



cryoSPARCv2 tutorial v0.1

Appion Part II
September 19, 2019

Tutorial setup:

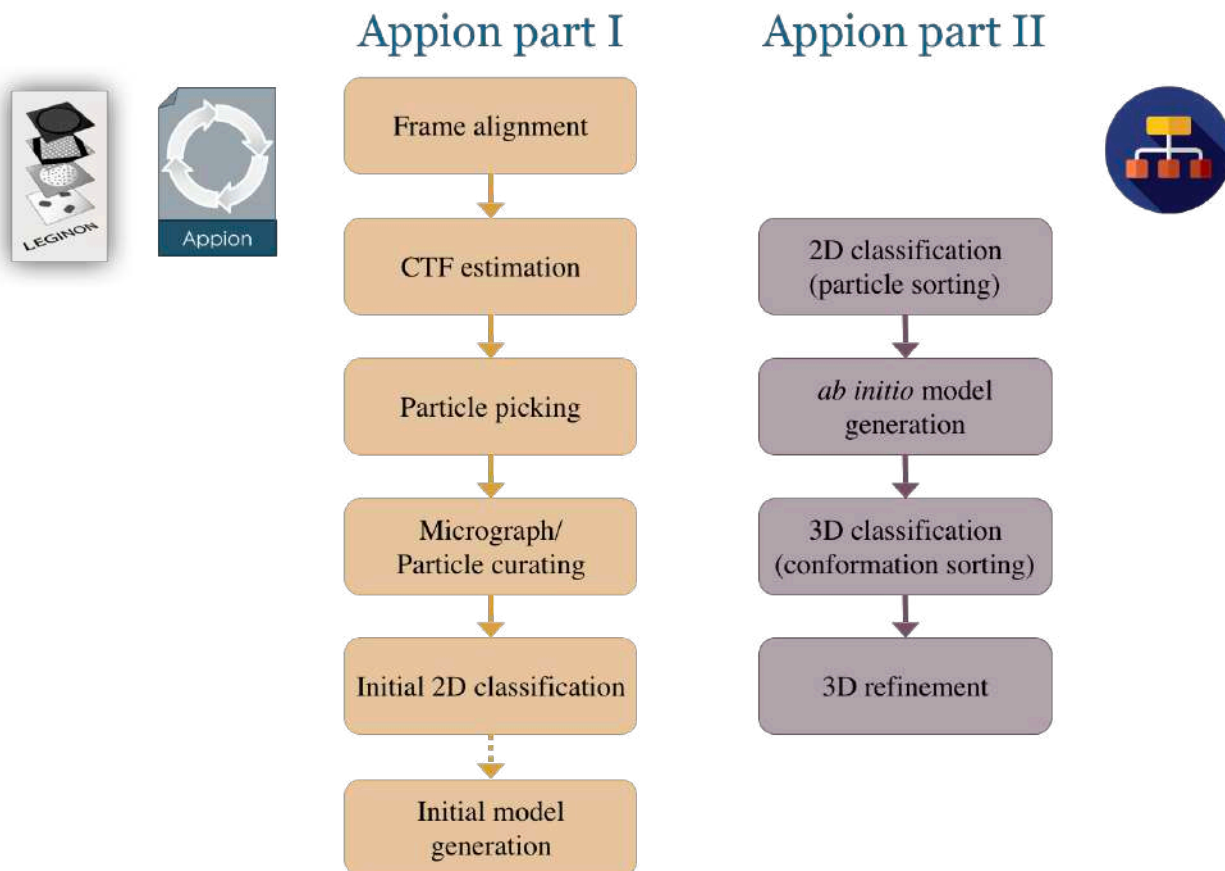
Requirements:

Web browser (Chrome recommended by CryoSPARC developers) for tutorial.

Recommended:

UCSF Chimera (<https://www.cgl.ucsf.edu/chimera/>) for visualization.

Workflow goals:



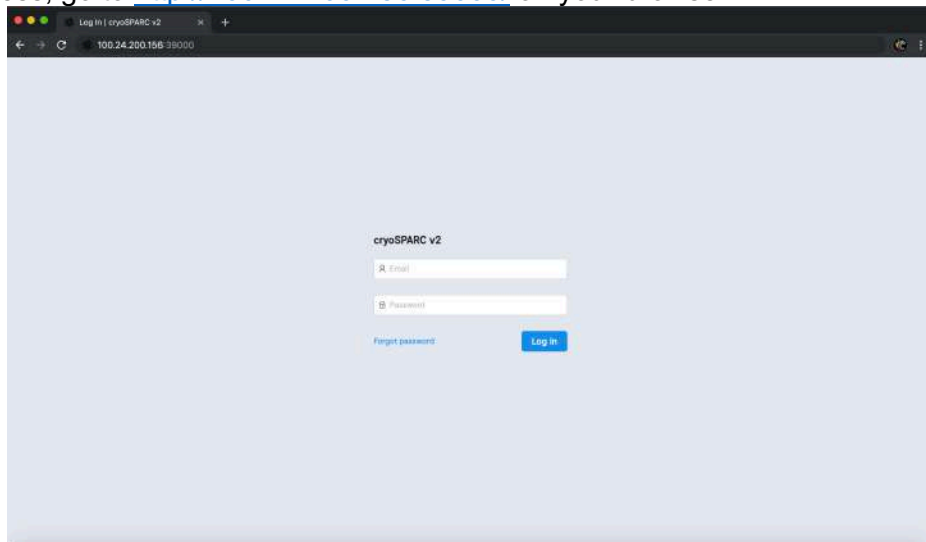
This tutorial assumes that students have taken “Appion part I: Basics and 2D processing” where the pre-processing pipeline has been covered and will focus “Appion part II: 3D reconstruction/refinement”.

Given standard SEMC pre-processing workflow within minutes of the last acquired image during data collection the user should have: 1) curated aligned micrographs, 2) CTF estimation, and 3) particle picks. If the user has also utilized skills learned in Appion part I they may also have: 4) initial particle curation and 5) an initial model.

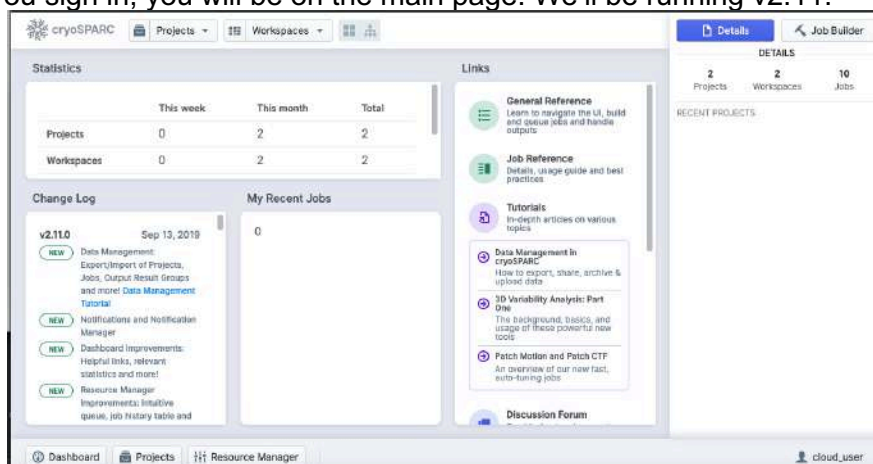
cryoSPARC in the cloud:

Accessing CryoSPARC 2 on the Cloud

1. What is cryoSPARC?
 - CryoSPARC (Cryo-EM Single Particle Ab initio Reconstruction and Classification) is a software package commonly used for computing high-resolution Cryo-EM structures (<https://cryosparc.com/>). We have installed it on the Amazon web services cloud for this practical.
 - To read more about running cryo-EM software on the cloud, go here: <http://cryoem-tools.cloud/>
 - A big advantage of the cloud is easy scale-up: We will have 16 GPUs for the workshop.
2. To access, go to <http://100.24.200.156:39000/> on your browser



3. Login using the following credentials.
 - username: *test@appion.org*
 - password: *appion2019*
4. Once you sign in, you will be on the main page. We'll be running v2.11.



Making Your Project and Dataset

- Today, we will processing bovine GDH (a 334 kDa D3 symmetry protein). Our goal is to compute a reconstruction of this protein while going over the basic workflow for single particle 3D construction. Note, upon familiarization with this workflow time will be allotted to test different workflows and/or the full single-particle pipeline.
- Now you can make a project and import your dataset. Click on projects, and make a new project named after yourself, and have the output directory as “/efs/projects/”. Then select the project.

The image shows two screenshots of the cryoSPARC web interface. The top screenshot displays the 'Projects' page with a navigation bar at the top containing 'Projects' and 'Workspaces'. A table in the 'Statistics' section shows the following data:

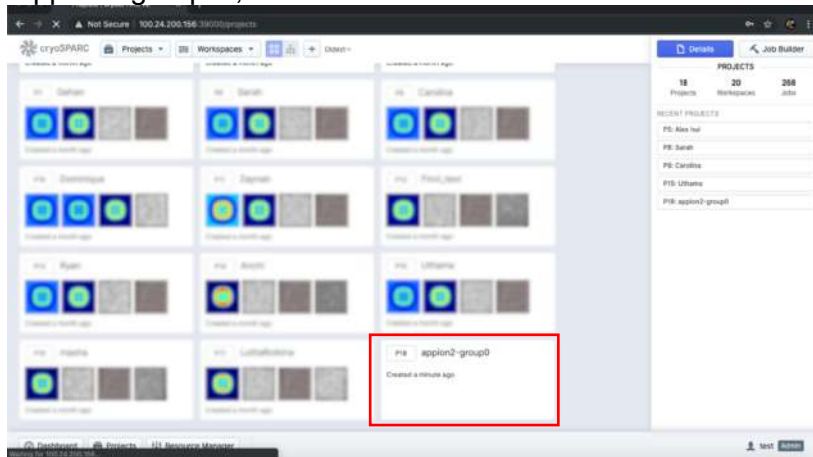
	This week	This month	Total
Projects	0	2	2
Workspaces	0	2	2

The bottom screenshot shows the 'New Project' dialog box with the following fields:

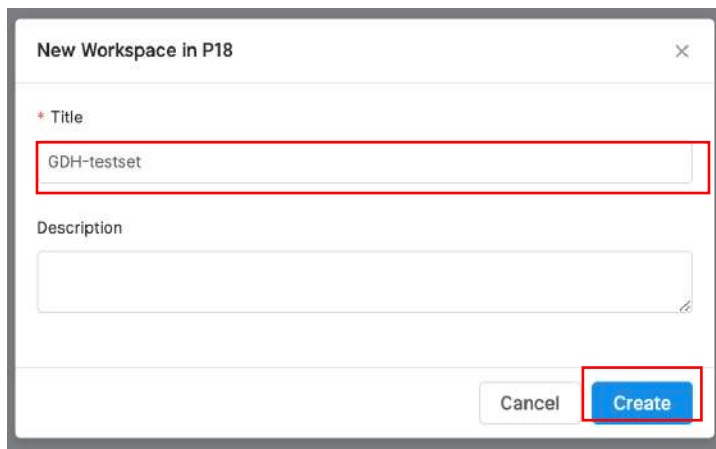
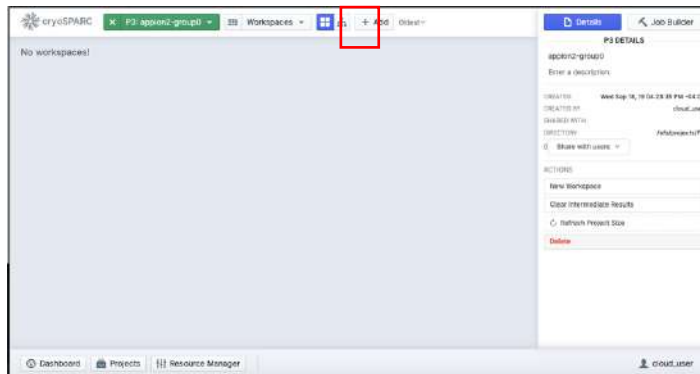
- Title:
- Directory in which to create new project directory:
- Description:

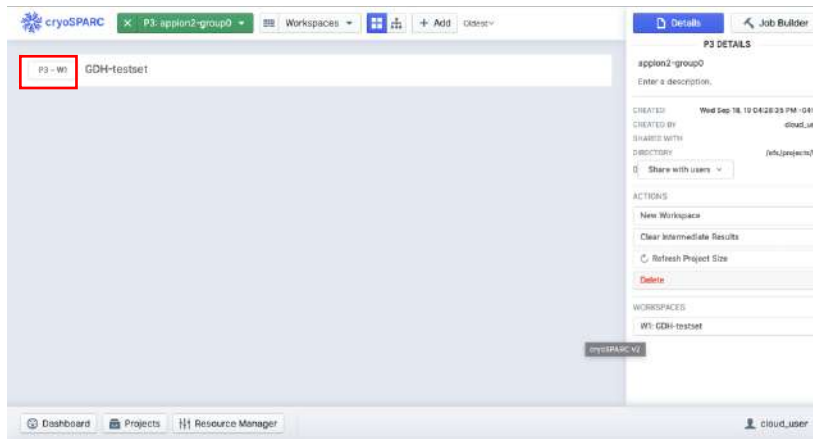
Buttons for 'Cancel' and 'Create' are visible at the bottom of the dialog box.

- Title your project in as per your group assignment. For example, “appion2-groupA”, “appion2-groupB”, etc...

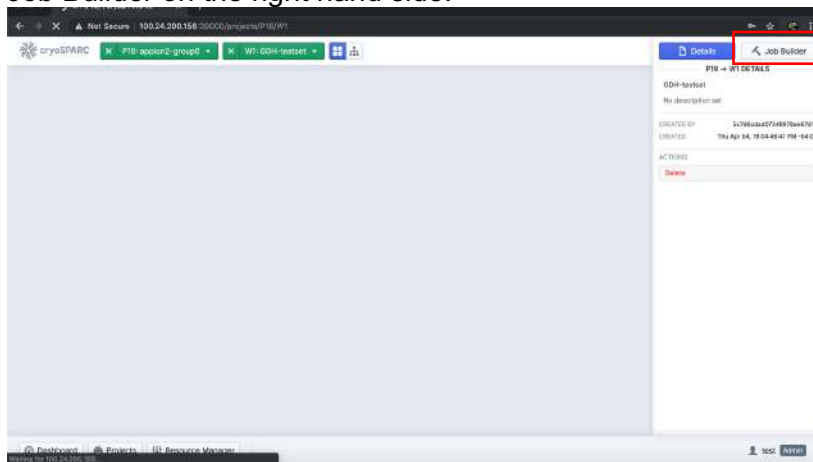


-
- 7. Select your project. In your newly created Project, no Workspace exists yet. Create one called GDH-testset and select it



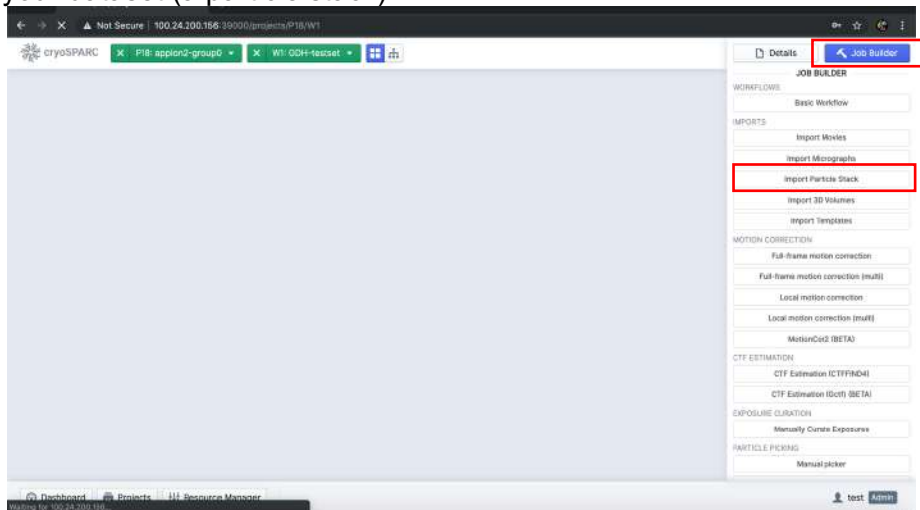


-
- Once inside your Workspace, we can now start our workflow. To build a workflow select Job Builder on the right hand side.

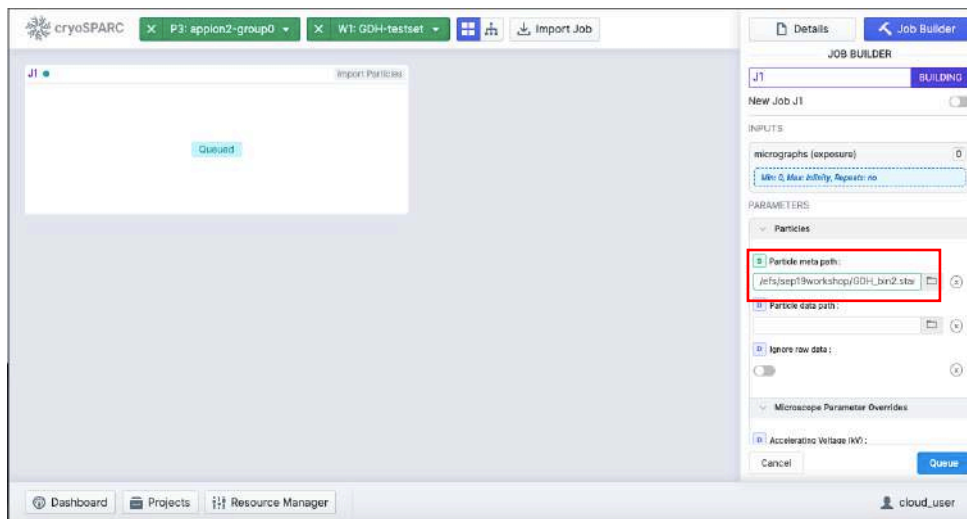


The job builder has many building blocks, but we will start the workflow assuming we have exported a particle stack with CTF estimation.

8. Import your dataset (a particle stack).

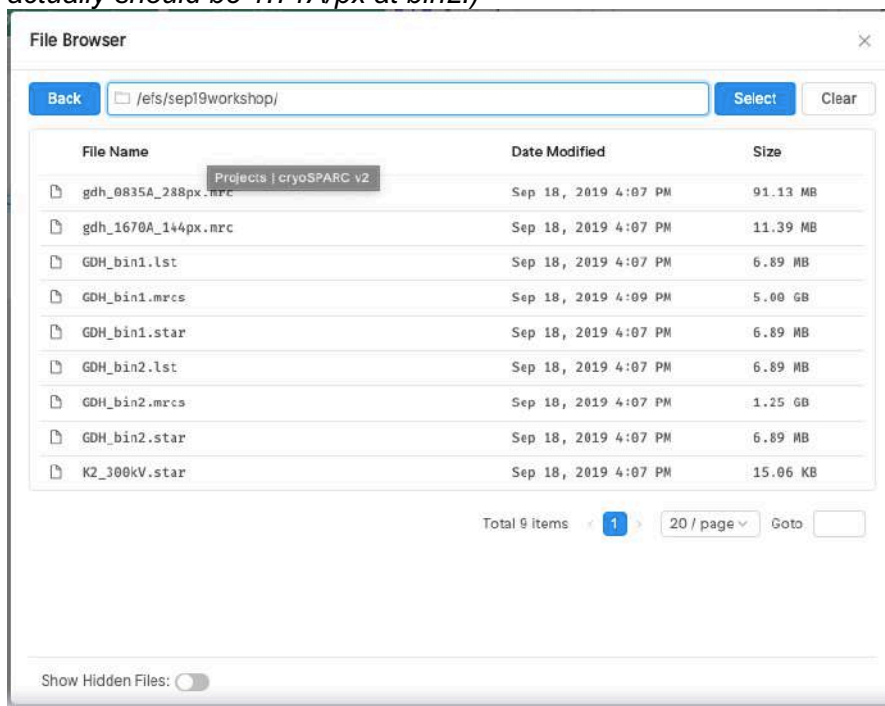


9. Load your dataset by selecting Particle meta path. The data is located in /efs/sep09workshop/



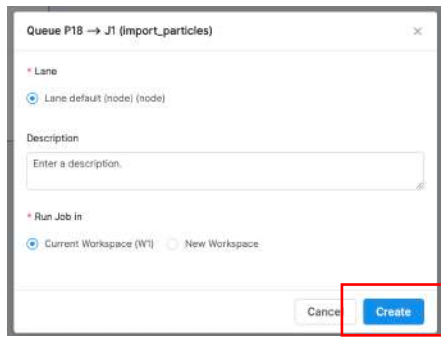
○

- These particle stacks have a list file with CTF estimation (.star file) and the particle stack in MRC format (.mrcs file). You may choose either the GDH_bin1.star or GDH_bin2.star file. The bin2 file is 144x144px box size @~1.77Å/px. (Note, the pixel size was calibrated subsequently by scaling the EM reconstruction to a pdb of the crystal structure and it actually should be 1.71Å/px at bin2.)

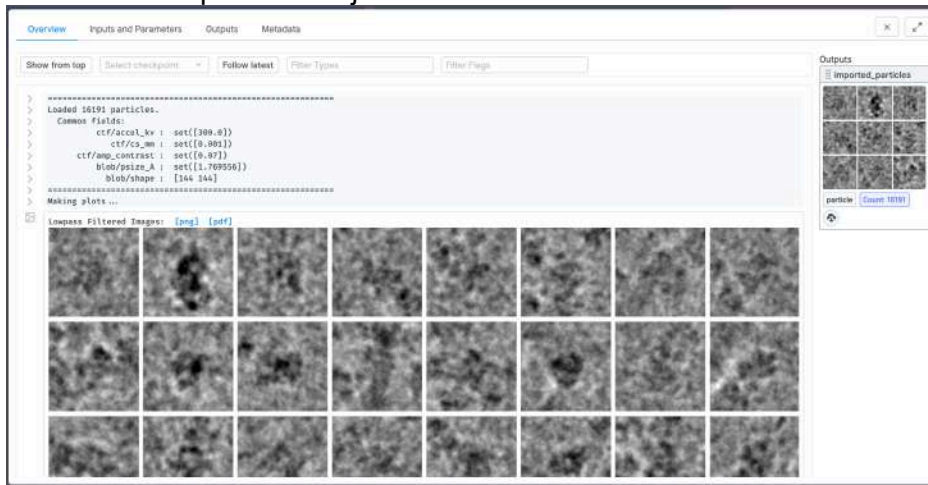


○

- A window will pop up. Click Create to submit the job.

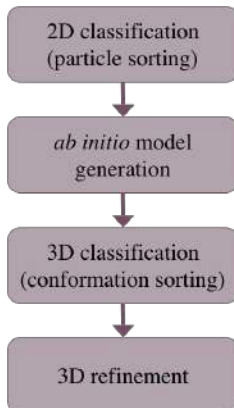


10. Look at the outputs as the job runs.



You can click on the job run to pop the overview window out again. Hit your spacebar to exit the window.

[Example SPA workflow for this tutorial]



We are going to focus on this type of single particle analysis (SPA) workflow for the tutorial. In terms of cryoSPARC it would be roughly these tasks in the Job Builder:

Particle curation

- 1) 2D Classification
- 2) Select 2D classes

3D reconstruction

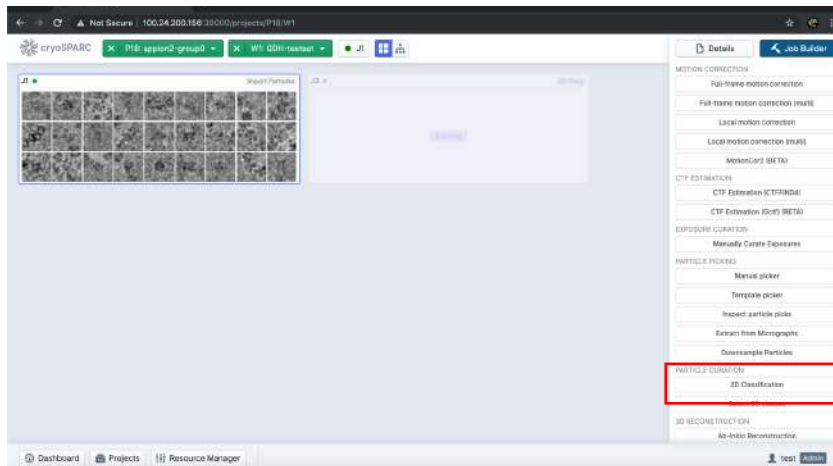
- 3) *Ab-Initio* Reconstruction

3D refinement

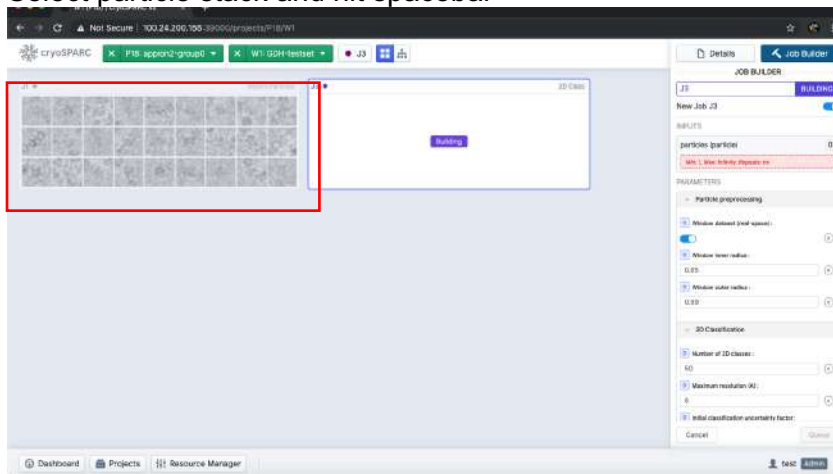
- 4) Heterogeneous refinement
- 5) Homogeneous refinement

2D classification

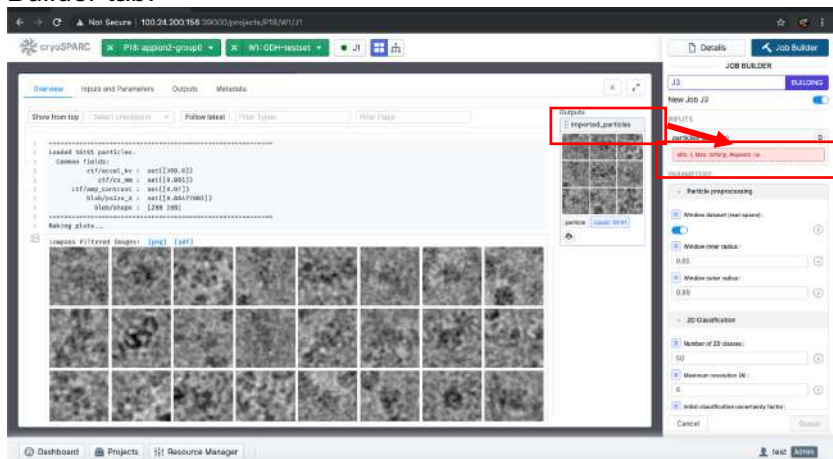
11. Create a 2D classification job in the Job Builder.



- 12. To associate the imported dataset to the job select the Import Particle job, hit spacebar to open the overview tab, and click and drag the particle stack to the Inputs in the Job Builder panel.
 - Select particle stack and hit spacebar

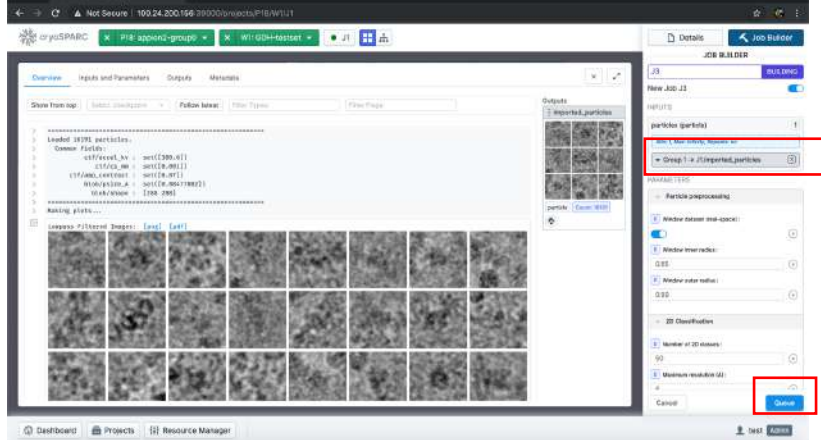


- Click and drag imported_particles in the imported particles to particles in the 2D class Job Builder tab.

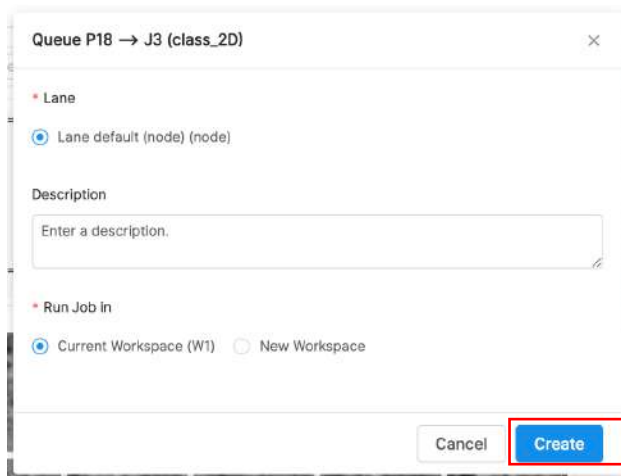


-

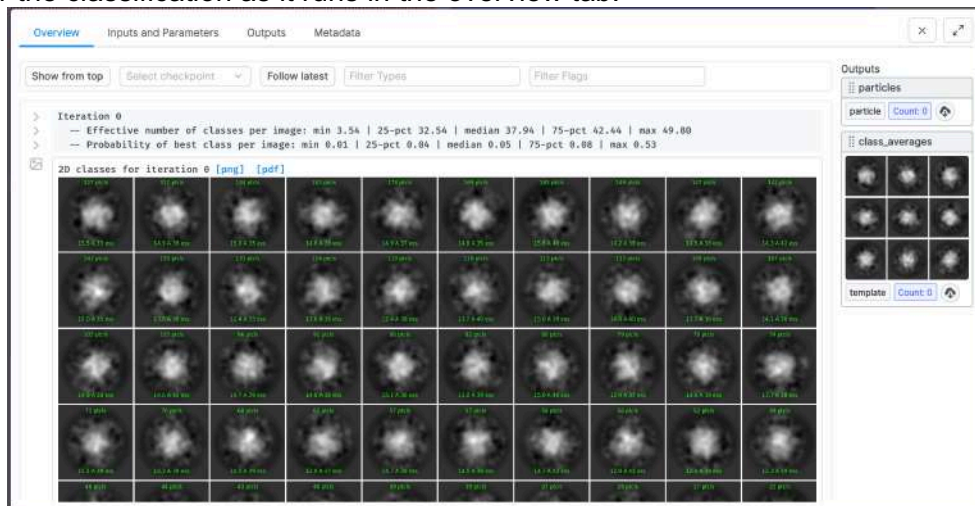
- If successful, then you should see this the following. Then queue the job.



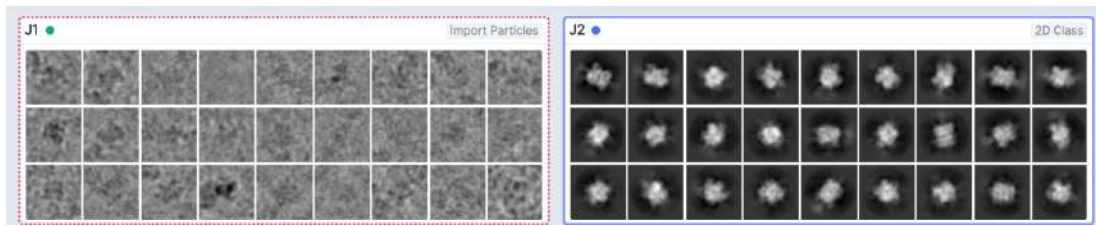
- You may decide to change the number of classes from 50 to a smaller number, but do not change the other parameters.



13. Monitor the classification as it runs in the overview tab.

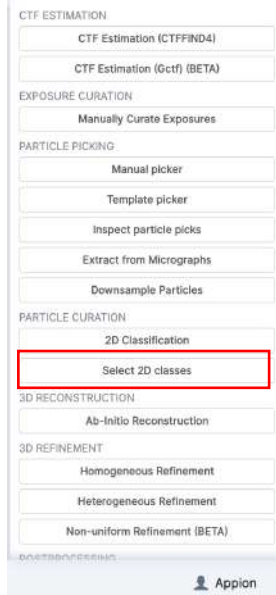


- In your workspace completed jobs have a green dot and in process jobs have a blue dot

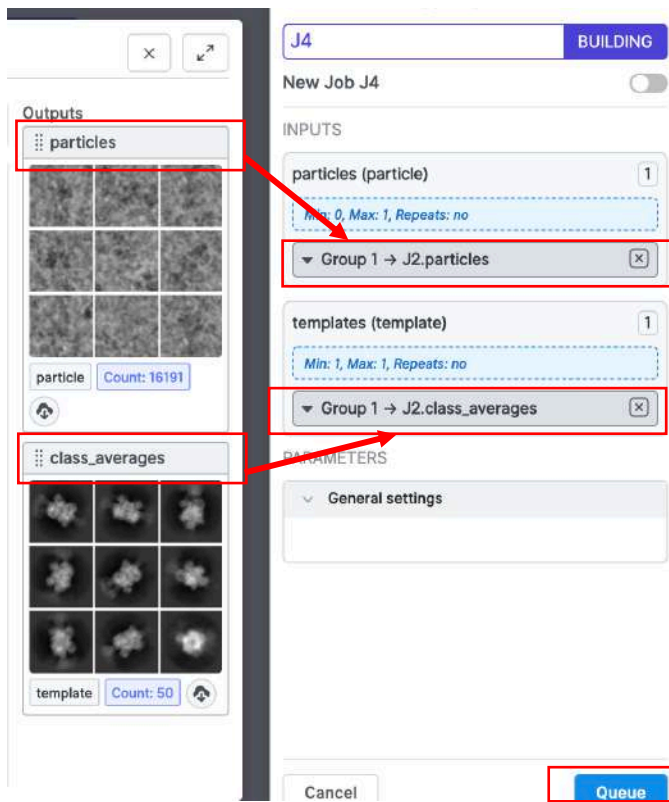


Select 2D classes

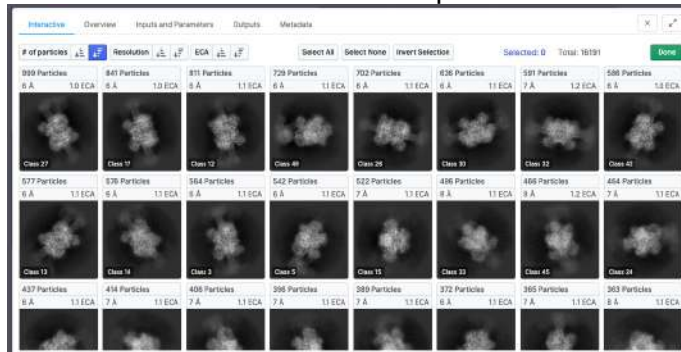
14. Start a new job to select 2D classes from your previous 2D Classification run.



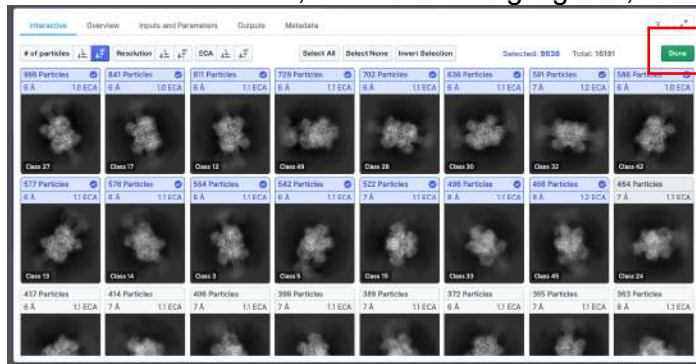
15. Choose your 2D classification run, press spacebar to open the overview tab, drag the the appropriate Outputs for Inputs fields, and submit the job.



16. In a moment an interactive window will open to allow selection of classes.



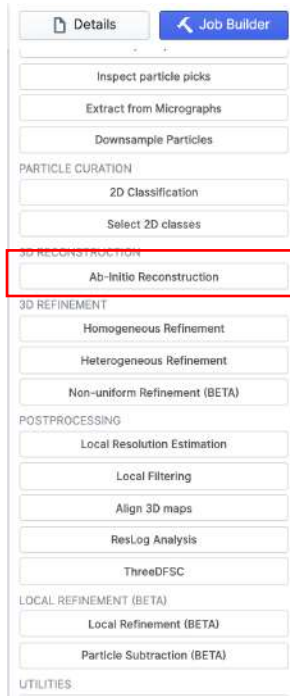
Select desired classes, which will be highlighted, and complete the job by selecting Done.



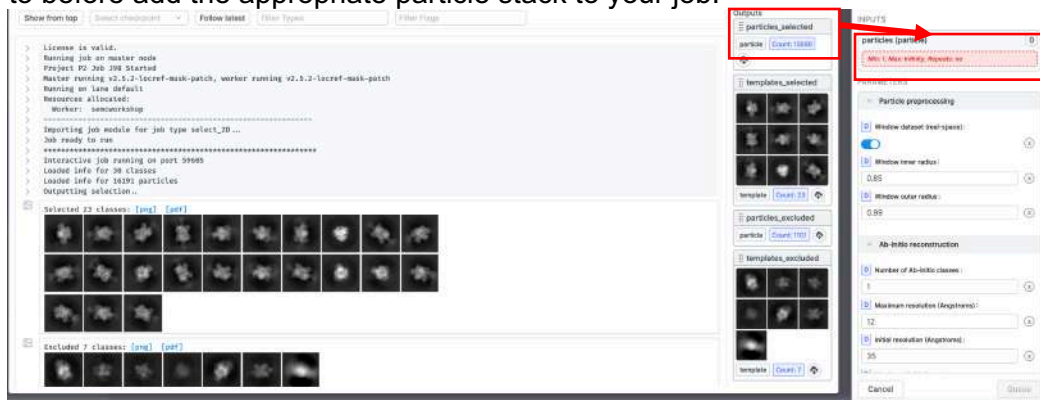
17. View the outputs in the overview window.

Ab-initio reconstruction

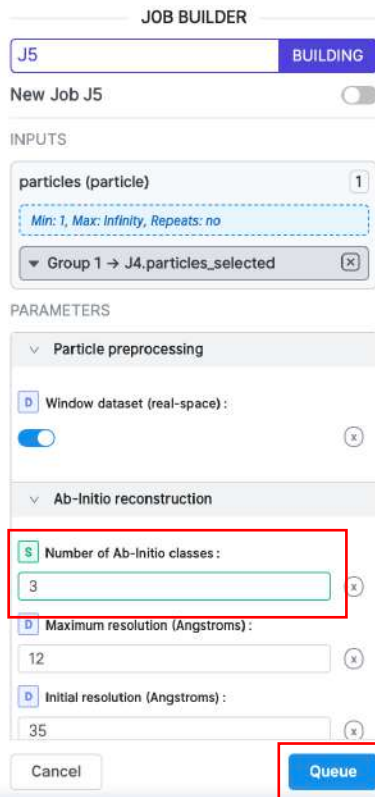
18. Create an *ab initio* structure from the curated particle set with ab-initio reconstruction.



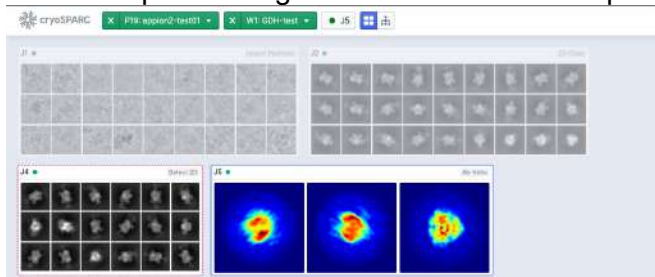
19. Similar to before add the appropriate particle stack to your job.



20. Choose to create multiple different potential models (for this tutorial choose a number between 1-4).

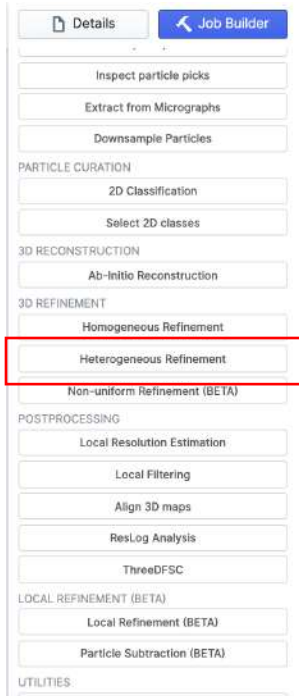


- - For this step start with C1 symmetry, but if time allows you may try again with different symmetries since we know this test data set has D3 symmetry.
21. Submit the job and monitor the outputs from each epoch.
- The workspace also gives an indication of the process of the run.

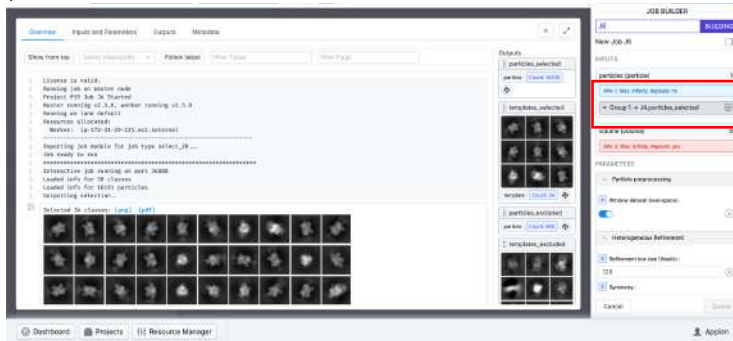


Heterogeneous refinement

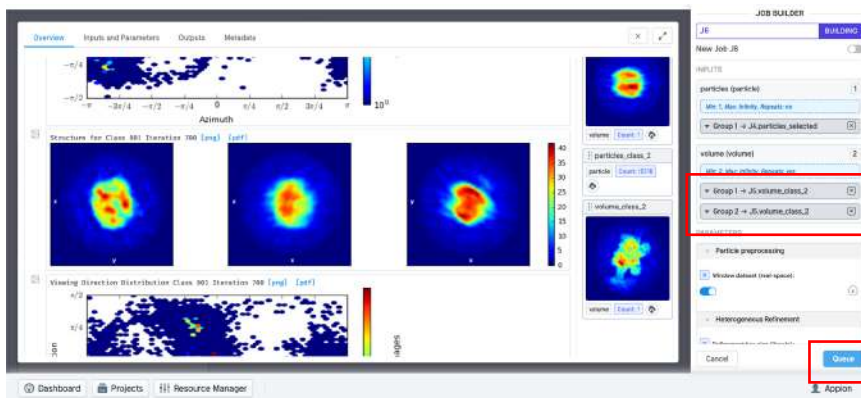
22. This test dataset has primarily one conformation, but in most datasets of macromolecular complexes there are multiple conformations, and/or varying state of completeness. One way to sort out these differences is to use heterogeneous refinement, however initial models are required. Since we have models via *ab initio* modelling we can potentially better curate the dataset with 3D models. Run a heterogeneous refinement run.



23. Add a particle stack. This may be the imported particle stack, particle stack from the selected 2D classes, or from the *ab initio* run. For this run choose either of the three.



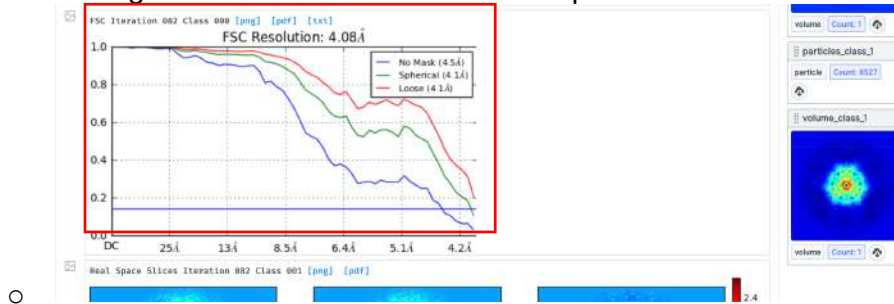
- Heterogeneous refinement is a reference based job and at least 2 models are required. Either the same model may be added multiple times, or different models may be added. For this tutorial add either 2, 3 or 4 volumes. For efficiency do not use more than 4 volumes.



- You may choose to either run this refinement with C1, C3 or D3 symmetry.

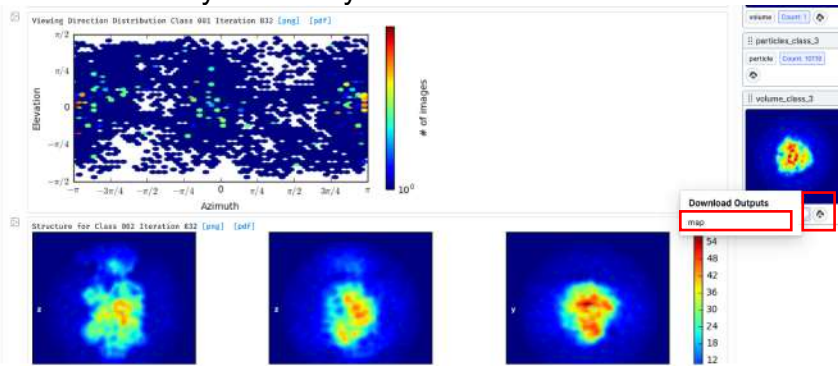
24. Submit the job and monitor the progress.

- As the iterations proceed a FSC and angular distribution graph will be displayed. Note that the FSC does not go to 0 because this refinement is binned for speed. Therefore, will not refine to the Nyquist frequency of the unbinned pixel size. When the job is complete a homogeneous refinement run will be required to do so



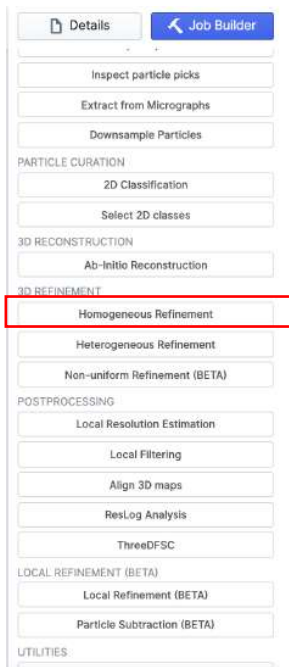
25. Use Chimera to inspect the resultant models.

- The models may be directly downloaded via the interface.

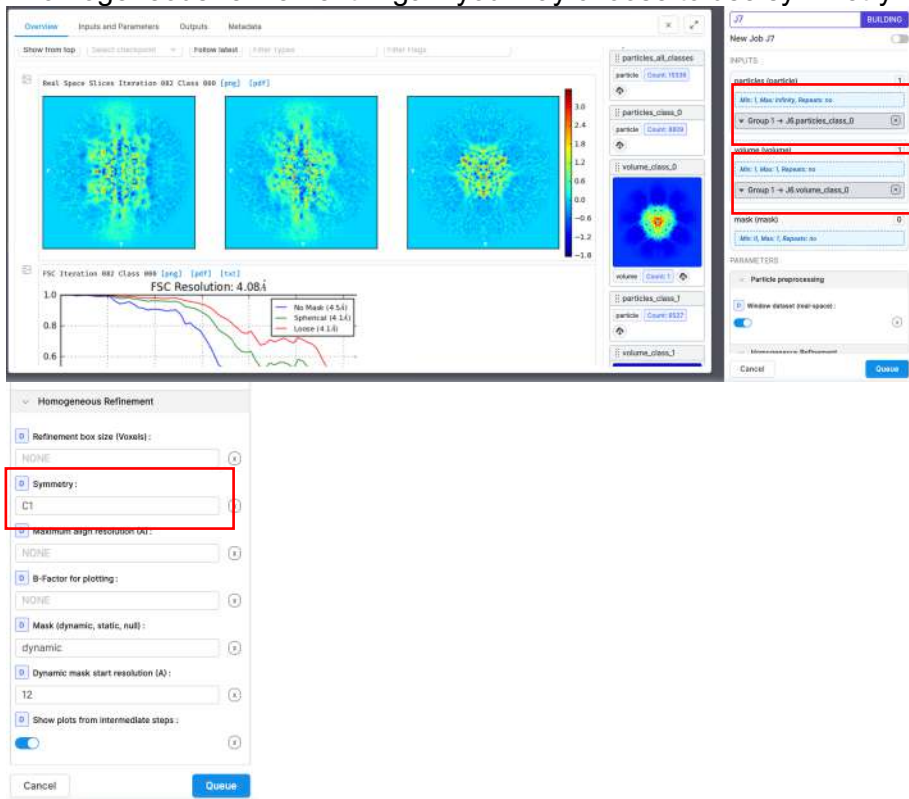


Homogeneous refinement

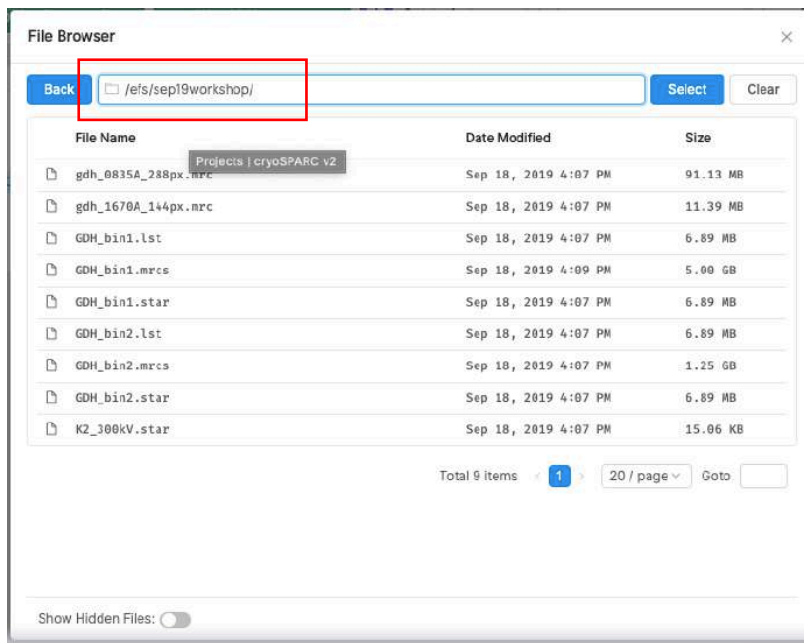
26. To create our final 3D reconstruction choose homogeneous refinement.



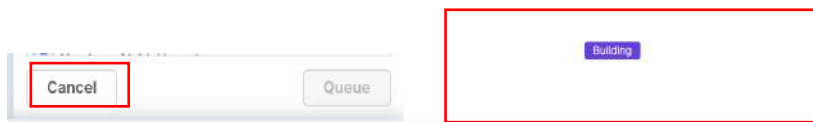
27. Add a particle stack and volume for the homogeneous refinement. Multiple particle stacks may be added if different 2 subsets from 2D or 3D are desired to be combined. Only 1 volume may be used for homogeneous refinement. Again you may choose to use symmetry.



28. For this tutorial an initial model is available in case ab initio modelling or heterogeneous refinement did not produce a reasonable model. To use one of these models Select Import Volume and import one of the following models.

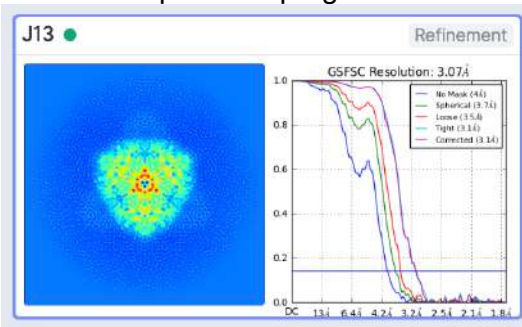


-
- In case you are already in the middle of building a job you may use the spacebar to exit the overview menu and exit the creation of the current job in the Job Builder by pressing “Cancel”. To return to building the job select the job that is still building and press “b” on the keyboard to resume building a selected job.

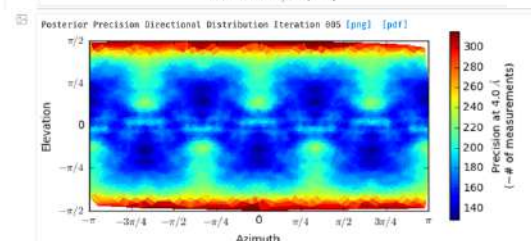
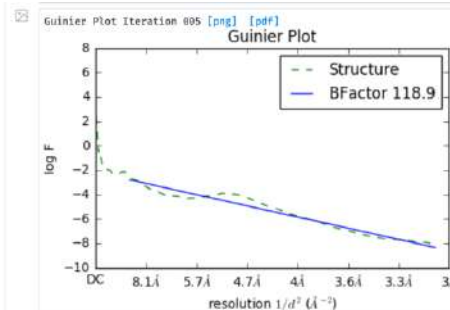
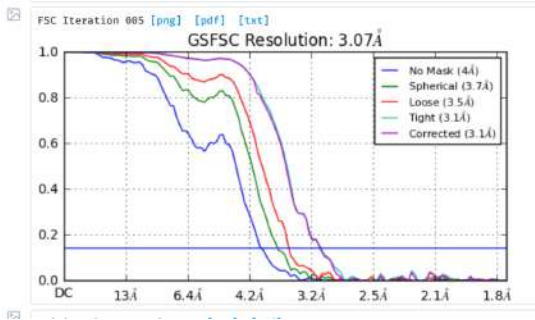


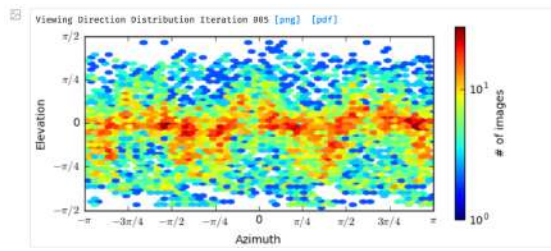
29. Submit the job and monitor the run.

- In the workspace the progress is shown.



- Analyse the feedback to ensure the quality of the reconstruction.





-
- Download the model and investigate with Chimera.

Additional tasks (if time allows)

30. Find a PDB of GDH and in Chimera check the handedness of your model.
31. Run any of the previous tasks with C1, C3 or D3 symmetry.
32. Try a different combination of particles or a different curation criteria to create a substack.
33. Try a different workflow.
34. Try additional tasks, such as ResLog Analysis, ThreeDFSC, etc...
35. Repeat the process with another dataset/movies.