What can electron microscopists and crystallographers learn from each other?

Chris Wood
Scientific Computing Department, STFC Rutherford Appleton Laboratory, Didcot, OX11 0FA
chris.wood@stfc.ac.uk / 01235 567864

The STFC have recently been awarded an MRC Partnership Grant to establish a Collaborative Computational Project for Electron cryo-Microscopy (CCPEM), in an effort to both support the work done by software developers within the UK EM community, and to provide assistance to EM users who require help with computational aspects of their work. The first of two developers funded by the grant (me!) started at the end of August 2012, and shares an office with the core CCP4 team in the Research Complex at Harwell. The hope is that I can use the lessons learnt by CCP4 over the years, and that the two projects can share some code. CCPEM has three principle aims:

- Build a UK community for computational aspects of cryo-EM. Provide a focus for the cryo-EM community to interact with CCP4 and CCPN (Collaborative Computational Project for NMR), and the broader international community.
- Support the users of software for cryo-EM through dissemination of information on available software, and directed training.
- Support for software developers including porting, testing, and distribution of software.

Neither I, nor Martyn Winn (Head of Computational Biology, STFC), have an EM background, so we have spent the first couple of months of the project familiarising ourselves with a range of EM software and the problems faced by both users and developers. I have visited users (Helen Saibil, Elena Orlova & Carolyn Moores at Birkbeck College, Louise Hughes at Oxford Brookes University, Kay Grünewald at Oxford, and Ariel Blocker's Group at Bristol University) and developers (Alan Roseman at Manchester, Maya Topf at Birkbeck, Sjors Scheres at the MRC-LMB, and Juha Huiskonen at Oxford) to get a feel for what CCPEM should deliver. We have also held a very useful one day community meeting in Leeds (for which Neil Ranson and Arwen Pearson need to be thanked for local organisation). We had a good attendance of 35 researchers, who between them represented users, developers, and modellers within structural biology. We also had excellent presentations from Ardan Patwardhan (EBI) on how the EBI/PDBe/EMDB and CCPEM could collaborate on validation and deposition, and David Bhella (Glasgow), Ariel Blocker (Bristol) & Ed Morris (Institute of Cancer Research) on the computational problems that they, as EM users, have encountered in their research, how they overcame the problems and how they think that CCPEM could assist them in future. As a result of the information gathered at the

meeting, Martyn and I will decide on specific outcomes for CCPEM and feed them back to the community, to ensure that we will deliver useful services; however, it seems likely that in the short term we will focus on supporting currently available software, and consolidate the software being written by UK-based developers into a single software suite that can be distributed. In the longer term, we will look at the feasibility of providing computational services on behalf of the community; this could either be in the form of a *sandbox*, whereby users can remotely test software before deciding which software would be most useful for their needs (which they would then download to a local service, and so carry out all image and data processing themselves), or a more complex setup of a small cluster, with an ability for users to upload their data. This would have the advantage that users would not need to maintain their own computational facilities. A brief set of minutes from the community meeting, and the presentations given by Martyn and me, are on the CCPEM website (www.ccpem.ac.uk/courses.php).

We have also started to support two software developers. Alan Roseman's Find-EM package – which is a set of Fortran programs to help with automated particle picking from electron micrographs, is now available from the CCPEM website (www.ccpem.ac.uk/download.php). We are assisting Alan by extending parts of the program, and helping him to rewrite parts of it so that it will be more cross-platform compatible, with the aim being that it will be able to be run on Windows, Mac OSX, and Linux computers. We are also in the early stages of assisting Maya Topf's group, who are developing a set of python libraries to help determine the structure of macromolecular assemblies using flexible fitting techniques. Initially, we aim to host the library on the CCPEM website, before developing a front end for it and incorporating it into a more generic software suite. In addition, we are involved in discussions regarding the MRC format, which is in principle

identical to the CCP4 map format. In cryoEM, the MRC format is used for images (or stacks of images), for 3D volume data, or for stacks of volumes (4D data). Due to perceived limitations in the format, a number of extensions have arisen which cause problems when files are transferred between software packages.

We are very happy to be contacted by anybody who has an interest in EM, and who thinks they could benefit from the work that CCPEM aims to do. We also have a mailing list (www.jiscmail.ac.uk/ccpem) to which anyone can subscribe. Information about future community meetings, and the minutes from working group meetings, will be distributed on the mailing list and posted on the website (www.ccpem.ac.uk).