SPA processing in cryoSPARC v5.0.0-privatebeta.8

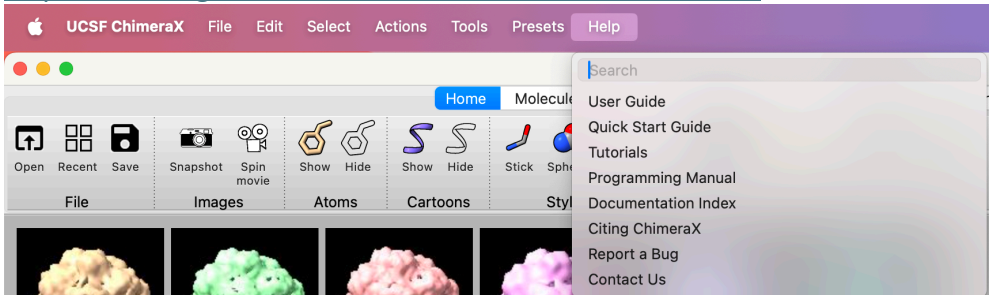
Goals & Objectives

- Become acquainted with the general SPA processing workflow by performing common processing tasks in cryoSPARC v5.0.0.
- Become acquainted with the cryoSPARC GUI and understand how it is used to perform key steps in SPA processing.
- In part 1, you will process a 200-micrograph dataset from movie alignment to 3D reconstruction. This is meant to simulate a small dataset one might obtain from a screening session. The goal is to assess micrograph and particle quality to determine whether a grid should go to large-scale data collection.
- In part 2, you will assess the conformational and compositional heterogeneity within a particle stack using 3D refinement and 3D classification tools. You will generate a clean consensus refinement at high resolution (Part 2A) and then assess conformational heterogeneity within that clean consensus particle stack (Part 2B).

Workflow and Setup

- Students will work in pairs
- Each student pair will have access 2 GPU for 3.5 hours
- Each group will present ~3minutes at the end. The content you share is totally up to you. Perhaps consider sharing:
    - Previous experience level (how often you use cryoSPARC)
    - 2D Classification results from Blob Picker (step 8) and Template Picker (step 13)
    - Best homogeneous refinement
    - What is something you found interesting about the data? Or did you perform a new job-type you have not performed before in cryoSPARC?
- If you find this tutorial elementary or find yourself with extra time at the end, please complete some step(s) in “further processing and analysis” section at the end
- If cryoSPARC becomes slow or you have trouble downloading files, you can access results here and inspect the volumes in Chimera on your computer
    - SPA Hemoglobin Results Part 1
    - SPA Hemoglobin Results Part 2

System requirements

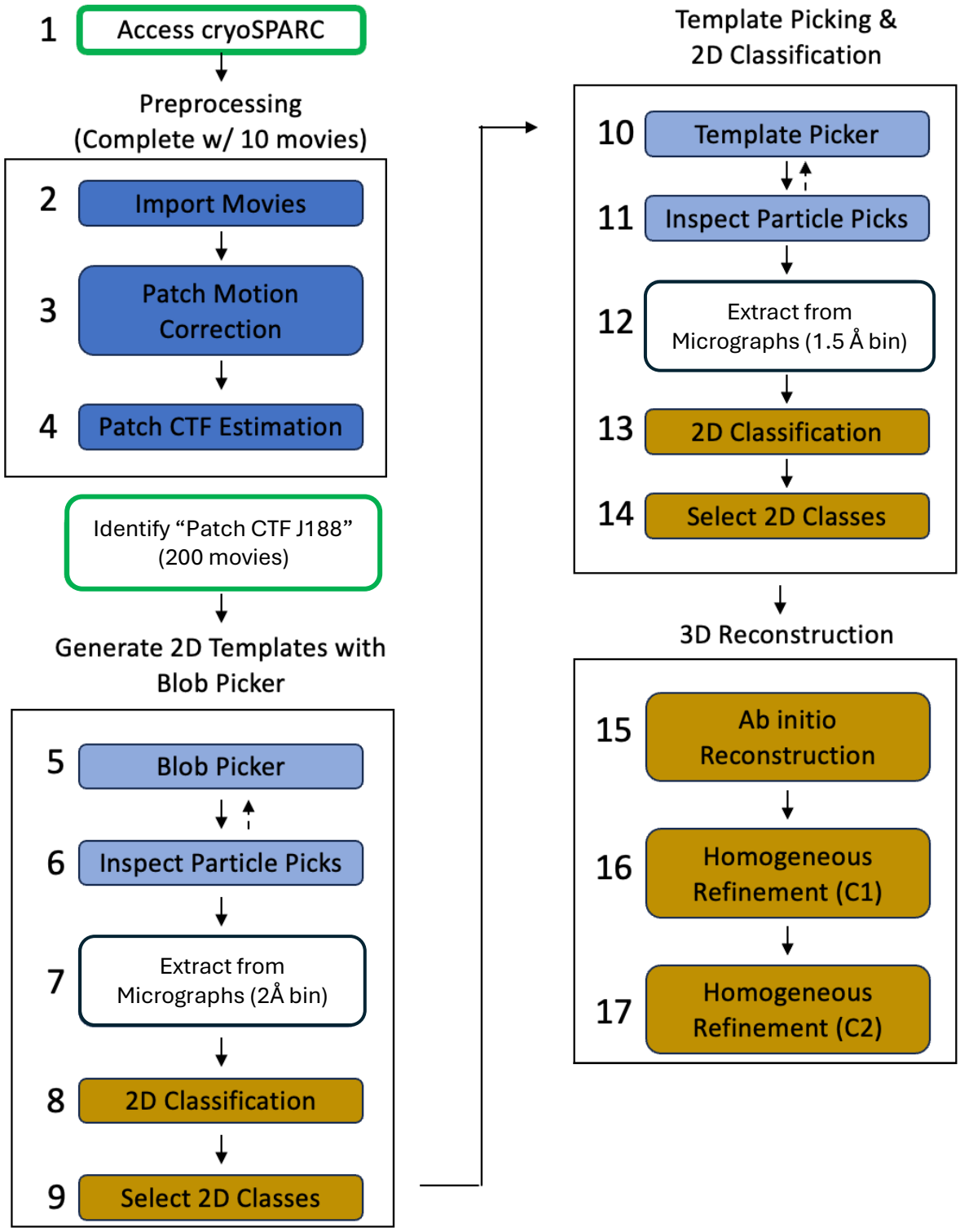
- Computer connected to an NYSBC Wi-Fi network and web browser
- UCSF ChimeraX – download and install from here (https://www.cgl.ucsf.edu/chimerax/download.html)
    - Many commands will differ from the earlier version UCSF Chimera, but both offer similar functions. Some functionalities which were available in Chimera are no longer available in ChimeraX, so it’s useful to have both installed.
    - The user guide for ChimeraX is available in the “Help” tab and online https://www.cgl.ucsf.edu/chimerax/docs/user/index.html
    - 

Additional Resources

The cryoSPARC Guide serves as an excellent resource for in-depth descriptions of each job type.
https://guide.cryosparc.com/processing-data/get-started-with-cryosparc-introductory-tutorial

The cryoSPARC Discussion Forum is an excellent community for troubleshooting, feature requests, and discussion of future developments. If you have a question that can't be answered by the cryoSPARC Guide, this is the next place to check. It's also a good place to search for solutions to specific error messages.
https://discuss.cryosparc.com

Part1: From movies to 3D consensus refinement



Accessing a Project and Workspace

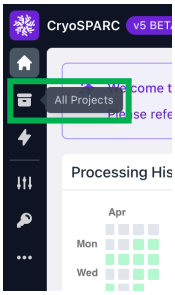
Background: The organizational hierarchy in cryoSPARC is Project -> Workspace -> Job.

- **Projects** are often used for different users, samples, or biological systems.
- **Workspaces** are often used for separate datasets. Each project can contain multiple workspaces
- **Jobs** are the basic processing operations of cryoSPARC and must be performed within a specific workspace. You can link jobs between workspaces, but you cannot link jobs between projects. Jobs can, however, be exported, manually moved to the destination project directory, then imported https://guide.cryosparc.com/guides-for-v3/tutorial-data-management-in-cryosparc#5.-ability-to-export-and-import-individual-jobs.

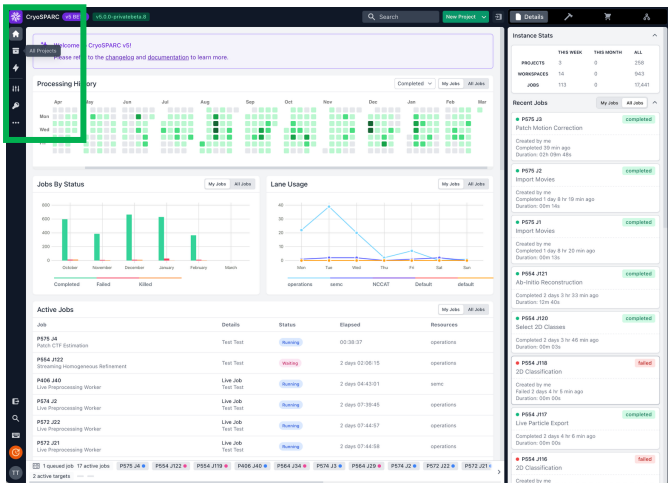
▶ Step 1 – System setup, accessing cryoSPARC and navigating to your workspace.

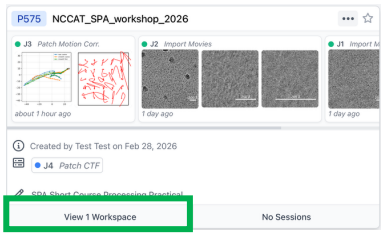
- 1.1 - Be at NYSBC on the Wi-Fi. In your web browser, enter: nccat-cryosparc.semc.nysbc.org
  The address may be flagged as untrusted by your browser. Proceed anyway.
- 1.2 - Sign in with the following information:
  - Email: test@test.test
  - Password: testtest
- 1.3 – Click the "All projects" icon in the top left. It looks like a file box.
  - Now you can see all the projects. These are different users' imaging sessions.
  - Please stay within the specified session for today (**P575**).
- 1.4 - Scroll down to **P575 NCCAT_SPA_workshop_2026**
- 1.5 - Click on the "**View # Workspaces**" button at the bottom of P575 card
- 1.6 – Find your group # and click "**View # Jobs**" at the bottom

Step 1 – Access cryoSPARC and navigate to your



1.3





1.4

1.5

Movie Preprocessing

You will import 10 movies (step 2) and run patch motion correction (step 3) and patch CTF estimation (step 4). You will then locate a completed Patch CTF Estimation job (J188) containing 200 micrographs that have undergone the same process and proceed to Blob Picker (step 5).

"Import Movies" job directs cryoSPARC to the location of movies on the working drive/database. This job also imports the gain file and requires input of data collection parameters (pixel size, spherical aberration, total dose, accelerating voltage). The Event Log will display unaligned summed frames that may appear streaky due to motion.

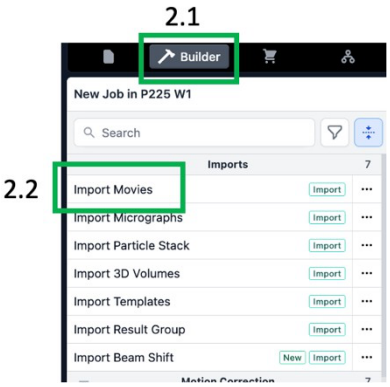
"Patch motion correction" aligns the raw movie frames in patches and applies gain correction. Patch motion correction creates diagnostic plots indicating the rigid motion trajectory and local trajectory for each patch. These aligned micrographs will be used for downstream jobs (patch CTF estimation, particle picking, and particle extraction).

"Patch CTF estimation" attempts to determine the defocus of aligned micrographs generated by Patch motion correction, and will measure additional parameters that vary from one micrograph to another; astigmatism, estimated resolution, and others.

You could import aligned micrographs generated from other software and directly run patch CTF estimation, skipping cryoSPARC's motion correction. This is helpful if you have run motion correction in a separate software or have transferred aligned micrographs generated during collection at a facility (e.g. *enn-a-DW.mrc files generated by motioncor2).

Import Movies

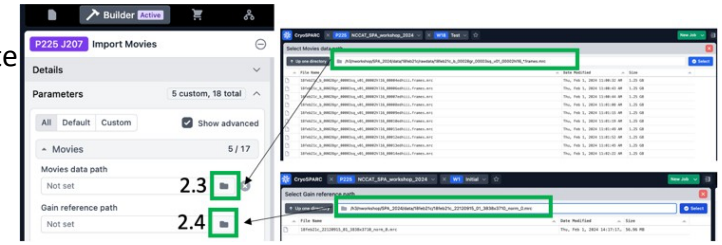
Step2 – Import Movies



- Step 2 - Import movies to cryoSPARC
  - 2.1 – Click "Builder" button
  - 2.2 – Click "Import Movies" at the top. A new job is built and you should have access to the "Details" and "Parameters" pages on the right.
  - 2.3 – Select Movies data path
    - click the folder icon for Movies data path & the Select movies data path page opens. From here you can navigate directories to point cryoSPARC to your movies. Copy and paste the following path into the directory bar at the top

/h3/nworkshop/SPA_2024/data/18feb21c/rawdata/18feb21c_b_00028gr_00003sq_v01_00002hl16_*.frames.mrc

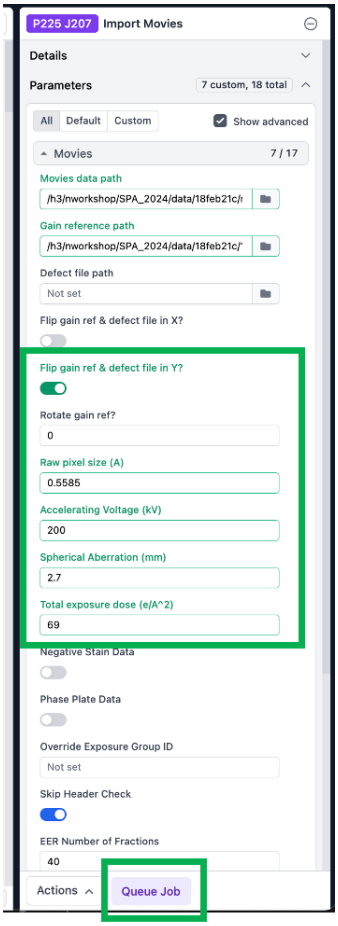
    - If this is done correctly, cryoSPARC should list the files. Ensure the files match the wildcard used. Sometimes multiple grids are present in a frames folder from a single microscope session and you must alter the wildcard to only include a single grid.
    - If it says, "Error, path must be an absolute path" then double check you entered it correctly.



- 2.4 – Select Gain reference path – Copy the following path:

  /h3/nworkshop/SPA_2024/data/small_subset/references/18feb21c_22120915_01_3838x3710_norm_0.mrc
- 2.5 – Set parameters
  - Flip gain ref and defect file in Y: ON
  - Raw pixel size (Å): 0.5585
  - Accelerating voltage (kV): 200
  - Spherical aberration (mm): 2.7
  - Total exposure dose (e-/Å2): 69
  - All others set to default. The text for parameters that have been changed from default turns green.
- A note about gain reference flipping and rotation. Different microscopes may produce gain reference that require different rotation or flipping – do not rely on a default rule of thumb as things can change as microscopes get upgraded. You can always ask your microscope operator if the gain reference needs to be flipped. You can also test the effects of flipping and rotation. When the gain is not properly applied, there will often be visible streaky artefacts in the gain-corrected image. View the effects of improper and proper gain-correction in jobs P575-W1 - J9, J10, and J1. It is quicker to run on a single micrograph or small subset than the full batch. The results are not always this obvious, so it can be helpful to view all the options in a PowerPoint slide together.

2.5



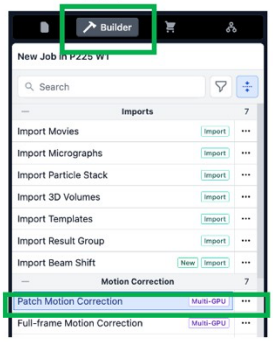
2.6 - Queue

Patch Motion Correction

- Step 3 – generate aligned micrographs
  - 3.1 – Click builder button
  - 3.2 – Click "Patch motion correction" job & a new job is built
  - 3.3 - Open the previously run Import Movies job (or exposure sets job) by clicking it and hitting the space bar on your keyboard – it should open. Alternatively, click the "four-corners square box" at the top of the job, to the right of "J## Import Micrographs" (this button is only visible once you hover the cursor over the job).
  - 3.4 – Drag "Imported movies" from the right side "2 output groups" section to the "inputs" Micrographs section of the Patch Motion Correction job builder.
  - 3.5 – Parameters:
    - Save results in 16-bit floating point: ON
    - Number of GPUs to parallelize: 2 (or more if available)
  - 3.6 – Click Queue, Choose semc lane and Queue

Step3 – Patch Motion Correction

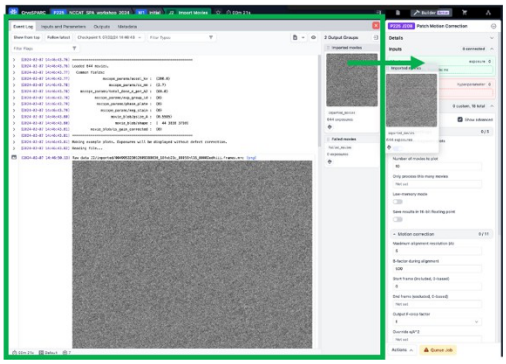
3.1



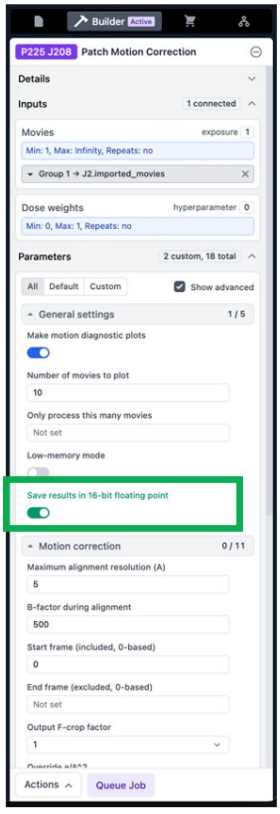
3.2

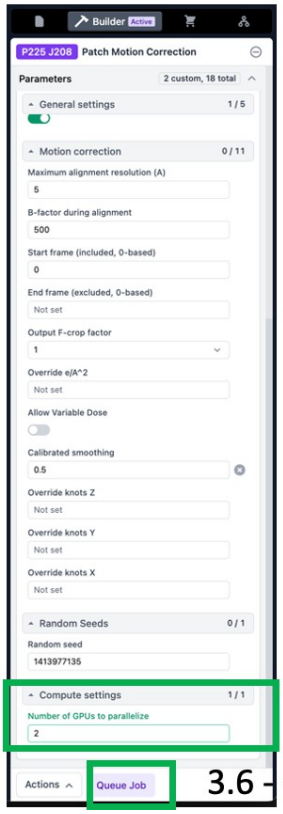
3.3 – Import movies open

3.4 – Drag input



3.5 - Parameters





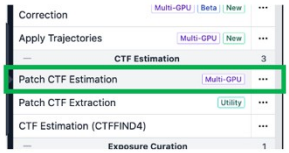
3.6 - Queue

Patch CTF Estimation

- Step 4 – Estimate the defocus for each micrograph
  - 4.1 – Click “Builder” button &
  - 4.2 – Click “Patch CTF Estimation” and the job builder opens
  - 4.3 – Open the previously run Patch Motion Correction job
  - 4.4 – Drag “Micrographs” from the right side “2 output groups” section to the “inputs” Micrographs section of the Patch CTF Estimation job builder.
  - 4.5 – Parameters – Only change the Compute Setting “Number of GPUs to parallelize” to 2 (from default of 1).
  - 4.6 – Queue Job on semc lane
  - Monitor diagnostic plots
    - 2D patch result visualizes the defocus across the whole micrograph. If you have a tilted sample, this should be tilted. If there are large features, like gold or carbon foil on the edge, large debris, or a very obvious ice gradient, these can often be visualized on this graph.

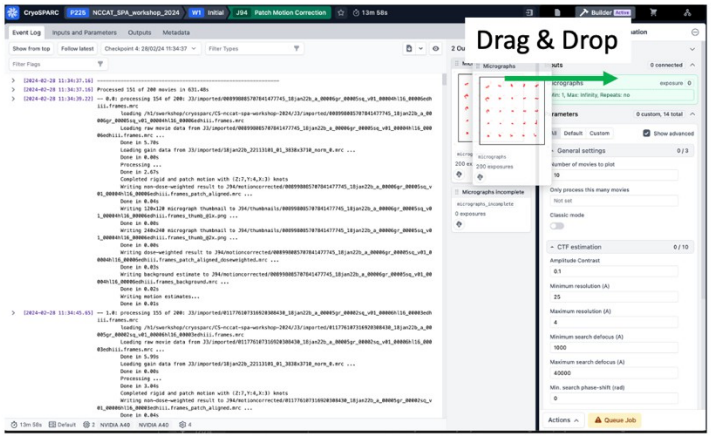
Step4 – Patch CTF Estimation

4.1 – Job Builder ->
4.2 - Patch CTF Estimation
Job builder opens on the right
(See 4.4 view)

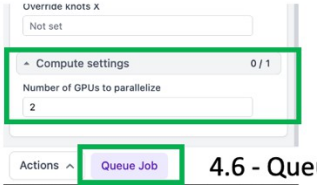


4.3 – open previously run Patch Motion Corr.

4.4 - Drag the Micrograph results output from Patch Motion to input



4.5 - Parameters



4.6 - Queue

Blob-based particle picking & Inspect Picks

Particle picking

There are various strategies for particle picking. When getting started on a new dataset, the shape and size of a particle may be unknown. Blob picker can be used to obtain initial particle picks, which will be extracted and 2D classified. The resulting 2D classes can be used as templates for Template Picking, which can be more accurate than Blob Picker.

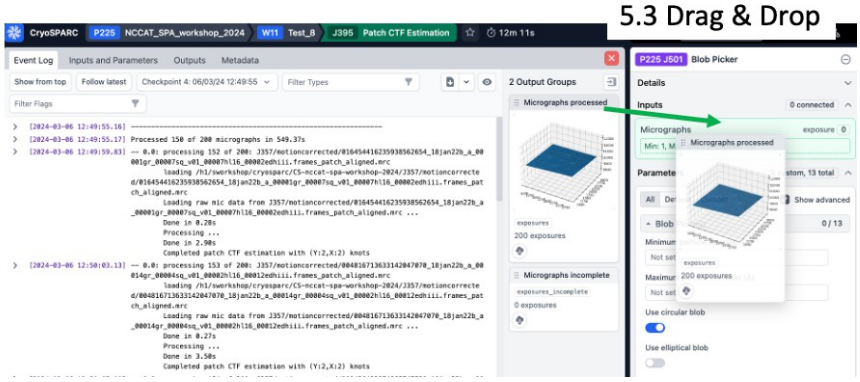
Inspect picks

Particle picking parameters often cannot be tuned to perfectly exclude off-target (non-protein) picks. The inspect particle picks job allows one to assess the quality of particle picks before proceeding to extraction and 2D classification. It also includes a filtering tool to remove particle picks based on a scoring system. We will use the filtering tool sliders to set thresholds (3 sliders to optimize) to exclude as many off-target non-protein picks as possible, while maintaining as many on-target picks as possible.

While 2D classification is useful to remove junk, removing as much junk as possible early on can give better results and an overall cleaner particle stack in the end.

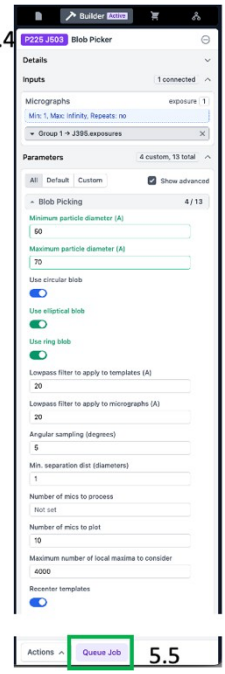
Step5 – Blob Picker

5.1 job builder
5.2 Blob picker
5.3 Drag & Drop



- Step 5 - Run Blob Picker
  - 5.1 - Click “Builder” button & then click “Blob Picker” & details and parameters pane opens
  - 5.2 - Open J188 “Patch CTF Estimation” which has been linked to your workspace
  - 5.3 – Drag “Micrographs Processed” into the Blob Picker input “Micrographs”
  - 5.4 - Parameters:
    - Minimum particle diameter (Å): 50
    - Maximum particle diameter (Å): 70
    - Use circular blob: ON
    - Use elliptical blob: ON
    - Not necessary to change other parameters
  - 5.5 - QUEUE JOB
  - Compute considerations – this job can only use 1 GPU (no parallelization option)
  - 5.6 – As the job runs, you can follow progress in the event log and see the red dot blob picks for each micrograph. Once the first 10 micrographs have completed, it will stop showing diagnostic plots. You can re-visit these by clicking “show from top” to go see the first 10 micrograph particle picks. You can also see the shape and size of the circular and elliptical blobs that were generated for picking.

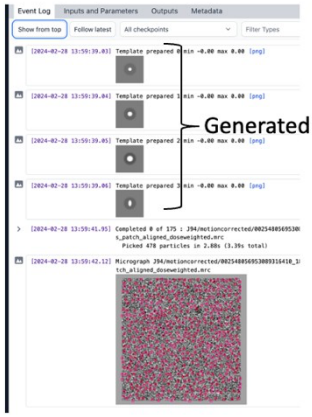
- Step 6 – Inspect particle picks
  - 6.1 - Click “Builder” button & then click “Inspect particle picks” & the details and parameters pane opens
  - 6.2 - Open the previously run “Blob Picker” job so the output groups are accessible
  - 6.3 – Drag “All particles” from the completed blob picker job into the inspect picks “Particles” input and drag “Micrographs” into the Micrographs input.
  - 6.4 – Queue – this job is interactive
  - 6.5 - The green circles depicting particle picks can get in the way of viewing the micrograph. Drag the “Particle diameter” slider to ~15 Å to put a small green circle on each particle pick.
  - 6.6 – click through the micrographs. You can sort them by number of picks, CTF estimated resolution, or other parameters
  - 6.7 – Drag the NCC score slider to the right and observe how many off-target picks are removed. Optimize this over 5 to 10 micrographs to remove off-target picks while maintaining on-target picks.
  - 6.8 – Adjust the power histogram sliders. Sort the micrographs by according to the Average defocus column. Use the highest defocus image to adjust the upper bound to remove picks from e.g. ice contamination. Use the lowest defocus image to adjust the lower bound to remove picks from background noise. Adjust these to maximize on-target picks while reducing off-target picks. Check these parameters on a few micrographs to verify they are good.
  - 6.9 - Click “Done Picking | Output Locations” in the top right



5.4

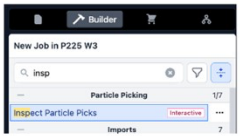
5.5

5.6 – inspect results



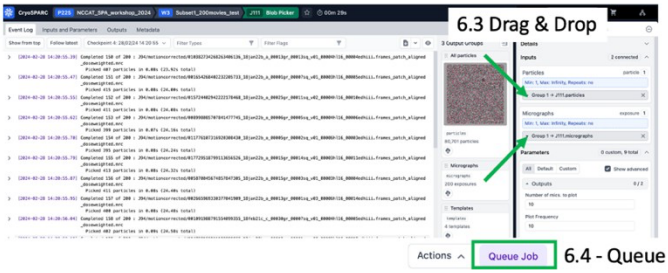
First 10 micrograph picks

Step6 – Inspect Particle Picks



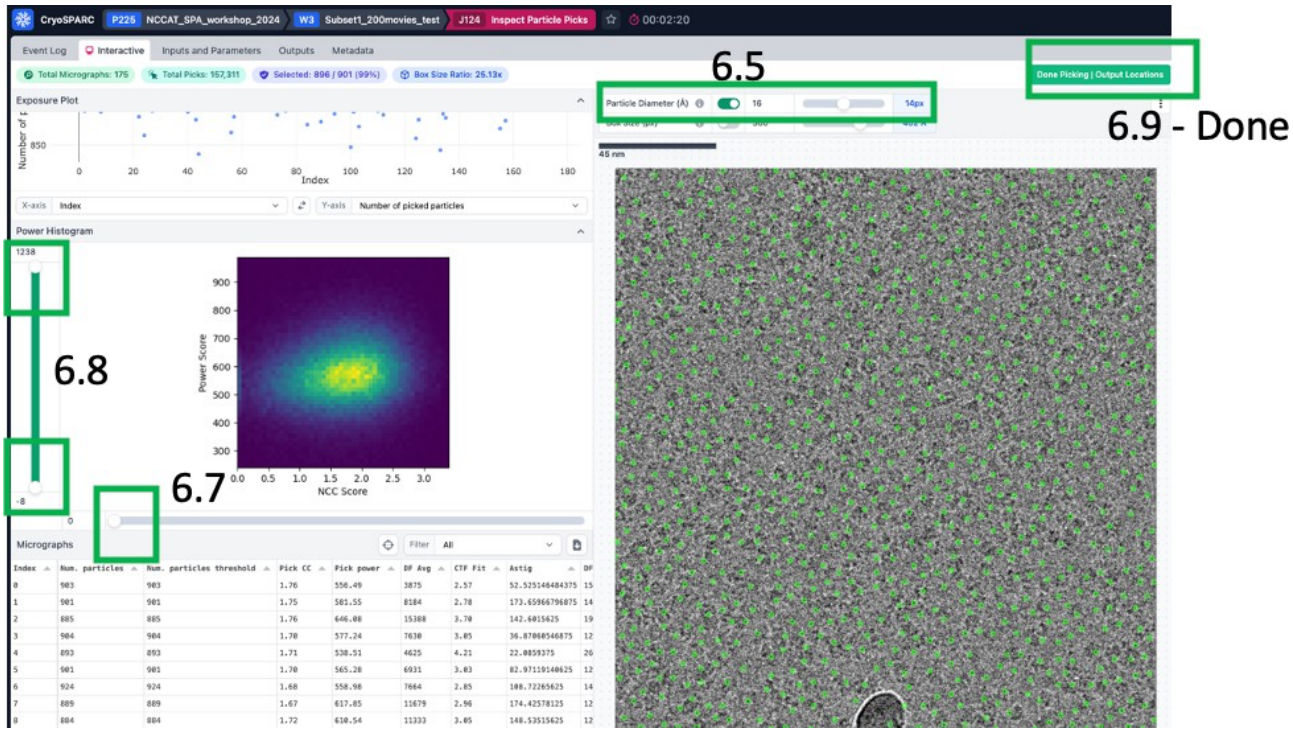
6.1 - job builder -> Inspect Particle Picks

6.2 - Blob picker is open

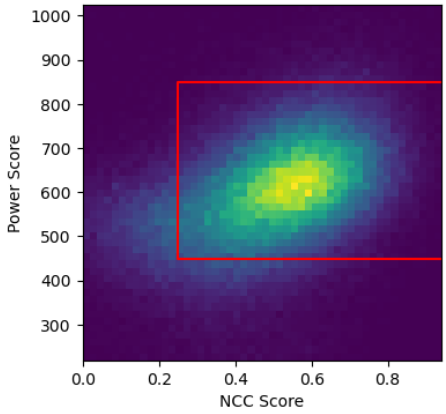


6.4 - Queue

- It is better to overpick than to underpick, because you can filter out many off-target picks with 2D classification. At this stage, our goal is to use these Blob-picked particles generate usable 2D class averages for template picking. These 2Ds will be lowpass filtered to low resolution during picking anyway, so the picks from this job just need to seem reasonable.



- Make sure that the thresholds chosen include most of the yellow-green peak in the Power Score vs NCC Score plot. A red box will appear in the Event Log after you click “Done” showing the particles which have been included. The example below shows a plot with an NCC of 0.25 and Power Score range of 450-850).



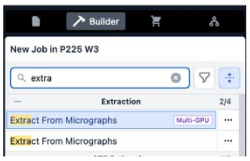
Particle Extraction

Extract From Micrographs (GPU)
This job generates a “particle stack” or a collection of images with specified box size taken from the micrograph. Key parameters to take into consideration are the extraction box size and the amount of binning (Fourier cropping) to apply. cryoSPARC has different job types for GPU-based extraction or CPU-based extraction and their output particle stacks will be the same.

Determining an appropriate box size
Generally, choose a box size that is 2-3x larger than the longest particle dimension to include high-resolution information in the Fourier transform of each extracted particle. If your particle is 50 x 50 x 70 Å, we would use a box size of 70 x 2 = 140 Å. The box size input parameter is in pixels, so convert 140 Å to pixels (140 Å / 0.5585 Å/pix = 250 pixels). Certain box sizes allow for faster processing, so we’ll choose the next highest value in this list (https://jianglab.science.psu.edu/2018/07/19/how-big-should-my-particle-box-be/), which is 256. Box sizes must be an even integer.
We can iteratively optimize this based on the 2D classification results, which can be helpful if we miscalculate or have unexpected features. For example, when particles are very close together, we might want a somewhat smaller box to exclude neighboring particles or to include a tighter mask.

Calculate the binning (Fourier crop) as follows:
Mic pixel size (Å/px) * Extract box size (px) = Fourier crop box size (px) * bin pixel size (Å/px)
let’s say we want to extract in a box of 256 pixels and bin to 2.5 Å. Calculate the Fourier crop box size.
(0.5585 Å/px * 256 px)/ (2.5 Å/pix) = Fourier crop box size
197.12 / 2.5 = 57.2 pixels
The box size must be an even integer and is best chosen from the list, so we’ll round up to 64 as Fourier crop box size. This is also a nice even divisor of 256. What’s the actual binned pixel size after a crop to 64 px?
(0.5585Å/pix x 256 pix)/64px = binned pixel size Å
2.234 Å/pix = binned pixel size (Nyquist of 4.468 Å)

Step7 – Extract From Micrographs (GPU) with binning

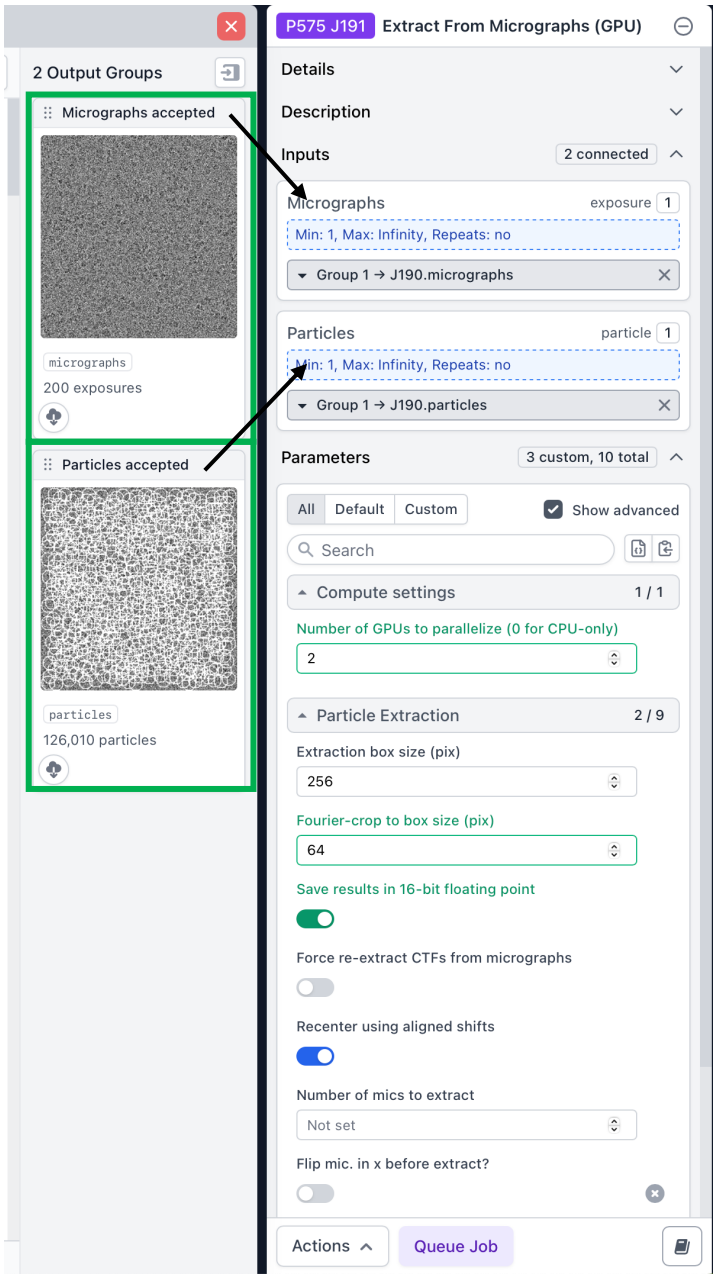


7.1 job builder -> Extract from Micrographs (GPU)*
* If GPU is not available, CPU-based job can be run and you can specify the number of threads

- Step 7 – Extract From Micrographs (GPU) with binning
  - 7.1 - Click “Builder” button & then click “Extract” & the details and parameters pane opens
  - 7.2 - Open the previously run “Particle Inspection” job so the output groups are accessible
  - 7.3 – Drag “All particles” from the completed blob picker job into the inspect picks “Accepted Particles” input and drag “Micrographs” into the Micrographs input.
  - 7.4 – Parameters
    - Extraction box size (pix): 256
    - Fourier Crop to box size (pix): 64
    - Save results in 16-bit floating point: ON
  - 7.5 Queue job

7.2 Inspect Particle Picks is open

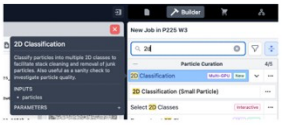
7.3



7.4

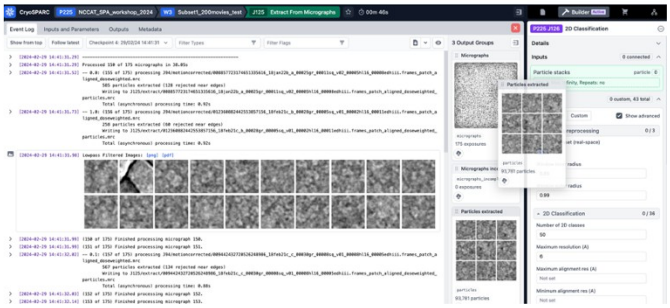
- Step 8 – 2D Classification
  - 8.1 - Click “Builder” button & then click “2D Classification” & the details and parameters pane opens
  - 8.2 - Open the previously run “Extract From Micrographs” job so the output groups are accessible
  - 8.3 – Drag the “Particles Extracted” from the completed Extract from Micrographs job into the 2D Classification “Particle Stacks” input
  - 8.4 – Parameters
    - Number of 2D classes: 50 (default)
    - Number of GPUs to parallelize: 2
    - Circular mask diameter: 80
    - Cache particle images on SSD: off
  - 8.5 Queue Job
    - The job runs in 3 minutes for 100 classes with ~100k particles binned in a ~64 pixel box.

Step8 – 2D Classification



8.1 Job builder -> 2D Classification
8.2 Extract From Mics job is open

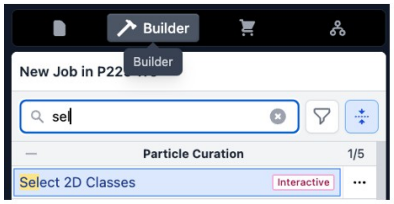
8.3 Drag & Drop



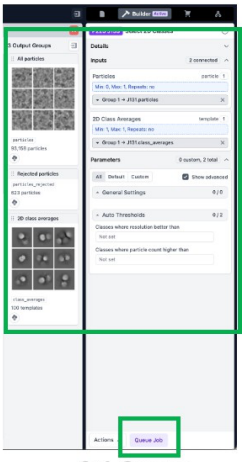
Step9 – Select 2D Classes

- Step 9 – Select 2D Classes
  - 9.1 - Click “Builder” button & then click “Select 2D Classes” & the details and parameters pane opens
  - 9.2 - Open the previously run “2D Classification” job so the output groups are accessible
  - 9.3 – Drag the “Particles” and 2D class averages” from the completed 2D classification job
  - 9.4 – Queue job – this job is interactive, wait for the “wait” step (it will turn pink)
  - 9.5 - Select 4-6 classes of interest which represent multiple views
  - 9.6 - Click “Done”

9.1 Job builder -> Select 2D Classes
9.2 2D Classification job is open



9.3 Drag & Drop

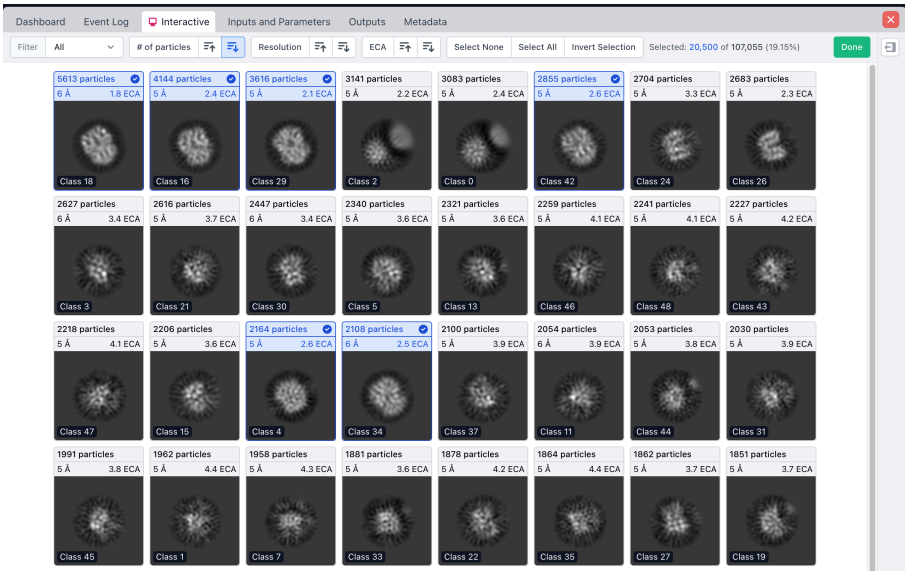


9.4 Queue

Choose 2D classes with protein-like features and expected shape to use for template picking

9.5 select

9.6 Done



Template Picker, Extraction, & 2D Classification

Template Picker

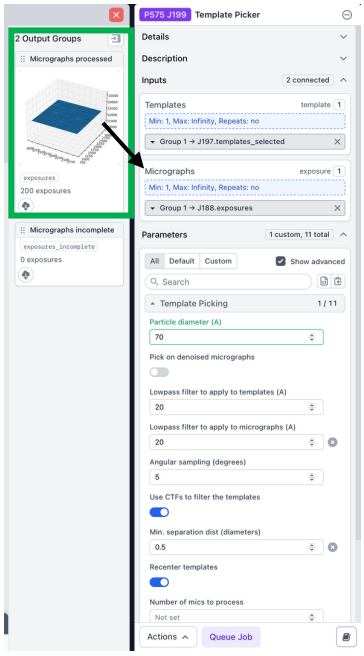
Now you have 2D classes that we can use as templates for template-based particle picking. Hopefully, template-based particle picking will give more on-target protein picks and increase the number of high-quality particles. These particle picks should also be better centered than blob-based particle picks.

Extraction

Again, we will extract with binning for the particles that will be used for 2D classification. We'd optimistically like our reconstruction to reach 3 Å, so we'll aim for a pixel size of 1.5 Å/px. We'll therefore bin by a factor of ~3, and again choose box sizes guided by the Jiang lab's list of good box sizes. Another useful tool is the Takanori CTF calculator (https://3dem.github.io/relion/ctf.html) which can show you the predicted CTF for a given box size, pixel size, and defocus. You can use this tool to inform your choice of box size and pixel size in order to minimize Moire pattern interference at the spatial frequency of your desired resolution.

2D classification

- Step 10 – Template Picker
  - 10.1 - Click "Builder" button & then click "Template Picker" & the details and parameters pane opens
  - 10.2 - Open the previously run "Select 2D Classes" job so the output groups are accessible
  - 10.3 – Drag the "Templates Selected" output from the completed "Select 2D Classes" job and drag it to the Templates input in the Template Picker builder
  - 10.4 – Open the previously run Patch CTF job
  - 10.5 - Drag the "Micrographs Processed" output from the completed "Patch CTF" job and drag it to the Micrographs input in the Template Picker builder
  - 10.6 – Parameters
    - Particle diameter (Å): 70
    - Min. separation dist (diameters): 0.5
  - 10.7 - Queue
  - 10.8 – follow the progress, click show from top and view the first 10 micrographs, as well as the low-pass filtered templates used for picking.
  - Compute considerations – this job requires 1 GPU and cannot be parallelized



10.4 Patch CTF job is open
10.5 Drag and drop micrographs processed
10.6 Parameters

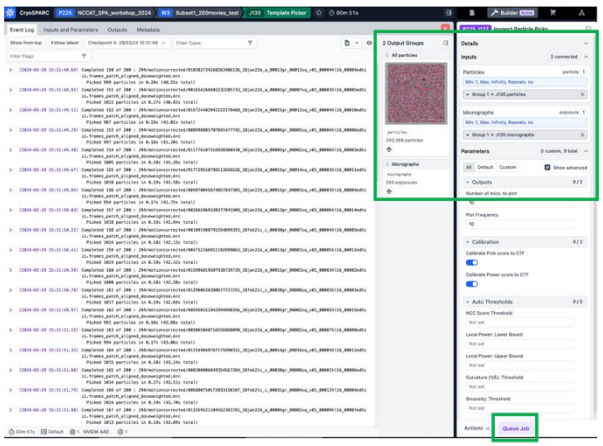
- Step 11 – Inspect particle picks
  - 11.1 - Click "Builder" button & then click "Inspect particle picks" & the details and parameters pane opens
  - 11.2 - Open the previously run "Template Picker" job so the output groups are accessible
  - 11.3 – Drag "All particles" from the completed template picker job into the inspect picks "Particles" input and drag "Micrographs" into the Micrographs input.
  - 11.4 – Queue – this job is interactive
  - 11.5 - The green circles depicting particle picks can get in the way of viewing the micrograph. Drag the "Particle diameter" slider to ~15 Å to put a small green circle on each particle pick.
  - 11.6 – click through the micrographs. You can sort them by number of picks, CTF estimated resolution, or other parameter
  - 11.7 – Drag the NCC score slider to the right and observe how many off-target picks are removed. Optimize this over 5 to 10 micrographs to remove off-target picks while maintaining on-target picks.
  - 11.8 – Adjust the power histogram sliders as with the blob picks to remove off-target picks while maintaining on-target picks.
  - 11.9 - Click "Done Picking | Output Locations" in the top right

Step11 – Inspect Particle Picks (repeat step 6)

11.1 job builder -> Inspect Particle Picks

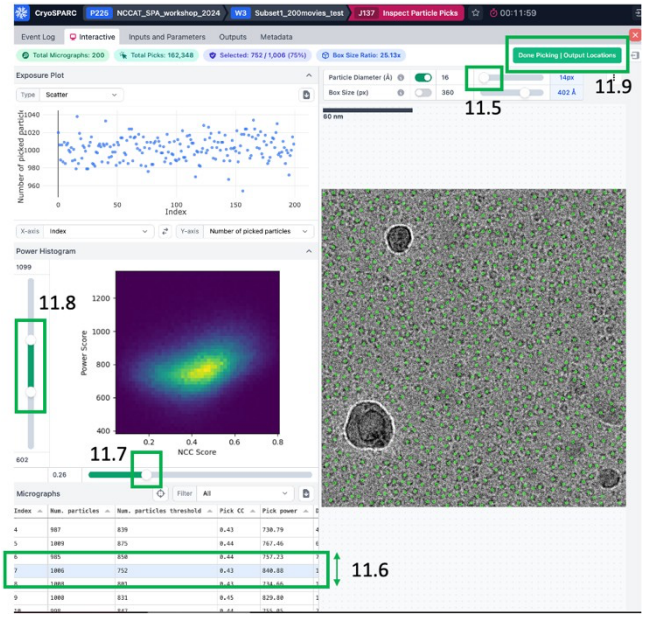
11.2 Template picker is open

11.3 Drag & Drop



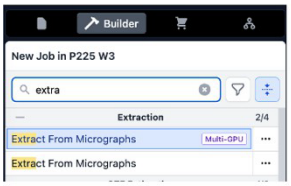
11.4 Queue

Step11 – Inspect Particle Picks



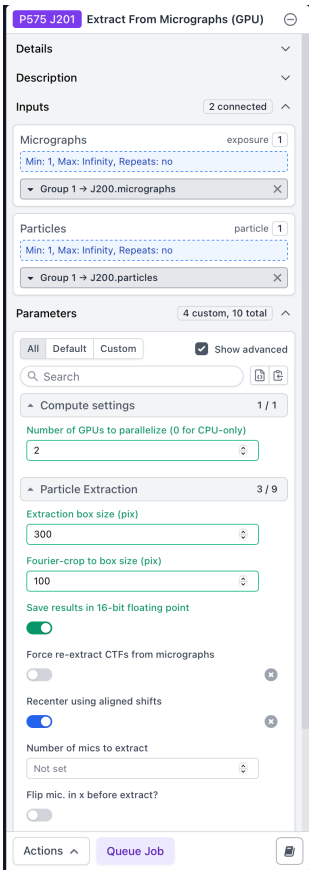
- Step 12 – Extract with binning
  - 12.1 - Click “Builder” button & then click “Extract From Micrographs” & the details and parameters pane opens
  - 12.2 - Open the previously run “Particle Inspection” job so the output groups are accessible
  - 12.3 – Drag “All particles” from the completed template picker job into the inspect picks “Accepted Particles” input and drag “Micrographs” into the Micrographs input.
  - 12.4 – Parameters
    - Number of GPUs to parallelize: 2
    - Extraction box size (pix): 300
    - Fourier crop to box size (pix): 100
    - Save results in 16-bit floating point: ON (saves space)
  - 12.5 Queue job

Step12 – Extract From Micrographs (GPU) with binning (repeat step 7)



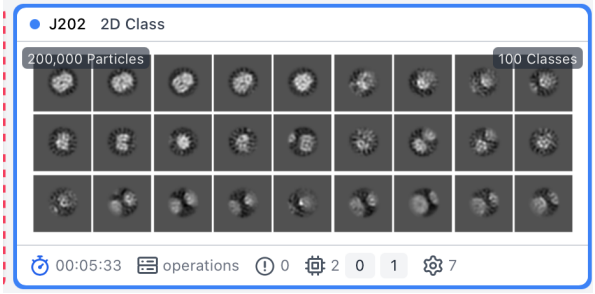
12.1 job builder -> Extract from Micrographs

12.4 Parameters

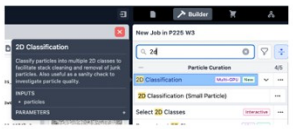


- Step 13 – 2D Classification(repeat Step 8 with the same settings)
  - 13.1 - Click “Builder” button & then click “2D Classification” & the details and parameters pane opens
  - 13.2 - Open the previously run “Extract From Micrographs” job so the output groups are accessible
  - 13.3 – Drag the “Particles Extracted” from the completed Extract from Micrographs job into the 2D Classification “Particle Stacks” input
  - 13.4 – Parameters
    - Number of 2D classes: 100
    - Maximum resolution: 10
    - Circular Mask Diameter: 80
    - Cache particles to SSD: OFF
    - Number of GPUs to parallelize: 2
  - 13.5 Queue Job
    - The job runs in 5-10 minutes for 100 classes with ~200 k particles.

Monitor progress

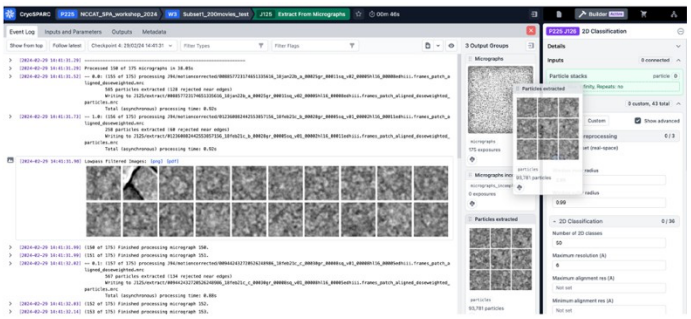


Step13 – 2D classification (repeat step 8)

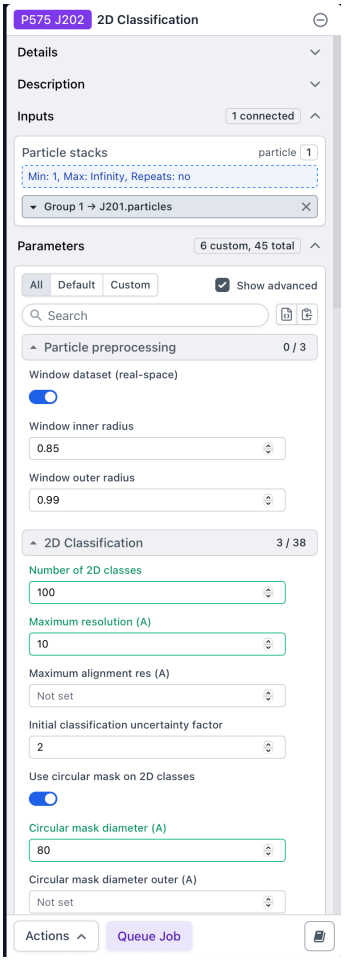


13.1 Job builder -> 2D Classification
13.2 Extract From Mics job is open

13.3 Drag & Drop



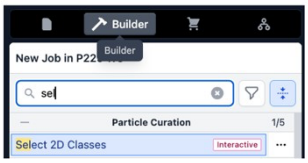
13.4 Parameters



- Step 14 – Select 2D Classes
  - 14.1 - Click “Builder” button & then click “Select 2D Classes” & the details and parameters pane opens
  - 14.2 - Open the previously run “2D Classification” job so the output groups are accessible
  - 14.3 – Drag the “Particles” and “2D class averages” from the completed Extract from Micrographs
  - 14.4 – Queue job – this job is interactive, wait for the “wait” step (it will turn pink)
  - 14.5 - Select all classes of interest
  - 14.6 - Click Done
  - This can be the final particle stack or you can run another select job followed by another round of 2D classification.

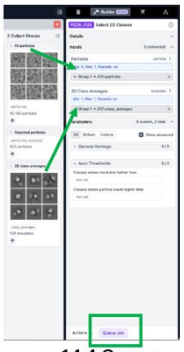
Step 14 – Select 2D Classes (repeat step 9)

14.1 - Job builder -> Select 2D Classes



14.3 Drag & Drop

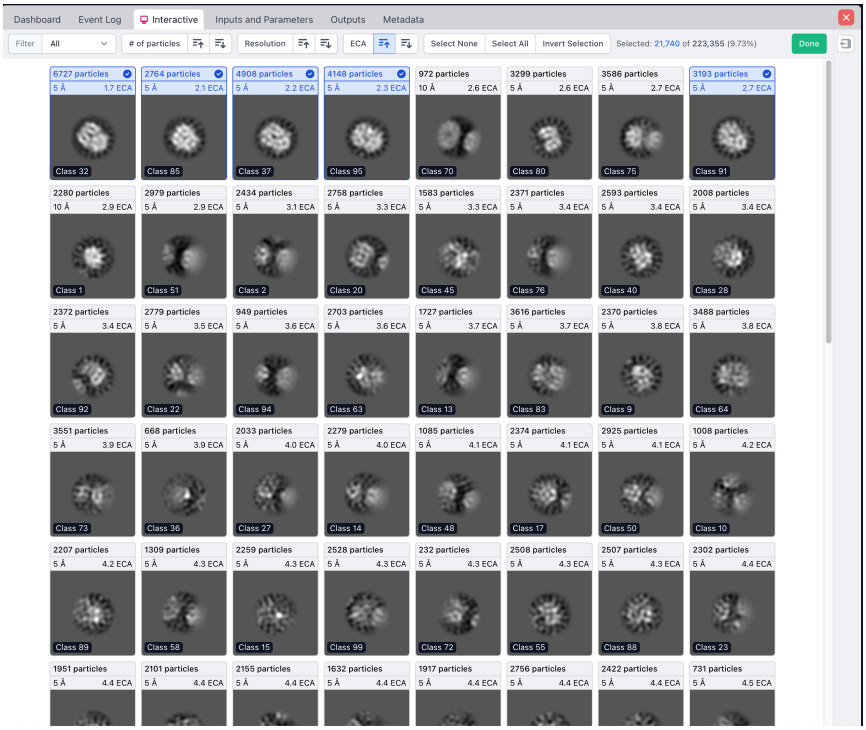
14.2 - 2D Classification job is open



14.4 Queue

14.5 Select Classes

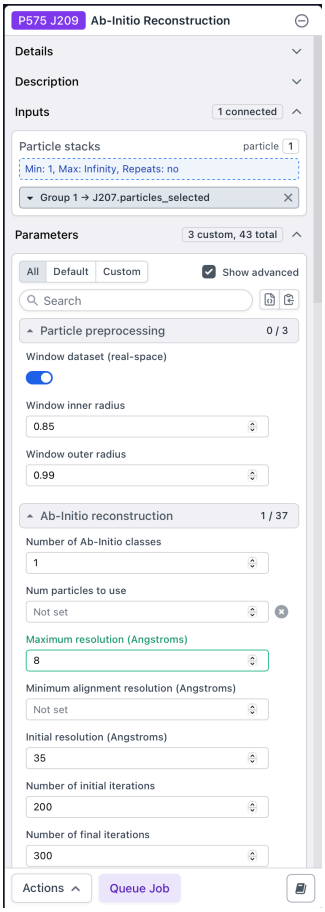
14.6 Done



Generate consensus refinement

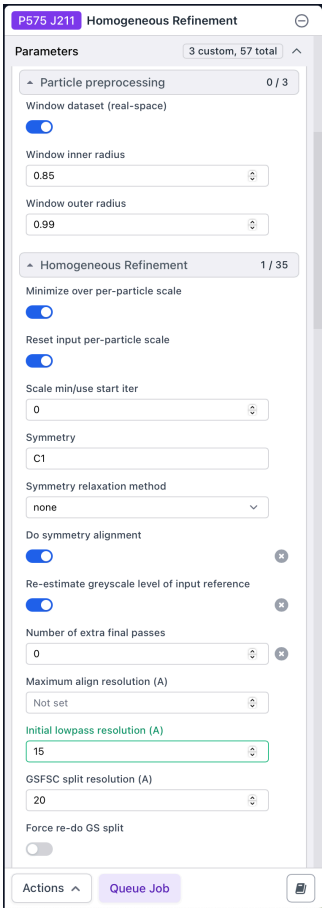
- Step 16 – *Ab Initio* reconstruction
  - 16.1a – Use the job builder to create the job and drag the particles from Step 15 as input OR
  - 16.1b – Right click the Extract job from Step 15 and click "Build Ab-initio reconstruction" to automatically populate the newly built job with the Extract particle output
  - 16.2 - Parameters:
    - Number of Ab-inito classes: 1
    - Maximum resolution: 8
    - Cache particle images on SSD: OFF

16.2 Parameters



- Step 17 – Homogeneous Refinement
  - 17.1a – Open as usual with the Job Builder OR
  - 17.1b – Right click the Ab initio job from Step 16 and "Build Homogeneous Refinement" – the new job auto populates with the ab initio particles and volume output
  - 17.2 – Parameters
    - Initial lowpass resolution: 8
    - Cache particles on SSD: OFF
    - Keep all others default

17.2 Parameters

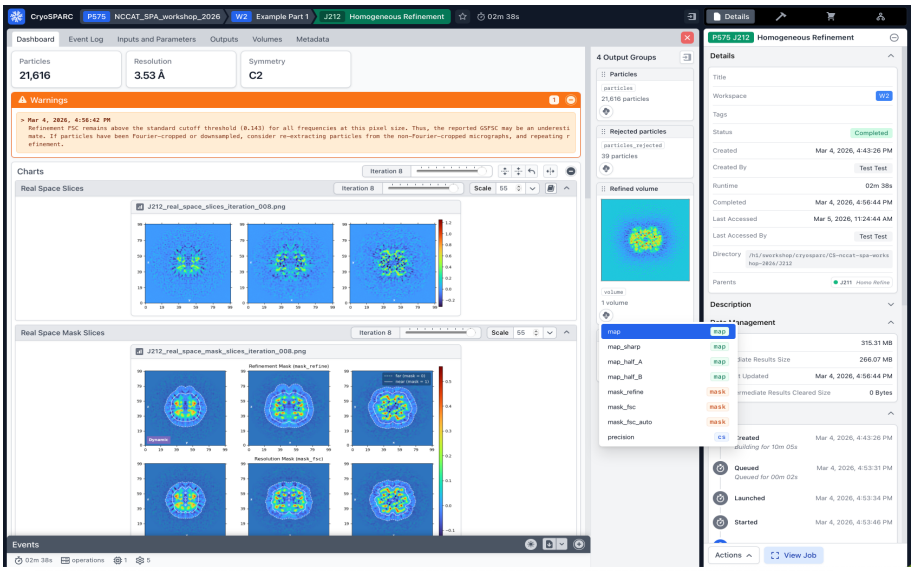


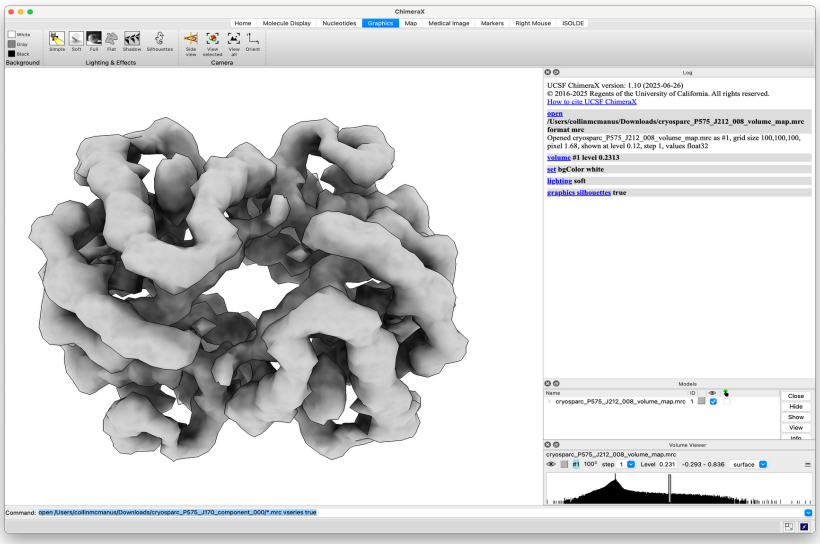
- Step 18 – Homogeneous Refinement with symmetry enforced. This uses the alignments established in your first homogeneous refinement but effectively multiplies your number of particles by the order of

symmetry.

- 18.1a – Open as usual with the Job Builder
- 18.1b – Use Particles and Volume from previous homogeneous refinement as input
- 18.2 – Parameters
  - Symmetry: C2
  - Initial lowpass resolution: 8
  - Cache particles on SSD: OFF
  - Keep all others default
- 18.3 – Queue
- 18.4 – Inspect Results either in the Volumes tab of the job or download the map and view in ChimeraX
  - In the volume below, background has been set to white, display settings to soft, and silhouettes turned on.

18.4 View



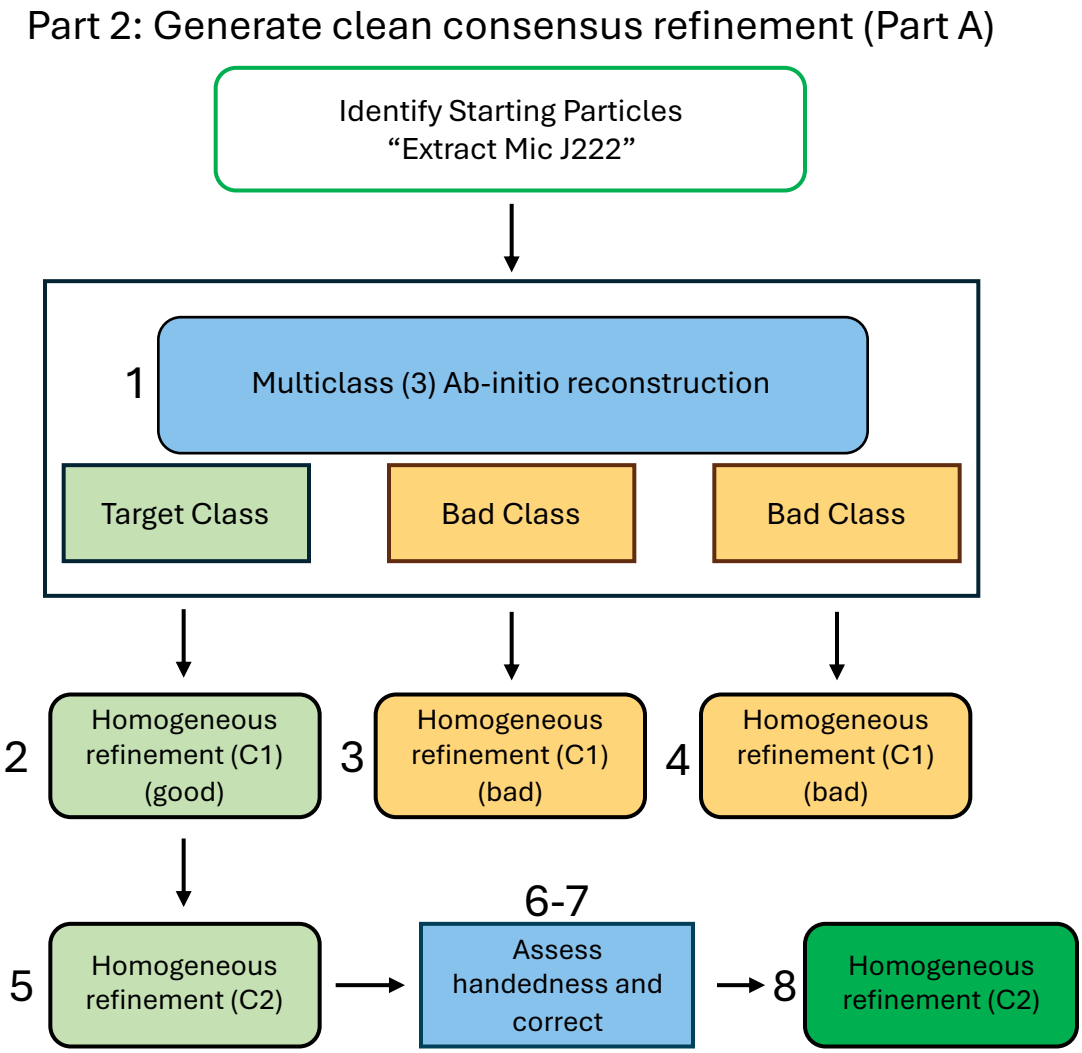


Part2

This part will use particles obtained by 2D classification for the whole dataset (1,663 micrographs). You will generate a 3D reconstruction (Part A) and assess the compositional and conformational heterogeneity (Part B) for the particles from this dataset using 3D refinement and classification tools in cryoSPARC. If time begins to run short, we can link the clean C2 refinement to your workspaces to begin Part B before you complete Part A.

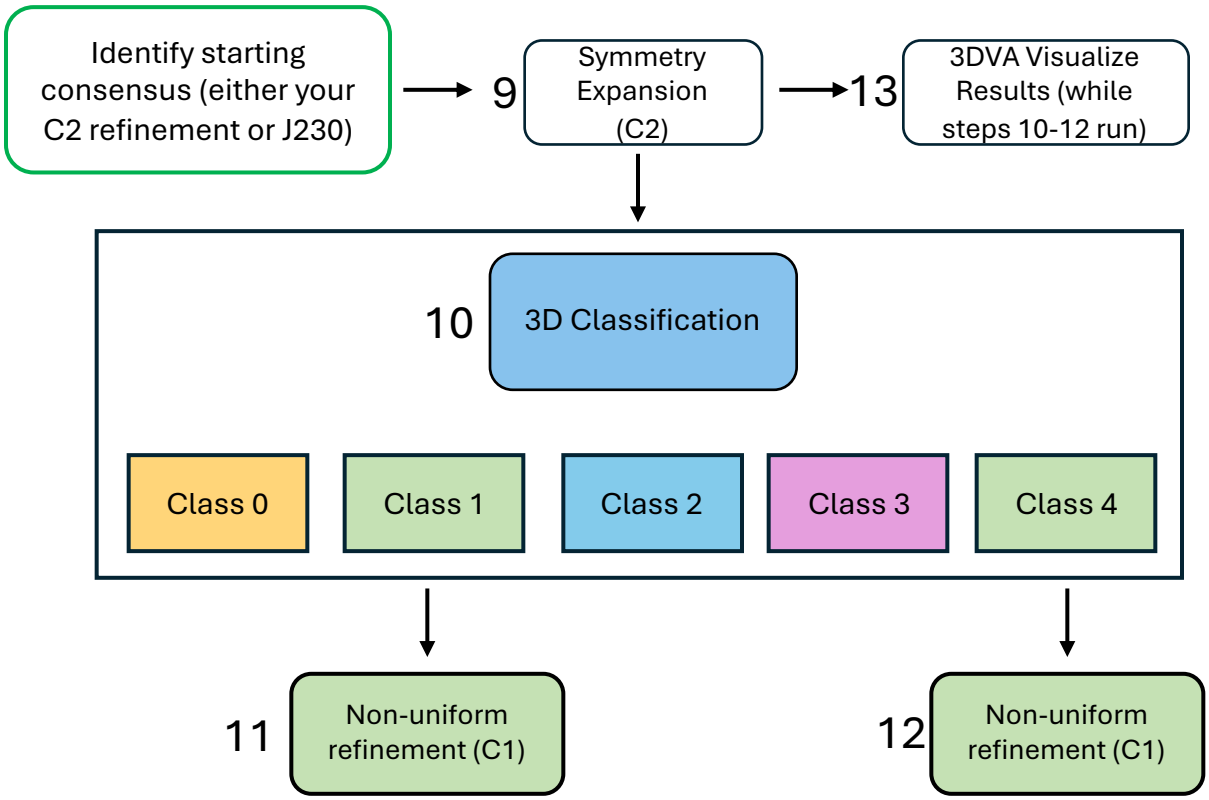
Part A - In Part A, you will generate a “clean” consensus refinement by removing compositional heterogeneity (broken particles) by multi-class *Ab initio* refinement (Step 1). Each of the three *Ab initio* volumes and their corresponding particles will undergo homogeneous refinement without symmetry (Steps 2-4). You will identify a “good” class and repeat homogeneous refinement with C2 symmetry enforced (Step 5). This generates a “clean consensus refinement” that can be used for conformational analysis. You will also assess the handedness of your reconstruction and correct it if need be.

Part 2 A visual guide



Part B – In Part B, you will run symmetry expansion on the C2 symmetrized clean consensus refinement, followed by 3D classification. The results from 3D classification should be inspected in chimera to identify different conformational states, which will then undergo Nonuniform refinement to generate final (hopefully) buildable maps corresponding to different conformations. You can compare these to the results of "pre-baked" 3D variability analysis already linked to your workspaces.

Part 2 B visual guide



14 – Analyze results in ChimeraX

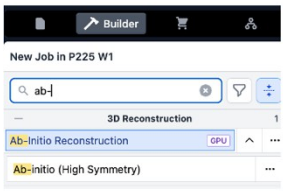
Multiclass *Ab initio* reconstruction

First start the job (steps 1.1a to 1.3 below) before reading the explanation.
The main purpose of Ab initio is to generate the initial volume for refinement (homogeneous refinement), which will be run in the next step. cryoSPARC Ab initio reconstruction can simultaneously generate multiple classes and separate particles into those classes, providing a 3D classification strategy. Let's use multi-class Ab initio reconstruction to generate three initial models, assess heterogeneity, and separate particles into the three classes. For this particle stack, this job should generate one minor class corresponding to broken hemoglobin particles, another minor class corresponding to "bad" hemoglobin particles and a final class corresponding to "good" hemoglobin particles. We will run homogeneous refinement on each class to assess its quality and the good class will be used for downstream steps.
A major benefit of Ab initio reconstruction is the lack of model bias. One drawback is the lack of gold-standard FSC due to the lack of half-maps so there is no quantitative resolution metric for the final map.
Note we will all use the same random seed to ensure we get similar results.

- ▶ Identify starting particles - Extract from Mics J222
    - The desired particles for 3D refinement and classification have already been selected and extracted for you. These came from 2D Classification(Step 14) from Part1 using the whole dataset.
    - Box size 300 -> 150 pixels ((300 pixels x 0.5585 Å/px)/150 pixels = 1.117 Å pixel size)

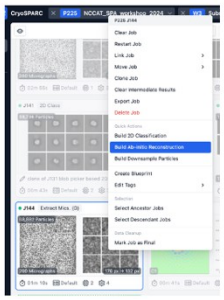
- Step 1 – *Ab Initio* reconstruction (Example J223)
  - 1.1a – Use the job builder to create the job and drag the particles from starting particles (Extract Mics J222) as input OR
  - 1.1b – Right click the starting particles (Extract Mics J204) and click "Build Ab-initio reconstruction" to automatically populate the newly built job with the Extract particle output
  - 1.2 - Parameters:
    - Number of ab initio classes: 3
    - Maximum Resolution: 10
    - Cache particles on SSD: OFF
    - Random seed: 1546673018
    - Note that it is often helpful to alter the initial and/or maximum resolution. Because hemoglobin so small, it helps subsequent refinements to allow the ab-initio to go to slightly higher resolution
  - 1.3 – Queue & wait for the job to finish
    - In the meantime, analyze the handedness of the provided 3D reconstruction and start the Symmetry Expansion and 3D classification jobs (Steps 6 & 7 below)
  - 1.4 – Interpret results
    - Inspect the volumes. How do they differ?
    - Try to predict which class corresponds to the expected particle shape and take it to the next step.
    - Can you identify the monomeric v dimeric class?
  - Note this job took ~20 minutes to run previously

1.1a – Job Builder -> Ab initio reconstruction + J204 particles

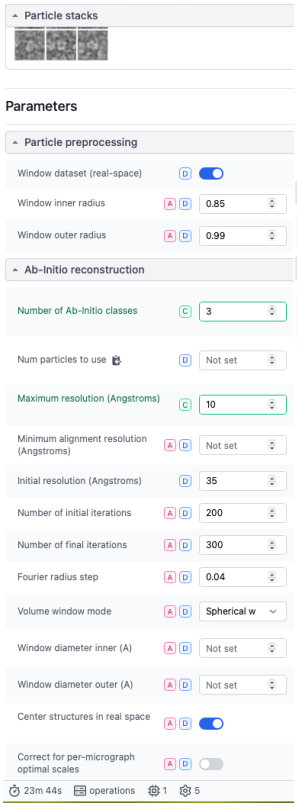


OR

1.1b



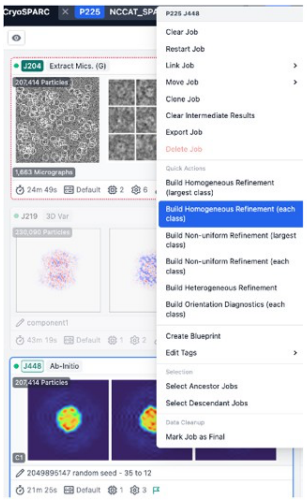
1.2 - Parameters



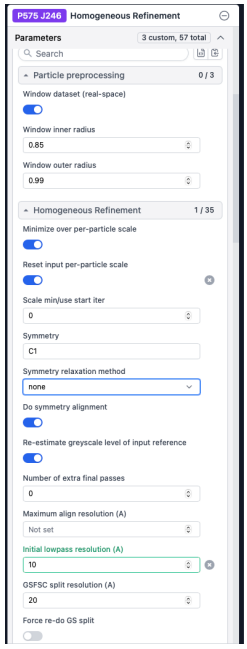
Homogeneous Refinement and Handedness Assessment

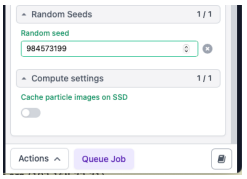
This job will generate a 3D reconstruction. The results can be highly dependent on the starting model and symmetry operator. We will run this job on all three ab initio volumes with their respective particle stacks with C1 symmetry (no symmetry). If we identify a good class that appears symmetric, we will repeat the refinement with that symmetry enforced. The C2 symmetrized refinement will be used for downstream processing and class separation.

2.1



2.2 - Parameters





2.3 Queue

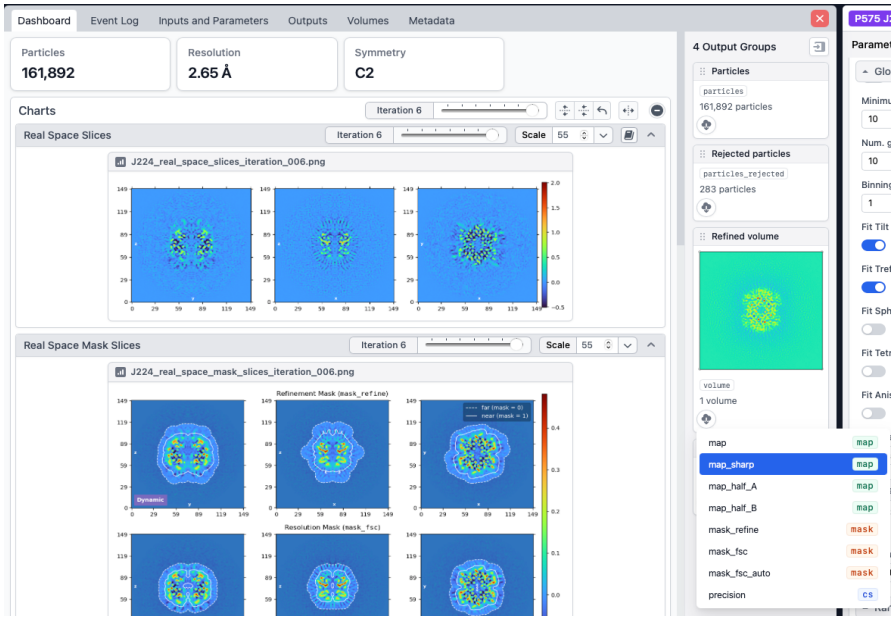
- Step 2 – Homogeneous Refinement of Class 0 (C1) (Example J238)
  - 2.1 – Right click the Ab initio job from Step 16 and "Build Homogeneous Refinement (each class)" – the new job auto populates with the ab initio particles and volume output.
  - 2.2 – Parameters
    - Symmetry: C1
    - Initial lowpass resolution: 10
    - Cache particles on SSD: OFF
    - Keep all others default
  - 2.3 – Queue one job with above parameters, wait for job to complete
- Step 3 – Homogeneous Refinement of Class 1 (C1) (Example J239)
  - Repeat Step 2 with class 1
  - If you clicked "Build homogeneous refinement (each class)" in Step 2, Queue the second job now
- Step 4 – Homogeneous Refinement of Class 2 (C1) (Example J240)
  - Repeat Step 2 with class 2
  - If you clicked "Build homogeneous refinement (each class) in Step 2, Queue the third job now
- Inspect the results – Is there an obvious good class? Use it for Step 5.

- Step 5 – Homogeneous Refinement of good class (C2) (Example J224)
  - 5.1 – Use the job builder to generate a new homogeneous refinement job.
  - 5.2 – Parameters
    - Symmetry: C2 (changed)
    - Initial lowpass resolution: 10
    - Cache particles on SSD: OFF
    - Keep all others default
  - 5.3 – Queue
- Step 6 – Inspect results in ChimeraX and determine the handedness of the reconstruction
  - 6.1 – Open the homogeneous refinement and download the sharpened map from the Refined volume tab. Open that map in ChimeraX
  - 6.2 – Adjust the threshold in the Volume Viewer so that you can clearly see helical features.

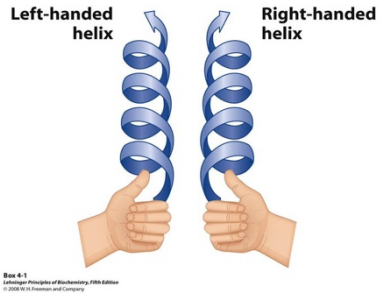
Visually trace the backbone of a helix. Is it right-handed or left-handed? If it's left-handed, this will have to be corrected. It's possible to do so simply in Chimerax with the volume flip #<model number> command, but we'll use CryoSPARC's tool.

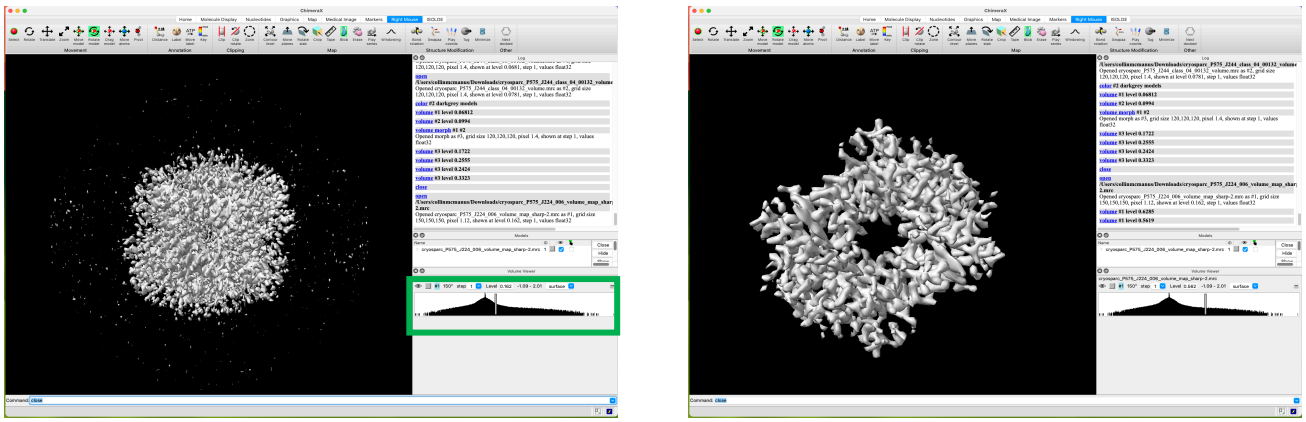
- Because the chirality of an object cannot be determined from its 2D projection, 3D reconstruction gets the handedness of a volume wrong half the time. It's best to check this and correct it if need be early in your processing so that refinements downstream are done with the correct hand.

- Step 7 – Use Volume Tools to flip hand of the C2 reconstruction.
  - 7.1 – Use the job builder to create a Volume Tools job and drag the Refined Volume from your C2 homogeneous refinement to the Input volumes section
  - 7.2 – Parameters
    - Flip hand: true
  - 7.3 – Queue. This job will run quickly, however it only changes the hand of the volume and does not change the alignments of its constituent particles. To ensure that downstream jobs using these particles have the correct hand, we have to run a new refinement with the corrected volume as a reference
- Step 8 – Repeat Homogeneous Refinement with C2 symmetry using correct reference volume. Parameters can be copied from Step 5. For the inputs in this step, use the particles output from the Step 5 refinement and the volume output from Step 6.



6.1





6.2

Part 2B - 3D Classification of Clean Consensus Refinement with Symmetry Expanded Particles

Now that we have a “clean consensus refinement” it is time to assess the conformational heterogeneity. We will run 3D Classification on symmetry expanded particles.

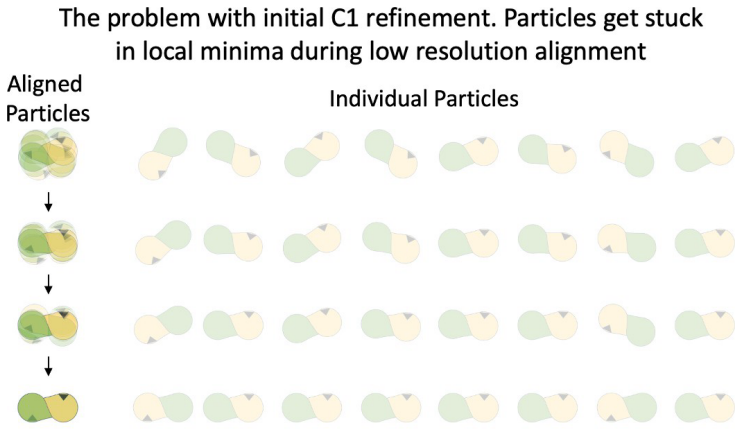
3D Classification

In cryoSPARC, the 3D classification job is similar to Relion’s “classification without alignments and shifts (without GPU)”. In this job type, the program does NOT refine individual particle 3D alignments and shifts, but relies on the 3D alignments and shifts from a previous refinement. This means that the results are highly dependent on the quality of the parent refinement used as input.

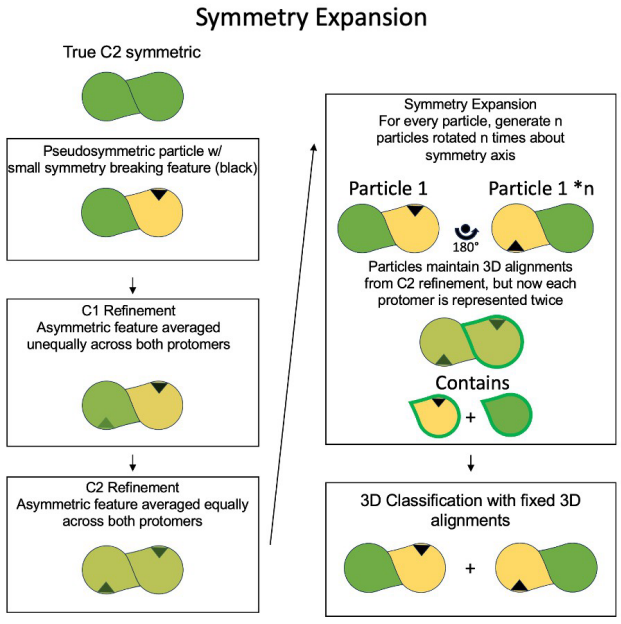
Dealing with symmetry - Symmetry Expansion

The “symmetry expansion” job is run on an “n” symmetrized refinement (C2 for hemoglobin) and outputs n-fold particles (2-fold in this case). Each particle image is duplicated n-fold and rotated along its symmetry axis. This results in alignment of each protomer.

Symmetry expansion allows each asymmetric unit within your particle stack to be treated as a separate particle. When subjected to 3D Classification or heterogeneous refinement, sorting individual asymmetric units enables discovery of asymmetric conformational or compositional differences that are generally “averaged out” in symmetrized refinements. These small asymmetric differences can be referred to as pseudosymmetry and it is difficult for most refinement programs to align particles based on these differences because particles get trapped in local minima during alignment at low resolution.

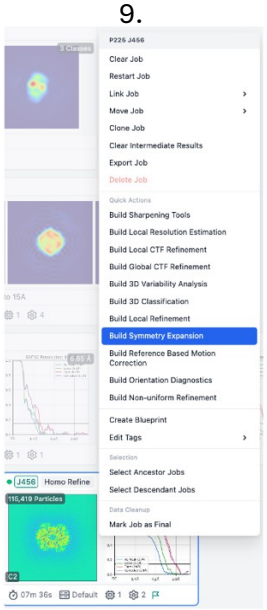


Symmetry expansion and subsequent 3D classification without alignment enables separation of previously aligned particles based on small differences. The resulting particle stacks can be refined with C1 symmetry to give asymmetric maps.

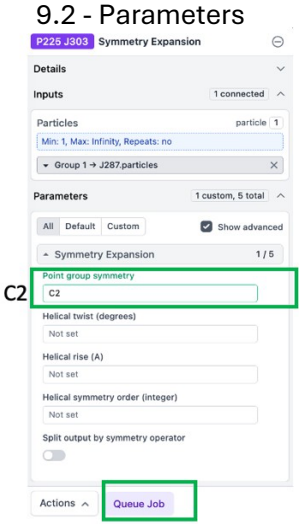


- Step 9 – Symmetry Expansion (Example J231)
  - 9.1 – Right click the Homogeneous Refinement (C2) job J456, choose "Build Symmetry Expansion" and the job auto-populates with the Homogeneous Refinement (C2) particles output.
  - 9.2 – Parameters:
    - Point group symmetry: C2
  - 9.3 – Queue, job runs fast

9.



9.2 - Parameters

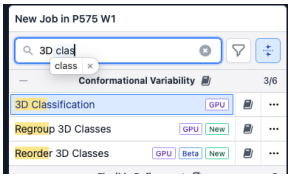


C2

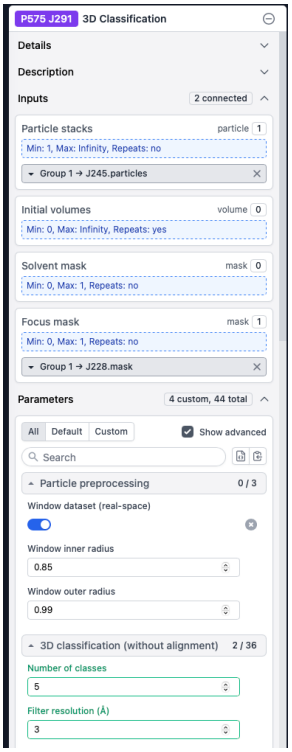
9.3 – Queue

- Step 10 – 3D Classification(Example J232)
  - This is focused 3D classification with a mask on a single hemoglobin dimer (Vol Tools J228)
  - 10.1 – Job Builder -> 3D Classification
  - 10.2 – Drag particles output from Symmetry Expansion to 3D Classification input
  - 10.3 – Drag mask from Vol Tools J228 to the "Focus Mask" input
    - Leave input Volume and Solvent Mask blank
  - 10.4 – Parameters:
    - Number of Classes: 5
    - Target Resolution (Å): 3
    - Random Seed: 1745344143
    - Cache Particles on SSD: OFF
  - 10.5 – Queue & wait
  - 10.6 - Inspect results (KEY DATA INTERPRETATION STAGE)
  - This job may take ~20 minutes to complete
  - While waiting, you can skip ahead to step 13.

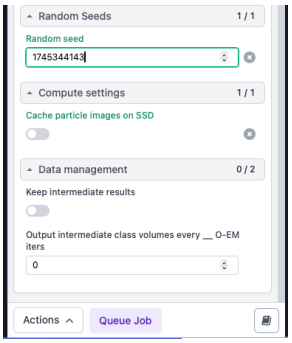
10.1



10.2 input particles and 10.3 focus mask

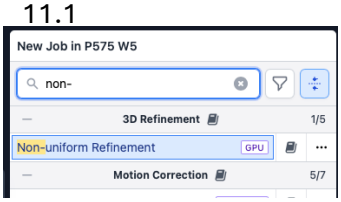


10.4 Parameters
5 classes
3 Å
Seed
SSD off

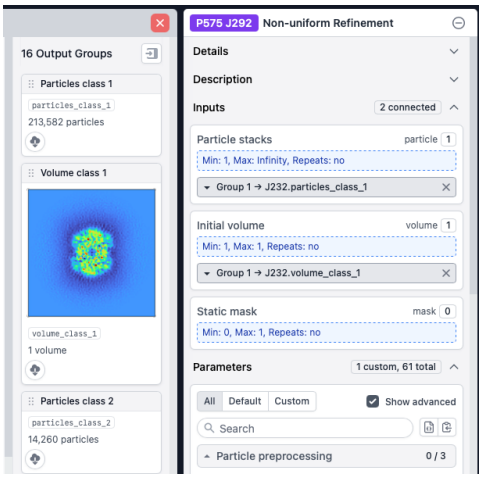


- Step 11 – Non-Uniform Refinement (Examples J472 and J473) of two classes
  - 11.1 – Job builder -> Local refinement
  - 11.2 – Drag the particle class 1 output and its volume from 3D Classification(Step 7) to Local Refinement input
  - 11.3 – Parameters:
    - Initial lowpass resolution: 10
    - Cache particle images on SSD: OFF
  - 11.4 – Queue and wait
- Step 12 – Repeat step 11 for class 4 (Examples J236 and J237). These will take 5-10 minutes to run.

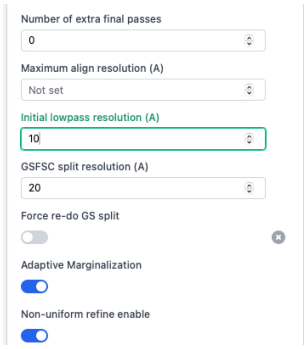
11.1

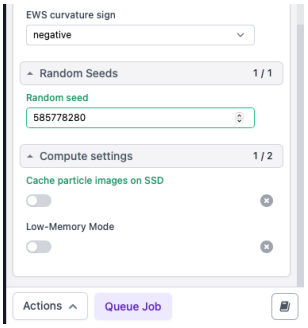


11.2



11.3 Parameters 10 Å initial lowpass



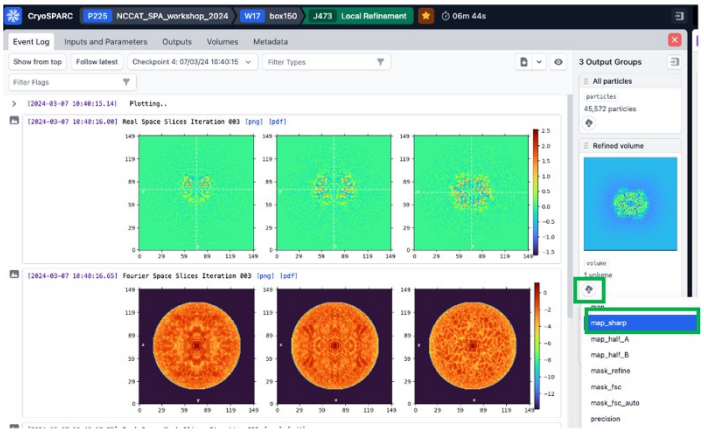


11.4 Queue

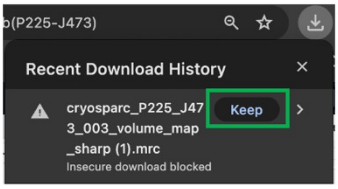
- Step 13 – Visualize 3DVA results in Chimera
  - 3DVA and 3DVA display have already been run on symmetry expanded particles with the refinement mask. This is a good method to identify global conformational change, similar to 3D classification. The output of the "simple" display mode is a volume series that can be viewed as a movie in chimera or chimeraX.
  - 13.1 – Download J235 Volume Series 1 and extract the .zip file. Repeat for the other two series.
  - 13.2 – Open chimera (the rest of the instructions are for Chimera)
  - 13.3- Open the volume series (3DVA display simple mode output results)
    - In the ChimeraX command line: open /Path/*.mrc vseries true
      - You can repeat this command for each of the three components and they will all open similarly and will all be controlled by the same playback controller. Display one volume at a time for clarity, or you can tile them to display them simultaneously using the "tile" command.
  - 13.5 - Change the threshold to reduce noise and visualize features of the hemoglobin
  - 13.6 – Adjust the value next to the snail to slow the playback to 4 or 5 and hit play.
  - 13.7 – Right click the playback menu and select Bounce Playback
  - View the movement

  The older version of Chimera is a bit better at handling volume series and provides more options. If you have it installed, you can try the following.
  - 10.7 – The visualization of volume series is a bit better in Chimera and can handle multiple series. If you have it installed you can follow the following:
    - Open Volume Series Panel
    - Tools, volume data, volume series (volume series panel opens)
    - Click "Open" in volume series panel
    - Choose folder path "cryosparc_P225_J465_component_000"
    - Click first item in the list "J465_component_000_frame_000.mrc"
    - Hold "shift" button and click "J465_component_000_frame_019.mrc"
      - Note that "cmd + a" to select all and then open DOES NOT WORK on chimera v1.17.3 (and likely other versions)
    - Click "open" – the volume series opens
  - - Change parameters in the Volume Series Panel:
    - Play direction - Oscillate
    - CHECK ON - Maximum playback speed (10) steps per second
      - Enter "10" into the field and hit the ENTER button on your keyboard
    - CHECK ON – Normalize threshold levels
    - CHECK ON – cache 30 renderings
    - Data cache size – Enter "1024" and hit ENTER on your keyboard
  - 10.8 - "Play"

- Step 14 – Analyze refinement results in Chimera
  - 14.1 – The NU-refinement job from Step 11 is open
  - 14.2 – Under “3 output groups” on the right, click the cloud download icon and a popup appears with volume download options. Click “map_sharp” to download. On mac with chrome, you may get a warning about an unsafe download and click “keep”
    - you may also want to download the full, unsharpened map (“map”) and/or half-maps to perform various types of post-processing
  - 14.3 – repeat for the other refinement (Step 12)
  - 14.4 - Open the two maps in chimera
  - 14.5 – You are ready to assess the conformational change and decide whether to proceed to model building
  - 14.6 Create a morph from one map reconstruction to the other
    - Open both map in ChimeraX
    - Use command: volume morph #1 #2
    - View the morph using the playback tool. How does it compare to the 3DVA results?



Chrome on mac



- An important next step is to consider what is acceptable to build into. When using symmetry expanded particles, it’s wise to build only into the asymmetric unit which was in the focus mask during classification steps. Otherwise, you must be certain that none of the original particles picked are not represented twice in your reconstruction.
- If more particles end up in a single class than you were in your original reconstruction, there’s probably real symmetry in your original particle stack. At this stage, you can go back to that original stack and try sorting symmetric particles from asymmetric particles for further processing.
- It can be good to further assess residual heterogeneity within any particle stack. Tools like 3D classification, 3D-Variability Analysis, and heterogeneous refinement can aid in further analysis. Processing to obtain the final states is iterative and can take multiple rounds of trial and error to identify optimal parameters for discrete state separation.

Further processing & analysis

Generate a focus mask using a published model

Further details and methods for map generation can be found here https://guide.cryosparc.com/processing-data/tutorials-and-case-studies/mask-selection-and-generation-in-ucsf-chimera

- At rcsb.org, search 6nbd and click on the entry
- Click Download Files on the right, click bottom option "Biological Assembly 1 (PDB – gz)"
- Open it in chimeraX with mrc files and/or 3DVA results
- Fit the model to the maps. You can select the model using the models panel, then navigate to the Right Mouse menu and choose move model or rotate model to roughly align the model with the map. Then execute the "fit #model inMap #map" command in the command line.
- Apply transparency to the maps (transparency #maps 50) and check that the fit is good. Tweak as necessary.
- Execute the "split #model" command to split 6nbd into its constituent chains. Close chain C and chain D. Then, execute the "combine #model.1 #model.2" command to rejoin subunits of a hemoglobin alpha beta dimer. This will create a new model.
- Execute the "molmap #model 8" command to generate a surface around the combined model lowpass filtered to 10 Å.
- Execute the "volume onesMask #surface" command to create a mask
- Execute the "volume resample #mask onGrid #map" command
- Save the final mask and ask a TA to help you transfer it to the cluster
- Run an import volumes job to import the mask
- Run a volume tools job to supply a threshold and add padding to avoid it being too tight.

Analyze result of Non-uniform refinement with symmetry relaxation (P225-J514)

- Download result J281 sharp or full map from W5
- h https://nccat-cryosparc.semc.nysbc.org/api/files/download_result/P575/J281.volume.map_sharp) (alternative map download link)
- Read more about symmetry expansion here
    - https://guide.cryosparc.com/processing-data/all-job-types-in-cryosparc/utilities/job-symmetry-expansion

Generate a figure using Chimera showing the local resolution of a final refinement

- local resolution estimation job was run using final refinement volumes as input, download map_sharp and map_locres from the following jobs in W5
    - J238 and J289 C1 consensus
    - J230 and J282 C2 consensus
    - J281 and J290 Non-uniform symmetry relaxes consensus
- View the map_locres file in chimera using Surface Color tool
    - Open the refinement volume
    - Open the map_locres
    - Tools -> Depiction -> surface color
        - Color surface: volume.mrc
        - By: volume data value
        - Volume file: locres.mrc
        - Options: color palette red-blue
        - Color
        - Optimize the number gradient

Generate a movie of 3DVA analysis simple mode using Chimera

- Follow instructions from the cryoSPARC tutorial
- https://guide.cryosparc.com/processing-data/tutorials-and-case-studies/tutorial-3d-variability-analysis-part-one
- Example commands:
    - windowsize 1200 1200
    - vseries open *.mrc
    - vol #0 step 1
    - movie record
    - vseries play #0 direction oscillate loop false normalize true cacheFrames 30
    - wait 77
    - vseries stop #0
    - movie stop
    - movie encode ~/Desktop/3DVA/result.mpg quality higher