EMBO Image processing for cryo-EM course
RELION practical 9th September 2019

Sjors Scheres

Single-particle analysis of electron cryo-microscopy (cryo-EM) images allows structure determination of macromolecular complexes. But when these molecules adopt many different conformations, traditional image processing approaches often lead to blurred reconstructions. By considering complexes to be comprised of multiple, independently moving rigid bodies, multi-body refinement in RELION enables structure determination of highly flexible complexes, while at the same time providing a characterisation of the motions in the sample. Here, we describe how to perform multi-body refinement in RELION using a publicly available example. We outline how to prepare the necessary files, how to run the actual multi-body calculation, and how to interpret its output. This method can be applied to any cryo-EM data set of flexible complexes that can be divided into two or more bodies, each with a minimum molecular weight of 150-200 kDa.
(From Nakane & Scheres 2019,
ftp://ftp.mrc-lmb.cam.ac.uk/pub/scheres/multibody_protocol.pdf)

We will be working with a dataset of pre-catalytic spliceosomal B-complex collected on a Titan Krios, K2 Summit GIF Quantum detector
Data deposited with EMPRIAR, entry 10180
Plaschka, Lin & Nagai, Nature, 2017

In this practical you will go through the steps for generating an initial model of the dataset, followed by 3D classification, 3D refinement and finally multi-body refinement, as well as pre-processing steps and mask generation.

Data for the complex were collected over multiple sessions. Movies were motion corrected, CTF estimation was performed, particles were picked and 2D classified to isolate those of high quality. In each folder Micrographs_2016XXXX you have the pruned particle stacks.

Open a terminal with a double click on the icon
cd Prac-6/relion/
relion
This will launch the relion 3 GUI

Part (A) Preparation: Import the star file which contains CTF information and coordinates for each particle.

1. On the left-hand list of job-types go to ‘Import’
2. On the ‘I/O’ tab, Input files > Browse > navigate to Example/consensus_data.star
3. On the ‘I/O’ tab, Node type, select Particles STAR file (.star) from list

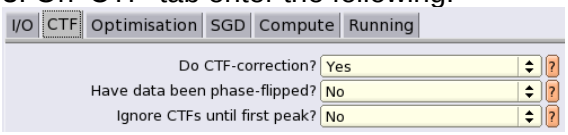
If you want to take a look at this file, open another terminal window and type the following:
cd Prac-6/relion/Example/
more consensus_data.star

Part (B) Initial Model: You can now generate an initial model using Stochastic Gradient Descent (SGD) methods.

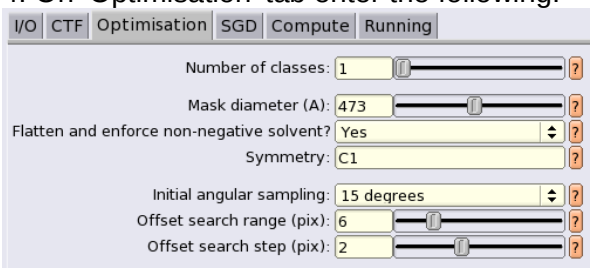
1. On the list of job-types go to ‘3D initial model’

2. On the ‘I/O’ tab, Input files > Browse > navigate to Import/job015/consensus_data.star

3. On ‘CTF’ tab enter the following:

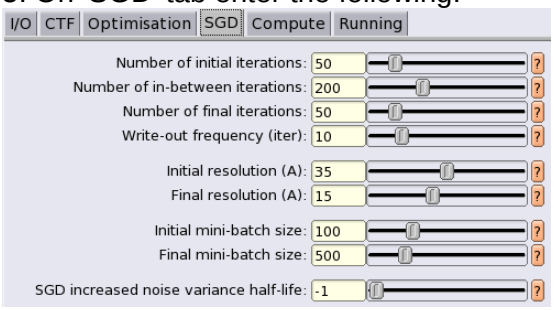


4. On ‘Optimisation’ tab enter the following:



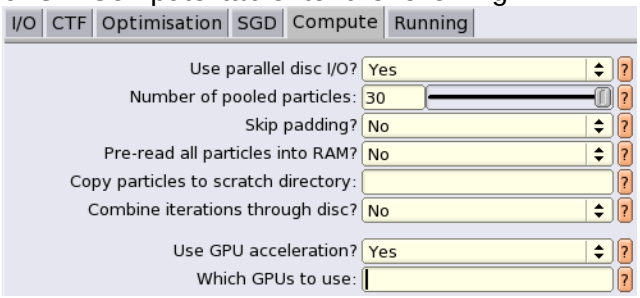
Notes:
Number of classes: Sometimes, using more than one class may help in providing a ’sink’ for sub-optimal particles that may still exist in the data set. The additional argument --sgd_skip_anneal may then also be useful. In this case, we will just use a single class in order to speed up things.
Mask diameter: Make sure this is wide enough to not cut-off any particle signal
Symmetry: If you don’t know what the symmetry is, it is probably best to start with a C1 reconstruction. Also, some higher-symmetry objects may be easier to solve by SGD in C1 than in their correct space group.

5. On ‘SGD’ tab enter the following:



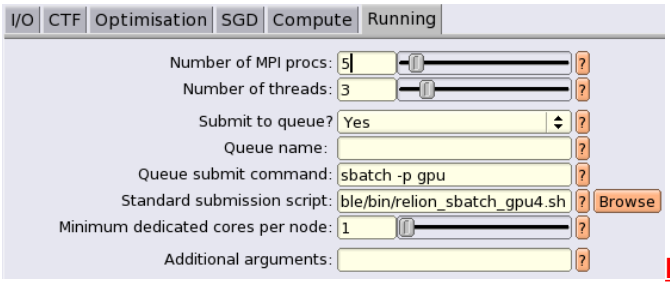
This will run for a total of 300 iterations and write out a map every 10 iterations. See ? for additional information on each value.

6. On ‘Compute’ tab enter the following:



Computing options are discussed in Part F: Multi-body refinement

7. On ‘Running’ tab enter the following:

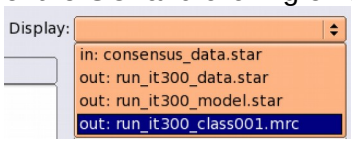


Don't hit Run!

**Standard submission script for Birkbeck:*
/s/emib/s/relion/v3.0/stable/bin/relion_sbatch_gpu4.sh

8. You would now be able to hit Run! – but don't – to save time we have pre-run it (~6 hrs on Birkbeck GPU). The results are in folder InitialModel/job002/

Analyse results:
9. Look at 2D slices through the 3D map by navigating to Display in the middle-right region of the GUI and clicking on the mrc file:



Then in the next window change the number of columns to 10 and hit Display

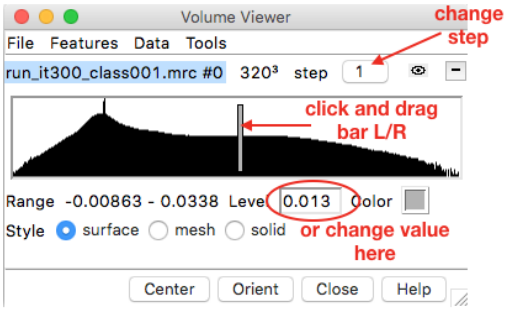
You can see regions of the structure are well defined whereas others are fuzzy/blurred and less white in intensity. Why do you think this is?

10. Look at the output volume of the final iteration with UCSF Chimera. Open a new terminal window and type chimera and open the file InitialModel/job002/run_it300_class001.mrc

• Click and drag the main window to change the size if display is small

To launch some essential controls in Chimera:
Tools > General Controls > Command Line (appears at the bottom of the display window)
Tools > General Controls > Model Panel
Tools > Viewing Controls > Side View
Tools tab > Volume Data > Volume Viewer

• You can change the threshold of the map (contour level of the map) by click and dragging the slider on the Volume Viewer or by entering a value into the Level box. High thresholds show strong regions of density, low thresholds start to show weaker regions:



Left mouse button rotates map, middle mouse button moves map, right mouse button zooms in/out
Change the step (sampling density) to 1 to display an unbinned view of the map

If you change the threshold in Chimera you can see that some regions of the molecule disappear more quickly than others at high threshold values. This tells you that portions of

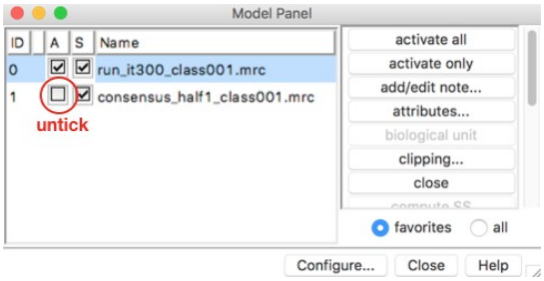
the structure are moving relative to each other (or if the structure is a complex of several proteins that some proteins may be falling off – incomplete occupancy). It is possible that the motions represent very discrete states, in this case 3D classification may separate those states and the next steps describe how you would do this, and in the case of a protein complex with partial subunit occupancy 3D classification may help sort particles of the intact complex and incomplete assemblies. It is possible, however, the motions are more continuous in nature. You would in theory need an infinite number of 3D classes to describe them. In this case multi-body refinement, that models dynamics in a macromolecular complex as being described by a discrete number of independently moving, rigid bodies would be more useful and is described in Part (F).

Part (C) Preparatory step - resampling a volume

During 3D image processing it is probable that you will need to change the orientation of a map to match that of a different map. For example, in this practical you will later be using pre-calculated models that have a particular 3D orientation. The initial model run you have just set-up generated a 3D map that has a different orientation to this (as initial Euler angle assignment in SGD is random). To see this:

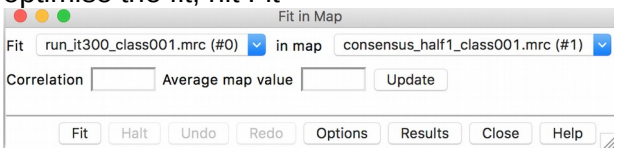
1. In Chimera open Example/consensus_half1_class001.mrc.
This is a half map from a pre-calculated consensus 3D refinement of the spliceosome dataset. The structure is in a different orientation to the SGD generated model you have just done. To get your model to match that of the consensus data (to save confusion when looking between the different maps) you need to:

2. In the Model Panel deactivate the consensus map so that only the initial model moves:

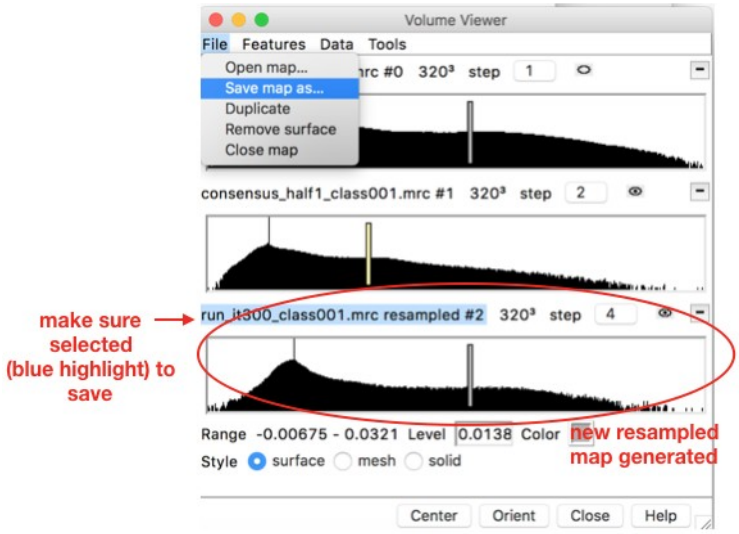


3. Rotate the initial model so that you get into roughly the same view as the consensus.
Left mouse button rotates map, middle mouse button moves map, right mouse button zooms in/out

4. In Tools > Volume Data > Fit in Map, set up the consensus map as the reference to optimise the fit, hit Fit



5. The maps are now aligned - but this alignment will not be saved to the initial model and if you type reset in the Command Line (bottom of UCSF Chimera window) the initial model will go back to it's original orientation (you can try this and then go to Tools > Movement > Undo Move to get you back to the aligned position). To save the new position you need to resample the volume in it's new orientation, in the Command Line
type vop resample #0 on #1
(#0 and #1 are the IDs of the two models, you can see this in the Model Panel under the ID column or in the Volume Viewer next to the file name for each model)
A new model appears in the Model Panel (ID #2) and Volume Viewer, save it in Volume Viewer: e.g. run_it300_class001_resampled.mrc in the parent 'relion' folder



Note: The resampling is an interpolation calculation therefore avoid doing this on high resolution maps as some information loss will occur, but as this is an initial model and will be low-pass filtered before use we do not need to worry.

Import the resampled map into the RELION framework
6. On the left-hand list of job-types go to 'Import'

7. On the 'I/O' tab, Input files > Browse > run_it300_class001_resampled.mrc

8. On the 'I/O' tab, Node type, select 3D reference (.mrc) from list

Part (D): 3D classification

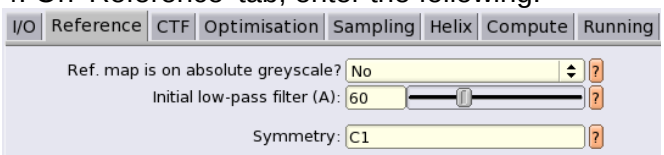
Now that you have an initial 3D model, you can use 3D classification to look for variations in the particles.

1. On the list of job-types go to '3D classification'

2. On the 'I/O' tab, Input files > Browse > navigate to Import/job015/consensus_data.star

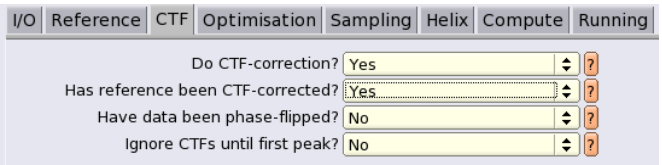
3. On the 'I/O' tab, Reference map > Browse > navigate to Import/job016/ run_it300_class001_resampled.mrc
Note: Leave Reference mask (optional) empty. This is the place where masks for focussed refinements can be added. If left empty, a spherical mask with the particle diameter given on the Optimisation tab will be used. This introduces the least bias into the classification.

4. On 'Reference' tab, enter the following:

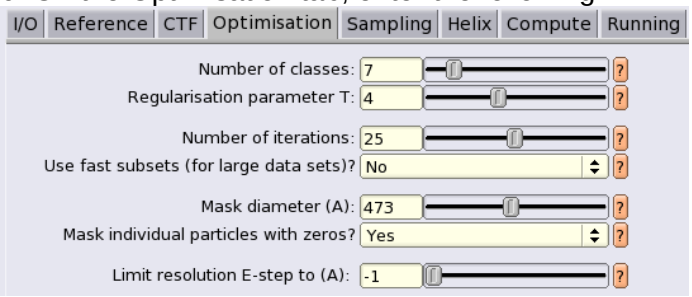


Note: as the reference map was resampled in Chimera put No for 'Ref. map is on absolute grayscale'

5. On the 'CTF' tab, enter the following:



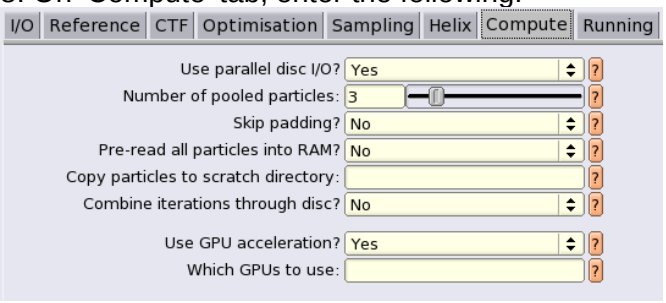
6. On the Optimisation tab, enter the following:



See ? of each parameter for a comprehensive description. For exact definition of T please refer to Sjors H W Scheres. Journal of Molecular Biology, 415(2):406–418, January 2012.

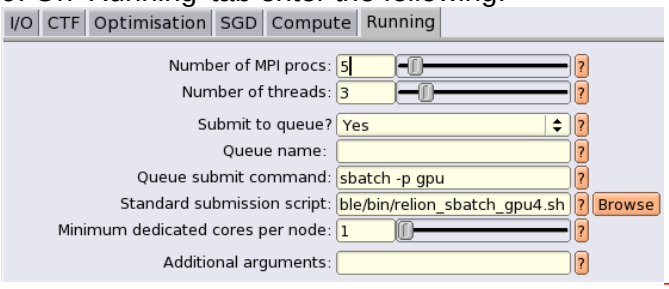
7. For ‘Sampling’ and ‘Helix’ tabs, use default values.

8. On ‘Compute’ tab, enter the following:



Computing options are discussed in Part F: Multi-body refinement

9. On ‘Running’ tab enter the following:



Don’t hit Run!

**Standard submission script for Birkbeck:*
/s/emib/s/relion/v3.0/stable/bin/relion_sbatch_gpu4.sh

You would now be able to hit Run! – don’t – it would take too long to run for the practical, we have pre-run it (~7 hrs on Birkbeck GPU). The results are in folder Class3D/job005/

Analyse results:
In the Class3D/job005 directory you can look at the run_it025_model.star file to see the class distribution and resolution of each class.
In a terminal window, navigate to /Class3D/job005 and type
nedit run_it025_model.star

Look for the data_model_classes section, you can see column 2 shows the class distribution as a fraction and estimated resolution is in column 5:

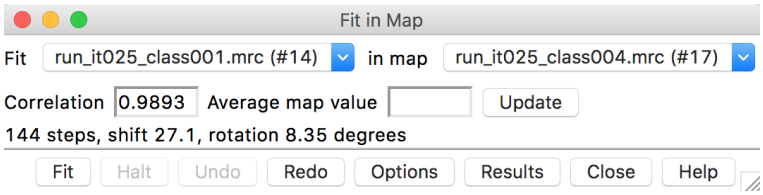
```
data_model_classes

loop_
_rlnReferenceImage #1
_rlnClassDistribution #2
_rlnAccuracyRotations #3
_rlnAccuracyTranslations #4
_rlnEstimatedResolution #5
_rlnOverallFourierCompleteness #6
Class3D/job005/run_it025_class001.mrc  0.061670  2.040000  1.353000  15.990587  0.999595
Class3D/job005/run_it025_class002.mrc  0.185721  1.451000  0.919000  11.567659  0.998216
Class3D/job005/run_it025_class003.mrc  0.113480  3.910000  2.820000  20.910768  0.999878
Class3D/job005/run_it025_class004.mrc  0.191160  1.484000  0.951000  11.819130  0.998249
Class3D/job005/run_it025_class005.mrc  0.145028  1.702000  1.126000  13.260487  0.998953
Class3D/job005/run_it025_class006.mrc  0.126411  1.492000  0.998000  12.356363  0.997067
Class3D/job005/run_it025_class007.mrc  0.176531  1.516000  0.993000  12.356363  0.998304
```

10. Open the 7 classes from the last iteration (itr. 25) in chimera (run_it025_class001-007.mrc).

11. Change the step to 1 for all volumes and the level to 0.015. A quick way is to type in the Command Line
volume all step 1 level 0.015

12. The classes will have drifted away somewhat from the starting orientation over the 25 iterations. This can make it harder to judge domain movement between the classes, therefore, it is sensible to pick a reference class and align the others to it.
• Go to Tools (main Chimera window) > Volume Data > Fit in Map.
• Align each class to the same reference structure, e.g. class004 has been used as the reference structure below:



Repeat for class002-007

13. Compare the structures. Show/hide the maps in Model Panel. You can see that some regions of the structure move between classes. It may help to switch between a tile and the superimposed view. To do this in the Command Line type savepos aligned, then type tile to switch to a tiled view. To go back to superimposed type reset aligned.
Note: savepos/reset are the commands, aligned is the name of the position and is user set.

Part (E): 3D refinement
Once you have performed 3D classification, you can use the 'Subset selection' job to select particles belonging to a single (or multiple) class/es and then perform 3D refinement on these particles. Multi-body refinement needs a 'consensus' 3D refinement so we will perform 3D auto-refine on all of the particles.

Preparatory step: We need a global mask that encompasses the entire molecule. This mask can be generated using the initial model (or a class from 3D classification) and the Mask creation job in RELION. We will practice making a mask in Part (F) and so will import a mask that has already been made.
• On the left-hand list of job-types go to 'Import'
• On the 'I/O' tab, Input files > Browse > navigate to Mask_and_Ref/global_mask.mrc
• On the 'I/O' tab, Node type, select 3D mask (.mrc) from list
Execute 3D refinement:

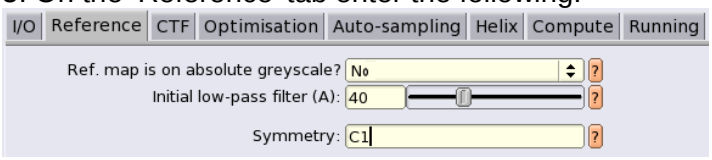
1. On the list of job-types go to ‘3D auto-refine’

2. On the ‘I/O’ tab, Input files > Browse > navigate to Import/job015/consensus_data.star

3. On the ‘I/O’ tab, Reference map > Browse > navigate to Import/job016/run_it300_class001_resampled.mrc

4. On the ‘I/O’ tab, Reference mask > Browse > navigate to Import/job017/global_mask.mrc

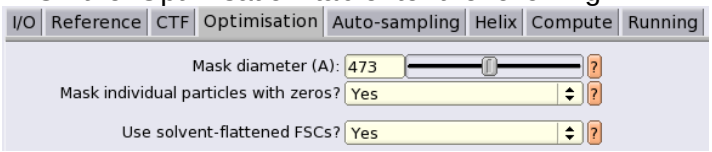
5. On the ‘Reference’ tab enter the following:



Note: as the reference map was resampled in Chimera put No for ‘Ref. map is on absolute grayscale’.

6. ‘CTF’ tab as for 3D classification

7. On the ‘Optimisation’ tab enter the following:



8. ‘Auto-sampling’ and ‘Helix’ tabs, leave defaults

9. ‘Compute’ and ‘Running’ tabs as for 3D classification **Don’t hit Run!**

10. You would now be able to hit Run – don’t – it would take too long to run for the practical, we have pre-run it (~13 hrs on Birkbeck GPU). The results are in folder Refine3D/job008/
Note: You can give each job an alias, e.g. job008 has alias consensus_3D_refine

Analyse results:
The final map is run_class001.mrc (in Refine3D/job008)
Look at 2D slices through the map using RELION Display
Loot at this map in Chimera
You can see beautiful detail of protein and nucleic acid has emerged but the map has weak blurry density in other regions.

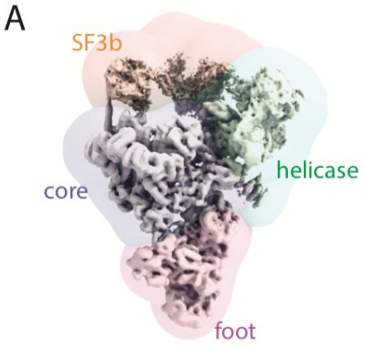
Part (F): Multi-body refinement

Protocol below has been adapted from
ftp://ftp.mrc-lmb.cam.ac.uk/pub/scheres/multibody_protocol.pdf
by Takanori Nakane & Sjors H.W. Scheres

Multi-body refinement models dynamics in a macromolecular complex as being described by a discrete number of independently moving, rigid bodies. This model builds on the observation that the structure of individual protein domains, or tertiary structure, often remains intact upon continuous changes in the overall (quaternary) structure of flexible complexes. During multi-body refinement, the signal from all-but-one body is iteratively removed from the experimental images using partial signal subtraction, and the signal from the remaining body is aligned with respect to reference projections of only that body. Thereby, alignments of the individual bodies are not compromised by differences in the

relative orientations between the bodies among the images in the data set, and reconstructions from the individual bodies are better defined than in a single refinement of the entire data set. Moreover, upon convergence of the multi-body refinement, principal component analysis on the relative orientations of all bodies for every experimental image in the data set is used to characterise the most dominant motions in the complex.

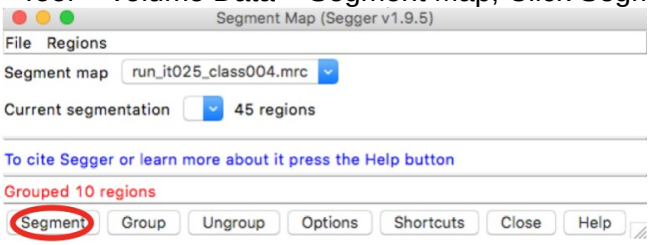
Multi-body refinement starts from a so-called consensus refinement, where all particles have been refined with respect to a single 3D reference. Use the consensus map to define a division of the complex into a discrete number of independently moving bodies (below). To ensure enough signal for accurate alignments of the individual bodies, each body should comprise a molecular weight of at least 100-150 kDa. In the spliceosome example, we chose to divide the complex into four bodies: the central core, the foot domain, the helicase domain and the SF3b domain (**Fig. 1A**)



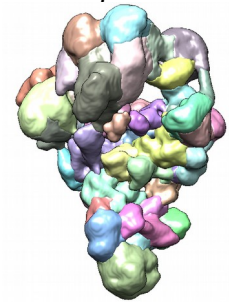
Preparatory step 1: Define a mask for each of the bodies. These can be made in Chimera (Segger or Volume eraser tool – see below). These have been pre-made and you can see them in 'Mask-and-Ref/*_mask.mrc'

Practice making a mask for the SF3b domain:

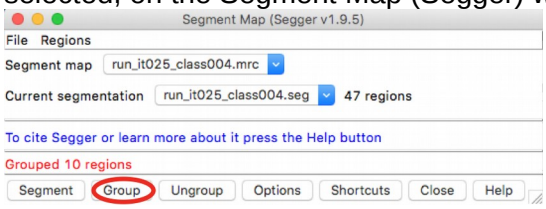
• Open a map that has good density for the SF3b domain, e.g. run_it025_class004.mrc from the 3D classification run in Chimera.
• Set threshold low e.g. 0.0096 so SF3b density nice and bulky
• Tool > Volume Data > Segment Map, Click Segment:



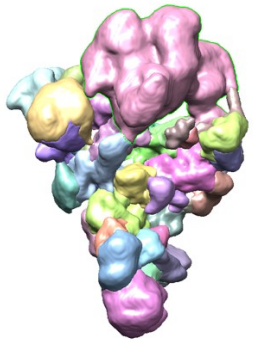
The map should be divided up into lots of multi-coloured regions: e.g.



• Using Ctrl + Shift and left mouse button, select the regions belonging to SF3b, they will get a green outline when selected (clicking again on a region will de-select it). Once all areas selected, on the Segment Map (Segger) window click Group:

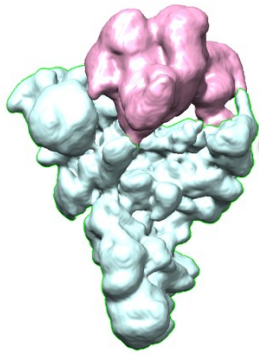


All the selected densities should now join and be the same colour: e.g.

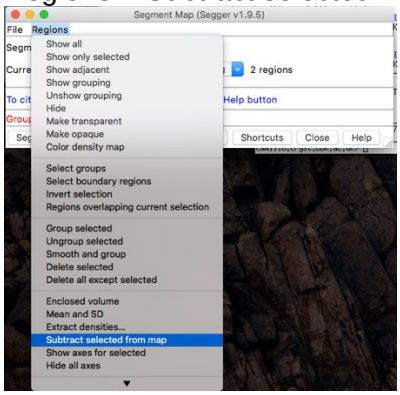


• Make sure the grouped region is selected (it should have a green outline: Ctrl and left mouse click on it if it is not). On the Segment Map window go to Regions > Invert selection. The rest of the structure should now have a green outline.

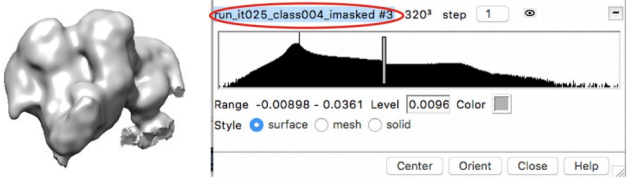
• On the Segment Map window, Group that selection. Your map should now have two colours, e.g



• Make sure the non-SF3b region is selected (has a green outline) and go to Segment Map > Regions > Subtract selected from map

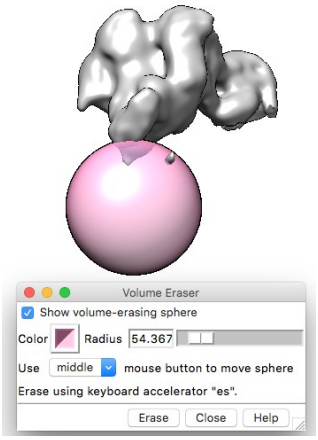


• A new map will appear in the Display window and in Volume Viewer, with an _imasked extension. Save the map: Volume Viewer > File > Same map as:



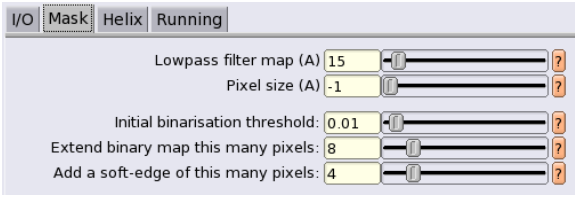
• Note the threshold where the density is well defined and there is no ‘ghost’ of the other region e.g. 0.01
Go to threshold 0.008 to see the ‘ghost’ density appear

• If needed you can also edit the map with the Volume Eraser tool (Tools > Volume Data > Volume Eraser). This can also be used to generate a mask as an alternative to Segger.



• Once you have saved your map go to the RELION GUI and Import the map as a 3D reference.

Now go to Mask creation on the jobs list in the RELION GUI
• On the ‘I/O’ tab, Input 3D map > Browse > navigate to your imported map
• On the ‘Mask’ tab, keep ‘lowpass filter’ and ‘pixel size’ as default and change the initial binarisation threshold to the value you noted down earlier. Set the ‘Extend binary map this many pixels’ to a generous amount as the domain is moving and you don’t want to cut any density off, e.g. 8 pixels and soft edge to e.g. 4 pixels. Helix and Running leave default. **Hit Run.**



The calculation will take about 10 minutes so progress to the next steps but you can come back and look at your mask and compare to the pre-made one in /Mask-and-Ref/SF3b_mask.mrc when it’s done.

Preparatory step 2:
Determine order of bodies. Earlier bodies are subtracted first so it’s best to place larger and more stable bodies earlier in the list.
Spliceosome body order: core > foot > helicase > SF3b

Preparatory step 3:
Determine rotation, which body relative to which other body it rotates. Spliceosome: we chose to have the foot, the helicase and the SF3B domains rotate relative to the central core domain. The central core domain was chosen to rotate relative to the second largest body, i.e. the foot domain

Preparatory step 4:
Determine the width, i.e. the standard deviation, of the Gaussian priors on the relative rotations and translations between pairs of bodies. Rotational and translational searches of the individual bodies will be performed locally, with a search range of three times these standard deviations. Therefore, these values should express the expectation by how much the bodies move relative to each other in the data set. For the spliceosome example, the fuzziness in the consensus structures indicated that the core and the foot are relatively stable, whereas the helicase, and especially the SF3b domains move more. We thus used widths on the angular priors of 10 degrees for the core and foot, 15 degrees for the helicase domain, and 20 degrees for the SF3b domain. Likewise, the widths on the translational priors were set to 2 pixels for the core and foot, to 3 pixels for the helicase and 5 pixels for the SF3b domain. For each project, these values should be determined empirically, but the values given here provide a useful starting point.

Put the information from preparatory steps 1–4 into a .star file, the spliceosome dataset is shown here:

In a terminal change directory to the Example folder:
cd /Example/
nedit 4-bodies-tight-mask.star

```
data_
loop_
_rlnBodyMaskName
_rlnBodyRotateRelativeTo
_rlnBodySigmaAngles
_rlnBodySigmaOffset
_rlnBodyReferenceName
Mask-and-Ref/CORE_mask.mrc 2 10 2 Mask-and-Ref/CORE_ref.mrc
Mask-and-Ref/FOOT_mask.mrc 1 10 2 Mask-and-Ref/FOOT_ref.mrc
Mask-and-Ref/HELICASE_mask.mrc 1 15 3 Mask-and-Ref/HELICASE_ref.mrc
Mask-and-Ref/SF3b_mask.mrc 1 20 5 Mask-and-Ref/SF3b_ref.mrc
```

Execute multi-body refinement

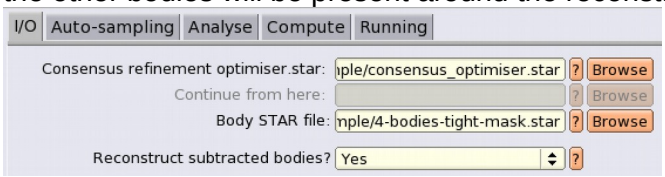
Note: You could use the results from your auto 3D refinement run as the consensus refinement, but the steps below are using a pre-calculated run found in Example/

5. On the left-hand, vertical list of job-types click on '3D multi-body'

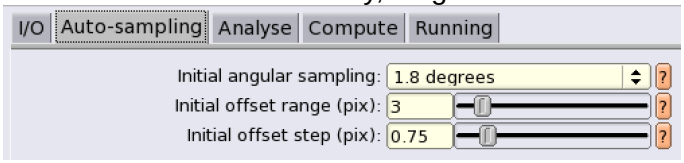
6. On the 'I/O' tab, set 'Consensus refinement optimiser.star' to 'Example/consensus_optimiser.star'.
Within a typical RELION project, the optimiser.star file is directly visible through the 'Browse' button, but this is not the case in this simplified example. Therefore, just type its location in the GUI text entry.

7. On the 'I/O' tab, set 'Body STAR file' to 'Example/4-bodies-tight-mask.star'. This is the STAR file that defines the bodies, which was made above.

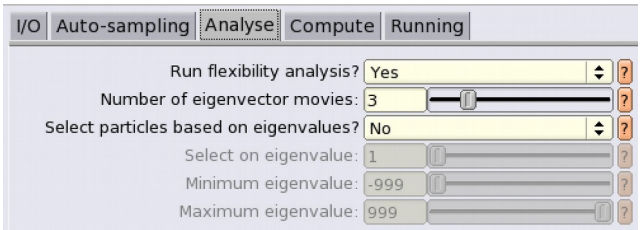
8. On the 'I/O' tab, set 'Reconstruct subtracted bodies' to 'True'. This will make the program write out reconstructions for each of the bodies where the density of the other bodies is subtracted. Thereby, the density around each body should become ever cleaner, and also provide feedback on the subtraction quality, as the alignments of the bodies improve during the subsequent iterations of multi-body refinement. When set to 'False', fuzzy densities for the other bodies will be present around the reconstructed density for each body.



9. On the 'Auto-sampling' tab, leave the default of 'Initial angular sampling: 1.8 degrees', 'Initial offset range (pix): 3' and 'Initial offset step (pix): 0.75'. In general, one sets the sampling rates to somewhat coarser values than the estimated accuracy as reported in the 'model_classes' table in the model.star file from the consensus refinement. For the spliceosome case, the values are 0.786 degrees for the angular accuracy and 0.535 pixels for the translational accuracy, as given in the file 'Example/consensus_half1_model.star'.



10. On the 'Analyse' tab, leave the defaults. 'Run flexibility analyses? ‘Yes' will perform the principal components analysis on the relative orientations of all bodies upon convergence. 'Number of eigenvector movies: 3' will select the first three principal motions in the data, and write out ten maps for each of them. 'Select particles based on eigenvalues? ‘No' will switch off the option to select particle subsets based on eigenvalues.



11. On the 'Compute' tab, set 'Use parallel disc I/O?' to 'Yes' if you have a fast, parallel storage system, or if you are working on a local hard drive. Set this option to 'No' if you are using a less parallel, shared storage system like NFS.

12. On the 'Compute' tab, set 'Number of pooled particles: 30' if you are using GPUs. Otherwise, use a smaller value like 3.

13. On the 'Compute' tab, set 'Skip padding? ‘Yes', as this will save RAM and GPU memory by up to eight-fold. Only set this option to 'No' if your particle occupies such a large volume of the box that artefacts are observed near the edges of the box in runs where this option was set to 'Yes'.

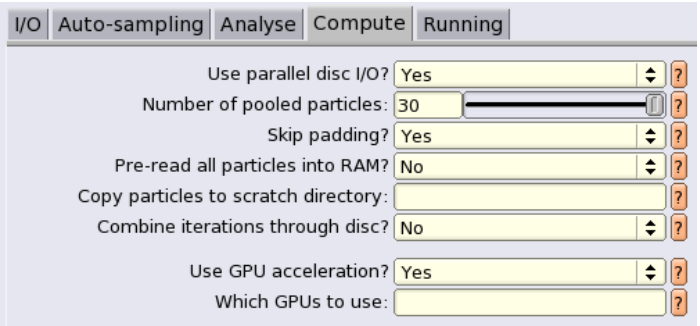
14. On the 'Compute' tab, set 'Pre-read all particles into RAM? ‘No', as the example data set occupies more than 120 GB, and each MPI slave would read the full data set into memory.

15. Only set this option to 'Yes' if your data set is small enough to fit into your computer's RAM.

16. On the 'Compute' tab, set 'Copy particles to scratch directory’ to a path for the scratch drive if you have one, leave blank here. If your data set is small enough to read into RAM (previous step), then copying the data to a scratch drive is not necessary.

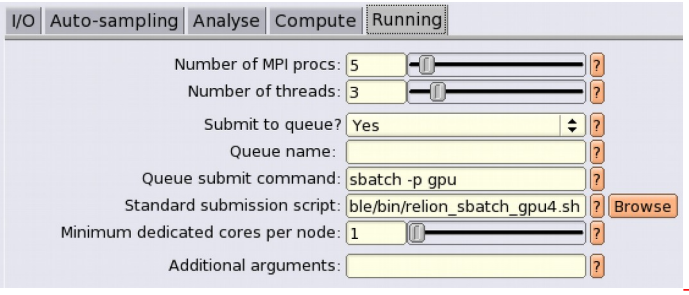
17. On the 'Compute' tab, set 'Combine iterations through disc? ‘No'. Only set this option to 'Yes' if the default option of combining iterations through the network gives problems, e.g. MPI-related errors at the end of an iteration.

18. On the 'Compute' tab, set 'Use GPU acceleration? Yes', since the program will run a lot faster on GPUs than on CPUs.



You can also define which GPUs to use, e.g. on a computer with four cards, can use '0,1:2,3', where the ':' sign is used to divide tasks among different MPI slaves and the ',' sign is used to divide tasks within a single MPI slave. If 3 MPI processes are used, i.e. one master and two slaves, GPUs 0 and 1 jointly do the work of the first slave, and GPUs 2 and 3 divide the work of the second slave.

19. On the 'Running' tab, enter the following:

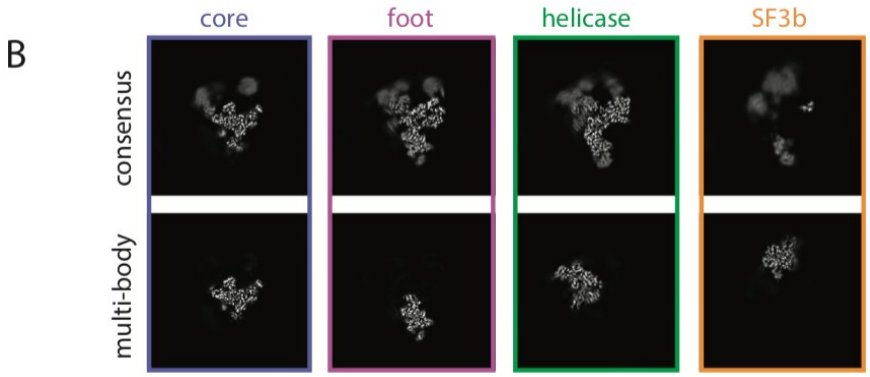


Don't hit Run!

You would now be able to hit Run – don't – it would take too long to run for the practical, we have pre-run it (~46 hrs on Birkbeck GPU). The results are in folder Multibody/job011/

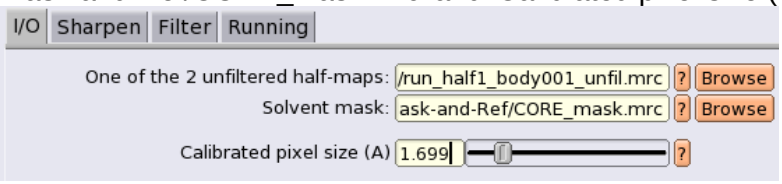
Analyse the body reconstructions

1. The multi-body refinement program writes out a reconstructed map that is filtered to its estimated resolution for each body as 'Multibody/job001/run_body00[1-4].mrc'. These maps can be directly visualised in a molecular volume viewer, or in slices from the 'Display!' pull-down menu on the GUI. **Fig 1B**, shows a comparison of regions from the consensus refinement and multi-body run.



Post-processing:

2. Besides the filtered maps, the program also writes out unfiltered maps for both independently refined half-sets. On the 'I/O' tab of the 'Post-processing' job-type, input the first half-map of the first body: 'One of the 2 unfiltered half-maps': 'Multibody/job001/run_half1_body001_unfil.mrc'. Also set 'Solvent mask: Mask-and-Ref/CORE_mask.mrc' and 'Calibrated pixel size (A) 1.699'.



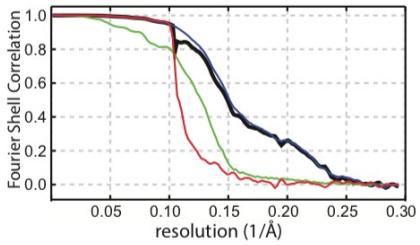
On the 'Sharpen' tab one can provide a STAR file with the modulation transfer function of the detector. Click Browse and find file K2_mtf.star

Leave all other options on the 'Sharpen', 'Filter' and 'Running' tabs to their defaults and execute the calculation by pressing the 'Run!' button.

Repeat **steps 2-4** for the remaining three bodies.

3. Inspect the Fourier Shell Correlation (FSC) plot (**Fig. 1C**) in the 'logfile.pdf' from the 'Display!' pull-down menu for each of the four bodies. Confirm that the red FSC curve is close to zero at the resolution estimate of each body. If this is not the case, their resolution may be under-estimated due to your masks being too tight or too sharp. In that case, repeat multi-body refinement with wider and/or softer-edged masks.

C

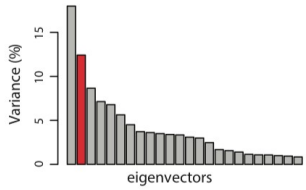


4. The post-processed maps of the four bodies obtained in **steps 2-5** can be opened in molecular volume viewers, and used for separate atomic modelling of the individual bodies, for example in COOT.

Analyse the body motions

1. Select 'MultiBody/job011' in the 'Finished jobs' window on the lower part of the GUI. Use the 'Display:' drop-down menu to visualise the 'analyse_logfile.pdf' file. The first plot (**Fig. 1D**) shows how much variance in the orientations is explained by each of the principal components. The subsequent plots show histograms of the projections of the relative orientations of the bodies of each experimental projection image onto the corresponding component. Mono-modal histograms are an indication of continuous motions in the complex, whereas bi- or multi-modal histograms indicate the presence of two or more discrete states. In the spliceosome example, all histograms are mono-modal, but in other cases it may be useful to select subsets of particles based on their projections along the principal components.

D



2. Make a movie that represents this first principal motion. Ten maps called 'analyse_component001_bin0??.mrc' are generated by rotating and translating each body along each principal motion. In Chimera, from the 'Tools' menu, select 'Volume Data' and then 'Volume Series'. Click on the 'Open...' button and browse to the directory 'Multibody/job001'. To make a movie for the first principal component, hold down the SHIFT key, and use the left-mouse button to select all ten maps called 'analyse_component001_bin0??.mrc'. Then, change 'Play direction' to 'oscillate' and press 'Play'. You can change the sampling rate and threshold of the maps through the 'Volume Viewer' tool, or by using the 'Command Line' from the 'Favourites' menu, and typing: 'vol all step 2' and 'vol all level 0.015'.

3. Write a movie to file by selecting from the 'Tools' menu 'Utilities' and then 'Movie recorder'. Repeat **steps 2-3** for the second and third components. **Fig. 1E** shows the third and eighth maps of the second principal motion (which is highlighted in red in Fig.1D).

E