Agenda

Meeting title:	CCP4 Working Group 2 meeting	
Date:	Wednesday 31st March 2021	**Time:** 9:30 – 13.00
Location:	zoom	
Circulation:	ccp4wg2@stfc.ac.uk	
Attending:	Jon Agirre, Emad Alharbi, Svetlana Antonyuk, Charles Ballard (CB), Arnaud Basle, Jon Berrisford (JB), Ben Bax, David Briggs, Ivan Campeotto, Lucrezia Catapano, Kevin Cowtan, Eleanor Dodson (ED), Paul Emsley (PE), Maria Fando, Daren Fearon, Ralf Flaig (RF), Vilmos Fulop (VF), Jonathan Grimes (JG), Misha Isupov (MI), Robbie Joosten, Ronan Keegan (RK), Nick Keep, Petr Kolenko, Eugene Krissinel, Andrew Lebedev (AL), Ed Lowe, Airlie McCoy, Karen McIntyre (KMT), Stuart McNicholas (SMN), Garib Murshudov (GM), Rob Nicholls (RN), Nikos Pinotsis, Randy Read (RR), Dan Rigden, Mark Roe, Alan Roseman, Paul Rowland (PR), Roberto Steiner (RS), Ivo Tews, Ville Uski, Isabel Uson, Sameer Valenkar, Melanie Vollmar (MV), Keith Wilson, Marcin Wojdyr, Keitaro Yamashita	

5 mins breaks at every full hour

1. Approval of the minutes from the previous meeting (Ivo Tews)
2. Chair's report (Ivo Tews)
3. Meetings and workshops (Charles Ballard)
4. Study Weekend 2022 – format (Ivo Tews, Charles Ballard, Ronan Keegan, Karen McIntyre)
    a. Can the CCP4SW go hybrid?
    b. Session structure, length of the meeting
    c. EDI – looking for a "champion"
5. Science talks
    a. Core group – CCP4 updates (Charles Ballard)
    b. Robbie Joosten – PDB-REDO tools for CCP4 - the prioritisation situation
    c. Petr Kolenko – PAIREF
    d. Emad Alharbi – Predicting performance of automated crystallographic model-building pipelines
    c. Lucrezia Catapano – Modelling and refinement of hydrogens in cryo-EM, electron and neutron diffraction experiments
    d. Helen Ginn – Vagabond
    e. Keitaro Yamashita – Structure refinement and map calculation for single particle analysis
6. Take note of the next meeting; suggested: June 9th 2021
7. AOB

Minutes

1. **Approval of minutes from the online WG2 meeting 20/1/21 (Ivo Tews)**

The minutes from the online WG2 meeting 20/1/21 were approved online.

2. **Chairs report (Ivo Tews)**

The CCP4WG2 meeting In the time since the last meeting, CCP4WG2 is delighted that Svetlanta Antonyuk, Melanie Vollmar and Rob Nicholls have agreed to organise SW2022 – "refinement". They will be alle to provide a sketch of the program in the next meeting.

The structure of the meeting can change to hybrid if possible, the length of the meeting has to be decided. The chair will ensure that lessons learned from the online-only 2021 study weekend will feed into organisation of the SW2022.

3. **Meetings and Workshops (Charles Ballard)**

All of this is impacted by COVID, but plans are currently in transition from virtual to hybrid or in person, but the meetings listeed here are still mainly planned as virtual events. Workshops use mainly CCP4 cloud (APS are using Amazon cloud). Generally speaking, virtual events are 50-100% longer than in-person meetings.

Meetings run or supported by CCP4:

- DLS Workshop (virtual) Dec 2020
- SW2021 (virtual) Jan 2021
- South Africa (virtual) Feb-March 2021
- APS (virtual) June 2021
- South America (virtual) Sept 2021
- DLS workshop (potentially attendance) Nov/Dec 2021
- SW2022 (hybrid?) Jan 2022
- Israel (potentially attendance) Feb 2022

Online meetings are reasonably popular – 60+ applicants for APS despite very little advertising.

Hybrid workshops have not been done so far, and it might be hard to manage interactions between in person and virtual attendees/speakers.

There may be an advantage for the online format as the chat function enables people to open up to ask any question; this lowers the threshold and is essential to learning.

Everyone misses the social elements in virtual/hybrid meetings – student interaction with each other and with speakers. Discord / Slack can do this to some extent.

KSW: Developers meeting at Cosener's is 6-8 July, we will have to revisit in May whether this is attendance or hybrid for UK based people or hybrid.

4. **Study Weekend 2022 – format (Ivo Tews, Charles Ballard, Ronan Keegan, Karen McIntyre)**

The decision has not been taken at this point, but discussions are ongoing how this is organised. This also respects inclusion criteria, for which hybrid may be an advantage,

a. **Can the CCP4SW go hybrid**

The meeting can go hybrid, that is attendance in person and via a virtual hub. The present venue is booked at Nottingham (5.-7.1.2022), and with "some" social distancing has a capacity of about 200, this would be limited attendance to UK based participants, only. Accommodation would be Hotel only. COVID H&S concepts must be in place, and we need to put liability on the agenda (Covid outbreak) – check with SHE. In a hybrid meeting, we would need to encourage speakers to be there in person if UK based as otherwise we would lose a lot of interaction with the audience. The organisers will require a steer also for planning the length and format of sessions.

RR: challenge in hybrid meetings is that the live part is going on for too many hours for an online participant to stay on.

David Briggs: Consider rolling back to full-virtual.

KMT: Yes that is in place. One thing: hybrid would be more work for chairs as it requires them to navigate questions from the chat and the live audience, but it has been done before.

b. **Session structure, length of the meeting**

Suggestion for a hybrid meeting, taking the points above into account: keep science sessions to the pm and put local meetings in the am, such as LTB, What's New, DLS User meeting etc. (these are recorded and streamed for everyone to access when outside the time-zone). It is probably a plan for a three-day hybrid meeting.

SMN: Would this mean that the "What's New" session does not get so much interactive use?

RN: Having "What's New" on 2nd day could get more people attending rather than seeing it as optional.

RF: DLS flexible, but it is in conjunction with the SW. Biggest constraint is DLS start up schedule. But we would need to know the structure of the Study Weekend first. The important thing that was missing last time is interaction.

For the LTB there were a number of lessons learned. For example, a moderated of the discussion would be good. We should consider a scientist and moderator in pair to run these.

Further lessons is to create space to meet, to have discussions, have interactive elements to enable interaction, do not leave out any of the other groups like on-line / different time-zones. Having a social "wind down" programme will require a social organisers team that supports the scientific team.

c. **EDI – looking for a "champion"**

Virtual and hybrid provide a new and more democratic meeting culture in which speakers and attendants are level, which helps EDI: it plays to equality, diversity, and inclusion (gender, family, early career). Online meetings do not replace "real" in attendance meetings, as they lack personal quality and interaction. CCP4 is looking for EDI champions!

KC: I have put some links in the chat to guide my own behaviour, like the IC conference policy. There are also links to ten easy rules and a more recent paper that goes wider than that. Having a policy and providing data to transparently show would be key.

SMN: We would get more inclusion from a hybrid meeting!

Jon Agirre and Karen McIntyre put themselves forward and were elected. Valerie Vollmar expressed interest to join as well.

5. Science talks

a. Core Team News (Charles Ballard)

CCP4 Current release v7.1 is the last to support python2. Update April (013) includes coot, crank2, arcimboldo, phaser and acedrg. Updates are ongoing for Cloud, monomers, and a few more.

Next release will be v8.0, but there are issues with Qt5.15.2 and Windows and switching over to MSVC. CCP4v8.0 is migration to Python3 with pyside2. Updates include i2 and ccp4mg, planned are Dials 3 series and pdb_redo.

Online offerings, moved to STFC/cloud/IRIS, sees increased usage with 8500+ jobs run with 800+ active users, with the most popular offerings being molecular replacement pipelines (Balbes, MrBump, MoRDa)

CCP4 cloud been used in multiple workshops and university courses. Many jobs submitted with little lag, really good service, approx. 1200 registered users.

EL: extremely good use of cloud with 50+ users.

CB: If you want help setting up on Cloud, contact Core Team who will be happy to help.

JB: how much work is required for CCP4 validation reports?

CB: 99% will stay same, just different python libraries.

ITE: Dials 3?

CB: Python 3 only, few changes only. Xia2 some changes about, DUI is more complicated.

b. PDB-REDO tools for CCP4 - the prioritisation situation (Robbie Joosten)

Robbie showed a number of tools and queried the order in which they should be ported into CCP4, there are now over 60 tools. There are presently >2000 active users and the PDB-REDO database has over 170,000 entries. The complexity of PDB-REDO is a challenge for CCP4 integration; simple tools are Fortran and Python based, but there are new C++ tools as well as third party software (needs replacing). Access to PDB-REDO is through a web-API with testing of CCP4i2 integration underway (SMN).

Tools covered were *DSSP* (new v4 with full mmCIF support), *Tortoize* as a validation tool for the Ramachandran plot and including PDB Rama-Z, the very popular *paired refinement workflow* (9% of all jobs) is adapted from Kay Dietrich's algorithm and uses a number of tools such as *binliner* to decide resolution steps and *resolute* for cutoffs under a wrapper, *Platonyzer* which generates restraints for metals including proper angles and can add LINK and remove SSBOND records and allows distinguishing metals such as Na^{+} and Mg^{2+}, *Hoder* which generates restraints from homology information using H-bond restraints weighted on local distance variance but requires the PDB-REDO database.

Tools under consideration are *DNATCO* (third party) for validation of DNA local structure, NucrmsZ for validation of Watson-Crick pairs that requires DSSR (third party), the mmCIF toolkit, and density-fitness for per-residue density metrics.

There was a lot of discussion on both *Platonizer*, in particular handling and deposition of restraints as well as communication with other programs, and on the implementation of the *paired refinement workflow*.

JB: density-fitness is high on our priority list for the PDB, with active work going on.

There was a long discussion on determination of density features / fitting methods (KC/ED)

c. PAIREF (Petr Kolenko)

Workflow introduced by Robbie and written ready for integration into CCP4. Pipeline to analyse data using cctbx, make cutoffs, takes (complex) decisions, and improves refinement using all data that contributes. It is fully cutsomisable, works with ligands and can do cross-validation. Graphical output, final decision depends on user. Further work will include model validation.

KSW: it would be nice for inexperienced users to give guidelines for decision taking?

KC: the report has a wealth of information that an expert will like. But there needs to be some intelligent decision-making approach that would make this complete.

RJ: we use four different criteria to cut resolution (Rfree, free weighted R, free log likelihood, free correlation coefficient) and that might smooth decision making. Another remark, the cross validation is really nice, and could be linked with Rcomplete which is a good metric.

d. Predicting performance of automated crystallographic model-building pipelines (Emad Alharbi)

A comparison of auto-building bucaneer, ARPwaARP, Phenix autobuild, Shelx. There is no one best solution, but there is usually a best for a certain protein, so the question becomes choosing of the right tool. Data was collected for research papers 2010-2020, then checked whether a better tool could be found for a given problem, and that was the case in a couple of examples.

e. Modelling and refinement of hydrogens in cryo-EM, electron and neutron diffraction experiments (Lucrezia Catapano)

NMX refinement is focussed on hydrogen structure. Interestingly, in EM hydrogens can be seen in rather low resolution (2 Å). For NMX, ideal distances are taken from neutron data analysis (Allen and Bruno, 2010) and QM calculation (Fei Long) and put into AceDRG, generating a new column in monomer files, available in next release.

f. Servalcat, Structure refinement and map calculation for single particle analysis (Keitaro Yamashita)

Single particle analysis real or reciprocal space refinement are available, the latter was covered. Principles of half-map and FSC were introduced, as well as map (over-) sharpening. The workflow for *Servalcat* with refmac5 was discussed, and map calculation explained, including local B weighting.

6. Take note of the date of the next meeting.

Proposed date: proposal 9th June 2021, online.

7. AOB

None.