A Light for Science



European Synchrotron Radiation Facility



Looking forward to the fully automated and optimal data collection using synchrotron radiation

Alexander Popov ESRF, MX group



European Synchrotron Radiation Facility



A light for Science

- Data quality is better
- Measurements are too difficult

Micro-crystallography



• Thermolysin, Space Group P6122; B-factor=11.5 Å²

• For a crystal $1x1x1 \mu m^3$ in dimensions partial data sets *from about 1000 crystals* would be needed to achieve a final data set resolution of $\underline{d}_{min} = 2.0 \text{ Å}$.



Why do we need automation?

• We are lazy and we don't want to do boring work. We don't know and we don't want to know a crystallography.





Core technology



RoboDiff

Nurizzo D, Bowler MW, et al. (2016), Acta Cryst D. accepted

European Synchrotron Radiation Facility



how we can realize automation?

Old available SB beam lines

	Energy [keV]	Beam size [mm²]	Flux [ph/s]	detector	Frame rate [Hz]			
ID23-1	6-20	10-40	3x1 - 10		50			
ID23-2	14.2	5x7	4x1 20		75			
ID29	6-20	10-50	5x1					
ID29S	optical spec	ctroscopy (CRYOB	ENCH; U					
BM29	7-15	500 (100)	2x1					
New SB beam lines								
MASSIF-1	12.8	20-150	10 ¹³	Pilatus3 2M	250			
MASSIF-2	12.8	20-100	10 ¹³	tbd	Not finished			
MASSIF-3	12.8	>10	5x10 ¹³	Eiger 4M	750			
ID30B	6-20	20-200	10 ¹³	Pilatus3 6M	100			



A light for Science



PILATUS3 S Series

ADVANTAGES S SERIES

- Maximum frame rate of 25 Hz
- Readout time of 2.03 ms

• Upgradeable to X series



PILATUS3 X Series

ADVANTAGES X SERIES

- Frame rates up to 500 Hz
- Readout time of 0.95 ms
- Region of interest readout

EIGER X SYNCHROTRON SERIES



The new EIGER X series provides ultimate performance for the most demanding synchrotron applications. Frame rates in the kilohertz range in combination with continuous readout open new horizons in time-resolved experiments and XPCS. Hitherto prohibitively slow scanning beam imaging techniques like ptychography become routinely possible. High resolution and coherent diffraction experiments benefit from the small pixel size and an outstanding point-spread function that is enabled by direct conversion of X-rays. Highest count rates per unit area ideally match the ever-increasing beamline brightness. Read more about the EIGER X series here.

PILATUS3 X detector series technical specifications

PILATUS3 X	200K-A	300K	300K-W	1M	2M	6M		
Number of detector modules	1×2	1×3	3×1	2×5	3×8	5 x 12		
Sensitive area: width × height [mm²]	83.8 × 70.0	83.8 × 106.5	253.7 × 33.5	168.7 × 179.4	253.7 × 288.8	423.6 × 434.6		
Pixel size [µm²]			172 ა	× 172				
Number of pixels: hor. × ver.	487 × 407	487 x 619	1475 × 195	981 x 1043	1475 × 1679	2463 × 2527		
Gap width: hor. / ver. [pixel]	- / 17	- / 17	7/-	7 / 17	7 / 17	7 / 17		
Dead area [%]	4.3	5.5	0.9	7.2	8.0	8.5		
Defective pixels			< 0.0	03%				
Maximum frame rate, full frame [Hz]	500	500	500	500	250	100		
Maximum frame rate, ROI [Hz]				500	500	500		
Readout time [ms]			0.	95				
Point-spread function	1 pixel (FWHM)							
Threshold energy [keV]	3.5 - 18	2.7 - 18	2.7 - 18	2.7 - 18	2.7 - 18	2.7 - 18		
Counter depth			20 bits (1,048	3,576 counts)				
Power consumption [W]	30	30	30	165	250	580		
Dimensions (WHD) [mm³]	156 × 155 × 284	158 × 193 × 262	280 × 62 × 296	265 × 286 × 455	384 × 424 × 456	590 x 603 x 455		
Weight [kg]	5.4	7.5	7.0	25	46	92		
Module cooling	Air-cooled	Water-cooled	Water-cooled	Water-cooled	Water-cooled	Water-cooled		
Electronics cooling	Air-cooled	Water-cooled	Water-cooled	Air-cooled	Air-cooled	Air-cooled		
	450 μm silicon sensor							
Standard configuration	detector, detector server	detector, detector server, water-cooling unit		detector, detector server, water-cooling unit, PPU mini				
			1000 µm sil	icon sensor				
Detector options	PPU mini, L or XL	PPU mini, L or XL	PPU mini, L or XL	PPU L or XL	PPU L or XL	PPU L or XL		
		vacuum co 320 µm sili	ompatibility con sensor					





Svensson, O., Monaco, S., Popov, A., Nurizzo, D. and Bowler, M. W. (2015), The fully automatic characterization and data collection from crystals of biological macromolecules. *Acta Cryst.* **D71** 1757-1767





Complications

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







Accuracy

Main uncertainties of the observed intensities are determined by counting statistics

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

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



Diffraction resolution vs. absorbed dose

A light for Science

14

Wilson plot

15

0

Radiation Damage

MGy

10² MGy

Absorbed Dose

Site-specific damage

KĞy

- 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:

 $\hat{\mathbf{J}}(h,D) \propto \operatorname{scale}(D) \cdot \hat{\mathbf{J}}_{u}(h) \cdot \operatorname{Exp}(-B(D) \cdot h^{2}/2)$

A light for Science

19

Insulin data collection

Semi-empirical model for diffraction intensity vs reciprocal space coordinate

$$\hat{\mathbf{J}}(\mathbf{h}) = \frac{1}{s} \cdot \hat{\mathbf{J}}_{u}(h) \cdot \operatorname{Exp}(-\mathbf{h} \cdot \mathbf{B} \cdot \mathbf{h}^{T})$$

Semi-empirical model of variance vs integrated intensity

 $\sigma_{l}^{2}(J)=k_{o}+k_{1}J+k_{2}J^{2}$

Integration over the scanned reciprocal space using Wilson distribution

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

Radiation-damage model

Resolution-dependent intensity decay:

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

Expected Intensity Variation

The 70 kDa membrane protein FtsH from Aquifex aeolicus 1222, a = 137.9 Å, b = 162.1 Å, c = 170 Å

Intensity Anisotropy

A light for Science

1 of 8 🕂

European Synchrotron Radiation Facility

Data collection strategy accounting radiation damage

Main Wedge

Resolution limit is set by the radiation damage

Space Group Cell Mosaicity	::	P 4 141.55 141.55 671.01 90.00 90.00 90.00 0.17 degree	
	dxv	v Line Length: 10.2 rm Lattice: 673.1 A Lattice: 673.1 A South State Stat	

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

Res	olution	limit =	3.59 Angst	rom 1	rans	mission	= 40.6°	bistan	ce = 562.	2mm	
	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 2 3 4	53.00 63.00 93.00 108.00	0.10 0.10 0.10 0.50	0.169 0.156 0.100 0.437	100 300 150 140	No Yes No No	10.00 30.00 15.00 70.00	16.9 46.9 15.0 61.2	16.9 63.8 78.8 140.0	8.711 24.156 7.729 31.530	8.711 32.867 40.596 72.126	32.9 76.0 90.2 100.0
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)											

European Synchrotron Radiation Facility

29

ESRF

Multi-positional and Helical data collection A Light for Science

FAE crystals ID23-1

E=12.75Kev, I=35 mA, Aperture=0.03 mm Flux=1.5x10¹¹ Photon/sec

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

M	ulti-positions data	collection		
Resolution limit is so Resolution limit = 1.73 Angs	et by the radiation trom Transmission	damage = 100.0% D	istance = 244.	6mm
WEDGE PARAMETERS		INFORMA	TION	
sub- Phi Rot. Exposure We- start width /image dge degree degree s	N.of Over sWedge ima- -lap width ges degree	Exposure Exp /sWedge to s	osure Dose tal /sWedge s MGy	Dose Compl total tenes MGy %
Wedge number = 1 Cryst 1 0.00 0.25 1.338 Wedge number = 2 Cryst	al position = 1 80 No 20.00 al position = 2	107.0	107.0 4.067	4.067 51.
1 20.00 0.25 1.338 Wedge number 3 Cryst. 1 40.00 0.25 1.338 Wedge number 4 Cryst. 1 60.00 0.25 1.338	80 NO 20.00 al position = 3 80 No 20.00 al position = 4	107.0	107.0 4.067 107.0 4.067	4.067 85.
Phi_start - Phi_finish Total rotation range Total N.of images Overall Completeness Redundancy R-factor (outer shell) I/Sigma (outer shell) Total Exposure time Total Data Collection time	: 0.00 - 80.00 : 80.00 degree : 320 : 98.6% : 3.18 : 5.6% (36.8%) : 22.9 (3.3) : 428.1 sec (0.119 : 1228.1 sec (0.34)	hour) L hour)		4.007 88.

..... 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

Se scattering factor estimates:

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

Estimate of signal from MAD experiment

8 Se atoms per 300 protein residues and 0 nucleic acid residues

Induced Burn Strategy

Example results from "burning strategy"

Express estimation of radiation damage by DOZOR Method to determine Mesh_and_Collect Exposure time and resolution

A Light for Science

One cycle 10 images, total rotation 2°

20 times

Spot number vs.

Average intensity vs.

A light for Science

σ (I_{peak}) =SQRT (I_{peak} + I_{background})

Diffraction resolution vs. absorbed dose

Achievable resolution vs. crystal size

Lactase, B-factor=9.6 Å²

P222

Crystal Size, µm

X-ray Meshscan – a solid technique for sample analysis

Experimental setup for X-ray crystallography

The data are accumulated during translational movement

X-ray centering

Evaluating diffraction signal with DOZOR A Light for Science

European Synchrotron Radiation Facility

40

DOZOR - indication on the ice or salt diffraction contaminations

0.04

0.05

0.06

0.01

👚 3 of 4 🤍

0.02

17 Resoluti

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

Ulrich Zander et al., Acta Crystallographica (2015) 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.

DOZOR result

The result of mesh clustering

Ellipse fit

Meshclustering gui

Some test examples

European Synchrotron Radiation Facility

Diffraction sample Modeling

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

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

ght for Science

X-ray Mesh Scans

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

Acknowledgements

- Gleb Bourenkov
- Igor Melnikov
- ESRF MX Group
- Olof Svensson & EDNA developers team