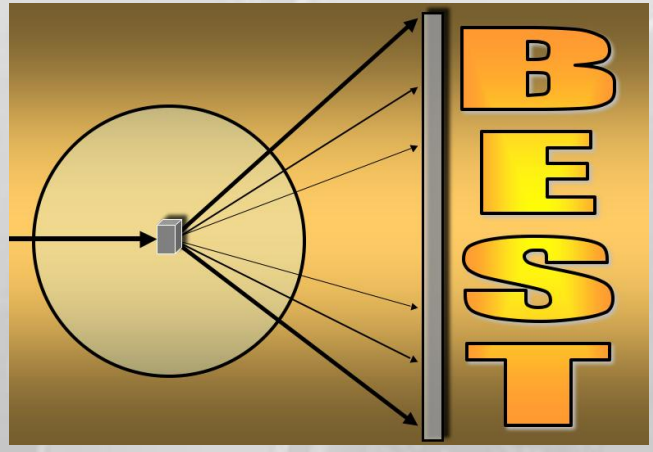
A Light for Science

European Synchrotron Radiation Facility

BEST

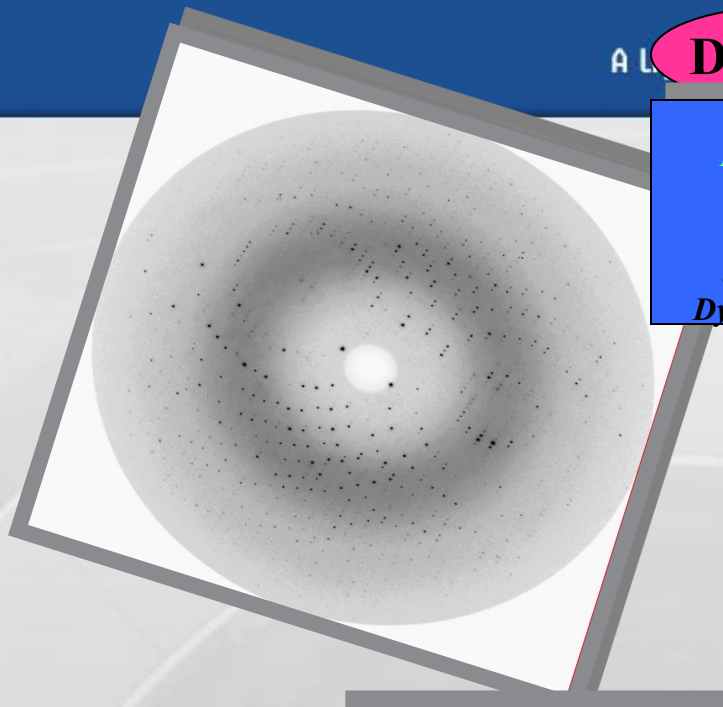
data collection strategies



Alexander Popov
ESRF, MX group

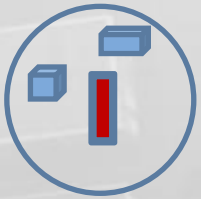
Data quality
Completeness, Resolution, Statistics

What the data is going to be used for?



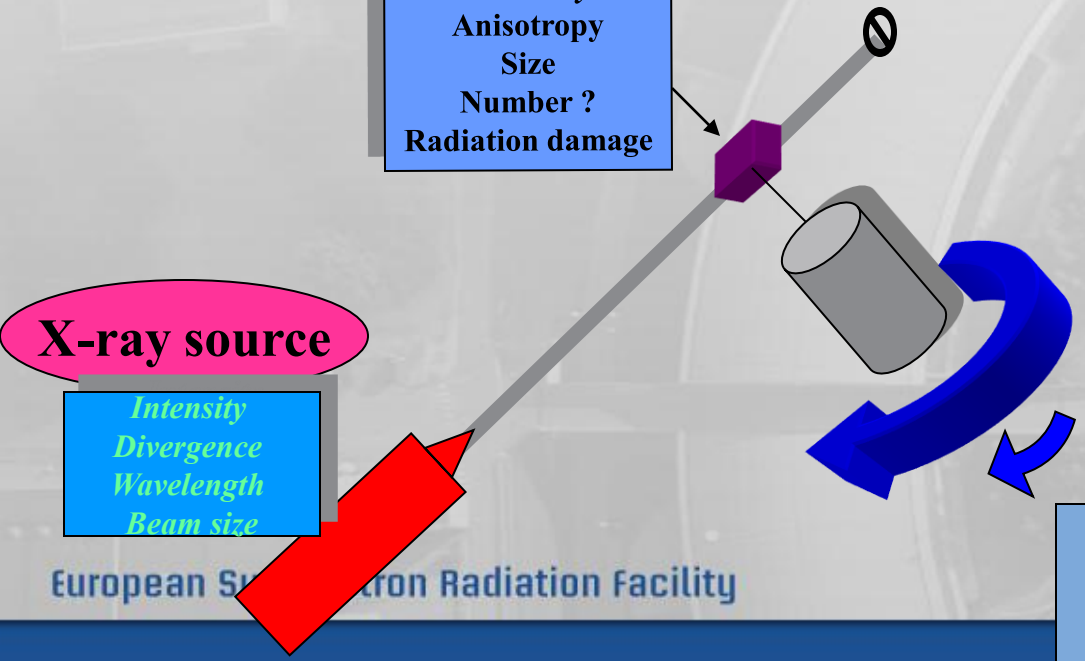
Detector

Size
N.of pixels
T readout
Noise
Sensitivity
Dynamic range



Crystal(s?)

Space group
Cell parameters
Orientation
B-factor
Mosaicity
Anisotropy
Size
Number ?
Radiation damage



X-ray source

Intensity
Divergence
Wavelength
Beam size

Crystal-to-detector distance
Resolution
I_p/I_b
spot overlap

Starting angle and total rotation range
Completeness
Multiplicity
d.c. time

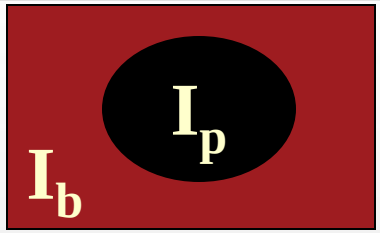
Rotation width
I_p/I_b,
spot overlap,
total d.c. time

Exposure time
statistics
overloads
total d.c. time

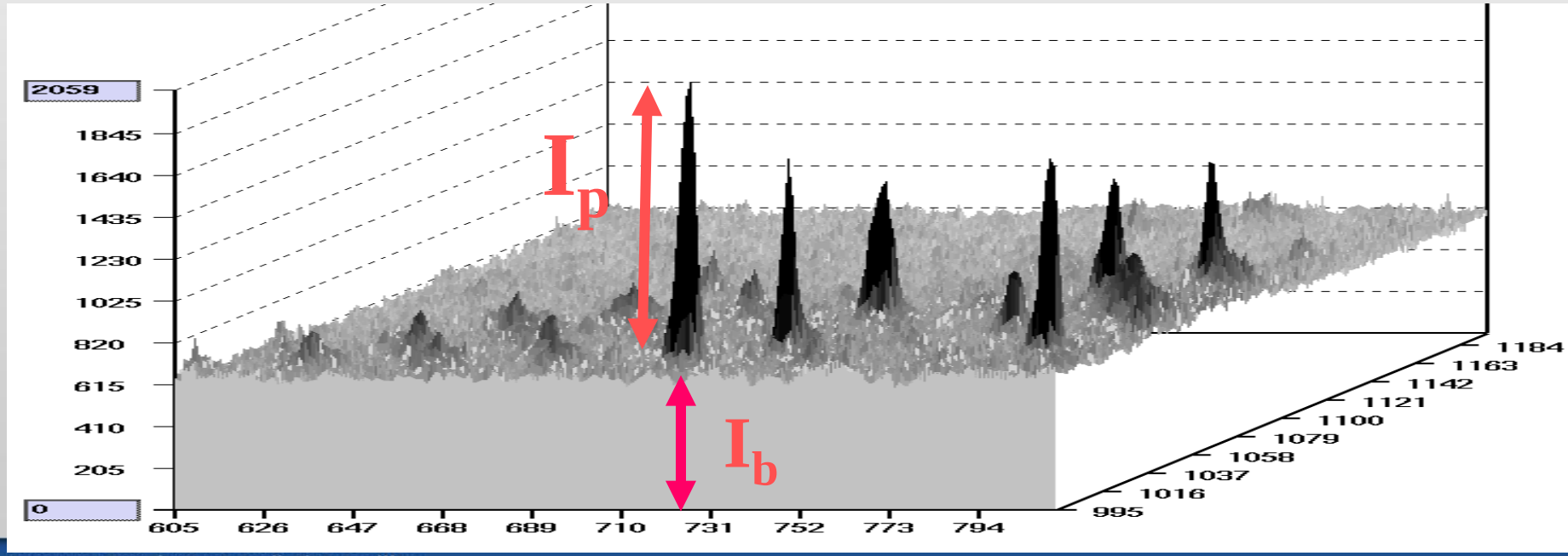
Accuracy

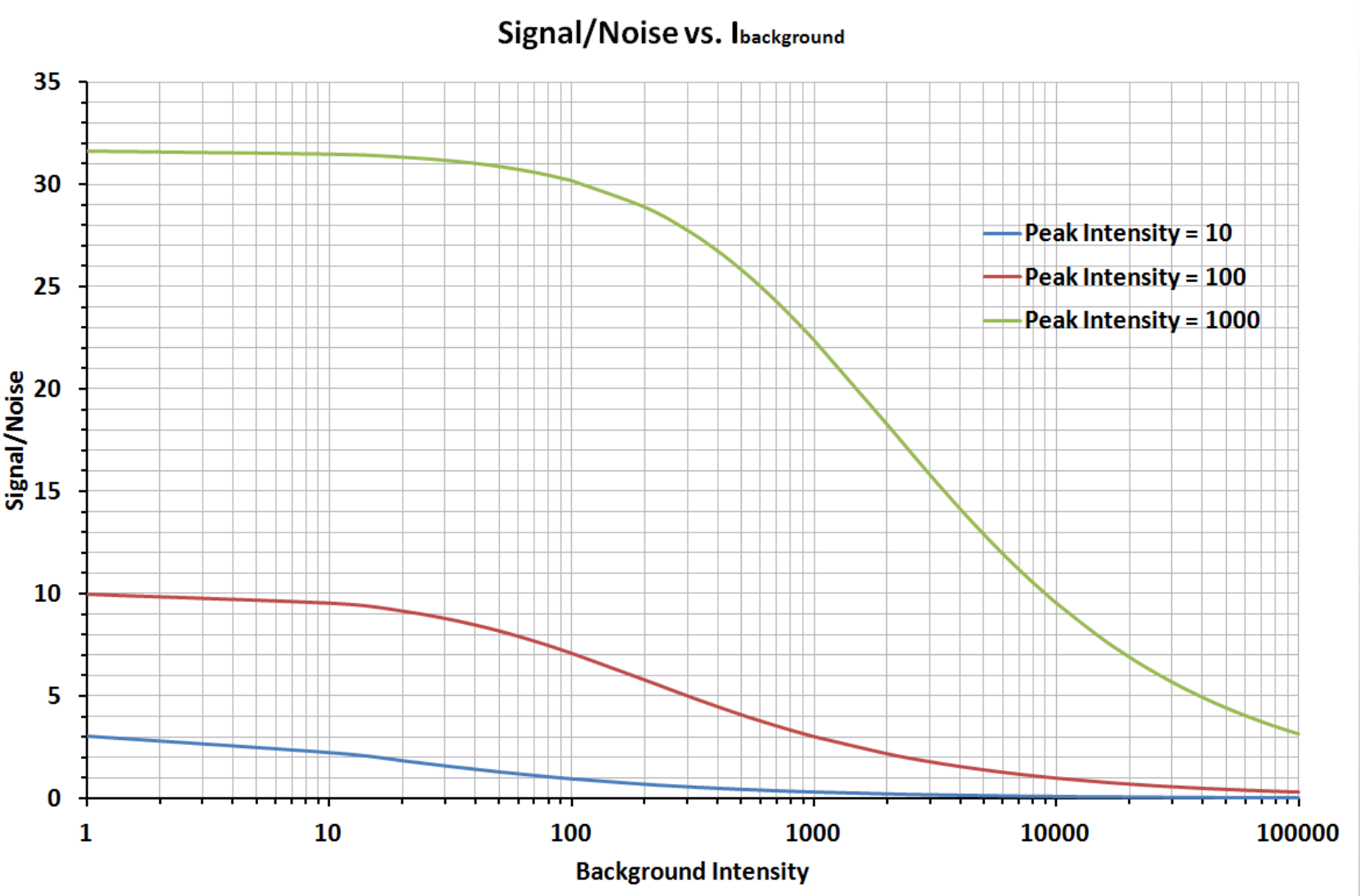
Main uncertainties of the observed intensities are determined by counting statistics

$$\sigma^2_{I_p} = \left(I_p + I_b \cdot \frac{m \cdot (m+n)}{n} \right) \cdot G$$



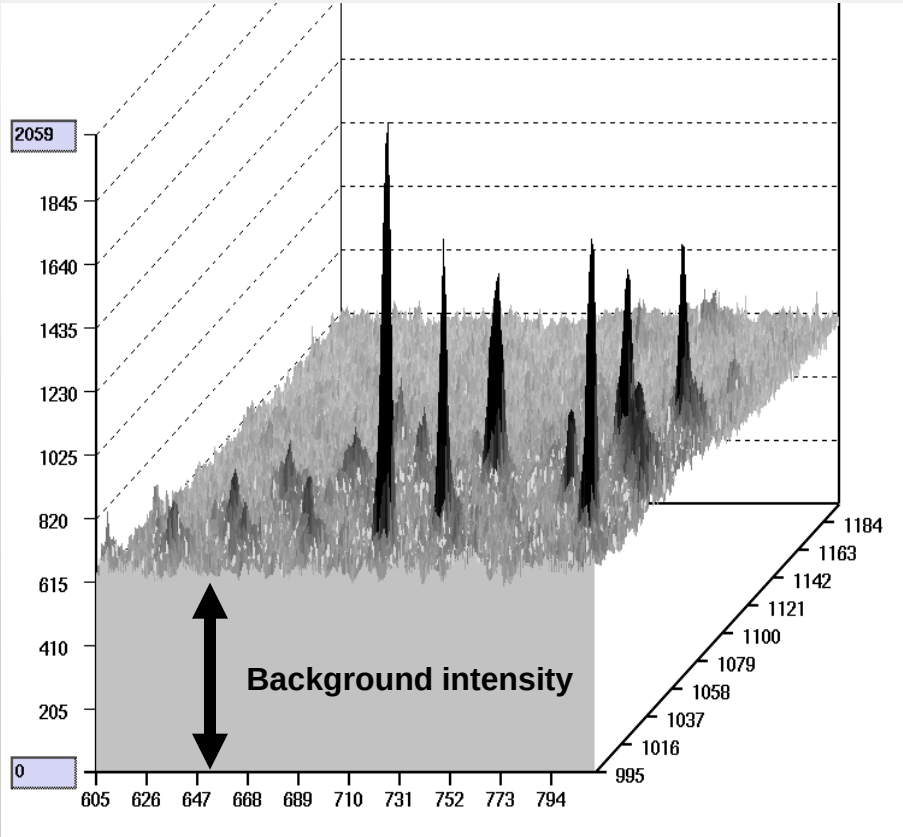
where ***m*** and ***n*** are number of pixels in the peak and background region of the measurement box respectively. ***G*** is the detector gain, which converts pixel counts to equivalent X-ray photons. K_{ins} is a proportionality constant for the instrument-error term

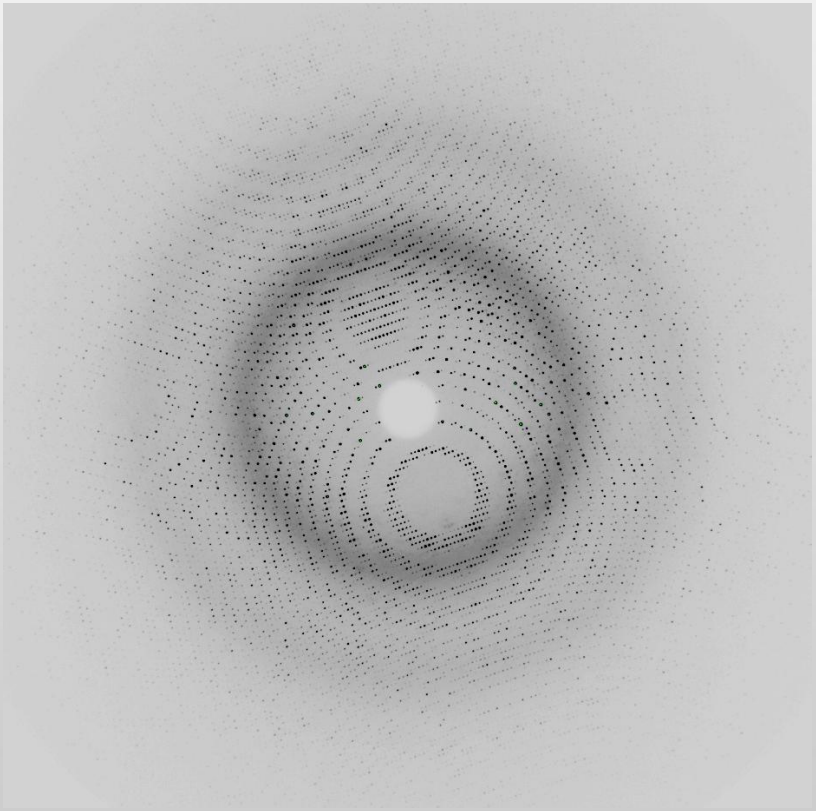



Signal/Noise vs. Ibackground
35
30
25
20
15
10
5
0
Signal/Noise
1
10
100
1000
10000
100000
Background Intensity
Peak Intensity = 10
Peak Intensity = 100
Peak Intensity = 1000

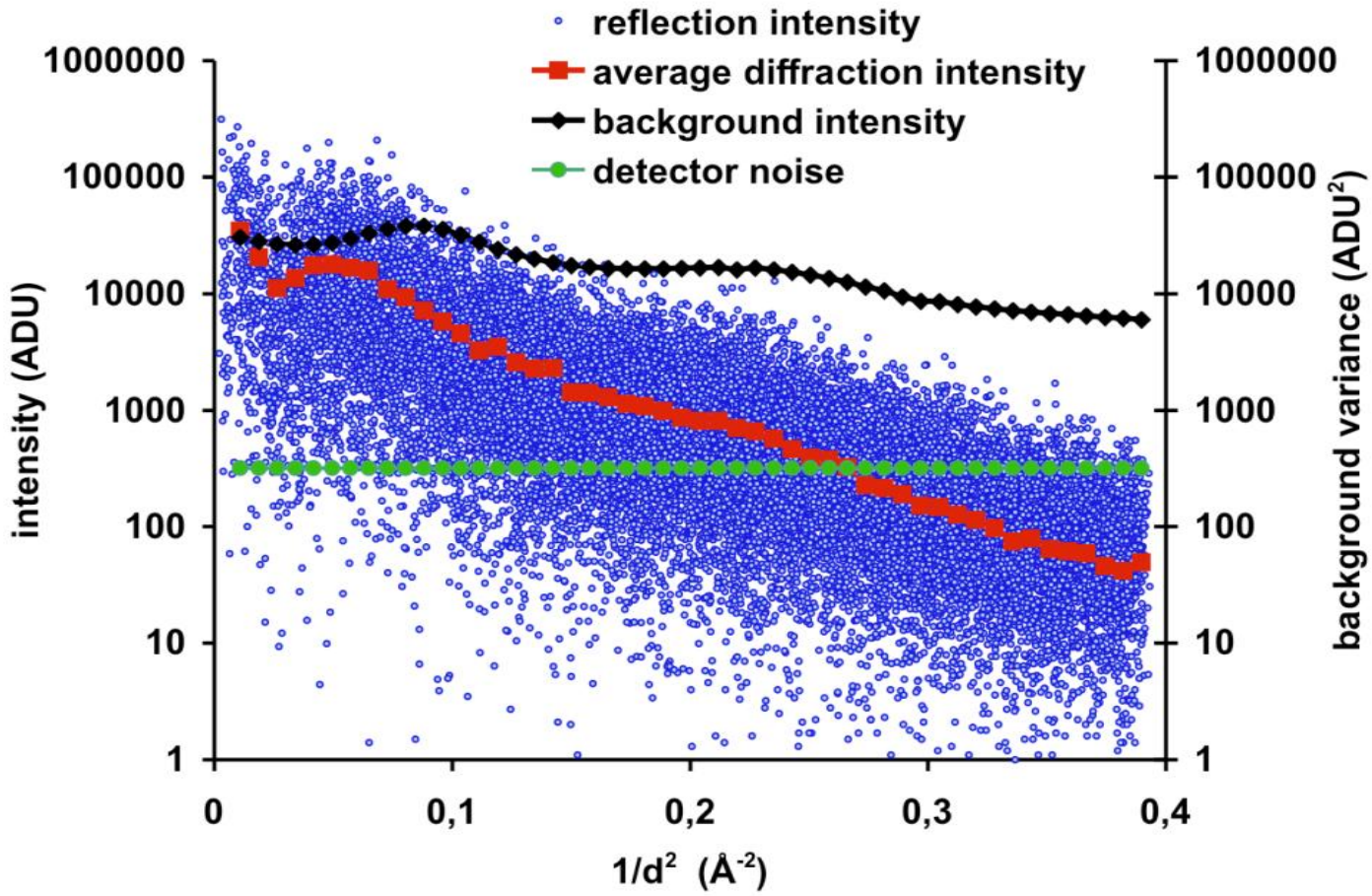
Complications

- Large cell parameters
- Weak diffraction intensity – light atoms
- Poor crystal quality – big B- factor
- Background intensity > diffraction intensity



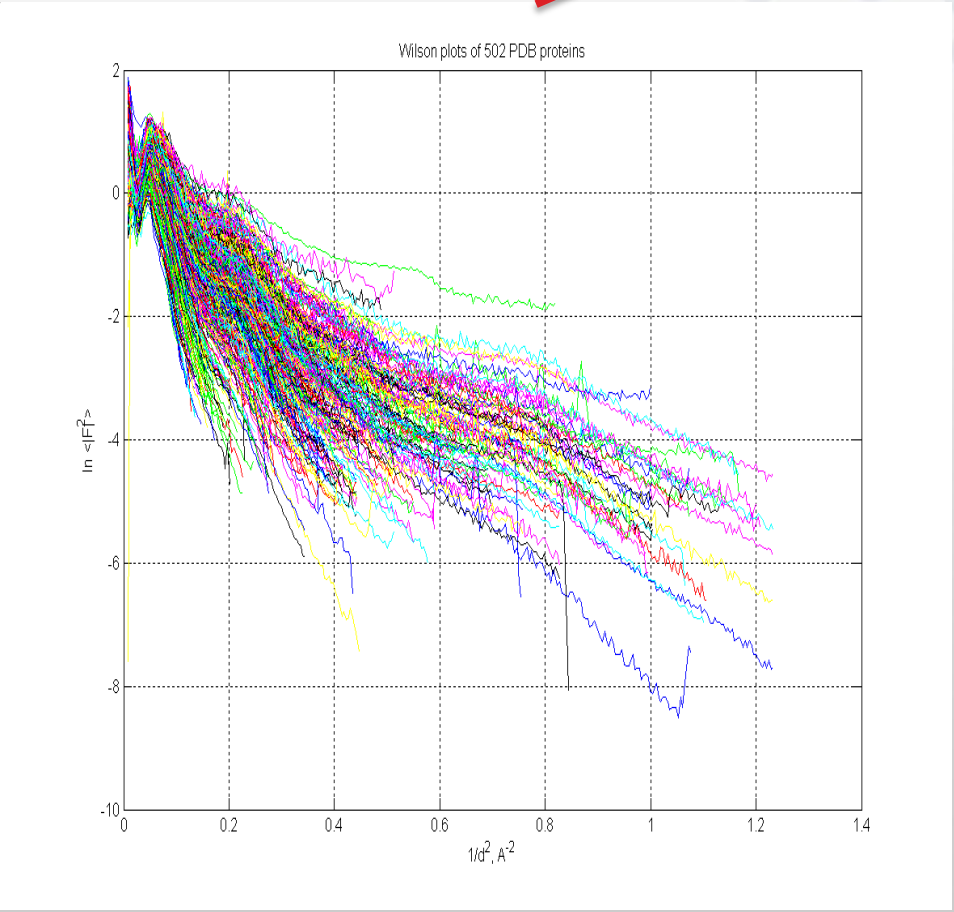


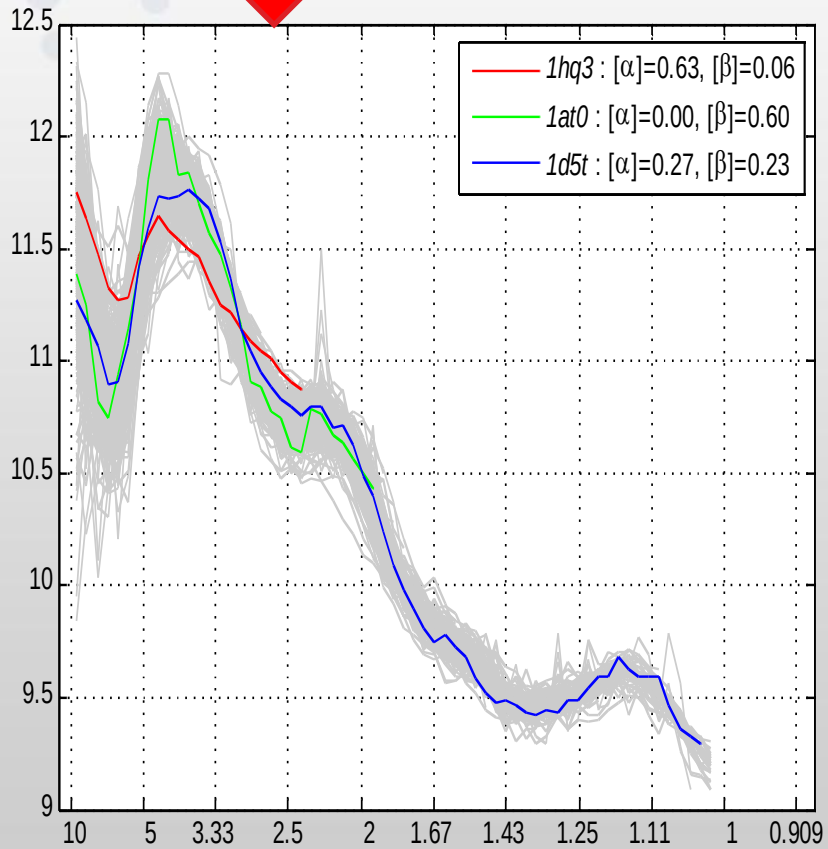
Statistics

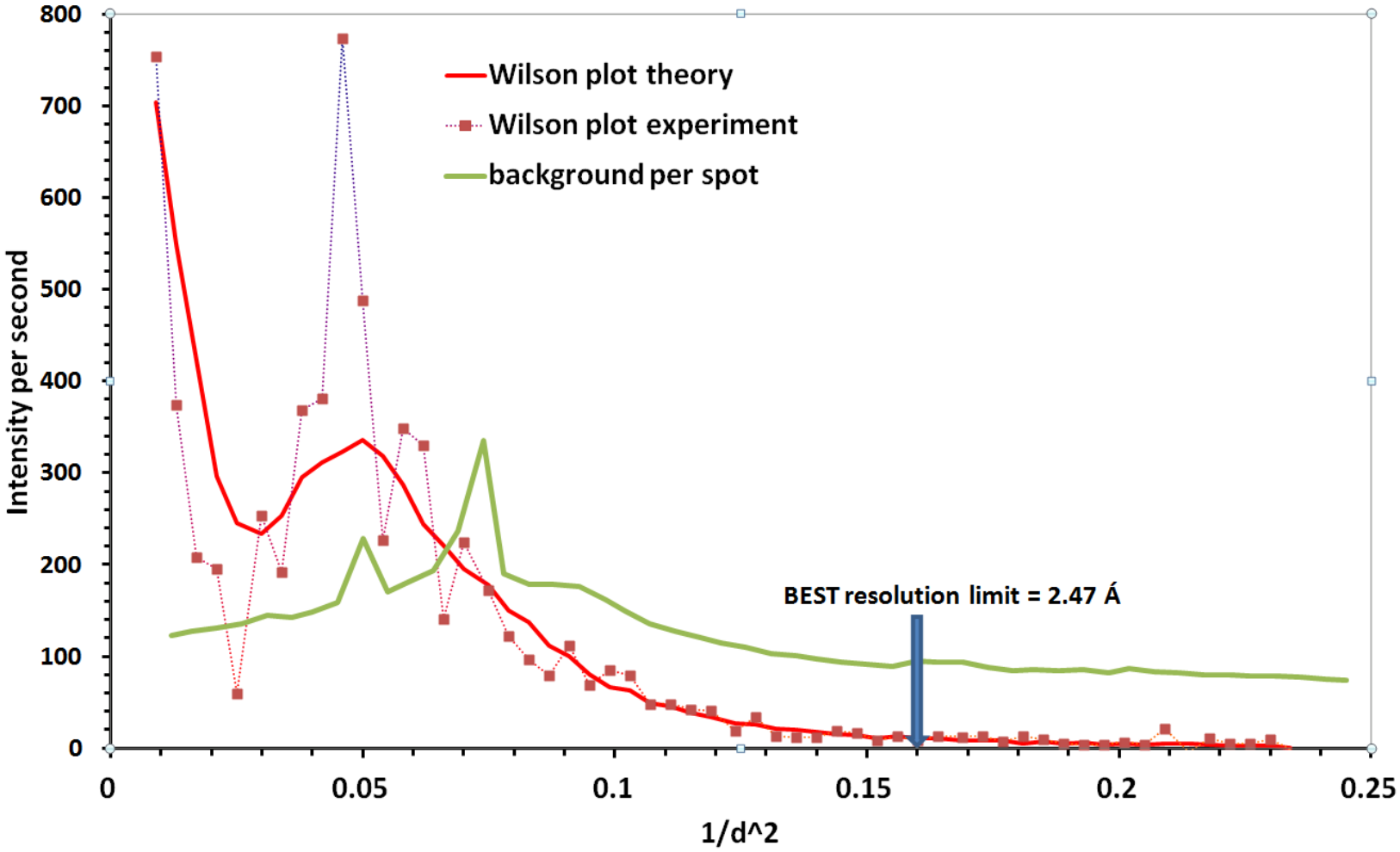

reflection intensity
average diffraction intensity
background intensity
detector noise
intensity (ADU)
background variance (ADU²)
1000000
100000
10000
1000
100
10
1
0
0,1
0,2
0,3
0,4
1/d² (Å⁻²)

Wilson plot

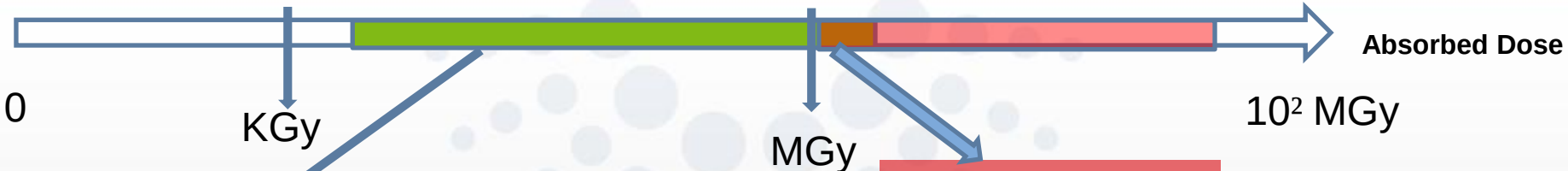
$$\hat{\mathrm{J}}(\mathbf{h}) \propto \cdot \hat{\mathrm{J}}_u(h) \cdot \mathrm{Exp}(-\mathbf{h} \cdot \mathbf{B} \cdot \mathbf{h}^T)$$





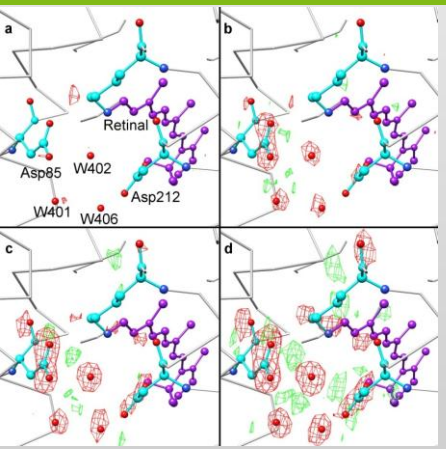

Wilson plot theory
Wilson plot experiment
background per spot
BEST resolution limit = 2.47 Á
Intensity per second
1/d^2
800
700
600
500
400
300
200
100
0
0
0.05
0.1
0.15
0.2
0.25

Radiation Damage



Site-specific damage

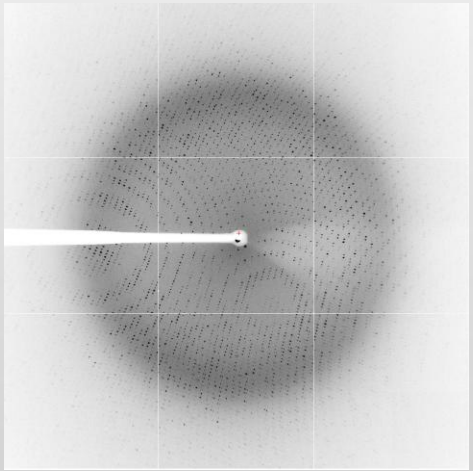
1. **Dose of 0.3 MGy - X-ray radiation damage effects are not detectable even at atomic resolution.**
2. **Doses above 2 MGy lead to partial decarboxylation of the most sensitive residues**
3. **Doses above 6 MGy may lead to wrong interpretation of chemistry for some protein residues**

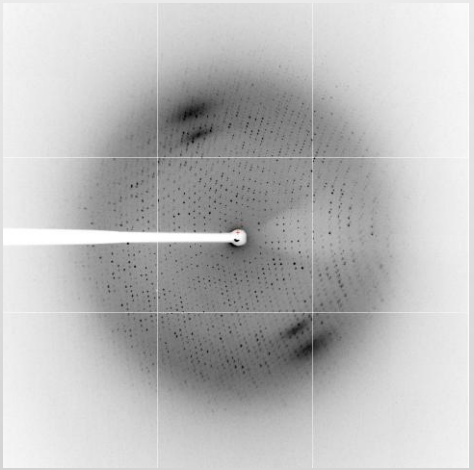


X-Ray-Radiation-Induced changes in Bacteriorhodopsin Structure
Borshchevskiy et al. *2011,*
***J.Mol.Biol.* V.409,813-825**

Global damage

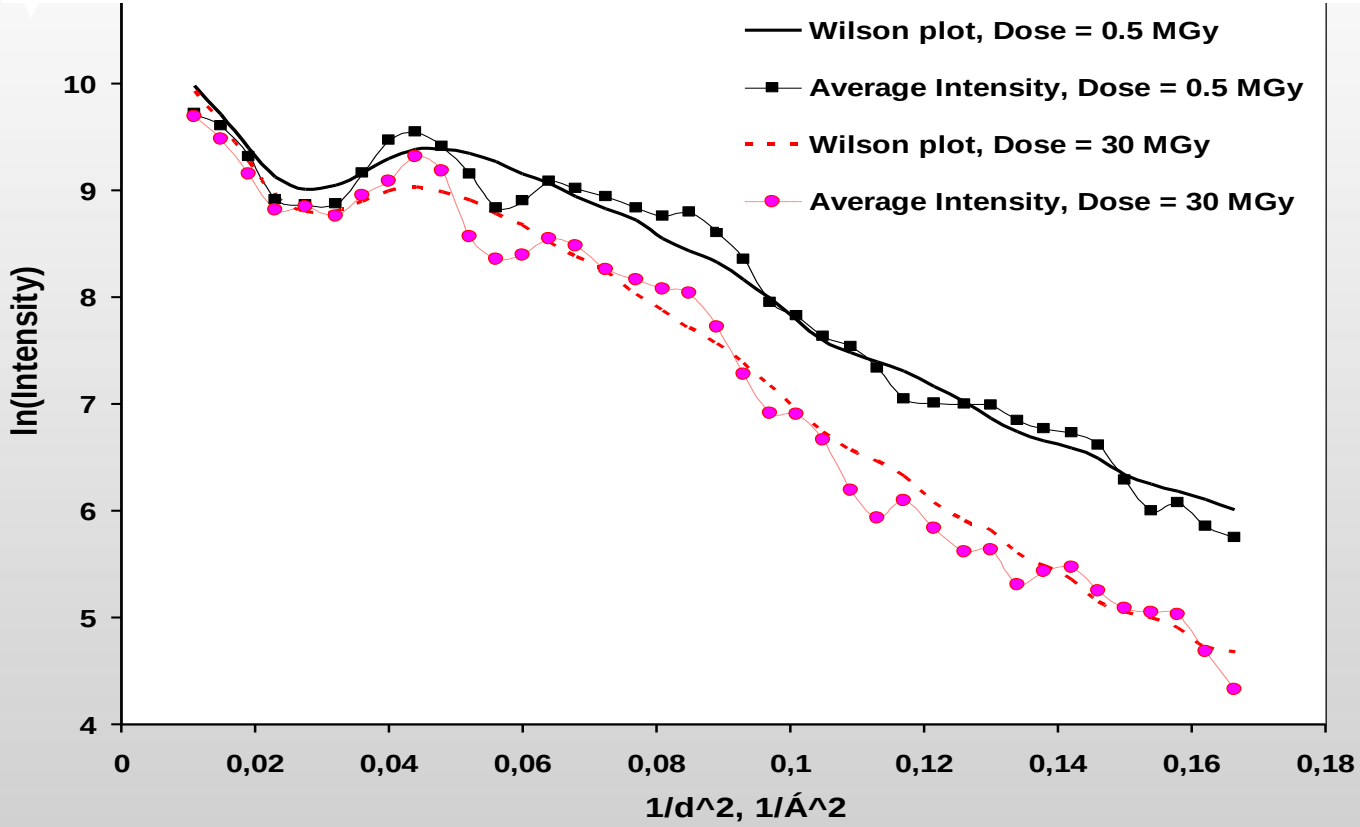
- ❑Overall and q-dependent loss of diffraction peak intensity
- ❑Non-specific non-isomorphism
- ❑Changes in unit-cell parameters
- ❑Increase in the mosaicity



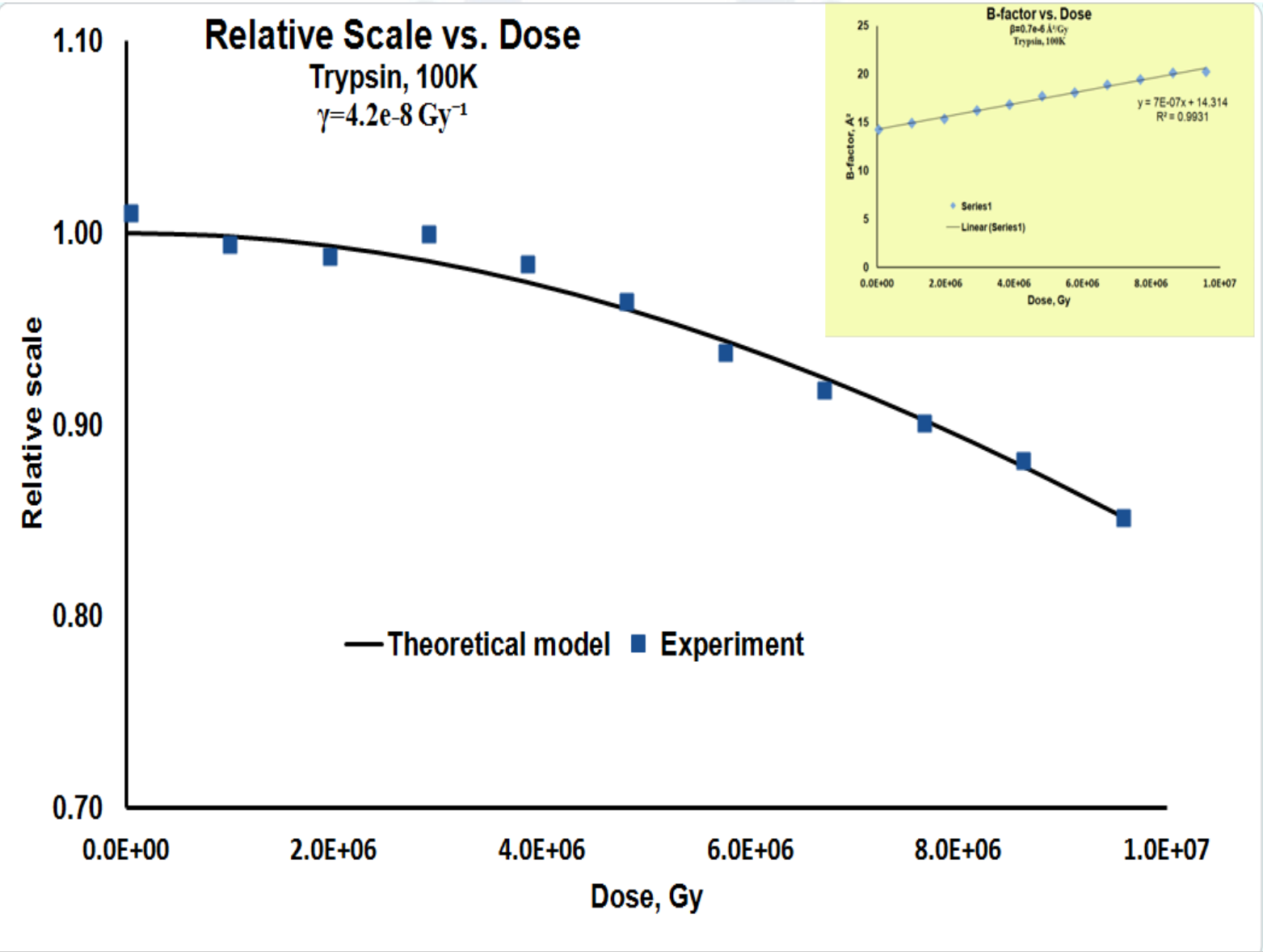


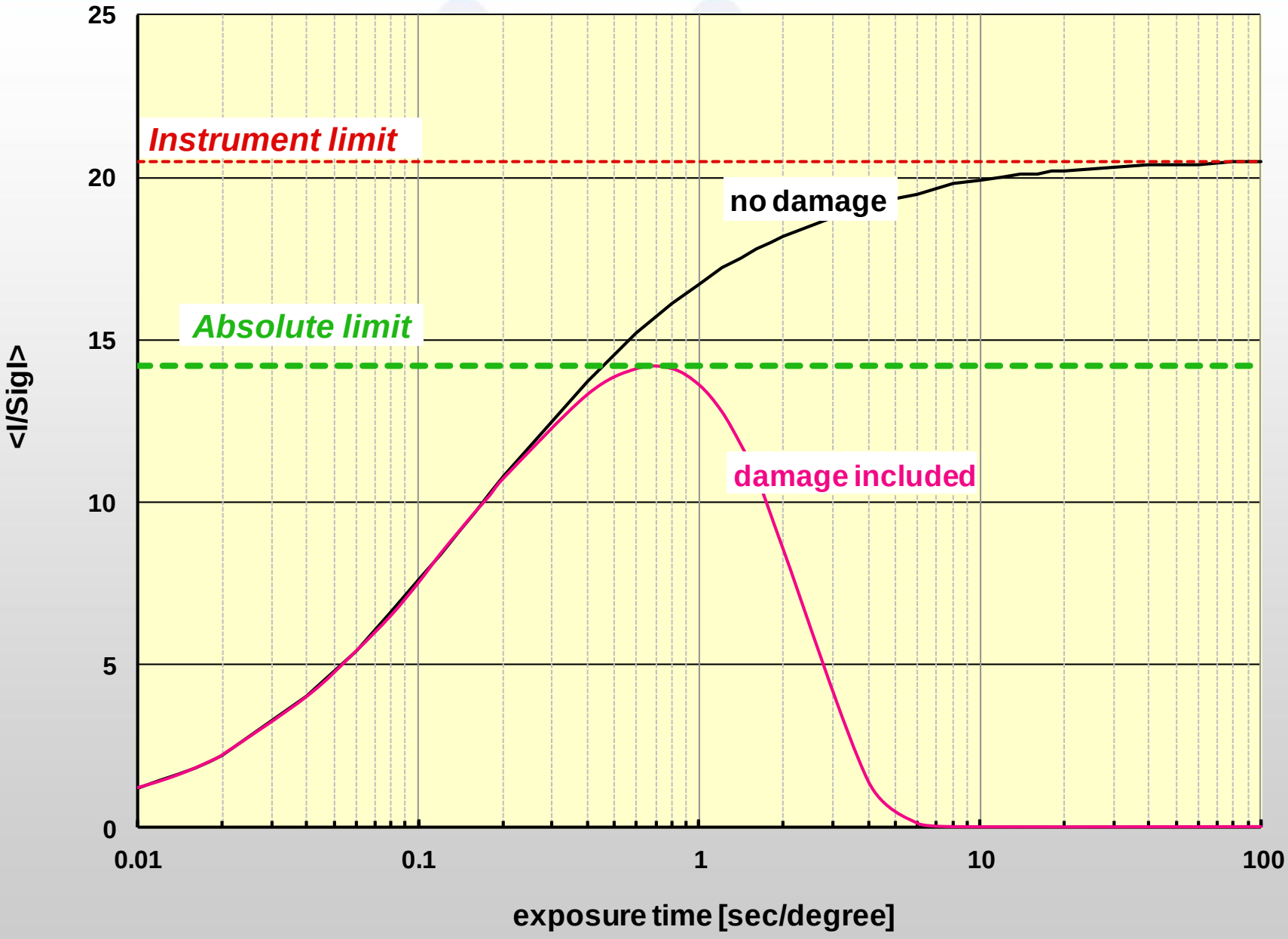
Intensity decay:

$$\hat{\mathrm{J}}(h, D) \propto \mathrm{scale}(D) \cdot \hat{\mathrm{J}}_u(h) \cdot \mathrm{Exp}(-B(D) \cdot h^2 / 2)$$

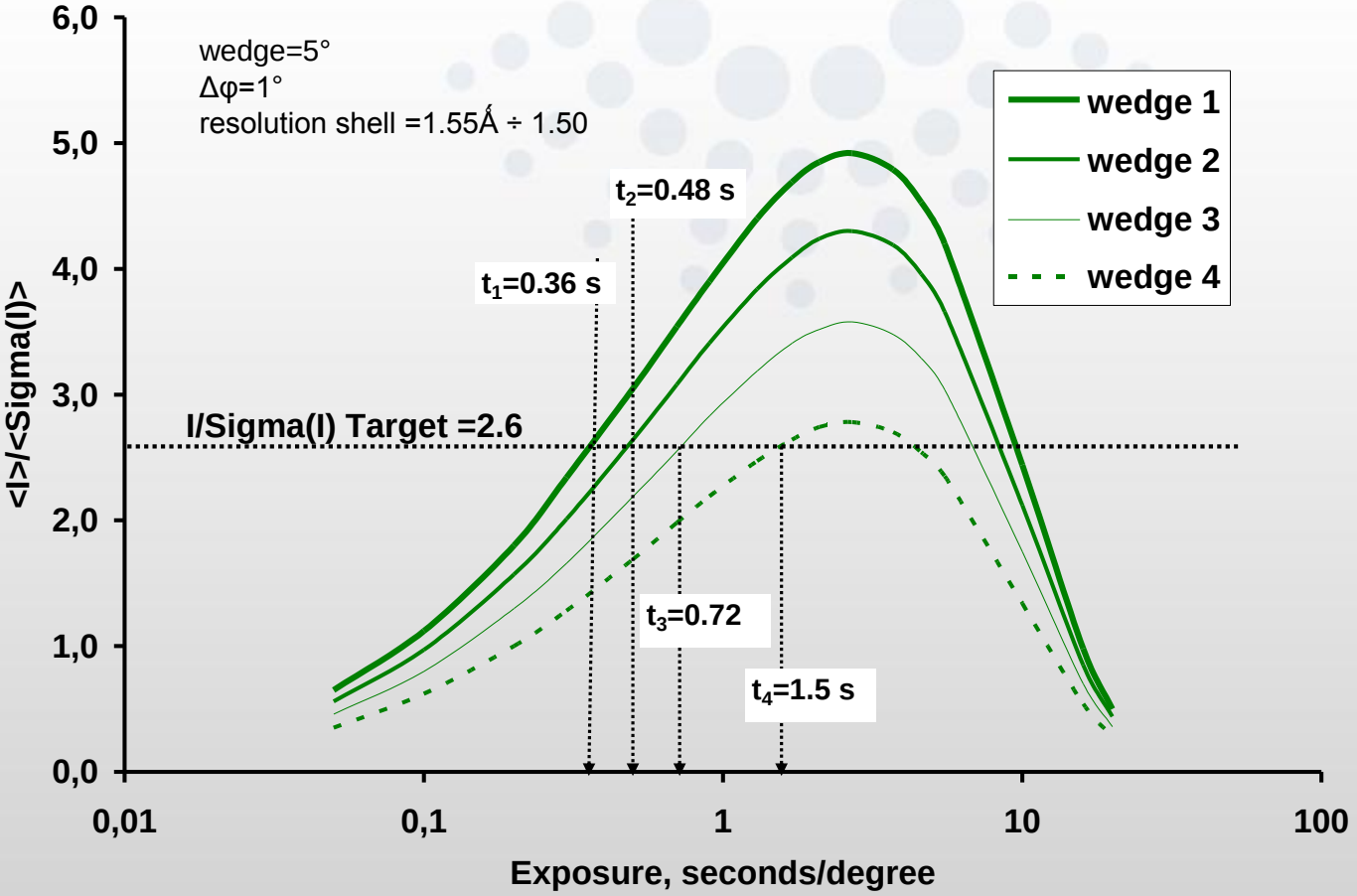


Global radiation damage



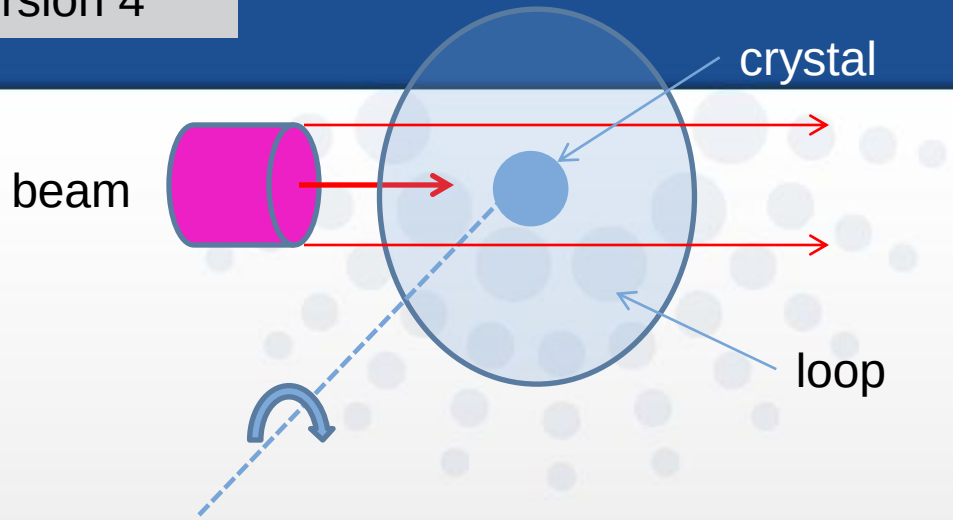

25
20
15
10
5
0
<I/SigI>
Instrument limit
Absolute limit
no damage
damage included
0.01
0.1
1
10
100
exposure time [sec/degree]

Insulin data collection



Graphical solution of the equation at resolution 1.5 Å and fixed rot. width for 5° sequential rotation wedges.

$$\frac{\hat{J}}{\hat{\sigma}_J}(t_{exposure}, D) = \frac{C}{\sqrt{M}} = \frac{4.0}{\sqrt{2.4}} = 2.58$$



$$\hat{J}(\mathbf{h}, D) = \hat{J}(\mathbf{h}, D = 0)\, scale(D) \exp(-\mathbf{h} \cdot \mathbf{B}(D) \cdot \mathbf{h}^T / 2)$$

Semi-empirical model of variance *vs* integrated intensity

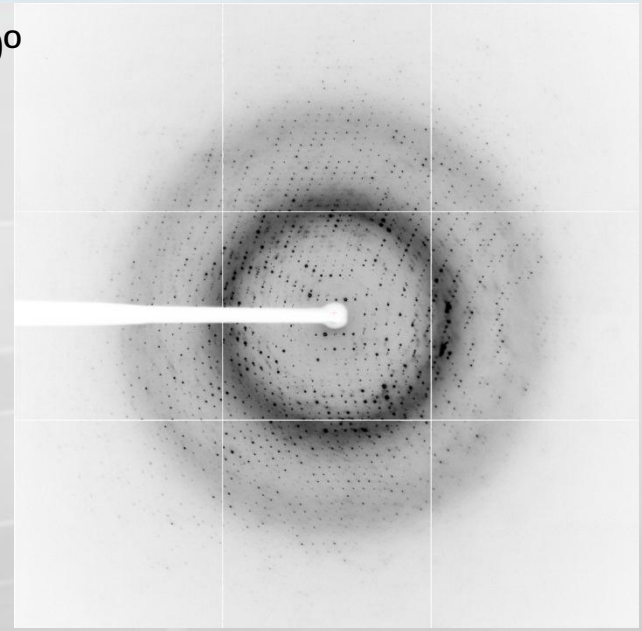
$$\sigma^2_I(J) = k_o + k_1 J + k_2 J^2$$

Integration over the scanned reciprocal space using Wilson distribution

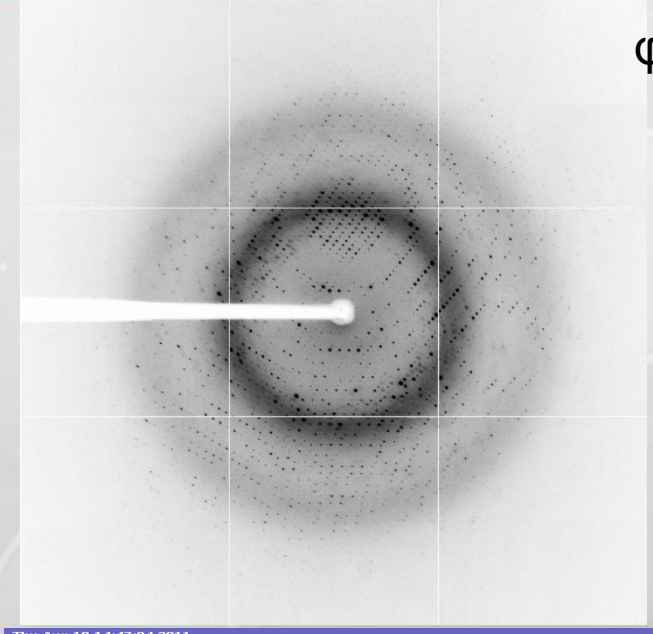
$$\hat{\sigma}_J(h,\varphi) = \frac{1}{2N} \sum_{i=1}^{N} \int_o^{\infty} \sqrt{k_{0i} + k_{1i} J + k_2 J^2} \left(p(J \mid \hat{J}(\mathbf{h}_{i1})) + p(J \mid \hat{J}(\mathbf{h}_{i2})) \right) dJ$$

Intensity Anisotropy

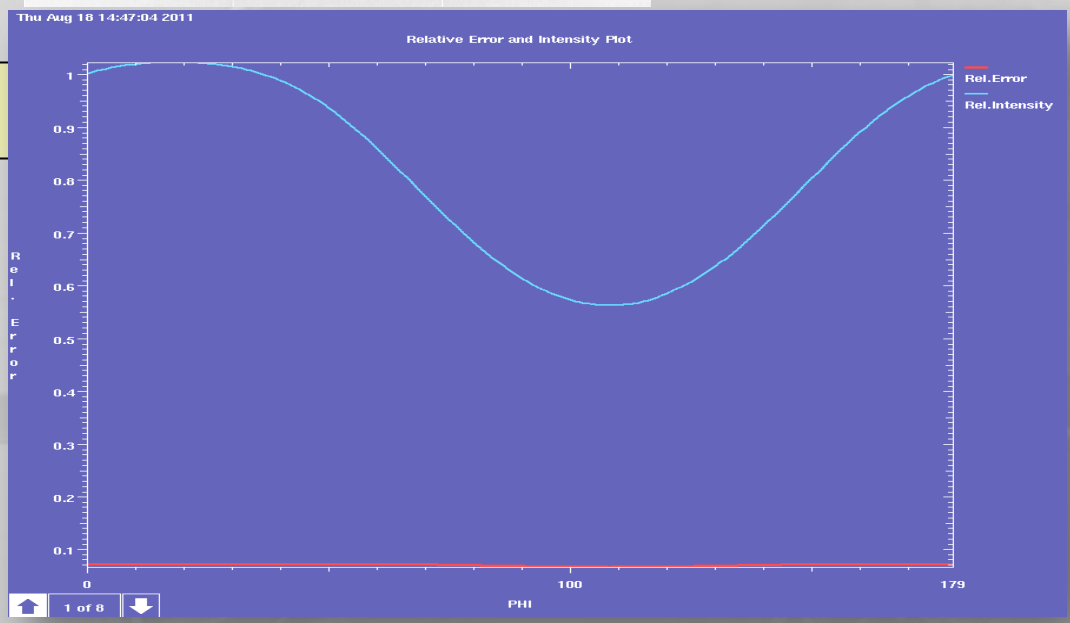
φ=0º



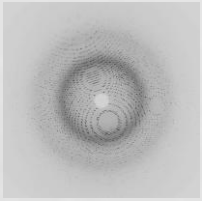
φ=90º



Intensity vs. crystal position



Data collection strategy accounting radiation damage



Initial Images

MOSFLM XDS

Beamline Flux
Crystal contents

RADDOSE

Absorbed dose rate

User choices

- Optimize data collection
- Optimize SAD data collection
- Find optimal crystal orientation
- Low-resolution optimal
- Rad. Damage sensitivity
- Multi-positional data collection
- Helical data collection
- Estimate data statistics

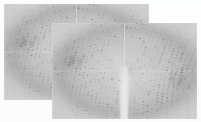
Dose (Time) limit
Geometry limits
Aimed statistics
Aimed completeness
Aimed redundancy
Aimed resolution

Detector parameters
Beamline parameters
and limitations

BEST 4.1

Optimal plan(s) of data collection
Statistics
B-factor

EDNA characterisation v1.3
A workflow written in Python



+ Xtal info
+ beam flux
+ diffraction plan

MOSLFM indexing

LABELIT DISTL

Indexing Evaluation

Ok

Failure

LABELIT indexing

Indexing Evaluation

Ok

Failure

MOSFLM Predictions

MOSFLM integration

XDS backg. estimation

[RADDOSE]

BEST

Data collection plan

File Instrumentation Help

User

Logout | opid-231 operator on ID23eh1 | (operator on ID23eh1 ID23eh1 Dates: 2013-09-05 to 2013-09-06)

Hutch | Collect | Energy scan | XRF spectrum | EDNA Log

Time and date	Prefix	Run number
2013-03-07 00:	propdiol_7	1
2013-03-28 19:	x2	2
2013 03 07 00:	mi_4	2
2013-03-06 23:	black_x9	2
2013-04-01 17:	xtal16	3
2013-04-01 16:	xtal7	2
2013-03-28 21:	x1	1
2013-04-01 14:	xtal2	3
2013-03-07 00:	propdiol_9	1
2013 03 28 21:	x1	2
2013-03-28 18:	x5	1
2013-04-01 15:	xtal2	5
2013-03-06 23:	mi_2	1
2013-04-01 15:	xtal5	5
2013-03-07 00:	mi_5	1
2013-03-28 18:	x6	1
2013 04 01 17:	xtal13	1
2013-04-01 15:	xtal4	1
2013-04-01 18:	xtal2	1
2013-04-01 17:	xtal14	5
2013-04-01 17:	xtal14	1
2013-04-01 17:	xtal16	1
2013-03-28 21:	x5	1
2013 03 06 23:	black_x9	3
2013-04-01 17:	xtal14	6
2013-03-28 20:	x4	1
2013-04-01 17:	xtal15	1
2013-03-06 23:	cf042a_x2	2
2013-03-06 21:	black_x8	1
2013-04-01 16:	xtal8	2
2013-04-01 19:	xtal10	2
2013-03-28 18:	x4	1
2013-03-06 21:	black_x9	1
2013-04-01 17:	xtal13	2
2013-03-06 21:	blue_x4	1
2013-03-28 15:	ins	1
2013-03-06 22:	cd468a_x5	1
2013-03-06 22:	cd468a_x2	1
2013-04-01 14:	xtal2	2
2013-04-01 18:	xtal7	1
2013-03-06 21:	blue_x3	1

Clear history

Back | Forward

Data collection info

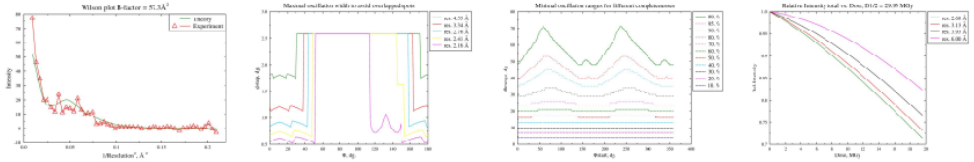
Data collection date	2013/Mar/06 21:26:18.859
Image prefix	ref-black_x9_1
Directory	/data/id23eh1/inhouse/opid231/20130306/RAW_DATA/daniele/j/black/x9

Diffraction Plan

Forced space group	Anomalous data	Aimed multiplicity	Aimed completeness	Aimed I/sigma at highest res.	Aimed resolution (Å)
None	False	Default (optimized)	Default (>= 0.99)	3.00	Default (highest possible)

Collection plan strategy (RADDOSE log file , BEST log file)

Resolution limit is set by the radiation damage								
Wedge	Subwedge	Start (°)	Width (°)	No images	Exp time (s)	Max res (Å)	Rel trans (%)	Distance (mm)
1	1	9.00	0.15	480	0.32	2.68	100.00	564.25



Indexing summary: Selected spacegroup: P3

Refined unit cell parameters (Å/degrees)					
a (Å)	b (Å)	c (Å)	alpha (°)	beta (°)	gamma (°)
133.844	133.844	38.022	90.000	90.000	120.000

Indexing log file

ref-black_x9_1_0002.cbf

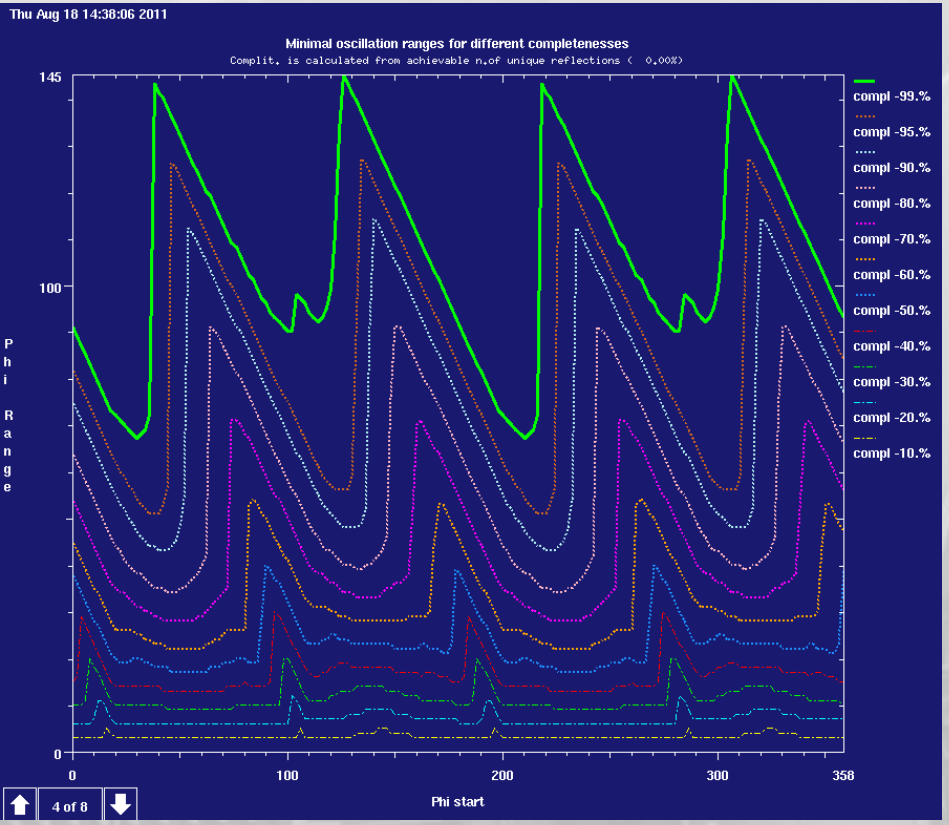
ref-black_x9_1_0001.cbf

Integration log file 1
Integration log file 2

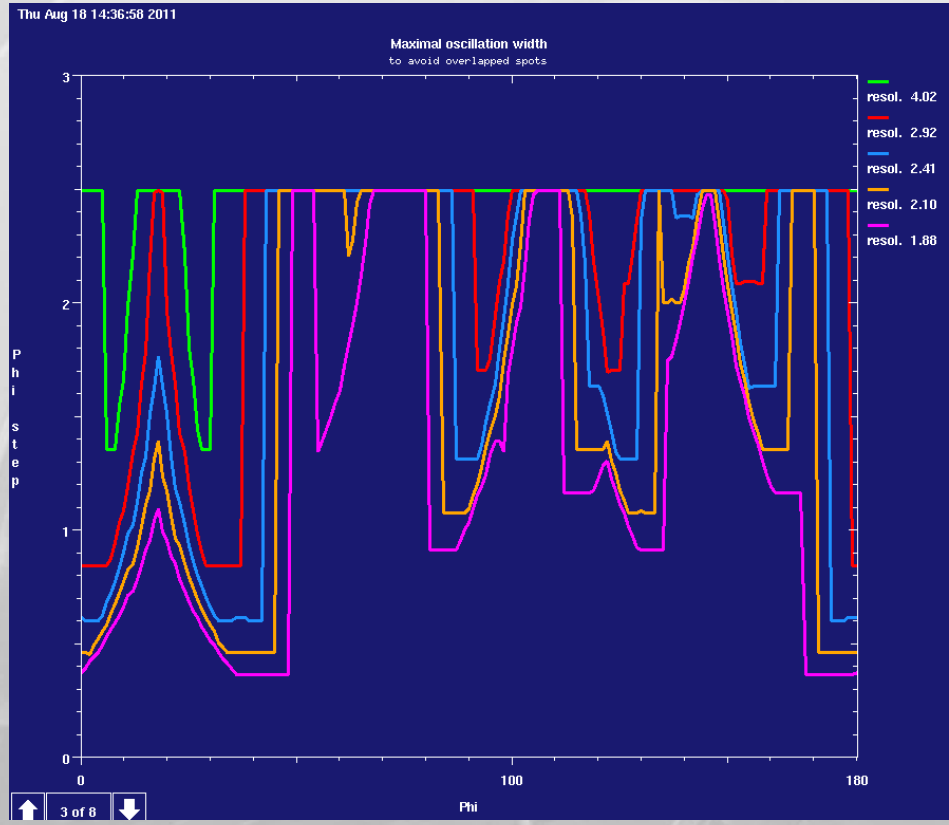
GEOMETRY

Space group, Cell parameters, Orientation, Mosaicity, Spot Size

Optimal starting spindle angle and scan range

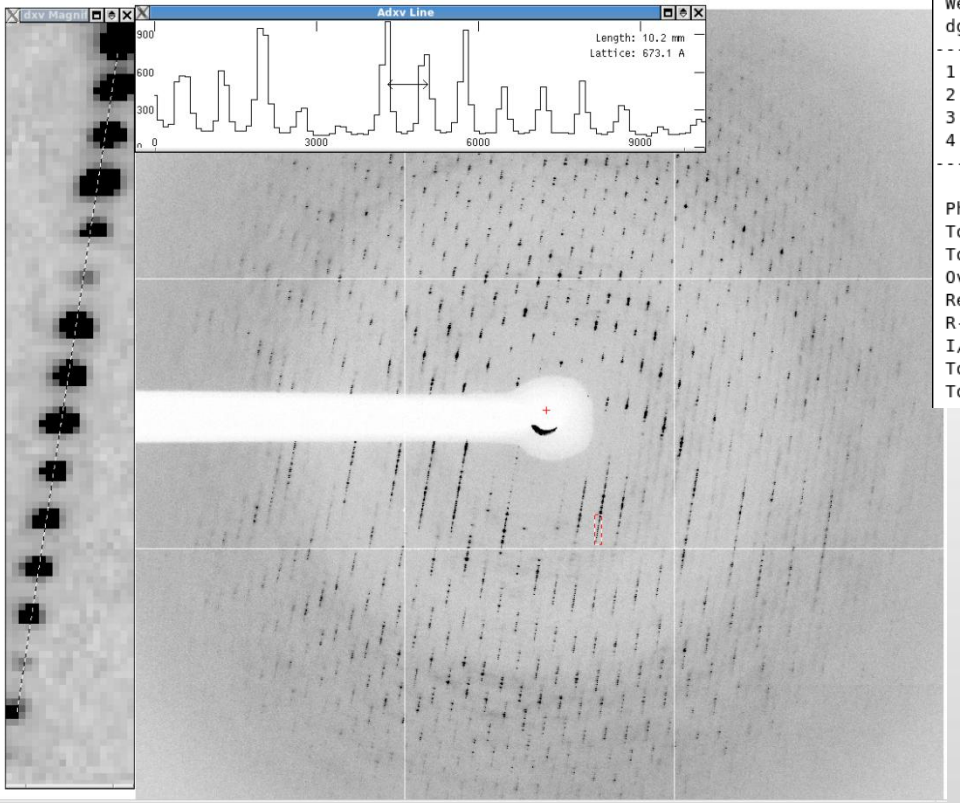


Maximum rotation angle without spot overlap




```
Crystal
 Space Group                    : P 2 2 2
 Cell                           :  53.52  75.64 123.21  90.00  90.00  90.00
 Mosaicity                      : 0.51 degree
```

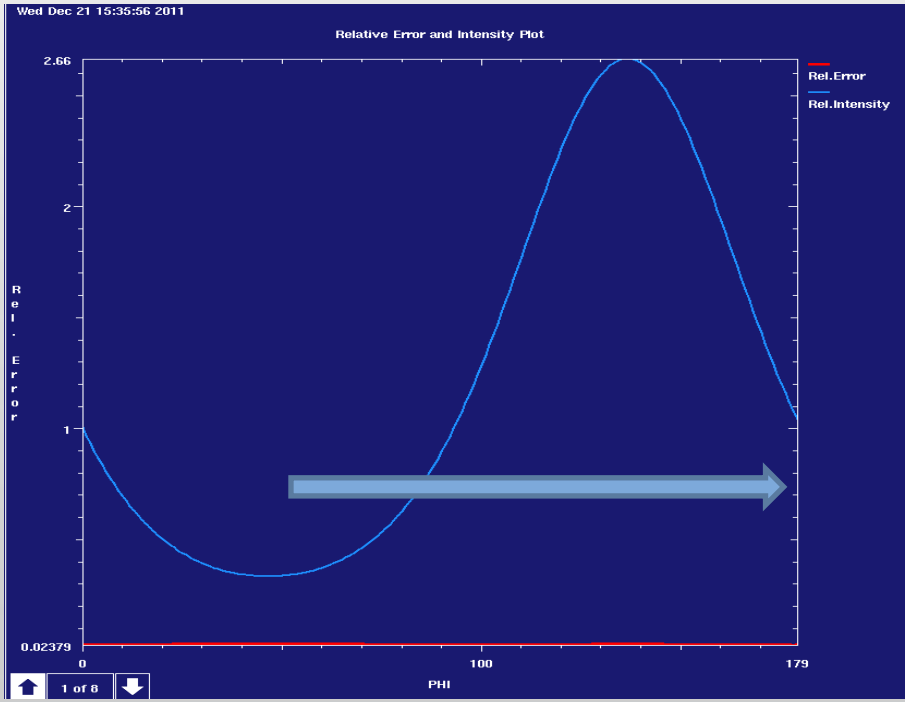
```
Crystal
 Space Group                : P 4
 Cell                       : 141.55 141.55 671.01  90.00  90.00  90.00
 Mosaicity                  : 0.17 degree
```




```
                                  Main Wedge
                                  ==========
      Resolution limit is set by the radiation damage
 Resolution limit =3.59 Angstrom   Transmission =   40.6%  Distance = 562.2mm
```

sub-Wedge	Phi start degree	Rot. width degree	Exposure /image s	N.of images	Overlap	sWedge width degree	Exposure /sWedge s	Exposure total s	Dose /sWedge MGy	Dose total MGy	Completeness %
1	53.00	0.10	0.169	100	No	10.00	16.9	16.9	8.711	8.711	32.9
2	63.00	0.10	0.156	300	Yes	30.00	46.9	63.8	24.156	32.867	76.0
3	93.00	0.10	0.100	150	No	15.00	15.0	78.8	7.729	40.596	90.2
4	108.00	0.50	0.437	140	No	70.00	61.2	140.0	31.530	72.126	100.0

```
Phi_start - Phi_finish      : 53.00 - 178.00
Total rotation range        : 125.00 degree
Total N.of images           : 690
Overall Completeness        : 100.0%
Redundancy                  : 5.46
R-factor (outer shell)      : 13.3% (   63.2%)
I/Sigma (outer shell)       : 17.3 (   3.4)
Total Exposure time         : 140.0 sec (0.039 hour)
Total Data Collection time  : 1948.2 sec (0.541 hour)
```



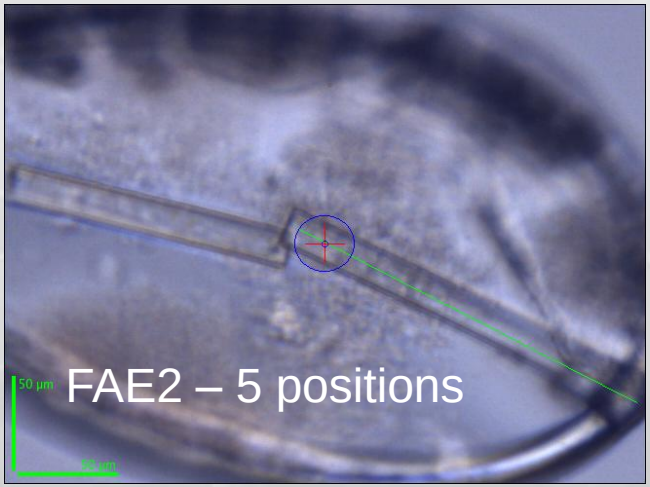

```
Scaling
 Relative scale         :   31.44
 Overall B-factor       :  82.17 Angstrom^2
 B-factor eigenvalues   :  51.11 118.04 118.04 Angstrom^2
 Scaling error          :   3% at the resolution limit
```

Multi-positional data collection

A Light for Science

FAE crystals
ID23-1

E=12.75Kev, I=35 mA, Aperture=0.03 mm
Flux=1.5x10^{11} Photon/sec



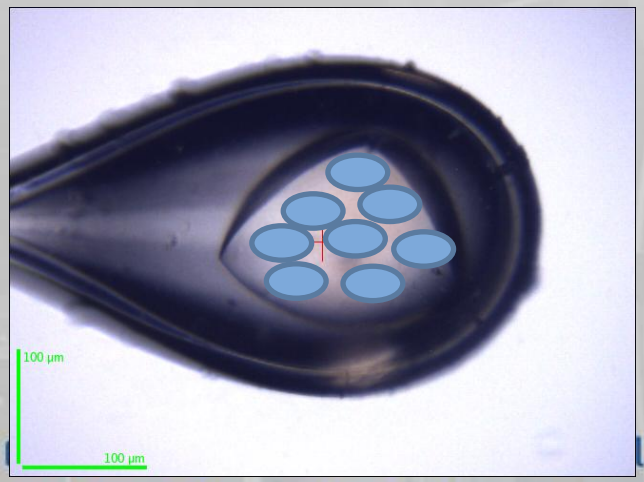
Multi-positions data collection
==========

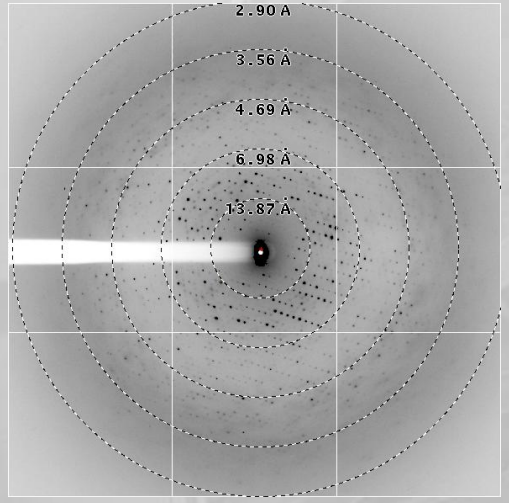
Resolution limit is set by the radiation damage
Resolution limit =1.73 Angstrom Transmission = 100.0% Distance = 244.6mm

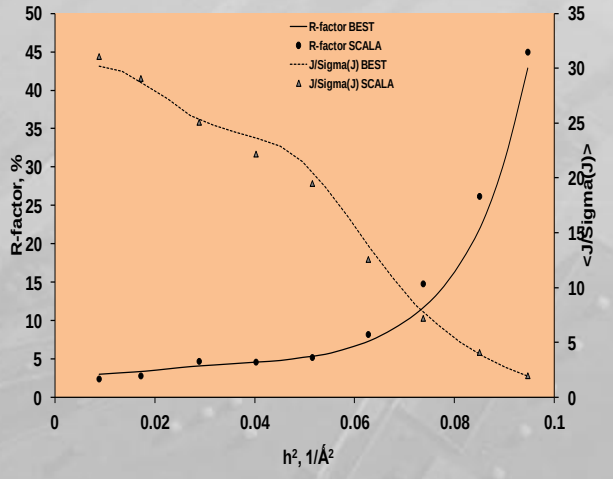
WEDGE PARAMETERS				INFORMATION						
sub-We-dge	Phi start degree	Rot. width degree	Exposure /image s	N.of ima-ges	Over-lap	sWedge width degree	Exposure /sWedge s	Exposure total s	Dose /sWedge MGy	Dose total MGy	Comple-teness %
	Wedge number = 1	Crystal position = 1									
1	0.00	0.25	1.338	80	No	20.00	107.0	107.0	4.067	4.067	51.9
	Wedge number = 2	Crystal position = 2									
1	20.00	0.25	1.338	80	No	20.00	107.0	107.0	4.067	4.067	85.6
	Wedge number = 3	Crystal position = 3									
1	40.00	0.25	1.338	80	No	20.00	107.0	107.0	4.067	4.067	75.2
	Wedge number = 4	Crystal position = 4									
1	60.00	0.25	1.338	80	No	20.00	107.0	107.0	4.067	4.067	88.1

Phi_start - Phi_finish : 0.00 - 80.00
Total rotation range : 80.00 degree
Total N.of images : 320
Overall Completeness : 98.6%
Redundancy : 3.18
R-factor (outer shell) : 5.6% (36.8%)
I/Sigma (outer shell) : 22.9 (3.3)
Total Exposure time : 428.1 sec (0.119 hour)
Total Data Collection time : 1228.1 sec (0.341 hour)

The 70 kDa membrane protein FtsH from Aquifex aeolicus I222, a = 137.9, b = 162.1, c = 170







SAD optimization

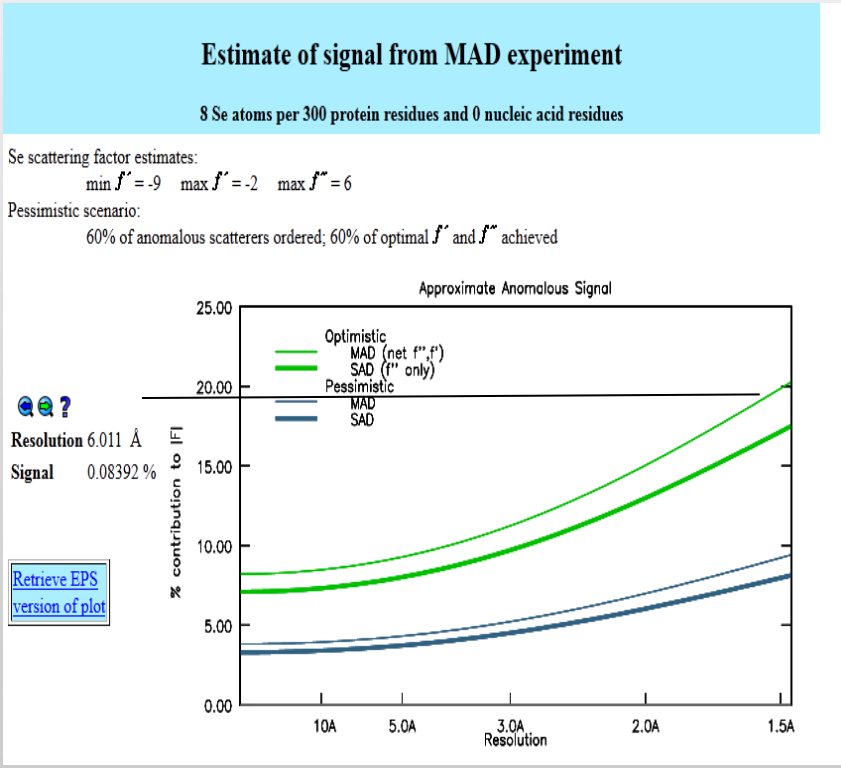
..... SAD data collection.....
-asad, strategy for SAD data collection, resolution selected automatically,rot.interval=360 dg.

-SAD {no|yes|graph}, strategy for SAD data collection if "yes", "graph" - estimation of resolution for SAD

Minimum of RFriedel = <|<E2+/w>-<E2-/w>|> is a target noise only, no anomalous scattering itself: decay, non-isomorphism exact pair-vice dose differences for Bijvoet mates

Resolution	RFriedel(%)	I/Sigma	Redundancy
10.12	0.8	74.1	23.7
6.90	0.8	43.6	23.7
5.34	1.1	48.4	23.0
4.51	1.2	47.5	23.5
3.98	1.6	34.5	20.6
3.60	2.5	22.4	13.9
3.31	4.0	14.0	11.9
3.08	6.6	8.3	7.0
2.89	10.5	5.2	6.1
2.73	15.6	3.7	2.5
2.60	23.0	2.4	3.8

http://skuld.bmsc.washington.edu/cgi-bin/MAD_power.pl



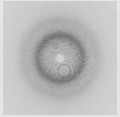
Induced Burn Strategy

Beamline Flux
Crystal contents
Crystal sizes

RADDOSE

Absorbed dose rate

Initial Images



MOSFLM XDS

User

Rad. Damage sensitivity

BEST

Plan of data collection

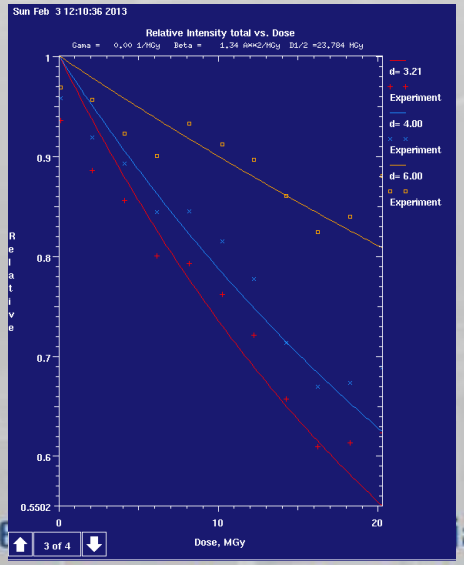
11 cycles for testing
10 cycles for burning

Measurements

XDS auto

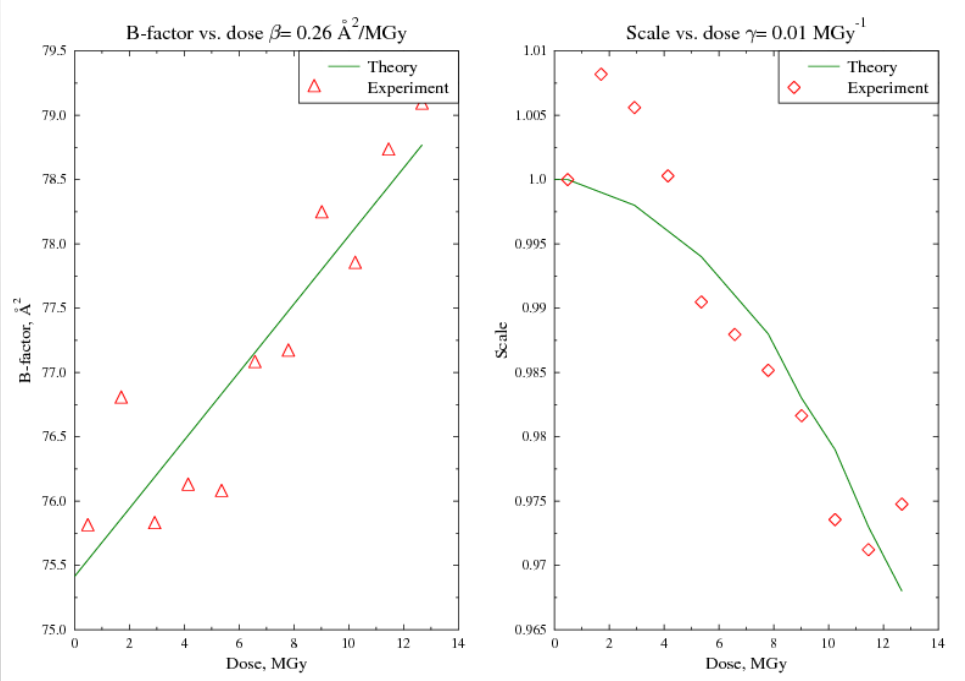
RDFIT

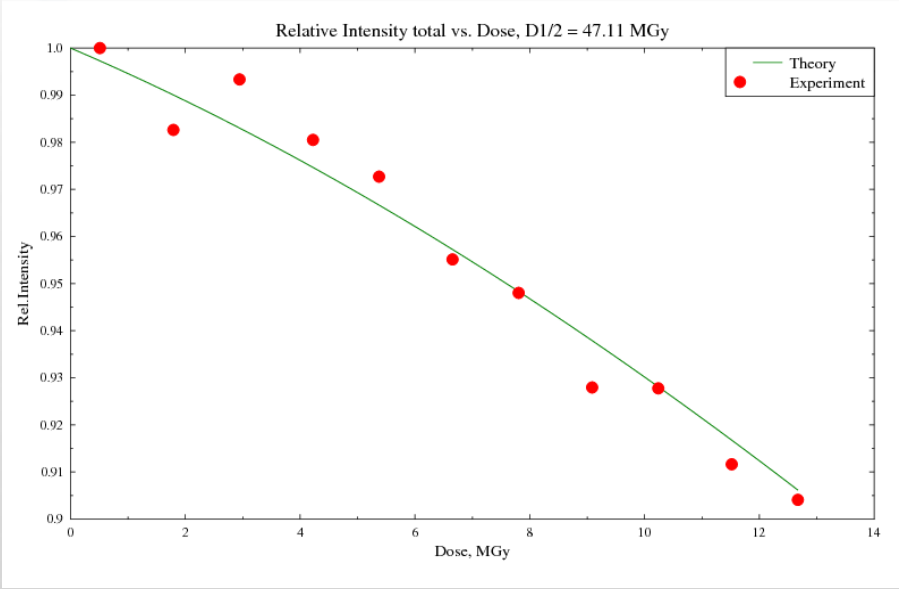
- Minimal RD inside the testing cycles
- Must induce significant changes in Intensity
- The intensity measurements remain statistically significant up to the last cycle of data collection



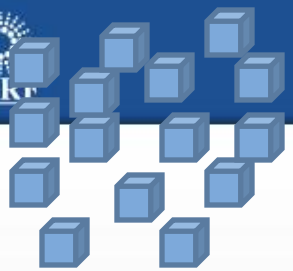

```
 Program RDFIT /A.Popov & G.Bourenkov/
 Version 1.1.0//  16.11.2012
 Copyright 2012 by Alexander Popov and Gleb Bourenkov

  PROJECT:     BEST plan

-------------------------------------------------------
Wedge|  dose   |max Resol.| B-factor| Scale|del_Volum
     |  MGy    |  Angstr. |Angstr**2|      | cell, %
-------------------------------------------------------
  0    0.0000      2.07     24.69   1669.62
  1    0.0081      2.07     24.24   1536.05   0.000
  2    0.0246      2.07     25.48   1625.30  -0.061
  3    0.0411      2.07     26.76   1669.59   0.008
  4    0.0576      2.07     26.82   1879.64  -0.026
  5    0.0741      2.07     27.49   2038.74   0.018
  6    0.0906      2.11     27.36   2207.69   0.059
  7    0.1071      2.11     27.90   2346.92   0.044
  8    0.1236      2.07     28.22   2466.71   0.071
  9    0.1401      2.11     28.18   2665.51   0.121
 10    0.1566      2.58     30.37   2618.73   0.133
 11    0.1731      2.58     30.33   2691.93   0.084
-------------------------------------------------------

 beta =    31.65 A**2/MGy   gama =    4.29  1/MGy
 Dose_1/2_th =  0.157 MGy    Dose_1/2 = 0.163 MGy
 Relative Radiation Sensitivity  =   59.12
```

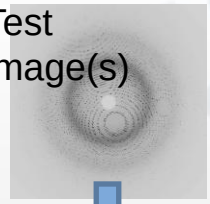
Example results from ”burning strategy”



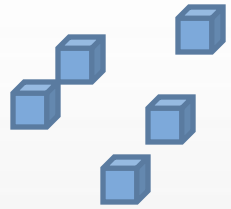


Two-dimension DC





X-ray

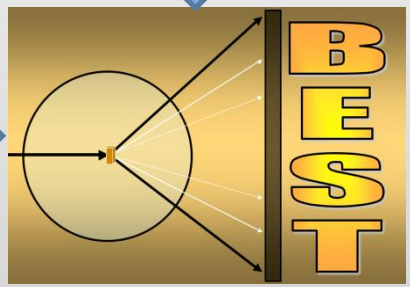


MOSFLM XDS

Space group, Cell parameters, Orientation

Already collected data

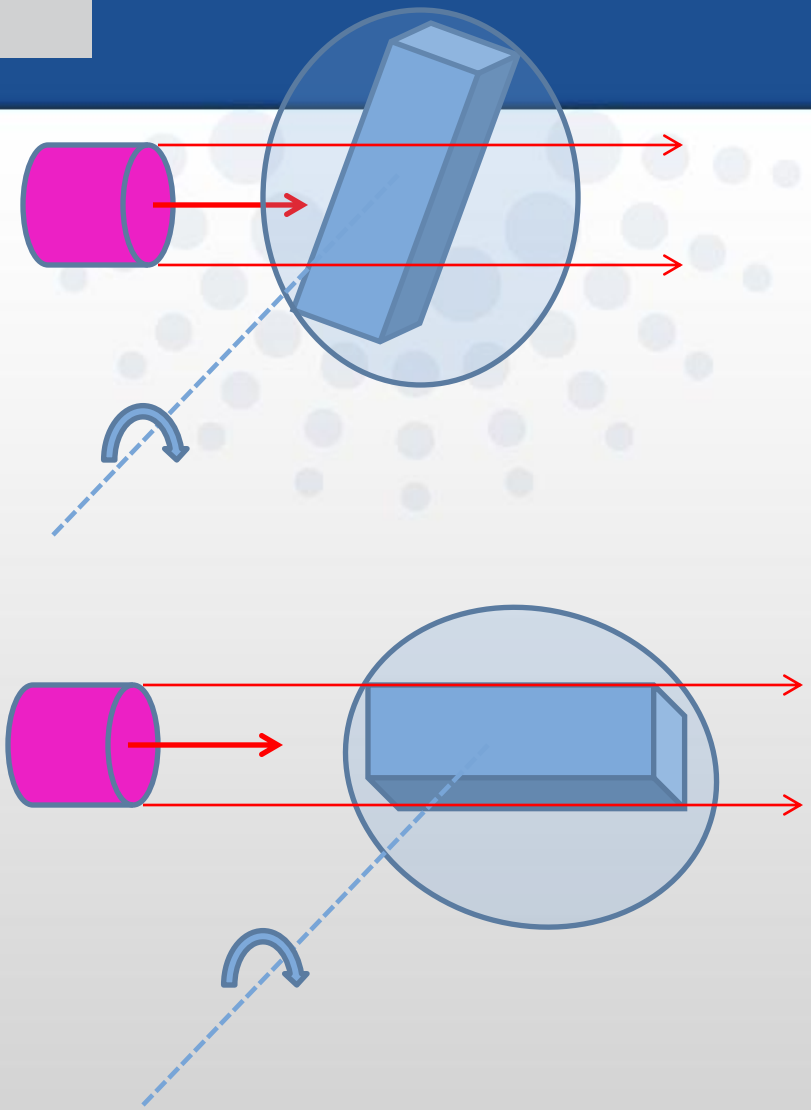
Number of crystals



Data Collection Strategy

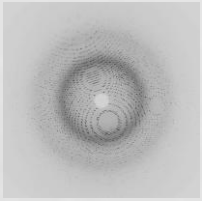
Data Collection

Auto Processing



$$\hat{J}(\mathbf{h}, D, \Omega) = \hat{J}(\mathbf{h}, D = 0)\, scale(D, \Omega) \exp(-\mathbf{h} \cdot \mathbf{B}(D, \Omega) \cdot \mathbf{h}^T / 2)$$

Data collection strategy accounting radiation damage



Initial Images

MOSFLM XDS

Beamline Flux
Crystal contents

Crystal shape and size
Beam profile and size

User choices

- Optimize data collection
- Optimize SAD data collection
- Find optimal crystal orientation
- Low-resolution optimal
- Rad. Damage sensitivity
- Multi-positional data collection
- Helical data collection
- Estimate data statistics

Dose (Time) limit
Geometry limits
Aimed statistics
Aimed completeness
Aimed redundancy
Aimed resolution

Detector parameters
Beamline parameters
and limitations

BEST 5.1

Optimal plan(s) of data collection
Statistics
B-factor

Diffraction sample Modeling

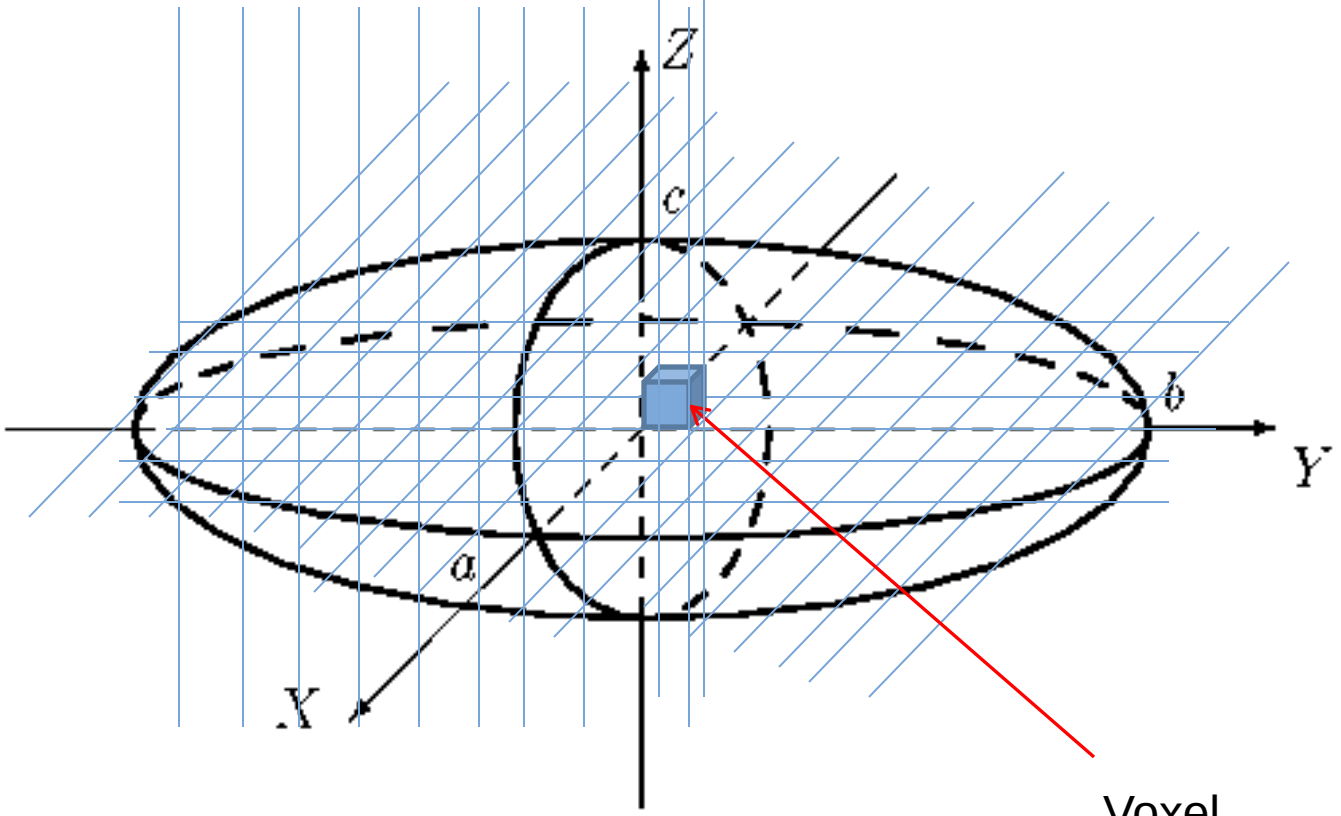
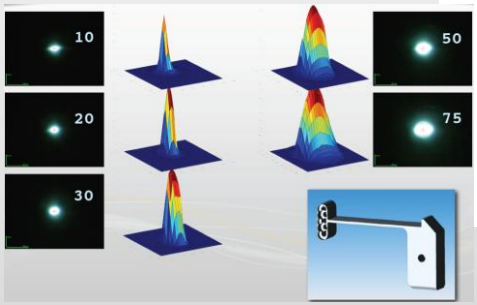


Рис. 1

Voxel
Volumetric Picture Element

$$Scale(\Omega) = Scale(voxel) \times NumberVoxel(\Omega)$$



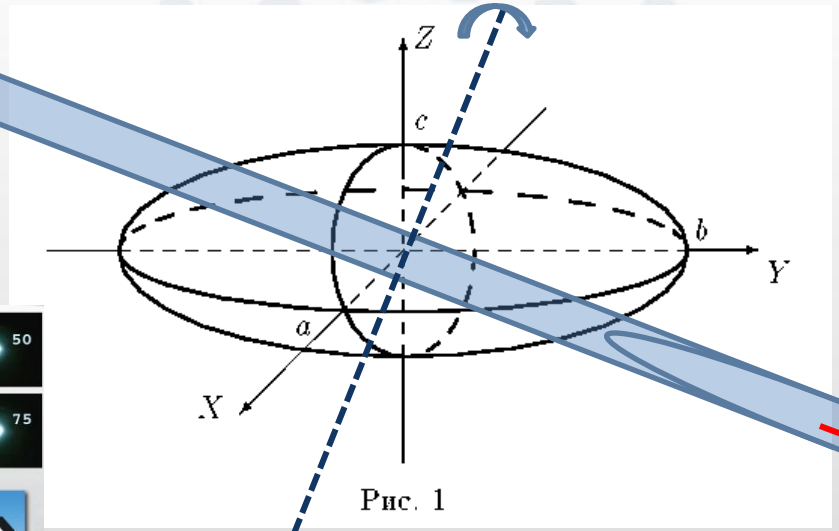
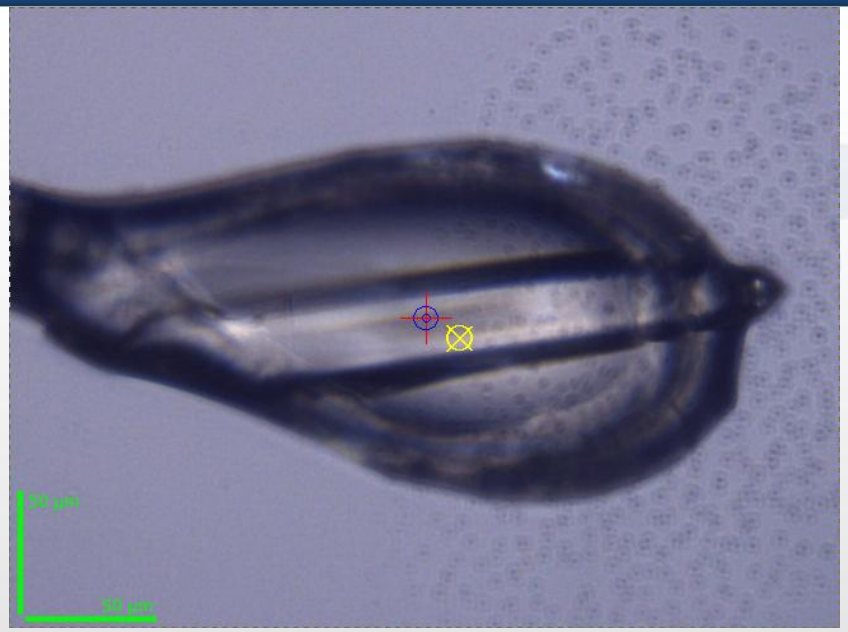
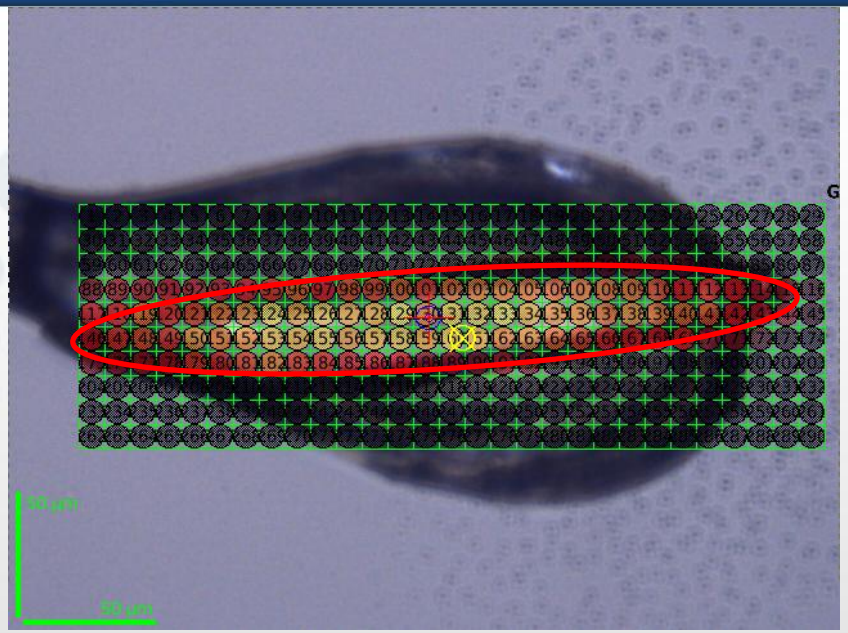


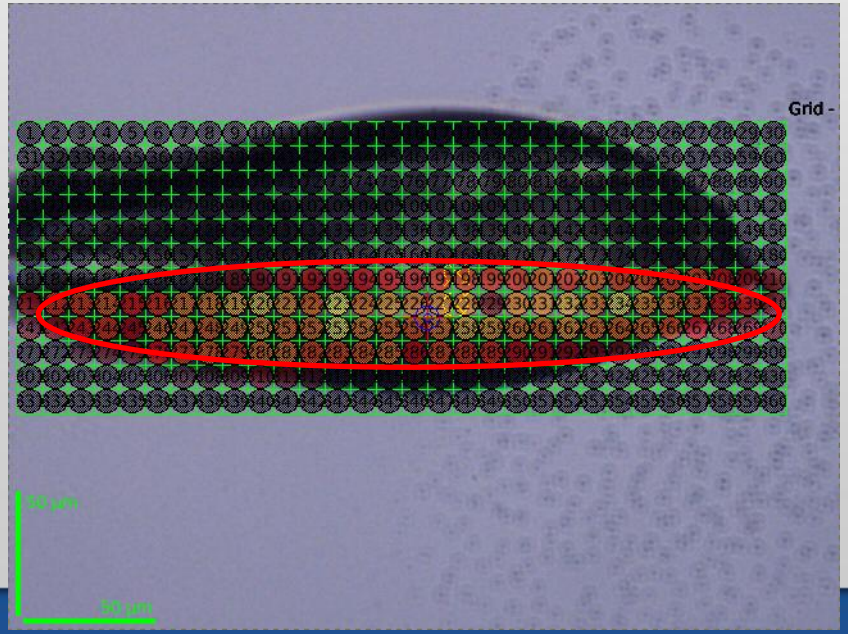
Рис. 1

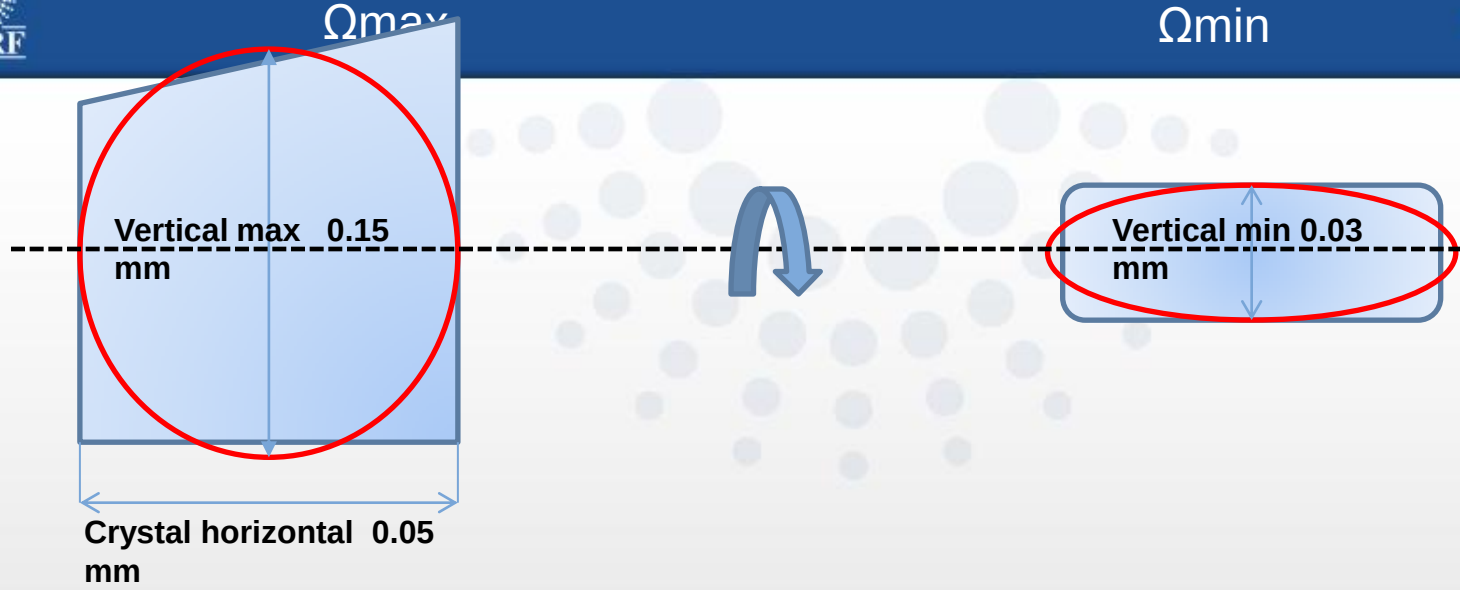
$$\hat{J}(\mathbf{h}, D) = \hat{J}_o(h) \sum_{voxel} \sum_{x,y} I_{x,y}(beam) \times scale(voxel, D_{voxel}) \exp(-\mathbf{h} \cdot \mathbf{B}(D_{voxel}) \cdot \mathbf{h}^T / 2)$$

Diffraction intensity at any moment of data collection is calculated as the sum of diffraction intensities from crystal voxels taking into account the profile of primary X-ray beam and the dose absorbed by each voxel.

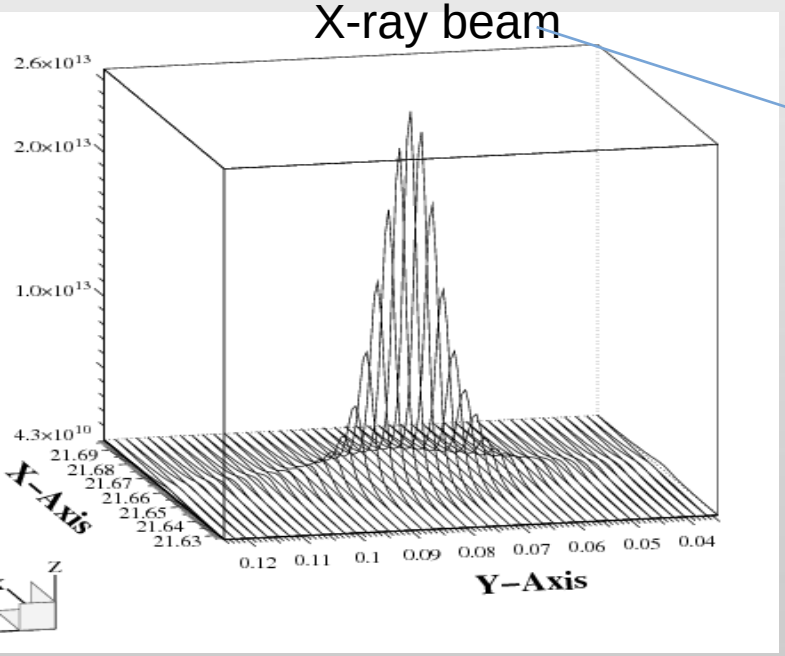





Grid -



X-ray beam



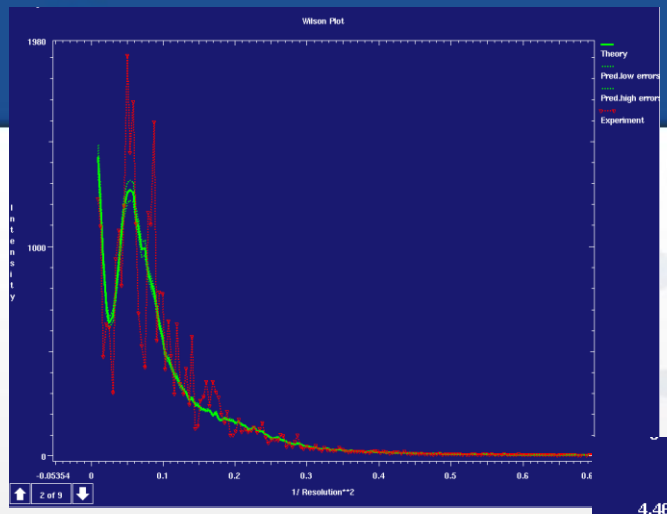
beam_crystal.dat - /mntdirect/_users/apopov/BEST4.1/T3D/

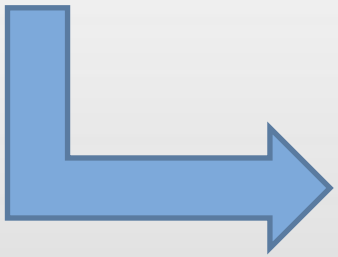
File Edit Search Preferences Shell Macro Windows

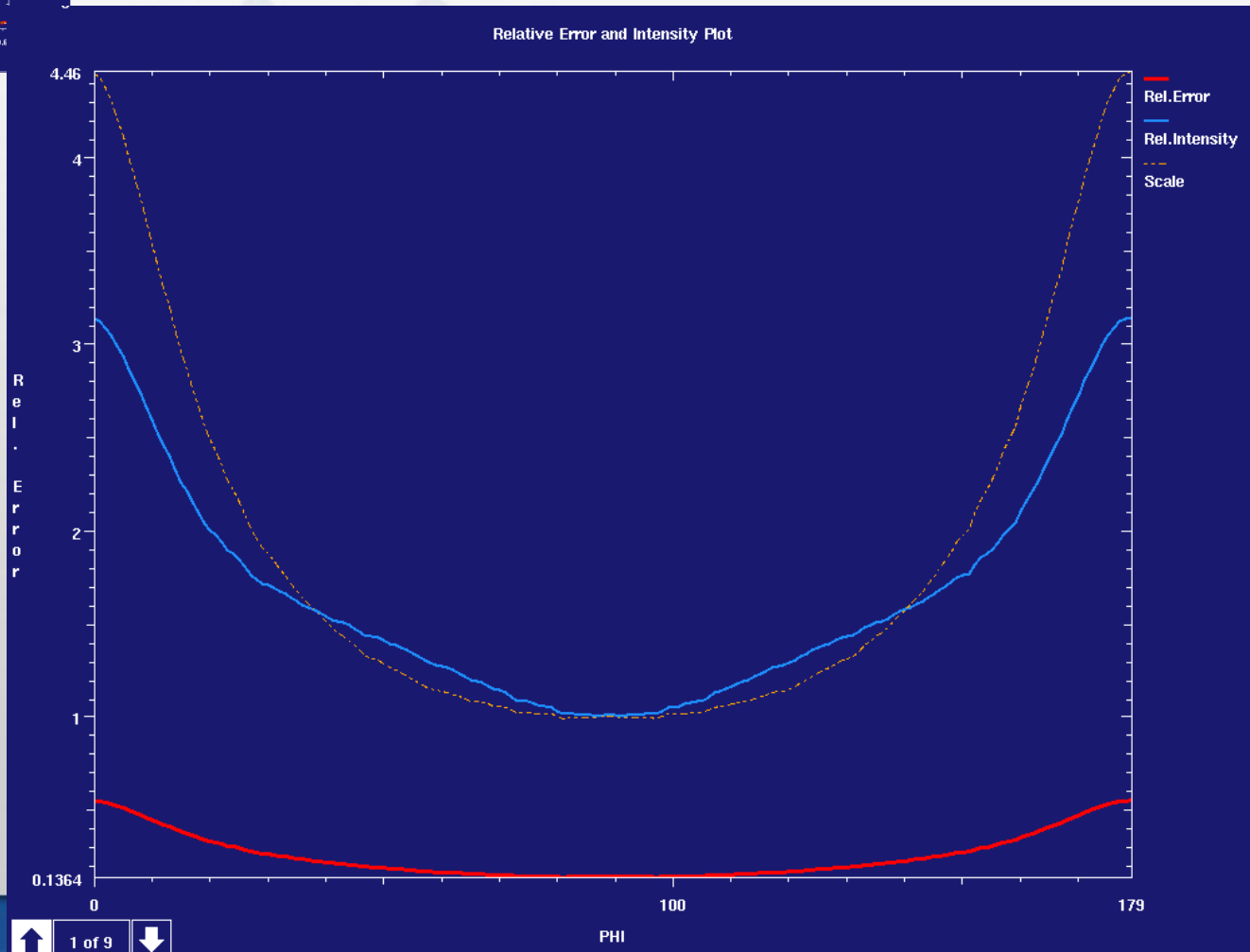
```
!   TRYPSIN TEST
! all sizes in mm
! beam size
horizontal_size 0.045
vertical_size 0.035
aperture_size 0.030
!defalt: no aperture
!horizontal_slit 0.1000
!vertical_slit 0.1000
beam_shift 0.0
!vertical shift relative to the rotation axis
!
crystal_vert_max 0.150
crystal_vert_min 0.03
crystal_hor 0.050
omega_min 0
!description of crystal shape and position- a,c,b
end
```

First step - scaling

```
Scaling
 Relative scale         :   77.02
 Overall B-factor       :  12.63 Angstrom^2
 B-factor eigenvalues   :   9.84   12.45   18.01 Angstrom^2
 Scaling error          :  13% at the resolution limit
```



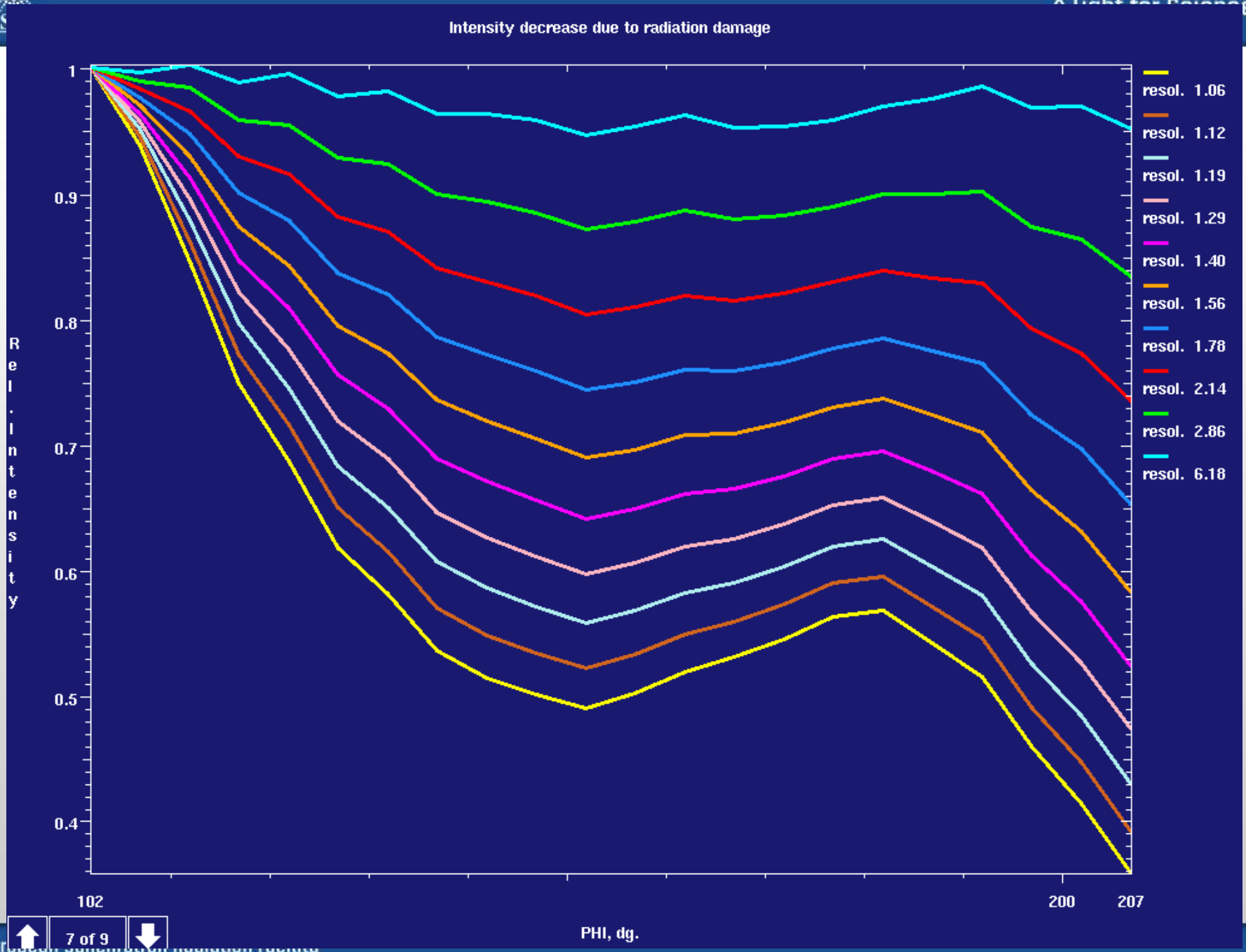


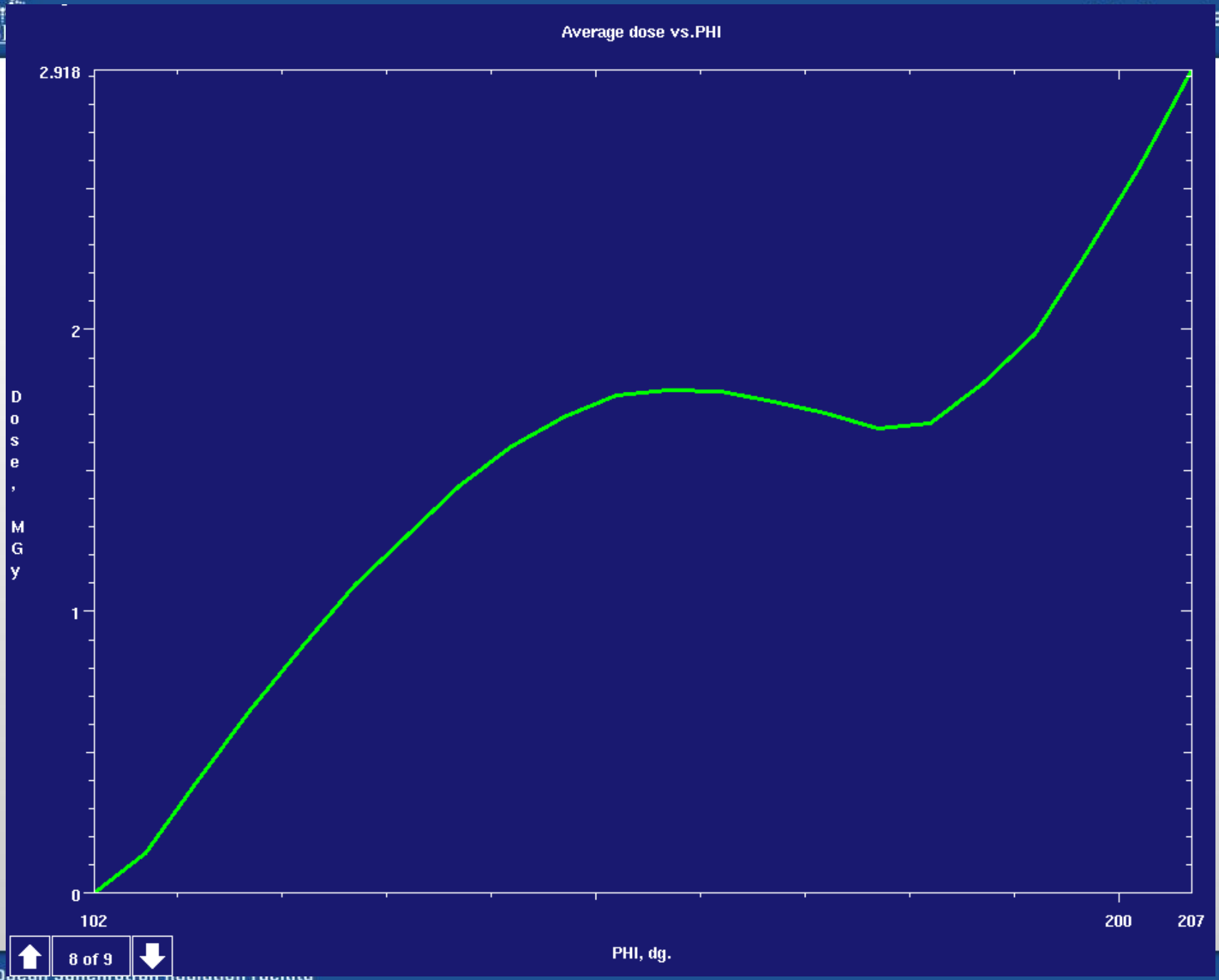


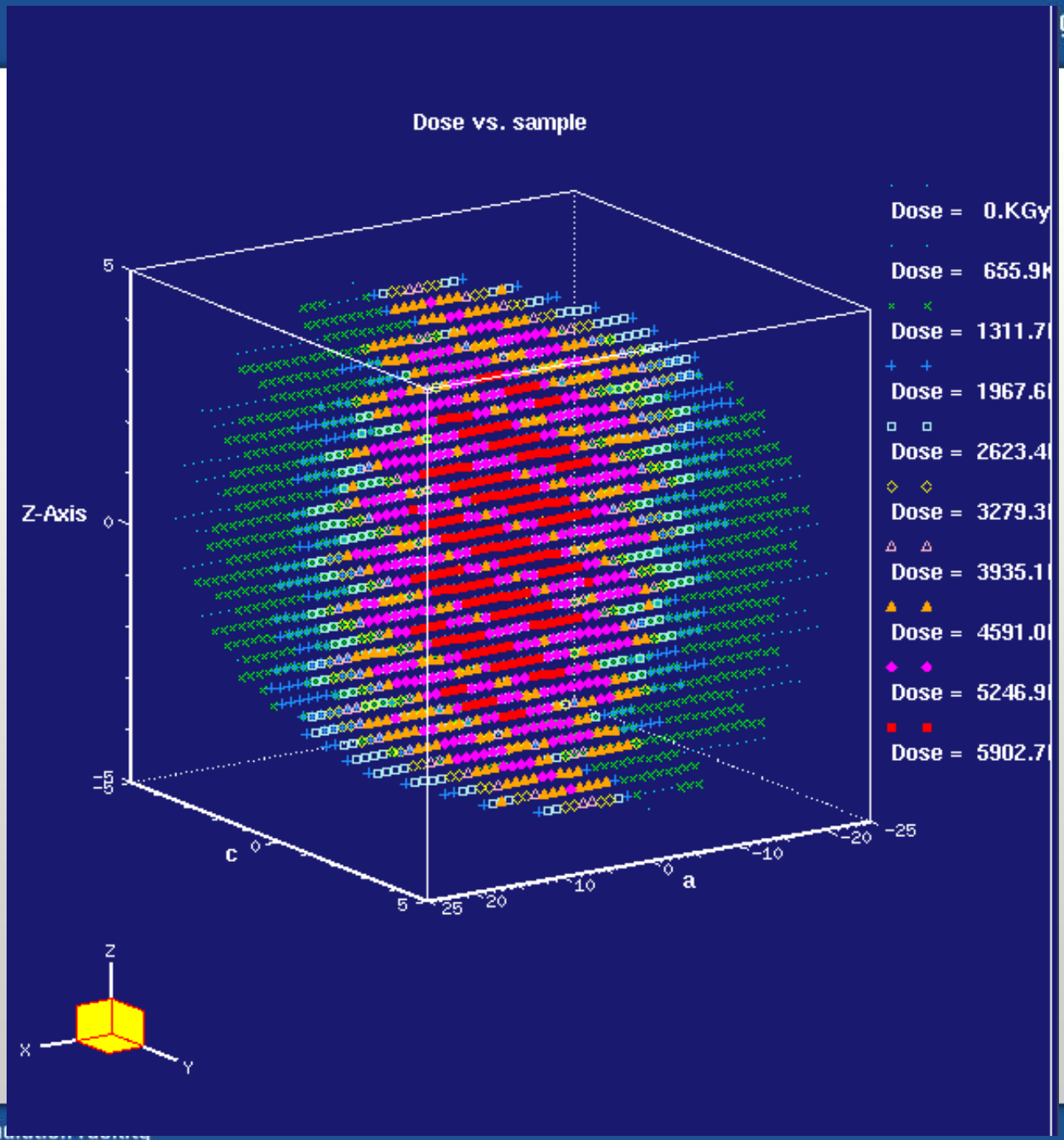
```
Program Best /A.Popov & G.Bourenkov/
Version 5.1.0 //  07.07.2015
Copyright 2004-15 by Alexander Popov and Gleb Bourenkov
<B><FONT COLOR="#FF0000"><!--SUMMARY_BEGIN-->

                                  Main Wedge
                                 ================
      Resolution limit is set by the radiation damage
 Resolution limit =1.06 Angstrom   Transmission =    8.4%  Distance = 150.3mm
```

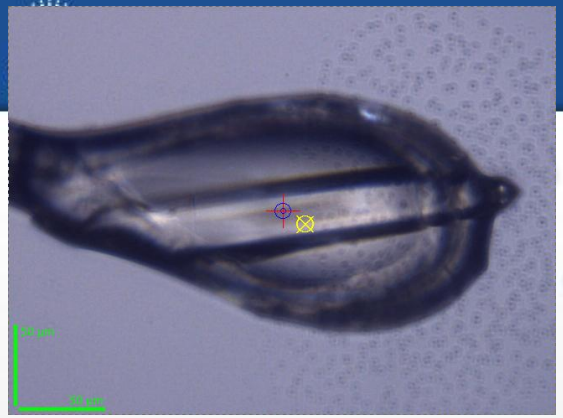
WEDGE PARAMETERS					INFORMATION						
sub-We-dge	Phi start degree	Rot. width degree	Exposure /image s	N.of ima-ges	Over-lap	sWedge width degree	Exposure /sWedge s	Exposure total s	Dose /sWedge MGy	Dose total MGy	Comple-teness %
1	102.00	0.15	0.059	467	No	70.05	27.6	27.6	3.666	1.701	98.0
2	172.05	0.15	0.037	233	No	34.95	8.6	36.2	1.144	2.918	99.1

```
 Phi_start - Phi_finish      : 102.00 - 207.00
 Total rotation range        : 105.00 degree
 Total N.of images           : 700
 Overall Completeness        : 93.3%
 Redundancy                  : 4.29
 R-factor (outer shell)      : 4.4% (   74.2%)
 I/Sigma (outer shell)       : 48.1 (   2.2)
 Total Exposure time         : 36.2 sec (0.010 hour)
 Total Data Collection time  : 72.2 sec (0.020 hour)
<!--SUMMARY_END--></FONT></B>
```

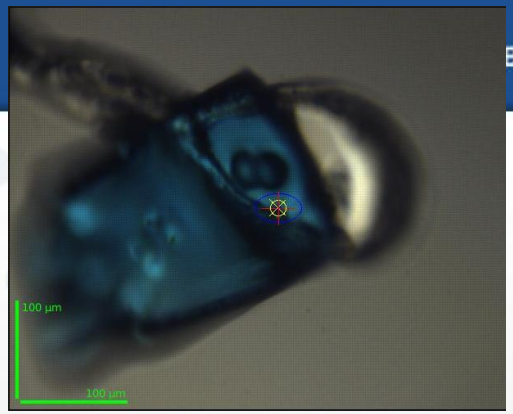

Intensity decrease due to radiation damage
Rel. Intensity
1
0.9
0.8
0.7
0.6
0.5
0.4
102
200
207
PHI, dg.
resol. 1.06
resol. 1.12
resol. 1.19
resol. 1.29
resol. 1.40
resol. 1.56
resol. 1.78
resol. 2.14
resol. 2.86
resol. 6.18

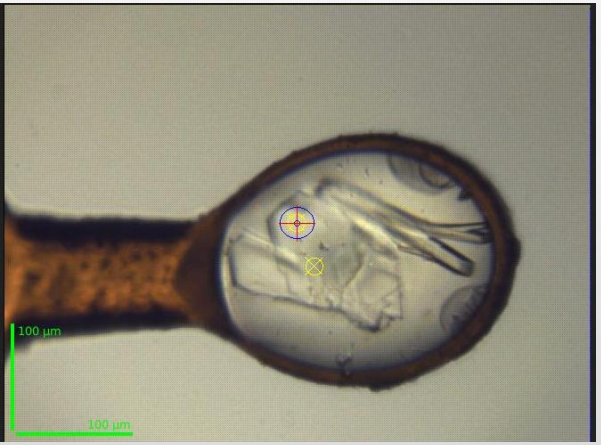

Average dose vs.PHI
2.918
2
1
0
Dose, MGy
102
200
207
PHI, dg.
8 of 9

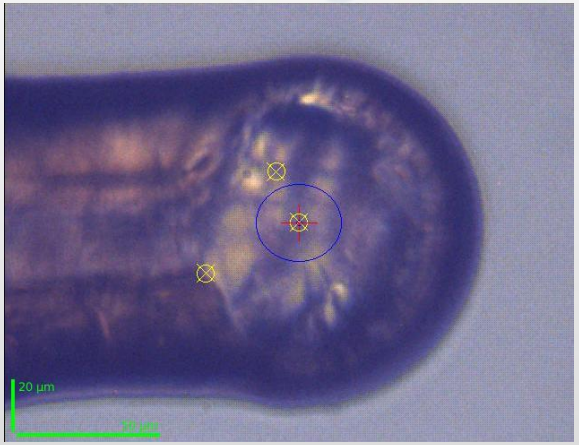

Dose vs. sample
Dose = 0.KGy
Dose = 655.9
Dose = 1311.7
Dose = 1967.6
Dose = 2623.4
Dose = 3279.3
Dose = 3935.1
Dose = 4591.0
Dose = 5246.9
Dose = 5902.7
Z-Axis
5
0
-5
c
0
5
a
25
20
10
0
-10
-20
-25
Z
X
Y

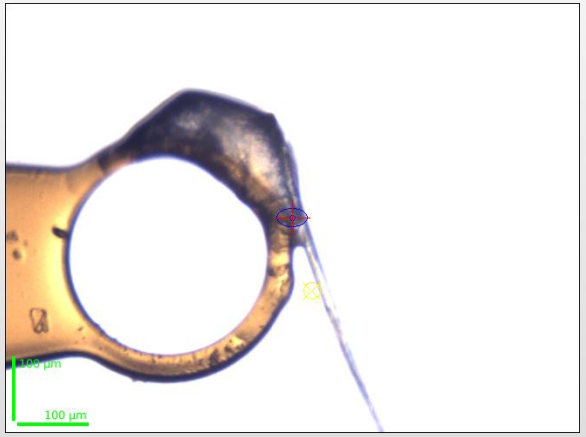

100 µm
100 µm

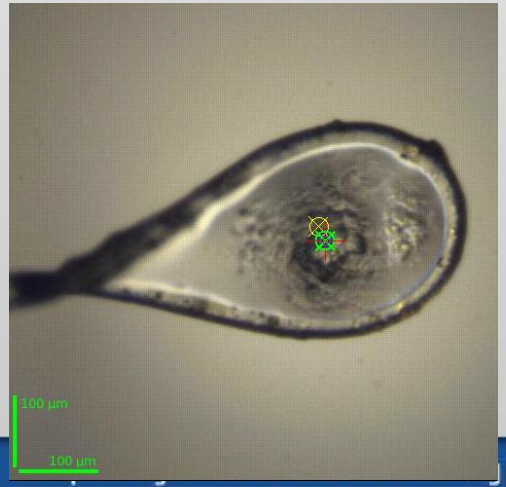

100 µm
100 µm

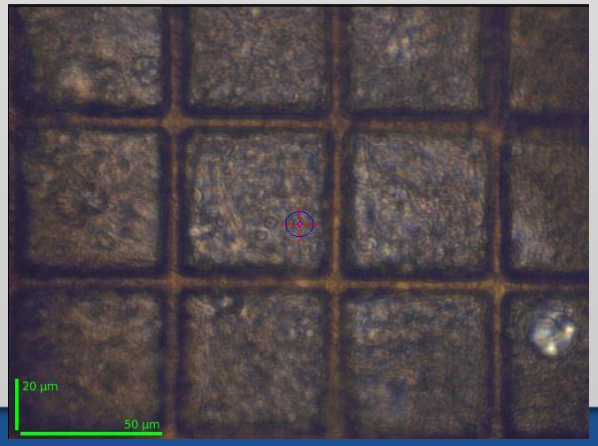

100 µm
100 µm

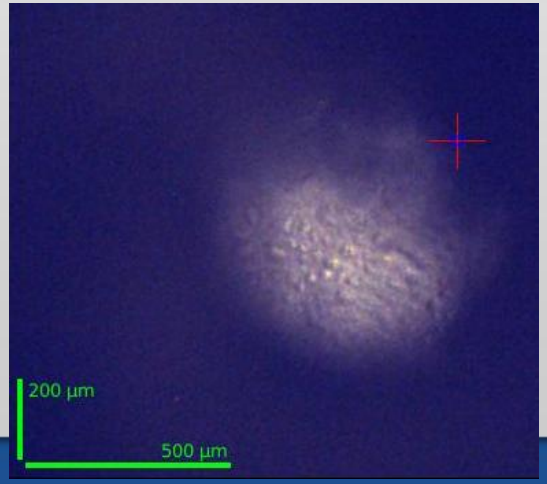

100 µm
100 µm


20 µm
50 µm


100 µm
100 µm


100 µm
100 µm


20 µm
50 µm

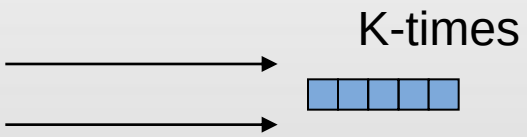

200 µm
500 µm

Multi-crystal and multi-positional strategy

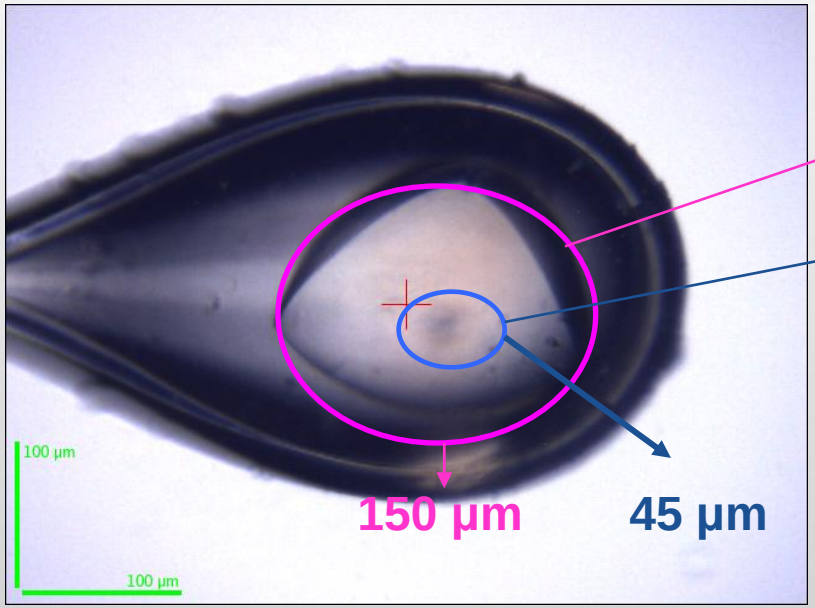
$$\sigma (I_{peak}) = SQRT (I_{peak} + I_{background})$$

$$I_{peak} / \sigma(I_{peak}) = n$$

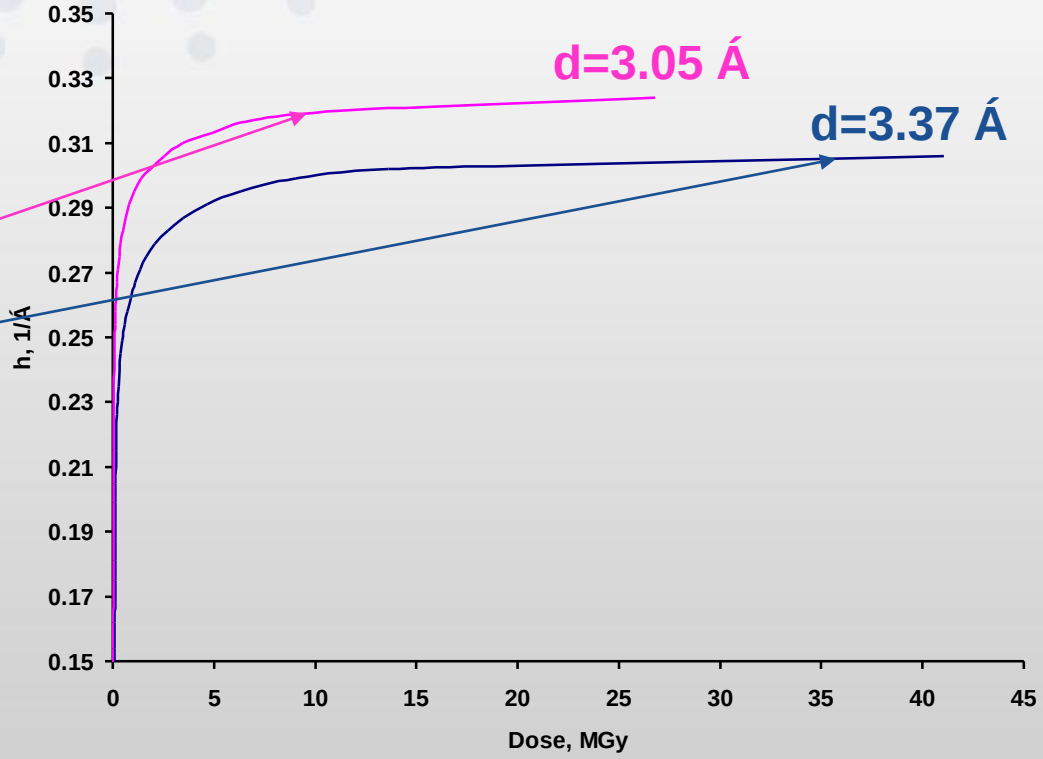
$$I_{peak} / \sigma(I_{peak}) = n*Sqrt(k)$$



The 70 kDa membrane protein FtsH from Aquifex aeolicus
I222, a = 137.9, b = 162.1, c = 170



Diffraction resolution vs. absorbed dose



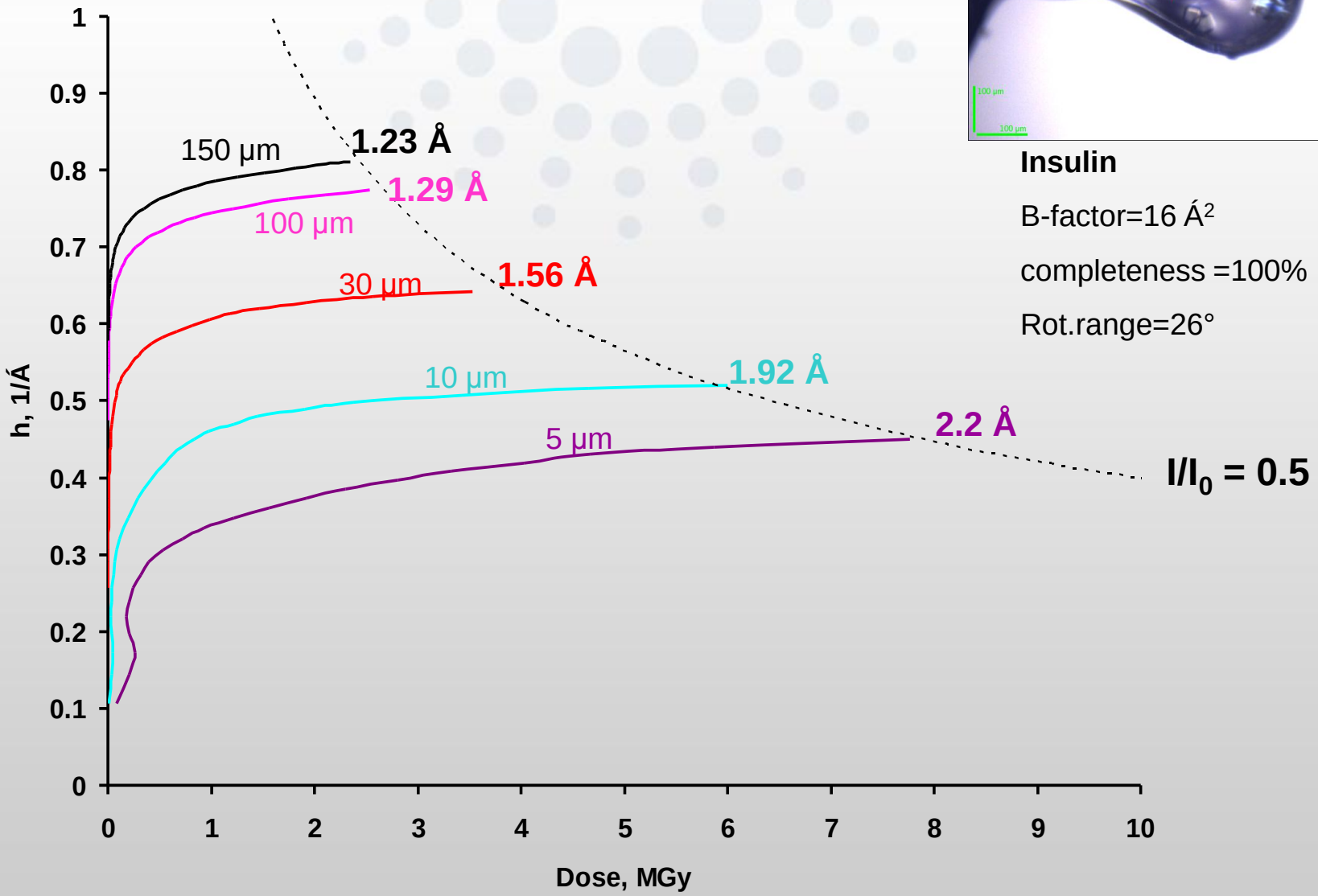
Diffraction resolution vs. absorbed dose

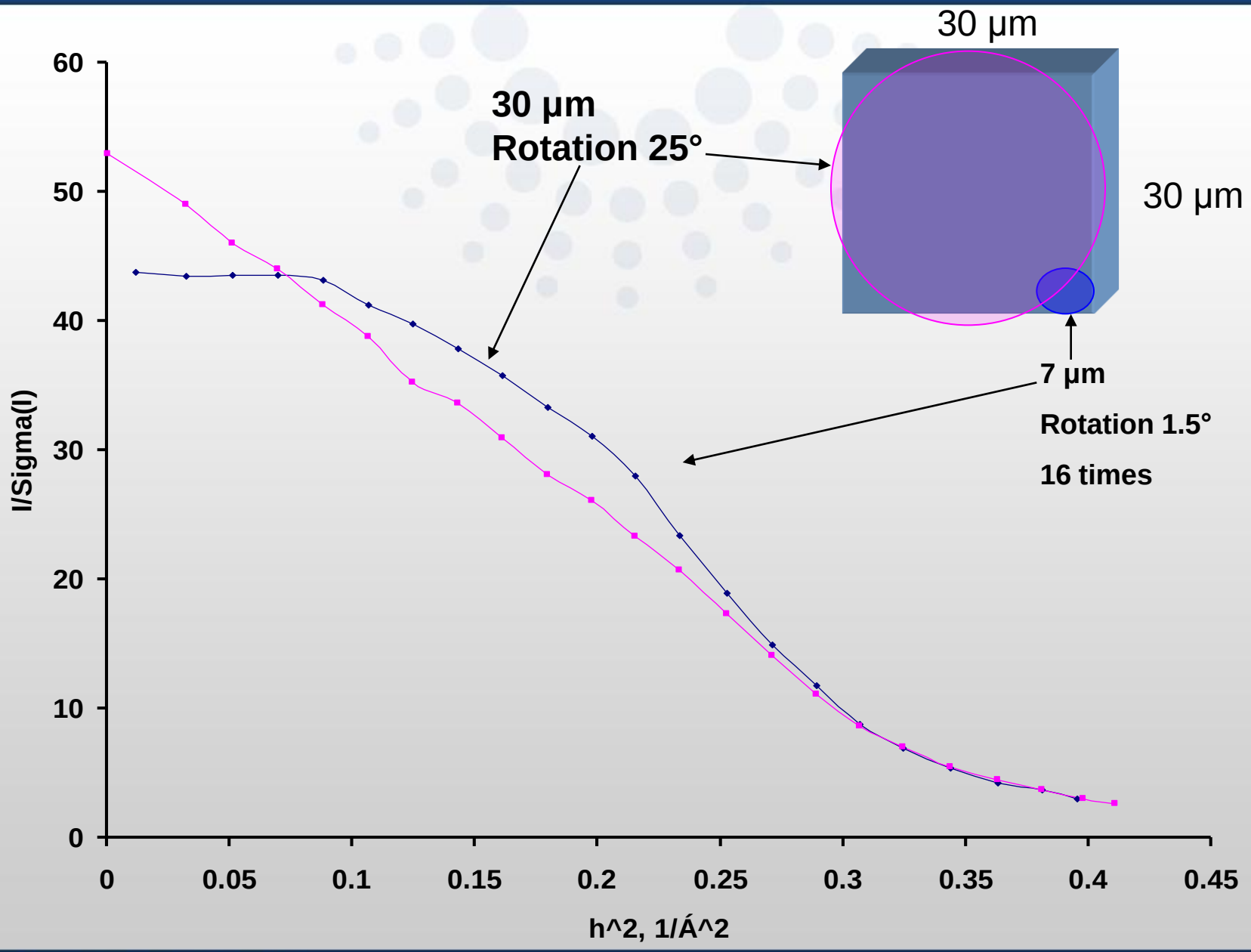
Insulin

B-factor=16 Á^2

completeness =100%

Rot.range=26°

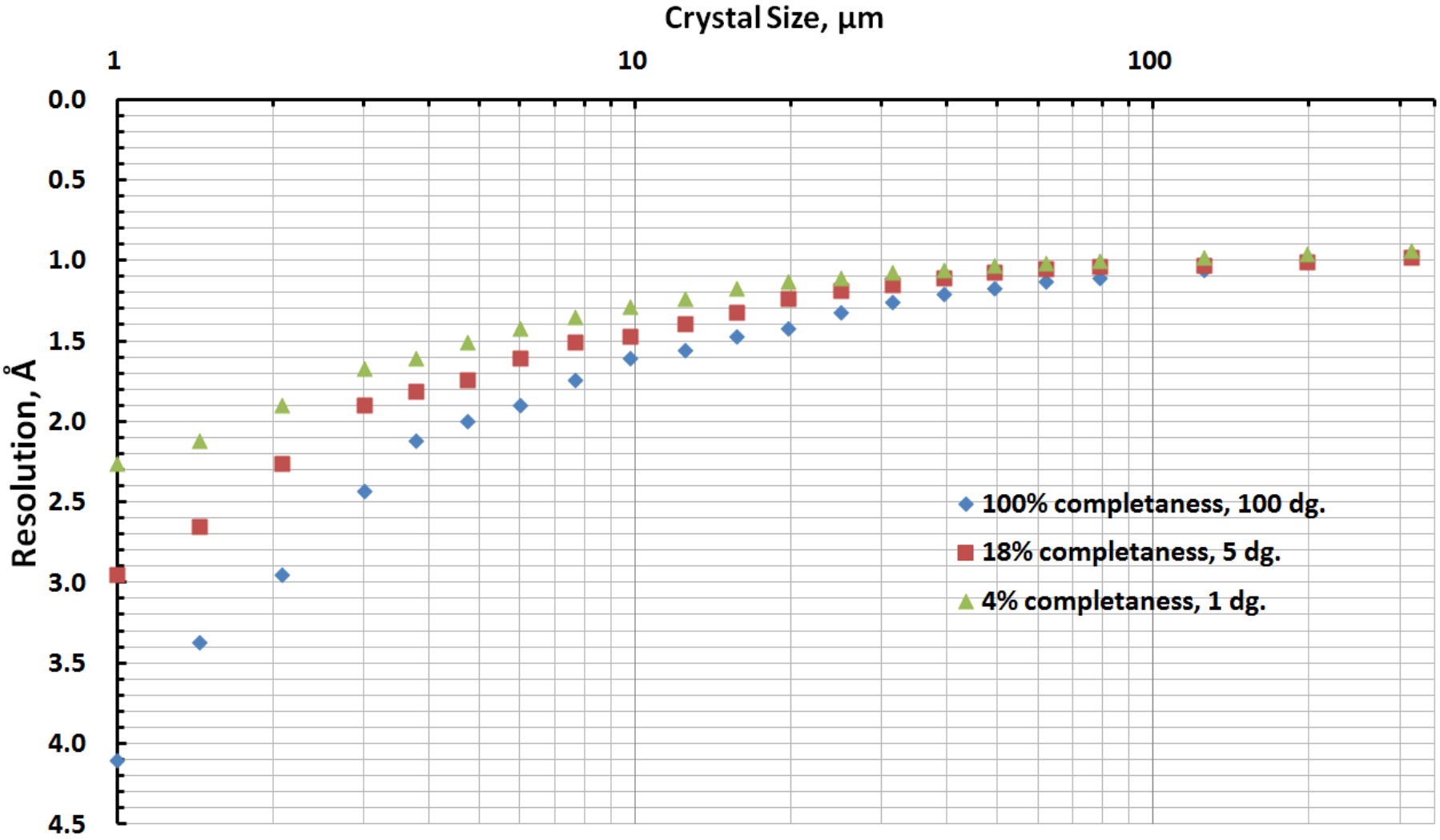



30 µm
30 µm
Rotation 25°
30 µm
7 µm
Rotation 1.5°
16 times
I/Sigma(I)
60
50
40
30
20
10
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
h^2, 1/Á^2

Achievable resolution *vs.* crystal size

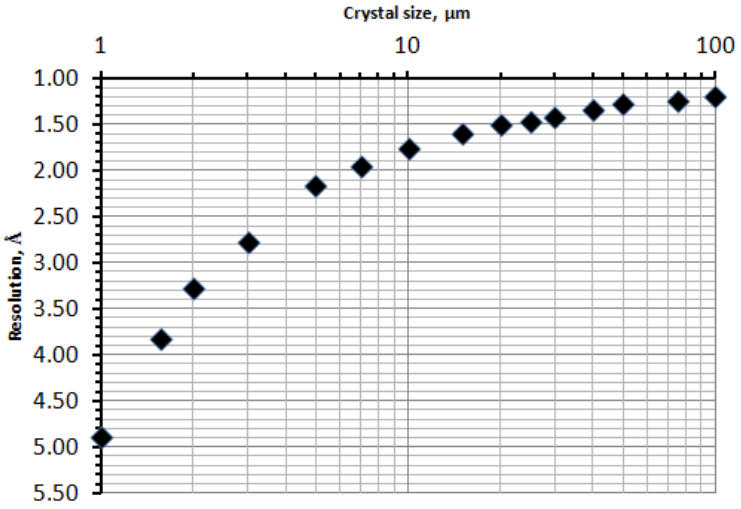
Lactase, B-factor=9.6 Å²

P222

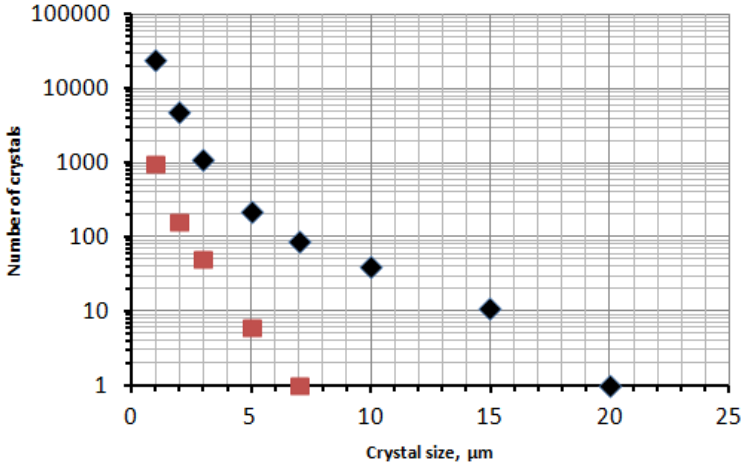


Micro-crystallography

• Thermolysin, Space Group $P6_122$; B-factor=11.5 Å^2



Complete data set resolution *vs.* crystal size



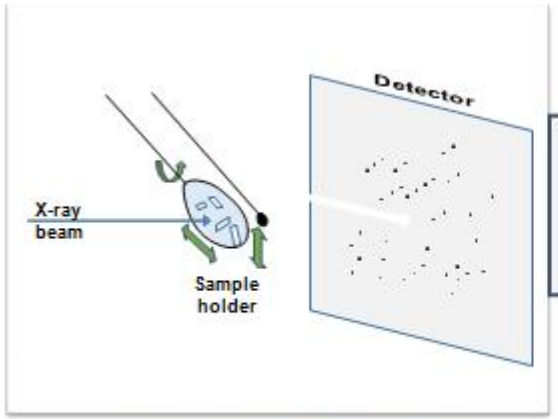
Number of cryocooled crystals of a given size required to achieve dataset resolutions of 1.5 Å (black) and 2.0 Å (blue).

• For a crystal 1x1x1 µm^3 in dimensions partial data sets *from about 1000 crystals* would be needed to achieve a final data set resolution of d_{min} = 2.0 Å .

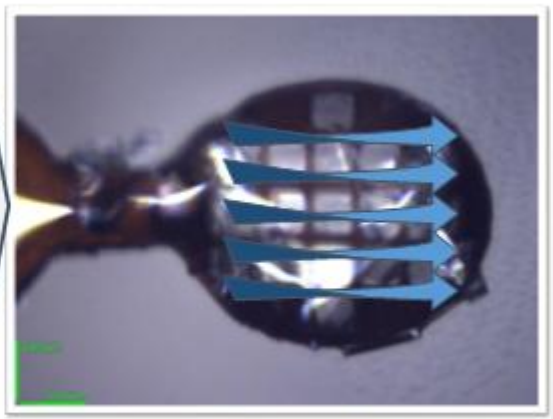
Use of micro-beams for diffraction data collection from the often in-homogeneously diffracting crystals of macromolecules allows improving the ratio of measured signal to noise and increases achievable resolution. In order utilize the advantages of micro-beam conditions, all available material of a crystal must to be used for measurements thus reducing the inevitable radiation damage effects.

X-ray Mesh Scans

micro X-ray beam, high-precision diffractometry , shutterless data acquisition with a pixel-array detector



Experimental setup for X-ray crystallography



Translation with slight rotation is the principle of Mesh Scan data collection

The Complex Analysis of the X-ray Mesh Scan

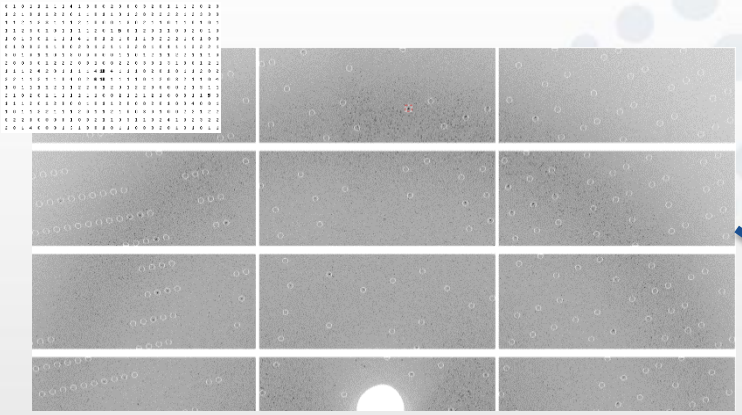
DOZOR

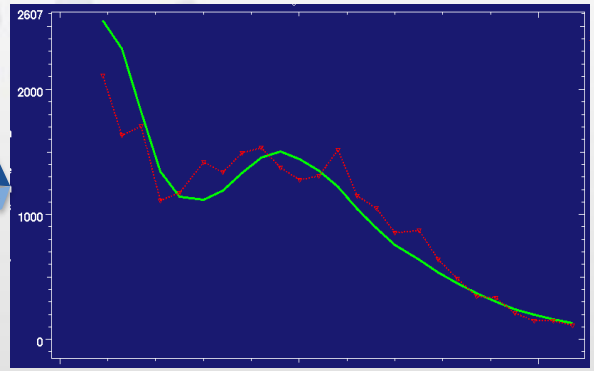
Crystals Maping – positions, size, diffraction quality

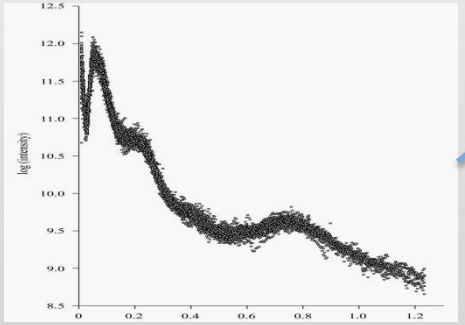
BEST – Data Collection Strategy accounting radiation damage

Data Collection and Processing

Evaluating diffraction signal with DOZOR







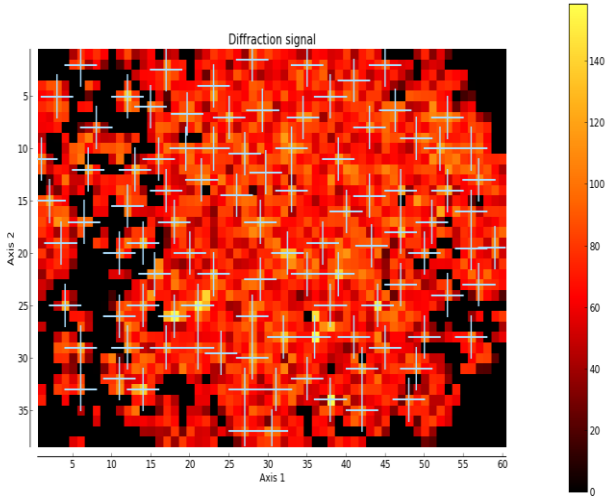
score =

total scattered intensity ×
radial shape similarity

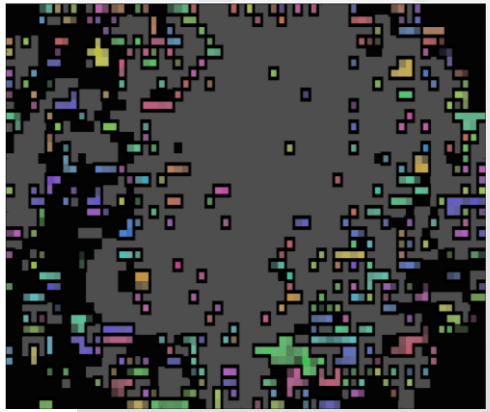
- Use Wilson plot as a prior
- Use all pixels, not just the local maxima

```
 Program dozor /A.Popov & G.Bourenkov/
 Version 1.3.3 //  18.05.2015
 Copyright 2014 by Alexander Popov and Gleb Bourenkov
 N    |            SPOTS    |        Powder Wilson                  |
image | num.of  INTaver Res.|   Scale B-fac. Res. Corr. R-factor|      Score
-------------------------------------------------------------------------------
  400 |     0        0. 50.0| ---------no results -----------   |       0.000
  401 |     0        0. 50.0| ---------no results -----------   |       0.000
  402 |     2        4. 10.0| ---------no results -----------   |       0.000
  403 |    20        7.  5.4| ---------no results -----------   |       0.005
  404 |    30       28.  5.9|   40.16    1.0  2.5  29.5    17.9 |       1.784
  405 |    14        6.  7.9| ---------no results -----------   |       0.001
  406 |    23        4.  6.7| ---------no results -----------   |       0.002
  407 |     3        2.  7.9| ---------no results -----------   |       0.000
  408 |    21       11.  7.2|  187.02    4.6  2.5  45.0    19.5 |       0.415
  409 |    31       69.  6.7|   47.83   -2.0  2.5  39.1    14.2 |       3.424
  410 |     1        1. 10.0| ---------no results -----------   |       0.000
  411 |    10        5.  6.7| ---------no results -----------   |       0.001
  412 |    31       60.  7.2|   29.01    1.6  2.5  38.3    15.8 |       2.430
  413 |    40       54.  7.2|   38.11    7.8  2.5  29.1    23.0 |       1.911
  414 |    19       14.  6.3|  115.30    0.7  2.5  39.2    16.3 |       0.889
  415 |    21       81.  6.7| ---------no results -----------   |       0.024
  416 |    35       81.  7.2|   20.12    4.1  2.5  27.2    22.1 |       3.321
  417 |    33       52.  7.2|   33.57    2.6  2.5  35.5    20.4 |       2.271
  418 |    12        7.  8.7| ---------no results -----------   |       0.000
  419 |    16        6.  7.9|  441.01   -3.2  2.5  34.9    16.4 |       0.156
  420 |    33       18.  7.2|  171.51   -0.8  2.5  30.6    17.2 |       0.633
  421 |    13       32.  7.2| ---------no results -----------   |       0.006
  422 |     0        0. 50.0| ---------no results -----------   |       0.000
  423 |     1        2. 10.0| ---------no results -----------   |       0.000
  424 |     9       18.  7.2| ---------no results -----------   |       0.003
```

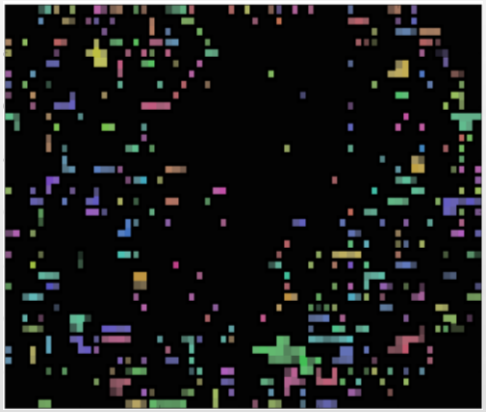
Dozor detects single crystals in the crystal mess as well as crystal overlapping



Initial diffraction map, showing regions with better protein crystal diffraction quality



Crystal map after Dozor analysis (different colors correspond to different homogeneities; gray regions correspond to pattern overlapping)



Crystal map omitting regions of overlapping

Thaumatin ~50 µm crystals measured on ESRF ID23_1, beamsize 20µm, oscillation 0.017°

Dozor determine regions of crystal homogeneity and size

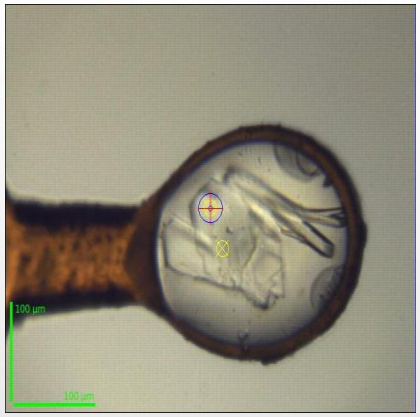
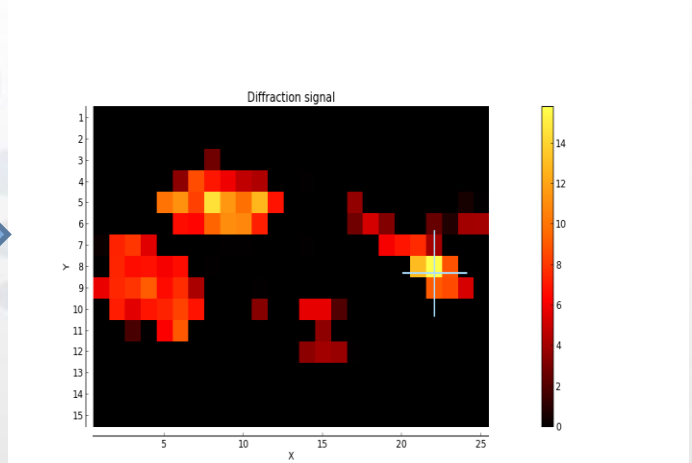
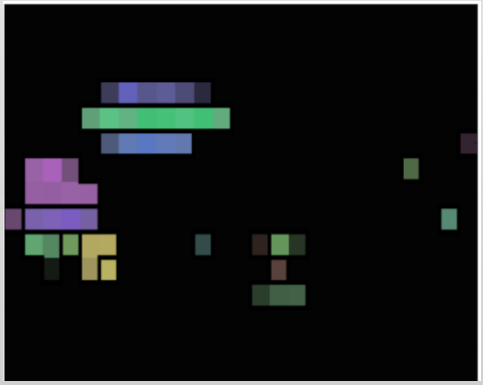


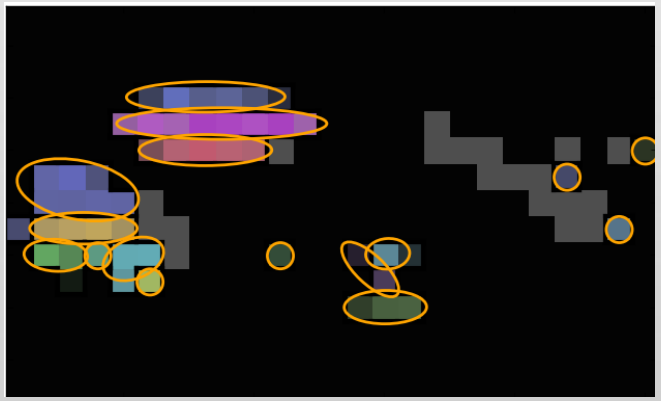
Photo snapshot of the sample



Initial diffraction map, showing regions with better protein crystal diffraction quality



Crystal map omitting regions of overlapping

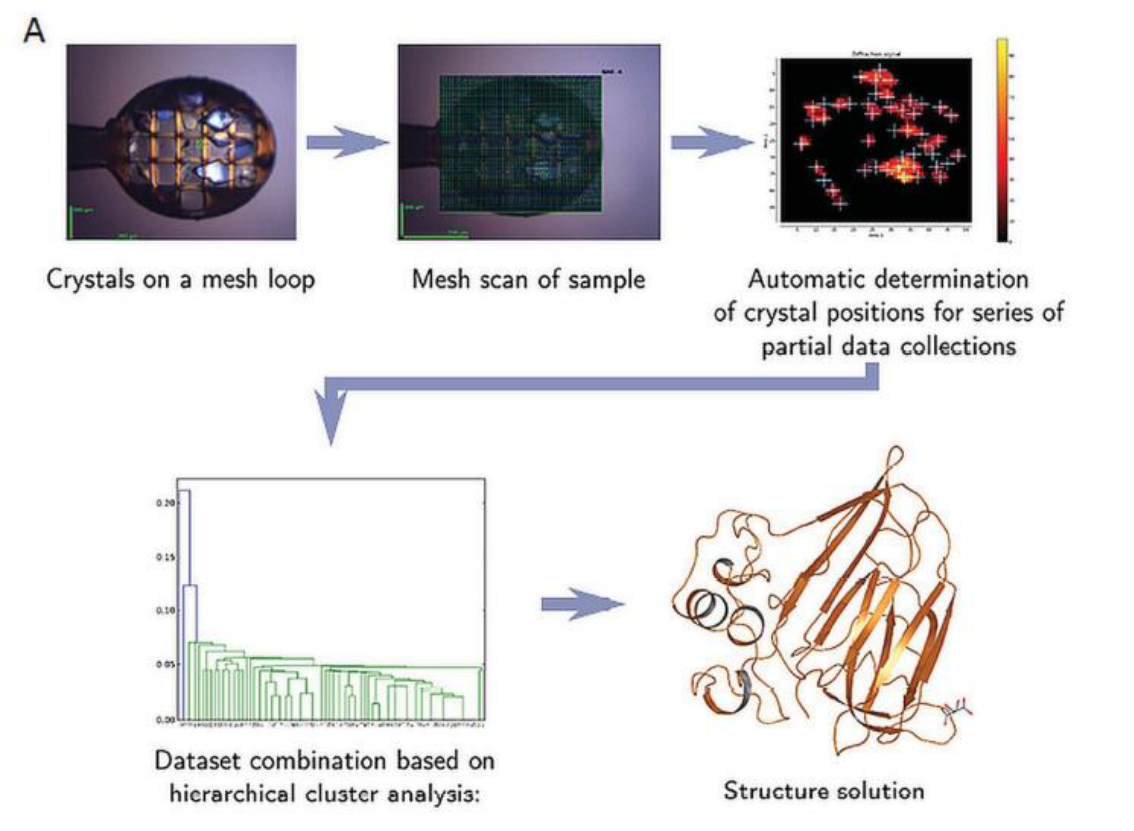


fit crystal shapes, corresponding to regions of crystal integrity.

(measured on ESRF ID23_1, beamsize 10µm, oscillation 0.1°)

MeshAndCollect: an automated multi-crystal data collection workflow for synchrotron macromolecular crystallography beamlines

Ulrich Zander *et al*., Acta Crystallographica Section D71, 2328-43



The 'MeshAndCollect' workflow for multi crystal data collection method.

mesh scan is performed on the sample. The resulting images are automatically inspected for protein diffraction and scored according to diffraction strength. A heat map is generated that represents the diffraction intensity where the positions for partial data collections are marked. After the user has selected the settings for the partial data collections, the MxCuBE2 data collection queue is automatically filled and all partial data sets collected. Once the partial datasets have been automatically processed, HCA can then be used to choose which data sets to merge to produce a final data set for structure solution.

Acknowledgements

- Bourenkov Gleb
- Igor Melnikov
- Olof Svensson
- Daniele de Sanctis
- Christoph Mueller-Dieckmann
- Gordon Leonard
- Ulrich Zander
- Andrew McCarthy
- Matthew Bowler