Modelling EM experiments

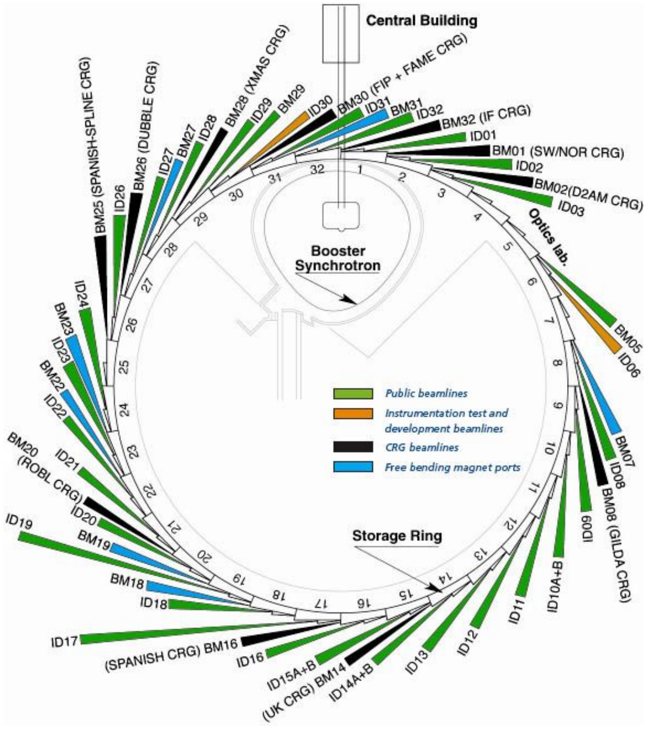
Alex de Maria
Software Group
Data Manager
European Synchrotron Radiation Facility

Outline

- **Context**
- **ISPyB**
  - Electron Microscopy for single particles
    - Data model
- **ICAT**
  - Electron Microscopy for tomography
    - Data model
- **User interfaces**
- **Conclusions**

Context

- **The ESRF: European Synchrotron Radiation Facility**

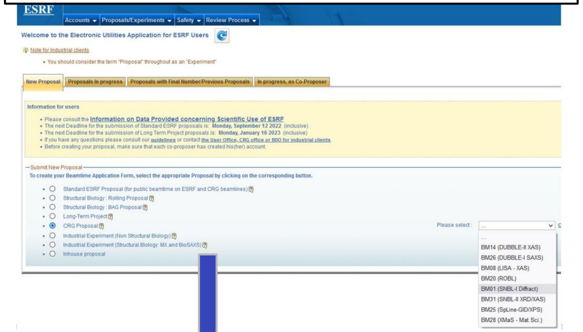


- Fourth generation high energy synchrotron
- **~40 beamlines**
  - **~10 for structural biology**
- **1 Titan Krios (CM01) and soon CM02**

- Proposal system
- Public and private beamtime
- Mail-in service and remote access

Data Management

1) Proposal and Sample Declaration



Welcome to the Electronic Utilities Application for ESRF Users

Note for Industrial clients

- You should consider the term "Proposal" throughout as an "Experiment"

Information for users

- Please consult the Information on Data Provided concerning Scientific Use of ESRF
- The next Deadline for the submission of Standard ESRF proposals is: **Monday, September 12 2022** (inclusive)
- The next Deadline for the submission of Long Term Project proposals is: **Monday, January 16 2023** (inclusive)
- If you have any questions please consult our guidelines or contact the User Office, CRG office or BDO for industrial clients.
- Before creating your proposal, make sure that each co-proposer has created his(her) account.

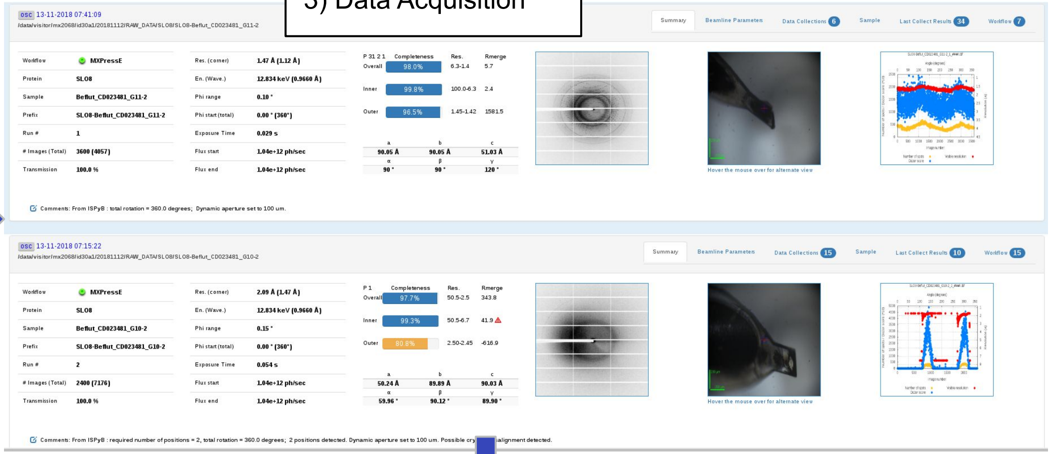
Submit New Proposal

To create your Beamtime Application Form, select the appropriate Proposal by clicking on the corresponding button.

- ○ Standard ESRF Proposal (for public beamtime on ESRF and CRG beamlines)
- ○ Structural Biology : Rolling Proposal
- ○ Structural Biology : BAG Proposal
- ○ Long-Term Project
- ● CRG Proposal
- ○ Industrial Experiment (Non Structural Biology)
- ○ Industrial Experiment (Structural Biology: MX and BioSAXS)
- ○ Inhouse proposal

Please select : ...

3) Data Acquisition



osc 13-11-2018 07:41:09
/data/visitor/mx2068/id30a1/20181112/RAW_DATA/SLO8/SLO8-Beflut_CD023481_G11-2

Workflow	MXPressE	Res. (corner)	1.47 Å (1.12 Å)
Protein	SLO8	En. (Wave.)	12.834 keV (0.9660 Å)
Sample	Beflut_CD023481_G11-2	Phi range	0.10 °
Prefix	SLO8-Beflut_CD023481_G11-2	Phi start (total)	0.00 ° (360°)
Run #	1	Exposure Time	0.029 s
# Images (Total)	3600 (4057)	Flux start	1.04e+12 ph/sec
Transmission	100.0 %	Flux end	1.04e+12 ph/sec

P 31 2 1	Completeness	Res.	Rmerge
Overall	98.0%	6.3-1.4	5.7
Inner	99.8%	100.0-6.3	2.4
Outer	96.5%	1.45-1.42	1581.5

a	b	c
90.05 Å	90.05 Å	51.03 Å
α	β	γ
90 °	90 °	120 °

Hover the mouse over for alternate view

Comments: From ISPyB : total rotation = 360.0 degrees; Dynamic aperture set to 100 um.

osc 13-11-2018 07:15:22
/data/visitor/mx2068/id30a1/20181112/RAW_DATA/SLO8/SLO8-Beflut_CD023481_G10-2

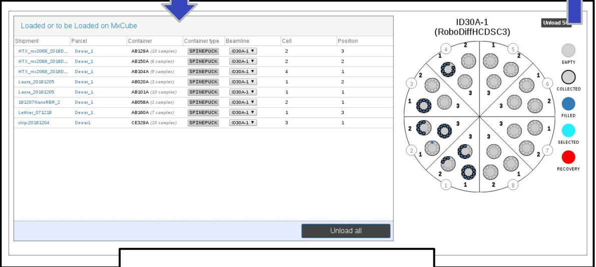
Workflow	MXPressE	Res. (corner)	2.09 Å (1.47 Å)
Protein	SLO8	En. (Wave.)	12.834 keV (0.9660 Å)
Sample	Beflut_CD023481_G10-2	Phi range	0.15 °
Prefix	SLO8-Beflut_CD023481_G10-2	Phi start (total)	0.00 ° (360°)
Run #	2	Exposure Time	0.054 s
# Images (Total)	2400 (7176)	Flux start	1.04e+12 ph/sec
Transmission	100.0 %	Flux end	1.04e+12 ph/sec

P 1	Completeness	Res.	Rmerge
Overall	97.7%	50.5-2.5	343.8
Inner	99.3%	50.5-6.7	41.9 ⚠
Outer	80.8%	2.50-2.45	-616.9

a	b	c
50.24 Å	89.89 Å	90.03 Å
α	β	γ
59.96 °	90.12 °	89.90 °

Hover the mouse over for alternate view

Comments: From ISPyB : required number of positions = 2, total rotation = 360.0 degrees; 2 positions detected. Dynamic aperture set to 100 um. Possible cry alignment detected.



Loaded or to be Loaded on MxCube

Shipment	Parcel	Container	Container type	Beamline	Cell	Position
HTX_mx2068_2018D...	Dewar_1	AB129A (10 samples)	SPINEPUCK	ID30A-1 ▾	2	3
HTX_mx2068_2018D...	Dewar_1	AB150A (6 samples)	SPINEPUCK	ID30A-1 ▾	2	2
HTX_mx2068_2018D...	Dewar_1	AB104A (9 samples)	SPINEPUCK	ID30A-1 ▾	4	1
Laura_20181205	Dewar_1	AB020A (3 samples)	SPINEPUCK	ID30A-1 ▾	1	2
Laura_20181205	Dewar_1	AB101A (10 samples)	SPINEPUCK	ID30A-1 ▾	1	1
181207NanoRBR_2	Dewar_1	AB058A (1 samples)	SPINEPUCK	ID30A-1 ▾	2	1
Lethier_071218	Dewar_1	AB160A (7 samples)	SPINEPUCK	ID30A-1 ▾	1	3
ship-20181204	Dewar1	CE329A (10 samples)	SPINEPUCK	ID30A-1 ▾	3	1

2) Sample Tracking

5) Data Publication

Collection plan strategy: resolution limit is set by the initial image resolution

Wedge	Subwedge	Start (°)	Width (°)	No images	Exp time (s)	Max res (Å)
1	1	0.00	0.05	7200	0.020	1.40

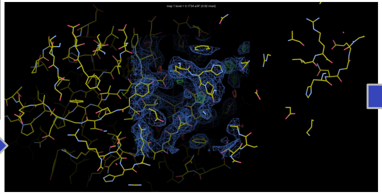
Diffraction Plan

Forced space group	Anomalous data	Aimed multiplicity	Aimed completeness	Aimed I/sigma at highest res.	Aimed resolution (Å)	Min osc. width
None	False	Default (optimized)	Default (>= 0.99)	1.00	Default (highest possible)	0.05

Image quality indicators

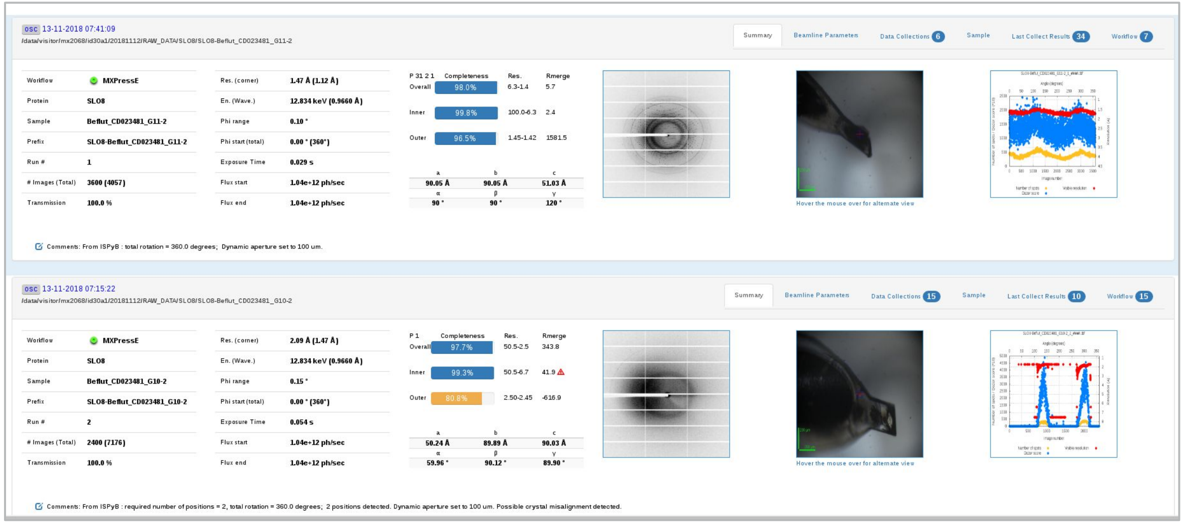
File	Dozor score (1)	Dozor visible res. [Å]	Tot integr signal (2)	Spot total	In-Res Total	Good Bragg	Ice Rings	Meth 1 Res	Meth 2 Res	Max unit cell
ref-SLO8-Beflut_CD023481_H10-2_4_0001.cbf	311.0	1.6	1265408	340	293	274	4	2.11	2.41	109.3
ref-SLO8-Beflut_CD023481_H10-2_4_0002.cbf	707.3	1.7	1369086	365	327	295	5	2.12	2.40	114.8
ref-SLO8-Beflut_CD023481_H10-2_4_0003.cbf	292.8	1.6	944824	272	238	221	3	2.12	2.45	109.5
ref-SLO8-Beflut_CD023481_H10-2_4_0004.cbf	435.8	1.6	1	336	297	265	5	2.13	2.43	112.9

4) Data analysis



Outline

- **ISPyB**
  - Information System for Protein crYstallographic Beamlines
  - Tailored-made LIMS solution for MX/SB
  - Born at the ESRF ~2004
  - Collaboration
    - Soleil
    - ALBA
    - EMBL
    - Desy
    - Elettra
    - MaxIV
    - Diamond Light Source



OSC 13-11-2018 07:41:09
/data/visitor/mx2068/id30a1/20181112/RAW_DATA/SLO8/SLO8-Beflut_CD023481_G11-2

Workflow	MXPressE	Res. (corner)	1.47 Å (1.12 Å)
Protein	SLO8	En. (Wave.)	12.834 keV (0.9660 Å)
Sample	Beflut_CD023481_G11-2	Phi range	0.10 °
Prefix	SLO8-Beflut_CD023481_G11-2	Phi start (total)	0.00 ° (360°)
Run #	1	Exposure Time	0.029 s
# Images (Total)	3600 (4057)	Flux start	1.04e+12 ph/sec
Transmission	100.0 %	Flux end	1.04e+12 ph/sec

P 31 2 1	Completeness	Res.	Rmerge
Overall	98.0%	6.3-1.4	5.7
Inner	99.8%	100.0-6.3	2.4
Outer	96.5%	1.45-1.42	1581.5

a	b	c
90.05 Å	90.05 Å	51.03 Å
α	β	γ
90 °	90 °	120 °

Comments: From ISPyB : total rotation = 360.0 degrees; Dynamic aperture set to 100 um.

OSC 13-11-2018 07:15:22
/data/visitor/mx2068/id30a1/20181112/RAW_DATA/SLO8/SLO8-Beflut_CD023481_G10-2

Workflow	MXPressE	Res. (corner)	2.09 Å (1.47 Å)
Protein	SLO8	En. (Wave.)	12.834 keV (0.9660 Å)
Sample	Beflut_CD023481_G10-2	Phi range	0.15 °
Prefix	SLO8-Beflut_CD023481_G10-2	Phi start (total)	0.00 ° (360°)
Run #	2	Exposure Time	0.054 s
# Images (Total)	2400 (7176)	Flux start	1.04e+12 ph/sec
Transmission	100.0 %	Flux end	1.04e+12 ph/sec

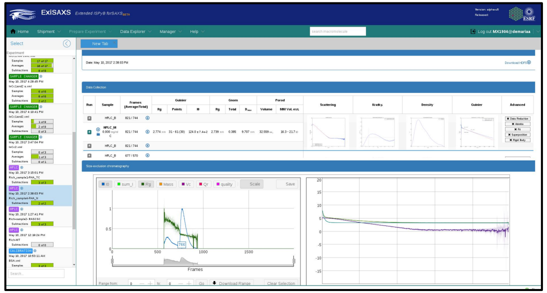
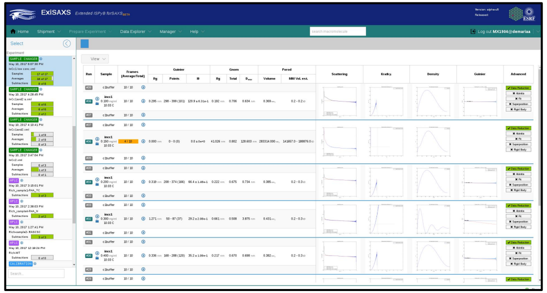
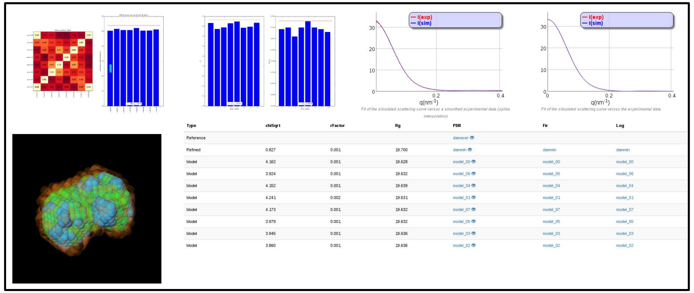
P 1	Completeness	Res.	Rmerge
Overall	97.7%	50.5-2.5	343.8
Inner	99.3%	50.5-6.7	41.9 ⚠
Outer	80.8%	2.50-2.45	-616.9

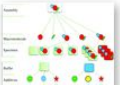
a	b	c
50.24 Å	89.89 Å	90.03 Å
α	β	γ
59.96 °	90.12 °	89.90 °

Comments: From ISPyB : required number of positions = 2, total rotation = 360.0 degrees; 2 positions detected. Dynamic aperture set to 100 um. Possible crystal misalignment detected.

ISPyB view of a MX experiment

ISPyB extension for BioSAXS (2014)



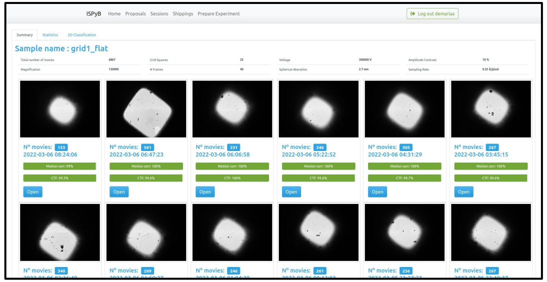
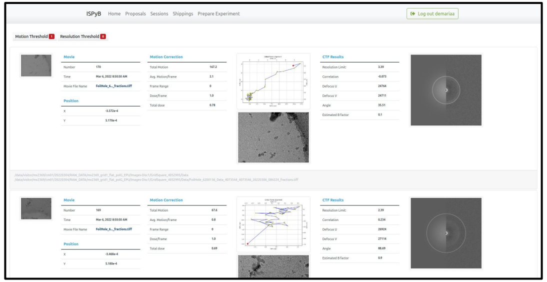
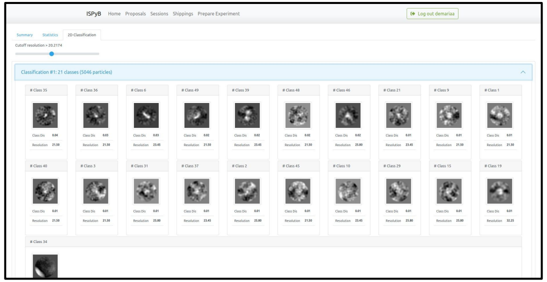
ISPyB for BioSAXS, the gateway to user autonomy in solution scattering experiments

A. De Maria Antolinos, P. Pernot, M. E. Brennich, J. Kieffer, M. W. Bowler, S. Delageniere, S. Ohlsso
and A. Round

Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution

P. Pernot, A. Round, R. Barrett, A. De Maria Antolinos, A. Gobbo, E. Gordon, J. Huet, J. Kieffer,
Surr, P. Theveneau, L. Zerrad and S. McSweeney

ISPyB extension for EM (2018)

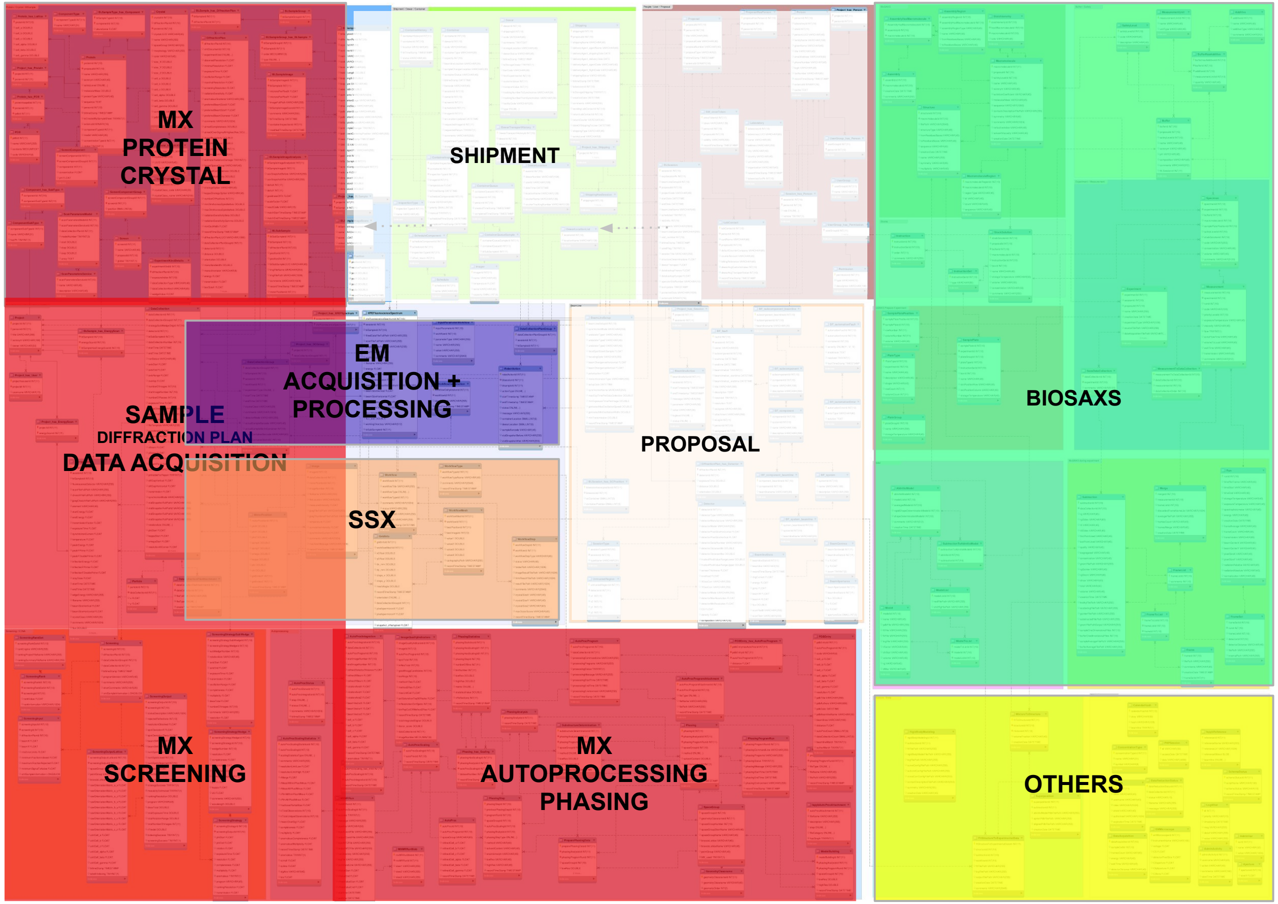
CM01: a facility for cryo-electron microscopy at the European Synchrotron

E. Kandiah, T. Giraud, A. de Maria Antolinos, F. Dobias, G. Effantin, D. Flot, M. Hons, G. Schoehn, J. Susini, O. Svensson, G. A. Leonard and C. Mueller-Dieckman

ISPyB Data Model

210 tables
40 Views

ISPyB Data Model



- 🟥 CRYSTALLOGRAPHY
- 🟩 BIOSAXS
- 🟪 EM
- 🟧 SSX

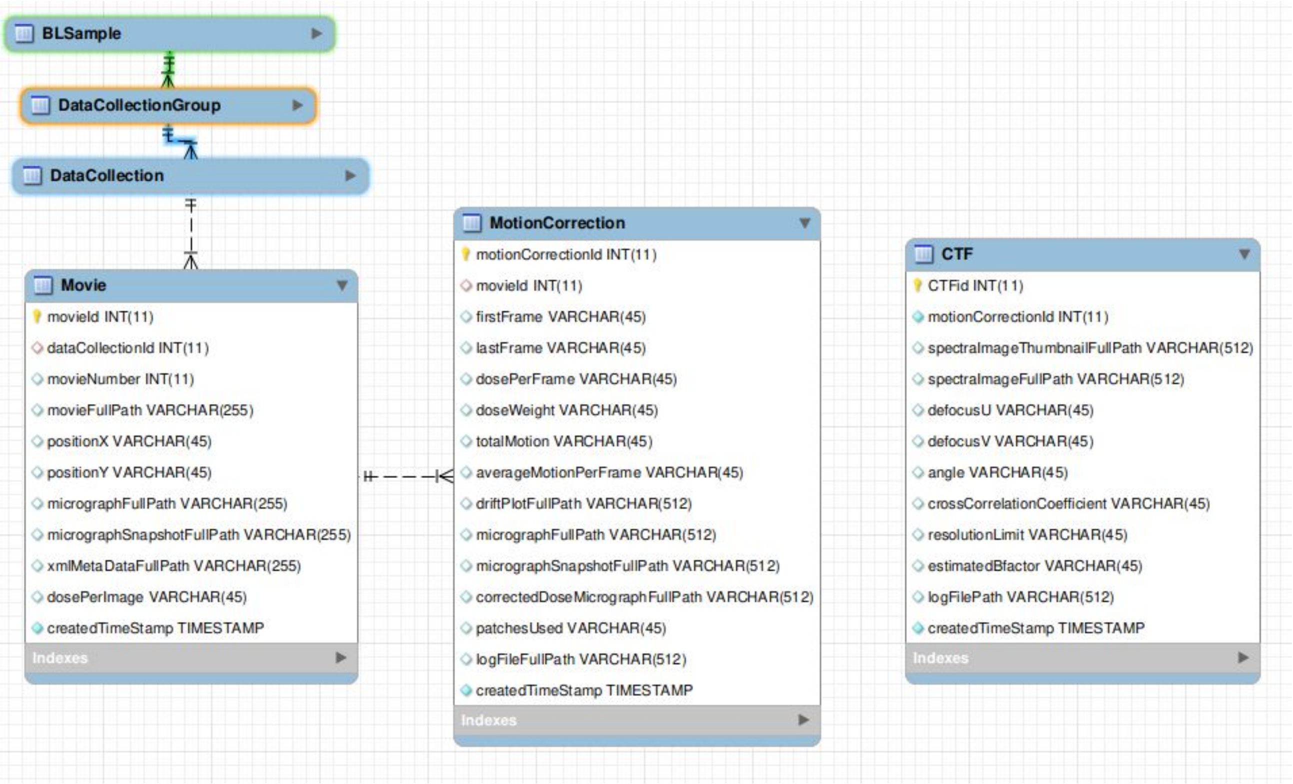
- ☐ COMMON TABLES
- 🟨 OTHERS

210 tables
40 Views

EM Tables

- 2018
  - Extension of MX tables
  - Collaboration
    - Diamond Light Source
  - Public discussions

https://github.com/ispyb/ispyb-database-modeling/issues/14



BLSample

DataCollectionGroup

DataCollection

Movie
- movieId INT(11)
- dataCollectionId INT(11)
- movieNumber INT(11)
- movieFullPath VARCHAR(255)
- positionX VARCHAR(45)
- positionY VARCHAR(45)
- micrographFullPath VARCHAR(255)
- micrographSnapshotFullPath VARCHAR(255)
- xmlMetaDataFullPath VARCHAR(255)
- dosePerImage VARCHAR(45)
- createdTimeStamp TIMESTAMP

Indexes

MotionCorrection
- motionCorrectionId INT(11)
- movieId INT(11)
- firstFrame VARCHAR(45)
- lastFrame VARCHAR(45)
- dosePerFrame VARCHAR(45)
- doseWeight VARCHAR(45)
- totalMotion VARCHAR(45)
- averageMotionPerFrame VARCHAR(45)
- driftPlotFullPath VARCHAR(512)
- micrographFullPath VARCHAR(512)
- micrographSnapshotFullPath VARCHAR(512)
- correctedDoseMicrographFullPath VARCHAR(512)
- patchesUsed VARCHAR(45)
- logFileFullPath VARCHAR(512)
- createdTimeStamp TIMESTAMP

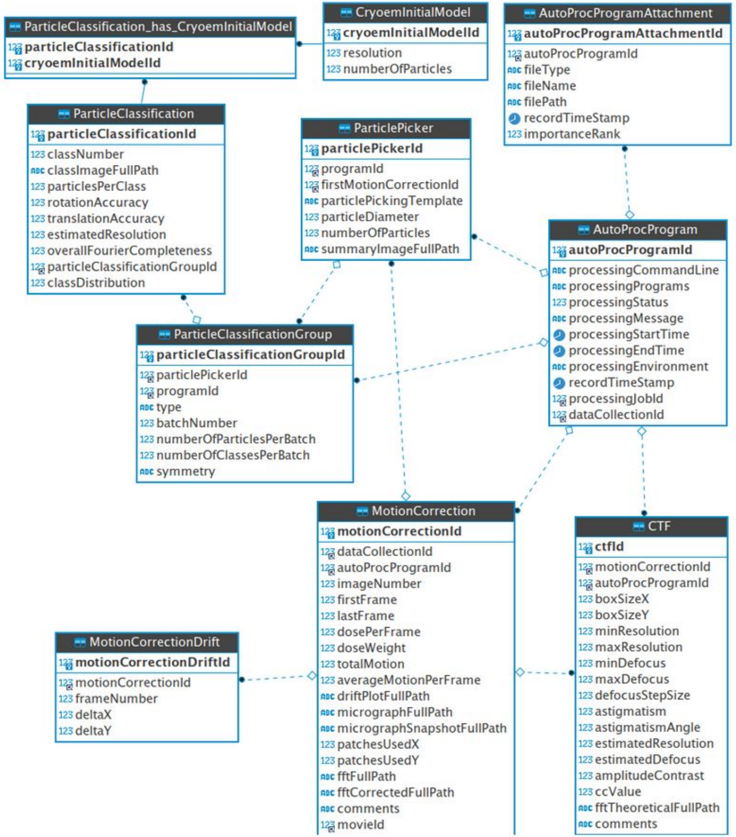
Indexes

CTF
- CTFid INT(11)
- motionCorrectionId INT(11)
- spectralmageThumbnailFullPath VARCHAR(512)
- spectralmageFullPath VARCHAR(512)
- defocusU VARCHAR(45)
- defocusV VARCHAR(45)
- angle VARCHAR(45)
- crossCorrelationCoefficient VARCHAR(45)
- resolutionLimit VARCHAR(45)
- estimatedBfactor VARCHAR(45)
- logFilePath VARCHAR(512)
- createdTimeStamp TIMESTAMP

Indexes

EM Tables

- 2018
  - Extension of MX tables
  - Collaboration
    - Diamond Light Source
  - Public discussions
    https://github.com/ispyb/ispyb-database-modeling/issues/14
- 2020
  - Particle picking included
    https://github.com/ispyb/ispyb-database-modeling/issues/64



ParticleClassification_has_CryoemInitialModel
particleClassificationId
cryoemInitialModelId

CryoemInitialModel
cryoemInitialModelId
resolution
numberOfParticles

AutoProcProgramAttachment
autoProcProgramAttachmentId
autoProcProgramId
fileType
fileName
filePath
recordTimeStamp
importanceRank

ParticleClassification
particleClassificationId
classNumber
classImageFullPath
particlesPerClass
rotationAccuracy
translationAccuracy
estimatedResolution
overallFourierCompleteness
particleClassificationGroupId
classDistribution

ParticlePicker
particlePickerId
programId
firstMotionCorrectionId
particlePickingTemplate
particleDiameter
numberOfParticles
summaryImageFullPath

AutoProcProgram
autoProcProgramId
processingCommandLine
processingPrograms
processingStatus
processingMessage
processingStartTime
processingEndTime
processingEnvironment
recordTimeStamp
processingJobId
dataCollectionId

ParticleClassificationGroup
particleClassificationGroupId
particlePickerId
programId
type
batchNumber
numberOfParticlesPerBatch
numberOfClassesPerBatch
symmetry

MotionCorrection
motionCorrectionId
dataCollectionId
autoProcProgramId
imageNumber
firstFrame
lastFrame
dosePerFrame
doseWeight
totalMotion
averageMotionPerFrame
driftPlotFullPath
micrographFullPath
micrographSnapshotFullPath
patchesUsedX
patchesUsedY
fftFullPath
fftCorrectedFullPath
comments
movieId

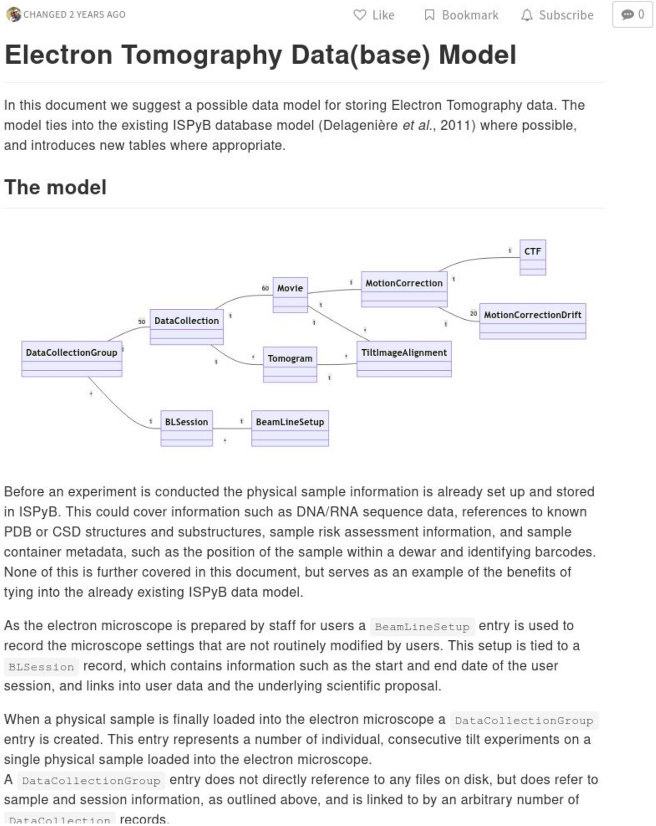
CTF
ctfId
motionCorrectionId
autoProcProgramId
boxSizeX
boxSizeY
minResolution
maxResolution
minDefocus
maxDefocus
defocusStepSize
astigmatism
astigmatismAngle
estimatedResolution
estimatedDefocus
amplitudeContrast
ccValue
fftTheoreticalFullPath
comments

MotionCorrectionDrift
motionCorrectionDriftId
motionCorrectionId
frameNumber
deltaX
deltaY

Credits: Karl Levik (DLS)

EM Tables

- 2018
  - Extension of MX tables
  - Collaboration
    - Diamond Light Source
  - Public discussions
    https://github.com/ispyb/ispyb-database-modeling/issues/14
- 2020
  - Particle picking included
    https://github.com/ispyb/ispyb-database-modeling/issues/64
- 2022
  - Cryo-ET and electron diffraction
    https://github.com/ispyb/ispyb-database-modeling/issues/73
    https://hackmd.io/RCfRD-FaTCagTxeTwVHIaA?view



Electron Tomography Data(base) Model

In this document we suggest a possible data model for storing Electron Tomography data. The model ties into the existing ISPyB database model (Delagenière *et al.*, 2011) where possible, and introduces new tables where appropriate.

The model

Before an experiment is conducted the physical sample information is already set up and stored in ISPyB. This could cover information such as DNA/RNA sequence data, references to known PDB or CSD structures and substructures, sample risk assessment information, and sample container metadata, such as the position of the sample within a dewar and identifying barcodes. None of this is further covered in this document, but serves as an example of the benefits of tying into the already existing ISPyB data model.

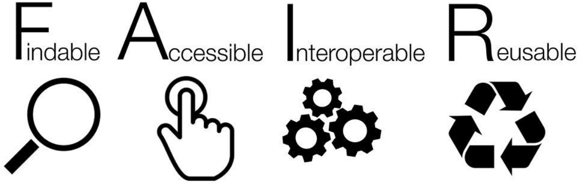
As the electron microscope is prepared by staff for users a `BeamLineSetup` entry is used to record the microscope settings that are not routinely modified by users. This setup is tied to a `BLSession` record, which contains information such as the start and end date of the user session, and links into user data and the underlying scientific proposal.

When a physical sample is finally loaded into the electron microscope a `DataCollectionGroup` entry is created. This entry represents a number of individual, consecutive tilt experiments on a single physical sample loaded into the electron microscope.
A `DataCollectionGroup` entry does not directly reference to any files on disk, but does refer to sample and session information, as outlined above, and is linked to by an arbitrary number of `DataCollection` records.

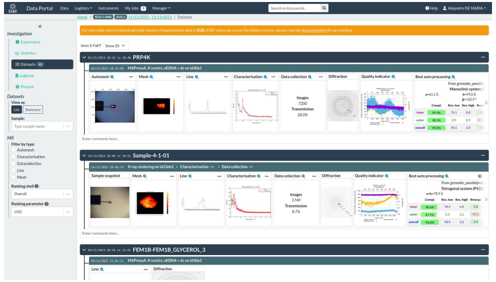
Credits: Markus Gerstel (DLS)

Data Policy (2015) and updated on 2024

- For **all peer-reviewed and in-house** experiments
  - Capture of raw and processed data and metadata
  - Preserve data for at least 10 years and metadata forever
  - **3 years embargo period** where data are only accessible to the experiment team
  - > 3 years data become **public and accessible** for everyone
  - Implementation of the FAIR principles

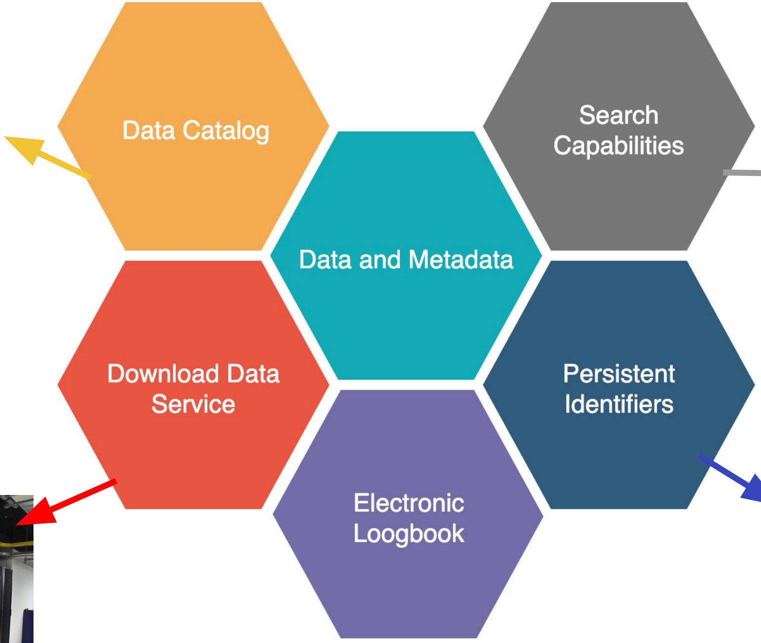


Data Services



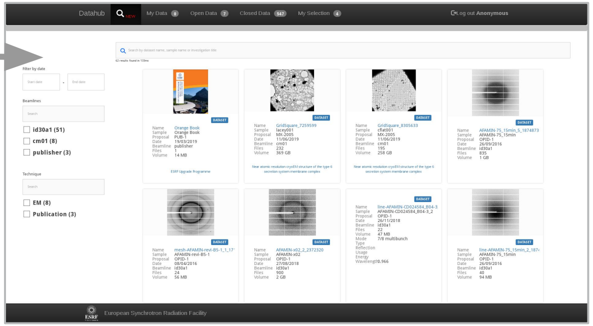
Data Catalog

https://data.esrf.fr



Search Service

Search engine for Big Data



DOI > 10.15151/ESRF-DC-186715732

Data collection

Dataset Open access

THE ORANGE BOOK: ESRF UPGRADE PROGRAMME PHASE II (2015 - 2022)

D. Andrault ; R. Barrett ; S. Bayat ; P. Berkvens ; J.-C. Biasci ; S. Billinge ; B. Boulanger ; J.-F. Bouteille ; A. Bravin ; E. Brun ; F. Carla ; N. Carmignani ; H. Castillo-Michel ; J.-M. Chaize ; J. Chavanne ; F. Cianciosi ; P. Cloetens ; P. Coan ; S. Corr ; M. Cotte ; D. Coulon ; Y. Dabin ; I. Daniel ; D. De Sanctis ; M. DiMichiel ; R. Dimper ; J. Drnec ; L. Dubrovinski ; L. Ducotté ; F. Ewald ; P. Fajardo ; L. Farvaque ; R. Felici ; C. Ferrero ; G. Fiquet ; A. Götz ; L. Goirand ; Y. Gouez ; M. Hahn ; J. Härtwig ; P.F. Henry ; O. Hignette ; J. Hudspeth ; S.D.M. Jacques ; J. Jacob ; N. Janvier ; J. Kieffer ; S.A.J. Kimber ; C. Knabe-Ducheyne ; M. Kocsis ; M. Krisch ; B. Lebayle ; G. Le Bec ; G. Le Duc ; G. Leonard ; S. Liuzzo ; P. Loubeyre ; P. Mackrill ; T. Marchial ; T. Mairs ; P. Marion ; D. Martin ; T. Martin ; M. McMahon ; J. Meyer ; M. Mezouar ; E. Mitchell ; C. Morawe ; J. Morse ; B. Müller ; C. Müller-Dieckmann ; T. Narayanan ; B. Nash ; B. Ogier ; S. Pascarelli ; H. Pedroso Marques ; P. Pernot ; T. Perron ; F. Pfeiffer ; E. Plouviez ; C. Ponchut ; A. Popov ; A. Poulin ; A. Rack ; P. Raimondi ; J.-L. Revol ; H. Reichert ; A. Royant ; L. Sánchez Ortiz ; M. Sanchez Del Rio ; M. Santoro ; M. Scandella ; K. Scheidt ; T. Schülli ; V. Serriere ; F. Sette ; A. Snigirev ; V.A. Sole ; J. Susini ; O. Svensson ; P. Taffoureau ; G. Vaughan ; F. Villar ; B. Winkler ; J. Wright ; L. Zhang.

DOI
DOI 10.15151/ESRF-DC-186715732
Licence (for files)

Abstract

The strength of modern X-ray science is its endeavour to provide unique tools and methods to understand and picture how atoms make and evolve the living matter and the materials that both surround us and of which we are constituted. Understanding materials for electronic devices and for storing and producing new forms of renewable and non-polluting energy, and unravelling the mysteries of biological reactions are examples of key challanges for which we need to improve our capacity to design future materials, processes, drugs, etc. These ambitious goals of 21st century science require fundamental understanding and visualisation of the hierarchical static and dynamic organisation of complex and functionalised matter with full continuity from macroscopic objects down to interactions among single atom pairs. Future X-ray science at modern synchrotrons and XFELs will bridge the gap between visible light and electron microscopy, i.e. it will give full access to study changes within heterogeneous atomic assemblies from ~1012 atoms down to only a few atoms; all of this with spatial and temporal resolutions that will be approaching a few atoms in space (nanometre) and inter-atomic motion in time (femtosecond) respectively. Such incredible resolution and range in space and time are made possible by the absolutely impressive and never-ending improvements of brightness, spectral range and degree of spatial and temporal coherence and resolution of accelerator-based X-ray sources. During the last ~50 years peak-brightness has increased up to 22 orders of magnitude when compared to conventional laboratory sources.

Proposals

PUB-1

Experimental report

There is currently no experimental report.

Experimental data

The data can be accessed by clicking on the link below

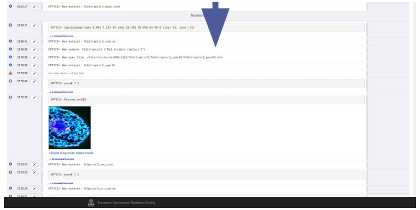
Access data

Citation

Below is the recommended format for citing this work in a research publication.

Data Service

Explore and Download data



Electronic Logbook

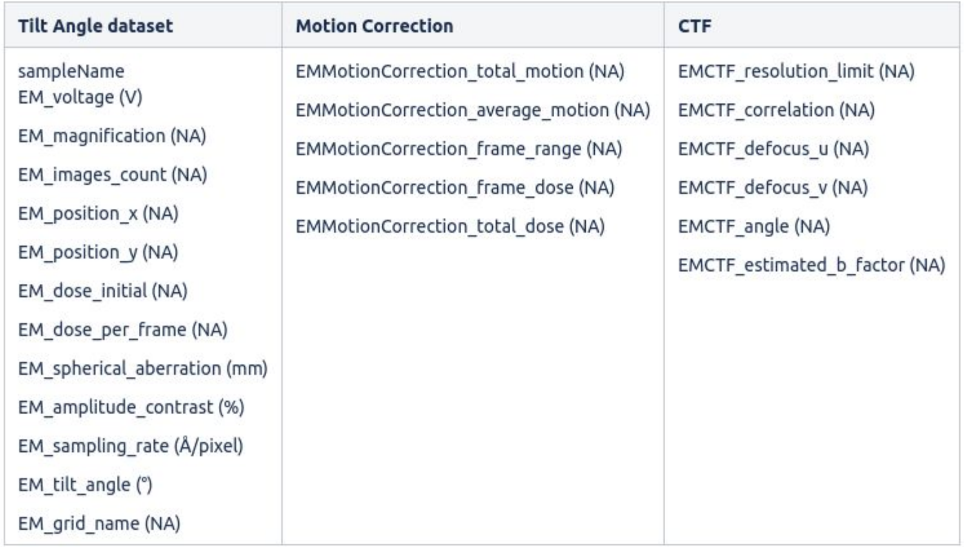
Annotate your experiments

Persistent Identifiers

Make your data findable and searchable

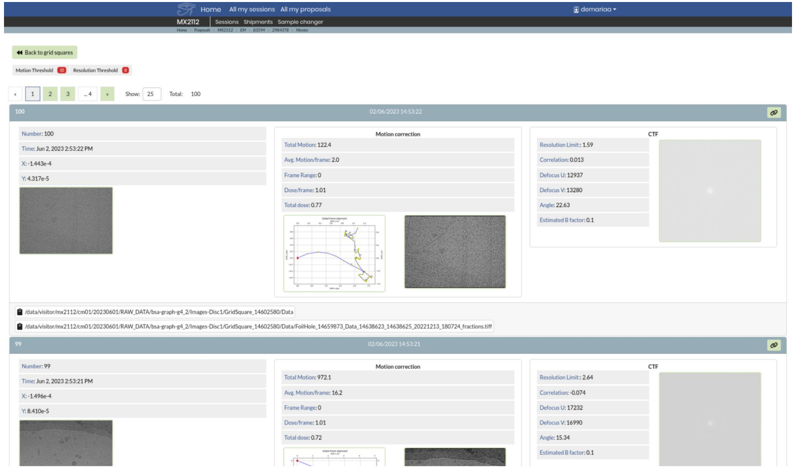
Electron Tomography in ICAT

- Evaluation of using ICAT for CryoET
  - Based on the work done on ISPyB
  - Relying on Nexus format (when possible)
  - No tables but datasets
  - No columns but metadata parameters



Tilt Angle dataset	**Motion Correction**	**CTF**
sampleName EM_voltage (V)	EMMotionCorrection_total_motion (NA)	EMCTF_resolution_limit (NA)
EM_magnification (NA)	EMMotionCorrection_average_motion (NA)	EMCTF_correlation (NA)
EM_images_count (NA)	EMMotionCorrection_frame_range (NA)	EMCTF_defocus_u (NA)
EM_position_x (NA)	EMMotionCorrection_frame_dose (NA)	EMCTF_defocus_v (NA)
EM_position_y (NA)	EMMotionCorrection_total_dose (NA)	EMCTF_angle (NA)
EM_dose_initial (NA)		EMCTF_estimated_b_factor (NA)
EM_dose_per_frame (NA)		
EM_spherical_aberration (mm)		
EM_amplitude_contrast (%)		
EM_sampling_rate (Å/pixel)		
EM_tilt_angle (°)		
EM_grid_name (NA)		

UI: ISPyB vs Data portal



100 02/06/2023 14:53:22

Number: 100
Time: Jun 2, 2023 2:53:22 PM
X: -1.443e-4
Y: 4.317e-5

Motion correction

Total Motion: 122.4
Avg. Motion/frame: 2.0
Frame Range: 0
Dose/frame: 1.01
Total dose: 0.77

CTF

Resolution Limit:: 1.59
Correlation: 0.013
Defocus U: 12937
Defocus V: 13280
Angle: 22.63
Estimated B factor: 0.1

/data/visitor/mx2112/cm01/20230601/RAW_DATA/bsa-graph-g4_2/Images-Disc1/GridSquare_14602580/Data
/data/visitor/mx2112/cm01/20230601/RAW_DATA/bsa-graph-g4_2/Images-Disc1/GridSquare_14602580/Data/FoilHole_14659873_Data_14638623_14638625_20221213_180724_fractions.tiff

99 02/06/2023 14:53:21

Number: 99
Time: Jun 2, 2023 2:53:21 PM
X: -1.496e-4
Y: 8.410e-5

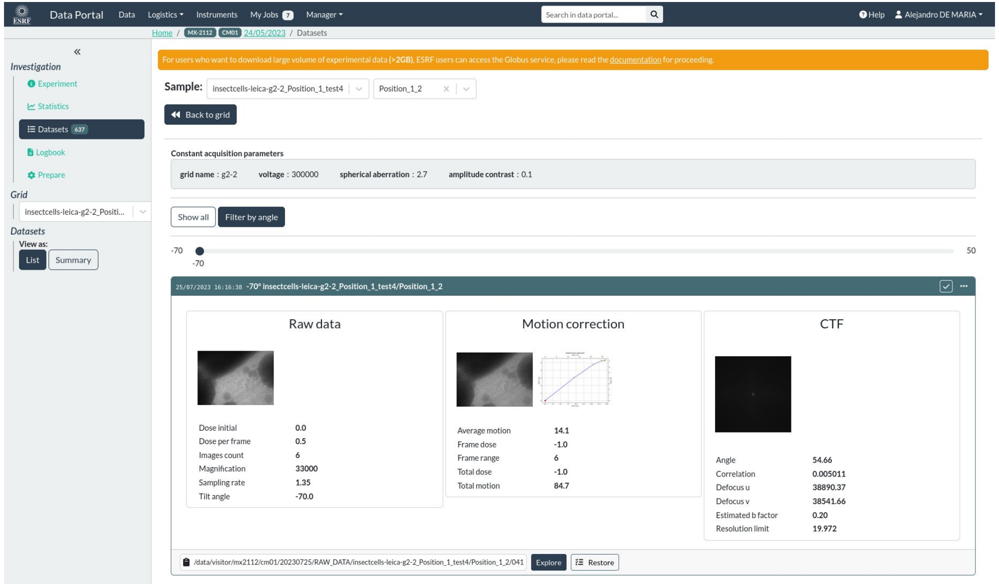
Motion correction

Total Motion: 972.1
Avg. Motion/frame: 16.2
Frame Range: 0
Dose/frame: 1.01
Total dose: 0.72

CTF

Resolution Limit:: 2.64
Correlation: -0.074
Defocus U: 17232
Defocus V: 16990
Angle: 15.34
Estimated B factor: 0.1

ISPyB



For users who want to download large volume of experimental data (>2GB), ESRF users can access the Globus service, please read the documentation for proceeding.

Sample: insectcells-leica-g2-2_Position_1_test4 Position_1_2

Grid: insectcells-leica-g2-2_Positi...

Constant acquisition parameters

grid name : g2-2 voltage : 300000 spherical aberration : 2.7 amplitude contrast : 0.1

-70 -70 50

25/07/2023 16:16:38 -70° insectcells-leica-g2-2_Position_1_test4/Position_1_2

Raw data

Dose initial	0.0
Dose per frame	0.5
Images count	6
Magnification	33000
Sampling rate	1.35
Tilt angle	-70.0

Motion correction

Average motion	14.1
Frame dose	-1.0
Frame range	6
Total dose	-1.0
Total motion	84.7

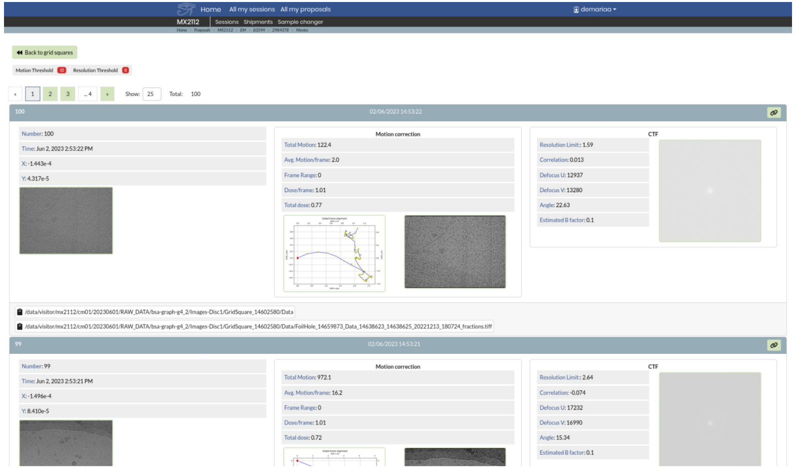
CTF

Angle	54.66
Correlation	0.005011
Defocus u	38890.37
Defocus v	38541.66
Estimated b factor	0.20
Resolution limit	19.972

/data/visitor/mx2112/cm01/20230725/RAW_DATA/insectcells-leica-g2-2_Position_1_test4/Position_1_2/041

ICAT

UI: ISPyB vs Data portal



100 02/06/2023 14:53:22

Number: 100
Time: Jun 2, 2023 2:53:22 PM
X: -1.443e-4
Y: 4.317e-5

Motion correction
Total Motion: 122.4
Avg. Motion/frame: 2.0
Frame Range: 0
Dose/frame: 1.01
Total dose: 0.77

CTF
Resolution Limit:: 1.59
Correlation: 0.013
Defocus U: 12937
Defocus V: 13280
Angle: 22.63
Estimated B factor: 0.1

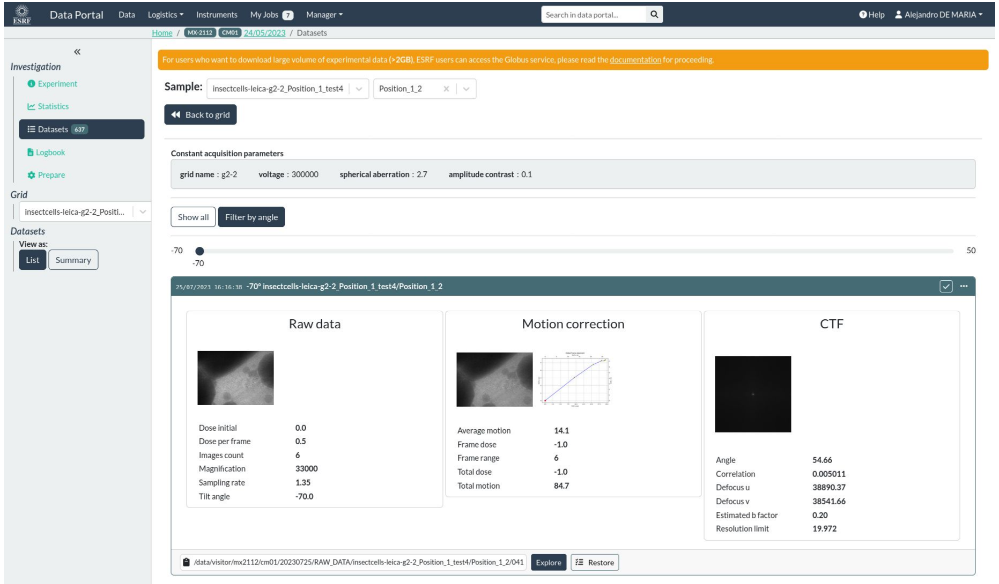
/data/visitor/mx2112/cm01/20230601/RAW_DATA/bsa-graph-g4_2/Images-Disc1/GridSquare_14602580/Data
/data/visitor/mx2112/cm01/20230601/RAW_DATA/bsa-graph-g4_2/Images-Disc1/GridSquare_14602580/Data/FoilHole_14659873_Data_14638623_14638625_20221213_180724_fractions.tiff

99 02/06/2023 14:53:21

Number: 99
Time: Jun 2, 2023 2:53:21 PM
X: -1.496e-4
Y: 8.410e-5

Motion correction
Total Motion: 972.1
Avg. Motion/frame: 16.2
Frame Range: 0
Dose/frame: 1.01
Total dose: 0.72

CTF
Resolution Limit:: 2.64
Correlation: -0.074
Defocus U: 17232
Defocus V: 16990
Angle: 15.34
Estimated B factor: 0.1



For users who want to download large volume of experimental data (>2GB), ESRF users can access the Globus service, please read the documentation for proceeding.

Sample: insectcells-leica-g2-2_Position_1_test4 Position_1_2

Constant acquisition parameters

grid name : g2-2 voltage : 300000 spherical aberration : 2.7 amplitude contrast : 0.1

25/07/2023 16:16:38 -70° insectcells-leica-g2-2_Position_1_test4/Position_1_2

Raw data

Dose initial	0.0
Dose per frame	0.5
Images count	6
Magnification	33000
Sampling rate	1.35
Tilt angle	-70.0

Motion correction

Average motion	14.1
Frame dose	-1.0
Frame range	6
Total dose	-1.0
Total motion	84.7

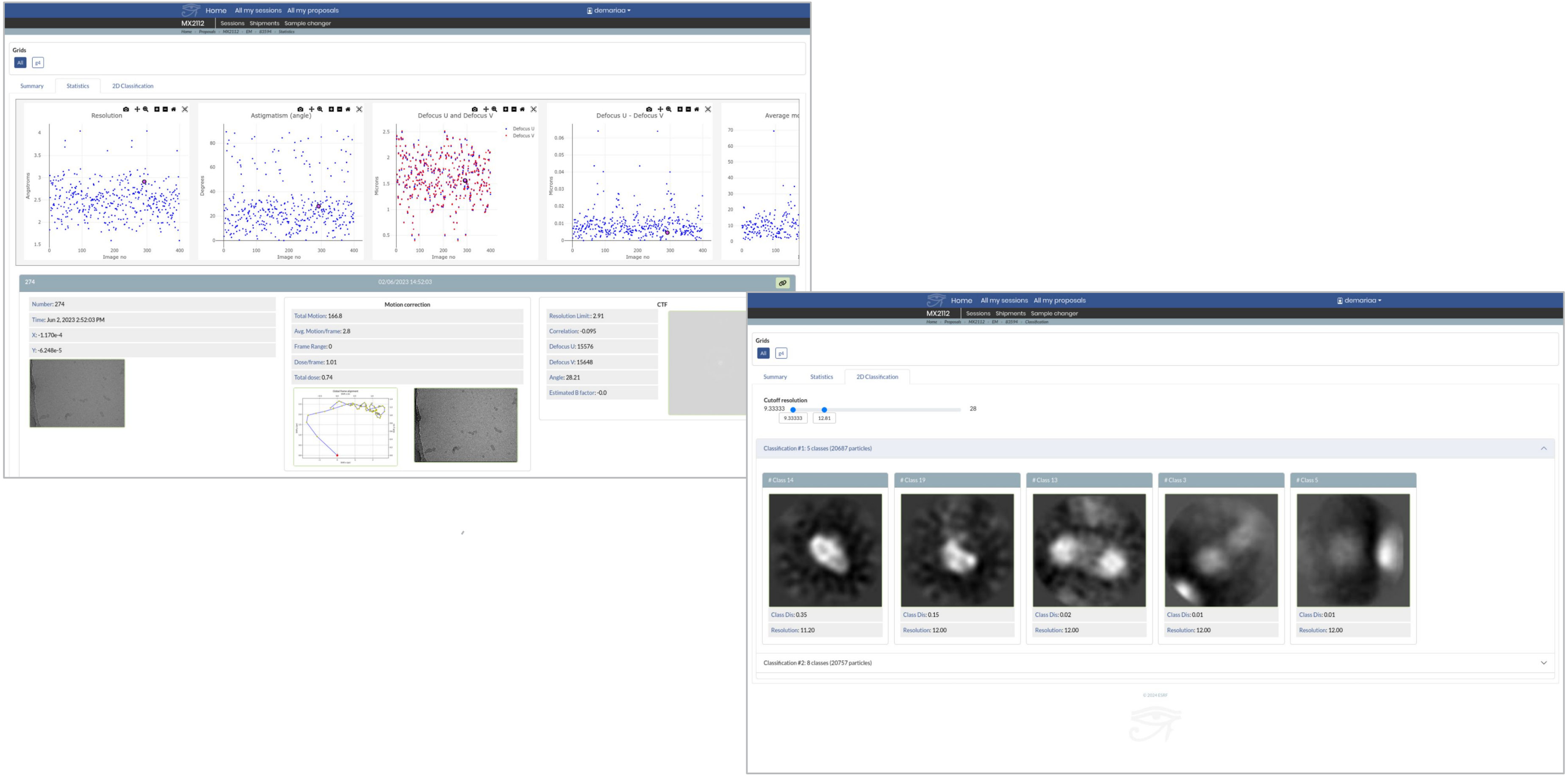
CTF

Angle	54.66
Correlation	0.005011
Defocus u	38890.37
Defocus v	38541.66
Estimated b factor	0.20
Resolution limit	19.972

/data/visitor/mx2112/cm01/20230725/RAW_DATA/insectcells-leica-g2-2_Position_1_test4/Position_1_2/041

- Final users do not care how data is persisted
  - Scicat, ICAT, ISPyB, …?
- They care about understanding their data
  - Standard vocabulary/ontology
- Experimental team needs to follow the experiment in real team
-

UI: ISPyB vs Data portal



Number: 274

Time: Jun 2, 2023 2:52:03 PM

X: -1.170e-4

Y: -6.248e-5

Motion correction

Total Motion: 166.8

Avg. Motion/frame: 2.8

Frame Range: 0

Dose/frame: 1.01

Total dose: 0.74

CTF

Resolution Limit:: 2.91

Correlation: -0.095

Defocus U: 15576

Defocus V: 15648

Angle: 28.21

Estimated B factor: -0.0

Cutoff resolution

9.33333 — 28

9.33333 12.81

Classification #1: 5 classes (20687 particles)

# Class 14	# Class 19	# Class 13	# Class 3	# Class 5
Class Dis: 0.35	Class Dis: 0.15	Class Dis: 0.02	Class Dis: 0.01	Class Dis: 0.01
Resolution: 11.20	Resolution: 12.00	Resolution: 12.00	Resolution: 12.00	Resolution: 12.00

Classification #2: 8 classes (20757 particles)

© 2024 ESRF

Conclusion

- ISPyB is a
  - Tailored-made trusted solution where **what** and **how** needs to be defined
  - Extension or evolution of the experiments require changes in the data model
- ICAT is a
  - Generic solution which need to define only **what** you need to store
  - New techniques or extensions does not imply changes in the data model (but its content)
  - Inline with FAIR principles

- In general, for further development on a model:
  - Take into account your needs and interaction with other systems: proposals, sample tracking, etc…
  - Define a set of metadata that need to be associated to each acquisition/processing
    - Lot of work has already been done and might be a good starting point
  - Make use of a standard ontology (Interoperability) when possible
  - Manage public, private and embargoed data/metadata
  - FAIRness
    - Handle persistent identifiers (e.g: doi)
    - Interface to tape system and sharing capabilities
    - E-Logbook
    - Manage private and public data with embargo period

Thanks!

Backup slides

Workflow for on-line data processing of CM01 tomography data collections

Automatic processing:

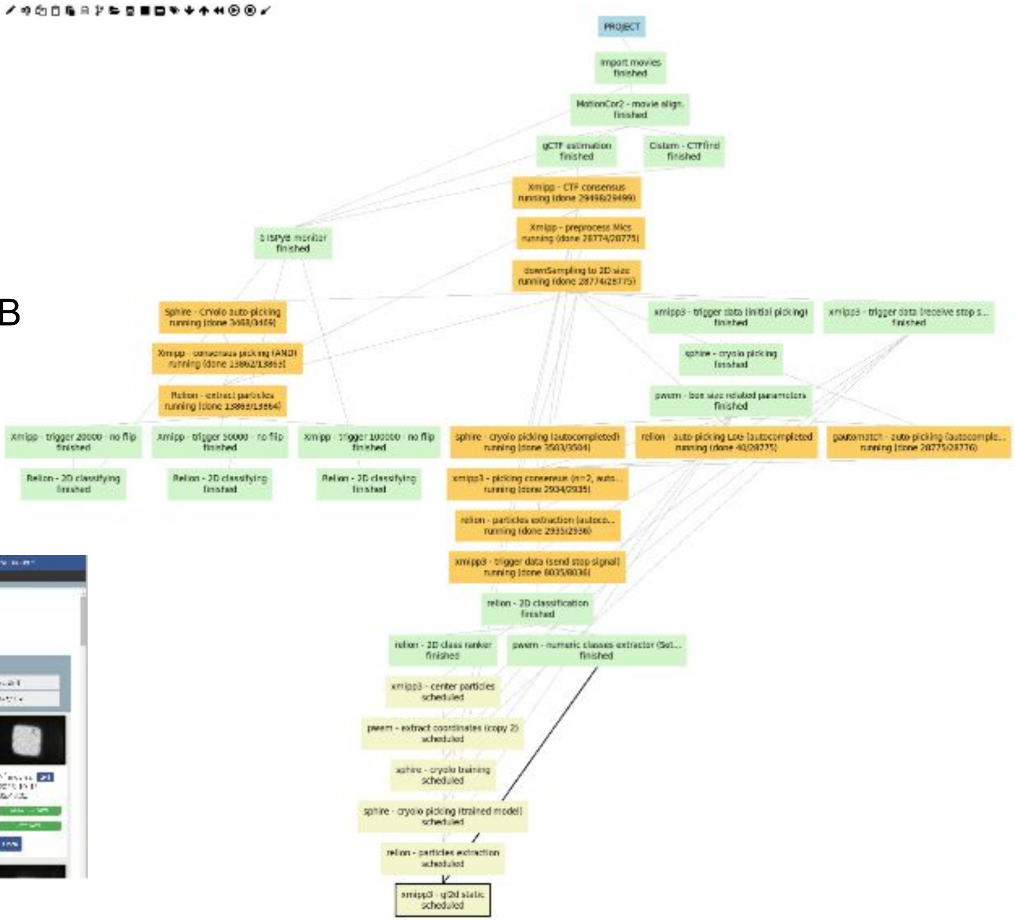
- Motion correction
- CTF
- Particle picking
- 2D classification

Raw data archived

- One experiment ~ 30000 movies / 10 TB

Processed data available in ISPyB

To do : 3D reconstruction



Outline

- **ISPyB**
  - Information System for Protein crYstallographic Beamlines
  - Tailored-made solution for MX/SB
  - Developed at the ESRF

- **ICAT**
  - Generic metadata catalog
  - Developed at STFC
  - Extended at the ESRF
  - Main features
    - Sample tracking
    - Electronic logbook
    - Tape interface
    - DOI