



Agenda

Meeting title:	CCP4 Working Group 2 meeting	
Date:	Wednesday 23 rd February 2022	Time: 09:00 – 11.00
Location:	Zoom	
Circulation:	ccp4wg2@stfc.ac.uk	
Attending:	Jon Agirre, Charles Ballard (CB), Arnaud Basle, Paul Bond (PB), David Briggs (DB), Eleanor Dodson (ED), Paul Emsley (PE), Maria Fando, Vilmos Fulop, Ronan Keegan (RK), Petr Kolenko, Eugene Krissinel (EK), Karen McIntyre, Stuart McNicholas (SMN), Rob Nicholls (RN), Peter Moody, Andy Purkiss (AP), Randy Read, Dan Rigden, Alan Roseman, Adam Simpkin, Ivo Tews (IT), Johan Turkenburg, Clemens Vonnrhein (CV), Graeme Winter (GW), Marcin Wojdyr, Keitaro Yamashita, Briony Yorke (BY)	

We will break at least 5 mins around the full hour

Starts 09:00.

1. Approval of minutes from the WG2 meetings 20/01/22
2. Chairs report (Ivo Tews)
3. Study weekend 2023 (Ivo Tews)
 - a. Proposal for the topic
 - b. proposal for organisers
4. Regional meetings
 - a. SWSBC (Ivo Tews)
 - b. Northern Meeting (Briony Yorke, Ivan Campeotto, Jon Agirre)
5. Take note of the date for the next meeting (proposal 16.03.2022)
6. AOB

Minutes

1. Approval of minutes from the online WG2 meeting 20/1/22 (Ivo Tews)

The minutes from the online WG2 meeting 20/1/22 were approved.

2. Chairs report (Ivo Tews)

Changes to the WG2 meeting structure and rota, went from 3-5 annually (half day) to 8-10 shorter meetings (2 hrs). Alternating between housekeeping meetings to deal with appointing organisers for the SW and for dissemination etc. Then science meetings for which we look for suggestions. Ronan Keegan volunteered to organise the next WG2 meeting together with Liverpool on alphafold2 and machine learning in March. A suggestion for end of April is for the DIALS team to organise a WG2 meeting.

The documentation project started at Southampton together with CCP4 core continues and migrate to Newcastle. There is also an effort to start a project on providing capability of interacting with serial data for CCP4, which has two elements: (a) data entry point for integrated serial data from any serial beamline; (b) pipelines with analytical tools and branch-points for user intervention for these data for downstream processing by the suite. A PostDoc is thought.

CCP4v8 rollout is on the way, and it include DIALS multiplex in update 1. CCP4 v8 is open for testing.

3. SW2023 (Ivo Tews)

a. Topic

CCP4 is normally on a 5-yr's rota with data processing / phasing / mol. replacement / refinement / model building, and then has an intermitted Integrative Structural Biology / From Crystal to structure, or another topic, which is out-of-rota. The discussion (WG1, WG2, online) first had a number of string voices that said we must move towards integrative meetings. However, there was then a growing number of voices then that said that we should not lose sight of the teaching and go in-depth on one topic, and perhaps the solution suggested here is to have a meeting with a broad topic as before but allow focus sessions.

The topic is clear and, following the rota, would be on data. The idea is to cover different modalities, which includes classic data processing, but then also fragments, MicroED, SSX and time resolved work, SAXS, and include possible future capabilities at SR and XFEL sources. Sample prep for MX, EM, XFEL and SAXS would complete this, covering crystallisation and crystallisability, or opportunity for structural work if you don't have xtals. The meeting would focus on software, instrumentation development and theoretical underpinnings.

Recognising the teaching remit, students will be given a chance to listen to discussion between experts in a way that doesn't happen in a training school/workshop environment. This also recognises that the PIs sending students may not be crystallographers. Increasingly, remote data collection and automated data processing distances users from the actual raw data, and we want to bridge that gap with knowledge.

ED: Schools and workshops look at simple things, that all you have time for. Study Weekends ought to look at more challenging things! Automation is almost a threat to data processing, focus on challenges.

PE: What fraction of data are good idea? What are bad data, how would you tackle / use them?

PM: Don't forget that a proper understanding of the process of data measurement etc is needed as framework for advanced discussion

Proposed working title of the meeting: **How Do Good Data Look Like and How Best to Use Them?**

- Introductory lectures covering some of the statistical aspects: expected data distributions (how derived) to cover optimal data collection/sample prep
- Pathologies (anisotropy, raddam, completeness, ice...) and optimal data processing (scaling - single/multi/serial datasets, expected statistical distributions, link back to data collection/sample problems)

This will put the meeting back to experiment (rather than prediction), which is what CCP4 is best at. Provides a link to machine learning, multi-dataset approaches, and considerations of non/isomorphism. And can explore how we best use linked but non-identical data sets such as time-series, temp series, ligand/fragment screens. How can we combine all these data into one coupled refinement that produces a model that captures and is consistent with all these variations.

KV: How you could combine different types of data would be a very interesting, but probably a topic for a meeting in itself.

IT: Maybe do this as a focus session.

b. Potential organisers

Candidates were shortlisted and the chair was asked to speak to the potential candidates.

4. Regional meeting.

a. SWSBC (Ivo Tews)

Members presently include Bath, Bristol, Cardiff, Exeter, Portsmouth, Reading, Southampton, Sussex. There are about 30 active PIs, and typically 70-100 in attendance, mostly students (free) and PostDocs. SWSBC covers X-ray, NMR, EM.

The meeting happens every year, which for a PhD means that in year 1 they would attend, in year 2 they present a poster, and in year 3 they give an invited talk. Further elements include ECR talks, exhibitor talks, keynotes. This year Bristol will host (Paul Race).

b. Northern (Ivan Campeotto)

A number of universities have been asked whether they would join this initiative, including Newcastle, Durham, York, Bradford, Leeds, Sheffield, Nottingham, Leicester. Suggestion to hold the first 'revived' meeting in York. Organisers JA, BY, IC (but open for others). Format: 2 (3 days). Now need to tie down a date.

AR: EM Manchester – we are interested to join.

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

Proposed date: proposal 16th March 2022, online.

6. AOB

None.