

Real-time cryo-EM structure determination for drug discovery

Ali Punjani

Structura Biotechnology Inc. / University of Toronto, United States

Single particle cryo-EM (cryo-electron microscopy) allows high-resolution imaging of macromolecular complexes in close-to-native state, at near-atomic resolutions. This provides a revolutionary new source of structural information for drug discovery. As cryo-EM based drug discovery efforts evolve and develop, there is a strong need for powerful methods and robust tools that enable the entire workflow of cryo-EM data collection and data processing to fit within the timelines that are relevant in drug discovery.

In this work, we describe cryoSPARC Live [1], the first end-to-end real-time streaming 2D and 3D reconstruction pipeline for single particle cryo-EM that is already in production use at industry and academic institutions worldwide. Prior to the introduction of cryoSPARC Live, it was common for cryo-EM data processing to take weeks after a collection session was completed before any insight about the sample or structure could be extracted. With cryoSPARC Live, we demonstrate the ability for practitioners to generate real-time feedback about data quality, grid and sample choice and imaging condition choices, enabling go/no-go decisions within hours, and multiple high-resolution 3D refinements well within the typical 24-hour timeframe of a data collection session. We provide experimental results demonstrating these capabilities of the cryoSPARC Live software system, as well as example case studies of published structures solved as part of drug discovery efforts at institutions that use cryoSPARC Live.

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 describe how cryoSPARC Live, that is built on top of the cryoSPARC software package, includes 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. 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.

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)