HHS Public Access

Author manuscript

Nat Methods. Author manuscript; available in PMC 2020 April 14.

Published in final edited form as:

Nat Methods. 2019 November ; 16(11): 1161–1168. doi:10.1038/s41592-019-0591-8.

A complete data processing workflow for CryoET and subtomogram averaging

Muyuan Chen[1], **James M. Bell**[1,2], **Xiaodong Shi**[1,3], **Stella Y. Sun**[4], **Zhao Wang**[1,5], **Steven J. Ludtke**[1,5,*]

[1]Verna Marrs and McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas, USA

[2]Quantitative and Computational Biosciences Graduate Program, Baylor College of Medicine, Houston, Texas, USA

[3]Jiangsu Province Key Laboratory of Anesthesiology and Jiangsu Province Key Laboratory of Anesthesia and Analgesia Application, Xuzhou Medical University, Xuzhou, Jiangsu, China

[4]Department of Bioengineering, Stanford University, Stanford, California, USA

[5]CryoEM Core at Baylor College of Medicine, Houston, Texas 77030, USA.

Abstract

Electron cryotomography (CryoET) is currently the only method capable of visualizing cells in 3D at nanometer resolutions. While modern instruments produce massive amounts of tomography data containing extremely rich structural information, the data processing is very labor intensive and results are often limited by the skills of the personnel rather than the data. We present an integrated workflow that covers the entire tomography data processing pipeline, from automated tilt series alignment to subnanometer resolution subtomogram averaging. Resolution enhancement is made possible through the use of per-particle per-tilt CTF correction and alignment. The workflow greatly reduces human effort and increases throughput and is capable of determining protein structures at state-of-the-art resolutions for both purified macromolecules and cells.

Introduction

Electron cryomicroscopy (CryoEM) is rapidly becoming the standard tool for near atomic resolution structure determination of purified biomolecules over 50 kDa. However, for studies of molecules within cells where the structure may be strongly influenced by the cellular environment or for purified molecules with high levels of variability, electron cryotomography (CryoET) is the preferred method[1]. In these experiments, the specimen is

*Corresponding author. sludtke@bcm.edu.

Author contributions

M.C., J.M.B and S.J.L. designed and implemented the protocol. X.S., Z.W. and S.Y.S. provided test datasets. M.C., J.M.B. and S.Y.S. tested and refined the protocol. M.C., J.M.B. and S.J.L. wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

tilted within the microscope providing 3D information for each molecule and permitting overlapping densities, such as those found in the crowded cellular cytosol, to be computationally isolated.

While recent microscope and detector advances have greatly boosted the throughput of CryoET data collection, substantial human effort and computational resources are still required to process recorded imaging data. Especially in cellular tomography projects, data processing has become a major bottleneck in studying high-resolution protein structures.

To expedite CryoET data processing, we present a complete tomography workflow as part of the EMAN2 environment that performs all steps, from raw tilt series alignment through high resolution subtomogram averaging. While many of these tools are based on knowledge gained from decades of development by many groups[2–10], numerous innovations have been introduced to reduce the need for human intervention and improve the resolution of the final averaged structure. These include a fully automated tilt-series alignment method not requiring fiducials, rapid 3D reconstruction using direct Fourier methods with tiling, an optimization-based strategy for per-particle-per-tilt CTF correction, robust initial model generation, and per-particle-per-tilt orientation refinement (Figure 1a). In addition to algorithm development, this protocol also includes a user-friendly graphical interface and a specially designed book-keeping system for cellular tomography that allows users to study multiple features/objects within one cell, and to keep track of particle location to correlate structural findings with their location in the cellular environment.

Our integrated pipeline significantly increases the throughput of CryoET data processing and is capable of achieving the state-of-the-art subtomogram averaging results on both purified and *in situ* samples. We demonstrate subnanometer resolution from previously published in vitro datasets[11], and cellular tomography of whole *E. coli* over-expressing a double-layer spanning membrane protein at 14 Å resolution.

Results

Automated tilt series alignment and tomogram reconstruction

The first stage of the tomogram processing workflow is tilt-series alignment. Our method uses an iterative 3-D landmark-based approach with progressive downsampling (binning) and outlier elimination (Figure 1b). It works well on a wide range of tomograms with or without fiducials and without any human intervention.

The method begins with a coarse cross-correlation based alignment of a downsampled tilt series, and a rough estimate of the orientation of the tilt axis. The input tilt series are downsampled to 512×512 pixels irrespective of their original size or sampling. Based on the coarse alignment, an initial tomogram is generated, despite the likelihood of significant alignment errors, and 3D landmarks are selected from the resulting volume to use in the next stage of alignment. These landmarks are simply the N darkest voxels in the downsampled map, with a minimum distance constraint (Figure 2b). When fiducials are present in the data, they will tend to be selected as landmarks, as long as they are sufficiently well-separated, but

they are not explicitly identified as fiducials. Selecting landmarks as localized high contrast points in 3-D rather than identifying them in 2-D is critical to the success of this procedure.

The next step is iterative alignment. This includes two stages: refinement of landmark coordinates and optimization of the tilt image transforms. First, 3D coordinates of the selected landmarks are projected back to the tilt series, and corresponding 2D patches are extracted from the tilt images. The 2D patches corresponding to each landmark form a subtilt series that is then reconstructed into a local 3D volume of the landmark, to provide a more accurate center of mass. Next, 2D patches are re-extracted from the tilt images using the refined landmark positions, and the translational alignment that centers each landmark in each extracted 2D patch is calculated. A global optimization algorithm is used to adjust the 3D tilt transforms such that center of all landmarks in 2D patches match the projected coordinates of the landmarks to the greatest possible extent. With these improved alignment parameters, a new tomogram is generated with better alignment which is used during the next round of reprojection and alignment. To improve convergence and increase the speed of alignment, the process begins with highly downsampled images and gradually increases sampling as alignment error decreases, finishing with the unbinned tilt series in the final iteration. A specified fraction of the worst matching landmarks is normally excluded in each iteration, and this is critical to obtaining a self-consistent consensus alignment.

In the past, center of mass based alignment algorithms have suffered problems with identifying fiducials from the tilt series directly, where overlapping densities can produce many false positives. By identifying only well localized features in 3-D, not 2-D, and including aggressive outlier elimination to identify only a set of self-consistent landmarks, this algorithm achieves a remarkable level of success. In testing it has been able to achieve accurate automated tilt-series alignments on virtually every tomogram. The only exceptions have been tomograms with severe data collection issues, such as ½ frame alignment shifts in the middle of the tilt-series, etc.

In most non-cellular tomograms it is convenient for slice-wise visualization and annotation if the X-Y plane is parallel to the ice surface. It is assumed that on average the landmarks will be coplanar with the ice, and thus this plane is rotated to become parallel to the X-Y plane, using principal component analysis of the landmark coordinates (Figure 2e).

Tomogram reconstruction is performed using direct Fourier inversion rather than real-space methods such as filtered back projection[2] or SIRT[12]. Fourier methods have gradually become the standard in single particle reconstruction[13]. Although there have been a few descriptions of Fourier methods for tomography data processing[14,15], most tomography software still uses real space methods due to the size of tomographic volumes, concerns about edge effects, and image anisotropy [9,16]. We have adopted a Fourier reconstruction approach using overlapping cubic tiles, which significantly reduces edge effects and memory requirements, while still remaining computationally efficient. For convenience, the tile size is defined by the reconstruction thickness, such that each reconstructed tile is a cube. While the algorithm can perform reconstructions on non-cubic volumes correctly, interpolation anisotropy can cause artifacts which are avoided by using cubic volumes. The overlapping

tiles are individually reconstructed, then averaged using a radial Gaussian weight (Supplementary Figure 1).

Although the tilt series alignment is performed using the original full-sized images, the reconstructed tomograms are normally downsampled to provide sufficient resolution for visual inspection, annotation, and particle selection, while dramatically improving interactivity and decreasing system requirements. For subtomogram averaging, the downsampled particle locations are automatically rescaled, and the particle data is extracted from the tilt series at full sampling. The entire iterative alignment and reconstruction algorithm is quite fast, typically requiring only ~10 minutes (12-core workstation) for full-resolution alignment of a 60 image 4k × 4k tilt series with a 2k × 2k × 512 reconstruction (Supplementary Table 1). Since this is comparable to the time required for tilt series acquisition, it would be straightforward to include automated tomogram reconstruction as part of the data collection process in real-time.

As a real-world example, we reconstructed a cellular tomogram of *E. coli* over-expressing Tolc-AcrAB (Figure 2a, Supplementary Video 1)[17] with gold fiducials. The improved alignment after the iterative process can be observed by comparing the reconstructions of fiducials before and after the iterative process. Internal cellular features are also clearly visible in the reconstruction. In fiducialless reconstructions, the program usually chooses small pieces of ice contamination or other localized high-density 3-D objects as landmarks (Figure 2d). However, even for a fiducialless apoferritin data set (EMPIAR-10171)[18], which contained few obvious features to use as landmarks, the program produced high quality reconstructions where individual proteins were clearly visible (Figure 2c, Supplementary Video 2).

Multiple methods for particle localization

Earlier versions of EMAN2 included a graphical program for manually selecting 3D particles using orthogonal slices[19]. This particle picking interface has now been reworked, enabling users to simultaneously select and visualize particles of multiple types and different sizes within each tomogram (Figure 3a,c). Each type is then extracted into a separate stack of 3D particles and accompanying 2D subtilt series, with the original location metadata retained for later per-particle processing.

In addition to the manual 3D picking interface, two semi-automatic tools are provided for annotation and selection. For purified macromolecules imaged by tomography, a template matching algorithm can be used to rapidly locate particles. For more complex tomograms, our convolutional-neural-network-based tomogram annotation tool can be used to identify features[20], followed by a second stage which converts annotations into subtomogram coordinates. For globular particles like ribosomes, the program locates and centers isolated annotations. For continuous structures like microtubules and protein arrays on membranes, the program randomly samples coordinates within the set of annotated voxels, with a specified minimum distance between boxes. The parameters of these semi-automatic tasks can then be tuned by visualizing results in the manual particle picking tool.

Per-particle-per-tilt CTF correction

Accurate CTF measurement and correction is critical for obtaining high-resolution structures through subtomogram averaging. The most commonly used method in tomographic CTF correction is the simple tiled CTF correction of rectangular strips within each tilt series[21]. This method is effective in getting past the first CTF zero-crossing when working with thin layers of purified macromolecules; however, when working with cellular data or other thicker specimens, the error in defocus due to the Z position of the particle within the ice becomes significant and requires more accurate correction on a per-particle per-tilt basis.

In our CTF estimation strategy, the entire tilt image is used to determine its central defocus, by splitting the image into tiles and summing the information from the entire image to estimate the defocus. Instead of fitting one defocus value for each strip, we calculate the relative defocus difference from the center of each tile to the center of image based on the geometry of the tilt, and search for one defocus value for the center of image that optimizes the fitting of theoretical CTF curves in all tiles. With this approach, the information in the full tilt image is used to estimate one scalar defocus value, and achieve more robust defocus estimation under low SNR conditions.

At high tilt, the SNR in an individual image is typically so low, and the thickness of each tile so large, that even using all information in the image is not sufficient to provide an unambiguous defocus estimation without a starting estimate. Thus, for the higher tilts, we limit the defocus search to three standard deviations around the mean defocus of the low tilt images. With this additional constraint, sufficiently accurate defocus values can still be determined at high tilt.

After CTF determination, fully sampled CTF corrected subtomograms are generated directly from the raw tilt series. Since we have the alignment parameters for each micrograph in the tilt series and the coordinates of particles in the tomogram, we can extract per-particle tilt series, which we henceforth refer to as a “set of subtilts”, from 2D micrographs. The center of each subtilt is determined by projecting the 3D coordinates of the particle using the transform of the micrograph calculated from tilt series alignment, so each subtilt series can be reconstructed to an unbinned 3D particle using the corresponding tilt image transforms. From these defocus values at the center of each tilt, the defocus of each tilt for each particle can be determined from the 3D location of the particle and the tilt-series geometry (Supplementary Figure 2). After subtilt images are extracted from the tilt series, we flip the phase of each subtilt according to its determined defocus before reconstructing the subtilt into CTF corrected 3D subtomograms.

Initial model generation via stochastic gradient descent

In many cellular tomography projects, the identities of extracted particles are unknown before subtomogram averaging. While it is possible to use catalogs of potential candidate structures and exhaustively compare particles to each of these for purposes of identification[22], there are many shortcomings to this approach, including the need for a complete catalog, the problem of model bias, and the difficulty of handling complexes. An unbiased approach would be to classify particles de-novo and generate independent initial

models for each class from the raw particles. Our previous subtomogram averaging methods offered several different strategies for handling this issue[19], as the failure rate was substantial, and user guidance was always required. We have found that similar problems plague other subtomogram averaging software. We have now developed a stochastic gradient descent (SGD) based initial model generation protocol[23], which produces reliable initial models even from cell-derived particles.

SGD is an optimization technique widely used in training for machine learning, offering advantages in both speed and avoidance of local minima. We begin with an effectively randomized map, produced by averaging a random subset of particles in random orientations, lowpass filtered to 100 Å. In each iteration, a batch of randomly selected particles are aligned to the reference map, and a new map is generated. This new map is used to update the reference using an adjustable learning rate. To avoid overfitting, the reference is filtered to a user-specified resolution (usually 30–50 Å) after each update. The alignment, average and map update steps are repeated until the reference map converges to a consistent initial model. As only a low-resolution initial model is needed, it is not critical that all particles be used. The program can typically produce good initial models within 1 hour on a typical workstation (Supplementary Table 1).

This method has performed well in testing on more than 20 structures with very distinct shapes from a variety of sources. This includes globular structures like ribosomes, linear structures such as microtubules, and even double-membrane spanning proteins (Figure 3b,d). While it is impossible to guarantee that any algorithm will be universally successful, we have found this method to work in the vast majority of cases.

Subtomogram alignment and averaging

There are two stages in producing a final high-resolution subtomogram average: traditional subtomogram alignment and averaging[5,19] and per-particle-per-tilt refinement (Figure 1c). The initial stage makes use of our existing subtomogram alignment and averaging algorithms which automatically detect and compensate for the missing wedge[6]. The alignment algorithm employs an extremely efficient hierarchical method, which scales well with particle dimensions. The overall refinement process follows “gold-standard” procedures similar to single particle analysis[24], in which even and odd numbered particles are processed completely independently with unique, phase-randomized starting models, with a Fourier shell correlation (FSC) used to filter the even and odd maps, assess resolution, and measure iteration-to-iteration convergence.

In the second stage, rather than working with subtomograms, we work instead with subtilt series. When full frame tilt series are aligned, we assume that each tilt is a projection of a single rigid body volume. With beam-induced motion, charging and radiation damage affects the assumption that the specimen remains globally rigid across a 1-μm span with the largest acceptable motion <10 Å is an extremely stringent requirement. Local deviations are common and can produce significant misalignments of individual objects in individual tilts. To compensate for this resolution-limiting effect, we have developed a strategy for refinement on a per-particle-per-tilt basis, where the alignment and quality assessment of each tilt of each particle are individually refined. Effectively this is a hybridization of

subtomogram averaging approaches with traditional single particle analysis. Some of these techniques are similar to those recently implemented in EMClarity[8].

Our subtilt refinement procedure starts from an existing 3D subtomogram refinement, preferably with a resolution of 25 Å or better. Subtilt series for each particle were already extracted as part of the CTF correction process above. The iterative refinement process is a straightforward orientation optimization for each tilt image of each particle. All 5 orientation parameters are refined independently per-particle-per-tilt. It is quite common for some images in a tilt series to be bad, either due to excessive motion or charging. To compensate for this, the quality of each tilt for each particle is assessed, and weighted correspondingly, with the very worst excluded entirely. All of the realigned particles are used to compute a new weighted average 3D map, which is then used for the next iteration of the refinement.

The subtilt refinement protocol significantly improves map quality and resolution for purified samples in thin ice, where relatively little density is present above and below each particle. In the EMPIAR-10064 dataset (purified ribosomes)[11], without subtilt refinement, subtomogram averaging achieved 13 Å “gold-standard” resolution (FSC>0.143) using 3000 particles from 4 tomograms. With subtilt refinement, the resolution improved dramatically, to 8.5 Å (Figure 4a–d, Supplementary Video 3). In the averaged map, the pitch of RNA helices is clearly visible and long alpha-helices are separated. Similarly, in the case of EMPIAR-10045[7], our standard subtomogram averaging algorithm resolved the 80S ribosome structure to 16Å, and subtilt refinement extended this to 9.3Å. A more detailed comparison between our results and those from other software packages is included in online Methods.

We did not initially expect subtilt refinement to work well in a cellular context, due to the presence of so much confounding cellular mass present in each subtilt image. Surprisingly, we found that an *in situ* dataset of the double-membrane spanning TolC-AcrAB complex in *E. coli*, reached 19Å in initial averaging, which improved to 14 Å resolution after subtilt alignment[17,25] (Figure 4e, Supplementary Video 4). We do not yet have sufficient test cases to set expectations for how well subtilt refinement will work in any given cellular system, but based on our preliminary studies, it may provide a significant improvement in a wide range of experimental situations.

Discussion

The entire protocol outlined above has been integrated into the graphical workflow in EMAN2.3 (e2projectmanager.py). This presents the process as a sequence of steps (Figure 1), and an online tutorial can be found at http://eman2.org/Tutorials. Graphical tools are also provided for evaluating tomogram reconstructions and subtomogram refinements, which are useful for managing projects involving a large amount of data. Unlike single particle analysis where it is possible to transition data from other tools into EMAN2 at virtually any stage of processing, the stringent requirements for all of the metadata generated at each stage of processing make it challenging to, for example, import a reconstructed tomogram from other software, then proceed. While some tools will be usable on imported data, such as the Deep Learning based annotation[20] and simple subtomogram alignment and averaging, the

new approaches, such as subtilt refinement, are simply not possible unless the complete EMAN2 pipeline is followed.

With per particle CTF correction and subtilt refinement, it is now relatively straightforward to achieve ~10Å resolution using 1000 – 2000 particles from a few good tilt series. This method can also be used with phase-plate data, though the difficulty of collecting Volta phase plate tilt series and determining per-tilt CTF parameters with continuously varying phase shift is significant. While we do optimize both the defocus and phase shift, particularly at high tilt, there is insufficient information available for simultaneous determination of both parameters. Our suggested approach is to target 0.5 – 1 μm underfocus with such tilt series, to put the first zero in a range where correcting beyond the second zero is not necessary to achieve slightly better than 10Å resolution. In this way, locating the first zero accurately is sufficient for subnanometer resolution.

One difficulty in subtomogram averaging *in situ* is masking and filtration of the averaged map after each iteration of refinement. In the cellular environment, proteins of interest are often surrounded by other strong densities and masking can have a strong impact on the final achieved resolution. To address this issue, we introduce the option of masking the averaged map with a large soft mask and filter it using the local resolution determined from even and odd sub-maps. This allows us to keep high-resolution information of the protein of interest for the next round of refinement and reduces misalignment caused by other densities surrounding the protein.

The algorithmic improvements we have discussed make it possible to perform data-driven cellular-structural biology research with CryoET. Researchers can take tomograms of cells or purified organelles, manually select a few features of unknown identity, and automatically annotate similar features in the whole dataset. Reliable, de novo initial models of the features of interest can be generated from raw particles without prior knowledge of the proteins. With per particle CTF correction and subtilt refinement, averaged maps at 10–15 Å resolutions can be achieved in a matter of days (Supplementary Table 1) with a few thousand subtomogram particles, so one can make reasonable hypotheses of the identity and composition of the proteins based solely on their structural features, and validate these hypotheses with biochemical experiments. Furthermore, the position and orientation of each protein particle can be mapped back to the tomogram to study the organization of proteins in cells (Figure 4f).

Online Methods

Tomogram reconstruction

To seed the iterative tilt-series alignment, a coarse alignment is first performed. First, the unaligned tilt series is downsampled to 512 × 512 pixels, subject to a real-space ramp filter, Fourier bandpass filter, and normalized to mean value of zero and standard deviation of one. A coarse alignment is then performed under a circular mask with soft Gaussian falloff. The alignment begins with the center tilt image (typically near untitled) and propagates sequentially in both directions. After the coarse translational alignment, the approximate tilt axis is identified by computing a coherent sum of the tilt series in Fourier space. The tilts

will approximately share one Fourier common-line without any rotational alignment. This can be readily identified by simply performing a coherent sum after translational alignment and looking for the axis with maximum intensity (Supplementary Figure 3). Only angles 0–180 degrees are permitted in this process to ensure no handedness flips occur between different tomograms within the same set. Although the handedness is thus consistent throughout the dataset, it is not necessarily correct, due to the 180-degree ambiguity in the tilt axis direction. If the correct orientation of the tilt axis in the images has already been determined for the microscope, it can be specified as a fixed parameter instead of performing a search.

Finally, the tilt series is reconstructed to produce the preliminary tomogram using EMAN2's normal direct Fourier inversion reconstruction algorithm[1]. The 512 × 512 box size is small enough that direct Fourier inversion can be used for this purpose without tiling. Since higher tilt images include information outside the frame of the zero tilt image, a soft mask is applied at the edges of each image parallel to the tilt axis as appropriate for each tilt just before reconstruction. The exact amount of required masking depends on the thickness of the specimen to achieve optimal results, but this can be ignored for purposes of tilt series alignment.

After the initial tomogram reconstruction, an iterative alignment-reconstruction process is performed beginning with 512 × 512 images gradually reducing downsampling until the fully sampled images are being used (typically 4k × 4k). Each iteration begins with landmark selection in the tomogram from the previous iteration, followed by multiple rounds of landmark location refinement and tilt parameter refinement as described above, and ends with the final downsampled tomogram reconstruction along with the optimized alignment parameters. By default, we perform 2 iterations at 512 × 512, and 1 iteration at 1024 × 1024, 2048 × 2048 and 4096 × 4096. When the input tilt series is larger than 4096 × 4096, such as DE-64 or K2 super-resolution images, we only perform alignments from 512×512 to 4096 × 4096. It is worth noting that in all iterations, reconstruction of the full tomogram is always done using the pre-filtered 512 × 512 tilt series. These tomograms are used only for selection of landmarks, whose locations are later refined in subtomograms using the appropriate downsampling.

To select landmarks, the 512 × 512 × 256 tomogram is further binned by 4 by taking the minimal value of each 4 × 4 × 4 cube and the result is highpass filtered to remove the impact of ice gradient in the tomogram. In this stage of processing, it is important to note that higher densities have lower values in raw tomograms, which is opposite from the normal EMAN2 convention. Voxel values in the tomogram are sorted and the program picks voxels separated by a minimal distance as landmarks. By default, 20 landmarks are selected with a minimum spacing between landmarks of 1/8 of the longest axis of the tomogram.

Multiple rounds of landmark location refinement and tilt parameter refinement are performed after landmark selection. In each round, we refine the 3D location of landmarks and one of the alignment parameters, including translation, tilt axis rotation, tilt angle and off-axis tilt. Because uncertainties vary for the determination of each parameter, we begin

with refinements of tilt image translation and global tilt axis rotation, then refine on and off-axis tilt angles.

In landmark location refinement, we first extract subtilt series of the landmarks from the tilt series and reconstruct the landmarks at the current level of binning. By default, we use box size of 32 for bin-by-8 and bin-by-4 tilt series, 1.5x box size for bin-by-2 and 2x box size for unbinned iterations. We locate the center of landmarks by the coordinate of the voxel with minimal value for bin-by-8 and bin-by-4 iterations and by the center of mass for bin-by-2 and unbinned iterations. This use of center-of-mass rather than aligning features within each landmark region might seem that it could reduce alignment accuracy. However, a common problem with tomographic alignments is that it is possible to have self-consistent alignments with an incorrect translation orthogonal to the tilt axis, producing distorted features in reconstructions when viewed along the tilt axis. Using of center-of-mass for alignment seems to largely avoid this problem, particularly when combined with exclusion of landmarks which are outliers in the alignment process.

To refine the alignment parameters, we first project landmark coordinates to each tilt using currently determined alignment, and extract 2D particles of the same box size at current binning. The center of each 2D particle is determined in the same way that 3D landmarks are centered, and the distance from the center of the 2D particle to the projection of 3D coordinates is computed. For each tilt, a local minimization routine (Powell optimizer from Scipy[2,3]) is used to refine alignment parameters and minimize the averaged distance from all landmarks. By default, 10% of landmarks with the highest average alignment distance in each tilt are ignored during the optimization. The averaged error per tilt is also used in the following round of landmark location refinement and tomogram reconstruction where 10% of tilt images with highest error are excluded.

After all the refinement iterations are finished, the final tomogram is reconstructed. When reconstructing the tomogram by tiling, we use a tile length of 1/4 the tomogram length and pad the 3D cube by an extra 40% during reconstruction. The step size between the tiles is 1/8 tomogram length, and overlapping tiles are shifted by half tile in x and y. 2-D tiles are subjected to an edge decay mask along the x-axis similar to the mask used in the full tomogram reconstruction. After reconstruction of each tile, a mask with Gaussian falloff is applied to subvolumes before they are inserted into the final reconstruction. The mask is described by

$$f = 1 + e^{-10\left(x^2 + y^2\right)} - e^{-10\left(\left(|x| - 0.5\right)^2 + \left(|y| - 0.5\right)^2\right)},$$

where x, y are the coordinate of the voxel from the center of tile, ranging from −1 to 1. This specific shape of mask is used so the summed weight in each voxel in the tomogram is 1, and the soft Gaussian falloff reduces the edge artifacts from the reconstruction of each tile. A Gaussian intepolation on a 2×2×2 voxel grid is used for tomogram reconstruction. After reconstruction, the tiles are clipped and added to the final volume to produce the final tomogram. This entire process requires on the order of 10 minutes per tomogram (Supplementary Table 1).

CTF correction

To determine the defocus for an image in a tilt series, overlapping 256×256 tiles are sampled on the raw micrograph and power spectrum of each tile is computed. Power spectrum curves from tiles with the same distance to the tilt axis are coherently averaged to increase the SNR. Using a global search, we find the defocus value d that maximizes $\sum_i S_i(p_i, d + x_i \sin(\theta))$, where x_i is the x position of the ith tile (y is the tilt axis), θ is the tilt angle, and $S_i(p,\Delta z)$ is the score function represented by the normalized dot product between a theoretical CTF curve with defocus Δz and the coherent, background subtracted power spectrum, p, of the ith strip of tiles parallel to the tilt axis.

Basic operations, including generation of theoritical CTF curves, computing power spetrum from image tiles and background substraction are implemented using the same strategy as in the single particle analysis protocol from the EMAN2 package.

Initial model generation for subtomogram averaging

Comparing to the classical gradient descent algorithm that calculates the gradient using the full dataset at every iteration, SGD breaks the training set into random small batches, calculates the gradient from each batch and update the model incrementally. The fluctuation introduced from the small batches makes it easier to get pass local minimums and achieves better convergence for high dimensional, non-linear functions.

In the SGD based initial model generation process, we use a very small batch size (12 particles per batch by default) and a learning rate of 0.1 to introduce enough fluctuations into the system. The list of input particles is shuffled before grouping into batches. Particles may be optionally downsampled and lowpass filtered before alignment. Particles in the first batch are averaged in random orientations to produce a map which is then filtered to 100Å and used as the initial alignment reference, which will have roughly the correct radial density profile, but meaningless azimuthal information. In each subsequent batch, particles are aligned to the reference and an average is generated. Any empty region remaining in Fourier space is filled with information from corresponding Fourier regions in the current reference. We calculate the per voxel difference between the current reference and the new averaged map and update the reference toward the average by the learning rate. The program goes through only 10 batches in each iteration by default, and does not necessarily use all the given particles before converging to a good initial model.

When symmetry is specified, the map needs to be aligned to the symmetry axis before symmetry can be imposed. Because of the way that the initial reference is generated, at the beginning of the initial model generation process, the reference is often radially symmetric. Applying the symmtry too early in the process will trap the refinement in a local minimum and make it more difficult to get to the correct structure. So the program always starts the initial model generation with C1 symmetry even when a higher symmetry is specified. After the first 20 batches, the program searches for the symmetry axis of the reference and alignes the reference to that axis. The symmetry search is performed in a similar way as a 3D alignment, except at each orientation tested, instead of computing the similarity between the rotated map and a reference, the similarity between the rotated map and its symmetrized

version is used as the objective function for the alignment. The symmetrized map is used as the reference for later batches and the orientation of symmetry axis is refined after every 10 batches.

Subtilt refinement

The first step of subtilt refinement is to compute the orientation of each subtilt using the orientation of the subtomogram and the alignment of tilt images in the tomogram. The refinement starts from 32 randomly distributed orientations centered around the previous orientation. One of the initial positions is always the previously determined orientation so the worst-case answer is no change. From these positions, an iterative search is performed starting from Fourier box size 64 to full box size, similar to the subtomogram refinement. During the refinement, the reference map is projected using Fourier space slicing with Gaussian interpolation. The comparison between the projection and the 2D particle is scored with CTF weighted Fourier ring correlation adapted from the single particle analysis protocol in EMAN2[1].

We refine even/odd particle sets independently in the subtilt refinement. By default, the program uses all tilt images and removes the 50% of particle images with the worst score. Generlally the low-scoring images correlate with tilt angle, such that higher tilts are most often excluded. There is also an option provided to explicitly exclude high angle tilt images. We also remove subtilt particles with scores beyond 2-sigma around the mean, because practically, particles with very high alignment scores often contain high contrast objects such as gold fiducials, and low score particles are often at the edge of the micrograph and has little signal. Before inserting the images to the 3D Fourier volume, we normalize their scores to (0,1) and weight the particles by their scores when reconstructing the 3D average. The 3D volume is padded by 2 to avoid edge artifacts, and reconstruction is performed with Gaussian interpolation with variable width with respect of Fourier radii. The averaged map is filtered by the “gold-standard” FSC.

Processing example data sets

We processed the 4 “mixedCTEM” tilt series from the EMPIAR-10064 purified ribosome dataset. The tomograms were reconstructed from the tilt series automatically, using default parameters. Defocus values were calculated using default options and the resulting defocus values range from 2.4 to 3.7 μm. CTF-corrected subtomograms were generated with a box size of 180. An initial model was produced using all particles as input, with 3x downsampling and a target resolution of 50Å. 3239 particles were selected via template matching using the initial model as a template, followed by manual bad-particle removal. Next, 4 rounds of subtomogram refinement and 3 rounds of subtilt refinement were performed to arrive at the final map, which was sharpened using a 1-D structure factor calculated from EMD-5592, masked via EMAN2 auto-masking, and filtered by the local gold-standard FSC.

Tomograms of the AcrAB-TolC pump in E. coli cells were collected on a JEOL3200 equipped with a Gatan K2 camera. Tomogram reconstruction and CTF determination were performed in EMAN2 using default parameters. The unbinned particle data had an Å/pix of

3.365, and a box size of 140 was used during particle extraction. Channel-shaped densities crossing both inner and outer memebranes that are not present in the control cells are identified as particles of AcrAB-TolC pumps. 25 high SNR particles were used for initial model generation. For structures with symmetry, applying the symmetry before the initial model generation converges tends to trap the SGD in a local minimum and not achieve the optimal result. So here a two-step approach was used to build the initial model. First 5 iterations of our SGD routine were performed imposing C1 symmetry. After aligning the result to the symmetry axis, we performed 5 more iterations with C3 symmetry. Subtomogram averaging was then performed using 1321 particles from 9 tomograms while applying c3 symmetry. To focus on the protein while preserving information from the membrane for improved alignment, a mask with values ranging from 0.5–1 around the pump and 0–0.5 covering a larger cylinder was applied to the map each iteration before alignment. The final map was filtered by local FSC and sharpened using a 1-D structure factor obtained from a high-resolution single-particle structure of the purified AcrAB-TolC complex.

Comparison with other software packages

Although it is difficult to find a good metric to compare the accuracy of an individual tilt series alignment from different software packages, when processing datasets of hundreds of tilt series using automated protocols, it is easy to spot “failed” cases of alignment from the successful cases visually. Very often there are some artifacts present in those “failed” cases (large area of ice contaimination, significant drift between tilt images, etc.), making the automated alignment protocol unable to align the tilt series at all. In many of those datasets, with some manual tracking, it is still possible to align the tilt series properly and retrieve meaningful information from the reconstructed tomograms. In our automated pipeline, our goal was to minimize the occurance of those problems in the tilt series alignment. To compare our automated approach with existing tilt series alignment pipelines, we obtained a sampling of ‘failed’ tilt series from three different research groups, and found that our new algorithm successfully aligned all of these tilt series, with the exception of a small handful with egregious data collection problems (useful tomograms cannot be reconstructed with even manual alignment). One example comes from the ETDB-Caltech database[4], which contains more than 11,000 publicly available tilt series. For the tilt series which were aligned successfully with IMOD, our alignment algorithm produces similar results. We also tested an assortment of tilt series where IMOD had failed to produce good results, and our algorithm also succeeded on these (Supplementary Figure 4).

Since they do not share a common set of markers, we have found no good way to make quantitative comparisons of alignment quality between different algorithms beyond global success/failure. While having an alignment quality metric would be desirable, it is a somewhat academic exercise, as in the end, the global alignment only needs to be good enough to permit 3-D localization of particles. The final alignments are redetermined on a per-particle basis as part of subtilt refinement, so even if present, slight inaccuracies in the initial alignments would not limit final resolutions.

To assess the quality of tilted Fourier reconstruction method, we compare the reconstruction results with other common techniques, including back projection (BP), SIRT and direct

Fourier transform without tiling. Here the alignment and reconstruction using BP and SIRT are performed with IMOD, while Fourier reconstruction with and without tiling are done in EMAN2. While the differences between different reconstruction methods are generally subtle, some visible differences can be observed at the edges of tomogram, both in real and Fourier space. In Fourier space, real space reconstruction methods like BP and SIRT create reflection artifacts at the edge of the Fourier tomogram, unless the tilt images are lowpass filtered before reconstruction. In real space, the reconstruction quality tends to decrease at the boundary of the tomogram or near high contrast objects. To better visualize the differences, we zoom in a small region containing cellular ribosome at the edge of tomogram next to a piece of carbon edge to compare the results (Supplementary Figure 5). Visually, the tiled Fourier method produces better features than the Fourier method without tiling and BP, similar to the result from SIRT. It also does not have the low resolution artifacts along the x axis that appear in the SIRT reconstruction.

To evaluate our subtomogram averaging results, we compare the structure we obtained with the EMPIAR datasets with published structures using other software on the same datasets. In both EMPIAR 10064 and 10045 datasets, we achieved structures with higher measured resolution and better real space features (Supplementary Figure 6, 7) than the original publication that produced the datasets. For EMPIAR-10064, the resolution reported in original publication[5] is 11.2 Å using PyTOM, and EMAN2 achieved 8.5Å; in EMPIAR-10045, the resolution reported by Relion[6] is 13 Å and the measured resolution of the EMAN2 structure is 9.3 Å. A more recent software package, EMClarity[7] also uses the two EMPIAR dataset as benchmarks, and reported 8.6 Å resolution for EMPIAR-10064 and 7.8 Å for EMPIAR-10045. Since the EMClarity result for EMPIAR-10064 is not publicly available, we can only make comparison with the EMPIAR-10045 dataset (Supplementary Figure 7). Although the front view of EMClarity structure on EMPIAR-10045 shows high-resolution features, significant anisotropic artifacts are visible from other orientations, showing the real space features to be clearly worse than the EMAN2 result. Accroding to the EMClarity authors, unpublished software improvements have largely resolved this anisotropy, but we have been unable to reproduce these results ourselves using EMClarity. As all of these software packages undergo continuous changes and improvements, any comparison among them has limited use beyond a narrow window of time.

It is also worth noting that in both cases, unlike in other software packages, tomogram reconstruction and most other processes in EMAN2 are performed automatically. That is, high resolution subtomogram averaging results can be achieved in EMAN2 with little manual effort in the entire process.

Finally, we compare the EMAN2 tomography pipeline with the currently commonly used IMOD/PEET pipeline using the EMPIAR-10064 dataset. We performed this test ourselves, and since the process has many manual steps, it is possible our results are typical but not optimal. The tilt series were aligned using automated fiducial tracking, and CTF corrected using the strip-wise CTF correction from IMOD. Tomograms were reconstructed using back projection, and ribosome particles were selected using template matching in EMAN2. The subtomogram refinement started from the same initial model used in and generated by EMAN2, using particles from bin4 tomograms. Three rounds of alignment search were

performed on the bin4 tomograms using PEET: a full rotational search with 30 degrees step size, followed by two rounds of local search with range of 30 and 10 degrees and step size 10 and 3 degrees. We then switched to the bin2 tomograms for two more rounds of refined rotation search with search range 3 and 1 degrees, starting from orientation determined from the bin4 tomograms. Finally, we perform two rounds of rotation search with searh range 3 and 1 degrees and step size 1 and 0.5 degrees on the unbinned tomogram. In the end, we are able to obtain a subtomogram average with ~13 Å resolution, close to the original result from pyTOM, again as compared to 8.5 Å resolution with clearly improved features in EMAN2. This comparison suggests that the EMAN2 pipeline can achieve better result from the same data and also requires much less human effort.

Visualization

Rendering of 2D images are performed in EMAN2, and rendering of 3D density maps is performed with UCSF Chimera and ChimeraX[8,9].

Data availability

The subtomogram averages are depositted in the EMDatabank: EMD-0529 and EMD-0530.

Code availability

EMAN2.3 is a free and open source software available from http://eman2.org with source code on GitHub (https://github.com/cryoem/eman2).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was partially supported by NIH grants R01GM080139, P01GM121203, Welch Foundation (Q-1967-20180324), BCM BMB department seed funds, and a Houston Area Molecular Biophysics Program (HAMBP) training grant from the Keck Center of the Gulf Coast Consortium (GCC, T32 GM008280-30). We also would like to thank early users for testing the workflow and providing valuable feedbacks.

References

1. Asano S, Engel BD & Baumeister W In Situ Cryo-Electron Tomography: A Post-Reductionist Approach to Structural Biology. J. Mol. Biol 428, 332–343 (2016). [PubMed: 26456135]
2. Kremer JR, Mastronarde DN & McIntosh JR Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol 116, 71–76 (1996). [PubMed: 8742726]
3. Frangakis AS et al. Identification of macromolecular complexes in cryoelectron tomograms of phantom cells. Proc. Natl. Acad. Sci 99, 14153–14158 (2002). [PubMed: 12391313]
4. Amat F et al. Markov random field based automatic image alignment for electron tomography. J. Struct. Biol 161, 260–75 (2008). [PubMed: 17855124]
5. Hrabe T et al. PyTom: A python-based toolbox for localization of macromolecules in cryo-electron tomograms and subtomogram analysis. J. Struct. Biol 178, 177–188 (2012). [PubMed: 22193517]
6. Galaz-Montoya JG et al. Alignment algorithms and per-particle CTF correction for single particle cryo-electron tomography. J. Struct. Biol 194, 383–394 (2016). [PubMed: 27016284]

7. Bharat TAM & Scheres SHW Resolving macromolecular structures from electron cryo-tomography data using subtomogram averaging in RELION. Nat. Protoc 11, 2054–2065 (2016). [PubMed: 27685097]
8. Himes BA & Zhang P emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging. Nat. Methods 15, 955–961 (2018). [PubMed: 30349041]
9. Mastronarde DN & Held SR Automated tilt series alignment and tomographic reconstruction in IMOD. J. Struct. Biol 197, 102–113 (2017). [PubMed: 27444392]
10. Noble AJ & Stagg SM Automated batch fiducial-less tilt-series alignment in Appion using Protomo. J. Struct. Biol 192, 270–278 (2015). [PubMed: 26455557]
11. Khoshouei M, Pfeffer S, Baumeister W, Förster F & Danev R Subtomogram analysis using the Volta phase plate. J. Struct. Biol 197, 94–101 (2017). [PubMed: 27235783]
12. Gilbert P Iterative methods for the three-dimensional reconstruction of an object from projections. J. Theor. Biol 36, 105–117 (1972). [PubMed: 5070894]
13. Penczek PA Image restoration in cryo-electron microscopy. Methods Enzymol. 482, 35–72 (2010). [PubMed: 20888957]
14. Taylor KA, Tang J, Cheng Y & Winkler H The Use of Electron Tomography for Structural Analysis of Disordered Protein Arrays. J. Struct. Biol 120, 372–386 (1997). [PubMed: 9441940]
15. Bernard Heymann J, Cardone G, Winkler DC & Steven AC Computational resources for cryo-electron tomography in Bsoft. J. Struct. Biol 161, 232–242 (2008). [PubMed: 17869539]
16. Chen Y et al. FIRT: Filtered iterative reconstruction technique with information restoration. J. Struct. Biol 195, 49–61 (2016). [PubMed: 27134004]
17. Du D et al. Structure of the AcrAB-TolC multidrug efflux pump. Nature 509, 512–515 (2014). [PubMed: 24747401]
18. Noble AJ et al. Reducing effects of particle adsorption to the air-water interface in cryo-EM. Nat. Methods 15, 793–795 (2018). [PubMed: 30250056]
19. Galaz-Montoya JG, Flanagan J, Schmid MF & Ludtke SJ Single particle tomography in EMAN2. J. Struct. Biol 190, 279–290 (2015). [PubMed: 25956334]
20. Chen M et al. Convolutional neural networks for automated annotation of cellular cryo-electron tomograms. Nat. Methods 14, 983–985 (2017). [PubMed: 28846087]
21. Xiong Q, Morphew MK, Schwartz CL, Hoenger AH & Mastronarde DN CTF determination and correction for low dose tomographic tilt series. J. Struct. Biol 168, 378–87 (2009). [PubMed: 19732834]
22. Bohm J et al. Toward detecting and identifying macromolecules in a cellular context: Template matching applied to electron tomograms. Proc. Natl. Acad. Sci 97, 14245–14250 (2000). [PubMed: 11087814]
23. Punjani A, Rubinstein JL, Fleet DJ & Brubaker MA cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296 (2017). [PubMed: 28165473]
24. Henderson R et al. Outcome of the first electron microscopy validation task force meeting. Structure 20, 205–214 (2012). [PubMed: 22325770]
25. Shi X et al. In situ structure and assembly of the multidrug efflux pump AcrAB-TolC. Nat. Commun 10, 2635 (2019). [PubMed: 31201302]

References

1. Tang G et al. EMAN2: An extensible image processing suite for electron microscopy. J. Struct. Biol 157, 38–46 (2007). [PubMed: 16859925]
2. Powell MJD An efficient method for finding the minimum of a function of several variables without calculating derivatives. Comput. J 7, 155–162 (1964).
3. Oliphant TE Python for Scientific Computing. Comput. Sci. Eng 9, 10–20 (2007).
4. Ortega DR, Oikonomou CM, Ding HJ, Rees-Lee P & Jensen GJ ETDB-Caltech: A blockchain-based distributed public database for electron tomography. PLoS One 14, e0215531 (2019). [PubMed: 30986271]

5. Khoshouei M, Pfeffer S, Baumeister W, Förster F & Danev R Subtomogram analysis using the Volta phase plate. J. Struct. Biol 197, 94–101 (2017). [PubMed: 27235783]
6. Bharat TAM & Scheres SHW Resolving macromolecular structures from electron cryo-tomography data using subtomogram averaging in RELION. Nat. Protoc 11, 2054–2065 (2016). [PubMed: 27685097]
7. Himes BA & Zhang P emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging. Nat. Methods 15, 955–961 (2018). [PubMed: 30349041]
8. Goddard TD, Huang CC & Ferrin TE Visualizing density maps with UCSF Chimera. J. Struct. Biol 157, 281–287 (2007). [PubMed: 16963278]
9. Goddard TD et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. 27, 14–25 (2018). [PubMed: 28710774]

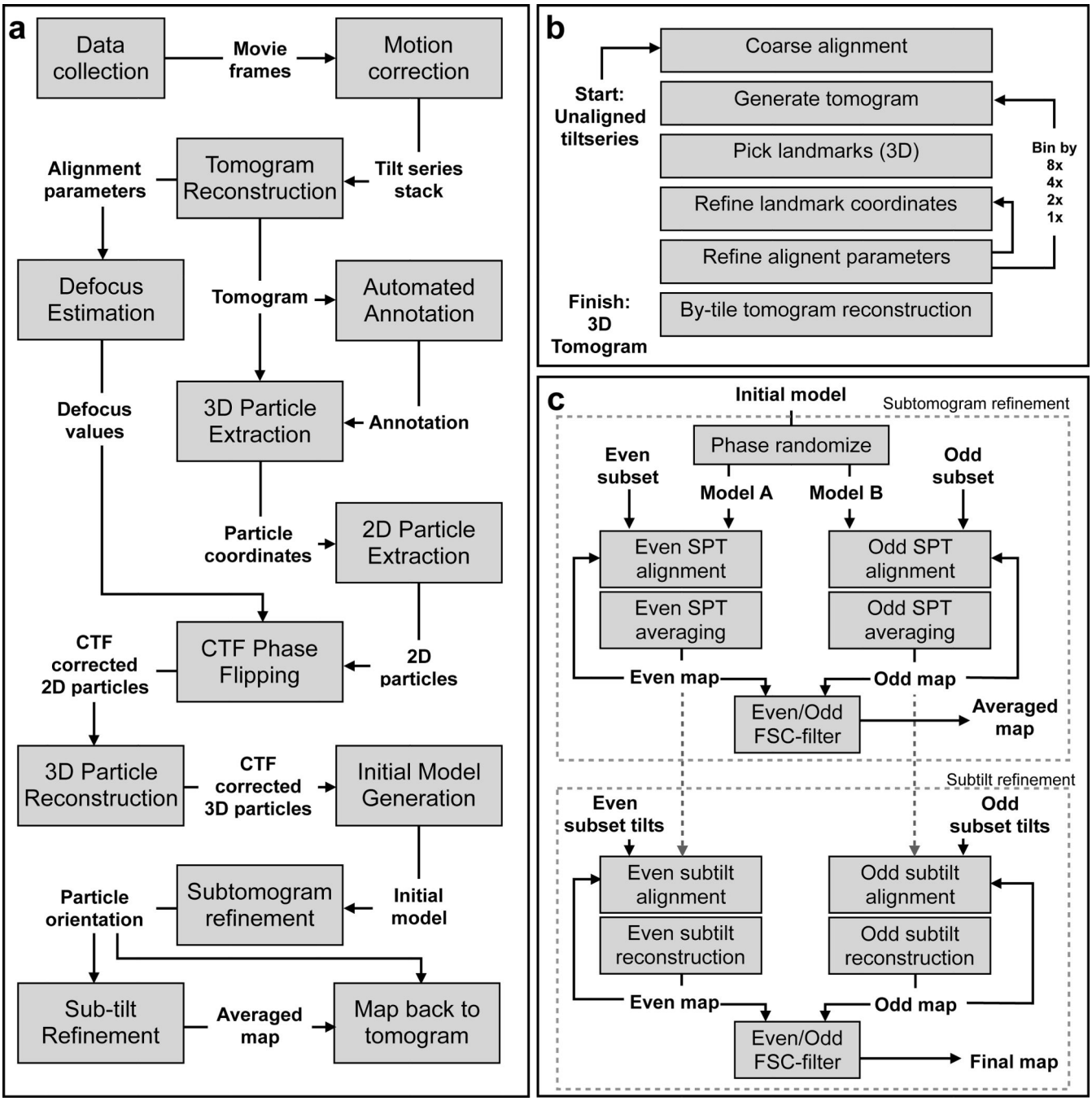

Figure 1. Diagram of CryoET data processing workflow.

(a) Main workflow diagram. (b) Workflow of tomogram reconstruction. (c) Workflow of subtomogram refinement and subtilt refinement.

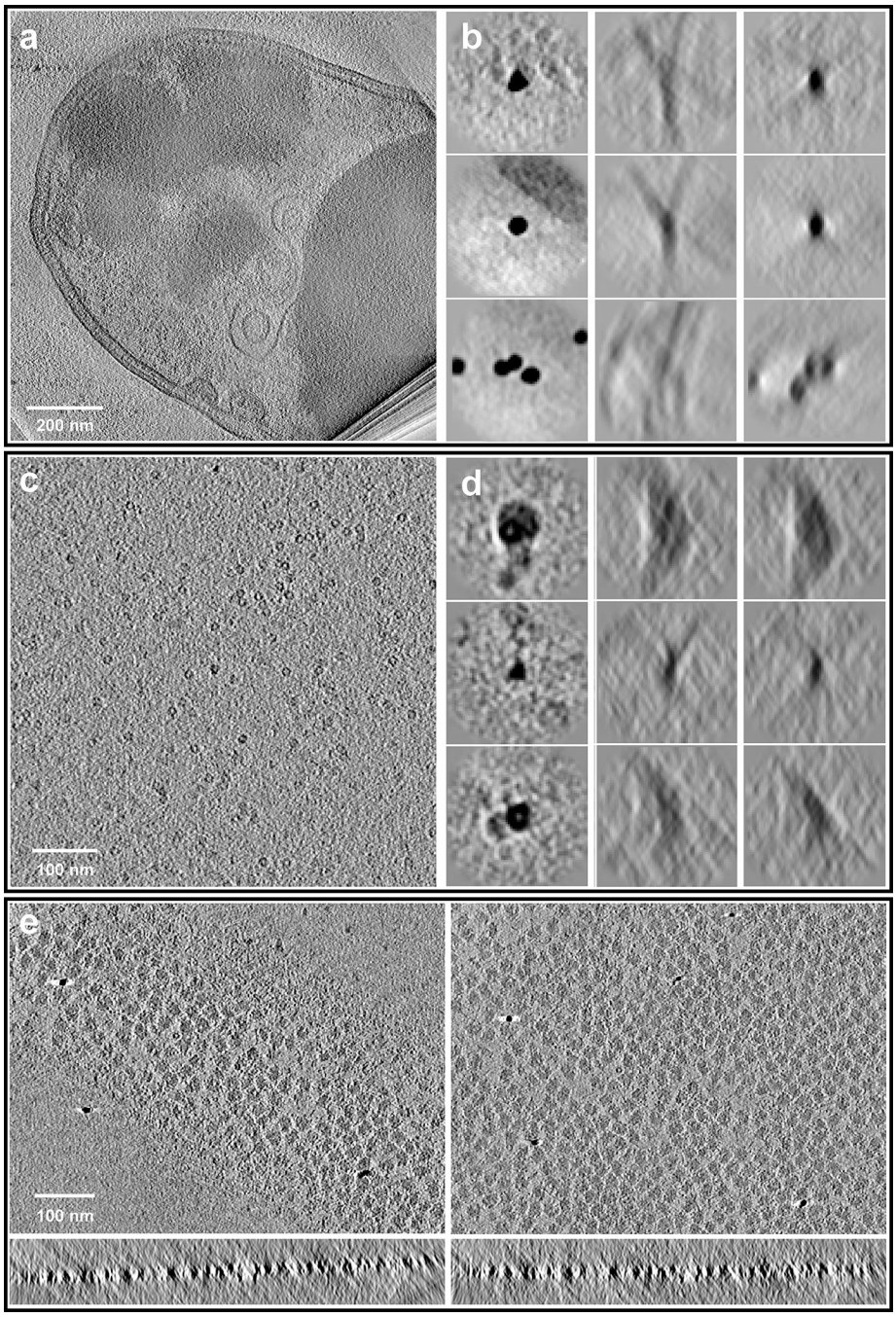

Figure 2. Results of iterative tomogram alignment and reconstruction.

(a) Cellular tomogram of an E. coli bacterium with gold fiducials. (b) Selected landmark projections from (a) (left) x-y plane; (mid) x-z plane after the first iteration of the iterative alignment; (right) x-z plane after iterative alignment. (c) Tomogram of purified apoferritin without fiducials (EMPIAR-10171). (d) Selected landmark projections from (c). (e) Automatic specimen plane positioning. Left: (top) x-y slice (bottom) x-z slice, both before specimen plane positioning; right: the specimen becomes flat in the tomogram after automated positioning.

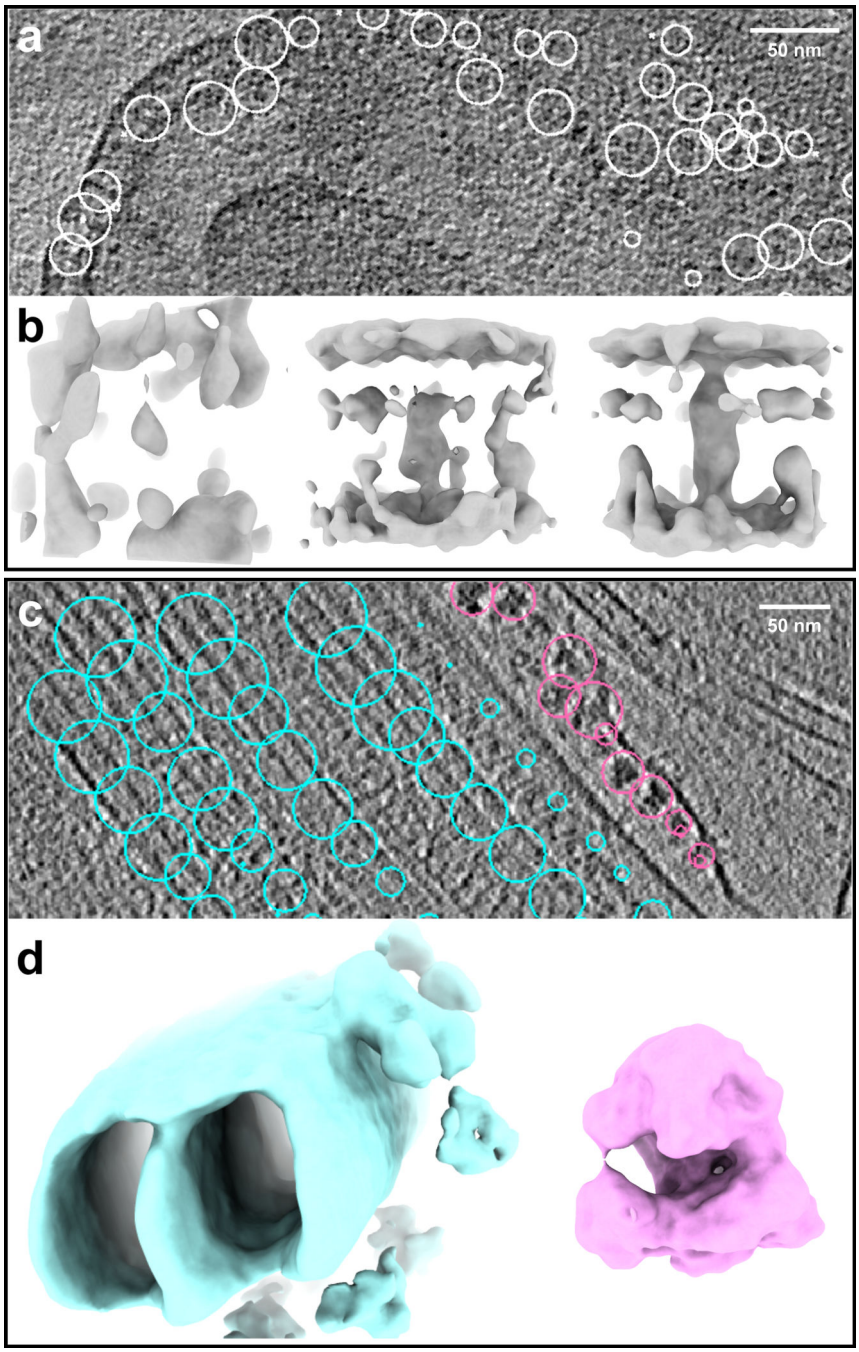

Figure 3. Particle extraction and initial model generation.

(a) Slice view of a E. coli tomogram with particles of Tolc-AcrAB pump selected. (b) Initial model generation from Tolc-AcrAB pump particles. From the left to right are density maps of the initial seed, after 5 iterations with c1 symmetry, and after 5 iterations with c3 symmetry. (c) A tomogram slice view of the flagellum of an anucleated *Trypanosoma brucei* cell, with cyan circles selecting microtubule doublets, and pink circle selecting ribosomes. (d) Initial model generated from microtubule (left) doubles and ribosomes (right).

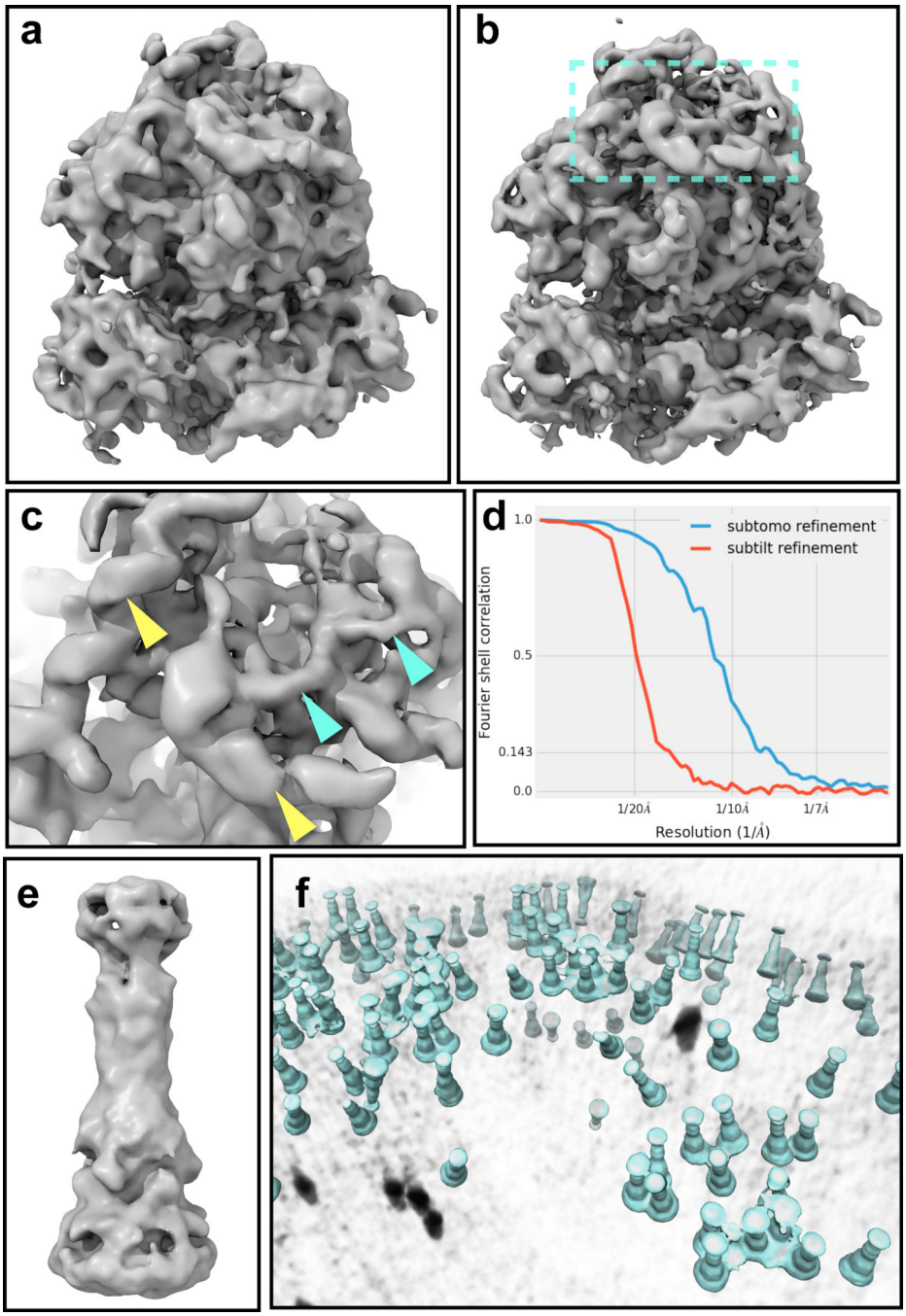

Figure 4. Subtomogram refinement.

(a) Subtomogram averaging of ribosome (EMPIAR-10064) before subtilt refinement. (b) Subtomogram averaging after subtilt refinement. (c) Zoomed-in view of (b) with yellow arrows pointing to RNA helices and cyan arrows pointing to resolved alpha-helices. (d) Gold-standard FSC curves of the ribosome subtomogram averaging before (red) and after (blue) subtilt refinement. (e) Subtomogram averaging of the tolc-acrAB drug pump. (f) Location and orientation of the drug pump particles mapped back to a tomogram.