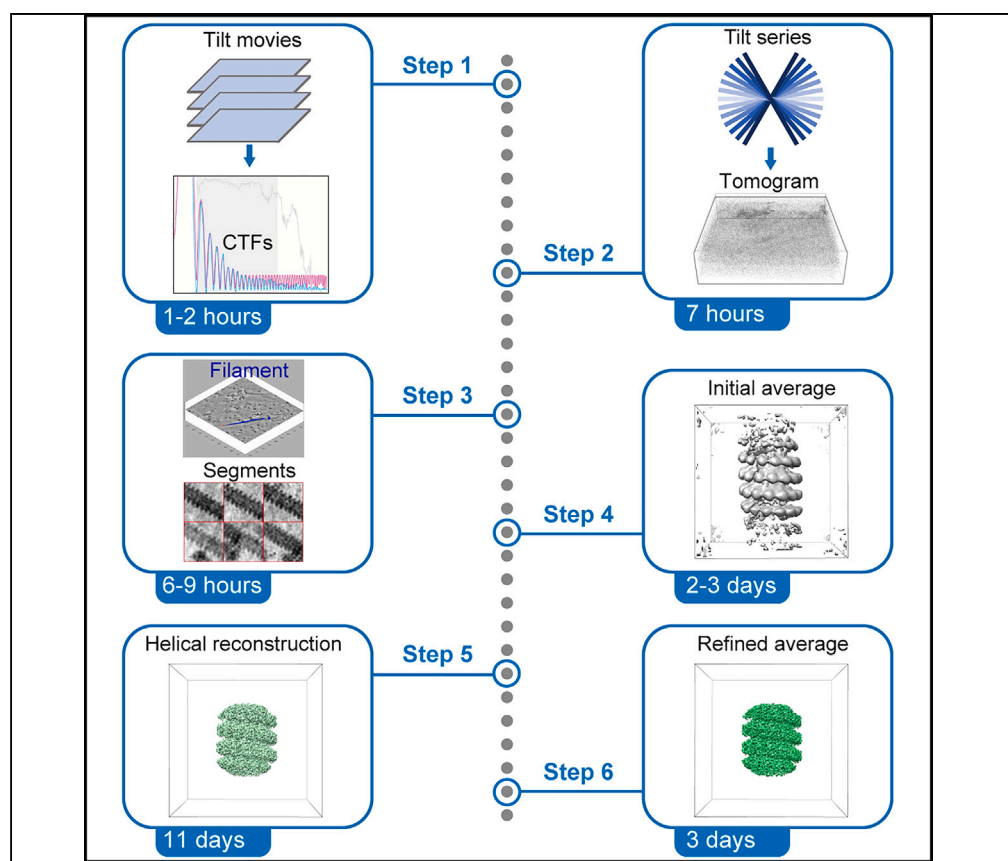


## Protocol

# Protocol for subtomogram averaging of helical filaments in cryo-electron tomography



Xiaojie Zhang, Julia Mahamid

zhangxj2@shanghaitech.edu.cn (X.Z.)  
julia.mahamid@embl.de (J.M.)

### Highlights

Processing of cryo-ET data using Warp, IMOD, Dynamo, RELION, and M

Transferable processing and analysis pipeline from raw movie frames to 3D maps

Generation of molecular-resolution subtomogram averages for filamentous assemblies

Unraveling of structures of helical viral nucleocapsids isolated from host cells

Helical filaments are essential macromolecular elements in cellular organization and dynamics. Recent advances in cryo-electron tomography allow faithful imaging of isolated or in-cell filaments. Here, we present a protocol to generate density maps at sub-nanometer resolution of helical filaments by subtomogram averaging, exemplified with isolated mumps virus nucleocapsids and their in-cell form as an extension of the protocol. We detail procedures from pre-processing of tilt-series movie frames to refinement of reconstructed averages for streamlined data processing of helical filaments.

**Publisher's note:** Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Zhang & Mahamid, STAR  
Protocols 5, 103272  
September 20, 2024 © 2024  
The Author(s). Published by  
Elsevier Inc.  
<https://doi.org/10.1016/j.xpro.2024.103272>



## Protocol

# Protocol for subtomogram averaging of helical filaments in cryo-electron tomography

Xiaojie Zhang<sup>1,3,4,\*</sup> and Julia Mahamid<sup>1,2,5,\*</sup>

<sup>1</sup>Structural and Computational Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Baden-Württemberg, Germany

<sup>2</sup>Cell Biology and Biophysics Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Baden-Württemberg, Germany

<sup>3</sup>Present address: iHuman Institute, School of Life Science and Technology, and Shanghai Key Laboratory of High-resolution Electron Microscopy, ShanghaiTech University, Shanghai 201210, China

<sup>4</sup>Technical contact

<sup>5</sup>Lead contact

\*Correspondence: [zhangxj2@shanghaitech.edu.cn](mailto:zhangxj2@shanghaitech.edu.cn) (X.Z.), [julia.mahamid@embl.de](mailto:julia.mahamid@embl.de) (J.M.)  
<https://doi.org/10.1016/j.xpro.2024.103272>

## SUMMARY

Helical filaments are essential macromolecular elements in cellular organization and dynamics. Recent advances in cryo-electron tomography allow faithful imaging of isolated or in-cell filaments. Here, we present a protocol to generate density maps at sub-nanometer resolution of helical filaments by subtomogram averaging, exemplified with isolated mumps virus nucleocapsids and their in-cell form as an extension of the protocol. We detail procedures from pre-processing of tilt-series movie frames to refinement of reconstructed averages for streamlined data processing of helical filaments.

For complete details on the use and execution of this protocol, please refer to Zhang et al.<sup>1</sup>

## BEFORE YOU BEGIN

The protocol below describes the specific steps required for determining structures of mumps virus nucleocapsid as an example of heterogeneous biological filaments with helical symmetry (raw micrographs in EMPIAR: 10751 and intermediate processing files in Mendeley Data <https://doi.org/10.17632/d95fvg4dnm.1>). The nucleocapsids were isolated from HeLa cells persistently infected with mumps virus, following 1 h of sodium arsenite stress treatment, imaged by cryo-electron tomography (cryo-ET) and computationally analyzed by subtomogram averaging (Figure 1). We have also used this protocol with slight modifications for determining the structure of nucleocapsids inside cells after 6 h of potassium arsenate or sodium arsenite treatment.<sup>1</sup>

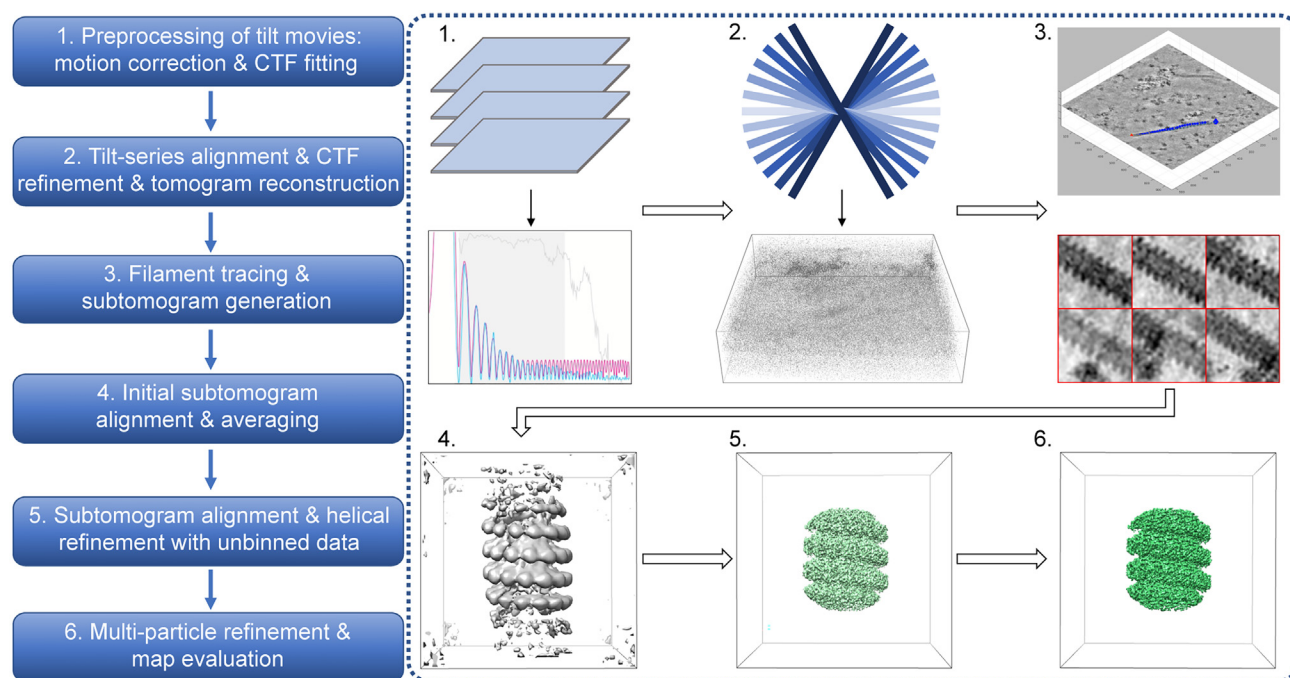
## Preparation of cryogenically preserved isolated nucleocapsids

⌚ Timing: 6–8 h

1. Isolate mumps nucleocapsids from HeLa cell lysate by first clearing of cell debris at 500 × g, 4°C for 5 min, followed by pelleting at 13,000 × g, 4°C for 20 min.

**Note:** The 13,000 × g pellet enriched with nucleocapsids was subjected to 50 µg/mL heparin treatment for 1–2 h in order to induce straightening of the nucleocapsids.<sup>1</sup> Straight nucleocapsids proved beneficial for obtaining high-resolution maps. Please refer to Zhang et al.<sup>1</sup> for the detailed steps for cell lysis.





**Figure 1. Overview of the computational pipeline for subtomogram averaging of filamentous helical assemblies described in this protocol**

Parts of figures in Zhang et al.<sup>1</sup> published under a CC-BY license are reused.

## 2. Plunge freeze the isolated nucleocapsids.

- a. Glow-discharge holey Quantifoil grids overlaid with an additional layer of continuous carbon (R2/1 + 2 nm C, Cu 200 mesh grid, Quantifoil Micro Tools) at the front side for 45 s with a PELCO easiGlow Glow Discharge Cleaning System (set at 0.39 mBar, 15 mA, with air as atmosphere).

**Note:** The additional carbon layer provides more uniformly oriented horizontal filaments on the grid, which helps in obtaining uniform ice thickness and in tracing the filaments in later steps.

- b. Concentrate Protein-A, 10-nm colloidal gold fiducials (Electron Microscopy Sciences) before applying to the grids.
  - i. Pellet gold fiducials at  $13,000 \times g$ ,  $4^\circ\text{C}$  for 20 min.
  - ii. Remove the supernatant carefully and resuspend the fiducials pellet in a 3:1 volume ratio with the same buffer used for nucleocapsids preparation.

**Note:** In our experiments, addition of  $1.5 \mu\text{L}$  of  $3 \times$  concentrated fiducials to every  $3 \mu\text{L}$  of nucleocapsid suspension per grid resulted in a low number of fiducials within the field of view of the tilt series. Therefore, tilt-series alignment was eventually performed with patch tracking (see sections below). Optimization of fiducial number might be beneficial for alignment of tilt series with lower signal-to-noise acquired on thicker specimens.<sup>2</sup>

- c. Set a Vitrobot Mark IV plunge freezer (Thermo Fisher Scientific) to  $22^\circ\text{C}$ , 100% humidity, blot force of 0, blot time of 2 s and drain time of 2 s, both sides blotting with filter paper.
- d. Store the frozen grids in sealed boxes in liquid nitrogen until imaging.

## Cryo-electron tomography data acquisition and computational setup

⌚ Timing: 10–20 h

3. Load vitrified grids onto a microscope with the following or similar configuration: Titan Krios microscope operated at 300 kV (Thermo Fisher Scientific) equipped with a field-emission gun, a Quantum post-column energy filter (Gatan) and a K2 direct detector camera (Gatan).
4. Record data using dose-fractionation mode in a dose-symmetric tilt scheme<sup>3</sup> (tilt range of  $-60^{\circ}$ – $60^{\circ}$ ,  $3^{\circ}$  increment) with automated acquisition procedures as implemented in SerialEM software (v3.7.2),<sup>4</sup> at a calibrated pixel size at the specimen level of  $1.69 \text{ \AA}$ , with a defocus range of 2.5– $3.5 \text{ }\mu\text{m}$ , and a total dose of up to  $120 \text{ e}^{-}/\text{\AA}^2$  for 41 tilts.

**Note:** In total, we acquired 23 tilt series for the relatively straight isolated nucleocapsids. Out of these, 3 tilt series did not contain filaments, whereas the remaining 20 contained 1–2 filaments.

△ **CRITICAL:** The filaments in the data must exhibit straight segments in order to achieve the final averages described in this protocol. Curved filaments are likely to exhibit larger local conformational heterogeneity, which in turn contributes to lower resolution reconstruction.<sup>1</sup>

5. Assess the data quality before proceeding to the extensive data processing steps by first processing a few tilt series in Warp and IMOD.

△ **CRITICAL:** If the estimated maximum resolution of the inspected tilts at  $0^{\circ}$ , determined by the Contrast Transfer Function (CTF) fitting in Warp, does not reach beyond  $5 \text{ \AA}$  and if the tilt-series alignment mean residual error determined in IMOD is on average larger than  $1 \text{ nm}$ , we expect the highest achievable resolution of the averages to be limited.

6. Ensure that the software and versions detailed in the [key resources table](#) are installed on your workstation, server or high-performance computing cluster. Ensure hardware with sufficient memory, CPU and GPU capacities.

**Note:** We run Warp and M on a virtual Windows system including 2 Intel Xeon Gold 6226R CPU (total 32 cores, 2.9 GHz), 2 Nvidia Grid V100S GPUs (32 GB memory), and 128 GB RAM; most Dynamo and MATLAB jobs were run on a server with Linux system including 2 AMD EPYC 7513 CPU (total 64 cores, 2.6 GHz), 2 Nvidia GeForce RTX 3090 GPUs (24 GB memory), and 256 GB RAM; RELION jobs were run on a high-performance computing cluster with multiple GPUs (e.g., 3090, A40). We used SSDs for SWAP storage when running jobs on the server or cluster. Overall, we required about 1 TB storage for the entire workflow (raw data and files generated in the Warp/M/IMOD/Dynamo/RELION steps, including those for testing parameters). All data for Warp/M processing were stored on a NetApp Tier-1 storage system.

## KEY RESOURCES TABLE

| REAGENT or RESOURCE  | SOURCE                                 | IDENTIFIER            |
|--|--|-----------------------|
| <b>Chemicals, peptides, and recombinant proteins</b>               |  |                       |
| Heparin sodium   | Sigma-Aldrich; manufactured in the USA | H3149; CAS: 9041-08-1 |
| Cu 200 mesh grid, holey carbon R2/1 + 2 nm continuous carbon layer | Quantifoil Micro Tools                 | Cat#C2-C15nCu20-01    |
| Protein-A/Gold, EM-grade 10 nm                                     | Electron Microscopy Sciences           | Cat#50-281-97         |
| <b>Deposited data</b>  |  |                       |
| EM map for isolated nucleocapsids, majority class                  | Zhang et al. <sup>1</sup>              | EMD: 13133            |
| EM map for isolated nucleocapsids, minority class                  | Zhang et al. <sup>1</sup>              | EMD: 13136            |
| Model for isolated nucleocapsids, majority class                   | Zhang et al. <sup>1</sup>              | PDB: 7OZR             |
| Raw micrographs for isolated nucleocapsids                         | Zhang et al. <sup>1</sup>              | EMPIAR: 10751         |

(Continued on next page)

**Continued**

| REAGENT or RESOURCE                            | SOURCE                           | IDENTIFIER  |
|--|----------------------------------|---|
| Intermediate files of data processing          | This paper                       | Mendeley Data: <a href="https://doi.org/10.17632/d95fvg4dnm.1">https://doi.org/10.17632/d95fvg4dnm.1</a>                  |
| <b>Software and algorithms</b>                 |                                  |   |
| SerialEM                                       | Mastronarde <sup>4</sup>         | <a href="https://bio3d.colorado.edu/SerialEM/">https://bio3d.colorado.edu/SerialEM/</a>                                   |
| IMOD 4.9.4 and 4.10.31 beta                    | Kremer et al. <sup>5</sup>       | <a href="https://bio3d.colorado.edu/imod/">https://bio3d.colorado.edu/imod/</a>   |
| Warp 1.0.7 beta and 1.0.9 (Windows; CUDA 10.0) | Tegunov and Cramer <sup>6</sup>  | <a href="http://www.warpem.com">http://www.warpem.com</a>   |
| M 1.0.9 (Windows; CUDA 10.0)                   | Tegunov et al. <sup>7</sup>      | <a href="http://www.warpem.com">http://www.warpem.com</a>   |
| MATLAB R2019a and R2019b                       | MathWorks                        | <a href="https://www.mathworks.com">https://www.mathworks.com</a>   |
| Dynamo 1.1.401 (CUDA 9.1.85)                   | Castano-Diez et al. <sup>8</sup> | <a href="https://www.dynamo-em.org/w/index.php?title=Main_Page">https://www.dynamo-em.org/w/index.php?title=Main_Page</a> |
| TOM  | Nickell et al. <sup>9</sup>      | N/A   |
| AV3  | Forster and Hegerl <sup>10</sup> | N/A   |
| RELION 3.0 (CUDA 11.1.1) and 3.1 (CUDA 11.5.1) | Zivanov et al. <sup>11</sup>     | <a href="https://github.com/3dem/relion">https://github.com/3dem/relion</a>   |
| Chimera 1.13.1                                 | Pettersen et al. <sup>12</sup>   | <a href="https://www.cgl.ucsf.edu/chimera/">https://www.cgl.ucsf.edu/chimera/</a>   |
| ChimeraX 1.1.1                                 | Pettersen et al. <sup>13</sup>   | <a href="https://www.cgl.ucsf.edu/chimerax/">https://www.cgl.ucsf.edu/chimerax/</a>                                       |
| dynamo2m                                       | Burt et al. <sup>14</sup>        | <a href="https://github.com/alisterburt/dynamo2m">https://github.com/alisterburt/dynamo2m</a>                             |

## STEP-BY-STEP METHOD DETAILS

In the following, we provide a step-by-step guide for the generation of medium to high resolution electron microscopy (EM) density maps of helical filaments by subtomogram averaging (Figure 1). This includes preprocessing of tilt-series frame movies, tilt-series alignment, CTF fitting, tomogram reconstruction, filament tracing, subtomogram generation, subtomogram alignment, averaging and classification at high to low data binning, multi-particle refinement and finally EM map evaluation. To illustrate these steps, we use the example dataset of mumps virus nucleocapsids isolated from HeLa cells following a 1-h stress treatment.<sup>1</sup> The data comprises 20 tilt series, each containing 1–2 filaments, acquired at a pixel size of 1.69 Å (EMPIAR: 10751). Generally, we recommend to create separate working directories for each major step and subdirectories as mentioned in the sub-steps.

### Preprocessing of tilt movies

⌚ Timing: 1–2 h

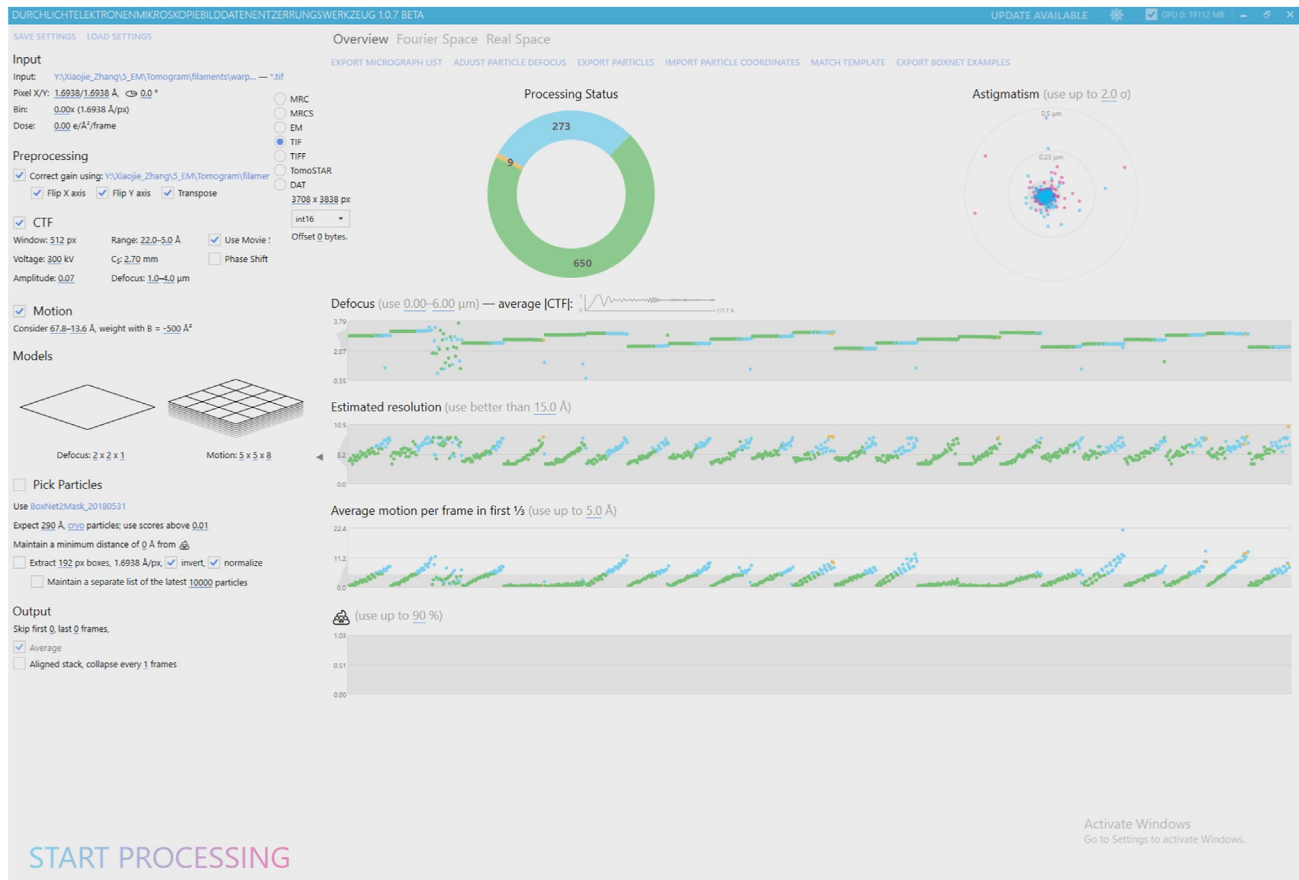
Here we describe steps for pre-processing of the raw frame movies of individual tilts in 2D mode in Warp and the generation of tilt-series stacks for subsequent alignment and tomogram reconstruction. We recommend referring to the original publications<sup>6,7</sup> and the software website ([http://www.warpem.com/warp/?page\\_id=378](http://www.warpem.com/warp/?page_id=378)) for further details about the settings for Warp.

1. Preprocess the raw frame movies of individual tilts in 2D mode.
  - a. Provide all frame movies and the corresponding gain reference file from a single microscopy session in a single folder (e.g., named “frames”).

**Note:** This will constitute the folder for Warp processing, and all output files from Warp will be written into this “frames” folder by default.

- b. Provide all mdoc files (metadata text files for each tilt series written by SerialEM during the acquisition) in a separate folder at the same level as the “frames” folder.

⚠ **CRITICAL:** The names of the mdoc files will be used as prefixes for tomograms in the next steps. Therefore, rename the mdoc files at this step if you wish to name the tomograms in a certain manner.



**Figure 2. Screenshot of the Warp image pre-processing graphical user interface (GUI) and parameters used for processing of the example data comprising 20 tilt series (raw micrographs in EMPIAR: 10751)**

Initially, 23 tilt series acquired during a single microscopy session were processed (shown in this screenshot), 3 of which were not processed further as they did not contain filaments.

- Launch Warp: execute the Warp.exe in the installation directory of Warp and add a desktop shortcut for future use.
- Set up input parameters for processing (Figure 2).
  - Under the **Input** section, first locate the input "frames" folder. Then choose \*.tif to select the TIF option, set the pixel dimensions as 3708 x 3838 px (for K2 data; px is used as an abbreviation of pixels in Warp/M while pix is used elsewhere throughout this protocol), and select int16 to match the raw, unbinned frame images. Set the pixel X/Y as 1.69 Å according to the unbinned pixel size of the data. Leave the **Bin** and **Dose** as default: 0.00x (unbinned) and 0.00 e/Å<sup>2</sup>/frame.
  - Under the **Preprocessing** section, set the correct path to the gain reference file (.dm file in this case; this needs the "Flip Y axis" operation).

△ **CRITICAL:** For the example dataset, additionally select "Flip X axis" and "Transpose" because the raw images were saved as "r/f = 3" in relation to the gain reference image (read from the header of the tif images; equivalent to 270° counterclockwise rotation of frame images relative to the gain reference image). Warp version 1.0.9 also allows input of a \*.txt or \*.tif defect map file to correct defects in camera sensors, which will be applied along with the gain correction. [Troubleshooting 1](#).

- iii. Under the **CTF** part of the **Preprocessing** section, set the following parameters:  
 Set the power spectrum size in the **Window** value as 512 px (a relatively small window).  
 Select the **Use Movie Sum** option as the fitting usually works better with the power spectrum of the movie average given the low dose per frame image in tilt-series data.  
 Set the **Voltage** and **Cs** (spherical aberration) values as 300 kV and 2.70 mm respectively based on the microscope configuration.  
 Set the **Amplitude** value to 0.07 (the fraction of amplitude contrast in the CTF; 0.07–0.10 recommended for cryo-electron microscopy (EM) data).  
 Set the **Defocus** range as 1.0–4.0  $\mu\text{m}$  based on the acquisition defocus range, with some margin.  
 To start with, set the spatial frequency **Range** to 25.0–5.0  $\text{\AA}$  (0.14–0.68 Nyquist (Ny), unbinned pixel size of 1.69  $\text{\AA}$ ) to test the CTF fitting.

**Note:** Adjust the spatial frequency range to include more of the lower or higher frequencies based on the quality of the CTF fitting test with a few frame movies before applying the fitting to all movies (switch to the **Fourier Space** tab and click **Process Only This Item's CTF** to see whether the fitting range represented by the gray rectangle is optimal to the signal in your data). A range of 22.0–5.0  $\text{\AA}$  (0.15–0.68 Ny) worked well for the example data.

- iv. Under the **Motion** part of the **Preprocessing** section, set the **spatial frequency range** for the motion estimation to 67.8–13.6  $\text{\AA}$  (0.05–0.25 Ny) and the **B-factor** to  $-500 \text{\AA}^2$ .
- v. Under the **Models** section, do the following.  
 Set the **Defocus** model resolution as  $2 \times 2 \times 1$ , with the first two defining the spatial resolution and the third as temporal resolution for the defocus estimation.

**Note:** The estimation is done on the sum of frames due to insufficient signal in a single frame. A value of 2 at the first two parameters means no local defocus estimation performed. The recommended settings in the Warp manual  $5 \times 5 \times 1$  may also work depending on the unbinned pixel size of your data.

Set the **Motion** resolution to  $5 \times 5 \times 8$  for this dataset, with the first two values defining the spatial resolution and the third defining the temporal resolution (8 frames per movie in the example dataset).

**Note:** As the manual suggests, set a temporal resolution value that matches the overall electron dose per pixel if a low dose rate ( $<1 \text{ e}/\text{px}/\text{frame}$ ) is used. The first two spatial resolution parameters can also be set to 1, if the signal in a single tilt movie average is suspected to be too low for alignment.

- vi. Skip the **Pick Particles** option in this case.
- vii. Leave the **Output** section as default. Do not skip any frames in this case. The aligned movie **Average** will be written out.
- e. Start processing and check the results.
  - i. With the parameters set in the previous steps, click the **START PROCESSING** button to start. The button will be switched to **STOP PROCESSING** once the procedure starts.

**Note:** Warp provides an overview tab to show the processing status of all movies and plots for the defocus value, the estimated resolution, and the average motion per frame in the first 1/3 of the processed movies. On the left side of each plot, a histogram provides additional evaluation of the data quality (Figure 2).



- ii. Click to **STOP PROCESSING** if parameters need be changed or the processing is completed.

**Note:** In this case, we only used one GPU, but Warp will try to use all compatible GPUs available on the system. Therefore, deselect the unwanted GPUs from the upper right corner of the window bar.

- f. When processing is done, check the processing results of individual tilt movies.
  - i. Collapse the settings panel on the left by clicking on the arrowhead between the settings and the Overview/Fourier Space/Real Space panel.
  - ii. Switch to the **Fourier & Real Space** tab, and inspect the results by clicking through the navigation bar (green/blue color in most cases) at the bottom of this tab (with the left and right arrow keys on the keyboard).
  - iii. For fine-tuning the CTF fitting parameters for individual tilt movies, use the **PROCESS ONLY THIS ITEM'S CTF** button after adjusting the spatial frequency range under the CTF section (switch back the settings panel).

**Note:** If this is done, the previously processed items with different settings will change their processing status to **Outdated** in yellow. Changing the spatial frequency range back to the previous values after individual tilt movies are fine-tuned should change the process status of the majority of movies back to green for proceeding to the next step. One can set the filter thresholds in each plot in the Overview tab so that movies with fitted parameters out of the selected ranges can be filtered out when proceeding with the Warp pipeline. Alternatively, one can manually exclude individual tilt movies by clicking on the check box in front of the name of each tilt movie at the navigation bar in the Fourier & Real Space window ("-" means automatic filters). In this example (Figure 2), movies were not excluded at this stage.

- 2. Export tilt-series stacks from the summed tilt frame movies generated in the last step.
  - a. Switch Warp to tilt-series mode by choosing the **TomoSTAR** input format at the top of the window.
  - b. In this mode, click on the **Import tilt series from IMOD** button at the Overview tab and a dialog pops up for choosing the path to the folder containing the mdoc files written by SerialEM and the path to the original input tilt movies ("frames" folder).
  - c. Leave the pixel size as 1.69 Å and dose per tilt as 0 e/Å<sup>2</sup>. Warp reads the dose value from the mdoc files.
  - d. Uncheck the Invert tilt angles in the Warp version 1.0.7 beta (option changed to invert tilt angles by default in the Warp version 1.0.9).
  - e. Click on **CREATE STACKS FOR IMOD**.

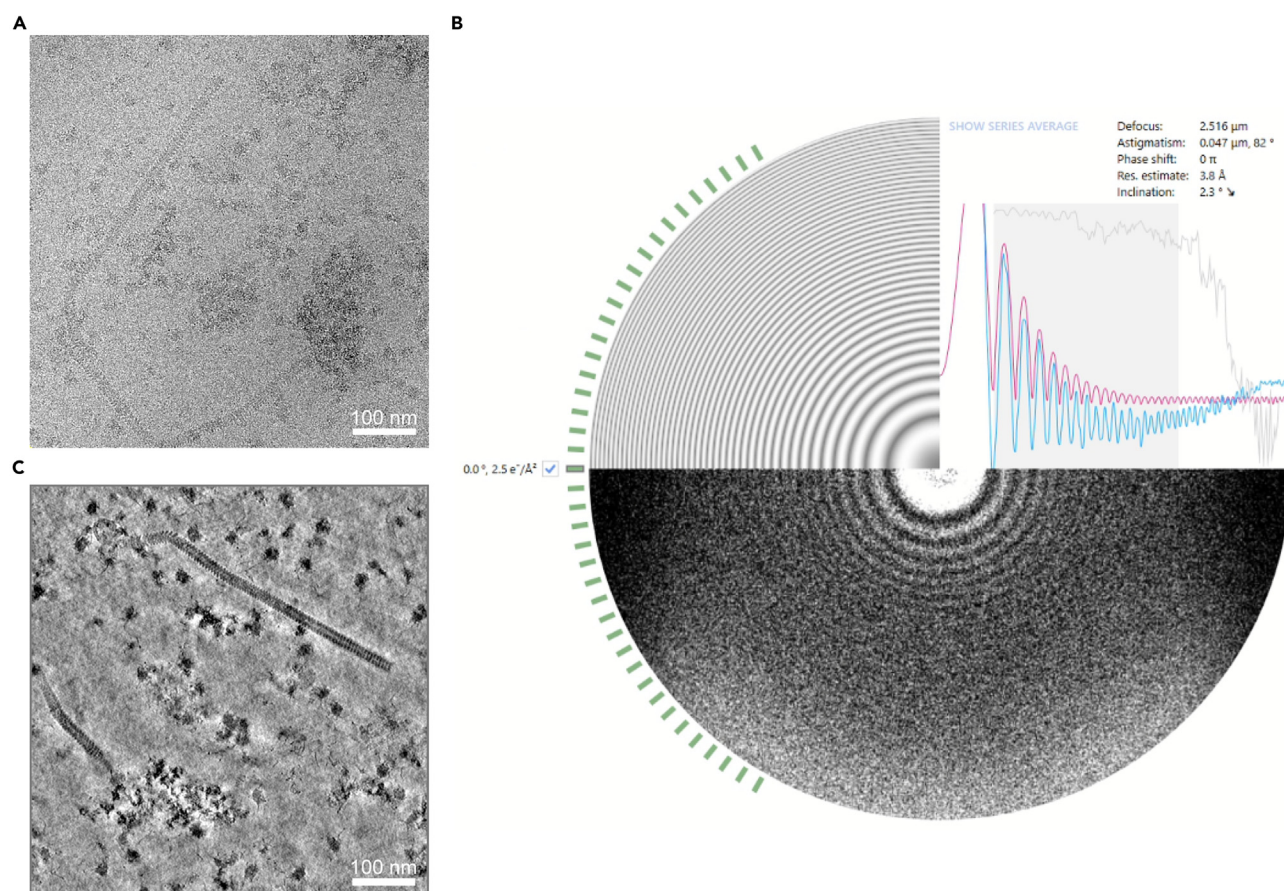
**Note:** Warp will write a new directory named "imod" containing separate folders for each tilt-series stack named according to the corresponding mdoc name.

### Tilt-series alignment and CTF refinement and tomogram reconstruction

⌚ Timing: 7 h

This section describes the alignment of tilt series generated by Warp with Etomo in the IMOD package. Processing of the tilt series until the fine alignment step (included) is sufficient to allow Warp to import the alignment parameters from IMOD. Continue with refining the CTF estimation for tilt series and reconstruct tomograms in Warp to the desired binning size for filament tracing in the next section.





**Figure 3. Tilt series CTF estimation and tomogram reconstruction in Warp**

(A) A tilt image from one of the tilt series in the example data.

(B) A screenshot from the Warp GUI showing the refined CTF processing result of the full tilt series in A (the screenshot shows details for the 0-degree tilt) after importing alignment parameters determined in IMOD. The CTF fitting quality is normally expected to be worse at higher tilt angles compared to that at lower angles.

(C) A tomographic slice of the data shown in (A) following Warp reconstruction and deconvolution. The tomogram was deconvolved with the default parameters during the Warp reconstruction procedure for better visual inspection and manual filament tracing in the next step. The filament is curved, but its relatively straight segments can be used for subtomogram averaging.

3. Perform tilt-series alignment with the Etomo module (IMOD version 4.9.4 described here) in the imod subfolders containing the individual tilt series written by Warp (Figure 3A). Leave parameters as default unless specified otherwise in the following steps.

**Note:** For further information on processing in Etomo, we recommend referring to the software website (<https://bio3d.colorado.edu/imod/doc/etomoTutorial.html>).

- a. With a Windows command prompt or Linux terminal, navigate to the subfolder for the first tilt series, and type “etomo” to launch the software.
  - i. Select **Build Tomogram** to enter the setup interface for single tomogram reconstruction.
  - ii. Specify the tilt series to be reconstructed in **Dataset name**.
  - iii. Select cryoSample.adoc in **Templates** and set **Axis Type** as Single axis while keeping **Frame Type** as Single frame.
  - iv. Click on **Scan Header** to extract the **Pixel size** and **Image rotation** from the input data.
  - v. Set **Fiducial diameter** as “0” as no fiducial markers were used for tracking with this dataset.

- vi. Select the “Extract tilt angles from data” option under the **Axis A** section.
- vii. Click on the **View Raw Image Stack** button to inspect whether some views of the tilt series need be excluded at this point due to bad quality, e.g., views of areas largely covered by ice.

△ **CRITICAL:** If views are excluded during the Etomo processing, edit the mdoc file accordingly to remove those views to avoid mismatch with the subsequent Warp processing that makes use of the alignment parameters from Etomo.

- viii. Continue to **Create Com Scripts** to start processing.
- b. At the **Pre-processing** window,
  - i. Click on the “Find X-rays (Trial Mode)” button to erase defects or “hot pixels” from tilt images.
  - ii. Click on “Create Fixed Stack” to create a fixed stack.
  - iii. Continue to use the newly generated stack by clicking “Use Fixed Stack”.
  - iv. Click on “Done” to proceed.
- c. At the **Coarse Alignment** window,
  - i. Click on the “Calculate Cross-Correlation” button to align the tilt series.
  - ii. In the **Newstack** section, set the “Coarse aligned image stack binning” to 4 (6.78 Å/pix) and select the “Reduce size with antialiasing filter” and “Float intensities to mean”. Deselect “Convert to bytes”.
  - iii. Click on “Generate Coarse Aligned Stack” to generate a new binned stack for the next step.
  - iv. Click on “Done” to proceed.
- d. At the **Fiducial Model Generation** window,
  - i. Select the “Use patch tracking to make fiducial model” option and set the **Size of patches** to 150 by 150 pixels and **Iterations** as 4 (the maximum value).
  - ii. Click on “Track Patches” to generate the patch fiducial model.
  - iii. Click on “Done” to proceed.
- e. Continue with **Fine Alignment**.
  - i. Leave settings at the **General** tab as default.
  - ii. In the **Global Variables** tab, set the **Rotation Solution Type** to “One rotation”, **Magnification Solution Type** to “Fixed magnification at 1.0”, **Tilt Angle Solution Type** to “Fixed tilt angles” (tilt angles pre-calibrated) and **Distortion Solution Type** to “Disabled”.
  - iii. Click on “Compute Alignment” and assess the alignment quality by reading the values corresponding to “Ratio of total measured values to all unknowns” and “Residual error mean and sd” in the Project Log.
  - iv. Right click at the Fine Alignment window to open the log file with the “Align log file” option. Check the mean residual of each tilt view under the **Solution** tab in the log file in order to decide whether some views with large residual errors should be excluded for improving alignment, or use the “robust fitting” option, both of which are possible under the General tab.

**Note:** Aim for a residual error mean value smaller than 1 pixel if possible. For example, the third tilt series in this dataset resulted in an initial value of 1.150 pixels (sd: 1.018 pixels). The mean value was reduced to 0.756 pixels (sd: 0.532 pixels) after removal of a few bad contours (use “View/Edit Fiducial Model”).

- v. Click on “Done” to proceed.

**Note:** For Warp reconstruction, the following steps f-j are not necessary. Warp makes use of the alignment parameters generated in Etomo up to this point. Proceed with the next steps for the first tilt series until the final tomogram is generated as a quality and consistency check.

- f. Skip the **Tomogram Positioning** step. Click “Done” to proceed.

**Note:** While this step can help in reducing the reconstructed file size and providing easier visualization for tilted specimens, the transformation applied in Etomo can create a mismatch of particle positions as downstream processing is performed on untransformed reconstruction.

- g. At the **Final Aligned Stack** window, only use the **Create** tab.
  - i. Enable “Use linear interpolation”, align image stack with a binning factor of 4, and enable “Reduce size with antialiasing filter”.
  - ii. Click on “Create Full Aligned Stack”.
  - iii. Click on “Done” to proceed.
- h. At the **Tomogram Generation** window,
  - i. Use the “Back Projection” algorithm.
  - ii. Set the “Tomogram thickness in Z” to 1000 (in unbinned pixels) and use the default Radial Filtering parameters.
  - iii. Click on “Generate Tomogram”.
  - iv. Click on “Done” to proceed.
- i. At the **Post-processing** window,
  - i. Deselect the “Convert to bytes” option under the **Scaling** section and select the “Rotate around X axis” option under the **Reorientation** section.

**Note:** This should produce a reconstruction (of the tomogram and map) with correct handedness given the microscope settings in the example dataset.

- ii. Click on “Trim Volume”.
  - iii. Click on “Done” to proceed.
- j. Skip the **Clean Up** step for now and click on “Done” to finish the reconstruction of the tilt series. Intermediate files of large sizes can be cleaned altogether when necessary.

△ **CRITICAL:** Complete the reconstruction of one tomogram per batch of data for determining optimal parameters for tomogram reconstruction (e.g. Z dimension size), assessing the data quality, and for checking the consistency between IMOD and Warp results.

**Optional:** Reconstructing tomograms in IMOD batch mode saves time especially with a larger dataset (<https://bio3d.colorado.edu/imod/doc/batchGuide.html>).

- k. Complete the reconstruction of one tomogram as detailed above and save the parameters in a “batch directive file” by using the **Export** option in the Etomo interface.
- l. In the batch reconstruction interface, set it as the “Starting directive file” and select “cryoSample.adoc” in the System template box.
- m. Import all the remaining tilt series in the “Stacks” tab, check parameters in the “Dataset Values” tab and set it to “Stop after Fine alignment”. Press “Run” to start. Assess the alignment results and adjust for individual tilt series if needed.

**Optional:** One can use AreTomo<sup>15</sup> (<https://github.com/czimagininginstitute/AreTomo2>) for automatic batch alignment and reconstruction of tomograms with GPU environment available. Warp also accepts alignment parameters/files generated by AreTomo.

4. Refine CTF estimation for tilt series and reconstruct tomograms in Warp.
  - a. Re-open Warp and click the **IMPORT TILT SERIES FROM IMOD** button to import the alignment results from IMOD in the TomoSTAR mode.

- i. Specify the folders for mdocs, original frame movies, and IMOD processing results (select the subfolder named "imod" created by Warp).

△ **CRITICAL:** If there are views excluded during the Etomo alignment, edit the mdoc file accordingly to remove those views for Warp to consistently take the corresponding alignment parameters from Etomo.

△ **CRITICAL:** Due to a difference in handling tilt angle direction in Warp compared to IMOD, there is an option to invert tilt angle when importing tilt series from IMOD. When the example data were processed with the Warp version 1.0.7 beta, the tilt angle convention was set inversely to that in IMOD. Thus, additional flipping of subtomogram averages was necessary to produce the correct map handedness.

**Note:** If all IMOD processing results are found, checkboxes in the "Aligned" column will automatically be ticked for all tilt series. Warp requires the xf and taSolution.log files in each of the "imod" reconstruction subfolders.

- ii. Click on **IMPORT** to continue. [Troubleshooting 2](#).
- b. Specify the **Unbinned tomogram dimensions** as 3838 × 3708 × 1000.

△ **CRITICAL:** According to the convention, the tilt axis in the raw movie frames is approximately aligned to the image X axis, while that in the tomogram is aligned to the Y axis. Thus, taking for example data recorded on a K3 detector, the unbinned image dimensions are 6000 × 4000 pixels and the tomogram dimensions should be 4000 × 6000 instead. The Z dimension with this example dataset was set to 1000 based on the sample thickness, leaving some margin at the border of tomograms to avoid problems later during subtomogram extraction.

- c. Leave all preprocessing parameters/settings as those for processing of raw movie frames.
  - i. Go to the **Fourier Space** tab and click on "PROCESS ONLY THIS ITEM'S CTF."
  - ii. Click on "CHECK TILT HANDEDNESS" to check if the defocus handedness is correct.

△ **CRITICAL:** This relates to the defocus convention. Improper handling will limit the maximum attainable resolution of the subtomogram averages. If the correlation value at this step is close to 1, it indicates correct defocus handedness, as in the case of the example data. If the value is close to -1, follow instructions to flip the defocus handedness.

- iii. Click on **START PROCESSING** to process the CTF for all tilt series in batch ([Figure 3B](#)).
- d. When the previous step is done, go to the Overview tab and click on the **RECONSTRUCT FULL TOMOGRAMS** button.
  - i. A new interface pops up that allows reconstruction of tomograms with user-defined parameters. Set the **pixel size** based on the binning factor required for the next step of data processing. Here, 6.78 Å is used to obtain 4× binned tomograms.
  - ii. Choose "Normalize input images" which is useful when data are processed in RELION<sup>11</sup> in the next steps.
  - iii. Use the "Also produce deconvolved version" option with the default parameters, which significantly improves the reconstruction contrast in the example data ([Figure 3C](#)).

**Note:** If denoising is needed to obtain even better contrast, use the "Reconstruct half-maps for denoising" option. One can further denoise the tomograms either with Warp or other software such as cryoCARE.<sup>16</sup>

- iv. For this data, use the “Include items outside of filter ranges”.
- v. Click on **RECONSTRUCT** to start. The output tomograms will be written to an automatically generated subfolder named “reconstruction” or “reconstruction/deconv” (for the deconvolved tomograms).

**Note:** Reconstruction of full tomograms at lower binning can result in slow processing, or even crashing of Warp. Also, the lower contrast in lower binning tomograms will not be beneficial for later steps of filament tracing.

**Optional:** Tomograms reconstructed in IMOD or AreTomo can also be used for filament tracing. But pay attention to possible pixel shifts or image rotation/flipping in comparison to the Warp reconstruction that will ultimately be used to generate subtomogram in the later steps.

## Filament tracing and subtomogram generation

⌚ Timing: 6–9 h

This section demonstrates how to use the filament model function in Dynamo<sup>8</sup> to manually trace the backbone of the isolated filaments in the 4× binned tomograms reconstructed in Warp, to generate positions for subtomograms based on the geometry of the filaments, and to crop subtomograms for next steps of alignment and averaging.

5. Create filament models in Dynamo.

**Note:** Our protocol was produced with an earlier version of Dynamo (1.1.401-foss-2017b-MCR-R2018a) together with MATLAB (2019b), but should be applicable to newer versions in most cases. We recommend to read the detailed instructions on usage of Dynamo in the software wiki ([https://www.dynamo-em.org/w/index.php?title=Main\\_Page](https://www.dynamo-em.org/w/index.php?title=Main_Page)).

- a. Activate Dynamo in MATLAB with this command:

```
>run /path-to-the-installation-folder-of-dynamo/dynamo/1.1.401-foss-2017b-MCR-R2018a/dynamo_activate.m
```

- b. Create and enter a new working directory to run the Dynamo project, e.g., “4bin\_dynamo”.
- c. Create a catalogue for 20 tomograms from this example data.
  - i. In a terminal window (outside of MATLAB), run the command:

```
>ls -d /path-to-the-Warp-preprocessing-folder/frames/reconstruction/deconv/*.mrc >> list_4b_deconv.vll
```

**Note:** This generates a text file recording the path and names of all deconvolved tomograms generated in previous steps by Warp.

- ii. Create a Dynamo catalogue named “filaments” by running:

```
>dcm -create filaments -fromvll list_4b_deconv.vll
```

- d. Generate filament models in the Dynamo graphical user interface (GUI).
  - i. Open the Dynamo GUI with MATLAB command:

```
> dcm filaments
```

- ii. In the dcm GUI, click on **Create binned versions** under the **Catalogue** menu to pre-bin all tomograms with a factor of 2 for easier visualization (target voxel size:  $6.78 \times 2^2 = 27.10 \text{ \AA}$ ) and filament tracing (will be annotated with the original  $4 \times$  binned pixel size  $6.78 \text{ \AA}$  by default in the next step by Dynamo).
- iii. Select the first tomogram at its **index** field and right click to open the binned tomogram in **tomoslice**.
- iv. Follow the tutorial video (<https://www.youtube.com/watch?v=hsZ-9xAYufl>) to generate filament model. In brief:  
Call the **dtmslice** interface from the **dpreview** GUI and manually mark the two ends ("anchors") of each filament with the [1] and [2] keyboard keys.

**⚠ CRITICAL:** Avoid including heavily curved regions at the tracing step already.

Create orthogonal sections along the line connecting each pair of two anchor points and click on the centers of these sections to trace the "backbones" of filaments (backbones are smoothed in later steps).

**Note:** For the example data, only modify the **sidelength** (the box size) for generating sections of 80 pixels (a bit over twice larger than the filament diameter; pixel size:  $6.78 \text{ \AA}$ ), a value large enough to fully cover the cross-section view of sometimes curved filaments.

**Note:** If needed, modify the step parameters in the **Traverse path** box (**Sections along path** in later Dynamo versions) in the GUI to adjust how many sections to generate along the traverse path (the default 1:10:100 means to divide the distance between two anchor points by 100 parts and create a section ever 10 parts).

In the slices view, use the [C] keyboard key to choose the center of the cross-section views on different slices.

- v. Convert these manually clicked points into coordinates for subtomogram/particle extraction: choose the **filamentWithTorsion** model (Figure 4A) to generate equidistant ( $dz: 8 \text{ pixels} = 54.24 \text{ \AA}$ ) segments of filaments with "artificial" torsion angle ( $dpsi: 36 \text{ degrees}$ ), in order to reduce the dominance of the missing wedge during alignment later.

**Note:** There are four filament model types available in Dynamo which can be selected to fit the type of filaments being processed: **filamentWithTorsion**, **filamentSubunitsOnHelix**, **filamentSubunitsOnRings**, or **filament** (with randomly distributed positions in the filament tube).

- vi. Save the models with the positions and angles for cropping subtomograms.

**Optional:** By incorporating loop scripting, it is also possible to input into Dynamo multiple series of points (along the traced filaments) as a matrix of their  $x/y/z$  coordinates created elsewhere (e.g. Amira) and generate geometrically computed sampling points for subtomogram extraction with MATLAB code provided in the Dynamo wiki ([https://wiki.dynamo.biozentrum.unibas.ch/w/index.php/Filament\\_types\\_code\\_example](https://wiki.dynamo.biozentrum.unibas.ch/w/index.php/Filament_types_code_example)).

6. Grep Dynamo tables and crop subtomograms from  $4 \times$  binned tomograms with Dynamo commands in MATLAB.
  - a. Follow the Dynamo wiki, and use the following command lines together with loop scripting to combine multiple model tables for each tomogram into one table per tomogram.



```
>dcmodels filaments -i 1 -ws o;

>m=dread(o.files);

>t_acc=[];

>for i=1:length(m);

>  n=m{i};

>  t=n.grepTable();

>  t_acc=cat(1,t_acc,t);

>end

>t_acc(:,1)=1:size(t_acc,1);

>t_acc(:,(16:17))=0;

>dwrite(t_acc,'table_1.tbl')

>clear
```

Here, the 1 after -i refers to the tomogram with index = 1 (also the first tomogram, explained in next steps).

- b. Crop particles in Dynamo for all tomograms together.
  - i. Create a new volume list for cropping particles from the non-deconvolved tomograms in a terminal.

```
>ls -d /path-to-the-Warp-preprocessing-folder/frames/reconstruction/*.mrc >> list_4b.vll
```

- ii. Modify this list file such that each tomogram gets an index number (1–20 in this case) and the absolute path and name of each table per tomogram in this format:

```
/path-to-the-Warp-preprocessing-folder/frames/reconstruction/*.mrc

*index = 1

>table_1.tbl
```

- iii. Crop all particles in MATLAB,

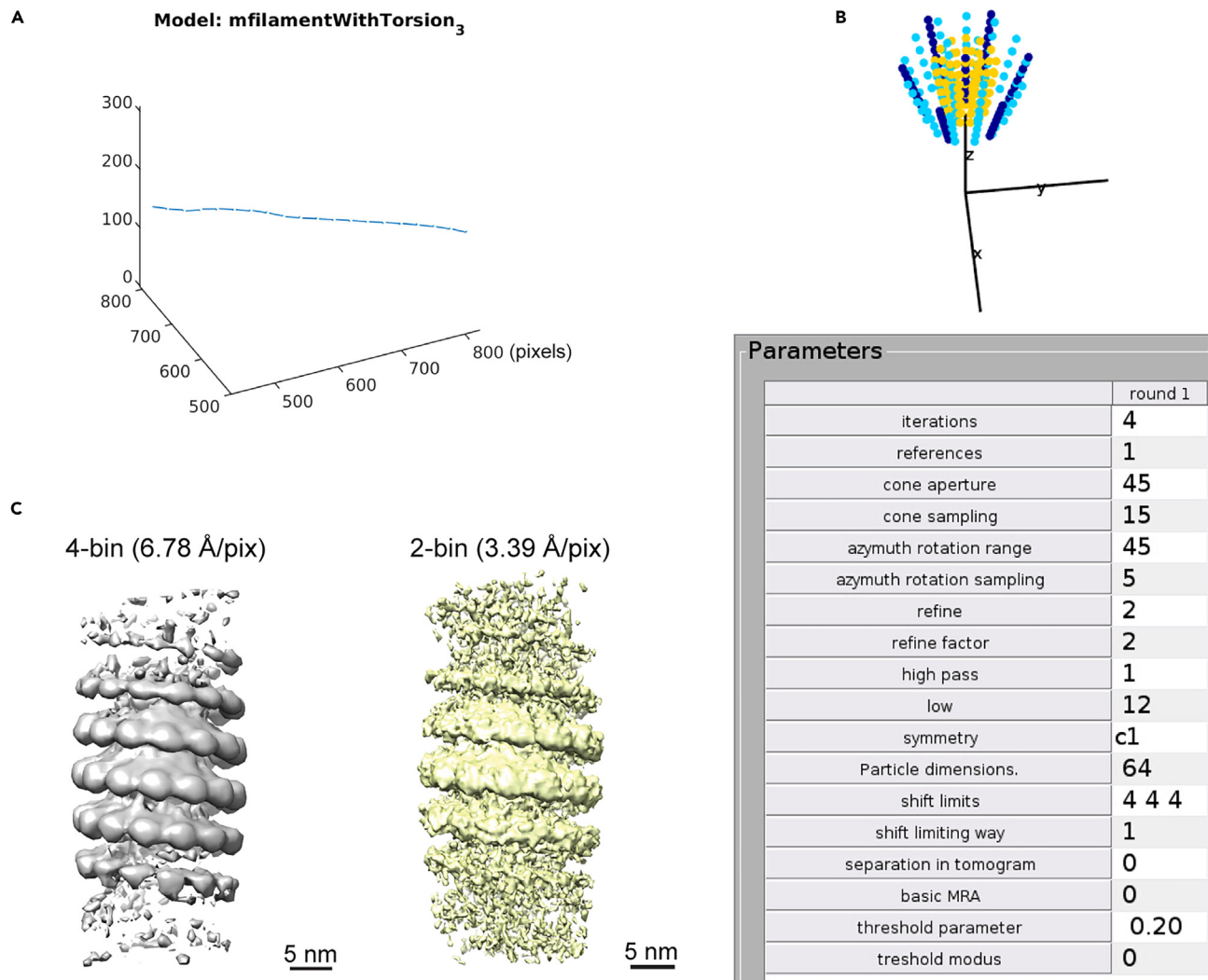
```
>dtcrop list_4b.vll reorder filamentsData 64
```

**Note:** Here, **reorder** re-indexes particles after cropping because some positions too close to borders of tomograms will be skipped automatically by Dynamo. Also, a new table with final indices corresponding to the cropped particles is included in the particle directory. One can split the table based on tomogram indices later in MATLAB if needed. 64 refers to the box size in 4× binned pixels (6.78 Å/pix), which is about twice larger than the filament diameter measured from the tomograms.

- iv. Use this newly generated table (here renamed to: “table\_combined\_reindex.tbl”) for Dynamo subtomogram averaging in next steps. 1297 particles were generated in this step for the example dataset.

**Note:** If MATLAB is not available, most Dynamo commands can also be run directly in the terminal as a standalone version: with Dynamo loaded as a module or path defined, one can type “dynamo” in terminal to activate it and then run Dynamo functions in a slightly modified syntax (see Dynamo wiki for more details).





**Figure 4. Filament model generation and initial subtomogram alignment in Dynamo**

(A) An example of “filament model with torsion” generated in Dynamo. Lines with short perpendicular lines at their ends indicate where and with what orientations subtomograms are generated (positions and angles are recorded in a Dynamo table).

(B) An illustration of the multilevel refinement of subtomogram orientations generated in Dynamo (top: dark blue, light blue, and yellow are colored according to the order of sampling grid in three consecutive levels). The corresponding numerical parameters used in the first Dynamo alignment project are shown at the bottom.

(C) Subtomogram averages after alignment at 4× binning and 2× binning of 1297 particles from all filaments in the example data. Y is flipped to the correct handedness. Pixel sizes are indicated. Parts of figures in Zhang et al.<sup>1</sup> published under a CC-BY license are reused.

**Optional:** Subtomograms can also be reconstructed in Warp. Use the dynamo2warp function in the dynamo2m package<sup>14</sup> to convert a Dynamo table to a STAR file for Warp reconstruction (detailed later at step 10). In our initial pipeline, we converted a Dynamo table to a motl file (for TOM/AV3<sup>9,10</sup> alignment) and/or a STAR file (for Warp/RELION) with MATLAB scripts in order to test alignment with different software at different steps.

### Initial subtomogram alignment and averaging

⌚ Timing: 2–3 days

While later high-resolution processing is performed in RELION, we found it beneficial to perform initial alignments, especially for non-globular structures, in subtomogram averaging-dedicated software, such as Dynamo. This section describes the strategy to align the 4× binned and 2× binned subtomograms (6.78 Å/pix and 3.39 Å/pix, respectively) with Dynamo for easy plots of aligned positions/angles and for ensuring uniform directionality of subtomograms positioned along the same filament, which is particularly important for filaments with polarity.

7. Create a project named “bin4\_align\_1” for aligning the 4× binned subtomograms (6.78 Å/pix; [Figure 4B](#)).

**Note:** For detailed explanation of the parameters used in Dynamo alignment, please refer to the Dynamo wiki mentioned above. We set up working directories for the following Dynamo alignment projects at the same level so that the tables, particles, and templates can be easily accessed by these projects at a working directory one level upper.

- a. Create an initial average of all subtomograms generated from the previous step as a template for Dynamo subtomogram alignment. With Dynamo loaded in a terminal, run the following in the working directory.

```
>oa = daverage('filamentsData', 't', 'table_combined_reindex.tbl');
>dview(oa)
>dwrite(oa.average, 'raw_template.em')
```

- b. Set up a Dynamo subtomogram alignment project with the following command:

```
>dcp.new('bin4_align_1', 'd', 'filamentsData', 'template', 'raw_template.em', 'masks',
'default', 't', 'table_combined_reindex.tbl');
```

- c. Modify the alignment parameters including the masks in the GUI and run the Dynamo project.

```
>dcp bin4_align_1
```

- i. Open the masks interface. Use the **Mask editor** to create a file, e.g., mask64\_r17\_h25\_g2.em, with the “raw\_template.em” as a reference.

**Note:** This corresponds to a cylinder mask with a box sidelength of 64 pixels, radius of 17 pixels, height of 25 pixels and Gaussian smoothened by a factor of 2; it should fit the template size in the overlaid view. Use this mask for the alignment, classification as well as calculating the Fourier Shell Correlation (FSC). The classification and FSC masks are not used here, but must be set for the project to run.

Set the Fourier mask on template as “fill with ones” by clicking on that button at the “Input masks for project” interface.

- ii. Set the alignment **numerical parameters** ([Figure 4B](#)). Key parameters are for example: iterations: 4; cone aperture and sampling: 45 and 15 degrees; azimuth rotation range and sampling: 45 and 5 degrees; refine and refine factor: 2 and 2; high and low pass: 1 and 12 Fourier pixels which correspond to a maximum achievable resolution of 36 Å at the first iteration of alignment ( $64 \text{ (box\_size)} / 12 \text{ (low\_pass)} \times 6.78 \text{ (pixel\_size)} = 36 \text{ Å}$ ); shift limits: [4, 4, 4] in pixels; shift limiting way: 1.

- iii. Set the **computing environment** depending your local settings for running Dynamo. Either the **Standalone** option with a few CPU cores or **GPU standalone** with the GPU identifiers properly set worked in our case.
- iv. Click on the **check** and then the **unfold** button to finish preparation of the project.
- v. Exit the GUI and submit the job either by running `./bin4_align_1.exe` in the Dynamo prompt, or submit the job to a cluster following the detailed instructions on the Dynamo wiki.

**Note:** Dynamo alignment parameters can also be set simply by command lines without the GUI.

**Note:** One can also use the average from a single long filament as an initial template/reference for alignment. As we have tried to keep uniform directionality during the filament tracing step, the average of all subtomograms gives better signal-to-noise ratio and should not be dominated by subtomograms from filaments with opposite directionality.

**Note:** One can also try using a large cone for angular search (e.g. 360/45). With the example data, such test resulted in a worse average showing fake 2-fold symmetry, likely due to averaging of filaments with mixed polarities. Thus, for later steps of alignment at lower binning, even smaller angular and shift limits than those in the first alignment stage are recommended.

**Note:** Use `dynamo_table_plot` function to plot orientation and positions of subtomograms after alignment. Dynamo has multiple flags regarding how the plots are displayed, e.g. colored according to cross-correlation (cc) values. The values in tables can be plotted in histograms, e.g. the shifts in the refined tables, which helps for evaluation and optimization of the refinement procedures beyond visualization of the resulting averages.

8. Check the consistency in the directionality of each filament, flip the ones with opposite directions, regenerate a new combined table, and align the subtomograms again with this new table.
  - a. Create a new working directory named “`direction_check`” (the same level as the alignment project).
  - b. Split the refined table from the project “`bin4_align_1`” into tables for individual filaments by reading the table into a variable and performing matrix operation in MATLAB based on the indices of tomograms in Column 20 of the table, or with the `dynamo_table_grep` function.
  - c. Generate averages for individual tomograms from the split tables with the `daverage` and `dwrite` functions as shown above. One can automate and combine these two steps with a loop script in MATLAB.
  - d. Inspect the averages for the individual filaments and takes note of the ones showing opposite directions to the first subtomogram average (used as a reference here).

**Note:** In our case, we found that Tomogram 1, 2, 7, 10, 13 gave averages of opposite directions (single filament in each tomogram). Flip the directions for these subtomograms by operating on their corresponding tables and adding 180 degrees to the values in Column 8 of these sub-tables in the version before the execution of any alignment, with the following MATLAB script.

```
>t= dread('table_combined_reindex.tbl');
>t_flip=t;
>t_flip(t(:,20)==1,8)=t(t(:,20)==1,8)+180;
>t_flip(t(:,20)==2,8)=t(t(:,20)==2,8)+180;
```

```
>t_flip(t(:,20)==7,8)=t(t(:,20)==7,8)+180;
>t_flip(t(:,20)==10,8)=t(t(:,20)==10,8)+180;
>t_flip(t(:,20)==13,8)=t(t(:,20)==13,8)+180;
>dwrite(t_flip, 'table_combined_reindex_flip.tbl')
```

- e. Run a new alignment project named “bin4\_align\_2” in the working directory where the project “bin4\_align\_1” is created.
  - i. Set the masks as for previous alignment projects
  - ii. Set the alignment parameters as the following: cone 60/15, in-plane 45/15, refine 2/2, high/low pass 1/12, iter 4, shift limits [4, 4, 4], shift\_limiting way 1.
9. Run a Dynamo adaptive bandpass (ABP) filtering project from the alignment results of “bin4\_align\_2” in the working directory where the project “bin4\_align\_2” is created.
  - a. Copy the refined table and average to the same working directory and rename the files if needed.

```
>t=dread('refined_table_ref_001_ite_0004.tbl');
>t_double=t(:,1:26);
>dwrite(t, 'table_all.tbl')
>dcg.new('abp_align', 'd', '
filamentsData', 'template', 'average_ref_001_ite_0004.em', 'masks', 'default', 't', 'table_all.tbl');
```

- b. Launch the project GUI with the dcp command and set the same masks as for previous alignment projects.

**Note:** Here, the FSC mask will be meaningfully applied to each of the averages of the half dataset at each iteration for estimating resolution and defining the lowpass filter parameter for the next iteration.

- c. Derive an ABP project from this single reference project:
  - i. In the GUI, choose the following options in the **Multireference** tab > **Adaptive filtering** (“golden standard”) > **Derive a project**.
  - ii. In the newly launched interface, rename the derived ABP project as “bin4\_align\_2\_eo” following the convention and click the Enter key.
- d. Generate the odd/even references by averaging the odd/even subtomograms in the eo project GUI (at the template interface).
- e. At the template interface, use the **visualize FSC** button to estimate the initial frequency given in Fourier pixels.

**Note:** The X value corresponding to a Y value of around 0.143 in the FSC plot when moving the cursor along the curve<sup>17</sup> is the initial frequency. This will be used for setting the lowpass filter parameter of the first round of alignment.

- f. Set the initial frequency value together with an FSC threshold of 0.143 and a resolution push-back of 2 Fourier pixels in the dcp GUI under the **Multireference** tab > **Adaptive filtering** (“golden standard”) > **Edit parameters for Adaptive Filtering run**.
- g. Set the numerical parameters and computing environment as for previous Dynamo projects.
- h. Run the project.
- i. When finished, launch the project GUI again and use the **Plot attained resolution FSC for all iterations** option under the **Adaptive filtering** (“golden standard”) option to check the evolution of FSC plots and resolutions in Fourier pixels in the different iterations.

**Note:** This project converged and resulted in an average (Figure 4C, left) with an estimated resolution of  $64/19 \times 6.78 = 22.8 \text{ \AA}$  ( $1 / \text{Fourier\_pixels} \times \text{box\_size} \times \text{pixel\_size}$ ; 19 Fourier pixels here).

10. Combine the tables from the bin4\_align\_2\_eo project and prepare for subtomogram reconstruction and alignment at 2× binning (pixel size: 3.39 Å).

**Note:** Due to the difficulty and unnecessary in reconstructing full tomograms at 2× binning for Dynamo reconstruction, we choose Warp for reconstructing subtomograms at lower binning.

- a. As mentioned in the optional section of step 6, use the dynamo2m package (dynamo2warp function) to convert the Dynamo table to a Warp STAR file.
- b. In the Warp GUI, use the TomoSTAR mode, and click on the **RECONSTRUCT SUB-TOMOGRAMS** button to create subtomograms and their corresponding 3D CTF models.

**Note:** CTF models and the output STAR file are not used at 2× binning, but those reconstructed at the unbinned pixel size will be important for the next step processing in RELION.

- i. In the pop-up options for exporting subtomograms, set the “Coordinates use 6.78 Å/px” if the coordinates are at 4× binning (pixel size: 6.78 Å) in the STAR file that is converted from a Dynamo table.

**Note:** In our case without the dynamo2m package, we already scaled the coordinates to 2× binning with a MATLAB script while converting them from a dynamo table to a Warp STAR file. Thus, set it to “Coordinates use 3.39 Å/px”.

- ii. Set the “Output will be scaled to 3.39 Å/px” for 2× binned subtomograms.
- iii. Set the “Box size is 128 px”.
- iv. Set the “Particle diameter is 433 Å” with 433 as the integer of  $3.39 \times 128 = 433.92$  to include most signal/noise in the box.
- v. Select the following options “Volumes”, “Invert contrast”, “Normalize input images”, “Normalize output volumes”, “Make paths relative to STAR”, and “Include items outside of filter range”.
- vi. Rename the subtomogram files from the Warp convention (e.g., f00003\_0000001\_3.39A.mrc) to the Dynamo convention (e.g., particle\_000100.em) and convert the mrc files to em files for Dynamo.

**Note:** Multiple methods such as the tom\_mrcread and tom\_emwrite functions in the TOM package<sup>9</sup> within MATLAB, the mrc2em module in PyTom,<sup>18,19</sup> or the e2proc3d.py in EMAN2<sup>20</sup> can be used to perform this conversion.

11. Align the 2× binned subtomograms and prepare for alignment of unbinned data.
  - a. Create a new Dynamo ABP project, following the same procedures at step 9.
    - i. Generate a new mask twice the size of that used for 4× binning.
    - ii. Decide on the initial lowpass frequency by calculating from the new odd/even averages at 2× binning.
    - iii. Use smaller angular and shift search ranges.

**Note:** When the ABP project at 2× binning (pixel size: 3.39 Å) converges, i.e. reaches a resolution of 13 Å in this case (Figure 4C, right), continue to prepare for alignment of unbinned subtomograms (pixel size: 1.69 Å).

**Note:** Originally, we aligned the 2× binned subtomograms in TOM/AV3 instead of Dynamo for speeding up the alignment in parallelized jobs essentially on the cluster (not working efficiently in Dynamo versions installed on our cluster with the corresponding setup).

- b. Clean up clashing subtomograms (particles that have shifted to occupy the same position in the helical filament) based on their distances to each other and their cross-correlation (cc) values within MATLAB.<sup>21</sup>

**Note:** Here, a low threshold (10 pixels at 2× binning, i.e. 3.39 nm) was used after checking the histogram of distances of subtomograms to their 20 nearest neighbors. For clashing subtomograms, the ones with higher cc values were retained. 1209 subtomograms were kept for the next step processing.

- c. Convert the cleaned Dynamo table to a STAR file for Warp and reconstruct the unbinned subtomograms in Warp following the same procedures at step 10.
- d. Perform one round of Dynamo alignment to make sure that the cleaned table/STAR file works before proceeding to RELION in the next steps.

### Subtomogram alignment and helical refinement with unbinned data

⌚ Timing: 11 days

This section describes settings for aligning the unbinned subtomograms (pixel size: 1.69 Å) with refined positions and orientations obtained from the steps above in RELION, including 3D auto-refinement, helical search, 3D auto-refinement with helical reconstruction, and further 3D classification to distinguish different conformational states of the filaments (Figure 5; RELION version: 3.1.0).

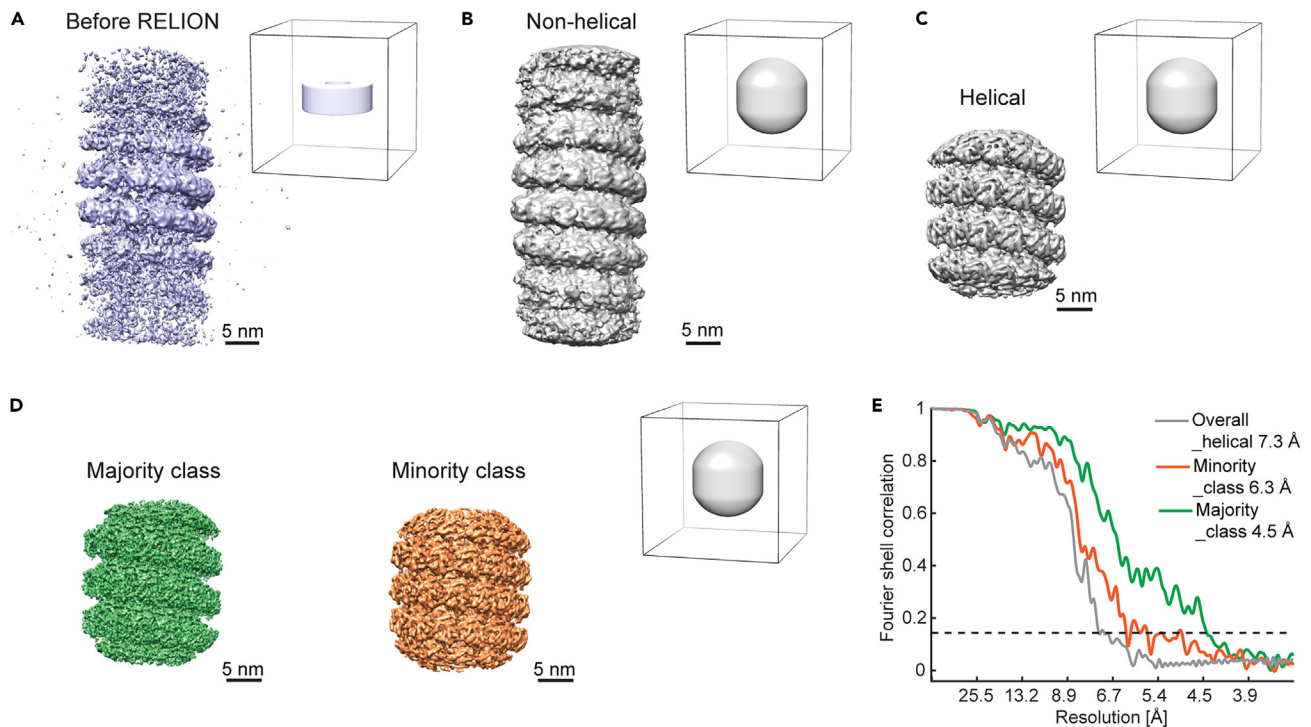
12. Run RELION 3D auto-refine without helical reconstruction.

**Note:** We recommend referring to the guidelines for RELION 3D auto-refine ([https://www3.mrc-lmb.cam.ac.uk/relion/index.php/Refine\\_a\\_structure\\_to\\_high-resolution](https://www3.mrc-lmb.cam.ac.uk/relion/index.php/Refine_a_structure_to_high-resolution)).

- a. The Warp-generated STAR file from the last step is in a format compatible with earlier RELION versions. For RELION version 3.1 or later, use the `relion_convert_star` command (with the RELION module loaded) to convert an old STAR file format to a new one whenever needed. [Troubleshooting 3](#).

⚠ **CRITICAL:** New-style RELION STAR files (version 3.1 onwards) mainly differ from old formats in: 1) the new-style includes the optics group part; 2) the new-style uses `OriginX/Y/ZAngst` columns to replace the `OriginX/Y/Z` and the units change from pixels to angstrom. Thus, these values in the new STAR files equal values in `OriginX/Y/Z` × pixelSize in angstrom. We used a MATLAB script for the conversion to old STAR for Warp. The `dynamo2m` package enables this with python scripting (`relion_star_downgrade` command).

- b. Convert the average from the last round of Dynamo alignments from a .em file to a .mrc file in Chimera.<sup>12</sup>
  - i. Use the Volume Viewer/Tools/Volume Filter to scale the volume by −1.
  - ii. As the Dynamo average likely does not contain pixel size information, set the correct pixel size (1.69 Å) with Chimera under Volume Viewer/Features/Coordinates.
  - iii. Save the map as a new map named `ref_unbin.mrc` (Figure 5A).



**Figure 5. Subtomogram averages imported into RELION for refinement and helical reconstruction**

Pixel size is 1.69 Å (unbinned data) for all averages.

(A and B) Half map, Gaussian filtered, without multiplication by the alignment mask (shown next to each map).

(C and D) Half map, filtered to the determined resolutions in RELION Post-processing jobs, with the alignment mask (next to the map) multiplied. The maps in (D) are results for each class, obtained by M refinement after classification. All maps are Y flipped to the correct handedness.

(E) Masked FSC curves for the maps in (C) and (D), using the same alignment mask. The dotted line indicates the 0.143 threshold. Parts of figures in Zhang et al.<sup>1</sup> published under a CC-BY license are reused.

- c. Generate a cylinder mask with RELION. Use the command as following:

```
>relion_helix_toolbox --cylinder --o cylmask256_o220_s55_w5.mrc --boxdim 256 --cyl_outer_diameter 220 --angpix 1.69 --sphere_percentage 0.55 --width 5
```

⚠ **CRITICAL:** Check the RELION instructions for detailed explanation about the mask parameters. In brief, measure the filament outer diameter based on ref\_unbin.mrc. The combination of boxdim (in pixels), cyl\_outer\_diameter (in Å), sphere\_percentage gives a central Z length of about 30%, which is optimal for RELION helical reconstruction in later steps. Width (in pixels) refers to the value for creating a soft-edged mask.

- d. Set up the RELION job in a GUI. Leave the setup parameters as default unless specified otherwise.
  - i. Specify the newly converted STAR file, the ref\_unbin.mrc, and the cylinder mask file in the I/O tab.
  - ii. In the Reference tab, leave the "Ref. map is on absolute grayscale?" as No because the map was not generated in RELION and thus may have a different absolute intensity grayscale from the RELION requirement.
  - iii. Set the Initial low-pass filter as 30 Å. In the Optimisation tab, set the Mask diameter as 300 Å and the "Use solvent-flattened FSC?" as Yes.
  - iv. In the Auto-sampling tab, set the Initial angular sampling as 3.7 degrees and the Local searches from auto-sampling as 3.7 degrees.



- v. For the Compute and Running tab, set parameters according to the computing environment.

**Note:** In our case, we copied the particles and their corresponding 3D CTF models to a directory on the cluster and set the **Number of MPI procs**, **Number of threads**, **GPU count**, **Memory**, **Walltime** and **Minimum dedicated cores per node** as 3, 7, 2, 120 GB, 3-23:00:00, 15, respectively.

- vi. Click on **Run** to submit the RELION job to a cluster. It took 33 iterations to converge in this case.
- e. In RELION Post-processing, provide **One of the 2 unfiltered half-maps** and the refinement mask as the **Solvent mask**, and a B-factor (e.g.,  $-500 \text{ \AA}^2$ ). When the job is done, check the estimated final resolution in the output window, as well as in the output files.

**Note:** The resulting resolution was  $10 \text{ \AA}$  with the refinement mask applied, showing clear separation of subunits along the helical turns (Figure 5B).

13. Perform helical parameter search in RELION.
  - a. With the output average from the Post-processing job in step 12, estimate the helical rise and twist parameters by measuring the helical pitch and counting the number of subunits per turn.

**Note:** Rise equals the helical pitch in  $\text{\AA}$  divided by the number of subunits per turn. Twist equals  $360$  degrees divided by the number of subunits per turn, with a positive sign for a right-handed helix, which is inverse to the actual handedness as explained above.

- b. Set the ranges in the following command line:

```
>relion_helix_toolbox--search--i xxx.mrc--cyl_outer_diameter 220--angpix 1.69--rise_min 3.5
--rise_max 4.5--twist_min 25--twist_max 35--z_percentage 0.3
```

**Note:** The output from this command is: rise  $4.46 \text{ \AA}$ ; twist  $27.08$  degrees.

14. Run RELION 3D auto-refine with helical reconstruction.

**Note:** Read the helical processing instructions for more details and conventions ([https://www3.mrc-lmb.cam.ac.uk/relion/index.php/Helical\\_processing](https://www3.mrc-lmb.cam.ac.uk/relion/index.php/Helical_processing)).

The settings are similar to those in step 12, except for the following.

- a. Use the output STAR file from the 3D auto-refine job in step 12.
- b. Use the output average from the Post-processing job in step 12 as reference map.
- c. Set up the RELION job in a GUI:
  - i. Specify the STAR file, the reference map, and the reference mask file (the same as in step 12) in the I/O tab.
  - ii. In the Reference tab, leave the "Ref. map is on absolute grayscale?" as Yes.
  - iii. In the Helix tab, change the following parameters:
    - Do helical reconstruction?* Yes;
    - Tube diameter – inner, outer ( $\text{\AA}$ ):*  $-1, 220$ ;
    - Keep tilt-prior fixed:* Yes;
    - Number of unique asymmetrical units:* 13;
    - Initial twist (deg), rise ( $\text{\AA}$ ):*  $27.08, 4.46$ ;
    - Do local searches of symmetry?* Yes;

*Twist search – Min, Max, Step (deg): 25, 29, 1;*

*Rise search – Min, Max, Step (Å): 3.8, 4.6, 0.05.*

- iv. In the Running tab, increase the **Walltime** as this job took longer than the refinement job without helical reconstruction with the RELION version used.
- v. Click on **Run** to submit the RELION job to a cluster.
- d. In RELION Post-processing, use one of the two unfiltered half-maps of the previous step and the refinement mask as the Solvent mask, and perform B-factor estimation automatically.

**Note:** The resulting resolution was 7.3 Å with an “Auto B-factor” of  $-382 \text{ Å}^2$ . Re-doing helical search gives a helical parameter set of 27.16 degrees and 4.17 Å (Figure 5C).

15. Run RELION 3D classification with helical reconstruction to check for heterogeneity. Settings are left as default, except for the changes in the following:

- a. Use the output STAR file from the 3D auto-refine job in step 14.
- b. Use the output average from the Post-processing job in step 14 as the reference map.
- c. Set up the RELION job in the GUI:
  - i. Specify the STAR file, reference map, and the cylinder reference mask file (the same as in step 12) in the I/O tab.
  - ii. In the Reference tab, leave the “Ref. map is on absolute grayscale?” as Yes. Set the Initial low-pass filter as 40 Å.
  - iii. In the Optimisation tab, set the **Number of classes** as 4, the **Mask diameter** as 300 Å.
  - iv. In the Sampling tab, set the **Angular sampling interval** as 1.8 degrees, the “Perform local angular searches?” to Yes and use a **Local angular search range** of 10 degrees.
  - v. In the Helix tab, change the following parameters:
 

*Do helical reconstruction?* Yes;

*Tube diameter – inner, outer (Å):* –1, 220;

*Keep tilt-prior fixed:* No;

*Number of unique asymmetrical units:* 13;

*Initial twist (deg), rise (Å):* 27.16 and 4.17;

*Do local searches of symmetry?* Yes;

*Twist search – Min, Max, Step (deg):* 26, 28, 1;

*Rise search – Min, Max, Step (Å):* 3.3, 4.5, 0.05.
- d. Inspect the results.

**Note:** The job resulted in 4 classes, of which Class 3 contained 77.7% of the particles (939 particles) and had a resolution of 5.9 Å, refined helical twist of 27.17 degrees, rise of 4.21 Å. Class 4 contained 21.3% of the particles (258 particles) and had a resolution of 6.6 Å, refined helical twist of 26.93 degrees, rise of 3.51 Å. The other two classes containing 12 particles in total did not result in any meaningful averages.

16. Run RELION 3D auto-refine with helical reconstruction for Classes 3 and 4 separately. Settings are similar to those in step 14, except for changes in the following:

- a. Use the STAR file for particles in Class 3 from the classification job, and that for Class 4 in a separate job.
- b. Use the output averages from the Classification job as Reference maps (averages from Class 3 and 4 separately for two jobs).
- c. Set up the RELION job in a GUI:
  - i. Specify the STAR file, the Reference map, and the cylinder Reference mask file (the same as in step 12) in the I/O tab.
  - ii. In the Reference tab, leave the “Ref. map is on absolute grayscale?” as Yes. Set the Initial low-pass filter as 20 Å.
  - iii. In the Helix tab, change the following parameters:

Do helical reconstruction? Yes;  
 Tube diameter – inner, outer (Å): –1, 220;  
 Keep tilt-prior fixed: Yes;  
 Number of unique asymmetrical units: 13;  
 Initial twist (deg), rise (Å): 27.17, 4.21 for Class 3 and 26.93, 3.51 for Class 4;  
 Do local searches of symmetry? Yes;  
 Twist search – Min, Max, Step (deg): 25, 29, 1;  
 Rise search – Min, Max, Step (Å): 3.8, 4.4, 0.05 for Class 3 and 3.4, 4.4, 0.05 for Class 4.

**Note:** The “Keep tilt-prior fixed” option is set to Yes in principle; if the RELION job aborts with the error “Segmentation fault”, set this to No and re-run it.

d. Run RELION Post-processing and inspect the results.

**Note:** The resulting resolution for Class 3 was 6.5 Å with an Auto B-factor of  $-85 \text{ Å}^2$ . Re-doing helical search for the Class 3 average gives a helical parameter set of 27.17 degrees and 4.21 Å. The resolution for Class 4 is 7.3 Å with an Auto B-factor of  $-161 \text{ Å}^2$ . Re-doing helical search for the Class 4 average gives a helical parameter set of 26.91 degrees and 3.51 Å (final refinement maps of these two classes are shown in [Figure 5D](#), with the corresponding FSC curves in [Figure 5E](#)).

**△ CRITICAL:** EM maps at a resolution of 8 Å or higher should exhibit some features of secondary structures such as alpha helices, which should be critically assessed along the alignment/refinement steps, rather than considering only the reported resolution.

## Multi-particle refinement and map evaluation

⌚ Timing: 3 days

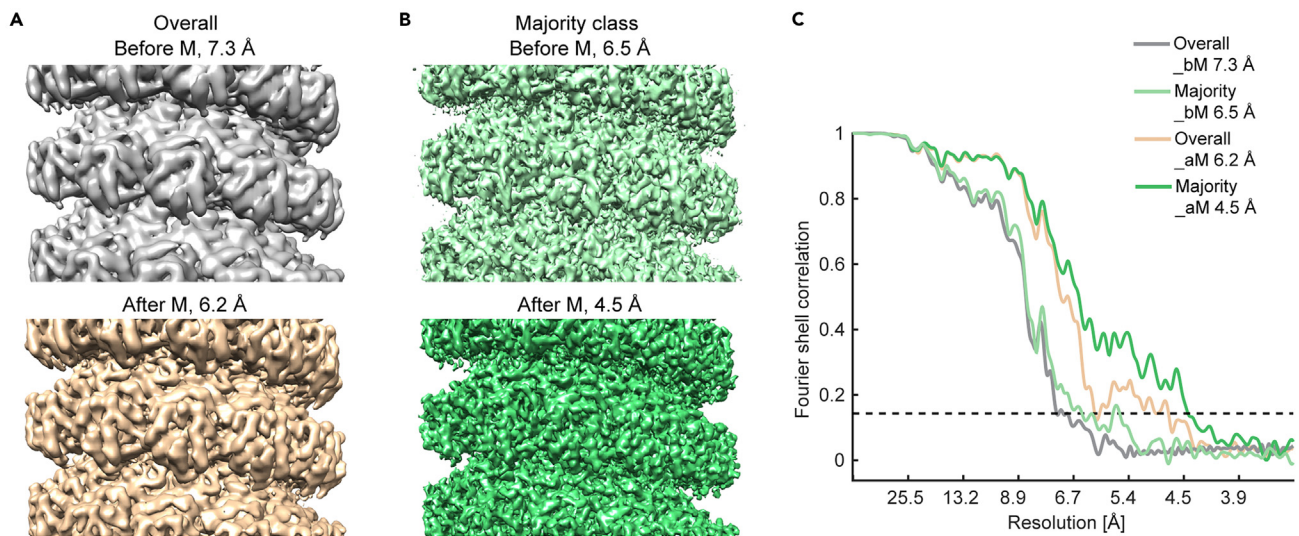
Based on the refinement results above, this section shows how to perform multi-particle refinement in M (following helical symmetry expansion with RELION, as M versions until 1.0.9 do not support helical symmetry operation) and rerun RELION 3D auto-refine with the refinement results from M. At the end, generate FSC plots and examine local resolution of maps.

17. Prepare files for M refinement.

- Make a copy of the original Warp preprocessing, especially the .xml and .tomostar files as a backup. M refinement will overwrite some of the files.
- Copy the 2 half maps and particle STAR file from the RELION 3D auto-refine job to a folder designated for M processing (ideally as a subfolder in the directory that contains the Warp frames folder): `it0xx_half1_class001_unfil.mrc`, `it0xx_half2_class001_unfil.mrc`, and `run_it0xx_data.star` (renamed to `run_it0xx_data_ori.star`)
- For an M version that does not support helical symmetry operation, perform helical symmetry expansion with the RELION command:

```
>relion_particle_symmetry_expand --i run_it0xx_data_ori.star --o run_it0xx_data_expanded.star --helix true --twist 27.17 --rise 4.21 --angpix 1.69 --asu 13
```

**Note:** This generates a new STAR file with an expanded particle list according to the symmetry (13-fold larger). Downgrade the RELION STAR file with the `relion_star_downgrade` command in the `dynamo2m` toolbox and rename the STAR file to `jobxxx_run_it0xx_data.star` for M refinement (M version 1.0.9 recognizes the suffix of the file as `_data.star`).



**Figure 6. Comparison of the subtomogram averages before and after M refinement**

Shown are the subtomogram averages before (top) and after (bottom) M refinement, from all the particles (A) or those belonging to the majority class (B), along with their masked FSC curves (C). The alignment mask shown in Figure 5D was used for refinement and post-processing of all averages shown here. The dotted line in (C) indicates the 0.143 criterion. bM, before M refinement; aM, after M refinement. Numbers refer to the determined map resolution. All maps are Y flipped to the correct handedness. Parts of figures in Zhang et al.<sup>1</sup> published under a CC-BY license are reused.

- d. Copy the cylinder mask that was used for RELION refinement to the M processing folder.
18. Run M refinement.
  - a. Open the Warp session before running M refinement to make sure the correct settings/input files can be imported by M.
  - b. Close Warp and open M.
  - c. Create a new population in the working directory that contains the Warp frames folder.
    - i. Link the new population to the data source by clicking "Manage data sources" and then "ADD LOCAL" at the interface.
    - ii. Choose and import the "previous.settings" file from the previous Warp processing folder.
    - iii. Set a name for the data source. Choose to "Include items outside of filter ranges" to include all data selected.
    - iv. Continue to click "Create" and close the pop-up interface when done. A .source file will be created within the frames folder by default.
  - d. Click the "+" icon to add species, for example the Class 3 average obtained in step 16.
    - i. Use the "From scratch" option for the first M refinement job and click on "ADD".
    - ii. Set a name for the species and use a parameter set of 220 Å diameter (filament diameter), 2600 kDa (the estimated molecular weight of the region included in the refinement mask; estimated based on 39 protein subunits plus  $6 \times 39 = 234$  nucleotides for RNA, with a small margin), C1 symmetry, 1 temporal sample for poses.
    - iii. Select the half-maps from the M processing folder above for the Class 3 average for example and input the correct pixel size. Use the "filter both to Å" option to visualize the input maps (filtered to a resolution lower than the estimated value, for inspection purpose).
    - iv. Set the mask file with the cylinder mask in the M processing folder.
    - v. Specify the symmetry expanded and downgraded STAR file. Input the pixel size for coordinates and shifts in the STAR file which relates to the binning factor at which alignments are performed.
    - vi. Select the source file under "Use these sources for name matching". M then displays the correct particle number read from the STAR file that can be matched to this data source file.

**Note:** All particles should be found if the input STAR file format is correct and the source file contains all corresponding tomograms.

- vii. Click on “FINISH” when done with these input settings. M starts preprocessing steps, including denoising, which may take 10 min.

**Note:** A GPU memory issue may cause the software crash at this step (GPU memory of about 24 GB or above is recommended).

- e. M reports an initial resolution comparable to that obtained from RELION auto-refinement job. Click on the “REFINE” button on the upper right corner in the M GUI, to set/enable the following refinement settings in the first round:  
*Refine for 3 sub-iterations, use 70% of available resolution in first sub-iteration.*  
*Image warp grid: 4 × 4.*  
*Particle poses.*  
*Stage angles.*  
*Volume warp grid: 3 × 3 × 2 × 10.*

**Note:** This provides an improvement in resolution from 6.5 Å to 6.1 Å for Class 3 as an example. Continue with a number of additional rounds of M refinement until the resolution stops improving.

- f. Add the following refinement options for a second stage of M refinement:  
*Anisotropic magnification.*  
*Tilt movies.*  
*Use species with at least 7.0 Å resolution.*  
*Defocus and grid search in 1<sup>st</sup> iteration.*
- g. Continue with a few additional rounds of M refinement until the resolution stops improving.

**Note:** When the refinement is done, M gives an estimated resolution of 4.8 Å in the refined map.

- h. Inspect the quality of the refined and postprocessed maps, as well as the corresponding FSC curves generated by M refinement in the species folder.

**△ CRITICAL:** The *fitsharp* and *denoised* maps should show improved details in the maps compared to the input map, e.g. better recognizable helices or even large side chains of residues at this resolution range. The FSC curve should drop to zero at the high frequency range and not show “bumps”.

19. Perform RELION refinement after M. This further improved the resolution of the EM map for the majority class from 4.8 Å to 4.5 Å.
  - a. Convert the STAR file from M (\_particles.star in the latest M species folder) for reconstruction of particle/subtomograms in Warp with the M-refinement-updated poses.

**Note:** We performed this with a MATLAB script that converts the coordinates unit from Å back to pixel, removes 12 extra positions of the symmetry-related particles per 13 particles. Modify the data columns in the STAR file using a previous STAR file for Warp as a template.

- b. Reconstruct subtomograms and their 3D CTF models in Warp with this STAR file and convert the STAR file again for RELION versions above 3.1.
- c. Prepare a new reference from the M refinement result.

**Note:** The M generated average (the filtered and sharpened *fitsharp* map in the species folder) does not use the same box size as the subtomogram particles for RELION.

- i. Open a previous RELION average (read as #0) and the M average (read as #1) in the same Chimera window.
- ii. Superpose the two averages by the “fit in map” option.
- iii. Use the following command to rescale the box size and set the origin for the M average:

```
>vop resample #1 onGrid #0
```

- iv. Save the resampled average in .mrc format and use it as an input reference in a new RELION 3D auto-refine job.
- d. Run a new RELION 3D auto-refine job with helical reconstruction, with settings similar to step 16.
    - i. Use a smaller auto-sampling angle as the particle orientations have been refined extensively in previous steps.
    - ii. Do post-processing for this new average to obtain a final refined map.
    - iii. Search for helical parameters again with the post-processed average.

**Note:** The helical parameters remained almost the same after M refinement compared to those before, but the calculated resolution with the 0.143 criterion and map quality improved substantially with M refinement (Figure 6).

**Note:** M version 1.1 beta could support helical refinement but needs to be tested. According to the software developer, one has to modify the “Symmetry” Value (e.g., Hel 27.17|4.21) in the .species file as setting the helical parameters is not yet supported directly in the GUI.

**△ CRITICAL:** In case M refinement does not improve the resolution of maps, but rather deteriorates it, make a copy of all files from an earlier version of that species (contained in [species folder]/versions, sort by date to get the version needed) except for the [species name]\_particles.star file, to get back to the maps with higher resolution and try other combinations of refinement options.

**Note:** For non-helical objects, one can use the export sub-tomograms function in M instead of resorting back to Warp with a converted STAR file.

**Note:** For M refinement of particles of a small number (e.g. Class 4 from step 16), set up a refinement job with two species using STAR files and half maps for both species (Classes 3 and 4 together). Refinement worked when including only the image warp grid, particle poses, stage angles, and volume warp grid options. We continued to reconstruct subtomograms after M refinement and run a new RELION 3D auto-refine job for the results described here.

20. Map quality evaluation and model building.
  - a. Run RELION Post-processing job to obtain FSC plots based on the unfiltered half maps, with or without a given mask.

**Note:** The resolutions we reported as 4.5 Å for the majority class average and 6.3 Å for the minority class average were derived with the cylinder refinement mask applied. The maps after M-RELION refinement show higher quality in that secondary structures can be easily visualized in the final maps.

- b. Run a RELION Local resolution job to calculate the local resolution and render the map accordingly in Chimera/ChimeraX.<sup>13</sup>
- c. Prepare for building a structural model based on the final map and check the model-to-map cross-correlation.

**Note:** As mentioned above, we generated the maps with inverted handedness using Warp 1.0.7 beta. Run the command line below to obtain maps with correct handedness:

```
>relion_image_handler --i xxx.mrc --o xxx_flipY.mrc --flipY
```

- d. For building a model, as detailed in Zhang et al.,<sup>1</sup>
  - i. Generate an initial model using homology with I-TASSER<sup>22</sup> or using AlphaFold.<sup>23</sup>
  - ii. Use the generated model as a starting point for iterative automated and manual refinement in Phenix<sup>24</sup> and Coot.<sup>25</sup>
  - iii. Generate a segmented map for the region corresponding to the tetramer model with the **volume zone** command in ChimeraX.
  - iv. Validate the real-space per-residue cross-correlation of the tetramer model to the map generated in Phenix.

**Note:** A tetramer model was built in order to better analyze the interaction interface among monomers of the same turn as well as that between the upper and lower turns.

## EXPECTED OUTCOMES

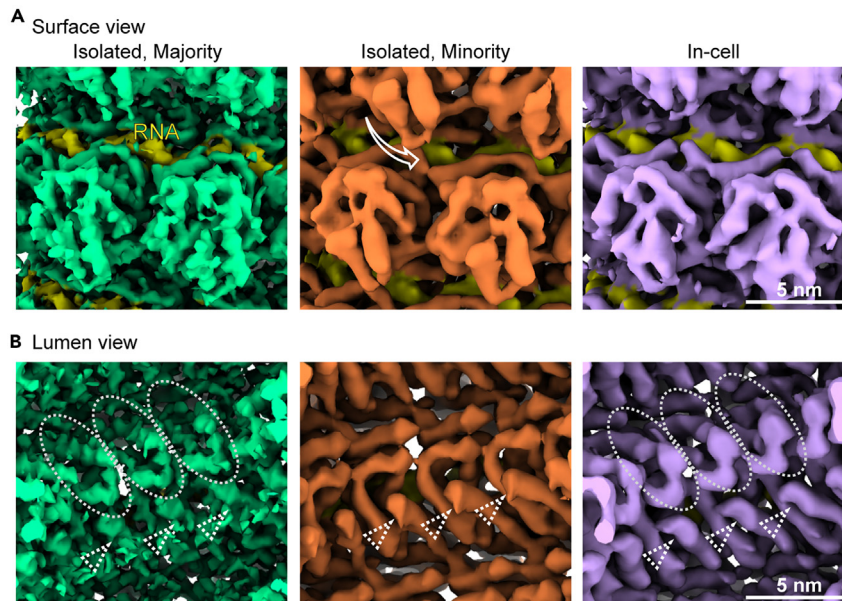
Based on 20 cryo-tomograms, which contained about 20 viral nucleocapsids, we extracted 1297 subtomograms from the relatively straight nucleocapsid segments and successfully obtained 2 reconstructions at 4.5 Å and 6.3 Å. The 4.5 Å map was of sufficient quality to serve as the basis for the building of an atomic model of the mumps viral nucleocapsid protein (N terminal 3–405 amino acids) with its RNA hexamer as one subunit for the assembly of the helical filamentous structure. We then used this model to interpret the 6.3 Å lower resolution minority class reconstruction by careful visual inspection in Chimera/ChimeraX, and provided information on the differences in helical pitch and assembly compactness between the two maps. Comparing the two models in detail revealed differences in the interaction interfaces between monomers leading to the different assemblies, as well as their conformations that indicate differences in the surface accessibility of the bound RNA (genome) and C-terminal regions (Figure 7).

Beyond the example dataset of extracted nucleocapsids described in this protocol, we have also processed cellular cryo-ET data of relatively straight nucleocapsids using steps similar to those discussed here, with the following adjustments: 1) the tracing of the helical filaments was performed in Amira using dedicated functions<sup>1</sup>; 2) the generation of equidistant sampling points along the helical segments was done with a MATLAB script<sup>26</sup>; 3) initial alignment of the subtomograms required more strict control of the Euler angles to avoid change of directionality of subtomograms along their respective filaments; 4) checking the directionality of each filament was complicated by the lower signal-to-noise ratio of the cellular cryo-ET data compared to that of the isolated filaments. Thus, nucleocapsids with unclear directionality based on the inspection of the per-filament average were largely discarded at the initial alignment steps for achieving high resolution maps. Nevertheless, a consensus map of 6.5 Å was finally obtained from 5 tomograms, 68 helical filaments and 2178 subtomograms.

## LIMITATIONS

Transfer of the workflow described in this protocol to the processing of different filament types and morphologies is expected to be successful, provided that the filaments are sufficiently straight or





**Figure 7. Comparison of the subtomogram averages of two classes of the isolated nucleocapsids and that of the in-cell ones**

The majority (green) and minority (orange) class averages of the isolated nucleocapsids and the average of the in-cell nucleocapsids (purple) are compared from the outer surface view (A) and lumen view (B). RNA is depicted in yellow in all images. The white arrow in (A) (middle) indicates movement of subunits in the upper turn towards the lower turn in the minority class average, in comparison to the other two averages. The white circles in (B) denote a region in the density which is not seen in the minority class average possibly due to flexibility, in contrast to density consistently seen in all averages (white arrowheads). Parts of figures in Zhang et al.<sup>1</sup> published under a CC-BY license are reused.

contain straight regions. Depending on the filament types and available software in your local settings, data processing workflows that integrate other software, such as those described for the processing of LRRK2-microtubule filaments,<sup>27</sup> the COPII coat reconstituted on lipid membranes,<sup>2,28</sup> the Ebola virus nucleocapsid,<sup>29</sup> or the skeletal thin filaments<sup>30</sup> can also be tested.

## TROUBLESHOOTING

### Problem 1

Uncertainty about the options for correcting gain reference during Warp preprocessing (related to step 1).

#### Potential solution

- The correct gain options need to be tested according to the camera setting and serialEM data saving options. Warp indicates whether the "Transpose" option is needed automatically. According to the "r/f" values of the raw tif images ([https://bio3d.colorado.edu/SerialEM/hlp/html/about\\_camera.htm#norm\\_frames](https://bio3d.colorado.edu/SerialEM/hlp/html/about_camera.htm#norm_frames)), different combinations of "Flip X axis" and "Flip Y axis" are needed.
- Using images with good contrast, test the flip combinations by processing a few movies without motion correction until a visible gain pattern disappears in the average images. Or, generate a sum of dozens of averages without motion correction. Check the gain pattern in this summed image and compare it to the gain reference image to decide on the flip options.

### Problem 2

Inconsistency between Warp and IMOD reconstruction (related to step 4).

### Potential solution

- Check the views excluded during IMOD reconstruction carefully and remove the corresponding views in the mdoc file accordingly. Even if the Warp 1.0.9 version does not report errors with the unmodified mdoc files (with the full tilt series) and the Etomo alignment files with excluded tilt views, artifacts can appear in the reconstructed tomograms due to dark regions in those views.
- Alternatively, after considering the views to exclude, remove them in the mdoc file and redo the tilt series generation step in Warp for a second time processing to avoid any mismatch.
- When using AreTomo or other software that performs automatic reconstruction, it is recommended to first test with one tilt series to solve possible inconsistency before batch processing.

### Problem 3

RELION refinement errors related to particle path not found or STAR file style (related to step 12).

### Potential solution

- It is important to update the path of subtomograms and 3D CTFs in the STAR files whenever the data path or folder name is changed. Using an absolute path instead of a relative path mostly avoids problems.
- The new STAR file style poses an issue of conversion when using Warp, RELION, and M. Warp/M until version 1.0.9 takes only the old-style STAR file, while RELION from version 3.1 onwards uses the new style. Although RELION 3.1+ versions may automatically upgrade old-style ones, errors reported as “not all necessary variables defined in \_optics.star file” could still occur. Pay special attention to the rlnPixelSize, rlnVoltage and rlnSphericalAberration values and manually add these lines if not found in the STAR file after using the relion\_convert\_star command.

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Julia Mahamid ([julia.mahamid@embl.de](mailto:julia.mahamid@embl.de)).

### Technical contact

Further technical questions on data processing should be directed to and will be answered by the technical contact, Xiaojie Zhang ([zhangxj2@shanghaitech.edu.cn](mailto:zhangxj2@shanghaitech.edu.cn)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- Raw micrographs are available in the Electron Microscopy Public Image Archive (EMPIAR) under accession code EMPIAR: 10751 (isolated nucleocapsids). Intermediate processing files are deposited to Mendeley Data (<https://doi.org/10.17632/d95fvg4dnm.1>). Cryo-EM maps are deposited in the Electron Microscopy Data Bank (EMDB) under accession codes EMD: 13133 (majority class, isolated nucleocapsids) and EMD: 13136 (minority class, isolated nucleocapsids). The associated atomic model is deposited in the Protein Data Bank (PDB) under accession code PDB: 7OZR.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) or [technical contact](#) upon request.

## ACKNOWLEDGMENTS

We thank the EMBL EMCF, central IT, and Thomas Hoffmann for computational support and Wim Hagen, Felix Weis, and the cryo-EM platform for support in data acquisition. We thank Mahamid

group members for discussions and help. We are grateful to Giulia Zanetti, Daniel Castano-Diez, and Beata Turonova for providing initial advice on subtomogram averaging strategies and Dimitry Tegunov for developing helical symmetry functionality in M. J.M. acknowledges funding from the EMBL and a European Research Council starting grant (3DCellPhase- 760067).

## AUTHOR CONTRIBUTIONS

X.Z. and J.M. conceived this work and wrote the manuscript. X.Z. generated data, performed processing and analysis, and prepared figures.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

- Zhang, X., Sridharan, S., Zagoriy, I., Eugster Oegema, C., Ching, C., Pflaesterer, T., Fung, H.K.H., Becher, I., Poser, I., Müller, C.W., et al. (2023). Molecular mechanisms of stress-induced reactivation in mumps virus condensates. *Cell* 186, 1877–1894.e27. <https://doi.org/10.1016/j.cell.2023.03.015>.
- Zanetti, G., Prinz, S., Daum, S., Meister, A., Schekman, R., Bacia, K., and Briggs, J.A.G. (2013). The structure of the COPII transport-vesicle coat assembled on membranes. *eLife* 2, e00951. <https://doi.org/10.7554/eLife.00951>.
- Hagen, W.J.H., Wan, W., and Briggs, J.A.G. (2017). Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. *J. Struct. Biol.* 197, 191–198. <https://doi.org/10.1016/j.jsb.2016.06.007>.
- Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* 152, 36–51. <https://doi.org/10.1016/j.jsb.2005.07.007>.
- Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* 116, 71–76. <https://doi.org/10.1006/jsbi.1996.0013>.
- Tegunov, D., and Cramer, P. (2019). Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* 16, 1146–1152. <https://doi.org/10.1038/s41592-019-0580-y>.
- Tegunov, D., Xue, L., Dienemann, C., Cramer, P., and Mahamid, J. (2021). Multi-particle cryo-EM refinement with M visualizes ribosome-antibiotic complex at 3.5 Å in cells. *Nat. Methods* 18, 186–193. <https://doi.org/10.1038/s41592-020-01054-7>.
- Castano-Diez, D., Kudryashev, M., Arheit, M., and Stahlberg, H. (2012). Dynamo: a flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. *J. Struct. Biol.* 178, 139–151. <https://doi.org/10.1016/j.jsb.2011.12.017>.
- Nickell, S., Förster, F., Linaroudis, A., Net, W.D., Beck, F., Hegerl, R., Baumeister, W., and Plitzko, J.M. (2005). TOM software toolbox: acquisition and analysis for electron tomography. *J. Struct. Biol.* 149, 227–234. <https://doi.org/10.1016/j.jsb.2004.10.006>.
- Forster, F., and Hegerl, R. (2007). Structure determination in situ by averaging of tomograms. *Methods Cell Biol.* 79, 741–767. [https://doi.org/10.1016/S0091-679X\(06\)79029-X](https://doi.org/10.1016/S0091-679X(06)79029-X).
- Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* 7, e42166. <https://doi.org/10.7554/eLife.42166>.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. <https://doi.org/10.1002/jcc.20084>.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. <https://doi.org/10.1002/pro.3943>.
- Burt, A., Gaifas, L., Dendooven, T., and Gutsche, I. (2021). A flexible framework for multi-particle refinement in cryo-electron tomography. *PLoS Biol.* 19, e3001319. <https://doi.org/10.1371/journal.pbio.3001319>.
- Zheng, S., Wolff, G., Greenan, G., Chen, Z., Faas, F.G.A., Bárcena, M., Koster, A.J., Cheng, Y., and Agard, D.A. (2022). AreTomo: An integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction. *J. Struct. Biol.* 6, 100068. <https://doi.org/10.1016/j.jsbx.2022.100068>.
- Buchholz, T.O., Krull, A., Shahidi, R., Pigino, G., Jékely, G., and Jug, F. (2019). Content-aware image restoration for electron microscopy. *Methods Cell Biol.* 152, 277–289. <https://doi.org/10.1016/bs.mcb.2019.05.001>.
- Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* 333, 721–745. <https://doi.org/10.1016/j.jmb.2003.07.013>.
- Hrabe, T., Chen, Y., Pfeffer, S., Cuellar, L.K., Mangold, A.V., and Förster, F. (2012). PyTom: a python-based toolbox for localization of macromolecules in cryo-electron tomograms and subtomogram analysis. *J. Struct. Biol.* 178, 177–188. <https://doi.org/10.1016/j.jsb.2011.12.003>.
- Chaillet, M.L., van der Schot, G., Gubins, I., Roet, S., Veltkamp, R.C., and Förster, F. (2023). Extensive Angular Sampling Enables the Sensitive Localization of Macromolecules in Electron Tomograms. *Int. J. Mol. Sci.* 24, 13375. <https://doi.org/10.3390/ijms241713375>.
- Chen, M., Bell, J.M., Shi, X., Sun, S.Y., Wang, Z., and Ludtke, S.J. (2019). A complete data processing workflow for cryo-ET and subtomogram averaging. *Nat. Methods* 16, 1161–1168. <https://doi.org/10.1038/s41592-019-0591-8>.
- Poge, M., Mahamid, J., Imanishi, S.S., Plitzko, J.M., Palczewski, K., and Baumeister, W. (2021). Determinants shaping the nanoscale architecture of the mouse rod outer segment. *eLife* 10, e72817. <https://doi.org/10.7554/eLife.72817>.
- Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. <https://doi.org/10.1038/nprot.2010.5>.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
- Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* 75, 861–877. <https://doi.org/10.1107/S2059798319011471>.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501. <https://doi.org/10.1107/S0907444910007493>.
- de Teresa-Trueba, I., Goetz, S.K., Mattausch, A., Stojanovska, F., Zimmerli, C.E., Toro-Nahuelpan, M., Cheng, D.W.C., Tollervey, F., Pape, C., Beck, M., et al. (2023). Convolutional networks for supervised mining of molecular patterns within cellular context. *Nat. Methods*

- 20, 284–294. <https://doi.org/10.1038/s41592-022-01746-2>.
27. Watanabe, R., Buschauer, R., Bohning, J., Audagnotto, M., Lasker, K., Lu, T.W., Boassa, D., Taylor, S., and Villa, E. (2020). The In Situ Structure of Parkinson's Disease-Linked LRRK2. *Cell* 182, 1508–1518.e16. <https://doi.org/10.1016/j.cell.2020.08.004>.
28. Zivanov, J., Oton, J., Ke, Z., von Kugelgen, A., Pyle, E., Qu, K., Morado, D., Castano-Diez, D., Zanetti, G., Bharat, T.A.M., et al. (2022). A Bayesian approach to single-particle electron cryo-tomography in RELION-4.0. *eLife* 11, e83724. <https://doi.org/10.7554/eLife.83724>.
29. Wan, W., Kolesnikova, L., Clarke, M., Koehler, A., Noda, T., Becker, S., and Briggs, J.A.G. (2017). Structure and assembly of the Ebola virus nucleocapsid. *Nature* 551, 394–397. <https://doi.org/10.1038/nature24490>.
30. Wang, Z., Grange, M., Pospich, S., Wagner, T., Kho, A.L., Gautel, M., and Raunser, S. (2022). Structures from intact myofibrils reveal mechanism of thin filament regulation through nebulin. *Science* 375, eabn1934. <https://doi.org/10.1126/science.abn1934>.