



Automated acquisition

High-resolution reconstruction

Directional local resolution

Biocomputing Unit,
Instruct Image Processing Center, CNB-CSIC
Carlos Oscar S. Sorzano

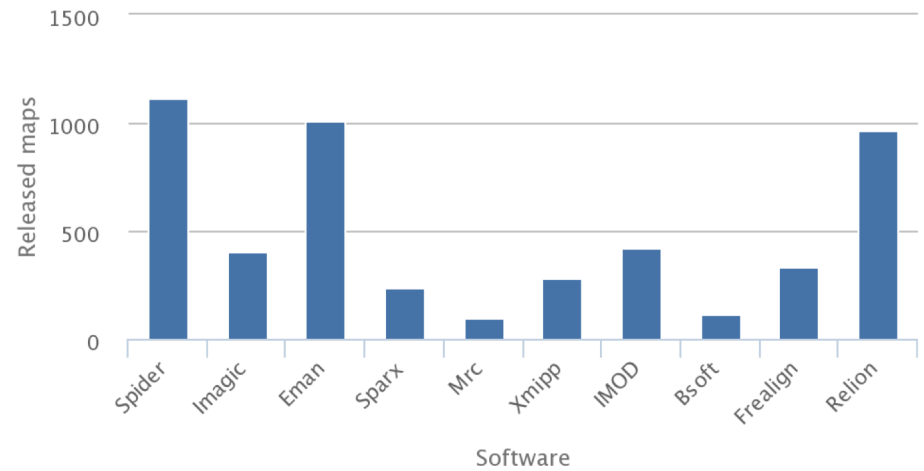


Scipion



Using different EM software packages is now like Babel's tower

Software package usage distribution



View: Protocols SPA

- Protocols SPA
 - Imports
 - import movies
 - import micrographs
 - import particles
 - import volumes
 - more
 - Micrographs
 - xmipp3 - optical alignment
 - grigoriefflab - unblur
 - grigoriefflab - summovie
 - xmipp3 - preprocess micrographs
 - CTF estimation
 - grigoriefflab - ctffind
 - xmipp3 - ctf estimation
 - more
 - Particles
 - Picking
 - eman2 - boxer
 - xmipp3 - manual-picking (step 1)
 - xmipp3 - auto-picking (step 2)
 - bsoft - particle picking
 - appion - dogpicker
 - more
 - Extract
 - Preprocess
 - Filter
 - Mask
 - 2D
 - Align
 - Classify
 - xmipp3 - c12d
 - reliion - 2D classification
 - mda
 - more
 - 3D
 - Initial volume
 - Preprocess
 - Refine
 - reliion - 3D auto-refine
 - grigoriefflab - frealign
 - xmipp3 - projection matching
 - eman2 - refine easy

Edit Copy Delete Steps Browse Db Collapse Labels
View: Tree Refresh

Summary Methods Output Log

METHODS:

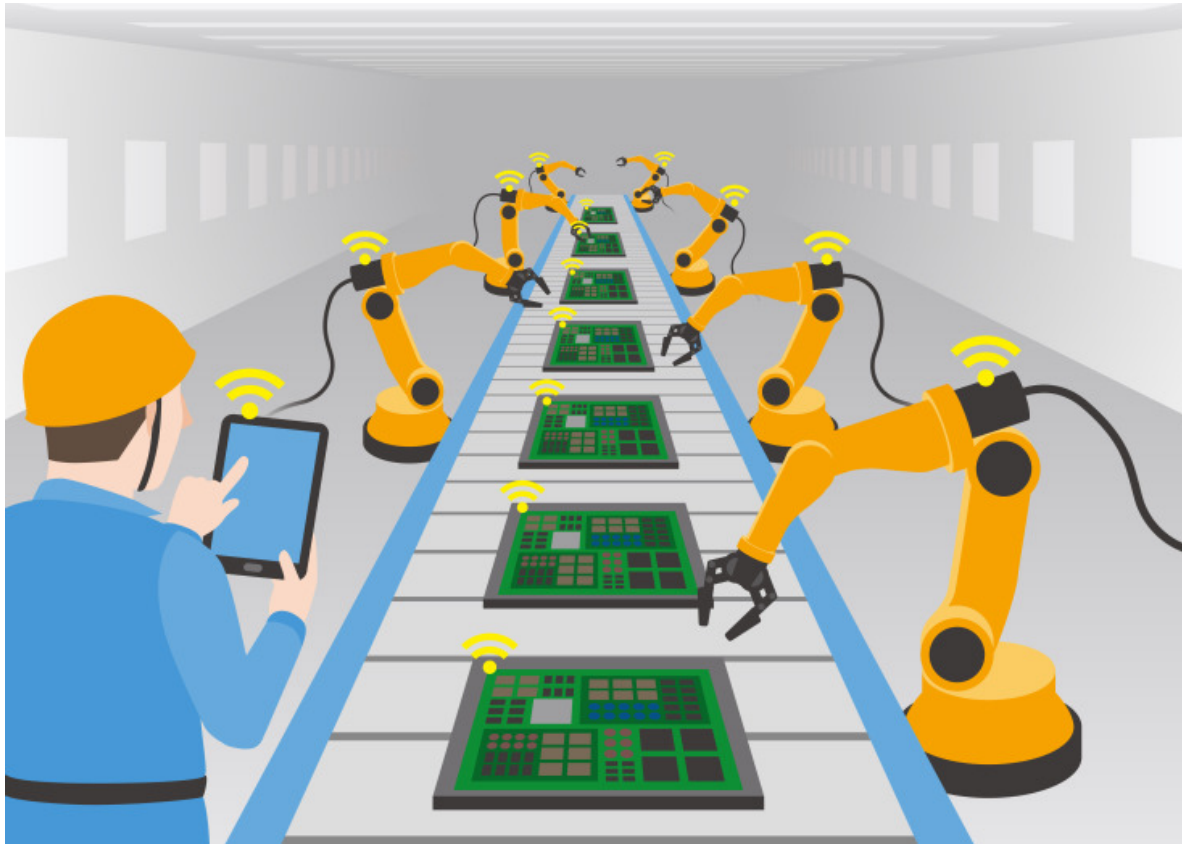
> *xmipp3 - preprocess micrographs*
 The micrographs in set [164.outputMicrographs.214](#) have
 The resulting set of micrographs is [232.outputMicrographs.278](#)

REFERENCES:

[Sorzano, et al, Proc. of IEEE Workshop on Intelligent Signal Processing, 2009](#)
[de la Rosa-Trevin, et al, JSB, 2013](#)
[Sorzano, et al, Methods in Molecular Biology, 2013](#)

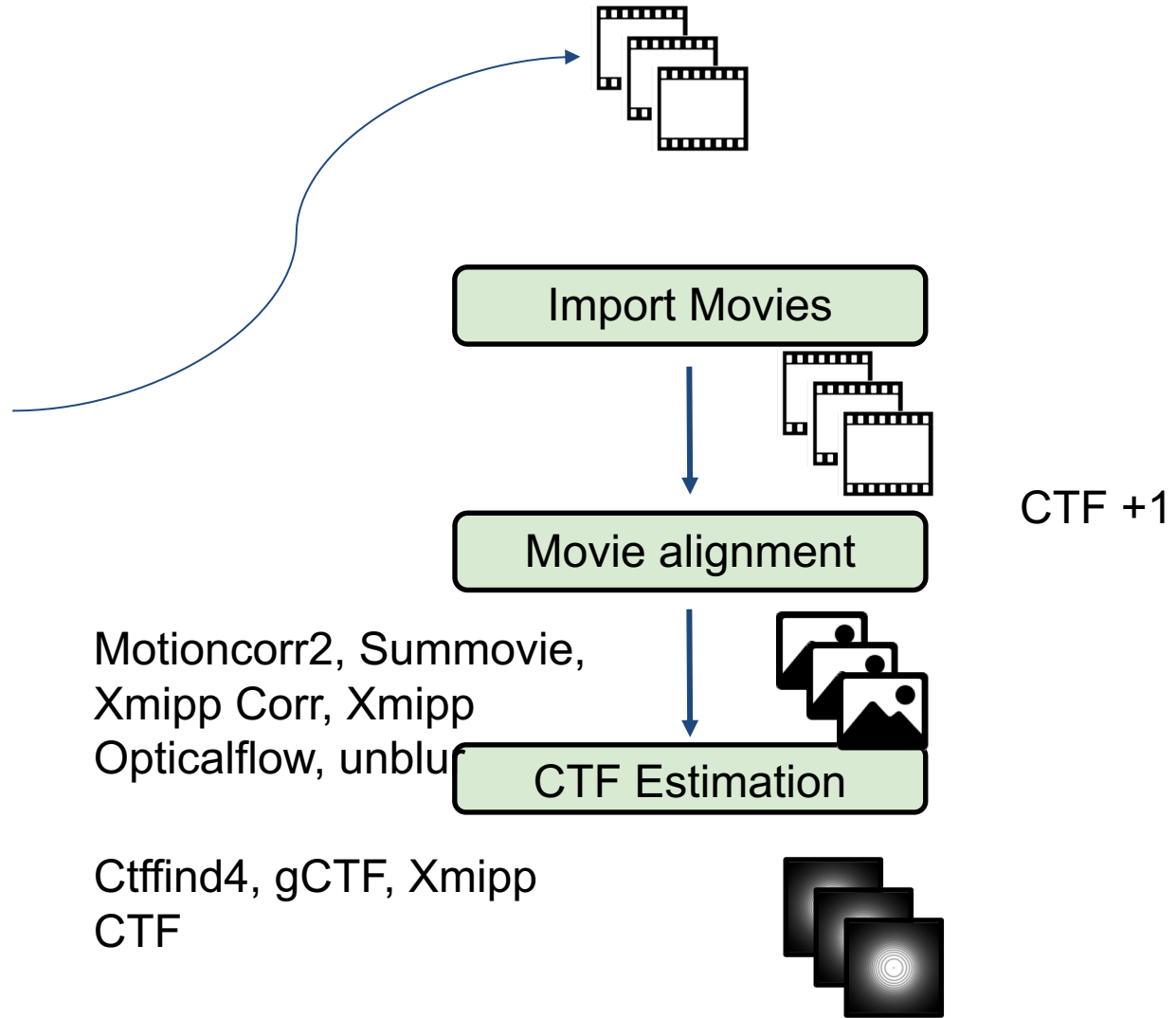
Export references

Automated acquisition

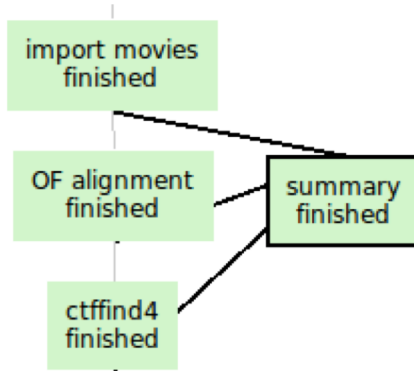


Scipion 1.1: Scipion box = Streaming = Facilities

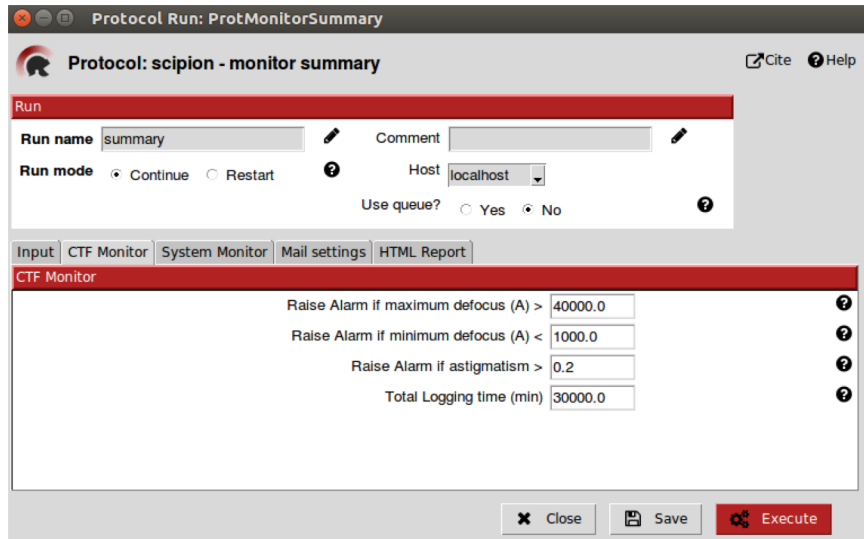
Acquisition



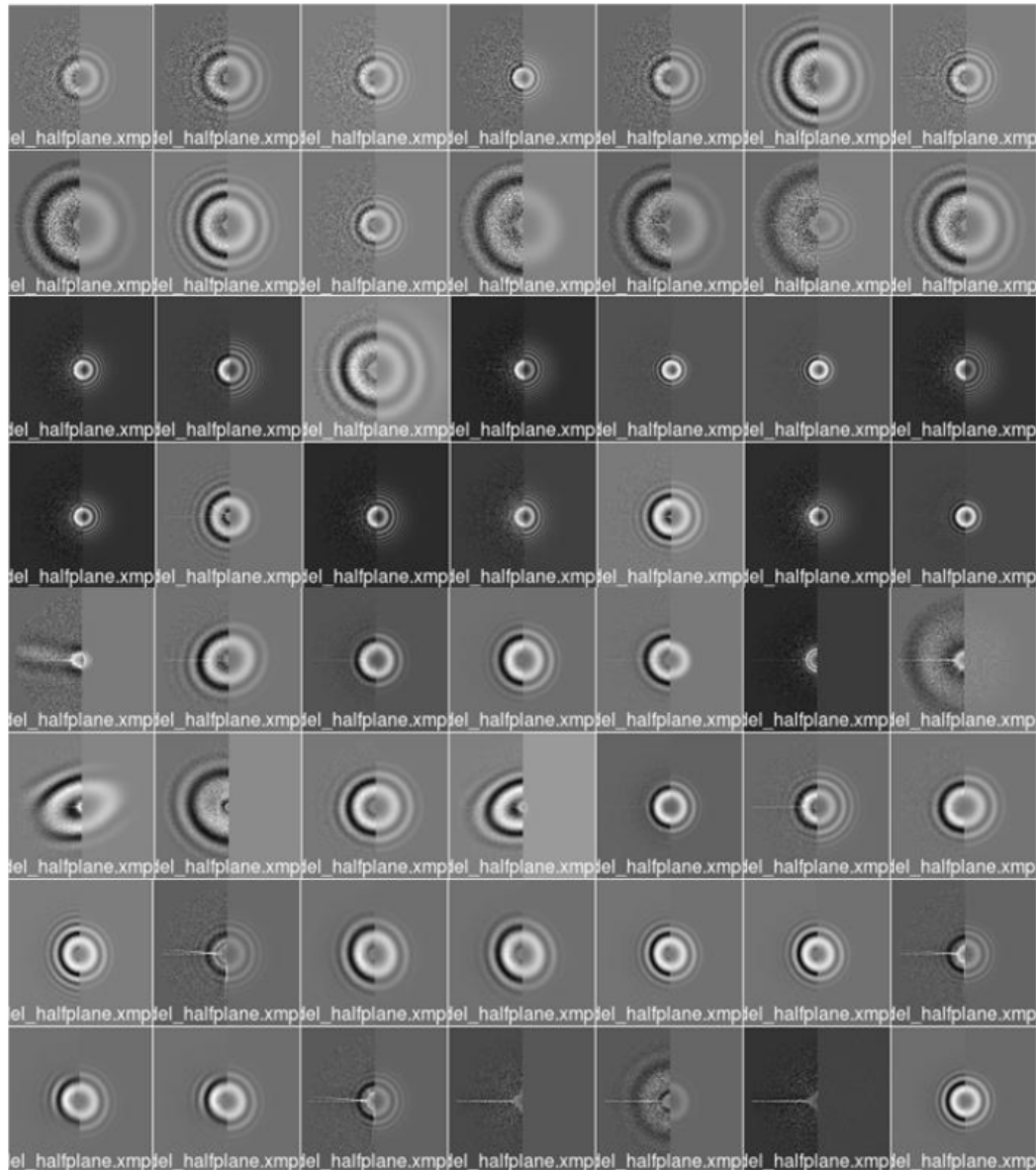
Monitoring



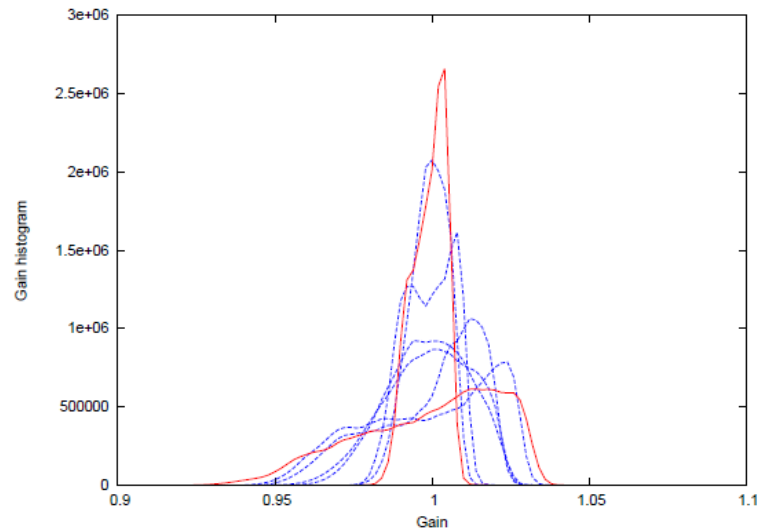
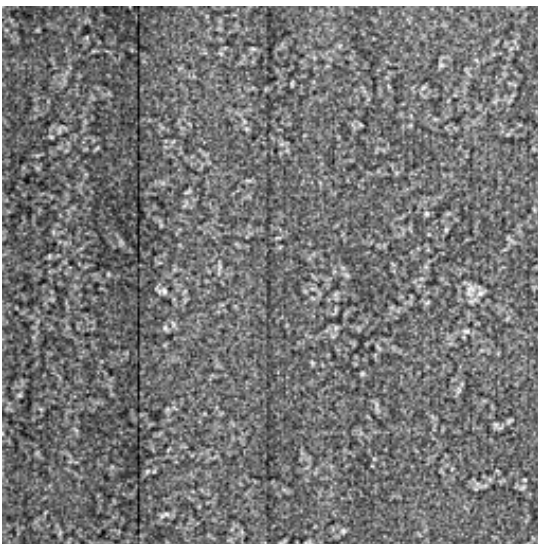
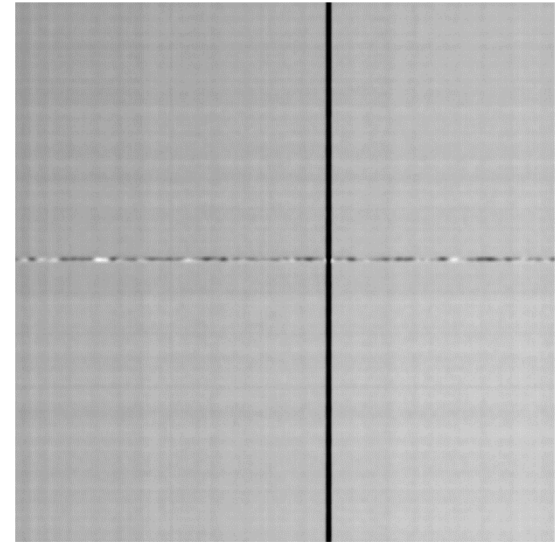
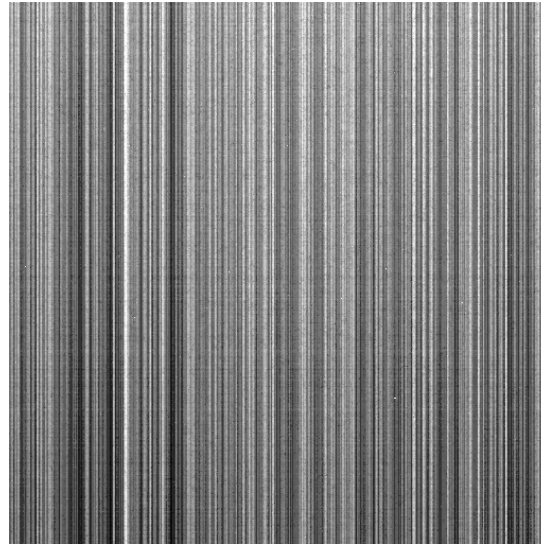
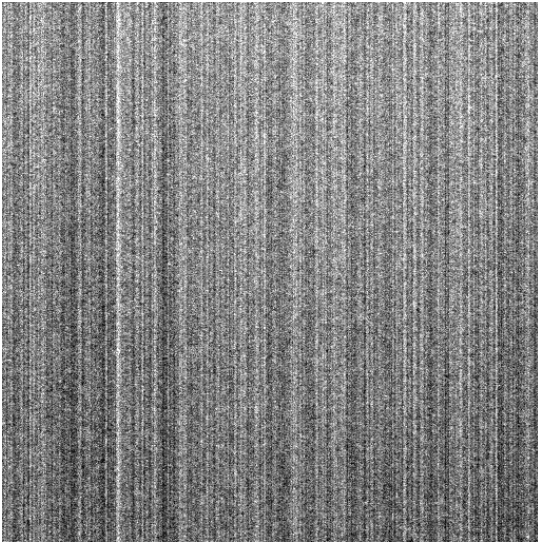
- Need feedback ASAP
- Report what has been done in the facility
- Track system status, memory, gpu, cpu, network
- Raise alarms when thresholds reached



Automatic CTF rejection



Residual gain estimate from a single movie



Scipion execution summary report

Project properties

Date: 29-07-2016 13:50:06
Project: Tails_ANA
Scipion version: scipion-box (2016-07-28) 9352cac

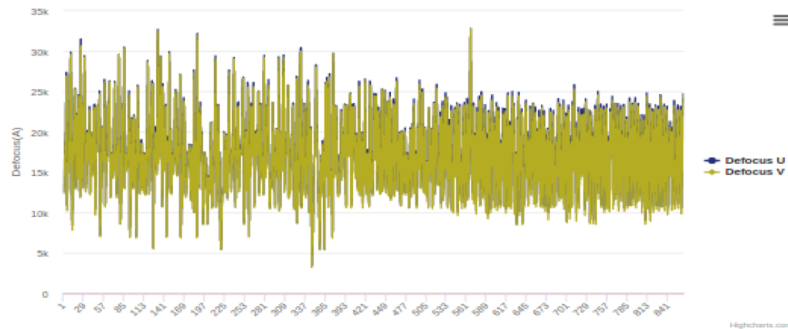
Acquisition

Microscope Voltage: 200.0
Spherical aberration: 2.7
Magnification: 73000
Pixel Size (A/px): 1.41

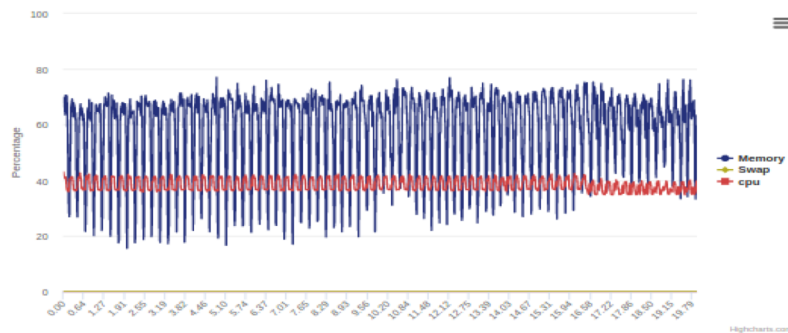
Runs summary

Name	Output	Number
Import movies (id=652)		
MotionCorr (id=696)	outputMovies	1079
xmipp3 - optical alignment (id=738)	outputMovies	869
	outputMicrographs	869
grigoriefflab - summovie (copy) (id=1436)	outputMicrographs	865
grigoriefflab - ctffind (copy) (copy) (id=1472)	outputCTF	864

CTF monitor



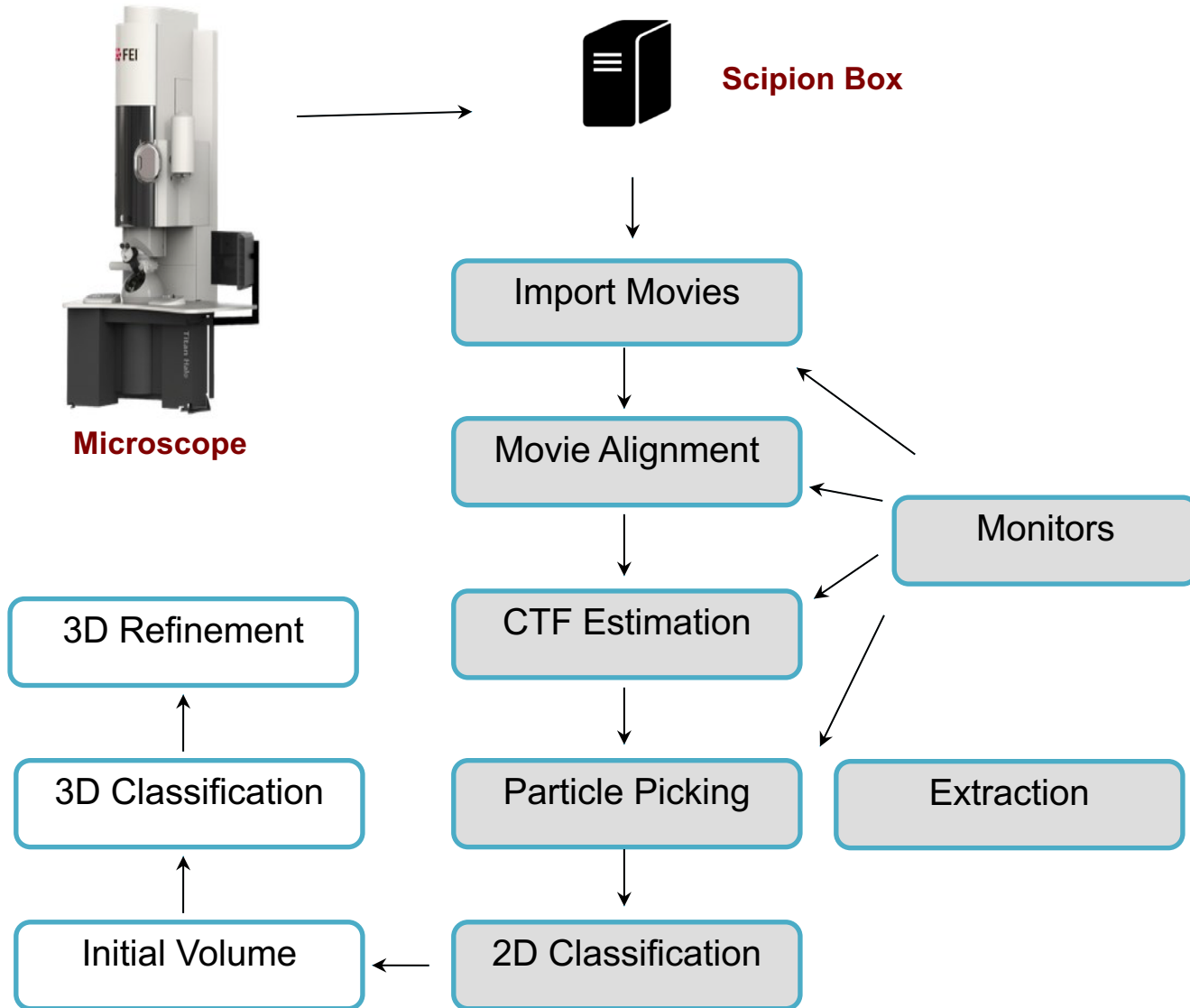
System monitor



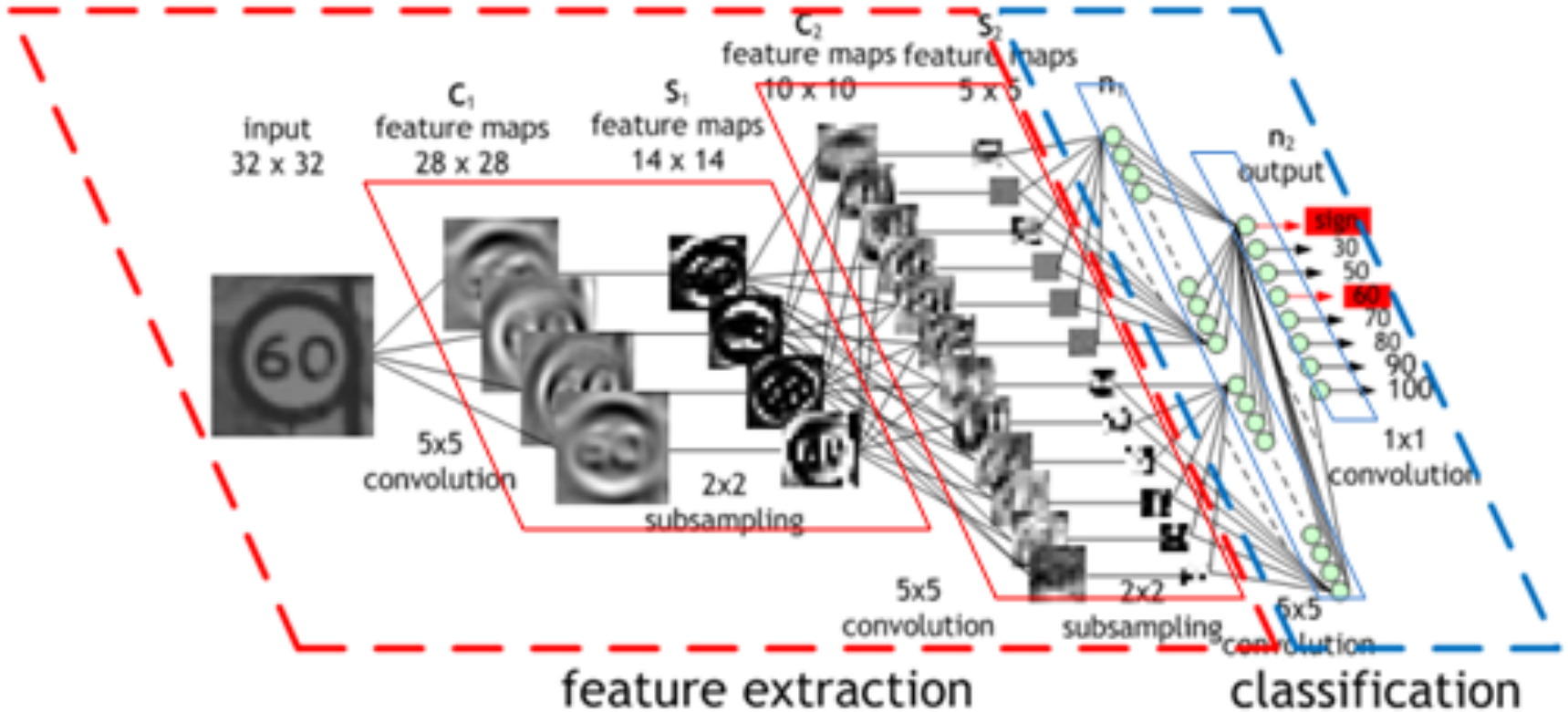
Powered by  Scipion

- Generic project info and items count (movies, ctf, micrographs)
- Defocus U and V changes
- System monitor: Memory, Swap, cpu
- HTML output and alarms

Particles Picking and Extraction



Deep consensus



False Negative Rate (Particles missed): 5-10%

False Positive Rate (False particles): 5-10%



ABOUT US

USERS & SCIENCE

INDUSTRY

EDUCATION & OUTREACH

JOBS

Home → Users & Science → Find a beamline → Structural biology → How to use our beamlines → ISPyB

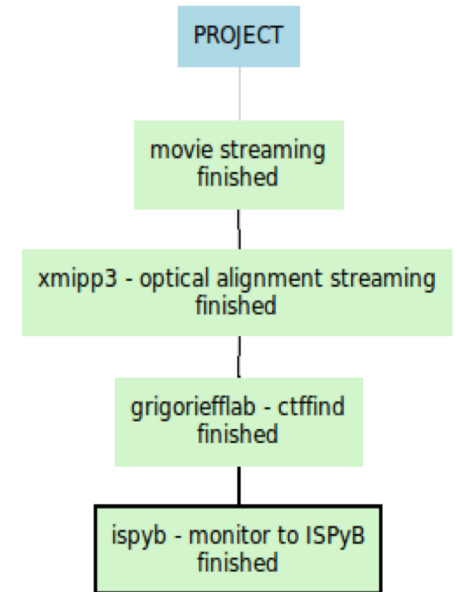
ISPYB

←	How to use our beamlines
	ISPYB
	ISPYB News
	ISPyB User Manual

ISPYB project

The ISPyB (Information System for Protein Crystallography) protein crystallography experiments on synchrotron beamlines to BioSaxs beamlines.

The ISPyB project was a joint development between ESRF/spine and PXWeb ESRF project.



Submission to public databases



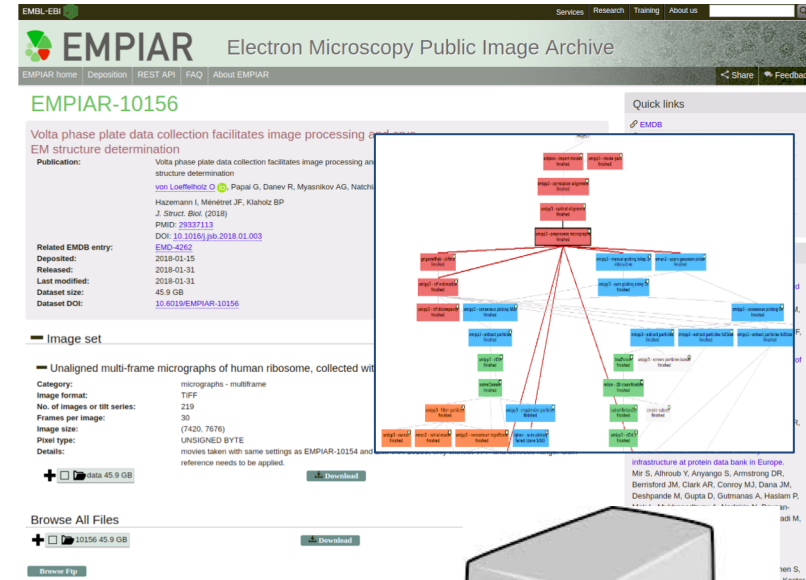
Download and reproduce



RAW data



.json



EMBL-EBI
EMPIAR Electron Microscopy Public Image Archive

EMPIAR-10156

Quick links
EMDB

Volta phase plate data collection facilitates image processing and structure determination

Publication:
von Loebenstein G, Pappal G, Danev R, Myasnikov AG, Natchez H, Hazemann L, Menzel JF, Klaholz BP, J. Struct. Biol. (2018)
PMID: 29337113
DOI: 10.1016/j.jmb.2018.01.003
EMD-4262

Related EMDB entry:
Deposited: 2018-01-15
Released: 2018-01-31
Last modified: 2018-01-31
Dataset size: 45.9 GB
Dataset DOI: 10.6019/EMPIAR-10156

Image set
Unaligned multi-frame micrographs of human ribosome, collected with Volta phase plate

Category: micrographs - multiframe
Image format: TIFF
No. of images or tilt series: 219
Frames per image: 30
Image size: (7420, 7670)
Pixel type: UNSIGNED BYTE
Details: moves taken with same settings as EMPIAR-10154 and reference needs to be applied.

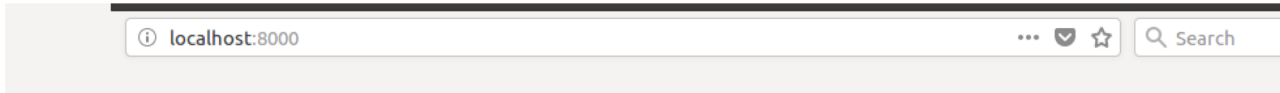
Download

Browse All Files
EMPIAR-10156 45.9 GB
Download

Browse File

infrastructure at protein data bank in Europe
Mr S. Alkhouf Y, Arayag S, Armstrong GR, Berrisford JM, Clark AR, Conroy MJ, Dana JM, Deshpande M, Gupta D, Gutmanas A, Haslam P, ...
ad M.
ven S, Kovdar

Workflow visualization



Scipion workflow viewer

Vestibulum id tincidunt Scipion. In lobortis facilis dignissim. Vestibulum nec pulvinar urna. Fusce condimentum sed tortor a consequat. Etiam id lacinia urna. Quisque id tempor metus, at tristique est. Donec at mollis lectus, quis vestibulum turpis. Integer vehicula sapien libero, lobortis rutrum diam suscipit eu.



Github:
<https://github.com/l2PC/web-workflow-viewer>

- Using webcomponents
- Easy to incorporate
- Already in use in our Facility

FAIR Data: Findable, Accessible, Interoperable, Reusable

The screenshot shows the FAIRsharing.org website. At the top left is the logo "FAIRsharing.org" with the tagline "standards, databases, policies". To the right is a search bar with the text "Search all of FAIRsharing" and a magnifying glass icon. Further right are navigation buttons for "Standards", "Databases", "Policies", "Collections", "Add/Claim Content", "Stats", and "Log in or Register".

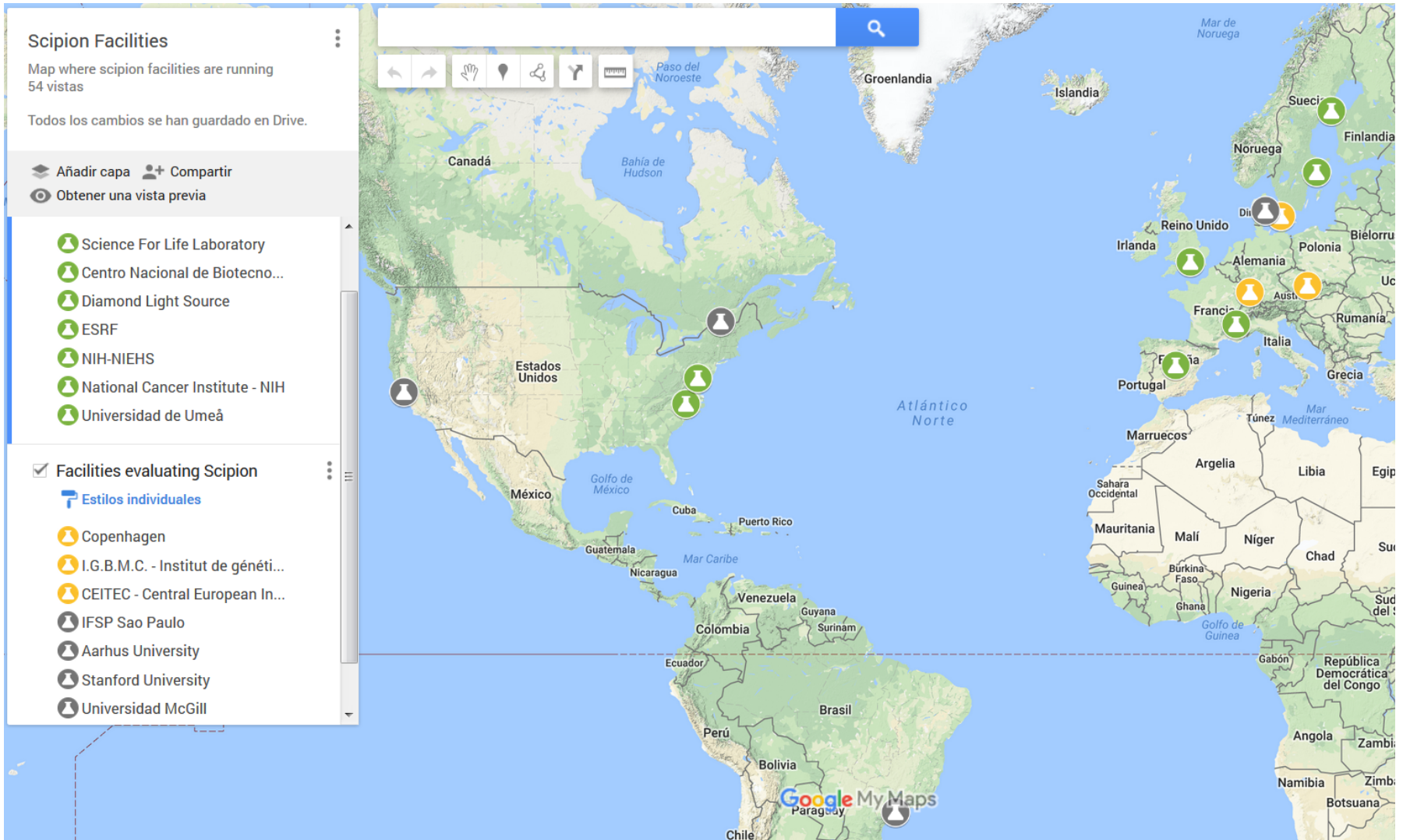
A yellow notification box contains the text: "FAIRsharing is here! From our first incarnation, BioSharing.org, which focussed on the life sciences, we are growing into FAIRsharing.org, to serve users across all disciplines and support Findable, Accessible, Interoperable and Reusable (FAIR) data." with a close button (X).

The "Standards" section features a dark blue header with the text "Standards" on the left and two buttons on the right: "Contribute by adding a standard" (light blue) and "Any problems? Please tell us!" (red). Below the header is a paragraph: "The standards in FAIRsharing are manually curated from a variety of sources, including [BioPortal](#), [MIBBI](#) and the [Equator Network](#)." To the left of this text are three icons: a document with a checkmark, a network diagram, and a grid with an arrow.

Below the paragraph is a search bar for standards. It contains the text "Search Standards" and a search input field with "emx" entered. To the right of the input field are three buttons: "Search" (light blue), "Reset" (white), and "Advanced" (green).

At the bottom of the search results area, it says "Showing records 1 - 50 of 1069." The number "1" in the pagination list is circled in red. The pagination list includes numbers 1 through 22, with a left arrow and a right arrow.





HighRes



HighRes philosophy

Global alignment until stable:

- You need only resolutions between 15-10 Å to correctly align
- Significant alignment

Then, local alignment

- as accurate as you can, refine:
 - Angles, shifts
 - Defocus
 - Scale (anisotropic)
 - Gray values

Remove noise “anchors”/unsignificant features:

- Reference volume should contain only significant information
- Output volume only with significant features

Multiresolution optimization:

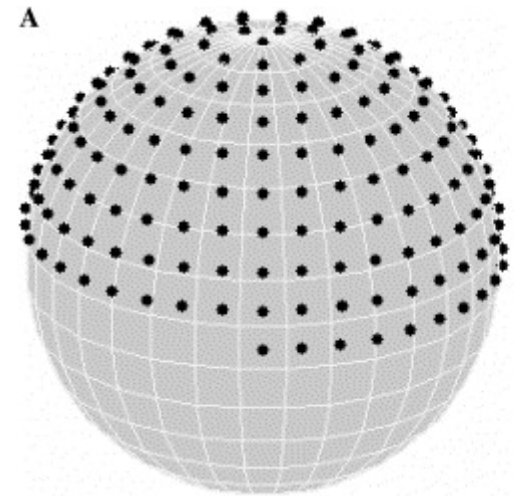
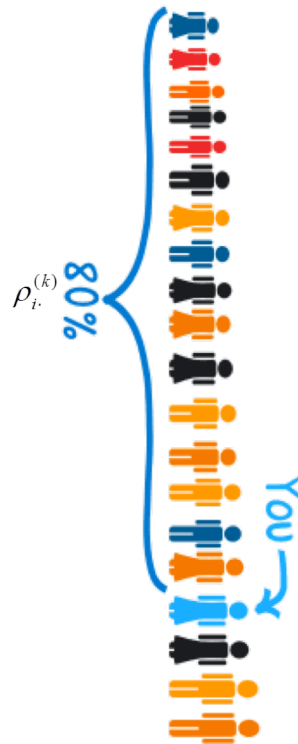
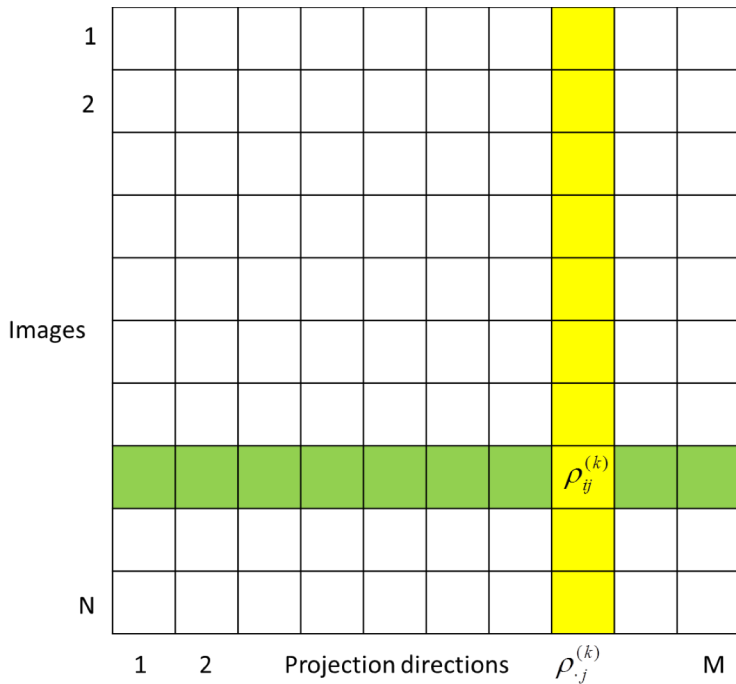
- Increase speed
- Smooth solution landscape
- Slow annealing to 5-7 Å



Significant Reconstruction

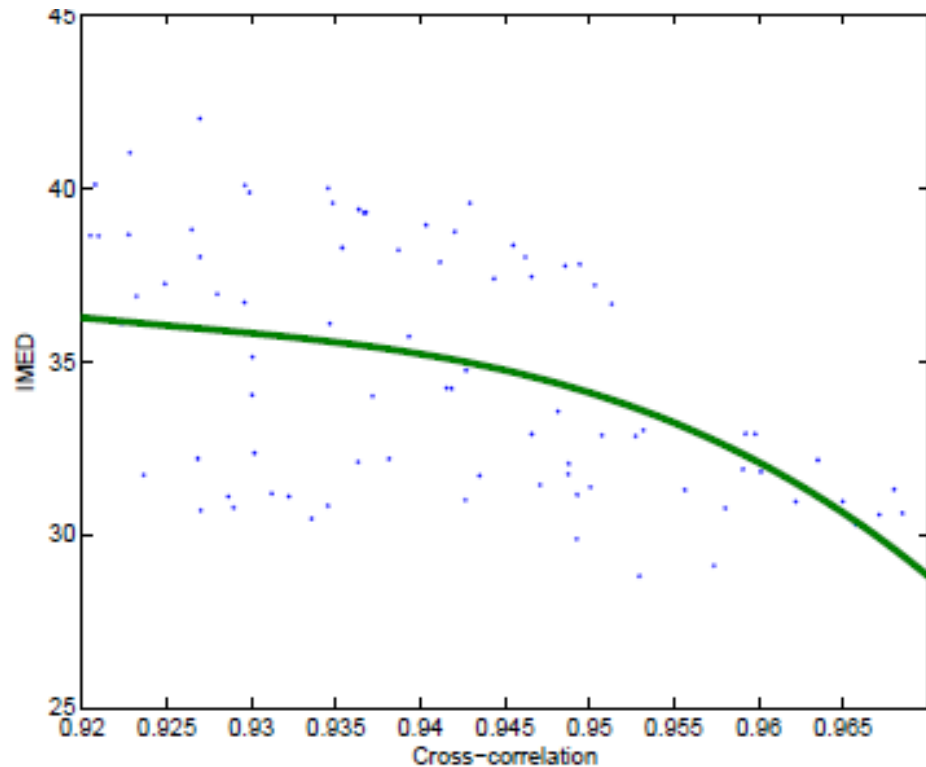
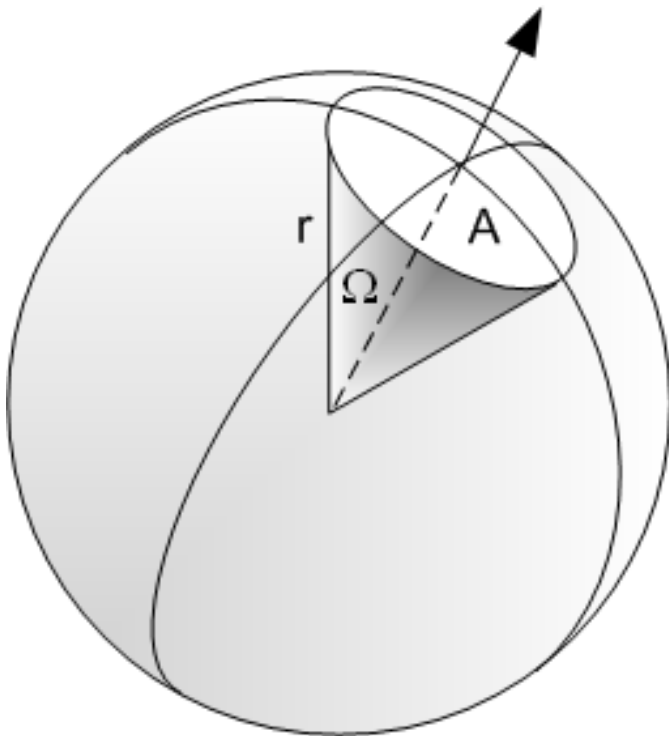


$$\rho \in \left[\tanh(\tanh^{-1}(\max_j \{\rho_{ij}^{(k)}\})) - \frac{z_{1-\alpha^{(k)}}}{\sqrt{N-3}}, \max_j \{\rho_{ij}^{(k)}\} \right]$$



Significant Reconstruction

$$w_{ij}^{(k)} = \left\{ \begin{array}{l} \frac{\min_{\substack{i' \in 1, \dots, N \\ j' \in \text{Neigh}_\theta(j)}} \eta_{i'j'}^{(k)}}{\eta_{ij}^{(k)}} \Pr \left\{ \eta_{i \cdot}^{(k)} \geq \eta_{ij}^{(k)} \right\} \Pr \left\{ \eta_{\cdot j}^{(k)} \geq \eta_{ij}^{(k)} \right\} \\ \frac{\rho_{ij}^{(k)}}{\max_{\substack{i' \in 1, \dots, N \\ j' \in \text{Neigh}_\theta(j)}} \rho_{i'j'}^{(k)}} \Pr \left\{ \rho_{i \cdot}^{(k)} \leq \rho_{ij}^{(k)} \right\} \Pr \left\{ \rho_{\cdot j}^{(k)} \leq \rho_{ij}^{(k)} \right\} \end{array} \right\}$$



Intuitively

- **Conventionally:** assign **ONE** orientation/classification, based on maximum CC.
- **Expectation-Maximization:** calculate the *probability-weighted average over ALL* possible assignments

So instead of choosing one of two very similar options (in terms of CC), both options are considered with similar weights

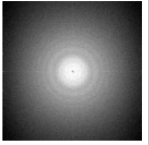
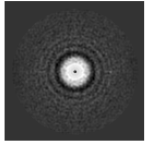
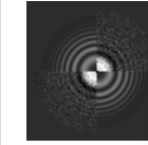
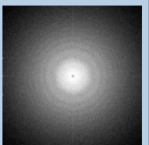
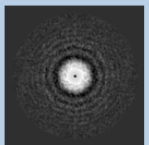
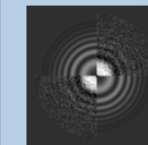
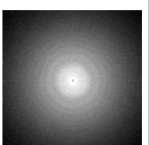
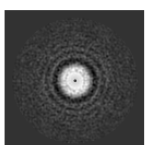
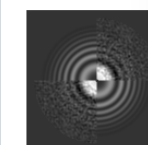
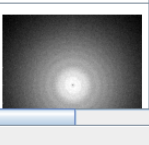
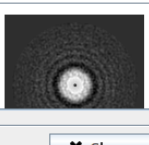
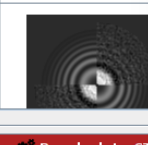
**EVERYONE
DESERVES
A SECOND
CHANCE
BUT NOT FOR
THE SAME
MISTAKE**

Envelope correction

Metadata: ctf.sqlite 15 items (256 x 256)

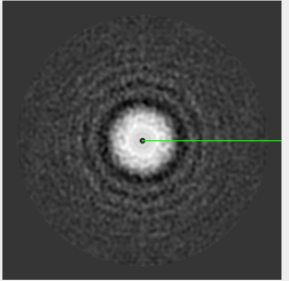
File Display Tools Help

Block CTF

id	enabled	comment	_psdFile	_xmipp_enhanced_psd	_xmipp_ctfmodel_quadrant	_xmipp_ctfmodel_ha
1	<input checked="" type="checkbox"/>					
2	<input checked="" type="checkbox"/>					
3	<input checked="" type="checkbox"/>					
4	<input checked="" type="checkbox"/>					

CTF Analyzer

PSD Image

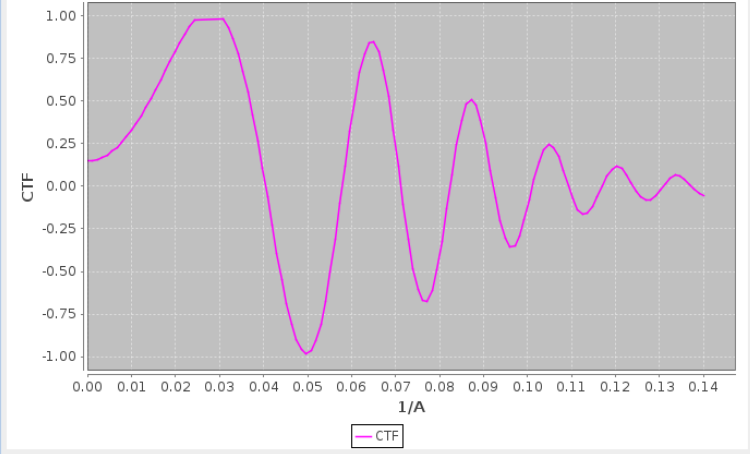


Display

- CTF
- Log(PSD)
 - Log(Theoretical Log(PSD))
 - Log(Envelope)
 - Log(BGNoise)
 - Log(PSD) - Log(BGNoise)

Export Graphics

Radial Profile Radial Average

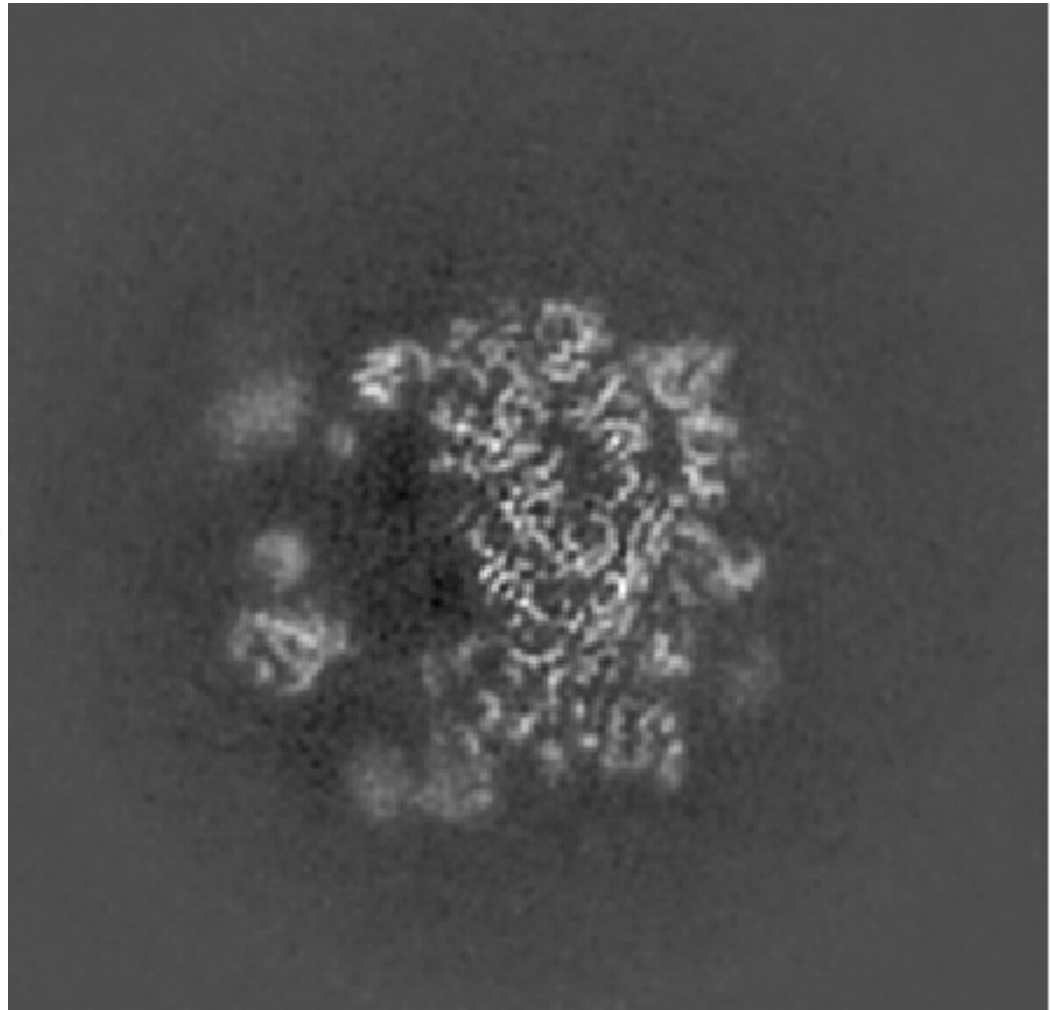
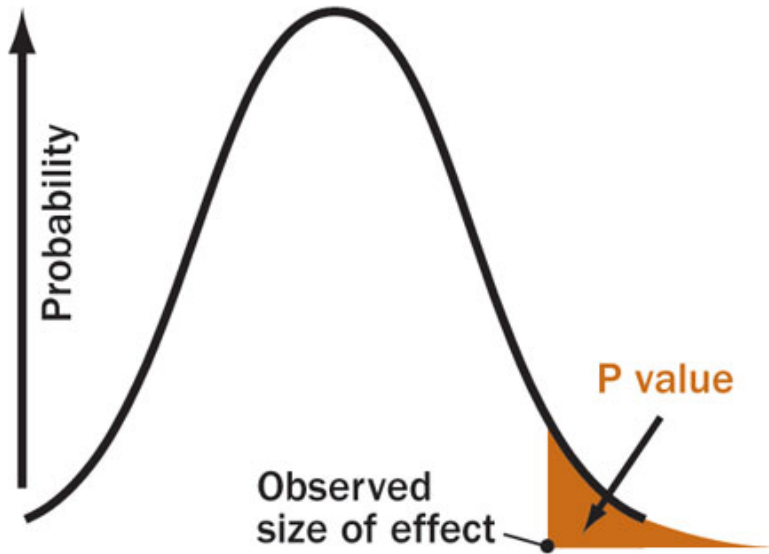


CTF

1/A

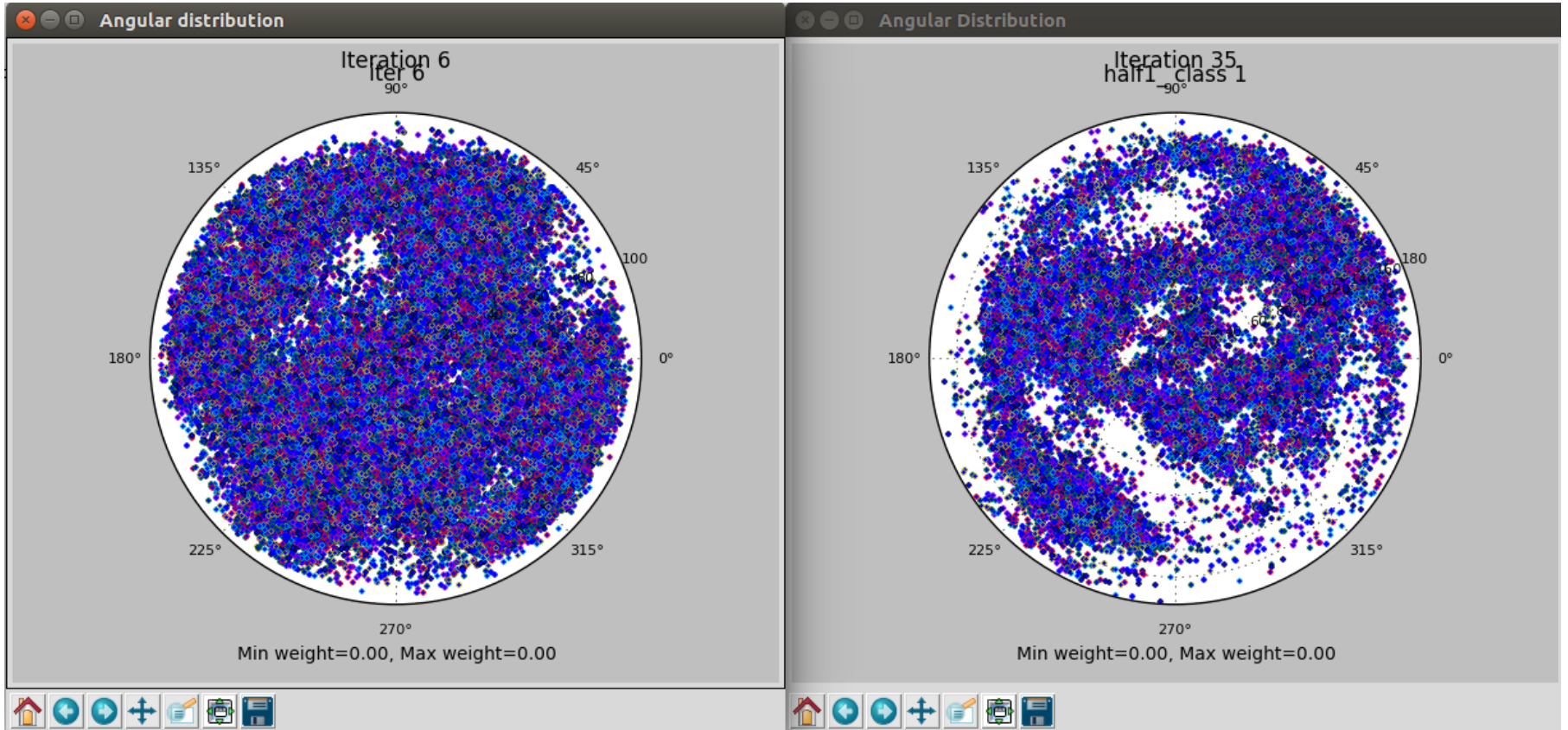
Close Recalculate CTFs Micrographs

Significant features

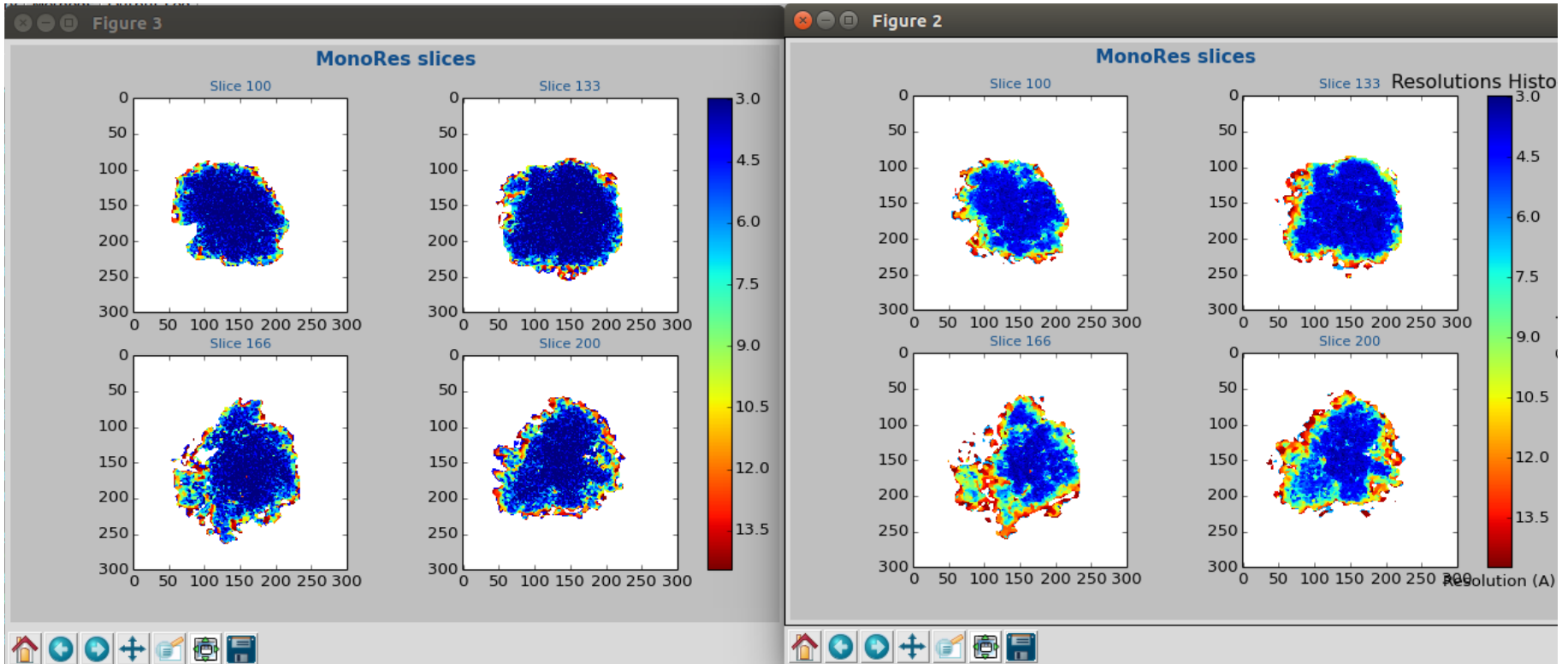


**A Researcher's
(Misguided)
Notion of Nirvana:
 $p < .05$**

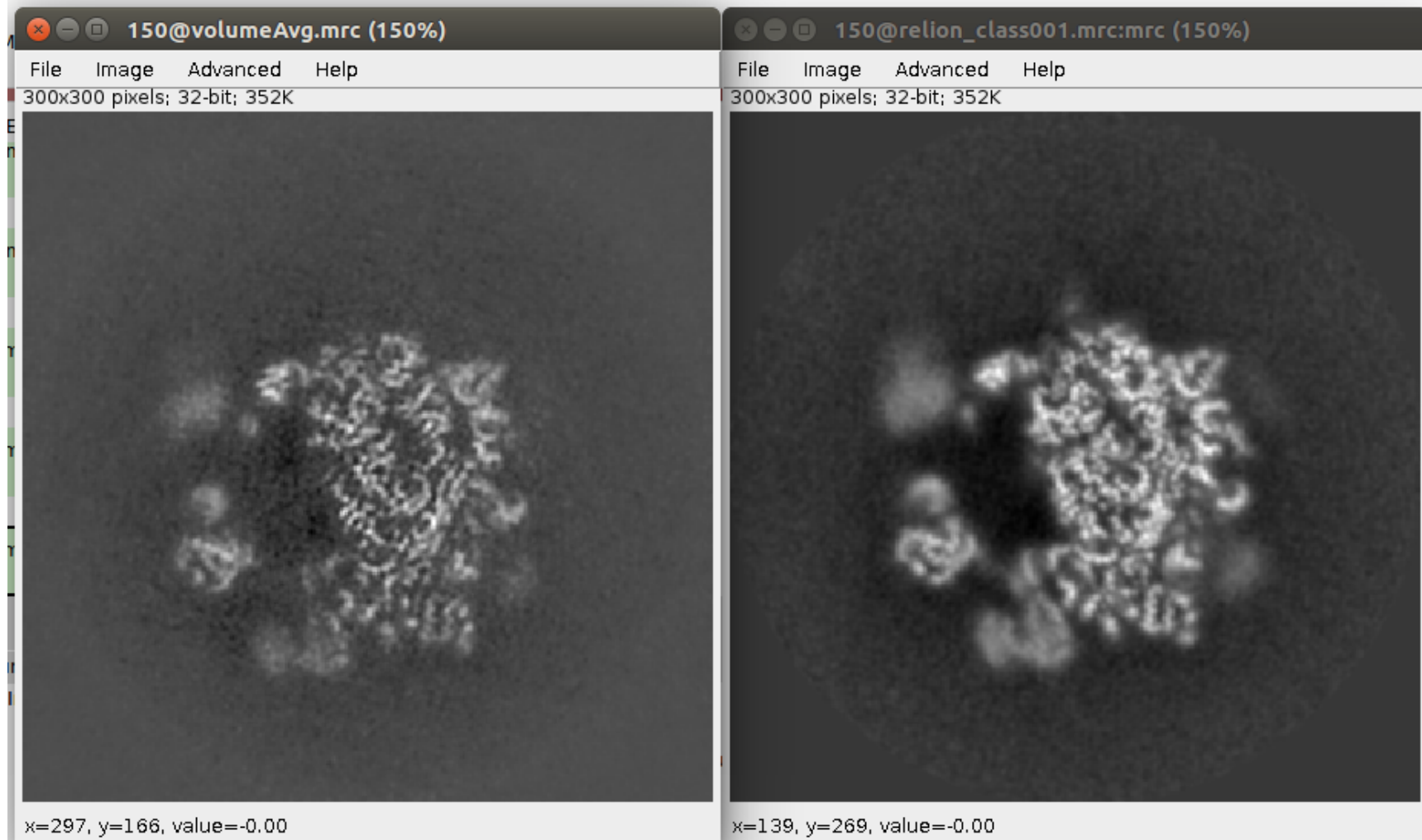
Results



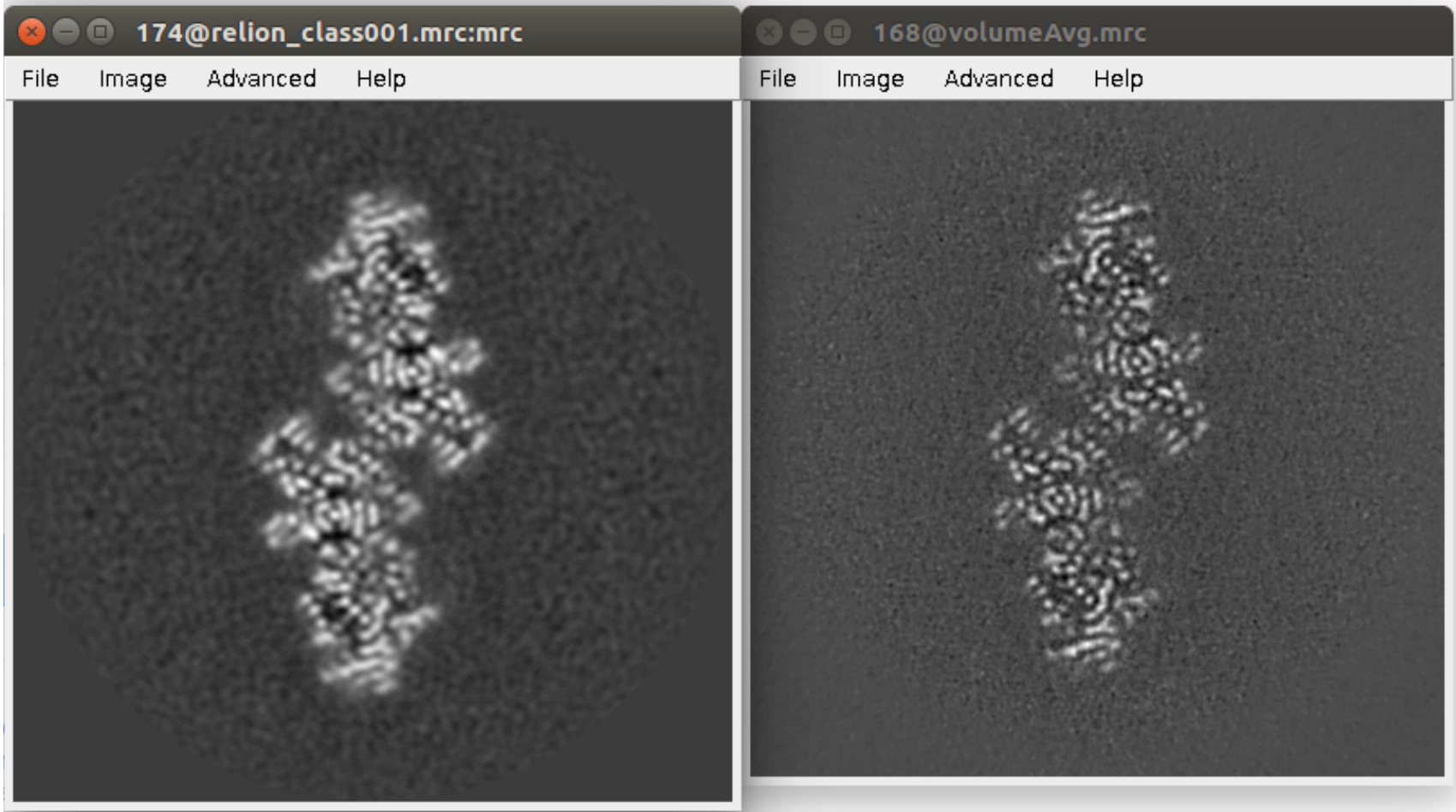
Results



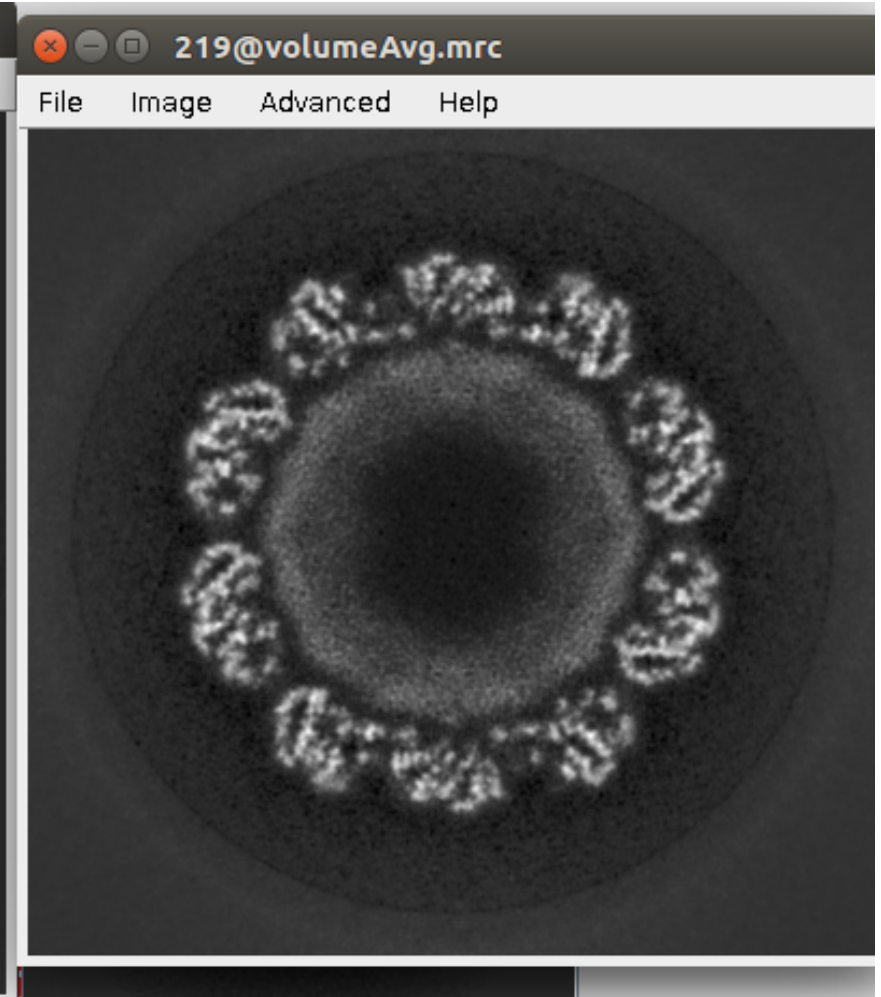
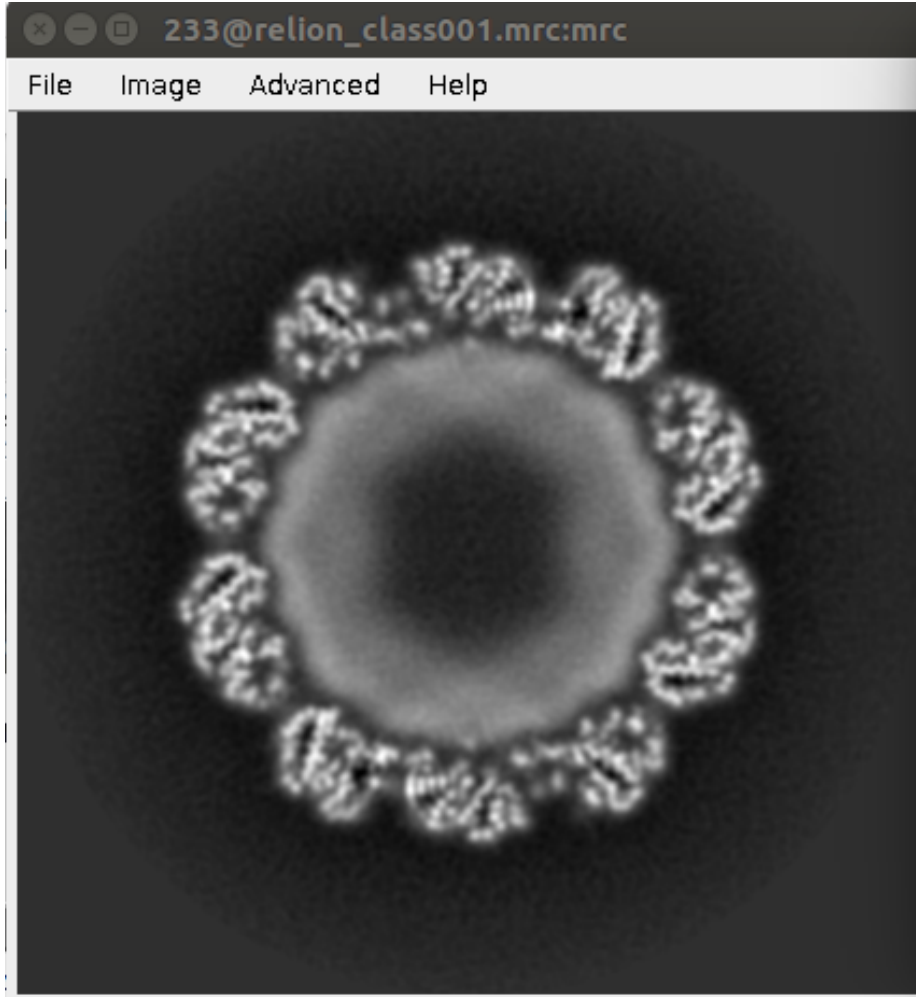
Results Ribosome



Results B-Galactosidase



Results Virus



Results X



Colorful Image Processing

xmipp3 - highres finished

Clean3D (70k coords)

xmipp3 - highres 2 finished

xmipp3 - highres 3 finished

highresSubset3 finished

xmipp3 - highres 4 finished

highresSubset4 finished

xmipp3 - highres 5 finished

highresSubset5 finished

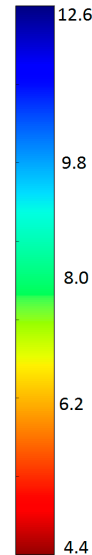
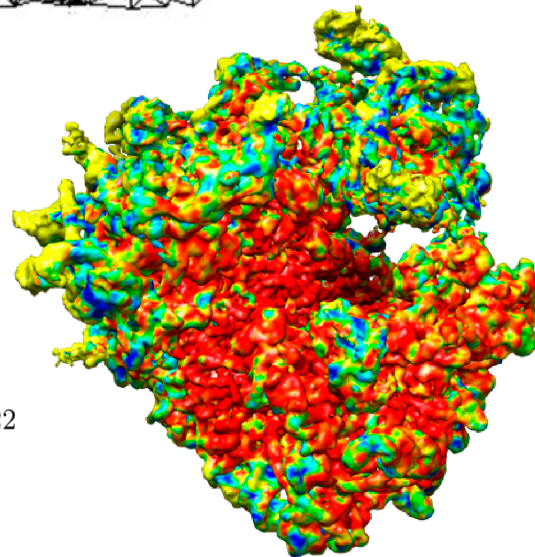
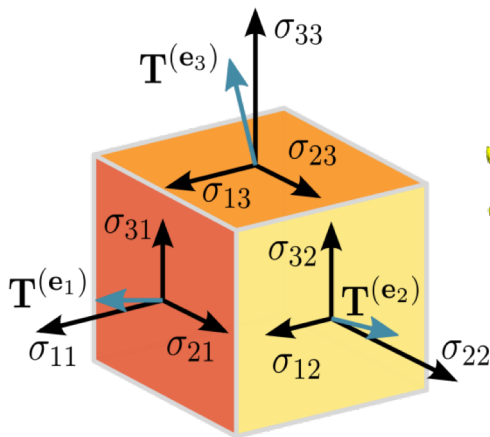
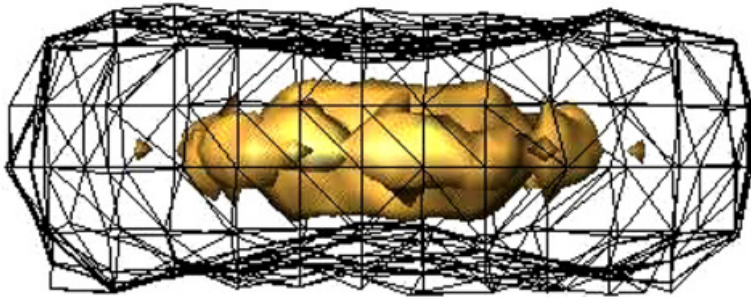
xmipp3 - highres 6 finished

xmipp3 - highres 7 finished

The screenshot shows the SCIPION software interface. The top bar indicates the user is 'co_laplacian*' and the session ID is '(2017-08-25) c648a63'. The left sidebar lists various processing steps, with 'xmipp3 - resolution 3D finished' and 'xmipp3 - filter volumes (finished)' highlighted. The main workspace contains three windows displaying 3D volumes. The top window shows a noisy volume with a coordinate readout of 'x=311, y=309, value=-0.04'. The bottom-left window shows a cleaner volume with a coordinate readout of 'x=331, y=1, value=-0.00'. The bottom-right window shows another clean volume with a coordinate readout of 'x=194, y=158, value=-0.01'. A green oval highlights the text 'They all report 2.2Å!!' overlaid on the middle window.

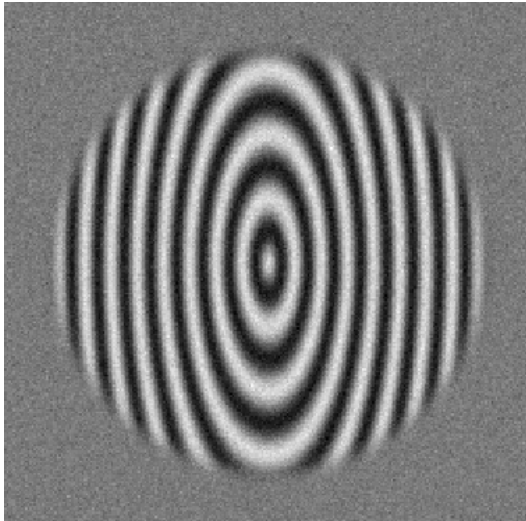
Clean3D (60k coords)

Resolution is local and directional

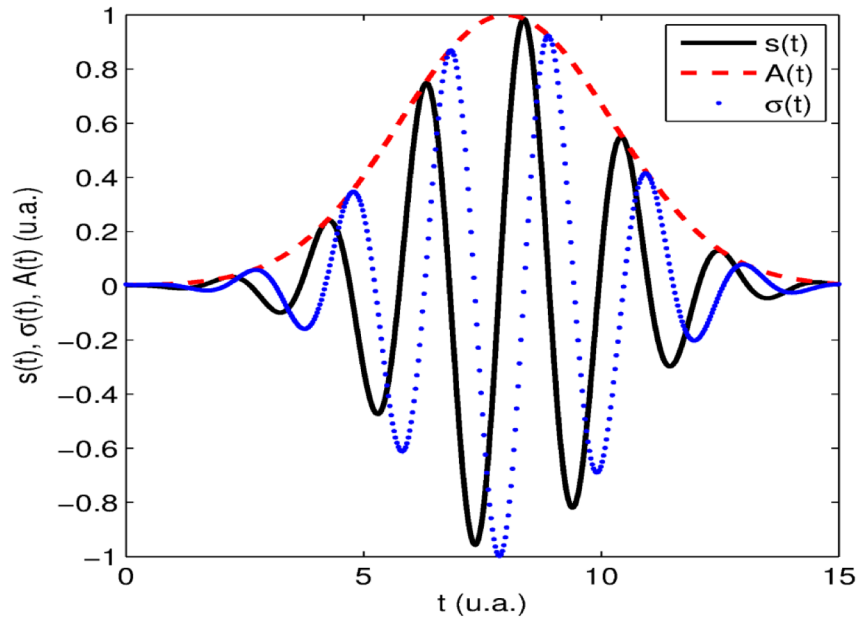


Resolution is local and directional

- An anisotropic fringe pattern presents different resolutions (wavelength) along different directions



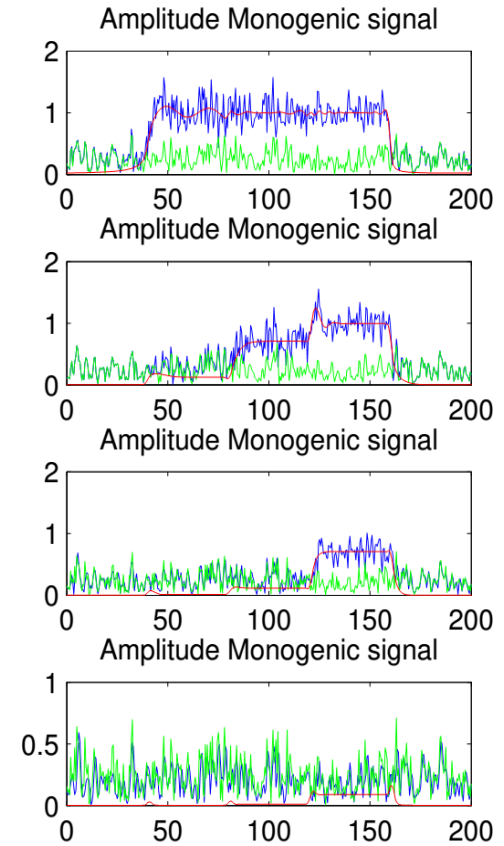
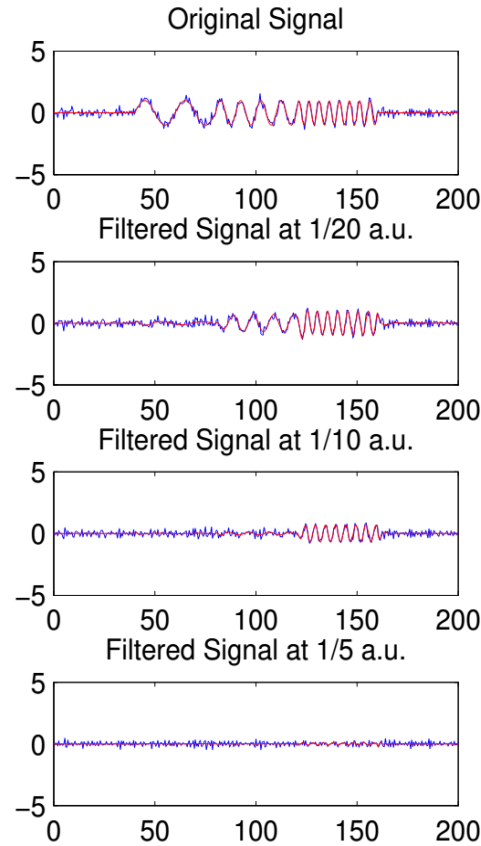
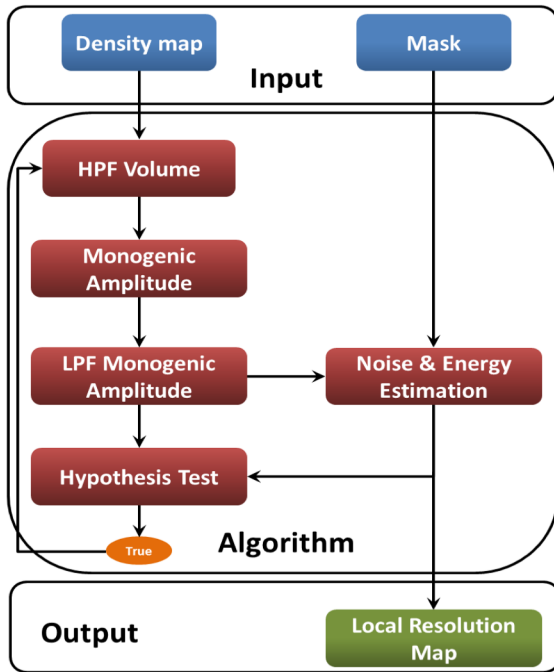
Monogenic signals



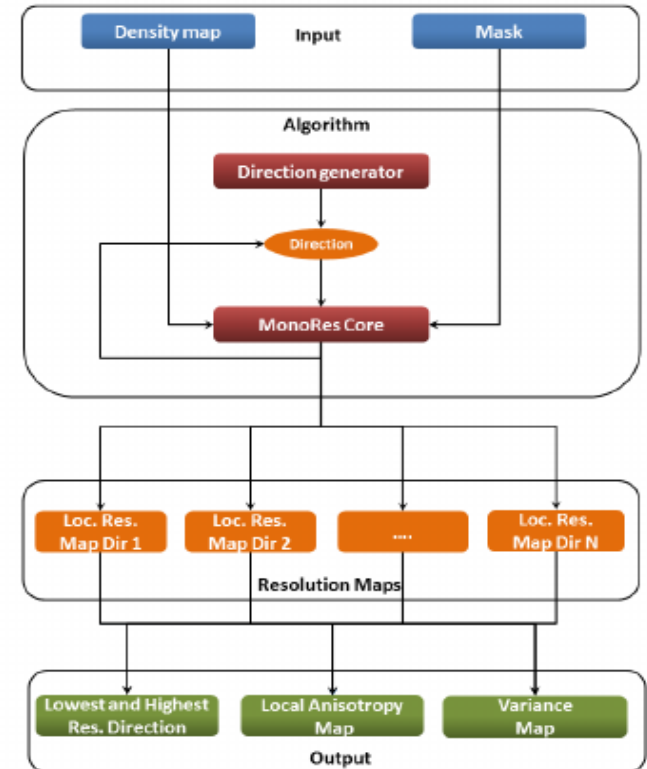
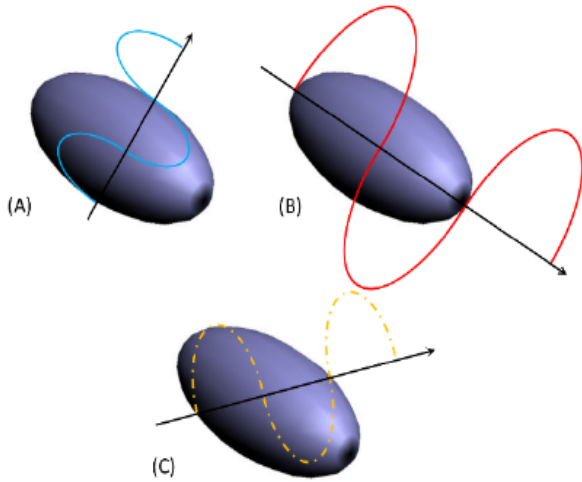
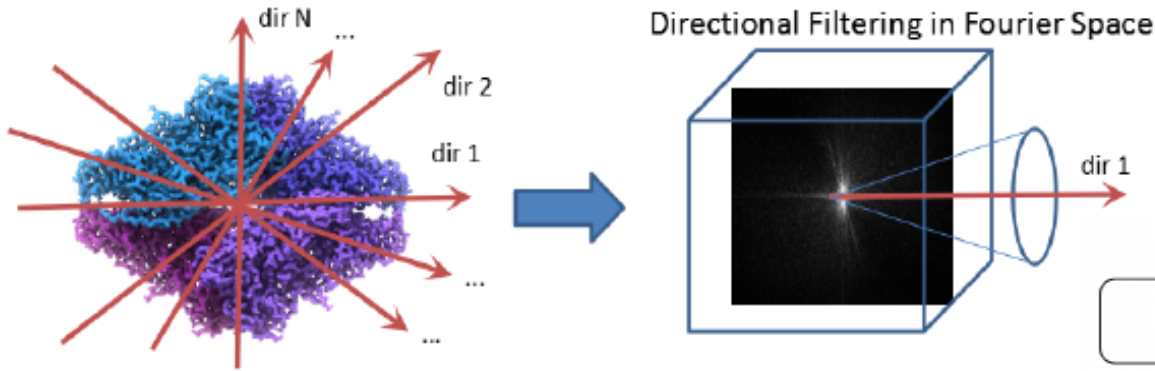
$$\hat{S}_R(\omega) = -\frac{\omega}{\|\vec{\omega}\|} \hat{S}(\vec{\omega}) = -\left(\frac{\omega_1}{\|\vec{\omega}\|} \hat{S}(\vec{\omega}), \frac{\omega_2}{\|\vec{\omega}\|} \hat{S}(\vec{\omega}), \dots, \frac{\omega_N}{\|\vec{\omega}\|} \hat{S}(\vec{\omega}) \right)$$

$$A(\vec{r}) = \sqrt{s^2(\vec{r}) + \sum_{j=1}^N s_{R_j}^2(\vec{r})}.$$

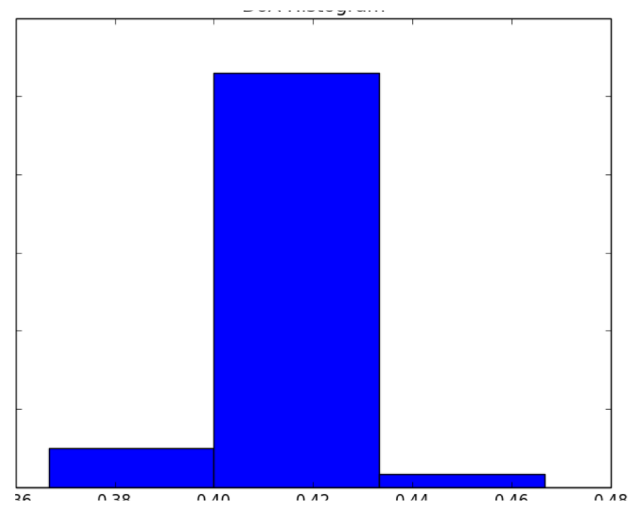
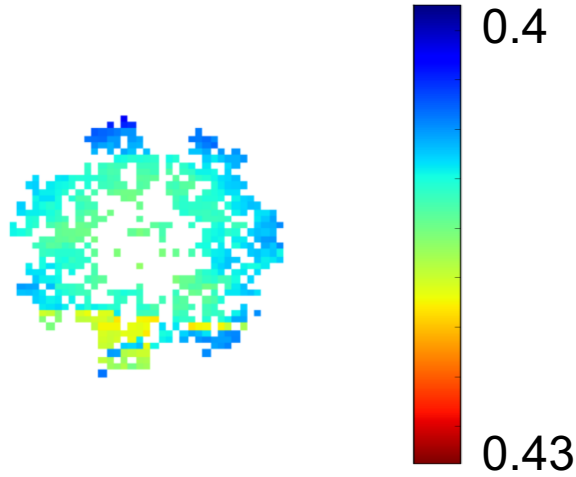
Monogenic local resolution



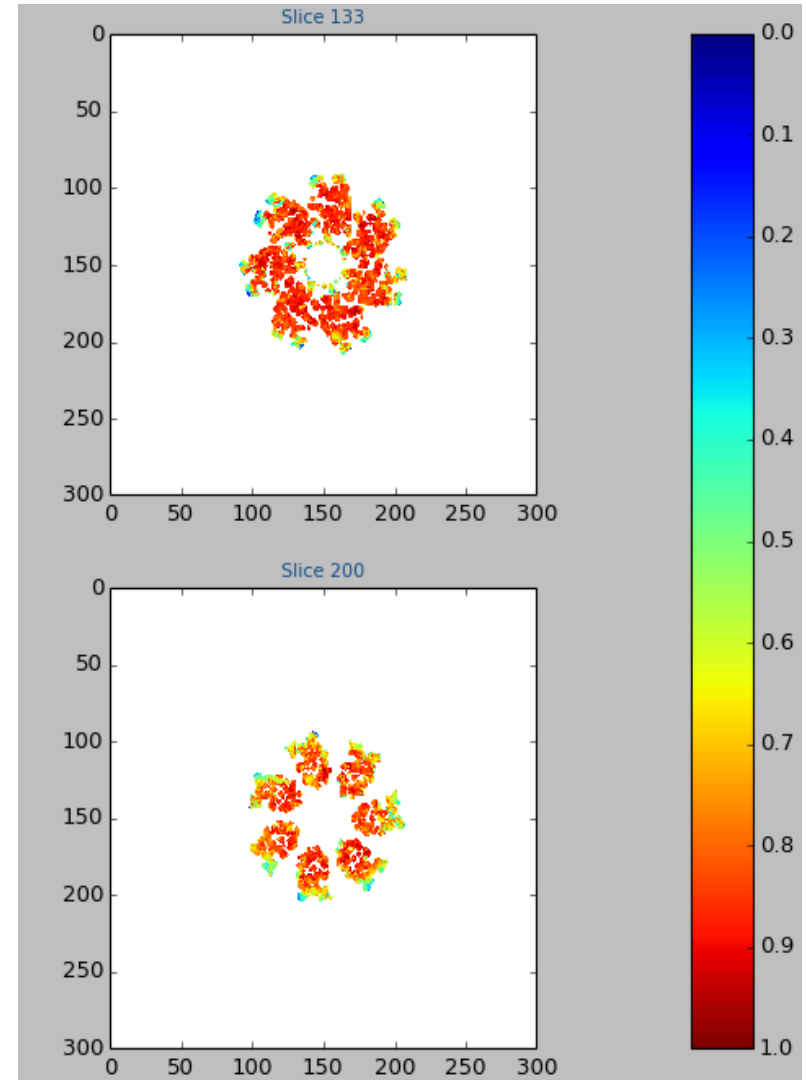
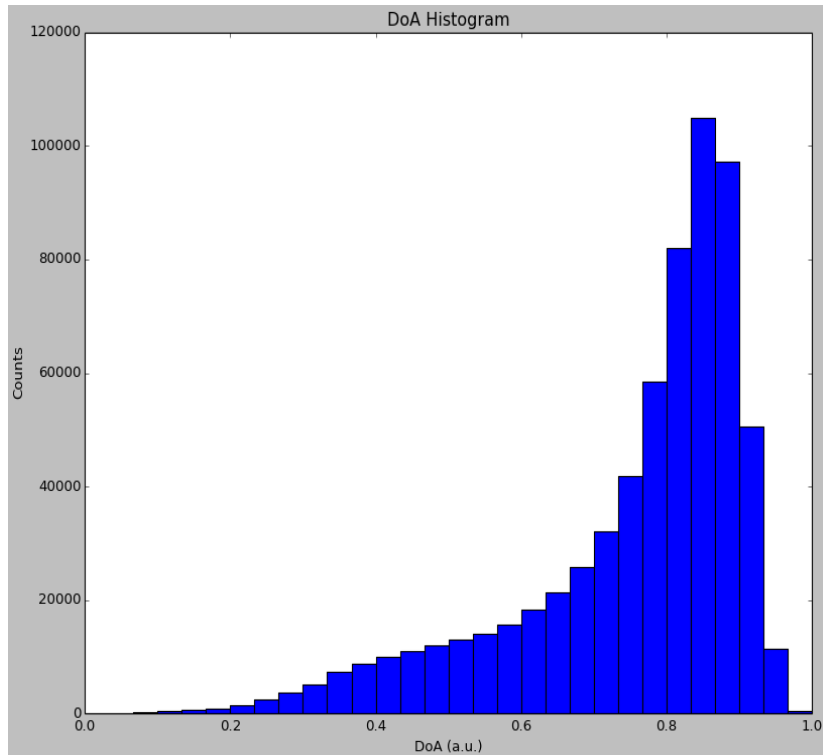
Monogenic local and directional resolution



Results missing wedge



Resolution FSC 2.8Å
High resolution present high isotropy





This work is supported by grants from the Madrid Regional Government (CAM S2010/BMD-2305), the Spanish Ministry of Economy and Competitiveness (AIC-A-2011-0638, BIO2013-44647-R, and BIO2016-76400-R).