Agenda

Meeting title:	CCP4 Working Group 2 meeting	
Date:	Wednesday 16th March 2022	**Time:** 10:00 – 12.00
Location:	Zoom	
Circulation:	ccp4wg2@stfc.ac.uk	
Attending:	Jon Agirre, Charles Ballard, Benjamin Bax, Paul Bond, David Briggs, Iracema Caballero Muńnoz, Albert Castellvi, Lucrezia Catapano, Kevin Cowtan, Paul Emsley. Maria Fando, Ralf Flaig, Luiso Fuentes Montero, Hanna, Yunyun Gao, Misha Isupov, Elisabeth Jiménez, Robbie Joosten, Ronan Keegan, Nick Keep, Petr Kolenko, Oleg Kovalevskiy, Eugene Krissinel, Andrew Lebedev, Ed Lowe, Martin Malŷ, Airlie Mc Coy, A Medina, Peter Moody, Nick Pearce, Randy Read, Dan Rigden, Filomeno Sanchez Rodriguez, Mark Roe, Adam Simpkin, Kyle Stevenson, Ivo Tews, Andrea Thorn, Dusan Turk, Ville Uski, Isabel Uson, Sameer Velankar, Clemens Vonrhein, Keith Wilson, Marcin Wojdyr, Keitaro Yamashita	

We will break at least 5 mins around the full hour

Starts 10:00.

1. Approval of minutes from the WG2 meetings 22/02/23
2. Chairs report (Ivo Tews)
3. Alphafold2 and Machine Learning (Adam Simpkin)
    a) ML validation methods (Filomeno Sánchez Rodríguez)
    b) ARCIMBOLDO Shredder (Isabel Uson)
4. Alphafold2 in CCP4 8.0 demonstration (Ronan Keegan)
5. CCP4 8.0 (Charles Ballard)
6. Take note of the date for the next meeting (proposal 27.4.22)
7. AOB

Minutes

1. **Approval of minutes from the online WG2 meeting 22/2/23 (Ivo Tews)**

The minutes from the online WG2 meeting 22/2/23 were approved.

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

The CCP4 Documentation is well on the way. CCP4WG2 further has identified the need for a post in serial crystallography, to add two elements to CCP4i2: (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 (following on from the WG2 28th July 2021).

The Study Weekend 2023 has the working title "What is good data and how to best use …". Appointing organisers is in progress.

DUI, the DIALS user interface, is actively being developed by Luiso. Efforts are to increase Dui1 acceptance and use, and Dui2 development (server architecture), to be featured in a future WG2.

3. **Alphafold2 and Machine Learning (Adam Simpkin)**
4. **Alphafold2 in CCP4 8.0 demonstration (Ronan Keegan)**
5. **CCP4 8.0 (Charles Ballard)**

The main access routes for users to get predicted models are:

- **EBI-AFDB (alphafold database)** - expected to reach > 100 million predictions by late 2022 (Sameer)
- **Google Colab servers** - access to Alphafold and Colabfold on Google's GPU servers
- **RoseTTAFold server**
- **Local installations** of Alphafold, Colabfold, Openfold

CCP4 and predicted models

We've set up a working party to coordinate predicted model developments in CCP4. The key developments to come from this are listed in the following.

A standardised command-line API for CCP4 applications to interact with local installations of Alphafold/Colabfold/Openfold. This has been led by Grzegorz, with Ville and Eugene contributing. The API is now used in CCP4_Cloud to make model predictions for use in CCP4 Cloud tasks. Users can make predictions as well as invoke the prediction generation as part of the automated workflows for molecular replacement. The API also allows for presentation of results of the predictions (e.g. pLDDT or PAE scores). The plan is to also utilise it in i2 (which will require linking to a local AF setup).

CCP4 core team to source resources for running AF at the RAL site (Ville, Ronan and Charles)**.** We had planned to procure GPU hardware for this but STFC have instructed us to make use of STFC Cloud resources if possible. AF has now been set up on various STFC cloud systems as well as one CCP4 own machine. Initially these were CPU-based installations, but we are now gaining access to more GPU-based machines, enabling predictions to be generated much more quickly. Some of these resources have already been made available to users through CCP4_Cloud. Ville is also working on an MR pipeline to be made available through CCP4-Online. It remains to be seen if the STFC Cloud resources are suitable, we have a requirement for "on-demand" access which may not be feasible for the GPU resources which can be utilised only by a single user at a time and are in high demand from other STFC Cloud users.

New predicted model developments to be incorporated into CCP4 interfaces include the following:

- **Process_predicted_model** task (Phaser/CCTBX team and Kyle) -- already in i2. Same functionality included in CCP4 Cloud seamlessly for user and not as a separate task.
- **SliceNDice** (Adam, Luc, Dan, Ronan) - automated pipeline to process predicted models and run MR - in CCP4Cloud and work ongoing for i2 incorporation
- **Search tools for EBI-AlphaFold Database** - (Adam, Dan, Stuart, Ronan) MrParse, MrBUMP and CCP4mg have both been updated to cater for this
- **ARCIMBOLDO** (Uson Group) - can also make use of predicted models for MR
- **New i2 task menu** (Kyle and Stuart) - dedicated menu for predicted model tasks (all above)
- **New AF Structure Prediction task** – in CCP4 Cloud (Ville, Maria, Grzegorz and Eugene)
- **Automatic MR Project Workflow based on AF structure prediction** – in CCP4 Cloud (Oleg (now transferred to Maria), Eugene)

Documentation and testing for predicted model tools (Maria, Kyle, Oleg) – this is work ongoing.

A **Slack workspace (CCP4-Fold)** was set up for interactions between the interested groups in the working party. Dan Rigden also uses this to post interesting new developments and news in the area of predicted models. The first working party meeting took place in March with the next being part of the Developers Meeting. The API development and AF resources group have been meeting more regularly.

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

Proposed date: proposal 27th April 22, online.

7. **AOB**

None.