CoSeC SOFTWARE SPOTLIGHT

CCP-EM Doppio

https://gitlab.com/ccpem
Dr Tom Burnley (ccpem@stfc.ac.uk)

The CCP-EM Software Suite 2.0 (aka Doppio) integrates software tools from 15+ collaborating laboratories to enable researchers to solve high resolution 3D structures of proteins by combining many thousands of 2D projections from electron microscopy images.

Value for Research

CryoEM structures allow researchers understand diseases by revealing the 3D structures of the proteins, viruses, and other molecular machines that cause or drive them, often in multiple functional states. This is especially useful in areas like neurodegeneration, cancer, and infectious disease, where seeing the disease-linked structures can explain mechanisms and guide drug design. CryoEM uses large data volumes and sophisticated multi-stage image processing algorithms to generate the 3D structures. The Doppio software enables users to efficiently process the data and analyse the outputs.

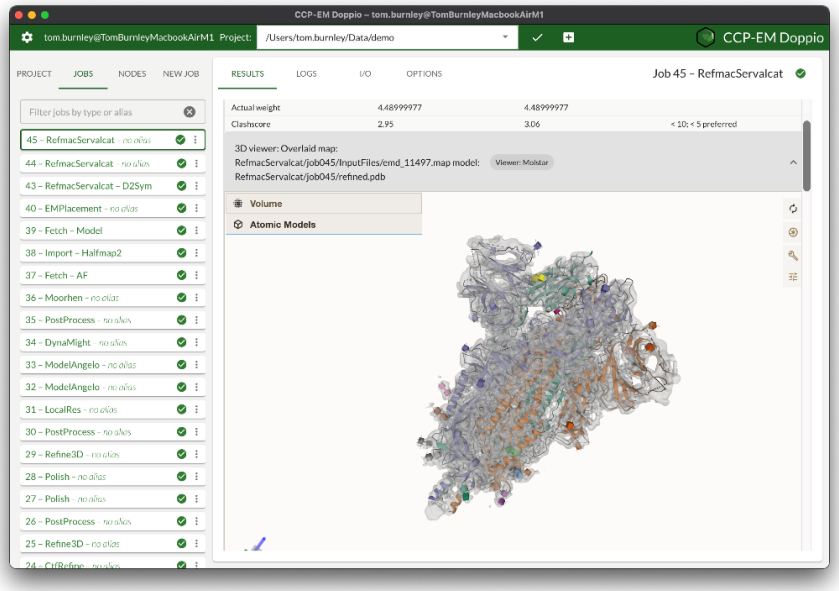
Development

CCP-EM was initiated in 2012 with core funding from the MRC. The project is based in the Molecular and Cellular Electron Microscopy Group in Scientific Computing of STFC. The development of Doppio was also supported by grants from BEIS, DSIT and the ALC. The user interface and workflow management tools were primarily developed by the MCEM group. This is used to integrate software tools developed by collaborating groups from multiple institutions including MRC-LMB, TU Delft, CSSB Hamburg and Universities of Cambridge, Liverpool, Princeton, Pittsburgh and Tokyo. The software was first released in 2024 and regular updates have been released since. CCP-EM software is citated in >600 public depositions in the EMDB and has multiple industrial license holders.

Functionality

The software provides an integrated pipeline for single-particle and subtomogram averaging cryoEM data processing, guiding users from raw micrographs through particle picking, classification, and high-resolution 3D reconstruction. It combines established tools such as RELION with machine-learning approaches (e.g. CryoDRGN) to improve particle selection and map quality. Downstream processing includes iterative

refinement, contrast transfer function and motion correction optimisation, and post-processing to produce high-resolution 3D density maps of proteins and other biomolecules. The pipeline then extends seamlessly into atomic model building and validation, integrating tools such as ModelAngelo, Servalcat, and Moorhen to generate and refine atomic models from cryoEM maps. The maps and models together enable researchers to understand proteins mechanism of action and can provide atomic models for structural based drug discovery.



CCP-EM 2.0: Doppio displaying Sars-Cov2 spike protein (PDB model 6zwv and EMDB entry emd11497

The Future

Future work will focus on improving the efficiency of data processing to accelerate structure determination, particularly for high-throughput facilities and pharmaceutical screening workflows. This includes streamlining pipelines, optimising resource usage, and reducing time-to-result without compromising map quality. In parallel, the software is being actively extended to further support cryo-electron tomography, enabling researchers to visualise macromolecular structures in their native cellular environments.

LICENCE

Academic, Micro-entity, Evaluation and Commercial

Programming Languages

Python, JavaScript, C++, Fortran

Repository URL

https://gitlab.com/ccpem

YouTube

https://www.youtube.com/playlist?list=PLo3ZBhRliQ92Wdbyi2mBwc-Ig_EpP1fzY