On-Line Data Analysis Workflows for Macro-Molecular Crystallography at the ESRF

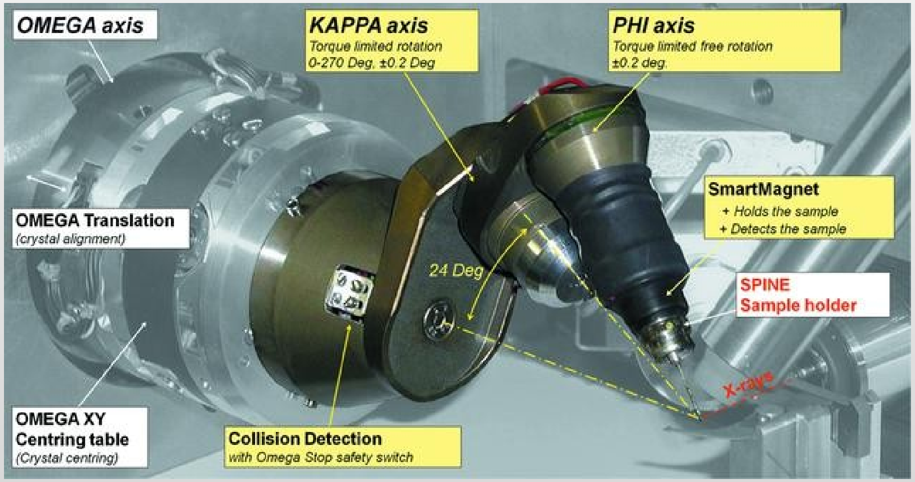
Olof Svensson[1], Sandor Brockhauser[2], Matthew Bowler[2], Max Nanao[2] and Matthew Gerring[3]

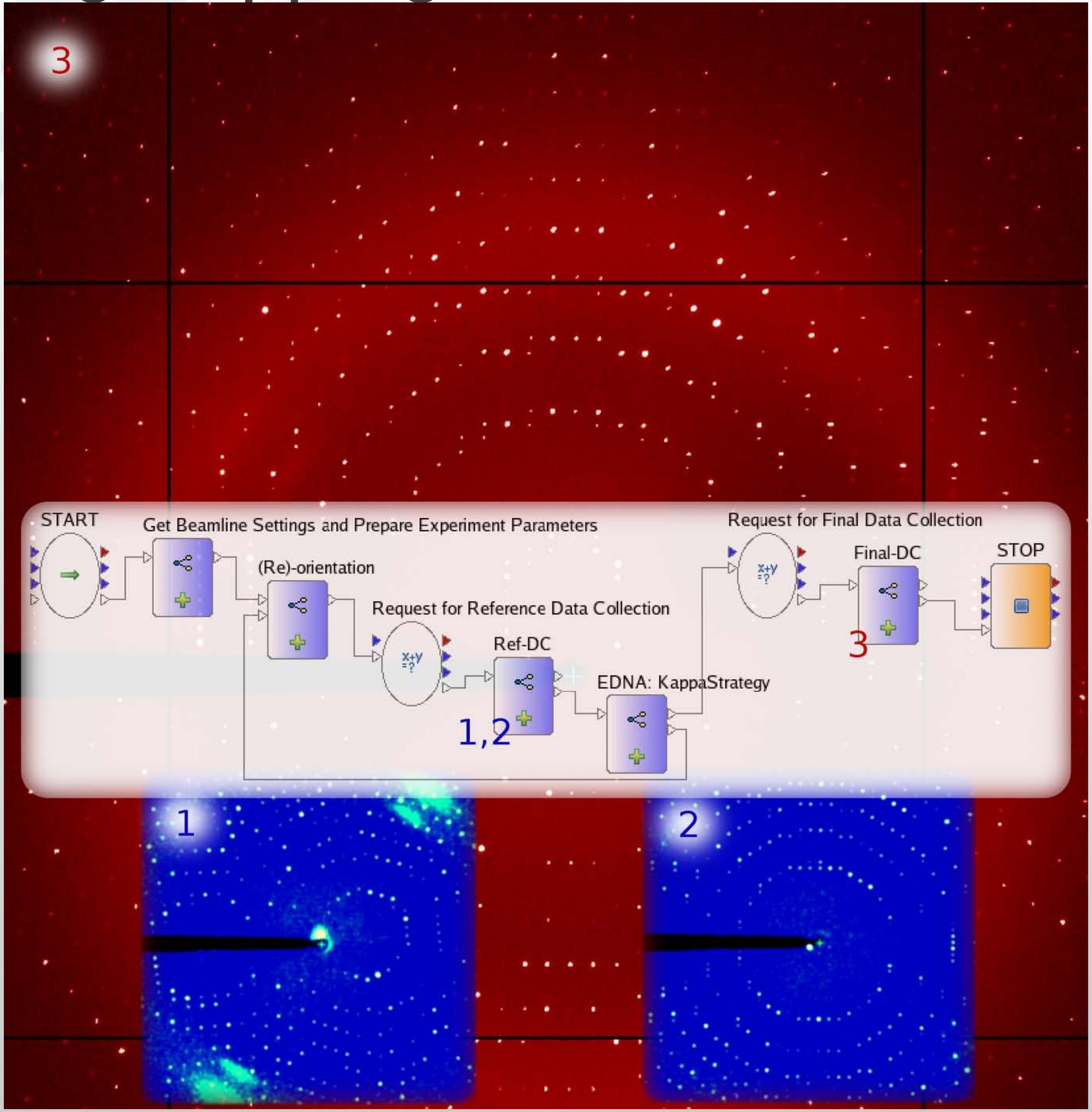
1) ESRF, France
2) EMBL Grenoble, France
3) DLS, UK

Overview

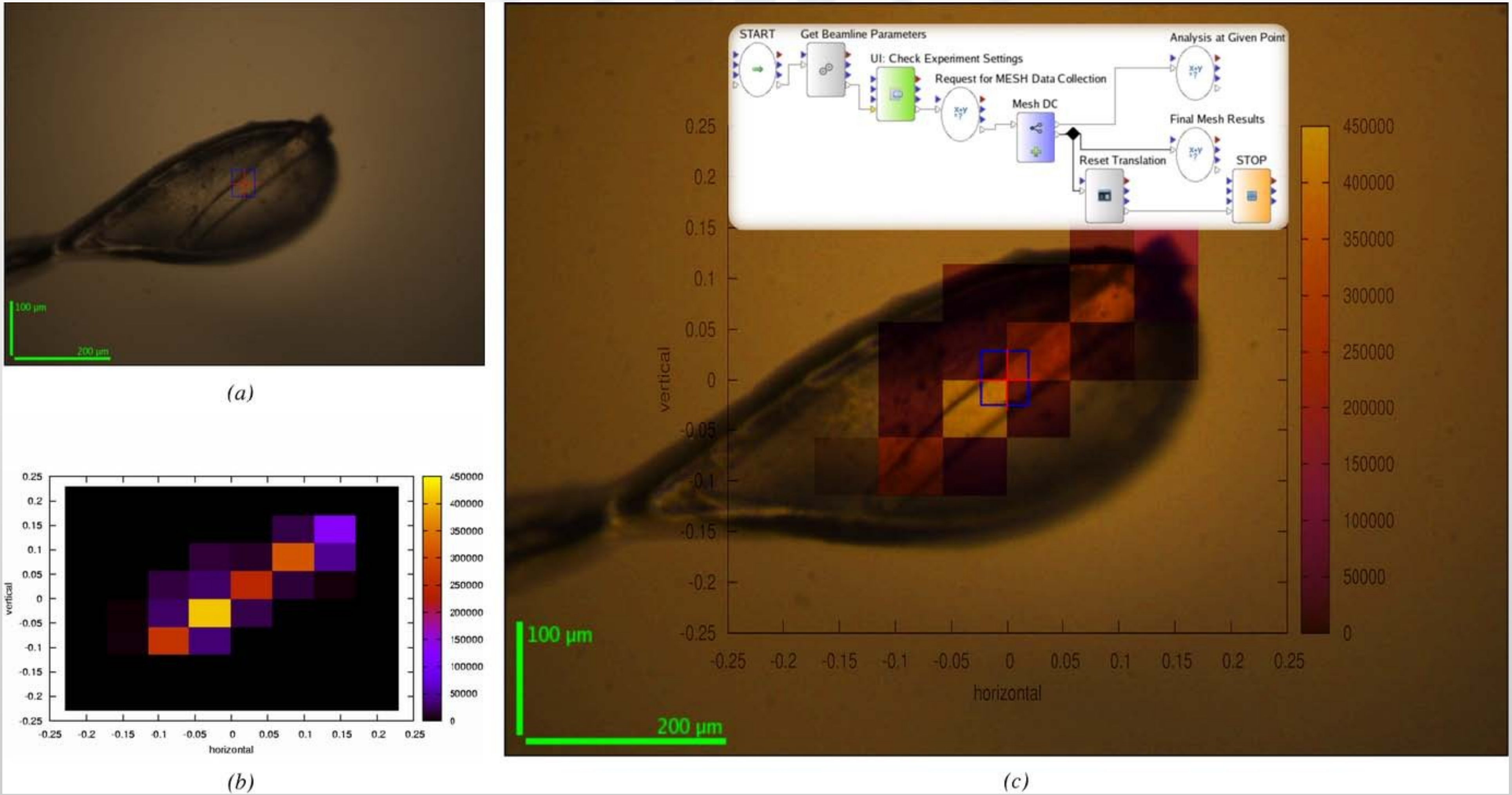
- On-line data analysis workflow examples:
  - Kappa goniostat reorientation
  - X-ray auto centring (mesh scan)
- Why use a workflow tool?
- DAWN vs EDNA
- Integration on ESRF MX beamlines
- Acknowledgenemnts

MX On-line data analysis 1 : Re-orientation using Kappa goniometer





MX On-line data analysis 2 : Mesh / grid scan

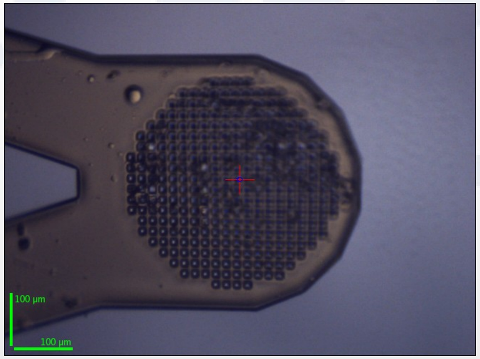


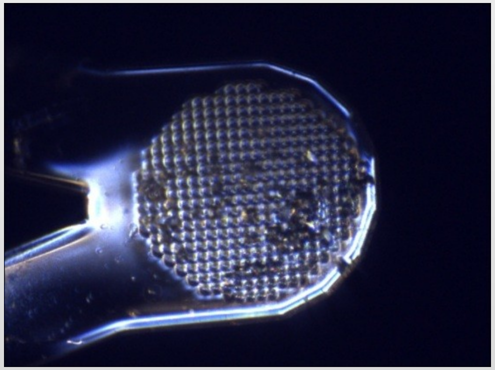
(a)

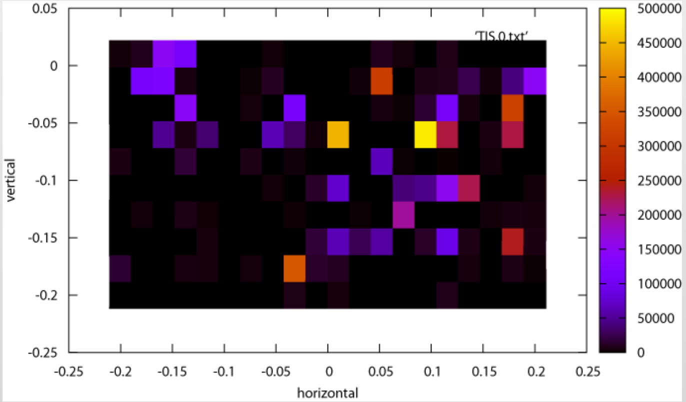
(b)

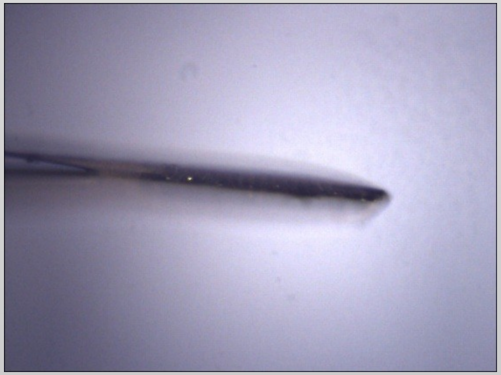
(c)

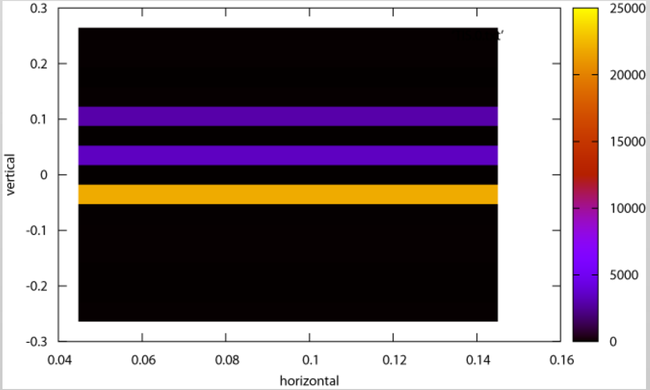
MX 3D Auto Centre



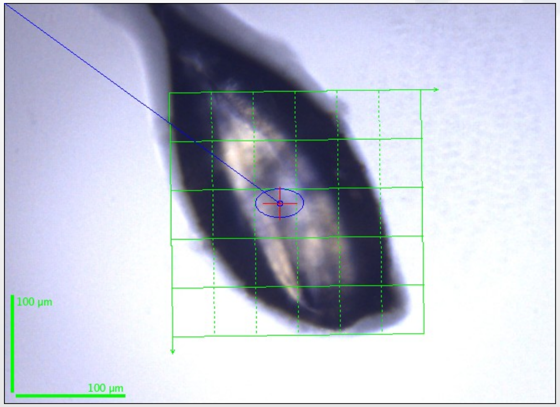


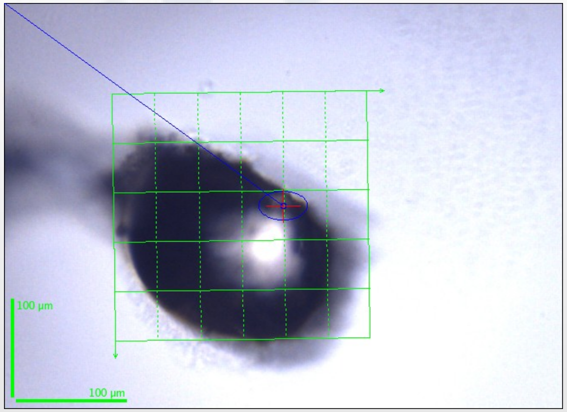


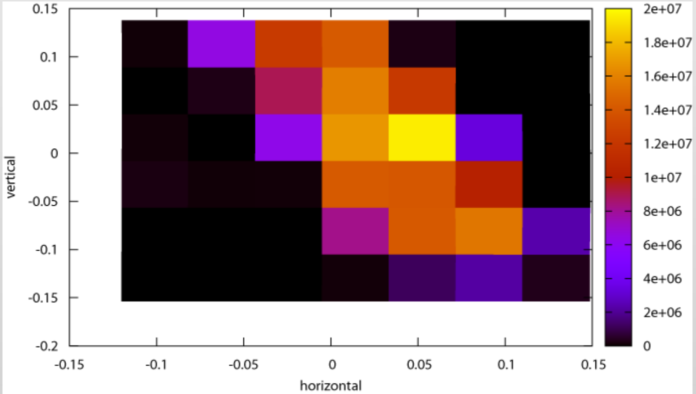


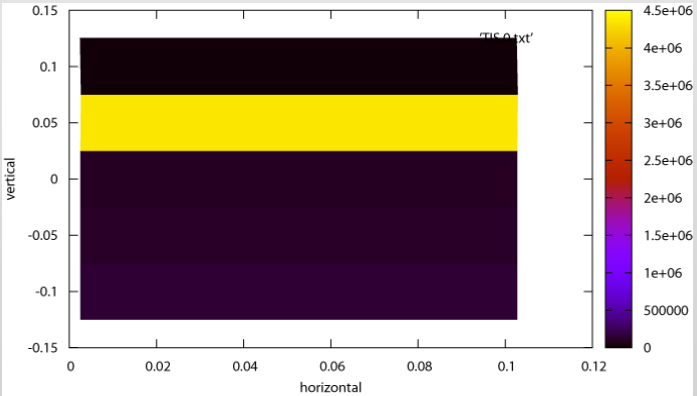


Combined workflows : Kappa reorientation and mesh scan







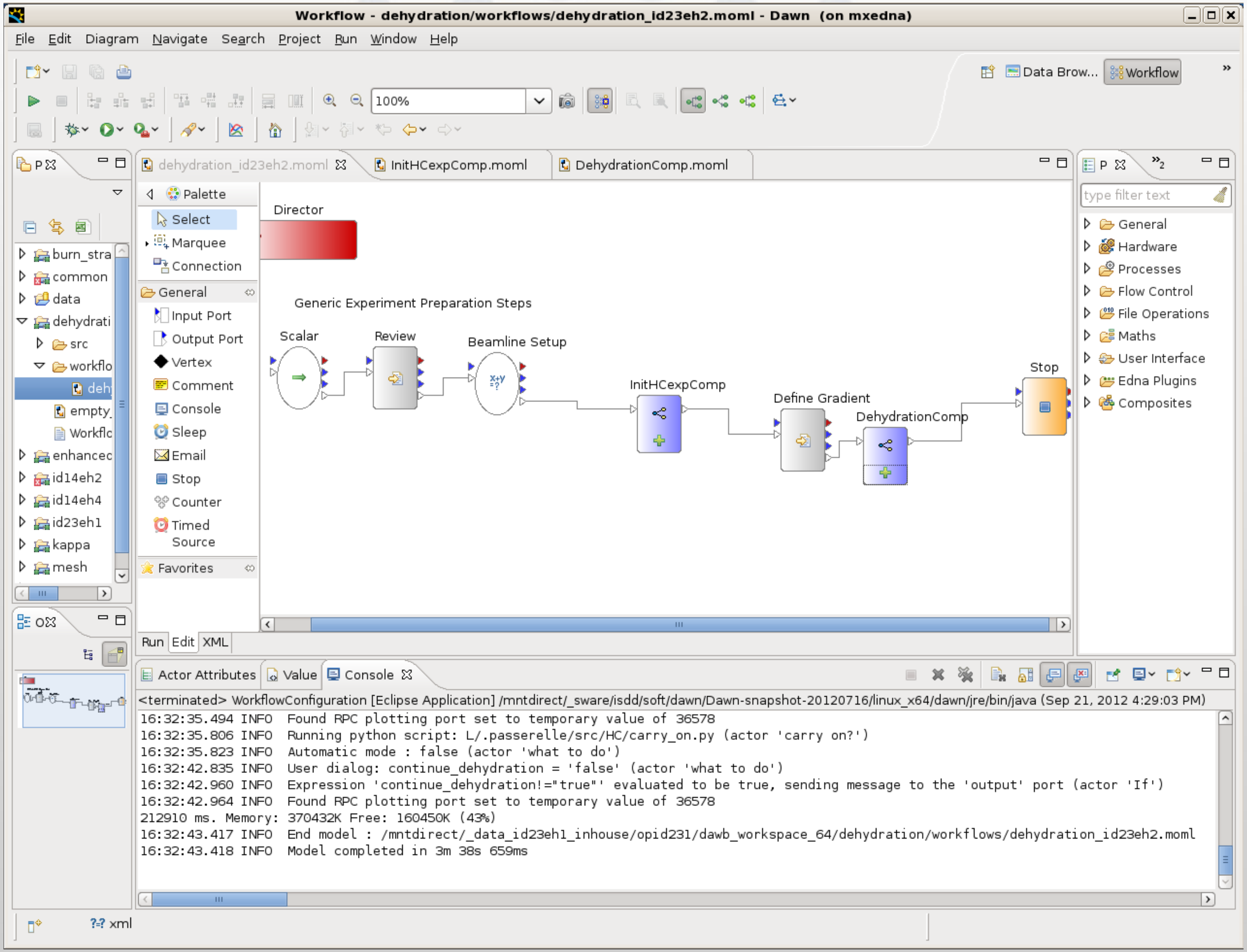


The orthorhombic sample was a very thin plate – difficult to visualise and very radiation sensitive – all standard data collections had resulted in low resolution at low completeness

Why use a workflow tool?

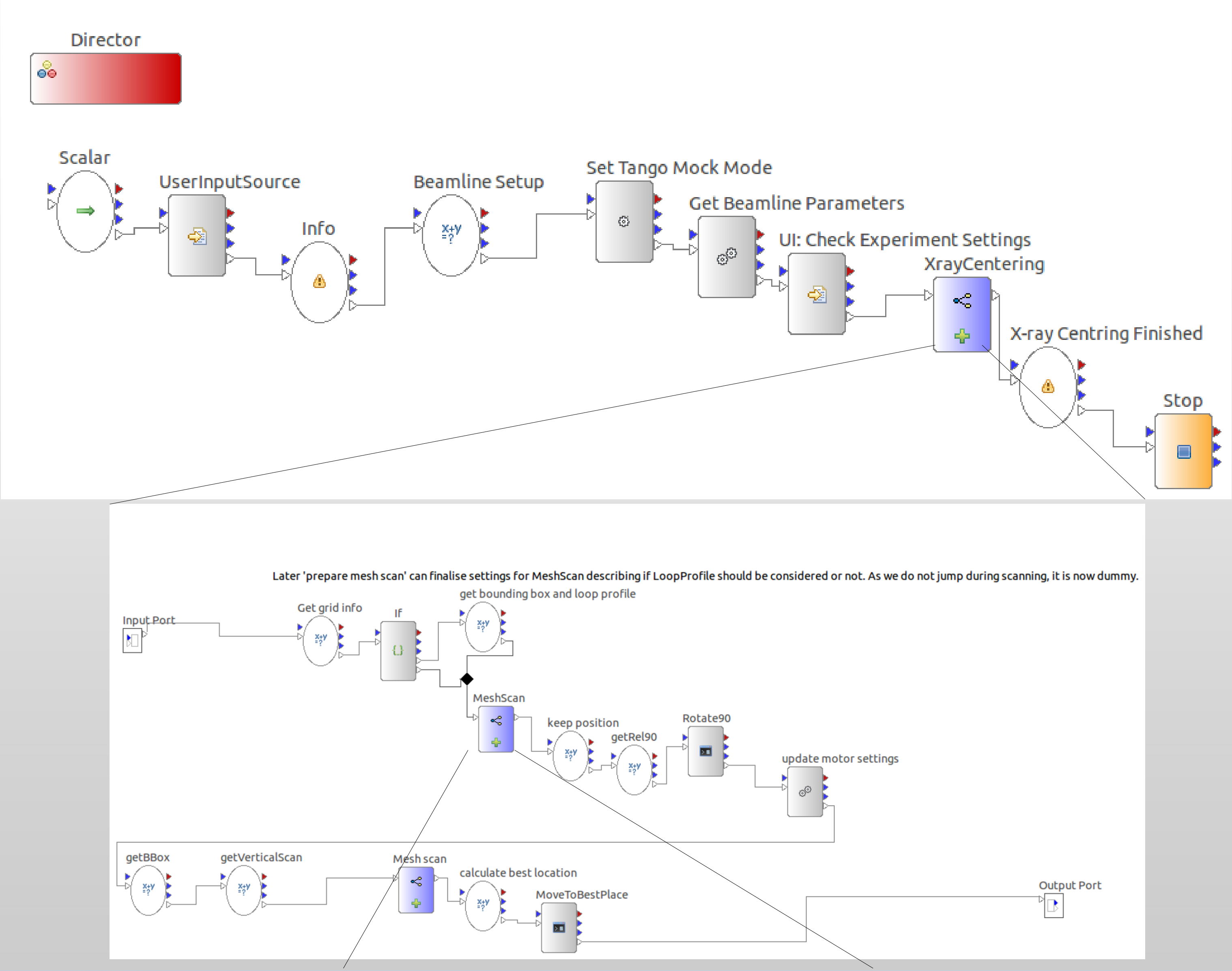
- Workflow tools for data analysis are new for synchrotron radiation facilities however they are widely used in other scientific fields e.g. biology
- They offer a higher level programming language than traditional languages like C, Python, Fortran etc
- The goal is not to replace these languages but to complement them
- Workflows facilitates development of the high level analysis:
  - Parallelism
  - Documentation
- Workflows ≠ Labview

The workflow tool : DAWN and Passerelle

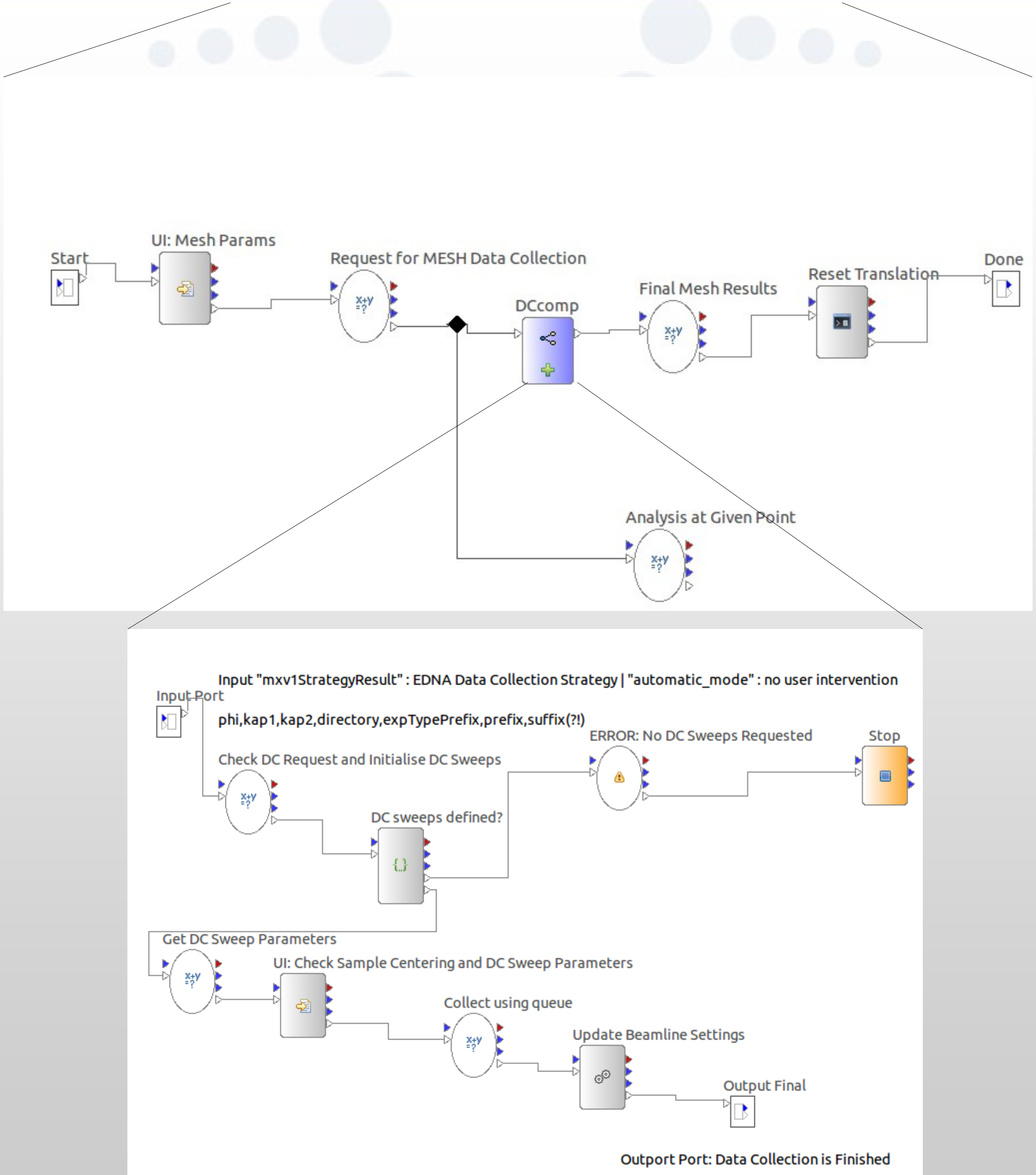


```
<terminated> WorkflowConfiguration [Eclipse Application] /mntdirect/_sware/isdd/soft/dawn/Dawn-snapshot-20120716/linux_x64/dawn/jre/bin/java (Sep 21, 2012 4:29:03 PM)
16:32:35.494 INFO  Found RPC plotting port set to temporary value of 36578
16:32:35.806 INFO  Running python script: L/.passerelle/src/HC/carry_on.py (actor 'carry on?')
16:32:35.823 INFO  Automatic mode : false (actor 'what to do')
16:32:42.835 INFO  User dialog: continue_dehydration = 'false' (actor 'what to do')
16:32:42.960 INFO  Expression 'continue_dehydration!="true"' evaluated to be true, sending message to the 'output' port (actor 'If')
16:32:42.964 INFO  Found RPC plotting port set to temporary value of 36578
212910 ms. Memory: 370432K Free: 160450K (43%)
16:32:43.417 INFO  End model : /mntdirect/_data_id23eh1_inhouse/opid231/dawb_workspace_64/dehydration/workflows/dehydration_id23eh2.moml
16:32:43.418 INFO  Model completed in 3m 38s 659ms
```

Workflow hierarchy

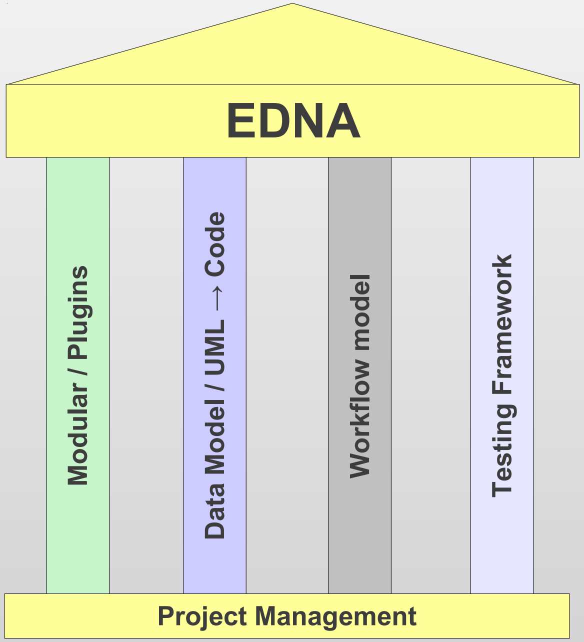


Later 'prepare mesh scan' can finalise settings for MeshScan describing if LoopProfile should be considered or not. As we do not jump during scanning, it is now dummy.

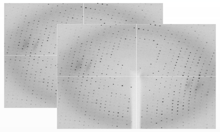

Start
UI: Mesh Params
Request for MESH Data Collection
DCcomp
Final Mesh Results
Reset Translation
Done
Analysis at Given Point
Input "mxv1StrategyResult" : EDNA Data Collection Strategy | "automatic_mode" : no user intervention
Input Port
phi,kap1,kap2,directory,expTypePrefix,prefix,suffix(?!)
ERROR: No DC Sweeps Requested
Stop
Check DC Request and Initialise DC Sweeps
DC sweeps defined?
Get DC Sweep Parameters
UI: Check Sample Centering and DC Sweep Parameters
Collect using queue
Update Beamline Settings
Output Final
Outport Port: Data Collection is Finished

Workflows can implemented using EDNA, but EDNA is not a Workflow tool!

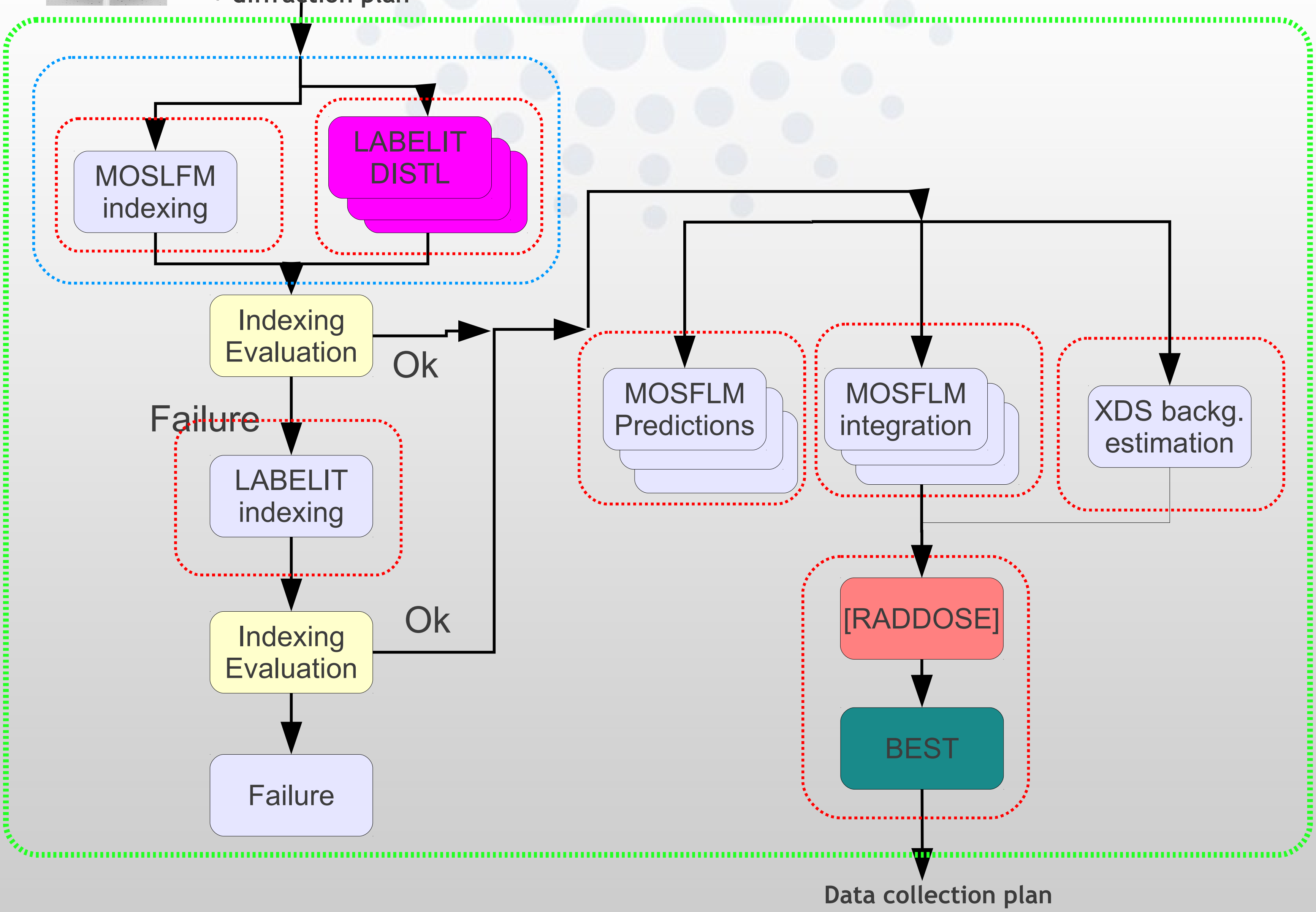
- EDNA is a framework aimed for online data analysis:
  - Python based
  - Data models
  - Modular / plugins
  - Workflows / pipelines
  - Strong testing
- No GUI!
  - All workflows must be programmed
  - No (buit-in) visualisation of results



EDNA Characterisation v1.3



+ 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

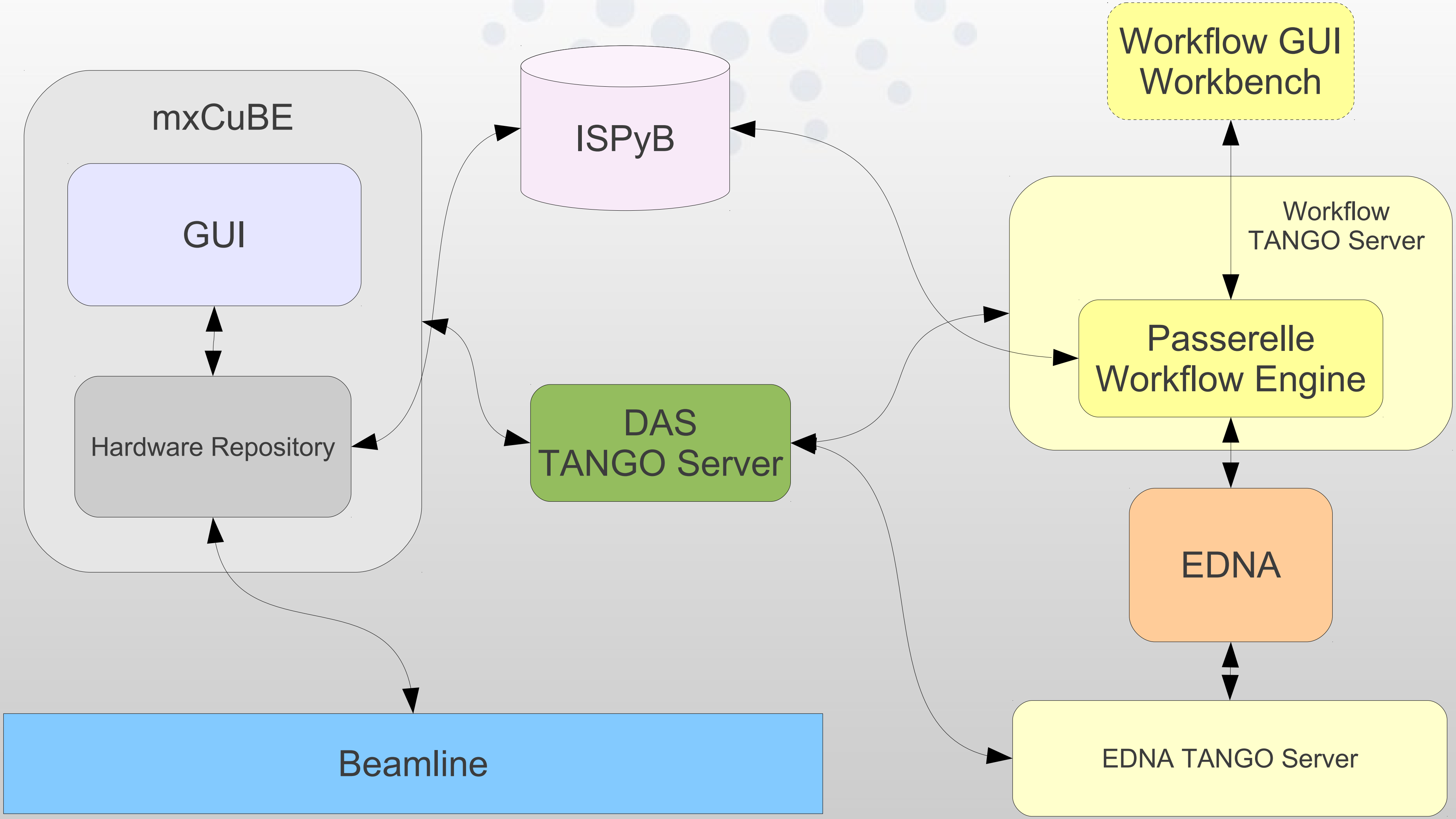
Data collection plan

Integration of the workflow tool in beamline control / data acquisition

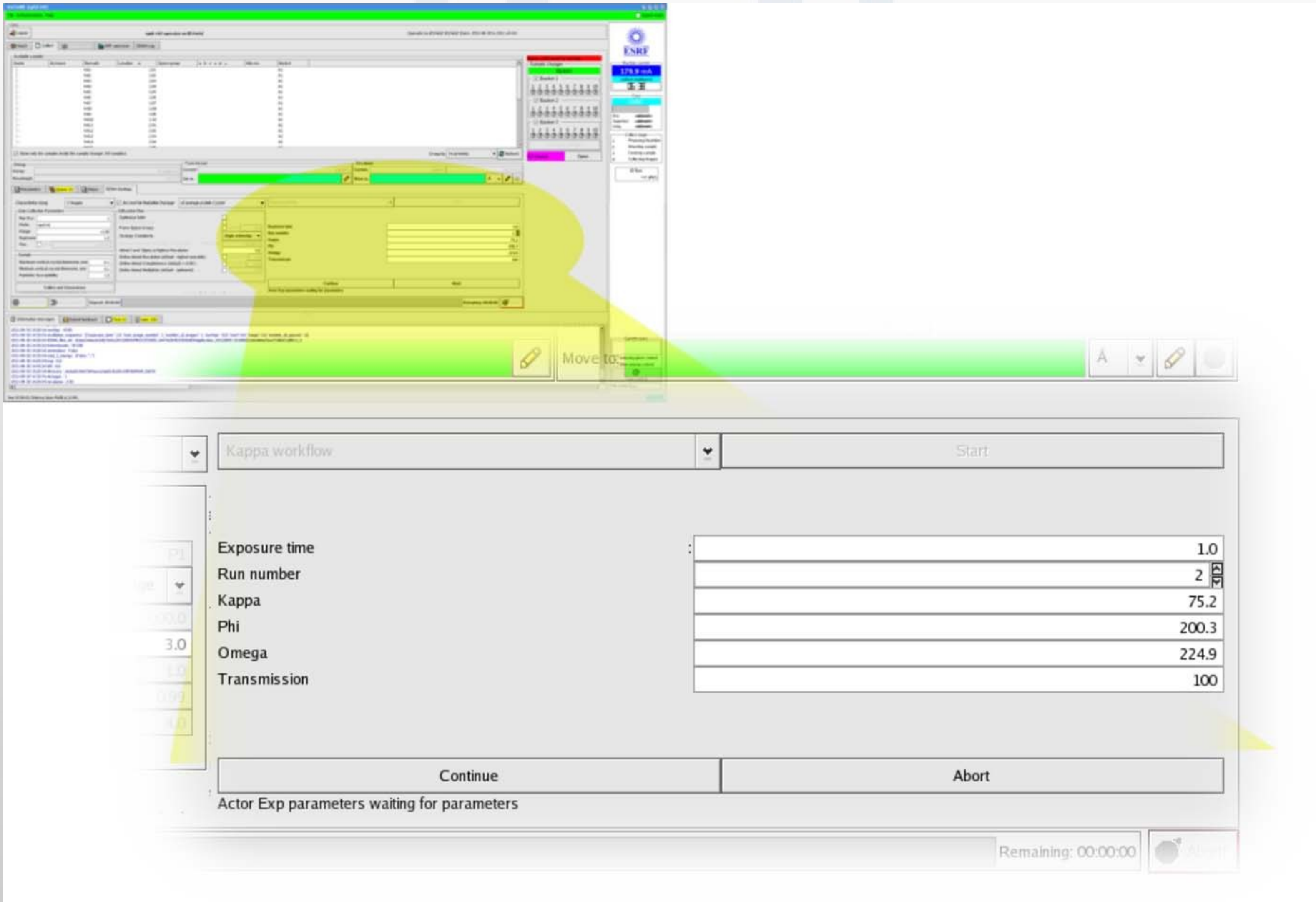
- Online without feedback to beamline control
  - Answers the question : ”Did my experiment succeed?”
  - Fast data processing – if necessary approximative
  - Examples ESRF : BM29 (BioSAXS), ID24 (EXAFS), ID15 (High speed diffraction)

- Online with feedback to beamline control
  - Strategy calculations
  - Fast data processing but no corners cut!
  - Examples ESRF : MX

Beamline integration of EDNA and the Workflow tool



Integration of the workflow tool into mxCuBE



Exposure time	1.0
Run number	2
Kappa	75.2
Phi	200.3
Omega	224.9
Transmission	100

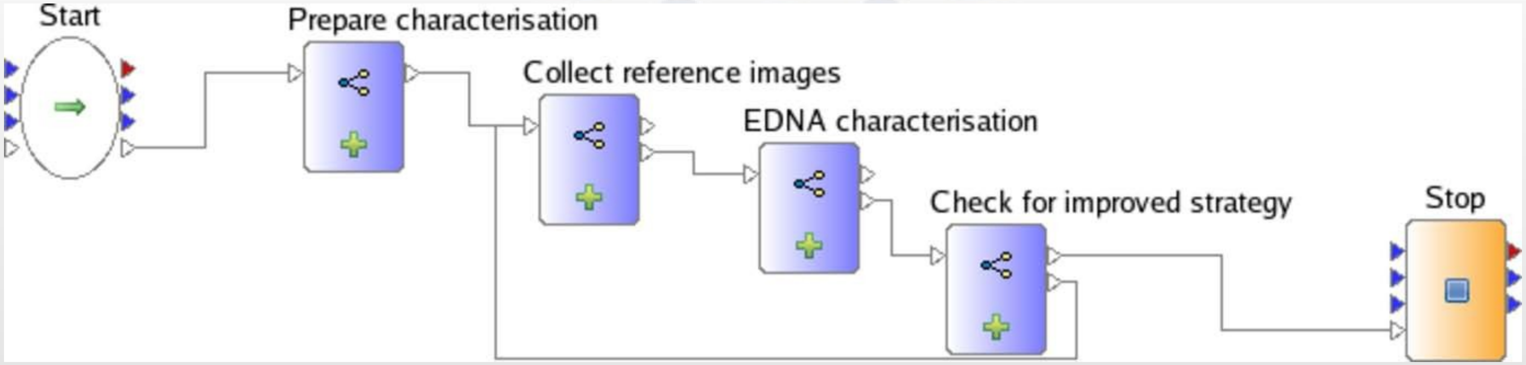
Actor Exp parameters waiting for parameters

Future work

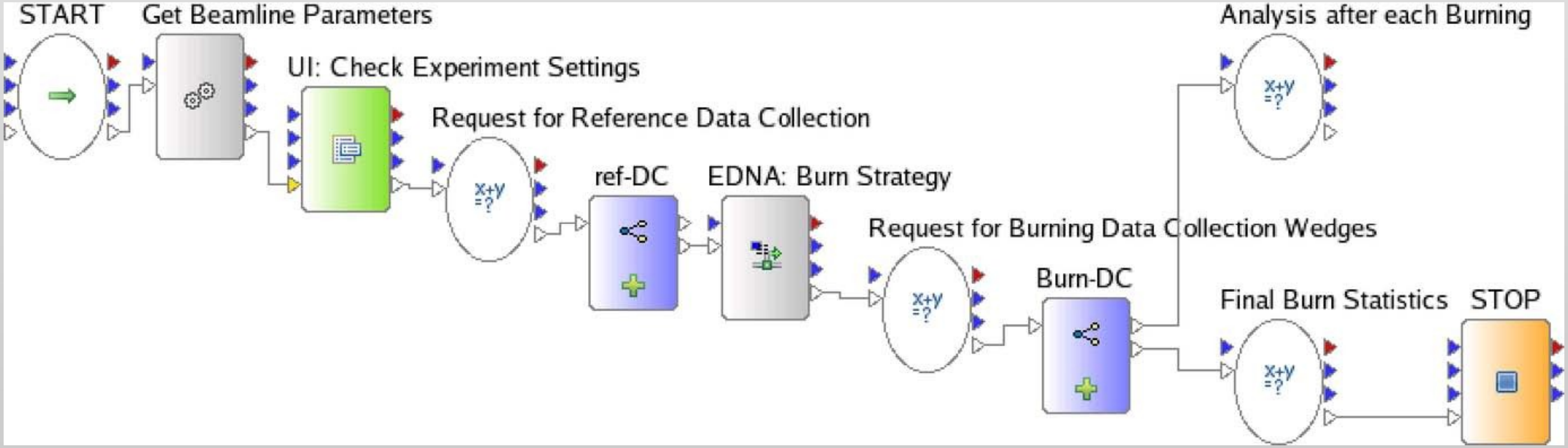
- DAWN / Passerelle :
  - Workflow-wide error handling
  - Abort
  - Debugging step execution
- MX :
  - Rapid development and deployment of more workflows
  - Improve result reports :
    - HTML pages
    - LIMS (ISPyB) integration
  - Integration into new mxCuBE

Other MX workflows

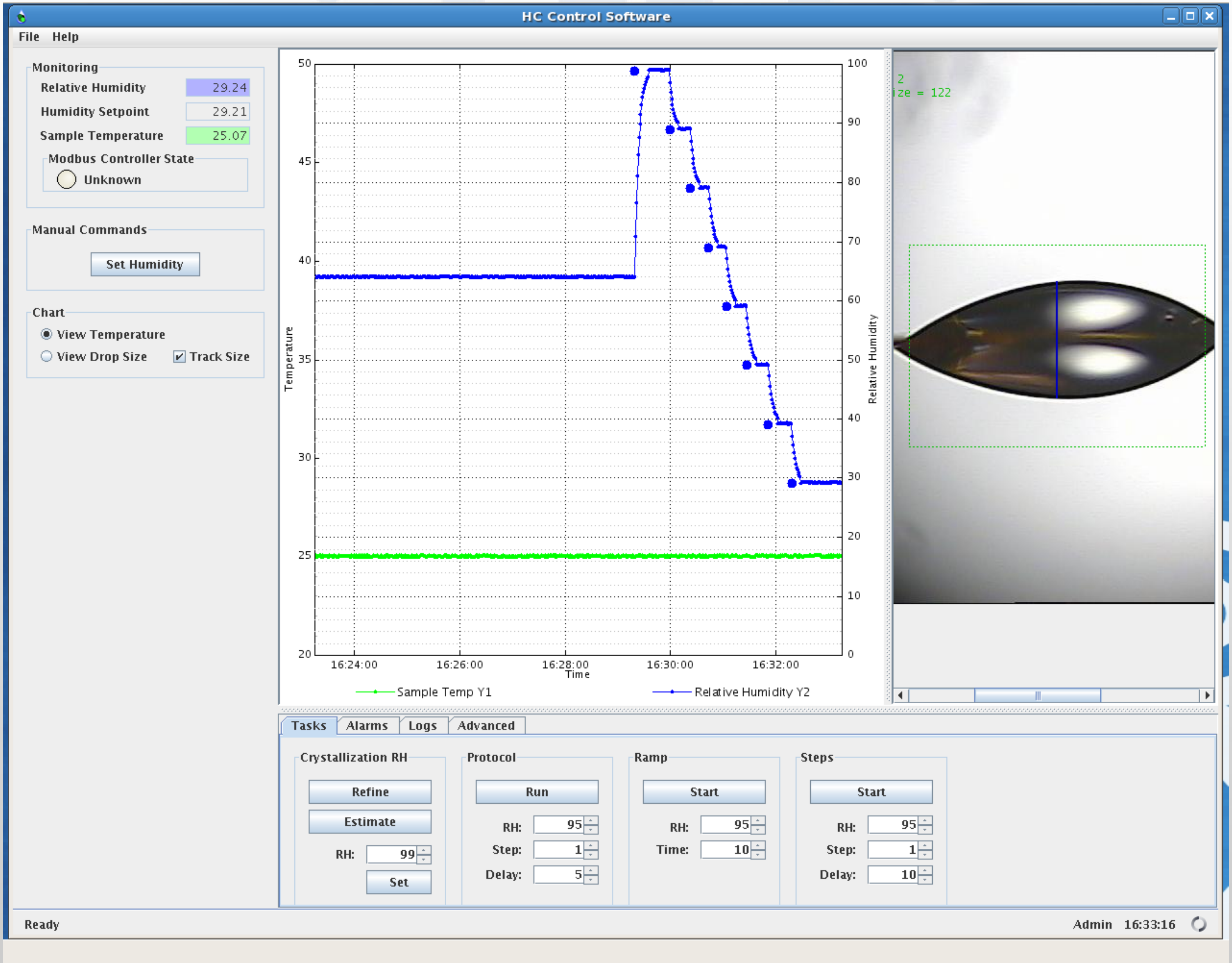
Enhanced characterisation



”Burning strategy” : Estimation of radiation damage susceptibiliity



Dehydration workflow



Relative Humidity 29.24
Humidity Setpoint 29.21
Sample Temperature 25.07

Modbus Controller State
○ Unknown

● View Temperature
○ View Drop Size ☑ Track Size

Crystallization RH
RH: 99

Protocol
RH: 95
Step: 1
Delay: 5

Ramp
RH: 95
Time: 10

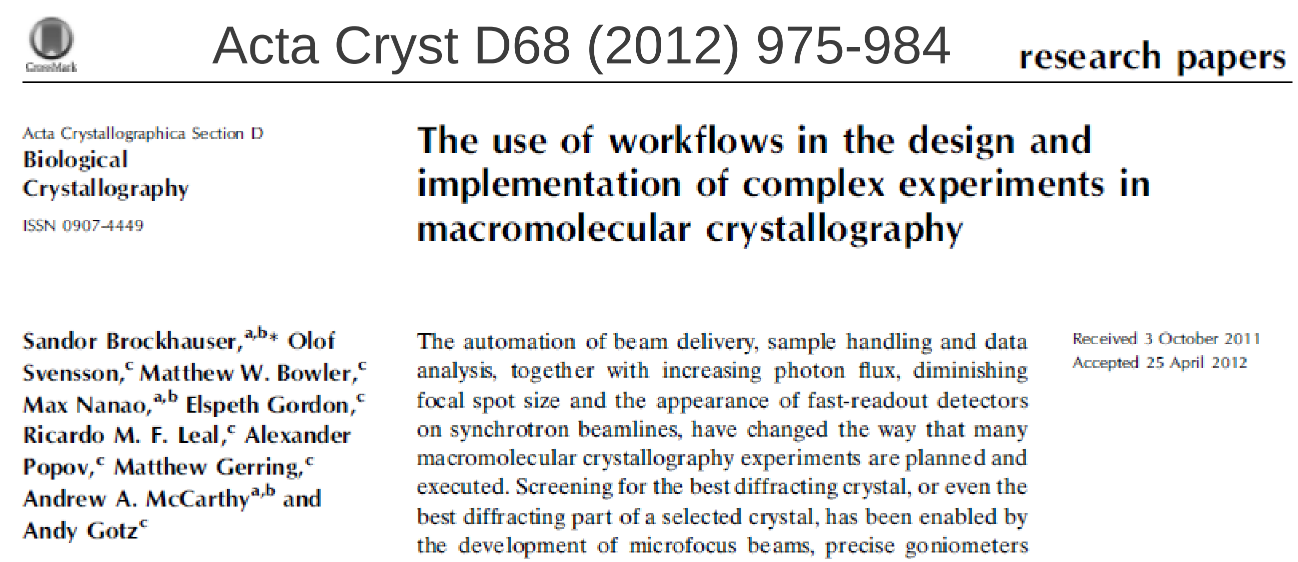
Steps
RH: 95
Step: 1
Delay: 10

Ready
Admin 16:33:16

Acknowledgements

- Elspeth Gordon and Sasha Popov (ESRF MX)
- Andy Gotz and Jonathan Pommier (ESRF software)
- The DLS Scisoft Team
- The ESRF Data Analysis Unit and Beamline Control Unit

Further reading



Acta Cryst D68 (2012) 975-984

research papers

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The use of workflows in the design and implementation of complex experiments in macromolecular crystallography

Sandor Brockhauser,[a,b]* Olof Svensson,[c] Matthew W. Bowler,[c] Max Nanao,[a,b] Elspeth Gordon,[c] Ricardo M. F. Leal,[c] Alexander Popov,[c] Matthew Gerring,[c] Andrew A. McCarthy[a,b] and Andy Gotz[c]

Received 3 October 2011
Accepted 25 April 2012

The automation of beam delivery, sample handling and data analysis, together with increasing photon flux, diminishing focal spot size and the appearance of fast-readout detectors on synchrotron beamlines, have changed the way that many macromolecular crystallography experiments are planned and executed. Screening for the best diffracting crystal, or even the best diffracting part of a selected crystal, has been enabled by the development of microfocus beams, precise goniometers