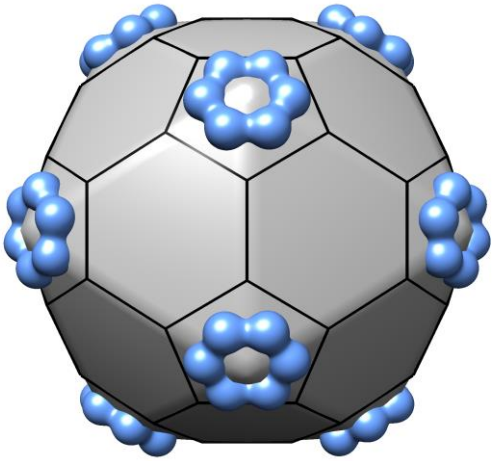


Helsinki Institute of Life Science HiLIFE
Laboratory of Structural Biology

A Scipion LocalRec plugin tutorial for handling symmetry mismatches

Please send your comments to Vahid via vahid.abrishami@helsinki.fi

Introduction

In this tutorial, we provide a step-by-step guide for handling symmetry mismatches in single-particle analysis using a new plugin called LocalRec [1] in Scipion [2], an image processing framework to obtain 3D models of macromolecular complexes using Electron Microscopy (EM). We cover the required steps to define and extract sub-particles and follow conventional SPA workflows using a range of state-of-the-art methods in order to resolve the symmetry-mismatched subunit. After this tutorial, you should be able to apply this plugin to your data.

For this purpose, we use data of human adenovirus D26 kindly provided by Vijay S. Reddy from the Scripps Research Institute. The particles were extracted and a refinement using the dominant symmetry (I1) has already been performed to calculate their alignment parameters. In this demo, we will use these particles to obtain 3D reconstructions of the adenovirus trimeric fibre and its four hexons in two separate Scipion projects.

1. Getting started 4
1.1. Add LocalRec plugin to Scipion 4
1.2. Download data 4
1.3. Workflow summary 4
2. Localized reconstruction of adenovirus fibre 6
2.1 Prepare data 7
2.1.1. Import icosahedral map 7
2.1.2. Import hexon-only icosahedral map 8
2.1.3. Define 3D mask for fibre subunit 8
2.2. Prepare particles 10
2.2.1. Import particles 10
2.2.2. Subtract projections 11
2.3. Prepare sub-particles 13
2.3.1. Define sub-particles 13
2.3.2. Filter fibre particles 15
2.3.3. Extract fibre particles 17
2.4. Reconstruct sub-particles 18
2.4.1. Fibre reconstruction 18
2.4.2. 3D classification of sub-particles 20
2.4.3. 3D refinement of sub-particles 26
2.5. Analysis 27
2.5.1. Fibre postprocessing 27
3. Localized reconstruction of hexons 1–4 of adenovirus 28
3.1. Prepare particles 28
3.1.1. Import particles 28
3.2. Prepare particles 28
3.2.1. Define sub-particles 28
3.2.2. Extract the sub-particles 29
3.3. Reconstruct sub-particles 30
3.3.1. Reconstruct hexons1-4 30
3.3.2. Refinement of hexons1-4 30
3.4. Analysis 32
3.4.1. Create hexons1-4 postprocessing masks 32
3.4.2. Postprocessing of hexons1-4 32
3.4.3. Combining reconstructions into a composite map 32

1. Getting started

1.1. Add LocalRec plugin to Scipion

In this tutorial, we assume you have already Scipion 2.0 installed (if not please check the Scipion 2.0 installation instructions). You can install LocalRec plugin by either use the following command

scipion installp -p scipion-em-localrec

or Install the plugin through the plugin manager GUI by launching Scipion and following Configuration -> Plugins.

After installation, you should be able to see the LocalRec protocols in Scipion.

1.2. Download data

The full data set (including particle images and Relion star file) is available at EMPIAR (ID 10455) and you can download it using one of the methods proposed by EMPIAR. In addition, you need to download the following tar file (including the JSON files for two workflows and a density map required for hexons' mask):

wget https://a3s.fi/swift/v1/AUTH_21dc6f465ca845b5ad55143b98e34e06/2000537-tutorial-data/localrec21_tutorial.tgz

tar -xf localrec21_tutorial.tgz

1.3. Workflow summary

In this tutorial, we will use LocaRec plugin along with other tools for localized reconstruction in Scipion to determine the structures of adenovirus trimeric fibre and its four hexons as two separate projects. Following is a summary of different steps in Scipion to perform localized reconstruction:

1. *Import particles*: Import particles with alignment and CTF information into the Scipion (can be the output of a refinement protocol).
2. *Define sub-particles*: Determine the coordinates, orientations, and defocus values of the sub-particles.

3. *Filter sub-particles (optional)***:** Excludes unwanted sub-particles based on alignment information.
4. *Extract sub-particles***:** Extracts sub-particle images from each particle image.
5. *3D reconstruction*: Determine an initial structure of the subunit, given an input set of sub-particles with angular assignments.
6. 3D classification/refinement: Determine the high-resolution structure of the subunit.

A typical workflow utilizing these protocols and their integration with standard SPA protocols in the Scipion framework is shown in Figure 1.

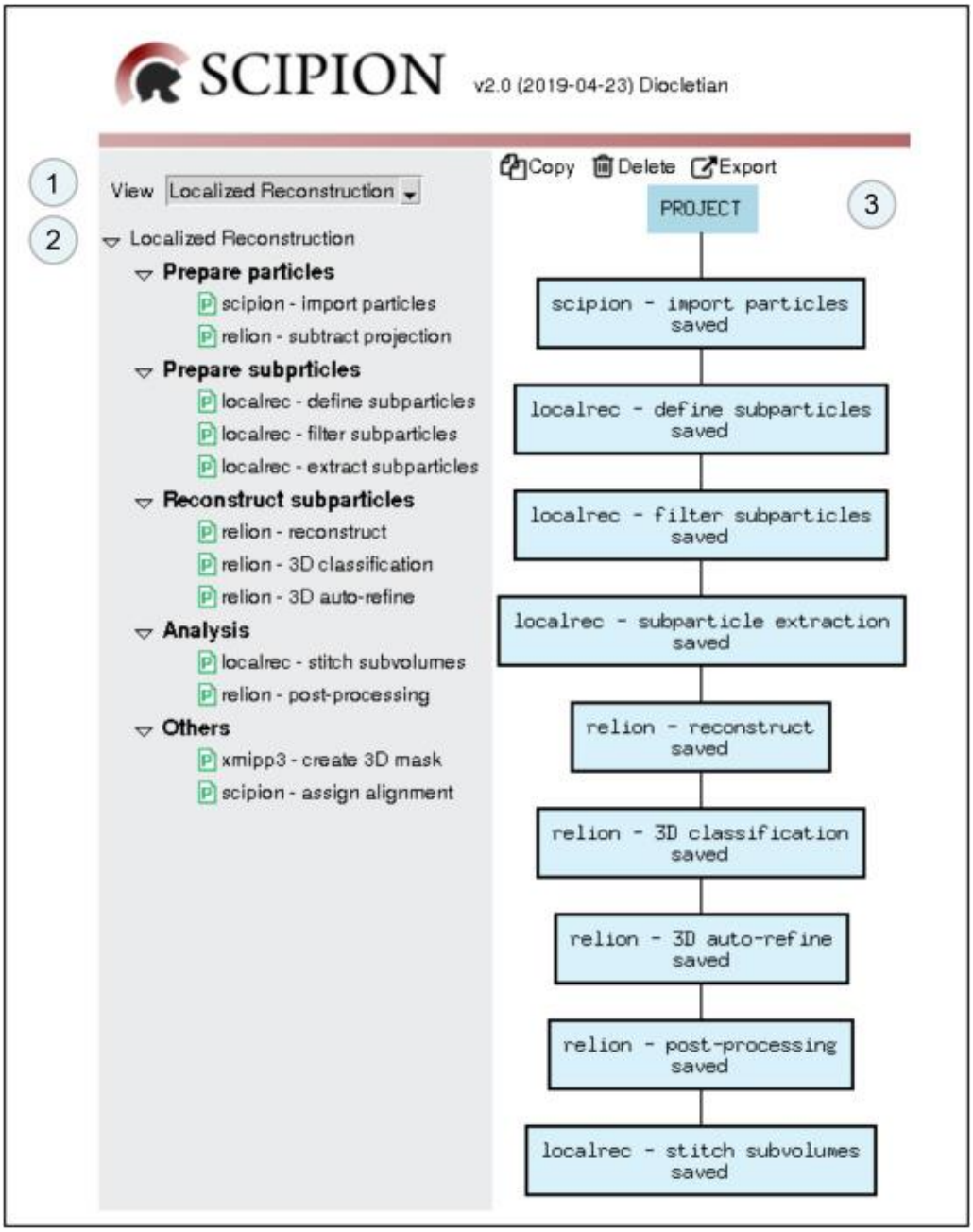


Figure 1: Localized reconstruction workflow in Scipion from Abrishami et al. (2020)

2. Localized reconstruction of adenovirus fibre

In this section, we will use localized reconstruction to determine the structure of HAdV-D26 trimeric fibre. First, you need to create a new Scipion project and name it "HAdV-TFibre-Tutorial" (Figure 2). Then after the project window popping up, you should open an existing workflow template (a json file) by going to Project -> Import workflow and selecting "HAdV-TFibre-Workflow.json" downloaded before. This way, we have a basic template that will make it easier to follow the processing pipeline.

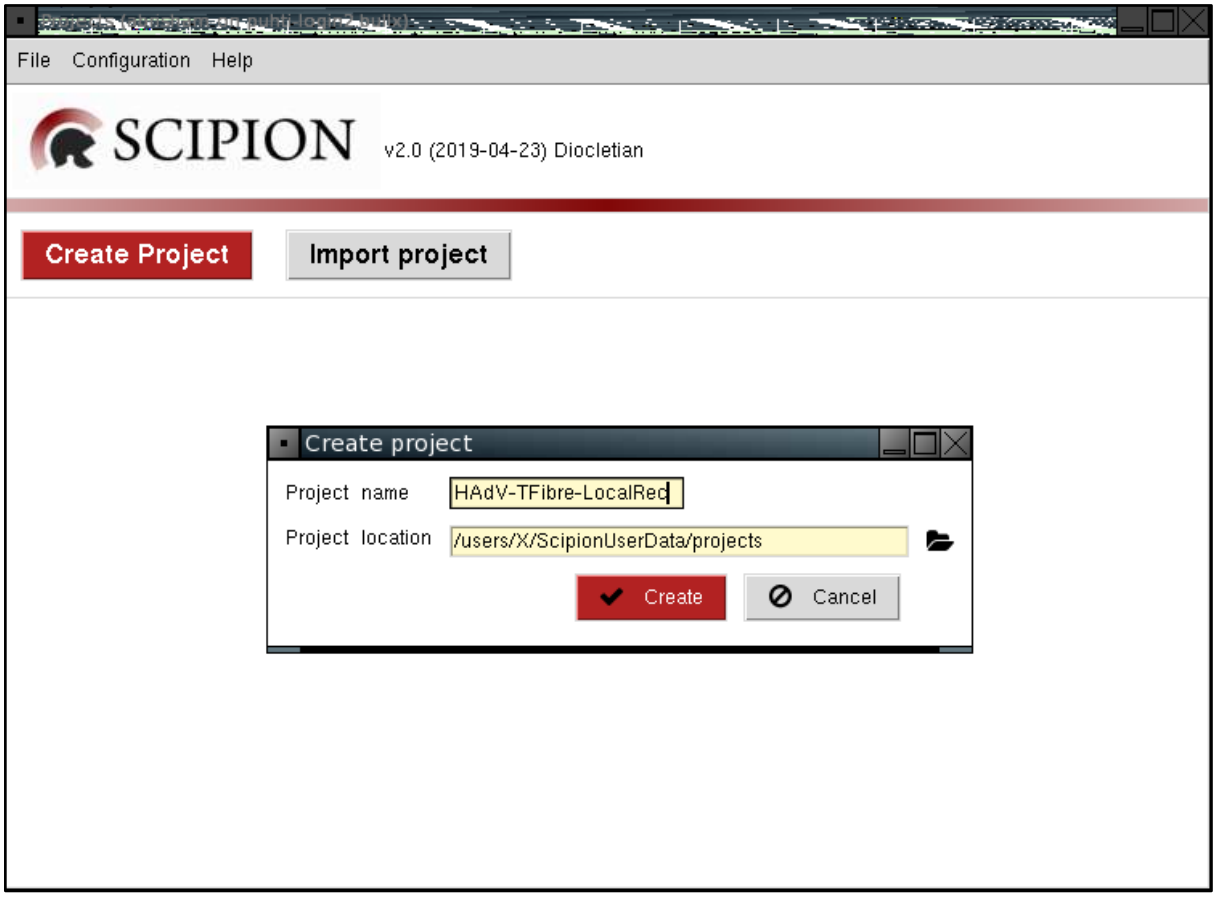


Figure 2: Create a new project in Scipion for the localized reconstruction of HAdV trimeric fibre.

2.1 Prepare data

After creating the project, we first import some data required for continuing with the rest of the workflow.

2.1.1. Import icosahedral map

The icosahedral map is required later for subtraction step (2.2.2). To import the map, double click on 1e. Import icos reconstruction box (or use the left menu Imports > import volumes) and set the parameters as below (Figure 3):

1. *Files directory*: Choose the file for icosahedral map (downloadable from EMDB with accession code 11008).
2. *Pixel size ("sampling rate")(Å/px)*: Set it to 1.31.

Click on the Execute button to run the protocol.

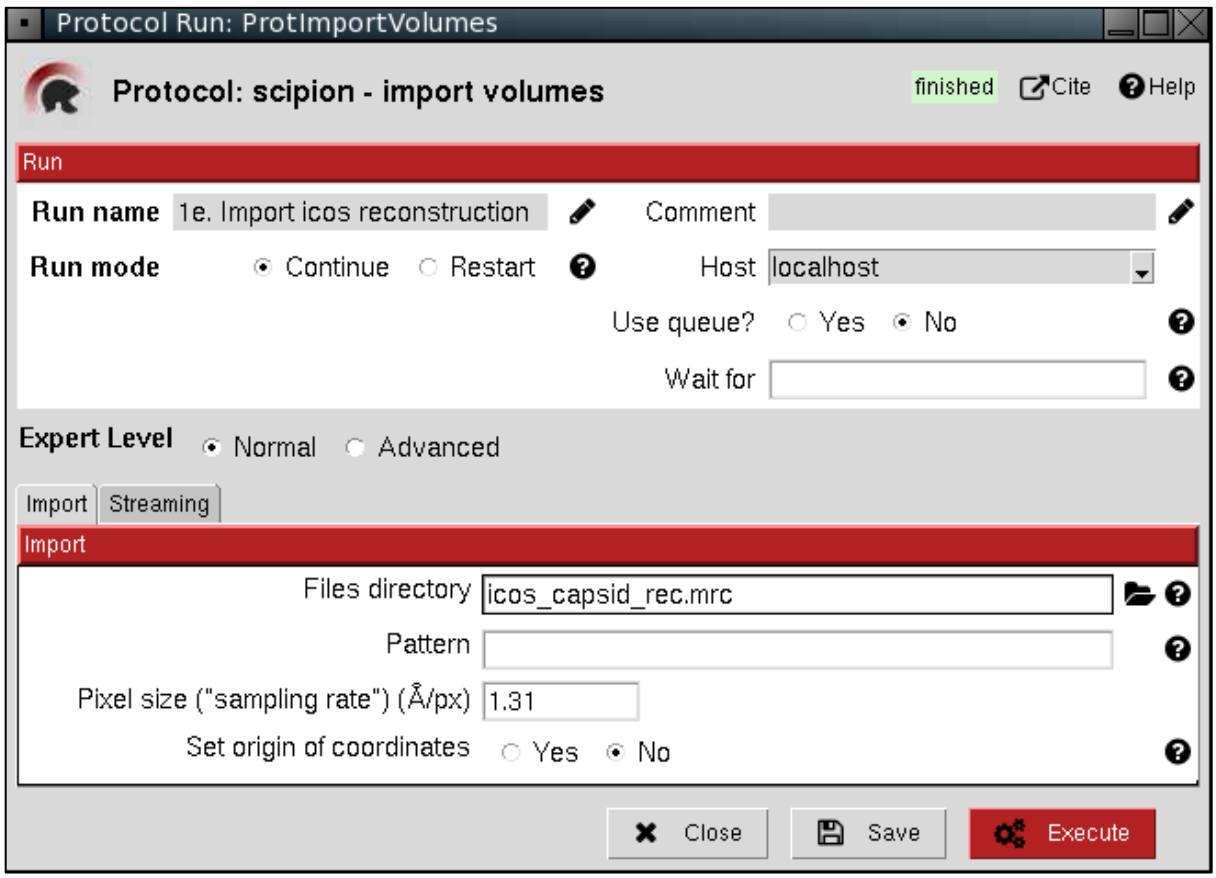


Figure 3: *Import volumes* protocol with set parameters.

2.1.2. Import hexon-only icosahedral map

The icosahedral map is required later for subtraction step (3.1.2). To import the map, double click on 1b. Import hexon only densitybox (or use the left menu Imports > import volumes) and set the parameters as below:

1. *Files directory*: Choose *adenovirus_capsid_nopenton.mrc* from the *localrec21_tutorial* folder.
2. *Pixel size ("sampling rate")(Å/px)*: Set it to 1.31.

2.1.3. Define 3D mask for fibre subunit

We generate a cylinder mask for the fibre subunit (shaft + knob) which will be used later for focused classification and postprocessing steps. Double click on the 1d. Mask defining single fibre box (or use the left menu 3D > Preprocess > xmipp3 – create 3d mask) and set the parameters as below (Figure 4):

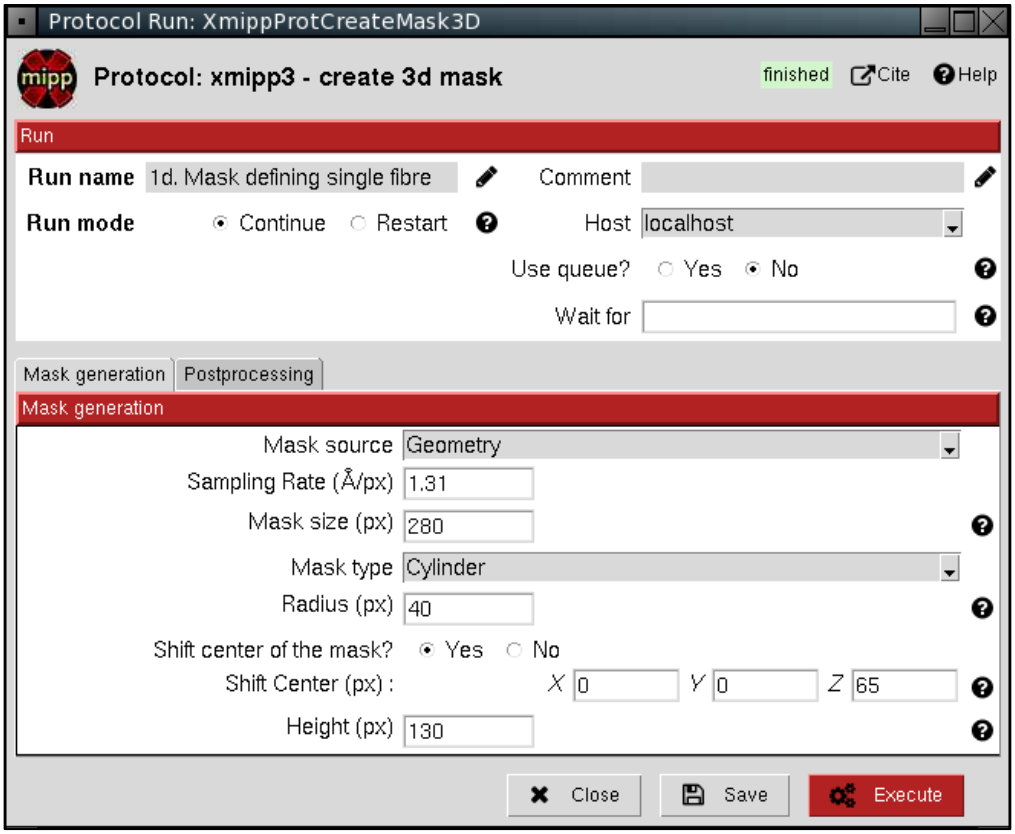
Mask generation tab:

1. *Mask source*: Select if you want to create a mask from an existing volume or geometry. Choose Geometry from the list since we want to generate a cylindrical mask.
2. *Sampling Rate (Å/px)*: Set it to 1.31.
3. *Mask size (px)*: This is the dimension of the output mask. Set this value to 280.
4. *Mask type*: Choose Cylinder from the list.
5. *Radius (px)*: This is the radius of the cylinder in pixels. Set it to 40.
6. *Shift center of the mask?*: Since we put the centre of the subunit at the bottom of the fibre, we should shift the mask in the *Z* direction to fit it to fibre density.
7. *Shift Center (px):* Here we specify shifts in different directions. Set *X* and *Y* to 0, and *Z* to 65.
8. *Height (px):* This is the height of the cylinder in pixels. Set it to 130.

Postprocessing tab:

1. *Smooth borders*: Set it to yes to smooth the borders.
2. *Gaussian sigma (px)*: Number of pixels to be added as soft edge. Set it to 3.

(A)



(B)

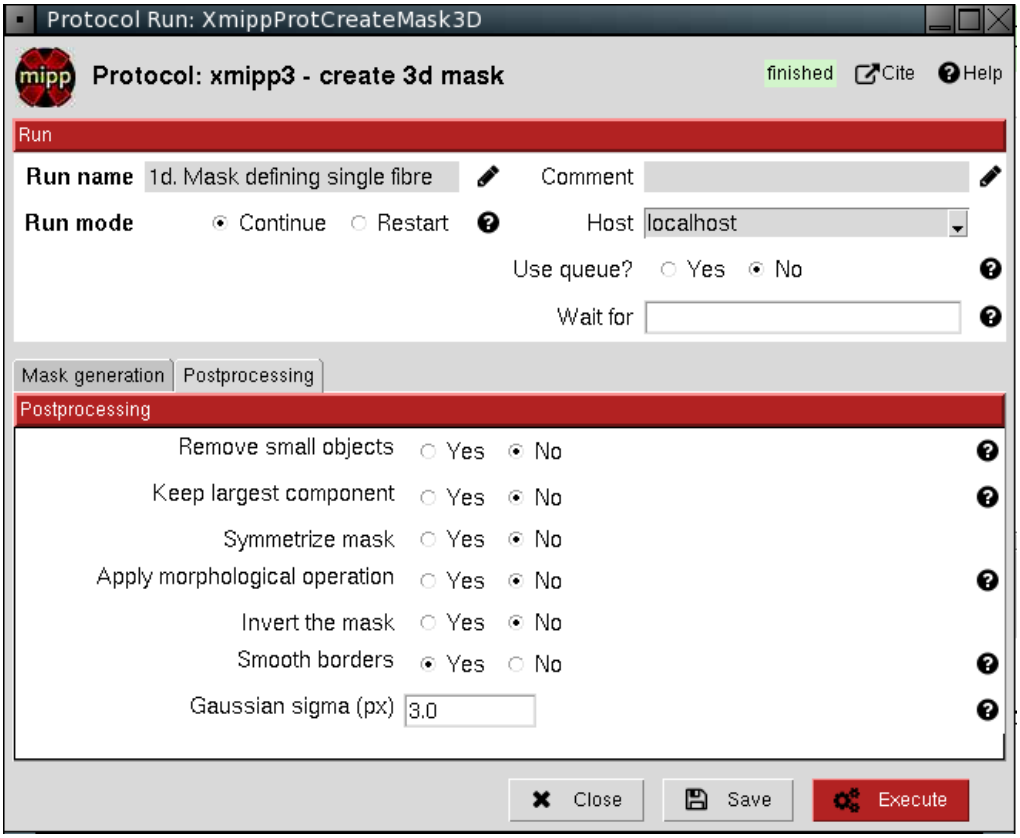


Figure 4: Filter sub-particles GUI. (A) Input tab. (B) Sub-particles tab.

2.2. Prepare particles

2.2.1. Import particles

The first step is to import the particles into your Scipion project. To do this, double-click the 1. Import refined virus particles (or use the left menu Imports > import particles) and set the protocol parameters as below (Figure 5):

1. *Import from*: Choose relion from the drop-down menu.
2. *Star file*: Select the "input_particles.star" by browsing to where you downloaded the data.
3. *Use the wizard button to import acquisition*: Click this button to set *Microscope voltage*, *Spherical aberration*, and *Amplitude contrast* parameters automatically.
4. *Magnification rate*: Set it to 50000.
5. *Pixel size ("sampling rate")(Å/px)*: Set it to 1.31.

In Figure 3, you can see this protocol with set parameters. Then, click on the Execute button to run the protocol. After execution, in Summary tab at the bottom, you can see the number of imported particles and other useful information. Additionally, by clicking on the *Analyze Results* you can visualize particle images.

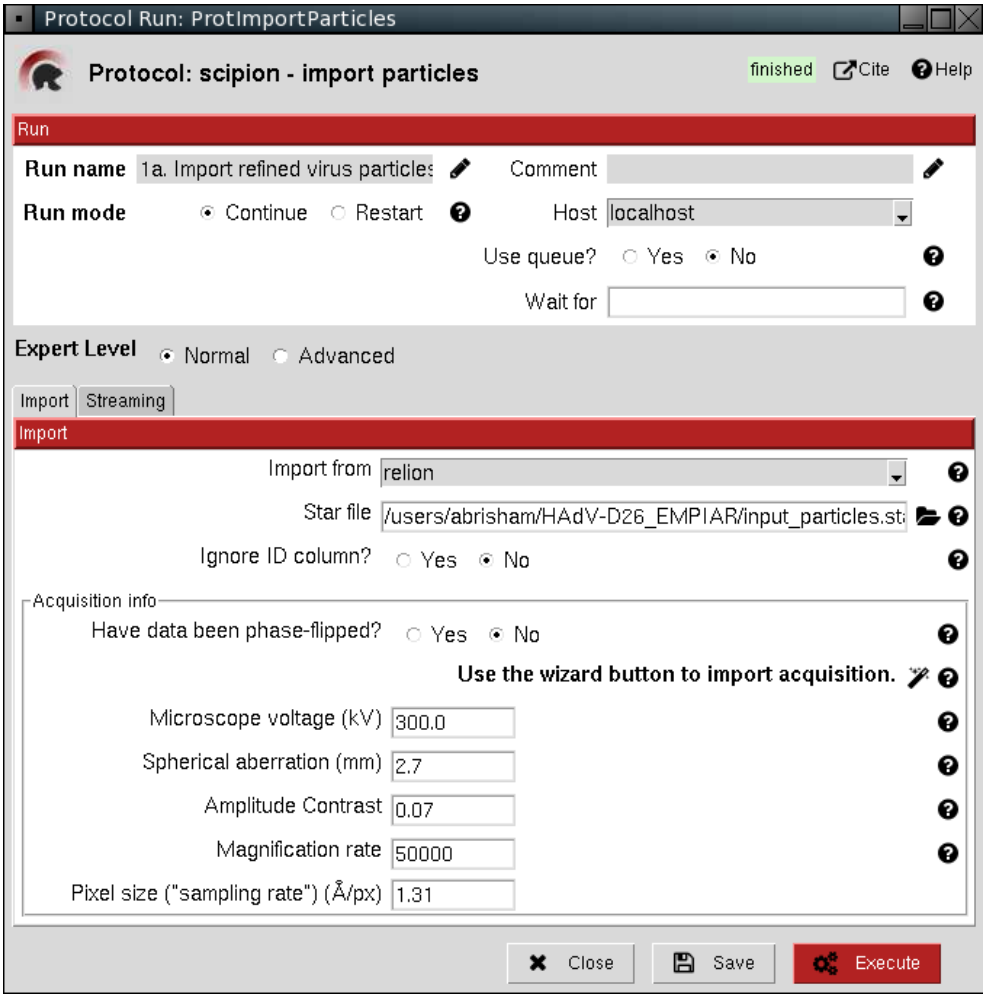


Figure 5: *Import particles* protocol with set parameters.

2.2.2. Subtract projections

After importing the particles, we subtract the signal contributed by the hexons from raw 2D particle images. For subtraction, a mask defining density to be subtracted is needed. To generate the mask, double click on 2. Mask defining hexons (or use the left menu 3D > Postprocess > relion – create 3d mask) box and set the parameters as following (Figure 6):

1. *Lowpass filter map by (A)*: Apply lowpass filter to the input volume to smooth the map and prevent from a noisy mask (with many white dots). Set this parameter to 20 A.
2. *Initial binarisation threshold*: Pixels with intensity values lower than this threshold will set to zero and the rest will set to 1. Set this threshold to 0.001.

3. *Extend binary mask by (px)*: To extend the initial binary mask with this number of pixels in all directions. This parameter is useful to fill holes in the mask. Set this parameter to 6 pixels.

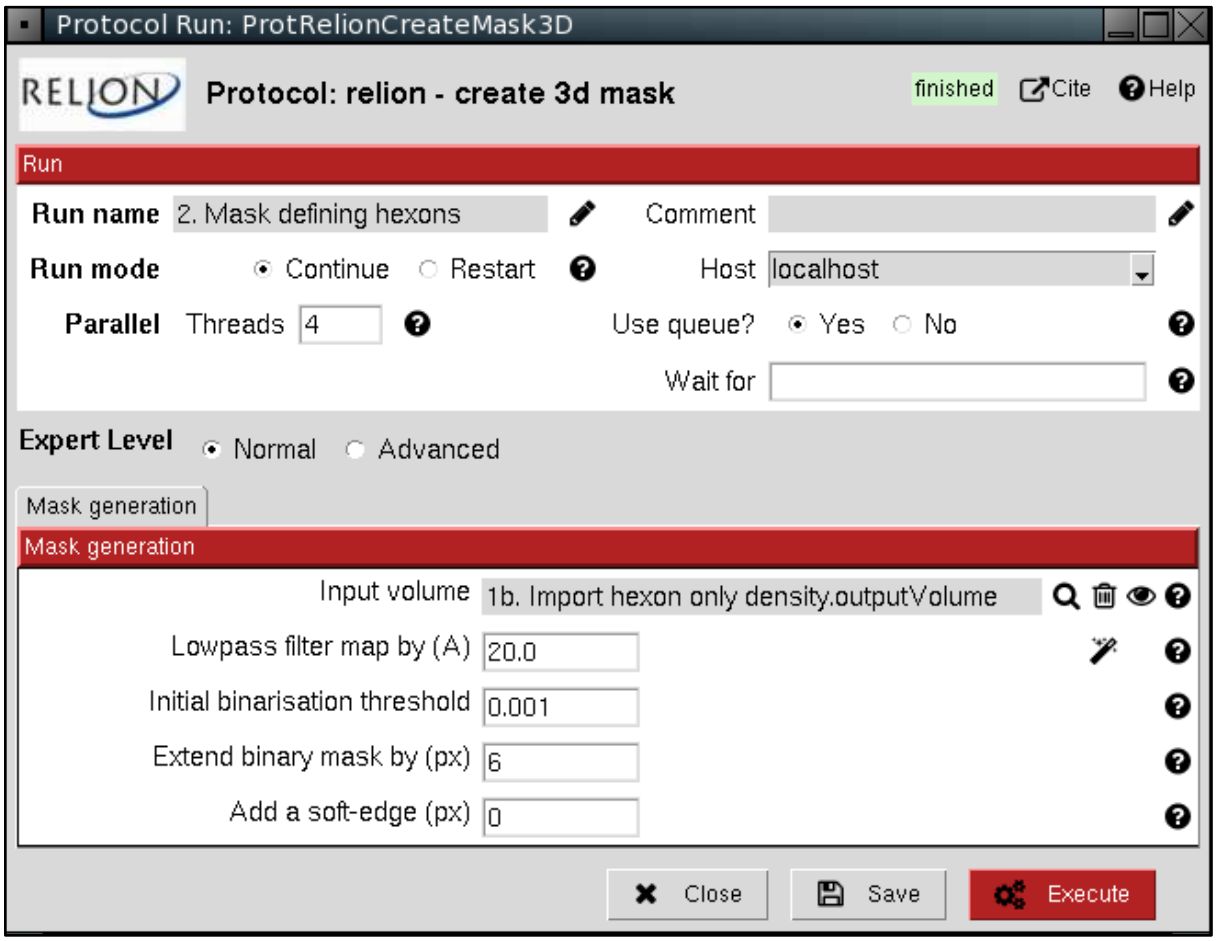


Figure 6: Create a mask defining the hexons density needed for subtraction

Run the protocol by clicking on the *Execute* button. After importing the mask, double click on 3. Subtract hexons signal (or use the left menu 3D > more > relion – subtract projection) and set the parameters as below (Figure 7):

1. *Input particles*: Particles with alignment parameters. Select 1. Import refined virus particles output particles.
2. *Input map to be projected*: This is the map to be projected for subtraction. You should select the 1e. Import icos reconstruction output volume which is a refined map where icosahedral symmetry applied.
3. *Mask to be applied to this map*: This is a mask defining the densities to be subtracted. Select 2. Mask defining hexons output mask for hexons.

After setting the parameters, click on the *Execute* button. When the protocol is finished, you can check the subtracted particle images by clicking on *Analyze Results*. You have now a set of particles almost without the contribution of hexons.

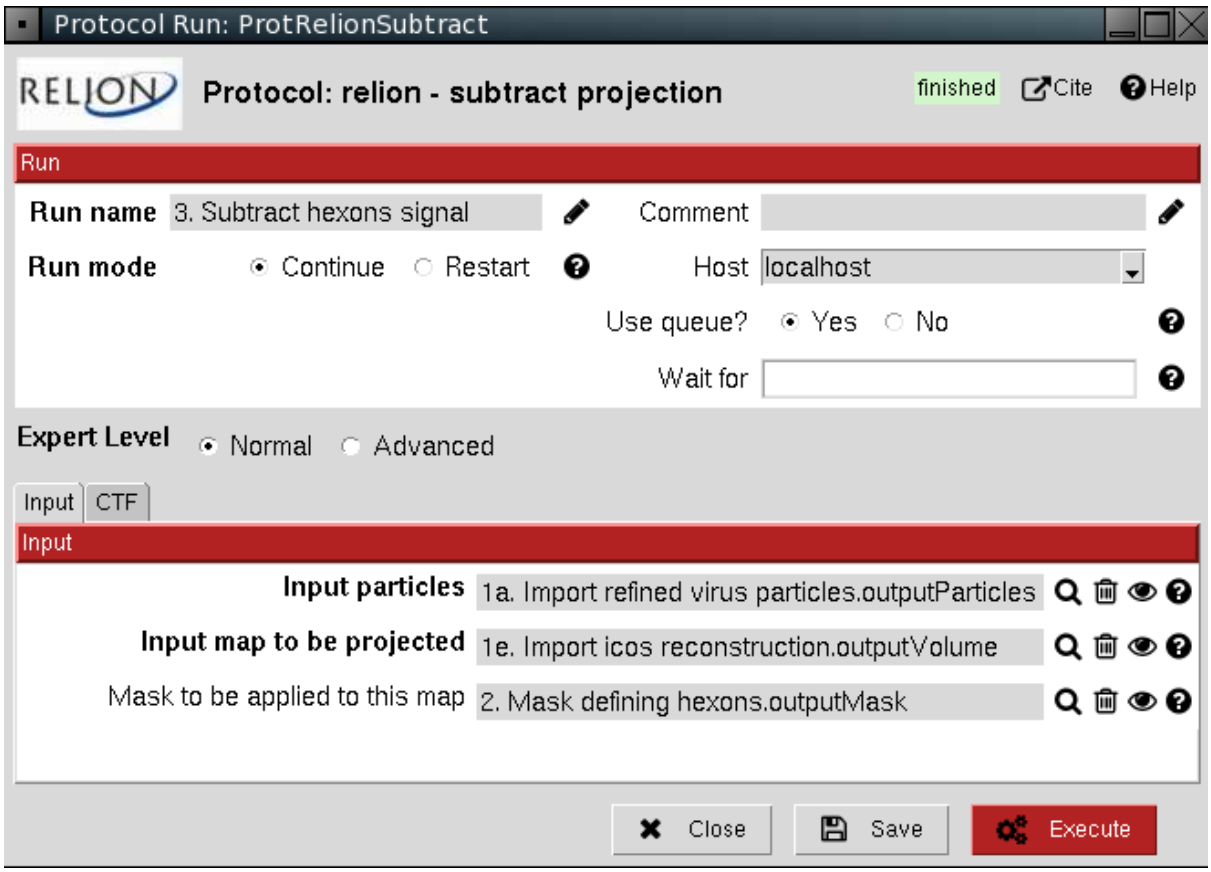


Figure 7: Subtract projection protocol parameters to subtract the contribution signal of hexons.

2.3. Prepare sub-particles

2.3.1. Define sub-particles

In this step, we define the coordinates of trimeric fibres for each virus particle. Double click on the 4. Calculate fibre coordinates box (or use the left menu View > Localized Reconstruction > Prepare subparticles > localrec – define subparticles) and then set the protocol parameters as the following (Figure 8):

1. *Input particles*: Particles with alignment parameters. Here we use the previously subtracted set of particles.
2. *Symmetry*: The symmetry used to refine the particles. Here we set it to *I1* convention.

3. *Is vector defined by*: The way you want to determine the vector to the center of the subunit. Here we choose string.
4. *Location vectors*: Vector defining the location of the sub-particles. The vector is defined by three values *x,y,z* separated by a comma. To create sub-particles for subunits at icosahedral symmetry axis, use these vectors (this assumes the so-called "I1 orientation"):

    - all 2-folds, use --vector 0.000,0.000,1.000
    - all 3-folds, use --vector 0.382,0.000,1.000
    - all 5-folds, use --vector 0.000,0.618,1.000

    since spikes are on 5-folds, and we follow the I1 convention, we set it to 0.000,0.618,1.000.
5. *Alternative length of the vector (Å)*: Use to adjust the sub-particle center. If it is <= 0, then the length of the given vector (as string) will be considered. You should set this value to 444 Å to center the sub-particles in the middle of the penton base.
6. *Align the subparticles*: To define sub-particle orientations relative to the *Z*-axis. Here we set this *Yes* to have trimeric fibre aligned on the *Z*-axis.

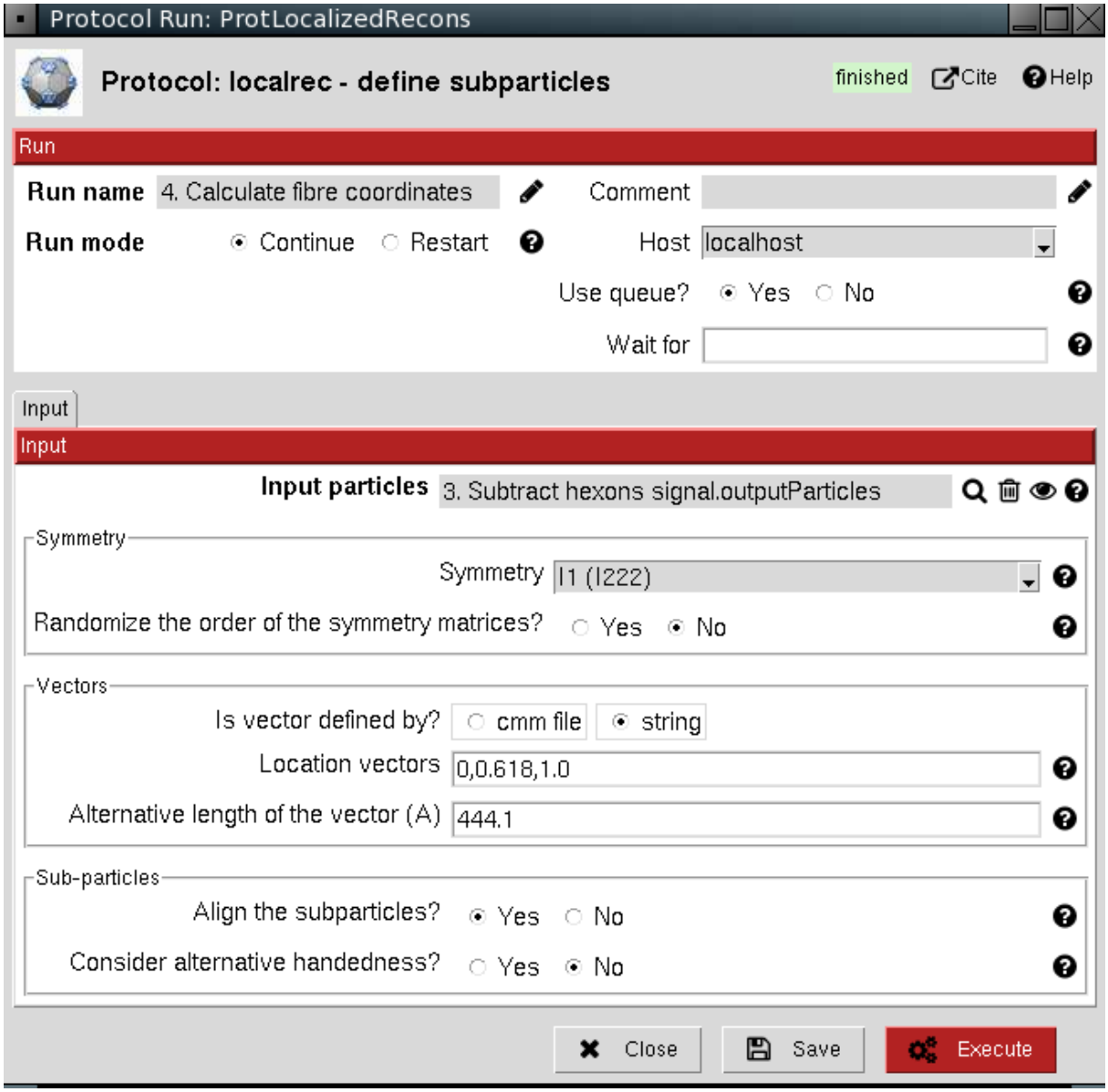


Figure 8: Parameters for define sub-particles protocol

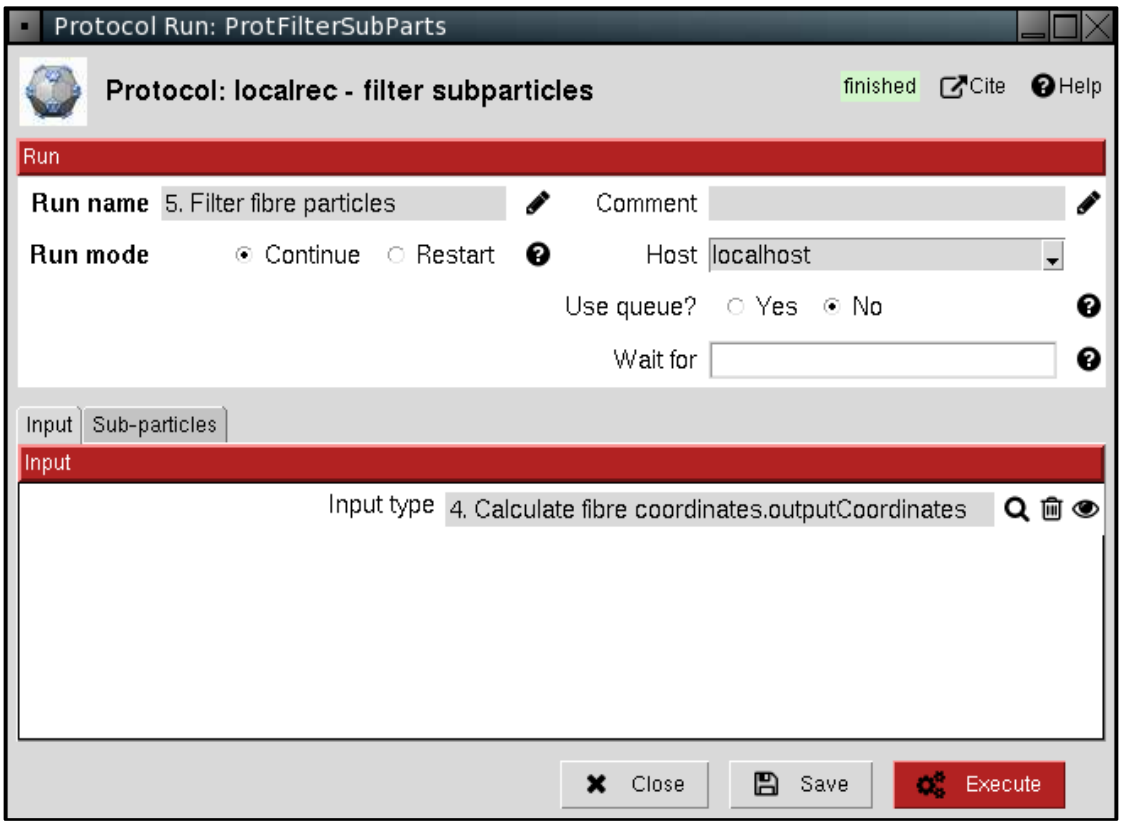
2.3.2. Filter fibre particles

At this step, we exclude unwanted sub-particles based on their alignment particles. Here we apply two filters:

1. Select just side-view sub-particles (top views can be overlapped).
2. Select a single sub-particle on the symmetry axis (instead of picking it together with its symmetry mates).

To apply this filter, double click on the 5. Filter fibre particles (or use the left menu View > Localized Reconstruction > Prepare subparticles > localrec – filter subparticles) and set the parameters as below (Figure 9):

(A)



(B)

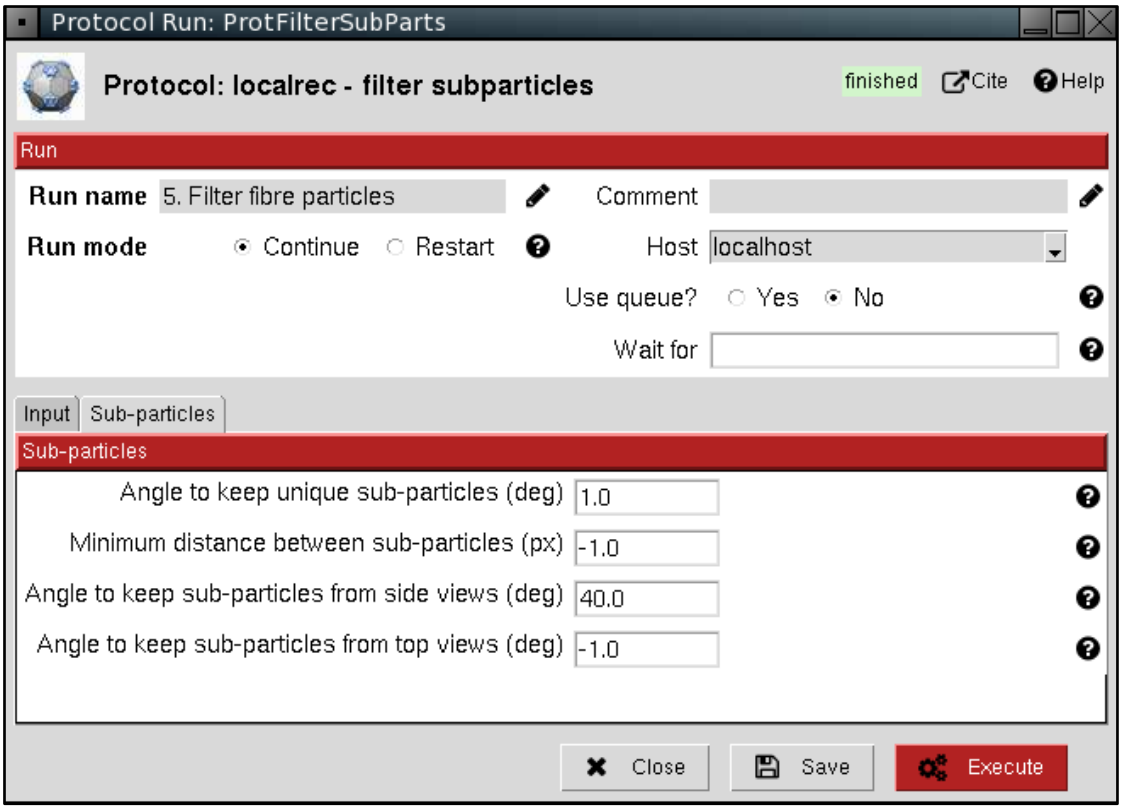


Figure 9: Filter sub-particles GUI. (A) Input tab. (B) Sub-particles tab.

Input tab:

1. *Input type*: Set of coordinates or particles to apply filter on. Select the output coordinates of 4. Calculate fibre coordinates.

Sub-particles tab:

1. *Angle to keep unique sub-particles (deg)*: This filter calculates and examines the angle between the view vectors calculated for each pair of sub-particles and discards one of them if the angle is less than this parameter value. Set this parameter to 1.
2. *Angle to keep sub-particles from side views (deg)*: To select side-view sub-particles within a specified angular distance and discard the rest of the sub-particles. Set this parameter to 40.

Click on the *Execute* button to run the protocol. This protocol will reduce the number of particles to 69,608 (it can be check in the *summary* tab at the bottom).

2.3.3. Extract fibre particles

After computing the sub-particle coordinates and filtering them, we can extract sub-particles from particle images. Click on 6. Extract fibre particles (or use the left menu View > Localized Reconstruction > Prepare subparticles > localrec – extract subparticles) and set the parameters as the following (Figure 10):

1. *Input particles*: Particle images with alignment parameters. Click on the button with magnifier icon and select output particles from 1. Import refined virus particles.
2. *Input coordinates*: Coordinates related to the sub-particles. Click the magnifier button and select output coordinates form 4. Calculate fibre coordinates.
3. *Subparticle box size (px)*: The box size in pixels for sub-particle extraction. Set the box size to 280.

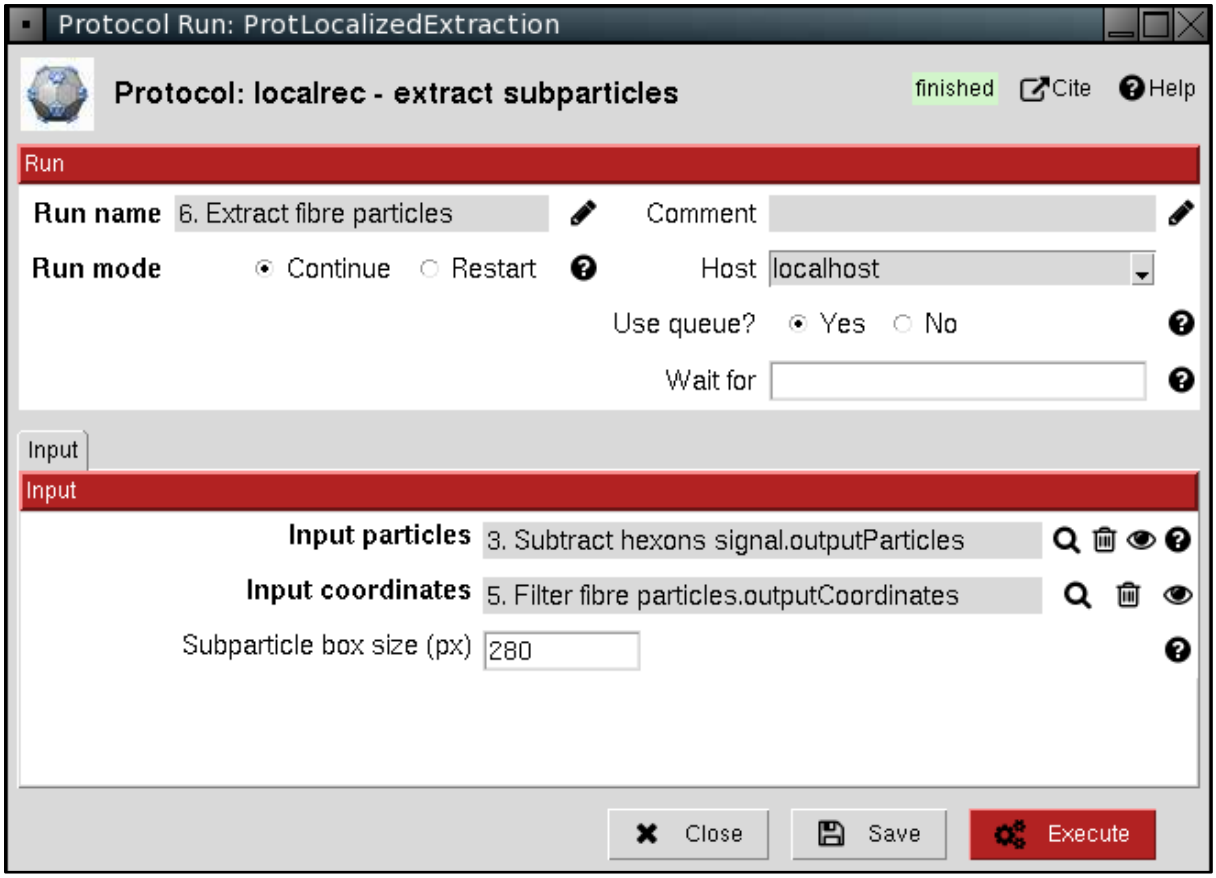


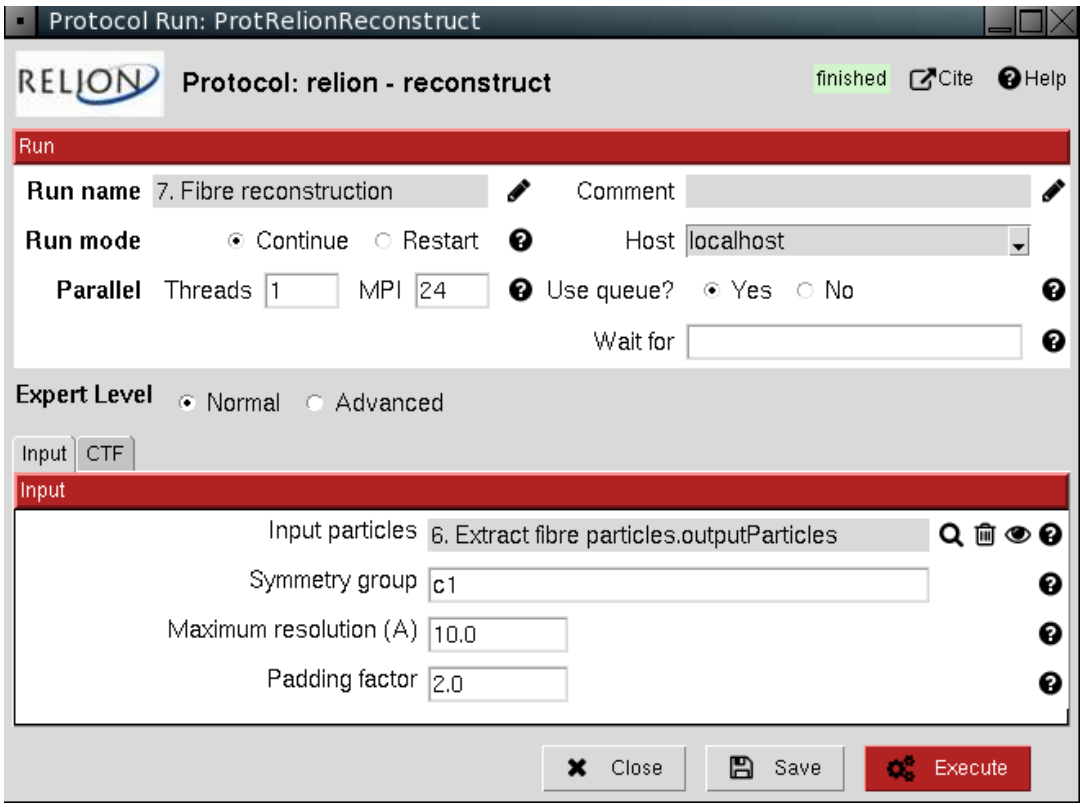
Figure 10: Extract sub-particles protocol GUI.

2.4. Reconstruct sub-particles

2.4.1. Fibre reconstruction

Having fibre sub-particles with alignment parameters from the previous protocol, now we can apply a 3-D reconstruction to have an initial estimation of the fibre structure. This step is not only for sub-particle picking validation, but it also provides an initial volume for 3-D refinement/classification. Double click 7. Fibre reconstruction box (or use the left menu 3D > Reconstruct > relion – reconstruct) and set the parameters as the following (Figure 11):

(A)



(B)

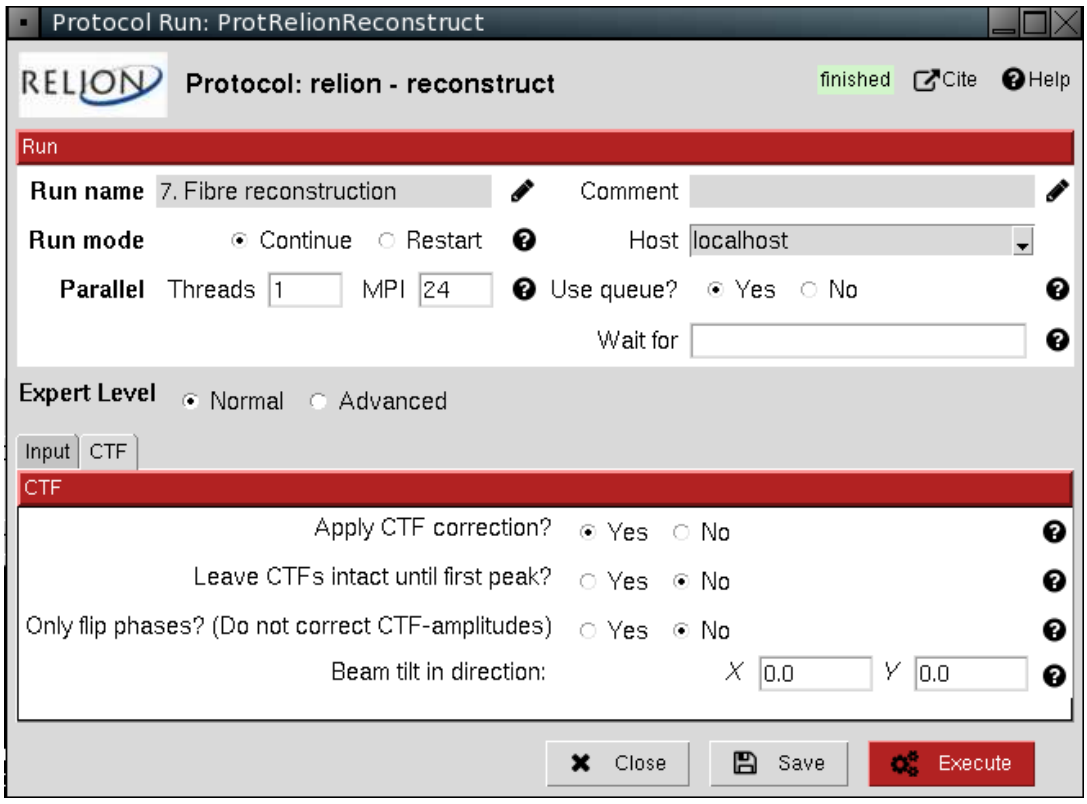


Figure 11: Reconstruction protocol GUI. (A) Input tab. (B) CTF tab.

Input tab:

1. *Input particles*: Select 6. Extract fibre particles output particles.
2. *Symmetry group*: Type c1 in this box to not apply any symmetry.
3. *Maximum resolution (A)*: This is the highest resolution for the reconstruction. Here it is set to 10 Å since we don't need a high-resolution structure at this level.

CTF tab:

1. *Apply CTF correction?:* Set this option to *Yes* to correct for the CTF.

Now click on the *Execute* button to run the protocol. When the protocol is finished, click on the *Analyze Results* button to check if you have selected the subunit correctly (Figure 12).

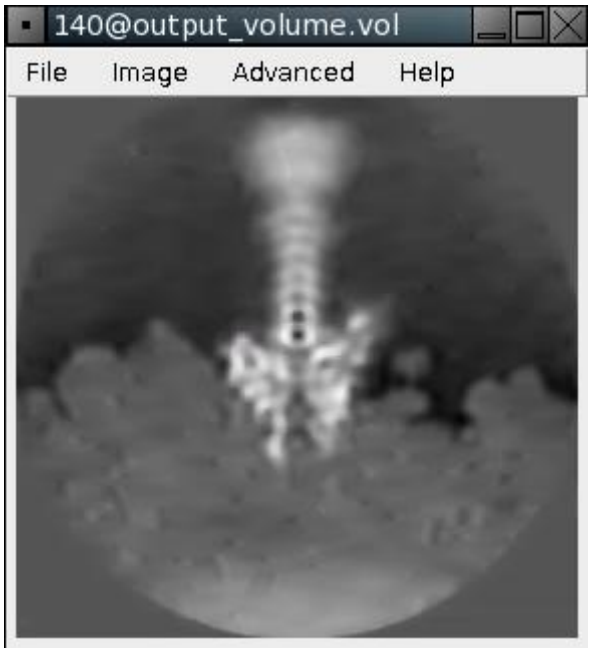


Figure 12: Centre slice of the fibre map after reconstruction (side view).

2.4.2. 3D classification of sub-particles

In this step, a 3D classification focused on the trimeric fibre is carried out using Relion 3D classification to classify the sub-particles into three classes. To determine which of the five possible C5-related views of the template matched the observed sub-particle the best, we relaxed the symmetry of the five-fold symmetric penton base (C5) to no symmetry (C1). To do this double click on 8. Fibre 3D classification box (or use the left menu 3D > Classify > relion – 3D

classification) and set the parameters as the following (select *Advanced* mode for the *Expert Level* of this protocol)(Figure 13):

Input tab:

1. *Input Particles*: Select output particles of 6. Extract fibre particles.
2. *Consider previous alignment*: Set it to *Yes* and then set *Consider alignment as priors* also to *Yes* to consider the previous alignment parameters as priors (this is necessary for doing a local search).

Reference 3D map tab:

1. *Input volume(s)*: Select the output volume from 7. Fibre reconstruction. This volume is used as the initial volume for the classification.
2. *Reference mask (optional)*: This mask is used for focused classification. Here we want to focus on the signal from fibre to prevent strong impact of penton base on the alignment. Set this to the output mask of 1d. Create mask for single fibre protocol.
3. *Is initial 3D map on absolute greyscale*?: Set this parameter to yes since the initial map is a Relion map.
4. *Initial low-pass filter (A)*: Use this parameter to filter an initial volume to prevent bias. Set this parameter to 40.

CTF tab:

1. *Do CTF-correction*?: Set this parameter to *Yes* to perform CTF correction.
2. *Has reference been CTF-corrected*?: Set this to *Yes* since our initial volume is CTF-corrected.

Optimisation tab:

1. *Number of classes*: Set this parameter to 3.

Sampling and compute tabs:

By relaxing symmetry, we want to limit the search space to *n* directions representing the “symmetry mates” of the particle (e.g., 60 symmetry mates in the case of icosahedral symmetry). In the compute tab:

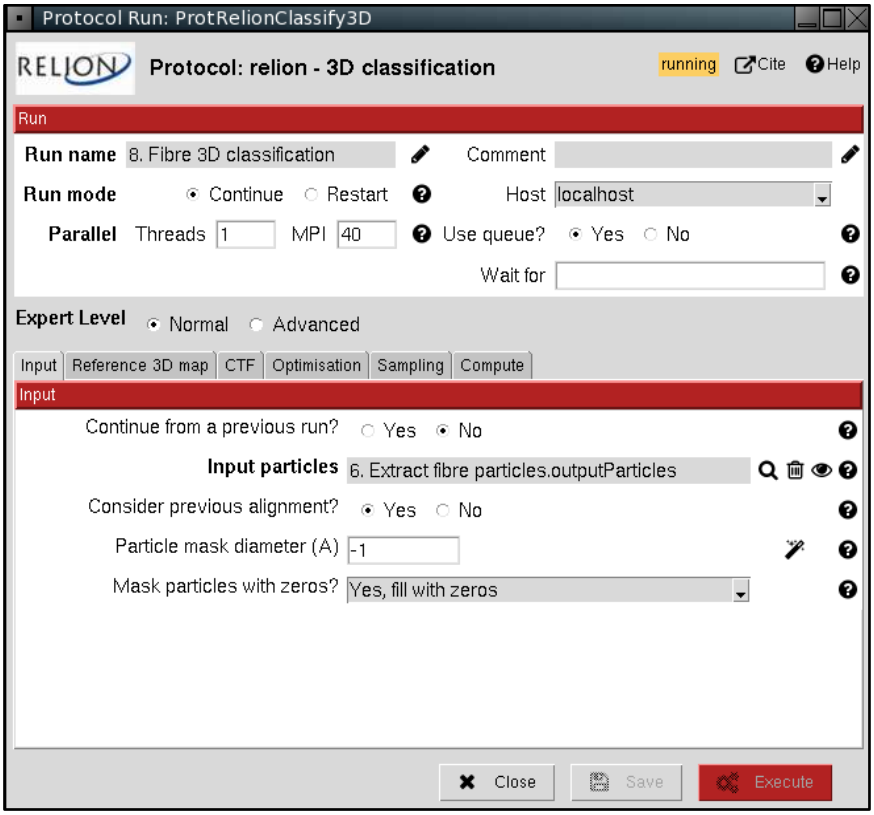
1. *Additional arguments*: Add --*relax_sym C5* to relax symmetry C5->C1. You can also add --sigma_ang X° to search a range of X° around each symmetry mate. Here we skip the --sigma_ang parameter and allow the program to set it to a very small default value (0.003) to consider exactly 5 symmetry mates.

When the relax_sym option is used, refinement performs a local search around the particle's symmetry mates. In the sampling tab:

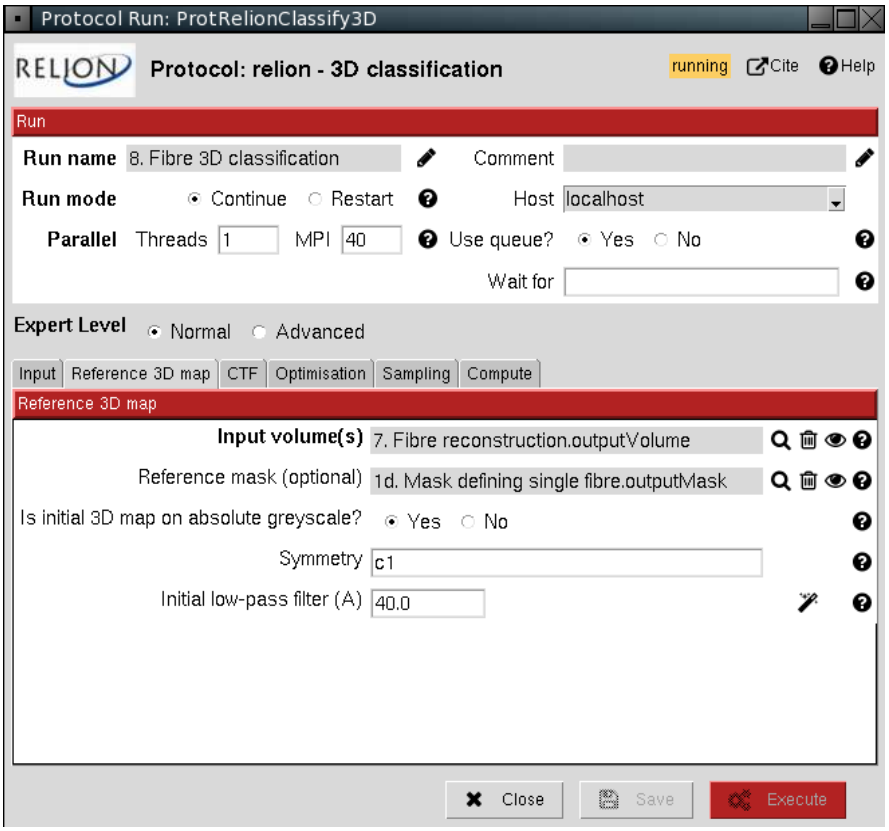
1. *Perform local angular search?*: Set this parameter to *Yes* to activate local search.
2. *Local angular search range*: Set proper sampling interval for --sigma_ang X° range. Here we use the value 0.01 since --sigma_ang is so small that no sampling can be done.

Then click on the Execute button at the bottom to run the protocol. After the classification is done, click on the *Analyze Results* button and then select *Show classification in Scipion* from the viewer to check three generated classes (Figure 12). As you can see, just class 1 with 40,000 particles looks reasonable. Select this class and click on the *Particles* button at the bottom. Name the subset as 9. 40,000 good fibre particles and press *ok* button.

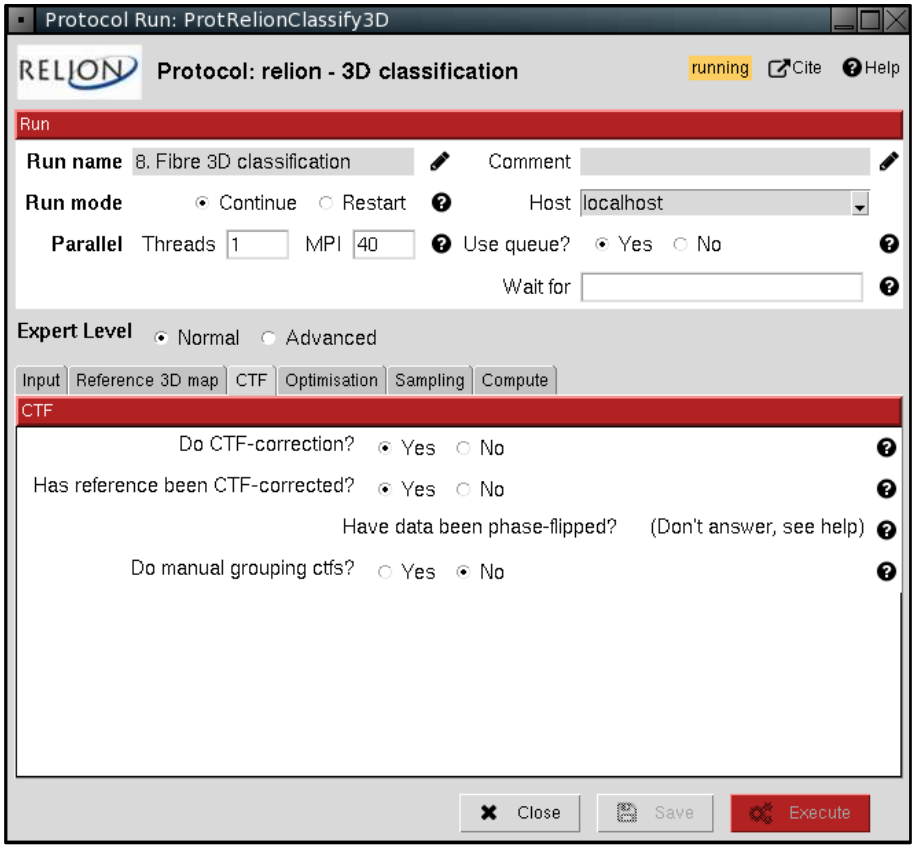
(A)


Protocol Run: ProtRelionClassify3D
RELION
Protocol: relion - 3D classification
running
Cite
Help
Run
Run name 8. Fibre 3D classification
Comment
Run mode Continue Restart
Host localhost
Parallel Threads 1 MPI 40
Use queue? Yes No
Wait for
Expert Level Normal Advanced
Input Reference 3D map CTF Optimisation Sampling Compute
Input
Continue from a previous run? Yes No
Input particles 6. Extract fibre particles.outputParticles
Consider previous alignment? Yes No
Particle mask diameter (A) -1
Mask particles with zeros? Yes, fill with zeros
Close
Save
Execute

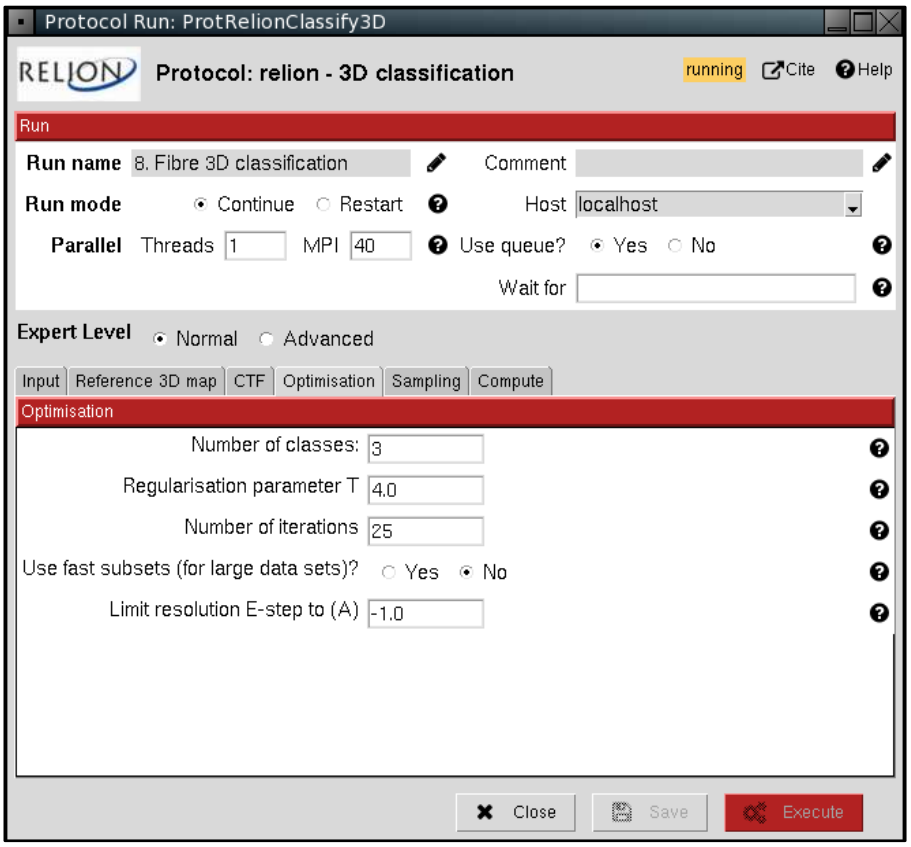
(B)


Protocol Run: ProtRelionClassify3D
RELION
Protocol: relion - 3D classification
running
Cite
Help
Run
Run name 8. Fibre 3D classification
Comment
Run mode Continue Restart
Host localhost
Parallel Threads 1 MPI 40
Use queue? Yes No
Wait for
Expert Level Normal Advanced
Input Reference 3D map CTF Optimisation Sampling Compute
Reference 3D map
Input volume(s) 7. Fibre reconstruction.outputVolume
Reference mask (optional) 1d. Mask defining single fibre.outputMask
Is initial 3D map on absolute greyscale? Yes No
Symmetry c1
Initial low-pass filter (A) 40.0
Close
Save
Execute

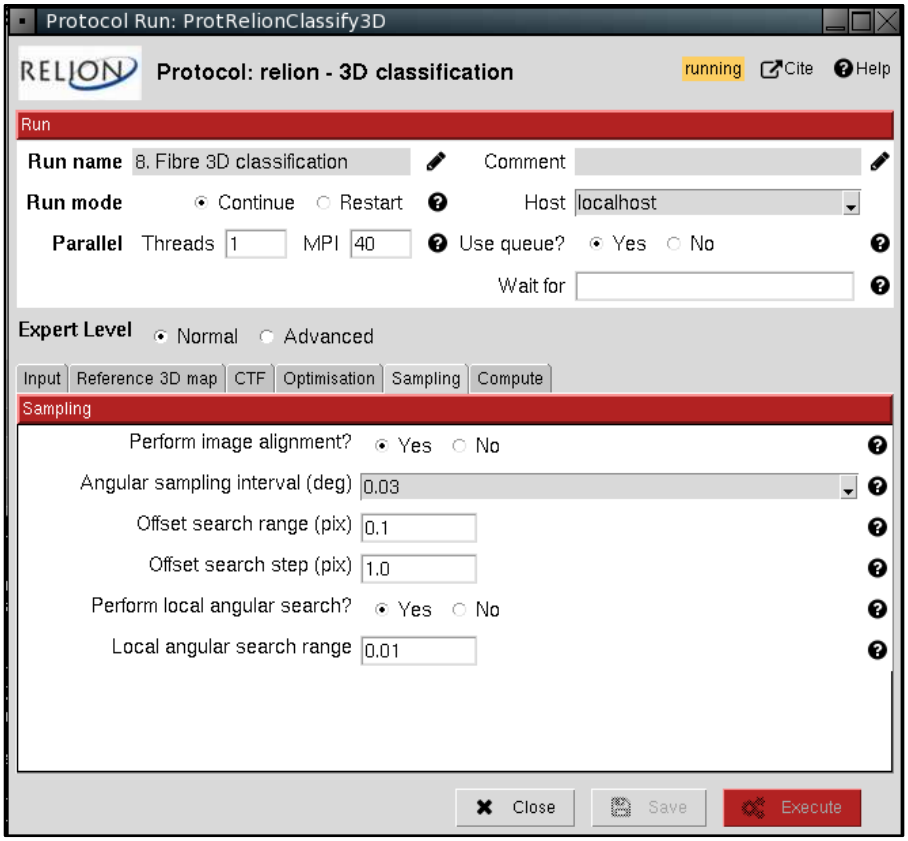
(C)


Protocol Run: ProtRelionClassify3D
RELION
Protocol: relion - 3D classification
running
Cite
Help
Run
Run name 8. Fibre 3D classification
Comment
Run mode Continue Restart
Host localhost
Parallel Threads 1 MPI 40
Use queue? Yes No
Wait for
Expert Level Normal Advanced
Input Reference 3D map CTF Optimisation Sampling Compute
CTF
Do CTF-correction? Yes No
Has reference been CTF-corrected? Yes No
Have data been phase-flipped? (Don't answer, see help)
Do manual grouping ctfs? Yes No
Close
Save
Execute

(D)


Protocol Run: ProtRelionClassify3D
RELION
Protocol: relion - 3D classification
running
Cite
Help
Run
Run name 8. Fibre 3D classification
Comment
Run mode Continue Restart
Host localhost
Parallel Threads 1 MPI 40
Use queue? Yes No
Wait for
Expert Level Normal Advanced
Input Reference 3D map CTF Optimisation Sampling Compute
Optimisation
Number of classes: 3
Regularisation parameter T 4.0
Number of iterations 25
Use fast subsets (for large data sets)? Yes No
Limit resolution E-step to (A) -1.0
Close
Save
Execute

(E)


Protocol Run: ProtRelionClassify3D
RELION
Protocol: relion - 3D classification
running
Cite
Help
Run
Run name 8. Fibre 3D classification
Comment
Run mode Continue Restart
Host localhost
Parallel Threads 1 MPI 40
Use queue? Yes No
Wait for
Expert Level Normal Advanced
Input Reference 3D map CTF Optimisation Sampling Compute
Sampling
Perform image alignment? Yes No
Angular sampling interval (deg) 0.03
Offset search range (pix) 0.1
Offset search step (pix) 1.0
Perform local angular search? Yes No
Local angular search range 0.01
Close
Save
Execute

(F)

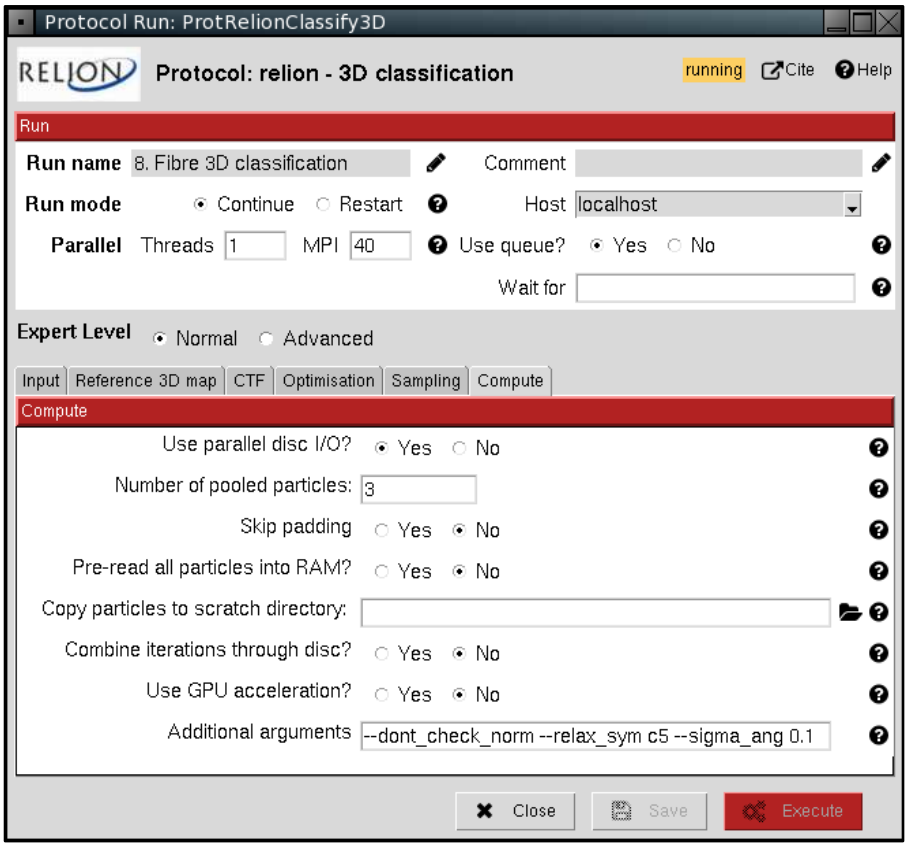

Protocol Run: ProtRelionClassify3D
RELION
Protocol: relion - 3D classification
running
Cite
Help
Run
Run name 8. Fibre 3D classification
Comment
Run mode Continue Restart
Host localhost
Parallel Threads 1 MPI 40
Use queue? Yes No
Wait for
Expert Level Normal Advanced
Input Reference 3D map CTF Optimisation Sampling Compute
Compute
Use parallel disc I/O? Yes No
Number of pooled particles: 3
Skip padding Yes No
Pre-read all particles into RAM? Yes No
Copy particles to scratch directory:
Combine iterations through disc? Yes No
Use GPU acceleration? Yes No
Additional arguments --dont_check_norm --relax_sym c5 --sigma_ang 0.1
Close
Save
Execute

Figure 13. 3D classification protocol GUI. (A) *Input* tab. (B) *Reference 3D map* tab. (C) *CTF* tab. (D) *Optimisation* tab. (E) *Sampling* tab. (F) *Compute* tab.

2.4.3. 3D refinement of sub-particles

To calculate two half-maps in addition to the final map at the appropriate resolution in one step, here we use Relion 3D refinement, where no shifts or rotations are allowed. Click on 9. Fibre 3D refinement box (or use the left menu 3D > Refine > relion – 3D auto-refine) and set the parameters as below:

Input tab:

1. *Input Particles*: Select output particles of 40,000 good fibre particles.
2. *Consider previous alignment?*: Set it to *Yes* and then set both *Consider alignment as priors?*, and *Consider random subset value?* also to *Yes* to consider the previous alignment parameters as priors (this is necessary for doing a local search).

Reference 3D map tab:

1. *Input volume(s)*: Select the output volume of class 1 from 8. Fibre 3D classification. This volume is used as the initial volume for the classification.
2. *Reference mask (optional)*: This mask is used for focused classification. Here we want to focus on the signal from fibre to prevent the strong impact of penton base on the alignment. Set this to the output mask of 1d. Mask defining single fibre protocol.
3. *Is initial 3D map on absolute greyscale*?: Set this parameter to *Yes* since the initial map is a Relion map.
4. *Initial low-pass filter (A)*: Use this parameter to filter an initial volume to prevent bias. Set this parameter to 40.

CTF tab:

1. *Do CTF-correction?*: Set this parameter to *Yes* to perform CTF correction.
2. *Has reference been CTF-corrected*?: Set this to *Yes* since our initial volume is CTF-corrected

Sampling tab:

1. *Offset search range (pix)*: Search for translation in a circle with this radius. We set it to 0.1, a value less than 1 pixel, thus no search for translations will be done.
2. *Local search from auto-sampling (deg)*: This is the sampling interval for local search. Select 0.005 from the drop-down list.

Compute tab:

1. *Additional arguments*: Put --sigma_ang 0.001. We set such a small --sigma_ang value to prevent 3D refinement algorithm to change current orientation parameters. In real cases, one would further refine the angles by adding the relax_sym option to arguments and set oversampling to 0. However, in this tutorial, we skipped this step.

Click the *Execute* button to run the protocol. After the protocol is finished, you should see 7.64 A resolution in the *summary* tab in the main project window.

2.5. Analysis

2.5.1. Fibre postprocessing

We calculate a postprocessed map from independent maps by sharpening with automatically estimated B-factors in the postprocessing step. Click on 10. Fibre postprocessing (or use the left menu 3D > Postprocess > relion – post-processing), and set the parameters as follow:

1. *Select a previous refinement protocol*: Select 9. Fibre 3D refinement protocol. The post-processing protocol uses the two half maps generated by the refinement algorithm.
2. *Solvent mask*: To remove the solvent noise surrounding a molecular envelope. Select the output mask of 1d. Mask defining single fibre protocol.

After setting the parameters, click on the *Execution* button to run the protocol. You can check the final resolution in the *summary* tab which should be equal to 7.3 Å.

3. Localized reconstruction of hexons 1–4 of adenovirus

In this section, we will reconstruct the HAdV-D26 asymmetric unit by combining four localized reconstructions; each centered on one of the four unique hexons not related by icosahedral symmetry. To start, create a new project in Scipion and name it "HAdV-D26 hexons". When the project window is open, choose the workflow json file by going to Project -> Import workflow and select "HAdV-Hexons-Workflow.json" that you have downloaded before. Now you can see the whole workflow for this project continue with the rest of this section.

3.1. Prepare particles

3.1.1. Import particles

Double-click the 1. import virus particles box and then follow the instructions in section 2.2.1 of this tutorial (Figure 3) to import virus particles into your project.

3.2. Prepare particles

3.2.1. Define sub-particles

In this step, we define the coordinates of four hexons. Double click on the 2a-2d. Calculate hexon#*X* coordinates (*X* is the hexon number) boxes one by one and set the protocol parameters for each hexon as the following (Figure 14):

1. *Input particles*: Particles with alignment parameters. Here we use the previously subtracted set of particles.
2. *Symmetry*: The symmetry used to refine the particles. Here we set it to *I1* convention.
3. *Is vector defined by*: The way you want to define the vector to the center of the subunit. Here we choose *cmm file*.
4. *file obtained by chimera*: A Chimera *cmm* file defines the initial and terminal points of the vector. The first marker, specifying the initial point, is typically placed at the centre of the particle map and the second marker, specifying the terminal point, is always placed on the sub-structure of interest. Choose one of the provided *hexonX.cmm* (hexon1.cmm for the first hexon) file depending on the hexon you are processing.

5. *Align the subparticles*: Here we set this *No* since we don't want the hexons to align on the *Z*-axis.

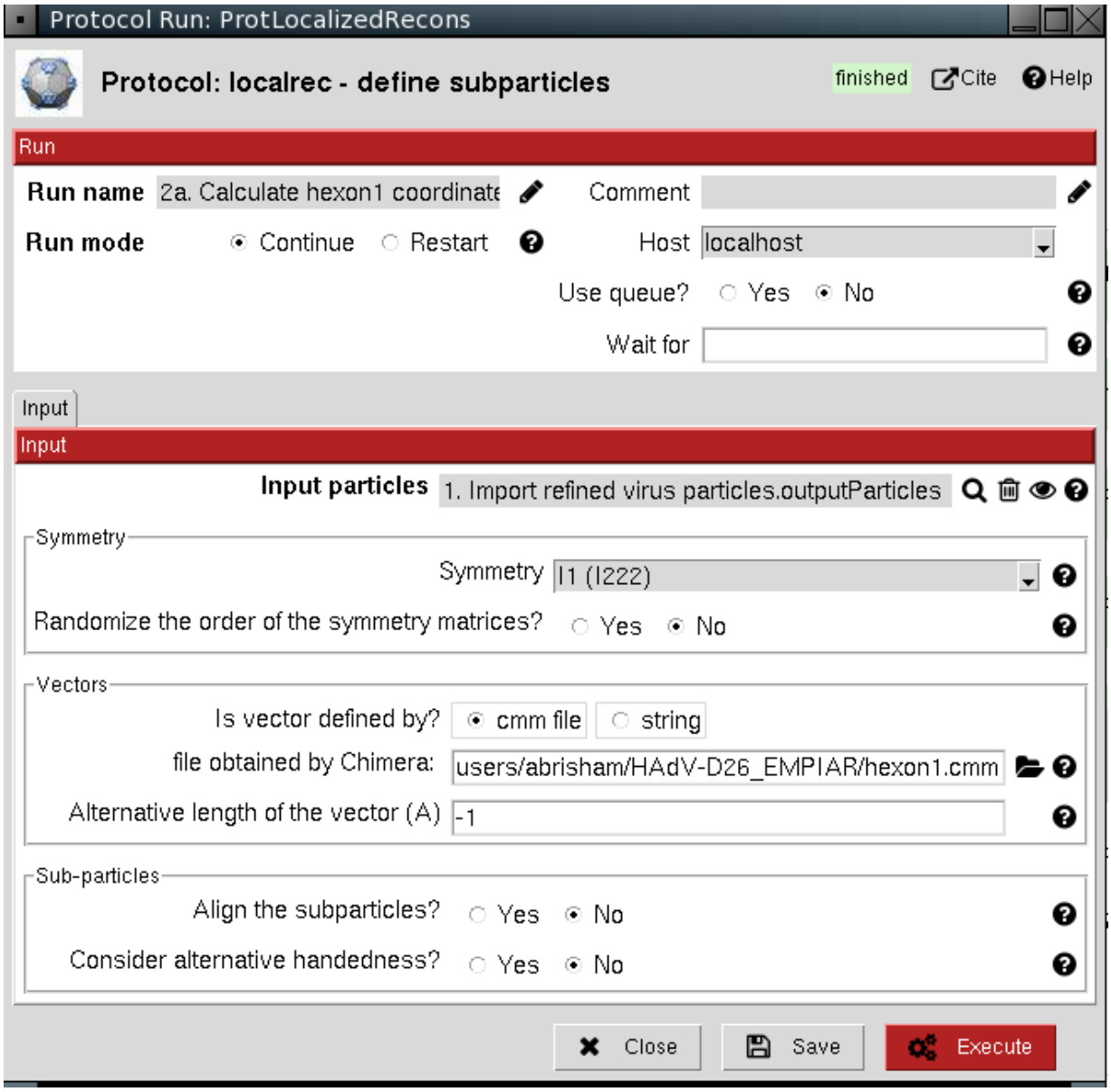


Figure 14. Define sub-particles protocol GUI for one hexon.

3.2.2. Extract the sub-particles

Having defined the sub-particle coordinates for the hexons, the sub-particle images can be extracted. Click on 3a-d. Extract hexon#*X* particles (*X* is the hexon number) and set the parameters as the following:

1. *Input particles*: Particle images with alignment parameters. Click on the button with a magnifier icon and select output particles from 1. Import refined virus particles.

2. *Input coordinates*: Coordinates related to the sub-particles. Click the magnifier button and select output coordinates form 2a-d. Calculate hexon#*X* coordinates depending on the hexon you are processing.
3. *Subparticle box size (px)*: The box size in pixels for sub-particle extraction. Set the box size to 256.

After setting the parameters, click *Execute* to run the protocol.

3.3. Reconstruct sub-particles

3.3.1. Reconstruct hexons1-4

Having hexon1-4 sub-particles with alignment parameters from the previous protocol, now we can calculate a 3-D reconstruction to have an initial estimation of each hexon structure. Double click on 4a-d. Hexon#*X* reconstruction box and set the parameters as the following:

Input tab:

1. *Input particles*: Select output particles of 3a-d. Extract hexon#*X* particles depending on the hexon under processing.
2. *Symmetry group*: Type c1 in this box to not apply any symmetry.
3. *Maximum resolution (A)*: This is the highest resolution for the reconstruction. Here it is set to 10 Å since we don't need a high-resolution structure at this point.

CTF tab:

1. *Apply CTF correction?:* Set this option to *Yes* to correct for the CTF.

3.3.2. Refinement of hexons1-4

In this step, we refine the hexons1-4 by finding the optimal shifts and orientations for each hexon. For this refinement, we search locally around the priors (current orientations of sub-particles) by defining a sigma value and also allow some shifts. To do the refinements, double click on 5a-d. Hexon#*X* refinement (depending on which hexon is under processing) and set the parameters as follow:

Input tab:

1. *Input Particles*: Select one of 3a-d. Extract hexon#X particles depending on the hexon you are processing.
2. *Consider previous alignment*: Set it to *Yes* and then set *Consider alignment as priors* also to *Yes* to consider the previous alignment parameters as priors (this is necessary for localized refinement).

Reference 3D map tab:

1. *Input volume(s)*: Select the output volume from 4a-d. Hexon #*X* reconstruction depending on the hexon you are processing. This volume is used as the initial volume for the refinement.
2. *Is initial 3D map on absolute greyscale*?: Set this parameter to yes since the initial map is a Relion map.
3. *Initial low-pass filter (A)*: Use this parameter to filter an initial volume to prevent bias. Set this parameter to 60.

CTF tab:

1. *Do CTF-correction*?: Set this parameter to *Yes* to perform CTF correction.
2. *Has reference been CTF-corrected*?: Set this to *Yes* since our initial volume is CTF-corrected

Sampling tab:

1. *Angular sampling interval (deg)*: Set this parameter to 0.9.
2. *Local search from auto-sampling (deg)*: Set this parameter to 0.9 as the *Angular sampling interval (deg)* to force refinement to enter local search directly. This interval will be used to take samples in the sigma range.

Compute tab:

1. *Additional arguments*: Put --sigma_ang 0.9. The refinement will search in 3 * sigma_ang around the priors.
2. *Oversampling*: Set the *Expert Level* of the protocol to *advanced* to see this parameter. Then set it to zero.

After refinements, we should have 3.29, 3.26, 3.22, and 3.22 resolution for hexon1-4, respectively (you can check it in the *summary* tab at the bottom).

3.4. Analysis

3.4.1. Create hexons1-4 postprocessing masks

Before postprocessing, we need to generate a mask for each of the hexons. Double click on 6a-d. Create mask for Hexon #*X* boxes and set the parameters as below:

4. *Lowpass filter map by (A)*: Apply lowpass filter to the input volume to smooth the map and prevent creating a noisy mask (with many white dots). Set this parameter to 20.
5. *Initial binarisation threshold*: Pixels with intensity values lower than this threshold will set to zero and the rest will set to 1. Set this threshold to 0.001.
6. *Add soft-edge (px)*: This prevents the mask from going rapidly to zero at the border but smoothly. It is essential to have some soft edges since postprocessing can be affected by sharp edges (shows wrongly higher resolutions). Set this to 6 pixels.

3.4.2. Postprocessing of hexons1-4

At this step, we calculate a postprocessed map from independent maps by sharpening with automatically estimated B-factors. Click on 7a-d. Hexon #*X* postprocessing, and set the parameters as follow:

1. *Select a previous refinement protocol*: Select the output of one of the 5a-d. Hexon #*X* refinement depending on the hexon under processing. The post-processing protocol uses the two half maps generated by the refinement algorithm.
3. *Solvent mask*: To remove the solvent noise surrounding a molecular envelope (and in this case neighbouring hexon density that may be less well ordered due to flexibility of the capsid). Select the output mask of 6a-d. Create mask for hexon #*X* protocol depending on the hexon under processing.

3.4.3. Combining reconstructions into a composite map

In this step, we want to build a composite map by combining four reconstructions (hexon1-4). Double click on 8. Build composite map box (or use

the left menu View > Localized Reconstruction > Analysis > localrec – stitch subparticles) and set the parameters as below (Figure 15):

1. *Use two half maps?*: By setting it to *Yes*, the protocol needs two half maps of each reconstruction and finally produces two independent half maps for the composite map. This option is useful when you want to calculate the resolution for the composite map. By setting this parameter to *No*, just one map is required for each subunit (normally postprocessed map), and the output is one composite map.
2. *Input sub-volumes*: Here you should provide a list of subunits reconstructions. You should choose 4 output volumes of 7a-d. Hexon #*X* postprocessing.
3. *Masks*: If you want to define regions of interest for your sub-volumes, here you can select a list of masks with the same order as sub-volumes. Here we leave it blank, so the program will automatically use a spherical mask.
4. *Method for interpolation*: How to interpolate the voxel values when combining the sub-volumes. Leave the default value (linear) since bspline method is more complex and the results are almost equal.
5. *Output volume size*: Size for the composite map. Set it to 1200.
6. *Use previous localrec run(s)*: By selecting *Yes*, you should provide a list of *localrec - define subparticles* protocols used to generate the sub-volumes in the same order as the sub-volumes. By choosing *No*, you should give information on symmetry and vectors for each sub-volume manually. Here we select *Yes* to get this information from previous *localrec - define subparticles* protocols.
7. *Localrec previous runs*: If you set 6 to *Yes*, then you should provide a list of localrec runs. Select 2a-2d. Calculate hexon#X coordinates protocols in the same order as sub-volumes.

Setting the parameters, now you can click *Execute* to run the protocol. After the execution, you can check the output volume by clicking on the *Analyze Results* button.

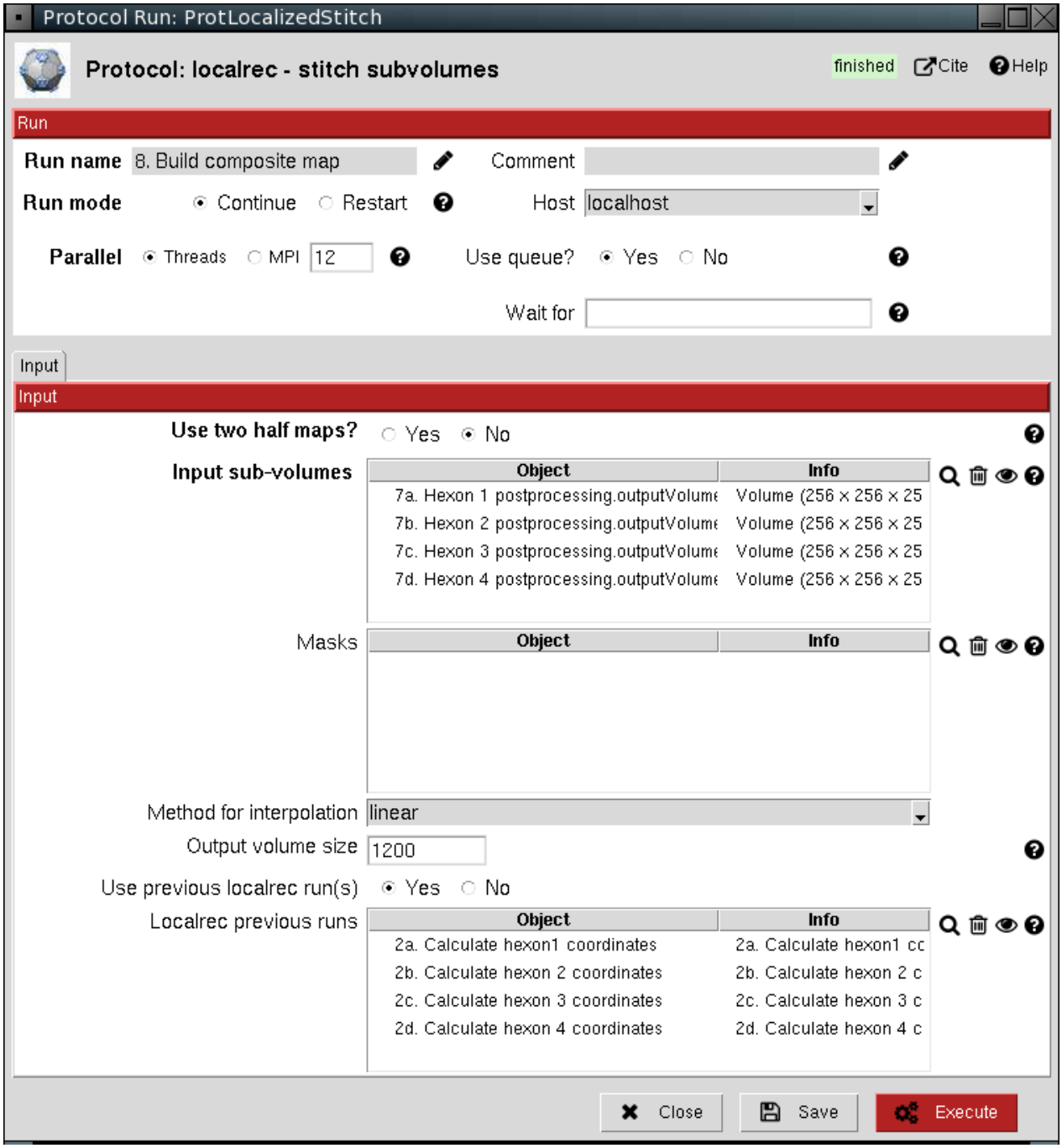


Figure 15. Stitch sub-volumes protocol GUI.

References

[1] S.L. Ilca, A. Kotecha, X. Sun, M.M. Poranen, D.I. Stuart, J.T. Huiskonen, Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes, Nat. Commun., 6 (Nov 4) (2015), p. 8843, 10.1038/ncomms9843.

[2] V. Abrishami, S.L. Ilca, J. Gomez-Blanco, I. Rissanen, J.M. de la Rosa-Trevín, V.S. Reddy, J.M. Carazo, J.T. Huiskonen, Localized reconstruction in scipion expedites the analysis of symmetry mismatches in Cryo-EM data, Progress in Biophysics and Molecular Biology (2020).