2322
doi:10.1017/S1431927620021194
Microsc. Microanal. 26 (Suppl 2), 2020


# Algorithmic Advances in Single Particle Cryo-EM Data Processing Using CryoSPARC

Ali Punjani

University of Toronto / Structura Biotechnology Inc., Toronto, Ontario, Canada

Single particle cryo-EM (cryo-electron microscopy) allows high-resolution imaging of macromolecular complexes in close-to-native state, at near-atomic resolutions. Cryo-EM has undergone several technological breakthroughs in microscopy, electron detectors, and image processing that have enabled its use recently in solving high-resolution structures of difficult proteins and complexes. As cryo-EM makes rapid progress, several key challenges remain in the quest for higher resolutions on challenging targets, as well as in the widespread and routine use of cryo-EM. Here, we introduce new methods in the cryo-EM data processing pipeline, along with implementations within the cryoSPARC software system [1], to address some of these challenges.

The image processing pipeline in single particle cryo-EM is required to solve the 3D electron density of a target molecule, in potentially many conformational states, from noisy 2D images collected using cryo-TEM. Each collected image is a movie of dose-fractionated frames that require motion estimation and correction. The corrected images (micrographs) are then used to estimate the microscope CTF during the exposure, as well as to find and pick out single particles. The single particles are extracted from the micrographs, and then are sorted and filtered using 2D classification methods. The resulting filtered particle stacks are used to perform *ab initio* 3D structure determination of potentially multiple discrete states or targets. These coarse structures are then further classified and refined in 3D to yield interpretable molecular density maps and achieve state of the art resolutions.

In this work, we introduce a new software system, cryoSPARC Live, that is built on top of the cryoSPARC software package, including modified reconstruction algorithms, data flow patterns, and GPU implementations, that together enable real time 3D reconstruction of single particle structures, at high resolutions, faster than data collection. Many projects (e.g., [2, 3, 4, 5]) have been created to address the challenge of processing EM data during collection. CryoSPARC Live is the first to enable real-time streaming 2D classification, 3D classification, and 3D refinement to high resolution.

CryoSPARC Live enables the real-time monitoring of sample and image quality, and optimization of data collection parameters and strategies. It also dramatically shortens the total time required to obtain a high resolution refined 3D structure, and is especially useful for centralized facilities that have many users collecting data on several instruments.

We show that cryoSPARC Live, using new methods for motion correction, CTF estimation, particle picking, streaming real-time 2D classification, 3D classification, and streaming refinement, can achieve state of the art resolutions for difficult datasets, including tilted data collection, in real time. We show results that have been obtained in our testing, as well as in production by centralized facilities operating on external user data.

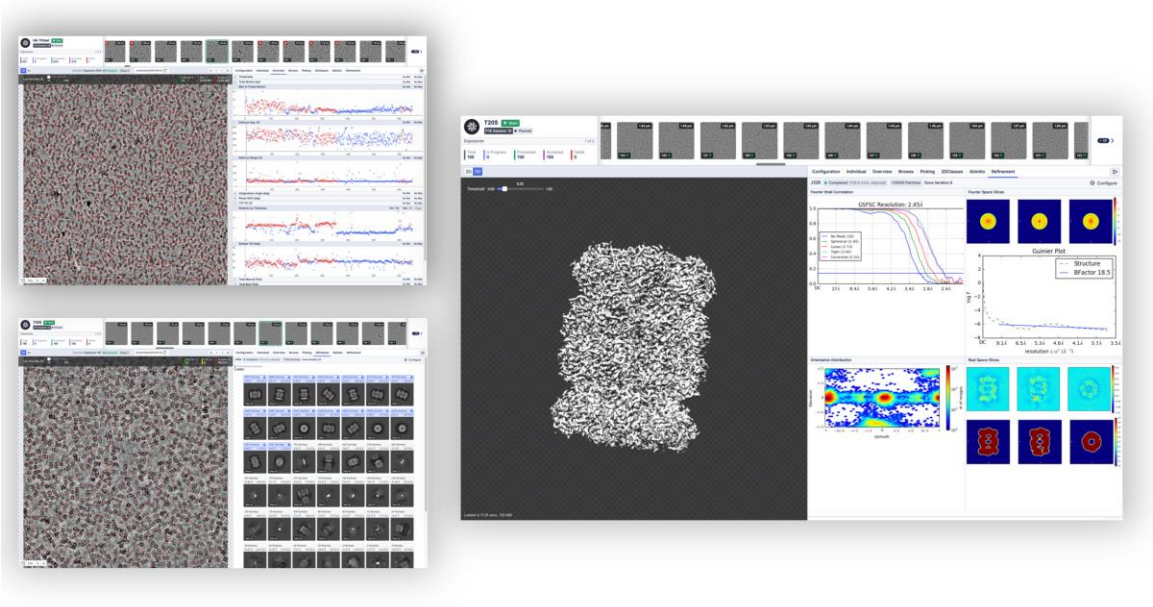

**Figure 1.** CryoSPARC Live performs real-time preprocessing (top left), streaming 2D classification (bottom left) and real-time streaming 3D refinement (right) faster than data collection for single particle cryo-EM.

References

[1] A Punjani *et al*, Nature Methods **14** (2017), p. 290–296
[2] D Tegunov & P Cramer, Nature Methods **16** (2019)
[3] J Gomez-Blanco *et al*, Journal of Structural Biology (2018)
[4] J Elferich *et al*, Microscopy and Microanalysis 25 (2019)
[5] N Biyani *et al*, Journal of Structural Biology (2017)

https://doi.org/10.1017/S1431927620021194 Published online by Cambridge University Press