



# Agenda

|                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                            |
|-----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|
| <b>Meeting title:</b> | CCP4 Working Group 2 meeting                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                            |
| <b>Date:</b>          | Thursday 20 <sup>th</sup> January 2022                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | <b>Time:</b> 13:00 – 15.00 |
| <b>Location:</b>      | Zoom                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                            |
| <b>Circulation:</b>   | ccp4wg2@stfc.ac.uk                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                            |
| <b>Attending:</b>     | Jon Agirre, Svetlana Antonyuk (SA), Charles Ballard (CB), Arnaud Basle, Ben Bax (BB), Paul Bond (PB), David Briggs (DB), Lucrezia Catapano, Kevin Cowtan (KC), Judit Debreczini, Eleanor Dodson (ED), Paul Emsley (PE), Maria Fando, Ralf Flaig (RF), Vilmos Fulop, Mike Hough (MH), Misha Isupov, Robbie Joosten (RJ), Ronan Keegan (RK), Petr Kolenko, Oleg Kovalevskiy (OK), Eugene Krissinel (EK), Andrey Lebedev, Ed Lowe, Airlie McCoy, Stuart McNicholas (SMN), Rob Nicholls (RN), Claudia Milan (CM), Peter Moody, Garib Murshudov (GM), Rob Nicholls (RN), Nick Pearce (NP), Andy Purkiss (AP), Dan Rigden, Mark Roe, Alan Roseman, Paul Rowland, Roberto Steiner, Kyle Stevenson, Ivo Tews (IT), Johan Turkenburg, Ville Uski, Isabel Uson, Melanie Vollmar (MV), Clemens Vonrhein (CV), Keith Wilson (KSW), Graeme Winter, Keitaro Yamashita, Briony Yorke (BY) |                            |

*We will break at least 5 mins around the full hour*

Starts 13:00.

1. Approval of minutes from the WG2 meetings 22/09/21
2. Chairs report (Ivo Tews)
  - a. Revision of the WG2 meeting structure
3. SW2022 (Svetlana Antonyuk, Melanie Vollmar, Rob Nicholls)
  - a. Lessons learned
  - b. ED&I and SW2022
  - c. Special issue
4. SW2023 (Ivo Tews)
  - a. Recommendations from WG1
5. CCP4 – core group news (Eugene Krissinel)
  - a. Meetings and workshops (Charles Ballard)
  - b. Updates, v.8.0 (Charles Ballard)
6. Take note of the date for the next meeting (proposal 23.02.2022)
7. AOB

# Minutes

## 1. Approval of minutes from the online WG2 meeting 22/9/21 (Ivo Tews)

The minutes from the online WG2 meeting 22/9/22 were approved.

## 2. Chairs report (Ivo Tews)

WG2 has reported to WG1 and CCP4-Exec on the 10<sup>th</sup> of January. Topics were the organisation of the study weekend, including feedback to SW2022 and planning of SW2023, advances on ED&I, and software featured in WG2 proposed for inclusion in the suite.

The report reviewed software presented in the WG2 meetings in 2021 and their implementation, as in:

WG2 January: SW review, no software featured

### WG2 March

- PDB-REDO tools for CCP4 (Robbie Joosten); RJ commented that DSSP and Tortoise were implemented; more to follow.
- PAIREF (Petr Kolenko); PK commented this is in process of being implemented (for CCP4v8).
- automated crystallographic model-building pipelines (Emad Alharbi); the development following this exercise is ModelCraft from the Cowtan group which CCP4 will implement in CCP4v8; this works and was shown and featured in the SW2022.
- hydrogens in EM and neutron diffraction (Lucrezia Catapano); LC, RS, GM commented these developments are already there and work and will be released soon.
- EM refinement and map calculation (Keitaro Yamashita); released as part of CCP4v8

### WG2 June

- FindMySequence for X-ray and EM (Grzegorz Chojnowski); being implemented for release in CCP4v8
- Refinement with VagaBond (Helen Ginn); technical requirements were already discussed with the core group. GM commented that this is an interesting development that should be included if there is user demand.
- Shift Field refinement (Paul Bond); this was a report of already implemented software; PB said there were no new developments for implementation, presently.
- Refinement / Ligand Fitting in CCP4 Cloud (Oleg Kovalevskiy); OK and EK reported these pipelines were implemented and work in CCP4cloud.
- Structure validation and analysis in coot (Paul Emsley); new features continue to be developed and implemented, though focus is on coot Python3 release for CCP4v8.
- Iris: validation metrics for protein models (Jon Agirre); depends on Python3 and will be part of CCP4v8.

### WG2 July

- Multi-crystal data – xia2.multiplex; this is being implemented in CCP4i2 and will be released with the first update of CCP4v8.
- Serial (single shot) crystallography; there is a dedicated post required here. IT to follow up.

### WG2 September: focussed on AlphaFold2

- Sequence search of MrParse against PDB/AFDB (Dan Rigden); implemented and being actively developed.
- Using the AlphaFold Database for MR in CCP4i2 (Ronan Keegan); implemented and available.

The open question remains whether CCP4 will require its own implementation. The chat discussed the just released CC BY 4.0 License from Deepmind for AF2 (post by CB). Implementation and use will be further discussed in the next WG2 meeting/s.

Following a doodle poll, participants opted overwhelmingly for shorter WG2 meetings (no more than 2 hrs). Shorter means (1) that the mixed programme of housekeeping AND science will be difficult to maintain in a

single meeting and will probably reflect in alternation of two different types of meetings. This may have the effect of dividing the audience into two groups of participants. Short also means (2) that the frequency of WG2 meetings necessarily must increase, roughly to eight meetings per year, to fit all activity WG2 hosts. Organising additional meetings, however, will require more time for organisation and minuting. In contrast, there are several advantages for short meetings (3) as WG2 may volunteer others to organise meetings on specific topics, which provides a path to be more inclusive and diverse – different people have different ideas. (4) Further, having new organisers supports ECRs and future talent. Based on these ideas, IT made a proposal for a new meeting rota:

- i. Meetings that deal with activities of organising the study weekend or workshops, and report on core team activities in general. These meetings will also allow discussion of software for/in the suite.
- ii. Meetings that deal with a specific topic, are based around science, or come from a particular group of developers who want to feature their efforts or strategy. These meetings will also allow discussion of novel software to be included the suite.

RN: I'd say to not have meetings for the sake of meetings - the WG2 "science meetings" should only happen if/when there are developments to communicate and discuss.

MH: Topics would be community driven.

GW: There are groups like the DIALS team who occasionally get invited to WG2 for individual talks. However, they may structure an entire meeting which would give better room to demonstrate the breadth.

KSW: Shorter meetings are better for online. When in-person meetings are held, of which we may have one a year, they should probably be in London, and give people enough time to travel in a day and back.

After the discussion, the proposal of having alternating meetings was accepted by vote.

### 3. SW2022 (Svetlana Antonyuk, Melanie Vollmar, Rob Nicholls)

#### a. Lessons learned

WG2 thanks the organisers Svetlana Antonyuk, Rob Nicholls, Melanie Vollmar for a brilliant meeting that was well received by the attendees! The structure of the meeting with 8 sessions was well received, although it was frequently said that longer breaks and an opportunity to meet was missed.

- Some time from the break was used for discussion, which, according to the organisers, was on purpose. WG2 suggests to formally declare a "buffer" time for discussions after the sessions to manage expectations.
- Technically, the chance to meet did exist e.g. through slack, but it is clear from the meeting that breakout rooms must be better facilitated in future meetings to allow spontaneos meetings and discussions.

The social party of the meeting has seen many proposals and ideas at the end of the social sessions. However, before the meeting it was difficult to get people to engage and organise social sessions.

- In the future, WG2 should appoint both, a scientific organisers team and a social organisers team for the study weekend.

The format of a hybrid meeting, although not possible this year, is expected to stay in future, i.e. we should continue to allow online participation. This will provide better opportunity for inclusion (visa etc.), reduce travel (cost, carbon footprint), and contribute to the success of the meeting in terms of more and more diverse participants.

The online format lends itself to shorter talks, which was criticised by some, but overall welcomed, as it promotes more and more diverse speakers, as well as giving younger speakers a chance (ECR). However, plenary lectures were very successful in terms of attendance, and we will always need introductory lectures of an educational nature, hence should have a mix of talks and lengths (as indeed we had in SW2022).

This time we have had three invited chairs per session, and additionally one of the organisers was present. This gave opportunity for diversity and gives ECRs a chance to chair. Having multiple people also helps with watching the chat and moderating questions. This was a success we may want to keep.

There was criticism on the technical side this year. The organisers reported embarrassing issues with the Cvent website (spelling mistakes, inconsistent formatting, speaker profile images) which, was felt, was unprofessional. Participants and organisers found the split between two web pages confusing.

- In a technical meeting after the event it was suggested to use Cvent in future only for registration / payment. However, several suggestions for the meeting website were also discussed.

Further, some participants had trouble using slack. While instructions were there and also send out in the mails, this help should perhaps be easier to find through better design of the web page and the email that had those instructions at the very end (where they were perhaps overlooked). Overall, however, the use of Slack was seen as a big plus, as conversations were kept, and could also continue outside sessions.

PROVISIONAL The number of registrations was with 1088 lower than last year, however actual participation at the meeting was similar to last year's, with 754, 587 and 571 unique viewers on the three days of the meeting. The Diamond Light Source satellite attracted 337 unique viewers.

Participant numbers for the Lunchtime Bytes (LTB) were also evaluated. The feedback by participants to LTB by participants was overwhelmingly positive.

## b. ED&I and SW2022

WG2 thanks the ED&I champions Karen McIntyre, Melanie Vollmar and Jon Agirre. They meet on a regular basis and have driven the agenda forward. Signs of their activity are obvious from the registration to SW2022, where attendees subscribe to a code of ethics and a code of conduct. For SW2022, we have seen the following (preliminary) figures:

- total participants: 1088 (male: 46%, female: 34%, n.r.: 20%)
  - Students: 399 --> 37%
  - PDRA and up: 689 --> 63%
- speakers:
 

|                                                                                                          |                                                                                                         |
|----------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>– male: 27 --&gt; 73%</li> <li>– female: 10 --&gt; 27%</li> </ul> | <ul style="list-style-type: none"> <li>chairs:</li> <li>13 --&gt; 54%</li> <li>11 --&gt; 46%</li> </ul> |
|----------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|

CCP4WG2 recognises that the lower number of females amongst developers has influenced the speaker list and will continue to drive these issues through the ED&I strategy.

Speaker invitations encouraged senior invited speakers to delegate presentation to an ECR, to provide opportunity and freshen the speaker demography. Several PIs took the opportunity to delegate to a suitable junior researcher, and WG2 is thankful for the positive response.

## c. Special issue

In the invitations sent by CCP4, prospective speakers have been told that contribution to the special issue is the preferred, but not a conditional requirement for acceptance. Several reasons lead to the change of policy away from a strict requirement. (1) Not making contribution a strict requirement gave more flexibility with selecting speakers, e.g. for case studies, unfinished developments, and younger speakers. (2) Some contributions have recently been published, so that the request for another publication seemed not justifiable. (3) The higher number of oral contributions in the online format of the study weekend meant that the special issue is easily filled: of a potential 34 oral contributions, about 18-20 are expected to contribute to the Special Issue. WG2 discussed and agreed to this proposal by the organisers.

CCP4 aims to give a fair overview of the Study Weekend in the special issue. WG2 remarked that the Special Issue should, along the guidelines given for the Study Weekend, also give room for an educational resource; and the organisers have confirmed that they did encourage educational articles. Offered alternatives to the

more typical research article were contributions in form of a topical review, a commentary, or an educational article.

## 4. SW2023 (Ivo Tews)

Study Weekends roughly follow a 5-yr rota with data processing / phasing / molecular replacement / refinement / model building, which are interrupted (about) every five years with a focus Study Weekend on a specific topic. However, there is an increasing number of voices in WG1 to suggest that the Study Weekend should be more general. Indeed, CCP4WG2 recognises that we may not need to follow a strict rota.

JA: Crystallography is being pulled in many directions. I believe now's the time to ditch the 5-year rota and go for a general meeting every year.

CM: I agree with Jon Agirre that kind of more 'general' topics, without losing the teaching part, but more open, could be a good idea.

RN: I agree with Claudia about Jon's comment - we shouldn't forget that the underlying objective is to inform students about what tools are available in CCP4, and things to think about them in practice.

CB: We may want to split the meeting across the days, as there is no need to have the same topic across all days of the study weekend.

RN: It's important to think about what's going on in the CCP4 suite. We should focus on giving exposure to modern developments.

DB: Yes, but we need to avoid the risk of turning the Study Weekend into one massive "What's new in ccp4".

Before the WG1 meeting, several topics have been mentioned, some carried over from the previous WG1 meeting in 2021. (1) Electron diffraction and EM were repeatedly mentioned. (2) The year 2021 has seen the advent of AlphaFold2, so this was mentioned as a Study Weekend topic various times. However, both topics were felt not make an "entire" Study Weekend, though they may contribute. (3) WG1 said that machine / deep learning would certainly be an interesting topic across the whole of Structural Biology, and this would include AlphaFold2, but also others such as crystallisation, data processing, structure solution and modelling. Indeed, crystallisation was mentioned several times. Finally, the question "what we do structures for?" was put forward, putting "Biology and Function" into focus. With respect to AlphaFold2 models it was clearly stated that crystallographic education is needed for structure interpretation, which is a core focus / activity of the Study Weekend.

The emerging topics of serial crystallography and time resolution have been taken up with a WG2 focus meeting in July 2021. These are also aligned with the 5 yrs rota that suggest the next Study Weekend would be on data and data processing. An interesting suggestion came from the Diamond Light Source, stating that sample delivery and preparation is a current challenge, and thus a Study Weekend on data may include crystallisation/crystallisability, fragment screening, serial synchrotron crystallography, time resolved crystallography, possible future capabilities at synchrotron and XFEL sources, and sample preparation for MX, SAXS, ED and EM.

The view of WG1 was that giving a general overview (for teaching) with special sessions would be a permissible format for a future Study Weekend.

JA: 'Integrative structural biology' is, I believe, the future for us.

ED: Critical thinking must not be neglected.

BY: From structure to function?

BB: What about chemistry and co-factors/ligands?

MH: As a session, perhaps jointly with ccpBioSym, what about dynamics and linking with molecular simulations - crystallography plus molecular dynamics / QM / MM?

DR: I like Structure -> Function but that's essentially a bioinformatics conference ;-)

RJ: Reading and writing structure data - from creation to interpretation.

JT: Data collection that does NOT use UDC is fast becoming the exception? Emphasis should be on data processing, assessing the output of pipelines, I imagine.

AP: Analysis of the data processing is still important. Lots of student's don't know 'how' to look at the results critically and look for problems once their structure solution has issues.

ED: Have we ever had a meeting to consider implications of enhanced data collection tools?

AP: Ligands and cofactors brings in Xchem and issues with non-isomorphism in the processing.

DR: How about a whole meeting "What AF2 can't do?"

CM: I would avoid to use a negative title, even if that is the idea behind the topic

NP: "What AF2 can't do?" — is that nothing more than just "what type of info isn't in the PDB yet"?

CV: I like the idea to bring data and AF together: importance of AF model validation against crystallographic data.

NP: Another example where the methods aren't mature yet.

RN: I think there needs to be a more solid pitch. A one-line title might not be sufficiently descriptive... Even though AF2 is utilised within CCP4, I don't think it's a good idea to focus solely on AF2 in an education-focused CCP4 meeting. AF2 will undoubtedly have a place, but to focus on it sounds more like a topic for a general scientific meeting on the latest novel thinking.

DR: Seems like we're not landing on an idea today.

CM: A kind of big-data in structural biology umbrella topic that would include from serial data to af2 to machine learning, etc.? It can fit quite a few of the things we have discussed.

DR: General Big Data is a nice idea.

JA: Big data/ AI good umbrella topic in this decade.

DR: Yes, ligands too. How about "Dealing with and exploiting Big Data in Structural Biology"?

KC: Big data is relevant, but dated as a term? Data driven? Data science?

CV: New challenges for structural biology - data, models and interpretation (to cover what AF2 can't yet do - but not as a competition, more working together towards new challenges).

JD: I like "Data driven" or "Data science" by Kevin. Broader than AI/machine learning and is a good umbrella.

RN: "Data driven" covers pretty much everything we do! (pretty much).

The topic for the next Study Weekend is now open to WG2 members for discussion through the mailing list.

## 5. CCP4 – core group news (Eugene Krissinel)

Current posts and assignments were discussed.

### a. Meetings and workshops (Charles Ballard)

The meeting schedule is severely affected by COVID restrictions. Meetings planned for 2022 are APS 13-24 June 2022 (virtual); ECM satellite 21-22 August 2022 at Soleil (in person?); DLS Nov/Dec at Diamond, and AsCA proposed as a 1-2 day meeting in Nov, Korea.

RN: There's also an in-person workshop in Lund Sweden scheduled for June, which may or may not go ahead, which will have a CCP4 presence?

### b. Updates, v.8.0 (Charles Ballard)

CCP4v8 is largely complete, and the Core team are tidying up and are in testing phase. Issues remain with ccp4i2 and the windows installer. Recent builds at <http://devtools.fg.oisn.rc-harwell.ac.uk/downloads/>. Test bundles to be available at <http://series-80.fg.oisn.rc-harwell.ac.uk/downloads> (pwd protected). Update 1 vor v8 is already planned with several developments that cannot be included in the initial rollout, for complexity and technical reasons.

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

Proposed date: proposal 23<sup>rd</sup> February 2022, online.

## **7. AOB**

Following the decision of having shorter meetings and appointing organisers for future WG2 meetings, suggestions were sought. Ronan Keegan volunteered to organise the next WG2 meeting together with Liverpool on alphafold2 and machine learning. A suggestion for summer was to invite the DIALS team to organise a WG2 meeting,