Automated acquisition
High-resolution reconstruction
Directional local resolution

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

CSIC

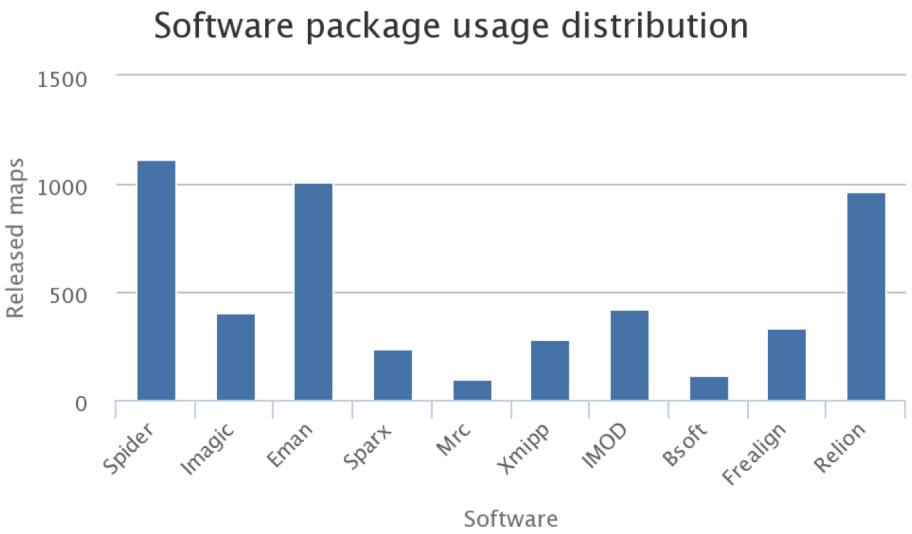
instruct
Integrating
Biology

BIO COMPUTING Unit

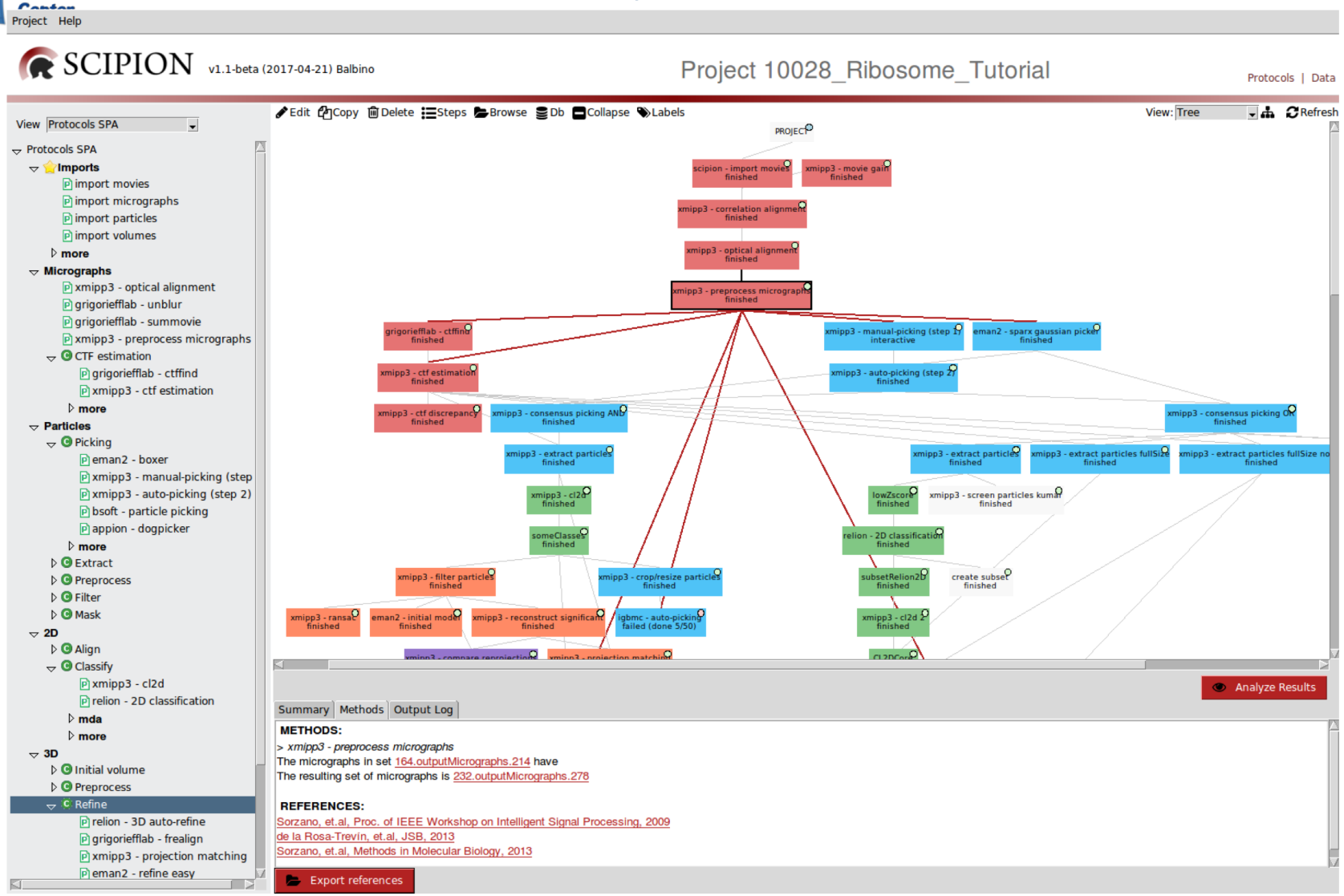
instruct
image
Processing
Center

Scipion

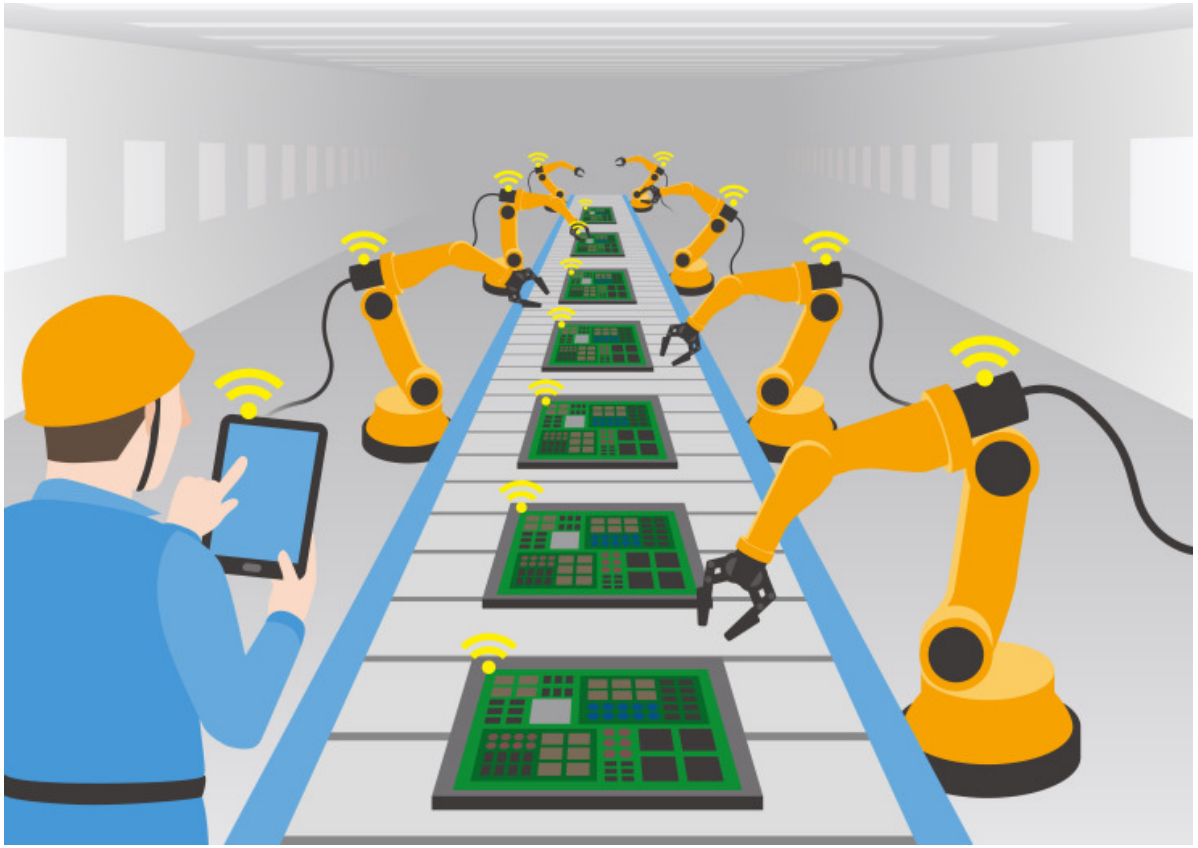
Using different EM software packages is now like Babel's tower



Scipion


instruct
image
Processing
Project Help
SCIPION v1.1-beta (2017-04-21) Balbino
Project 10028_Ribosome_Tutorial
Protocols | Data
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
xmipp3 - auto-picking (step 2)
bsoft - particle picking
appion - dogpicker
more
Extract
Preprocess
Filter
Mask
2D
Align
Classify
xmipp3 - cl2d
relion - 2D classification
mda
more
3D
Initial volume
Preprocess
Refine
relion - 3D auto-refine
grigoriefflab - frealign
xmipp3 - projection matching
eman2 - refine easy
Edit Copy Delete Steps Browse Db Collapse Labels
View: Tree Refresh
PROJECT
scipion - import movies finished
xmipp3 - movie gain finished
xmipp3 - correlation alignment finished
xmipp3 - optical alignment finished
xmipp3 - preprocess micrographs finished
grigoriefflab - ctffind finished
xmipp3 - manual-picking (step 1) interactive
eman2 - sparx gaussian picker finished
xmipp3 - ctf estimation finished
xmipp3 - auto-picking (step 2) finished
xmipp3 - ctf discrepancy finished
xmipp3 - consensus picking AND finished
xmipp3 - consensus picking OR finished
xmipp3 - extract particles finished
xmipp3 - extract particles finished
xmipp3 - extract particles fullSize finished
xmipp3 - extract particles fullSize no finished
xmipp3 - cl2d finished
lowZscore finished
xmipp3 - screen particles kurtosis finished
someClasses finished
relion - 2D classification finished
xmipp3 - filter particles finished
xmipp3 - crop/resize particles finished
subsetRelion2D finished
create subset finished
xmipp3 - ransac finished
eman2 - initial model finished
xmipp3 - reconstruct significant finished
igbmc - auto-picking failed (done 5/50)
xmipp3 - cl2d 2 finished
Analyze Results
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-Trevín, et.al, JSB, 2013
Sorzano, et.al, Methods in Molecular Biology, 2013
Export references

Automated acquisition



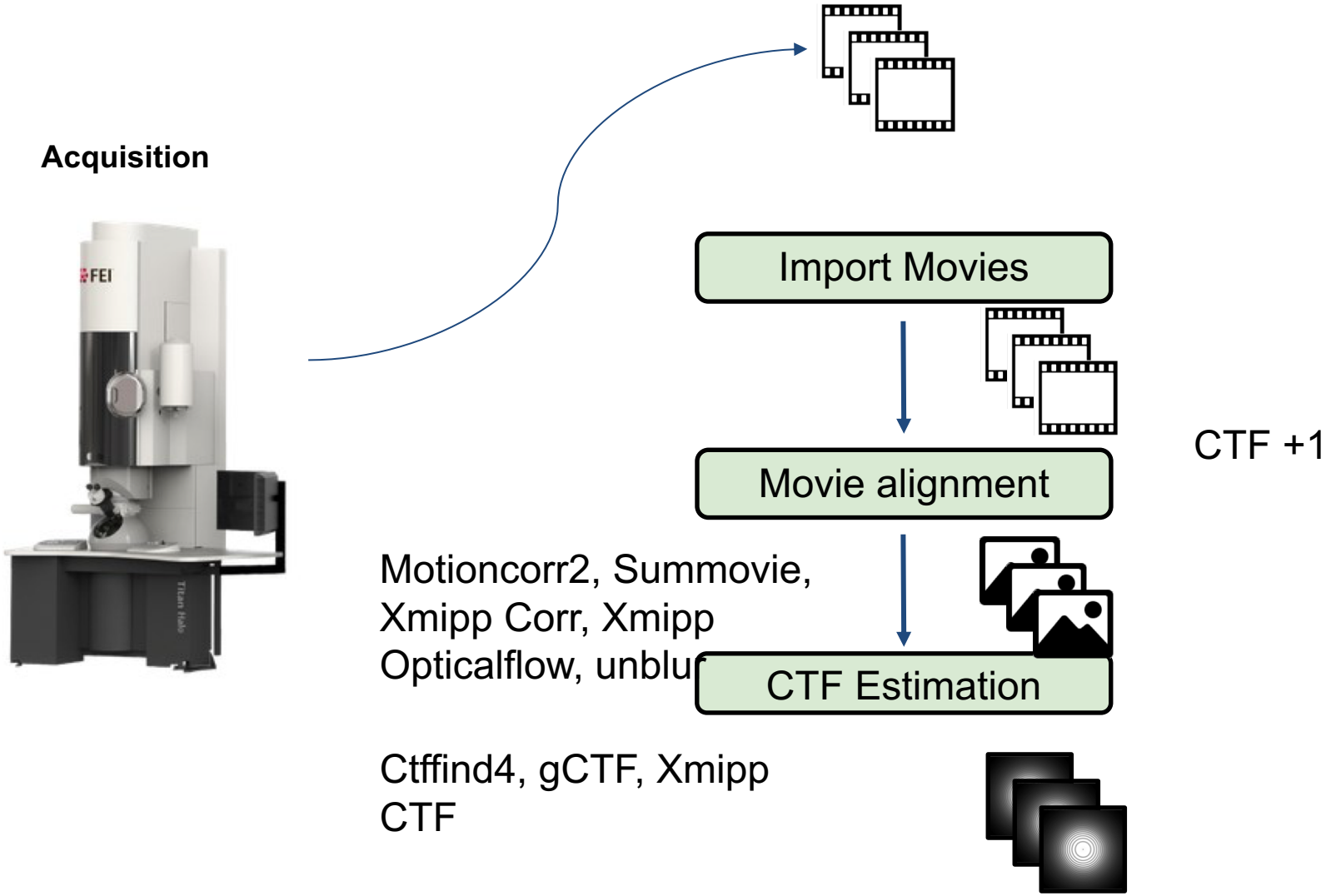
CSIC

instruct
Integrating
Biology

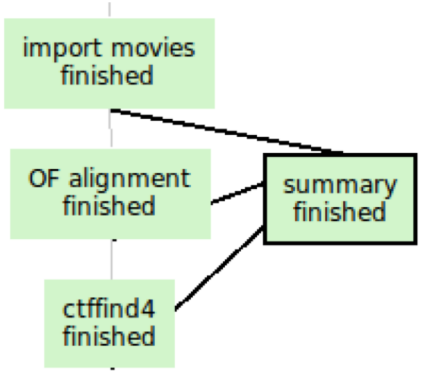
BO COMPUTING Unit

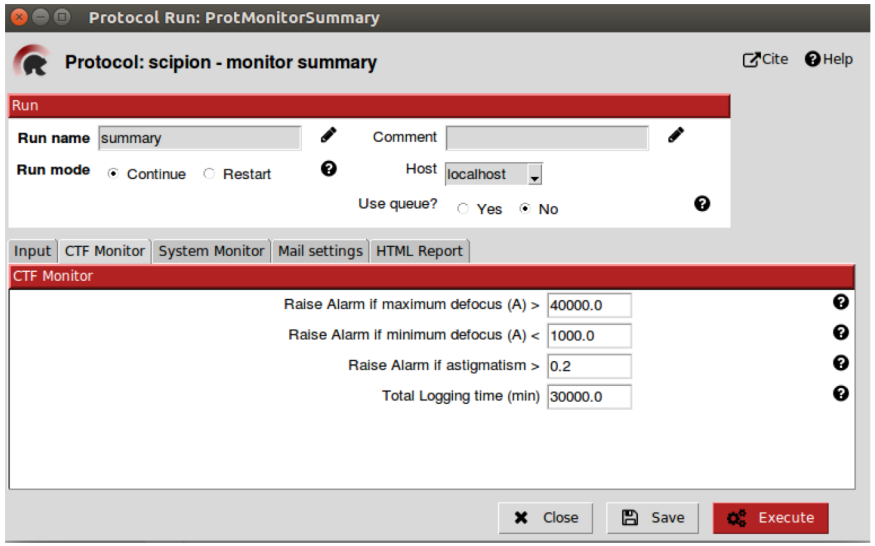
instruct
image
Processing
Center

Scipion 1.1: Scipion box = Streaming = Facilities



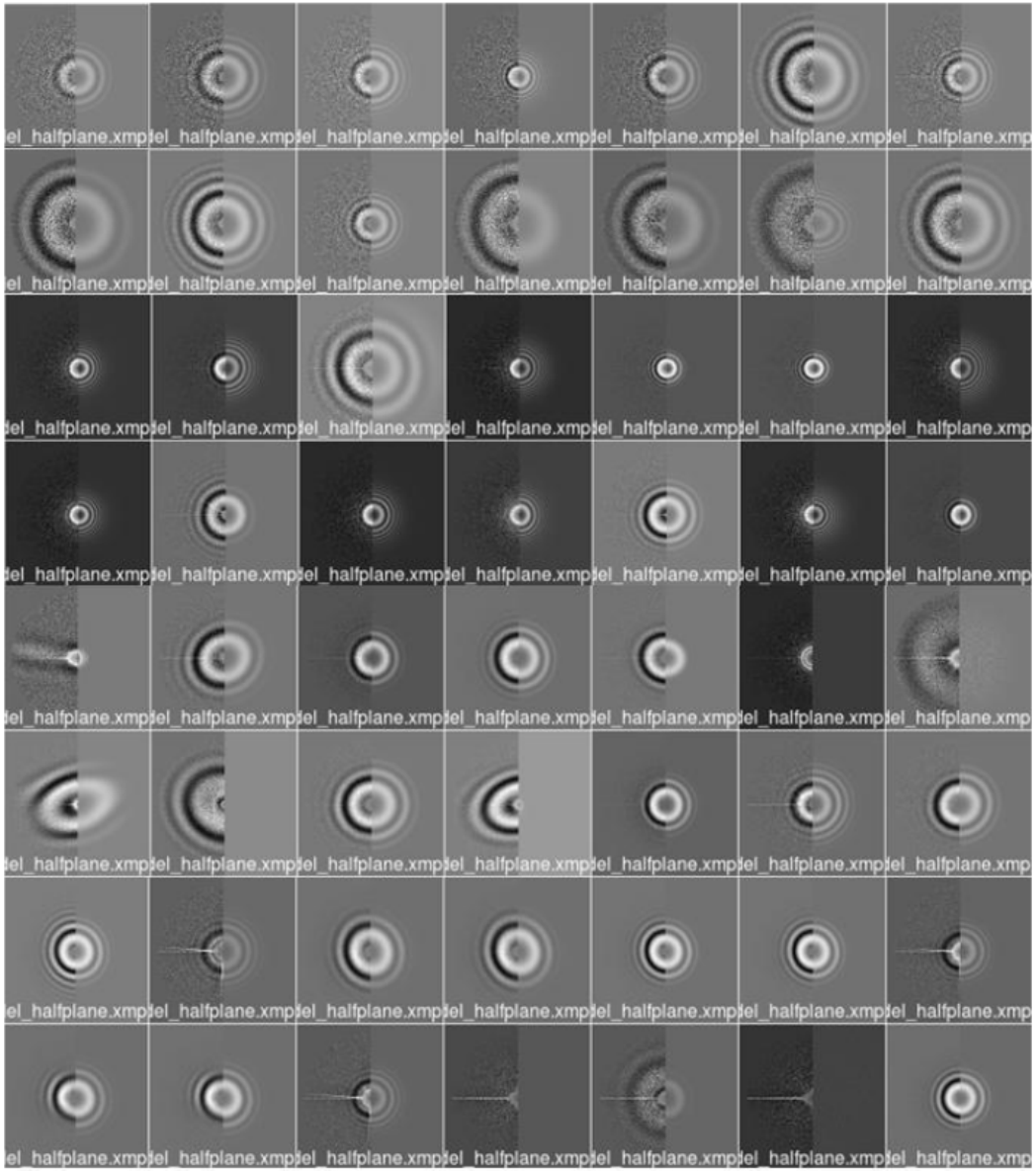
Monitoring



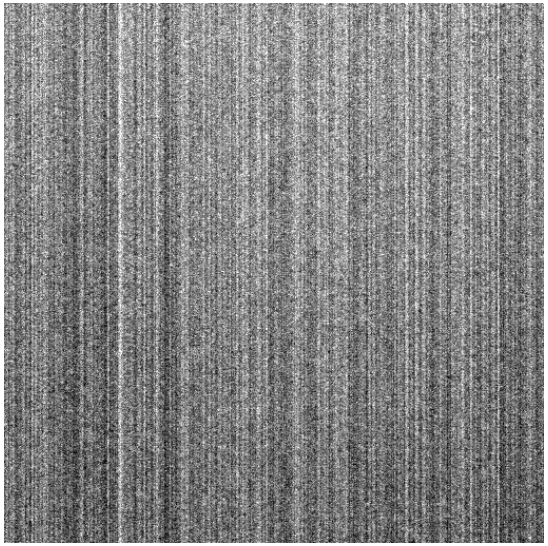


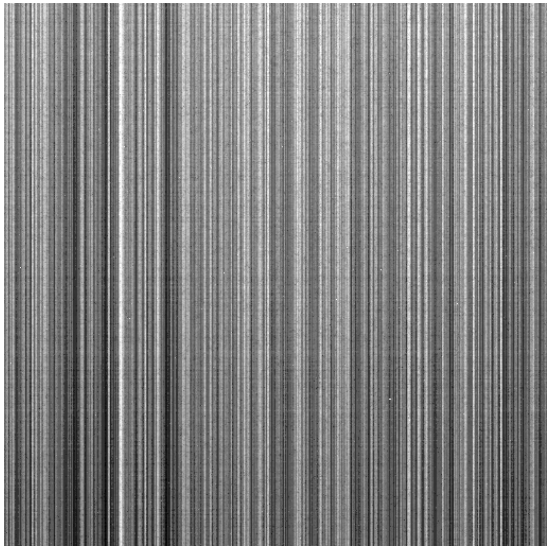
- Need feedback ASAP
- Report what has been done in the facility
- Track system status, memory, gpu, cpu, network
- Raise alarms when thresholds reached

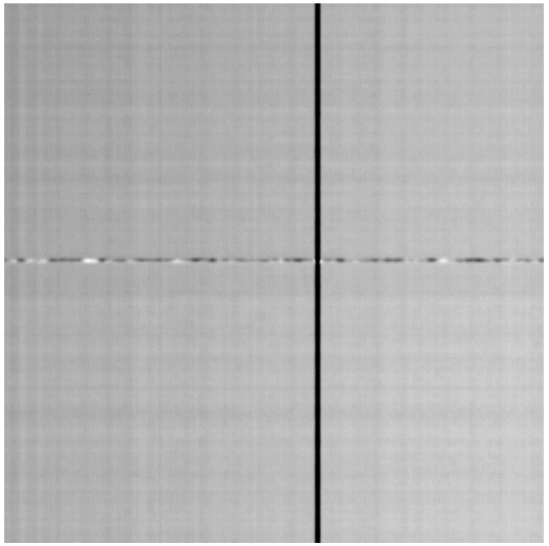
Automatic CTF rejection

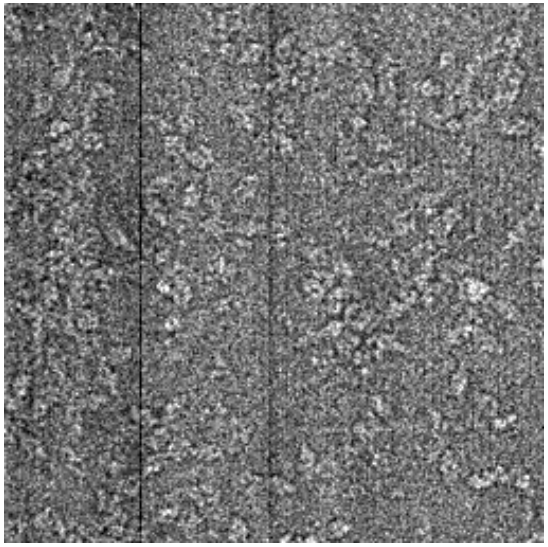


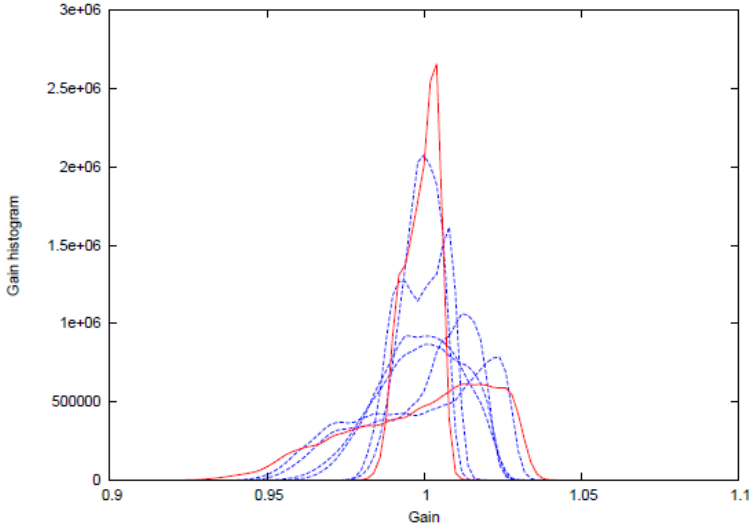
Residual gain estimate from a single movie



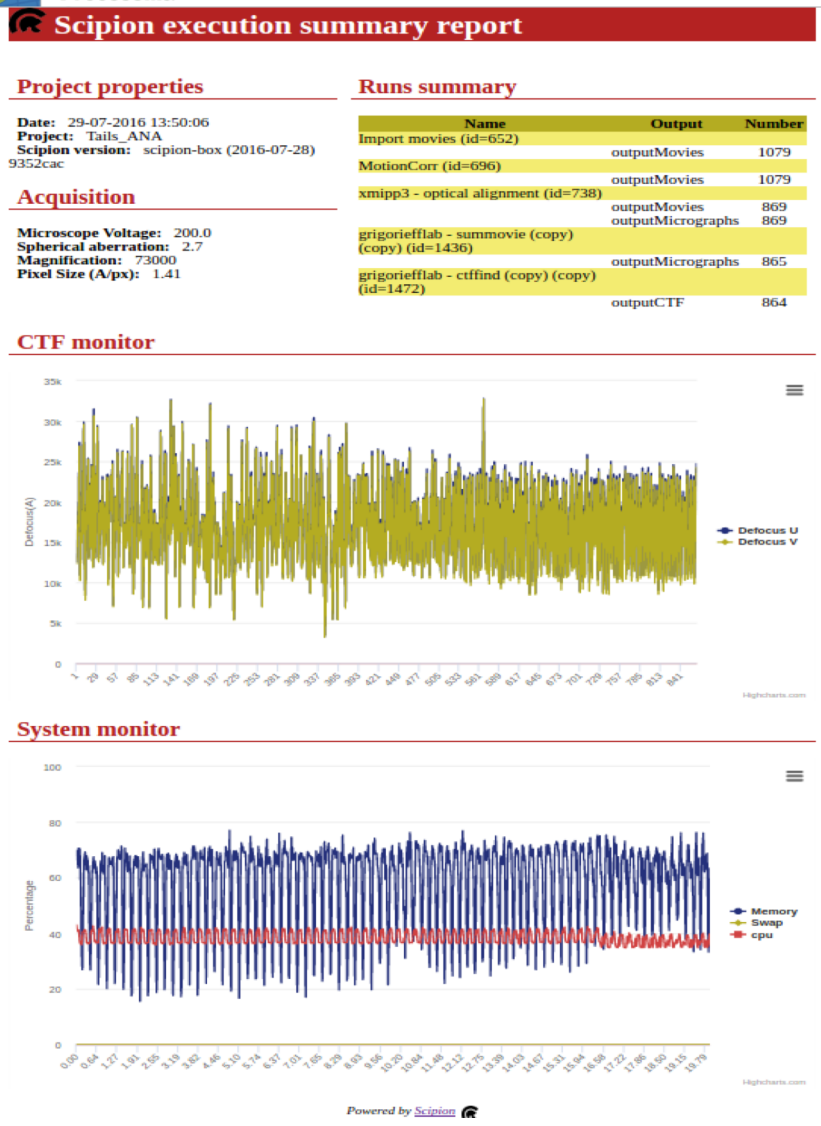








Report



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)		
	outputMovies	1079
MotionCorr (id=696)		
	outputMovies	1079
xmipp3 - optical alignment (id=738)		
	outputMovies	869
	outputMicrographs	869
grigoriefflab - summovie (copy) (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

Microscope

Scipion Box

Import Movies

Movie Alignment

Monitors

CTF Estimation

Particle Picking

Extraction

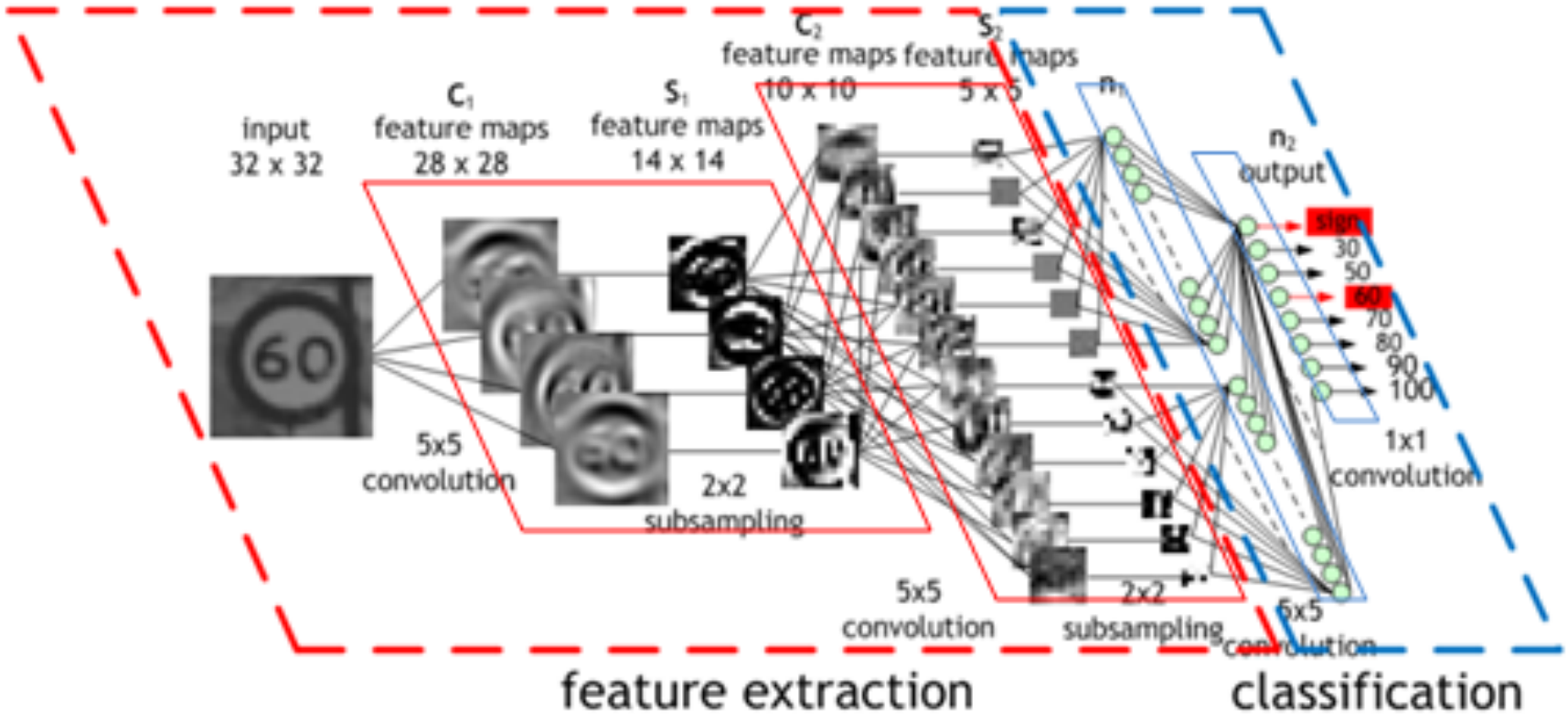
2D Classification

Initial Volume

3D Classification

3D Refinement

Deep consensus



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

Submission to facility databases

ESRF
The European Synchrotron

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

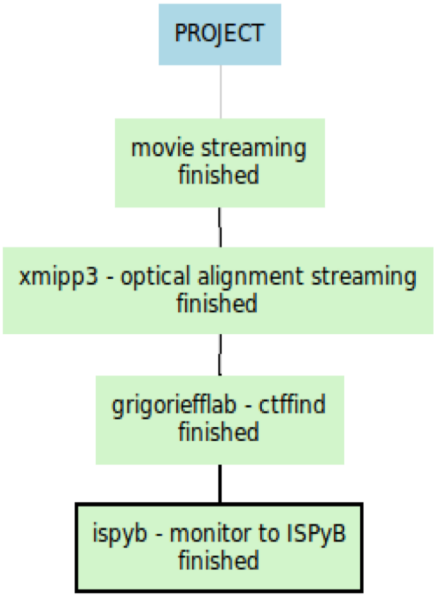
ISPYB News

ISPyB User Manual

ISPYB project

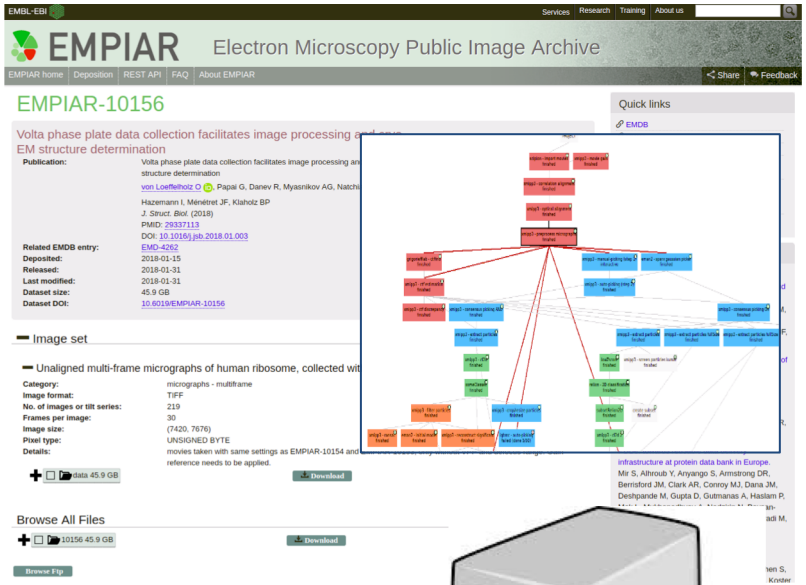
The ISPyB (Information System for Protein Crystallograp
protein crystallography experiments on synchrotron bear
to BioSaxs beamlines.

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



Submission to public databases

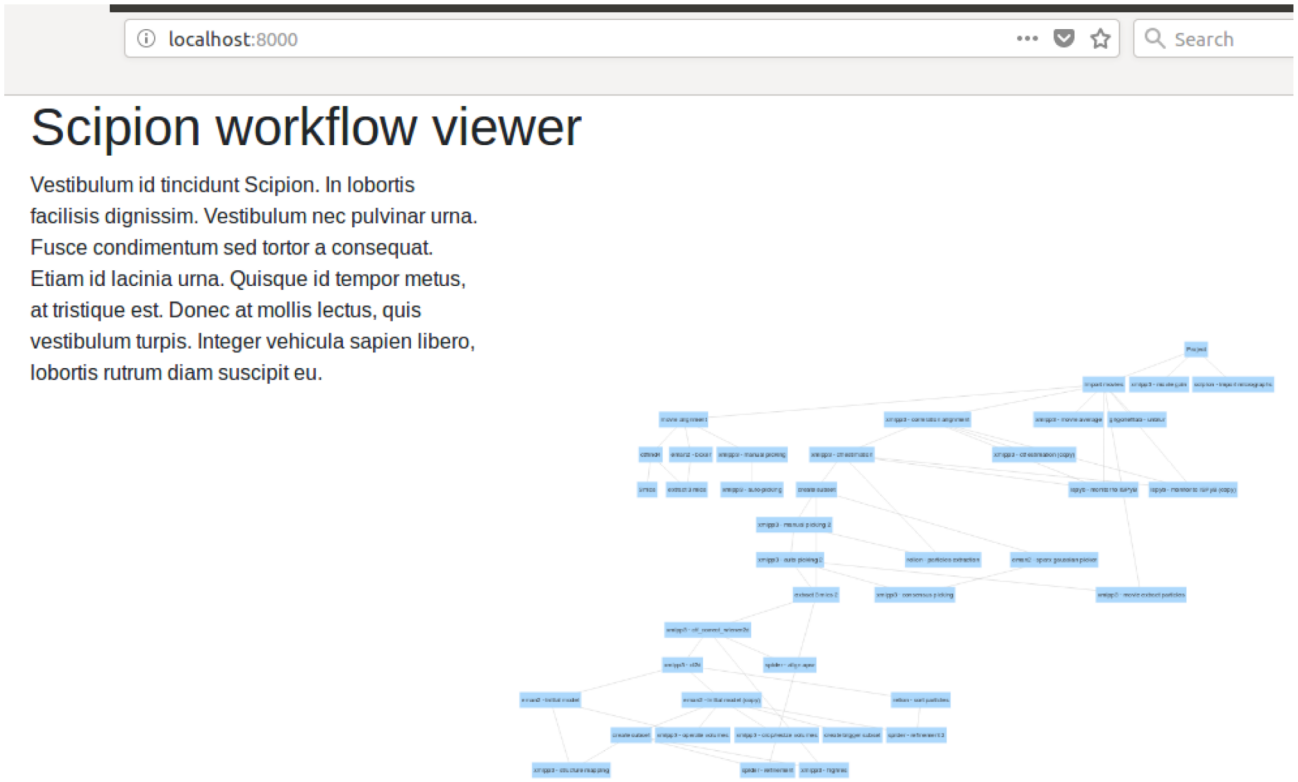
Download and reproduce



RAW data +

.json

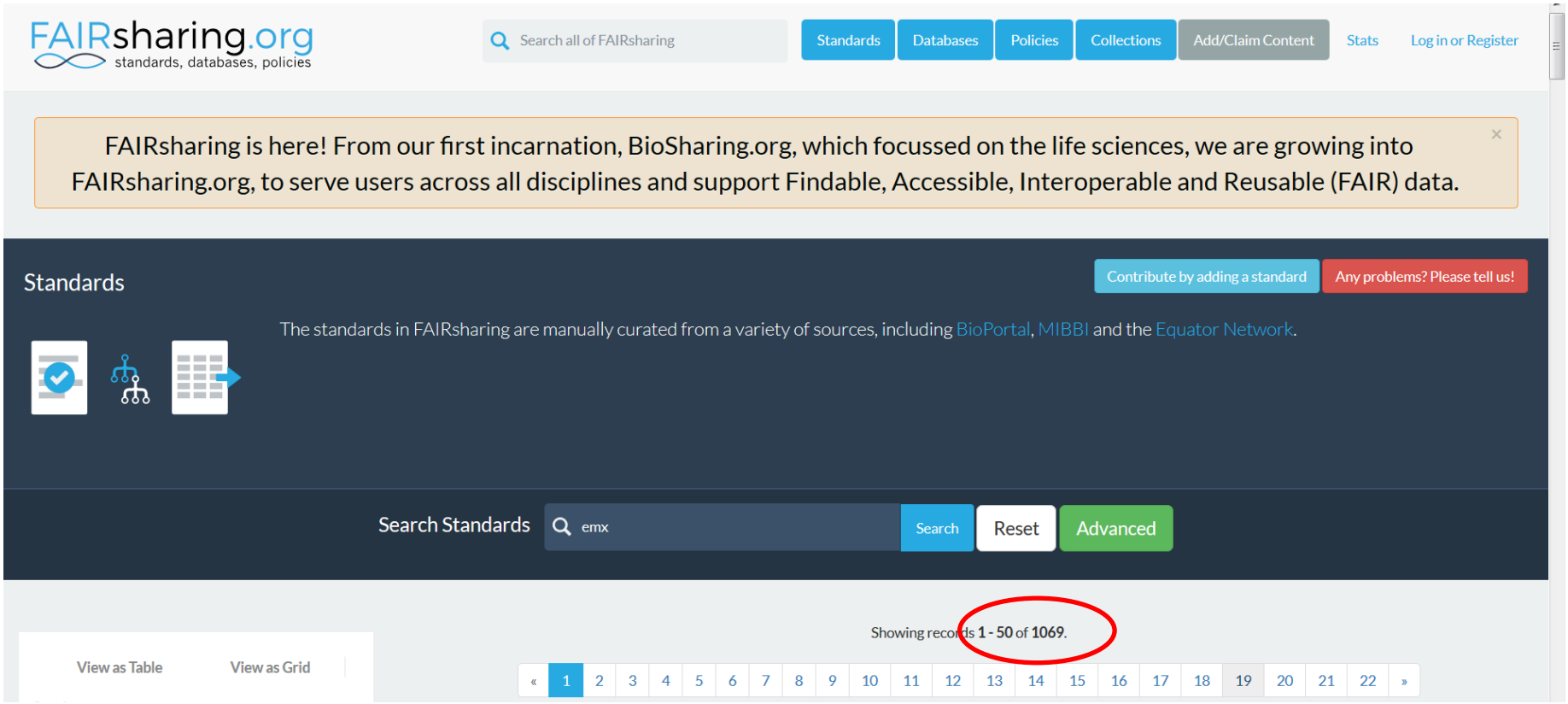
Workflow visualization



Github: https://github.com/I2PC/web-workflow-viewer

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

FAIR Data: Findable, Accessible, Interoperable, Reusable

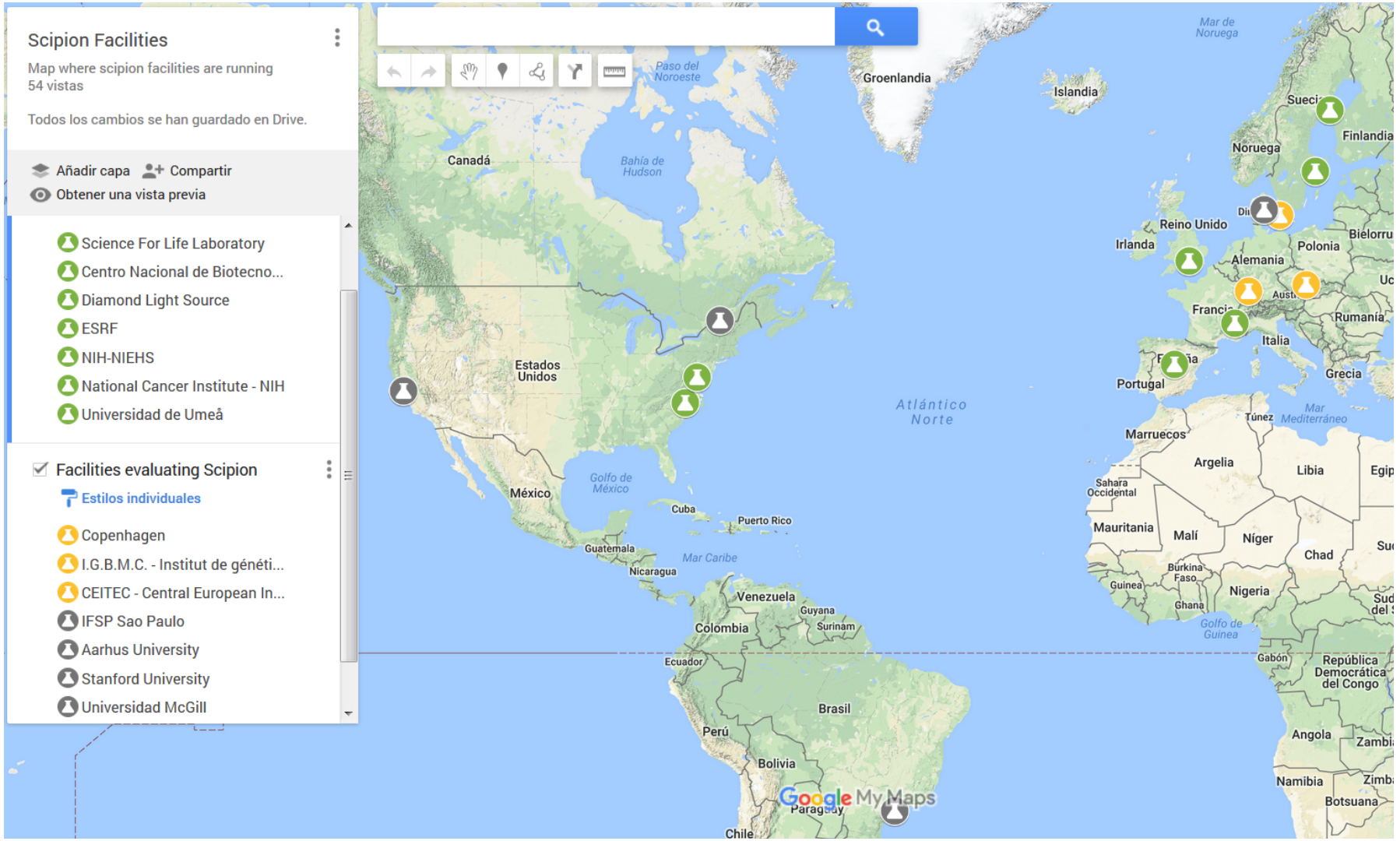


CSIC

instruct Integrating Biology

BIO COMPUTING Unit

instruct image Processing Center



CSIC

instruct Integrating Biology

BIO COMPUTING Unit

instruct image Processing Center

instruct
image
Processing
Center

HighRes

HighRes philosophy

Global alignment until stable:

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

Then, local alignment

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

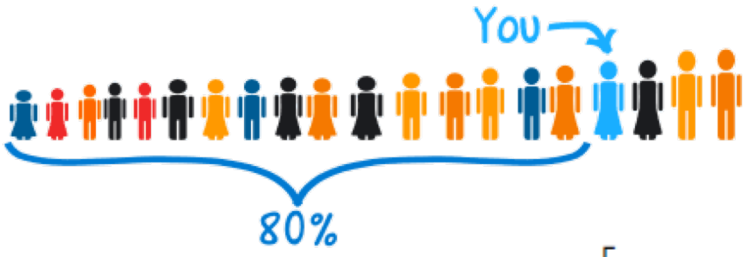
Remove noise “anchors”/unsignificant features:

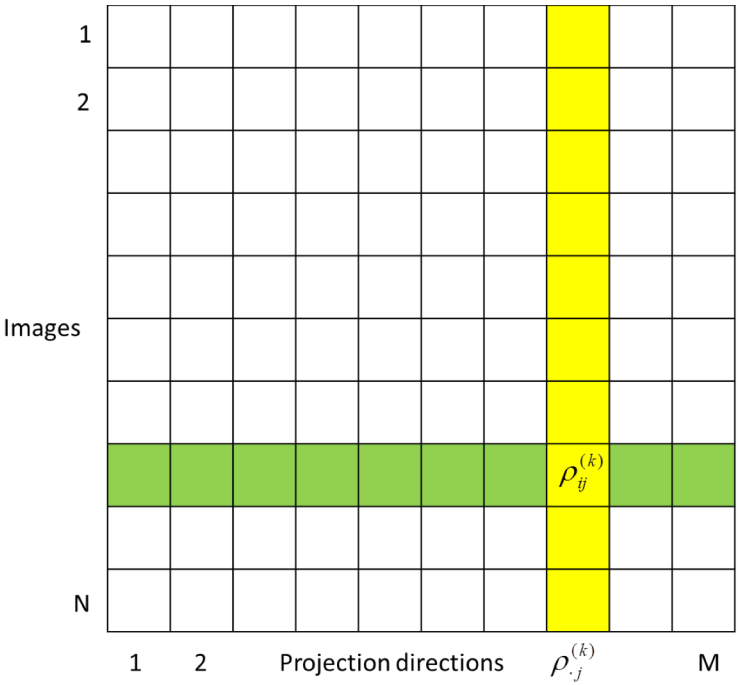
- Reference volume should contain only significant information
- Output volume only with significant features

Multiresolution optimization:

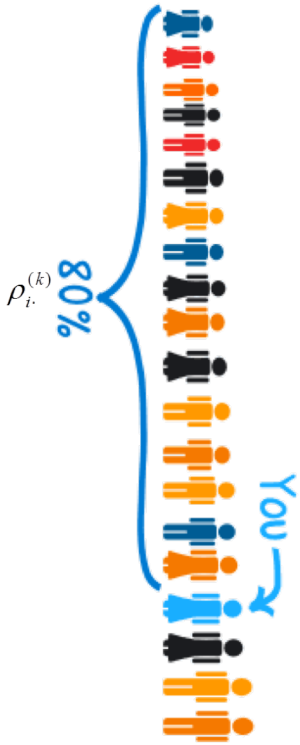
- Increase speed
- Smooth solution landscape
- Slow annealing to 5-7 A

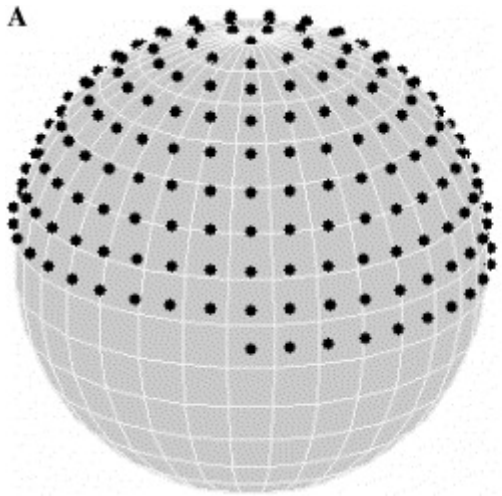
Significant Reconstruction





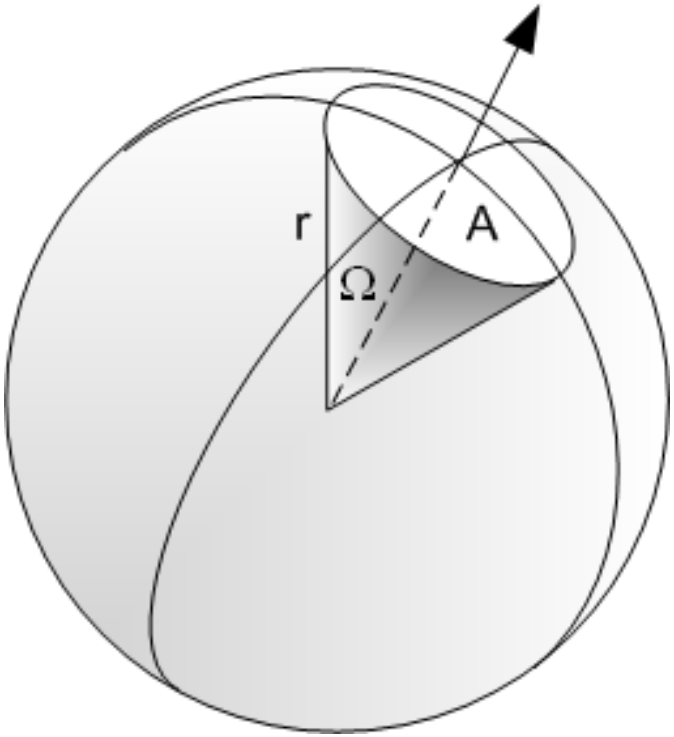
$$\rho \in \left[\tanh(\tanh^{-1}(\max_j\{\rho_{ij}^{(k)}\}) - \frac{z_{1-\alpha^{(k)}}}{\sqrt{N-3}}, \max_j\{\rho_{ij}^{(k)}\}\right]$$

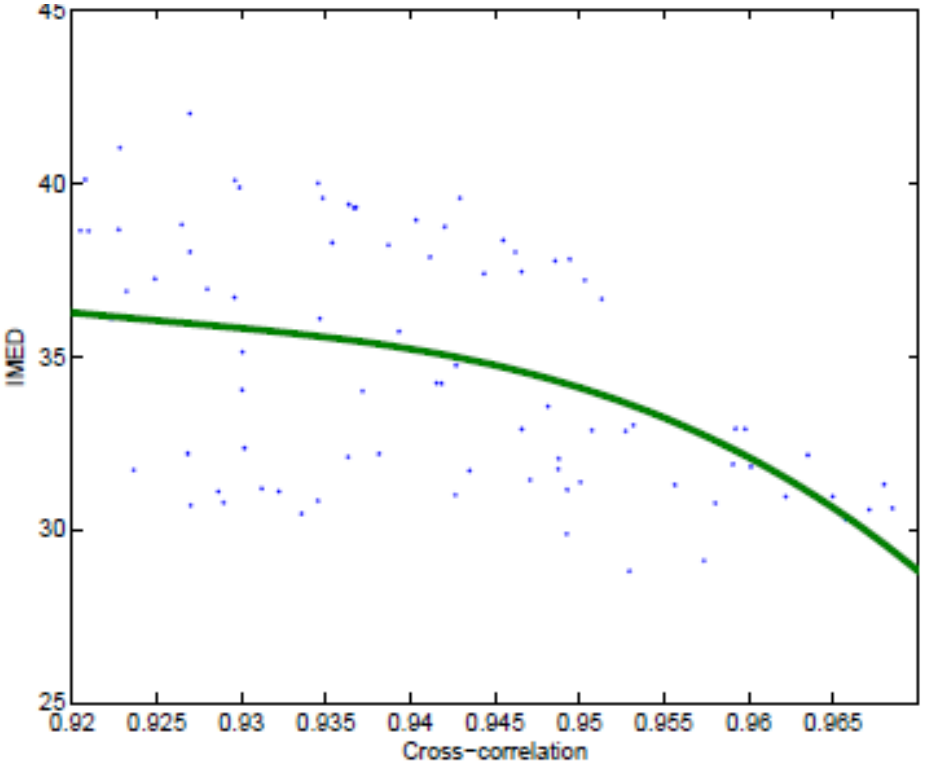




Significant Reconstruction

$$w_{ij}^{(k)} = \begin{cases} \dfrac{\min\limits_{\substack{i' \in 1,\ldots,N \\ j' \in \mathrm{Neigh}_\theta(j)}} \eta_{i'j'}^{(k)}}{\eta_{ij}^{(k)}} \Pr\left\{\eta_{i\cdot}^{(k)} \geq \eta_{ij}^{(k)}\right\} \Pr\left\{\eta_{\cdot j}^{(k)} \geq \eta_{ij}^{(k)}\right\} \\ \dfrac{\rho_{ij}^{(k)}}{\max\limits_{\substack{i' \in 1,\ldots,N \\ j' \in \mathrm{Neigh}_\theta(j)}} \rho_{i'j'}^{(k)}} \Pr\left\{\rho_{i\cdot}^{(k)} \leq \rho_{ij}^{(k)}\right\} \Pr\left\{\rho_{\cdot j}^{(k)} \leq \rho_{ij}^{(k)}\right\} \end{cases}$$





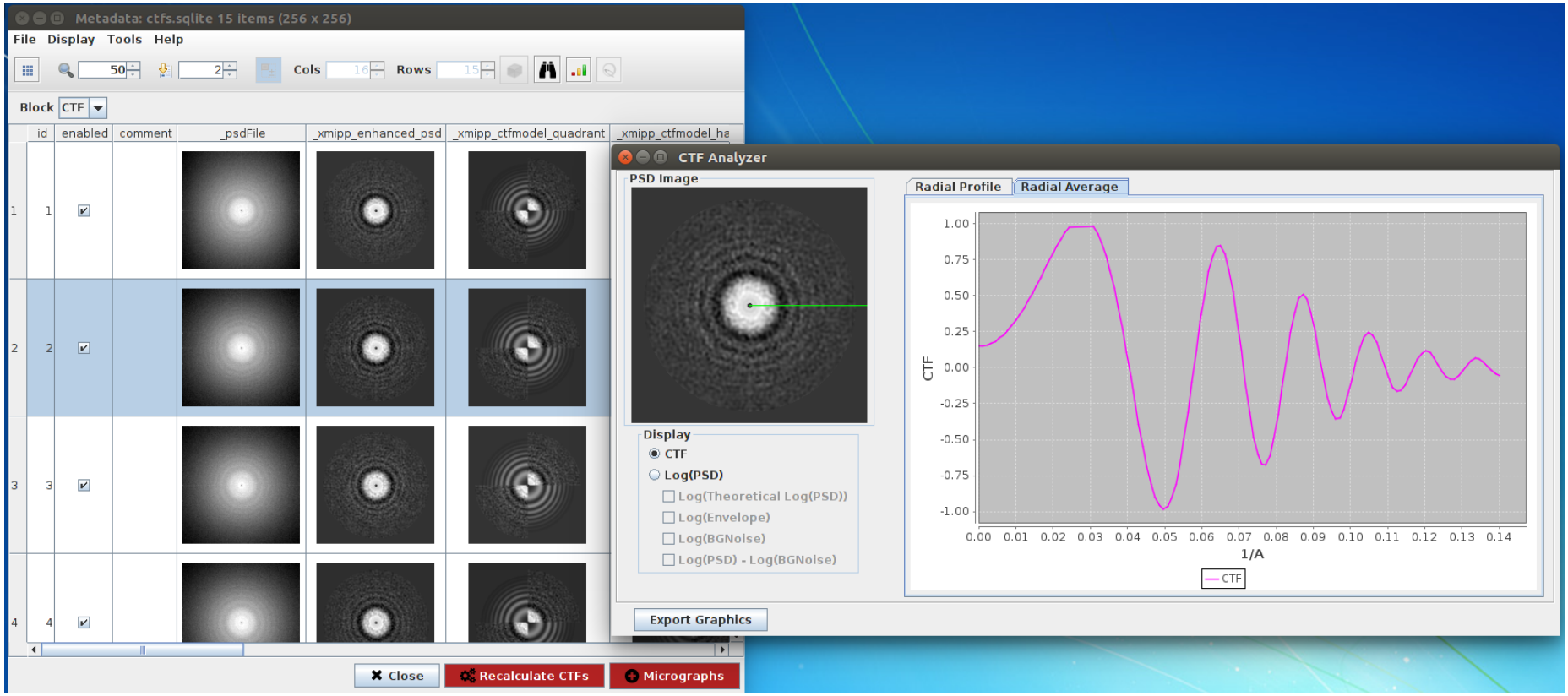
Not a 2nd chance

Intuitively

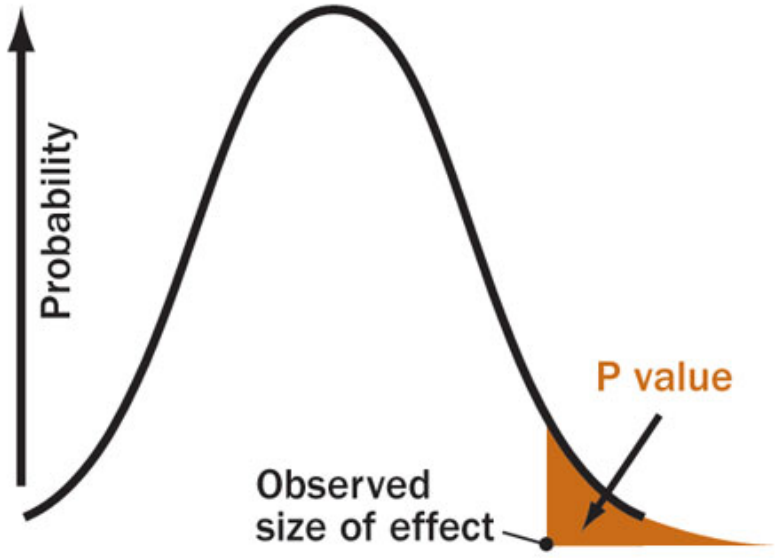
- Conventionally: assign **ONE** orientation/classification, based on maximum CC.
- Expectation-Maximization: calculate the *probability-weighted average over* **ALL** possible assignments

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

Envelope correction

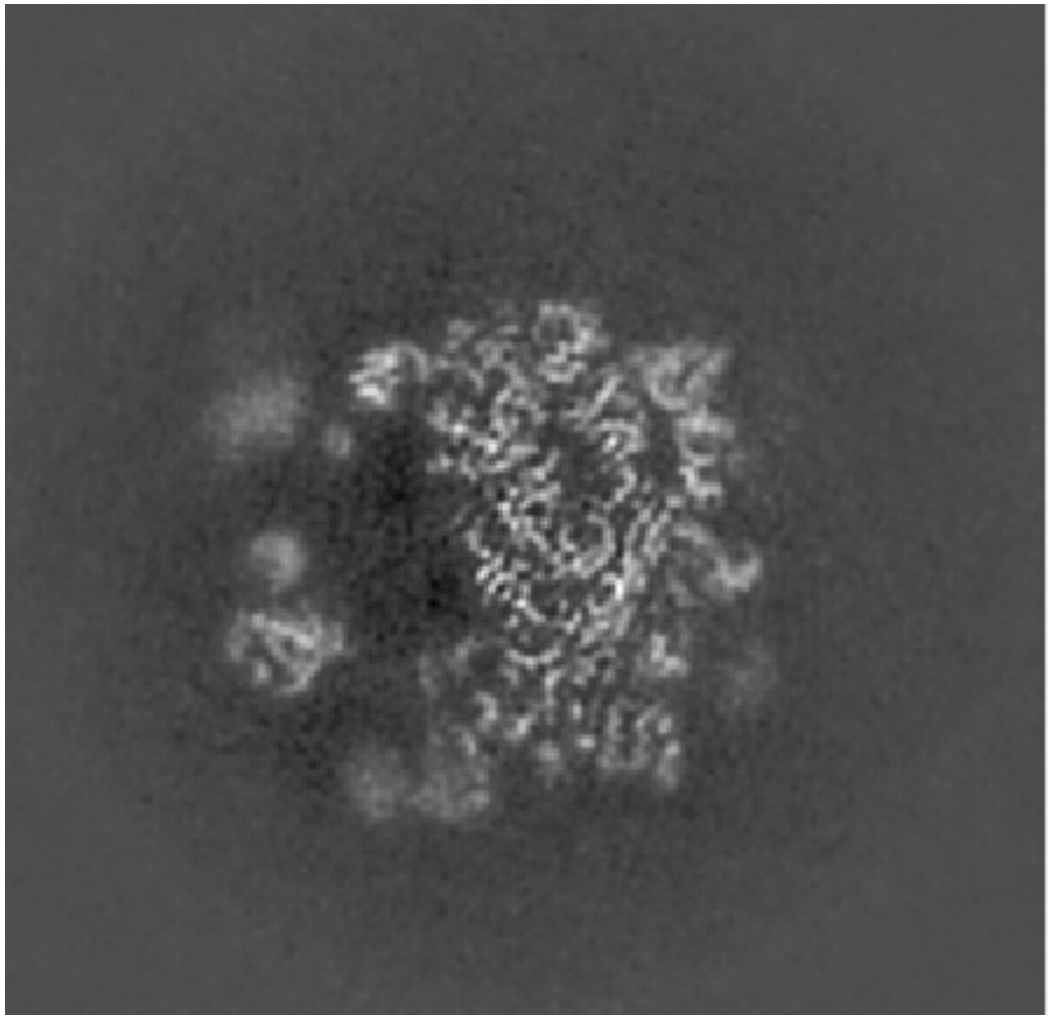


Significant features

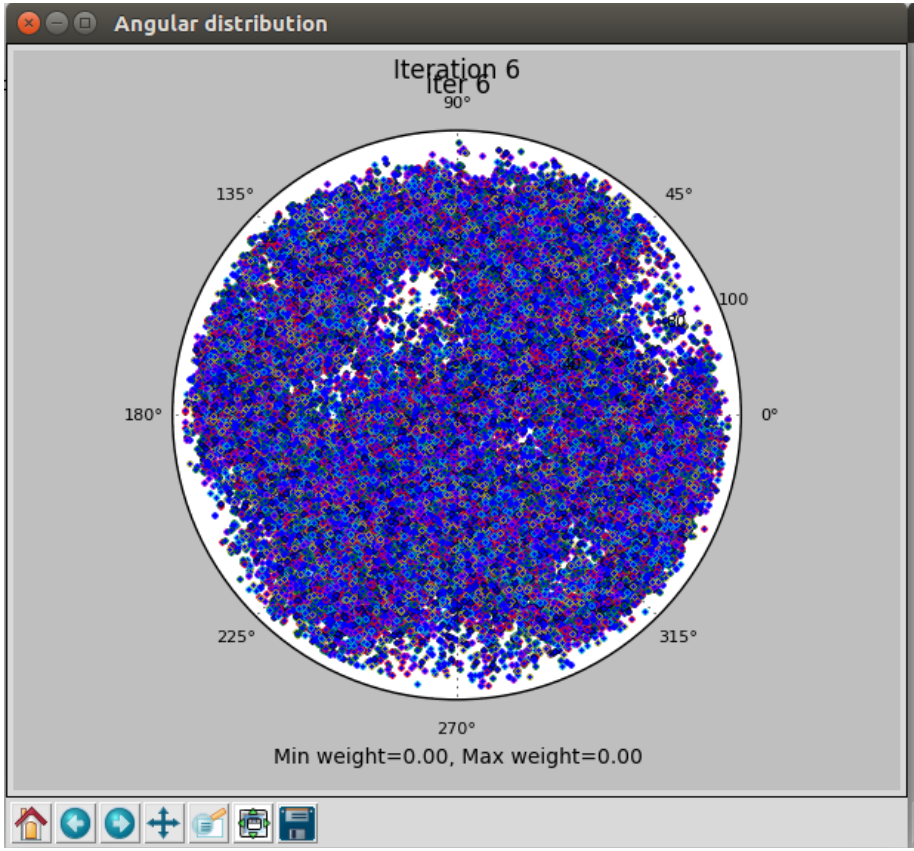


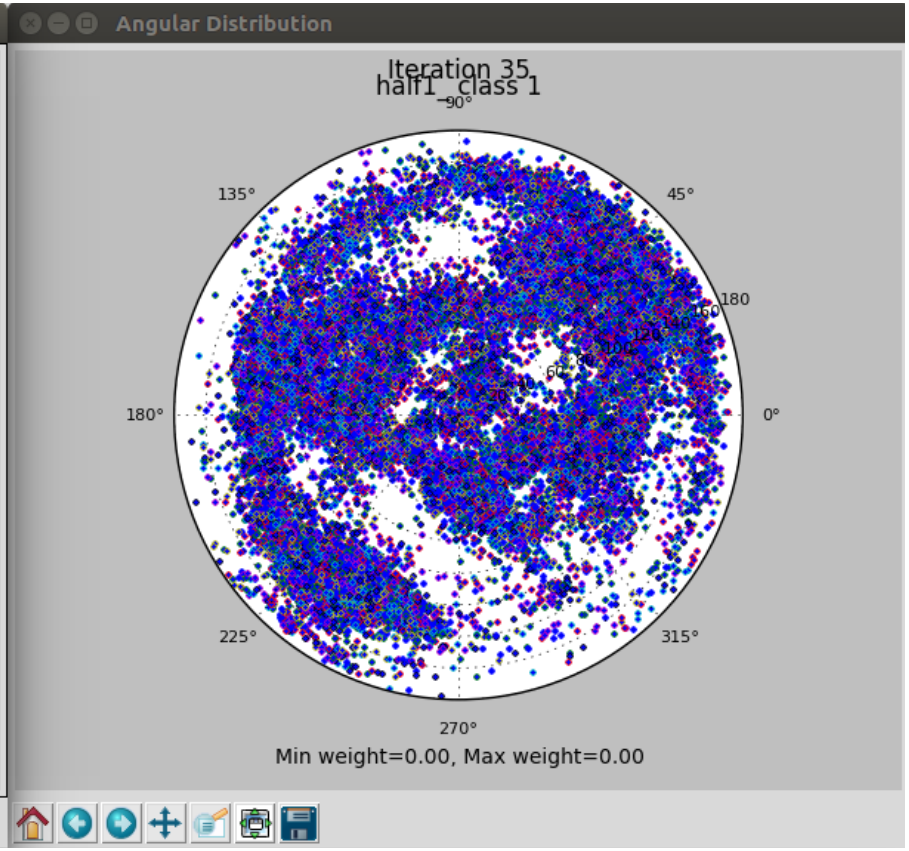
A Researcher's (Misguided) Notion of Nirvana:

$p < .05$



Results


Angular distribution
Iteration 6
Iter 6
90°
135°
45°
180°
0°
225°
315°
270°
Min weight=0.00, Max weight=0.00


Angular Distribution
Iteration 35
half1_class 1
90°
135°
45°
180°
0°
225°
315°
270°
Min weight=0.00, Max weight=0.00

Results

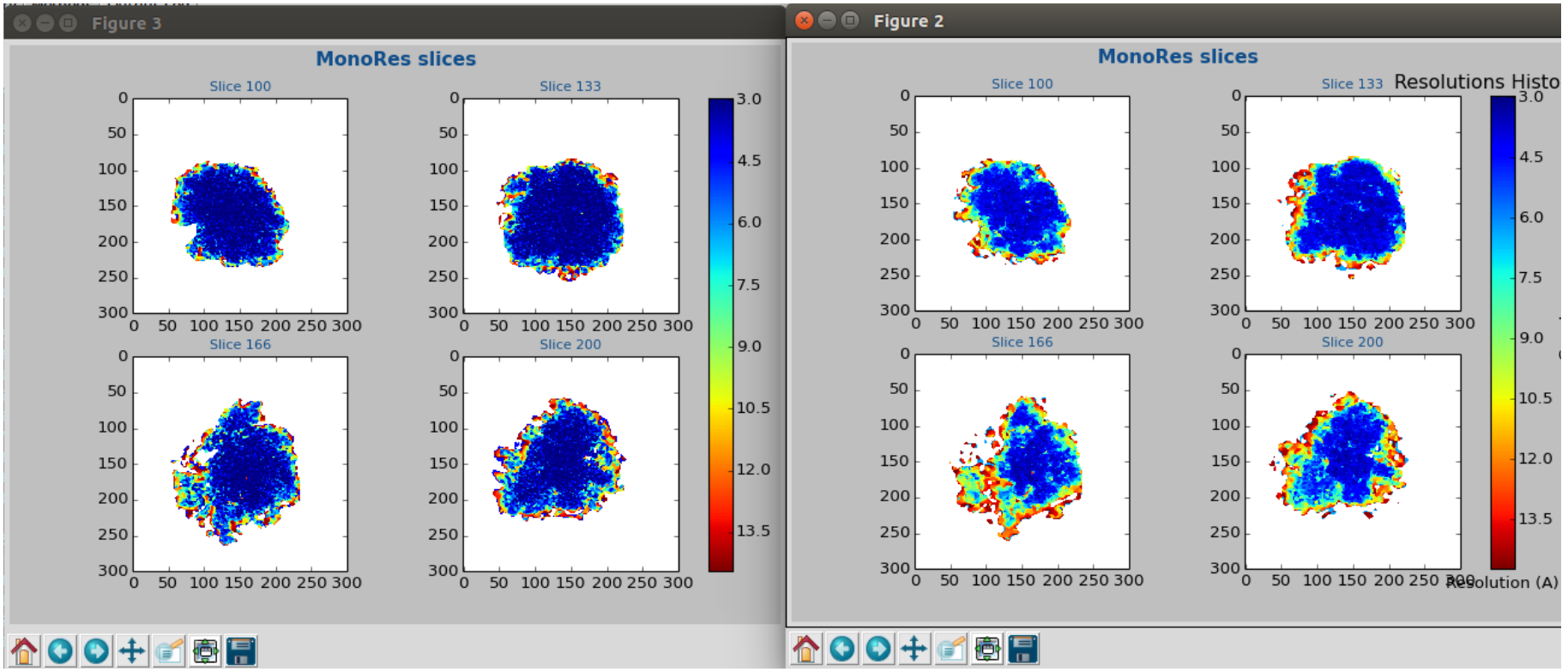
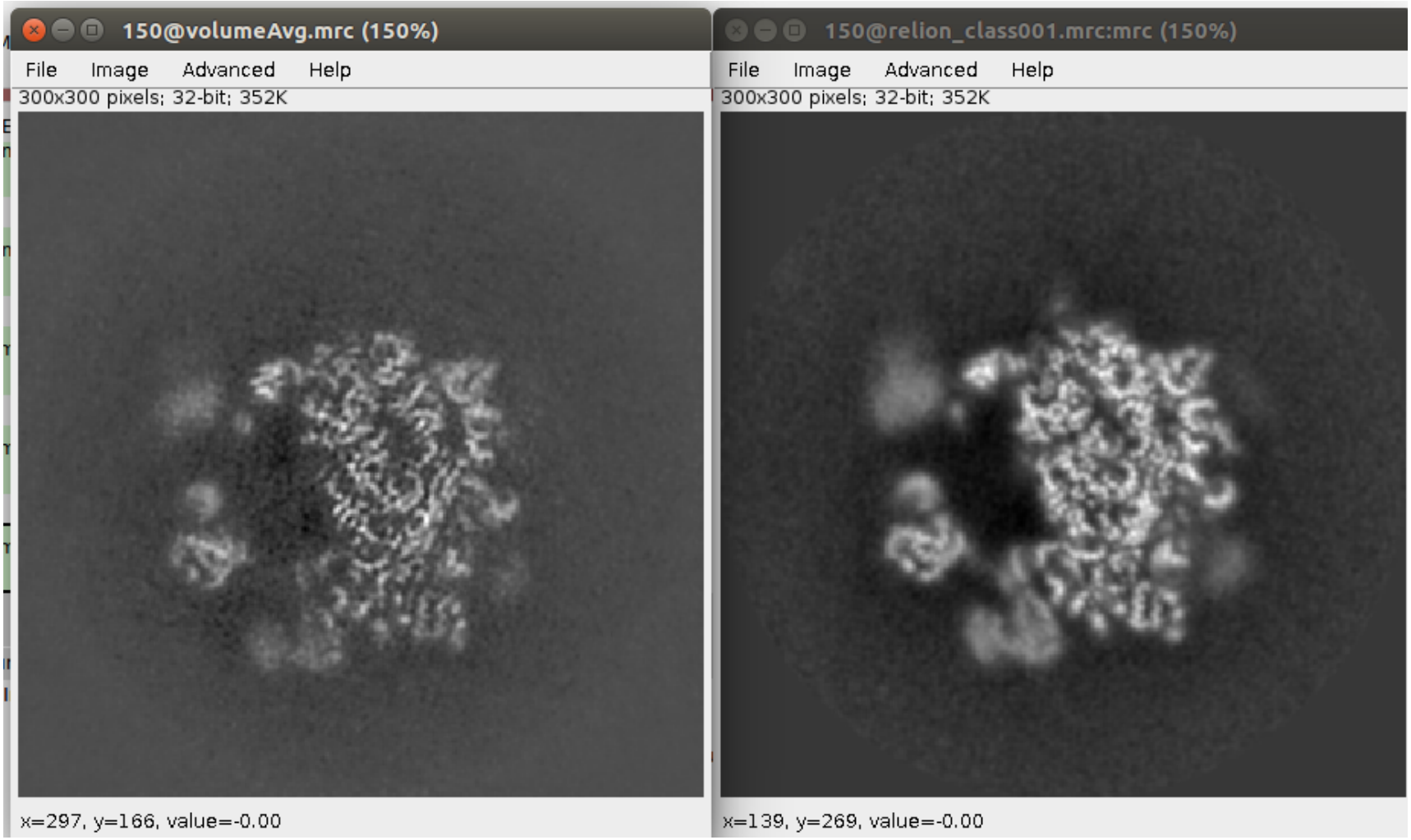
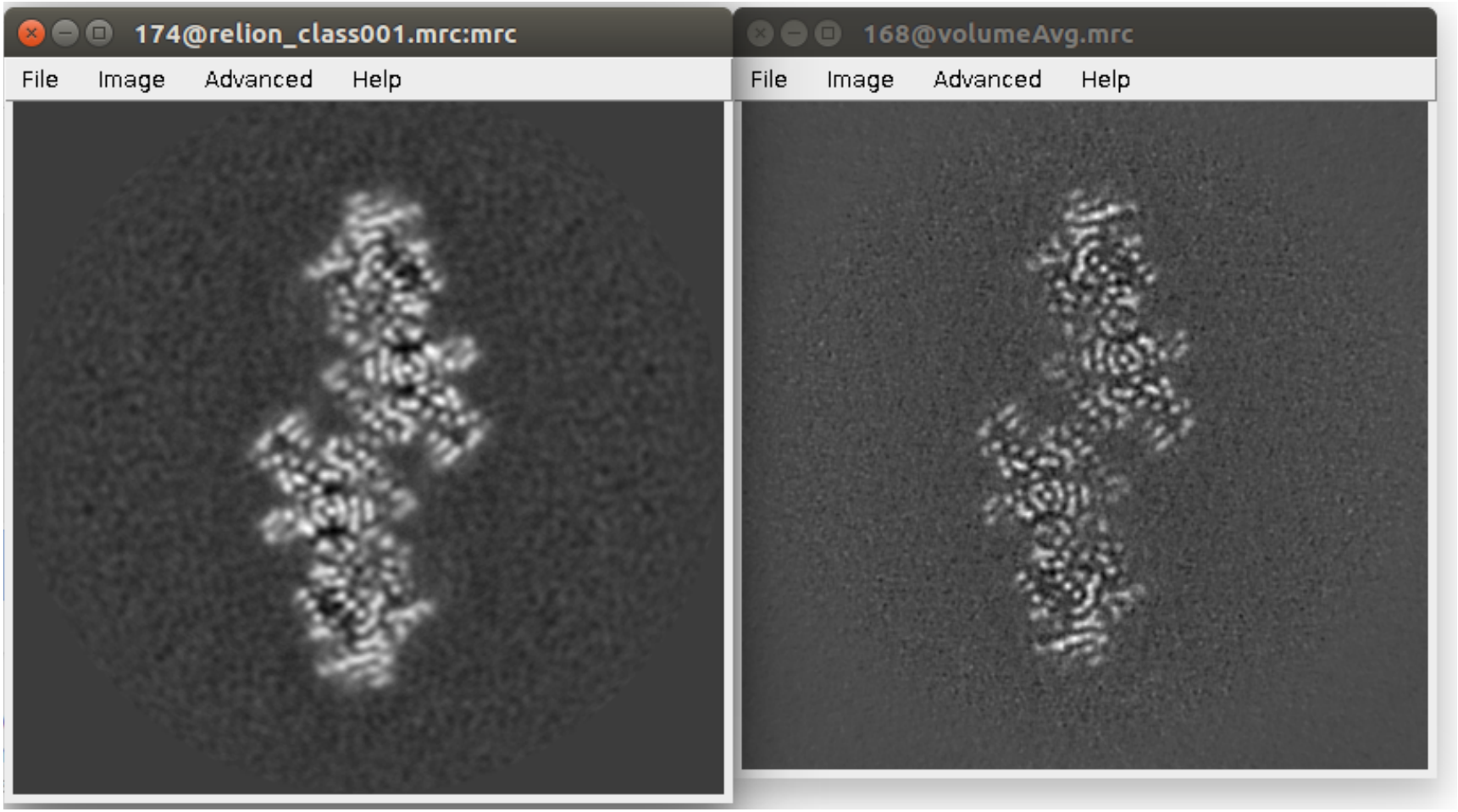


Figure 3
MonoRes slices
Slice 100
Slice 133
Slice 166
Slice 200
Figure 2
MonoRes slices
Slice 100
Slice 133
Resolutions Histo
Slice 166
Slice 200
Resolution (A)

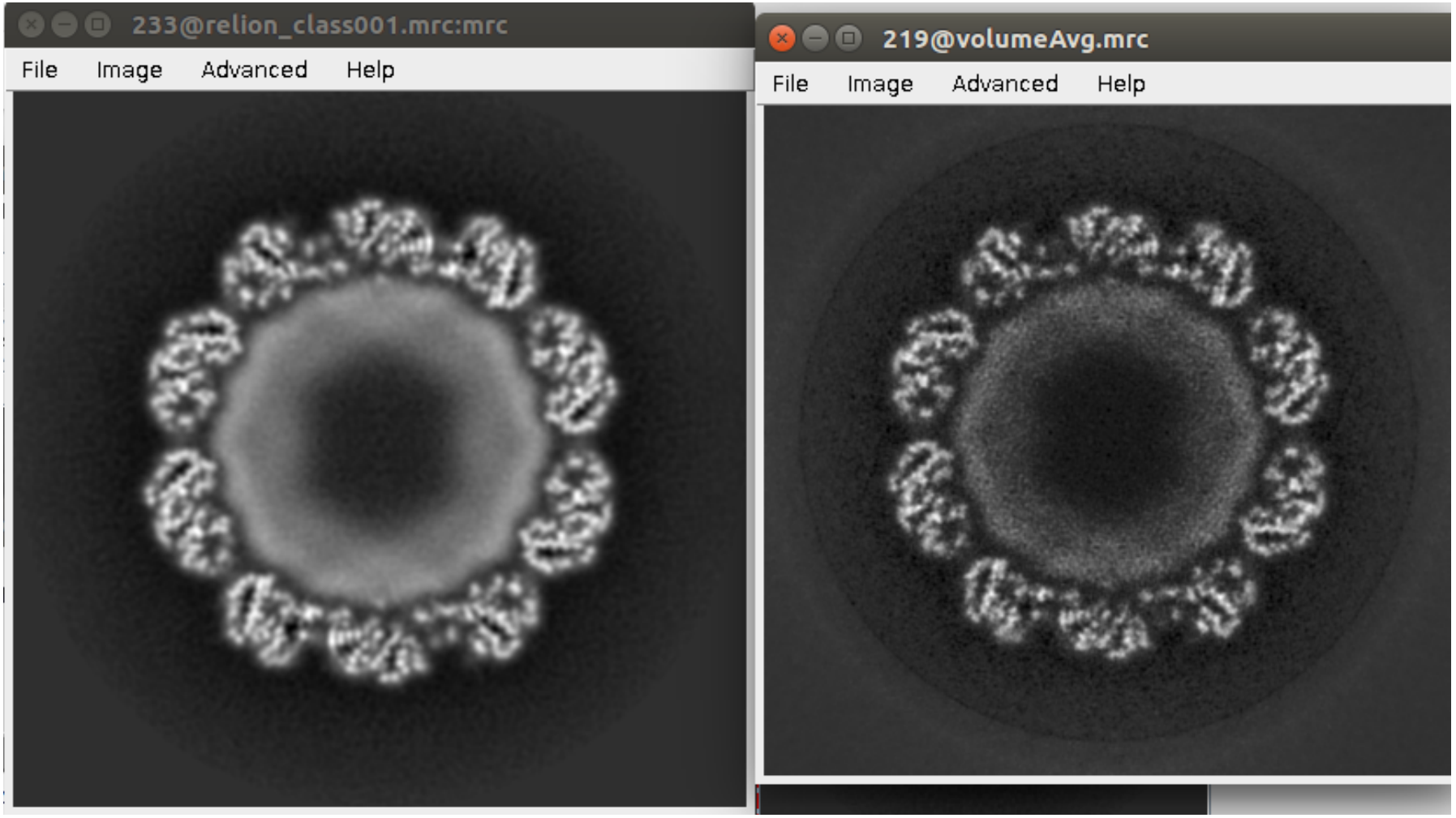
Results Ribosome



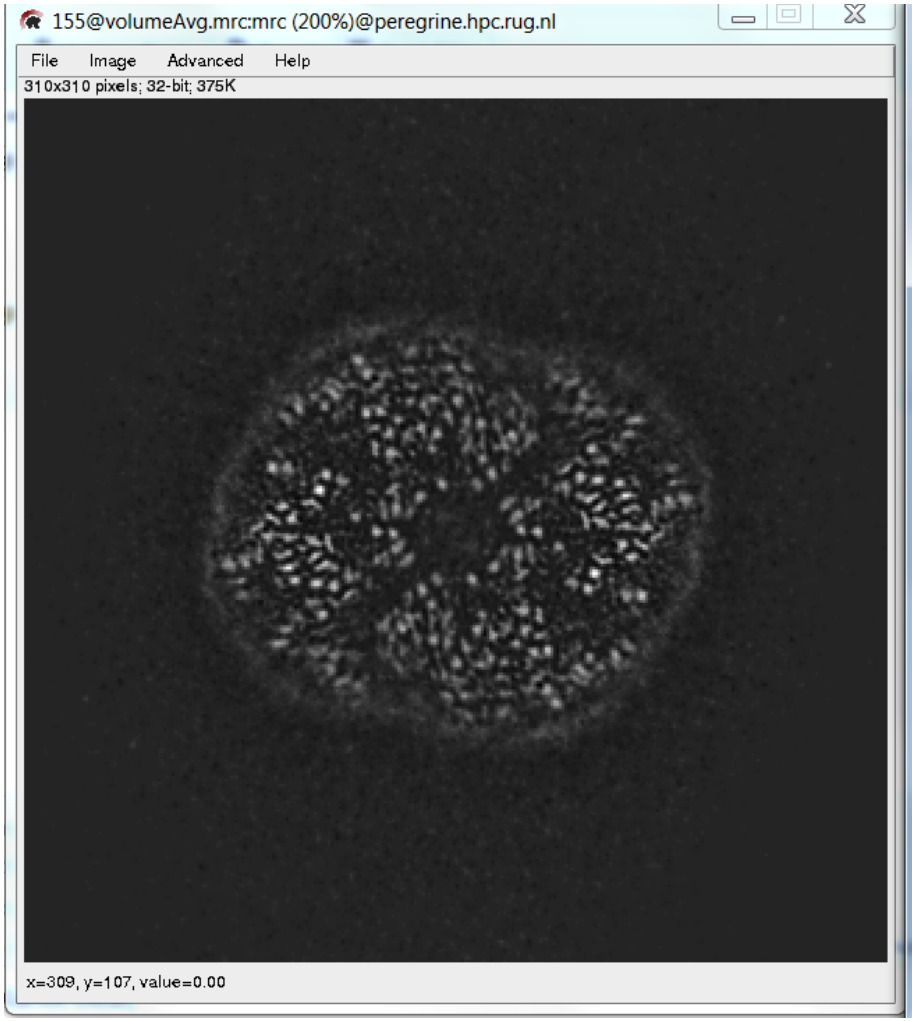
Results B-Galactosidase

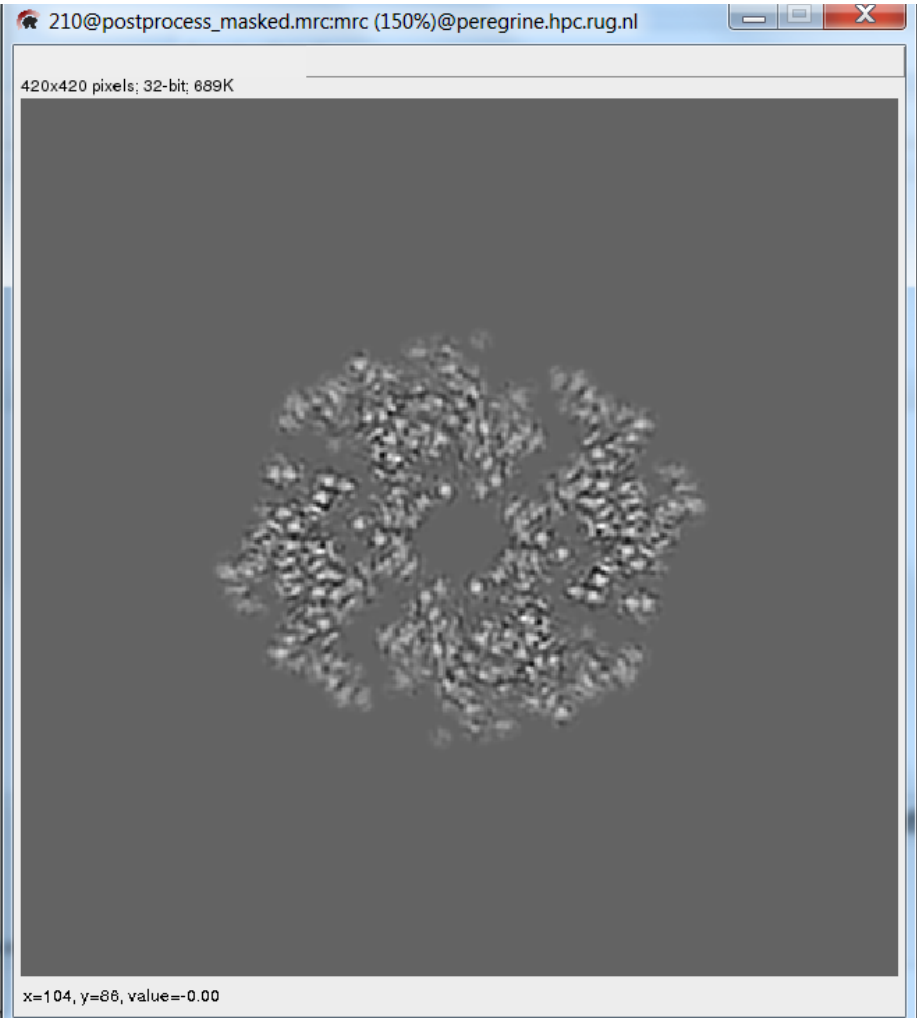


Results Virus



Results X


155@volumeAvg.mrc:mrc (200%)@peregrine.hpc.rug.nl
File
Image
Advanced
Help
310x310 pixels; 32-bit; 375K
x=309, y=107, value=0.00


210@postprocess_masked.mrc:mrc (150%)@peregrine.hpc.rug.nl
420x420 pixels; 32-bit; 689K
x=104, y=86, value=-0.00

Colorful Image Processing

xmipp3 - highres finished

xmipp3 - highres 2 finished

xmipp3 - highres 3 finished

highresSubset3 finished

xmipp3 - highres 4 finished

highresSubset4 finished

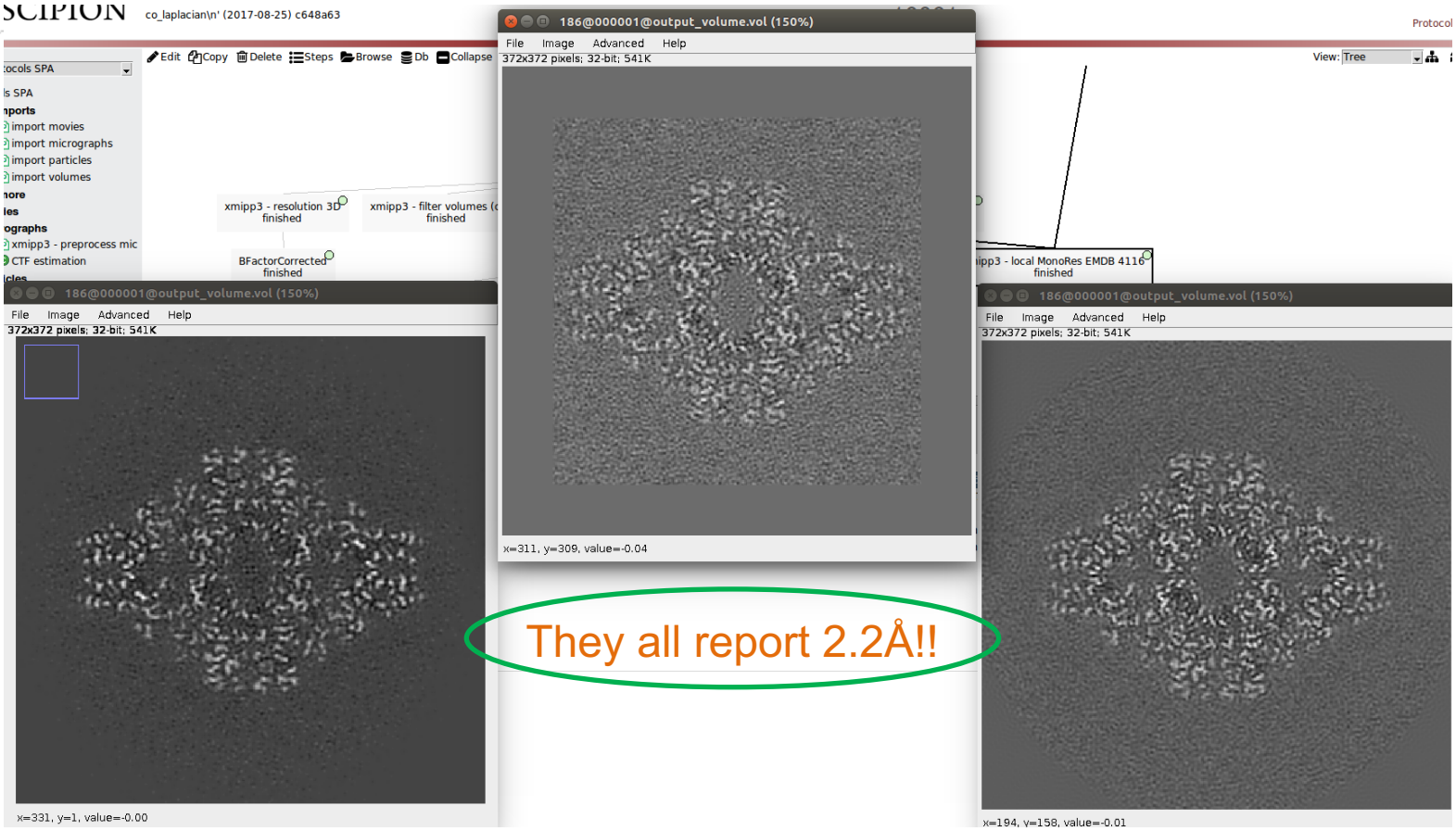
xmipp3 - highres 5 finished

highresSubset5 finished

xmipp3 - highres 6 finished

xmipp3 - highres 7 finished

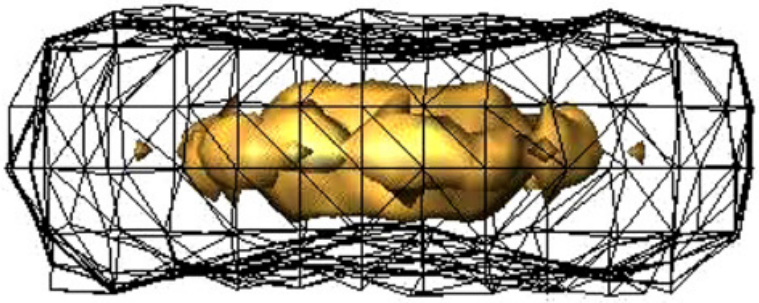
Clean3D (70k coords)

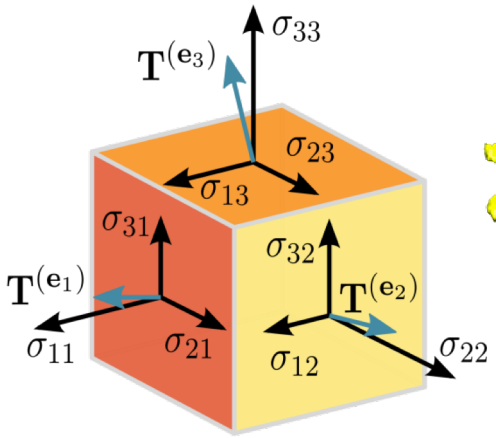


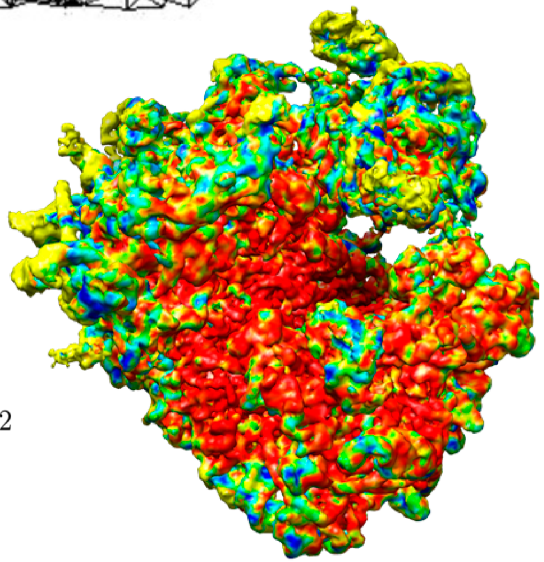
They all report 2.2Å!!

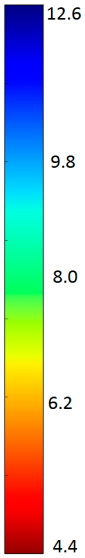
Clean3D (60k coords)

Resolution is local and directional



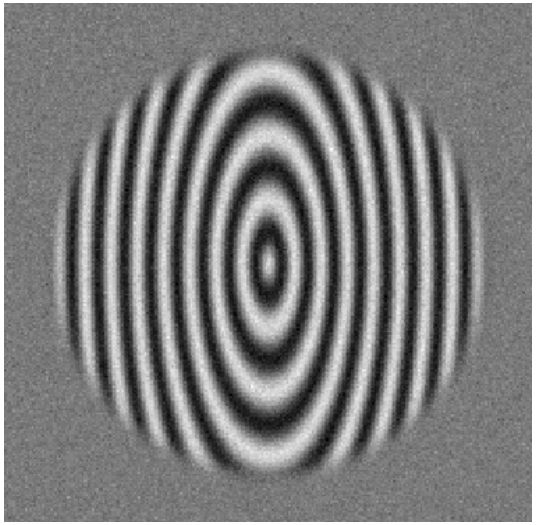


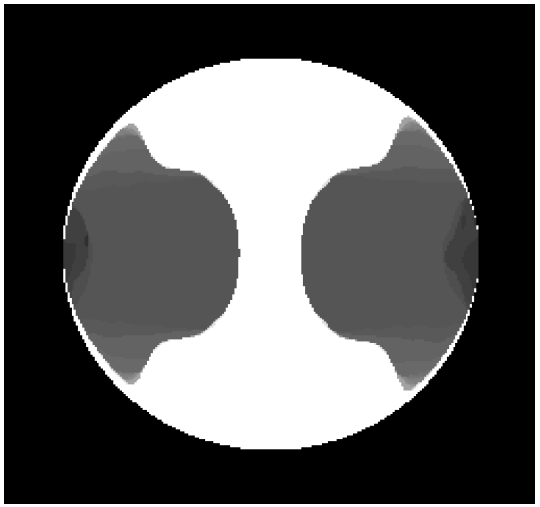




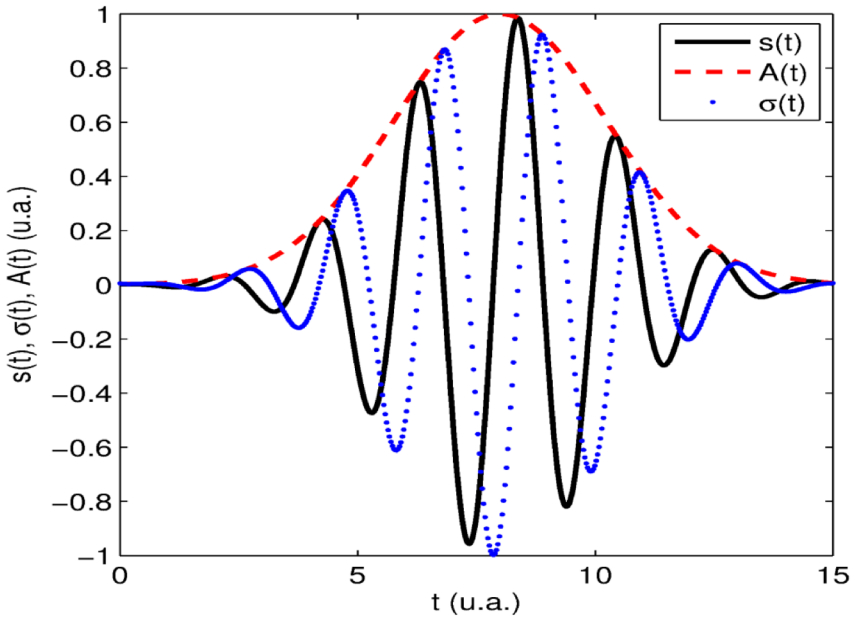
Resolution is local and directional

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





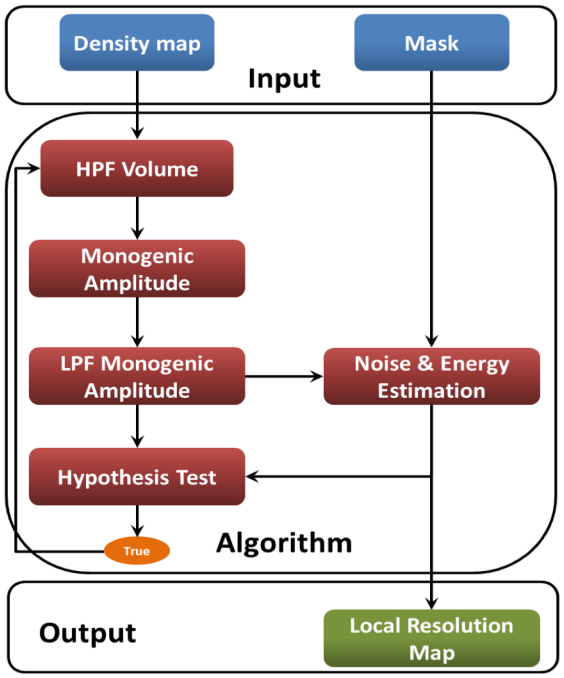
Monogenic signals

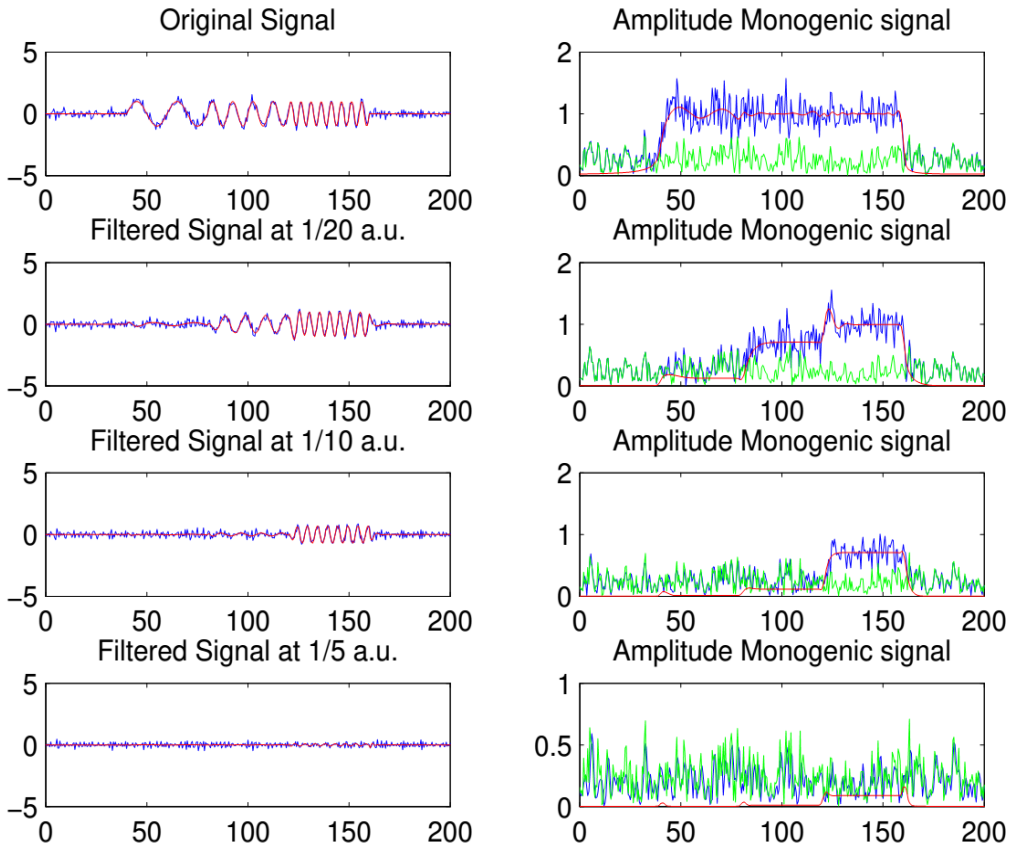


$$\widehat{s}_R(\omega) = -\frac{\omega}{\|\vec{\omega}\|}\widehat{s}(\vec{\omega}) = -\left(\frac{\omega_1}{\|\vec{\omega}\|}\widehat{s}(\vec{\omega}), \frac{\omega_2}{\|\vec{\omega}\|}\widehat{s}(\vec{\omega}), \ldots, \frac{\omega_N}{\|\vec{\omega}\|}\widehat{s}(\vec{\omega})\right)$$

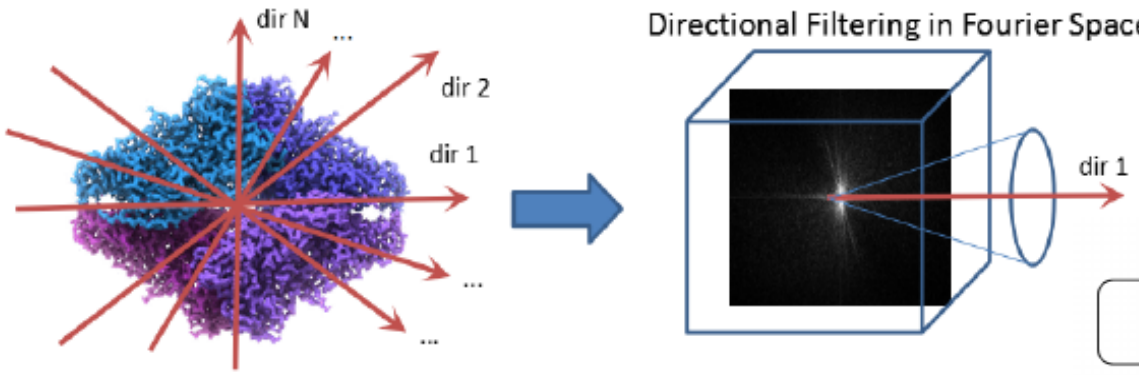
$$A(\vec{r}) = \sqrt{s^2(\vec{r}) + \sum_{j=1}^{N} s_{R_j}^2(\vec{r})}.$$

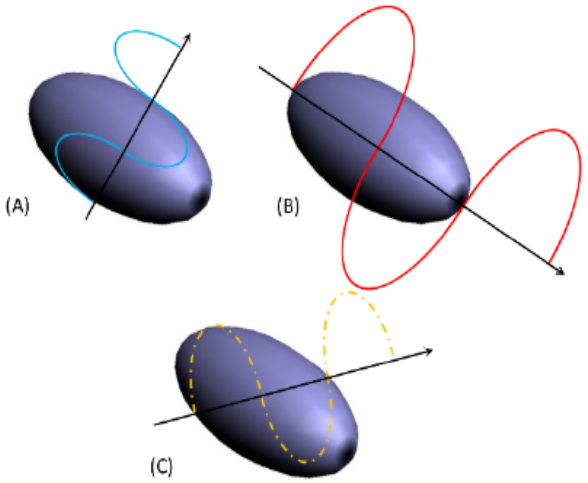
Monogenic local resolution

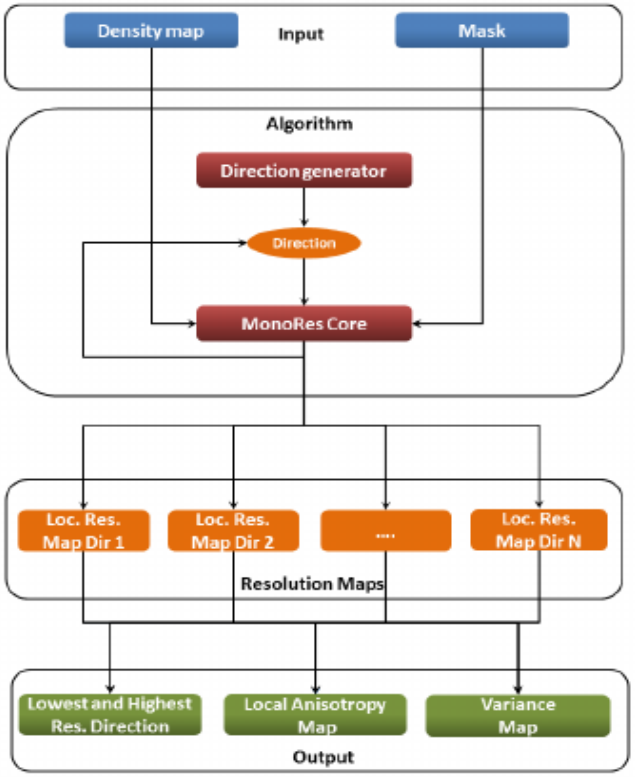




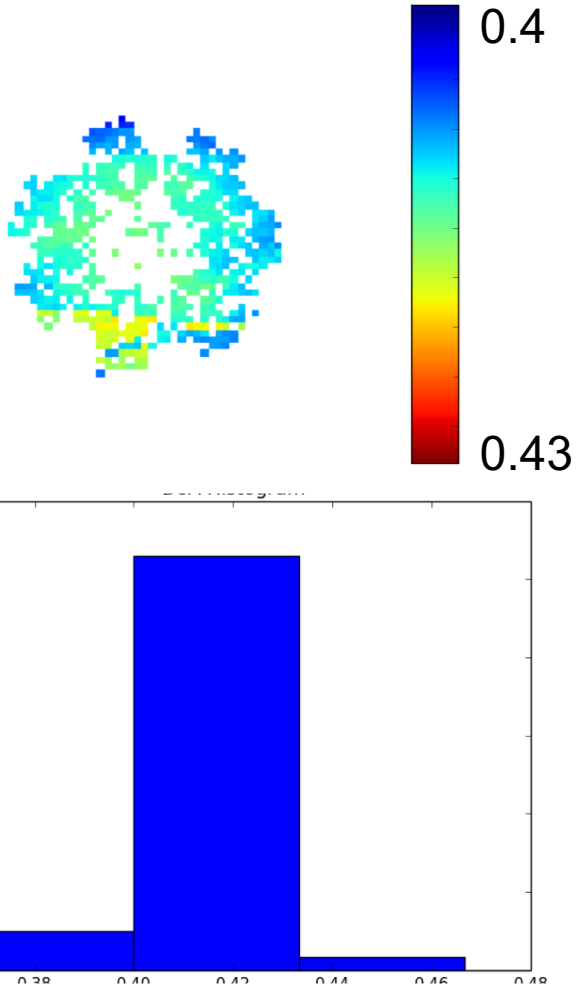
Monogenic local and directional resolution





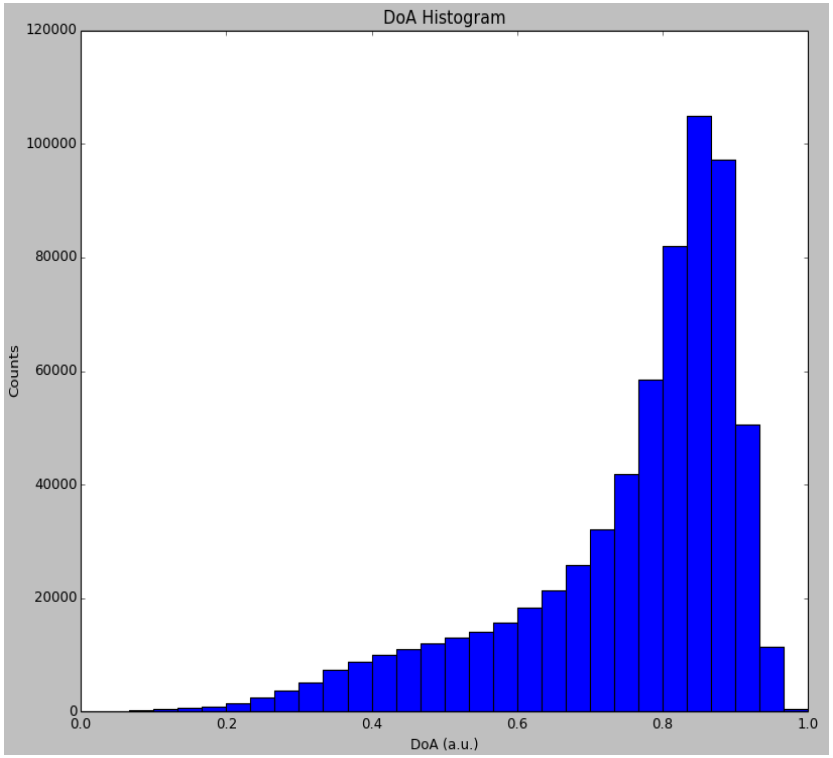


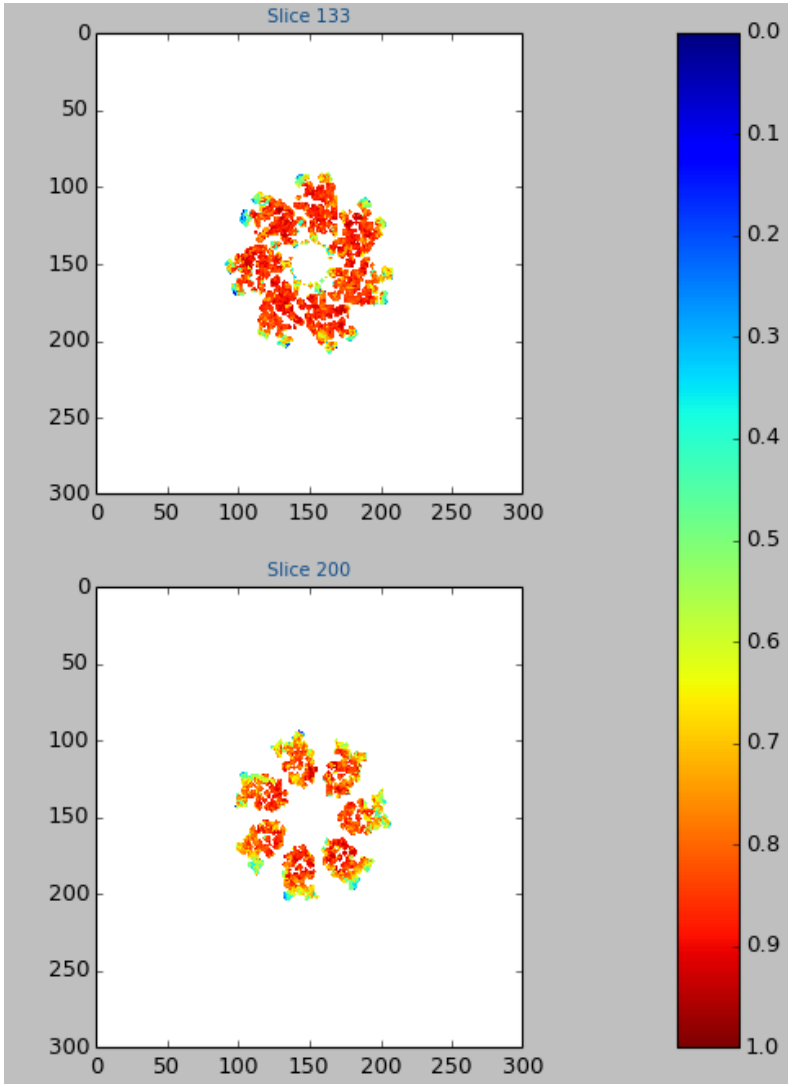
Results missing wedge



Results Proteasome 20S EMDB 6287

Resolution FSC 2.8A
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).