

HOTSPUR: A Real-time Interactive Preprocessing System for Cryo-EM Data

Johannes Elferich¹, Rich Posert¹, Craig Yoshioka² and Eric Gouaux^{1,3*}

¹ Vollum Institute/Oregon Health and Science University, Portland, OR, USA.

² Biomedical Engineering/Oregon Health and Science University, Portland, OR, USA.

³ Howard Hughes Medical Institute, Portland, OR, USA.

* Corresponding author: gouauxe@ohsu.edu

Single-particle cryo-electron microscopy enables structure determination of biological molecules and complexes at near atomic resolution with unprecedented speed. Yet in many instances, acquisition of a sufficient number of particles requires collection of thousands of micrographs, collected using direct electron detectors and comprised of dose-fractionated “movie” frames. To achieve the highest resolution reconstructions, the user must search the grid for optimal regions, typically ones with thin ice and dense yet non overlapping particle distribution. While software packages such as SerialEM [1] or EPU facilitate automated image acquisition, user intervention is required to choose the best acquisition areas. However, images require preprocessing- motion correction and contrast-transfer function (CTF) estimation – before they can be visually inspected. In addition to enabling visual inspection of the images, preprocessing yields quantitative information on image quality. To allow one to judge image quality in near real time we developed a software package, called HOTSPUR, that performs preprocessing without user intervention and that presents the results in an easy-to-use interface. HOTSPUR allows researchers to rapidly judge data quality and to focus on decision-making to obtain the highest quality data from limited microscope time.

HOTSPUR has two components: a flexible processing system for micrographs that stores processing results in a database and a web interface driven by this database where the user can inspect results and annotate images. Processing starts with a module that monitors a data folder for images and directs them to the appropriate processing pipeline. The processing pipelines rely on established software, such as motioncor2 [2] for motion correction and Gctf [3] and CTFFIND4 [4] for CTF estimation. For particle picking we modified DOG-picker [5] to allow for interactive selection of parameters.

The web interface of HOTSPUR consists of multiple panels. Micrographs can be inspected in the “Micrograph” panel (Figure 1A). In addition to the micrograph this view also shows information regarding motion correction and CTF estimation. Based on these elements of information, users can annotate micrographs with a quality score. In the “Statistics” panel (Figure 1B) key values for every micrograph are plotted as a function of acquisition time. Individual micrographs are color coded according to user annotation. This overview is useful for detection of problems during data collection. The “Montage” panel (Figure 1C) shows low- and medium-magnification montages as used in SerialEM for data collection. The location of micrographs is annotated in the medium-magnification montage together with the user annotation of micrograph quality. Because holes for data collection are chosen based on their appearance in medium-magnification images, this can help to correlate appearance of holes to micrograph quality.

In addition to micrograph inspection and curation, HOTSPUR can also be used for manual or automated particle picking (Figure 1D). Automated particle picking using DOG-picker requires two parameters: particle size and intensity thresholds. Both of these can be adjusted in HOTSPUR, with immediate

feedback on the results. Manual picking can be used to amend the automated picking by adding picks or disabling automated picks. At any point during data collection an export program can create a particle stack based on the current micrograph curation and particle picking parameters.

At the OHSU Multiscale Microscopy Core roughly 300,000 micrographs have been processed using HOTSPUR between February 2017 and February 2019. This has allowed the structure determination of a wide variety of transmembrane proteins [6,7,8]. An additional benefit is that the web interface allows rapid access to the results of past microscope sessions. In summary, HOTSPUR allows users to completely focus on directing the microscope to the ideal areas of their EM grid while still having all relevant information at their disposal.

References:

- [1] DN Mastrorade, *Journal of Structural Biology* **152** (2005), p. 36.
- [2] SQ Zheng et al., *Nature Methods* **14** (2017), p. 331.
- [3] K Zhang, *Journal of Structural Biology* **193** (2016), p. 1.
- [4] A Rohou and N Grigorieff, *Journal of Structural Biology* **192** (2015), p. 216.
- [5] NR Voss et al., *Journal of Structural Biology* **166** (2009), p. 2015.
- [6] J Ge et al., *Elife* **7** (2018).
- [7] F Jalali-Yazdi et al., *Cell* **175** (2018), p. 1520.
- [8] N Yoder, C Yoshioka and E Gouaux. *Nature* **555** (2018), p. 397.

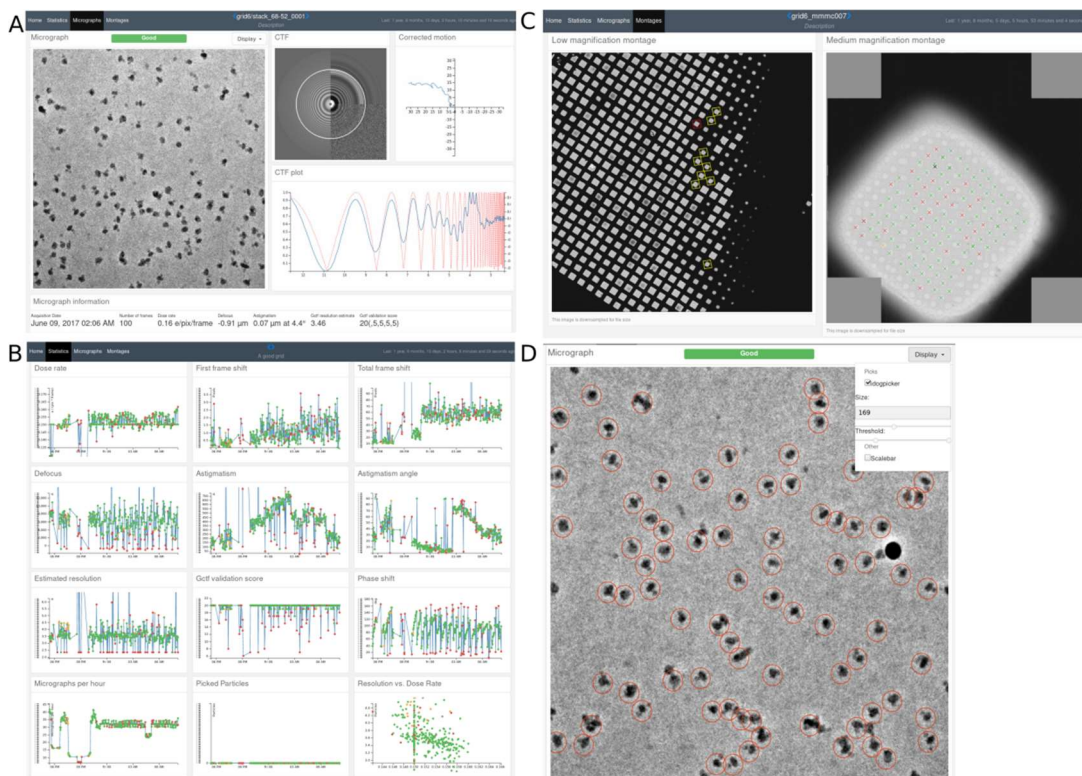


Figure 1. Screenshots of HOTSPUR web interface (A) Micrograph view (B) Statistics view (C) Montage view (D) Interactive particle picking