Article:
Thompson, RF, Iadanza, MG, Hesketh, EL et al. (2 more authors) (2018) Collection, pre-processing and on-the-fly analysis of data for high-resolution, single-particle cryo-electron microscopy. Nature Protocols, 14. pp. 100-118. ISSN 1754-2189

https://doi.org/10.1038/s41596-018-0084-8

Collection, pre-processing, and on-the-fly analysis of data for high-resolution, single-particle cryo-electron microscopy

Rebecca F. Thompson*, Matthew G. Iadanza, Emma L. Hesketh, Shaun Rawson[1], & Neil A. Ranson*

The Astbury Biostructure Laboratory, School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, UK.

[1] Present address Harvard Medical School, Boston MA 02115
* For correspondence: r.f.thompson@leeds.ac.uk; n.a.ranson@leeds.ac.uk

KEYWORDS Cryo-electron microscopy; cryo-EM; single particle analysis; Titan Krios; Falcon 3; F3EC; K2; EPU

EDITORIAL SUMMARY This Protocol describes a pipeline for data collection, pre-processing, and on-the-fly analysis for single-particle cryo-electron microscopy using EPU software and two direct electron detectors; Thermo Fisher Scientific Falcon 3 (F3EC) and Gatan K2

TWEET A pipeline for single-particle cryo-EM data collection from @Astbury_BSL, pre-processing and on-the-fly analysis for #singleparticle #CryoEM

COVER TEASER Pipeline for cryo-EM data collection and analysis

Up to three primary research articles where the protocol has been used/developed:

1. Hesketh, E.L, Saunders, K., Fisher, C., Potze, J., Stanley, J., Lomonossoff, G.P. & Ranson, N.A. (2018). How to build a geminate virus capsid. Nature Communications 2369
2. Baggen, J.*, Hurdiss, D.L.*, Zocher, G., Mistry, N., Roberts, R.W., Slager, J.J., van Vliet, A.L.W., Casasnovas, J.M., Arnberg, N., Stehle, T., Ranson, N.A., Thibaut, H.J., & van Kuppeveld, F.J.M, (2018). The role of evolutionary enhancement of receptor binding in the emergence of pandemic viral conjunctivitis. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 392-402. DOI:10.1073/pnas.1713284115
3. Agip A., Blaza J.N., Bridges H.R., Viscomi C., Rawson S.D., Muench S.P., Hirst J. (2018) Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states. *Nature Structural Molecular Biology, 25,* 548–556

Abstract

The dramatic growth in using cryo-electron microscopy to generate high resolution structures of macromolecular complexes has changed the landscape of structural biology. The majority of structures deposited in the electron microscopy data bank at higher than 4 Å resolution were collected on Titan Krios microscopes. While the pipeline for single particle data collections is becoming routine, there is much variation in how sessions are set up. Furthermore, when collection is underway there are a range of approaches to efficiently move and pre-process this data. Here we present a standard operating procedure for single particle data collection with Thermo Fisher Scientific EPU software, using the two most common direct electron detectors (Thermo Fisher Scientific Falcon 3 (F3EC) and Gatan K2), and a strategy for structuring this data to enable efficient pre-processing and on-the-fly monitoring of data collection. This protocol takes 3-6 hours to set up a typical, automated data collection session.

Introduction

The use of cryo-electron microscopy (cryo-EM) to determine near-atomic resolution structures of macromolecular complexes has grown dramatically in the last 5 years, led by improvements in microscope hardware, detector technology and image processing methods [1-4]. Two companies manufacture high end cryo transmission electron microscopes (cryo-TEMs) aimed at the life science market (Thermo Fisher Scientific (Formerly FEI) and JEOL) and several manufacture direct electron detectors (Thermo Fisher Scientific, Gatan & Direct Electron) [5-7], suitable for visualisation of frozen hydrated specimens in this

resolution range. A variety of hardware combinations can thus be used to obtain high resolution data. However, the vast majority of structures deposited in the electron microscopy data bank (EMDB) at higher than 4 Å resolution have been achieved using either the Gatan K2 or Thermo Fisher Scientific F3EC direct electron detectors on Titan Krios microscopes[8].

Increasing numbers of institutions are investing in cryo-EM technology, but the operation of these instruments, and the management of the enormous data flows they create, present a series of challenges. Many Titan Krios sites are operational 24h/day, normally ~ 80 % or more of the time. With current detector technology this can mean producing ~ 5 Tb of data per 24 h per microscope, and this figure may increase dramatically as new detector technologies come online. These datasets present enormous practical challenges including how and when to move them around, and how they should be organised, but such considerations are essential for improving the efficient use of machine time, as they facilitate on-the-fly pre-processing of data (including motion correction and contrast transfer function (CTF) estimation) in parallel with data collection. Through such approaches, the quality of data can be assessed in near real time, allowing data collections unlikely to yield meaningful results to be halted early or have their parameters altered. Pre-processing also allows essential steps in cryo-EM structure determination to be executed in parallel to data collection, reducing the time and therefore cost from data collection to structure.

Based on the experiences of the Astbury Biostructure Laboratory (ABSL), here we present our complete protocol to go from optimised, frozen, cryo-EM grid to pre-processed single particle dataset. In our view this is generally the least time consuming and best characterised section of the cryo-EM workflow, as specimen preparation is still a major bottleneck for many cryo-EM projects [9,10], and full processing of the data to get the most out of the dataset can be months of dedicated, expert work. The protocol describes the standardised operating procedures developed and used at the ABSL by all users: internal and external; academic and industrial. The procedures are robust and tested, and have been used to successfully collect data on very wide range of macromolecular complexes, including small protein complexes (<150kDa), membrane proteins (in both detergent and nanodiscs), viruses, and amyloid fibrils. Using these procedures, we (and others) have determined multiple structures in the 2.5-5 Å resolution range (Table 1), and published their structures in peer reviewed publications [11-14]. The protocol is aimed at a broad audience, from non-expert users looking to set up automated data collection using EPU software, experienced users as an aide-mémoire, to new Facility Managers looking for information on implementations of on the fly image processing.

Overview of the Procedure

The procedure consists of twelve main stages (Figure 1). First, grids are transferred to the microscope (Steps 1-11). Initial checks are then performed to ensure the correct software is loaded and that the microscope is in the correct mode (Steps 12-15). The grids are then checked to identify the most suitable for data collection (Steps 16-17). The beam setting parameters in the EPU software that are to be used during automated collection are then set (Box 1, Steps 18-21), and the image shift calibrations performed (Steps 22-25). A low magnification atlas is then taken of the grid to help identify regions of appropriate ice thickness for data collection (Step 26). The areas for automated collection are then selected (27-30). Direct alignments are checked (Step 31), a gain reference is the taken (if required), and the final imaging parameters set (Steps 32-37). Final checks are then performed to ensure no key steps have been missed (Step 38-39), then the automated collection is started (Steps 40-41). As data is generated, it goes through organisation and pre-processing to enable monitoring of data quality during the session and decrease time from specimen to structure (Steps 42-51).

The protocol uses Thermo Fisher's Titan Krios electron microscope, EPU automation software with either F3EC or Gatan K2 detectors to collect data, saves the data to storage systems, and pre-processes the data using RELION2.1 [15,16], MotionCor2 [17] and Gctf [18]. It also performs statistical analyses to assess and maximise data quality. It is essential that users choose the most appropriate microscope hardware,

imaging conditions and data collection schemes to answer the scientific question, as these choices are sample specific. We offer generic advice on this, with information on direct electron detector choice (Box 2) and examples of imaging conditions for a variety of samples (Table 1).

This protocol is for single particle data collections and is not restricted to any particular sample type. The protocol for can be readily adapted to any Thermo Fisher Scientific microscope with an autoloader, including Talos Arctica and Glacios models with EPU installed. The data pre-processing workflow described in this protocol can be adapted for any electron micrograph movies created by any direct electron detector from any microscope. It is also modular, so that different software packages can be readily interchanged. This workflow specifically describes data collection for processing by single particle analysis, but changes to the protocol would allow it to be adapted for other types of data collection, notably tilt series collection.

Limitations

In this protocol we describe how we utilise our specific hardware setup at ABSL using the Titan Krios microscope equipped with a F3EC or integrated K2 with our specific data storage systems and processing hardware. This will need to be adapted for each individual hardware setup at new facilities. This protocol describes a 'standard' single particle data collection. In some cases where the specimen has specific challenges, such as preferred orientation, the method could be easily altered to include collection of tilted data [19].

There are a range of other software packages that can be used for automated data collection for single particle analysis, notably SerialEM [20] and Leginon [21]. These are both attractive options which can offer benefits including increased speed of data collection and greater flexibility. For data transfer and pre-processing there are a huge variety of options many of which could achieve comparable outcomes. Many Titan Krios sites use programs such as SCIPION [22] or Focus [23], or their own scripts as wrappers to call external programs for CTF estimation or motion correction. A range of programs for both motion correction and CTF estimation are available, including Unblur and Summovie [24], and CTFFIND4 [25].

Experimental design

One of the first decision points is the choice of hardware, particularly electron detector (Box 2) and whether to use a phase plate (Box 3). This choice may be predetermined based on the user's access to equipment. For clarity, the protocol described below is for F3EC. For K2, the full protocol can be found in Supplementary Methods 1. A flowchart of the individual protocols for each detector, and their estimated timings is shown in Figure 1.

In this protocol we assume the microscope is properly aligned, therefore microscope column alignment is not discussed. We do assume that any user looking to follow this protocol and set up data collection on high-end TEMs will have had training in the basic operation of TEM's and be familiar with the principles of microscopy. Accessing training in the operation of (high-end) TEM's is a major challenge for cryo-EM as a rapidly expanding field, and the publication and sharing of training resources and protocols is one contribution that can be made to tackling this challenge.

This workflow assumes that grids have been thoroughly pre-screened and identified as suitable for automated data collection, and so we include only limited information on grid screening. Without a sample of sufficient quality (clean, vitreous ice, good particle distribution showing range of orientations, minimal heterogeneity or aggregation of sample), even with the most optimal microscopy a high-resolution structure cannot be achieved. Many variables can be altered when optimising sample preparation, which have been reviewed elsewhere [9,10].

Materials

Equipment

Clipping grids

- Liquid nitrogen (LN_2) (Caution can cause cryogenic burns and asphyxiation in confined spaces).
- Negative pressure tweezers (Dumont N5)
- Positive pressure tweezers (Dumont SS140)
- Cryo-EM grids of specimen of interest
- C Clip ring (Thermo Fisher Scientific; 1036173)
- C Clip (Thermo Fisher Scientific; 1036171)
- C clip insertion tool (Thermo Fisher Scientific; 1115575)
- Autogrid container (Thermo Fisher Scientific; 1084591)
- Autogrid assembly station (Thermo Fisher Scientific; 1130697)
- Nanocab (Thermo Fisher Scientific; 1121822)
- Krios loading station (Thermo Fisher Scientific; 1130698)
- Autogrid tweezers (Thermo Fisher Scientific; 1121750)
- Cassette Tweezers (Thermo Fisher Scientific; 1121751)
- Krios Cassette (Thermo Fisher Scientific; 1121816)

Microscopy

- Thermo Fisher Scientific Titan Krios microscope
- TEM software (Thermo Fisher; 2.9.1)
- TEM Imaging & Analysis (TIA) (Thermo Fisher; 4.17)
- FluCam (Thermo Fisher; 6.9.1)
- EPU software (Thermo Fisher; 1.11.0)
- Electron detector: F3EC and Buddy camera with live view such as Ceta (Thermo Fisher) or BioQuantum-K2 summit direct electron detector (Gatan)
- TEM software (Thermo Fisher; 2.9.1)
- TEM Imaging & Analysis (TIA) (Thermo Fisher; 4.17)
- FluCam (Thermo Fisher; 6.9.1)
- EPU software (Thermo Fisher; 1.11.0)
- GMS 3/Digital Micrograph (DM) (Thermo Fisher; 3.22)

Optional

- AutoCTF (Thermo Fisher; 0.6.9)
- Volta Potential phase plate (Thermo Fisher)

Computing

Critical: All computing described in this protocol is connected with 10Gbit fibre ethernet and operating on microscope UPS (Uninterruptible Power Supply).

- Falcon 3EC direct electron detector with manufacturer provided 60Tb offload server or Gatan K2 with manufacturer provided PC with RAID and 60Tb offload server (Dell, PowerEdge R730)
- On-the-fly GPU server (Dell PowerEdge 7910) with 4x QUADRO M4000 8Gb GPU (NVIDIA)
- GPFS storage node (Dell, PowerVault MD3860i) on storage servers (Dell, PowerEdge R430)
- Scripts for data processing. The scripts described in the manuscript, along with a modified copy of pipeliner.cpp for Relion are available for download at https://github.com/Leeds-ABSL/ABSL_pipeline. A description of the scripts used is available in Supplementary Note 1.

Procedure

Critical- This procedure is designed to act as an aide-mémoire for more experienced users, and a more complete guide for the non-expert. All microscope users should expect local rules or procedures to be in place, and local variations in the layout of software. Please check with local facility management before using this protocol.

Critical- Steps 12 onwards describe a protocol specific for the F3EC workflow. When using the K2-summit with energy filter, follow the procedures in Supplementary Methods 1.

Transfer of cryo-EM grids to the microscope (30 minutes)

Critical- Steps 1-11 are visualised in Supplementary Video 1.

Critical- To reduce contamination on grids steps 1-9 should ideally be carried out in a dehumidified environment using clean liquid nitrogen (LN_2), freshly-decanted from a pressure vessel or clean onion dewar. All tools should be dry, and at room temperature (20-21 ℃) before cooling in LN_2. Tools must be warmed back to room temperature and dried in a warming cabinet or heat block in between uses. Throughout the process levels of LN_2 should be kept topped up to appropriate levels to reduce on grid contamination and reduce the risk of grid devitrification.

1. Retrieve cryo-EM grids from liquid nitrogen storage.

2. Load the C-clips into the C-clip insertion tool using tweezers. With the base of the tool on a flat, clean surface press down on the C-clip insertion tool so the C-clip becomes positioned at the rim of the tool.

Critical Step- ensure anything which will enter into the vacuum of the microscope is not touched with bare hands, as oils will deteriorate the condition on the vacuum. This includes C-clips, C-clip rings and the cassette.

Critical Step- inspect the C-clip inside the tool to ensure it has not become deformed during this process. The C-clip should sit flush around the rim of the tool. If any perturbations are seen, for example an end of the C-clip is bent, or it is not sat fully flush around the rim of the tool, reject these C-clips.

3. Cool the autogrid assembly station to liquid nitrogen temperature. Leave for several minutes for the temperature to equilibrate. Following equilibration, the level of the liquid nitrogen should either allow the transfer of the cryo-EM grids (see Step 4) to be completed under nitrogen vapour, or in a thin layer of liquid nitrogen. The station may need to be topped up throughout the procedure to ensure samples remain vitrified.

4. Transfer the cryo-EM grids to be clipped to the autogrid assembly station.

5. Insert C-clip rings into each of the four positions of the autogrid assembly station, ensuring they are orientated so the flat side is positioned against the base of the station. Cool the autogrid tweezers for manipulating grids and C-clip insertion tools (pre-loaded with C-Clips, as described in Step 2). Gently transfer the grid to the C-clip ring. Move the station around to the ‘closed’ position and insert the C-clip ring tool over the top, ensuring it is straight. Press down the button on the top of the tool to release the C-clip. Remove the C-clip ring tool and turn the station to the open position.

6. Use the autogrid tweezers to flip the autogrid assembly 180^{o} to ensure the grid is properly and securely clipped. It is recommended to repeat this twice for each grid. Once clipped grids are now referred to as autogrids.

Critical Step- it is vital that grids are securely clipped. If the autogrid assembly falls apart within the microscope, either the autoloader or the octagon, this can cause serious problems. If grids are bent they may not be securely clipped and so these should be rejected.

?Troubleshooting

7. Repeat steps 5-6 for each grid that is to be clipped.

8. Transfer the autogrids to the autogrid container. In the autogrid container, autogrids should be orientated so the flat base of the C-clip ring faces away from the notch.

9. Take a clean, room temperature autoloader cassette and transfer to the loading station. Cool down the loading station and a NanoCab to liquid nitrogen temperature. Once cooled, transfer the autogrid container containing the autogrids to the station and use the autogrid tweezers to position the autogrids in the slots of the cassette so the flat side of the autogrid is facing the gold band of the cassette. Once loaded, visually inspect to check if the grids are all fully inserted down into the slot. Use the autogrid tweezers to press very gently against the top side of the autogrid and check that it springs back into position, indicating it is seated properly against the springs.

Critical Step- If you apply too much force the grid will become dislodged from the cassette slot, so apply only slight pressure.

?Troubleshooting

10. Dock the pre-cooled NanoCab, full of LN_2, against the loading station. Slide the cartridge arm over the cassette and hold down the button on the arm to grasp the cassette. Slide the cassette into the NanoCab, release the button on the arm and withdraw. Undock the NanoCab. Use a pair of blunt tweezers to press down on the cassette to ensure it is properly seated against the bottom of the NanoCab (the cassette should not move upon pressing firmly down).

11. Take the NanoCab and insert it into the autoloader of the microscope. Click ‘Dock/Undock’ in the User Interface (UI) software. When the loading process is complete, remove the NanoCab and close the microscope doors. It is useful to monitor the Autoloader menu and autoloader vacuum display to ensure this completes successfully.

?Troubleshooting

Perform initial microscope checks (10 minutes)

12. Check that all required software is open (UI, FluCam viewer, TIA and EPU).

13. Ensure you are in TEM mode (UI> Setup > Beam settings > TEM mode should be yellow), Autozoom is off (UI> Setup > Beam settings > Tab out > Autozoom should be grey), the autoloader turbo is auto off (UI> Autoloader > options > check Turbo Auto Off). Check the UI for any error messages, as indicated by red icons with a black cross. If errors are present consult with a member of facility staff.

14. Check that the gun lens, high tension and extraction voltage in UI are set to the values recommended by your facility management (based on recent alignment of the microscope).

15. Check in UI software that vacuum values are green and all autoloader temperatures are colder than -170 ^{o}C (this may take > 10 minutes after loading specimens).

Identify grid for data collection (0.5-4 hours)

16. Start grid inventory (UI > Autoloader >Tab out > Inventory) and check the number and position of the grids in the microscope correspond to those loaded. Once the inventory has finished, select the desired grid in the autoloader and press load. When complete, 'cartridge successfully loaded' will be displayed in the UI, the autoloader menu will not display anything in the log area, and the grid will be shown as located in the column if you hover the mouse over that slot in the autoloader menu. The slot position in the autoloader will be coloured yellow.

?Troubleshooting

17. Confirm that the grid is suitable for automated data collection, i.e. its appearance is consistent with your screening images. If the grid has been previously loaded into a microscope, assess the quality of the ice to ensure there is no increase in surface ice contamination or any evidence of devitrification.

Critical Step- Here we assume that the grids loaded have been pre-screened and identified as suitable for data collection. The 'ideal' ice conditions vary dramatically for different samples, for more guidance see references [9,10].

?Troubleshooting

Determine beam setting pre-sets (30 minutes)

18. Navigate to an intact square which you are unlikely to use for data collection (a dry square works best, and areas of thick ice should be avoided). Set the eucentric height manually at low magnification to bring the specimen roughly to eucentric height. In EPU, navigate to Preparation > Acquisition Settings and Optics > Presets and push the 'Grid Square' settings to the microscope by clicking 'Set' and view on the FluCam viewer. Use the joystick to centre the image (5mm circle on the FluCam Viewer) on a visible feature like ice contamination. Tilt the stage to 20° by navigating to UI> Stage > Set Alpha (this becomes yellow when activated) and use the Z axis buttons on the right-hand panel to move the feature back to the centre of the screen. Click 'Set Alpha' again to return the stage to 0° tilt. If this has been done correctly the visible feature should still be centred. If not repeat the process.

?Troubleshooting

19. In EPU software, each beam setting preset should be checked to ensure variables are optimal for collection, as these will vary experiment to experiment (Box 1, Supplementary Table 1 and 2). These are set in EPU > Preparation > Acquisition and Optics Settings > Presets dropdown menu. For each setting, on the FluCam viewer ensure that the beam fully illuminates the entire detector, that the beam is parallel (this can be checked in UI > Beam Settings > 'Parallel', 'spreading' or 'condensing') and that you are using the dimmest beam needed to complete the Atlas, Grid Square and Hole/Eucentric tasks. This exposes your specimen to as little electron dose as possible prior to the data acquisition exposure. Small adjustments to variables such as the illuminated area can be made later in the session, but changes to spot size can mean repeating steps and can cause delays to data acquisition.

20. Check that the electron dose is appropriate for your detector mode choice and data collection goals (see Box 2 for information of detector dose and Table 1 for example parameters). To set the detector mode, in Preparation > Acquisition and Optics Settings > Presets > Data acquisition select either 'Counted' or 'Linear'. To calculate electron dose, move to an area where there is no specimen, i.e. a broken grid square. Push the data acquisition beam settings to the microscope by clicking 'Set' in EPU. Press 'Eucentric focus' on the hand panel. With the FluCam viewer, check that the beam is centred over the detector, if not use the beam shift direct alignment to correct (Supplementary Methods 2). The condenser apertures should be set

in the way you intend to use them during data acquisition. In the UI, go to the Camera > Check F3EC is selected and inserted (yellow) > Bias/Gain tab > Reference image manager. Make sure the the beam is unblanked.

A) For a F3EC integrating mode data collection session, select 'normal' from the 'Available Reference Images' in Falcon reference image manager and press 'Measure Dose'. We typically use 40-100 e^-/pixel/second.

B) For a F3EC counting mode data collection session, select one of the two electron counting reference images from the 'Available Reference Images' in Falcon reference image manager and press 'Measure Dose'. 0.5-0.7 e^-/pixel/second is recommended.

This will give a reading in electrons/physical pixel/second.

?Troubleshooting

21. To work out the dose per $Å^2$, first calculate your $Å^2$. i.e. at 75 k magnification, each physical pixel represents 1.065 Å (calibrated for ABSL krios 1, F3EC combination), and 1.065 x 1,065 = 1.13 $Å^2$. Now divide your dose per physical pixel per second by the $Å^2$ to get e^-/ $Å^2$/second, e.g. if you have 50 e^-/pixel/second, 50/1.13= 44.3 e^-/ $Å^2$/second. To get the total dose, multiply this figure by the length of exposure in seconds, for example for a 2 second exposure the total accumulated does per exposure would be 44.3 x 2 = 88.6 e^-/ $Å^2$. If you need to, alter data acquisition parameters to achieve the desired dose. Be mindful to keep the illuminated area as small as possible to a) allow multiple exposures per hole, if appropriate and b) to reduce the likelihood of the beam pre-exposing neighbouring acquisition areas. However, the beam should be large enough to prevent any beam fringing appearing at the edges of the image. Beam spot size, illuminated area and, if needed, condenser 2 aperture can be changed to fine-tune the desired dose.

Image shift calibrations (10 minutes)

22. Navigate to an intact grid square. Centre on a feature of interest that will be visible at the 'Atlas' and 'Data Acquisition' magnifications, for example a piece of ice contamination. Roughly set eucentric height, as described in step 18. Run the auto function 'set eucentric height by stage tilt'.

?Troubleshooting.

23. In EPU 'Set' data acquisition beam settings to the microscope. On the FluCam viewer check that the feature is visible in the centre of the screen. Ideally centre a distinctive feature such as the point of an ice crystal. Press 'Eucentric focus' on the right-hand panel. In EPU, navigate to Preparation > Calibrate Image Shifts and click Start Calibration.

24. In EPU, an image at data acquisition magnification will appear with a marker in the centre. If the marker is placed on your recognisable feature, click 'proceed'. If you want to move it, double click in the image and click 're-acquire' until the feature is in the centre of the image.

25. The microscope will now shift to the next highest magnification in the beam presets and take an image. In the second image, double click so the same feature is centred identically to the first image and click 're-acquire'. Repeat until the feature is identically centred and then click 'proceed'. Repeat this step until it says 'image shift calibration finished successfully' in the EPU log on the right.

Atlas the grid (15 minutes)

26. Ensure the objective aperture is not inserted (Apertures > Objective > none). In the EPU, navigate to > Atlas > Session Setup, click on 'Create new sample' and name your session. We recommend a format such as Username_myprotein_date_ATLAS. Save images in MRC format and store atlas directly onto the F3EC offload server (usually designated Z: in a standard Krios installation). Click on 'Acquire' to acquire the atlas.

Select square for data acquisition and define template (1-2 hours)

27. In EPU, navigate to EPU > Session Setup and click on 'New Session'. We suggest a session name such as Username_myprotein_date_EPU. Select 'manual selection', save images in MRC format and save data directly to the F3EC offload server (usually Z: a directory with the session name will be created and all data will write here. If you choose a subdirectory data will be split between different directories). Choose the type of grid from the drop-down menu i.e. Quantifoil and the size of the holes, or Lacey grids, then click 'Apply'.

28. Go to EPU > EPU > Square Selection, if all squares are green, click 'Unselect all' in the top left. Open tiles (hover over atlas image > right click > open tile) and look to see which areas of the grid have appropriate ice thickness for data acquisition. For most specimens you can tell at this low magnification if an area is dry (no ice) or the ice is too thick (Figure 2a). When you have identified a square, select it by hovering over the square followed by right click > 'add', then right click > 'move stage to grid square'.

29. Bring grid square to eucentric height by following step 18 for manual adjustment, and run the auto function 'set eucentric height by stage tilt' as described in step 22. Press 'Eucentric focus'. Go to EPU > EPU > Hole Selection and click 'acquire'. This will save your x, y and z coordinates.

30. Set the template for automated acquisition, including areas and variables for autofocus, drift measurement and image acquisition. This process varies depending on the type of grid hole, which can be either regular such as Quantifoil ® and C-flat ™ (Option A) or irregular such as lacey carbon (Option B).

?Troubleshooting

A) For grids with a regular array of holes such as Quantifoil ® and C-flat ™

i. Go to EPU > EPU > Hole Selection and click 'acquire'. The first time you do this, click 'measure hole size'. Move and adjust the yellow circles so they match the hole size then click 'find holes'. Repeat this until the software finds the hole sizes correctly (use the zoom function to see more accurately). This only needs to be done once per session.

ii. The 'filter ice quality' histogram on the right can be adjusted by moving the red histogram lines in order to adjust and refine hole selection. Use the 'Select' tools menu at the top to fine-tune hole selection. For example, remove holes which are empty or in which the ice is too thick/thin/contaminated and holes that are close to the grid bar (Figure 2c).
?Troubleshooting

iii. Go to EPU > EPU > Template Definition. Click 'Acquire' and 'Find and Centre Hole'.
?Troubleshooting

iv. Change the 'Delay after Stage Shift' and the 'Delay after Image Shift' times to between 1-5 seconds, depending on the stability of the stage. 1 second is generally sufficient, but this time can be increased if unacceptable drift is observed (as measured during motion correction in step 42-51).

v. Click ‘Add acquisition area’, then click anywhere on the template image. The outer circle represents the illuminated area, the inner square represents the exposure area (Figure 2 e,f). On the top right, add your defocus range for the acquisition (add the defocus you would like, making sure it’s a negative number if you wish to work in standard defocused mode imaging) see Table 1 for examples.

vi. Move the acquisition area to desired location. Depending on the sample, hole size and user preference you can do one exposure in the middle of the hole, ideally with the illuminated area covering the whole hole (this may help reduce the effect of charging), or add multiple exposures around the hole, taking care not to overlap illuminated areas into neighbouring exposure areas (Figure 2e,f).

vii. If you want to add additional exposure areas, click ‘Add Acquisition Area’, click the template image and move the exposure to the desired location (double checking they have retained the defocus list). When choosing the number of acquisitions to take around a hole, bear in mind that the beam diameter shown in EPU can vary by +/- ~10 %, depending on the accuracy of alignment, so it’s safer to leave some space between the acquisition areas or check the physical beam and the virtual beam in EPU coincide, by burning a hole on the carbon to confirm true size of the illuminated area.

viii. Click ‘add Autofocus Area’ and click anywhere on the image. Move the autofocus area to the carbon surrounding your hole. Standard practice is to autofocus every 5-15 µm, depending how large the variation in height is across the grid square (with more uneven grids we recommend to autofocus more often). Focus using the objective lens and ensure autostigmate is set to ‘no’.

ix. Click ‘add Drift Measurement area’. Perform a drift measurement once per grid square, set the threshold to 0.05 nm/s. If your microscope has known stage stability issues you may have to relax this threshold and/or perform drift measurement more often. The drift measurement area should overlap directly with the autofocus area. Make sure neither drift nor autofocus area overlap with an acquisition area either in this hole, or neighbouring holes (Figure 2e,f).

x. Check the template layout by running the ‘template execution’ function. This is a good idea in order to see if you need to move your acquisition areas (e.g. too much/not enough carbon in images) or would like to assess particle distribution.

B) For irregular arrays of holes such as lacey carbon

i. In EPU > EPU > Area Selection, choose the spacing between acquisitions. It is recommended that you add ~ 300-400 nm to the data acquisition illuminated area (e.g. illuminated area 0.9 µm, choose spacing of 1.3 µm), to ensure the illuminated areas do not overlap. Click ‘View Pattern’. Use area selection tools to add or remove desired acquisition areas (Figure 2d).

ii. Change the stage shift delay to 5 seconds. A shorter or longer time may be used depending on the stability of the stage. As collection on irregular carbon typically means a greater number of stage shifts, a longer stage shift delay is recommended compared with regular arrays.

iii. Set the autofocus recurrence to ‘after distance’. Depending on the height variation of your grid, every 8-15 µm is recommended. Focus using the objective lens, ensure autostigmate is set to ‘no’.

iv. In ‘Data Acquisition Area Settings’ add your defocus range for the acquisition (add the defocus you would like, making sure it’s a negative number if you wish to work in standard defocused mode imaging. You only have to do this for one grid square, it remembers for subsequent squares.

v. Use the bottom histogram (filter ice quality) on the right and the exposure area selection tools to optimise hole selection in order to exclude areas of sub-optimal/no ice.

Check direct alignments (30 minutes)

31. Perform direct alignments on the grid (Supplementary Methods 2).

! CAUTION: In many electron microscopy facilities, direct alignments are only carried out by Facility Staff. We provide a guide to performing the basic direct alignments in Supplementary Methods 2 but you should check local procedures in your EM Facility before attempting.

Critical Step- Some direct alignments cannot be adequately performed on UltrAuFoil ® grids [26], and so if using these, perform direct alignments on a carbon grid or cross grating before starting data collection.

Perform gain reference and set final imaging settings (10-120 minutes)

32. The gain reference should be of sufficient quality that when an image is taken using the parameters chosen for data acquisition, with no specimen obstructing the beam, a completely featureless image is obtained. The procedure for collecting a new gain reference in integrating mode is very quick to perform, and so this can be done for each data collection run. In counting mode, the procedure takes 1.5 h so collection of a new gain reference is not recommended unless required. In order to collect a new gain reference, make sure the FluScreen is lifted, and that you are over a hole with absolutely no obstruction in the field of view. Ensure that the beam is not blanked

33. Go to EPU > Preparation> Acquisition and Optics Settings > Presets > Data acquisition and click ‘Set’.

34. Check that the beam is centred and there is no beam fringing visible in the image. If necessary, adjust using the beam shift direct alignment (Supplementary Methods 2).

Critical Step- If beam fringes are present in the gain reference they will be seen in every image of the data collection, even if the illuminated area is subsequently expanded. Take great care to ensure that the beam diameter is sufficiently large so no beam fringing is seen.

35. In the UI > Camera (check F3EC is selected and inserted) > Bias/Gain tab > Reference Image Manager. When using F3EC in integrating mode, follow Option A, when using F3EC in counting mode, follow option B.

A) F3EC in integrating mode - 10 minutes

i. In ‘Falcon Reference Image Manager’, select ‘normal’ from the ‘Available Reference Images’, check that the exposure time is 10 s and images to average is 1.

ii. Click ‘Measure Dose’. You might have to click ‘Measure Dose’ twice to get a reliable dose reading. Check this is consistent with the earlier value checked at step 20.

iii. Select ‘normal’ gain reference, click ‘Acquire selected gain reference’.

iv. Once complete, acquire a test image in UI > Camera > Acquire (use 2 s exposure). Inspect the image for signs of beam fringing at the edges. Check FFT has no features. If the

inspection reveals something wrong (e.g. beam clipping), rectify the problem and retake the gain reference until the flat field image is completely featureless.

B) F3EC in counting mode - 1.5 hours.

i. In 'Falcon Reference Image Manager', select "pre-EC" from the Available Reference Images and click 'measure dose'. The reading should be 0.5-0.7 e^-/pix/second. Change the spot size and/or beam intensity in order to correct the dose.

i. Take a counting mode image in TIA to ensure there is no beam clipping in the image (UI > Camera > select 'counting' tick box > 60 s exposure > Acquire) if there is expand illuminated area. If the image appears as a flat field image with no features there is no need to take a gain reference.

ii. If a new gain reference is required, in 'Falcon Reference Image Manager', click 'pre-EC' from the Available Reference Images, check that the exposure time is 60 s and images to average is 45 and click 'Acquire selected gain reference'.

iii. Once complete, select 'post-EC' and 'Acquire selected gain reference'.

iv. Once complete, acquire a test image (UI > Camera > ensure that the 'counting' box is checked > Acquire with 60 s exposure). Inspect to ensure flat field image. If the inspection reveals something wrong (e.g. beam clipping), rectify the problem and retake the gain reference until the flat field image is completely featureless.

36. With no obstruction in the field of view, take a dose measurement and use this to calculate your final electron dose parameters as in step 21.

37. In EPU > Preparation > Acquisition and Optics Settings> Presets > Data Acquisition check you are collecting fractions. In EPU > Preparation > Direct Detector Dose Fractions, set the number of factions you would like to split your exposure into and click 'Equal Doseage' and 'Validate' to ensure values are compatible with software. We recommend a number of fractions that results in between 1-2 e^-/A^2/fraction, as this seems to represent a good trade-off between signal in each frame to perform motion correction but still sufficiently frequent to correct for beam induced movement.

Final checks (10 minutes)

38. Before automated acquisition begins, perform a final check of variables that can affect data quality or prevent common mistakes. In UI> Autoloader > Turbo > Options > click Turbo 'auto off'. Ensure the Titan Krios enclosure is shut.

39. Check the microscopy parameters. Ensure that the aperture series is as desired, the beam is centred in the data acquisition beam settings and no beam fringes appear in the image. Check that you are collecting fractions, if desired. Ensure that the disk you are writing to has sufficient space available for the entire planned data collection.

Start automated collection (5 minutes- data typically collects for 24-72 h)

40. In EPU > EPU > Automated Acquisition, 'Start' the EPU run. We suggest you now move on to start data organisation and on the fly image processing steps, then come back to select more grid squares as described in steps 28-30, but note the template only needs to be set once for each data collection.

?Troubleshooting

41. Fill out a data acquisition report (Supplementary Table 4).

Data transfer, organisation and on-the fly processing (20 minutes- carries on for the duration of data collection)

CRITICAL: Data must be moved from the microscopes limited offload server (at ABSL, ~60Tb) to a storage system (ABSL ~ 5Pb). To do this we use a custom copying script (ABSL_OTF.sh; see Supplementary Note 1). The following steps (42-43) describe the usage of this script which were designed for the computational set up at ABSL but can be readily adapted to other hardware set ups.

42. Open a terminal window and navigate to a directory in which the user will process their data. Create a directory for processing i.e. 'myprotein_date' (i.e. run command: `mkdir myprotein_date`).

43. Run the script 'ABSL_OTF.sh' with the appropriate arguments. Run command:

```
sh ABSL_OTF.sh <runtime in mins> <which microscope (krios1 or gatan)>
<what you want directory to be called (date is automatically added)>
<name of EPU directory>
```

e.g. `sh ABSL_OTF.sh 2880 krios1 myprotein Username_myprotein_date_EPU`

Allow the script to run, data will be transferred over in blocks every 30 seconds.

Critical: The network implementation at ABSL is based on 10Gb ethernet, allowing sustained transfer speeds of ~1Tb per hour. Slower network speeds will create a backlog of untransferred/ unprocessed data that might take several hours/days to be cleared after the data collection finishes.

44. Begin on-the-fly processing of the data. In our workflow a slightly modified version of RELION 2.1 (ref. [16]) (Supplementary Note 2) is used for on-the-fly processing. ABSL_OTF.sh copies data to a storage location and then a Raw_data directory in a separate, designated processing directory, then it creates symbolic links to the raw images (unaligned frame images in mrc stack format, written by EPU). This serves to protect the original data so users cannot accidently delete or modify the original files.

45. Open RELION. Set up an import job. Select Raw_data/*.mrc as the input files. Schedule, but do not run, this job.

46. Set up a motion correction/dose weighting job in RELION. Select 'Import/job001/movies.star' as the input files. Set the other motion correction parameters, including dose weighting, as desired. Schedule, but do not run, the job.

47. Set up a CTF determination job in RELION. Select 'MotionCor/job002/corrected_micrographs.star' as the input file. Schedule, but do not run, the job.

48. Use the autorun function to run the scheduled jobs for the duration of the data collection run by navigating to Autorun > 'Run scheduled jobs'. Set the 'Run the jobs how many times?' and 'Wait at least in between (in minutes)?' parameters. Begin running scheduled jobs with the execute button.

49. Assess the quality of the data. ABSL_OTF.sh will prepare a diagnostic image named 'micrograph_analysis_0.png' every time a cycle of CTF estimation is completed (Figure 4).

50. Finish the on the fly processing. After the data collection has finished allow ABSL_OTF.sh to run until all data has been transferred. Once all files have been motion corrected and CTF estimated stop the scheduled jobs in RELION by navigating to Autorun > 'Stop running scheduled jobs'. Allow any active RELION jobs to finish running.

51. The data are now ready for downstream processing steps. As part of this workflow users can also use RELION's automated particle picking and 2D classifications as described in ref [16].

Timing

The time taken to set up automated data collection is variable depending on the hardware used, the length of the automated data collection session and experience of the user. While this protocol may take as little as 1 h for a short collection set up by an experienced user, typically 3-6 hours would be standard (see Figure 1). For some samples, especially where there is significant variation in particle distribution across a single grid, much more time may be needed picking acquisition areas.

Steps 1-11, Transfer of cryo-EM grids to the microscope: 30 minutes
Steps 12-15, Perform initial microscope checks: 10 minutes
Steps 16-17, Identify grid for data collection: 0.5-4 hours
Steps 18-21, Determine beam setting pre-sets: 30 minutes
Steps 22-25, Image shift calibrations: 10 minutes
Step 26, Atlas the grid: 15 minutes
Steps 27,30, Select square for data acquisition and define template: 1-2 hours
Step 31, Check direct alignments: 30 minutes
Steps 32-37, Perform gain reference and set final imaging settings: 10-120 minutes
Steps 38-39, Perform final checks: 10 minutes
Steps 40-41, Start automated collection: 5 minutes- data typically collects for 24-72 h
Steps 42-51, Data transfer, organisation and on-the-fly processing: 20 minutes- carries on for the duration of data collection

Troubleshooting

Troubleshooting guidance can be found in **Table 2**.

Table 2. Troubleshooting Table

Step	Problem	Possible reason	Solution
6	Autogrid assembly repeatedly comes apart upon checking	Grid is bent, putting pressure on the assembly	Make sure only visibly flat grids are clipped, do not attempt to clip grids which have been bent or folded over.
6	Autogrid assembly repeatedly comes apart upon checking	Manufacturing faults in C-clips	When there are repeated problems with a single batch of C-clips, there may be an issue with the manufacturing of the C-clips. We recommend keeping track of what batch of C-clip and C-clip rings are being used, and contacting the manufacture if this is suspected.
9	Autogrid assembly comes out of position in cassette	Autogrid not loaded into correct position in the cassette	Make sure when the autogrid is picked up with the autogrid tweezers they fully cover the grid,

			this helps to ensure when loading into the cassette it is sufficiently inserted. When in the cassette, the autogrid tweezers can be used to gently grip and push the autogrid directly downwards to ensure it is fully in position before pushing to ensure it is correctly seated agains the springs of the cassette
11	NanoCab does not dock properly to microscope	Cassette is not properly seated in NanoCab	Use the back of a pair of tweezers to press firmly against the cassette to ensure it is fully seated against the base of the NanoCab. The pin on the top of the NanoCab should be protruding.
16	Grid inventory does not match up with samples that were loaded, or when grid is loaded no grid can be seen	Autogrid assembly has come apart	Safest option is to unload all grids from the microscope octagon (if loaded) and then autoloader and assess if autogrid assembly has come apart. Try to account for all components (grid, C-clip, C-clip ring) as these can cause problems in the octagon or autoloader of the microscope, for example by blocking valves.
17	Grids appear to have surface ice contamination	Crystalline ice in LN_2	Use LN_2 freshly decanted from a pressure vessel into a clean, dry dewar in steps 1-11
17	Grids appear to have surface ice contamination	Crystalline ice forming in LN_2 during clipping procedure	Where possible, perform steps 1-11 in a dehumidified environment ideally < 20 % RH. Cool down the clipping station immediately before using and perform the clipping as quickly as possible to minimise time for water in the air to condensate on the cold LN_2.
18 (Supplementary Methods 1, 17)	Autofunction ‘eucentric height by stage tilt’ fails.	Too far from eucentric height.	Set ‘grid square’ magnification and manually set eucentric height using stage tilt and z axis buttons on hand panel.
18 (Supplementary Methods 1, 17)	Autofunction ‘eucentric height by stage tilt’ fails.	Not enough signal in images	Ensure there is carbon in the image. Bin the image to 2 in ‘hole/eucentric’ preset. Increase the brightness of the beam by decreasing illuminated area, using brighter (lower number) spot size.
20	Measure dose in reference image manager gives no value or very low value	Software bug/dose measure while screen is retracting	Ensure the FluScreen is retracted, the beam is not blanked, and the F3EC inserted. Press the measure dose button again

30Aii (Supplementary Methods 1, 25Aii)	Hole selection tools are greyed out	Incorrect session settings	In EPU ‘session set up’ ensure ‘manual’ is selected.
30Aiii (Supplementary Methods 1, 25Aiii	Software unable to find hole	If yellow circle doesn’t match up to hole size, measured hole size is wrong.	Go back to grid square selection > measure hole size
30Aiii (Supplementary Methods 1, 25Aiii)	Software unable to find hole	Signal is too low to reliably find hole.	Bin the image to 2 in ‘hole/eucentric’ preset. Increase the exposure time, or brightness of the beam by decreasing illuminated area, using brighter (lower number) spot size (recommended to keep this the same as data acquisition where possible)
30Aiii (Supplementary Methods 1, 25Aiii)	Hole appears different size to yellow circle	Hole size is incorrect	Go back to step 30 ii and ensure hole size is correctly measured. Sometimes ice halo effects can produce a misleading image on the lower magnification square image so use of the zoom tool is recommended. Ideally this should be done on a thin ice/dry area for the most accurate results.
40 (Supplementary Methods 1, 29)	Images coming off/by motion correction analysis, have unacceptable drift	Source of vibration present	If the enclosure of the Titan Krios is open more drift than usual may be seen. Ensure all doors are properly closed. Another common source is the autoloader turbo, ensure this is off (step 38)
40 (Supplementary Methods 1, 29)	Images being collected are seen to have motion and motion correction analysis shows unacceptable drift	Mechanical drift because of grid or stage.	If the support film of the grid is cracked, more drift may be seen in the images. Try to pick grid squares which have no cracks or broken areas. Some stages are less stable than others. If stage is the problem, try increasing the stage settling time (Step 30 A iv, B ii) or perform drift measurement more often (Step 30 A ix).
40 (Supplementary Methods 1, 29)	Images appear to have charging (often seen as localised areas of ‘drift’ or where image appears blurred)	Reasons for charging often unclear.	One strategy that can be tried is to use whole hole illumination. Set up beam settings and template (Step 30A) so at data acquisition magnification the beam illuminates the whole hole. This appears to improve charging in some cases.
40 (Supplementary Methods 1, 29)	Images being collected are seen to have (by CTF analysis) defocus outside desired range	Changes in Z height across grid	Autofocus more frequently (Step 30 A viii, B iii)

40 (Supplementary Methods 1, 29)	Images being collected are seen to have (by CTF analysis) defocus outside desired range	Autofocus procedure resulting in slightly different defocus compared with calculated results	Change the autofocus range set (Step 30 A viii, B iii)
Supp. Methods 1 Step 6	Unable to find Zero Loss Peak	Not enough dose on K2	Check K2 is inserted and beam is centred over GIF. With the slit out, in Digital Micrograph take an image in linear mode and ensure you have 1000's of counts on the detector. Use a larger C2 aperture and brighter spot if needed.

Anticipated results

Following this protocol, the user should be able to produce high quality electron micrographs for single particle data analysis and pre-process the results in close to real time. Monitoring the micrograph analysis output permits visualisation of the estimated defocus, resolution in the micrographs, astigmatism, and where relevant the phase shift. An example micrograph analysis is shown in Figure 4. The user can then make informed decisions about alterations to defocus range, objective stigmation. For phase plate data, useful phase shifts are between 20 and 120 degrees and so on-the-fly analysis of the phase shift allows a user to ensure data is collected in the optimal range by altering the frequency with which the phase plate position is changed.

The micrograph analysis script uses the resolution estimate from Gctf. Gctf V.1.06, estimates the resolution of a micrograph as the resolution at which the cross correlation coefficient (CCC) between the eqi-phase average and actual micrograph power spectrum falls to 0. We feel this overestimates the resolution and prefer to use a CCC cut off of 0.5, as implemented in the ABSL_EPA_CC_threshold.py script. The estimated resolution allows a user to determine if their dataset has high resolution features, which can be a good general indicator of data quality. However, this has two caveats. Firstly, that the resolution estimate is reliant on signal in the images, so images of grids with continuous carbon film will appear to have higher resolution than images of vitreous ice containing a small protein. The second caveat is that this estimate is based on signal transferred through the imaging system and recorded on the detector. This is not necessarily signal from the biological specimen. It should also be noted therefore that while this resolution estimate is an excellent indication of the quality of the micrograph (and the dataset), and thus a metric of microscope performance, the structure of the macromolecular complex being imaged might not be solvable to high resolution.

These on-the-fly analyses can also act as an early warning if there is a deterioration in microscope performance as a result of instability in air temperature or chilled water to the lenses. This is typically seen as a grid square-independent deterioration in the resolution over time and/or changes in objective stigmation.

An example dataset of Cowpea Mosaic Virus has been uploaded to EMPAIR (EMPAIR-10205; https://www.ebi.ac.uk/pdbe/emdb/empiar/entry/10205/), with the corresponding output from micrograph analysis in Supplementary Figure 1. The corresponding reconstruction has been uploaded under accession number EMD-3952 (http://www.ebi.ac.uk/pdbe/entry/emdb/EMD-3952)[12].

Concluding remarks

The aim of this protocol is to permit collection of high-quality single particle data and its facile organisation, storage and pre-processing as a prelude to 3D structure determination. Assuming a high quality, stable,

and homogeneous macromolecular complex is imaged. This protocol combined with single particle image processing techniques will lead to a high resolution cryoEM 3D reconstruction**.**

Competing financial interest statement

The authors declare that they have no competing financial interests.

Acknowledgements

The Titan Krios microscopes were funded by the University of Leeds (UoL ABSL award) and Wellcome Trust (108466/Z/15/Z). We are grateful to the electron microscopy community at Leeds and our external user base for continuing feedback on our procedures and example data collection parameters in Table 1. We would like to thank the Faculty of Biological sciences IT team at UoL for their help with setting up and maintaining the data processing and storage servers and data transfer scripts particularly Pasquale Pelliccia, Aidan Richmond and Michael Beck. ELH is partially funded by BBSRC (BB/L021250/1). MGI received funding from the European Research Council (ERC) under European Union's Seventh Framework Programme (FP7/2007-2013) ERC grant agreement no. 322408 and the MRC (MR/P018491/1). The script EPA_CC_threshold.py is a modified version of a script kindly provided by Rado Danev, who we also thank for helpful discussions around optimal use of the phase plate.

Author contributions

RFT and ELH wrote the EPU setup protocol. MGI and SR wrote the scripts. All authors contributed text to the manuscript.

Data and code availability statement

An example dataset of Cowpea Mosaic Virus has been uploaded to EMPAIR (EMPAIR-10205; https://www.ebi.ac.uk/pdbe/emdb/empiar/entry/10205/) and corresponding reconstruction EMD-3952 (http://www.ebi.ac.uk/pdbe/entry/emdb/EMD-3952)[12], an example output from the micrography analysis script can be seen in Supplementary Figure 1. The scripts described in the manuscript, along with a modified copy of pipeliner.cpp for RELION are available for download at https://github.com/Leeds-ABSL/ABSL_pipeline. A description of the scripts used is available in Supplementary Note 1. RELION 2.1 is available on a GPLv2 license.

References

1. Cheng, Y., Grigorieff, N., Penczek, P. A. & Walz, T. A primer to single-particle cryo-electron microscopy. Cell **161,** 438–449 (2015).
2. Renaud, J.-P. et al. Cryo-EM in drug discovery: achievements, limitations and prospects. Nat Rev Drug Discov **17,** 471–492 (2018).
3. Thompson, R. F., Walker, M., Siebert, C. A., Muench, S. P. & Ranson, N. A. An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. Methods **100,** 3–15 (2016).
4. Kuehlbrandt, W. The Resolution Revolution. Science **343,** 1443–1444 (2014).
5. Ruskin, R. S., Yu, Z. & Grigorieff, N. Quantitative characterization of electron detectors for transmission electron microscopy. Journal of Structural Biology **184,** 385–393 (2013).
6. Kuijper, M. et al. FEI's direct electron detector developments: Embarking on a revolution in cryo-TEM. Journal of Structural Biology **192,** 179–187 (2015).
7. McMullan, G., Faruqi, A. R., Clare, D. & Henderson, R. Comparison of optimal performance at 300 keV of three direct electron detectors for use in low dose electron microscopy. Ultramicroscopy **147,** 156–163 (2014).
8. Patwardhan, A. Trends in the Electron Microscopy Data Bank (EMDB). Acta Crystallogr D Struct Biol **73,** 503–508 (2017).
9. Noble, A. J. et al. Routine single particle CryoEM sample and grid characterization by tomography. eLife **7,** 32 (2018).

10. Drulyte, I. et al. Approaches to altering particle distributions in cryo-electron microscopy sample preparation. Acta Crystallogr D Struct Biol **74,** 560–571 (2018).
11. Amporndanai, K. et al. X-ray and cryo-EM structures of inhibitor-bound cytochrome bc1 complexes for structure-based drug discovery. IUCrJ **5,** 200–210 (2018).
12. Meshcheriakova, Y., Durrant, A., Hesketh, E. L., Ranson, N. A. & Lomonossoff, G. P. Combining high-resolution cryo-electron microscopy and mutagenesis to develop cowpea mosaic virus for bionanotechnology. Biochem. Soc. Trans. **45,** 1263–1269 (2017).
13. Rawson, S. et al. Elucidating the structural basis for differing enzyme inhibitor potency by cryo-EM. Proceedings of the National Academy of Sciences **115,** 1795–1800 (2018).
14. Baggen, J. et al. Role of enhanced receptor engagement in the evolution of a pandemic acute hemorrhagic conjunctivitis virus. Proceedings of the National Academy of Sciences **115,** 397–402 (2018).
15. Scheres, S. H. W. A Bayesian View on Cryo-EM Structure Determination. Journal of Molecular Biology **415,** 406–418 (2012).
16. Fernandez-Leiro, R. & Scheres, S. H. W. A pipeline approach to single-particle processing in RELION. Acta Crystallogr D Struct Biol **73,** 496–502 (2017).
17. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nature Methods **14,** 331–332 (2017).
18. Zhang, K. Gctf: Real-time CTF determination and correction. Journal of Structural Biology **193,** 1–12 (2016).
19. Tan, Y. Z. et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nature Methods **14,** 793–796 (2017).
20. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. Journal of Structural Biology **152,** 36–51 (2005).
21. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. Journal of Structural Biology **151,** 41–60 (2005).
22. la Rosa-Trevin, de, J. M. et al. Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. Journal of Structural Biology **195,** 93–99 (2016).
23. Biyani, N. et al. Focus: The interface between data collection and data processing in cryo-EM. Journal of Structural Biology **198,** 124–133 (2017).
24. Grant, T. & Grigorieff, N. Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. eLife **4,** e06980 (2015).
25. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. Journal of Structural Biology **192,** 216–221 (2015).
26. Russo, C. J. & Passmore, L. A. Electron microscopy. Ultrastable gold substrates for electron cryomicroscopy. Science (New York, N.Y.) **346,** 1377–1380 (2014).
27. Hesketh, E. L. et al. The 3.3 Å structure of a plant geminivirus using cryo-EM. Nat Commun **9,** 2369 (2018).
28. Hurdiss, D. L., Frank, M., Snowden, J. S., Macdonald, A. & Ranson, N. A. The Structure of an Infectious Human Polyomavirus and Its Interactions with Cellular Receptors. Structure **26,** 839–847.e3 (2018).
29. Plaschka, C., Lin, P.-C., Charenton, C. & Nagai, K. Prespliceosome structure provides insights into spliceosome assembly and regulation. Nature **559,** 419–422 (2018).
30. Conley, M. J. et al. Calicivirus VP2 forms a portal to mediate endosome escape. bioRxiv 397901 (2018). doi:10.1101/397901
31. Danev, R., Buijsse, B., Khoshouei, M., Plitzko, J. M. & Baumeister, W. Volta potential phase plate for in-focus phase contrast transmission electron microscopy. Proceedings of the National Academy of Sciences (2014). doi:10.1073/pnas.1418377111

Figure Legends

Figure 1. Flowchart describing the main steps of the procedure for F3EC (A) and energy filtered K2 Summit (B), with approximate timings. We expect this procedure to take 3-6 hours, although timings will be specimen dependent.

Figure 2. EPU setup. A) Typical Atlas view with thick (purple, appropriate thickness (blue) and dry (white) areas indicated (which will vary by sample). Scale bar 50 μm B) Square selection on an Atlas.

Each square should be inspected to ensure it is not broken (data collected on broken squares may contain significant motion) and ice thickness is appropriate for specimen. Grid squares which have been collected on are in blue, orange is collection in progress and green is areas to be collected. Scale bar 50 µm C) Grids with regular holes selection. Holes close to the grid square bars (which typically are poorly vitrified), are deselected. Scale bar 3.5 µm D) Lacy carbon with thin continuous film acquisition area selection. Note large contaminants, areas at the edge of the square and areas where carbon support:hole ratio is poor are deselected. Scale bar 10 µm E) Template with single shot per hole and whole hole illumination. Scale bar 1 µm F) Template with multiple shots her hole. In E and F, autofocus and drift measurement areas (purple) are overlaid. Scale bar 1 µm

Figure 3. On-the-fly data processing pipeline (Steps 42-51). Data is copied from its write on F3EC (A) and K2 (B) location to storage location. Symbolic links are then made to the processing directory, where RELION batch jobs are used to motion correct and perform CTF estimation. The outputs from this are plotted by micrograph analysis for the user to inspect.

Figure 4. Example output of the micrograph analysis script. A scatter plot of the two orthogonal defocus measurements. A) Provides a quick visual assessment of the range of defocus values in the dataset. Histograms in the left column describe the overall dataset estimated resolution (B), astigmatism (C) and phase shift (D). The right-hand column shows estimated resolution (E) and astigmatism (F) values for each micrograph in order as they were acquired, expressed as a percentage of the mean values for the entire dataset. Large changes in these values over time suggest a problem may have occurred during the data acquisition run. Phase shift for each microscope in order of acquisition (G) allows the tracking of the change in phase shift as the plate becomes charged and the microscope moves to new phase plate positions. The non-phase shift version of the script produces identical output, minus the bottom two panels.

Tables

Table 1 - Detector choices and electron doses for example samples.

Sample (kDa)	Buffer	Detector (pixel size)	Total dose ($e^-/Å^2$) (dose per frame)	Defocus range	Resolution (Å)	Reference
Virus- Cowpea mosaic virus (4080 KDa)	10 mM sodium phosphate, pH 7.0.	F3EC (integrating) (1.065)	67.5 (1.5)	-0.5 to -3.5	2.7	[12]
Cytochrome bc_1 membrane protein (480 kDa)	25 mM Tris pH 7.5, 100 mM NaCl, 0.5 mM, 0.015% DDM	K2 (Counting) (1.065)	44 (2.2)	-1 to -4	4.4	[11]
Cytochrome bc_1 membrane protein (480 kDa) with inhibitor bound	25 mM Tris pH 7.5, 100 mM NaCl, 0.5 mM EDTA and 0.015% DDM	F3EC (integrating) (1.065)	66 (1.13)	-1 to -4	3.7	Unpublished
Polyketide synthase module (230 kDa)	200 mM HEPES, 200 mM NaCl	F3EC (integrating) (1.065)	111 (1.4)	-0.75-3	3.7	Unpublished

E.coli ribosome (2.5 mDa)	10 mM HEPES-KOH pH 7.5, 50 mM KCl, 10 mM NH4Cl, 10 mM Mg(OAc)2, 1 mM DTT	F3EC (integrating) (1.065)	63 (1.4)	-0.8 -2.9	3.2	Unpublished
Virus – Ageratum Yellow Vein Virus (3110 KDa)	100mm sodium phosphate buffer pH7.0	F3EC (integrating) (1.065)	110 (1.4)	-0.5 to -2.5	3.3	[27]
Amyloid fibrils	25 mM sodium phosphate 25 mM sodium acetate pH 2.5	K2 (counting) (1.065)	49.9 (1.2)	-1.25 to -3	3.9	In press
Coxsackievirus A24v:ICAM1 complex	TBS	F3EC (integrating) (1.065)	60 (1.5)	-0.5 to -3	3.9	[14]
BK Polyomavirus + GT1b oligosaccharide	10 mM Tris, 50 mM NaCl, 0.01 mM CaCl2 (pH 7.8)	F3EC (jntegrating)(1.06 5)	50 (1.3)	-0.6 to -5	3.4	[28]
Bacterial nutrient transporter (320kDa	10 mM HEPES, 100 mM NaCl, 0.03 % (vol/vol) DDM, pH 7.5	K2 (counting) (1.065)	77.9 (1.6)	-1.2 -- 2.6	3.7	Unpublished
Prespliceosome	See reference	K2 (counting) (1.065)	63.52 (3.176)	-0.8- -4	4	[29]
Saccharomyces cerevisiae imidazole glycerol phosphate dehydratase	50 mM Tris, 30 mM NaCl (pH 8.0	F3EC (integrating) (1.065)	50 (1.3)	-0.6 to -5.0	3.2	[13]
Feline calicivirus	Phosphate Buffered Saline	F3EC (integrating (1.065)	63 (1.26)	-1.2 – 3.5	3	[30]

50S large ribosome subunit from S. aureus (~ 1.5 MDa)	10 mM Tris-HCl pH 7.51, 60 mM NH4Cl, 12 mM Mg(OAc)2, 200 mM NaCl	F3EC (integrating)(1.0 5)	46 (1.3))	-1.1 to – 2.9	2.9	Unpublished
Protein complex 400 KDa	20 mM MES (pH 6.5), 500 mM NaCl and 2 mM TECP	F3EC (integrating) (1.065)	83 (1.2)	-1.5 - - 3.5	4.1	Unpublished
Chaperone complex of RuvB-like AAA+ ATPase (390 kDa)	25 mM Hepes 140 mM NaCl, 10 mM 2-Mercaptoeth anol	K2 (counting) (1.065)	48 (1.2)	-1.5— 2.5	3.6	Unpublished
Beta Galactosidase (440 kDa)	Not disclosed	F3EC (counting)(0.66)	61 (0.8)	-0.7 – 2.3	2.82	Unpublished
Dynein/dynactin/HOOK3 complex (2.6 MDa)	25 mM HEPES pH 7.2; 150 mM KCl; 5 mM DTT; 1 mM ATP; 0.005% Tween-20	K2 (counting) (1.065)	48 (1.2)	-1.8--3	5.7	Unpublished

Boxes

Box 1 Beam setting presets

Beam setting presets in EPU are a convenient way of setting beam setting parameters to perform specific tasks and switching between the magnifications used during automated data collection. They are a set of parameters relating to the beam, including spot size and illuminated area. You can push these settings to the microscope by selecting it in EPU > Preparation > Acquisition and Optics Settings > Presets (Select option from drop down) > 'Set'. These will need to be altered depending on properties of the grid you are using. Here is a description for each;

- Atlas- This is usually optimised to be as fast as possible and does not need to be varied according to different grid types.
- Grid Square- This is adjusted according to mesh size. It should show one entire grid square within the field of view with no neighbouring squares.
- Hole/Eucentric- Where grids with a regular array of holes are being used, this magnification should contain a single, complete hole, and at most small fractions of neighbouring holes. Choose a magnification where it is not possible to take an image just of carbon (i.e. too high magnification) as this may lead to the software skipping holes. For lacey carbon grids this parameter is not required for automated data collection. Where possible, the spot size should be kept the same as that used for data acquisition, to minimise changes to lens settings during acquisition. As the specimen will be pre-exposed to the electron beam during hole imaging, the total dose should be kept to a minimum by reducing exposure time, and limiting the illuminated area, to prevent neighbouring holes being inadvertently exposed to the beam.

- Data acquisition- These settings depend on your desired settings for data acquisition, but usually a magnification resulting in 1-1.35 Å/pixel is chosen for most single particle projects (although some choose a smaller physical pixel size), with illuminated area/spot size chosen to deliver the desired amount of dose to the detector.
- Autofocus and Drift measurement- Where possible, focus and drift measurement presets should be kept the same as data acquisition to reduce the number of changes to lens power. In F3EC counting mode, these as focus/drift measurement calculations can be carried out in integrating mode for speed.

Example beam setting pre-sets using different detectors and modes can be found in Supplementary Tables 1, 2 and 3.

- END OF BOX 1 -

Box 2 Choosing your detector

The most appropriate detector choice, mode and parameters will be sample specific, and in reality, users choice may be limited by the hardware available at their facility. All of the direct electron detectors currently on the market, in all modes, are capable of producing sub 3 Å structures of a range of biological specimens. Most of the time, it is the sample, not the detector choice, that limits the final resolution. There are more comprehensive accounts of direct electron detector characterisation and performance [5,7], here we offer a quick guide to choosing a detector. This protocol is written for F3EC and K2 and so we focus on choosing between these models, while mentioning other detectors.

Integrating/counting

During integrating mode, a signal on the detector is summed, while in counting mode, each incident electron is individually detected with pixel (counting) or sub pixel (super resolution) accuracy. For data collection, the F3EC and Direct Electron DE-64 operates in integrating and counting modes, while the K2 operates in counting and super resolution modes. Counting mode detectors have better detective quantum yield (DQE) curves compared with integrating, especially at low spatial frequencies. This essentially means for the same number of electrons applied to the specimen, a counting mode image will have more contrast, meaning it is easier to see and align the particles. However, counting mode collection is typically 1.5-3 x slower compared with integrating mode collection due to longer exposure times.

Optimal doses for each detector in each mode.

In counting mode, performance of the detector is highly dose dependent, with lower doses resulting in better DQE curves. To avoid very long exposure times, which may bring their own problems such as mechanical movement of the stage, F3EC is optimally used with 0.5-0.7 e^-/pix/second, while the K2 3-10 e^-/pix/second in counting mode, and 1-3 e^-/pix/second in super resolution mode. To get a reasonable total signal in the final image (> 35 e^-/ $Å^2$) typically 60-90 or 8-13 second exposures respectively are needed. This results in the collection of ~ 25-30 (F3EC) and 45-60 (K2) micrographs per hour for each detector in counting mode.

In integrating mode dose rates can be much higher, and performance of the detector is linear across a greater range. We generally use the F3EC in integrating mode between 40-100 e^-/pix/second, with a 1-2 second exposure. Depending on other data collection parameters, this results in 70-150 images an hour.

General rules

- Choosing your detector and mode involves a data quantity/quality trade off.
- For > 400 kDa specimens with compact globular structures, where you are expecting better than 3.5 Å, such as ribosomes or icosahedral viruses, we recommend F3EC in integrating mode.

- For 100-400 kDa specimens, we recommend (if possible) taking images in different modes on different detectors to compare contrast and make an informed decision. Typically, these smaller specimens benefit from counting mode increase in contrast, but integrating mode combined with high dose can be used (Table 2) especially if your protein preparation is hererogeneous and data processing will benefit from the faster collection which will in turn result in more micrographs and subsequently more particles.
- If you are tilting the stage during data collection, you will benefit from using an energy filtered K2.
- Anything < 100 kDa probably requires a phase plate (see Box 3).
- In your final movie, aim for a dose of 1-2 $e^-/Å^2$/frame, the bigger the specimen/if you are pushing for sub 4 Å resolution, use a number closer to 1 $e^-/Å^2$/frame.

- END OF BOX 2 -

Box 3 Using a Volta potential phase plate

A Volta potential phase plate (VPP) is a thin amorphous carbon film positioned in the back focal plane of the objective lens [31]. The beam in a parallel state is brought to crossover at the VPP, where beam interactions with the carbon surface cause an 'on-the-fly' local Volta potential, creating a phase shift between the scattered and unscattered electrons, which increases continuously with accumulated dose on the VPP. This dramatically increases contrast of the specimen, which can enable the visualisation and alignment of smaller particles. Where a specimen has ordered density of ~< 100 kDa, use of a phase plate is usually the only way to create enough contrast to align and classify particles with sufficient accuracy to get a high-resolution structure. Useful phase shifts for single particle work are between 20-120°. The protocol presented here for can be easily modified to include use of the VPP (Supplementary Methods 3).

- END OF BOX 3 -

Supplementary information

Supplementary Figure 1- Micrograph analysis output associated with example data (EMPAIR-10205)
Supplementary Methods 1- Mirrored protocol for EPU set up with K2 summit direct electron detector.
Supplementary Methods 2- Direct alignments.
Supplementary Methods 3- Use of a Volta potential phase plate.
Supplementary Note 1- Description of scripts used in this protocol.
Supplementary Note 2- Modifications of RELION for on-the-fly processing
Supplementary Table 1- Example EPU beam setting pre-sets for F3EC integrating mode
Supplementary Table 2- Example EPU beam setting pre-sets for F3EC counting mode
Supplementary Table 3- Example EPU beam setting pre-sets for K2 counting mode
Supplementary Table 4- Data acquisition report
Supplementary video 1- Steps 1-6 of protocol, clipping grids for loading into a Thermo Fisher Scientific autoloader microscope.