

Enhancing Density Maps by Removing the Majority of Particles in Single Particle Cryogenic Electron Microscopy Final Stacks

Mengjia Cai^{*1}, Jianying Zhu^{*2}, Qi Zhang^{*3,4,5,6}, Yu Xu⁷, Zuoqiang Shi^{2,8}, Chenglong Bao^{2,8,9}, Mingxu Hu^{1,5,6}

¹Institute of Bio-Architecture and Bio-Interactions (IBABI), Shenzhen Medical Academy of Research and Translation ²Yau Mathematical Sciences Center ³Key Laboratory of Protein Science (Tsinghua University), Ministry of Education ⁴School of Life Science, Tsinghua University ⁵Beijing Advanced Innovation Center for Structural Biology ⁶Beijing Frontier Research Center for Biological Structure ⁷Department of Mathematical Sciences, Tsinghua University ⁸Yanqi Lake Beijing Institute of Mathematical Sciences and Applications ⁹State Key Laboratory of Membrane Biology, School of Life Science

*These authors contributed equally

Corresponding Authors

Chenglong Bao

clbao@tsinghua.edu.cn

Mingxu Hu

humingxu@smart.org.cn

Citation

Cai, M., Zhu, J., Zhang, Q., Xu, Y., Shi, Z., Bao, C., Hu, M. Enhancing Density Maps by Removing the Majority of Particles in Single Particle Cryogenic Electron Microscopy Final Stacks. *J. Vis. Exp.* (2024), e66617, doi:10.3791/66617 (2024).

Date Published

May 10, 2024

DOI

10.3791/66617

URL

jove.com/video/66617

Introduction

Cryogenic electron microscopy (cryo-EM) single-particle analysis (SPA) has become a dominant method to determine high-resolution three-dimensional density maps of

Abstract

Over the past decade, advancements in technology and methodology within the field of cryogenic electron microscopy (cryo-EM) single-particle analysis (SPA) have substantially improved our capacity for high-resolution structural examination of biological macromolecules. This advancement has ushered in a new era of molecular insights, replacing X-ray crystallography as the dominant method and providing answers to longstanding questions in biology. Since cryo-EM does not depend on crystallization, which is a significant limitation of X-ray crystallography, it captures particles of varying quality. Consequently, the selection of particles is crucial, as the quality of the selected particles directly influences the resolution of the reconstructed density map. An innovative iterative approach for particle selection, termed CryoSieve, significantly improves the quality of reconstructed density maps by effectively reducing the number of particles in the final stack. Experimental evidence shows that this method can eliminate the majority of particles in final stacks, resulting in a notable enhancement in the quality of density maps. This article outlines the detailed workflow of this approach and showcases its application on a real-world dataset.

biological macromolecules. Due to a series of technological innovations^{1,2,3,4,5,6}, named resolution revolution⁷, cryo-EM has the capability to determine the structures of

biological macromolecules with up to atomic resolution at an unprecedented rate. This breakthrough marks the beginning of a new era in molecular insights, overtaking X-ray crystallography as the predominant technique and answering longstanding biological questions.

Cryo-EM SPA diverges from X-ray crystallography by not requiring the crystallization of biological macromolecules. Instead, a solution containing the target biological macromolecules is rapidly frozen in vitreous ice. It is then imaged with an electron beam to produce a series of micrographs, bypassing the need for crystallization⁸. Subsequently, particle-picking algorithms are utilized to extract individual raw particles from these micrographs^{4,9,10,11,12}. As cryo-EM does not depend on crystallization, it is natural that extracted particles are predominantly damaged or in undesired conformational states, necessitating multiple rounds of particle selection to achieve a high-resolution density map. In cryo-EM SPA image processing, particle selection is therefore crucial for obtaining high-resolution density maps¹³.

In cryo-EM SPA, standard particle selection methods include two-dimensional (2D) and three-dimensional (3D) classification¹⁴. 2D classification categorizes particles into a predefined number of groups, yielding an average image and an estimated 2D resolution for each class. Researchers can then visually inspect these classes, removing particles from lower resolution groups to use the remaining ones in reconstructions aimed at achieving higher resolution. Once particle poses are established using refinement algorithms, researchers will proceed with 3D classification, clustering particles into multiple classes. This enables visual inspection of the reconstructed density map for each class, allowing for the exclusion of undesirable particles, such as those

from undesired conformations. Following multiple rounds of classification, a final stack comprising relatively high-quality particles is obtained. These final stacks are instrumental in producing atomic or near-atomic resolution density maps.

Zhu and her colleagues have demonstrated that further particle selection can be conducted on these final stacks¹⁵. CryoSieve¹⁵, an innovative iterative method for particle selection, can be applied to enhance the quality of the final density map by significantly reducing the number of particles. While other particle sorting criteria and software, such as the normalized cross-correlation (NCC) method¹⁶, the angular graph consistency (AGC) approach¹⁷, and non-alignment classification⁵, are currently in use within the field, this method has been shown to outperform these algorithms in terms of effectiveness.

In this study, we present a detailed guide to the entire process. As a case study, we applied this new method to the dataset of the influenza hemagglutinin trimer (EMPIAR entry: 10097)¹⁸, which includes 130,000 particles in its final stack. Our procedure successfully discarded about 73.8% of the particles from the final stack of this dataset, improving the resolution of the reconstructed density map from 4.11 Å to 3.62 Å. In addition to the influenza hemagglutinin trimer, results from multiple datasets are presented in earlier publication¹⁵, showcasing a variety of resolutions and molecular weights of biomolecules.

Protocol

1. Installation

1. Check and configure GPU-acceleration environment
 1. Open the terminal and enter the command: `nvidia-smi`. Make sure the command successfully displays

all information about the GPU card(s) and the CUDA version is higher than 10.2. Execute the command: `conda -V` to check if Conda is installed (**Supplementary Figure 1**).

2. Configure virtual environment

1. Enter the following command to set up the virtual environment, replacing `CRYOSIEVE_ENV` with your desired environment name: `conda create -n CRYOSIEVE_ENV python=3.8 cudatoolkit=10.2 cupy=10.0 pytorch=1.10 -c pytorch -c conda-forge`. Wait for a few minutes until the environment is successfully configured (**Supplementary Figure 2**).

NOTE: Users have the flexibility to modify the environment name as needed. The provided command is specific to CUDA 10.2. If a different CUDA version is desired, adjust the version number for `cudatoolkit`.

3. Install CryoSieve

1. Activate the environment by executing the command: `conda activate CRYOSIEVE_ENV`. Install the software by running: `pip install cryosieve` or `conda install -c mxhulab cryosieve` (**Supplementary Figure 3**). Enter `cryosieve -h` and ensure that help information is correctly displayed (**Supplementary Figure 4**).

2. Particle sieving

1. Retrieve the data

1. Download the EMPIAR-10097 final stack dataset from EMPIAR (see **Table of Materials**). Download the star file, the mask file (`mask.mrc`) and the initial model (for the re-estimate step; `initial.mrc`) from

Github (see **Table of Materials**). Place all these files in a folder together (**Supplementary Figure 5**).

NOTE: The repository at <https://github.com/mxhulab/cryosieve-demos> utilizes Git Large File Storage (Git LFS). Installing Git LFS is essential for cloning the entire repository. Alternatively, access the file via the GitHub link and click the **Download raw file** button to download an individual file.

2. Process particle sieving

1. Open the terminal and use the command: `cd FILEPATH` to navigate to the folder where the dataset is located. Activate the Conda environment by: `conda activate CRYOSIEVE_ENV`.
2. Enter the following command to start our particle sieving experiment: `cryosieve --reconstruct_software relion_reconstruct --postprocess_software relion_postprocess --i T40_HA_130K-Equalized_run-data_CryoSPARC_refined.star --o output/ --mask mask.mrc --angpix 1.3099979 --num_iters 10 --frequency_start 40 --frequency_end 3 --retention_ratio 0.8 --sym C3 --num_gpus 1 --balance` (**Supplementary Figure 5**). During the execution, the terminal will display the output logs for each iteration.

NOTE: Detailed instructions for each option can be found in **Supplementary File 1**. Processing time and the minimum requirements for execution are detailed in **Supplementary File 2**. `T40_HA_130K-Equalized_run-data_CryoSPARC_refined.star` was refined by CryoSPARC from `T40_HA_130K-Equalized_run-data.star` (downloaded from EMPIAR) to mitigate the effects brought about by advancements in orientation estimation techniques.

3. Finding the optimal iteration

1. Check resolutions

1. Use the command: `grep "+ FINAL RESOLUTION:" output/_postprocess*.txt` to print resolution results for the 10 iterations of sieving (**Figure 1**). Since the particle stack filtered in the 7th iteration has the highest resolution with the fewest particles, it is likely to provide the optimal result.

NOTE: To avoid unintentional information transfer from discarded to retained particles¹⁵ and to ensure that the particle stack posts the 7th iteration is indeed optimal, users are required to execute a re-estimation step for nearby iterations. In this protocol, iterations 4, 5, 6, 7, and 8 are subjected to verification.

2. Import sieved particles

1. Open the CryoSPARC web interface and follow these steps: Enter a workspace and click on the **Builder** button at the top right of the panel. In the panel, select and click on the **Import Particle Stack** option. In the Parameters section of the Particle Stack Import panel, specify the Particle meta path as the `_iter{n}.star` file located in the output folder of the completed results and the particle data path to the folder where the mrcs file is stored. Click the **Queue Job** button, then click the **Queue** button to initiate the process. Use the same way to import the remaining iterations that need re-estimation (**Supplementary Figure 6A**).

3. Import initial model

1. Click on the **Builder** button at the top right of the panel. In the panel, select and click on the **Import 3D Volumes** option.
2. Specify the Volume data path as the initial.mrc file. Click the **Queue Job** button, then click the **Queue** button to initiate the process (**Supplementary Figure 6B**).

NOTE: The initial model can also be generated through *ab initio* reconstruction (**Supplementary File 3**).

4. Homogeneous refinement (Build job)

1. Click on the **Builder** button at the top right of the panel. In the panel, select and click on the **Homogeneous Refinement** option.

NOTE: Non-uniform refinement is also applicable.

5. Homogeneous refinement (Import particles)

1. In the main panel on the left, open the job for importing the particle stack of the 5th iteration (or the desired iteration). Drag the imported particles module from the right side of the main panel and drop it into the Particle stacks section of the Builder on the right. Close the Import Particle Stack job by clicking on the red **X** in the top right corner of the main panel.
2. Open the job for importing 3D volumes. Drag the imported volumes module from the right side of the main panel and drop it into the Initial volume section of the Builder on the right.

6. Homogeneous refinement (Modify the parameters)

1. Under the Parameters fold, locate the Symmetry option and set it to C3. Find the Force re-do GS split option and disable it. Click the **Queue Job**

button, then click the **Queue** button to initiate the Homogeneous Refinement. Perform Homogeneous Refinement for the remaining iterations using the same method (**Supplementary Figure 6C-D**).

NOTE: The Force re-do GS split option is critical. Disabling this option ensures that CryoSPARC retains the gold-standard split given by the star file, avoiding overfitting. A detailed rationale for disabling Force Re-do GS Split can be found in **Supplementary File 4**.

7. Wait until all the jobs finish running to obtain the results. Based on the results, it is confirmed that the particle stack filtered in the 6th iteration is the actual optimal result.

NOTE: It is normal for the results obtained to have minor random deviations from the results provided in this protocol. These deviations do not affect the overall conclusion.

Representative Results

In this protocol, we utilized the influenza hemagglutinin trimer dataset (EMPIAR entry: 10097) as a demonstration of the efficacy of this process. Due to the preferred orientation of the sample, data acquisition required tilting at 40°. The protein exhibits C3 symmetry and has a molecular weight of 150 kDa.

We have implemented the protocol described earlier to process the final particle stack. It progressively removed

20% of the particles in each iteration, resulting in a retention ratio of 80.0%, 64.0%, 51.2%, and so on. As depicted in **Figure 1** and **Figure 2**, the resolution of the retained particles initially improved but eventually decreased. Among the iterations, the 6th iteration was identified as the most optimal subset, containing the fewest particles yet achieving the highest resolution. Our algorithm successfully identified a subset of particles comprising only 26.2% of the original stack, resulting in an improved resolution from 4.19 Å to 3.62 Å (re-estimated by CryoSPARC), shown in **Figure 2**. Furthermore, density maps before and after using CryoSieve were compared in **Figure 3**. Model-to-map Fourier Shell Correlation (FSC) curve and half-maps FSC curve of the reconstructed density maps before and after the method are also shown (**Figure 3A-B**). Raw density maps and sharp density maps obtained were also compared, with the equivalent contour level applied (**Figure 3C**). The side chains of sharp density maps were compared, showing the enhancement of reconstructed density maps. The estimated Rosenthal-Henderson B-factor was also adopted for the criteria of particle quality¹⁹. After removing the majority of particles in the final stack, the Rosenthal-Henderson B-factor raised from 226.9 Å² to 146.2 Å² (**Figure 3D**). Local resolution, local B-factor²⁰, and ResLog²¹ were also utilized for comparison, indicating that CryoSieve indeed enhances both the quality of the density maps and the particles (**Figure 4**).

```

(CRYOSIEVE_ENV) [humx@gpu05 cryosieve_jove]$ grep "+ FINAL RESOLUTION:" output/_postprocess*.txt
output/_postprocess_iter0.txt: + FINAL RESOLUTION: 4.19199
output/_postprocess_iter1.txt: + FINAL RESOLUTION: 4.14024
output/_postprocess_iter2.txt: + FINAL RESOLUTION: 4.14024
output/_postprocess_iter3.txt: + FINAL RESOLUTION: 4.08975
output/_postprocess_iter4.txt: + FINAL RESOLUTION: 4.08975
output/_postprocess_iter5.txt: + FINAL RESOLUTION: 4.04048
output/_postprocess_iter6.txt: + FINAL RESOLUTION: 4.04048
output/_postprocess_iter7.txt: + FINAL RESOLUTION: 3.99237
output/_postprocess_iter8.txt: + FINAL RESOLUTION: 4.04048
output/_postprocess_iter9.txt: + FINAL RESOLUTION: 4.04048
(CRYOSIEVE_ENV) [humx@gpu05 cryosieve_jove]$
  
```

Figure 1: Resolutions of each iteration. Resolutions that have been reported are highlighted in red boxes. [Please click here to view a larger version of this figure.](#)

#	Project	Job	Type	Info	Status	Elapsed	Lane	Priority
1	P33	36	Import Particle Stack	53,248 Particles 1.31 Å, 300 kV, 256 x 256	Completed	14 days ago	00m 21s	0
2	P33	39	Import Particle Stack	42,998 Particles 1.31 Å, 300 kV, 256 x 256	Completed	14 days ago	00m 19s	0
3	P33	312	Import Particle Stack	34,078 Particles 1.31 Å, 300 kV, 256 x 256	Completed	14 days ago	00m 17s	0
4	P33	315	Import Particle Stack	27,262 Particles 1.31 Å, 300 kV, 256 x 256	Completed	14 days ago	00m 16s	0
5	P33	318	Import Particle Stack	25,810 Particles 1.31 Å, 300 kV, 256 x 256	Completed	14 days ago	00m 15s	0
6	P33	326	Import 3D Volumes		Completed	14 days ago	00m 04s	0
7	P33	327	Homogeneous Refinement	53,248 Particles: 3.77 Å C3	Completed	14 days ago	09m 55s	SP_V100 0
8	P33	328	Homogeneous Refinement	42,998 Particles: 3.77 Å C3	Completed	14 days ago	08m 11s	SP_V100 0
9	P33	329	Homogeneous Refinement	34,078 Particles: 3.66 Å C3	Completed	14 days ago	06m 47s	SP_V100 0
10	P33	338	Homogeneous Refinement	27,262 Particles: 3.79 Å C3	Completed	14 days ago	06m 58s	SP_V100 0
11	P33	331	Homogeneous Refinement	25,810 Particles: 3.77 Å C3	Completed	14 days ago	05m 45s	SP_V100 0

Figure 2: Resolutions of each iteration. Resolutions identified by homogeneous refinement jobs are highlighted in red boxes. [Please click here to view a larger version of this figure.](#)

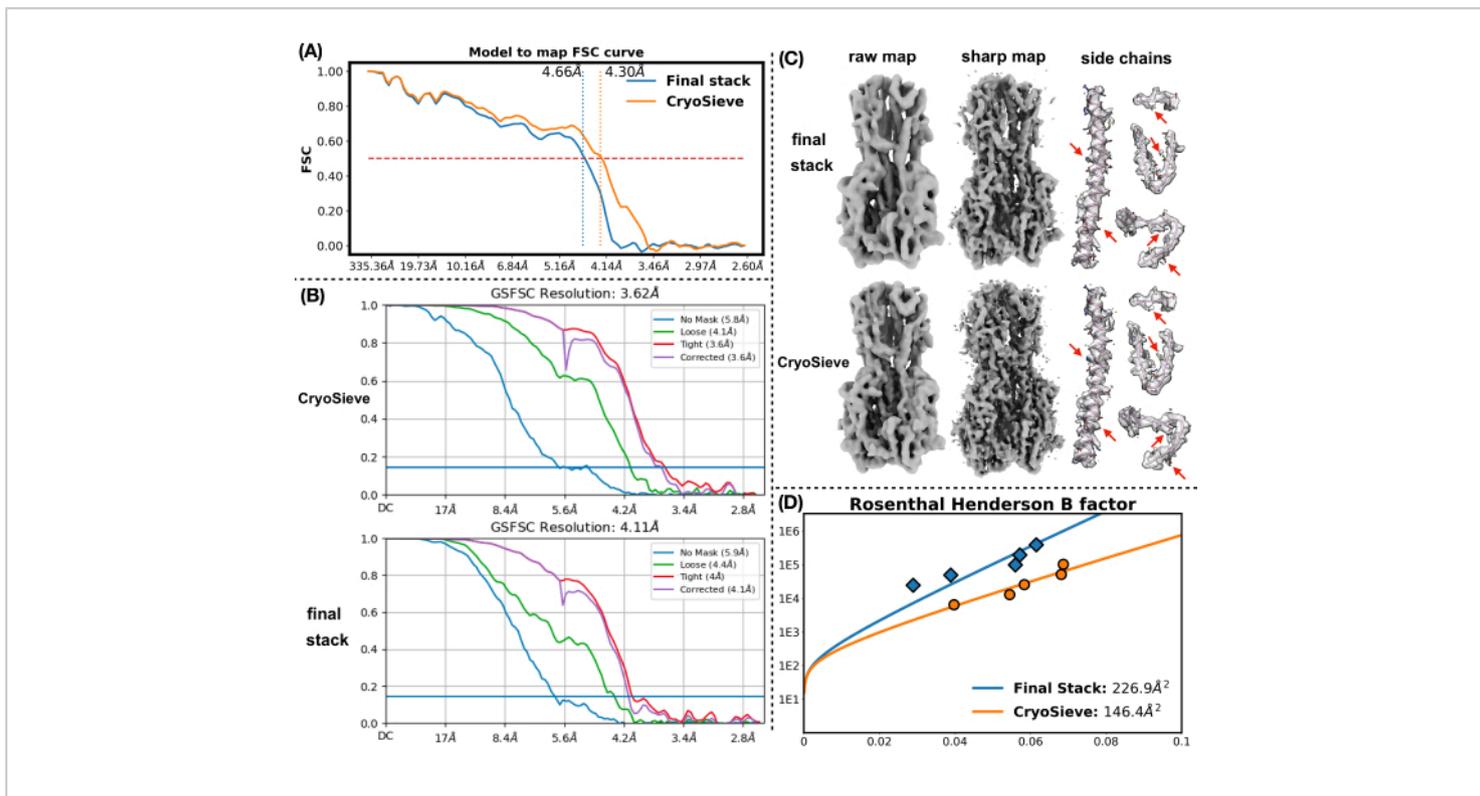


Figure 3: Density maps. (A) Comparison of model-to-map FSC curve of reconstructed density maps before and after using CryoSieve. The y-axis represents FSC, while the x-axis represents resolution. The red dashed line marks the threshold of 0.5 for the FSC. The vertical dashed line illustrates the resolution of the density maps obtained under a threshold of 0.5. (B) Half-maps FSC curve were obtained from reconstructed density maps before and after using CryoSieve via CryoSPARC. The y-axis represents FSC, while the x-axis represents resolution. (C) Raw density maps and sharp density maps were shown for both the CryoSieve-retained particles and the complete set of particles in the final stacks. The equivalent contour level of 0.65 was applied for raw density maps. The equivalent contour level of 0.84 was applied for sharp density maps. Sharp density maps were directly obtained by CryoSPARC. The sharp density maps were auto-postprocessed, first FSC-weighted (based on FSCs given by CryoSPARC). Then, the B-factor was sharpened using the auto-determined B-factors (232.0 \AA^2 for all particles in the final stack and 160.8 \AA^2 for CryoSieve). The side chains in the sharp density maps were compared, incorporating atomic models for reference. Red arrows highlight the improved regions. (D) The estimated Rosenthal-Henderson B-factor was shown for both the CryoSieve-retained particles and the complete set of particles in the final stacks. The y-axis represents the number of particles used, and the x-axis represents the reciprocal of the square of the resolution. Moving from top to bottom, each point represents half the particles of the previous one. The resolutions were determined by refinement. B-factors were determined using a least-squares approximation of the measured points, as shown by the fitting curves. The estimated Rosenthal and Henderson's B-factors are indicated in the legends: orange represents particles

retained by CryoSieve, while blue denotes all particles in the final stack. [Please click here to view a larger version of this figure.](#)

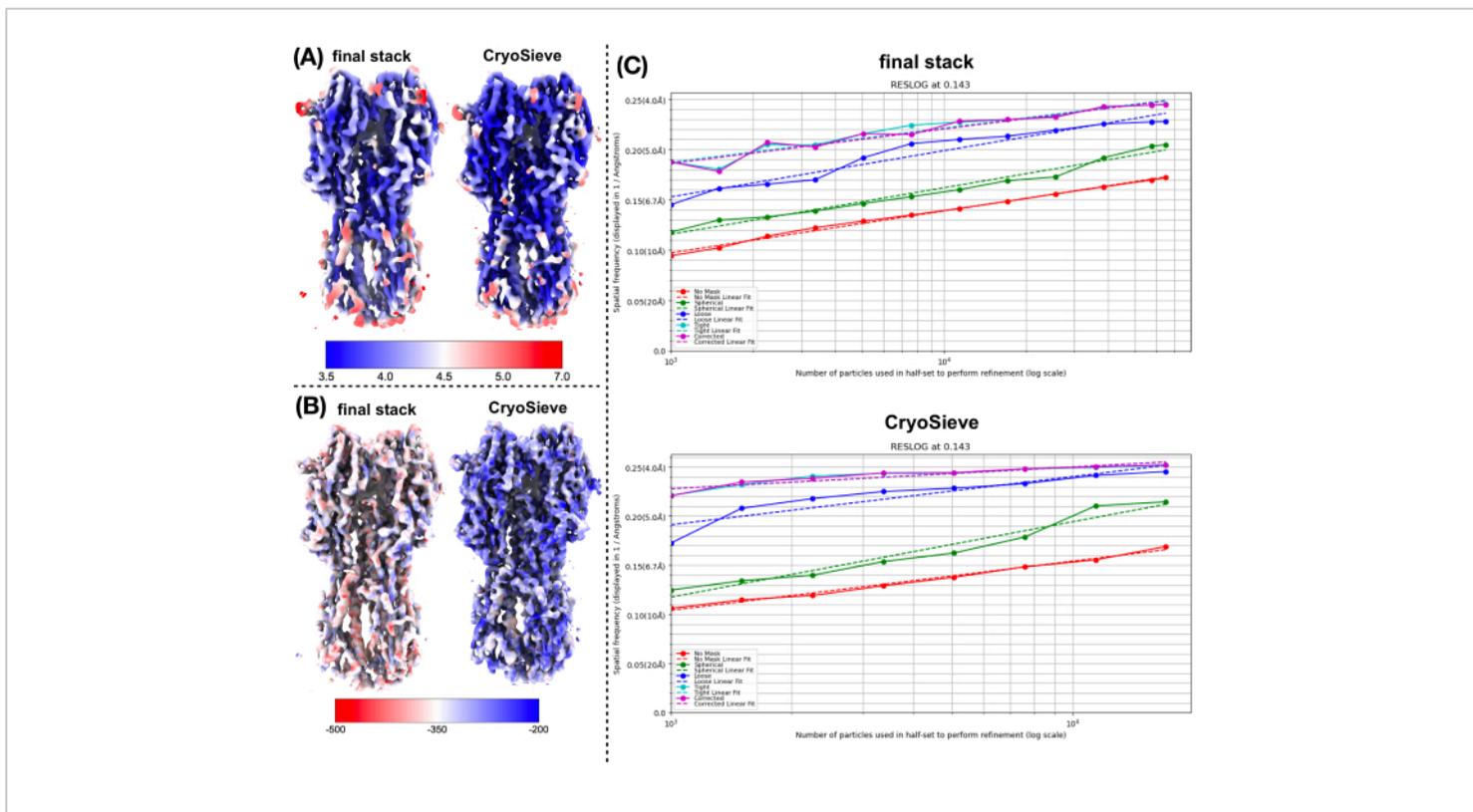


Figure 4: Comparison of various metrics of density maps. (A) Comparison of local resolution maps before and after using CryoSieve obtained by CryoSPARC. The local resolution ranges between 7 Å (red) and 3.5 Å (blue). (B) Comparison of density maps before and after using CryoSieve, colored with the local B-factor map obtained by LocBFactor using a resolution range of [20-3.5] Å. (C), Comparison of ResLog plots before and after using CryoSieve obtained by CryoSPARC. [Please click here to view a larger version of this figure.](#)

Supplementary Figure 1: Using commands nvidia-smi and conda -V to verify the prerequisites. If the prerequisites are met, typing the command nvidia-smi will display the GPU driver version, the CUDA version, and the status of the GPU cards. Similarly, entering the command conda -V should correctly display the installed version of Conda. [Please click here to download this File.](#)

Supplementary Figure 2: The process of creating new GPU-acceleration environments. The screen displays the

output generated by the command used to create the Conda environment. [Please click here to download this File.](#)

Supplementary Figure 3: Installation of CryoSieve in the GPU-acceleration environment. After activating the newly created Conda environment, the screen displays the output resulting from executing the command to install CryoSieve using Pip. [Please click here to download this File.](#)

Supplementary Figure 4: Help information. Please click [here to download this File](#).

Supplementary Figure 5: Running process. Upon executing CryoSieve through the command line, the screen then displays information regarding the running process. Please click [here to download this File](#).

Supplementary Figure 6: The configuration of CryoSPARC's jobs. (A) Import particle stack. (B) Import 3D volumes. (C-D) Homogeneous refinement. Please click [here to download this File](#).

Supplementary File 1: Options of CryoSieve. Please click [here to download this File](#).

Supplementary File 2: Processing time and minimal requirement for running Cryosieve. Please click [here to download this File](#).

Supplementary File 3: Generation of initial model by CryoSPARC. Please click [here to download this File](#).

Supplementary File 4: Rationale for disabling force re-do GS split. Please click [here to download this File](#).

Supplementary File 5: Options of cryosieve-csrefine. Please click [here to download this File](#).

Supplementary File 6: Options of cryosieve-csrhbfactor. Please click [here to download this File](#).

Discussion

Cryo-EM stands as a pivotal technique for elucidating the structures of biological molecules. In this process, after data collection via microscopy, particle extraction from micrographs is essential, followed by their classification in multiple stages to compile the final stack. A common challenge is the predominance of damaged or undesirably

conformed particles, underscoring the need for repeated particle selection to attain high-resolution density maps. This makes particle selection a critical step in cryo-EM SPA for achieving high-quality density maps. Existing particle selection techniques include the statistical non-tilt validation algorithm²², the z-score-based approach²³, and the angular accuracy estimation method²⁴.

CryoSieve emerges as a valuable tool in this context, adept at eliminating a significant number of extraneous particles from the final stack. This reduction not only enhances the reconstruction's computational efficiency but also streamlines the process. It offers a comprehensive suite for particle selection, where the extent of particle discard and the consequent improvement in resolution largely hinge on the initial data quality and the methodologies employed in data processing.

In this manuscript, we have presented a complete workflow of particle sieving using the real case dataset of influenza hemagglutinin trimer (EMPIAR entry: 10097). The steps covered and discussed here can be summarized as particle sieving and pose re-estimation. The final 3D reconstructed volume achieved a resolution of 3.62 Å, and side chains in alpha-helices were clearer in the post-processed volume compared to the published density map.

CryoSieve is an open-source method which is available on GitHub (<https://github.com/mxhulab/cryosieve>). A detailed tutorial can also be found on its homepage. Users can install and use it by following the tutorial. Additionally, two modules, cryosieve-csrefine and cryosieve-csrhbfactor, are provided. The cryosieve-csrefine module is specifically crafted to automate the sequential execution of various operations within CryoSPARC (**Supplementary File 5**). These operations include importing particle stacks and conducting

ab initio, homogeneous refinement, or non-uniform refinement jobs. On the other hand, the cryosieve-csrhbfactor module is designed to automate the determination of the Rosenthal-Henderson B-factor by leveraging the capabilities of cryosieve-csrefine (**Supplementary File 6**).

Presently, this method's application is confined to single conformation scenarios. Consequently, in instances where particles represent multiple conformations, their capabilities are limited. Users are advised to initially engage in 3D classification to segregate particles of disparate conformations before employing it for refined particle selection. Moreover, although the method demonstrates proficiency in filtering out over 50% of particles from the final stack, the origins of these discarded particles and the underlying reasons for their negligible contribution to reconstruction quality remain unclear. This gap in understanding necessitates additional research to comprehensively address and potentially rectify this limitation.

There are three possible existing methods of particle sorting or particle sieving. First of all, cisTEM⁴ can report a score for each single particle image after 3D refinement. Users could sort particles using the cisTEM score to discard particles. The angular graph consistency (AGC) approach¹⁷ is also a method to discard misaligned particles. Furthermore, the non-alignment classification⁵ is a traditional way to discard particles using 3D classification. We compared the quality of particles retained by these methods with CryoSieve and found that the retained particles of CryoSieve are of higher quality¹⁵. The method presented here significantly outperforms alternative methods and achieves the smallest number of particles at the same resolution.

As demonstrated in the result, the majority of particles in a cryo-EM final stack do not contribute to density map reconstruction. In other words, among all particles gathered during image acquisition, only a select few, namely the finest subset, actually contribute to the final reconstruction. Consequently, the ratio of this final subset to the total number of collected particles could serve as a quantitative metric for assessing sample quality. The higher this ratio, the better the sample quality. Despite technical advancements that have made cryo-EM more accessible to structural biologists, sample preparation remains a major bottleneck in the workflow. Scientists and engineers are thus focusing their efforts on this challenge²⁵. In single-particle analysis (SPA), sample preparation consists of two crucial steps: sample optimization and grid preparation. The former involves purifying the specimen while maintaining its optimal biochemical state. The latter entails preparing the sample for analysis in the microscope, including chemical or plasma treatment of the grid, sample deposition, and vitrification. Numerous techniques have been proposed to address macromolecular instability, but the efficacy of one approach over another depends on the sample's characteristics^{25,26}. Currently, grid preparation results are heavily influenced by the user's expertise and experience, which can make the process time-consuming and challenging^{27,28}. The numerous variables encountered in sample and grid preparation pose challenges in establishing cause-and-effect relationships, as researchers can only assess the sample at the molecular level using the microscope. As a result, quantitative statistics from comparisons of different sample and grid preparation protocols are still lacking, and a systematic approach is necessary to investigate trends and comprehend the fundamental mechanisms of sample behavior²⁹.

Disclosures

All other authors declare no competing interests.

Acknowledgments

This work was supported by Shenzhen Academy of Research and Translation (to M.H.), the Advanced Innovation Center for Structural Biology (to M.H.), the Beijing Frontier Research Center for Biological Structure (to M.H.), the National Key R&D Program of China (No.2021YFA1001300) (to C.B.), the National Natural Science Foundation of China (No.12271291) (to C.B.), and the National Natural Science Foundation of China (No.12071244) (to Z.S.).

References

- Bai, X. C., Fernandez, I. S., McMullan, G., Scheres, S. H. Ribosome structures to near-atomic resolution from thirty thousand cryo-em particles. *elife*. **2**, e00461 (2013).
- Campbell, M. G. et al. Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure*. **20** (11), 1823-1828 (2012).
- Li, X. et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-em. *Nat Meth*. **10** (6), 584-590 (2013).
- Grant, T., Rohou, A., Grigorieff, N. Cis tem, user-friendly software for single-particle image processing. *elife*. **7**, e35383 (2018).
- Scheres, S. H. Relion: Implementation of a bayesian approach to cryo-em structure determination. *J Str Biol*. **180** (3), 519-530 (2012).
- Punjani, A., Rubinstein, J. L., Fleet, D. J., Brubaker, M. A. Cryosparc: Algorithms for rapid unsupervised cryo-em structure determination. *Nat Meth*. **14** (3), 290-296 (2017).
- Kühlbrandt, W. The resolution revolution. *Science*. **343** (6178), 1443-1444 (2014).
- Dubochet, J. et al. Cryo-electron microscopy of vitrified specimens. *Quart Rev Biophys*. **21** (2), 129-228 (1988).
- Wagner, T. et al. Sphire-cryolo is a fast and accurate fully automated particle picker for cryo-EM. *Comm Biol*. **2** (1), 218 (2019).
- Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Meth*. **16** (11), 1153-1160 (2019).
- Wang, F. et al. Deeppicker: A deep learning approach for fully automated particle picking in cryo-em. *J Str Biol*. **195** (3), 325-336 (2016).
- Heimowitz, A., Andén, J., Singer, A. Apple picker: Automatic particle picking, a low-effort cryo-em framework. *J Str Biol*. **204** (2), 215-227 (2018).
- Glaeser, R. M. How good can single-particle cryo-em become? What remains before it approaches its physical limits? *Ann Rev Biophys*. **48**, 45-61 (2019).
- Diiorio, M. C. Kulczyk, A. W. A robust single-particle cryo-electron microscopy (cryo-em) processing workflow with cryosparc, relion, and scipion. *J Vis Exp*. (179), e63387 (2022).
- Zhu, J. et al. A minority of final stacks yields superior amplitude in single-particle cryo-em. *Nat Comm*. **14** (1), 7822 (2023).
- Zhou, Y., Moscovich, A., Bendory, T., Bartesaghi, A. Unsupervised particle sorting for high-resolution single-particle cryo-em. *Inv Probl*. **36** (4), 044002 (2020).

17. Méndez, J., Garduno, E., Carazo, J. M., Sorzano, C. O. S. Identification of incorrectly oriented particles in cryo-em single particle analysis. *J Str Biol.* **213** (3), 107771 (2021).
18. Tan, Y. Z. et al. Addressing preferred specimen orientation in single-particle cryo-em through tilting. *Nat Meth.* **14** (8), 793-796 (2017).
19. Rosenthal, P. B. Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol.* **333** (4), 721-745 (2003).
20. Kaur, S. et al. Local computational methods to improve the interpretability and analysis of cryo-em maps. *Nat Comm.* **12** (1), 1240 (2021).
21. Stagg, S. M., Noble, A. J., Spilman, M., Chapman, M. S. Reslog plots as an empirical metric of the quality of cryo-em reconstructions. *J Str Biol.* **185** (3), 418-426 (2014).
22. Vargas, J., Otón, J., Marabini, R., Carazo, J. M., Sorzano, C. Particle alignment reliability in single particle electron cryomicroscopy: A general approach. *Sci Rep.* **6** (1), 21626 (2016).
23. Vargas, J. et al. Particle quality assessment and sorting for automatic and semiautomatic particle-picking techniques. *J Str Biol.* **183** (3), 342-353 (2013).
24. Vargas, J., Melero, R., Gomez-Blanco, J., Carazo, J.-M., Sorzano, C. O. S. Quantitative analysis of 3d alignment quality: Its impact on soft-validation, particle pruning and homogeneity analysis. *Sci Rep.* **7** (1), 6307 (2017).
25. Carragher, B. et al. Current outcomes when optimizing 'standard' sample preparation for single-particle cryo-em. *J Microsc.* **276** (1), 39-45 (2019).
26. Drulyte, I. et al. Approaches to altering particle distributions in cryo-electron microscopy sample preparation. *Acta Crystallographica Sec D: Str Biol.* **74** (6), 560-571 (2018).
27. Glaeser, R. M. How good can cryo-em become? *Nat Meth.* **13** (1), 28-32 (2016).
28. Kim, L. Y. et al. Benchmarking cryo-em single particle analysis workflow. *Front Mol Biosci.* **5**, 50 (2018).
29. Weissenberger, G., Henderikx, R. J., Peters, P. J. Understanding the invisible hands of sample preparation for cryo-em. *Nat Meth.* **18** (5), 463-471 (2021).