S2C2 cryoEM Image Processing Workshop June 11, 2020
Q&A
Webinar ID 948 6863 3490

Q	Does Cryo-SPARC use the Fourier-slice theorem in an-initio model building or refinement?
A	we may not be able to get to this in-depth today, but please see the supplementary materials to the 2017 cryoSPARC Nature Methods paper, here: http://www.cs.toronto.edu/~fleet/research/Papers/cryoSPARC-Suppl-1.pdf
Q	how compatible is cryoSPARC with Relion? how to export results such as particles etc to relion?
A	We will be going over data management including exporting results for use in other programs, later this afternoon. In general you can export and import fairly easily using available scripts. https://github.com/asarnow/pyem/wiki/Export-from-cryoSPARC-v2
Q	If I want to install CryoSPARC in a laptop for personal use, what is the requirement for the laptop?
A	https://guide.cryosparc.com/setup-configuration-and-management/hardware-and-system-requirements
Q	Comment: I did the ab-initio reconstruction using all 50 2D classes and then refinement. It yielded a 3.99 Å GSFSC resolution vs. the 2.93 Å resolution from the Select 2D classes proceedure.
A	How many particles were used for the ab-initio/refinement job? Because each particle serves as 14x as many views due to symmetry being enforced in refinement, if you chose too few particles it could impact the final reconstruction resolution.
Q	Hi All, a question: I've tried to consider 2 or 8 classes before the Ab initio reconstruc. The final results are 2.95A with 2 classes and 2.93 with 8 classes. Of course the numebr of particles/time of data processing is much higher with 8. In general, how many classes are better needed for Ab initio? 2 could be enough?
A	I'm trying to make sure I understand this correctly. For the T20S dataset, where we are very certain there is only one major "class", it would not be helpful to do a multi-class ab-initio reconstruction to reconstruct multiple 3D structures from the 2D data since then we will just be splitting particles across multiple classes when they actually belong to one class, thereby limiting the final resolution we can achieve if refining any given class. So if you have a "homogeneous" sample, you should use 1 class for Ab-Initio, and only use multiple classes (and their outputs for further processing) where you expect or are exploring heterogeneity.
Q	Sometimes jobs in cryosparc fail midway through the run (usually due to having too many jobs running/memory problems), is there a way to continue the job instead of having to completely restart? For example a hetero refinement job
A	Unfortunately, there isn't a way to continue processing a job that has failed mid-way. You will need to clear the job then queue it again, which will start the job from scratch.
Q	can one continue any "stuck" jobs e.g., if long 3d refinement jobs terminates prematurely (to save computation time)? how?
A	In some cases, a job could become stuck due to system memory limits being reached. In that case, you can try to free up some RAM (by killing other jobs, for example) and will see the job in question to resume processing. It's not possible to resume a refinement that was in progress and has stalled or failed.
Q	Following CTF estimation, is there a way to select dose-weighted micrographs for particle picking?
A	Hi Melesse, by default, cryoSPARC particle picking jobs use the dose-weighted micrographs when both non-dose weighted and dose-weighted micrographs are available. You can always override this by manually selecting one micrograph over another using the low-level results feature of the job builder: https://cryosparc.com/docs/tutorials/job-builder#individual-results
Q	What is the "corrected" plot on the fsc plots?
A	https://discuss.cryosparc.com/t/tight-corrected-and-loose-gsfsc-curves/201/5?u=stephan
Q	There is a downsample particles option in the job builder, but how to scale back the partilces if I didn't import the micrographs. (Cannot use extract from micrographs)
A	Hi, you can only fourier crop (bin) particles. You cannot "upscale" particles without unfortunately re-extracting the particles at a larger box size.
Q	What is MW of HA protein?
A	MW = molecular weight. You can take a look at the source paper: https://www.nature.com/articles/nmeth.4347
Q	Hi I haven't registered to do the practical, where can I find the data if I want to process at the same time?
A	Hi, See: https://www.ebi.ac.uk/pdbe/emdb/empiar/entry/10097/ to download the data (via aspera or directly from the browser). Note that the size is about 1.8TB so it may not finish downloading in time to follow the presentation
Q	What exactly is meant by super-resolution data? Could you explain that a little more?
A	https://www.gatan.com/improving-dqe-counting-and-super-resolution
Q	do we need to reextract the particles if we have downsample to get the best resolution later after 2D classification

	A	If you are noticing that the box size of your particles is limiting your reconstruction, (i.e. the FSC is hitting the Nyquist rate), then yes, you can consider using a larger particle size. If you had to run the downsample particles job to get a smaller particle size, you shouldn't need to re-extract- those particles should be available from the extract from particles job that initially extracted them. If you did however extract them at a smaller box size (i.e. Fourier cropped), then you will indeed need to re-extract at a larger box size.
Q		What is degree of the tilt or tilts?
	A	This data is a subset of the EMPIAR 10097 dataset (https://www.ebi.ac.uk/pdbe/emdb/empiar/entry/10097/) which was collected at 40 degree tilt
Q		in the ctf estimation job, do we use the motion correction as the input, or the import job as input?
	A	You would use the motion correction's outputs, as the CTF Estimation job takes in aligned micrographs as inputs.
Q		Is the cryoSPARC program free?
	A	CryoSPARC is free for academic use and paid for commercial use. https://cryosparc.com/download
Q		what's the different color in the plat of GSFSC?
	A	https://discuss.cryosparc.com/t/tight-corrected-and-loose-gsfsc-curves/201/5?u=stephan
Q		Can you explain how to read GSFSC Resolution graph again? thanks
	A	We may not be able to repeat it right now but it will come up again later today, and in the meantime, please see this post where Ali discusses the plot on our discussion forum: https://discuss.cryosparc.com/t/tight-corrected-and-loose-gsfsc-curves/201/5?u=stephan
Q		How is the dataset split into half? random? or odd vs even number? How fine is the range of the Fourier shell? every 1/A or so?
	A	there is a detailed session tomorrow morning on CryoEM map resolution assessment, validation, visualization and segmentation, hopefully the speakers can answer more detailed questions about validation at that time
Q		For the AAA+ 2D classification job it says to input a circular window diameter of 150 Å. There is no window diameter setting. There is a window inner radius and window outer radius (both are unitless); there is also Circular mask diameter (Å) and circular mask diameter outer (Å). Which setting are we to set at 150Å?
	A	Hi, Ali means the "Circular mask diameter (A)" in the 2D Classification parameters.
Q		There are two sets of FES curves reported by cryosparc. One says before FSC-mask auto-tightening and the other one doesn't. What are the differeces between them, and which is the one we should report as our result?
	A	Here's a detailed overview: https://discuss.cryosparc.com/t/tight-corrected-and-loose-gsfsc-curves/201/5?u=sdawood
Q		In GSFSC, what does GS stand for?
	A	Gold Standard
Q		input mask, or provide mask? how to you draw the mask
	A	You do not need to input a mask - it is dynamically generated.
Q		Regarding masks, for T20S in C1 sym there are artifacts due to gold particles that I couldn't clear up with 2D classification. Would you suggest generating a mask and providing it for refinement in this case?
	A	spoke extensively about symmetry yesterday- you can re-watch the recording when it goes live for his explanation. Though you should be using D7 symmetry for T20s- you'll get a much better result since one image will contribute to 14 different viewing angles.
Q		For the HA project. we did not pick or inspect any particles yet. so which particles can i use as input for the "manually curate exposure" job?
	A	At the moment we are only curating the micrographs, so you dont need to connect the (optional) particles input to the Manually Curate Exposures job
Q		For the HA project P8. my patch motion job J23 is still running.It was started at ~9:10. Is that ok?
	A	The Ptach Motion job should take around 14 minutes to complete. To confirm, you should be processing 100 movies from the HA Trimer dataset
Q		My motion correction is still running, now at ~90/100. It'll end soon but I'm worried It might be to slow for next parts of the day.
	A	The Patch Motion job should take around 14 minutes. You can choose to run it on 2 GPUs
Q		How different is hetero-refinement from NU-refinement? Are there specific cases to use hetero/NU refinement?
	A	Heterogeneous refinement is more akin to 3D classification (multiple structures/classes). Non-uniform refinement is useful for automatically regularizing 3D density maps during refinement to account for spatial variability. Non-uniform refinement will be discussed in more detail later on today. Preprint: https://www.biorxiv.org/content/10.1101/2019.12.15.877092v1
Q		can you monitor CTF / ice thickness for on the fly data collection ?
	A	This is one of the use cases that we accounted for when developing cryoSPARC Live. It's possible to set thresholds similar to in the exposure curation job, but they persist over time as new exposures are processed. More details here: https://guide.cryosparc.com/about-cryosparc-live
Q		Is curate database for bad images?

A	Manually Curate Exposures (or Exposure Curation) is a a tool for removing “bad” images according to the various statistics that you choose. More info here: https://cryosparc.com/docs/tutorials/exposure-curation
Q	When I clicked on the selecting it minimized the Curate Exposures window and doesn’t allow me to select done. It just says waiting.
A	The minimize button is next to the close button. Re-open the dialog by tapping the spacebar once the job is selected, then press the minimize button at the top right
Q	What if you know already the symmetry of the molecule and don't have at all top views and having only differet side views? Does this helpto get meaningful reconstruction?
A	In general, if you have only a few viewing directions, you will likely not get a good reconstruction; it may exhibit streaking or be apparently missing density corresponding to missing views. Single particles adopting a preferred orientation in the ice (leading to a lack in diversity of viewing directions) is a common problem in single particle cryo-EM and sample optimization, collecting the data at tilt, etc are all strategies for addressing this. Perhaps others in the audience will have more insights to share!
Q	Is mask diameter same as the particle diameter?
A	If you mean “Circular mask diameter (A)” in the 2D Classification job, no. This parameter is “The outer diameter (in Angstroms) of the window. If None, outer diameter is 20 percent larger than inner diameter. The window mask transitions smoothly between inner and outer diameters.”
Q	If we manually provide a mask for the refinement. During the refinement whether the dynamic mask will be ignored automatically or still contribute to the refinement?
A	If you provide a mask during Refinement, this mask is what will be used during Refinement. You can choose whether to use a dynamic mask or static mask by changing the “Mask (dynamic, static, null)” parameter. Changing it to “static” while providing a mask will ensure no changes to the mask are made.
Q	can we use exposure curation after particle extraction
A	Yes, you can add the ‘particles’ output of the extraction job into the exposure curation job
Q	When I tap the spacebar I get a white window. Should I kill the job?
A	For the exposure curation, the job dialog should look like this (note the ‘Interactive’ tab) when you press spacebar after selecting the job card: https://cryosparc.com/static/images/curate_overview.png
Q	Sorry if its already been stated, but could you describe how to intrepret the blue patch CTF curve shown?
A	in case we don’t get to this again today, please see https://cryosparc.com/docs/tutorials/patch-motion-ctf
Q	Is it possible to import particles from eman2?
A	cryoSPARC’s Import Particles job supports these file formats: https://cryosparc.com/docs/reference/jobs#import-particle-stack
Q	Could you please suggest how to deal with small closed packed particles, less than 100 kDa in size? Majority 2D classification jobs results in significant over fitting and blurring in just a few iterrations.
A	hopefully this AAA+ unfoldase example with masking is relevant since its a ~83kDa and also highly concentrated. Please feel free to raise any other questions you may have about this though!
Q	what is the particle is off center, what would the circular mask do?
A	If the particle is off center, then the circular mask may not be applied properly. So we need to recenter the classes by flipping on that switch in the job builder
Q	what is circular mask diameter outer?
A	Here, we used 150A
Q	Hi, what do I use for template input into the select 2D job?
A	The ‘2D Class Averages’ from the 2D Classification job go into the ‘Template’ input of the Select 2D job
Q	Why the stacked particles in J47 disappear in J48 and J50?
A	In J47, we did not apply a circular mask. In J48, we did so and in J50 I believe we also recentered the particles. All params available in the Job Builder for 2D Classification
Q	In addition to giving exposures, if I give 'particles' to a Manual Picker job, what is it doing with that information?
A	Supplying particles to the Manual Picker job will tell the job to plot the particles on the images. You can then “curate” the particles by right-clicking on them to deselect them. You can also add particle picks by using the left-click button. This is useful for example if you want to feed a set of curated picks identified by the blob picker to a deep picker like Topaz for training (as Topaz requires really good particle locations for training).
Q	Is it possible to export and save the tree of data processing workflow?
A	In the current version of cryoSPARC it’s not possible to export or save the image of the tree

Q		If I understand correctly, if you are doing a reconstruction with negative stained data at low resolution, set the Maximum resolution to 15 -20 Å and it will save time and energy.
	A	Yes this would be a good strategy to save time
Q		can cryosparc run with cuda 8.0?
	A	You can run cryoSPARC with CUDA 8.0, but you will notice some jobs will fail. It is recommended to run cryoSPARC with the latest version of CUDA, and the minimum requirement is CUDA 9.2.
Q		Hi, why my ab-inition reconstruction job NOT taking the selected 2D classes; it requires particles not classes; what is wrong I am doing?
	A	The Ab-initio requires particles, it doesn't use templates. You can use the particles output group from the Select 2D (or 2D Classification) job as an input
Q		Do we need CUDA 9.2 for all the versions of cryosparc or just for the latest version?
	A	You can assume any version of cryoSPARC after v2.9.0 requires CUDA 9.2+
Q		Can you use the 2D classes to create figures for group meetings and such? or is that done in other programs by exporting… meaning, is there a job to create a summery of the data processing?
	A	There's no job to specifically create a report of processing, but you can use the job log's text and figures to create a summary of processing steps. You can download PNG, PDF and text outputs (for applicable plots) to use elsewhere. Specifically for 2D classes, there's a post on our forum that outlines how to independently generate high resolution images of each 2D class: https://discuss.cryosparc.com/t/high-resolution-image-of-2d-classification/3395/2
Q		How do we know our PXX number for the apoferritin data?
	A	This is your project number - it is in the top left project selector and also in the URL
Q		Hi, I do not see the "Do symmetry alignment" parameter in the homogeneous refinement (new). also, no Initital lowpass resolution" parameter.
	A	can you please ensure the 'Advanced Mode' toggle is on?
Q		AAA+: there are 3 classes. which one to be used for homogeneous refinement?
	A	Please use the *inactive* state for one homogeneous refinement, and the *active" state for the other (steps 6 and 7)
Q		what is an appropriate contour for veiwing the map in chimera? How does one decide this?
	A	First, ensure the step size is set to 1 as it will ensure all the voxels are being shown. The threshold should be adjusted to a point where there are not too many artifacts and the map isn't over sharpened. For published datasets, authors will typically note the contour level they used on EMDB.
Q		can the resolution of the flexible portion of the map be improved in cryosparc?
	A	This depends on the amount and quality of data for the flexible part that is contained within the dataset. We'll cover local refinement and non-uniform refinement later on in the session
Q		Do I need change the number of 3D classes, or 3 is always good?
	A	We are covering this now. If you think there is some heterogeneity in your dataset, its best to experiment with different numbers of classes in Ab-Initio Reconstruction until you can determine how many classes or conformations you have
Q		Do all of the ab-initio classes use the same particles? Or do they split these particles into the different ab-initio classes?
	A	The particles will be assigned to one of the N classes, by the end of the Ab-Initio job
Q		What is the input for the ab-initio?
	A	Particles
Q		For steps 6 and 7 of unfoldase tutorial (homogeneous refinement), did you use the same "particles" for the two different abinito models?
	A	No you need to use the particle set that corresponds to each volume separately. So for the inactive state, which may be labelled class 0, 1, or 2, take the particles of the same class. For the active state, same thing - the particles that correspond to the active state volume
Q		If we did homogeneous refinement and then heterogenous refienment, after that, should we do homegenous refinement against each class seperatd by heterogenous refinemnt?
	A	If you have multiple different classes, all of which you are trying to refine to high resolution, then yes exactly
Q		May I ask you where I can find all the changable parameters for homogeneous Ref? many of the parameters that are suggested to change, I do not have
	A	In order to view advanced parameters as shown in the screencast, enable the toggle that is under the pruple 'building' tag in the job builder sidebar (where you can see the other parameters)
Q		Dynamic masking creats a general overall mask. In case of micelles or nanodisc, how to mask that out to focus on the trans-membrane part?
	A	We can use non-uniform refinement (will be discussed shortly on this GPCR dataset) - no need to mask out the micelle
Q		In which way, the order of classes of Ab-inito will be different ? class0 vs class1, is there any significant difference? for exampe the number of particles in the group or the most convinced group?

A	Since Ab-Initio Reconstruction is a stochastic algorithm, it is initialized with a different random seed each time you run it (even on the same set of particles). You will see an output for particles_all_classes, which is all the particles (not separated at all). Then you will see particles_class_0 through to particles_class_n where n is the number of Ab-Initio classes you asked for. Once the job is finished, you can look to the number of particles in each class, the volume output (inspect visually e.g. in Chimera) and the orientation distribution plot to understand whether each class is a true class (i.e. different states or conformations) or if you have asked for "too many" classes and need to reduce the number as you continue to explore heterogeneity
Q	in the most realystic case when your 2d class is not very clean - can you repeat the stages of how to cleanup the data
A	If after running one 2D classification (and Select 2D to view and select good classes), there is still a lot of junk, you can run additional 2D classifications on the particles you have remaining to try and cleanup. We would also recommend a multi-class ab-initio reconstruction (with or without a "clean" particle stack) as shown in the AAA+ unfoldase example where you can remove entire junk classes (and all the particles associated with those classes) by choosing which ab-initio classes to keep after you've finished running that job. Finally, there is also heterogeneous refinement, which simultaneously classifies particles across classes and refines them to higher resolution
Q	for the GPCR tutorial, for the refinment input, do I use "particles_all_classes", or "particles_class_0"?
A	Since they contain the same number of particles, you can use either. particles_all_classes is fine
Q	is it a regular practice to deposit all the raw cryoEM movies to a database?
A	It is highly encouraged and very useful, but not always feasible due to size of raw data. EMPIAR is the main public repository. https://www.ebi.ac.uk/pdbe/emdb/empiar/ I believe this will also be covered in detail tomorrow.
Q	is the cryoSPARC mainly a web-based program, or is it usually downloaded to the user's computer?
A	The primary method of interacting with cryoSPARC is through the web interface - you just need a modern web browser to access the UI, no additional downloads
Q	for flexible molecules can we always use Non Uniform refinements?
A	Yes we would recommend this (but you can also run homogeneous refinement as well to check)
Q	How about the privacy of the data/projects/file names I use in the cryoSPARC? Are they private or public?
A	A cryoSPARC instance can have more than one user- if you are an "Admin" user, you will be able to see all projects/data. If you are not, you will only be able to see the projects/data you own (i.e. the projects/data you've created). Only system administrators should be "Admins". CryoSPARC is usually hosted privately, so it is not available on the world-wide web.
Q	What do you mean install on a node? I thought this is a mainly web-based program. Is this installation for the data storage?
A	The cryoSPARC interface/web application is just one component of the 'master' node/installation of cryoSPARC. The compute code runs on a workstation with sufficient CPU, memory and GPU requirements. More details: https://guide.cryosparc.com/setup-configuration-and-management/hardware-and-system-requirements
Q	Do you have sdandalone version of cryoSPARC to install to local Unix workstation?
A	Yes, please follow the instructions for a single workstation/standalone installation on our guide: https://guide.cryosparc.com/setup-configuration-and-management/hardware-and-system-requirements#single-workstation
Q	One 16 Gb GPU vs two of 8 Gb GPU, which is better for performance?
A	One 16GB GPU would usually be better for performance because NVIDIA's higher-memory GPU options are usually compute-focused, meaning they have higher memory bandwidth and usually more cores. Having more memory will not only allow you to reconstruct larger box sizes, but also allow you to run more than one GPU process at the same time.
Q	General question- if I downsample my images at certain job, how do i upscale it the orginal during refinement.
A	If you mean you downsampled your job in Refinement using the "Refinement box size (voxels)" parameter, the original particles will remain the same- it's only if you used Fourier cropping during Extract from Micrographs or the actual Downsample Particles job you would need to re-extract your particles.
Q	are you going to show sometime how to create a mask?
A	Unfortunately we won't have time to do this today in the workshop but you can use Volume Tools in cryoSPARC to create a mask from another volume, or you can use programs like UCSF Chimera to hand-create a mask
Q	How can this algorithm apply to map with symmetry but with heterogeneity.
A	Hi Yanyan, do you mind clarifying the question please? do you mean, can this algorithm apply to a map with symmetry but without heterogeneity?
Q	If the volume was reconstructed with D2 symmetry (for example), can we do the 3D variability anagysis? thanks,
A	Yes, you should be able to provide the consensus refinement (even one done with non-C1 symmetry) and all the corresponding particles, as inputs to a 3D Variability job

Q		Is it possible to connect the different motion components together in sequence?
	A	See this tutorial for how to visualize the components in a sequence, in Chimera: https://cryosparc.com/docs/tutorials/3d-variability-analysis#visualization (Part 1 of the 3D Var Tutorial) Part 2 of the 3D Var Tutorial is here: https://cryosparc.com/docs/tutorials/3d-variability-analysis-2
Q		I do not have volume in my options in chimera…
	A	Please refer to the tutorial here: https://cryosparc.com/docs/tutorials/3d-variability-analysis#visualization

LIVE ANSWERED

I am referring to the proteins for later session like GPCR, AAA+ unfoldase etc
When you ab-initio classify into several classes taking the initial particle stack, what similarity do you choose between the classes?

For CTF estimation, did you use dose weighted images, or non-dose weighted images? Does it have the option to choose the non-weighted images for CTF estimation?

In the thon ring image. what is the difference between the left and the right half of the image?
Can I sort images or automatically removed images with too strong ice ring?
I think I missed it but what is the units of 'ice thickness values
' shown in the table…?
If we have both unliganded and liganded proteins in our sample, is heterogeneous refinement better than homogeneous refinement?
Can we provide same model as multiple inputs for heterogenous refinement?
If motion more than 5 angstrom, those images will be slectively removed or not processed in later steps?
How many movies do we have in the HA dataset?
number of movies to plot is always 10?
what is the significance of the fourier shell diameter (the yellow circle)
percentages of data at each resolution shell?
sometimes you see a small dip in the corrected FSC at lower resolution, say, like at 4.2 A in the map you show, what cause that?

what is y axis on the GSFSC graph
how do you validate the correctness of masks
At what resolution we should use dynamic mask, and in what. conditions we should supply a mask as input for refinement
is the huge bump saw in FSC normal (I mean the bump presented in the middle of FSC curve)? And how would you interpret the bump?
Are we trying to get the tightest mask at the last itiration, since then we get less "noise" from the solvent, which is usually not contributing to the resolution estimate?

Is there any parameters that suggest final tightening of mask is correct?
can the user still manually adjust mask tightening after the final refinement instead of using the auto mask-tightening?
Is the auto-tightening done for all data sets?
Do we trust images with -ve avg. intensity?
how did you meansure ice thickness in cryosparc
in this case, what is the pitfall of using the "bad" images in the next steps? would you suffer a loss in resolution in the final model from including the "thick ice" images that were excluded from the HA trimer dataset?
If the sample has preferred orientation especially in thin ice. Would removing the thicker movies make it worse?
If we have two data set, one collected at one plane, the other at tilted plane, what's your strategy to do ctf and refinement?
is there a resolution from where you can start aplying 3D variability. lets say you can do it at 3-4 A but can you do it in some model where you have 6-7 A resolution?

Can we discuss a bit what exactly an ab-initio class is? Is it multiple probably sructures from the same data, or is it splitting the data into multiple conformations based on their electron density?
What if you are doing a low resolution resconstruction? What should you set your Maximum resolution to?
what if you have rather featureless particles, protien in a micelle, would you change the max resolution or max alignment resolution?
maximux resolution for 2D classification: means highest number or loweres number for resolution?
what is the rule of thumb to decide the radius of circular mask?
How narrow should the circular mask be in 2D classification?

In 2d after recentering does it need to be reextracted for next rounds?
What is the meaning of the number and ECA in the class averages?
what is ECA and how to use it in the interactive session of 2D selection job
Wow, in the 2D class averages with the small inner radius, how did you know that the fuzz was protein and not that the particle picking was done poorly to choose particles off-center?

Are there any ways (numbers) that can help to select 2D class apart from visual inspection
Re: 2D classification paramaeters, what about batch size? How do we assign that rationally? (or did I miss it?)
When switching off the Force Max over poses/shifts, should much more number of classes be taken?
Would it make any sense to increase the circular mask on 2D classes, more than 150 Å, so that a larger region of the second molecule of unfoldase may be included, or will it introduce too much noise background?
What I don't understand about using a crowded grid of particles is that even though the picked particles you showed at the start contained multiple protein particles, then the 2D class averages still contain information from neighboring protein particles, but are just masked for classification sake. Don't we need to re-pick from the dataset with a smaller box size?

I'm still struggling with when to use ab initio (multiple classes) vs. heterorefinement for 3d classification. When multiple ab initio look similar (maybe slight conformation changes), should we put those classes and volumes together for a hetero refinement? Or should all ab initio classes be put together for hetero refinement?

Is the inactive C6? there might be a confusion in the instructions email that the active is C6 and inactive C1
Is there a way to check for symmetry computationally, before imposing symmetry?
so for proteins in micelles would you change the class similarity to a higher value?
How to find about Symmetry for a new structure? Is there any parameter that suggests about symmetry?
It sounds like Cryo SPARC does what is called "3D classification" in RELION while Cryo-SPARC does ab-initio reconstruction. Is that the case?
how to check the quality of the ab intio classes apart from viewing direction distribution
How big does the particles need to be in order to have an improvement with the CTF per particle refinement?
does it makes sense to do CTF refinement for a sructure with a that refines to 4-5 A?
for proteins having a disordered region what CTF refinement do you suggest?
for homogeneous Vs heterogeneous particles do we nee to do different CTF refinements?
At what point exactly do you use CTF refinement?
what kind of resolution structure are worth to do CFT refinement? How about one structure at the resolution of 3.5A?
if I am using standalone CTF refinement, do they need to be done in certain order, i.e. first do local and then Global CTF refinement and then use those particles for refinement.

If my data was collected with beam image shift, would running CTF refinement during homogeneous refinement be a bad idea (before splitting by exposure groups). Should I split them into groups and then run stand alone CTF refinements and recombine again for a Homogeneous refinement?
Does the per particle defocus optimization and CTF refinement work for imported particles? For instance, if the motion correction was carried out in Relion.
Can non-uniform refinement be applied to a heterogeneous refinement?
Can we use NU with outputs from homogeneous refinement?
Is nonuniform refinement more computationally expensive than homogeneous refinement, as a rule?
What's the difference between NU-refienment strategy and applying a real-space mask to exclude unnecessary background?
What parameters for non-uniform refinement would be different (to be optimized) for membrane proteins in detergent v.s. in nanodiscs?
Would a small membrane protein (59 KDa) be completely inside a micelle and impossible to process?
Can you do CTF and defocus refinement in non-uniform refinement?
If there's a flexible region that you think could break your molecule's symmetry, should you be using C1 in non-uniform refinement?
Is it possible to installed cryoSPARC on HPC node?
Is that the way to export particle or map for other software like relion?
Is it relatively straight forward to add a work node to an existing installation? Let's say we add a microscope and a GPU node to do on the fly processing for the microscope.

Is it often that obne bumps into 'out of memory' errors if you only have a system with two GPUs each 8Gb memory? If you only have this setup, when encountering such errors, what can one do to work around?
Is there a minimum number of particles for 3D Variability to be reliable?

How to create a mask for this 3D Var job?
What makes a ‘volume’ a ‘mask’?
What is the best approach to utilize 3D variability in a sample that has both conformational variation and compositional variation?
Could you please explain why 8 Å was chosen in this case as a Filter Resolution parameter?
Are the Vk modifications always “positive” or can Vk indicate removing some of the model?
What numer for z do you use, pdb coordinhates?
Is the number of variability components determined without any input from the user?
Would the continuous 3D variability algorithm produce similar results as a multibody refinement in other softwares like RELION?
Is it better not to run heterogeneous refinement before 3DVA?
How does this compare with cryoDRGN?
can I see artifacts in 3Dvariability? how can i identify them?
I there is should i impose symetry in 3d variability? how?
At what resolution does it make more sense to perform 3DVAR?
What are approaches for performing “validation” of 3D variability analysis outputs? Something similar/analogous to FSC?
Is 3D variability job very memory consuming?
Is there a way to use 3D variablity on a specific region of the protein instead of the entire structure?
how to record that paly movie in Chimera?
can we do 3DVA on small protein components (~40kDa)?
How long does it take to run a 3D variability analysis on a standard-sized dataset?
Can you please explain again what is the right half image of the Thon rings, when looking at the CTF correction? The left half are the thon rings calculated by an FT of the EM data. Is the right half a calculation too? is the right half a diffraction pattern?

NO ANSWERS RECORDED IN TRANSCRIPT

Can a small angle x-ray scattering map be used as a template for particle picking?
what are three different views on the refinement result screen?
I’d like to know if there are special tricks for smaller particles, like monomeric, flexible, elongated proteins ~15 nm in one direction.
How to interpret the particle orientation plot? For TS20 set, most spots are in middle. Does it mean it has limited views?
When should I discard an image based on "too much" local motion, and can you do it through curation job?
Did you tried or was thinking to see experimently, the correlation about ice-thinkness estimation from CTF against some experimental ice thickness measured from tomography?

Re: CTF estimation: Is there a big practical difference between using the CTFFIND4 job versus the patch CTF estimation job?
In case of such background noise due to water , can it be corrected by changing parameters in CTF correction ?
how would one differentiate in the T20S a C7 or D7 symmetry? what happens if a wrong symmetry is applied in the homogeneous refienement? should it be expect an improvement in resolution as well?
I want to recap on yesterday question, about patch motion correction. How is it happened in details, pathes are alingned independantly across movie frames? What happened to particles which are at the borders of patches?
Hi, in the homogeneous refinement, there is a plot called guinier plot. What does this plot mean? What information can we get from it? Thanks
this is a general question but from a user perspective, can you speak to the advantages of using cryoSPARC over other SP programs (such as Relion)? are the algorithms used in cryoSPARC more efficient/faster? does it boil down to personal preference and using a program that one feels comfortable with?
Not sure if it mentioned yesterday, what is minimum dose per frame that is reasonable for a successful alignment?
Do we need to have a priori knolewdge of protein structure in order to use a specific mask?
Is mask diameter same as the particle diameter?
for resolution estimation- should a mask be used to cover the compact part of protein if there are flexible regions present in the strcuture with lower EM density coverage?

How do you define the tight mask for FSC calculation, if one doesn’t have an atomic/interpreted model but have only a map. What about “loose”/flexible domains ?

How does local resolution depend on the reference point in the structure?
So, the cryoSPARC creates small boxes and calculate s the FSC for each to estimate the local resolution?
and if you structure is being overfitted what should you do?
Re: orientation distribution: Will you please provide a few examples of plots of elevation v azimuth that would show preferred/too-few number views?
What if you know already the symmetry of the molecule and don't have at all top views and having only differet side views? Does this helpto get meaningful reconstruction?

For local resolution estimation, what does it mean when resolution is lower around the edge of your molecule/mask? What are the best and worst types of 3D volumes (resolution/data amount or quality) to use local resolution to analyze?
how do you calculate the local resolution
Does the color in this shell mean anything?
This fourier shells are always real numbers? Or they can be complex?
When comparing the FSC, are they divided? or substracted? or somthing else mathematically?
What does negative density mean in the real space slices?
How would this masking change in the case of a sample with multiple different conformations? Would it just form a “looser” mask?
Is there holes inside the mask? why?
Could we adapt pseudo sym. when we run refinement ?
Does the cryosparc output contain something like phase-randomized FSC curves?
Do you have Mask A and Mask B for validation?
how to deal if there is any discrepancies between these masks?
is there a parameter where we can prove that the resultant model is truly the oligomer that we see or is a forced symmetry bias? especially for high resolution data where the particles can be very hazy to decipher that we are seeing a monomer or fragmented oligomers?
so how would those parameters change if you had a membrane protein in a micelle or nanodisc?
Does the mask at 6 ang mean excluding tata beyond 6 ang?
How do you decide the threshold for the mask?
What are the best mask parameters to change to search densities of a molecules that are connected/adjacent to a refined input volume?
I mean measure....
Can you please explain the right side of the CTF again? is that scattering of the electrons measured by baclk scattering in a different detector?
is this diffraction pattern of crystalline water is spill over from diffraction plane to image plane?
Can we use chemical chaperones (low molecular weight organic molecules) to stabilize the protein if it is unstable under invitro conditions and if we use them what will be the contribution of these molecules in the background noise of cryo micrographs?
what is the origin of cross parallel lines on the micrograh top right corner previous slide
regarding particle density, there are examples of apoferritin data collections (Herzik et. al. Nat Comm 2019 i think) where the particles are tightly packed. apoferritin is very homogenous so perhaps the crowding is not a huge problem but could you comment further on the issues with a crowded particle set?
What happens if I miss a poor viewing class from 2D? Can we get the 3D construction from that missing class?
how can one configure masking parameter to focus on internal structures? ie the genome of virus
What about analyzing data for proteins that form homo-oligomers, especially those with an unknown number of particles? Any specific settings to when performing 2D classification?

can you explain step size and mini batch size?
For Ab initio : what is a good amount of 2D classes ?
Can you please explain again what is the right half image of the Thon rings, when looking at the CTF correction? The left half are the thon rings calculated by an FT of the EM data. Is the right half a calculation too? is the right half a diffraction pattern?
Hello, I have a dataset, where one part of the particles has an additional peptide/small density showing up. I would like to classify the dataset in classes with and without this peptide. How can this be done in cryoSPARC? Thanks.
Does the maximum resolution only fit ab-initio classes at 12A, so any data that are at higher resolution aren’t considered in the reconstruction? Is there a benefit to changing this value to higher resolution to see other conformations? Or is it just better to change the class similarity score?
can you please explain a bction it about Viewing direction Distribution and its significance?

Is there a way to apply a mask radius during ab initio reconstruction to try to get particles to converge on part of the protein closer to the center (similar to the mask shown for 2D classificiation)?
Would you then put all particles in and the respective classes from ab initio?
Does uncertainty in pixel size (ex. 1.11 vs. 1.18 A) affect refinwment results? by how much?
What if, the still heterogeneous refinement cannot separate heterogeneity ?
Can CTF Refinement be applied to tomography the same as regular single particle analysis?
can you go over the parameters to review to see if we are leaving data on the table that could be potentially useful for improving resolution refinement? ie. how do we know we have the best possible map / structure for the data provided?
So exposure group will correct defects coming from the beam tilting?
Sorry, a bit out of order. Regarding Peferred orientation, some papers have reported that removing particles from over present views helps in improving quality of the final structure. Is there any way of doing that in CS2?
Can cryosparc model the residues of the protein in the final EM map?
Why the non-uniform refinement is called "cross-Validation"
In Cryosparc, which would be the best approach to solve the structure of a flexible domain that is moving in a different way in each subunit on the surface of a symmetric protein cage?
So, how would you deal with 12 flexible identical tails that are moving differently and indepedently in each particle?
nonuniform after homogenous refinement cause adecrease in quality of map, why is that
What if a molecule in nanodisc having very elongated morphology and only side view can be seen in images, how to address this? Can we use tilt series collection?
With a flexible hinge on the elongated axis
Does the non-uniform refiment will be better suited for a full length antibody with the issues of the hinge? or the mask? (FL vs FAB)
what causes “spike kike” features to appear in refinement?
is there any way to improve the alignment of a small membrane protein completely inside the micelle in 2D classification?
how do you define the 3D variability area?
will that induced objective artifacts?
molecular motion can be time dependent, will that be also considered for the linear model?
Is there a static reference volume in the whole electron potential map when comparing images, also is there a limit to size/area whose dynamics is being assessed given variable s/n ratio ?
Can the particles corresponding to certain datapoint in this variabilty analysis be extracted and used for construction alone?
A question about heterogeneous refinement that I just run through. The dafault parameter for Refinement box size (Voxels) is 128. Does that mean all the volume will be cropped to this size to accelerate GPU? Will this number affect the final GSFSC value reported from cryosparc?
Can I provide the mask from the consensus refinement?
why did you see different motions for ATPase? What's the difference concerning data processing?
Is there any relationship between 3D variability and mutiple occupacy?
Question from practical group. The GPCR result from NU-refinement seems has a prepfered orientation problem which looks worse than the regular refinement (not about the resolution but the shape)?