

Negative-Stain Transmission Electron Microscopy of Molecular Complexes for Image Analysis by 2D Class Averaging

John R. Gallagher,¹ Alexander J. Kim,¹ Neetu M. Gulati,¹
and Audray K. Harris^{1,2}

¹Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases,
National Institutes of Health, Bethesda, Maryland

²Corresponding author: harrisau@mail.nih.gov

Negative-stain transmission electron microscopy (EM) is a technique that has provided nanometer resolution images of macromolecules for about 60 years. Developments in cryo-EM image processing have maximized the information gained from averaging large numbers of particles. These developments can now be applied back to negative-stain image analysis to ascertain domain level molecular structure (10 to 20 Å) more quickly and efficiently than possible by atomic resolution cryo-EM. Using uranyl acetate stained molecular complexes of influenza hemagglutinin bound to Fab 441D6, we describe a simple and efficient means to collect several hundred micrographs with SerialEM. Using RELION, we illustrate how tens of thousands of complexes can be auto-picked and classified to accurately describe the domain level topology of this unconventional hemagglutinin head-domain epitope. By comparing to the cryo-EM density map of the same complex, we show that questions about epitope mapping and conformational heterogeneity can readily be answered by this negative-stain method. © 2019 The Authors.

Keywords: electron microscopy • image analysis • image processing • negative stain

How to cite this article:

Gallagher, J. R., Kim, A. J., Gulati, N. M., & Harris, A. K. (2019). Negative-stain transmission electron microscopy of molecular complexes for image analysis by 2D class averaging. *Current Protocols in Microbiology*, 54, e90. doi: 10.1002/cpmc.90

INTRODUCTION

Technical advances in cryo-electron microscopy (cryo-EM) have drawn attention for atomic resolution structure determination of macromolecular complexes at a revolutionary pace (Kuhlbrandt, 2014). Yet, even as infrastructure supporting atomic resolution cryo-EM has expanded rapidly, obtaining time for data collection on high-resolution (field emission gun operated at 300 kV with a direct electron detector) transmission electron microscopes (TEM) can be a bottleneck to progress for a research project. Meanwhile, low-voltage (80-120 kV) TEMs have already been widely procured. Low-voltage TEM designs, such as the Philips CM10 and Tecnai 12, have been workhorses since the 1980s. Continued improvements in microscope components and hardware updates have improved vacuum robustness, stage stability, and filament brightness. Upgraded software has made microscope operation and alignment more accessible to non-experts.

Improvements in image acquisition have also been enormous, summarized by how digital cameras replaced film. Automatic drift correction, recently available in complementary metal-oxide-semiconductor (CMOS) cameras, can now compensate dramatically for stage instability. Automatic drift correction allows a staggering amount of data to be acquired quickly since it permits the stage stabilization delay between image acquisitions to be inconsequentially small. Negative-stain TEM is perfectly suited to a low-voltage TEM microscope, where the absolute resolution is limited by the grain size of the negative stain. Grain sizes range from 4 to 5 Å for uranium-based stains, which typically results in structural information on a biological complex to about 10 to 20 Å resolution (Scarff, Fuller, Thompson, & Iadaza, 2018).

In general, lower voltage microscopes (e.g., 100 kV) used for negative stain are easier to use and maintain than their higher voltage cryo-EM equivalents (e.g., 300 kV). As a result, low-voltage microscopes are often underutilized and undervalued compared to their high-voltage counterparts. Additionally, negative-stain images are inherently higher contrast than cryo-EM due to the heavy metal stain, and thus require fewer particles for data analysis. Even conformationally heterogeneous samples can be well suited to negative-stain analysis, as the image contrast for individual particles is sufficient to guide proper image classification (Bianchi et al., 2018; Ohi, Li, Cheng, & Walz, 2004). By using fewer particles, negative-stain data analysis can commonly be performed on a workstation computer, instead of a computer cluster commonly utilized for cryo-EM. Together, these attributes permit a streamlined rendition of the cryo-EM data collection and processing workflow to be applied to negative-stain data collection and analysis from low-voltage TEM microscopes already in service. The result of this revised negative-stain workflow is improved resolution of the biological complex, and a better understanding of the diversity of structures among the population of imaged complexes.

Information from 2D-image analysis of macromolecules from negative-stain EM images are important for a number of reasons. First, information about the molecular morphology, organization, and heterogeneity of samples can be obtained. Second, negative-stain structural data can be obtained for a larger number of samples in a given amount of time than possible with cryo-EM, allowing many samples to be screened before pursuing atomic resolution cryo-EM. Lastly, structural information from negative-stain imaging can be coupled with other information such as antibody competition data or functional data, to gain insights into the structure of macromolecules.

In this article, we discuss the steps following sample purification through generation of representative 2D class averages that visualize the most common orientations of the particle. This article aims to depict the workflow moving from particles in buffer to a large number of electron micrographs, and then analyzing these micrographs by 2D class averaging. 2D class averages enable robust conclusions founded upon the full dataset, rather than relying on individual particles in single micrographs to be representative. Furthermore, the increased signal-to-noise of negatively stained sample facilitates dramatically faster data acquisition and data analysis times than cryo-EM, enabling categorically different experimental questions to be addressed, such as epitope mapping of Fabs derived from polyclonal sera (Bianchi et al., 2018). This article targets existing users of low-voltage TEM microscopes who wish to add 2D classification to an existing workflow, as well as individuals who may have access to an underutilized low-voltage TEM microscope, yet do not have an established workflow. A primary motivation to implement a workflow such as described in this article is to draw overarching conclusions about the entire population of particles captured by EM imaging.

As an example for this article, we use the molecular complex of the antibody fragment antigen-binding (Fab) 441D6 in complex with influenza hemagglutinin (HA) A/New

York/653/1996. The epitope of the 441D6 antibody targets the immunodominant head region of influenza hemagglutinin, yet 441D6 is unique in that it has broad activity against viruses spanning 90 years of evolution by the H1 subtype of influenza. The structure of this complex was previously published as a cryo-EM map at 8 Å resolution (Kanekiyo et al., 2019), and this map is available in the EM data bank as entry EMD-7021. The HA ectodomain in this complex is just 10 nm tall, but the 3 Fabs bound to HA add substantial information to the images that aid in particle alignment during 2D classification. Basic Protocol 1 addresses the collection of images in a semi-automated fashion to arrive at sufficiently high particle counts for analysis. Support Protocol describes preparation of an EM grid by negative stain before images can be collected. Basic Protocol 2 enumerates the computational steps for analyzing a set of images, resulting in a set of representative 2D images summarizing all particles observed in the sample. Using the protocols presented here, structural information on different macromolecules could be obtained and coupled to other data from techniques such as microbiology, biochemistry, virology, and immunology. This would facilitate understanding of molecular structure and function for a broad-range of projects of interest to scientists across disciplines.

DATA ACQUISITION

With an EM grid prepared (see the Support Protocol), the first milestone is to collect a large number of micrographs, containing approximately 50,000 particles in total. Towards this goal, the first step is to evaluate if the grid is good, or, at the least, better than previously acquired data. If the grid is not good, a new grid should be remade before continuing. A good grid has good contrast for individual particles, and good particle dispersion (see Critical Parameters and Troubleshooting). When a grid of suitable quality is made, the objective becomes to collect a much data as possible in a short amount of time. The data should be collected close to focus to minimize the importance of the contrast transfer function (CTF) correction, yet with enough defocus to properly fit the defocus value. Using a camera that supports motion correction can greatly reduce the dwell time for stage settling between camera shots, such as implemented in the OneView camera (Gatan, Warrendale, PA).

Classically, a “micrograph” referred to the picture taken on film from the electron microscope. Now with digital cameras, the term “image” is commonly used, and often “image” and “micrograph” can be used interchangeably. “Micrograph” may be more appropriate when describing the immutable raw data for analysis, as compared to the processed images during analysis. Meanwhile “image” more commonly refers to everything coming from the electron microscope camera, especially data which is used only transiently while configuring the microscope.

Generally, collecting more micrographs is better for data analysis, but available compute resources will impose a limit on the dataset size that can be efficiently processed. Looking forward to data processing, 2D classification in RELION may behave differently with when 2D classifying large numbers of particles, compared to the same data processed in smaller batches. Yet working with multiple batches of 2D classifications is primarily recommended for curating giant datasets for 3D reconstruction by cryo-EM, not necessarily negative stain. If more than 50,000 particles are used in a single 2D classification, it can become difficult to parameterize the classification to create the desired separation between classes. For the high contrast of negative stain, 20,000 particles can be very effective, and as few as 2,000 to 5000 particles in a dataset can sometimes reveal important insights beyond what is visible in single micrographs. While sample-specific factors will influence the count of micrographs desired, if one assumes 150 particles per micrograph and a target of 20,000 total particles, then roughly 130 micrographs are needed.

This article assumes the EM microscope alignments and the software package SerialEM (Mastronarde, 2005) have been properly parameterized previously. Many resources for SerialEM are available at varying levels of engagement (see Internet Resources). We note that in a context of EM software that typically fluctuate in stability as they are developed, SerialEM software and support has categorically exhibited sustained excellence.

Materials

Liquid nitrogen

Negatively stained sample prepared on an EM grid (see Support Protocol)

FEI Tecnai 12 electron microscope (operated at 80, 100, or 120 kV), or equivalent TEM microscope

Digital camera containing:

Preferably with a 4k sensor size, or larger, to maximize particles per micrograph

Preferably with drift compensation, to avoid significant stage settling delays

SerialEM installation (requires FEI scripting interface software add-on)

Prepare the microscope

1. Chill the microscope cold finger with liquid nitrogen.

Starting and ending a microscope session should be well described by the operating protocol for the EM microscope at your facility. The microscope-specific steps are named here as a reminder, or to inspire the microscope operator to investigate their microscope documentation further if these steps are not already familiar to the operator.

2. Turn on high voltage to the filament and ramp up the current to the filament to bring it to operating temperature.
3. Wait 30 min to 1 hr for the microscope cold finger to chill.
4. Open the column valves and confirm the electron beam is visible on the phosphor screen.
5. Close the column valves in preparation to load an EM grid.
6. Load EM grid in the microscope sample holder and insert the sample into microscope.

Among these microscope-specific steps, inserting the grid holder into the microscope is one of the most critical. Be sure to coordinate with the individual responsible for microscope maintenance that this step is executed according to the established practice for your facility.

7. After the microscope vacuum recovers and stabilizes, open the column valves to reveal the electron beam.
8. Find the electron beam.

The beam is commonly blocked by a copper crossbar of the EM grid, and the grid must be moved using the microscope stage control in order to see the beam on the phosphor plate.

Parameterize the sample

9. Bring the sample to a Z-height previously known to be near eucentric for this microscope sample holder.
10. Find a visible feature on the carbon surface of the grid at an intermediate magnification (e.g., 9.7 kX).

Dust on the grid, a negative-stain aberration, or any visible feature at the current magnification can be used. Inserting the objective aperture in the path of the electron beam can increase contrast if needed.

11. Set the proper eucentric height for this sample (e.g., use the alpha-wobbler on a FEI Tecnai 12).

Setting the eucentric height can be easier if performed first at lower magnification, then increasing the magnification to approach the desired acquisition magnification.

12. Find focus on the sample, using the visible feature on the carbon surface.
13. Increase the magnification to that desired for acquisition [e.g., 59 kX at the phosphor screen, which becomes 67 kX (1.8 Å/pixel) at the camera for a bottom-mount Gatan OneView].
14. Adjust the beam to a brightness appropriate to the camera.

CAUTION: Exposing the camera to too much electron beam intensity can be detrimental to the camera.

Selecting appropriate beam intensity is important, and involves these considerations: (1) a beam that is overly converged will become unfocused and resolution will decrease, (2) beam intensity is related to incident beam size on the grid, which is important for parameterizing low-dose mode, and (3) a more intense beam often provides more contrast, reduces exposure time, and reduces drift affects, as long as the beam intensity is within the camera's operating range.

A beam size slightly smaller than the phosphor viewing screen is often a good choice, and then the microscope spot size can be used to tune brightness appropriate for the camera.

15. Find focus on the molecules of interest using the camera.

If available, interactive mode on the acquisition camera is ideal for this step.

Sometimes molecular complexes can be seen on the phosphor screen, but sometimes the added sensitivity of the camera is necessary to visualize the particles.

CTF fitting programs such as CTFFIND will be used to determine exact defocus but recognizing approximate defocus values directly at the microscope is important. An overfocus particle will have a halo (Fresnel fringe) of white signal, while an underfocus particle (the desired defocus offset) will have a halo of black signal. We show various defocus values for a micrograph series of the HA-441D6 complex stained with uranyl acetate (UA; Fig. 1A-H), and defocus effects have been nicely illustrated previously for bacteriophage Basilisk (Belnap, 2015).

16. Confirm distinct molecules of the samples are visualized.

Determine if the sample is too dense or too sparse, or if the stain is too dark (almost opaque) or too light (almost clear). Record an image to annotate this EM grid, even if the sample or grid is poor.

If the sample density or the stain thickness is undesirable, first try other locations across the grid. Stain may be thicker where the grid was gripped by tweezers, and thinner at the point of blotting. Sample density may vary across the grid as well.

If numerous places across the EM grid have undesirable stain depth or sample density, remake the EM grid addressing the problem.

EM imaging setup

17. Confirm “tune” alignments of EM microscope.

Check beam astigmatism (condenser stigmator), rotation center, and image astigmatism using the Fourier transform of the live camera image (objective stigmator).

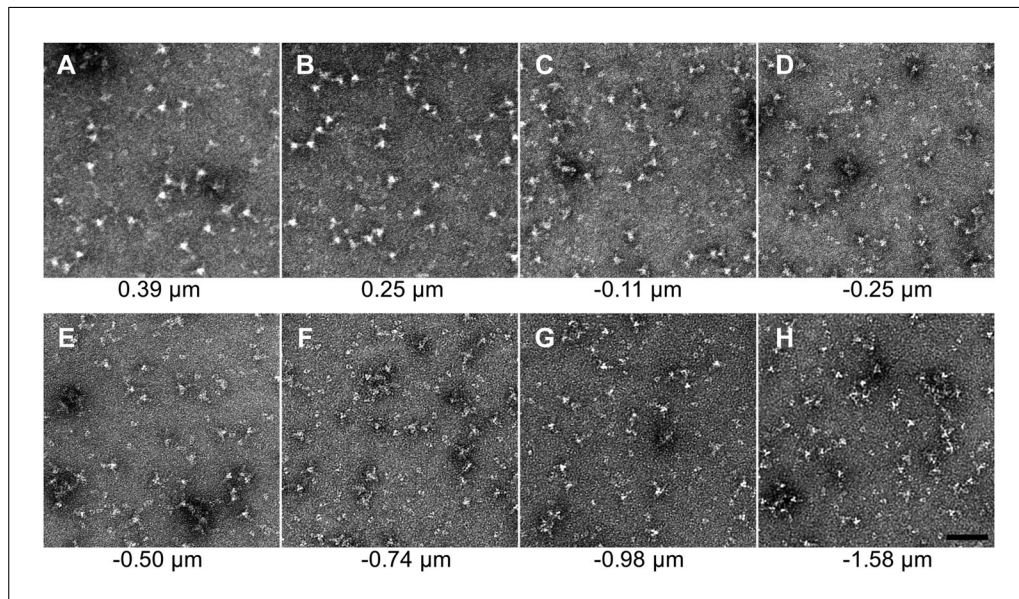


Figure 1 Micrograph defocus series. Defocus values for micrographs acquired at 100 kV were fit with CTFFIND4. (A-H) Representative micrographs near focus were selected, and the CTFFIND4 defocus value is stated below the micrograph. While the effect of defocus is a function of microscope voltage, 100 kV images acquired at focus up to about $-0.3 \mu\text{m}$ can have a granular appearance, since high-frequency noise is retained. Defocus values from -0.75 to $-1.5 \mu\text{m}$ often appear crisp because the particles are well defined by a black halo, but high-frequency noise is suppressed as the increased size of the Fresnel fringes interfere destructively. At defocus values greater than $-2 \mu\text{m}$, particles may readily appear blurry as the outline of the particle is suppressed by destructive interference of the Fresnel fringes. Desired defocus values are likely to be between -0.1 and $-0.4 \mu\text{m}$, leaving for particle averaging to differentiate signal from noise, while retaining enough defocus for proper CTF fitting of Thon rings. Scale bar is 50 nm.

Inaccurate alignments will blur micrographs and defeat the purpose of performing 2D class averaging to improve resolution.

18. Open the software application SerialEM, if not already open.

Some hardware configurations may not allow multiple acquisition software to be open simultaneously, therefore the manufacturer's camera software may need to be closed.

19. Acquire a "Record" image ("Camera and Script Controls" Panel → "Record") with SerialEM to confirm camera and SerialEM operability (Fig. 2A).
20. Enable SerialEM low-dose mode (Fig. 2A).

Electron dose can perturb negative stain and cause rounding of object edges in the micrographs, resulting in loss of resolution and creating artifacts in the sample. Low-dose mode "blanks" the beam (takes the beam off the sample) whenever an image is not being acquired, dramatically reducing the total dose to the region of interest on the EM grid.

It is recommended that low-dose mode be initially configured at a time prior to production data collection. Considerations important to low-dose mode parameterization are discussed in steps 14, 23, and 29. Settings for each SerialEM low-dose acquisition mode can be adjusted by toggling "Continuous update of mag & beam" in the "Low-Dose Control" panel, choosing the desired acquisition mode in the "Go to area/show when screen down" region of the "Low-Dose Control" panel, and then lowering the phosphor screen and making the desired changes.

21. Acquire a "Record" image with SerialEM in low-dose mode, performed as in step 19.

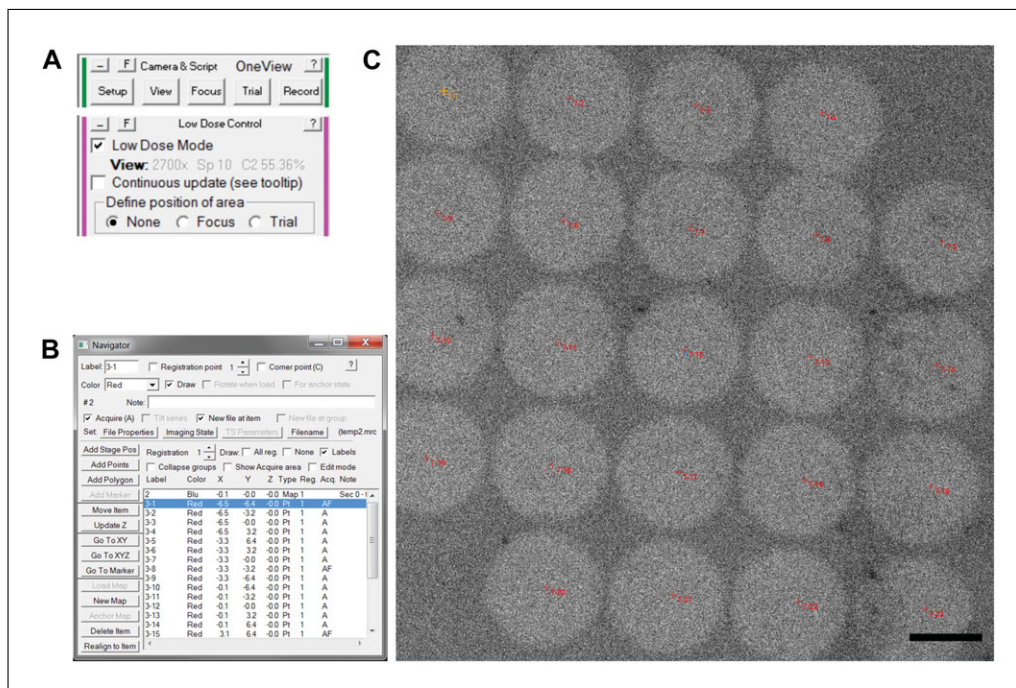


Figure 2 SerialEM data acquisition. **(A)** Four preset acquisition types are offered in SerialEM (top panel). By toggling on the low-dose mode (bottom panel), special meaning is imparted to “View,” which is used for montages, and “Record,” which is used for data acquisition. **(B)** The “Navigator” window manages points for data acquisition, here showing the first 15 points to be acquired and stored in two distinct MRC stacks, annotated by the “F” for filename. **(C)** The electron beam can cause radiolysis of the negatively stained sample, with some grid preparations more susceptible than others. After acquiring the set of “Record” images, the lower-magnification montage “View” was reacquired, and the experimental acquisition locations were plainly visible where the beam impinged on the sample. The red labels indicated coordinates where images were marked to be acquired, and they correspond well with the circular beam footprint. The effect from a 1 s exposure illustrates the importance of low-dose mode to manage dose on the sample. Scale bar is 2 μm .

Low-dose mode maintains its own image acquisition settings, which are recalled each time low-dose mode is enabled.

22. Confirm the beam is centered on the camera, and beam intensity is correct by observing the beam on the phosphor screen at the “Record” settings.

If the beam is not correct, update the low-dose settings for “Record” as indicated in step 20.

Beam intensity should lead to a beam size approximately covering the phosphor imaging plate. Beam brightness can be tuned with microscope spot size.

An over-condensed beam results in loss of high-resolution information, while an under-condensed beam causes radiolytic damage on an unnecessarily large grid area.

23. Acquire a “View” image (“Camera and Script Controls” Panel → “View”) with SerialEM to acquire a lower magnification micrograph (Fig. 2A).

The “View” acquisition will be automatically used for the montage acquisition in low-dose mode. We use a “View” magnification of 2.7 kX, resulting in a micrograph with 40.4 $\text{\AA}/\text{pixel}$. “View” settings are set by the toggle for “Go to area/show when screen down” to “View,” and adjust the beam parameters as in the annotation of step 20.

Confirm that the appropriate lower magnification is selected, that the beam remains centered about the camera, and that very low-dose is applied to the grid by spreading the beam and/or increasing the spot size.

Acquire the data

24. Initiate a new montage in SerialEM by selecting “File” menu → “New Montage . . .”.

A montage image is required for the navigator map and semi-automated data collection. For simplicity, we recommend using a montage of 1 image. Montaging multiple images can be powerful, but requires increasingly accurate alignments and microscope parameterization, and is likely to run in to the copper grid cross bars.

25. Acquire the montage: “Montage Controls” Panel → “Start” button.
26. Open the navigator window: “Navigator” menu → “Open”.
27. Load the montage into the navigator: select “New Map” in the navigator window (Fig. 2B).

Note that only montages, not regular images, can become “maps” in the navigator interface.

28. Add locations for acquisition of “Record” images: “Navigator” menu → “Montaging & Grids” → “Add Grid of Points.”
29. When prompted by the “Add Grid of Points” menu, enter the point spacing in micrometer, which is appropriate for your magnification and beam settings. This can be determined empirically, as shown in Fig. 2C.

For a bottom mounted OneView Camera (15 μm detector pixel size), the following two choices have worked well: (1) 3.2 $\mu\text{m}/\text{shot}$ spacing for 67 kX acquisition magnification (1.8 $\text{\AA}/\text{pixel}$) from a 2.7 kX montage, and (2) 2.2 $\mu\text{m}/\text{shot}$ spacing for 110 kX acquisition magnification (1.1 $\text{\AA}/\text{pixel}$) from a 2.7 kX montage. The authors commonly use 1.8 $\text{\AA}/\text{pixel}$ acquisition settings, which corresponds to a theoretical maximum resolution of 3.6 \AA .

30. When prompted by the “Add Grid of Points” menu, enable acquisition for all points.

Acquisition can be toggled on or off for each navigator point by selecting it in the navigator, then checking the box for “Acquire (A).”

31. Set acquisition file names. In the navigator window, select the first point to be acquired, then check the box: “New file at item.”

If not all points have been assigned a filename, then images will be added to the previous image stack from this acquisition, which we recommend. If no output filenames have been set, the acquired data may not be written to a file, so the first point should always be assigned a filename.

*Enable optional output of *.mdoc files to log image acquisition parameters.*

Ensure that all points of interest have “acquisition” toggled on, as indicated by the “A” in the “Acq” field of each point in the navigator window.

32. Start the acquisition: “Navigator” Menu → Acquire at points” → “Just acquire and save image.”

The “Record” image setting is used during automated acquisition. The beam settings should be high-dose for acquisition, within the constraints described in steps 14 and 23.

Interactively adjust focus while the stage moves between image acquisitions, targeting a very slight negative focus (e.g., between -0.3 and -0.6 μm). This drastically saves time by skipping autofocus. Autofocus, while appropriately rigorous, requires more accurate microscope alignments, and it must be properly parameterized within SerialEM. If the goal is to setup the microscope to acquire data independently, such as overnight, autofocus will be essential. If the microscope will be used interactively, such described here, autofocus may not be warranted. Autofocus may work within a certain range of offsets, but it may fail when those are exceeded, forcing the operator to manually intervene anyway.

Without autofocus, some micrographs will be unusable because they are too far from focus, but most will be usable, and this is sufficient. Having a diversity of defocus values is helpful for averaging of particles during image analysis, since a particular defocus results in loss of information at a specific spatial frequency.

Repeat data acquisition on a fresh area of the grid

33. Pan the stage to a fresh location on the grid by acquiring a view image, then right-clicking and dragging the image in SerialEM to pan the stage.

As an alternative to using SerialEM to pan the stage, the microscope controls can be used to directly pan the stage.

Acquire a new “View” image after panning to see the new area and repeat until a desirable area is found.

34. Collect more data repeating from step 25.

Collect a new Montage (Montage panel → start).

Add the montage as a new map (Navigator panel → New Map).

Add new acquisition points to the map using “Add Grid of Points” as described in step 28.

Set new filenames for files to be saved.

Initiate the acquisition, then observe and manually tune the defocus between acquisition of each micrograph.

35. Collect as much data as possible during a fixed time interval, such as 30 min or 1 hr.

Estimate how many particles per micrograph, then estimate how many micrographs are needed for a target particle count, such as 10,000 particles. For a particle less than 30 nm in diameter on a 4k detector, 100-200 pickable particles per micrograph would be anticipated when the particle density is favorable.

A set of 5,000-20,000 particles is likely enough data to draw conclusions from the sample of interest. Up to 50,000 particles may be used to increase resolution of the class averages. Beyond 50,000 particles for a single negative-stain data collection, incremental gains in 2D class resolution may become negligible, but this can be sample dependent.

Finish the EM session

36. Close the EM column valves, retract the camera, remove the grid holder and store the EM grid, then replace the grid holder in the microscope. If the microscope is done for the day, remove the cold finger liquid nitrogen Dewar and cryo-cycle the microscope.
37. Transfer images from the data acquisition computer to a location where data will be processed.

NEGATIVE-STAIN SAMPLE PREPARATION

Preparation of negatively stained sample is both an art and a science, and it has been described well elsewhere (Booth, Avila-Sakar, & Cheng, 2011; Burghardt & Droleskey, 2006). A carbon-coating machine can be used to apply a carbon coat to a thin layer of plastic on the copper grids or pre-made carbon coated grids can be purchased as done in this protocol. Many approaches exist to apply sample and negative stain to an EM grid (Kastner et al., 2008; Scarff et al., 2018), but here we highlight a single method that has proved effective the majority of the time for the authors. The quality of the negatively stained EM grid determines the difficulty of the data analysis. Often, increased effort and computational resources during data analysis cannot compensate for a poorly stained

EM grid. Thus, it is often desirable to screen several negatively stained grids before acquisition of a dataset, and when revisiting a sample, to collect data only on grids that offer an improvement to previous data collections.

Sample purity can be paramount to the success of a project. Some generalities about producing a pure sample and avoiding aggregation are described in Critical Parameters.

Materials

Distilled water (dH₂O)

0.5% EM-grade UA, diluted in dH₂O from a 2% stock (see recipe)

15-30 µg/ml sample, 5 µl per EM grid, “pure,” monodisperse

Formvar/carbon (F/C) 300 mesh grids (e.g., Ted Pella, cat. no. 01753-F)

Glass microscope slide

Fine-tipped self-closing tweezers (e.g., Electron Microscopy Sciences, cat. no. 0304-N1-PO)

EM grid glow discharger (e.g., PELCO easiGlow model 91000, Ted Pella)

Parafilm

Improvised cover for the EM grid and solutions: e.g., petri dish lid or plastic container for 96-well plate

Blot paper (e.g., Whatman grade 1 qualitative filter paper, standard grade, 90 mm circle)

EM grid storage box (e.g., Electron Microscopy Sciences, cat. no. 71150 or use empty boxes from purchased grids)

Glow discharge the EM grid

1. Select one formvar/carbon EM grid and place on the glass slide with the carbon side up (dark side of the grid up) using the fine tweezers.

Using the wrong side of the grid (formvar side) will result in terrible staining of the sample and uninterpretable results.

Grip the grid at the edge with the tweezers to avoid tearing the carbon coating.

The formvar coating provides stability for the carbon, while adding a small amount of noise to the images. While pure carbon grids can be made or purchased, the added stability of formvar or collodion support is often a desirable tradeoff.

2. Place the glass slide (containing the EM grid) in the PELCO glow discharger.

NOTE: Operate the PELCO glow discharger in accordance with the manufacturer's instructions.

Multiple grids can be glow discharged at once, but the sample should be applied to grids within 30 min of glow discharge.

3. Place the glass bell over the sample pedestal.

Take care to make a good seal on the rubber gasket to ensure a good vacuum.

4. Run the PELCO easiGlow glow discharge program.

Parameters that have worked well include a 10 s hold equilibration time for vacuum stabilization, followed by a 15-mA glow for 25 s. The chamber should glow a purple/pink hue when using atmospheric gas composition. A successful glow discharge renders the EM grid hydrophilic, as judged by the spreading of sample on the grid (Fig. 3A and B).

Allow the vacuum to fully dissipate before lifting the glass bell. Prematurely lifting the glass bell causes stress to the glass, leading to chipping of the bell.

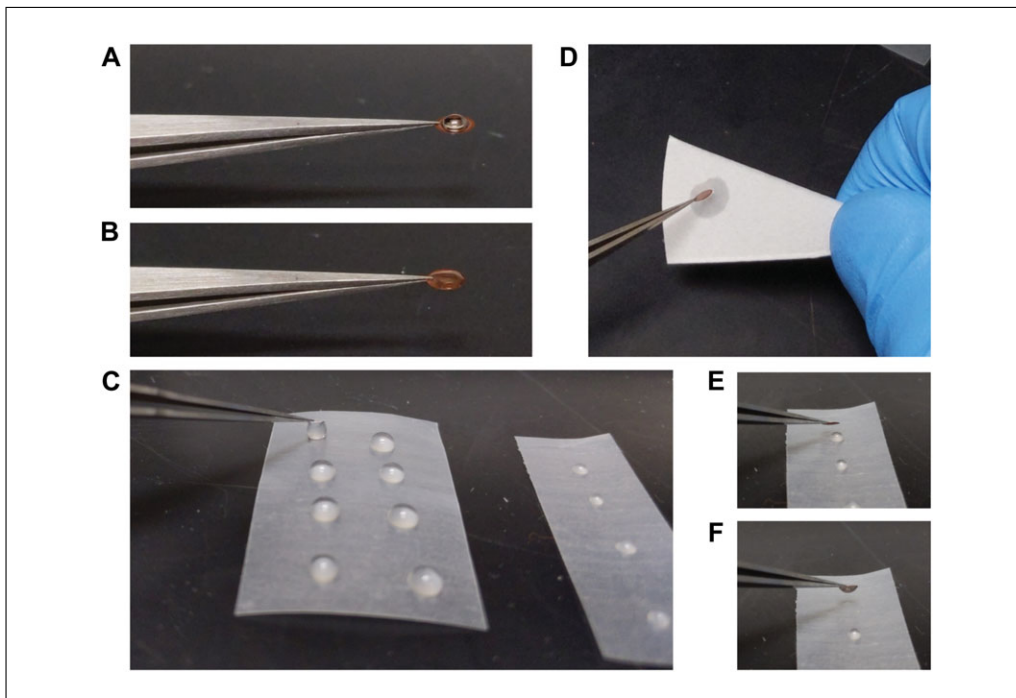


Figure 3 Grid preparation. **(A)** A thin carbon-film grid that has not been glow discharged is hydrophobic, and it will exhibit a high contact angle between the droplet of sample and the grid surface. **(B)** A successfully glow discharged carbon surface will be hydrophilic and yield a low contact angle of the sample droplet with the grid surface, causing the sample to fully wet the grid surface. Most molecular complexes adhere better to the hydrophilic carbon surface. **(C)** Rinsing the grid can be done by dipping the grid in 30- μ l droplets of H₂O arrayed on Parafilm rectangles. **(D)** Reproducible blotting of a wet grid can be done by touching the side of the grid to blotting paper. **(E, F)** Negative stain droplets containing 8 μ l are small enough that they will adhere to the grid by surface tension, allowing the tweezers to be set on the working surface during the remaining incubation time.

If a glow discharge machine is not available, the glow discharge step may be skipped, but sample deposition on the carbon surface will likely be less efficient and may require higher samples concentrations or a larger volume of sample.

5. Pick up the EM grid with the self-closing tweezers and place the tweezers with the grid on the laboratory bench.

Prepare EM grid rinsing and staining solutions

6. Cut Parafilm into a square or rectangle and place on a working surface, such as the laboratory bench (Fig. 3C).
7. For each EM grid: place two droplets of dH₂O, approximately 30 μ l each, and place one droplet of 0.5% UA negative stain, approximately 8 μ l each.

To minimize solid waste that has touched uranium, use a separate smaller Parafilm square for the uranium-based stain. Similarly, uranium-contaminated blotting paper can be minimized when staining multiple samples by using blot paper exclusively to blot the uranium stain droplets, touching each grid to a different region of the blot paper.

Cover the droplets to reduce evaporation. Be aware that evaporation from a single droplet can lead to significant changes in stain concentration. To cover the droplets, several simple options include the lid of a petri dish, the lid from a pipet tip box, or the lid of the sterile wrapping from a 24-well cell culture plate.

Consult with your institute's appropriate office concerning the use and proper disposal of materials related to uranyl stains.

8. Apply 5 μ l sample to the carbon side (dark side) of the grid (Fig. 3B), and then cover the grid.

The amount of solution needs to be enough to prevent drying during the incubation period. For precious samples, less solution can be used, but it is critical that the grid does not dry.

9. Allow sample to settle on the EM grid for 1 min.

Both shorter and longer times can work, resulting in slightly less/more sample deposited on the grid surface. Ensure the sample does not begin to dry on the grid.

Stain the EM grid

10. Pick up the tweezers with the EM grid in one hand and blot paper in the other hand.

Have all staining and rinsing droplets ready and immediately accessible and be prepared to move quickly between blotting and rinse steps to ensure that there is no intervening time for the grid to dry.

The blot paper can be cut into 1/8 round wedges, for ease of handling. One wedge can be used per grid prepared, to avoid cross-contamination between different samples.

11. Blot the sample from the grid by touching the grid to the blot paper briefly, approximately 0.5 s.

12. Immediately touch the sample side of the grid to the first dH₂O droplet to initiate the first rinse (Fig. 3C).

13. Blot the first rinse from the grid by touching it to a fresh location on the blot paper (Fig. 3D).

Rinsing removes sample buffer from the EM grid, for example, preventing formation of uranium phosphate salt crystals on the grid.

14. Immediately touch the sample side of the grid to the second dH₂O droplet for the second rinse.

15. Blot the second rinse from the grid by touching it to a fresh location on the blot paper.

16. Immediately touch the sample side of the grid to the UA droplet (Fig. 3E and F).

Due to the small size of the negative-stain droplet, it will completely adhere to the grid and be picked up from the Parafilm. This permits the tweezers and grid to be set down while the stain incubates with sample.

17. Wait 1 min: allow staining of the sample by UA.

18. Blot the excess negative stain from the grid by gently touching/sliding the side of the grid along fresh blotting paper for 0.5-1 s.

The amount of stain solution remaining on the grid during drying determines the darkness/thickness of the stain, which impacts how the particle is imaged.

Other methods such as blotting the face of the grid can be utilized, but may best be paired with other protocol adjustments, such as using 1% UA rather than 0.5% UA.

19. Initiate drying of the grid by gently waving the grid in the air, using the tweezers, for up to 5 s to promote initial drying.

This step is optional and does increase the risk of mishandling the grid.

20. Set down the tweezers with the EM grid on the working surface and wait 10 or more minutes for the EM grid to completely dry.

Using a still-wet EM grid can cause difficulty during loading of the EM grid into the microscope grid holder, because the grid will stick to surfaces on the grid holder, and to the tweezers. In addition, a wet grid can contribute to the accumulation of contamination inside the EM microscope column.

If multiple grids are made at the same time, they may be “labeled” by placing them on labeled filter paper in a petri dish, which could be divided up into four quadrants by simply drawing lines between the areas. Beware that EM grids can become statically charged and may move about in a surprising fashion, particularly in low-humidity environments. A reliable way to keep grids identified is by placing them directly in an EM grid storage box with numbered slots. The grid box slot number is then recorded so that it can be associated with the corresponding sample description.

A dry grid can be safely stored in an EM grid storage box for extended periods of time prior to imaging.

DATA ANALYSIS

While negative-stain EM of molecular complexes is a classic technique, substantial improvements in data analysis, recently driven by single-particle cryo-EM, can be leveraged to arrive at conclusions about molecular structure from large datasets that are far superior to conclusions drawn from single micrographs. By employing a contrast transfer function (CTF) correction to correct for varying micrograph defocuses, destructive interference during particle averaging can be minimized. By averaging across thousands of particles, less-common views and conformations can still be well defined. If multiple conformations are present, the relative proportions of these conformations can be estimated by the population of particles assigned to each of the different class averages. A foundational description of the interpretation of negative-stain EM micrographs and the impact of the CTF has been well described previously (Belnap, 2015). This protocol focuses on RELION3 with assistance from other software such as EMAN2, yet we note that EMAN2 is itself a full package that itself could be used to complete all the steps described here, and as noted in the background section of this article, many other packages exist as well. Many software packages are freely available and provide example data sets and online tutorials.

Much or all of the analysis can be done with the graphical user interface (GUI) of these software packages, but the command line interface at the Linux (or macOS) command terminal must be utilized to explore the full set of runtime parameters. In addition to the GUI buttons, we describe the text-based terminal commands, some of which are provided in the appendix, with the goal of familiarizing a user with commands they may later elaborate on. A general familiarity with the Linux command line (see Internet Resources) facilitates organization of data between analysis steps, as well as troubleshooting of data analysis programs.

One such Linux command line tool is “xargs.” While usages can be complicated, we use it in a single methodology, which is more straightforward. Given a command that creates a list of words, such as “ls”, which creates a listing of the files in the current directory, then this list of files can be “piped” into xargs with the Linux syntax of a vertical bar, “|”. The command xargs then runs a given command, provided as an xargs argument, on all the text piped into xargs. For example, “ls * | xargs -I{} cp {} {}.back” will first create a list of files with the list command, “ls”. The argument to “ls” is the wildcard, “*”, which matches all file names. The resulting filename list from “ls” is passed to “xargs” with the pipe command (vertical bar). Then “xargs” will use those filenames to substitute into the “{}” characters in the copy command, “cp”, and execute it. The result is each of those files being copied to a new file of the same name, but with “.back” appended to the destination filename. In this protocol, we use xargs to unstack all the MRC image stacks

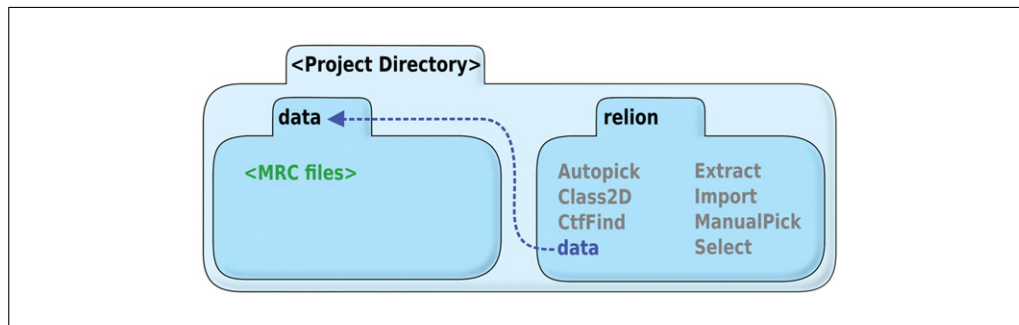


Figure 4 Directory structure for RELION data processing. Directories made by the user are named in black, directories created by the RELION GUI are shown in gray, symbolic links are shown in blue, and micrograph files are shown in green. This directory structure permits multiple RELION analysis directories to sit within a project folder, while keeping the raw data outside of the “reliion” processing folder. Critically important is that the path from the RELION processing directory to the micrograph’s directory avoids a parent directory reference (“..”), which would otherwise cause havoc for RELION’s auto-generated directory structures.

in a directory or make JPG thumbnails of all MRC files in a directory. It is essential to be capable of managing data as a batch rather than manually on a per-file basis when working with larger datasets, but batch operations open the possibility of errors going unnoticed. An important follow-up to each command issued in a batch format is to always spot-check a few results. This is to ensure the desired action indeed occurred.

Terminal commands are listed as bold annotations under their respective steps. Commands spanning more than one line will be broken across lines as could be done at the Linux command line using the “\” character.

Materials

- Digital micrographs of negatively stained complexes (see Basic Protocol 1)
- Linux system—(e.g., CentOS or Ubuntu): either remote or local access to a Linux operating system (macOS may also be used, but may require additional effort to compile software)
- Data storage of 50-200 GB
- RELION3 software (e.g., RELION 3.0)
- CTFFIND4 software (e.g., CTFFIND 4.1.10)
- EMAN2 software (e.g., EMAN 2.12, although newer versions are likely a great choice)
- FIJI software (e.g., based on ImageJ 1.52i, but continually updated)
- 3-button mouse, or capability to emulate a 3-button mouse

Review raw micrographs

1. Within a new (empty) project directory, make a “reliion” directory, and a “data” directory. Copy the set of micrographs into the “data” directory.

An example directory hierarchy for data analysis is shown in Figure 4.

2. Create a symbolic link from the “data” directory above the “reliion” directory to inside the “reliion” directory (Fig. 4).

```
cd reliion
ln -s ../data ./data
```

It is good to keep the micrographs outside the RELION directory so that multiple independent analyses be performed on a dataset. Having a reference to a parent directory (“..”) in RELION’s relative path to the data can cause havoc for RELION’s auto-generated directory structures. Therefore, the symbolic link is chosen to allow RELION to use the raw data as if it was inside the RELION directory hierarchy, even though it is not.

3. If the micrographs exist as an MRC stack containing multiple micrographs per file, convert them to a single micrograph per file using EMAN2 as follows:

```
ls *.mrc | xargs -I{} e2proc2d.py --threed2twod {} {}-split.mrc --unstacking
```

After the individual micrographs are created, move the MRC stacks to another directory for archiving.

4. Create a full set of JPG thumbnails using EMAN2, since it is commonly more practical to review a large set of JPGs than a large set of raw micrographs in MRC format as follows:

```
find . -maxdepth 1 -name "*.mrc" | xargs -I{} e2proc2d.py --process  
normalize \  
--scale 0.25 --clip 1024,1024 {} {}_THUMB.jpg  
mkdir jpg; mv *.jpg ./jpg/
```

Here, we assume the original images are 4k resolution, thus a scale factor of 0.25 results in a 1k image (i.e., 1024 pixels).

If the “-maxdepth” argument is omitted from “find,” then thumbnails will be made for all MRC files in all subfolders of the current directory.

5. Review the JPG files in their respective subdirectory, and discard or set aside corresponding raw micrographs that are excessively out of focus or exhibit negative staining artifacts.

Micrographs that are undeniably blurry due to excessive defocus, or due to lateral stage drift, should not be included in the analysis. These micrographs will also exhibit fewer Thon rings than expected in the Fourier transform of the image.

RELION micrograph processing

6. Use the Linux command terminal to navigate into the “relion” directory, and start the RELION GUI:

relion

For a given project, the RELION GUI needs to be started from the same directory each time. RELION terminal commands also must be run from the main RELION directory, since all references to image data are relative to this path.

Only one instance of the RELION GUI should be run at a time for a given project, otherwise the job tracking and “default_pipeline.star” can become corrupted.

By default, the RELION GUI will timeout and close after 1 hr of idle time. This timeout can be extended by starting RELION with the “-idle” argument, such as the following that extends the timeout to 1 day.

relion --idle 86400

7. Import data into RELION. Select “Import” in the GUI, and the input files can be specified as “data/*.mrc”. Choose “Node type: 2D micrographs (*.mrc)”. (See Appendix for corresponding RELION terminal commands.)

The product of a successful run is an enumeration of all input micrographs in the “micrographs.star” text file, within the “Import” subdirectory.

Note that every job submitted through the RELION GUI receives a job number that is incremented by one from the previous job in the project. Results for that job number are found inside the respective directory for that job class, such as “Import.”

8. Run CTFFIND4 to fit defocus values. Select “CTF estimation” through the RELION GUI. Set the spherical aberration, voltage, and micrograph pixel size. Select “Yes”

for “Use CTFFIND-4.1.” Default values are otherwise likely good enough to begin. (see Appendix)

The spherical aberration (CS) is characteristic to each microscope model and can be found in the microscope specifications.

An alternative software to CTFFIND4 that can also be run through the RELION GUI is gCTF, which is GPU accelerated (Zhang, 2016).

To facilitate rapid testing of CTF fitting parameters, it is helpful to work with a subset of micrographs at first. Create a copy of the “micrographs.star” file named “micrographs-test.star,” then use a text editor to delete all but four micrographs listed within the copy and save the changes. Use this “micrographs-test.star” to validate CTF fitting parameters.

```
cd Import/job001/  
cp -i micrographs.star micrographs-test.star  
gedit micrographs-test.star
```

“gedit” is a graphical text editor commonly available in Linux environments. For a text-based text editor, “nano” can be used instead, and is widely available on Linux environments.

9. Review the CTF fitting results by checking for correspondence of actual and predicted Thon rings (Fig. 5A). Either (1) start “Manual picking” through the GUI using the same “micrographs.star” file, then select the “CTF” button for each micrograph to be examined, or (2) convert the CTFFIND output files to JPGs and review them directly:

```
cd Ctffind/job002/data  
ls *.ctf | xargs -I{} e2proc2d.py {} {}.jpg
```

RELION particle processing

10. Manually pick 500 to 1000 particles (Fig. 5B). Select “Manual picking” in the GUI. Choose the input file as the CTFFIND result, micrographs_ctf.star. Input the particle diameter. Default values are otherwise a reasonable start. (see Appendix)

With a three-button mouse, the left click picks a particle, the middle click erases a picked particle location, and the right click offers a menu that allows the current picked-particle list to be saved, which is critical.

The actual number of picked particles needed depends on your data, and on how rigorous you elect to make the initial classification to be. Once suitable templates are created for auto-picking, the initial set of manually picked particles is no longer needed.

Center the particles as best as possible when picking them. If the particles are already centered, the 2D classification steps later on will not need to use compute power to figure out where center is.

11. Extract particles. Select “Particle extraction” in the GUI. Choose the “micrographs_ctf.star” from the CTFFIND result, and the particle coordinates picked from manual picking. Choose a box size (in pixels) that is 1.5-2× the diameter of the particle. Choose “No” for “Invert Contrast.” Otherwise, default values are reasonable to start. (see Appendix)

Larger particle boxes allow for better background estimation during classification, but the resulting particle stack requires more disk space to store, and, to a lesser degree, more RAM to compute on. The particle diameter for classification can later be fine-tuned with an outer mask, but the mask can only make the particle size smaller, not bigger.

“Invert Contrast” is set to “No” because RELION, and the majority of EM analysis programs, expect light particles on a dark background. Negative stain already presents

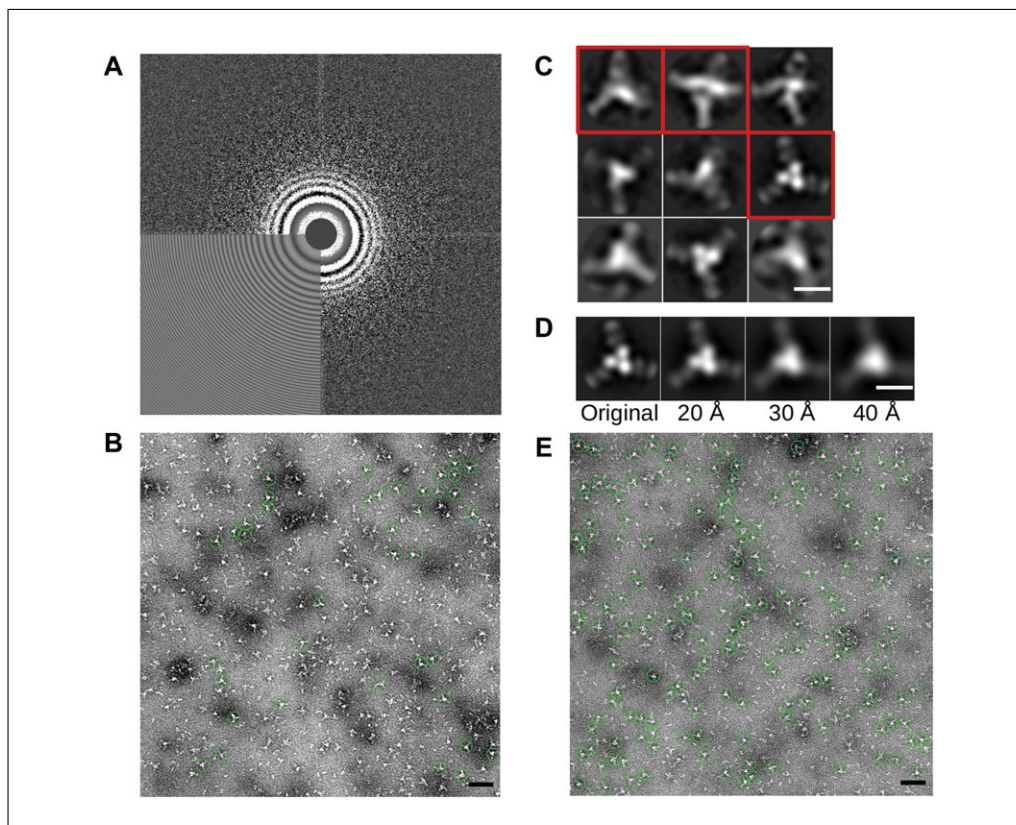


Figure 5 Data processing. **(A)** Thon rings in the Fourier transform of the micrograph describe the defocus of that micrograph. The lower-left of the image overlays the predicted Thon rings from CTF fitting, while the remaining quadrants are derived from the raw micrograph. A proper CTF fit results in continuous Thon rings across the experimental and predicted quadrants. **(B)** A subset of particles were manually picked, to be used for initial 2D classification. **(C)** From nine classes of manually picked particles, a subset of the classes (red boxes), representing top, side, and oblique views, could be selected as templates for auto-picking. **(D)** From the original image at left, a series of search templates were generated by low-pass filtering to 20, 30, and 40 Å resolution, representing the actual templates employed by auto-picking. **(E)** Auto-picking results were visualized on the original micrograph for inspection. Representative micrographs should be manually inspected to evaluate proportion of false-positive and false-negative results. Scale bars are 50 nm in panels B and E, and 10 nm for panels C and D.

light particles on a dark background. Cryo-EM datasets present dark particles on a light background, and therefore must be inverted to meet this convention.

- Run 2D classification to generate auto-picking templates (Fig. 5C). Select “2D Classification” in the GUI. Choose the “particles.star” produced from particle extraction. Set “Number of Classes” to between 4 and 10. Select a mask diameter in angstroms that is about half-way between the actual particle diameter, and the particle box size. Increase the number of threads to a reasonable number for the CPUs at your disposal. Default values are otherwise a reasonable start. (see Appendix)

Note that the effectiveness of 2D classification can be very sensitive to the mask diameter.

Runtimes are dramatically affected by the sampling parameters. 2D classification for up to a few thousand particles should proceed rapidly.

Expect at least one class to be a repository for particles that do not resolve into a homogeneous class. This class is often termed the “junk” class, and typically appears as a low-resolution blur. The junk class may contain many particles, some of which could make their own well-resolved class if there were more classes available, or if the search parameters were different.

2D classification is likely the first processing step in RELION where there is any concern of crashing the computer due to memory use. Threads share memory while applying multiple CPUs to a computation. Multiple threads must reside on the same computer node. MPI processes don't share memory, so running two MPI processes will require twice the memory to engage two CPUs in a single computation. MPI processes can be run on different computer nodes, and MPI is the mechanism by which a large RELION computation is run on a computer cluster with many nodes. The RELION FAQ contains much useful information on managing a large RELION computation (see Internet Resources).

13. If only a subset of classes from 2D classification will be used for auto-picking, such as when excluding a “junk” class, the subset needs to be generated explicitly. Select “Subset selection” and select a model.star file from the final iteration of the 2D classification. Use the RELION GUI to select the classes to keep (Fig. 5C). To save the particle selection, right click on the images and select “Save selected classes.” A “class_averages.star” file will be written that can be used as a reference for auto-picking.
14. Auto-pick particles using templates from 2D classification. Select “Auto-picking” from the GUI. Choose the micrographs_ctf.star from CTFFIND. Select the 2D references as the mrcs file from the conclusion of 2D classification, or the “class_averages.star” from subset selection. The parameters “Picking threshold,” “Minimum inter-particle distance,” and “Maximum stddev noise” will need to be manually optimized. Otherwise, default values are reasonable to start. (see Appendix)

The “Picking threshold” parameter relates to how well the template must match the picked particle. The “Minimum inter-particle distance” is used to avoid double-picking the same particle, as well as avoiding picking closely overlapping particles, which are inherently problematic for 2D classification. “Maximum stddev noise” sets a ceiling value for the standard deviation of the background for a picked particle. The noise in the background of a particle can be used for cryo-EM datasets to identify particles on carbon support which have a higher background noise than particles in carbon holes where background is purely frozen buffer. The utility of this parameter for negative-stain data is less clear, and it can prevent particles from being picked that otherwise would be desirable particles.

It is strongly recommended to make a subset of the total micrographs for testing of auto-picking parameters, before auto-picking the entire dataset. See step 8 in this protocol to make a “micrographs_ctf-test.star.”

When troubleshooting auto-picking, it is best to start from conditions where too many particles are picked, and then refine the picking back to the ideal amount. If no particles are picked, it is difficult to judge if the revised picking parameters are an improvement, or not.

Note that particles can readily be culled from a picked dataset, but auto-picked particles cannot be added to a picked dataset without the concern of double-picking particles. Thus, it is better to over-pick than under-pick particles. Still, if too many false-positive particles are picked, RELION may be unable to key in on features of the real molecular complex, resulting in ineffective 2D classification.

Runtimes for auto-picking with RELION are linearly related to the number of templates. It is advisable to select just a few templates, then rely on low-pass filtering of the templates and sufficiently generous picking thresholds to capture the full spectrum of views presented by a molecular complex. Low-pass filtering facilitates fuzzy matching to chosen templates, so only a few views of the complex are needed (Fig. 5D).

15. Confirm auto-picked particles are accurate by clicking on the auto-pick output via the RELION3 GUI. With the auto-picking job selected, select the output field indicated by “Display: out: coords_suffix_autopick.star.” The particle picking GUI

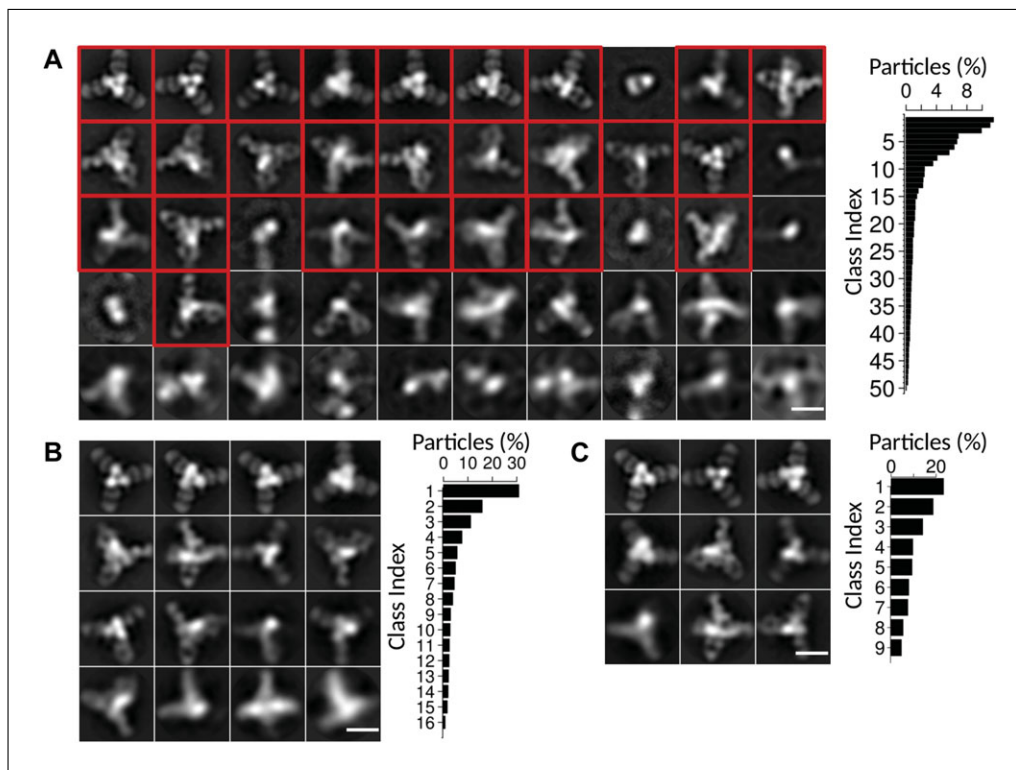


Figure 6 Arriving at final 2D classification results. For all panels, 2D classification results are at left, and class occupancy plots are shown at right. **(A)** Following preliminary particle curation, 2D classification was run with 50 classes to spread particles among unique classes. Classes in panel A were culled that appeared inconsistent with the HA-Fab complex, such as Fab alone in class #8. **(B)** In search of a representative minimal set of classes, culled particles from panel A were reclassified in 16 classes. Several classes have very similar views of the complex, suggesting 16 may be too many classes. **(C)** Culled particles from panel A were classified into nine classes. Class #9 appears to be a combination of top and side views, suggesting that nine classes may be too few. The ideal number of classes is sample dependent and is also at the author's discretion. Class indices increase left-to-right, then top-to-bottom. Scale bar is 10 nm.

will spawn, and the picked particles will be those resulting from auto-picking (Fig. 5E).

16. Extract auto-picked particles. Proceed as in step 11 in this protocol, except select the particle coordinates from auto-picking.

If many particles are picked and the particle dataset will be unwieldy to perform 2D classification, it is advisable to start with 2× binned particles. For the extraction parameters, select “Rescale particles: Yes.” Choose a size in pixels to be half of the original box size. Later, a smaller, refined particle set can be extracted again at 1× to achieve the best possible resolution.

17. Run 2D classification. Proceed as in step 12 in this protocol, except select the “particles.star” from the particle extraction from auto-picking. Set “Number of Classes” to be between 8 and 50 (Fig. 6A).

The number of classes needed is highly dependent on the data. It can be good to try several different numbers of classes, such as 10, 20, and 40. Twenty classes may be a good place to start.

2D classification of the full auto-picked particle set is the most compute intensive step of this protocol. If possible, utilize a computer cluster with MPI to run the computation on multiple computer nodes. GPUs can vastly accelerate this step, if available, although additional setup and command line arguments are required (see RELION in Internet Resources). The search parameters can be adjusted as well, such as coarsening the

sampling, or reducing the number of classes, or working with higher binning of particles (i.e., 2× binned, or 1.5× binned, instead of 1×). If your dataset is giant, divide the “particles star” file into sets of 40,000 particles and classify each set separately. After “junk” particles have been removed, the dataset subsets can be merged for a final 2D classification run.

18. Cull the particles in poorly resolved classes from the particle set. Similar to step 13, select “Subset selection” and select a model.star file from the final iteration of the 2D classification, select the classes to keep, and then save the particle selection. Differing from step 13, the output that will be used for future 2D classification is the “particles.star” file.

Some classes may be poorly resolved, not because the particles are outliers, but because the search parameters did not allow them to become a distinct class. Focus on culling the classes containing fewer particles. Often there are poorly resolved classes that have more particles than the well-resolved classes, and those classes should be retained because their constituents have not yet properly partitioned into a class.

Particle culling is accomplished over multiple rounds, but the number of rounds required can be sample dependent. Sometimes a cautious approach yields a better final selection of particles, while other times a cautious approach is just slower but not better.

Particle classification is imperfect, and each culling step removes some good particles with the bad. The RELION GUI can be used to display individual particles belonging to a class, and it is insightful to browse the diversity of particles within a class before choosing to cull or keep a class.

19. Multiple rounds of 2D classification and subset selection can be repeated until the particle set consists of well resolved classes, and the classes are representative of the particles in the sample as best as possible.

20. Perform a final 2D classification. The output becomes the final result for this sample (Fig. 6B and C).

Particles should be extracted at 1× binning for this final step.

The number of classes, angular sampling, translation search, and particle diameter all may be tuned to arrive at an ideal partition of particles into distinct, well-resolved classes.

Presentation of final 2D classification results can be made more clear if classes are sorted by an attribute such as percent of the particle contributing to the class (Gallagher, Torian, McCraw, & Harris, 2017), or the diameter of the particle in a varied population (Gallagher & Harris, 2018), or other pertinent descriptors. Two software that provide tools to construct these sorted image stacks are IMOD (Kremer, Mastronarde, & McIntosh, 1996), and FIJI (Schindelin et al., 2012). Plots of class occupancy were generated using the “R Project for Statistical Computing” software package using data from the RELION star files of the final iteration.

A full 3D reconstruction is possible from negative stain, but full 3D analysis is beyond the scope of this article. See Understanding Results for further discussion and an example of 3D analysis.

REAGENTS AND SOLUTIONS

Phosphotungstic acid (PTA)

Prepare 3% PTA stock can be made by dissolving PTA (e.g., Electron Microscopy Sciences, cat. no. 19500) in water

Titrate the pH to 7 using sodium hydroxide

Store up to 1 year at room temperature

Uranyl acetate (UA)

Prepare 2% UA stock by dissolving UA (e.g., Electron Microscopy Sciences, cat. no. 22400) in water

Filter the solution with a 0.22- μm syringe filter (or centrifuge and retain the supernatant)

Store in an aluminum foil wrapped container up to 1 year at 4°C (UA is light sensitive)

COMMENTARY

Background Information

Negative-stain electron microscopy (EM) has been widely used in identification and evaluation of virus particles (Gelderblom, Renz, & Ozel, 1991; Horne & Wildy, 1979). To improve resolution of biological complexes that were even smaller than viruses, computational averaging of images was first performed on negatively stained glutamine synthetase (Frank, Goldfarb, Eisenberg, & Baker, 1978). 2D classification of particles into distinct sets was first performed on negative-stain micrographs of hemocyanin (van Heel & Frank, 1981), and facilitated statistical analysis of particle conformations. Automation of EM data collection initiated with tomography, where multiple micrographs of a tilt series needed to be acquired precisely to enable reconstruction of the 3D volume (Dierksen, Typke, Hegerl, & Baumeister, 1993; Dierksen, Typke, Hegerl, Koster, & Baumeister, 1992; Mastronarde, 2005). Development for acquisition of single particle data as described in this article, but with full automation, required integrated effort managing multiple magnifications, stage control, focus, and machine vision algorithms (Suloway et al., 2005).

Image processing advances have been driven by the massive amounts of cryo-EM data generated by automated image acquisition by tools such as EPU (FEI, Eindhoven, The Netherlands), or Leginon (Suloway et al., 2005). A variety of software are commonly employed by each lab working in cryo-EM image analysis, but major packages include RELION (Kimanius, Forsberg, Scheres, & Lindahl, 2016), EMAN2 (Tang et al., 2007), BSOFT (Heymann, 2018), and FREALIGN (Lyumkis, Brilot, Theobald, & Grigorieff, 2013). New software suites are coming into use as well, such as cryoSPARC (Punjani, Rubinstein, Fleet, & Brubaker, 2017). These major packages are commonly supplemented by smaller software with very specific purposes, such as CTFIND4 (Rohou & Grigorieff, 2015) for CTF fitting, and GAU-TOMATCH (developed by Kai Zhang, MRC

Laboratory of Molecular Biology, Cambridge, UK) for GPU accelerated particle picking. A data-processing pipeline for cryo-EM may utilize substantial computational resources and require hundreds of computer cores. Both the data acquisition and data processing steps of cryo-EM analysis are resource intensive for personnel, and for infrastructure. Because of the increased signal-to-noise of negative-stain micrographs, the requirements for data collection and data analysis are much less stringent, all the while offering a substantial fraction of the information available from an atomic resolution cryo-EM reconstruction.

Critical Parameters

Negative-stain EM grid preparation

Pure sample greatly facilitates 2D image classification. Obtaining pure and monodisperse sample for negative-stain sample preparation can be difficult and may require developing a customized protocol. Nonetheless, the following observations may guide the process. Protein samples can be synthesized in many different expression systems, but generally all will require purification of the target protein away from other components (Wingfield, 2015). Contaminating proteins not only confound analysis by outnumbering the protein of interest, contaminants can imperil the target protein's integrity via the presence of trace proteases. At least two purification steps are typically required to attain high purity. Inclusion of a third purification step may lead to ideal purity, but also may be the point of diminishing returns on time manipulating the sample. Categorically, those steps are an initial affinity purification step, followed by an ion exchange purification step, and a final gel filtration step. This combination of selections by an expression tag, the neutral point of the protein, and the molecular size often leads to either pure protein or a clear understanding when aggregation prevents further purification. Nucleic acids and other samples also benefit directly from purification methods such as gel purification.

Pure samples may still suffer from aggregation. A monodisperse sample is important for image analysis since particle overlap confounds the shape recognition that is essential for particle alignment and averaging. EM is tolerant to detergent, and addition of detergents such as *n*-dodecyl- β -D-maltoside (DDM) and lauryldimethylamine oxide at concentrations below the critical micelle concentration may increase sample dispersion on the EM grid. These detergents may also be employed in cryo-EM, so knowledge gained from negative-stain analysis may be applied directly to ensuing atomic-resolution cryo-EM investigations.

When studying a complex of multiple components, such as our example of influenza hemagglutinin in complex with Fab 441D6, it is preferred that the complex can be purified away from its un-complexed constituents. This is not a strict requirement; however, since the unbound constituents may often be computationally identified and deleted from the analysis. Therefore, it may be most practical to first image the mixture of bound and unbound components of the complex, while also creating a plan to purify the bound complex. The purification plan for the complex may be invoked either when analysis of the mixture indicates it is necessary or as soon as sufficient reagents have been generated to pursue the purified complex in parallel.

Both the stain depth and molecular complex concentration are critical to producing useful images for generating 2D class averages. Stain depth plays a critical role in image contrast to define the features of the molecular complex. Concentration of the complex on the grid has a dramatic effect on the rate particles are collected, and a sparse sample could require an untenable number of micrographs to reach a practical particle count. Meanwhile, an overly dense sample will suffer from neighboring particles confusing the particle alignment during averaging.

Data analysis

Particle box size and radial mask can have large and sometimes surprising effects on generating 2D class averages. Ideally, the box of a picked particle in a micrograph should have no other particles in the same boxed out region. Using either the box size or the radial mask to disregard information at larger radial distances can avoid the particle alignment from being distracted by instances of neighboring particles, which could dominate the alignment. Less intuitively, increasing the box size or ra-

dial mask size can sometimes permit the 2D classification to discern more subtle details of the particle, particularly at the outer edges. In cases where a particle has a core which is especially easy to align, such as the round core of a virus capsid or an encapsulating complex such as ferritin, the round core can dominate the alignment at early stages while further structural details at the surface of the particle remain unresolved.

Runtimes can be dramatically affected by the angular sampling of the search space, the number of classes chosen for 2D classification, or the number of templates chosen for auto-picking. RELION output will provide runtime estimates. Monitor these runtimes, and strongly consider killing the job (see Troubleshooting) and restarting it with increasingly coarse search parameters if the runtime will be prohibitively long. Empirically determining the coarseness of sampling required to create class separation for particular dataset will likely recoup any time invested if multiple rounds of subclass selection are performed (see Basic Protocol 2, step 19).

Troubleshooting

Negative-stain contrast is low

When negative-stain depth is too thin, both particles and background appear light, and a subtle dark staining of the particle outline is seen. When negative-stain depth is too thick, the background is uniformly dark, while particles are very light, but the particle edges are poorly resolved. At even thicker stain depths, particles begin to darken and are hard to discern from the dark background, and the entire beam intensity will be reduced. Remaking a grid is the solution to incorrect stain depth, and it may require slight modifications to the standard protocol to consistently obtain stain at a desired stain depth (Booth et al., 2011; Scarff et al., 2018).

Some negative stains may provide better contrast for a particular sample, for example if the stain is bound by a complex with unusual affinity compared to average protein or nucleotide complexes. While UA reliably provides good contrast with most samples, phosphotungstic acid (PTA) is another stain that is widely successful at generating sample contrast. PTA is stable and is used at neutral pH, whereas UA requires acidic pH. Uranyl formate, not discussed here, can offer excellent contrast at neutral pH, but is unstable and must be made fresh each day. To use PTA stain, start from a 3% stock (see Reagents and Solutions),

prepare a 1.5% working solution diluted in water. 3% PTA is commonly used to stain samples, but the authors have had excellent results with 1.5% PTA, and thus recommend it. Staining with PTA can be performed by substituting the PTA solution for the UA solution in the Support Protocol, step 16, and by inserting an additional two steps of dH₂O rinse following the negative staining at step 18, performed as described by steps 12 to 15.

If no features attributed to the molecular complex of interest are visible on the EM grid, it is useful to have a control sample to confirm functionality of the negative-stain reagents and of the microscope. Protein complexes such as keyhole limpet hemocyanin (KLH) are an excellent positive control. KLH forms a 390 kDa complex that is a hollow cylindrical structure which provides good contrast with negative stain (Fig. 7A). KLH is commercially available (e.g., Sigma-Aldrich, cat. no. 9013-72-3) making it straightforward to obtain, though any large molecular complex less than 100 nm in diameter could be used as an alternative, under the constraints that they are readily available in pure form, and that negative stain of the complex produces good contrast.

If the sample concentration is wrong by being either too dilute, or too concentrated, then it is possible that no complexes may be discerned in the micrographs. Confirming the sample concentration is important when no complexes are observed by negative-stain EM. Also, confirm sample is being applied

to the carbon side of the grid rather than the plastic support side (formvar or collodion). Complexes may not adhere to the plastic, giving the appearance that the sample was excessively dilute.

Salt crystals and other semi-crystalline artifacts are found in the sample

Some buffers can react with the negative stain. An example of one such common pitfall of UA is that it forms precipitate with phosphate. Rinsing of the sample with dH₂O while adhered to the EM grid, but before negative staining, commonly avoids the interaction of stain and buffer. Air drying of the grid at any intermediate steps of staining can create strange artifacts. Remaking the EM grid is a good first step in understanding the source of inexplicable sample features. Some lipoparticles are known to bind to PTA and aggregate, presenting artifactual morphology (Zhang et al., 2011). Thus, another negative stain should be used when artifacts such as these are observed.

Acquired images are blurred

Check if stage drift is causing image blurring by acquiring consecutive images of the same location of the grid, as indicated by transposition of the particles in one direction between consecutive images. Drift in one direction also manifests in the Fourier transform of the image (typically available in the camera's image acquisition software, or in SerialEM's "Process" menu). Drift will cause the

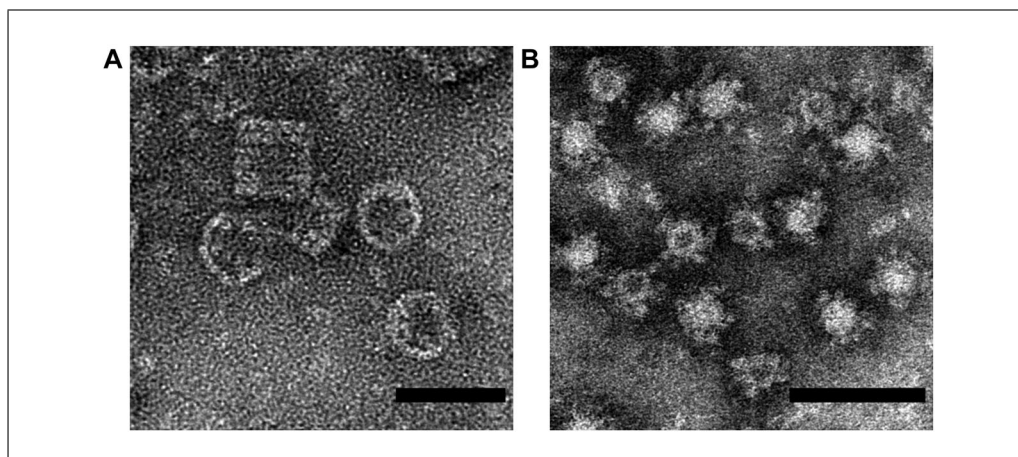


Figure 7 Sample preparation. **(A)** Having a reliable positive control sample can be helpful in diagnosing EM grid preparation and imaging problems. Shown here is keyhole limpet hemocyanin (KLH), which is a commercially available. KLH typically stains well due to its size and structure and produces top views and side view of the cylindrical particle. **(B)** Some negative stain issues can be left for data processing, such as stochastic penetration of negative stain into these ferritin nanoparticles, which happen to be derivatized with a mosaic of influenza hemagglutinin head domains. Data processing was used to separate particles with dark centers from those with light centers so that those features did not dominate particle alignment. Scale bars are 50 nm for both panels.

Thon rings to disappear along one axis of the Thon ring ellipse, indicating loss of resolution on that axis. If stage drift is manifest during automated image acquisition in SerialEM, increase the stage settling delay before image acquisition by setting the SerialEM property “StageMoveDelay” in the properties file (see SerialEM in Internet Resources).

Vibrations from the room or microscope can also cause blurring of the micrograph. Listening or feeling for vibration at the microscope can identify problems. For nanometer resolution, all discernable vibrations of the microscope are bad.

RELION programs are not running properly

RELION is available on some high-performance computer clusters. However, it can be more complicated to troubleshoot a RELION job submitted to a cluster than running RELION interactively, or locally. Consider building a separate RELION directory tree on a local computer solely for RELION testing, i.e., using just two micrographs. Use this testing RELION directory to generate skeletal input files, such as by picking only 10 particles. This directory can be used to confirm that RELION runs successfully in the skeletal setup, and that the command syntax is correct. Differences between the production run of particles and the testing RELION directory tree can be investigated until the expected behavior occurs on the production run. The most useful information in debugging a RELION run is often available in the STDOUT and STDERR terminal output. When run through the GUI, STDOUT and STDERR are saved as “run.out” and “run.err” files in the job directory. Direct inspection of the text form of the RELION input star files and the output star files can also help resolve problems such as a star file missing an entire field (column) of data.

Information on RELION program arguments can be found either at the terminal by running the program with no arguments (i.e., “relion_refine”) and observing the help output, or by clicking on the “?” in the GUI for the box that corresponds to that argument. Information from these two places can be complementary, so both should be checked for relevant information.

Currently running RELION jobs are not shown in the GUI

The RELION GUI can become disconnected from, or unaware of the number of actual RELION processes running. Use Linux system commands to interrogate which RE-

LION processes are running, and then kill any undesired RELION processes using its system process id (PID):

```
ps -u $USER  
kill -9 <PID>
```

RELION 2D classification is not generating distinct classes

There are some circumstances where RELION converges all of the particles into a few, or even just a single 2D class, while other classes are black and contain no particles. Instead of working from an unconverted iteration earlier in the 2D classification, use the following additional arguments to 2D classification that affect the probabilities of particles being considered as the same class.

```
--sigma_psi 0.001 --fix_sigma_noise  
--fix_sigma_offset
```

Some particle sets are not well separated during 2D classification, but they can be separated by using search templates. At the manual picking and auto-picking step, particle classes can be separately picked to create distinct “particles.star” files that can be independently 2D classified. Instances where negative stain penetrates some particles, but not others, can create large differences in particle appearance that may not be of biological relevance, as depicted by ferritin-based nanoparticles in Fig. 7B. Particles with dark centers (penetrated by negative stain) and those with light centers (not penetrated by negative stain) can be processed independently in these cases (Kanekiyo et al., 2019). The RELION “Particle Sorting” job type also offers the ability to sort out subsets of particles subsequent to particle auto-picking.

Understanding Results

2D class averages represent the average of the constituent particles. Perhaps counter-intuitively, instead of representing the most commonly observed structure within a class, the class average represents all structures that comprise it. Therefore, the final image for a class average may actually be a physically impossible conformation. A single class average can be composed of different conformations which may be vastly different, and they have been grouped by virtue of being more similar to each other than to the other particles included in the refinement. Therefore, multiple rounds of 2D classification are desirable. Each round culls false-positive particle picks, overlapping particles, or other artifacts in the region of the picked particle that do not reflect the particle’s structure, and can influence

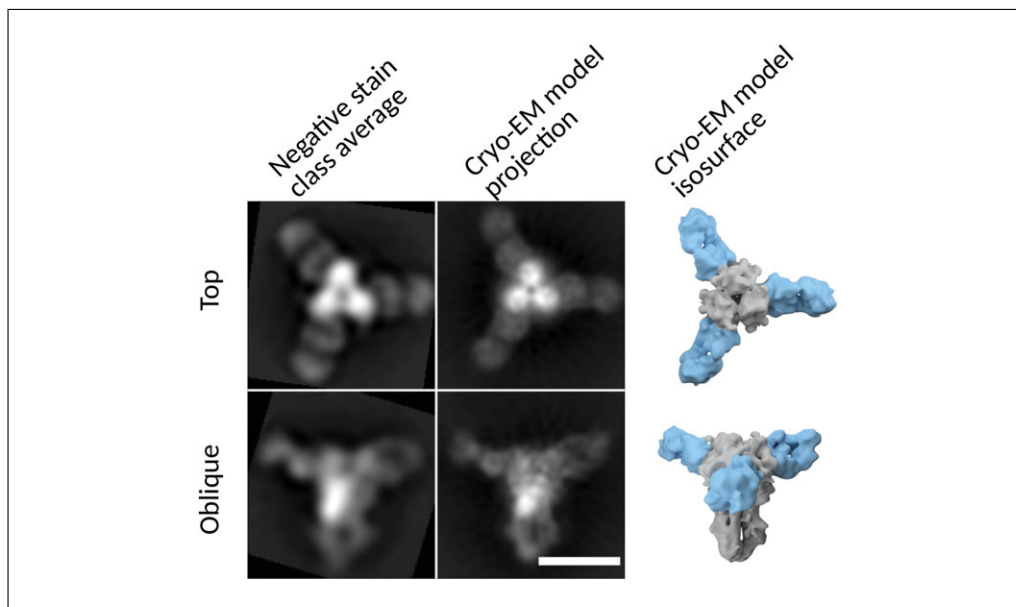


Figure 8 Negative stain results compared to cryo-EM. The left column highlights two well-resolved classes from negative stain 2D class averaging. The center column shows projections from the cryo-EM 8 Å 3D reconstruction of the same complex (EMDB-7021), showing good agreement with the negative stain results. The right column shows isosurface contours of the interpreted 3D cryo-EM map. Fab density is shown in light blue, and influenza hemagglutinin glycoprotein density is depicted in gray. Scale bar is 10 nm.

the alignment of its image. As the differences between particles across the dataset decreases (i.e., staining artifacts and partially resolved particles are removed), the refinement can increasingly focus on details within the complex of interest. A series of rounds of 2D classifications will show that a highly populated single class of particles can give rise to multiple distinct, well-resolved classes. After curation of picked particles derived from a well-prepared grid, final results from negative-stain 2D analysis can be in excellent agreement with a final map derived cryo-EM (Fig. 8).

A full 3D reconstruction is possible from negative stain, but there are important considerations before proceeding to 3D, which we merely introduce here. The stain depth may not cover the entire complex, causing the portion of the complex farthest from the carbon surface to be invisible. The carbon surface may impart preferred orientations to the complex, causing some views of the complex to be ill-defined, and resulting in poor resolution on the axes of these views of the complex. The particle may also be compressed towards the carbon surface during drying of the negative-stain sample, distorting a 3D reconstruction.

To illustrate that 3D reconstruction from negative stain can be successful, we used the culled particle set depicted in Figure 6A as in-

put to RELION “3D Classification” with one class, followed by RELION “3D auto-refine” to arrive at a plausible 3D model of the HA-441D6 complex (Fig. 9A and B). Atomic coordinates were available for the Fab and HA molecules, and these could be readily docked into the negative-stain 3D reconstruction of the complex to depict the Fab epitope. To enumerate how to begin to assess a 3D reconstruction, two simple criteria are helpful. (1) Direct inspection of the map in molecular visualization software such as Chimera (Pettersen et al., 2004) should reveal the map does not appear stretched or compressed along any particular axis or radially from the center, and that the density map is not discontinuous or disjoint when using an appropriate threshold level. (2) The Fourier shell correlation (FSC) plot resulting from a “gold-standard” refinement (Scheres, 2012; Scheres & Chen, 2012), which depicts the agreement of two independently refined sets of data according to resolution shell (i.e., Fourier space), should begin near 1.0 at low resolution, and then smoothly decline to 0 at some higher resolution (Fig. 9C). The exact resolution of the map is indicated by the FSC curve crossing the threshold 0.143 (Rosenthal & Henderson, 2003). Maps meeting these two criteria are on a good start to indicating a robust result.

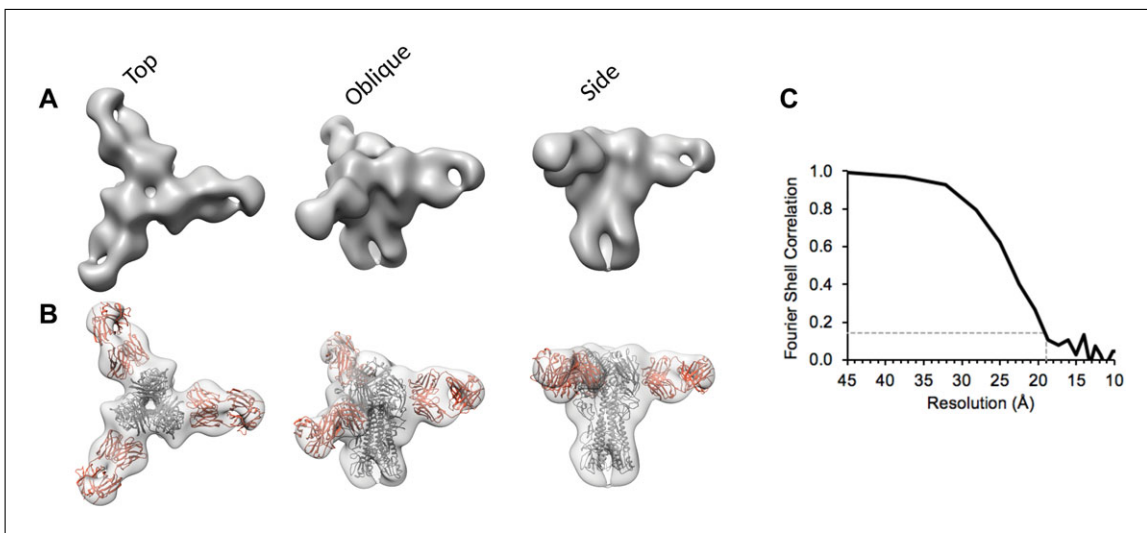


Figure 9 3D reconstruction from negative stain. **(A)** Views of the 3D map reconstructed from the final particle set of 2D classification. **(B)** Atomic coordinates for the Fab (PDB 5TR8), shown in orange, and a homology model of the HA (GenBank ABK79959) ectodomain, shown in gray, were fit into the negative stain reconstruction. The epitope is definitively localized to the HA head domain, but at a site distal to the known receptor binding pocket. **(C)** A Fourier shell correlation plot indicates that negative stain reconstruction is accurate to 19 Å resolution, as given by a 0.143 cutoff.

Time Considerations

Data collection

Preparation of negatively stained EM grids can take approximately 30 min, although multiple grids can be made concurrently, as long as one takes care not to mix up the unlabeled grids. Often, multiple iterations of grids will need to be prepared before an appropriate “good” grid is found, but once made, that grid can be stored and imaged as needed.

Operating the EM microscope may encounter unpredictable delays, but here we provide some optimistic times. Chilling the microscope and validating its functionality may take approximately 1 hr. Inserting a sample and setting up operating parameters can take 10 min, then browsing the grid before deciding if a particular EM grid deserves to have a dataset collected may take 10–30 min. Once collecting data, we have averaged just over six micrographs per minute, collecting about 180 micrographs in 30 min. With good particle density on the EM grid and a 4k resolution detector, 30 min of collection is likely to already be enough data (e.g., 180 micrographs with 200 particles per micrographs yields 36,000 particles). Even while it is tempting to simply embark on the additional 30 min of data collection after an hour has already been invested in the grid, the additional data may not be of any utility if the grid quality is not better than previously imaged. Concluding the EM session may take just 10 min, although copying the acquired data to a data analysis system

may take significant time. Setting up acquisition parameters for the first dataset in this semi-automated fashion with SerialEM may be a slow process, taking between a day and a week. Once a protocol is established that is optimized for the available equipment, it is not unreasonable to collect good data on four different samples in a single afternoon.

Data analysis

Manual particle picking may be as short as 20 to 60 min if one is confident the chosen particles are representative, and then auto-picking is relied upon to bootstrap picking of a full dataset. CTFFIND runtimes may be fast or slow depending on the chosen parameters and available computational power but may take several hours. Auto-picking via template matching in RELION3 can have runtimes of several hours. Thus, it is advisable to develop and test parameters on just one, or only a few micrographs, and ensure the results are as desired before queuing the processing of the entire dataset. Alternative auto-picking programs such as Gautomatch (GPU based) are available, and these can process the whole dataset in minutes to an hour once the correct picking parameters are discovered. It is worthwhile to spend extra time tuning the auto-picking parameters to avoid problems induced by overabundant false-positive particles during 2D classification. RELION3 2D class averaging will run quickly on the manually picked dataset, but it can be slow for a full dataset of auto-picked particles. For production 2D

classification runs on a full particle set, the availability of a computer cluster will greatly improve refinement runtimes. RELION3 scales well and can commonly be run on several hundred CPU cores via MPI. 2D classification can still be performed on a desktop workstation computer, but the search parameters will require optimization, and the dataset may need to be processed in subsets. If extending the analysis to 3D, runtimes on a desktop workstation may be excessively long, thus access to a compute cluster becomes more desirable. The full gamut of negative-stain data analysis, up to definitive 2D classification, may be completed in a week. The manual effort of setting up processing parameters, organizing files and directories, and curating the data is not negligible, which becomes more apparent when compute cluster resources are available and compute job runtimes are shorter.

Acknowledgment

This work was supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, National Institutes of Health, and the Biomedical Advanced Research and Development Authority.

Literature Cited

- Belnap, D. M. (2015). Electron microscopy and image processing: Essential tools for structural analysis of macromolecules. *Current Protocols in Protein Science*, 82, 17.2.1–17.2.61. doi: 10.1002/0471140864.ps1702s82.
- Bianchi, M., Turner, H. L., Nogal, B., Cottrell, C. A., Oyen, D., Pauthner, M., ... Hangartner, L. (2018). Electron-microscopy-based epitope mapping defines specificities of polyclonal antibodies elicited during HIV-1 BG505 envelope trimer immunization. *Immunity*, 49(2), 288–300.e8. doi: 10.1016/j.immuni.2018.07.009.
- Booth, D. S., Avila-Sakar, A., & Cheng, Y. (2011). Visualizing proteins and macromolecular complexes by negative stain EM: From grid preparation to image acquisition. *Journal of Visualized Experiments: JoVE*, (58), pii: 3227. doi: 10.3791/3227.
- Burghardt, R. C., & Droleskey, R. (2006). Transmission electron microscopy. *Current Protocols in Microbiology*, 3, 2B.1.1–2B.1.39. doi: 10.1002/9780471729259.mc02b01s03.
- Dierksen, K., Typke, D., Hegerl, R., & Baumeister, W. (1993). Towards automatic electron tomography II. Implementation of autofocus and low-dose procedures. *Ultramicroscopy*, 49(1), 109–120. doi: 10.1016/0304-3991(93)90217-L.
- Dierksen, K., Typke, D., Hegerl, R., Koster, A. J., & Baumeister, W. (1992). Towards automatic electron tomography. *Ultramicroscopy*, 40(1), 71–87. doi: 10.1016/0304-3991(92)90235-C.
- Frank, J., Goldfarb, W., Eisenberg, D., & Baker, T. S. (1978). Reconstruction of glutamine synthetase using computer averaging. *Ultramicroscopy*, 3(3), 283–290. doi: 10.1016/S0304-3991(78)80038-2.
- Gallagher, J. R., & Harris, A. K. (2018). Observed structural heterogeneity of human hepatitis B virus surface antigen particles by cryo-electron microscopy. *Microscopy and Microanalysis*, 24(Suppl 1), 1246–1247. doi: 10.1017/S1431927618006712.
- Gallagher, J. R., Torian, U., McCraw, D. M., & Harris, A. K. (2017). Structural studies of influenza virus RNPs by electron microscopy indicate molecular conformations within NP supra-structures. *Journal of Structural Biology*, 197(3), 294–307. doi: 10.1016/j.jsb.2016.12.007.
- Gelderblom, H. R., Renz, H., & Ozel, M. (1991). Negative staining in diagnostic virology. *Micron and Microscopica Acta*, 22(4), 435–447. doi: 10.1016/0739-6260(91)90061-4.
- Heymann, J. B. (2018). Guidelines for using Bsoft for high resolution reconstruction and validation of biomolecular structures from electron micrographs. *Protein Science*, 27(1), 159–171. doi: 10.1002/pro.3293.
- Horne, R. W., & Wildy, P. (1979). An historical account of the development and applications of the negative staining technique to the electron microscopy of viruses. *Journal of Microscopy*, 117(1), 103–122. doi: 10.1111/j.1365-2818.1979.tb00234.x.
- Kanekiyo, M., Joyce, M. G., Gillespie, R. A., Gallagher, J. R., Andrews, S. F., Yassine, H. M., ... Graham, B. S. (2019). Mosaic nanoparticle display of diverse influenza virus hemagglutinins elicits broad B cell responses. *Nature Immunology*, 20(3), 362–372. doi: 10.1038/s41590-018-0305-x.
- Kastner, B., Fischer, N., Golas, M. M., Sander, B., Dube, P., Boehringer, D., ... Stark, H. (2008). GraFix: Sample preparation for single-particle electron cryomicroscopy. *Nature Methods*, 5(1), 53–55. doi: 10.1038/nmeth1139.
- Kimanius, D., Forsberg, B. O., Scheres, S. H., & Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *Elife*, 5, e18722. doi: 10.7554/eLife.18722.
- Kremer, J. R., Mastronarde, D. N., & McIntosh, J. R. (1996). Computer visualization of three-dimensional image data using IMOD. *Journal of Structural Biology*, 116(1), 71–76. doi: 10.1006/jsbi.1996.0013.
- Kuhlbrandt, W. (2014). Biochemistry. The resolution revolution. *Science*, 343(6178), 1443–1444. doi: 10.1126/science.1251652.
- Lyumkis, D., Brilot, A. F., Theobald, D. L., & Grigorieff, N. (2013). Likelihood-based classification of cryo-EM images using FREALIGN. *Journal of Structural Biology*, 183(3), 377–388. doi: 10.1016/j.jsb.2013.07.005.
- Mastronarde, D. N. (2005). Automated electron microscope tomography using robust

- prediction of specimen movements. *Journal of Structural Biology*, 152(1), 36–51. doi: 10.1016/j.jsb.2005.07.007.
- Ohi, M., Li, Y., Cheng, Y., & Walz, T. (2004). Negative staining and image classification—Powerful tools in modern electron microscopy. *Biological Procedures Online*, 6, 23–34. doi: 10.1251/bpo70.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612. doi: 10.1002/jcc.20084.
- Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, 14(3), 290–296. doi: 10.1038/nmeth.4169.
- Rohou, A., & Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, 192(2), 216–221. doi: 10.1016/j.jsb.2015.08.008.
- Rosenthal, P. B., & Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology*, 333(4), 721–745. doi: 10.1016/j.jmb.2003.07.013.
- Scarff, C. A., Fuller, M. J. G., Thompson, R. F., & Iadaza, M. G. (2018). Variations on negative stain electron microscopy methods: Tools for tackling challenging systems. *Journal of Visualized Experiments: JoVE*(132). doi: 10.3791/57199.
- Scheres, S. H. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, 180(3), 519–530. doi: 10.1016/j.jsb.2012.09.006.
- Scheres, S. H., & Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination. *Nature Methods*, 9(9), 853–854. doi: 10.1038/nmeth.2115.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., . . . Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, 9(7), 676–682. doi: 10.1038/nmeth.2019.
- Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., . . . Carragher, B. (2005). Automated molecular microscopy: The new Leginon system. *Journal of Structural Biology*, 151(1), 41–60. doi: 10.1016/j.jsb.2005.03.010.
- Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., & Ludtke, S. J. (2007). EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology*, 157(1), 38–46. doi: 10.1016/j.jsb.2006.05.009.
- van Heel, M., & Frank, J. (1981). Use of multivariate statistics in analysing the images of biological macromolecules. *Ultramicroscopy*, 6(2), 187–194. doi: 10.1016/0304-3991(81)90059-0.
- Wingfield, P. T. (2015). Overview of the purification of recombinant proteins. *Current Protocols in Protein Science*, 80, 6 1 1–35. doi: 10.1002/0471140864.ps0601s80.
- Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *Journal of Structural Biology*, 193(1), 1–12. doi: 10.1016/j.jsb.2015.11.003.
- Zhang, L., Song, J., Cavigliolo, G., Ishida, B. Y., Zhang, S., Kane, J. P., . . . Ren, G. (2011). Morphology and structure of lipoproteins revealed by an optimized negative-staining protocol of electron microscopy. *Journal of Lipid Research*, 52(1), 175–184. doi: 10.1194/jlr.D010959.

Internet Resources

- <https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=CCPEM>
CCPEM listserv is for discussion related to the RELION software package.
- <https://www2.mrc-lmb.cam.ac.uk/relion/>
RELION Web site includes descriptions of software installation and best practices, as well as discussion of errors and software bugs.
- <https://www3.mrc-lmb.cam.ac.uk/relion/index.php/FAQs>
RELION FAQ is an important resource for strategies on CPU and memory usage by RELION.
- http://linuxcommand.org/lc3_learning_the_shell.php
General Linux command line tutorials, which can be used to solve many problems that arise in batch processing large sets of data files.
- <https://vim.fandom.com/wiki/Tutorial>
The VI command-line text editor is powerful, especially for the “search,” and “search and replace” functionality.
- <https://shapedshed.com/unix-xargs/>
“xargs” often requires working examples to learn from, such as offered here.
- <https://bio3d.colorado.edu/SerialEM/>
SerialEM Web site includes usage and installation guides.

APPENDIX: EXAMPLE RELION COMMANDS

Commands spanning more than one line are broken across lines as could be done at the Linux command line using the “\” character.

Import:

```
relion_star_loopheader rlnMicrographName > Import/job002/micrographs.star
ls -rt data/2019-01-18_NY96-441D6_montage_01-shot_*.mrc \
>> Import/job002/micrographs.star
```

CTF estimation:

```
`which relion_run_ctffind_mpi` --i Import/job002/micrographs.star \  
  --o CtfFind/job005\  
  --CS 2.2 --HT 100 --AmpCnst .1 --XMAG 10000 \  
  --DStep 1.757 --Box 512 --ResMin 40 --ResMax 2 \  
  --dFMin -30000 --dFMax 40000 --FStep 100 --dAst 500 \  
  --ctffind_exe ../shared/ctffind --ctfWin -1 --is_ctffind4 \  
  --fast_search --only_do_unfinished
```

The defocus step size (FStep) can be increased to reduce runtimes.

Manual picking:

```
`which relion_manualpick` --i CtfFind/job005/micrographs_ctf.star \  
  --odir ManualPick/job006/--pickname manualpick \  
  --allow_save --fast_save \  
  --selection ManualPick/job006/micrographs_selected.star --scale .3 \  
  --sigma_contrast 3 --black 0 --white 0 --lowpass 20 --ctf_scale 1 \  
  --particle_diameter 250  
echo CtfFind/job005/micrographs_ctf.star > \  
  ManualPick/job006/coords_suffix_manualpick.star
```

Particle extraction:

```
`which relion_preprocess_mpi` --i CtfFind/job005/micrographs_ctf.star \  
  --coord_dir ManualPick/job006/--coord_suffix_manualpick.star \  
  --part_star Extract/job007/particles.star \  
  --part_dir Extract/job007\  
  --extract --extract_size 128 --scale 128 --norm --bg_radius 48 \  
  --white_dust -1 --black_dust -1 --set_angpix 1.757
```

2D classification:

```
`which relion_refine_mpi` --o Class2D/job008/run \  
  --i Extract/job007/particles.star \  
  --dont_combine_weights_via_disc --pool 3 --pad 2 \  
  --iter 25 --tau2_fudge 1.2 --particle_diameter 200 \  
  --K 6 --flatten_solvent --zero_mask --oversampling 1 \  
  --psi_step 10 --offset_range 10 --offset_step 2 \  
  --norm --scale --j 4 --maskedged 10 \  
  --bimodal_psi --strict_highres_exp 8
```

Auto-picking:

```
`which relion_autopick_mpi` --i CtfFind/job005/micrographs_ctf.star \  
  --odir AutoPick/job009/--pickname autopick \  
  --ref Select/job061/class_averages.star --ctf --ctf_intact_first_peak \  
  --ang 10 --shrink 1 --lowpass 20 \  
  --angpix 1.757 --angpix_ref 1.757 --threshold .9 \  
  --min_distance 150 --max_stddev_noise 1.5 \  
  --skip_optimise_scale --skip_side 20  
echo CtfFind/job005/micrographs_ctf.star > \  
  AutoPick/job009/coords_suffix_autopick.star
```

Subset selection:

```
`which relion_display` --gui --i Class2D/job010/run_it025_model.star \  
  --allow_save \  
  --fn_parts Select/job011/particles.star \  
  --fn_imgs Select/job011/class_averages.star \  
  --recenter
```