



An easy-to-follow tutorial for anyone new to *cisTEM*

Introduction

Welcome to *cisTEM*! *cisTEM* is user-friendly software to process cryo-EM images and obtain high-resolution 3D reconstructions from them. Its name derives from “computational imaging system for transmission electron microscopy” and it comprises a number of tools to process image data including movies, micrographs and stacks of single-particle images. It implements a complete “pipeline” of processing steps to obtain high-resolution single-particle reconstructions. The tutorial should take about two hours to complete (depending on your workstation) and takes you from movies of frozen-hydrated apoferritin particles to a 3-Å reconstruction. We assume that you know basic concepts in cryo-EM, including movie mode data, exposure filtering, contrast transfer function (CTF), Fourier shell correlation (FSC), low-pass and high-pass filtering, and B-factor sharpening. If you are unfamiliar with these terms we recommend reading some of the basic texts in the cryo-EM field. We further assume that you do not want to read extensive texts before you can get going with your own data, so we will keep this tutorial to the essentials. For more information, please read the help pages included in *cisTEM* or consult the online documentation at cistem.org.

Installation

We have optimized *cisTEM* to run on CPUs. GPUs are currently not supported because we found in tests that CPU-optimized code runs as fast on a modern multi-core CPU as GPU-optimized code runs on a modern GPU. This means that you can expect fast processing without the need for specialized hardware (recompiling *cisTEM* on your local machine may produce slower code, see below). It also means that the speed with which *cisTEM* can process data depends on the number of CPUs available. *cisTEM* has its own parallelization scheme that is independent of common architectures such as openMPI. Furthermore, it can run on a local workstation and utilize high-performance computing environments (e.g. a computer cluster) by remotely executing compute jobs. All that is required is that all machines (local workstation, cluster login and compute nodes) have access to the same file system (i.e. paths to data files are the same everywhere) and that the login node for the cluster is accessible via ssh.

To install *cisTEM*, we recommend downloading one of the compiled versions of *cisTEM* from cistem.org to take advantage of optimizations that may not be available when compiling on a local machine. Select the archive that is appropriate for the architecture used and, after downloading it to a local disk, unpack it using the command line

```
tar -xzf cistem-<version>-<architecture>-<OS>.tar.gz
```

This will create a directory called `cistem-<version>` that contains all the necessary programs to run *cisTEM*. Type the command line

```
cistem-<version>/cisTEM
```

This should start the *cisTEM* graphical user interface (GUI), indicating that *cisTEM* runs on your system. Close the *cisTEM* windows and, optionally, include the `cistem-<version>` directory in your `PATH` environment variable for executables. Then, after logging out and back in, you should be able to start *cisTEM* by typing

```
cisTEM
```

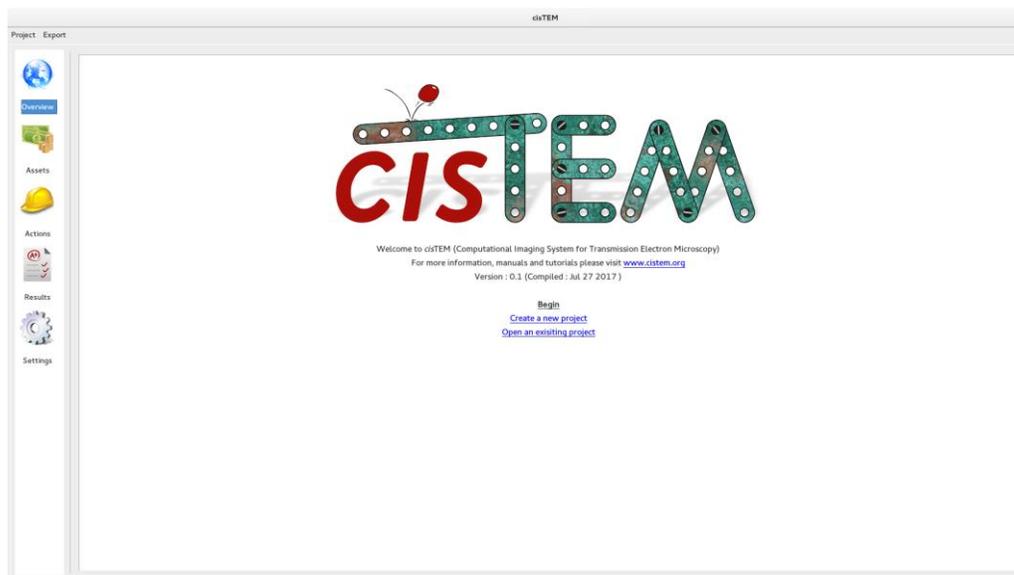
on the command line. If you did not set the `PATH` variable, you have to provide the full path to your *cisTEM* executable. If everything was installed correctly, this will start the *cisTEM* GUI.

The GUI

Once installed, you can launch the *cisTEM* GUI by typing

```
cisTEM
```

on the command line.



The GUI is organized into “Assets” and “Actions.” Assets are data files (Movies, Images), results of processing steps (Particle Positions, 3D Volumes) and user-specified “Refinement Packages.” The different Asset types will be displayed when clicking on “Assets.” Actions are processing steps that can be applied to Assets, producing results and new Assets that can be processed further. You can access different Actions by clicking on “Actions.” Once results are obtained, you can display them by clicking on “Results.” You can also change the way *cisTEM* runs by clicking on “Settings.” If installed correctly, Settings will list a number of “Run Profiles”, each specifying a workstation and/or computer cluster to run *cisTEM* on. For the tutorial, we will run *cisTEM* on your local workstation using the “Default Local” profile that is available if other profiles were not defined. For this, you do not need to make changes in “Settings.”

Start a New Project

For the tutorial, you will need to download the apoferritin test dataset from the *cisTEM* web page. It comes as another archive called `apoferritin_data.tar.gz` and is attached to the “Getting Started” Documentation page. Store this archive on a fast local disk where you have about 15 GB of space and unpack it by typing

```
tar -xzf apoferritin_data.tar.gz
```

This will create a new directory called `apoferritin_data` that contains 20 movies, recorded on an aberration-corrected Titan Krios microscope (300 kV, $C_s = 0$, effective pixel size = 1.5 Å, exposure/frame = 2 e⁻/Å²). The movies are the only data required for this project.

Now launch *cisTEM*. On startup, the GUI presents a list of previously opened projects, as well as options to create a new project or open an existing project. To continue a previous project, click on the provided link. For the tutorial, we start fresh and click on “Create a new project.” A dialog will be displayed:

In this tutorial, you will calculate a 3-Å reconstruction of apoferritin, so enter “Apoferritin” as the project name. You must also provide a path to the directory where your new project will reside. For best results, this should be on a fast local disk, for example the disk you chose to store and unpack the apoferritin dataset. Enter the path directly or use the “Browse” button to navigate to the directory. Click “Finish” to create the project. *cisTEM* keeps track of files and results in a database (SQLite) that is stored in a file called `<project name>.db` inside the project directory (here `Apoferritin`). Because file names are stored in the database with their absolute paths, the project directory and database should not be moved once created. If it is moved, *cisTEM* will attempt to update the paths to files within the project directory. However, paths to imported files will not be updated and, therefore, these files will have to remain in their original location to remain accessible within *cisTEM*.

Importing Movies and Images

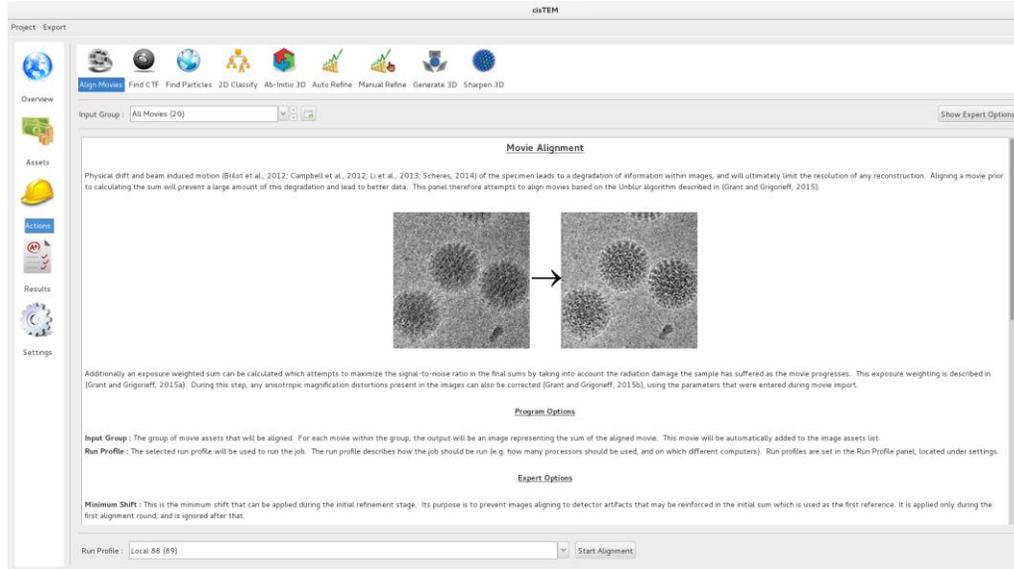
Once a project is open or has been newly created, Assets can be imported. These will usually be Movies or Images but can also be Particle Positions, 3D Volumes and Refinement Packages. For the tutorial, you should

import the movies from the test dataset. Click on Assets, then Movies and Import. This will open a dialog:

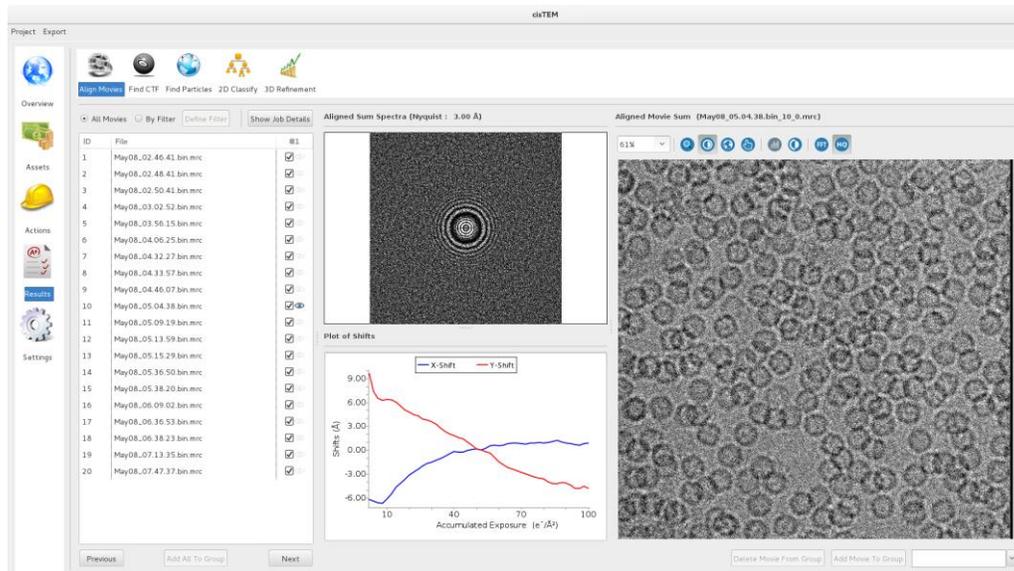
In the dialog, select “Add Directory” and navigate to the directory containing the 20 apoferritin movies. Then click “Open.” This should list all the movies in the directory. Now enter the required information describing the data (see figure above): Voltage (kV): 300; Spherical Aberration (mm): 0.0; Pixel Size (Å): 1.5; Exposure per frame ($e^-/\text{Å}^2$): 2.0. There are also options to import compressed tiff files and separate gain references, to change the pixel size (binning) and apply magnification distortion on the fly. For the tutorial, these options should remain unchecked. Click “Import” and *cisTEM* will show a list of the 20 imported movies. The movies are all part of a group called “All Movies.” Additional groups can be created using “Add” to select subsets of a dataset for further processing. You should continue with all the data for now. If images are available instead of movies, these can be imported as Image Assets in the same way as Movies, by clicking “Images.”

Movie Alignment

Movie data collection and frame alignment have been part of the single-particle image processing pipeline since it was first introduced by Brilot et al. in 2012. The original software Unblur was developed further by Grant & Grigorieff (2015) when exposure weighting was added to take into account the radiation-dependent signal loss when adding movie frames, yielding signal-optimized frame sums. *cisTEM* implements the Unblur algorithm in the Align Movies panel, which also provides some background to the method. Click “Actions” and select “Align Movies” to call up the panel:



Actions panels display parameters that you can change. Some of these are shown on the main panel while others are only accessible when “Show Expert Options” is selected. Movie alignment usually works with the default parameters and you should simply click “Start Alignment” near the bottom of the panel. You will notice that next to the start button a menu is shown that allows you to select different run profiles. The Local profile should be selected by default but can be changed to other profiles if these were previously set up under Settings. For the tutorial, you should run all steps locally on your workstation. On a 16-CPU workstation, the alignment of all 20 movies takes less than a minute. While the job is running, X,Y traces are displayed for some of the movies and a progress bar indicates the time left until completion of the job. After termination (you must click on “Finish” at the end of all jobs), you can inspect the results by clicking “Results” and selecting “Align Movies”:

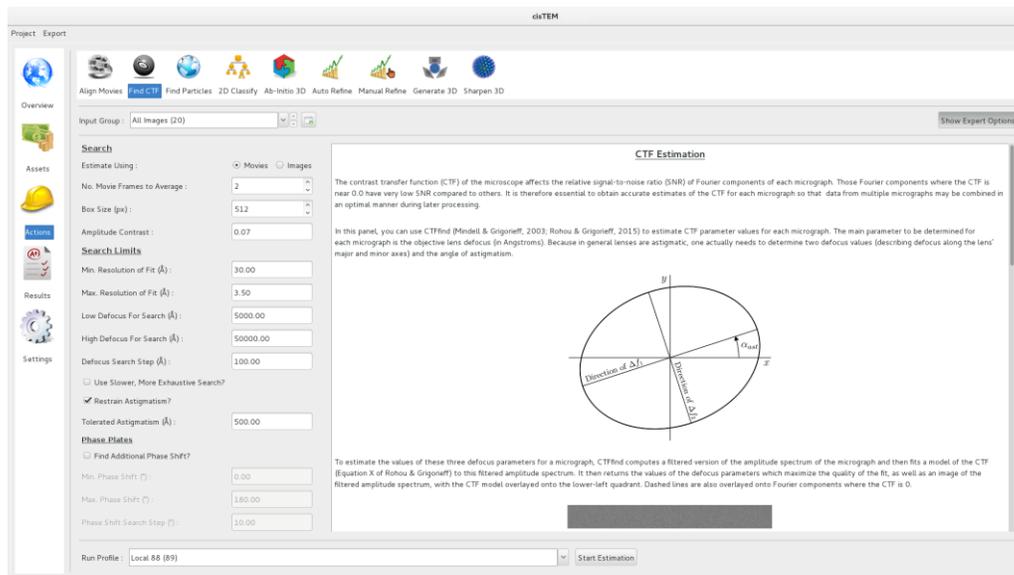


For each movie, the X,Y trace and the power spectrum and image of the final aligned and exposure-filtered frame sum are shown. You can move between results from different movies by clicking on the eye symbols in the list (next to the ticks), by clicking “Previous” and “Next”, and by using the “P” and “N” keys on the keyboard. The ticks cannot be selected/deselected until movie alignment is run additional times. Most Actions

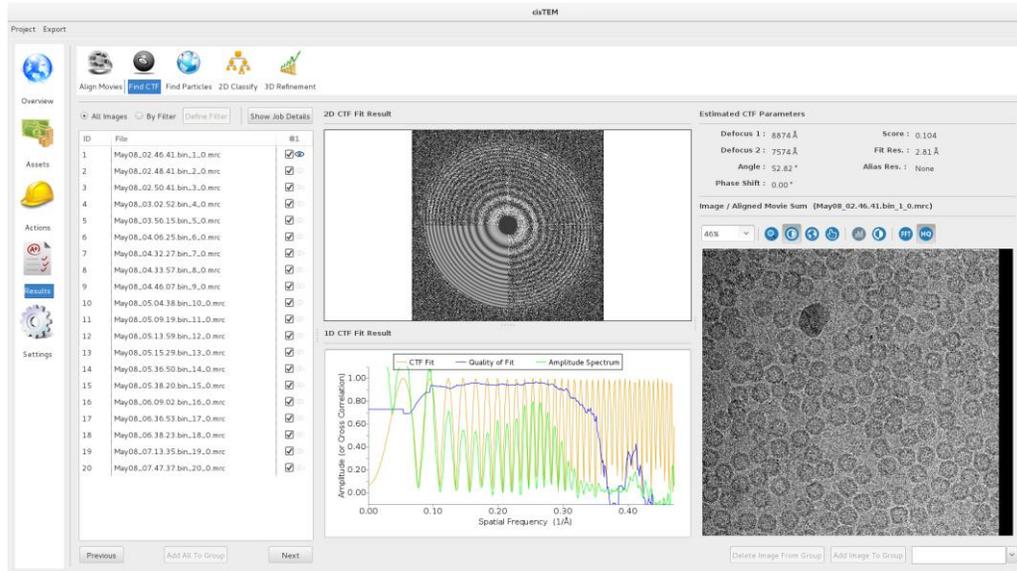
in *cis*TEM can be re-run to test different sets of parameters. *cis*TEM stores the results for each run and allows you to select the best results using the tick boxes in the Results panels. There is also a “Show Job Details” button to display the parameters used for a particular job (try it by clicking on it, then click it again to hide it). For the tutorial, keep the current results and scroll through them to spot movies that did not align well. The apoferritin dataset should include two movies that yield blurry images and/or power spectra with poor Thon rings. You can add successfully aligned movies to a new Assets group but for the tutorial you should delay the selection of good data until the next step where you will determine the CTFs of each image. Finally, Align Movies produced new Image Assets, which can be listed by clicking “Assets”, then “Images.”

CTF Estimation

The image defocus can be estimated by fitting a CTF to the Thon rings visible in an image power spectrum. CTFFIND is a commonly used tool to fit CTFs and has been implemented in *cis*TEM using the latest version developed by Rohou & Grigorieff (2015). Click on “Actions”, then “Find CTF” to access the panel. Here you have the option to estimate the CTF from the movie frames or the summed frames (Images). Click on “Show Expert Options” to see all parameters:



Estimation from Movies is selected by default and usually gives the best results. The algorithm will average a few frames at a time to reduce noise in the power spectra used for the fit. The number can be adjusted and defaults to 3. As a rule of thumb, select the number of frames that amount to a combined exposure of $4 e^{-}/\text{Å}^2$ (Rosenthal & Henderson, 2003). In the case of the apoferritin dataset, this equals 2 frames. You may also want to adjust the assumed amplitude contrast (default is 0.07, i.e. 7%), resolution range to be considered in the fit and search range. Additionally, for phase plate data, a phase shift can be estimated. For the tutorial, you should only adjust the Max. Resolution of Fit to 3.5 Å , then click “Start Estimation” near the bottom of the panel. The job should take less than a minute to complete. As before, the results can be found in the Results panel:



You should inspect all the results using the buttons and “P”/”N” keys as before. For each Movie/Image, the power spectrum (this time without exposure filter) and fit are shown as 2D and 1D plots, as well as the image itself. A useful additional plot shows a “Quality of Fit” that indicates to what resolution Thon rings were detectable. The detection limit is displayed as the Fit Resolution and can be taken as a rough measure of image quality. 19 out of the 20 apoferritin images should have a Fit Resolution better than 3.5 Å. To sort the results according to the Fit Resolution (and other criteria), click on “By Filter.” This will show a dialog with many options of which we will select the “Detected Fit Res.” in the Filter By window:

Filter / Sort Images (on bern)

Filter By :-

Asset ID : 0 To 0

Estimation ID : 0 To 0

Date of Run : 09/26/17 To 09/26/17

Job ID : 0 To 0

Defocus : 0.0 To 0.0

Astigmatism : 0.0 To 0.0

Astigmatism Angle : 0.0 To 0.0

Score : 0.0 To 0.0

Sort By :-

Asset ID Estimation ID Date of Run

Job ID Defocus Astigmatism

Astigmatism Angle Score Detected Fit Res.

Detected Alias Res. Phase Shift Voltage

Spherical Aberration Pixel Size Amplitude Contrast

Box Size Min. Resolution Max. Resolution

Min. Defocus Max. Defocus Defocus Step

Tolerated Astigmatism Min. Phase Shift Max. Phase Shift

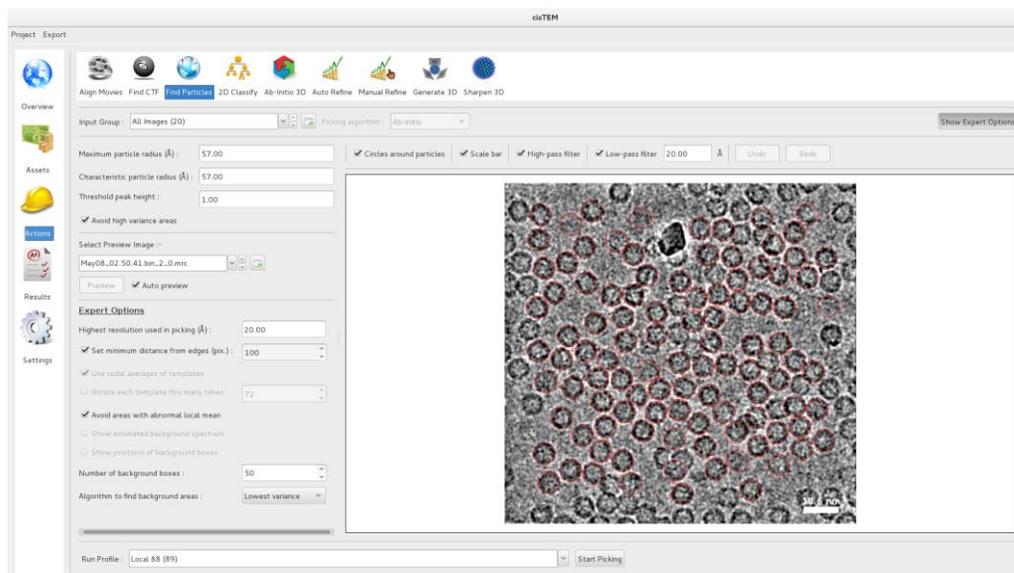
Phase Shift Step

Cancel Filter/Sort

Applying the sorting filter will move the worst images to the end of the list. To create a new Assets group on the 19 good images, click “Assets”, then “Images” and “Add.” Rename the new group to “Bad Images”, then switch back to the CTF estimation results, select the eye symbol next to the bad image, then click “Add Image To Group” (the “Bad Images” group should be shown next to the button). This should be repeated with all images that show poor Thon rings. For the tutorial dataset, only image #16 should be excluded. Switch back to Assets/Images and select the “Bad Images” group. Now click “Invert” to generate a new group with all the good images and rename it “Good Images.” Images can also be added and removed within the Assets panel by drag/drop, clicking “Add To Group” and “Remove.” The good images are now ready for the next step: particle picking.

Particle Picking

Finding particles in EM images used to be a labor-intensive step, costing many hours or days of someone’s time to click on potential candidates. The development of good algorithms to pick particles is still ongoing and different strategies exist. Some algorithms are based on templates to identify particles. While they can be quite accurate depending on the quality of the templates, they also bear the danger of template bias, i.e. the selection of noisy areas of the images that do not contain particles but will reproduce the templates when aligned and averaged. Furthermore, template-based algorithms are computationally expensive because each image has to be searched with each template, usually in different rotational views. The current version of *cisTEM* implements a faster algorithm, termed “Ab initio”, that is based on an algorithm described by Sigworth (2004) and uses a simple low-pass filtered disk with a radius close to the average particle radius to search for matching features in the images. It excludes areas of low and abnormally high variance to avoid areas that have no particles and areas that contain high-contrast features that are not particles (ice contamination, carbon). Apart from the speed, this algorithm also largely avoids the problem of template bias since no specific templates are used.



To get started, click “Actions”, then “Find Particles.” For the Input Group, select “Good Images”, then check the radio button “Auto preview.” This will show picking results for the currently selected image and since the picking algorithm is fast, these results are displayed almost instantly. Scroll through a few of the images to get a

feeling for how well your parameters work. When previewing, you can enable additional low-pass filtering that will make it easier to see particles and some display features can be turned on and off. Selected particles are indicated by red circles that have the size of the maximum particle radius. Using preview mode, you should try different values for the maximum and characteristic particle radii, as well as threshold for the peak to be accepted as a particle. The maximum particle radius determines how closely spaced particles can be picked while the characteristic radius determines the size of the disk used for picking. Clicking “Show Expert Options” will reveal additional parameters and you should also play with “Highest resolution used in the picking.” In previous runs of the tutorial, the following values worked well: Maximum radius = Characteristic radius = 50 Å, Threshold peak height = 4.0, High-resolution limit = 15.0 Å, “Avoid high variance areas” enabled. Later, you will generate a particle image stack from the picked coordinates with a box size of 192 x 192 pixels. You should therefore also make sure that the number displayed in “Set minimum distance from edges” is not too small. A value of 100 works well. When all parameters are set, click “Start Picking.”

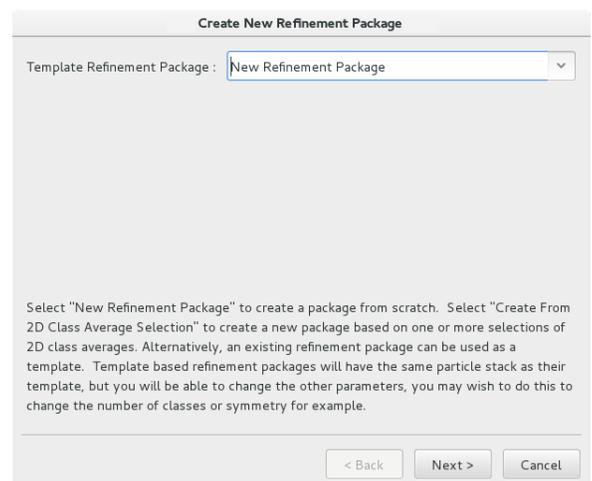
The job should run for only a few seconds and results will be viewable under Results. The display offers the same features as the preview mode. You can now select additional particles or delete particles that you deem unacceptable by left-clicking with the mouse in empty areas or red circles, respectively. In addition to mouse clicks, you can press the Ctrl key and rubberband an area to delete all the selected particles in it. Finally, pressing the Shift key, the mouse pointer turns into an eraser and running it across an image will delete all particles it encounters. If you are uncertain about one set of picking parameters, you can switch back to Actions/Find Particles and re-run picking with a different set of parameters. You can then select the best results for each image by using the ticks in the file list. Using a combination of automatic picking and manual editing, you should be able to annotate a dataset in a relatively short time. For the tutorial, you do not need to edit the particle picking results manually as apoferritin is easy to pick and the automatic picking results are sufficient to obtain a 3-Å structure. Bad particles will be filtered out in the next step, 2D classification. Using the parameters listed above, about 2000 particles should be picked and added as new Assets to Particle Positions.

2D Classification

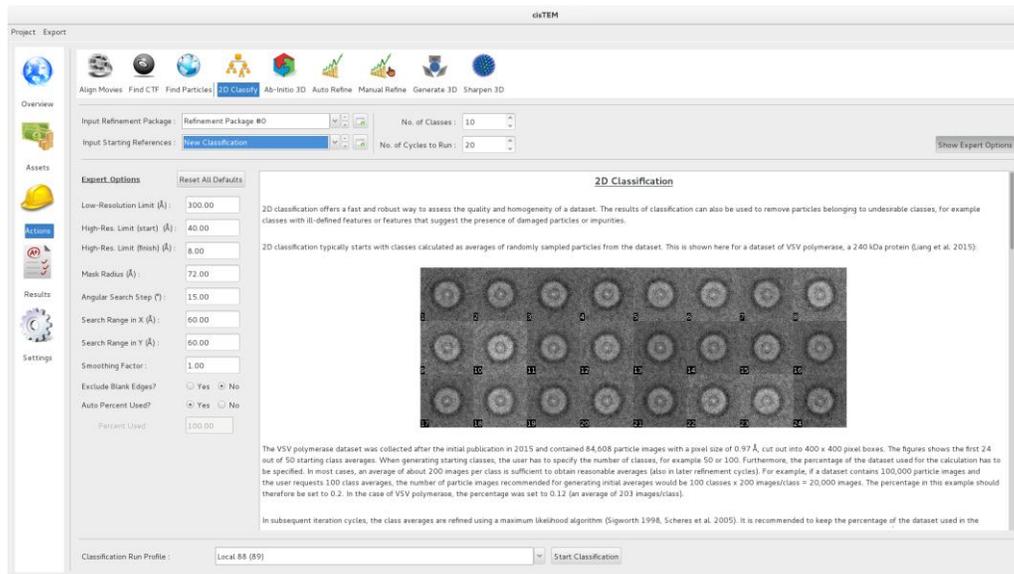
Classification of single particle images is often used as a “cleaning” step to remove damaged particles and contaminants. It is also a quick way to assess the quality of a dataset. For a homogeneous set of particles, class averages should show well-defined views, ideally with high-resolution features. *cisTEM* implements a maximum-likelihood algorithm that is based on the original work by Sigworth (1998) with further developments for multiple references by Scheres et al. (2005).

Before you can run 2D classification, you need to create a Refinement Package, which is a type of Asset and simply defines a combination of particle image stack, alignment parameters and a 3D reference structure. Click on “Assets”, then “Refine Packages” and “Create.” A wizard will take you through the creation process.

Select “New Refinement Package” for the template, “All

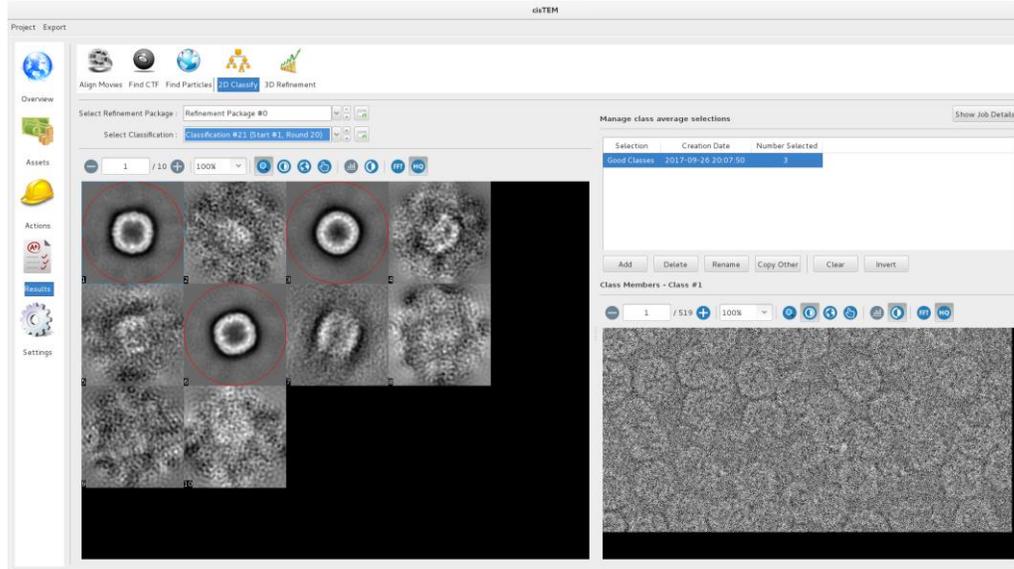


Positions” for the found particles, “440.0” for the particle weight (apoferritin), “120.0” for the particle diameter, “192” for the box size used to cut the particles out of the images, “0” for octahedral symmetry, “1” for the number of requested classes (this number does NOT refer to the number of 2D classes), and “Generate From Params.” for the 3D reference. Finally, click “Finish” and *cisTEM* will start cutting out the particles to generate a stack. This will only take a few seconds because only about 1900 particles were picked. For larger datasets, this can take up to a few minutes. You will see a new Asset called “Refinement Package #0” that you can rename if you wish. If, at a later stage, you would like to change the 3D reference associated with a Refinement Package you can select a new reference with a double-click on the listed reference inside the “Active 3D References” section towards the lower right.

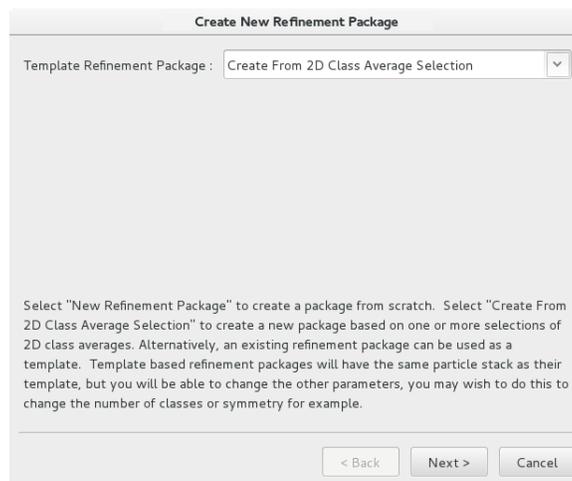


You are now ready to start 2D classification. Select “Actions”, then “2D Classify” and choose the Refinement Package you would like to use (for now, only the one you just created will be available). You can also select starting references but again, for now only the “New Classification” option is available. The default for the number of classes depends on the size of the dataset and should default to 5. Try changing this to 10 to make room for a few more views of the particle that will be represented in the class averages. Keep the default number of cycles to run (20). There are other parameters to change under “Show Expert Options” but you should leave this unchanged for now. Click “Start Classification” to run your job.

The classification of the small apoferritin dataset will take only a few minutes. You will see updated class averages after every iteration, giving you a sense of how the classification progresses. You will also notice plots of Image Mobility, Likelihood and average Sigma values (Sigma indicates the reciprocal of the estimated signal-to-noise ratio in the particle images) that grow with every iteration. For successful classification, you should see decreasing Sigma values. Image Mobility will remain high (between 25 and 50%) until the final iterations because the algorithm initially uses only a small changing fraction of the data in each iteration.



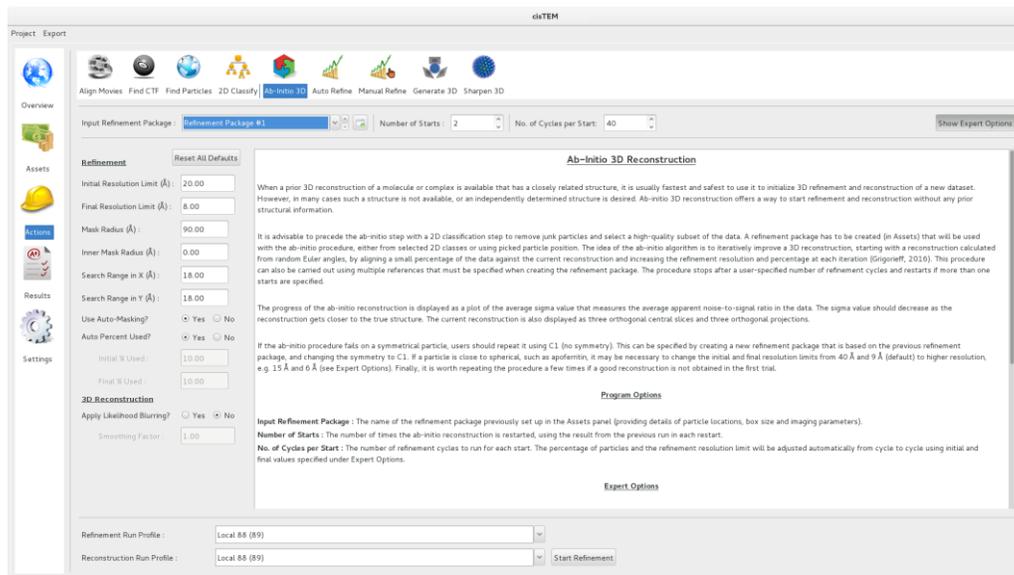
The classification results can be inspected in the Results section, “2D classify.” Using the right mouse button, you can select different class averages and see the particles assigned to this class in the lower right panel. Create a selection of good class averages by clicking “Add” and renaming the selection to “Good Classes” using the “Rename” button or double-clicking on it. Now, using the left mouse button, select good class averages. Since there are only about 2000 particles, it may be best to exclude only averages that are clearly “junk.” You can now switch back to “Assets” and “Refine Packages” and create a new package using the particles in your Good Classes. Click on “Create” and in the wizard, select “Create From 2D Class Average Selection” for the template:



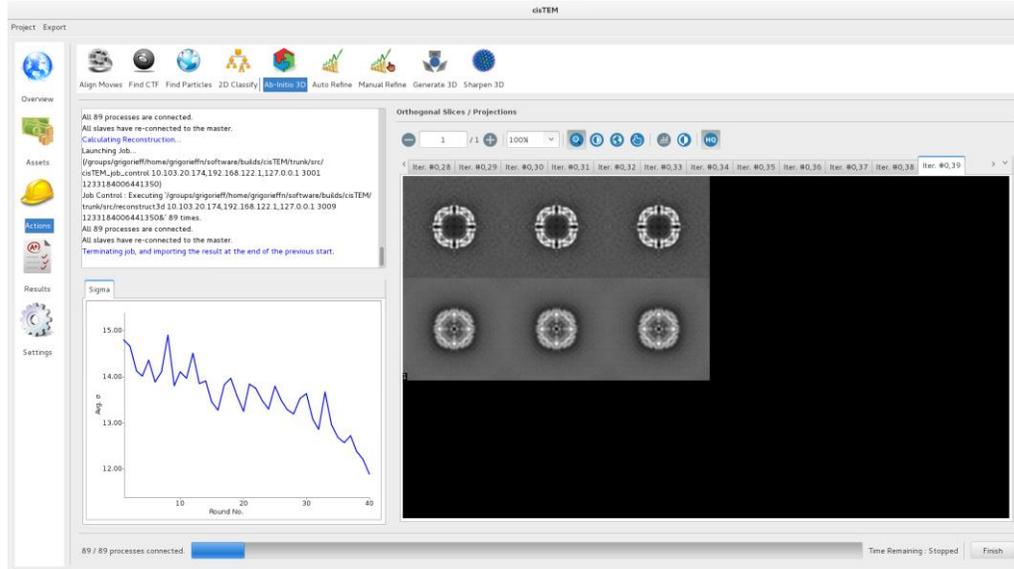
Clicking “Next”, you will see your “Good Classes” selection. Select it and click “Next.” The remaining questions should be answered as before (box size = 192 pixels, molecular weight = 440 kDa, largest dimension = 120 Å, symmetry = “O”, number of classes = 1, initial reference = Generate From Parameters, Finish). When the Refinement Package has been created, check how many particles it contains and compare this with your first Refinement Package. Excluding the particles in your bad classes should have reduced the particle number by about 100 in your new Refinement Package. You now have a “cleaned” particle stack and are ready to obtain an ab-initio 3D reconstruction.

Ab-Initio 3D Reconstruction

Obtaining an initial reconstruction is an important step to get right because the initial structure serves as a reference for initial particle alignment. If it contains significant errors (artifacts), these may bias the alignments and get reproduced in subsequent refinement cycles, preventing convergence of the refinement onto the correct structure. One way to obtain an initial reconstruction is to find a closely related structure in a database (PDB, EMDB) and use this for the initial alignment (after adjusting the pixel size and dimension of the reconstructed volume). If such a structure is not available or you think that there are major differences to your particle, such as conformational differences or additional/missing subunits, you can generate an initial reference ab initio. Please keep in mind, however, that the handedness of an ab-initio structure is undetermined and might have to be corrected later on (see section on Finalizing Reconstruction).



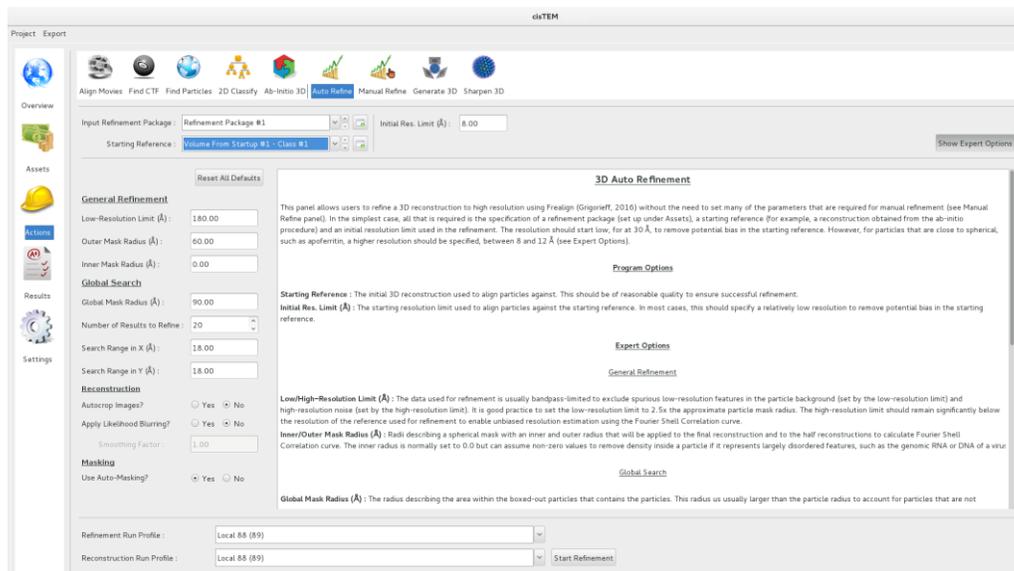
The algorithm implemented in *cisTEM* was first described by Grigorieff (2016) and uses multiple “Starts” of refinements, each performing a user-defined number (usually 40) of global particle alignment parameter searches at successively increasing resolution, starting from random alignment parameters. Go to the “Ab-Initio 3D” action panel, select the correct Refinement Package (this is the one you created from the Good Classes in the previous step) and click “Start Refinement.” The calculation will take about 15 minutes to complete one Start with 40 refinement cycles. The default is to run two Starts. In the case of this apoferritin dataset, one Start is sometimes sufficient and it should always converge to the correct structure after two Starts. You can follow the evolution of the 3D reconstruction while the job is running. After each refinement cycle, orthogonal central sections through the volume (top), as well as orthogonal projections (bottom) are shown that will display increasingly detailed features of the structure. There is also a plot of Sigma values that will initially show quite variable values from cycle to cycle but should decline on average as the refinement converges. You can terminate the calculation at any point in time by clicking “Terminate Job” or “Take Current.” After the first Start has finished, you can also click “Take Last Start”, the option you should select if the reconstruction from the first start displays clear secondary structure (i.e. α -helices in the case of apoferritin):



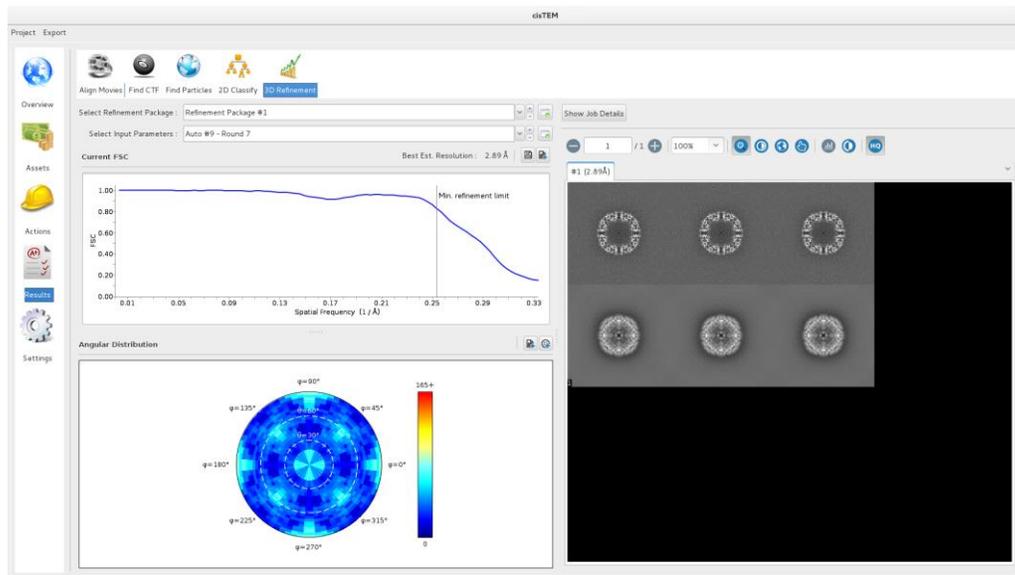
The 3D reconstruction from the first Start will then be added as a new Asset to 3D Volumes. Click “Assets”, then “3D Volumes” and “Display”, or double-click on the volume name (it should be called “Volume From Startup #1 – Class #1”) to display sections through the reconstruction. If the ab-initio step fails to produce a correct reconstruction, you should run it again. All files containing Volume Assets are contained in the `Assets/Volumes` directory inside the project directory, stored in MRC format that can also be displayed using software such as UCSF Chimera (Pettersen et al., 2004). The ab-initio reconstruction can be used to align the complete dataset and refine it to high resolution.

3D Refinement

cisTEM offers two ways to refine a 3D reconstruction, Auto Refine and Manual Refine. The Manual Refine panel exposes most of the control parameters necessary to customize a refinement using FrealignX, the latest version of Frealign (Grigorieff, 2016). For this tutorial, you should use simplified Auto Refine panel:



It implements procedures that make use of Frealign but sets most of the control parameters automatically. For apoferritin, only the **“Initial Res. Limit (Å)”** parameter has to be changed from its default (30 Å) to 8 Å. This is necessary because apoferritin does not have clearly recognizable features at low resolution and starting refinement at 30 Å resolution would lose most of the features contained in the ab-initio reconstruction. After setting “Initial Res. Limit (Å)”, click “Start Refinement.” After a few refinement iterations (about 7), refinement should be completed, resulting in a 3-Å reconstruction. The result can be viewed in the 3D Refinement Results panel:



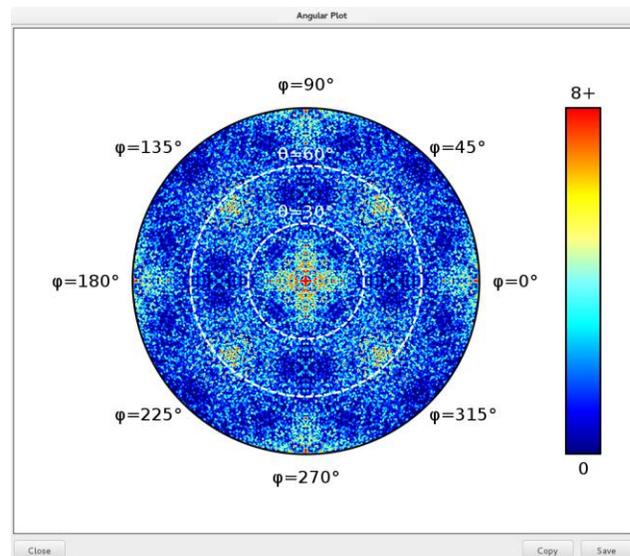
The Results panel displays the FSC curve together with the resolution limit used during the last refinement iteration, an angular distribution plot and orthogonal central sections through, and projections of the map. The FSC and angular distribution plots can also be displayed in more detail and saved by clicking on the small icons above the respective plots:

FSC Info

Class 1 - Estimated Res = 2.89 Å (Refinement Limit = 3.95 Å)

Shell	Res. (Å)	Radius	FSC	Part. FSC	√Part. SSNR	√R
2	288.00	0.01	1.000	1.000	10.585	
3	144.00	0.01	1.000	1.000	17.515	
4	96.00	0.02	1.000	1.000	12.062	
5	72.00	0.02	1.000	1.000	5.904	
6	57.60	0.03	1.000	1.000	1.864	
7	48.00	0.03	1.000	1.000	1.262	
8	41.14	0.04	1.000	1.000	1.898	
9	36.00	0.04	1.000	1.000	1.241	
10	32.00	0.05	0.999	0.999	0.611	
11	28.80	0.05	1.000	1.000	0.977	
12	26.18	0.06	0.999	0.999	0.541	
13	24.00	0.06	0.996	0.998	0.343	
14	22.15	0.07	0.996	0.998	0.374	
15	20.57	0.07	0.996	0.997	0.378	
16	19.20	0.08	0.992	0.995	0.301	
17	18.00	0.08	0.992	0.995	0.309	
18	16.94	0.09	0.995	0.997	0.415	
19	16.00	0.09	0.996	0.998	0.544	
20	15.16	0.10	0.995	0.997	0.611	
21	14.40	0.10	0.994	0.996	0.625	
22	13.71	0.11	0.996	0.998	0.691	
23	13.09	0.11	0.998	0.998	0.705	
24	12.52	0.12	0.997	0.998	0.589	
25	12.00	0.12	0.996	0.998	0.519	
26	11.52	0.13	0.995	0.997	0.499	
27	11.08	0.14	0.994	0.996	0.488	
28	10.67	0.14	0.994	0.996	0.528	
29	10.29	0.15	0.994	0.996	0.532	
30	9.97	0.16	0.994	0.996	0.452	

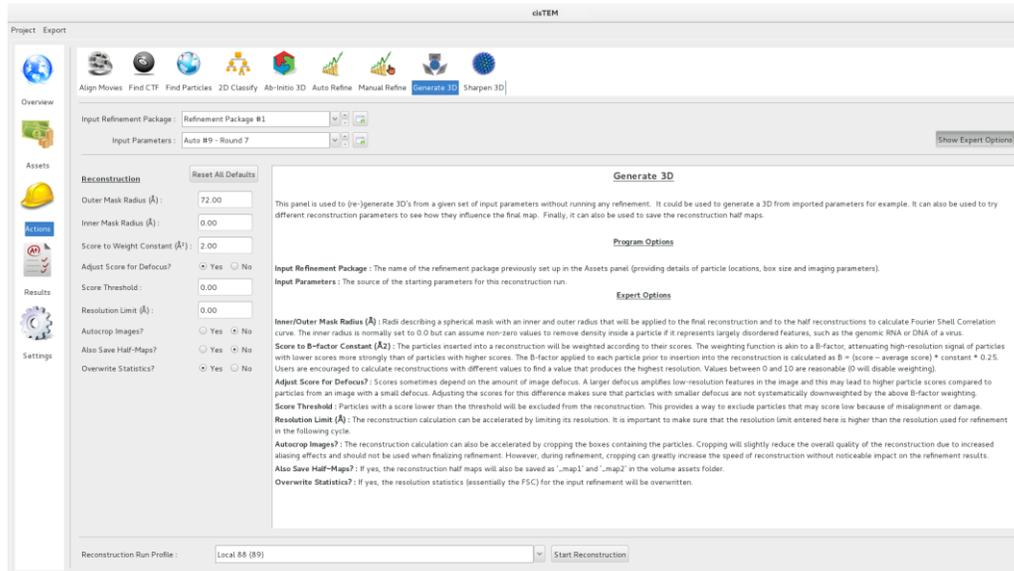
Buttons: Close, Copy All, Save



The MRC-formatted reconstructions can again be found in the Assets/Volumes directory inside the project directory. To see a full list of Frealign control parameters used for the refinement iteration that generated a

particular Volume, click on “Show Job Details.” This is particularly useful when using Auto Refine since many of the parameters are set automatically but could be informative to a more experienced user.

If the half-maps used to calculate the FSC are required, you must recalculate the reconstruction(s) using the “Generate 3D” panel:



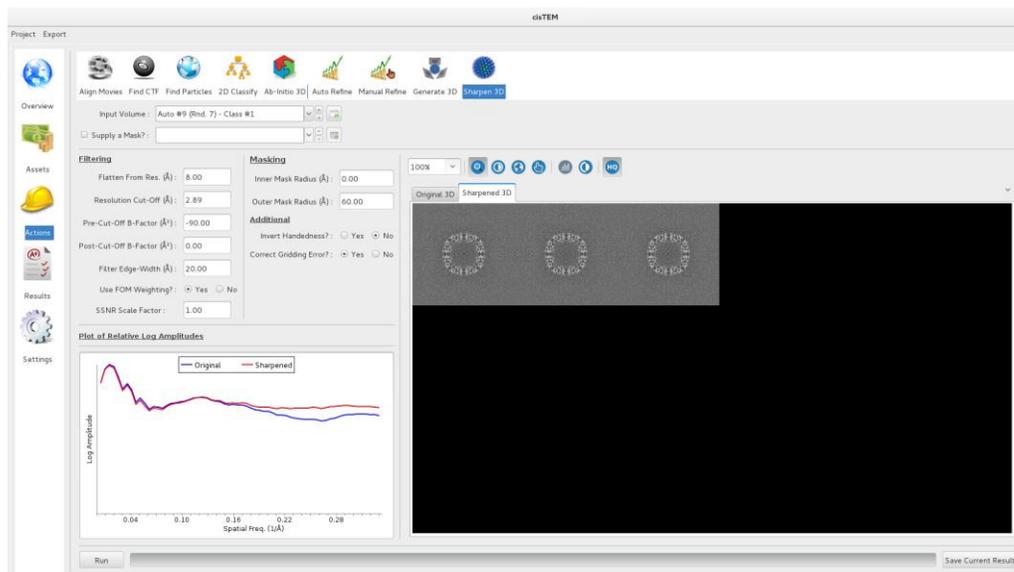
The panel recalculates the map and provides an option to save the half-maps as new Volume Assets where they are accessible for further analysis.

Please note that the Run Profiles used for refinement and reconstruction can be different. For the tutorial, a single profile can be used but if larger jobs are run on computer clusters, it may be necessary to define separate profiles in the “Settings” panel. Refinement can be efficiently parallelized and run on hundreds of CPUs. However, parallelization of the reconstruction process requires significant disk space per CPU used and therefore, using more than 100 CPUs for a reconstruction job is rarely advisable due to excessive disk I/O and potential disk quota limitations. A typical selection of Run Profiles for a job that runs on a cluster may look like this:

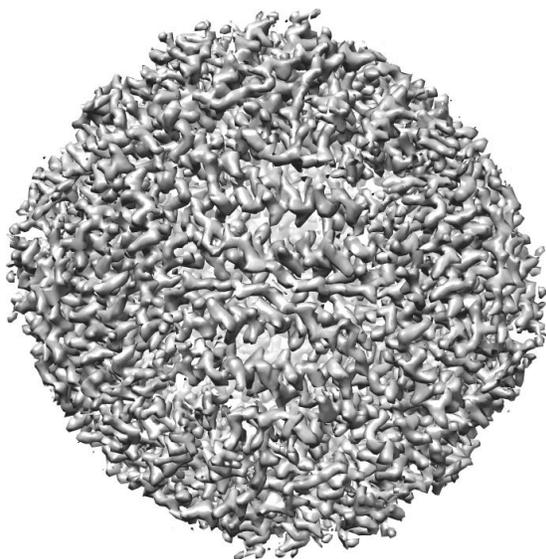


Finalizing Reconstruction

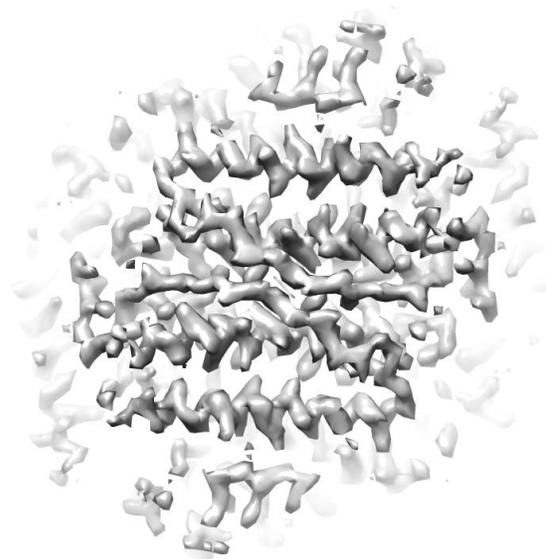
Once refinement is complete, the map requires a few additional corrections. These include correction of interpolation artifacts (gridding), B-factor sharpening and optional figure-of-merit (FOM) filtering (Rosenthal & Henderson, 2003), and inversion of handedness if necessary (the handedness of a reconstruction cannot be determined in the ab-initio procedure). All correction steps can be performed using the “Sharpen 3D” panel. After selecting the Volume to correct (the final reconstruction from Auto Refine), click “Run” (in the lower left corner). After a few seconds, orthogonal central sections of the original and sharpened maps are displayed, as well as amplitude spectra (on a log scale) of the original and sharpened maps:



Although the default settings usually yield good results, you can play with different parameters. For example, you could set the “Post-Cut-Off B-factor” to a negative value, e.g. -30 \AA^2 , to amplify the high-resolution terms, or you could set the inner mask radius to 30 \AA . To save the sharpened map, click “Save Current Result” and enter the location for the MRC-formatted map to be saved. Sharpened maps are not automatically added to the project as Volume Assets. You should display the sharpened map using UCSF Chimera (Pettersen et al., 2004) or similar to check that the handedness is correct and repeat the sharpening with reversed handedness if necessary. The final sharpened apoferritin map should show clear side chain density:



Apoferritin at 3 \AA resolution



Zoomed view showing side chains

References

- Brilot, A. F., Chen J. Z., Cheng, A., Pan, J., Harrison, S. C., Potter, C. S., Carragher, B., Henderson, R., Grigorieff, N. (2012) Beam-induced motion of vitrified specimen on holey carbon film. *J. Struct. Biol.* **177**, 630-637.
- Grant, T. & Grigorieff, N. (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *eLife* **4**, e06980.
- Grigorieff, N. (2016) Frealign: An exploratory tool for single particle cryo-EM. *Methods Enzymol.*, **579**, 191-226.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004) UCSF chimera – a visualization system for exploratory research and analysis. *J. Comp. Chem.* **25**, 1605–1612.
- Rohou, A. & Grigorieff, N. (2015) CTFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221.
- Rosenthal, P. B. & Henderson, R. (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745.
- Scheres, S. H. W., Valle, M., Rafael Nuñez, R., Sorzano, C. O. S., Marabini, R., Herman, G. T. & Carazo, J.-M. (2005) Maximum-likelihood multi-reference refinement for electron microscopy images. *J. Mol. Biol.* **348**, 139–149.
- Sigworth, F. J. (1998) A maximum-likelihood approach to single-particle image refinement. *J. Struct. Biol.* **122**, 328–339.
- Sigworth, F. J. (2004) Classical detection theory and the cryo-EM particle selection problem. *J. Struct. Biol.* **145**, 111–122.