Grand Challenges

The Collaborative Computational Project for Electron cryo-Microscopy (CCP-EM)

In the next 5-10 years, CCP-EM will enable software development and large-scale computing support to create state-of-the-art CryoEM/T methods to develop life-changing drug discoveries

The Community

CCP-EM www.ccpem.ac.uk

The Collaborative Computational Project for Electron cryo-Microscopy (cryoEM) provides support in computational areas for users and developers in biological cryoEM. CCP-EM serves the worldwide cryoEM community by providing software and training to users, researchers and developers of cryoEM software; it also provides a focal point for the cryoEM community to interact with related communities via CoSeC. It's UKRI/STFC licensed software suite enables end users to efficiently process their data and solve structures whilst supporting the software developer community to develop and distribute novel tools using both classical and ML/AI algorithms.

The Challenge

To develop new computational tools and techniques to aid drug discovery, improve storage efficiency and speed of data processing

Cryogenic-sample Electron Microscopy (cryoEM) is one of the principal methods for solving structures of proteins, and other biomolecules, to near atomic resolution. The field has seen an exponential increase in use over the previous decade due to improvements in microscope hardware and image processing software, crowned by the award of the Nobel Prize for Chemistry in 2017. CryoEM is now a primary driver in structural biology and >£50M has been invested by many, including Wellcome, MRC, BBSRC to ensure access to essential high-end equipment for the UK life science sector. During the COVID-19 pandemic, cryoEM played an essential role in understanding the molecular basis of the disease, aiding the development of effective vaccines and other therapeutics. This method is now routinely used in novel drug discovery and allows biomolecules not amenable to other biophysical techniques to be studied, which is of particular importance for neurodegenerative diseases such as Alzheimer's and Parkinsons. CCP-EM serves the worldwide cryoEM community by providing software and training. As the field has expanded CCP-EM has taught over 1600 researchers from across the world in the last 10 years and the annual 'CCP-EM Spring Symposium' conference is now the principal cryoEM meeting in the UK with over 1000 delegates in 2025. Much of work described above has focused on studying macromolecules *in vitro* i.e. isolated from their cellular environment but in recent years breakthroughs in electron tomography (cryoET) have allowed researchers to

observe protein at high resolution in their native environment. This promises to bring a new dawn in drug discovery and, more broadly, fundamentally enhance our understanding of life at the molecular level. CCP-EM is actively developing computational tools in this area though there are many challenges to overcome, such as object identification and heterogenous reconstruction. CryoEM is highly data intensive image processing with large demands for compute power and storage. The field produces ~400Pb of data per year, a vast amount on par with the Square Kilometre Array. The Electron Bio-Imaging Centre (eBIC) at the Diamond Light Source (DLS) currently generates ~10Pb data per year and CCP-EM is helping develop new compression protocols to improve storage efficiency, which will be of benefit to the wider community. CCP-EM is also evolving GPU accelerated image processing routines to improve the speed of data processing and both initiatives will help reduce the carbon cost of data analysis and storage.

The Solution

Develop state-of-the-art computational support for essential biophysical methods

To keep its world leading position, the UK life science sector must have state-of-the-art computational support for essential biophysical methods. CCP-EM are targeting the following areas for development: 1) improved workflows for resolving protein:ligand complexes for drug discovery; 2) develop heterogenous reconstruction algorithms to yield molecular movies allowing researchers capture different states of molecular machines; 3) develop new methods for cryo-tomography using AI/ML where appropriate and 4) data/metadata management to handle the increasing data volumes, and ensure that database entries are well annotated and FAIR. CCP-EM will continue to support key facilities at Harwell Campus as well as its extensive training programme. Long-term external contributions are essential for CCP-EM, with an extensive network including amongst others, the Crick, MRC-LMB, EMBL-EBI, TU Delft and the Universities of Cambridge, Tokyo, and Princeton. CCP-EM works closely with the current related life science CCPs within CoSeC: CCP4, CCPN and CCPBioSim, which was recently strengthened by the DRIIMB (Digital Research Infrastructure for Integrated Molecular Biology) initiative, helping CCPs that support highly specialised fields to work together.

The Outcome

Facilitating innovative cryoEM/T methods enabling new therapeutics in structural biology

Over the next decade CCP-EM will continue drive software development and large-scale computing support to facilitate state of the art cryoEM/T methods for the next generation of structural biologists that will in turn enable new therapeutics by providing essential tools and training for researchers in both the academic and industrial sectors.

Grand Challenges

More Information

CoSeC
www.CoSeC.ac.uk
CoSeC@stfc.ac.uk

CCP-EM
www.ccpem.ac.uk
ccpem@stfc.ac.uk