CCP-EM Working Group 1 2023 minutes

Virtual meeting on Tuesday 2nd May 10:00 to 12:00

Present: Martyn Winn, Tom Burnley, Colin Palmer, Agnel Joseph, Matt Iadanza, Neil Ranson, Maya Topf, Vicki Gold, Ste Muench, Corinne Smith, Jason Van Rooyen, Gerard Kleywegt, Giulia Zanetti, Lori Passmore, Martin Singleton, Michael Grange, Nobuhiro Morone, Peter Rosenthal, Pamela Williams, Randy Read, Robert Deller, Sjors Scheres, Alan Roseman, Dan Clare, Dan Rigden, Dave Lawson, Yuriy Chaban

Actions

All: Contact Corinne with feedback for the SAB.
All: Make suggestions for Scientific Organisers for next year’s Spring Symposium
All: Contact CCP-EM team if want to be part of the Doppio roadshow.

1. Welcome

Martyn summarised the remit and membership rules for CCP-EM Working Group 1. Membership is restricted to PI-level and UK-based, and the meeting agreed that this was still appropriate.

Corinne Smith (Warwick) briefly introduced the Scientific Advisory Board for the EMBL-EBI Molecular and Cellular Structure resources https://www.ebi.ac.uk/pdbe/scientific-advisory-board, of which she is currently chair. Some of the SAB are also members of the wwPDB Advisory Committee. She would welcome any input from the community ahead of the next meeting - please email her directly.

2. Plan for Pipeliner / Doppio

Matt gave a brief demo of the new Doppio interface (more details can be found in the video of his Spring Symposium talk). Coding of the Pipeliner toolkit began in April 2020, and of the Doppio interface in Jan 2022. Having announced the first public beta release at the Spring Symposium (available from https://www.ccpem.ac.uk/download.php), we are now planning to make approximately biweekly updates. The top priority now is to implement asynchronous job running, allowing the user to run more than one task at a time. Other major developments include EMPIAR deposition, interfacing with OneDep deposition, and pipelines for subtomogram averaging.

Despite not being strongly advertised, the Pipeliner toolkit has been adopted by Dan Hatton for the single particle pipeline at eBIC, and by a Dutch startup company CryoCloud. Pipeliner is Open Source (Mozilla Public Licence v2) to encourage wide take-up.

Giulia suggested that the CCP-EM core team should do workshops and site visits to encourage adoption of Pipeliner/Doppio, and this is indeed planned now that the beta release is out. Agnel confirmed that Doppio will be the platform used for the Icknield model building workshop in October. We plan to install CCP-EM/Doppio on the Jade2 Tier 2 high performance compute cluster, and use this to support workshops.

3. Tomography in CCP-EM

As one of its grant deliverables, CCP-EM will look at pipelines for sub-tomogram averaging. These will use the Pipeliner framework, and will be a natural extension of the single particle work in Relion. Rangana joined the CCP-EM team in March to focus on tomography pipelines.

We will need to interface with external software to reconstruct the tomograms. Relion uses the original tilt series, and the tomogram is only required for locating the particles. There was a discussion of how good the tomogram needs to be in this context. For other applications, more interested in cellular volumes, the tomogram quality will be important. In addition, it is often useful to map sub-tomogram averages back to their original location. Michael Grange offered to provide some input to our development plans.

Peter stated that CCP-EM could play an important role migrating those used to single particle analysis towards trying tomography. There are currently many codes and many workflows for tomography. We need to be aware of these, though not necessarily implement all. There are queries over the long-term stability of some software, and CCP-EM can help to provide a more robust backbone of code. While there are many directions we could take, it is important to make a start. If we are careful to stay modular, then we can expand and adapt later.

4. Industrial feedback

Pamela Williams (CCP-EM industrial representative) tried to solicit feedback from companies on their use of CCP-EM via an open email to the CCP-EM list and direct emails to the licencees. There was one response, as well as local feedback that Pamela asked for from her team.

Although limited, the feedback was clear that there was a need for automation, speed and smaller files. It was felt that cryoSparc do better here. The speed of motion correction was mentioned in particular. Automation would be important for ligand/fragment screening, but not all companies are there yet. Nevertheless, there will be a need for ligand tools, as in CCP4.

Jason reported the view from eBIC. Some, but not all, companies invest in cryoSPARC. Many want help in data processing, either at eBIC or via a cloud platform.

Sjors commented that the work with Erik Lindahl optimising Relion and porting to GPUs is now a few years ago, and it would be timely to revisit this, e.g. re-implementing relion_refine using PyTorch.

We discussed whether the CCP-EM core team could visit companies and help with installation and set-up. This will vary a lot between companies. Many have their IT departments on a different continent and/or would be unwilling to have outsiders access their systems. We need an installation that is straightforward for an IT team (Neil commented that this would be useful for universities too). STFC has a trial licence to enable commercial users to try the software before committing. Many companies develop internal pipelines that connect to corporate databases. Modular suites such as CCP4 / CCP-EM are useful for this.

Tom confirmed that we have been invited to the Protein Structure Determination in Industry (PSDI) meeting in November.

5. Finances

CCP-EM receives funding from the MRC Partnership Grant, commercial licensing, and occasional additional grants. The MRC grant funds the core team Tom, Colin and Agnel, and a small component for collaborative activities. The latter was reduced on the assumption that the commercial income can provide this. In the past year, most of the commercial income has gone towards development of the Pipeliner/Doppio software. This will continue to be the case, but we can nevertheless support additional small projects or studentships. The commercial fees have been raised for the first time, by approximately 50% from April 2023. In future, we aim to make smaller but more frequent increases, to the extent necessary to maintain the growing suite.

6. Training

We have run fewer workshops recently, as developer effort has been focussed on Pipeliner/Doppio. Nevertheless, we ran the first post-pandemic Icknield workshop in September 2022, and two developer meetings. We have also contributed talks or practicals to several events organised by others.

As above, we now plan to advertise and provide training for Pipeliner/Doppio now that it is released. There was a discussion whether we can attach a workshop to the next Spring Symposium. We had a couple of satellites on the morning of the first day this year. Alternatively, we could consider a drop-in Doppio clinic.

7. Spring Symposium 2023 and 2024

Thanks to Vicki and Ste for organising this year’s programme, and to Martin Walsh and Lorna Malone for the Biological Cryo-Imaging user meeting.

Ste and Vicki reported back on the organisation of last week’s Spring Symposium. Some of the desired speakers were double booked by the time they got in touch (November) so could not make it. This affected the diversity, with replacements often being male: the percentage of female speakers dropped from 70% to 40%. We need to start the process earlier next time.

Feedback from one company is that they are happy to do other things such as demos, which would add to the sponsorship level. In principle, this would be helpful, as long as it doesn't affect the nature of the meeting.

The poster session seemed to go very well, and great to see lots of people there. Slido for voting worked well. There was a suggestion to have flash talks for poster presenters, but that would take a long time (we had 65 posters this year). Conversely, there seemed to be low usage of Slack for questions and parallel discussions.

There were a few comments that the extra time worked well in terms of long breaks, with lots of time for networking and catching up. Randy commented that this worked better than the equivalent CCP4 Study Weekend. Gerard felt that there was still a lack of big names present, that the students would like to meet.

There was some feedback appreciating the mixture of technical and biological within one session, rather than separating topics into different sessions.

This year, we had the potential to fund a number of bursaries and invited people to contact us. The take-up was disappointing. This may be a consequence of the pandemic, with travel budgets still being used up. It may also not have been advertised early enough. There was support to continue a targeted approach, rather than a reduction for all. The bursary didn't include travel, and that might be a problem for those from far away. Giulia reported that EMBO have childcare grants to help parents to travel.

Next year's Spring Symposium will be the 10th edition and we plan to mark this in some way. The first step is to appoint scientific advisors for next year. We would welcome any ideas or volunteers for additional elements of the programme.