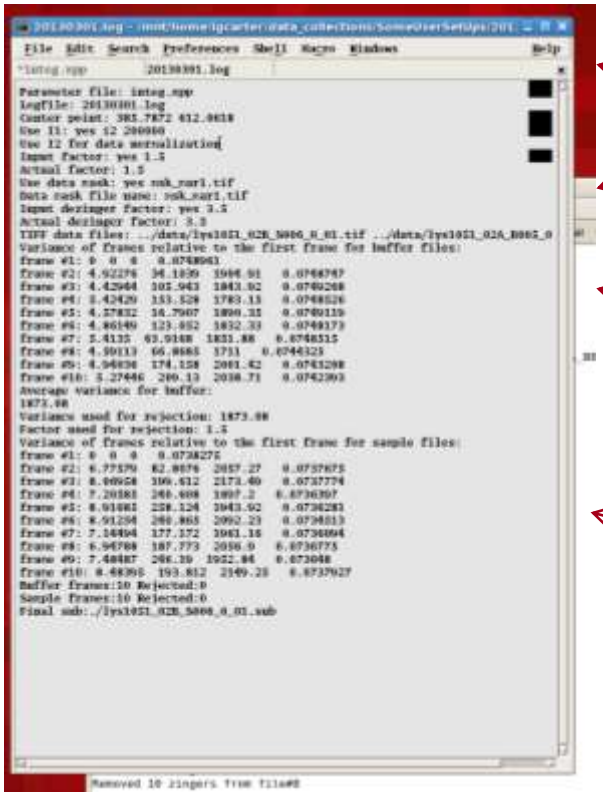
- ls –lrt ../data/*01.tif

Save the file and run sastool

- sastool integ.mpp



Sastool log file



Input parameters

Buffer files

Samples files

Frame information

```
Data mask file name: msk_mar1.tif
Input dezinger factor: yes 3.5
Actual dezinger factor: 3.5
TIFF data files: ../data/lys1051_02B_S006_0_01.tif ../data/lys1051_02A_B005_0
Variance of frames relative to the first frame for buffer files:
frame #1: 0  0  0   0.0748961
frame #2: 4.92276  34.1039  1904.91   0.0748747
frame #3: 4.42944  105.943  1843.92   0.0749268
frame #4: 5.42429  153.528  1783.15   0.0748526
frame #5: 4.57832  16.7907  1890.35   0.0749119
frame #6: 4.86149  123.052  1832.33   0.0748173
frame #7: 5.4135  63.9168  1851.88   0.0748515
frame #8: 4.59113  66.8685  1711   0.0744325
frame #9: 4.94036  174.158  2001.42   0.0743208
frame #10: 5.27446  209.13  2038.71   0.0742393
Average variance for buffer:
1873.08
Variance used for rejection: 1873.08
Factor used for rejection: 1.5
Variance of frames relative to the first frame for sample files:
frame #1: 0  0  0   0.0738275
frame #2: 6.77579  82.8076  2057.27   0.0737675
frame #3: 8.06958  199.612  2173.49   0.0737774
frame #4: 7.20585  240.608  1697.2   0.0736397
frame #5: 8.91685  258.124  1943.92   0.0736281
frame #6: 8.91254  240.865  2092.23   0.0734513
frame #7: 7.14494  177.172  1961.16   0.0736094
frame #8: 6.94788  187.773  2056.9   0.0736775
frame #9: 7.48487  246.19  1952.84   0.073648
frame #10: 8.48395  193.812  2149.25   0.0737927
Buffer frames:10 Rejected:0
Sample frames:10 Rejected:0
Final sub:./lys1051_02B_S006_0_01.sub
```

The first number is frame variance

The second number is the summation of difference (normalized by standard error) between the frame and the first frame

The third number is the summation of the square of the difference (normalized by standard error) between the frame and the first frame

The forth number is the ratio of I2/I1, i.e. transmission

The APD and beamstop photodiode reading

The photodiode is located in the beamstop

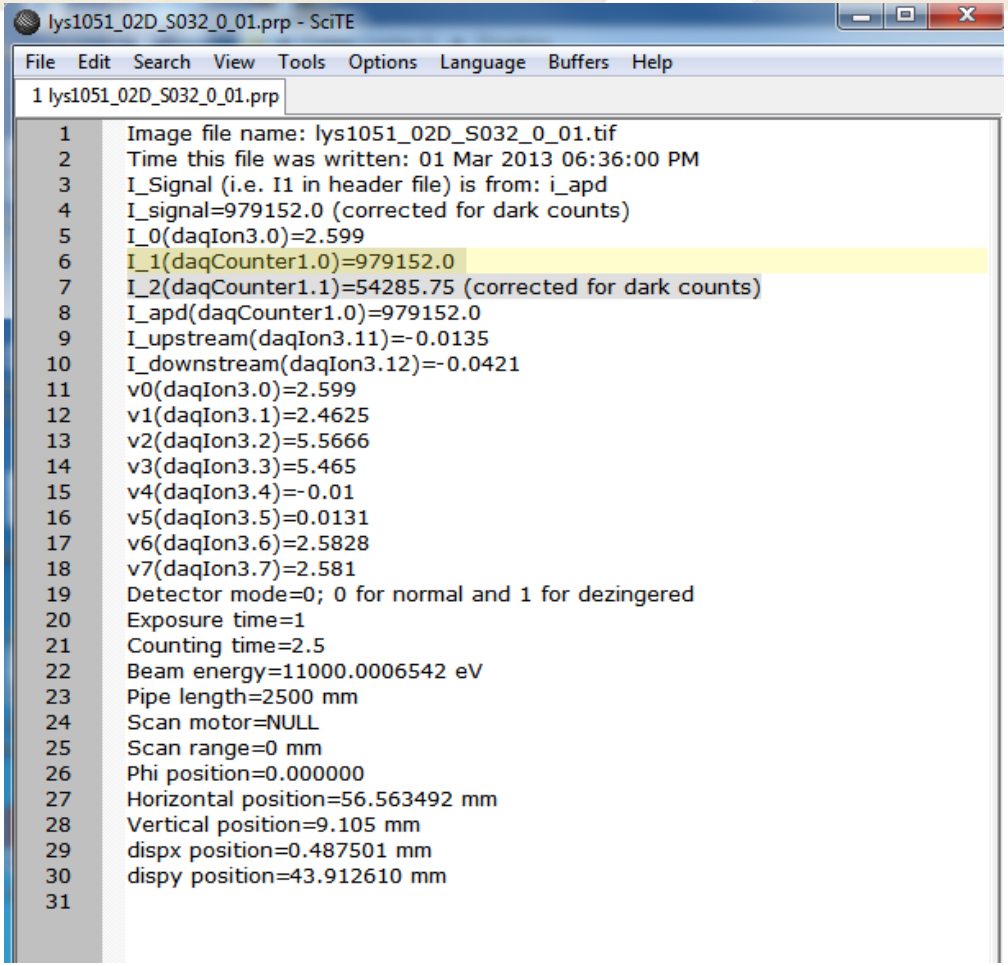
The photodiode value is readout and written to the image header.
There is usually an offset, so the value will never be 0. If the value falls significantly during data collection (on similar samples) this could be a sign of a problem

You can check the values through the run by using grep:

grep "I_1" *.prp
grep "I_2" *.prp

Data .prp files

Contain the recorded intensities from the photodiodes and information about the experimental setup



$$\frac{\text{Photodiode (i2)}}{\text{Avalanche Photodiode (i1)}} = \text{transmission}$$

Sastool outputs

A .dat file for each individual image

A .tot file for the buffer and the sample

A .sub subtracted file

The final .sub file

q Int Int_{std}

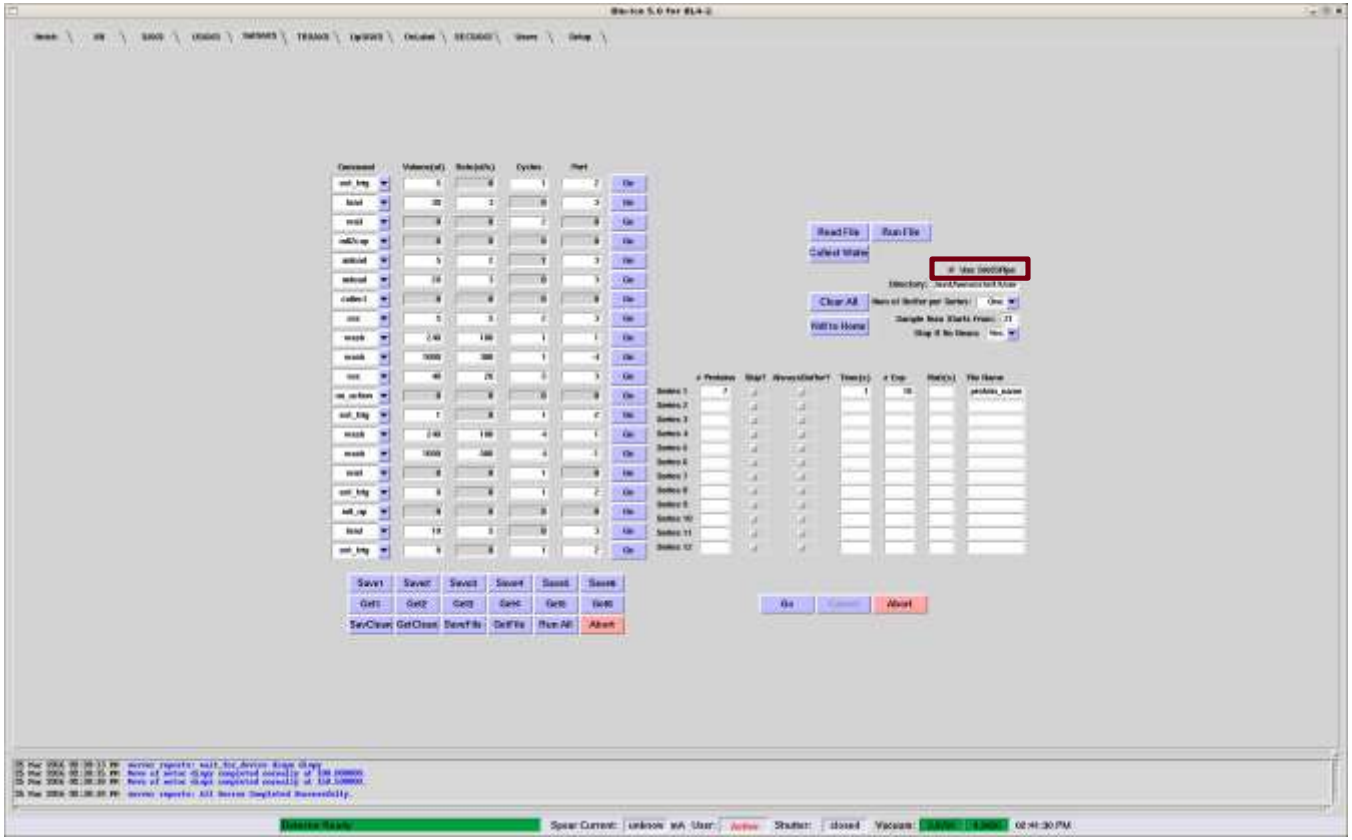
bsa2_135_02D_S048_0_01.sub - SciTE

File Edit Search View Tools Options Language Buffers Help

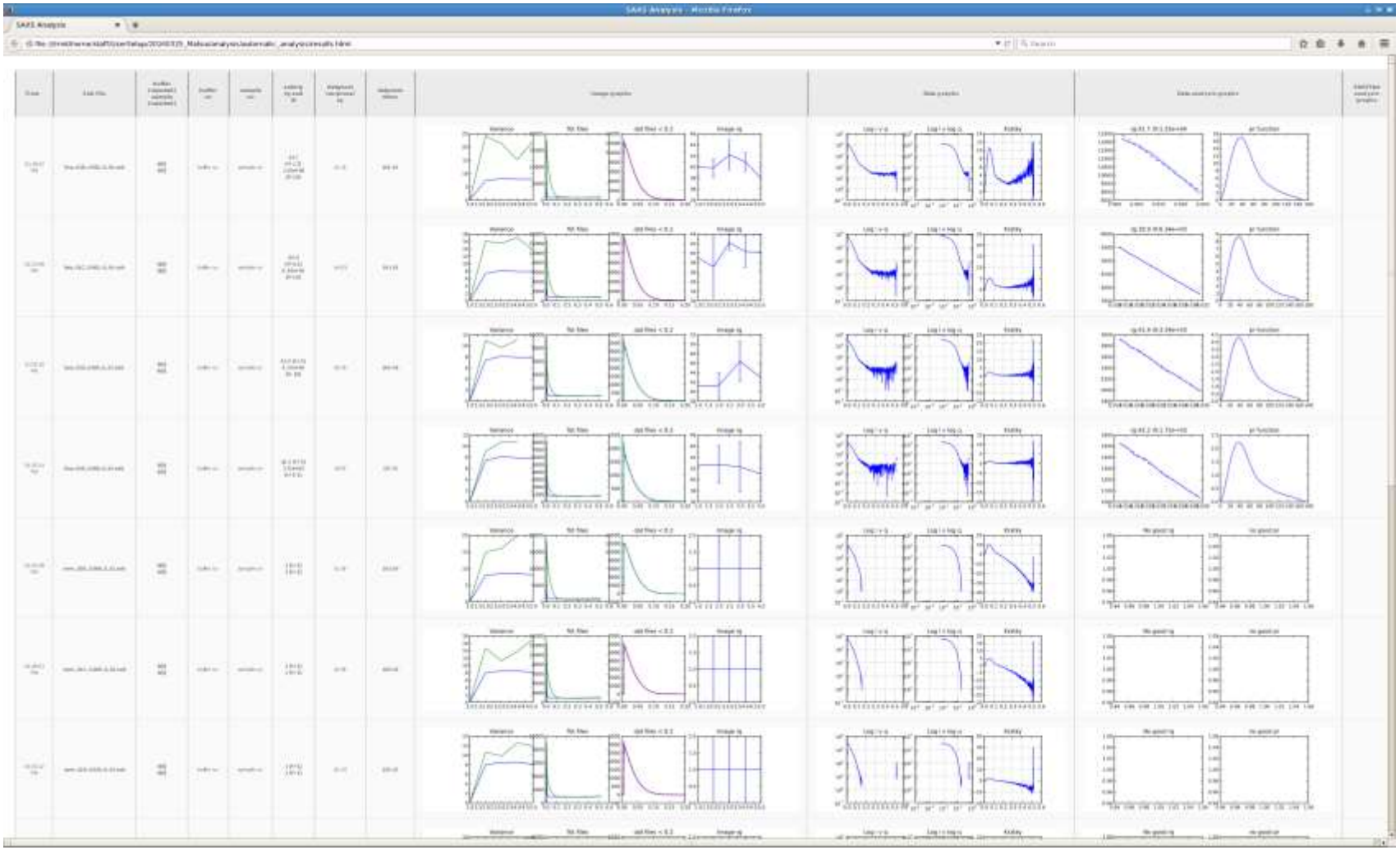
1 bsa2_135_02D_S048_0_01.sub

```
0.000308826 0 0
0.00123902 0 0
0.00216922 0 0
0.00309941 0 0
0.00402961 0 0
0.0049598 0 0
0.00589 0 0
0.00682019 0 0
0.00775039 0 0
0.00868058 7211.85 129.286
0.00961078 7083.24 110.458
0.010541 6857.91 67.8044
0.0114712 6816.67 59.1862
0.0124014 6760.16 44.5936
0.0133316 6626.99 47.7453
0.0142618 6525.03 30.9577
0.0151919 6437.85 29.5986
0.0161221 6322.71 24.2302
0.0170523 6224.81 25.1823
```

SAXSPipe - automatic analysis pipeline

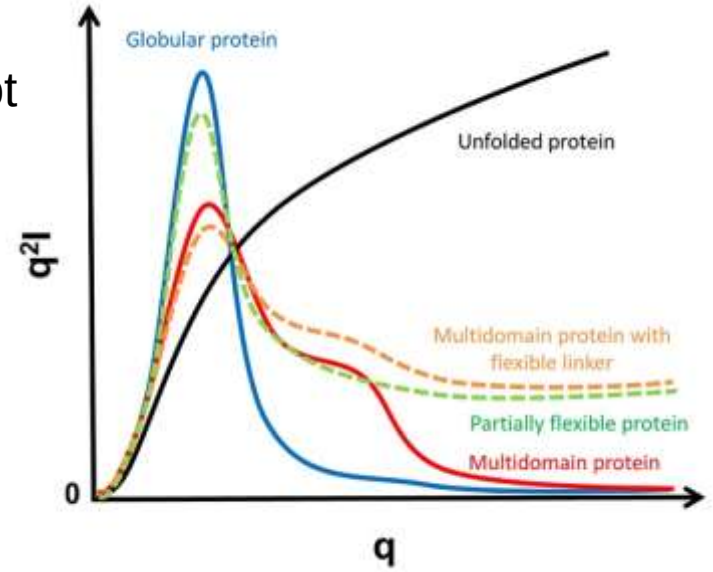


SAXSPipe - automatic analysis pipeline

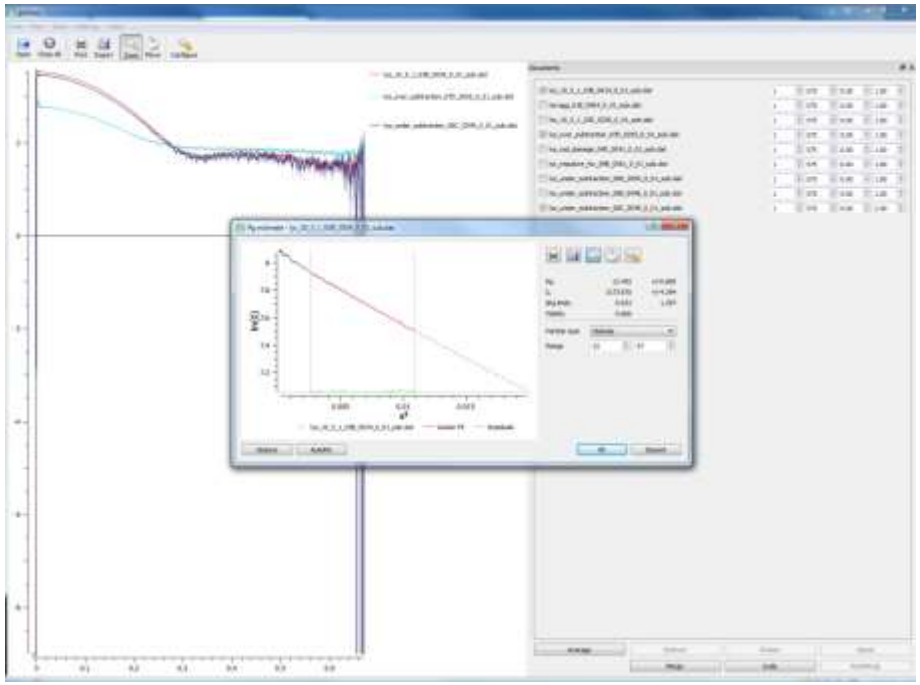


Data quality check

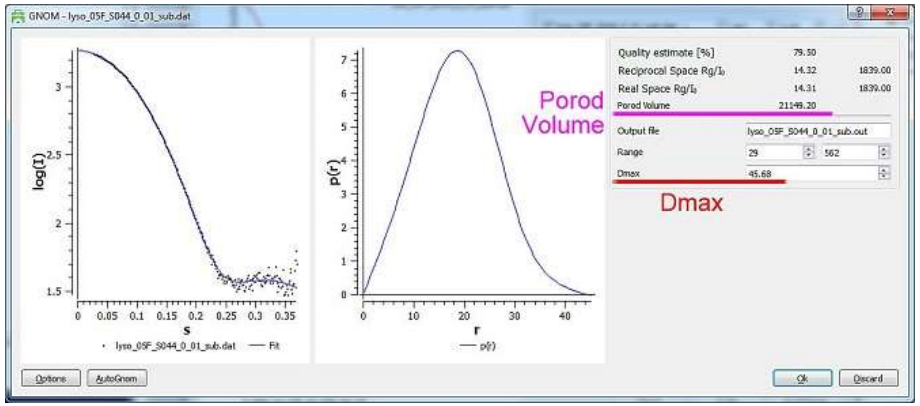
Kratky plot



Guinier plot



Pair-distribution function



Basic analysis process

Check images for issues

Run sastool – check for rejected images

- Air bubbles
- Radiation damage

Plot data

Look for upturn or down turn (aggregation and inter-particle interference)

Check Guinier plot

Check Pr function

Check Kratky plot

Carry out concentration series – are there any concentration effects?

Then move on with analysis