

SEMC Appion workshop series

September 19, 2019

PART II : SINGLE PARTICLE ANALYSIS



SIMONS ELECTRON
MICROSCOPY CENTER

New York Structural
Biology Center

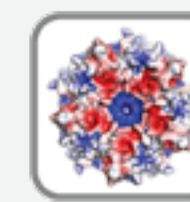




Department of Health



NMR
CoMD/
NMR



X-ray
NYX
@NSLS-II



Protein Production
COMPPA



NRRMM

Schedule

September 19, 2019 10am-2pm

A practical workshop that continues the single-particle analysis workflow from the Part 1 workshop. The topics include Appion advanced features, moving into commonly used packages and validation methods.

- **1. Workshop intro**
 - Overview of SPA processing and web tools
 - Advanced Appion features
 - *Photo*
- **2. Going from Appion to other packages**
 - RELION
 - cryoSPARC
- **3. Validation & additional tools** *[if time allows]*
 - FSC & 3D-FSC / local res & ResMap / other methods
- **4. Closing** *[Q&A with SEMC staff]*

What is Appion?

Appion is a "pipeline" for processing and analysis of EM images.

NRAMM

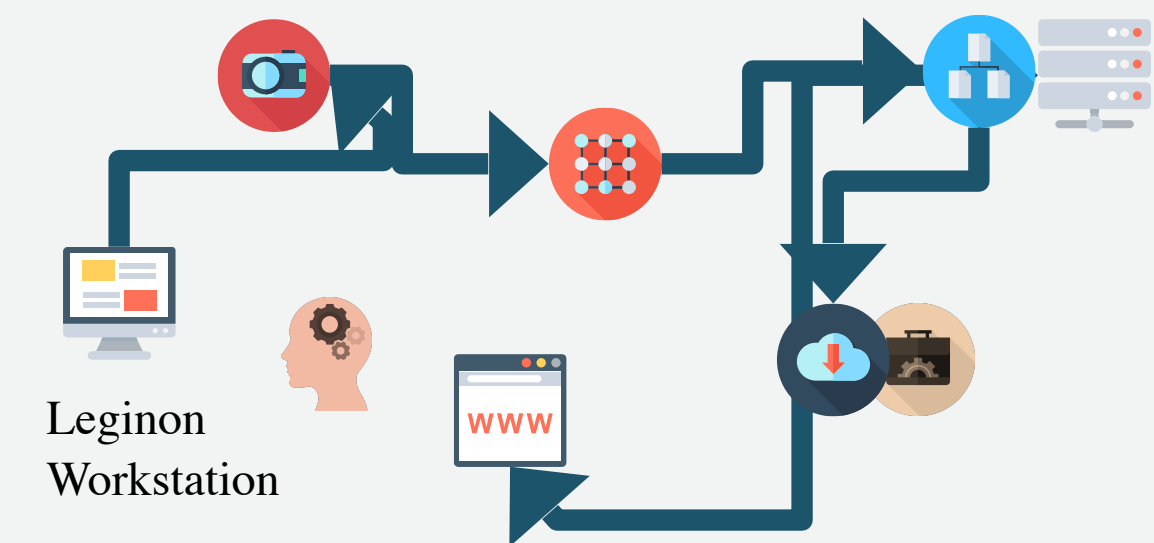
The overall mission of NRAMM is to develop, test and apply technology for automating and streamlining cryo-electron microscopy (cryoEM) for structural biology.



Automated sample prep



Automated image acquisition



Automated image processing

Simons Electron Microscopy Center

Acknowledgements

Some of this work was performed at the Simons Electron Microscopy Center at the New York Structural Biology Center that is supported by a grant from the Simons Foundation (Grant Number: SF349247) with additional support from: NIH S10 OD019994-01, the Agouron Institute (Grant Number: F00316), NIH S10 RR029300-01, NIH S10 RR017291-01, NYSTAR, and NIH C06 RR017528-01-CEM.

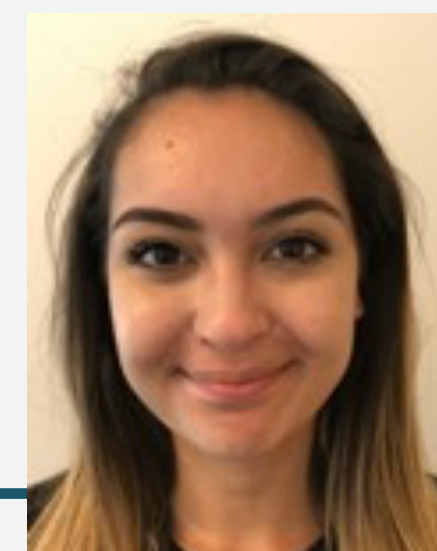
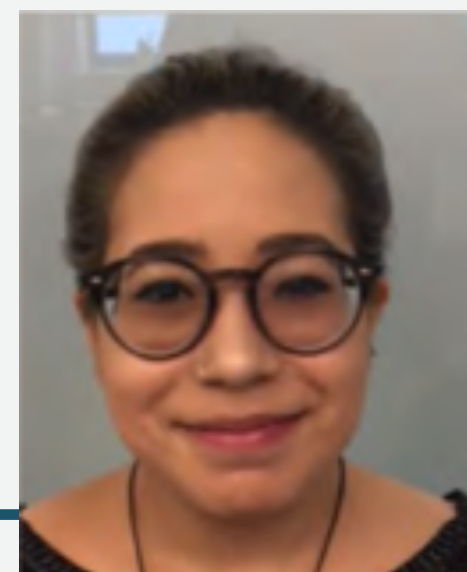
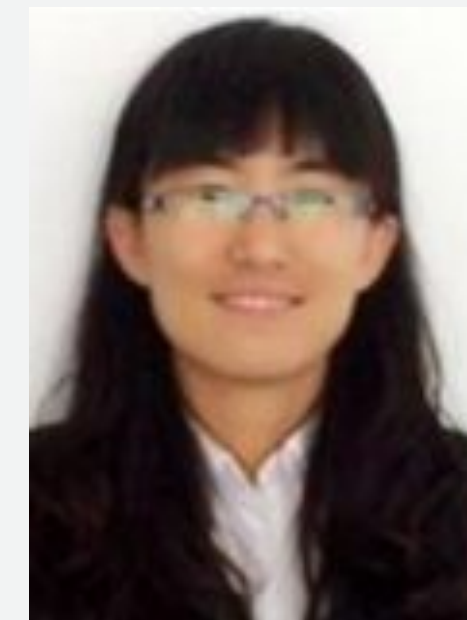
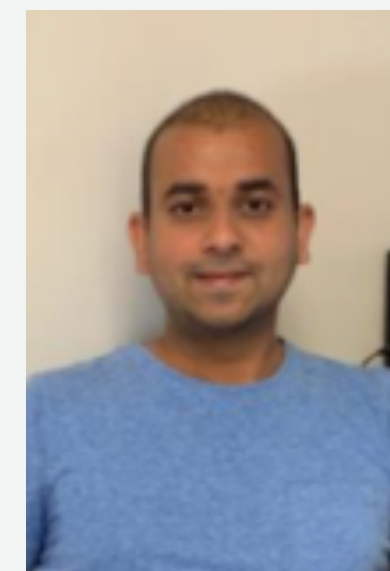
SEMC operations team

Directors



 @nysbc.org

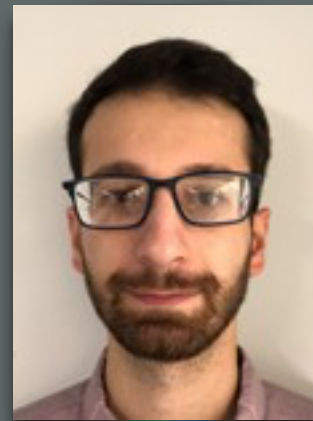
Staff



SEMC Appion development team



Swapnil
Bahktar



Shaker
Krit



Anchi Cheng



Sargis Dallakyan



Alex Noble

Tomography

SEMC Training

Type	Frequency	Title	Description
 feedback	yearly	General user meeting	State of the center
 knowledge	yearly	SEMC EM Course	Theory behind EM <i>Graduate level course</i>
 computation	quarterly	SEMC Appion workshop	Data processing workshops
 access	monthly	SEMC New User Orientation	Facility use and Safety Training Sample preparation [neg-stain & cryo] <i>Leginon intro/use of screening microscopes</i>
 troubleshooting	weekly	User Project Discussion Meetings	<i>Appointments: Tue @3pm, Thurs @3pm & @3:30pm</i>
 instrumentation	daily	Advanced Leginon use / 1-1 training appointments	Training for independent use of the microscopes and other SEMC resources

appion.org

http://emg.nysbc.org/redmine/projects/nysbc-semc-user-group/issues

Appion

Search: Appion

OverviewActivityRoadmapIssuesNew issueNewsDocumentsWikiForumsFilesRepositorySettings

Issues

Filters

☒ Status

open

☒ Tracker

is

Add filter

Bug

Options

Apply

Clear

Save

	#	Project	Tracker	Status	Priority	Subject	Author	Assignee
☐	3811	Appion	Bug	Assigned	Normal	dd alignment with motioncorr fails with larger movies or larger shifts	Anchi Cheng	Anchi Cheng
☐	3809	Appion	Bug	New	Normal	When users are identified in the path, upload image is put into base leginon directory, not user directory if leginon.cfg is not set in user home directory	Anchi Cheng	
☐	3808	Appion	Bug	Assigned	Normal	uploadImages.py does not set "frame path" field in SessionData	Anchi Cheng	Anchi Cheng
☐	3804	Appion	Bug	Assigned	Normal	DogPicker fails to display result images when testing through job submission	Carl Negro	Carl Negro
☐	3802	Appion	Bug	New	Normal	slow aligned particle insertion caused time out in uploadMaxlikeAlignment.py	Anchi Cheng	

Issues

View all issues

Summary

Custom queries

Appion 2.2 Test Matrix

Appion 3.2 bug fixes

Appion 3.2 feature

Appion Known Bugs

Appion Test Suite

Unassigned issues

appion.org

Register an account

Appion

Search: Appion

Overview

Activity

Roadmap

Issues

New issue

News

Documents

Wiki

Forums

Files

Repository

Settings

Appion Forums

Edit

★ Watch

🔒 Lock

🔥 Rename

🗑 Delete

⏪ History

Appion and Leginon use a shared forum

Appion is closely related to its sister product, Leginon. They are built on the same code base, require a similar installation procedure, and share user, project and administration management tools. Because of these similarities, Appion and Leginon also share a single user Forum.

Link to the forum

The Forum can be found in the [Leginon product's Forums tab](#).

Link to user registration

To post or reply to a message, you must be logged into this site. If you have not yet registered, first go to the [Registration page](#).

New file

Wiki

Start page

Index by title

Index by date

appion.org

http://emg.nysbc.org/redmine/projects/nysbc-semc-user-group/issues

Appion

Search: Appion

Overview

Activity

Roadmap

Issues

New issue

News

Documents

Wiki

Forums

Files

Repository

Settings

Issues

Filters

☒ Status

open

☒ Tracker

is

Add filter

Bug

Options

Apply

Clear

Save

	#	Project	Tracker	Status	Priority	Subject	Author	Assignee
☐	3811	Appion	Bug	Assigned	Normal	dd alignment with motioncorr fails with larger movies or larger shifts	Anchi Cheng	Anchi Cheng
☐	3809	Appion	Bug	New	Normal	When users are identified in the path, upload image is put into base leginon directory, not user directory if leginon.cfg is not set in user home directory	Anchi Cheng	
☐	3808	Appion	Bug	Assigned	Normal	uploadImages.py does not set "frame path" field in SessionData	Anchi Cheng	Anchi Cheng
☐	3804	Appion	Bug	Assigned	Normal	DogPicker fails to display result images when testing through job submission	Carl Negro	Carl Negro
☐	3802	Appion	Bug	New	Normal	slow aligned particle insertion caused time out in uploadMaxlikeAlignment.py	Anchi Cheng	

Issues

View all issues

Summary

Custom queries

Appion 2.2 Test Matrix

Appion 3.2 bug fixes

Appion 3.2 feature

Appion Known Bugs

Appion Test Suite

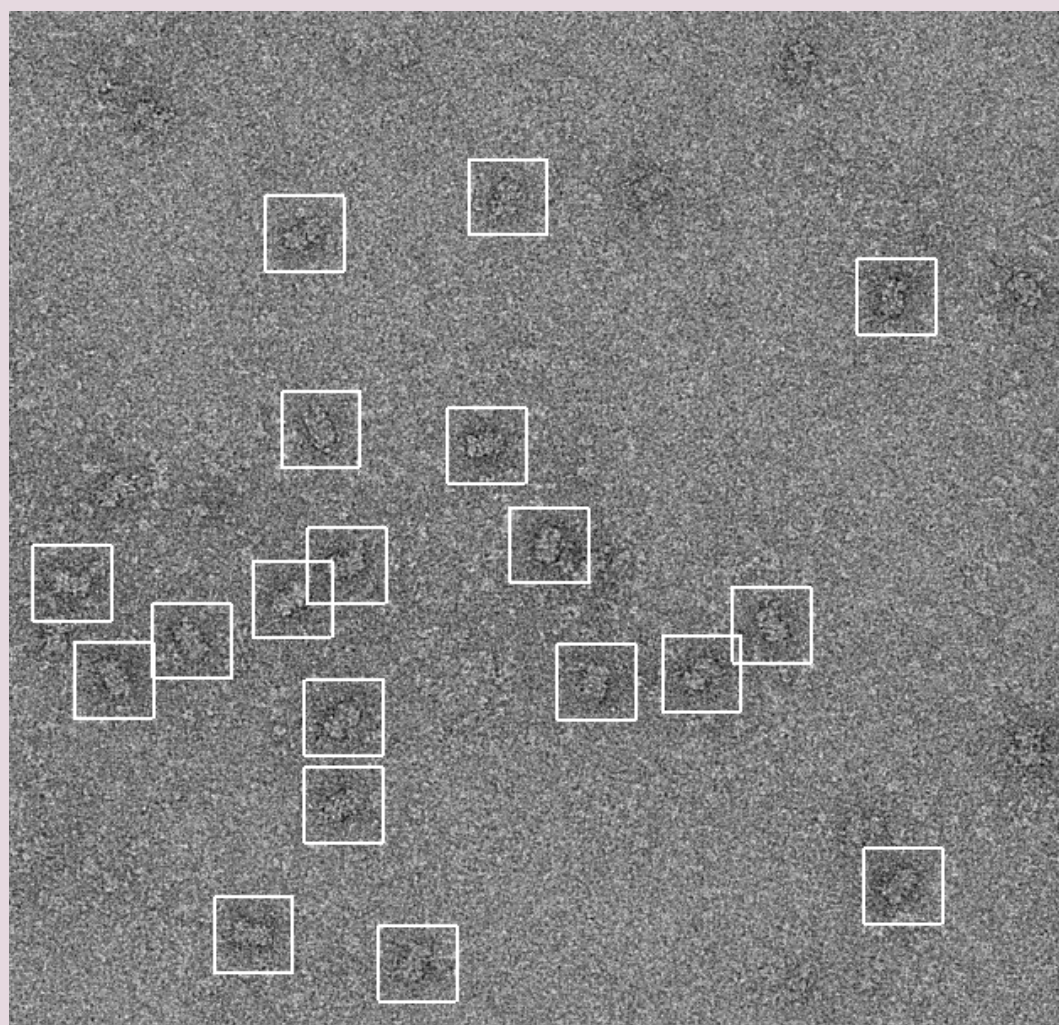
Unassigned issues

Appion workshops

Single particle

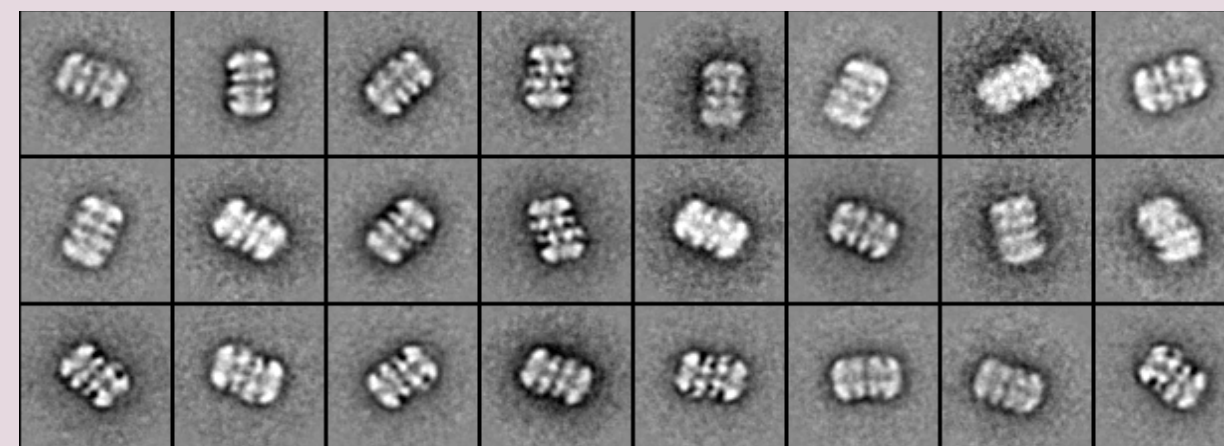
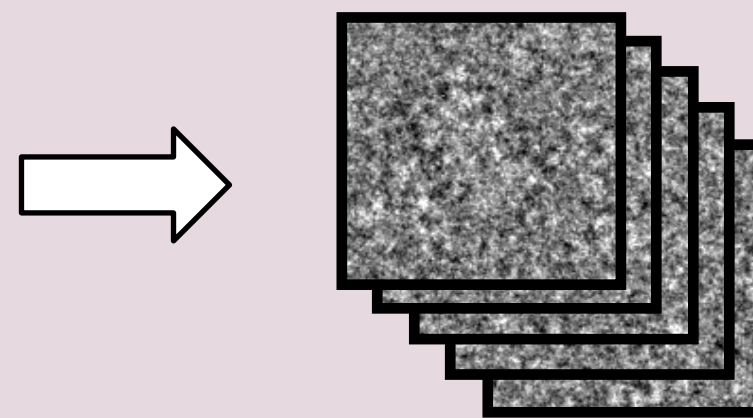
Part I

Tools to generate a particle stack and an initial model from micrographs.



Part II

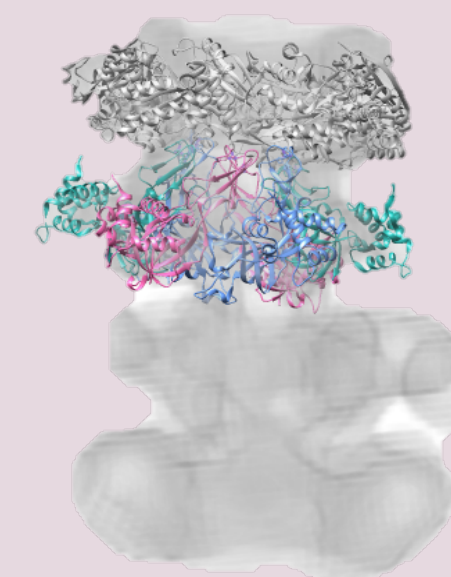
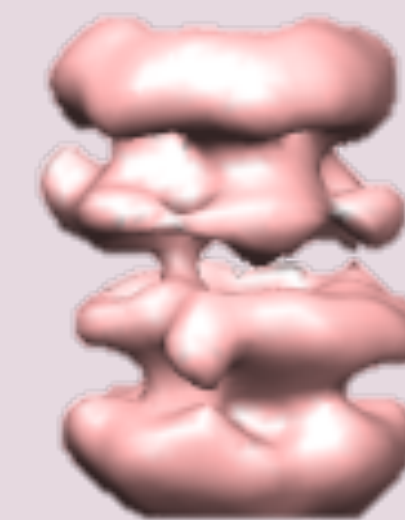
**Advanced Appion tools,
3D tools and validation.**



Tomography

Part III

Tomography tools.



Data generation at SEMC

EM instrumentation

FEI Titan Krios#1 / #2 / #3
Falcon3 x3 | K2 x3



FEI Tecnai F20
DE20 | TVIPS 4K CMOS



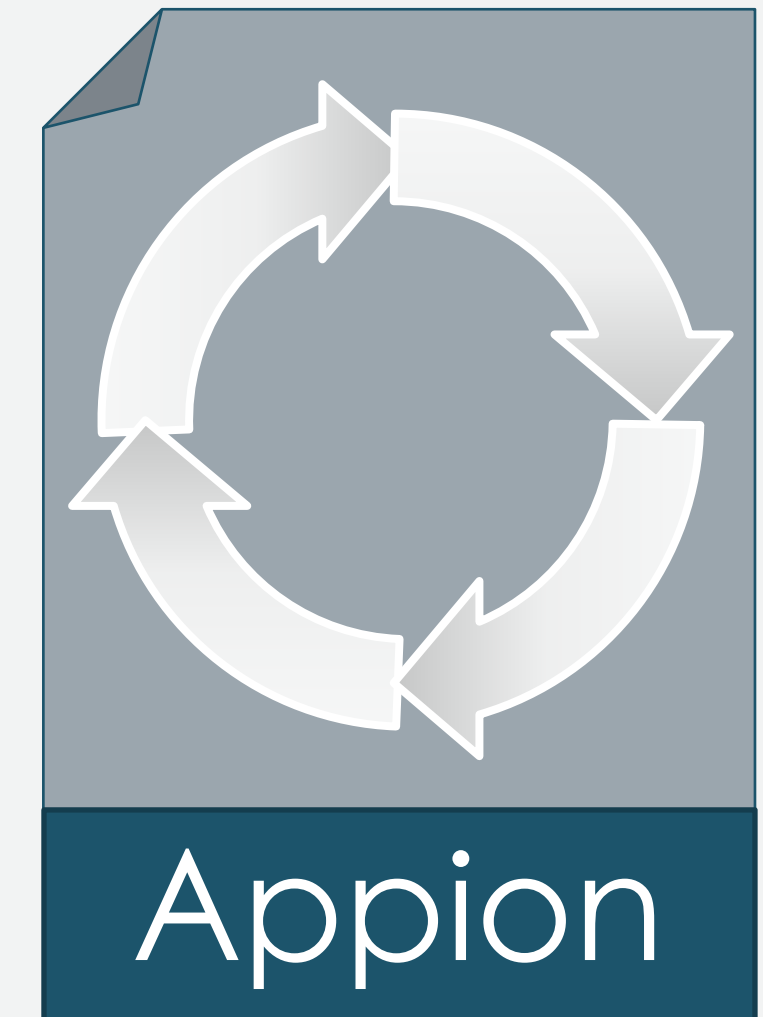
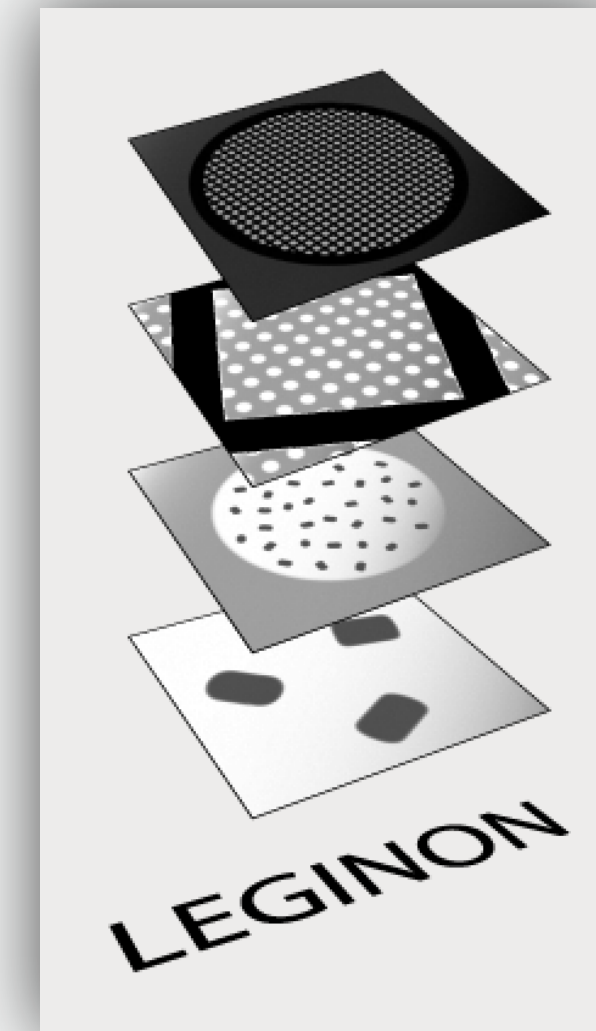
FEI Tecnai Biotwin
TVIPS 4K CMOS



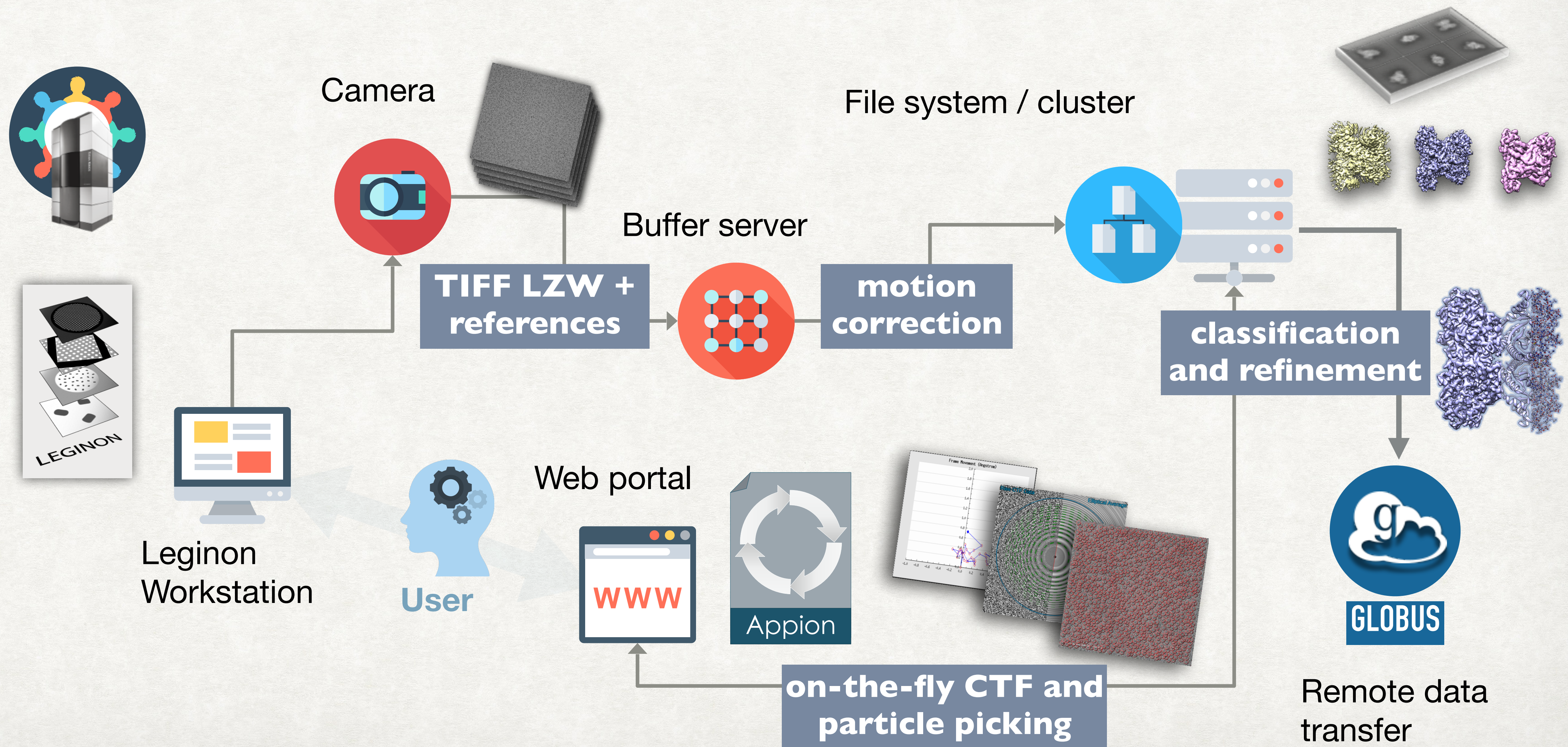
JEOL 1230
Gatan US4000 CCD



FEI Helios 650
Quorum cryostage



DATA PIPELINE



Remote data transfer



Create an account at
globusid.org

Login at
globus.org

Transfer files

Remote
data
transfer

The collage consists of four overlapping screenshots from the Globus website:

- Top Left:** A screenshot of the 'Create a Globus ID' page. It features the Globus logo and the text 'Research data management'. Below this, there are input fields for 'E-mail' (with 'user@example.edu' entered) and 'Organization'. A red box highlights the radio button selection for 'non-profit, research or educational purposes' under the heading 'This account will be used for'. At the bottom, there is a 'Create ID' button.
- Top Right:** A screenshot of the 'Transfer Files' interface. It shows a navigation bar with 'Manage Data', 'Publish', 'Groups', 'Support', and 'Account'. Below the navigation bar, there are tabs for 'Transfer Files', 'Activity', 'Endpoints' (which is highlighted with a red box), 'Bookmarks', and 'Console'. The main area contains input fields for 'Endpoint' and 'Path' with 'Go' buttons.
- Bottom Left:** A screenshot of the 'Manage Endpoints' page. It shows a section titled 'Add Globus Connect Personal Endpoint'. Step 1 is 'Create & Copy Your Globus Connect Personal Setup Key', which includes a 'Display Name' field (with 'Laura's macbook' entered) and a 'Generate Setup Key' button. Step 2 is 'Download & Install Globus Connect Personal', with buttons for 'for Mac OS X', 'for Linux', and 'for Windows'.
- Bottom Right:** A screenshot of an 'Endpoint' details view. It shows the endpoint name 'nysbc', a link to 'NYSBC - SEMC', and the owner 'nysbc@globusid.org'.

Start your web browsers

emgweb.nysbc.org

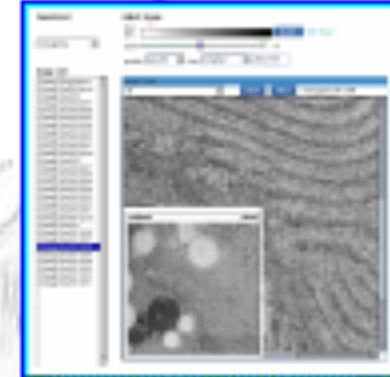
Login:	test
Password:	appion2019

Appion & Leginon webtools

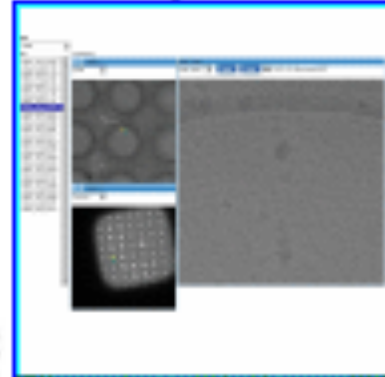
emgweb.nysbc.org

EMGWEB Appion and Leginon Tools

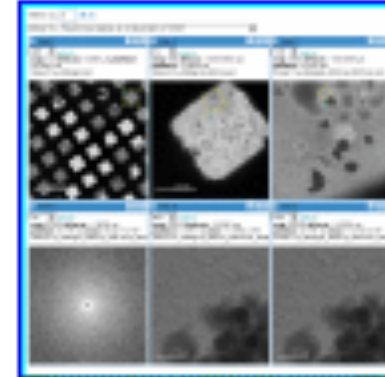
Image Viewer



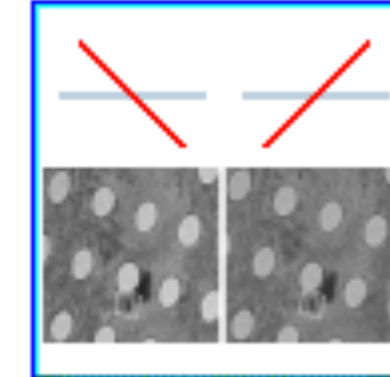
3 Way Viewer



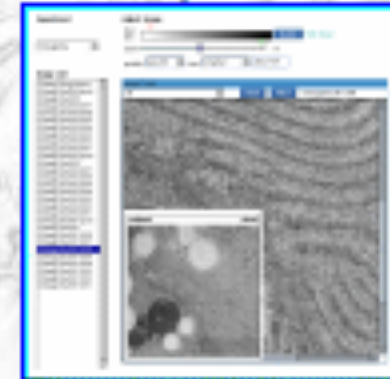
LOI



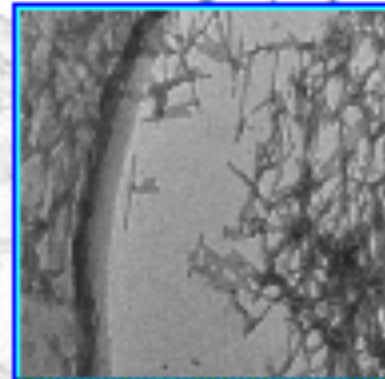
RCT



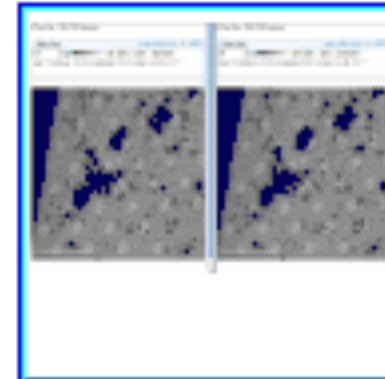
2 Way Viewer



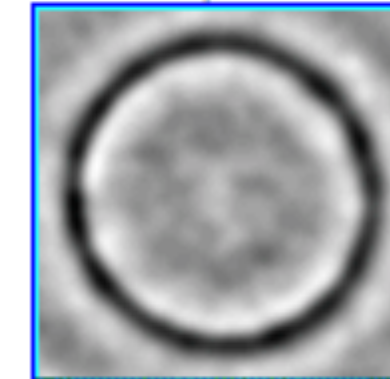
Tomography



Dual Viewer



Hole Template viewer



Administration



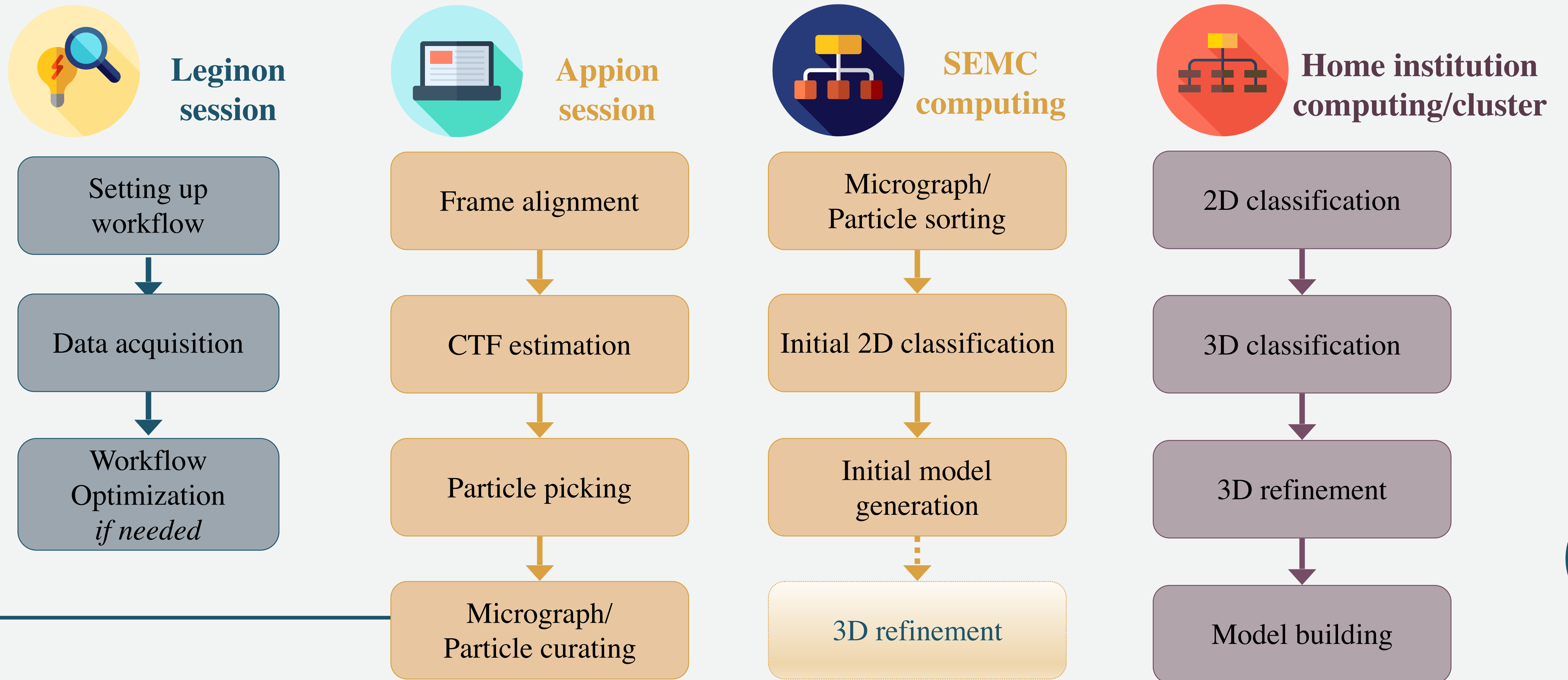
Project DB



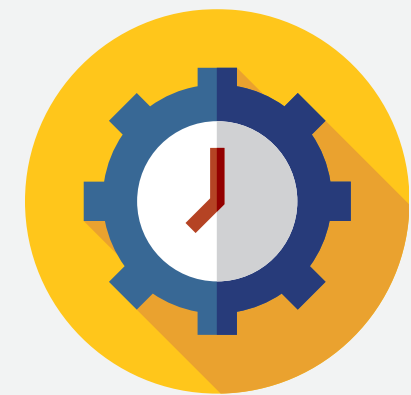
Example: Single-particle workflow

During EM session

After EM session



Processing workflow



During EM session

 **Appion Data Processing**

Project: SEMC - Glutamate Dehydrogenase (C
Session: 17jan26c - Krios2 gdh phase plat
Image Path: /gpf/legion/rice/17jan26c/rawdata

Hide | Expand | Contract

- ▶ Object Selection : 1
- ▶ CTF Estimation
- ▶ Stacks : 3
- ▶ Particle Alignment
- ▶ Ab Initio Reconstruction
- ▶ Refine Reconstruction
- ▶ Helical Processing
- ▶ Tomography
- ▶ Direct Detector Tools
- ▼ Import tools

Upload particles

Upload template

23 available

Upload CTF

Upload stack

Upload reconstruction

PDB to Model

EMDB to Model

Upload model

1 available



Appion
session

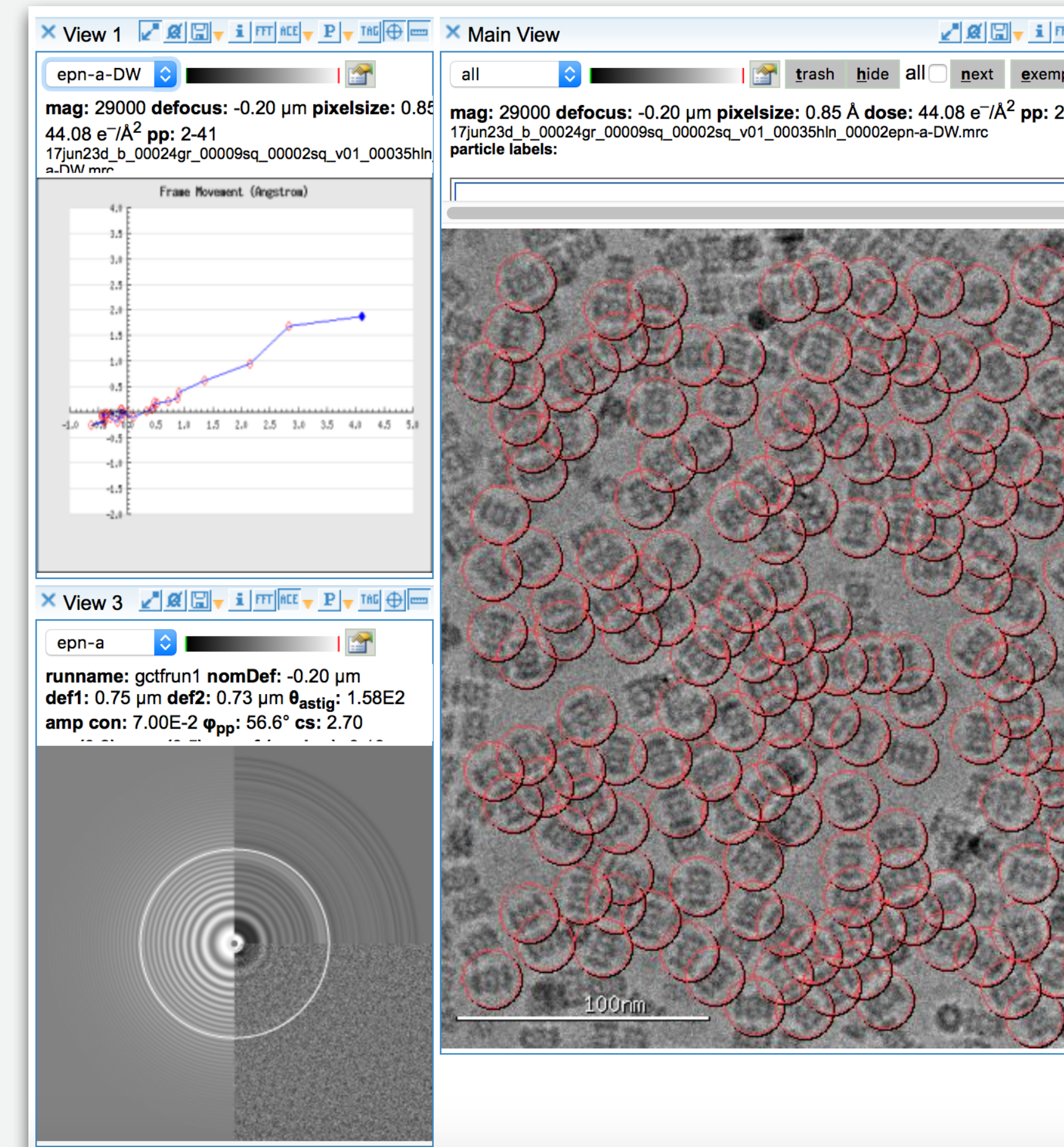
Frame alignment

CTF estimation

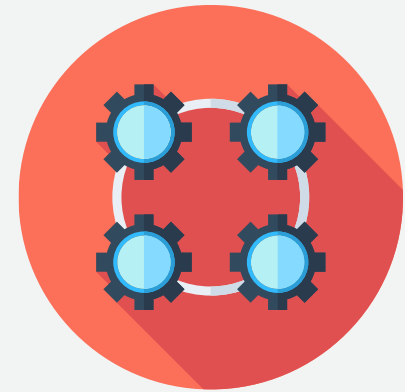
Particle picking

Micrograph/
Particle curating

3/4 users who use Legion also have Appion sessions



Who does the processing workflow?

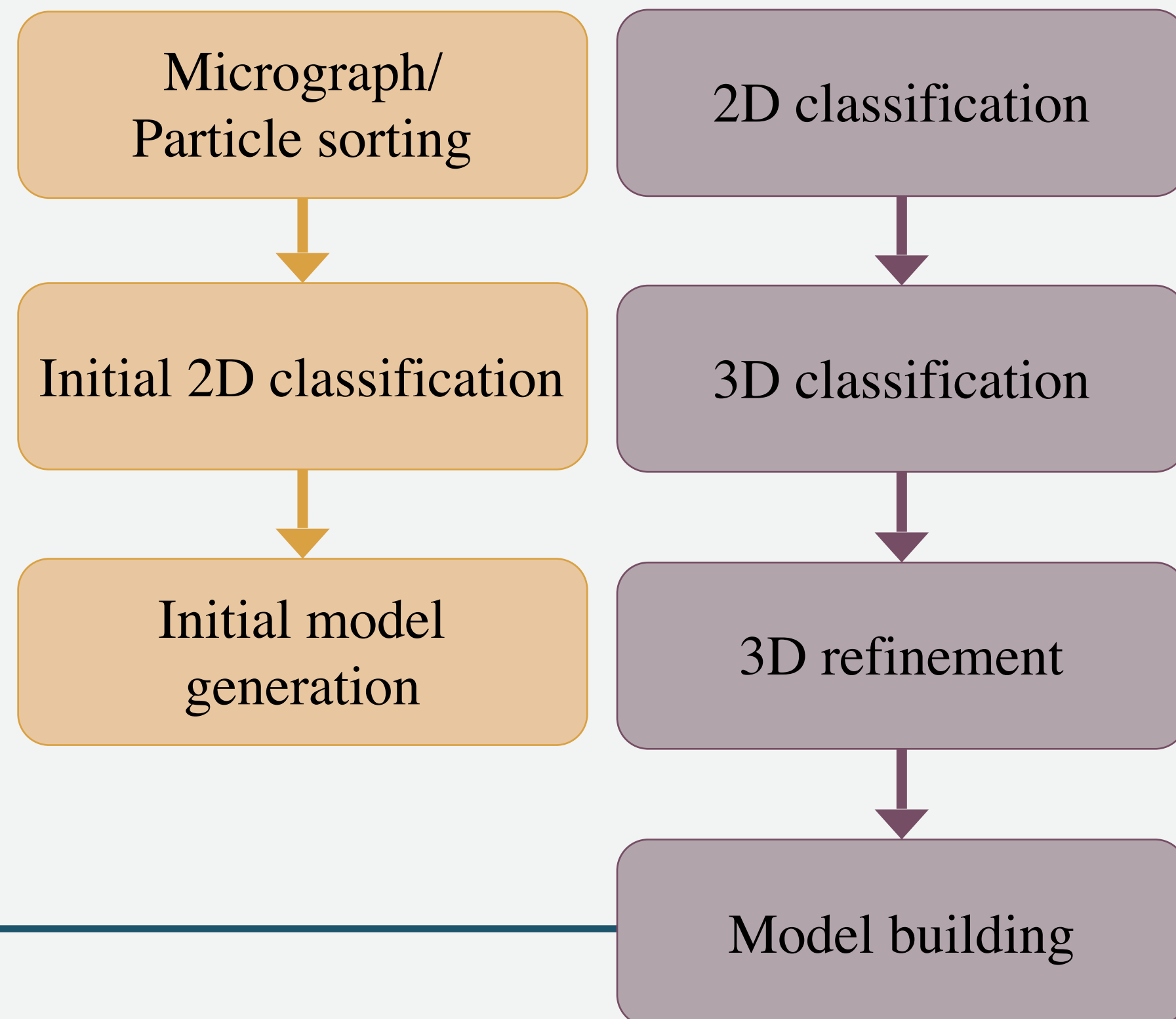


After EM session

Appion web interface

SEMC computing

Home institution
computing/cluster



Cluster Stack List

Project: SEMC - Glutamate Dehydrogenase (GDH) (256)
Session: **16dec06e** - 1.68 mg/mL GDH + 0.001% DDM (1.3/1.4 UltrAu 300mesh)
Image Path: /gpfs/legionon/lkim/16dec06e/rawdata

Hide | Expand | Contract

▼ Object Selection : 3

Select Particle Picker...
4 complete

Repeat an image loop run

▼ CTF Estimation

Estimate the CTF...
5 complete
1 running

Transfer results to another preset

Repeat an image loop run

▼ Stacks : 9

Stack creation
11 complete

more stack tools

▼ Particle Alignment

Select Particle Alignment...
2 complete

Run Feature Analysis...
3 complete

Run Particle Clustering...
4 complete

Run MaskItOn

Template Stacks

▼ Ab Initio Reconstruction

OptiMod Common Lines

EMAN Common Lines

SIMPLE Common Lines

▼ Refine Reconstruction

Run Single-Model Refinement...
1 jobs ready to run

Run Multi-Model Refinement...

▼ Helical Processing

Helical Image Processing

▼ Direct Detector Tools

Select Frame Alignment

Cluster Stack List

Show Composite Page

Clustering Info: **kerden3** (ID: 13) with 1 clusters

Type: Xmipp KerDen SOM

View montage of self-organizing map
View montage as a stack for further processing (ID 39)

Clustering Info: **coran1** (ID: 12) with 2 clusters

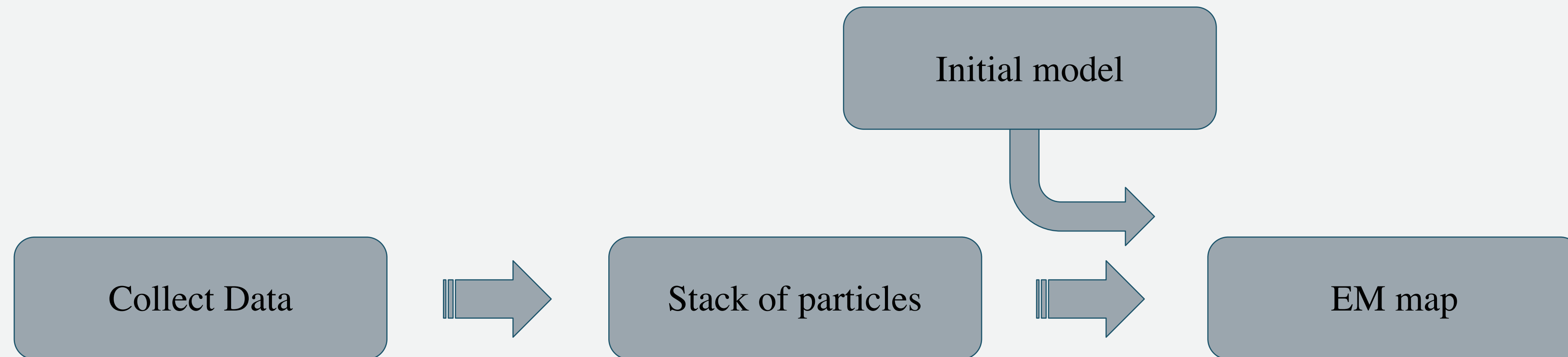
Type: SPIDER Coran
Method:
Factor list:

160 Class Averages [variance] (ID 41)
80 Class Averages [variance] (ID 40)

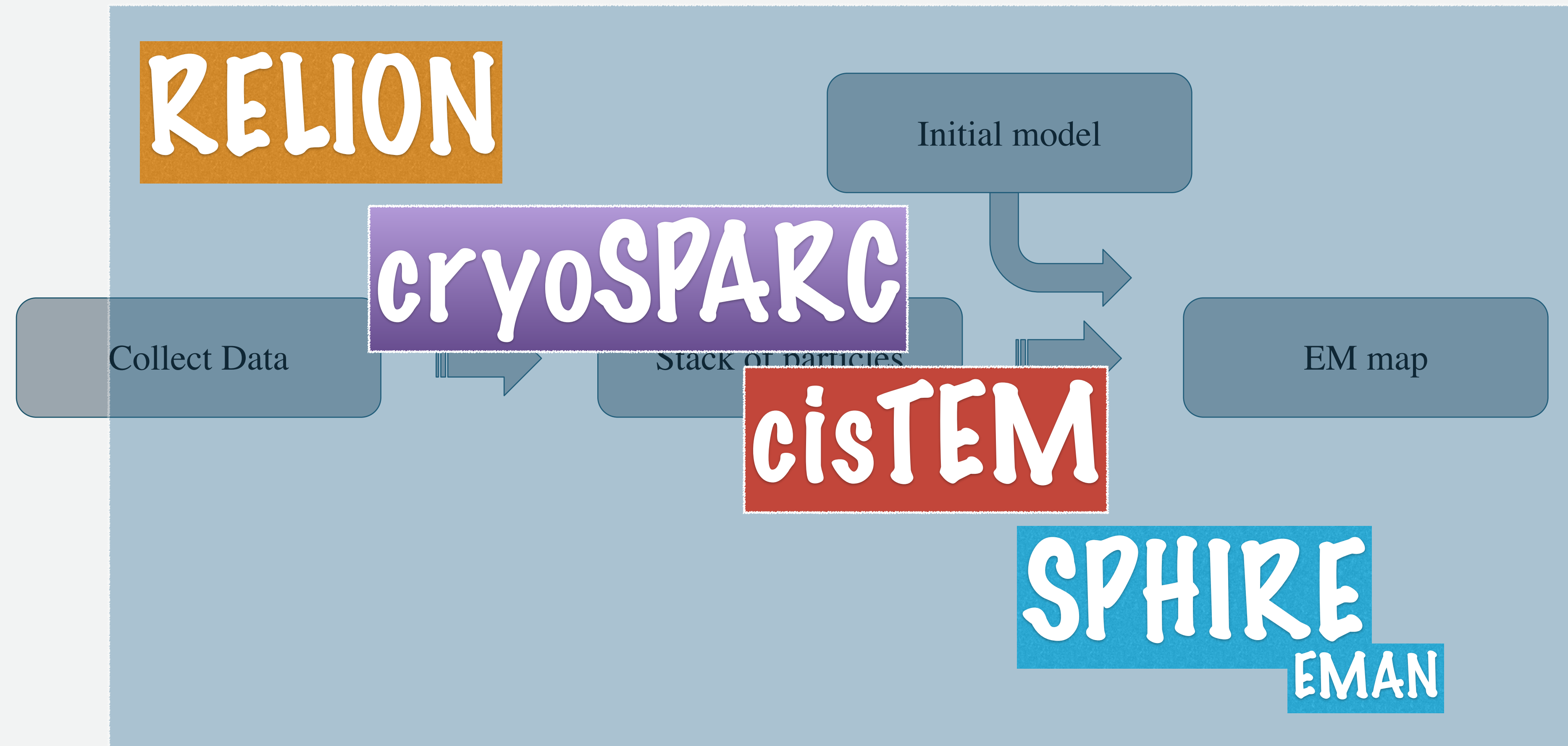
Clustering Info: **kerden1** (ID: 11) with 1 clusters

Type: Xmipp KerDen SOM

Single particle analysis

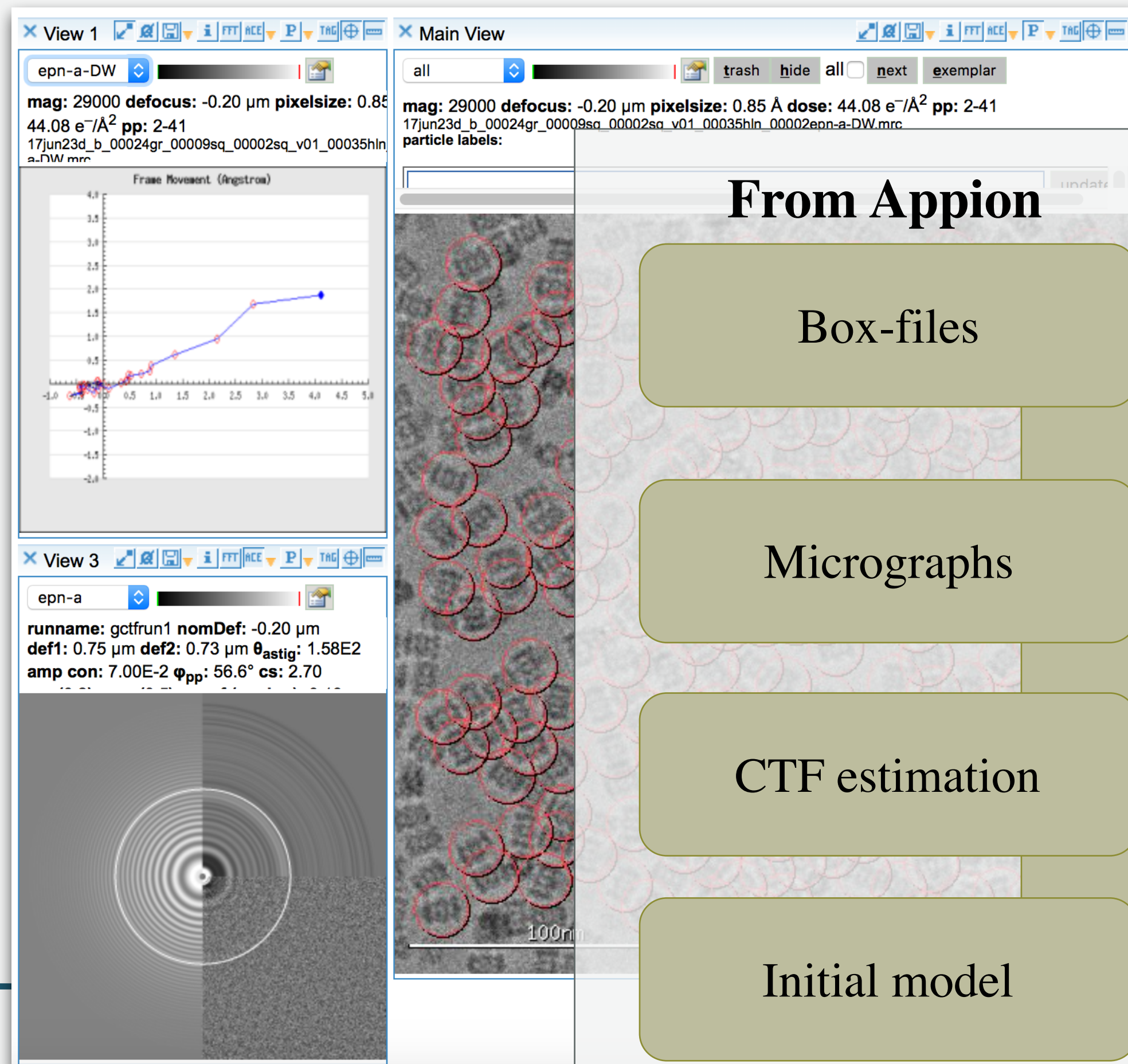


Single particle analysis



Single particle analysis

3/4 users who use Leginon also have Appion sessions



From Appion

Box-files

Micrographs

CTF estimation

Initial model

Initial model

RELION/cryoSPARC

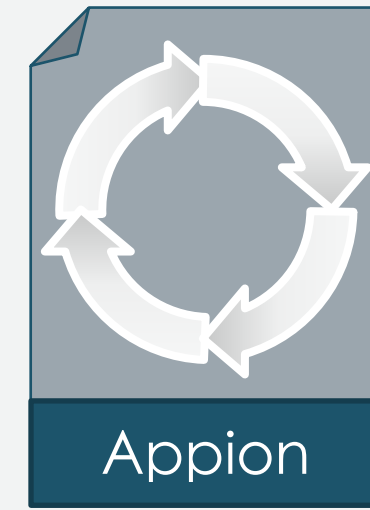
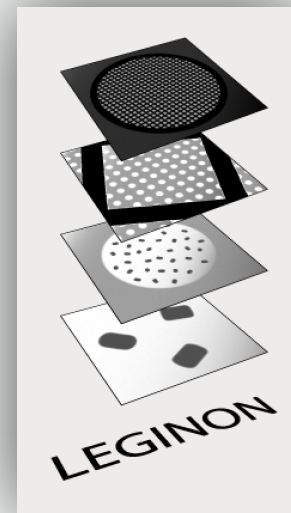
Particle extraction

EM map

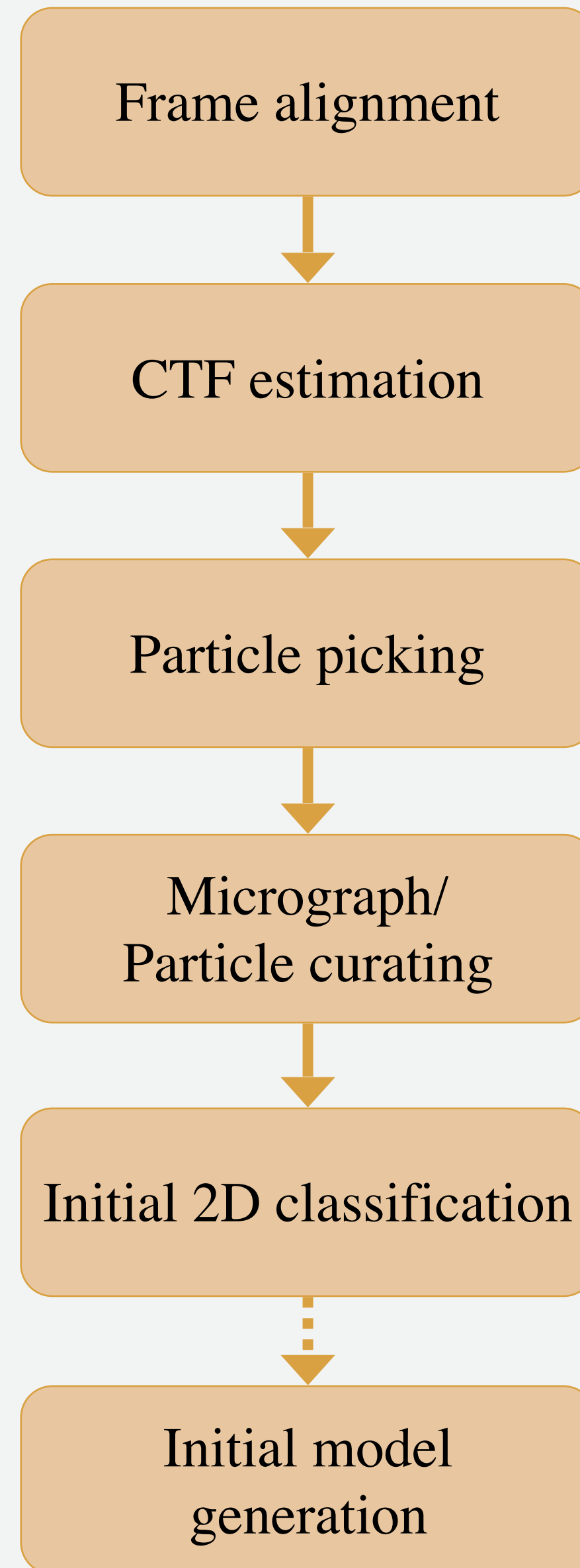
2D classification

3D refinement

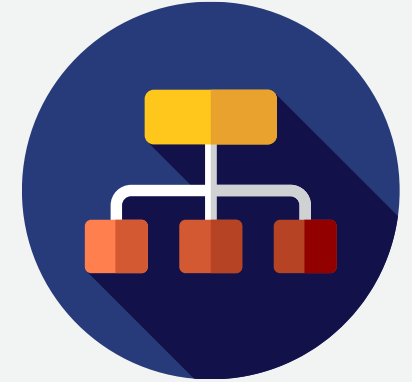
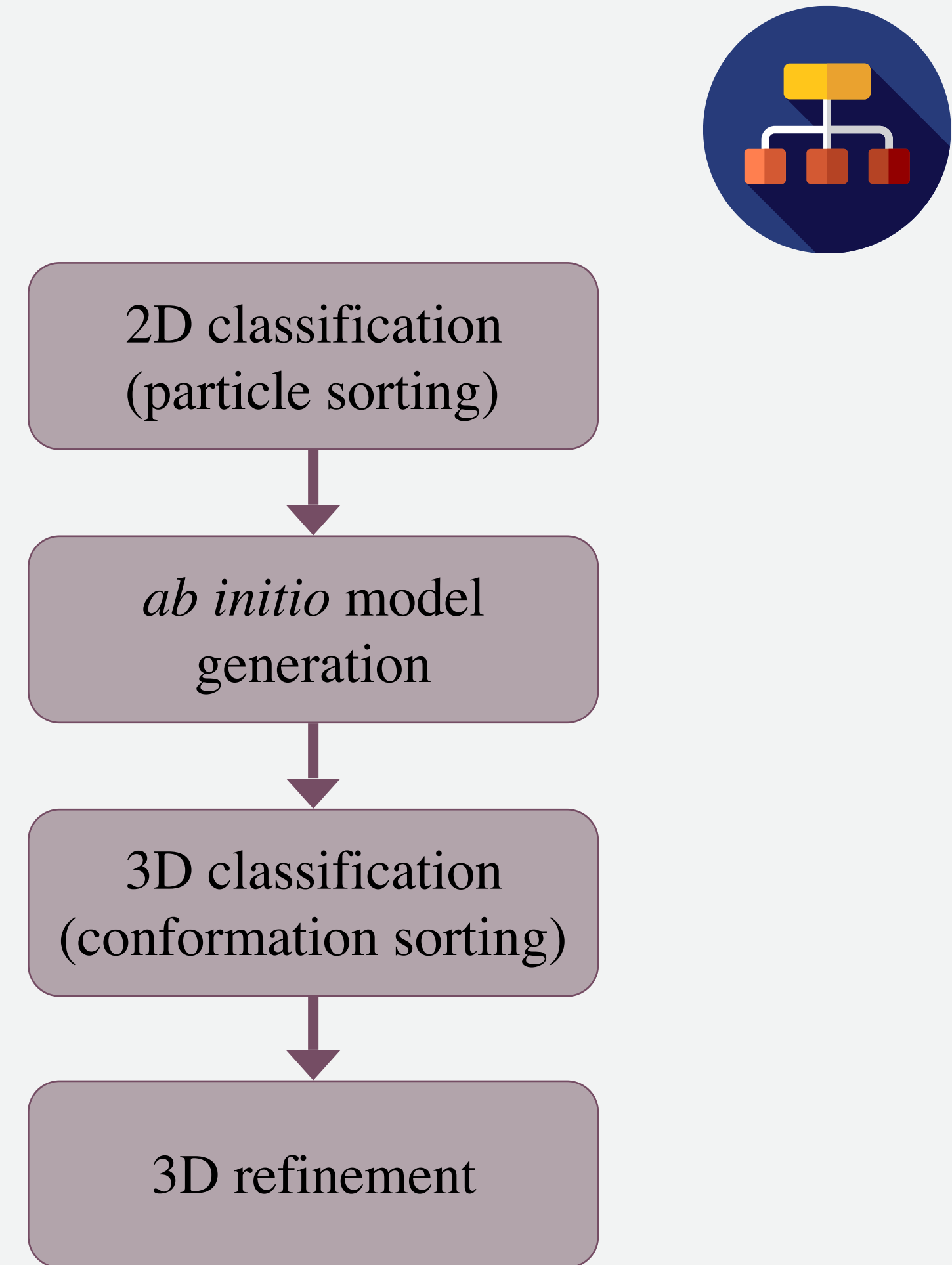
3D classification



Appion part I



Appion part II



Going from Appion to RELION/cryoSPARC

Raw frames

/gpfs/frames/[username]/[sessionname]/rawdata/

Aligned Movies

/gpfs/legion/[username]/[sessionname]/rawdata/

Particles

/gpfs/appion/[username]/[sessionname]/stacks/

CTF Estimation

/gpfs/appion/[username]/[sessionname]/CTF/

Selecting particles – box-files

Hide | Expand | Contract

▼ Object Selection : 1

Dog Picking
1 complete

Template Picking

Manual Picking

Object Tracing

Signature

Repeat an image loop run

▼ CTF Estimation

Estimate the CTF
2 complete

Repeat an image loop run

▼ Stacks : 5

Stack creation
6 complete

more stack tools

▼ Particle Alignment

Run Alignment
2 complete

Run Feature Analysis
2 complete

Run Particle Clustering
6 complete

Run MaskItOn

Template Stacks

▼ Ab Initio Reconstruction

OptiMod Common Lines
2 complete

EMAN Common Lines

SIMPLE Common Lines

▼ Refine Reconstruction

Run Single-Model Refinement

Run Multi-Model Refinement

▼ Helical Processing

Helical Image Processing

▼ Direct Detector Tools

Create frame stack
1 running

Launch Alignment Catchup

Launch DE Frame Alignment

Launch Particle Polishing

▼ Import tools

Run name
stackFINAL

Output directory
/gpts/appion/eeng/16jul26b/stacks

Description
"just box-files"

esn-a ▾ Preset
☐ Wait for more images after finishing

☐ Max number of images to process
Images to process:
☒ Do not process hidden or rejected images
☐ All images independent of status
☐ Exemplar and keep images only
Image order:
☒ Forward
☐ Reverse
☐ Shuffle
Continuation:
☒ Continue unfinished run (default)
☐ Reprocess all images

☒ Commit results to database

Density modifications:
☒ Invert image density
Normalization Method
edgenorm: normalize by mean 0 stdev 1 in based on edge pixels ▾

Particles: None ▾

Stacks: stack4-centered4 (4,696 parts) ▾

128 Box Size (Unbinned, in pixels)
☐ (override boxsize)

1 Binning

☒ Fast Insert

☐ Ctf Correct Particle Images

☐ CTF Confidence Cutoff
Use Values Above: 0.8 (between 0.0 - 1.0)

CTF Correction Method
Ace 2 Wiener Filter Whole Image ▾

CTF Sorting Method
Resolution 0.8 criteria ▾
-OR- choose a ctf run: all (0) ▾

☐ CTF Resolution Cutoff
Resolution range at 0.8 criteria:
between and (in Angstroms)
Resolution range at 0.5 criteria:
between and (in Angstroms)

Defocus Limits
Min -3.9e-7 & Max -6.71e-6

No Mask Assessed for this Session

Raw frame processing if available:
start frame: start frame total frame: total frames

Filter Values:
☐ Low Pass (in Angstroms)
☐ High Pass (in Angstroms)
4 Pixel Limit (in Standard Deviations; 0 = off)
☐ Particle Correlation Cutoff
Remove particles with CCC below: 0.5
Remove particles with CCC above: 1.0

Defocal pairs:
☐ Use defocal pairs

Limit # of particles to:

☒ Only create EMAN boxfiles

Helical Alignment:
Apply helical alignment analysis

Selecting CTF

▼ Object Selection : 7

Select Particle Picker...
7 complete

Repeat an image loop run

▼ CTF Estimation

Estimate the CTF...
4 complete

Transfer results to another preset

Repeat an image loop run

▼ Stacks : 15

Stack Creation...
16 complete

▼ Particle Alignment

Select Particle Alignment...
1 complete

Run Feature Analysis...

Run Particle Clustering...
7 complete

Run MaskItOn

Template Stacks
1 complete

[show all statistics]

Summary of confidence values from all runs

esn-a-DW ▾ Preset

Select Run: all ▾ Ctf Run

Update Preset/Run Selection

↓ download star file for RELION 1.4

↓ download star file for Relion 2.0 or VLION

↓ download best ctf data

↓ download best ctf EMX file

Ctf Run: 50 tctf2 hide delete (↓ download ctf data)

Time: 2017-02-18 15:39:38

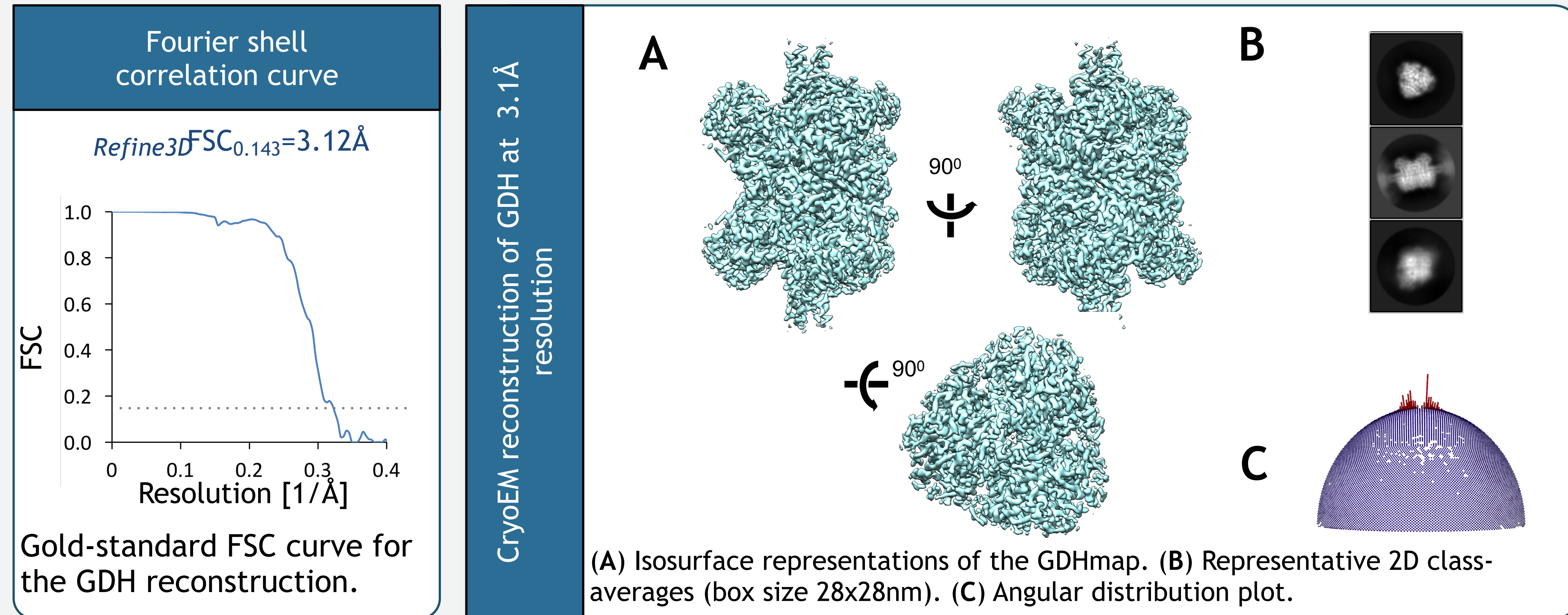
Path: /gpfs/appion/eeng/17feb14h/ctf/tctf2

	nb	min	max	avg	stddev
defocus1	443	0.02 μm	7.50 μm	1.74 μm	0.66 μm
defocus2	443	0.11 μm	7.56 μm	1.80 μm	0.66 μm
angle astigmatism	443	-89.991	83.662	-42.236	30.574

Test set

Sample: Glutamate Dehydrogenase (GDH)

Reconstruction: Relion 2 3D refine



Imaging statistics: 20 eV slit, stage shift

17feb14h experiment : 528 esn images, 523 esn-a-DW images used

443 enn-a-DW images used from Appion

16,191 particles = 3.16Å (using a CTF conf 0.8 >8Å CTF cut off)

Imaging parameters:

Grid loaded: 17feb13

Session: 17feb14h

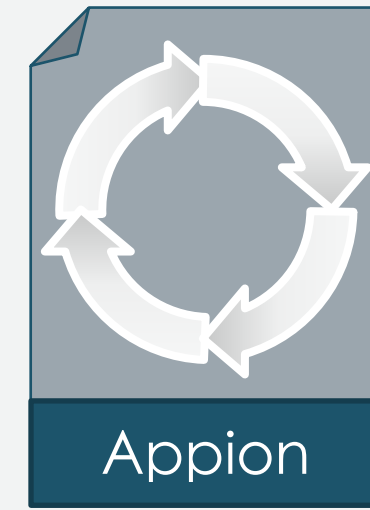
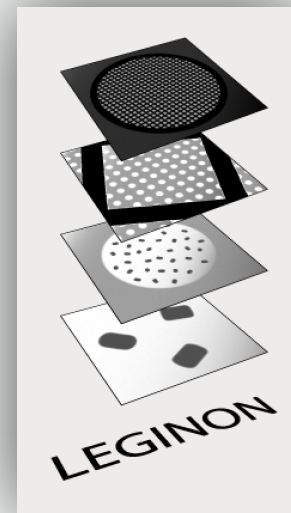
Pixel size: 0.43Å superres
bin2 - 0.854905Å/px

Camera: K2 counting, 400 ms
- 32 frames

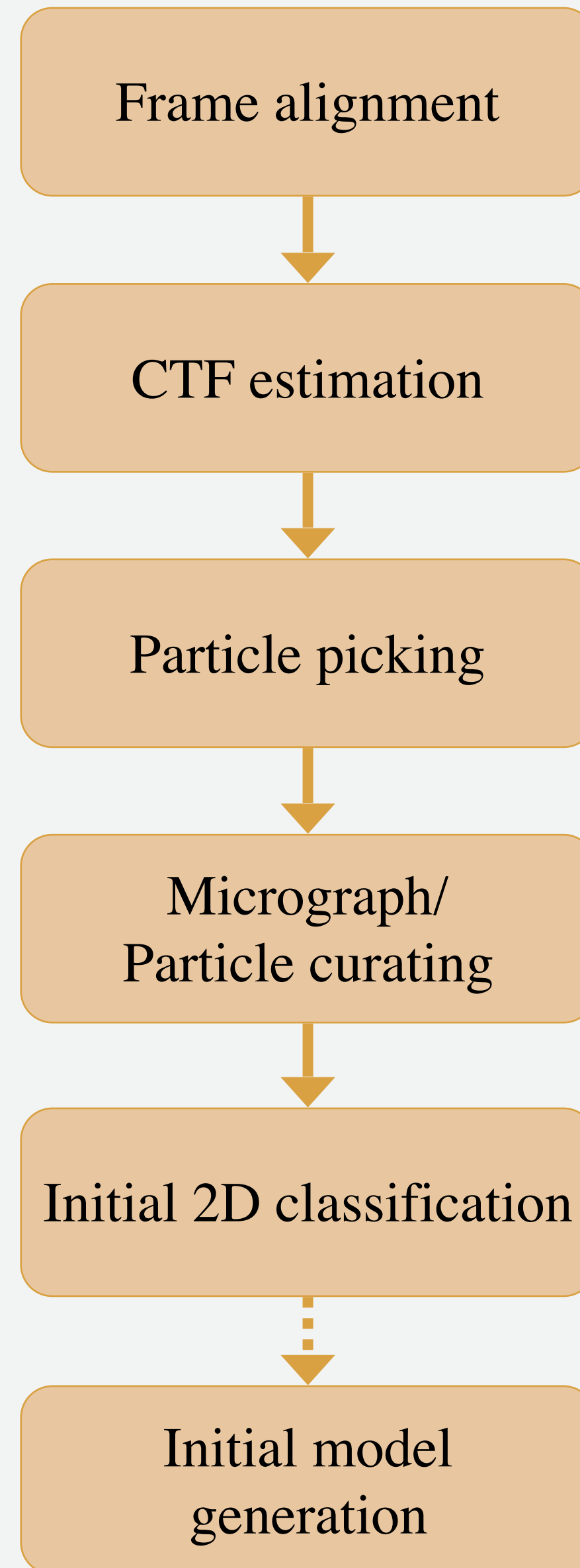
Camera settings: 12,800 ms
exp

Dose: 35.73 e⁻/Å²

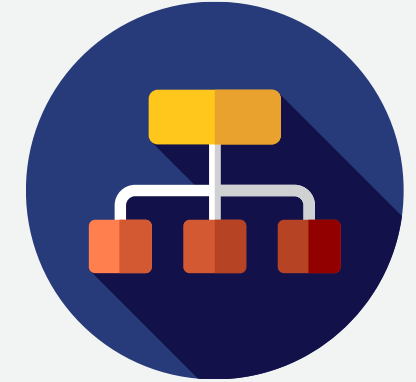
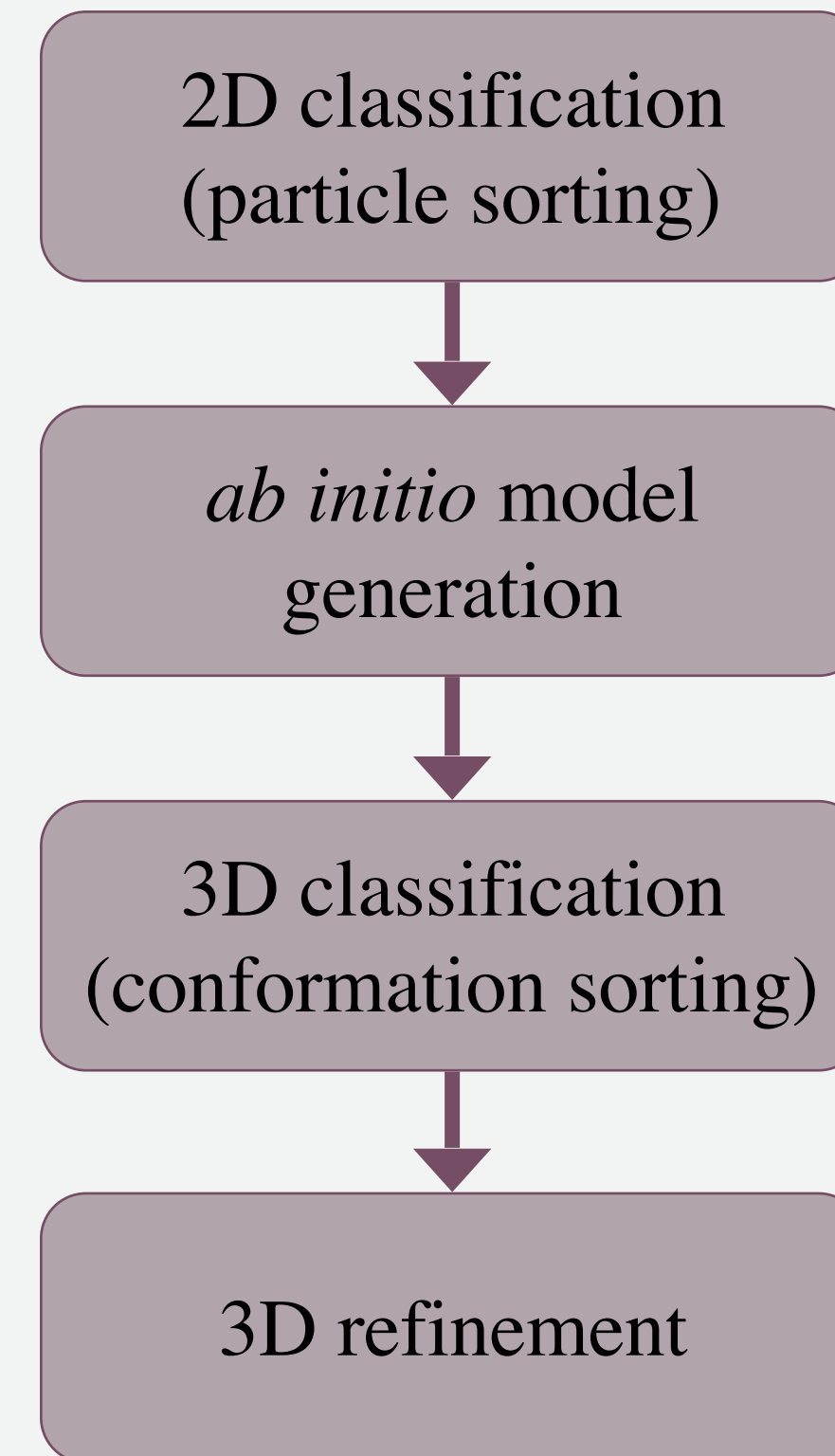
Photo time



Appion part I



Appion part II



Schedule

September 19, 2019 10am-4pm

A practical workshop that continues the single-particle analysis workflow from the Part 1 workshop. The topics include Appion advanced features, moving into commonly used packages and validation methods.

- **1. Workshop intro**
 - Overview of SPA processing and web tools
 - Advanced Appion features
 - *Photo*
- **2. Going from Appion to other packages**
 - RELION
 - cryoSPARC
- **3. Validation & additional tools** *[if time allows]*
 - FSC & 3D-FSC / local res & ResMap / other methods
- **4. Closing** *[Q&A with SEMC staff]*

2D classification

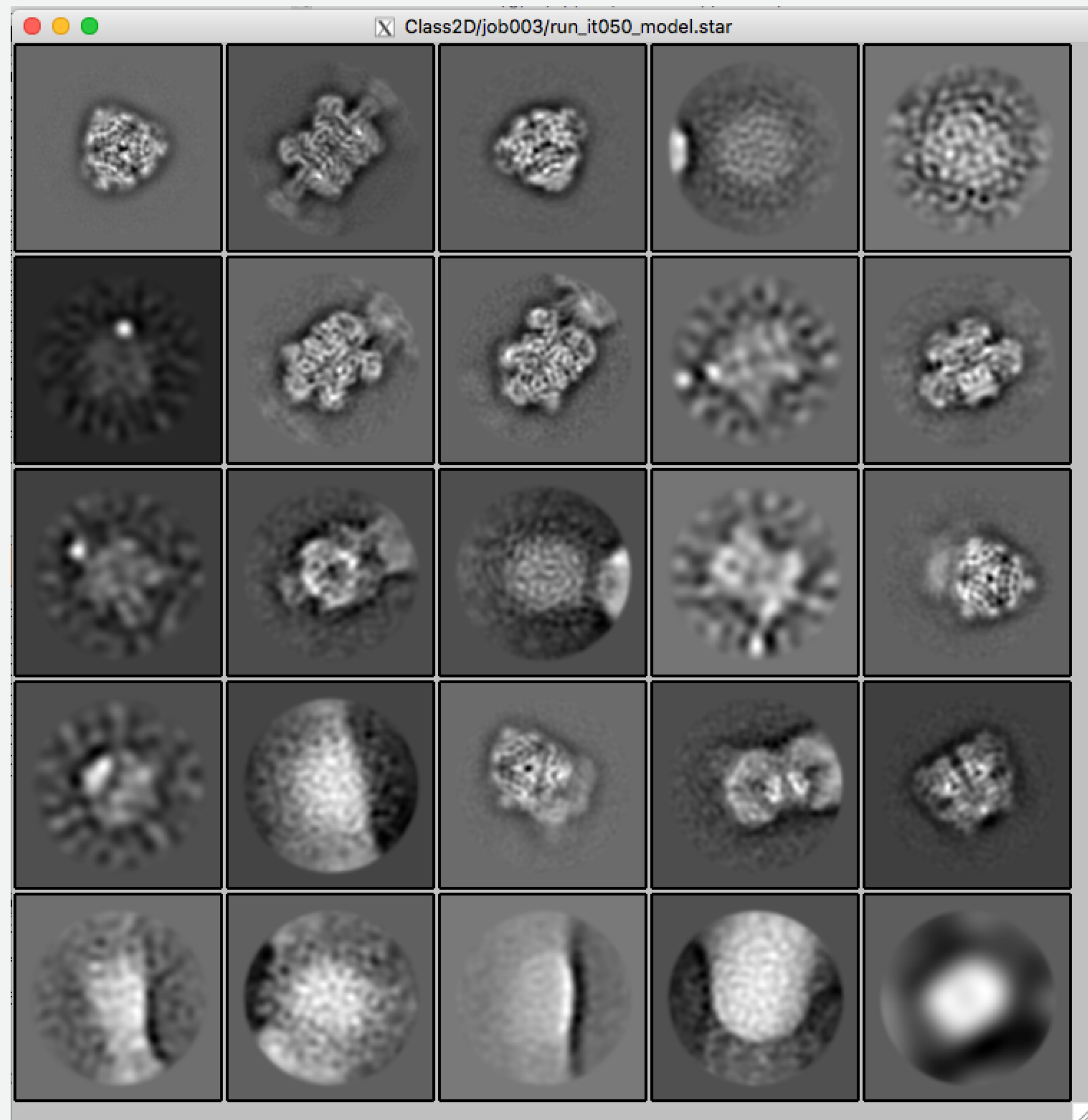
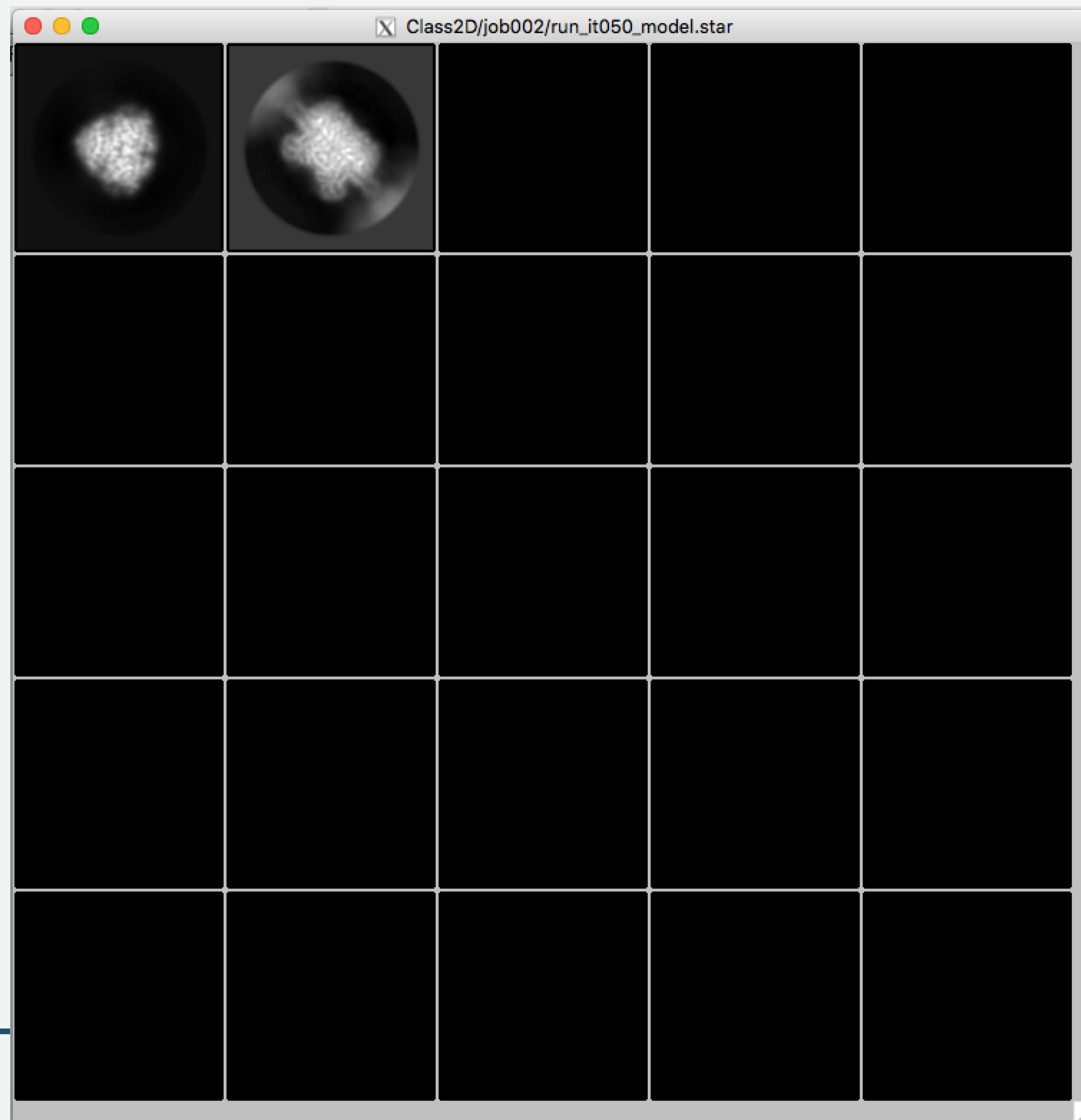
The screenshot displays the RELION-3.0-beta-2 software interface. The main window title is "RELION-3.0-beta-2: /gpfs/appion/sworkshop/relion3/run0". The top menu bar includes "File", "Jobs", and "Aautorun". Below this, a series of tabs are visible: "I/O", "CTF", "Optimisation", "Sampling", "Helix", "Compute", and "Running". The "Optimisation" tab is currently selected, showing settings for "Do CTF-correction" (checked), "Have data been phase-flipped" (unchecked), and "Ignore CTFs until first peak" (checked).

On the left side, a vertical list of job actions is shown, with "2D classification" highlighted. Below this list, a "Job actions" button is visible. The "Current job:" field displays "Give_alias_here".

The bottom section of the interface is divided into three panels: "Finished jobs", "Running jobs", and "Scheduled jobs". The "Finished jobs" panel lists several completed tasks, including "009: Select/run0_288-144px_2D-it42_3D-it42/", "007: Refine3D/run0_288-144px_2D-it42/", "006: Class3D/run0_288-144px_2D-it42/", "004: Select/run0_288-144px_2D-it42/", "003: Class2D/run0_288-144px_ignoreCTF1", "002: Class2D/run0_288-144px/", and "001: Extract/run0_288-144px/". The "Running jobs" panel shows "005: InitialModel/run0_288-144px_2D-it42/" and "008: Class3D/run0_288-144px_2D-it42_ign". The "Scheduled jobs" panel lists "010: Refine3D/run0_288-144px_2D-it42_3D".

On the right side, a detailed view of the "Optimisation" tab settings is shown. It includes sliders and input fields for various parameters: "Number of classes" (set to 25), "Regularisation parameter T" (set to 2), "Number of iterations" (set to 50), "Use fast subsets (for large data sets)?" (set to Yes), "Mask diameter (A)" (set to 200), "Mask individual particles with zeros?" (set to Yes), and "Limit resolution E-step to (A)" (set to -1). Each parameter has a corresponding help icon (?) and a slider control.

2D classification



Initial model

RELION-3.0-beta-2: /gpfs/appion/sworkshop/relion3/run0

File Jobs Autorun

I/O CTF Optimisation SGD Compute Running

Import
Motion correction
CTF estimation
Manual picking
Auto-picking
Particle extraction
Particle sorting
Subset selection
2D classification
3D initial model
3D classification
3D auto-refine
3D multi-body
CTF refinement
Bayesian polishing
Mask creation
Join star files
Particle subtraction
Post-processing
Local resolution

Number of classes: 1

Mask diameter (A): 200

Flatten and enforce non-negative solvent? Yes

Symmetry: C1

Initial angular sampling: 15 degrees

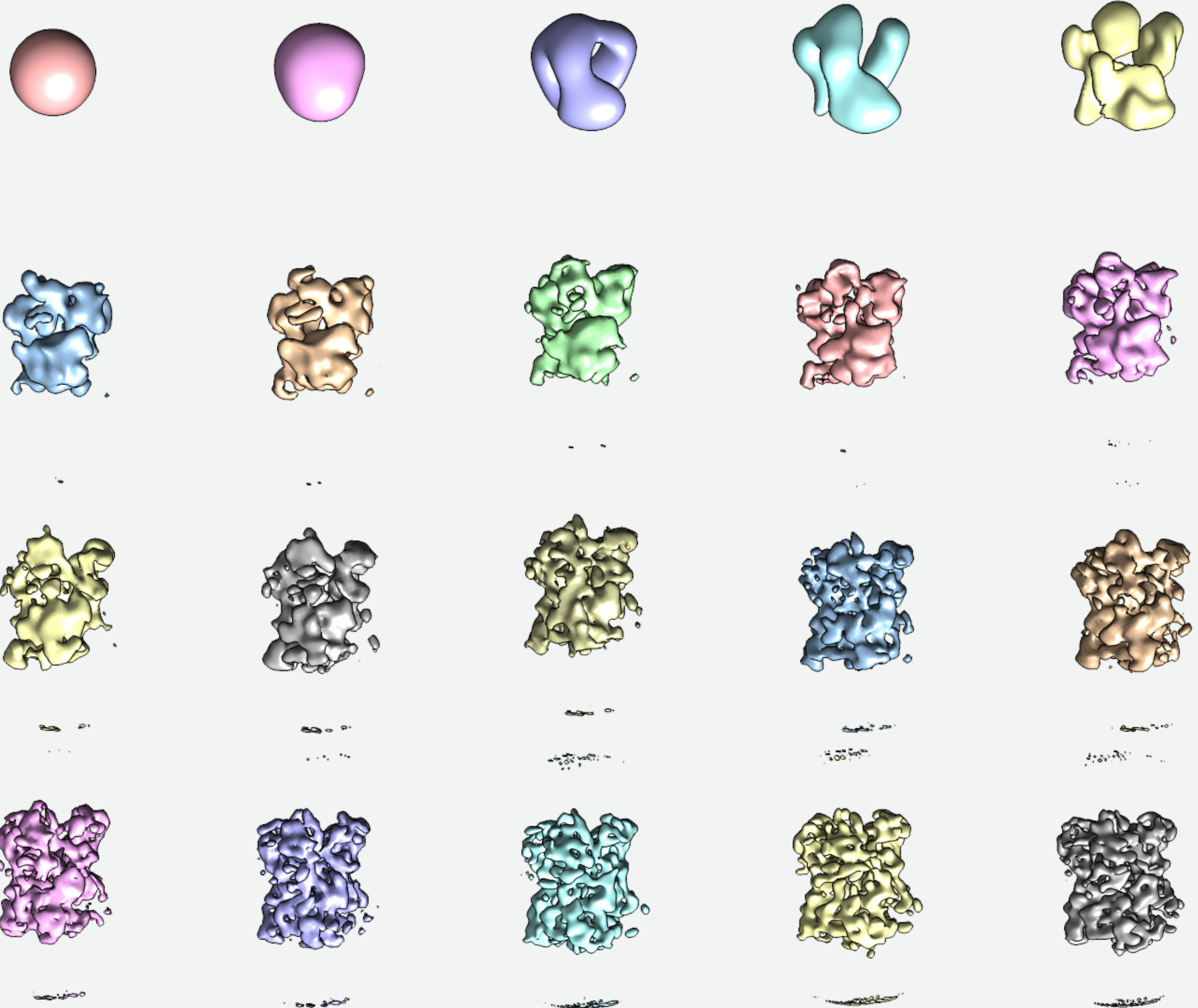
Offset search range (pix): 6

Offset search step (pix): 2

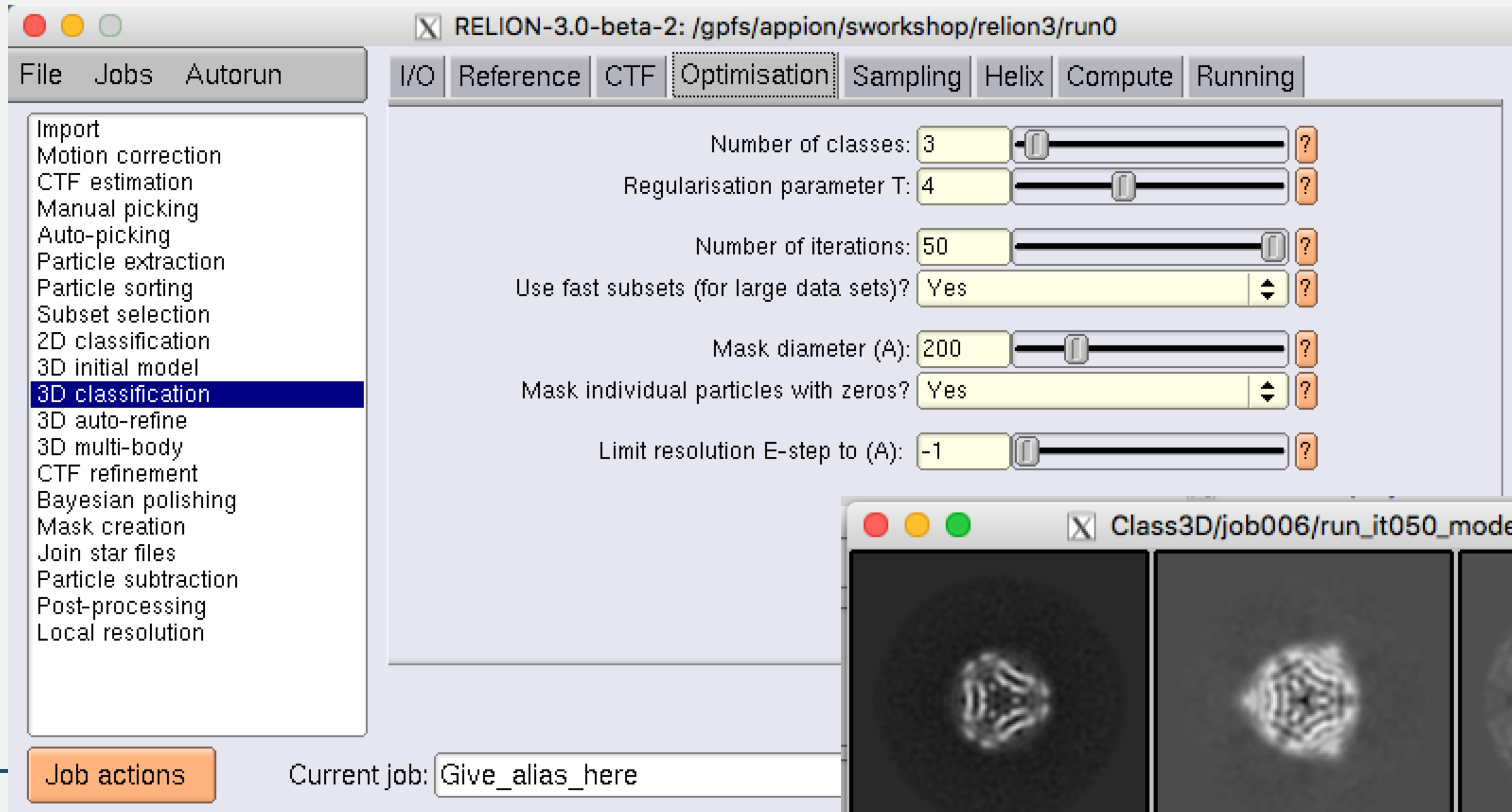
Print command Schedule Run!

Initial model

Stochastic Gradient Descent
algorithm for initial 3D model
generation



3D classification



RELION 3 workflow

Workshop workflow : Real time demo test set – 17feb14h (~1/2 day run)

Particle extraction

Extract/job000 folder: 16,191 particles
(144 px box size – bin4 @ 1.70981 Å/px)

2D classification

16,124 particles
(used 10 classes from K=25, 12 iterations)

3D classification

12,131 particles
(used 2 classes from K=3, 7 iterations)

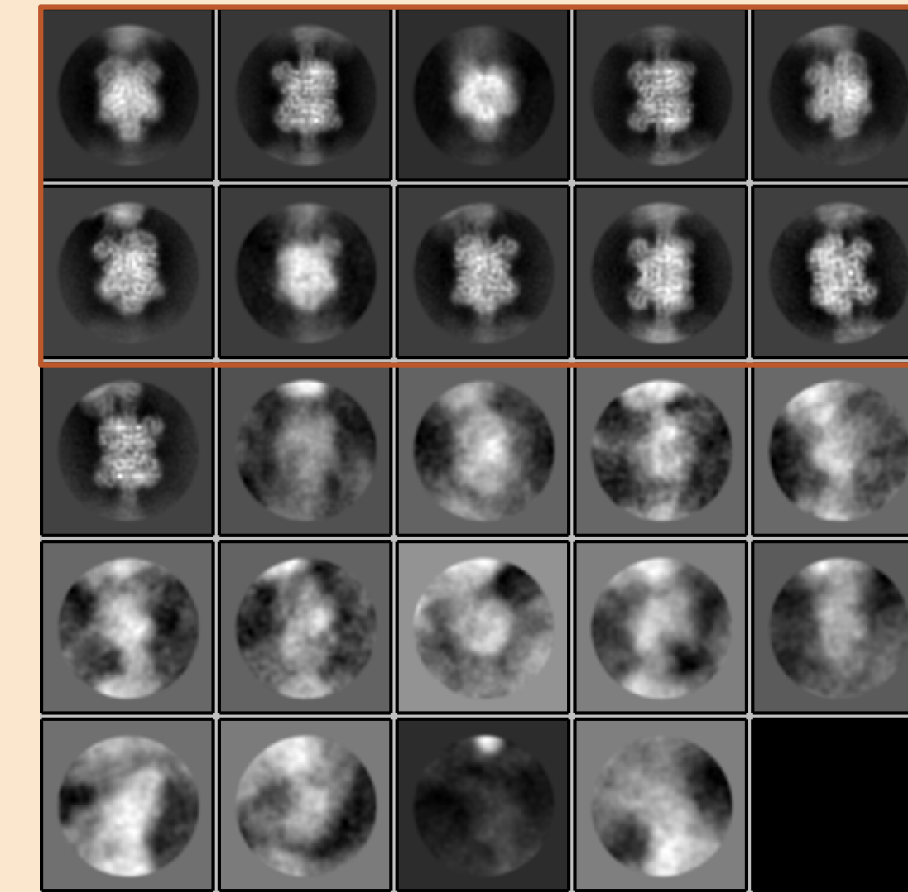
3D refine

3.7473Å res
(17 iterations, ~1 hr on semcgpu01)

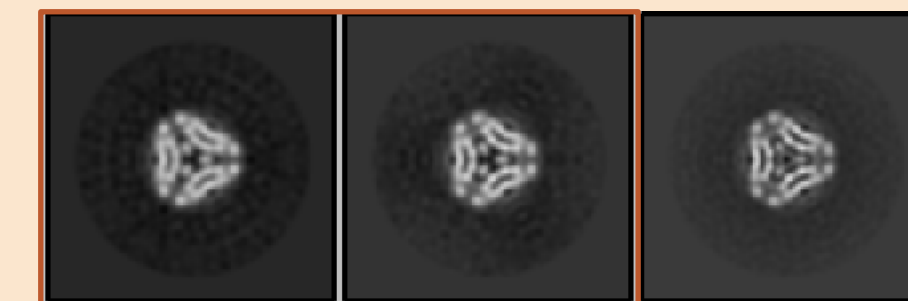
Postprocessing

3.41962Å res
(using 0.02 mask, 3 px ext, 3 px soft mask)

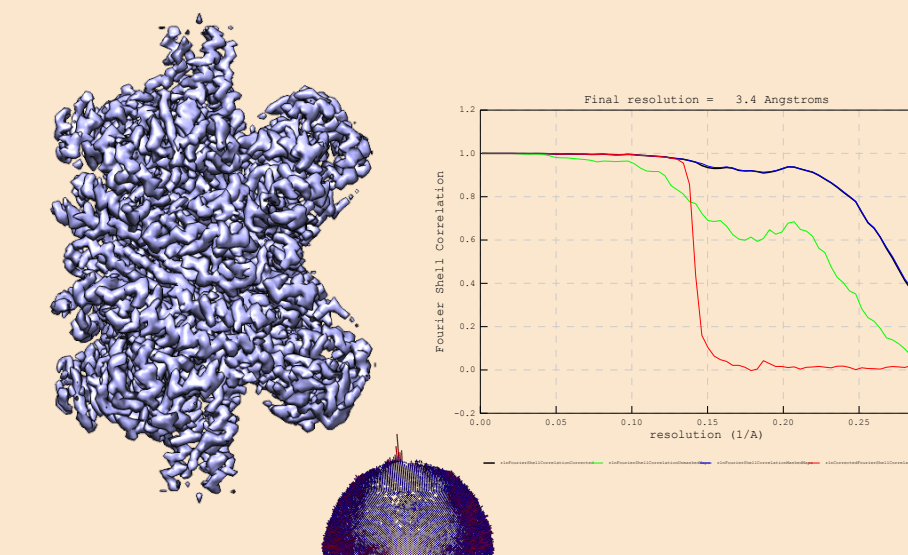
2D classification



3D classification



3D refinement



The image features a solid teal background. In the upper and lower portions, there are decorative white dotted lines forming swirls and loops. A light gray rectangular box with slightly irregular, hand-cut edges is positioned in the center. Inside this box, the word "Questions?" is written in a black, elegant serif font.

Questions?