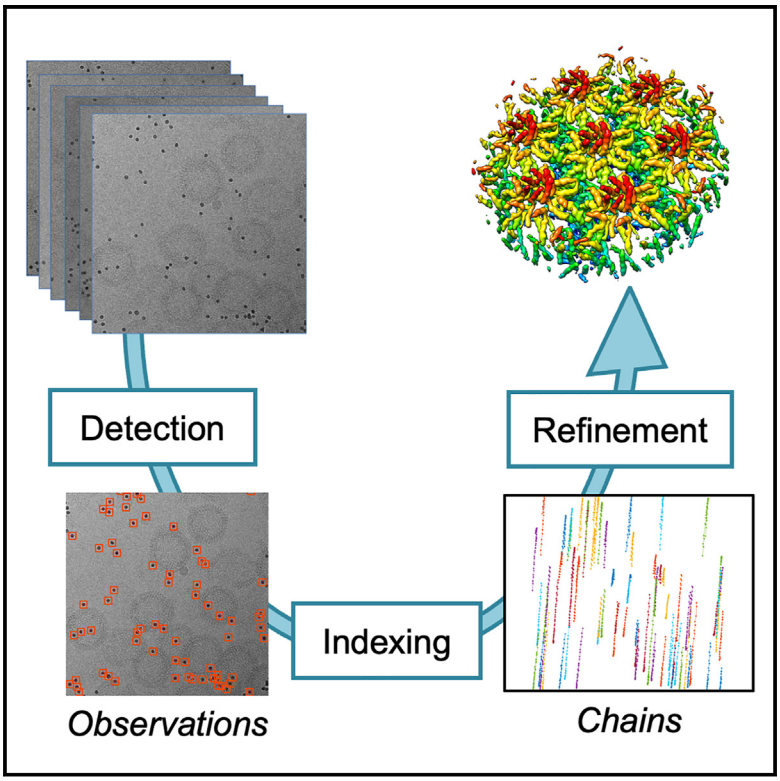
Resource

Structure

Automated fiducial-based alignment of cryo-electron tomography tilt series in Dynamo

Graphical abstract



Authors

Raffaele Coray, Paula Navarro, Stefano Scaramuzza, Henning Stahlberg, Daniel Castaño-Díez

Correspondence

daniel.castano@csic.es

In brief

Coray et al. present the software implementation of an algorithm for alignment of cryo electron tomography tilt series, necessary to correct the mechanical imperfections of data acquisition and bring all acquired micrographs to a common geometric reference. The software has been tested on different samples, assessing its robustness and accuracy.

Highlights

- Automated gold-based alignment of cryo-ET tilt series in Dynamo
- Robustness validated in 356 tilt series, aligning 99.7% micrographs out of 14387
- Automated alignment shown to be suitable for high resolution subtomogram averaging

Coray et al., 2024, Structure *32*, 1808–1819
October 3, 2024
https://doi.org/10.1016/j.str.2024.07.003

CellPress

CellPress
OPEN ACCESS

Structure

Resource

Automated fiducial-based alignment of cryo-electron tomography tilt series in Dynamo

Raffaele Coray,[1] Paula Navarro,[2,3] Stefano Scaramuzza,[2] Henning Stahlberg,[2,4] and Daniel Castaño-Díez[1,2,5,*]
[1]Instituto Biofisika (Consejo Superior de Investigaciones Científicas, Universidad del País Vasco), University of Basque Country, 48940 Leioa, Spain
[2]Center for Cellular Imaging and NanoAnalytics (C-CINA), Biozentrum, University of Basel, Mattenstrasse 26, CH-4058 Basel, Switzerland
[3]Department of Fundamental Microbiology, Faculty of Biology and Medicine, University of Lausanne, 1015 Lausanne, Switzerland
[4]Laboratory of Biological Electron Microscopy, Institute of Physics, School of Basic Science, EPFL, and Department of Fundamental Microbiology, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland
[5]Lead contact
*Correspondence: daniel.castano@csic.es
https://doi.org/10.1016/j.str.2024.07.003

SUMMARY

With the advent of modern technologies for cryo-electron tomography (cryo-ET), high-quality tilt series are more rapidly acquired than processed and analyzed. Thus, a robust and fast-automated alignment for batch processing in cryo-ET is needed. While different software packages have made available several approaches for automated marker-based alignment of tilt series, manual user intervention remains necessary for many datasets, thus preventing high-throughput tomography.

We have developed a MATLAB-based framework integrated into the Dynamo software package for automatic detection of fiducial markers that generates a robust alignment model with minimal input parameters. This approach allows high-throughput, unsupervised volume reconstruction. This new module extends Dynamo with a large repertory of tools for tomographic alignment and reconstruction, as well as specific visualization browsers to rapidly assess the biological relevance of the dataset. Our approach has been successfully tested on a broad range of datasets that include diverse biological samples and cryo-ET modalities.

INTRODUCTION

Alignment of tilt series in cryo-ET

Cryo-electron tomography (cryo-ET) is the method of choice for visualizing the near-native arrangement of cellular organelles and macromolecules within their biological context[1–3] The cryo-ET pipeline has been classically recognized as a time-consuming workflow with identified bottlenecks in sample preparation and image processing, so far preventing high-throughput cryo-ET. In brief, a series of micrographs are collected at several tilt angles (producing a tilt series) by physically rotating the specimen around a single axis perpendicular to the electron beam.[4,5] Subsequently, each tilt series is aligned and reconstructed into a 3D-volume, the tomogram. Thus, tilt series alignment and tomogram reconstruction are essential steps in cryo-ET with a direct impact on the attainable resolution of the resulting 3D structure.[6] To accurately determine the alignment parameters for a tilt series for high-resolution cryo-ET, fiducial markers are often used.[7] Spherical gold beads are the most common fiducial marker, as gold scatters electrons more strongly than biological material and the projections of these spheres appear as high-contrast circles in the images, regardless of tilt angle.[8,9] Markerless alignment methods can be used when the introduction of gold beads is not viable, or simply to skip the sample optimization attached to their introduction.[10,11]

Advances in cryo-EM, such as the development of energy filters and direct electron detectors, have been essential to overcome common challenges, including the low signal-to-noise ratio (SNR) and resolution limits.[12,13] In parallel, important improvements in microscope hardware and software for automated acquisition have recently allowed rapid collection of high-quality tilt series.[14–19] In combination with high-performance CPU (central processing unit) and GPU (graphics processing unit) algorithms,[20,21] high-resolution 3D structures below the nanometer have been determined.[22–26] Accordingly, several software packages have implemented tilt series alignment and tomogram reconstruction in their workflows, being IMOD the most popular.[6] Other software packages including such workflows are Protomo,[27–29] AuTom,[30] TOM,[31] UCSF tomography,[32] EMAN2,[33] AreTomo,[11] and RAPTOR[34] among others. These software packages provide a well-defined pipeline, addressing tilt series alignment and tomogram reconstruction.

Subtomogram averaging

An often needed further processing step is subtomogram averaging (STA), necessary to bring cryo-ET to high resolution.[35] STA exploits the fact that biological features are present in multiple copies within a tomogram and provides a framework to determine structural information by the numerical analysis of many noisy, randomly oriented subtomograms. STA

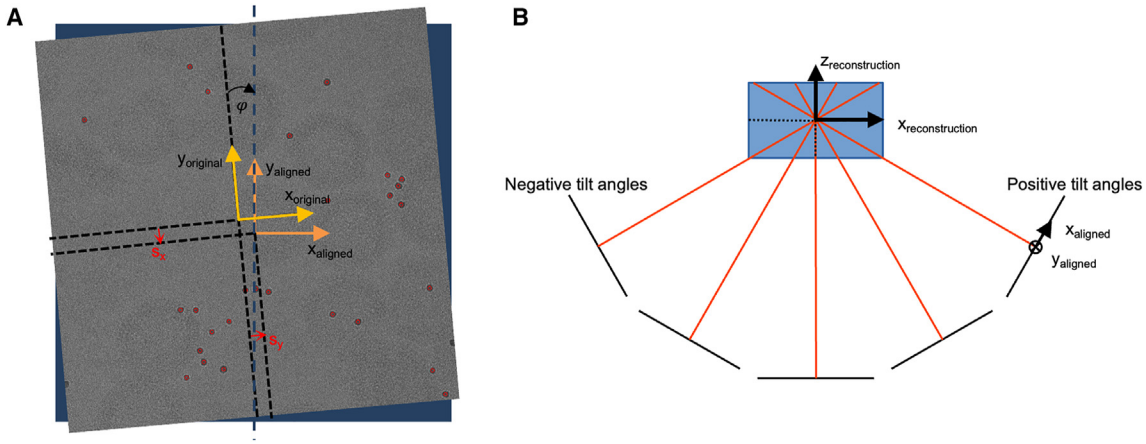


Figure 1. Schematic representation of misalignment

(A) The micrograph is presented after alignment. The real tilt axis of the sample is noted with a dashed blue line. The fiducials are noted with red circles. The micrograph is initially described in the system of reference of the microscope detector, noted by the $x_{original}$ and $y_{original}$ axes. The required system of reference for alignment is described by the $x_{aligned}$ and $y_{aligned}$ axes. Note the collinearity between the $y_{aligned}$ axis and the tilt axis. The transform from the original reference system to the aligned reference system is then described as a mis-centering and a rotational correction. The mis-centering of the sample is defined by the shifts s_x and s_y in the original reference system and a rotation of angle ϕ after shifting. The shifts are noted by the red arrows. The rotational correction is noted by the black arrow.

(B) Schematic representation of the general projection geometry of a tilt series. The origins of the alignment system of reference of each micrograph correspond to projections of the center of the tomogram. As such, the centering shifts of each micrograph can only be determined by exploiting the positions of the fiducials of the other micrographs in the same tilt series.

computationally extracts subtomograms containing the structure of interest from a set of tomograms. After subtomogram extraction, STA performs a 3D alignment to compute an average that retrieves a 3D signal-enhanced structure to overcome the low signal-to-noise ratio characteristic of cryo-ET data; beyond the so obtained consensus structure, further classification steps can be used to discover different conformational states. In such a workflow, the quality of structures attained by STA is directly linked to the quality of the alignment of the tilt series. While newer trends[36] allow for further refining the tilt series alignment *a posteriori* using information gained during the STA procedure itself, a reliable initial alignment is still necessary. In fact, despite efforts toward streamlining processing,[37,38] many high-resolution STA structures rely on manual-aided tilt series alignment.[7,22,24,26,35] Currently, a variety of packages offer different implementations of the technique, such as Dynamo,[20] RELION,[39] PEET,[40] Jsubtomo,[41] EMAN2,[21] PyTom,[42] M,[43] and emClarity.[36]

The Dynamo software package

Dynamo, in particular, specializes in image processing for STA and is adapted for various high-performance computing platforms, including GPU or CPU clusters. The backbone of Dynamo is written in the programming languages MATLAB, CUDA, and C++, so that Dynamo is distributed both as MATLAB scripts as well as precompiled executables for different computing environments (Linux and macOS).[44] Dynamo is also characterized by user-friendliness and flexibility to create new algorithms and combine them with other external tools. Moreover, it includes, among others, rich modules for data visualization and management, tomogram annotation, STA experiments, subtomogram classification and, as detailed in this paper, a new module for marker-based tilt series alignment.

While existing software packages provide reliable results for tilt series alignment, we aim at enhancing the automation necessary for high-throughput and precise alignment.

Consequently, many high-resolution STA structures rely on manual-aided tilt series alignment,[7,22,24,26,35] with recent efforts in streamlining processing.[37] The principal task that demands manual intervention is the precise detection of bead centers and tracking through all tilts. Taking into account the increasing speed of data acquisition,[14] we are in the need for a robust and fast automated alignment in cryo-ET that enables batch processing and aims to be integrated with high-resolution STA strategies.

Here, we present the implementation of an automatic tilt series alignment algorithm in Dynamo. The core of the algorithm is an elaborate method for automatic fiducial detection and tracking. Its implementation is optimized for high throughput, minimal user intervention, and minimal input parameters. The alignment pipeline can be controlled through a graphical user interface (GUI) or the command line. Along the core utilities in the pipeline, we have also developed additional features for localized reconstruction and the handling of large datasets and a broad range of diagnostic tools.

We present various numerical experiments performed on publicly available data to demonstrate the software's diverse features. The automated analysis of four different EMPIAR (electron microscopy public image archive) entries (10453, 10499, 10393, and 10889), comprising 356 tilt series, illustrates the tool's robustness in handling a large number of tilt series across different scenarios. A STA computation based on EMPIAR entry 10164 highlights the method's capability to yield high-resolution structures, reproducing results attained with carefully curated tilt series alignments. Finally, the analysis of EMPIAR entry 10226

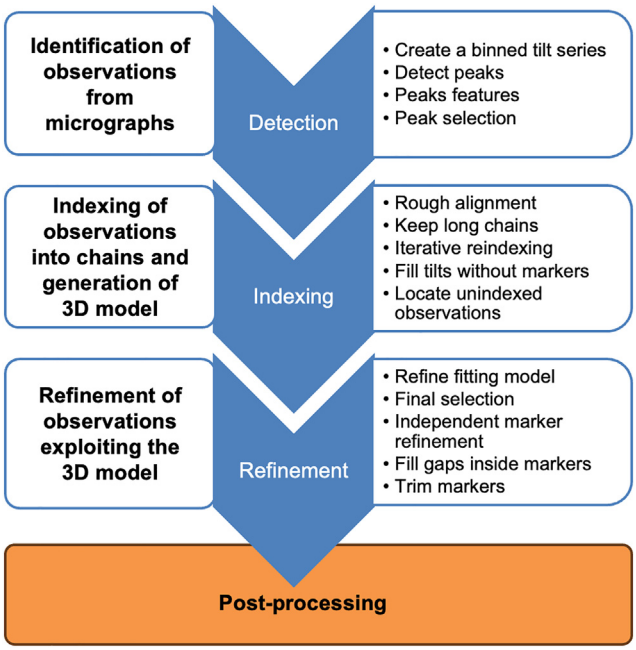


Figure 2. Algorithm workflow

The flowchart presents most of the separate algorithmically steps of the implemented algorithm. Each step is executed sequentially. In total, there are 21 steps that have been assigned to six main blocks: Detection, Indexing, Refinement, Alignment, CTF, and Reconstruction. The names of the steps included in a block are presented on the right, while the overall goal of the block is presented on the left. The last three blocks are considered collectively as a post-processing block because they consist of merely applying the estimated 3D model to the micrographs to generate a tomogram. The description of the most important steps of the blocks is detailed in the appendix provided in the supplementary material.

demonstrates the algorithm's suitability for processing data acquired through a fast tilt series acquisition scheme.

RESULTS

General workflow

Figure 1 illustrates the image formation model underpinning our marker-based alignment algorithm.

We posit that each micrograph observes a different, random shift experienced by the sample due to mechanical imperfections during the tilting process. Additionally, the apparent axis of rotation of the sample is not expected to be coincident with the axis of the detector. The goal of our algorithm is thus to determine those shifts and rotations that lead to an optimal match between the theoretical projections of the 3D locations of the physical gold beads, to the actual trails of gold beads observed in the micrographs. The optimality criterion is coded as a least squares problem, a full mathematical formulation of which can be found in the appendix provided in the supplementary material. Therein, once chains of reliable observations of gold beads are available, solving the equations that describe the image formation model is relatively straightforward (even the non-linearity implied by the entanglement of shifts and rotations can be safely and efficiently handled by approximated solution methods). The challenging task is precisely to fulfill the premise for the numerical solution of the least squares problem: the identification of chains of gold beads.

We state here a set of consistent definitions of the key concepts (highlighted in cursive) that appear in our approach. An *observation* is a position in a micrograph that we assume to correspond to the projection of a fiducial gold bead (a three-dimensional entity). *Detection* is the process of finding the *observations* in a micrograph. *Indexing* of observations is the process of identifying which -two dimensional- observations correspond to a single -three dimensional- fiducial. A *chain* is a set of unique *indexed observations* in different micrographs that correspond to the same fiducial, while the *reprojection trail* is the theoretical positions of where a fiducial would appear in each micrograph given its position in the tomogram. A fiducial that corresponds to a chain is a *marker*. With this terminology, aligning a tilt series corresponds to solving for the alignment parameters that generate the smallest fitting residual between chains and reprojection trails of markers. The final fitting residual is estimated as a root-mean-square error of the final chains, and it will be noted as RMS.

Figure 2 shows the steps of the algorithm that we have incorporated into the Dynamo package to solve the alignment problem as defined previously. The overall flow seen by the user of the software is a sequence of five operations. Following the naming convention in the accompanying GUI, the first block, "Detection," creates a set of unindexed observations that represent the putative positions of gold beads in the whole set of micrographs. The second block, "Indexing," generates an initial set of chains, by testing for subsets of gold bead observations that are compatible with a 3D projection model. The third block, "Refinement," chooses an initial model as a seed to iteratively improve the assignment of gold bead observations to chains, then solve for the final alignment parameters. Further modules called "Alignment" and "CTF" are used to apply the mathematical alignment model previously computed on the micrograph set and are considered as post-processing blocks.

The different mechanisms used to initiate and improve the model as the iteration procedure progresses (by detecting markers with missing observations, locating undetected markers, revealing false positives, and fine-tuning the centering of observed gold peaks) are described in detail in the STAR methods and depicted in Figures S1–S3, with a formal mathematical description in the appendix provided in the supplementary information.

Post processing options

Even after determination of a final set of chains, the set of alignment parameters that produce a coherent projection model is not uniquely determined: As explained in the following, both the orientation of the micrographs relative to the electron beam, and the definition of the center of the reconstructed tomogram induce multiple possible solutions, This is handled by the last two blocks in the workflow, which do not seek to improve the alignment quality but to make explicit the constraints that determine an unique, physically meaningful solution to the alignment problem.

(1) Reconstruction centering. As described in Figure 1, the shifts necessary to register each micrograph coherently depend on the definition of the tomogram center. As

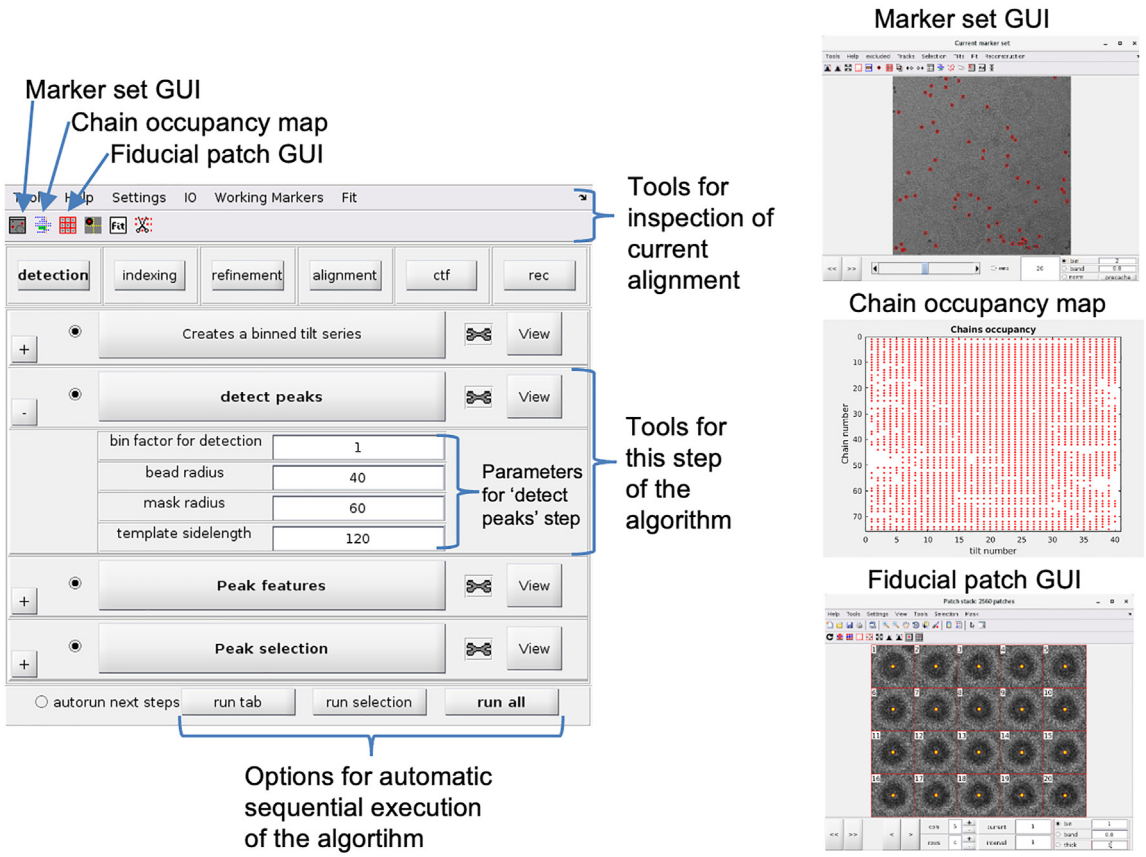


Figure 3. General GUI and different inspection GUIs

The general GUI is presented on the left. The header includes tools for the inspection and manipulation of the last computed alignment geometry. The main viewing tools are the marker set GUI, the chain occupancy map, and the fiducial patch GUI. These GUIs are presented on the right part of the figure. The main body of the GUI presents the alignment algorithm step-by-step following the description of Figure 2. Each step is represented as a sub-box. The "detect peak" step has been chosen as an example. Each step includes a parameter window, a main execution button, an additional tools button with a wrench logo, and a view button. The algorithm can be executed by pressing the buttons with the step names manually or using the option at the bottom of the general GUI.

such, an indexing model can generate equivalent but different alignment models for different tomogram centers. Our convention to resolve this ambiguity is to assign a shift of zero pixels to the micrograph whose tilt angle is closest to zero degrees, designating it as a fixed reference, then select the alignment model that produces the overall smallest possible shifts on the entire tilt series. We call this process z-centering. This strategy produces the smallest possible alignment shifts while unchanging the residuals, thus limiting the generation of aligned micrographs with large empty regions. In the general case, this strategy ensures that most of the material present in the original micrographs appears in the reconstruction. However, this does not ensure that the material of interest will be located at the center of the tomogram; thus the code allows the option of selecting different tomogram centers if requested.

(2) CTF coherence. One of the tasks of the alignment procedure is the determination of the angle of the tilt axis ϕ, which customarily will be assigned to the y direction in the aligned tilt series and posterior reconstruction. The election of this angle entails a mathematical ambiguity: if a value of ϕ generates a valid solution, then an additional rotation of 180° generates an equally valid solution. However, only one of them is coherent with the physics of the image formation model: in our convention (following the default in SerialEM[16,45]), a positive tilt angle is meant to represent a clockwise rotation around the y axis of the aligned micrograph. Thus, for a positive angular rotation during underfocus image acquisition, the part of the sample with positive x axis values should move downward, away from the front focal plane when in underfocus image acquisition. To accommodate this effect, the CTF coherence step automatically tests the ϕ initially computed by the algorithm by applying the alignment parameters and then estimating separately the defocus of sections along the x axis in all the aligned micrographs in a tilt series. If the tested ϕ follows the geometrical convention, the defocus computed on the sections will increase along the x axis more rapidly for larger positive tilts angles and will decrease for negative tilt angles. The algorithm computes a robust regression to model the trends of the local defocus values for all of the tilt angles. If the estimation does not match the convention, ϕ will be corrected by an additional 180°.

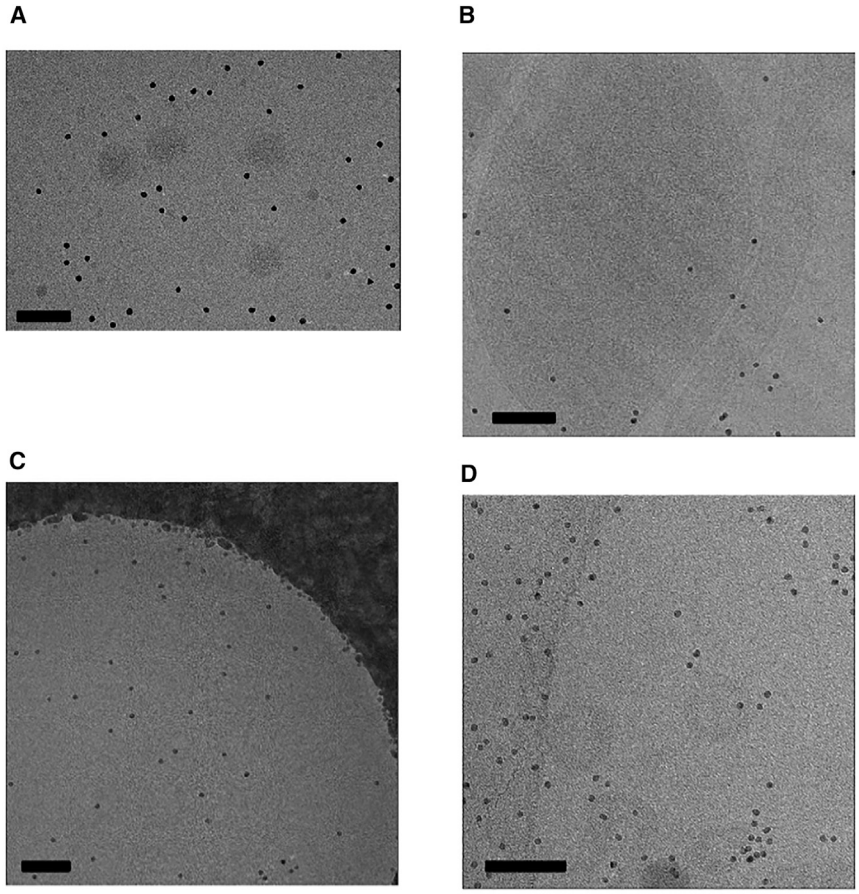


Figure 4. Example micrographs of each of the selected EMPIAR datasets

All micrographs are at a tilt angle of 0°. Scale bars are 100 nm.

(A) Micrograph from the EMPIAR 10453 dataset.

(B) Micrograph from the EMPIAR 10499 dataset.

(C) Micrograph from the EMPIAR 10393 dataset.

(D) Micrograph from the EMPIAR 10889 dataset.

Software package

The described algorithm has been implemented in the Dynamo software package. The Dynamo software is open-source and freely distributed. Dynamo is written mainly in MATLAB but can also be used without a MATLAB installation using the standalone modus. The algorithm can be executed through a step-by-step GUI-based approach or directly through the command line, enabling straightforward scripting in MATLAB or from a Unix terminal. This allows for interoperability with other cryo-ET packages as IMOD or RELION, as specific tools for format conversion are provided and documented online (see STAR methods). Both these approaches will generate a Dynamo workflow object that allows the user to modify, execute and inspect each step of the algorithm. This allows the user to develop a custom, fine-tuned version of the algorithm for their specific dataset if required. Conversely, the algorithm can be executed in its default configuration with only minimal parameter tuning. The only necessary parameters requested from the user are the gold bead radius and the gold bead masking radius.

Workflow GUI

the general Dynamo tilt series alignment workflow GUI has been developed as the basic GUI for the execution and inspection of the implemented algorithm. Figure 3 presents the general GUI as well as the main inspection GUIs. Each step of the algorithm corresponds to a subsection of the GUI containing an "execution" button, an extendable parameter window, as well as a "tool" and "view" functions. The "tool" button can be used to access auxiliary tools that are not included in the default pipeline, such as tools to manually correct the results. The "view" button is used to access the visualization tools for inspection of the results of that specific step of the algorithm. Each step of the algorithm can be executed by pressing the button with the name of the step, or all sequentially by pressing the "run all" button. Each algorithm execution step updates an object encoding the current status of the alignment. This allows the user to compare the current alignment status with the status at each of the previous algorithm steps and identify possible optimal parameter choices. Additionally, the current alignment can be reverted to the alignment of a previous step to rapidly test different parameters. The current status of the alignment can be inspected with the inspection tools in the top part of the GUI. The main inspection tools are the marker set GUI, the chain occupancy map, and the fiducial patch GUI. The fiducial patch GUI presents a gallery of all gold beads found during the detection part of the algorithm, while the marker set GUI will highlight the positions of the indexed observations in the full micrographs. On the other hand, the chain occupancy map is a graphical summary of the populations of chains on the different micrographs.

Command line

while the GUI implementation of the algorithm is suitable to investigate each step of the algorithm and determine an optimal set of parameters, the most practical approach to align numerous tilt series is though the Dynamo command line. The Dynamo command line can be used to achieve the same results as with the GUI. In fact, a Dynamo workflow object generated using the command line can at any time be opened and inspected using the GUI. Detailed descriptions of the most common operations with the command line and with the GUI are available in the Dynamo online wiki.[46,47]

Auxiliary tools

a set of auxiliary tools are provided in addition to the generalist tools used for running and inspecting the automatic alignment.

These tools allow for the manual identification and correction of misaligned micrographs. All of them can be accessed through the general GUI or independently through the command line.

(1) *Gold bead revision*: this tool allows the user to manually correct the generated chains. It consists of an interactive gallery of micrographs where the user can select, remove,

Table 1. General characteristics of validation datasets

Name	Sample	Number of samples	Size (px)	Pixel size (Å)	Tilt step (°)	Tilt range (°)
10453	SARS-CoV-2	231	5760 × 4092	1.33	3	−60,60
10499	*M. pneumoniae* cells	36	3838 × 3710	1.7	3	−60,60
10393	Tobacco mosaic virus	4	7420 × 7676	1.1	3	−60,60
10889	Equine infectious anemia virus	85	4096 × 4096	1.18	3	−60,60

Table 2. Algorithm parameter applied for each validation dataset

Name	Gold bead radius (px)	Mask radius (px)	Template sidelength (px)	Binning
10453	40	60	120	2
10499	30	50	100	2
10393	60	80	160	2
10889	40	50	100	2

reposition and add observations to the chain. The tool is dynamically linked to other inspection tools to provide a good contextual view of the chain models.

(2) *Tilt line revision:* this feature allows for the selection of a chain as a reference to inspect all other chains. The indexed observations of the other chains are shown using the observations in the chosen chain as reference. This allows the user to quickly identify possible inconsistencies of relative positions caused by improper indexing.

(3) *Selection of features*: this tool allows for the inspection of the gold beads as gallery of patches. A user can utilize this tool to detect false positive gold beads and manually remove them to increase the quality of the observation set before indexing.

(4) *CTF fitting estimation*: this feature allows the user to inspect the estimated defocus of the entire tilt series. This information can be used to evaluate if the hand of the micrographs is consistent with the geometrical model.

(5) *Tomogram generation*: the final step of the algorithm is the application of the final alignment parameters to the tilt series. The algorithm will generate the set of rotated and shifted micrographs to create the alignment tilt series. The algorithm will also generate a log of the corresponding tilt angles. This is necessary because the aligned tilt series will present fewer micrographs than the unaligned tilt series if the algorithm or the user rejected some micrographs. The aligned tilt series can then be dose weighted and CTF corrected by the algorithm and the tomogram can be reconstructed. By default, the algorithm will reconstruct the tomogram using weighted back-projection or SIRT (simultaneous iterative reconstruction technique) as well as binned tomograms. Additionally, optional local reconstructions of the tomogram only around given positions are possible

Computational validation of robustness on 356 tilt series

Our tool aims at automating the task of aligning large sets of tilt series, pursuing the 2-fold objective of performing the alignment both robustly and accurately.

Here, *robustness* implies the capacity of the algorithm to use minimal user input to assign viable alignment parameters to a large subset of micrographs in a given tilt series, avoiding apparent jumps in the transition between aligned adjacent projections, and minimizing the number of micrographs that are rejected by the algorithm. *Accuracy* entails the capacity of the produced alignment parameters to deliver tomograms that preserve high-resolution features of the data.

To quantify the extent to which these goals are met by the presented algorithm, we have devised two types of numerical experiments. In the first set of experiments, we aim at assessing the robustness of the method for different profiles of spatial distribution of fiducial markers. For this, we have first visited several sets of tilt series registered in the public database EMPIAR, selecting the entries (10453, 10499, 10393, and 10889),[43,48–50] which jointly amount to a total of 356 unaligned tilt series. This selection was driven by an initial visual inspection that indicated a qualitatively distinct behavior of the gold beads in each of the EMPIAR entries.

A micrograph from each dataset is presented in Figure 4. In particular, each group was chosen to present different pixel sizes, micrograph dimensions and samples. Table 1 summarizes these characteristics of each of the selected EMPIAR datasets.

Each tilt series has been processed by the algorithm using the parameters selected for the particular EMPIAR dataset and default parameters when not specified otherwise. Table 2 presents the parameters chosen for each EMPIAR dataset. The parameters gold bead radius, gold bead mask radius and the template sidelength are the set of template matching parameters used in the initial step of the algorithm, the detection by cross-correlation. The parameters were assigned by visual inspection of some gold beads in the first tilt series of each EMPIAR dataset.

The results of the algorithm have been inspected for all tilt series to determine if the alignment was viable. The alignment of each micrograph was determined to be not satisfactory if a visual inspection showed an obviously incorrect positioning of the aligned micrograph. Table 3 presents the results of the visual inspection of the automatic alignment. In total, the algorithm was requested to align 14,387 micrographs divided into 356 tilt series. The algorithm automatically discarded a total of 218 micrographs from the alignments because they were considered as not consistent with the alignment geometry of the tilt series. The algorithm proceeded to generate alignment parameters for the remaining 98.5% of the full number of micrographs. Only 11 tilt series presented at least one micrograph that was identified as misaligned (3% of the total number of tilt series). The number of misaligned micrographs was 43, 0.3% of all micrographs that were not discarded. Some causes of errors in the alignments as well as strategies to correct misaligned tilt series are provided in the next section. The approximate computing time required to generate an aligned tilt series is about 3.5 min. The computing time was evaluated only as the time required to estimate the

Table 3. General statistics of the automatic alignment

Name	Number of tilt series	Number of input micrographs	Number of rejected micrographs	Number of misaligned micrographs	Number of correctly aligned micrographs	Computing time per tilt series (min)
10453	231	9305	126	25	9,154	3.5
10499	36	1470	8	0	1,462	2.5
10393	4	164	9	0	155	10
10889	85	3448	75	18	3,355	4.2

alignment geometry and generate the aligned tilt series. As such, this estimation does not include the time required to correct the CTF and backproject the aligned tilt series to generate a tomogram. The computations were done with the MATLAB 2023a parallel computing toolbox and using 20 parallel workers, on a workstation with an Intel(R) Xeon(R) Gold 6240R CPU at 2.40 GHz processor.

The quality of the final alignments has been measured using the RMS value and the number of frames rejected by the alignment algorithm in each tilt series. The overall distribution of these quality scores for each EMPIAR dataset is provided in Figures 5A and 5B.

In addition, the characteristics of the fiducials in each dataset are also shown in Figures 5C and 5D. The alignment algorithm has been used to estimate the relative positions of the fiducials in the reconstructed samples. The number and relative positions of fiducials have been estimated from these positions. It can be noted that the datasets present different trends of distributions of gold beads. The alignment algorithm achieved reasonable alignments with very limited user intervention on a set of samples that presented a large spectrum of different fiducials.

It is possible that more user intervention to fine-tune the algorithm parameters in each tilt series would improve the quality of the alignments in a case-by-case approach. While the algorithm achieved robust alignment results with general parameters, some particular cases of misalignment caused by suboptimal algorithm parameters are presented in the following section, as well as suggestions on how to improve alignment by choosing better algorithm parameters.

Particular cases

The main indicators of a poor alignment are a large number of dropped micrographs and a large RMS. In addition, a low number of chains compared to the number of expected gold beads in the micrographs can be an indicator of a suboptimal alignment. While the integration of every visible observation of every physical marker into the final indexing model is not a prerequisite for the quality of the alignment model and subsequent reconstruction, a low number of chains increases the chances of the solution of the alignment algorithm being suboptimal.

There are some particular cases that can generate a lower-than-expected number of chains. Missing micrographs can cause the coarse alignment by pairwise shifting (described as Step 2.1 in the STAR methods) to be computed between two tilt angles that differ by multiple angular tilt steps. This can render the matching of observations more imprecise and, by consequence, generate fewer chains. A way to reduce this effect is to increase the matching threshold parameter to allow the algorithm to compensate for the perspective change caused by a larger angular tilt gap between micrographs.

Similarly, the presence of occluded or motion-blurred micrographs can be problematic. Since no correct alignment parameters exist for such micrographs, their presence will only decrease the quality of the chains and may cause the algorithm to drop neighbor tilt micrographs. The most straightforward solution to remove this possibility is to either directly remove poor quality micrographs from the tilt series before alignment or to manually indicate the identity of micrographs to be ignored by the alignment algorithm.

Additionally, we have identified instances of tracking loss in other automated alignment systems, such as IMOD, and implemented corrective measures. Figure S4 provides an example, demonstrating Dynamo's capability to identify and discard micrographs that exhibit significant shifts and are incompatible with the global alignment model. By detecting and removing these problematic micrographs, Dynamo prevents subsequent tracking failures.

Finally, poor template matching parameters (used in the detection stage) can reduce the number of chains and, in extreme cases, lead to misalignment. Here, the two parameters the user needs to input are the radius of the gold bead and the radius of the circular mask used for identifying gold beads through template matching. The radius of the gold bead can be measured on screen with a specific tool in Dynamo. It is suggested to set the radius of the mask to a maximum value of one and a half times the radius of the gold bead. A mask that is too large can cause issues for gold beads located near high-intensity features (such as membranes, carbon, or other gold beads), while a mask that fits the gold bead too tightly will produce a template that does not effectively include the contrast between the body of the gold bead, the characteristic CTF halo around it, and the background of the image. If needed, the GUI visualization tools allow the user to inspect the results produced by the template matching on sample micrographs. Once it is verified that the parameters yield coherent results, they can be applied to all other tilt series from the same acquisition session.

Accuracy assessment on HIV-1 virus-like particles

The second experiment illustrates the accuracy attained by the method by performing a STA computation on a set of tilt series previously aligned by Dynamo and measuring the attained resolution. EMPIAR entry 10164 was chosen for this task, a dataset frequently used to gauge the performance of new numerical approaches in the field.[26] This dataset consists of compressed, unaligned, multi-frame micrographs of tilt series containing immature HIV-1 dMACANC virus-like particles assembled in the presence of Bevirimat. This dataset shows how the major structural protein of HIV-1, Gag, assembles in highly structured lattices close to a spherical geometry.

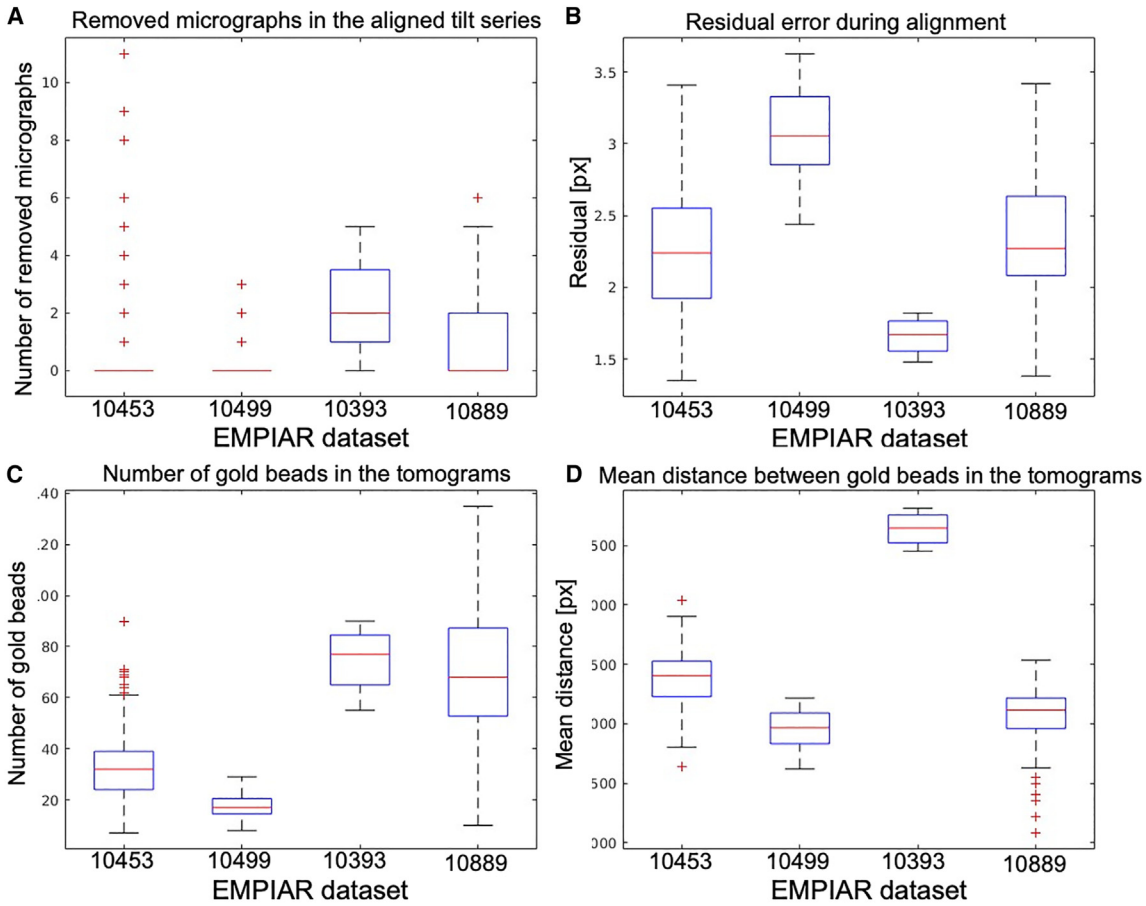


Figure 5. Quality scores of the automatic alignment and statistical features of the four tested datasets represented with boxplots

On each blue box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the black whiskers extend to the most extreme datapoints not considered as outliers, and the outliers are plotted individually as red crosses.

(A) The number of rejected frames in each tilt series. A complete tilt series contains generally 41 micrographs.

(B) Residual errors in each aligned tilt series.

(C) Number of gold beads in the tomograms.

(D) Estimated mean distance between of gold beads in the tomograms.

For our study, we considered a subset of tilt series, selecting the five tilt series with identifiers 01, 03, 43, 45, and 54 of the full EMPIAR dataset 10164, an example set that we have characterized in a previous study. In a study by Scaramuzza and Castaño-Diez,[51] we described a step-by-step procedure that starts in a manual alignment of each tilt series and leads to a final average of a hexamer of the CA (capsid) and SP1 (spacer peptide 1) region of Gag, the building piece of the lattice. Now, we have tested the capacity of the tilt series alignment algorithm presented here by reproducing exactly the full protocol that we used in the study by Scaramuzza and Castaño-Diez[51] with a single change: at the stage of the alignment of the tilt series, the human intervention used in our previous study has now been replaced by the application of the automatic system now implemented in Dynamo. Therefore, after generating the five tomograms with our new tilt series alignment algorithm, we analyzed them in exactly the same way as the tomograms produced by manual alignment with IMOD in our previous study. To extract and align 14,878 particles containing ca-SP1, we manually segmented a total of 45 virus-like particles in the tomograms. We then ran separate alignments on the particles from each tomogram and finally merged and aligned the particles from all the tomograms using a gold standard approach. The full procedure is described and made reproducible through the Dynamo wiki.[52] The resolution was estimated at 4.5 Å by gold standard Fourier shell correlation (FSC) in RELION. Figure 6 presents the FSC curve and the density map of the final average. This resolution estimate matches the one we previously attained using the same protocol with manual tilt series alignment.

Accuracy assessment on fast tilt series acquisition

Lastly, we validated the performance of the algorithm in conjunction with a fast tilt series acquisition scheme. We used the tilt series of *Bdellovibrio bacteriovorus* from EMPIAR 10226, acquired by the fast-incremental method, employing a bi-directional tilt scheme with 1-degree tilt steps. Dynamo successfully achieved an alignment that allowed the identification of all features described in

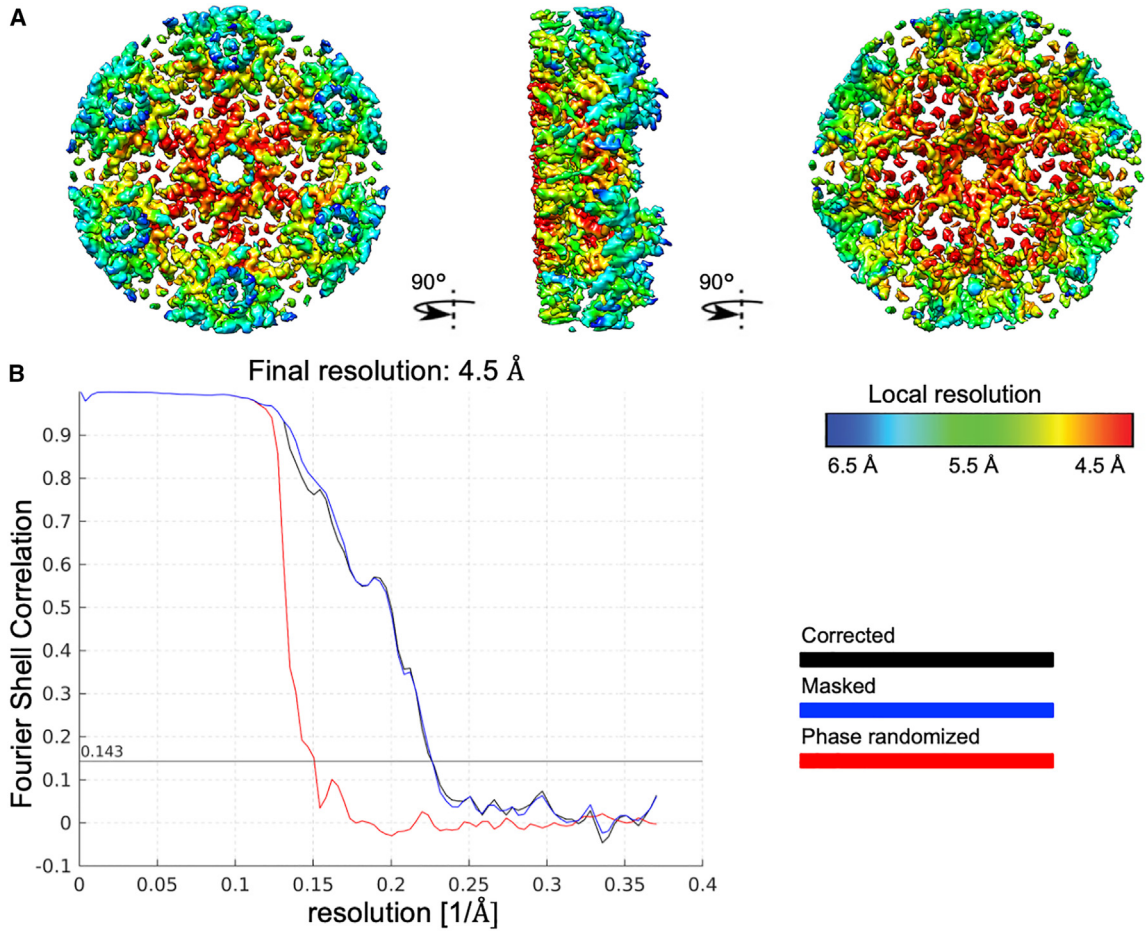


Figure 6. Cryo-EM structure validation of HIV-1 immature particles with automatic tilt series alignment

(A) Local resolution result for the density in the reconstruction estimated with the Bsoft software.

(B) Global resolution estimation using gold standard FSC curves computed with the RELION software.

the original publication (see Figures S5A–S5C), where the authors used IMOD for alignment. Additionally, a subtomogram average of 273 particles of a methyl-accepting chemotaxis protein (MCP) array was generated (see Figures S5C and S5D).

DISCUSSION

An implementation of an automatic fiducial-based alignment of tilt series for cryo-ET has been implemented in the Dynamo software package. The algorithm's robustness and scalability have been tested on multiple datasets from the public domain. The algorithm proved to be able to correctly align 99.7% of the tested micrographs, while at the same time requiring only minimal parameter tuning. Additionally, the aligned procedure was completed in a few minutes per tilt series on a low-range computer. The testing datasets consisted of 356 tilt series presenting a range of different samples, micrograph aspect ratios, pixel sizes and gold bead distributions. Moreover, it has been demonstrated that a resolution of 4.5 Å can be achieved in a structure using tomograms generated with the new algorithm, matching the results previously attained through laborious manual alignment. Additionally, a discussion of specific tilt series cases that may lead to suboptimal alignment results has been included, offering guidelines to optimize the use of the provided tools.

The main steps of the algorithms have been described, along with the most important parameters. Additionally, a set of tools has been provided for inspecting the quality of the alignment and allowing users to manually improve the alignment if necessary.

A beta version of the algorithm has been used by third-party users to solve structures, present a set of protocols[53,54] and develop powerful computational tools[55] in a cross-software framework.

The modularity of the software allows for an easy integration of possible enhancements by the user. For instance, the initial detection of gold beads can be handled by third party software, then directly fed into Dynamo or used to provide the workflow with an automated, data-based estimate of the radius of the gold beads in the sample.

Finally, the introduction of the automatic tilt series alignment workflow in the Dynamo toolbox opens the door for the exploitation of per-particle refinements, such as local CTF refinements

and corrections of non-rigid sample motions,[39,53] in the context of STA. In particular, the core mechanism in the approach presented here is the identification of a coherent projection model on the basis of an initially unindexed set of observation. This could be used as a framework to build an algorithm for markerless alignment, where observations of gold beads could be replaced by salient features detected in the micrographs.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Main step 1: Identification of gold beads
  - Main step 2: Indexing of observations
  - Main step 3: Trail refinement
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.str.2024.07.003.

ACKNOWLEDGMENTS

We thank Alma Vivas Lago for her assistance in the reproduction of the STA pipeline described for the analysis of EMPIAR dataset 10164 from immature HIV-1 virus-like particles. H.S. acknowledges support by the Swiss National Science Foundation (grants CRSII5_177195 and 310030_188548). D.C.-D. acknowledges support by the Human Frontiers Science Program (HFSP) (grant RGP0017/2020), the Swiss National Science Foundation (grant 205321 179041), and the Spanish Ministry of Science and Innovation (grant PID2021-127309NB-I00 funded by AEI/10.13039/501100011033/FEDER, UE). P.N. acknowledges support from the Swiss National Science Foundation (Starting Grant TMSGI3_218251).

AUTHOR CONTRIBUTIONS

D.C.-D. conceived the study; and D.C.-D. and H.S. co-supervised the study. D.C.-D. designed methods and D.C.-D. and R.C. implemented methods and software. R.C., P.N., and S.S. ran analysis experiments. D.C.-D. and R.C. analyzed the results. R.C. and D.C.-D. wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 15, 2024
Revised: June 6, 2024
Accepted: July 3, 2024
Published: July 29, 2024

REFERENCES

1. Asano, S., Engel, B.D., and Baumeister, W. (2016). In Situ Cryo-Electron Tomography: A Post-Reductionist Approach to Structural Biology. J. Mol. Biol. *428*, 332–343. https://doi.org/10.1016/j.jmb.2015.09.030.
2. Gan, L., and Jensen, G.J. (2012). Electron tomography of cells. Q. Rev. Biophys. *45*, 27–56. https://doi.org/10.1017/S0033583511000102.
3. Lučić, V., Rigort, A., and Baumeister, W. (2013). Cryo-electron tomography: The challenge of doing structural biology in situ. J. Cell Biol. *202*, 407–419. https://doi.org/10.1083/jcb.201304193.
4. Beck, M., and Baumeister, W. (2016). Cryo-Electron Tomography: Can it Reveal the Molecular Sociology of Cells in Atomic Detail? Trends Cell Biol. *26*, 825–837. https://doi.org/10.1016/j.tcb.2016.08.006.
5. Tocheva, E.I., Li, Z., and Jensen, G.J. (2010). Electron Cryotomography. Cold Spring Harb. Perspect. Biol. *2*, a003442. https://doi.org/10.1101/cshperspect.a003442.
6. Mastronarde, D.N., and Held, S.R. (2017). Automated tilt series alignment and tomographic reconstruction in IMOD. J. Struct. Biol. *197*, 102–113. https://doi.org/10.1016/j.jsb.2016.07.011.
7. Wan, W., and Briggs, J.A.G. (2016). Chapter Thirteen - Cryo-Electron Tomography and Subtomogram Averaging. In Methods in Enzymology The Resolution Revolution: Recent Advances In cryoEM, R.A. Crowther, ed. (Academic Press), pp. 329–367. https://doi.org/10.1016/bs.mie.2016.04.014.
8. Amat, F., Castaño-Diez, D., Lawrence, A., Moussavi, F., Winkler, H., and Horowitz, M. (2010). Chapter Thirteen - Alignment of Cryo-Electron Tomography Datasets. In Methods in Enzymology Cryo-EM, Part B: 3-D Reconstruction, G.J. Jensen, ed. (Academic Press), pp. 343–367. https://doi.org/10.1016/S0076-6879(10)82014-2.
9. Hou, G., Yang, Z., Zang, D., Fernández, J.-J., Zhang, F., and Han, R. (2024). MarkerDetector: A method for robust fiducial marker detection in electron micrographs using wavelet-based template. J. Struct. Biol. *216*, 108044. https://doi.org/10.1016/j.jsb.2023.108044.
10. Sorzano, C.O.S., de Isidro-Gómez, F., Fernández-Giménez, E., Herreros, D., Marco, S., Carazo, J.M., and Messaoudi, C. (2020). Improvements on marker-free images alignment for electron tomography. J. Struct. Biol. X *4*, 100037. https://doi.org/10.1016/j.yjsbx.2020.100037.
11. Zheng, S., Wolff, G., Greenan, G., Chen, Z., Faas, F.G.A., Bárcena, M., Koster, A.J., Cheng, Y., and Agard, D.A. (2022). AreTomo: An integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction. J. Struct. Biol. X *6*, 100068. https://doi.org/10.1016/j.yjsbx.2022.100068.
12. Kühlbrandt, W. (2014). Cryo-EM enters a new era. Elife *3*, e03678. https://doi.org/10.7554/eLife.03678.
13. Kühlbrandt, W. (2014). The Resolution Revolution. Science *343*, 1443–1444.
14. Chreifi, G., Chen, S., Metskas, L.A., Kaplan, M., and Jensen, G.J. (2019). Rapid tilt-series acquisition for electron cryotomography. J. Struct. Biol. *205*, 163–169. https://doi.org/10.1016/j.jsb.2018.12.008.
15. Hagen, W.J.H., Wan, W., and Briggs, J.A.G. (2017). Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. J. Struct. Biol. *197*, 191–198. https://doi.org/10.1016/j.jsb.2016.06.007.
16. Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51. https://doi.org/10.1016/j.jsb.2005.07.007.
17. Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: The new Leginon system. J. Struct. Biol. *151*, 41–60. https://doi.org/10.1016/j.jsb.2005.03.010.
18. Zheng, Q.S., Braunfeld, M.B., Sedat, J.W., and Agard, D.A. (2004). An improved strategy for automated electron microscopic tomography. J. Struct. Biol. *147*, 91–101. https://doi.org/10.1016/j.jsb.2004.02.005.
19. Eisenstein, F., Yanagisawa, H., Kashihara, H., Kikkawa, M., Tsukita, S., and Danev, R. (2023). Parallel cryo electron tomography on in situ lamellae. Nat. Methods *20*, 131–138. https://doi.org/10.1038/s41592-022-01690-1.
20. Castaño-Díez, D., Kudryashev, M., Arheit, M., and Stahlberg, H. (2012). Dynamo: A flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing

environments. J. Struct. Biol. *178*, 139–151. https://doi.org/10.1016/j.jsb.2011.12.017.

21. Galaz-Montoya, J.G., Flanagan, J., Schmid, M.F., and Ludtke, S.J. (2015). Single particle tomography in EMAN2. J. Struct. Biol. *190*, 279–290. https://doi.org/10.1016/j.jsb.2015.04.016.
22. Hutchings, J., Stancheva, V., Miller, E.A., and Zanetti, G. (2018). Subtomogram averaging of COPII assemblies reveals how coat organization dictates membrane shape. Nat. Commun. *9*, 4154. https://doi.org/10.1038/s41467-018-06577-4.
23. Khoshouei, M., Pfeffer, S., Baumeister, W., Förster, F., and Danev, R. (2017). Subtomogram analysis using the Volta phase plate. J. Struct. Biol. *197*, 94–101. https://doi.org/10.1016/j.jsb.2016.05.009.
24. Kovtun, O., Leneva, N., Bykov, Y.S., Ariotti, N., Teasdale, R.D., Schaffer, M., Engel, B.D., Owen, D.J., Briggs, J.A.G., and Collins, B.M. (2018). Structure of the membrane-assembled retromer coat determined by cryo-electron tomography. Nature *561*, 561–564. https://doi.org/10.1038/s41586-018-0526-z.
25. Pfeffer, S., Burbaum, L., Unverdorben, P., Pech, M., Chen, Y., Zimmermann, R., Beckmann, R., and Förster, F. (2015). Structure of the native Sec61 protein-conducting channel. Nat. Commun. *6*, 8403.
26. Schur, F.K.M., Obr, M., Hagen, W.J.H., Wan, W., Jakobi, A.J., Kirkpatrick, J.M., Sachse, C., Krausslich, H.-G., and Briggs, J.A.G. (2016). An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. Science *353*, 506–508. https://doi.org/10.1126/science.aaf9620.
27. Noble, A.J., and Stagg, S.M. (2015). Automated batch fiducial-less tilt-series alignment in Appion using Protomo. J. Struct. Biol. *192*, 270–278. https://doi.org/10.1016/j.jsb.2015.10.003.
28. Winkler, H., and Taylor, K.A. (2006). Accurate marker-free alignment with simultaneous geometry determination and reconstruction of tilt series in electron tomography. Ultramicroscopy *106*, 240–254. https://doi.org/10.1016/j.ultramic.2005.07.007.
29. Winkler, H., and Taylor, K.A. (2013). Marker-free dual-axis tilt series alignment. J. Struct. Biol. *182*, 117–124. https://doi.org/10.1016/j.jsb.2013.02.004.
30. Han, R., Wan, X., Wang, Z., Hao, Y., Zhang, J., Chen, Y., Gao, X., Liu, Z., Ren, F., Sun, F., and Zhang, F. (2017). AuTom: A novel automatic platform for electron tomography reconstruction. J. Struct. Biol. *199*, 196–208. https://doi.org/10.1016/j.jsb.2017.07.008.
31. Nickell, S., Förster, F., Linaroudis, A., Net, W.D., Beck, F., Hegerl, R., Baumeister, W., and Plitzko, J.M. (2005). TOM software toolbox: acquisition and analysis for electron tomography. J. Struct. Biol. *149*, 227–234. https://doi.org/10.1016/j.jsb.2004.10.006.
32. Zheng, S.Q., Keszthelyi, B., Branlund, E., Lyle, J.M., Braunfeld, M.B., Sedat, J.W., and Agard, D.A. (2007). UCSF tomography: An integrated software suite for real-time electron microscopic tomographic data collection, alignment, and reconstruction. J. Struct. Biol. *157*, 138–147. https://doi.org/10.1016/j.jsb.2006.06.005.
33. Chen, M., Bell, J.M., Shi, X., Sun, S.Y., Wang, Z., and Ludtke, S.J. (2019). A complete data processing workflow for cryo-ET and subtomogram averaging. Nat. Methods *16*, 1161–1168. https://doi.org/10.1038/s41592-019-0591-8.
34. Amat, F., Moussavi, F., Comolli, L.R., Elidan, G., Downing, K.H., and Horowitz, M. (2008). Markov random field based automatic image alignment for electron tomography. J. Struct. Biol. *161*, 260–275. https://doi.org/10.1016/j.jsb.2007.07.007.
35. Briggs, J.A. (2013). Structural biology in situ—the potential of subtomogram averaging. Curr. Opin. Struct. Biol. *23*, 261–267. https://doi.org/10.1016/j.sbi.2013.02.003.
36. Himes, B.A., and Zhang, P. (2018). emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging. Nat. Methods *15*, 955–961. https://doi.org/10.1038/s41592-018-0167-z.
37. Liu, H.-F., Zhou, Y., Huang, Q., Piland, J., Jin, W., Mandel, J., Du, X., Martin, J., and Bartesaghi, A. (2023). nextPYP: a comprehensive and scalable platform for characterizing protein variability in situ using single-particle cryo-electron tomography. Nat. Methods *20*, 1909–1919. https://doi.org/10.1038/s41592-023-02045-0.
38. Balyschew, N., Yushkevich, A., Mikirtumov, V., Sanchez, R.M., Sprink, T., and Kudryashev, M. (2023). Streamlined structure determination by cryo-electron tomography and subtomogram averaging using TomoBEAR. Nat. Commun. *14*, 6543. https://doi.org/10.1038/s41467-023-42085-w.
39. Bharat, T.A.M., Russo, C.J., Löwe, J., Passmore, L.A., and Scheres, S.H.W. (2015). Advances in Single-Particle Electron Cryomicroscopy Structure Determination applied to Sub-tomogram Averaging. Structure *23*, 1743–1753. https://doi.org/10.1016/j.str.2015.06.026.
40. Heumann, J.M., Hoenger, A., and Mastronarde, D.N. (2011). Clustering and variance maps for cryo-electron tomography using wedge-masked differences. J. Struct. Biol. *175*, 288–299. https://doi.org/10.1016/j.jsb.2011.05.011.
41. Huiskonen, J.T., Parsy, M.-L., Li, S., Bitto, D., Renner, M., and Bowden, T.A. (2014). Averaging of viral envelope glycoprotein spikes from electron cryotomography reconstructions using Jsubtomo. J. Vis. Exp. *xx*, e51714. https://doi.org/10.3791/51714.
42. Hrabe, T., Chen, Y., Pfeffer, S., Kuhn Cuellar, L., Mangold, A.-V., and Förster, F. (2012). PyTom: A python-based toolbox for localization of macromolecules in cryo-electron tomograms and subtomogram analysis. J. Struct. Biol. *178*, 177–188. https://doi.org/10.1016/j.jsb.2011.12.003.
43. Tegunov, D., Xue, L., Dienemann, C., Cramer, P., and Mahamid, J. (2021). Multi-particle cryo-EM refinement with M visualizes ribosome-antibiotic complex at 3.5 Å in cells. Nat. Methods *18*, 186–193. https://doi.org/10.1038/s41592-020-01054-7.
44. Castaño-Díez, D. (2017). The Dynamo package for tomography and subtomogram averaging: components for MATLAB, GPU computing and EC2 Amazon Web Services. Acta Crystallogr. D Struct. Biol. *73*, 478–487. https://doi.org/10.1107/S2059798317003369.
45. Mastronarde, D.N. (2003). SerialEM: A Program for Automated Tilt Series Acquisition on Tecnai Microscopes Using Prediction of Specimen Position. Microsc. Microanal. *9*, 1182–1183. https://doi.org/10.1017/S1431927603445911.
46. Walkthrough on GUI based tilt series alignment - Dynamo. https://www.dynamo-em.org/w/index.php?title=Walkthrough_on_GUI_based_tilt_series_alignment.
47. Walkthrough on command line based tilt series alignment - Dynamo https://www.dynamo-em.org/w/index.php?title=Walkthrough_on_command_line_based_tilt_series_alignment.
48. Turunova, B., Sikora, M., Schurmann, C., Hagen, W.J.H., and Beck, M. (2020). In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. Science *370*, 203–208.
49. Sanchez, R.M., Zhang, Y., Chen, W., Dietrich, L., and Kudryashev, M. (2020). Subnanometer-resolution structure determination in situ by hybrid subtomogram averaging - single particle cryo-EM. Nat. Commun. *11*, 3709. https://doi.org/10.1038/s41467-020-17466-0.
50. Obr, M., Hagen, W.J.H., Dick, R.A., Yu, L., Kotecha, A., and Schur, F.K.M. (2022). Exploring high-resolution cryo-ET and subtomogram averaging capabilities of contemporary DEDs. J. Struct. Biol. *214*, 107852. https://doi.org/10.1016/j.jsb.2022.107852.
51. Scaramuzza, S., and Castaño-Diez, D. (2021). Step-by-step guide to efficient subtomogram averaging of virus-like particles with Dynamo. PLoS Biol. *19*, e3001318.
52. High resolution walkthrough with HIV1 - Dynamo. https://www.dynamo-em.org/w/index.php?title=High_resolution_walkthrough_with_HIV1.
53. Pyle, E., and Zanetti, G. (2021). Current data processing strategies for cryo-electron tomography and subtomogram averaging. Biochem. J. *478*, 1827–1845.

54. Pyle, E., Hutchings, J., and Zanetti, G. (2022). Strategies for picking membrane-associated particles within subtomogram averaging workflows. Faraday Discuss *240*, 101–113. https://doi.org/10.1039/D2FD00022A.

55. Burt, A., Gaifas, L., Dendooven, T., and Gutsche, I. (2021). A flexible framework for multi-particle refinement in cryo-electron tomography. PLoS Biol. *19*, e3001319. https://doi.org/10.1371/journal.pbio.3001319.

56. Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221. https://doi.org/10.1016/j.jsb.2015.08.008.

57. Grant, T., and Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. Elife *4*, e06980.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Cryo electron tomography - tilt-series of Sars-Cov-2	Turunova et al.[48]	EMPIAR: 10453
Tilt series of native M. pneumoniae cells treated with chloramphenicol	Tegunov et al.[43]	EMPIAR: 10499
Subnanometer-resolution structure determination *in situ* by a hybrid subtomogram averaging - single particle cryoEM - workflow - on TMV	Sánchez et al.[49]	EMPIAR: 10393
Tilt series of EIAV CASPNC VLPs acquired on Krios G4 with Selectris X and Falcon 4	Obr et al.[50]	EMPIAR: 10889
Cryo-electron tomography of immature HIV-1 dMACANC VLPs	Schur et al.[26]	EMPIAR: 10164
Bdellovibrio bacteriovorus electron cryotomography tilt-series acquired by fast-incremental method	Chreifi et al.[14]	EMPIAR: 10226
Software and algorithms		
Dynamo software package	Castaño-Díez et al.[20]	https://www.dynamo-em.org/w/index.php?title=Downloads
Other		
Documented tutorial to reproduce the subtomogram averaging experiments on HIV-1 EMPIAR dataset 10164	This paper	https://www.dynamo-em.org//w/index.php?title=High_resolution_walkthrough_with_HIV1
Tutorial that documents the use the *Dynamo* tools for manual alignment.	This paper	https://www.dynamo-em.org//w/index.php?title=Walkthrough_on_manual_marker_clicking
Tutorial that documents the use of the *Dynamo* tools for automated alignment through the GUI.	This paper	https://www.dynamo-em.org//w/index.php?title=Walkthrough_on_GUI_based_tilt_series_alignment
Tutorial that documents the use of the *Dynamo* tools for automated alignment through the command line.	This paper	https://www.dynamo-em.org/w/index.php?title=Walkthrough_on_command_line_based_tilt_series_alignment

RESOURCE AVAILABILITY

Lead contact

Further information should be directed to and will be fulfilled by the lead contact, Daniel Castaño (daniel.castano@csic.es).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data used in this study stems from the EMPIAR public repository. Accession numbers are listed in the key resources table.
- All original code is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

We detail here the implemented algorithm, as explained in Figure 2.

Main step 1: Identification of gold beads

Given an initial tilt series, this stage of the algorithm aims at creating a set of *observations* defined across all micrographs in the tilt series, which represent the putative positions of gold beads. An initial set of observations is created by selecting the peaks of the cross-correlation images resulting from comparing the micrographs to a gold bead template. Then, the neighbourhoods of the peaks

are analyzed in a process called *feature extraction* in order to eliminate false positives (i.e., cross-correlation peaks that do not correspond to actual gold beads).

At this stage, the final set of observations is not fully accurate, as the putative gold bead positions are not precisely centered. Additionally, the set is not pure, as false positives may be included, and it is not exhaustive, as some actual gold bead projections might have been missed or incorrectly removed from the observation set.

Step 1.1 Bin tilt series

A simple step required to ease later computations. By default, the algorithm will select a binning factor in which gold beads appear with a radius of at least 8 pixels in the binning image, ensuring detectability in the next stages.

Step 1.2 Peak detection

A cross-correlation score is used to identify the initial set of gold bead observations. Each micrograph is compared to a synthetic template of a gold bead, based on the user-provided radius in pixels. The peaks of the normalized cross-correlation are located and used to create an averaged image of a gold bead in the tilt series. The cross-correlation map of each micrograph against this new template is then computed, providing a finer estimation of putative gold bead centers. This two-step process deliberately overestimates the number of possible gold beads in each micrograph. The algorithm uses a default value of 300, which is significantly higher than the typical number of gold beads in a normal micrograph. This approach aims to generate an initial set of observations that includes most, if not all, of the real gold beads present. Subsequent steps will focus on removing false positive observations.

Step 1.3 Extraction of peak features

Each observation corresponds to a micrograph patch around a peak of the cross-correlation computed in the previous. In this step, features are computed for each patch, to be optionally used for the selection of patches that actually contain gold beads. The most consistently useful feature is the rotational merit of each patch, defined in this context as the minimum cross correlation attained by a patch when it is compared against a series of rotated versions of itself. This arbitrary definition is conceived as a rough figure of merit to estimate the rotational symmetry of each patch. Additionally, non-default patch quality features are computed by applying different filters to the template and patch (as Sobel filters, bandpass filters or combinations of both) before computing their cross-correlation, by estimating the inertia moment of each patch, and by calculating the components of each patch along the basis of the principal components of the set of all patches, along the usual lines of Principal Component Analysis (PCA).

Step 1.4 Peak selection

The features computed in the previous step are used to detect false positive observations and reject them. The default selection method uses only the rotational merit and the normalized cross-correlation. The observations on each micrograph are ordered by their cross-correlation merit in descending order and the ones above the threshold computed by Otsu's method are selected. This procedure is repeated by reordering the detected observations in terms of their rotational merit and selecting the ones above Otsu's threshold. The default selection method outputs the set of observations that score above Otsu's threshold, both for cross-correlation and rotational merit. The union of all such observations for all micrographs is the *initial (unindexed) observation set*. Alternative options for peak selection include clustering in the bidimensional space of rotational merit and cross-correlation value, or N-dimensional k-means clustering in the space of PCA components. The manual selection of the initial (unindexed) observation set is also possible in the visualization GUI.

Main step 2: Indexing of observations

The previous step of detection generates the initial observation set. However, the observations do not carry explicit information on how some may stem from different views of the same gold bead. As such, the observations are now indexed into chains.

Step 2.1 Rough alignment

The first step of the indexing of the initial observation set is based on the analysis of the observations of each subsequent pair of micrographs. Subsequent micrographs present the smallest relative tilt angle difference and, as such, the smallest apparent deformation of the cloud of gold bead projections. The initial indexing is based on the procedure called cloud shifting between micrographs and consists of exhaustively computing the number of *matches* between the observations of two subsequent micrographs for all possible relative shifts between the two micrographs. A match is defined as two observations in subsequent micrographs whose relative distance falls below a predetermined matching threshold after shifting one of the micrographs. The relative shift that produces the maximum number of matches is stored, repeating this procedure for each pair of subsequent micrographs. Figure S1 summarizes the pairwise shifting procedure. The parameters to be provided by the user are the matching threshold and the maximal shifts to be tested, while the result is a set of matched pairs of observations between subsequent micrographs. After the collection of all possible matched pairs, the algorithm constructs chains of matched observations across the tilt series. The initial chain is constructed from a concatenated subset of indexed observations from different micrographs. Observations from two subsequent micrographs are matches. The chain is said to be maximal if no more observations from the initial set can be added to the chain while still respecting the matching rule. A maximal chain is constructed by selecting a random observation in a micrograph, then adding the matched observation in the next micrograph if a matched operation is available, and repeating this operation along the tilt series, until the last added observation has no matches on the next micrograph, or until the last micrograph has been reached. The algorithm computes all maximal chains of matches attainable in the initial observation set, and they constitute the initial set of chains.

Ideally, each chain should represent a single marker (trail purity), and all projections of a given marker would be contained in the same chain (trail completeness). At this stage, the chains will not have either of these characteristics. Wrong matches of observations can easily happen when the view in a micrograph presents the projections of two gold beads within a close distance, thus preventing

the purity of the chains. Also, if the process of building a chain fails to identify a match in a given micrograph, that chain will be interrupted. As a result, observations of one single marker might be assigned to different experimental chains, each being defined on a different set of micrographs, creating gaps in a given chain.

Step 2.2. Chain selection

The chains generated by the previous steps have been based on only comparing the micrographs with their neighbors in the tilt series. However, the fact that all micrographs are projections of the same sample can be used to keep the chains that present the best purity and completeness.

Firstly, chains composed of too few observations are rejected. The default policy is to exclude chains that include less than 15 observations. This default value aims at ensuring the stability of the computations of the least squares problems underpinning the projection models in subsequent steps, ensuring a safe ratio of unknowns to observed data even for low numbers of indexed chains. The remaining chains are used to solve the alignment equation, giving a first estimation of the position of the markers, the rotational angle and the shifts.

At this step, a further selection of chains is performed based on the estimated 3D markers positions. When two or more markers are closer than a threshold, only the one that presents the best-fitting residual will cause its corresponding chain to be kept.

Step 2.3 Iterative reindexing

The surviving markers in the model fitted to the selected chains are reprojected. Each marker induces a set of positions from its reprojection trail, each on a different micrograph. The reprojection trail of a marker may not appear in every micrograph, as the coordinates of its reprojection might fall out of scope in some views. The chains created in previous steps are then discarded and recomputed using a new indexing: each reprojection trail is analyzed by visiting sequentially its reprojected positions on each micrograph. The experimental observation on that micrograph that is closest to a reprojected position is assigned to the marker index of that reprojection trail, provided this distance is smaller than a given threshold parameter. Figure S2 presents this reindexing process schematically.

The newly created set of chains can then be used to produce a refined set of markers, which can then be used to generate a further set of chains. By default, this procedure is repeated until no further improvement in the fitting quality is measured, or until a maximal number of iterations is met.

Step 2.4 Tilt gap filling

The set of markers created by the iterative refinement is used to locate imperfections inherited from previous processing steps. In particular, the pairwise shifting procedure applied to generate the first temptative indexing of the experimental observations did not incorporate any assumptions on the shifts among the micrographs, nor on the general projection geometry. This might lead to wrong matches or low numbers of correct matches. If all the matches in a micrograph are wrong, the reindexing step will reject most or all of the experimental observations in that micrograph, as they will not be coherent with the model induced by the majority of the observations. This will thus produce a "full tilt gap": a micrograph in which no observation has been indexed.

Micrographs that have not been indexed can be revisited after iterative reindexing has generated an initial geometry model. Similarly to the pairwise shifting procedure described above, the set of all predicted reprojection trails of the markers on one micrograph is compared to the set of all available observations on that micrograph. For this given micrograph, if a shift is found in which a number of projected markers is sufficiently close to a set of experimental observations, those observations are indexed, each one with index of the closest projected marker. The criterion for sufficient closeness between an individual observation and a projection is defined as a given threshold in pixels, or in terms of the typical residuals estimated on other micrographs.

Finally, after indexing a gap micrograph, the alignment model is recomputed and the process is repeated for all other remaining missing micrographs.

Step 2.5 Marker set extensor

The final step in the construction of the chains addresses the recovery of 3D markers whose real observations have never been indexed in any chain. To attempt to recover missing markers and their chains, the updated alignment model is exploited. The rotations and shifts corresponding to the current alignment model are applied onto all non-indexed observations. This allows to check for new *putative trails*. In this step, a putative trail is defined as a set of aligned observations from different micrographs whose positions correspond to the projections of a *putative marker* (or *putative gold beads*).

Putative markers are extracted from the tomogram generated by backprojecting the observations using the current alignment model. The candidate putative gold beads are then reprojected on the micrographs. A putative trail is defined as a set of aligned observations from different micrographs whose distance is below a given threshold. The putative trails are then used to index the observations in a way so that as the generated chains best match the putative trails. Figure S3 presents the marker set extension schematically. After the generation of the new set of chains, a new alignment model is solved.

Main step 3: Trail refinement

All of the previously computed chains are based on observations computed through cross-correlation in the initial steps of the algorithm.

In this section, the identities of chains are fixed and exploited to refine the observations of the gold beads. Indeed, the image patches around the observations in chains are analyzed in order to refine both the position of observations and the composition of the chains.

Step 3.1. Position refinement

The position of each gold bead of each observation is refined to subpixel precision by alignment of the patches cropped around the original observations. In this procedure, an average image is produced for each chain, and the observations are refined using this average image as template. The new position of each observation is computed by locating the closest cross-correlation peak.

Step 3.2 Chain gap filling

Missing observations inside a chain are called chain gaps. Chain gaps might be caused by the physical absence of the projection of the given marker in a particular tilt. This happens typically for markers located far away from the tilt axes, since projections of such markers might be possible at high tilt angles but not in the low tilts.

This step analyzes the gaps in the chains that are not explainable by the previously described geometrical reason. Other causes of a chain gap might be the occlusion of the gold bead by other objects of high intensity, such as other gold beads or sample features. This interference might have caused spurious cross-correlation peaks that caused the center of the gold bead projection to not be detected as an observation. This step adopts the strategy of scanning the behavior of the cross-correlation in the neighborhood of the reprojection: peaks detected inside a given radius of the reprojection are indexed into the trail if no further peaks are found inside an outer threshold radius. The inner threshold radius is set by default to the mean residual plus two standard deviations. The outer threshold radius prevents indexing cross-correlation peaks whose location might belong to other objects.

QUANTIFICATION AND STATISTICAL ANALYSIS

In section "Computational validation of robustness on 356 tilt series": Figure 5 depicts the statistics of four quantities (number of rejected frames, residual error per tilt series, number of gold beads, estimated mean gold bead distance) observed each on a different set of tilt series. Each set of tilt series is identified by its EMPIAR entry number (10453, 10499, 10393, 10889), and the number of tilt series (the n in the statistics) is the value reported in Table 3 for each dataset (231, 36, 4 and 85, respectively). The plot was generated with MATLAB R2024a through the function "boxplot", that reports median values and 25th and 75th quartiles.

In section "Accuracy assessment on HIV-1 virus-like particles". The dataset used consists of the five tilt series 01, 03, 43, 45 and 54 of the EMPIAR dataset 10164 from immature HIV-1 virus-like particles. All frames have a size of 3710 × 4092 pixels and a pixel size of 1.35 Å. The defocus estimation was done using CTFFIND4[56] and the CTF was corrected using the ctfphaseflip function of IMOD 6. Dose weighting was done using a MATLAB implementation of the algorithm introduced in Unblur.[57] The tilt series alignment and tomogram reconstruction were done using the Dynamo implementation of the automatic fiducial-based tilt series alignment algorithm using the same parameters in all tilt series. The algorithm parameters are: gold bead radius of 30 pixels with a corresponding mask and a tomogram thickness of 2000 pixels. The gold bead parameters were chosen by inspection and the tomogram thickness was chosen such that the samples would be completely inside the tomograms. The other optional parameters are left at their default values. The final average RMS score of the alignment is 2.04 pixels and no micrographs were dropped. The automatically generated final five tomograms were reconstructed using weighted back-projection.