

Beyond Ribosomes: In Situ Structural Biology of Diverse Targets in *C. reinhardtii*

Dr Ron Kelley¹, Dr Xianjun Zhang¹, Dr. Martin Obr¹, Mr. Sagar Khavnekar¹, Dr Jessica Heebner¹, Dr Saikat Chakraborty¹, Dr Ricardo Righetto², Dr Florent Waltz², Dr Alicia Michael², Dr Wojciech Wietrzynski², Dr Grigory Tagiltsev³, Prof John Briggs³, Prof Juergen Plitzko³, Dr Ben Engel², Dr. Abhay Kotecha¹
¹Thermo Fisher Scientific, Eindhoven, Netherlands, ²Biozentrum, University of Basel, Basel, Switzerland, ³Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, Martinsried, Germany

Understanding high resolution protein structure in the context of the whole cell is the vision of visual proteomics. With the advent of high-throughput cryo-electron tomography and cryo-FIB milling, paired with cutting edge computational techniques, achieving such an ambitious goal is no longer a far-reaching dream. The new generation of cryo-FIB from Thermo Fisher Scientific uses plasma to generate focus ion beam, which reduces redeposition and ion beam damage, substantially improving throughput. In preparation for a large-scale effort towards visual proteomics of the model organism *Chlamydomonas reinhardtii*, we have prepared enough lamellae to allow for acquisition of more than 1800 tomograms of different cellular compartments (Fig 1).

Initial results indicate that the data is of sufficient quality to achieve sub-nanometer resolution for selected complexes using a fraction of the dataset. Six tomograms from the same lamella were used to determine the structure of 80S ribosome at 6Å using template matching¹ and subtomogram averaging. Encouraged by this, a number of other targets were selected among the collaborators for further study such as ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO), ATP synthase, nucleosomes, photosystem II, and microtubules.

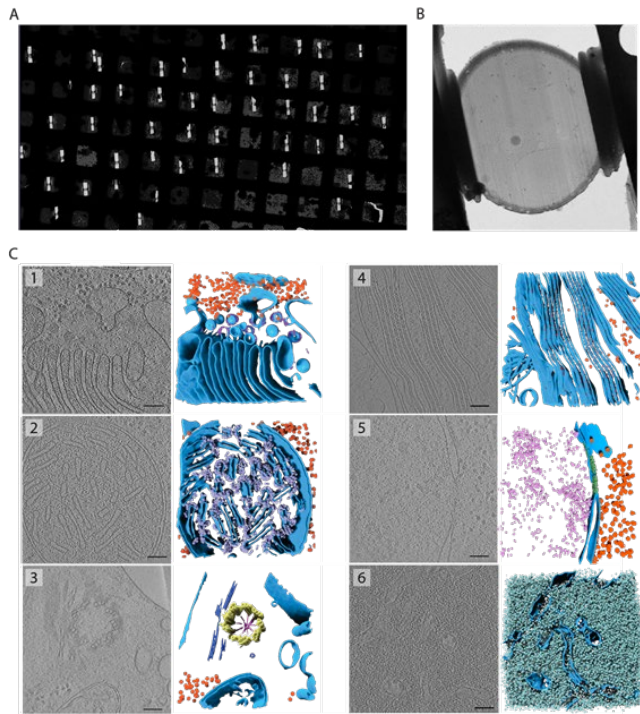
While some proteins were in such high abundance that template matching alone was sufficient to obtain a sub-nanometer structure, as in the case of Rubisco where we were able to determine a 7Å structure from within the pyrenoid compartment, other targets proved more of a challenge and necessitated a combination of computational techniques to achieve results. The denoising neural network cryoCARE was implemented on all datasets to increase template matching accuracy. In the case of microtubules, which are rare in *C. reinhardtii*, filament tracing in Amira (Thermo Fisher Scientific) was necessary to determine initial centerlines for reconstruction along individual protofilaments.

Photosystem II (PSII), which is present in very high abundance in the chloroplast, proved to be a challenging target as it is a membrane embedded

protein with only a small region protruding into the lumen of the thylakoid membrane. Multiple template matching methods failed to identify most of the visible particles and unfortunately noise2noise denoising tended to blur or erase the small densities along the membrane. We therefore employed a novel approach that paired two deep learning networks to generate PSII candidate coordinates which could be further refined for classification and subtomogram averaging. The first network is a regression UNet trained on purely synthetic data² to denoise cryo-tomographic data. The denoised data was then used to train a 2.5D UNet to segment membrane, ribosomes, and all densities that protrude into the thylakoid lumen as PSII (Dragonfly 2022.2, Object Research Systems). Coordinates (without any angular information) were extracted from the PSII class and used as particle picks, which were then aligned using the surrounding membrane. Work is ongoing, but initial results from 24 datasets are very promising.

The scale of this dataset is exciting, but the huge number of molecular complexes within living cells makes it difficult to identify, confirm the identity of, and determine each structure by just one group. Achieving a full visual proteome of *C. reinhardtii* will necessitate a large collaborative effort. To that end, we would like to create an open access database for *C. reinhardtii* to accelerate annotation and curation, enable further cell biology research, and develop new computational tools for in situ cryo-ET. Along with sharing the raw data, reconstructed tomograms, denoised datasets, and structural determinations, we will provide high quality segmentations of selected datasets created using 2.5D (Dragonfly) and 3D (MemBrain Seg3) UNets. This project has the potential to provide invaluable insights into cellular processes and will hopefully lay the foundation for other large-scale studies of other species.

Graphic:



Keywords:

CryoET, FIB-milling, plasmaFIB, visual proteomics

Reference:

1. Wan, W., Khavnekar, S. & Wagner, J. bioRXIV (2023).
<http://biorxiv.org/lookup/doi/10.1101/2023.12.20.572665> (2023)
doi:10.1101/2023.12.20.572665.
2. Purnell, C. et al. bioRXIV (2023).
<http://biorxiv.org/lookup/doi/10.1101/2023.04.28.538636> (2023)
doi:10.1101/2023.04.28.538636.
3. Lamm, L. et al. bioRXIV(2024).
<http://biorxiv.org/lookup/doi/10.1101/2024.01.05.574336> (2024)
doi:10.1101/2024.01.05.574336.