

## STOPGAP: A Software Package for Subtomogram Averaging and Refinement

William Wan<sup>1</sup>, Sagar Khavnekar<sup>2</sup>, Jonathan Wagner<sup>2</sup>, Philipp Erdmann<sup>2</sup> and Wolfgang Baumeister<sup>2</sup>

<sup>1</sup>Vanderbilt University, Nashville, Tennessee, United States, <sup>2</sup>MPI Biochemistry, Martinsried, Bayern, Germany

While cryo-electron microscopy (cryo-EM) and single particle analysis have become widely used techniques for structure determination, they are generally limited to specimens that consist of purified particles suspended in thin layers of ice. When molecules begin to overlap, single particle methods become unable to distinguish their individual features, preventing alignment and structure determination. This limitation precludes the analysis of many important biological systems, including membrane-associated assemblies, pleomorphic viruses, or cellular sections, where heterogeneous molecules exist in multi-layered mixtures. In such complex cases, the method of choice is cryo-electron tomography (cryo-ET), which allows for direct 3D reconstruction of the field of view, overcoming the problem of overlapping densities. For high-resolution structure determination, subtomogram averaging is an analogous approach to single particle methods, where repeating units in the tomographic data can be extracted, aligned, and averaged.

STOPGAP is an open-source, MATLAB-based, software package focused specifically on subtomogram averaging: i.e. particle picking via template matching, subtomogram alignment, and classification. It consists of a number of executables as well as a MATLAB toolbox for data analysis. The goal of STOPGAP is to provide a set of robust and efficient algorithms for subtomogram averaging. As part of its design philosophy, STOPGAP is not intended to be run as a black box, and provides users direct access to a large number of low-level parameters, allowing for more flexible workflows to optimize the algorithms for specific specimens or biological problems. However, many default parameters are also available, allowing for novice users to also use STOPGAP. At the core of STOPGAP is an optimized missing wedge filter, which allows for improved template matching, alignment scoring, and weighted averaging. Rescaling approaches have been added to improve computational speed, while additional approaches such as subset processing or stochastic hill climbing are also implemented to improve computational times. We find that STOPGAP is able to produce well-defined density maps while balancing efficient CPU usage.