

**ISPyB Developer's meeting
December 3, 2018**

Present:

DLS: Karl Levik (KL), Neil Smith (NS), James Hall (JH)
GP: Peter Keller (PK), Gerard Bricogne (GB), Rasmus Fogh (RF),
Clemens Vornrhein (CV)
EMBLHH: Ivars Karpics (IK)
ELETTRA: Milan Prica (MP)
ESRF: Stephanie Monaco (SM), Olof Svensson (OS), Solange Delageniere (SD)

Agenda:

- Follow-up of actions from last VC
- Presentation of MR status at ESRF
- Autoprocessing and display of anom/no anom results at ESRF

1. Follow-up of actions from last VC

There is nothing new on the production of a mock-up for the display of STARANISO results in ISPyB viewers. In discussion it was agreed that there could be up to three different diffraction limits, but that there should be no need for displaying the directions of the diffraction limits in this context. There should also be slots for completeness statistics and for an “operational resolution” (which should give values comparable between spherical and anisotropic data cut-offs). It was discussed whether the diffraction limits could be shown as a graphical figure.

There were no news on Cryo-EM data model merging.

Follow up of Trieste meeting: all the presentations, and a photo were put on the web site.

2. Presentation of MR status at ESRF

OS reported on the display of automatic MR in EXI.. This involved a single change in database tables, using the Phasing table, with the rest in EXI (see attached screenshot in Annexe 1), and included displays of maps, e.g. with the highest difference density peak.

It is established that Synchweb is under active development.

The point is made that there is already a table for MR data, so using the Phasing table may not be ideal. OS pointed out that they had needed an extremely quick result to have something to show before the December shutdown. ESRF will have a look after the shutdown to decide if we move the development to MXMR tables or if we try to rationalize tables and reduce the nb of tables used.

3. Autoprocessing and display of anom/no anom results at ESRF

The processing at the ESRF has since the beginning produced results and statistics for both anon and no-anom processing (Friedel's pairs true or false). For some pipelines (Grenades, EDNA_proc) these results can be obtained with little extra processing power as a part of the pipeline is common, however, for other pipelines (autoPROC, XDSAPP and XIA2_DIALS) the whole pipeline is executed twice.

ESRF are now in the process of switching over to processing always as anomalous='yes' and providing both types of statistics afterwards. This will halve the number of lines in the summary table, at the cost of increasing the number of columns. This clearly intersects with the STARANISO discussion, which will add more columns such as anisotropic completeness.

CV pointed out that anomalous versus non-anomalous only makes a difference for the R-values. Otherwise the statistics are the same, it is only the calculation and results that differ. His proposal is to always do calculations (scaling etc.) non-anomalously, and then decide during processing whether the data justify anomalous treatment. One effect of running anomalous processing is that it reduces the multiplicity, which in itself improves (spuriously) the R-values. It is obvious to see in extreme cases, but intermediate situations require more consideration. GB concurs, pointing out that scaling on non-anomalous data cannot interfere with later anomalous signal detection and exploitation if it is present, as it varies over a different spatial scale than the effects of scaling parameters. Scaling on the assumption of non-anomalous data involves fewer free parameters, which is an advantage. Giving both types of results is mostly for historical reasons, and we should not be slaves to tradition.

OS points out that this is a user requirement, not his idea. The person most on top of the details within ESRF (Max) is unfortunately away sick at the moment.

NS says that Diamond also does anomalous-only processing, but that the pipeline needs are changing, and they may want to cut the numbers they show.

OS points out that both auto-processing procedures and statistics have been discussed since the meeting at Diamond, but that the people needed to decide have never been present at the same time. This should involve scientists, and should maybe be put to the upcoming workshop as well. GB agrees.

KL suggests that it is the purpose of ISPyB is simply to display the data that are available, but GB counters that the choice of data to present is not a neutral act and carries a responsibility to assess the potential to mislead users by e.g. sticking to discredited criteria such as Rmerge in the ranking of datasets.

It is discussed, and agreed, that a developers' meeting would be beneficial. NS suggests that the MXCuBE/ISPyB meetings are not ideal and suggests a web meeting instead. But GB, with wide agreement, suggests that a physical meeting is necessary to establish agreement and truly put heads together. It is agreed to organise one. NS proposes refactoring as a theme, like that of the recent (mid-November) MXCuBE developers' meeting. It will be necessary to do preparations ahead of time, and people should start thinking about discussion ideas now. SD proposes, and it is agreed, that there should be a github issue to carry this conversation forward. GB will propose that developers should be consulted in planning the next MXCuBE/ISPyB meeting.

It is agreed that the metrics/autoprocessing workshop should be held after the next meeting with MXCuBE, before Easter, at ESRF. It should be held over three days, from lunch to lunch, should invite key developers, and should be open to XFEL, ED and neutron diffraction practitioners as well. In summary:

- The workshop should gather the relevant people and involve them
- It appears that the workshop should not only be about the metric and what and how to present but also about the content of the metrics, and how we should process data, recommended protocols.
- Example from the Trieste meeting : bad usage of RMerge.
- We should reconsider what to display.
- Have scientists and experts make presentations and then have a F2F discussion
-> after next ISPyB F2F meeting, that's why the dates of April 8th has been proposed.
- Invite guests speakers : ILL, XFEL, ...

4. Preparation of next F2F meeting

Ideas from Neil:

- Break-down to small web applications ?
- Microservices ?
- Have a F2F developers meeting like MXCube one ?

Solange will open an issue to start a discussion to gather ideas for next Meeting

**Next VC : January 7th 15h (UTC+1) on appear.in/ispyb
Everyone is invited to propose topics before**

Annexe 1 : EXI MR screenshots

01-12-2018 20:47:01
data/visitor/.../id50a120181201RAW_DATA/

Summary Beamline Parameters Data Collection 1 Sample Last Collect Results 28 Workflow 7 MR 2

Phasing	PHASING	REFINEMENT	Download	Program	Method	Resolution	Electron density MR	Electron density REFINEMENT
1222	✓	✓	 	mr_phasing	MR	1.8 - 63.246		

