

ISPyB@ESRF

Alex de Maria
Data Manager@Data Analysis Unit
Software Group
ESRF

CryoEM

- Microscope to be installed in June
- Users by the end of the year with automatic pipeline and ISPyB up and running



- Use of SCIPION
- Use as much as possible of the software provided by DLS
- Use of ISPyB webservice
- Use of EXI for data visualization



Achievements

- **CM01** (CryoEM) was inaugurated on 10th of November
 - First users on 24th of November
- Data was collected and stored in **ISPyB** from the first experiment
 - Online data analysis by using **Scipion**
 - Data was ingested into ISPyB by using **web services**
 - Data was displayed on **EXI**



Partners



CM O1



EMBL



Thanks!!



Developments Timeline

- **Started July 2017**
 - Requirements analysis
 - Data Model
 - Pipeline using Scipion
- **October 2017**
 - Web Services
 - User interface in EXI



Data Model

16 Open	3 Closed	Author	Labels	Projects	Milestones	Assignee	Sort
1		Pins with multiple samples					4
		#20 opened on Nov 28, 2017 by KarlLevik					
1		Tables for processing jobs					
		#19 opened on Oct 8, 2017 by KarlLevik					
1		New column ScreeningOutput.alignmentSuccess?					
		#18 opened on Oct 2, 2017 by KarlLevik					
1		DataCollection column for identifying the internal data path in the HDF5 file					5
		#17 opened on Sep 25, 2017 by KarlLevik					
1		Recording the goniometer orientation of data collections					
		#16 opened on Sep 5, 2017 by KarlLevik					
1		Deleted column for Shipping, Dewar, Container and BLSample					
		#15 opened on Sep 1, 2017 by KarlLevik					
1		EM Data Model					10
		#14 opened on Aug 23, 2017 by stufisher					
1		Measure marked sample					
		#13 opened on Jul 18, 2017 by stufisher					
1		Data collection limits					
		#12 opened on Jul 11, 2017 by KarlLevik					
1		A role column for ProposalHasPerson					
		#11 opened on Jul 3, 2017 by KarlLevik					
1		Table for auto processing analysis					4
		#10 opened on Jun 28, 2017 by stufisher					
1		Changes for automatic shipment creation					5
		#8 opened on Jun 18, 2017 by stufisher					
1		Further changes for XPDF					1
		#6 opened on Jun 15, 2017 by KarlLevik					
1		Puck database schema changes					3
		#5 opened on Jun 15, 2017 by KarlLevik					
1		New "Notifications" table					10
		#3 opened on Mar 10, 2017 by olotvensson					
1		Add ISA to ISPyB strategy statistics					3
		#1 opened on Feb 6, 2017 by olotvensson					



<https://github.com/ispyb/ispyb-database-modeling/issues>



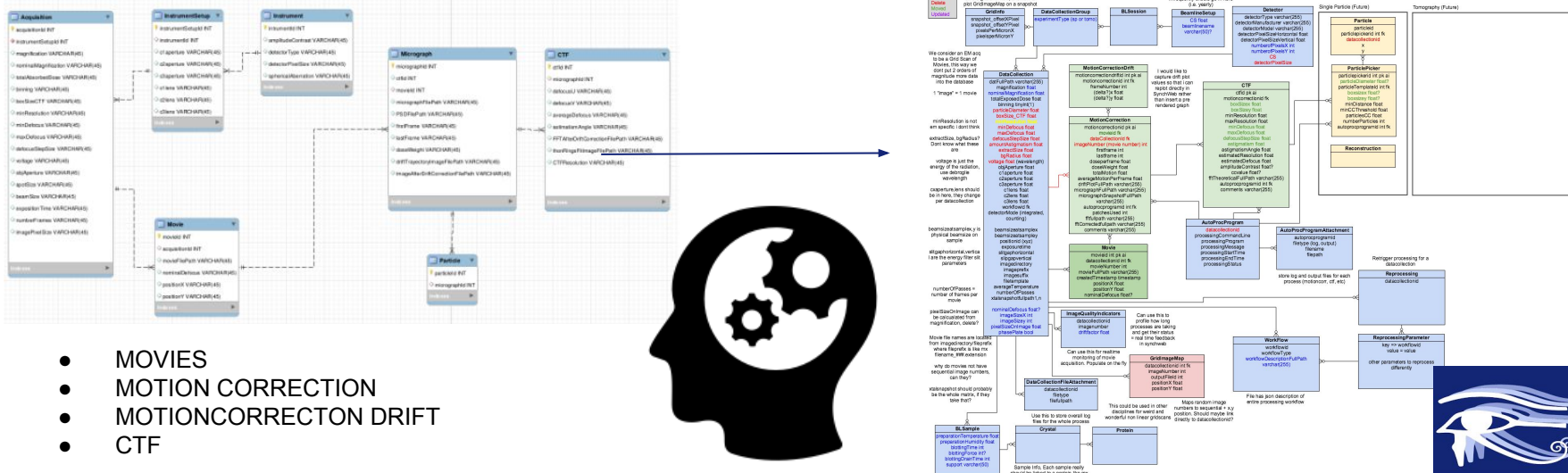
EM Data Model

#14 opened on Aug 23, 2017 by stufisher



10

Data Model



- MOVIES
- MOTION CORRECTION
- MOTIONCORRECTON DRIFT
- CTF

New
Updated
Deleted
Moved

can use this to tell us where 0.0 is on the grid overview, and how to plot GridImageMap on a snapshot

Only things that change very infrequently should go in here (i.e. yearly)

We consider an EM acq to be a Grid Scan of Movies, this way we dont put 2 orders of magnitude more data into the database

1 "image" = 1 movie

minResolution is not em specific i dont think

extractSize, bgRadius? Dont know what these are

voltage is just the energy of the radiation, use debroglie wavelength

cxaperture, lens should be in here, they change per datacollection

beamsizeatsample, y is physical beamsize on sample

slitgaphorizontal, vertical are the energy filter slit parameters

numberOfPasses = number of frames per movie

pixelSizeOnImage can be calculated from magnification, delete?

GridInfo
snapshot_offsetXPixel
snapshot_offsetYPixel
pixelsPerMicronX
pixelsPerMicronY

DataCollectionGroup
experimentType (sp or tomo)

BLSession

BeamlineSetup
CS float
beamlineName varchar(50)?

Detector
detectorType varchar(255)
detectorManufacturer varchar(255)
detectorModel varchar(255)
detectorPixelSizeHorizontal float
detectorPixelSizeVertical float
numberOfPixelsX int
numberOfPixelsY int
CS
detectorPixelSize

Single Particle (Future)

Tomography (Future)

Particle
particleId
particlePickerId int fk
datacollectionid
x
y

ParticlePicker
particlePickerId int pk ai
particleDiameter float?
particleTemplateId int fk
boxsizeX float?
boxsizeY float?
minDistance float
minCCThreshold float
particlesCC float
numberParticles int
autoprocprogramid int fk

Reconstruction

MotionCorrectionDrift
motionCorrectionDriftId int pk ai
motionCorrectionId int fk
frameNumber int
(delta?)x float
(delta?)y float

I would like to capture drift plot values so that i can replot directly in SynchWeb rather than insert a pre rendered graph

MotionCorrection
motionCorrectionId pk ai
movieId fk
dataCollectionId fk
imageNumber (movie number) int
firstFrame int
lastFrame int
dosePerFrame float
doseWeight float
totalMotion float
averageMotionPerFrame float
driftPlotFullPath varchar(255)
micrographFullPath varchar(255)
micrographSnapshotFullPath varchar(255)
autoprocprogramid int fk
patchesUsed int
fftFullPath varchar(255)
fftCorrectedFullPath varchar(255)
comments varchar(255)

CTF
ctfId pk ai
motionCorrectionId fk
boxSizeX float
boxSizeY float
minResolution float
maxResolution float
minDefocus float
maxDefocus float
defocusStepSize float
astigmatism float
astigmatismAngle float
estimatedResolution float
estimatedDefocus float
amplitudeContrast float?
cvalue float?
fftTheoreticalFullPath varchar(255)
autoprocprogramid int fk
comments varchar(255)

AutoProcProgram
datacollectionid
processingCommandLine
processingProgram
processingMessage
processingStartTime
processingEndTime
processingStatus

AutoProcProgramAttachment
autoprocprogramid
filetype (log, output)
filename
filepath

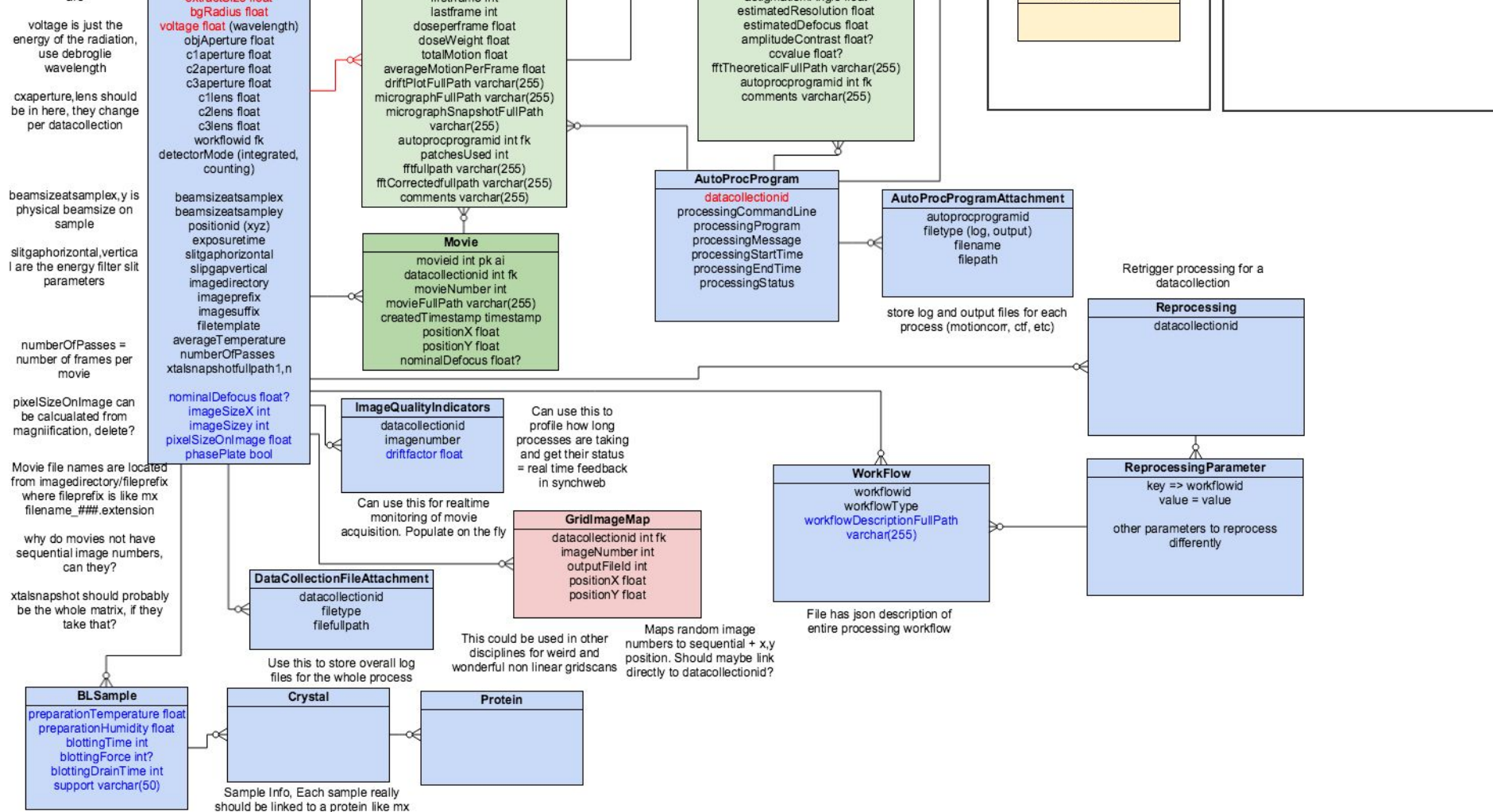
store log and output files for each process (motioncorr, ctf, etc)

Retrigger processing for a datacollection

Reprocessing
datacollectionid

ImageQualityIndicators
datacollectionid

Can use this to profile how long

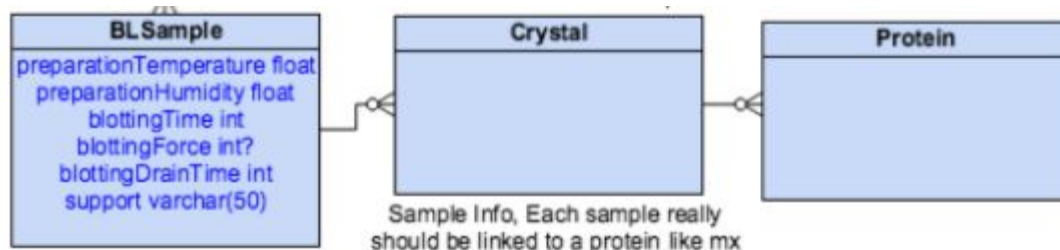


Potential Issues

- Data Collection Table
 - **Grid square** snapshot stored as **xtalSnapshot1**
 - **Voltage** stored as **wavelength**
- DataCollectionGroup Table
 - Grid snapshot stored as xtalSnapshotFullPath

Potential Issues

- BLSample



BLSample	
blSampleId	INT(10)
crystalId	INT(10)
containerId	INT(10)
diffractionPlanId	INT(10)
name	VARCHAR(45)
code	VARCHAR(45)
location	VARCHAR(45)
holderLength	DOUBLE
loopLength	DOUBLE
loopType	VARCHAR(45)
wireWidth	DOUBLE
comments	VARCHAR(1024)
completionStage	VARCHAR(45)
structureStage	VARCHAR(45)
publicationStage	VARCHAR(45)
publicationComments	VARCHAR(255)
blSampleStatus	VARCHAR(20)
isInSampleChanger	TINYINT(1)
lastKnownCenteringPosition	VARCHAR(255)
SMILES	VARCHAR(400)
Indexes	

Movie
movieid INT(11)
dataCollectionId INT(11)
movieNumber VARCHAR(45)
movieFullPath VARCHAR(255)
positionX VARCHAR(45)
positionY VARCHAR(45)
micrographFullPath VARCHAR(255)
micrographSnapshotPath VARCHAR(255)
xmlMetaDataFullPath VARCHAR(255)
dosePerImage VARCHAR(45)
createdTimeStamp TIMESTAMP
Indexes

ESRF

Movie
movieid int pk ai
datacollectionid int fk
movieNumber int
movieFullPath varchar(255)
createdTimestamp timestamp
positionX float
positionY float
nominalDefocus float?

DLS

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(45)
micrographFullPath VARCHAR(45)
micrographSnapshotFullPath VARCHAR(45)
correctedDoseMicrographFullPath VARCHAR(45)
patchesUsed VARCHAR(45)
logFileFullPath VARCHAR(45)
createdTimeStamp VARCHAR(45)
Indexes

ESRF

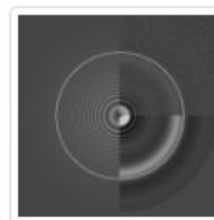
MotionCorrection
motioncorrectionid pk ai
movieid fk
dataCollectionId fk
imageNumber (movie number) int
firstframe int
lastframe int
doseperframe float
doseWeight float
totalMotion float
averageMotionPerFrame float
driftPlotFullPath varchar(255)
micrographFullPath varchar(255)
micrographSnapshotFullPath varchar(255)
autoprocprogramid int fk
patchesUsed int
fftFullPath varchar(255)
fftCorrectedFullPath varchar(255)
comments varchar(255)

DLS

CTF	
ctfId	INT(11)
motionCorrectionId	INT(11)
spectralImageThumbnailFullPath	...
spectralImageFullPath	VARCHAR(255)
defocusU	VARCHAR(45)
defocusV	VARCHAR(45)
angle	VARCHAR(45)
crossCorrelationCoefficient	VARCHAR(45)
resolutionLimit	VARCHAR(45)
estimatedBfactor	VARCHAR(45)
logFilePath	VARCHAR(45)
createdTimeStamp	VARCHAR(45)

CTF	
ctfId	pk al
motionCorrectionId	fk
boxSizeX	float
boxSizeY	float
minResolution	float
maxResolution	float
minDefocus	float
maxDefocus	float
defocusStepSize	float
astigmatism	float
astigmatismAngle	float
estimatedResolution	float
estimatedDefocus	float
amplitudeContrast	float?
ccvalue	float?
fftTheoreticalFullPath	varchar(255)
autoprocprogramId	int fk
comments	varchar(255)

CTF	
Resolution Limit	3.402
Correlation	0.088252
Defocus U	13639.72
Defocus V	14075.71
Angle	56.29
Estimated B factor	87.71



Scipion Monitors

- **Development of Scipion monitors**
 - Based on **DLS** work
 - Automatic capture of metadata
 - Grid and grid squares
 - Movies
 - Motion correction
 - CTF
 - Use of web services



Web Services

```
@WebMethod(operationName = "addMovie")
public Movie addMovie(
    @WebParam(name = "proposal") String proposal,
    @WebParam(name = "proteinAcronym") String proteinAcronym,
    @WebParam(name = "sampleAcronym") String sampleAcronym,
    @WebParam(name = "movieDirectory") String movieDirectory,
    @WebParam(name = "movieFullPath") String movieFullPath,
    @WebParam(name = "movieNumber") String movieNumber,
    @WebParam(name = "micrographFullPath") String micrographFullPath,
    @WebParam(name = "micrographSnapshotFullPath") String micrographSnapshotFullPath,
    @WebParam(name = "xmlMetaDataFullPath") String xmlMetaDataFullPath,
    @WebParam(name = "voltage") String voltage,
    @WebParam(name = "sphericalAberration") String sphericalAberration,
    @WebParam(name = "amplitudeContrast") String amplitudeContrast,
    @WebParam(name = "magnification") String magnification,
    @WebParam(name = "scannedPixelsSize") String scannedPixelsSize,
    @WebParam(name = "imagesCount") String imagesCount,
    @WebParam(name = "dosePerImage") String dosePerImage,
    @WebParam(name = "positionX") String positionX,
    @WebParam(name = "positionY") String positionY,
    @WebParam(name = "beamlineName") String beamlineName,
    @WebParam(name = "gridSquareSnapshotFullPath") String gridSquareSnapshotFullPath,
    @WebParam(name = "startTime") String stringStartTime
```


Developments@EXI

Alex de Maria
Data Manager@Data Analysis Unit
Software Group
ESRF

EXI for CryoEM

Common ISPyB access for all Techniques

Choose a Date

Choose a period of time

enter search term (proposal or title)

search by local contact

BM29

ID23-1

ID23-2

ID29

ID30A-1

ID30A-2

ID30A-3

ID30B

BM14U

BM30A

CM01

Filter by term (proposal or title) or comment

MX

BIOSAXS

EM

Start

Beamline

A-form

Proposal

Local Contact

En. Scans

XRF

Samples

Tests

Collects

Calibration

SC

HPLC

Grid

Movie

Comments

28-01-2018

BM29

PERNOT P

2

13

6

29-01-2018

ID30A-1

BOWLER M

29-01-2018

ID29

SOLER LOPEZ M

21

31

17

29-01-2018

ID23-2

ZANDER U

14

63

29-01-2018

ID30B

MUELLER DIECKMANN C

1

8

40

27

29-01-2018

BM29

VAN LAER B

3

1

4

28-01-2018

ID23-1

Session created by the BCM

28-01-2018

ID30A-1

BOWLER M

24

11

71

MXPressAUTO continues

28-01-2018

ID29

1

Session created by the BCM

28-01-2018

ID30A-1

BOWLER M

EXI for CryoEM

Common ISPyB access for all Techniques

Choose a Date

Choose a period of time

enter search term (proposal or title)

search by local contact

BM29

ID23-1

ID23-2

ID29

ID30A-1

ID30A-2

ID30A-3

ID30B

BM14U

BM30A

CM01

Filter by term (proposal or title) or comment

					MX					BIOSAXS			EM		
Start	Beamline	A-form	Proposal	Local Contact	En. Scans	XRF	Samples	Tests	Collects	Calibration	SC	HPLC	Grid	Movie	Comments
29-01-2018	ID30A-1	<div></div>	<div></div>	<div></div>											<div></div>
28-01-2018	BM29	<div></div>	<div></div>	<div></div>							<div>6</div>				<div></div> session created by ISPyBB
28-01-2018	ID23-2	<div></div>	<div></div>	<div></div>				<div>90</div>	<div>25</div>						<div></div>
28-01-2018	ID29	<div></div>	<div></div>	<div></div>	<div>1</div>		<div>70</div>	<div>103</div>	<div>81</div>						<div></div>
28-01-2018	ID30B	<div></div>	<div></div>	<div></div>	<div>1</div>		<div>20</div>	<div>141</div>	<div>88</div>						<div></div>
28-01-2018	ID30A-3	<div></div>	<div></div>	<div></div>				<div>85</div>	<div>95</div>						<div></div>
28-01-2018	ID23-2	<div></div>	<div></div>	<div></div>											<div></div> Session created by the BCM
27-01-2018	ID30A-1	<div></div>	<div></div>	<div></div>			<div>1</div>	<div>1</div>	<div>5</div>						<div></div> MXPressAUTO started
27-01-2018	ID30A-1	<div></div>	<div></div>	<div></div>			<div>44</div>	<div>22</div>	<div>145</div>						<div></div> MXPressAUTO continues
26-01-2018	CM01	<div></div>	<div></div>	<div></div>			<div>1</div>		<div>8</div>				<div>8</div>	<div>160</div>	<div></div>

New Developments on EXI

BioSAXS

- It is used in daily basis
- Still many minor issues to be solved
- No progress done last semester

EXI for CryoEM

MX

New Developments on EXI



New Developments on EXI

CryoEM



ACKNOWLEDGEMENTS



- Alun Ashton
- Karl Levik
- Stuart Fisher
- Kevin Savage
- Daniel Clare



- Gregory Effantin



- Michael Hons



- Jose Maria Carazo
- Pablo Conesa
- Carlos Oscar Lozano
- Roberto Marabini
- Jose Maria de la Rosa



- Olof Svensson
- Eaazhisai Kandiah
- Stephanie Monaco
- Christoph Mueller

- Solange Delageniere
- Gordon Leonard